Published in Journals: Aerospace, Energies,
Materials, Sensors and Applied Sciences

Topic Reprint

Structural Health Monitoring and Non-destructive Testing for Large-Scale Structures

Volume I

Edited by
Phong B. Dao, Lei Qiu, Liang Yu, Tadeusz Uhl and Minh-Quy Le

mdpi.com/topics

MDPI

Structural Health Monitoring and Non-destructive Testing for Large-Scale Structures—Volume I

Structural Health Monitoring and Non-destructive Testing for Large-Scale Structures—Volume I

Editors

Phong B. Dao
Lei Qiu
Liang Yu
Tadeusz Uhl
Minh-Quy Le

MDPI

Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Editors

Phong B. Dao
AGH University of Krakow
Krakow, Poland

Lei Qiu
Nanjing University of Aeronautics and Astronautics
Nanjing, China

Liang Yu
Northwestern Polytechnical University
Xi'an, China

Tadeusz Uhl
AGH University of Krakow
Krakow, Poland

Minh-Quy Le
Hanoi University of Science & Technology
Hanoi, Vietnam

Editorial Office
MDPI
St. Alban-Anlage 66
4052 Basel, Switzerland

This is a reprint of articles from the Topic published online in the open access journals *Aerospace* (ISSN 2226-4310), *Energies* (ISSN 1996-1073), *Materials* (ISSN 1996-1944), *Sensors* (ISSN 1424-8220), and *Applied Sciences* (ISSN 2076-3417) (available at: https://www.mdpi.com/topics/shmndtlss).

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

Volume I
ISBN 978-3-0365-9933-5 (Hbk)
ISBN 978-3-0365-9934-2 (PDF)
doi.org/10.3390/books978-3-0365-9934-2

Set
ISBN 978-3-0365-9931-1 (Hbk)
ISBN 978-3-0365-9932-8 (PDF)

Contents

About the Editors

Phong B. Dao

Phong B. Dao received an Engineering degree in Cybernetics in 2001, an M.Sc. degree in Instrumentation and Control in 2004, both from Hanoi University of Science and Technology in Vietnam, and a Ph.D. degree in Control Engineering in 2011 from the University of Twente, the Netherlands. In May 2020, Dr. Dao received a D.Sc. degree (Habilitation) in Mechanical Engineering from the AGH University of Science and Technology, Poland. He is currently an Associate Professor at the Department of Robotics and Mechatronics of the AGH University of Krakow. His research interests include structural health monitoring (SHM), non-destructive testing (NDT), wind turbine condition monitoring and fault diagnosis, statistical time series methods for SHM and NDT, advanced signal processing, and intelligent control. He has been the recipient of several awards for scientific achievements and has published over 50 articles, mostly as the first, single, or corresponding author, in major international peer-reviewed journals and conferences. He is the co-inventor of a European Patent on the homoscedastic nonlinear cointegration methodology for SHM. He is the author of the monograph entitled "Cointegration-Based Approach for Structural Health Monitoring: Theory and Applications" published in 2018. He delivered five invited lectures on topics about structural damage detection and wind turbine condition monitoring. He serves as an active reviewer for many international Top-10 journals such as IEEE Transactions on Industrial Electronics, Expert Systems with Applications, Applied Energy, and Mechanical Systems and Signal Processing.

Lei Qiu

Lei Qiu is a doctoral supervisor at the State Key Lab of Mechanics and Control for Aerospace Structures, College of Aerospace Engineering, Nanjing University of Aeronautics and Astronautics, and the Director of the Institute of Smart Materials and Structures. His interests include aircraft smart structures and structural health monitoring (SHM) technology, and he has made innovative advancements in SHM principles and methods of complex structures, reliable diagnostic mechanisms for structural damage under the influence of service, aircraft smart skin based on flexible electronics, and equipment development for aerospace applications. Prof. Qiu has presided over more than 40 projects and is currently undertaking projects such as the general program of the National Natural Science Foundation of China, the Outstanding Youth Foundation of Jiangsu Province of China, key project subjects of Foundation Strengthening of Military Commission of Science and Technology (173) and the pre-research program of the Army Armament Department. He has received nine important honors and technological awards, such as the China Youth Science and Technology Award, two first prizes in Science and Technology from Jiangsu Province, the Jiangsu Province Youth May Fourth Medal, the Youth Science and Technology Award of Chinese Society of Aeronautics and Astronautics, and the best paper award of the 8th European Workshop on SHM, etc.

Liang Yu

Liang Yu received a bachelor's degree in communication engineering from Hubei University of Technology, Wuhan, China, in 2007, a double M.S. degree in communication and information systems from Shanghai University, Shanghai, China, and Waseda University, Fukuoka, Japan, in 2011, and a Ph.D. degree in acoustics from INSA de Lyon, Lyon, France, in 2015. He is currently a Full Professor with Northwestern Polytechnical University, Xi'an, China. His research interests include fault diagnosis of mechanical equipment, intelligent signal processing methods and advanced acoustic measurement, and experimental aeroacoustics.

Tadeusz Uhl

Tadeusz Uhl graduated from the AGH University of Science and Technology, obtaining his doctorate in 1983 and the title of professor in 1997. He worked in research centers and universities abroad in the Netherlands, France, Belgium, Japan and the USA. Since 2020, he has been the Director of the AGH UST Space Technology Center. His research interests include mechatronics, structural health monitoring, hydrogen vehicles, robotics, and space technologies. He is the author of over 1,000 publications, with over 400 in JCR journals. He has a citation index of h = 32 (according to SCOPUS). He is the promoter of 49 defended doctoral theses and over 150 master's theses. He is the Coordinator of many European projects in topics related to structural diagnostics, transport and energy with partners such as Alstom, Siemens, Airbus, Dassault, EDF, FIAT, RENAULT, LMS, Bombardier, and ABB. He carries out extensive cooperation with universities and research centers around the world. He has 51 national and foreign patents. He was a member of HLG in the field of Key Enabling Technology at the European Commission in 2014 – 2017. In 2020, Stanford University placed him on the prestigious list of "World's Top 2% Scientists". He effectively implements the latest technologies in the economy, and he is the founder of 28 start-ups, several of which have grown to the size of a large company.

Minh-Quy Le

Minh-Quy Le received his BS degree from Hanoi University of Science and Technology, Vietnam, in 1995, and his MS degree in Mechanical Engineering from University of Poitiers, France, in 1997. He obtained his Ph.D. in Mechanical Engineering in 2004 from Kyungpook National University, Republic of Korea. In 2007–2008, he worked as a Postdoctoral Fellow at the Technical University of Dresden, Germany. Currently, he is a Full Professor at the School of Mechanical Engineering, Hanoi University of Science and Technology, Hanoi, Vietnam. His research interests include computational mechanics, solid mechanics, nanomechanics, fracture mechanics, finite element method, molecular dynamics simulation, and material testing methods.

Preface

Structural health monitoring (SHM) and non-destructive testing (NDT) have gained significant importance for civil, mechanical, aerospace, and offshore structures. Nowadays, we can find SHM and NDT applications being used on various structures with very different requirements. The SHM-NDT field involves a wide range of transdisciplinary areas, including smart materials, embedded sensors and actuators, damage diagnosis and prognosis, signal and image processing, wireless sensor networks, data interpretation, machine learning, data fusion, energy harvesting, etc.

Since the 1970s, there has been a large and increasing volume of research on SHM and NDT; a great deal of this effort has focused on developing cost-effective, automatic, and reliable damage detection technologies. However, few industrial and commercial applications can be found in the literature. The practical implementation of strategies for the detection of structural damage to real structures outside of laboratory conditions is always one of the most demanding tasks for engineers. One reason for the rare transfer of research into industrial practice is that most of the methods that have been developed have been tested on simple beam and plate structures in the laboratory, while many practical problems only manifest themselves in complex structures. Another reason is the influence of environmental and operational variations (EOVs) on damage-sensitive features. Thus, for the successful development of SHM and NDT for large structures, techniques should be enhanced to have the capability of dealing with the influence of EOVs. In addition, signal/data processing plays an important role in the implementation of SHM and NDT technologies. The processing and interpretation of the massive amount of data generated through the long-term monitoring of large and complex structures (e.g., bridges, buildings, ships, aircrafts, wind turbines, pipes, etc.) has become an emerging challenge that needs to be addressed by the community.

This Topic brings together the most established as well as newly emerging SHM and NDT approaches that can be used for the detection and evaluation of defects and damage development in large-scale or full-scale structures. After a strict peer-review process, 44 papers were published, which represent the most recent progress in SHM and NDT methods/techniques for aerospace, civil, mechanical, and offshore infrastructures.

Phong B. Dao, Lei Qiu, Liang Yu, Tadeusz Uhl, and Minh-Quy Le
Editors

sensors

MDPI

Article

Improvement of Performance for Raman Assisted BOTDR by Analyzing Brillouin Gain Spectrum

Qiang Huang [1,2], Junqiang Sun [1,*], Wenting Jiao [1] and Li Kai [1]

[1] Wuhan National Laboratory for Optoelectronics, School of Optical and Electronic Information, Huazhong University of Science and Technology, Wuhan 430074, China; qiangh0115@hust.edu.cn (Q.H.); wtjiao@hust.edu.cn (W.J.); kaili@hust.edu.cn (L.K.)

[2] Hunan Provincial Key Laboratory of Grids Operation and Control on Multi-Power Sources Area, School of Electrical Engineering, Shaoyang University, Shaoyang 422000, China

* Correspondence: jqsun@hust.edu.cn

Abstract: We propose a simplified partitioned Brillouin gain spectrum (BGS) analysis method to enhance the spatial resolution and measurement accuracy of a Brillouin optical time-domain reflectometer (BOTDR) assisted by a first-order Raman pump. We theoretically derive the mathematical model of the partitioned BGS and analyze the superposition process of sub-Brillouin signals within a theoretical spatial resolution range. We unified all the unknown constant parameters of the calculation process to simplify the partitioned BGS analysis method and the value of the uniform parameter is attained through the system test data and numerical analysis. Moreover, to automate data processing, the starting point of the temperature/strain change is determined by the first occurrence of the maximum Brillouin frequency shift (BFS), then the position where the partitioned BGS analysis method calculation begins is obtained. Using a 100 ns probe pulse and partitioned BGS analysis method, we obtain a spatial resolution of 0.4 m in the 78.45-km-long Raman-assisted BOTDR system, and the measurement accuracy is significantly improved. In addition, we achieve a strain accuracy of 5.6 με and a spatial resolution of 0.4 m in the 28.5-km-long BOTDR without Raman amplification.

Keywords: BOTDR; spatial resolution; Brillouin gain spectrum analysis; Raman amplification; optical fiber

Citation: Huang, Q.; Sun, J.; Jiao, W.; Kai, L. Improvement of Performance for Raman Assisted BOTDR by Analyzing Brillouin Gain Spectrum. *Sensors* **2022**, *22*, 116. https://doi.org/10.3390/s22010116

Academic Editors: Phong B. Dao, Lei Qiu, Liang Yu and Zahra Sharif Khodaei

Received: 29 November 2021
Accepted: 23 December 2021
Published: 24 December 2021

1. Introduction

Distributed optical fiber sensing has attracted the attention of researchers in recent years due to its unique advantage in realizing multi-point temperature, strain, and other physical measurements. The Brillouin optical time-domain reflectometer (BOTDR) with probe pulse and signal processing equipment connected at the same end has favorable flexibility [1,2], and thus has been widely applied in power cables, oil and gas pipelines, large-scale structure health monitoring [3–6], etc. However, the performance of the distributed optical fiber sensing system in terms of spatial resolution (SR), measurement accuracy, measurement speed, and maximum sensing distance are mutually restricted. Increasing the width of the probe pulse can attain a better signal-to-noise ratio (SNR), sensing distance, and measurement accuracy of the system, but this method will reduce the SR [7].

To solve the tradeoff between these performance indicators, researchers have proposed various methods. Adopting appropriate pulse coding technology can increase the average power of the probe light and enhance the SR, which is a special practical method to improve the system performance [8–11]. The advanced data processing method can raise the SNR and SR without increasing the system cost [12–15]. The iterative subdivision method [14] has been used to improve the SR, where the author exploited the probe pulse with a width of 100 ns to achieve SR of 1.5 m in a 50 km BOTDR. However, this method needs to be combined with energy density distribution (EDD), and thus the

complexity of the algorithm is increased. A spatial resolution of 0.3 m was obtained in a 5 m BOTDR using 20 ns probe pulse by analyzing the Brillouin gain spectrum [15], but like the iterative subdivision, the position where the data processing method calculations begin is not given. Raman amplification is an effective approach to prolong sensing distance and realize quasi-transparent transmission [16,17]. The ultra-long sensing distance of 150 km can be accomplished by using coherent detection and Raman amplification [18], but the temperature resolution and SR of the system are 5.2 °C and 50 m, respectively. Meanwhile, by adjusting parameters such as Raman pump power and Erbium-doped optical fiber amplifier (EDFA) gain, the sensing distance of 100 km can be achieved [19], but the temperature accuracy and SR are ±3 °C and 40 m, respectively.

SR is one of the most important performance parameters of the BOTDR, indicating the minimum optical fiber length accurately measured by the system. The experimental spatial resolution is defined as the fiber length of the temperature transition region between 10% and 90% of the peak amplitude [20]. We propose a simplified partitioned analysis method through decomposition of BGS, combined with Raman-assisted amplification, which extends the sensing distance and improves the measurement accuracy and SR of the system. To automate data processing, the beginning calculation point of the proposed method is obtained by analyzing the superposition characteristics of the BGS. We experimentally demonstrate a temperature accuracy of 5.7 °C and a spatial resolution of 0.4 m in the 78.45-km-long Raman-assisted BOTDR system through the partitioned BGS analysis; the sub-meter level SR is thus achieved in the long-distance optical fiber distributed sensing system. Moreover, we verify the effectiveness of a simplified partitioned BGS in a 28.5-km-long BOTDR without the Raman pump.

2. Fundamentals of the System

2.1. First-Order Raman Assisted BOTDR Theory

We mainly analyze first-order co-directional Raman pumping which maintains the advantage of BOTDR single-ended access. When the incident power exceeds the stimulated Raman scattering (SRS) threshold, stimulated inelastic scattering will occur in some nonlinear medium, which will lead to the pump energy transfer to Raman scattering light. First-order Raman-assisted BOTDR utilizes this nonlinear effect to amplify the probe light and Stokes light (Brillouin scattering light) at ~1550 nm with a higher Raman pump at ~1455 nm, then enhances the SNR of the system and increases the measurement accuracy. In the case of the continuous pump, the coupling equation describing the first-order Raman process is written as follows [21]:

$$\frac{\mathrm{d}P_R}{\mathrm{d}z} = -\alpha_R P_R - g_R \frac{\nu_R}{\nu_s} P_R(P_\mathrm{s} + P_B) \tag{1}$$

$$\frac{dP_s}{dz} = -\alpha_s P_s + g_R P_R P_s \tag{2}$$

$$\frac{dP_B}{dz} = \alpha_s P_B - g_R P_R P_B \tag{3}$$

where P_R, P_s and P_B are the power of Raman pump, probe light and Stokes light, respectively. ν_R and ν_s are the corresponding frequencies of Raman pump and probe signal. α_R and α_s are the optical fiber attenuation coefficients corresponding to the wavelengths of the Raman pump and probe pulse, respectively. g_R is the gain coefficient of the Raman pump to the probe and Stokes.

2.2. Simplified Partitioned BGS Analysis Method

To boost the spatial resolution of the system, we divide the optical fiber into m segments within L, as shown in Figure 1; the length of each segment is

$$\Delta z = \frac{L}{m} \tag{4}$$

where L represents the theoretical spatial resolution of the BOTDR system, which is given by [14]

$$L = \frac{c(\tau + \tau')}{2n_{eff}} \tag{5}$$

where c is the light speed in vacuum; n_{eff} is the fiber-core effective refractive index; τ is the width of probe pulse; τ' is the response time of the BOTDR detection system, which is determined by the bandwidth of photoelectric devices, such as detectors, filters, and amplifiers [22]. Although the narrow pulse width can improve the spatial resolution of the BOTDR system, it will reduce the frequency resolution and measurement accuracy, which is limited by the phonon lifetime (~10 ns).

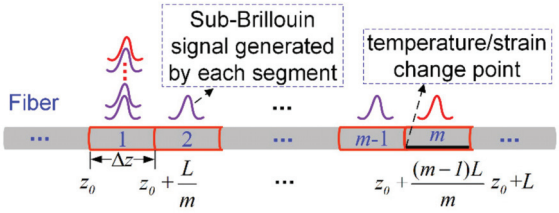


Figure 1. Superposition of sub-Brillouin signal, black section represents the temperature/strain change area. z_0 is an arbitrary position of optical fiber.

As shown in Figure 1, the probe pulse width is set to 100 ns in this experiment, which is greater than the phonon lifetime; the sub-Brillouin signals generated by each segment can thus be approximated by the Lorentzian shape [23]. Hence, the sub-Brillouin signal generated by the ith segment is

$$G_i(\nu, \nu_{Bi}) = \frac{g_i(\Delta\nu_{Bi}/2)^2}{(\nu - \nu_{Bi})^2 + (\Delta\nu_{Bi}/2)^2} \tag{6}$$

where ν is the frequency detuning round the BFS; g_i, ν_{Bi} and $\Delta\nu_{Bi}$ are the Brillouin peak gain coefficient, BFS, and the full width at half maximum (FWHM) in the ith segment, respectively.

The BGS measured by the system is the overlapping result of sub-Brillouin signals within L, thus the BGS at the point z_0 is obtained as

$$G(\nu, z_0) = \sum_{i=1}^{m} a_i G_i(\nu, \nu_{Bi}) = \sum_{i=1}^{m} a_i \frac{g_i(\Delta\nu_{Bi}/2)^2}{(\nu - \nu_{Bi})^2 + (\Delta\nu_{Bi}/2)^2} \tag{7}$$

where $G(\nu, z_0)$ is the BGS at point z_0 measured by the BOTDR system; a_i represents the constant of proportionality.

Assuming that the temperature/strain change occurs in the mth section, as shown in Figure 1, the sub-Brillouin signal of the mth can be obtained from (7), which is given by

$$G_m(\nu, \nu_{Bm}) = \frac{1}{a_m} G(\nu, z_0) - \frac{1}{a_m} \sum_{i=1}^{m-1} a_i \frac{g_i(\Delta\nu_{Bi}/2)^2}{(\nu - \nu_{Bi})^2 + (\Delta\nu_{Bi}/2)^2} \tag{8}$$

$G(\nu, z_0)$ is known, thus the BGS of the mth segment can be attained by solving the sub-Brillouin signal from 1st to $(m-1)$th segment. The sub-Brillouin signal generated by each segment from position z_0 to $z_0 + (m-1)L/m$ is approximately equal to that from $z_0 - (m+1)L/m$ to $z_0 - 2L/m$ because the temperature/strain of these sections have not changed, as shown in Figure 2.

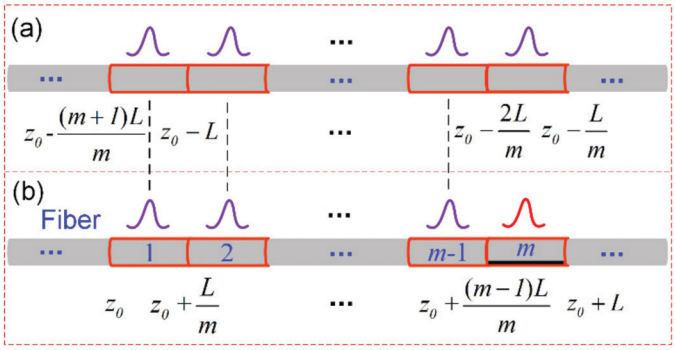


Figure 2. Distribution of sub-Brillouin signal. (**a**) Sub-Brillouin signal within L length before z_0 point. (**b**) Sub-Brillouin signal within L length after z_0 point.

We analyze the sub-Brillouin signal generated within an L length before the z_0 point; the sub-Brillouin signal from $z_0 - (m+1)L/m$ to $z_0 - 2L/m$ and z_0 to $z_0 + (m-1)L/m$ have the same ν_{Bi} and $\Delta\nu_{Bi}$, which satisfies (9).

$$G(\nu, z_0 - \frac{(m+1)L}{m}) = \sum_{n=1}^{m} a_n G_n(\nu, \nu_{Bn}) = \sum_{n=1}^{m} a_n \frac{g_n(\Delta\nu_{Bn}/2)^2}{(\nu - \nu_{Bn})^2 + (\Delta\nu_{Bn}/2)^2} \tag{9}$$

where a_n represents the constant of proportionality.

Using the scan data detected by the BOTDR system from $z_0 - (m+1)L/m$ to $z_0 - L/m$ point and performing Lorentzian curve fitting, the BFS and FWHM of the BGS at each point can be obtained. Thus, the sub-Brillouin gain spectrum of the mth segment can be calculated if we can attain the gain peak coefficient g_n. Assuming that g_n is equal to 1, the two intermediate variables A and B can then be expressed as

$$A = \sum_{n=1}^{m} \frac{(\Delta\nu_{Bn}/2)^2}{(\nu - \nu_{Bn})^2 + (\Delta\nu_{Bn}/2)^2} \tag{10}$$

$$B = G(\nu, z_0 - \frac{(m+1)L}{m})/A \tag{11}$$

From (8), (10), and (11), we can get:

$$G_m(\nu, \nu_{Bm}) = \frac{1}{a_m} G(\nu, z_0) - \frac{FB}{a_m} \sum_{i=1}^{m-1} a_i \frac{(\Delta\nu_{Bi}/2)^2}{(\nu - \nu_{Bi})^2 + (\Delta\nu_{Bi}/2)^2} \tag{12}$$

where F is a constant. We unify all the constants to k without affecting the final fitting results of BFS to simplify the calculation, then (12) becomes

$$G_m(\nu, \nu_{Bm}) = G(\nu, z_0) - kB \times \sum_{i=1}^{m-1} \frac{(\Delta\nu_{Bi}/2)^2}{(\nu - \nu_{Bi})^2 + (\Delta\nu_{Bi}/2)^2} \tag{13}$$

As shown in Figure 3, the BGS generated in the $(m+1)$th segment can be calculated by (14). By analogy, we can get the sub-Brillouin signal of the following multiple segments.

$$G_{m+1}(\nu, \nu_{B(m+1)}) = G(\nu, z_0 + \frac{L}{m}) - kB \times \sum_{i=2}^{m-1} \frac{(\Delta\nu_{Bi}/2)^2}{(\nu - \nu_{Bi})^2 + (\Delta\nu_{Bi}/2)^2} - G_m(\nu, \nu_{Bm}) \tag{14}$$

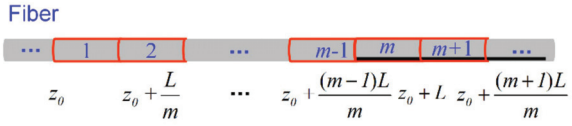


Figure 3. Multi-segment temperature/strain change area. Black indicates the unknown sub-Brillouin signal.

Now, the key problem is to find z_0 in Equation (13), and then the signal processing algorithm can be automated. Assuming that there are three continuous sections where temperature/strain changes, Figure 4 depicts the process of sub-Brillouin signal generation and superposition. In Figure 4b, the sub-Brillouin signal generated in the $(m-1)$th and mth segments after the probe pulse propagates forward for a distance of L/m. Since the probe pulse and the Brillouin scattering light propagation are opposite the optical fiber normally, the sub-Brillouin signal generated in the mth segment in the previous stage overlaps with the generated in the $(m-1)$th segment this time, which is shown in Figure 4c. The sub-Brillouin signal completely generated by the temperature/strain sections is superimposed on the mth segment (point $z_0+(m-1)L/m$) for the first time with the propagation of the light, that is, the BFS appeared as the maximum value. According to the distance between the first BFS maximum point fitted by the system measurement data and the actual temperature/strain starting point L, then the following relationship can be obtained:

$$z_{TSC} = z_0 + \frac{(m-1)L}{m} = z_{MBFS} - L \tag{15}$$

where z_{TSC} is actual the temperature/strain starting point. z_{MBFS} is the first maximum value point of BFS fitted by the system measurement data. The start change point of temperature/strain and z_0 can thus be obtained using Equation (15).

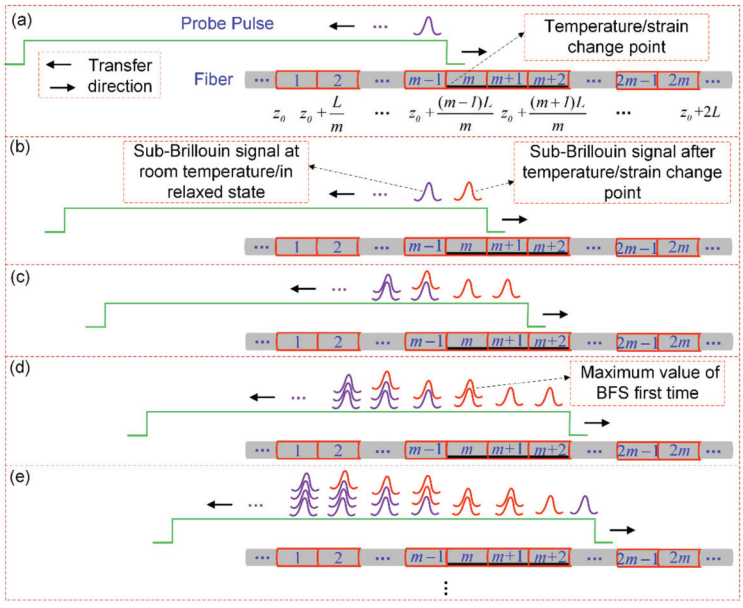


Figure 4. Brillouin gain spectrum superposition process. The black indicates the temperature/strain change area. (**a**) The head of the probe pulse generates a sub-Brillouin signal in the $(m-1)$th section. (**b**–**e**) is the generation and superposition process of the sub-Brillouin signal of each segment for each L/m long distance of the probe pulse propagation, respectively.

Combining (15), (13), (14) and its deformations, we can solve the sub-Brillouin signals generated in each segment. Only addition and subtraction are involved in the process of these calculations, thus the partitioned BGS analysis method has no significant effect on the measurement time of the system compared with the process of the data fitting and frequency scanning.

3. Experimental Setup and Results

To verify the correctness of the simplified partitioned BGS analysis method, we design the experimental test scheme as shown in Figure 5. The total length of the optical fiber under test (FUT) is 78.45 km, which is composed of two spools with different parameters, with its 36 m section heated and the remaining section kept under room temperature and slack. A distributed feedback (DFB) laser emits continuous-wave light with a wavelength of 1550.074 nm, which is split into the probe and reference light by the 3 dB optical coupler. The probe light passes through the polarization controller (PC), electric-optic modulator (EOM1 bandwidth: 2.5 GHz, extinction ratio $\geq$25 dB), Erbium-doped fiber amplifier (EDFA1), filter, circulator, and wavelength division multiplexer (WDM), and then enters the FUT. After transiting the EOM2 (bandwidth: 40 GHz, extinction ratio $\geq$30 dB), EDFA2, and the polarization scrambler (PS) with the 700 MHz scrambling frequency, the reference light beats with backward Brillouin scattering light in the balanced photodetector (BPD) with the 800 MHz bandwidth. A Raman fiber laser (RFL) launches a ~1455 nm Raman pump through a tunable attenuator (TA) and isolator to enter the WDM, then amplifies the probe and scattered light simultaneously. The sampling rate of the data acquisition and processor (DAP) is f_s = 250 MSa/s, and the average times is 1000. The output power of the DFB laser is set to 5 dBm, and the pump power of the RFL after passing the tunable attenuator is 26.43 dBm. Meanwhile, the arbitrary waveform generator (AWG) generates a square wave with a frequency of 1 kHz and a pulse width of 100 ns to modulate the probe light. The ambient temperature is 23 °C, corresponding to the BFS of 11.245 GHz. The temperature of the oven is heated to 80 °C, which corresponds to the BFS of 11.302 GHz according to the temperature correlation coefficient of 1 MHz/°C; in this experiment, and other system parameters such as EDFA pump current are set to the optimal value. The microwave source (MS) frequency scan range is 11.15–11.36 GHz with 10 MHz intervals, and the normalized Brillouin scattering Power–BFS–Distance three-dimensional map is shown in Figure 6 after a five-layer wavelet transforms denoising. In Figure 6, the abrupt drop in power at 50 km is caused by the inconsistent parameters of the two optical fiber spools.

Direct Lorentzian curve fitting is performed for the measured Brillouin power and the BFS distribution curve along the optical fiber is then obtained, as shown in Figure 7. The fluctuations of the BFS profile are mainly caused by the coiling strain in the fiber, inducing BFS oscillations along the entire sensing range.

It can be seen from the illustration in Figure 7 that the distance from the temperature starting change point (78.3508 km) to the first maximum value of BFS point (78.362 km) equates to the L (11.2 m), which is consistent with the result calculated by (15). Moreover, the measured BGS at the heated front end is generated by the superposition of the sub-Brillouin signals from both the unheated and the heated segments, resulting in rather low BFS fitting accuracy.

Since the sampling interval of DAP is $s = c/2f_s n_{eff}$ = 0.4 m, we set $\Delta z = s$ to maintain the continuity of calculation. The optical fiber can then be divided into $m = L/\Delta z = 28$ segments, thus (13) can be transformed into:

$$G_{28}(\nu, \nu_{B28}) = G(\nu, z_0) - kB \times \sum_{i=1}^{27} \frac{(\Delta\nu_{Bi}/2)^2}{(\nu - \nu_{Bi})^2 + (\Delta\nu_{Bi}/2)^2} \tag{16}$$

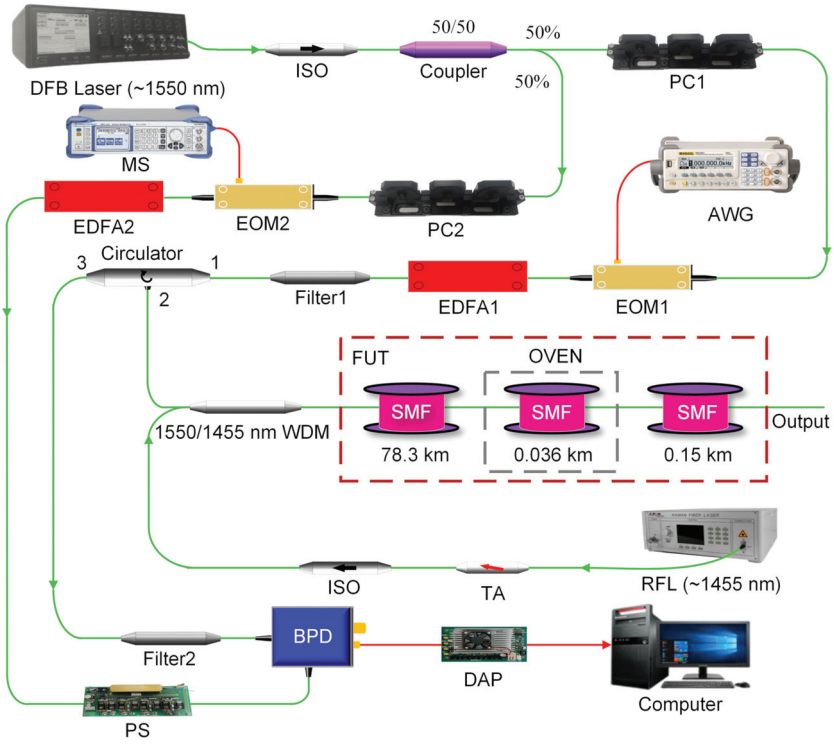


Figure 5. Schematic depiction of Raman-assisted BOTDR system. EOM: electro-optical modulator; WDM: wavelength division multiplexer; TA: tunable attenuator; DFB: distributed feedback; ISO: isolator; PC: polarization controller; BPD: balanced photodetector; EDFA: erbium-doped fiber amplifier; FUT: fiber under test; SMF: single-mode fiber; RFL: Raman fiber laser; PS: polarization scrambler; DAP: data acquisition and processing; AWG: arbitrary waveform generator; MS: microwave source.

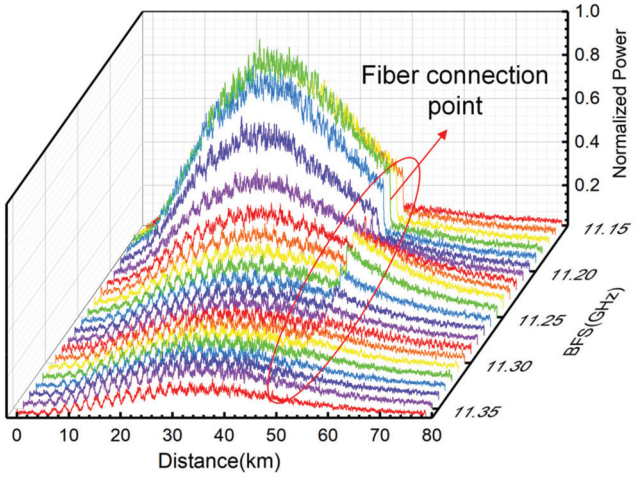


Figure 6. Normalized Brillouin scattering Power–BFS–Distance three-dimensional diagram.

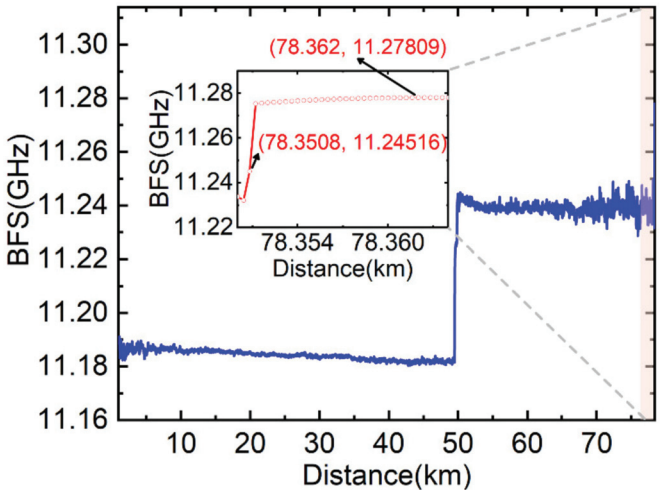


Figure 7. Directly fitting BFS distribution curve.

The constant $k = 0.17$ is attained by numerical analysis using known short-distance BOTDR system experimental data. The partitioned BGS analysis method is exploited to calculate the sub-Brillouin signal within the L length range (from 78.3508 km to 78.362 km) at the heated front end, and then the BFS distribution in this range is obtained by Lorentzian fitting. Additionally, the BFS distribution beyond this range is obtained by direct fitting to avoid the accumulation of errors and to reduce the calculation time. Finally, the total BFS distribution curve is shown in Figure 8. It can be seen from the illustration in Figure 8 that the spatial resolution and measurement accuracy of the system are improved.

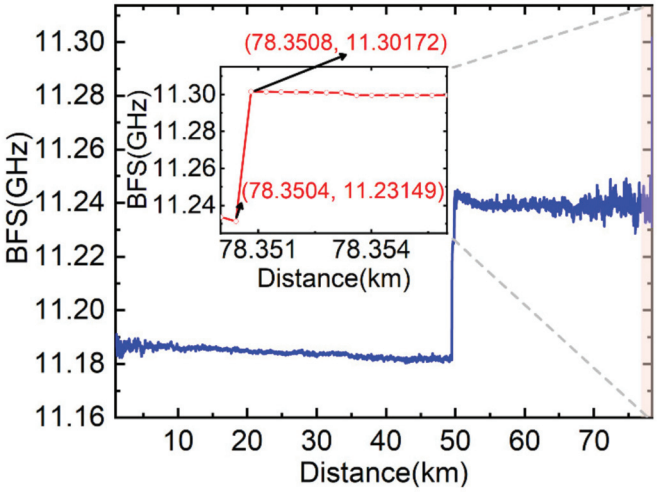


Figure 8. Lorentzian fitting BFS distribution curve after partitioned BGS analysis.

Table 1 shows the comparison between the direct fitting and the fitting result after simplified partitioned spectrum analysis within the L length range at the heated front end. As can be seen from Table 1, the temperature accuracy of the direct fitting is 24.2 °C in this experiment, obtained from the difference between the measured mean temperature from 78.3508 km to 78.362 km and heating temperature 80 °C, which is reduced to 5.7 °C after partitioned BGS analysis, and the spatial resolution is improved to 0.4 m.

Table 1. Comparison of test results.

Results	Direct Lorentzian Curve Fitting	Lorentzian Fitting after Partitioned BGS Analysis
Mean BFS amplitude (GHz)	11.2778	11.2963
Corresponding temperature (°C)	55.8	74.3
Accuracy (°C)	24.2	5.7
Spatial resolution (m)	11.2	0.4

To validate the robustness of the simplified partitioned BGS analysis method in the BOTDR system without the Raman pump and strain sensing, we set up an experimental scheme as shown in Figure 9. A total of 28.5 km length SMF is connected in the BOTDR, and the 3 m section is coiled on two micro-positioners, with one fixed and the other movable for applying uniform strain.

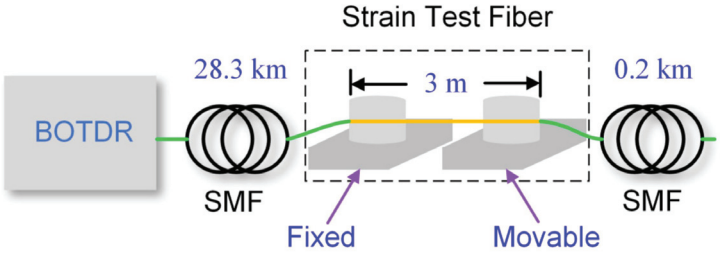


Figure 9. Test scheme of the BOTDR system without Raman pump.

The EDFA pump current and the EOM bias voltage are adjusted to the optimal state, while other system parameters remain unchanged. The axial strain of 1400 με is applied to the 3 m length fiber by moving the precision displacement platform. Then, the BFS distribution curve is obtained by fitting the experimental data shown in Figure 10. Since the measured BGS is a superposed spectrum of the strain and the unstrained section within the spatial resolution and the strained optical fiber length is much less than L, the shape of BGS deforms that of the Lorentzian shape. It is obvious in Figure 10 that the results of direct Lorentzian fitting have a significant error and do not reflect the true axial strain in the strain test region. After processing by the partitioned BGS analysis method and then fitting, the mean BFS amplitude of the strain test region is 11.3098 GHz, which is equivalent to the strain of 1405.6 με according to the strain correlation coefficient of the BFS is 4.61 MHz/100με in this experiment. The strain measurement accuracy is 5.6 με, and the rising edge of the BFS distribution curve is 0.4 m, that is, the spatial resolution is 0.4 m. In addition, it can be seen from the blue curve that the applied strain area is 2.8 m, which is close to the actual value of 3 m.

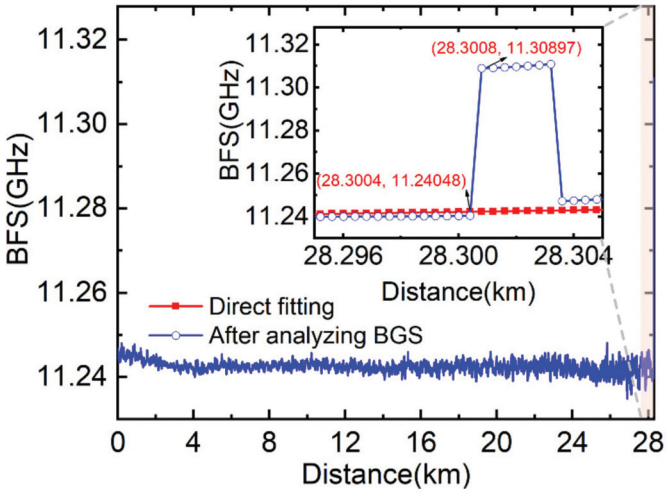


Figure 10. BFS distribution curve of the BOTDR system without Raman pump.

4. Conclusions

In summary, we propose an analysis method based on simplified partitioned Brillouin gain spectrum to improve the spatial resolution and the measurement accuracy of the BOTDR system. According to the superposition characteristics of the sub-Brillouin signals, the position where the partitioned BGS analysis method begins to calculate is determined to automate data processing. Moreover, the method of the partitioned BGS analysis is simplified by unifying all the constants to the parameter k, which is obtained by the short-distance BOTDR system experiment data and numerical calculation. We set up the first-order co-directional assisted Raman amplification BOTDR system with a 78.45 km length optical fiber to demonstrate the correctness of the proposed method. We utilize the simplified partitioned BGS analysis method to deal with the experimentally obtained Brillouin scattering power in the heated area. The experimental results show that we have achieved a temperature accuracy of 5.7 °C and a spatial resolution of 0.4 m, that is, the sub-meter spatial resolution of the long-distance distributed fiber sensing is realized. Meanwhile, we also designed the strain test scheme of a 28.5-km-long BOTDR system without Raman amplification, in which a 1400 με is applied to strain test fiber with 3 m length. We obtain the spatial resolution of 0.4 m and the strain accuracy of 5.6 με at the far-end fiber by exploiting the proposed method, which is not limited to BOTDR but can also be applied to other fiber optical distribution sensing systems based on Brillouin scattering. Hence, distributed optical fiber temperature/strain sensing based on the method of the partitioned BGS analysis, which has obvious practical application significance, can realize more effective health monitoring for large-scale structures such as pipelines, bridges and power lines.

Author Contributions: Conceptualization, Q.H.; methodology, Q.H., J.S.; software, Q.H.; formal analysis, W.J. and L.K.; writing—original draft preparation, Q.H., W.J. and L.K.; writing—review and editing, Q.H., W.J. and L.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 61875063 and Wuhan Project of Applied Basic Frontier, grant number 2018010401011272.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the reviewers for their time and efforts.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Horiguchi, T.; Shimizu, K.; Kurashima, T.; Tateda, M.; Koyamada, Y. Development of a Distributed Sensing Technique Using Brillouin Scattering. *J. Lightwave Technol.* **1995**, *13*, 1296–1302. [CrossRef]
2. Bai, Q.; Wang, Q.; Wang, D.; Wang, Y.; Gao, Y.; Zhang, H.; Zhang, M.; Jin, B. Recent Advances in Brillouin Optical Time Domain Reflectometry. *Sensors* **2019**, *19*, 1862. [CrossRef]
3. Zhao, L.; Li, Y.; Xu, Z.; Yang, Z.; Lü, A. On-line monitoring system of 110 kV submarine cable based on BOTDR. *Sens. Actuators A Phys.* **2014**, *216*, 28–35. [CrossRef]
4. Feng, X.; Wu, W.J.; Li, X.Y.; Zhang, X.W.; Zhou, J. Experimental investigations on detecting lateral buckling for subsea pipelines with distributed fiber optic sensors. *Smart Struct. Syst.* **2015**, *15*, 245–258. [CrossRef]
5. Seo, H. Monitoring of CFA pile test using three dimensional laser scanning and distributed fiber optic sensors. *Opt. Lasers Eng.* **2020**, *130*, 106089. [CrossRef]
6. Yuan, Q.; Chai, J.; Ren, Y.W.; Liu, Y.L. The Characterization Pattern of Overburden Deformation with Distributed Optical Fiber Sensing: An Analogue Model Test and Extensional Analysis. *Sensors* **2020**, *20*, 7215. [CrossRef] [PubMed]
7. Soto, M.A.; Thévenaz, L. Modeling and evaluating the performance of Brillouin distributed optical fiber sensors. *Opt. Express* **2013**, *21*, 31347–31366. [CrossRef]
8. Soto, M.A.; Bolognini, G.; Di Pasquale, F. Analysis of optical pulse coding in spontaneous Brillouin-based distributed temperature sensors. *Opt. Express* **2008**, *16*, 19097–19111. [CrossRef]
9. Yang, Z.S.; Soto, M.A.; Thevenaz, L. Increasing robustness of bipolar pulse coding in Brillouin distributed fiber sensors. *Opt. Express* **2016**, *24*, 586–597. [CrossRef]
10. Horiguchi, T.; Masui, Y.; Zan, M.S.D. Analysis of Phase-Shift Pulse Brillouin Optical Time-Domain Reflectometry. *Sensors* **2019**, *19*, 1497. [CrossRef]
11. Hao, Y.Q.; Ye, Q.; Pan, Z.Q.; Cai, H.W.; Qu, R.H.; Yang, Z.M. Effects of modulated pulse format on spontaneous Brillouin scattering spectrum and BOTDR sensing system. *Opt. Laser Technol.* **2013**, *46*, 37–41. [CrossRef]
12. Bashan, G.; Diamandi, H.H.; London, Y.; Preter, E.; Zadok, A. Optomechanical time-domain reflectometry. *Nat. Commun.* **2018**, *9*, 2991. [CrossRef] [PubMed]
13. He, Q.; Jiang, H.; Wang, Z.Q.; Ye, S.B.; Shang, X.J.; Li, T.; Tang, L.J. Spatial resolution enhancement of DFT-BOTDR with high-order self-convolution window. *Opt. Fiber Technol.* **2020**, *57*, 102188. [CrossRef]
14. Wang, F.; Zhan, W.W.; Zhang, X.P.; Lu, Y.G. Improvement of Spatial Resolution for BOTDR by Iterative Subdivision Method. *J. Lightwave Technol.* **2013**, *31*, 3663–3667. [CrossRef]
15. Murayama, H.; Kageyama, K.; Shimada, A.; Nishiyama, A. Improvement of spatial resolution for strain measurements by analyzing Brillouin gain spectrum. In Proceedings of the 17th International Conference on Optical Fibre Sensors, Bruges, Belgium, 23–27 May 2005; Volume 5855, pp. 551–554.
16. Soto, M.A.; Bolognini, G.; Di Pasquale, F. Optimization of long-range BOTDA sensors with high resolution using first-order bi-directional Raman amplification. *Opt. Express* **2011**, *19*, 4444–4457. [CrossRef] [PubMed]
17. Martin-Lopez, S.; Alcon-Camas, M.; Rodriguez, F.; Corredera, P.; Ania-Castanon, J.D.; Thevenaz, L.; Gonzalez-Herraez, M. Brillouin optical time-domain analysis assisted by second-order Raman amplification. *Opt. Express* **2010**, *18*, 18769–18778. [CrossRef]
18. Alahbabi, M.N.; Cho, Y.T.; Newson, T.P. 150-km-range distributed temperature sensor based on coherent detection of spontaneous Brillouin backscatter and in-line Raman amplification. *J. Opt. Soc. Am. B—Opt. Phys.* **2005**, *22*, 1321–1324. [CrossRef]
19. Song, M.P.; Xia, Q.L.; Feng, K.B.; Lu, Y.; Yin, C. 100 km Brillouin optical time-domain reflectometer based on unidirectionally pumped Raman amplification. *Opt. Quantum Electron.* **2016**, *48*, 30. [CrossRef]
20. Wu, H.; Wan, Y.Y.; Tang, M.; Chen, Y.J.; Zhao, C.; Liao, R.L.; Chang, Y.Q.; Fu, S.N.; Shum, P.P.; Liu, D.M. Real-Time Denoising of Brillouin Optical Time Domain Analyzer With High Data Fidelity Using Convolutional Neural Networks. *J. Lightwave Technol.* **2019**, *37*, 2648–2653. [CrossRef]
21. Nuno, J.; Martins, H.F.; Martin-Lopez, S.; Ania-Castanon, J.D.; Gonzalez-Herraez, M. Distributed Sensors Assisted by Modulated First-Order Raman Amplification. *J. Lightwave Technol.* **2021**, *39*, 328–335. [CrossRef]
22. Maughan, S.M.; Kee, H.H.; Newson, T.P. Simultaneous distributed fibre temperature and strain sensor using microwave coherent detection of spontaneous Brillouin backscatter. *Meas. Sci. Technol.* **2001**, *12*, 834–842. [CrossRef]
23. Alem, M.; Soto, M.A.; Tur, M.; Thevenaz, L. Analytical expression and experimental validation of the Brillouin gain spectral broadening at any sensing spatial resolution. In Proceedings of the 2017 25th International Conference on Optical Fiber Sensors (OFS), Jeju, Korea, 24–28 April 2017.

sensors

MDPI

Review

A Review of Infrared Thermography for Delamination Detection on Infrastructures and Buildings

Ko Tomita * and Michael Yit Lin Chew

Department of the Built Environment, National University of Singapore, Singapore 117566, Singapore; bdgchewm@nus.edu.sg
* Correspondence: e0441600@u.nus.edu

Abstract: This paper provides a comprehensive review on the use of infrared thermography to detect delamination on infrastructures and buildings. Approximately 200 pieces of relevant literature were evaluated, and their findings were summarized. The factors affecting the accuracy and detectability of infrared thermography were consolidated and discussed. Necessary measures to effectively capture latent defects at the early stage of delamination before crack formation were investigated. The results of this study could be used as the benchmarks for setting standardized testing criteria as well as for comparison of results for future works on the use of infrared thermography for detection of delamination on infrastructures and buildings.

Keywords: infrared thermography; delamination; building; infrastructure; time window; environment; infrared camera; target object; thermal property

Citation: Tomita, K.; Chew, M.Y.L. A Review of Infrared Thermography for Delamination Detection on Infrastructures and Buildings. *Sensors* **2022**, *22*, 423. https://doi.org/10.3390/s22020423

Academic Editors: Phong B. Dao, Lei Qiu, Liang Yu and Zahra Sharif Khodaei

Received: 3 December 2021
Accepted: 4 January 2022
Published: 6 January 2022

1. Introduction

With the aging of civil infrastructures and buildings, those deterioration has become an important social issue that can threaten public safety. The American Road & Transportation Builders Association reported in 2020 that 36% of bridges in the US need replacement or rehabilitation due to their aging [1]. Similarly, in Singapore, the age of 74% of high-rise residences exceeds 20 years old, and more than 90 incidents of falling parts of facades from high places occurred in recent three years [2]. To ensure public safety, governments introduced mandatory periodic inspection schemes of infrastructures and buildings. For civil infrastructures, long highway bridges in the US are required to be inspected every 24 months [3]. For buildings, Singapore [2], Japan [4], Hong Kong [5], and 13 cities in the US and Canada [6] enacted periodic inspection laws to prevent falling objects from building facades.

Defects in infrastructures can be diverse and include delamination, cracks, staining, and spalling, caused mainly by water penetration, reinforcement corrosion, thermal/moisture movements, differential settlement/loading, poor construction practices, etc. [7,8]. Among them, delamination, the condition in which the surface and inside are unbonded or unintegrated properly, are crucial because they lead to further deterioration, such as crack formation and element falling [9]. In concrete infrastructures, delamination arises in concrete cover near the surface because of the expansion of corroded embedded rebars as well as cyclical traffic load stress and environmental changes [10]. In building facades, delamination generally occurs at the interface between a finish layer, such as tiles or render, and a substrate, such as concrete or bricks [8,11–13]. Delamination constitutes a significant part of defects occurring on tile façades, accounting for 27% of facade defects in Singapore [8] and 71% in Brazil [14]. Since delamination arises under the surface, it is to be detected via nondestructive testings (NDTs).

In recent decades, various NDTs were developed to detect defects in multiple fields since they can evaluate object characteristics [15–17]. Each NDT has different principles and features, so that it is necessary to select appropriate NDTs according to the purpose

and conditions of inspection [18]. Several NDTs can identify delamination, e.g., tapping tests, chain drag tests, hammer sounding tests, ground-penetrating radar, and infrared thermography (IRT) [19,20]. Among them, IRT especially drew increasing attention due to its advantages of real-time [21], contactless [22], and wide-area measurements [23]. Another advantage is that the price of an infrared (IR) camera has recently become affordable [21]. Therefore, IRT can serve as a suitable NDT for civil infrastructures and buildings.

IRT is defined as a process of measuring surface temperature distribution using IR cameras and processing and interpreting the data of IR images [24]. For infrastructures and buildings, IRT is used not only for delamination detection but also a wide range of inspections: moisture [25–31], thermal insulation [7,32–35], internal structure [36], cracks [37–39], air leakage [40,41], and cultural heritage [42–46]. In terms of delamination on infrastructures and buildings, IRT generally employs a passive analysis scheme, which uses surrounding environments as heat sources to stimulate temperature distribution [7,47,48]. However, passive IRT has some limitations at the step of data acquisition. The most critical limitation is that the detectability of passive IRT depends on uncontrollable environmental conditions, such as solar irradiation, ambient temperature, and wind [49–52]. Even in the same infrastructure, microclimates around surfaces differ depending on surface directions [53]. Other factors that may affect the detectability include delamination properties, target object [54], and IR camera [52]. If IR images are measured without due consideration of these conditions, delamination may be overlooked or misinterpreted. Understanding the mentioned conditions is hence crucial for planning and conducting passive IRT. Thus, many studies were conducted on the effects of environmental conditions, delamination properties, target objects, and IR cameras [55]. However, inconsistent results were often observed because of different conditions of experiments [56].

This paper focuses on the use of IRT to detect delamination on infrastructures and buildings to prevent falling objects from heights that endanger public safety. It provides a comprehensive review of the use of IRT by providing backgrounds, principles, and state-of-the-art knowledge on affecting factors and desirable conditions. This paper will contribute to increasing the reliability of IRT and facilitating further research.

Section 2 of this paper presents related review papers of IRT inspection on infrastructures and buildings. Section 3 explains the theory of temperature measurement and classifications of IRT. In Section 4, the principle and analysis methods of IRT for delamination detection and existing standards and guidelines are described. Additionally, the performance of IRT in detecting delamination is compared with that of other NDTs. Section 5 compiles and discusses some of the latest case studies on the impact of the various factors and investigates the different methodologies adopted. Section 6 compares and synthesizes relevant literature on factors affecting detectability. Lastly, Section 7 states conclusions.

2. Related Review Works on IRT

This section investigates review papers on IRT within the last decade. Recent review papers on IRT were conducted from perspectives of applications, methodologies, and research trends.

The first perspective is IRT applications, which are commonly used in reviews. Application reviews range from the level of introducing case studies in industrial fields to the level of in-depth investigation of a specific application. IRT applications were developed in many fields, including medical [57], aerospace [58], plant [59], electronic component [52,60], gas [61], machine [57,62], metal corrosion [63], photovoltaic panels [64], composite materials [65–67], and cultural heritage [68,69]. Similarly, various IRT applications were proposed for infrastructure and building inspection. Garrido et al. [22] introduced past studies in terms of inspected subjects: buildings, civil infrastructures, and heritage sites. Among these types, applications for civil infrastructures and buildings are the main subject of review papers.

Several review papers focused on the energy audit of building envelopes using IRT to evaluate building energy performances [7,33,48,50,70]. Lucchi [70] reviewed detailed

applications of energy audit: detection of thermal bridge, insulation defects, air leakage, and moisture; indoor temperature and U-value measurements; and human comfort assessment. Among those applications, Nardi et al. [71] focused on quantitative IRT for the U-value measurement of building walls, representing building energy efficiency. Similarly, Bienvenido–Huertas et al. [72] focused on IRT as one of the in-site methods for assessing U-value. Apart from energy audits, another important application of IRT in civil infrastructures and buildings is the detection of delamination, void, and high moisture content. Lourenço et al. [9] investigated state-of-the-art techniques of IRT to detect delamination and moisture beneath ceramic claddings facades to reveal efficient quantitative and qualitative survey methods. Meanwhile, Sirca Jr. & Adeli [73] focused on experiment conditions and examined previous studies on IRT for concrete defect detection in laboratory tests and in field surveys. These reviews indicate the effectiveness of IRT to detect delamination on infrastructures and buildings.

The second perspective is the methodologies of conducting IRT. A wide variety of methodologies was developed and is classified based on their features and principles [63]. One classification of IRT is an analysis scheme including passive IRT, active pulsed IRT, and active lock-in IR [21]. Among them, Milovanović & Pečur [51] focused on active IRT for concrete infrastructures and described physical backgrounds, equipment, and postprocessing methods. Furthermore, Garrido and coauthors reviewed IRT methodologies for infrastructures during data acquisition [55] and postprocessing [74], respectively. During the data acquisition step, the authors introduced the IRT approaches for data collection and compared the latest studies regarding experimental setups, target materials, IRT modes, and analysis schemes in each defect type and application [55]. At the postprocessing step, the authors introduced the theories and representative studies on analysis algorithms and discussed those advantages and disadvantages. These reviews provide overviews of traditional and latest IRT methodologies.

The final perspective is the analysis of research trends based on statistical data of the number of past studies. For example, Fox et al. [75] analyzed research trends about IRT of energy-related building defects detection and discussed the correlation between the types of methodologies. Similarly, Kylili et al. [21] statistically analyzed research trends of IRT in building facades regarding measurement methods, analysis schemes, and analysis types. These statistical reviews objectively indicate the increase of literature on IRT for infrastructures and buildings.

As mentioned above, previous reviews about IRT were conducted from various perspectives. However, no review focuses on the characteristics of detectable delamination and measurement conditions affecting reliabilities and detectability of IRT for infrastructures and buildings.

3. Infrared Thermography

3.1. Theory of Temperature Measurement

This section explains the principle of temperature measurement by IR cameras. Heat energy can be transferred in three ways: conduction, convection, and radiation. Temperature measurement by IR cameras utilizes radiation transfer. All objects with an absolute temperature greater than 0 K emit electromagnetic waves, mainly in infrared spectra. According to Stefan–Boltzmann's law, radiant energy from a black body is as follows:

$$W_b = \sigma {T_{obj}}^4 \ (\mathrm{W/m^2}), \tag{1}$$

where W_b is the total radiant flux emitted per unit area ($\mathrm{W/m^2}$), σ is the Stephan–Boltzmann constant, and T_{obj} is the absolute temperature of the object (K). The black body is defined as an ideal object that absorbs all the radiation that collides with it at any wavelength. However, an actual object, called a grey body, is not a black body because objects have some reflection and transmission. The radiant energy emitted by a grey body (W_g) is as follows:

$$W_g = \varepsilon W_b = \varepsilon \sigma {T_{obj}}^4 \ (\mathrm{W/m^2}), \tag{2}$$

where ε is emissivity, defined as the ratio of the radiant energy emitted from the object to the energy emitted from the black body at the same temperature.

IR cameras capture radiant energy in a specific infrared spectrum region emitted from an object and convert the energy into a temperature value. Figure 1 shows the elements of thermal radiation captured by an IR camera when measuring the surface temperature of an opaque object. Infrared radiation received by an IR sensor consists of three sources: emission from the target object (W_{obj}), assuming the object as the black body, emission from surroundings reflected on the object (W_{refl}), and emission from the atmosphere (W_{atm}). The following formula expresses the total heat radiation detected by the IR camera (W_{total}):

$$W_{total} = \varepsilon\tau W_{obj} + (1 - \varepsilon)\tau W_{refl} + (1 - \tau)W_{atm}, \quad (3)$$

where τ is the transmittance of the atmosphere. The reflected radiation assumes that reflection temperature T_{refl} is the same for all reflections from surroundings, and the emissivity of surrounding surfaces ε_{refl} is assumed to be one. Radiation from the object and reflected radiation are absorbed by the atmosphere during traveling. Atmospheric radiation is emission from the atmosphere between the object and the camera at ambient temperature T_{atm}. "$1 - \tau$" indicates the atmosphere's emissivity, and τ depends on T_{atm}, relative humidity, and measurement distance between the target object and the IR camera. Therefore, accurate surface temperature measurements need to be compensated for the effects of emissivity ε, ambient temperature T_{atm}, relative humidity, reflection temperature T_{refl}, and distance [76].

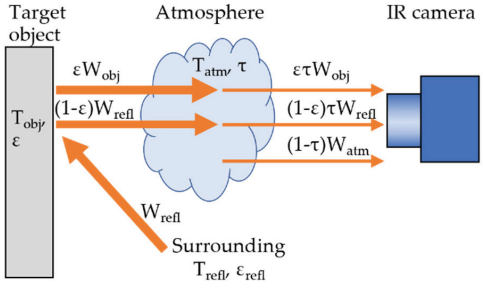


Figure 1. Diagram of temperature measurement by infrared (IR) camera.

In particular, the emissivity of target objects has a significant influence on temperature measurement [77]. The emissivity is a value from 0 to 1 and varies depending on materials, surface texture, angle, wavelength, and surface temperature [78]. Materials generally used in infrastructure, such as concrete, plaster, and general paint, have a high emissivity of 0.70–0.95 [78,79]. Thus, qualitative evaluation of defects can use emissivity values listed in emissivity libraries, whereas accurate temperature measurements for quantitative evaluation require the measurement of emissivity of target materials [50].

3.2. Classification of IRT

There are various methodologies of conducting IRT. They are classified from several viewpoints: analysis scheme, mode, and measurement method [50]. When assessing defects or thermal performances on infrastructures and buildings, it is crucial to select appropriate methodologies based on the purpose of the survey and conditions.

3.2.1. Analysis Scheme

An analysis scheme is a classification criterion based on the origin of the energy input to a target object to generate temperature distribution on the surface. There are passive and active IRT.

Passive IRT utilizes natural heat sources as stimuli to generate a thermal gradient inside an object, causing thermal contrast on the object's surface between sound and defect area [50]. The primary heat sources are generally solar irradiation and ambient temperature [22]. The natural heat sources heat large areas uniformly, so that passive IRT can inspect an extensive infrastructure at one time. Additionally, passive IRT does not require artificial heat sources, resulting in low cost. However, the detectability and accuracy of passive IRT significantly rely on various factors such as weather, surface orientation, and sunlight direction [50], so that the detectability of passive IRT may be limited. Additionally, passive IRT is not suitable for quantitative evaluation because the natural heat sources cannot be controlled. Therefore, passive IRT is mainly applicable for identifying defect locations before conducting advanced NDTs and is often used to inspect civil infrastructures and building facades [50,51,75].

Active IRT uses artificial heat sources to heat a target object to generate a thermal gradient [51]. Active IRT typically captures clear visualization of thermal anomalies compared to passive IRT [80] and can survey under conditions difficult for passive IRT. Traditional artificial heat sources are heat guns and hot water jets and bags [55]. Advanced thermal excitations include thermal induction, laser, ultrasonic, and microwave [22,81]. The mainstream for infrastructure inspection is optical excitation, such as halogen lamps and xenon lamps. Furthermore, controlled energy input allows quantitative assessment of defects, for example, defect depth [21]. Thus, active IRT is suitable for investigating specific areas in detail, such as heritage sites [75]. However, artificial heat sources are difficult to heat large areas uniformly, so that active IRT is not suitable for surveying large areas, such as buildings and infrastructures.

Active IRT is further classified according to heating processes: pulsed IRT (PT), step heating thermography (SH), and lock-in IRT (LT) [9]. PT provides a short pulse thermal stimulus of milliseconds and analyzes decreasing temperature curves [22]. SH is a method of applying a long-term thermal excitation pulse, called long-pulsed IRT, square-pulse IRT, or conventional IRT [9,55,82]. LT supplies a modulated sinusoidal wave energy, synchronizes an IR camera with energy input, and measures its thermal response's phase difference and amplitude [55]. Laboratory tests or field surveys for infrastructures by active IRT usually adapt SH. This is probably because common construction materials, such as concrete, have lower thermal diffusivity than metals, hence a long heating time is required to cause thermal response [82].

3.2.2. Mode

IRT is also classified into two modes according to the relative position of an IR camera and a heat source: transmission mode and reflection mode [55,66]. These modes require different environmental conditions for measurement.

The transmission mode places a heat source on one side and an IR camera on the opposite side of a target object [66]. The temperature difference between both surfaces generates heat flow passing through the target object. Defect areas have different thermal properties from sound areas and disturb the heat flow, leading to nonuniform thermal distribution on the opposite surface. Thus, this mode can detect deep defects and internal structure differences, so that it is commonly used for energy audits to diagnose insulation

defects, moisture, and air leakage [48,55]. Additionally, this mode can quantitatively evaluate the U-value based on heat flux and the temperature difference between both surfaces [83]. However, The transmission mode needs to access both sides of the object. Moreover, as heat flow may take a long time to pass through an object, transmission mode requires keeping the difference between the inside and outside temperature for a long period to achieve a thermal equilibrium state in walls [7,33,48,83]. For instance, the British standard [7] states test requirements of a stable ambient temperature for at least 24 h before the measurement and no exposure to direct sunlight for at least 12 h. Therefore, thin building walls are appropriate for this mode.

The reflection mode places a heat source and an IR camera on the same side of a target object [66]. In this mode, radiation detected by the camera comes from heat flow reflected by defects [55]. Thus, this mode is suitable for detecting subsurface delamination at shallow places [55]. The advantage of the mode is that IRT can be conducted with access to only one side of the object. Therefore, the reflection mode is often applied to delamination inspection for infrastructures and buildings. However, it demands dynamic energy input into the surface from the outside by radiation or convection [51].

3.2.3. Measurement Method

IRT has two measurement methods: qualitative and quantitative IRT [50]. Qualitative IRT evaluates defects from color patterns indicating temperature in IR images [75]. Qualitative IRT does not require measuring accurate temperature values [22]. The primary aim of the survey for infrastructures is generally the investigation of the presence and location of defects. Thus, qualitative IRT is commonly employed as standards and guidelines [7,26,33,47,48,79] due to its simplicity. However, it is not easy to provide information on defect properties or severity levels [50].

Quantitative IRT is a numerical evaluation method by comparing temperature values on IR image pixels between identical items or baselines [50]. The quantitative method can assess defect properties or levels of severity. Various quantitative methods were studied, for example, the thermal resistance of walls [84], depth of delamination [85], and moisture content in lightweight concrete [86]. The challenge of this IRT is the requirement to measure accurate temperature. Hence, IR images need to be compensated for emissivity, atmospheric attenuation, and reflected temperature [22], in addition to the thermal properties of the tested object [50].

4. Delamination Detection

4.1. Principle of Delamination Detection

IRT for detecting subsurface delamination on infrastructures and buildings generally adopts passive IRT of the reflection mode [47]. The principle of passive IRT is capturing thermal contrast between delamination area and sound area due to nonuniform heat flow [87]. Figure 2 shows the heat transfer in a target object with delamination during (a) a heating cycle and (b) a cooling cycle. Figure 2c illustrates typical daily changes of surface temperature and thermal contrast on a sunny day.

From early morning to noon, solar radiation increases, and the ambient temperature rises. Solar irradiation and warm ambient temperature heat the surface of a target object, creating heat flow to the inside of the object. The thermal conductivity of concrete is approximately 1.6 W/mK, while that of air, filling delamination, is significantly low at 0.024 W/mK [88]. Thus, delamination acts as insulation and disturbs heat flow. As a result, the surface temperature above delamination becomes higher than the temperature of the surrounding area. Delamination areas appear as positive thermal contrast or hot spots in IR images, as shown in Figure 3. This period during daytime is called a heating cycle [89].

On the other hand, during nighttime, the surface temperature declines due to radiative cooling and low ambient temperature [90]. The heat energy stored in a target object during daytime transmits toward the surface, while this heat flow is obstructed by delamination. As a result, the surface temperature above delamination becomes lower than that of the

surroundings. Delamination appears as negative thermal contrast or a cold spot. This period during nighttime is called a cooling cycle [89].

IRT survey can be conducted during both the heating cycle and the cooling cycle. However, the periods when the two cycles exchange in the early morning and the evening, called interchange times [90], are not recommended for IRT surveys. The reason is that interchange time has lower thermal contrast than the two cycles, as shown in Figure 2c.

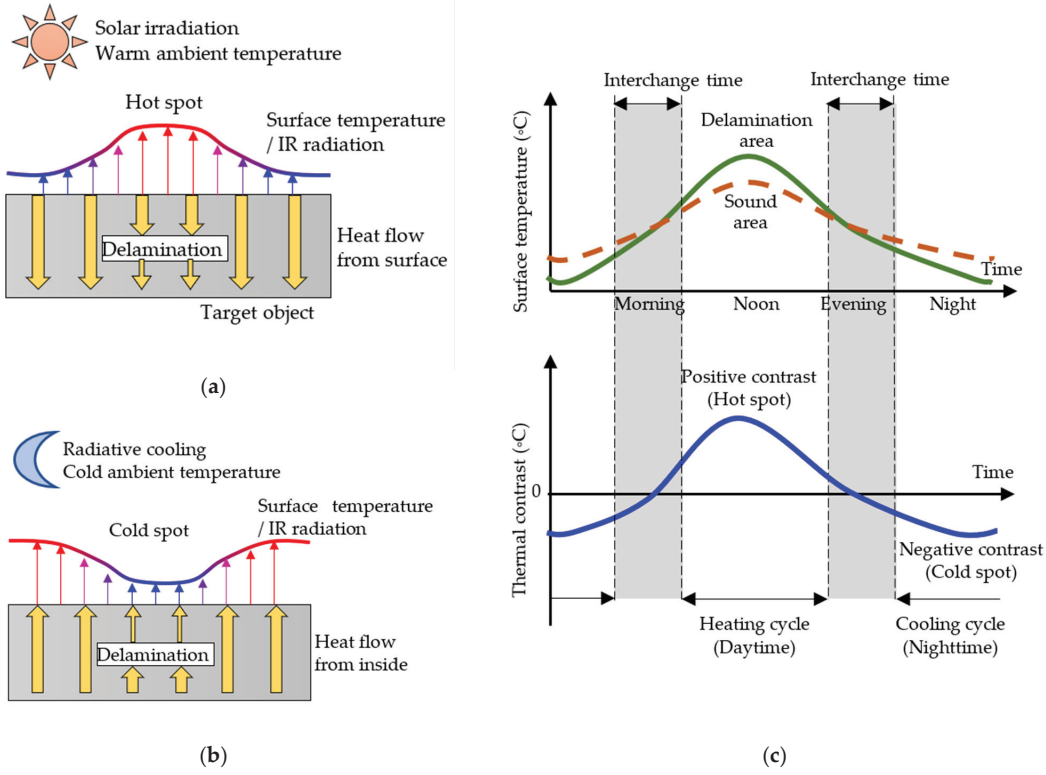
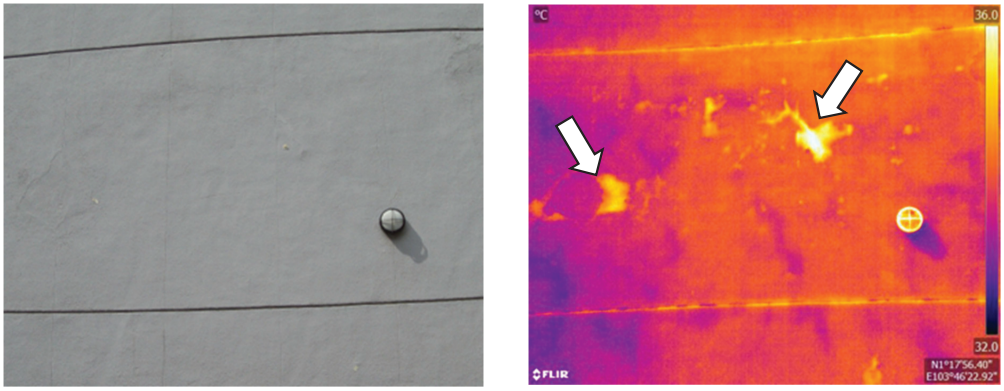


Figure 2. Principle of thermal contrast generation due to delamination: (**a**) diagram of heat flow during the heating cycle (daytime); (**b**) diagram of heat flow during the cooling cycle (nighttime); (**c**) daily changes of surface temperatures and thermal contrast.



(**a**) (**b**)

Figure 3. An example showing images of delamination on a building wall during heating cycle: (**a**) visual image; (**b**) IR image. Arrows indicate delamination areas.

4.2. Analysis Method

Analysis methods for delamination detection from IR images were developed as postprocessing procedures [74]. The analysis methods are divided into two groups based on the number of IR images used for analysis: one-time data analysis and time-series data analysis.

4.2.1. One-Time Data Analysis

One-time data analysis, called single-frame image processing [91], processes a single IR image at a specific moment. This analysis requires only one captured IR image of target objects, so that an inspector can survey large areas efficiently with one IR camera. Thus, the analysis is widely used for infrastructure and building surveys [47,92,93]. However, the analysis tends to be subject to noise due to surroundings and nonuniform heatings [94]. The one-time data analysis includes (a) visual evaluation, (b) thermal contrast, and (c) image processing.

(a) Visual Evaluation

Visual evaluation is a method that an inspector interprets temperature distribution patterns in an IR image by comparing surroundings and assesses the presence and location of delamination [95–97]. This evaluation is practical and has also been adopted in surveys [7,48]. One problem is that its accuracy and detectability depend on the inspector's experience, intuition, and judgment [98]. Therefore, interpretation should be conducted by a qualified inspector of IRT [99] to ensure inspection qualities. Another problem is that color scales representing temperature values need to be set in proper temperature ranges to avoid overlooking delamination [95,100–102]. For example, Washer et al. [100] suggested the range of 2.2–4.4 °C for shaded areas.

(b) Thermal Contrast

Thermal contrast, called ΔT, is referred to the surface temperature difference between the delamination area and the sound area [88,100]. Thermal contrast is a primary quantitative indicator to evaluate delamination in previous studies [10,88,103]. Thermal contrast may be due to causes other than delamination, such as surface conditions, subsurface materials, or object shape.

(c) Image Processing

Image processing was developed to extract temperature abnormalities automatically, quantitatively, accurately, and sensitively. The processing mainly utilizes threshold temperature values and temperature gradients.

Threshold temperature values are generally used for image processing. This image processing sets a threshold temperature value to judge areas as delamination and converts an IR image into a binary image based on the value. The processing has the advantages of simple evaluation; however, this processing primarily has two challenges.

The first challenge is determining threshold values because the values are affected by environmental conditions. Therefore, various methods to decide threshold values were proposed. The primitive method is that an inspector decides a threshold value that gives clear contrasts between sound and delamination areas by changing the value. The disadvantages of this method are subjective and time-consuming [104]. Japanese guideline of tile façade inspection [96] proposed that a delamination area is confirmed by the tapping method in advance, and the temperature difference between the delamination area and surrounding area is used as the value. Another approach is analyzing the temperature histogram of an IR image to determine the threshold value objectively. Garrido et al. [105] assumed that the histogram was a bimodal distribution composed of sound and delamination area. They acquired the temperature of the modal overlapping point by the Otus method as the threshold value. Meanwhile, Omar et al. [106] employed a k-means clustering method, an unsupervised machine learning method, to divide temperature values in an IR image into multiple clusters. They considered the boundary temperature values of clusters as the threshold values.

The second challenge of threshold values is difficult to evaluate the entire target object by one global threshold value. The reason may be that the entire surfaces of infrastructures or buildings are not under the same conditions, and each local area has a different average temperature and gradient [104]. Thus, methods for detecting temperature anomalies in local areas rather than in a global area were proposed. For example, Oh et al. [104] simply divided the IR image of a bridge deck into 16 local areas and used different threshold values for each area. Park et al. [107] extracted wall areas from building facades in visual images using a convolutional neural network (CNN) and analyzed the threshold values within wall areas. Cheng et al. [108] developed a delamination segmentation technique that extracts regional maximum temperature by a weight decay function. In these ways, it is necessary to limit the region of interest by some methods.

Temperature gradients are also employed for image processing. The processing identifies the areas of thermal anomalies based on the significant temperature changes at the edge of delamination. The advantages of the gradient are that measuring accurate temperature values is not required [109], and a slight temperature gradient over the entire surface may not be judged as delamination. For example, Lia et al. [109] identified delamination areas precisely by a spatial pixel differentiation algorithm even under unfavorable measurement conditions. In addition, Cheng & Shen [110] proposed temperature gradient-based level set method (LSM) and showed that LSM was more accurate and stable detection than the k-means method.

Overall, A substantial number of image processing methods using one IR image were developed. However, they may be designed to be optimized under specific conditions. To improve the accuracy of detectability and applicability for field inspection, further research is needed.

4.2.2. Time-Series Data Analysis

Time-series data analysis collects courteous IR images over time and analyses time-series temperature data. It is also called time-lapsed thermography [36,75], time-dependent IRT [9], or continuous multiframe image processing [91]. The advantages of this analysis are robust to noise by nonuniform environment conditions [94] and high detectability [82]. It also allows conducting the quantitative assessment of delamination depth [85]. Thus,

various image processing methods using multi-IR images were developed [22,74,91]. For example, simple image subtraction (SIS), also known as the computation of image differences, subtracts temperatures between two IR images at the same pixel location [111]. Principal component thermography (PCT), advanced processing based on principal component analysis (PCA) to summarize high-dimensional data [22], transforms a temperature 3D matrix in a combination of space and time into a 2D matrix by singular value decomposition to extract features and reduce noise [112,113]. Pulsed phase thermography (PPT), a method based on active IRT with one-dimension discrete Fourier transform, converts time-domain temperature data into frequency-domain data [112]. PPT has the advantage of suppressing the effects of spatially nonuniform heating and emissivity distribution [114]. Additionally, Cotič et al. [82] stated that PPT increased the maximum detectable depth by 50% over thermal contrast of one-time data analysis. In addition to the above, other methods were proposed including nonnegative matrix factorization (NMF) [111,115] and wavelet transformation [116]. Although time-series data analysis tends to be superior to one-time data analysis about detectability, the analysis requires fixing IR cameras and measuring the same object for a long duration. Therefore, time-series analysis is suitable for detailed inspection of a specific area, such as heritage sites, rather than the overall survey of infrastructures and buildings.

4.3. Standards and Guidelines

Table 1 shows existing standards and guidelines of IRT for delamination detection for infrastructures and buildings. These documents employ the passive analysis scheme and the reflection mode. Target objects include bridge decks [47], concrete structures [79,117,118], and tile and render finish façades [96,118]. The documents describe recommendations or requirements for environmental conditions and IR cameras.

Table 1. Existing standards and guidelines of infrared thermography (IRT) for delamination detection and recommended/required environmental conditions.

Document	Target Object	Recommended/Required Environmental Conditions			
		Solar Irradiation	**Ambient Temperature**	**Wind**	**Weather**
ASTM D47888-03 [47]	Bridge deck	A minimum direct solar irradiation for 3 h	An air temperature rise of 11 °C with 4 h of sun for concrete in winter An air temperature rise of 11 °C with 6 h of sun for asphalt in winter	Wind speed of less than 15 mph (6.7 m/s)	Dry for at least 24 h before the survey
Japan Public Work Research Institute [117]	Concrete infrastructure	A minimum direct solar irradiation of 350 Wh/h for 2–3 h	Daily temperature change of more than 10 °C in shaded areas Not suitable for 3–4 h after the maximum or minimum air temperatures	Wind speed of less than 5 m/s	Fine weather
British Instiute of Non-Destructive Testing [79]	Structural finishes	Strong solar exposure		Low wind speed	Fine weather

Table 1. *Cont.*

Document	Target Object	Recommended/Required Environmental Conditions			
		Solar Irradiation	**Ambient Temperature**	**Wind**	**Weather**
Japanese Society for Non-Destructive Inspection [118]	Concrete infrastructure, Tile façade, Shotcrete	A minimum direct solar exposure for 2 h			Fine or partly cloudy weather
Japan Building and Equipment Long-Life Cycle Association [96]	Tile façade, Render façade	Around the period of maximum solar irradiation on each elevation 2–4 h after sunset	Daily temperature change of more than 10 °C for shaded elevations	Wind speed of less than 5 m/s	No rain from one day before

Regarding environmental conditions, four factors are generally stipulated: solar irradiation, ambient temperature, wind, and weather. All documents recommend the survey with direct sunlight exposure because solar irradiation has significant energy input and generates high thermal contrast. ASTM [47] for bridge decks and the Japanese IRT standard [118] for tile or plaster finishes require continuous solar irradiation for 2–3 h before and during the measurement. Additionally, the Japanese Public Work Research Institute [91] defines the minimum intensity of solar irradiation. Regarding nighttime inspection, Japan Society for Non-Destructive Inspection [118] recommends the time window for the survey of 9 p.m. to 5 am, while Japan Building & Equipment Long-Life Cycle Association [70] recommends 2–4 h after sunset. The daily change of ambient temperature is considered as another stimulus. Thus, some documents mentioned recommended values, for example, a daily change of at least 10 °C for shaded areas [96,117]. In terms of wind, low wind speed is considered a suitable condition because wind removes heat from the surface. Thus, several documents stipulate that wind speed is less than 5 or 6.7 m/s [47,96,117]. These wind speeds correspond to 3 "Gentle Breeze" or 4 "Moderate Breeze" in the Beaufort wind force scale [119], respectively. Regarding weather, a fine day is recommended in all the documents since it provides direct sunlight and high daily ambient temperature change. Additionally, some documents [47,96] require no rain for one day and dried surfaces. In summary, long-duration solar irradiation, high daily ambient temperature change, low wind speed, and fine weather are commonly recommended conditions.

IR camera specifications, distance from a target object to an IR camera, and observation angle are also mentioned in the documents. One of the specifications is temperature resolution, represented as noise equivalent temperature difference (NETD) [120]. The NETD indicates a temperature difference that can be distinguished from noise. ASTM [47] requests an IR camera with the NETD of 0.2 °C or less, and other guidelines [117,118] demand that of 0.1 °C or less. With the recent development of IR camera technologies, even affordable cameras can commonly satisfy NETD of 0.1 °C or less [60,121]. Regarding distance, a short distance is preferable due to less infrared attenuation by the atmosphere. However, documents set a wide range of distances, such as 5–20 m [118] and 5–50 m [117]. Concerning angle, the limitations of observation angle vary depending on the documents, such as 30° [96] and 60° [117,118]. Large tolerances about distance and angle may be due to limitations of accessibility and surroundings of infrastructures and buildings.

4.4. Comparison with Other NDTs

In addition to IRT, several NDTs were developed to detect delamination on infrastructures and buildings: audio methods, stress wave methods, and electromagnetic methods [17–19,122,123].

Audio methods are based on a feature that when a mechanical impact is applied on a target object from outside, delamination areas produce impact sound with a frequency significantly different from intact areas (hollow sound); an inspector listens to the impact sound and evaluates delamination areas. The suitable method of giving mechanical impact

relies on target objects. For vertical surfaces such as buildings or tunnel linings, coin tapping testing, which uses coins, steel rods, or lightweight hammers, is widespread [96,122]. For bridge decks, chain dragging testing was standardized by ASTM [124]. The disadvantage of these audio methods is that interpretation depends on inspectors. Thus, a method of analyzing sound with fast Fourier transformation (FFT) to evaluate objectively was proposed [125].

Stress wave methods utilize characteristics of stress-wave propagation in a target object. Among the methods, impact-echo testing (IE) and ultrasonic testing (UT) can detect delamination. IE is a method that a mechanical impact is applied to a target object, then the frequency of the wave reflected on delamination is analyzed with FFT [126,127]. UT is a method in which a transducer emits ultrasonic pulses into a target object. An adjacent transducer receives the pulses reflected on delamination, rebar, or the object's boundary (pitch-catch method). The travel time of the pulses determines path length [123]. Additionally, a synthetic aperture focusing technique (SAFT) using multiple transducers can image the position and depth of delamination in 3D [126]. Although the stress wave methods require contact with an object, they can measure delamination depth.

Ground penetration radar (GPR) employs electromagnetic pulses [128]. Electromagnetic pulses propagate through a target object from an antenna. A receiver captures the pulses reflected on the boundary between media having different dielectric constants. GPR is widely used to inspect the inside of structures or bridge decks because it can detect delamination, voids, rebars, and buried objects [129,130].

IRT was compared with these NDTs about delamination detection: coin tapping testing [126], chain dragging testing [19,131], IE [19,126,132–134], UT [19,126,135], and GPR [19,126,128,129,132,135]. In the case of bridge deck inspection, IRT is as accurate as or slightly less accurate than IE [19,135] and more than as accurate as GPR [19,129]. Additionally, IRT is more suitable for detecting shallow delamination than GPR and UT, while IRT cannot detect deep delamination [126,132,135]. The advantages of IRT are that it can collect data without contact, inspection speed is the fastest among these NDTs, and the inspection cost is relatively low [19,135]. The disadvantage is that IRT is more sensitive to environmental conditions than other NDTs, so that the reliability of IRT is not high [132,135]. Therefore, a method that combines IRT with other NDTs to enhance accuracy, reliability, and measurable depth was investigated [89,128,131,133,134,136].

5. Recent Studies of Affecting Factors on IRT for Infrastructures and Buildings

The detectability of IRT is affected by many factors, including environmental conditions, delamination properties, target objects, and IR cameras [52]. Thus, the existing standards and guidelines of IRT state recommended conditions as mentioned in Section 4.3. However, these recommendations are not sufficiently quantitative and explicit. Furthermore, it is not practical for all the surfaces of an infrastructure to meet these recommendations, such as solar irradiation for a long duration. In addition, environments differ depending on the survey region. Therefore, affecting factors and these impacts on detectability were studied. Table 2 covers studies over the last 20 years on these factors using different environmental conditions, delamination properties, target object, and IR camera. This section compiles and discusses experimental methodologies adopted.

Table 2. Recent studies on affecting factors of delamination detectability of IRT for infrastructures and buildings.

Author	Year	Test Method	Target Object	Test Location (Outdoor Test or Field Survey)		Study Factors												
				Environmental Conditions							Delamination Properties			Target Object			IR Camera
				Region	Direction	Time window	Irradiation	Ambient Temperature	Wind	Others	Size * (cm)	Depth (cm)	Thickness (cm)	Material	Thermal Property	Others	
Chew [137]	1998	Laboratory test, Outdoor test	Concrete + Tile	Singapore	Vertical		✓	✓	✓		1–15	1	1–1.5				✓
Maierhofer et al. [138]	2002	Laboratory test, Numerical simulation	Concrete, Concrete + CFRP				✓				10–20	1–10	10				
Clark et al. [95]	2003	Field survey	Concrete, Stonemasonry	UK	Vertical, Soffit						No detail				✓		
Maierhofer et al. [139]	2004	Laboratory test, Numerical simulation	Concrete				✓				10–20	1–10	10		✓		
Maierhofer et al. [140]	2005	Laboratory test, Numerical simulation	Concrete				✓				10–20	1–10	10		✓		
Meola et al. [141]	2005	Laboratory test, Field survey	Brick/Marble/Tuff + Render	Italy	Vertical						4–10	1–5.5	0.1–0.2	✓	✓		
Maierhofer et al. [142]	2006	Laboratory test, Numerical simulation, Field survey	Concrete, Concrete + CFRP/Stone, Asphalt,	Germany	Horizon						10–20	2–8	10		✓	Rebar	
Meola [54]	2007	Laboratory test	Brick/Marble/Tuff + Render, Concrete								2–10	1–1.5	0.1–3	✓	✓	Water	
Maierhofer et al. [143]	2007	Laboratory test, Numerical simulation	Concrete								10	6–10	5	✓	✓	Concrete age, Rebar	
Yehia et al. [132]	2007	Outdoor test	Concrete	USA	Horizon						3.8–10.2	1.9–10.2	1.3–5.1	✓			
Cheng et al. [133]	2008	Laboratory test	Concrete, Concrete + Tile								5–16	0.5–3	7–9.5				
Washer et al. [144]	2009	Outdoor test	Concrete	USA	South	✓	✓		✓		30	2.5–12.7	1.3				
Washer et al. [145]	2010	Outdoor test	Concrete	USA	South	✓	✓				30	2.5–12.7	1.3				
Washer et al. [146]	2010	Outdoor test, Field survey	Concrete	USA	North			✓	✓		30	2.5–12.7	1.3				

Table 2. *Cont.*

Author	Year	Test Method	Target Object	Test Location (Outdoor Test or Field Survey)		Study Factors										
						Environmental Conditions					Delamination Properties			Target Object		IR Camera
				Region	Direction	Time window	Irradiation	Ambient Temperature	Wind	Others	Size * (cm)	Depth (cm)	Thickness (cm)	Thermal Material Property	Others	
Gucunski [19]	2012	Outdoor test	Concrete	USA	Horizon	✓					30–61	6.4–15.2	0.03–0.2			
Kee et al. [89]	2012	Outdoor test	Concrete	USA	Horizon	✓					30–61	6.4–15.2	0.03–0.2			
Scott et al. [147]	2012	Outdoor test	Concrete	South Africa	North	✓	✓		✓		15–40	1–6.5	1		Rebar	✓
Edis et al. [97]	2013	Field survey	Tile finish	Portugal	Vertical		✓			Reflection	No detail				Color, Texture, Moisture	✓
Washer et al. [100]	2013	Outdoor test, Field survey	Concrete	USA	North, Soffit			✓	✓		30	2.5–12.7	1.3			
Freitas et al. [148]	2014	Laboratory test, Field survey, Numerical simulation	Concrete + Render	Portugal	South	✓	✓			Weather	No detail					
Rumbayan & Washer [49]	2014	Numerical simulation	Concrete	USA	South, North		✓	✓	✓		30	2.5–12.7	1.3			
Scott & Kruger [149]	2014	Outdoor test	Concrete	South Africa	North	✓					25–52	5–10	5			
Alfredo-Cruz et al. [150]	2015	Outdoor test	Concrete	Colombia	Horizon	✓					15	2.5–7.5	1			
Bauer et al. [14]	2015	Laboratory test	Concrete + Render								No detail					✓
Cotič et al. [82]	2015	Laboratory test, Numerical simulation	Concrete								1.2–10	0.5–12.5	0.5	✓		
Edis et al. [20]	2015	Field survey, Numerical simulation	Brick + Tile	Portugal	Vertical	✓	✓	✓		Season	10	1	1–2			
Edis et al. [111]	2015	Field survey	Brick + Tile	Portugal	South, West	✓					No detail				Moisture content	
Khan et al. [151]	2015	Laboratory test, Numerical simulation	Concrete masonry				✓				20–142	No detail			Size	
Lai et al. [109]	2015	Outdoor test	Concrete + Tile/Render	Hong Kong	East		✓				7.5	0.3–2	0.3–2			
Vaghefi et al. [131]	2015	Field survey	Concrete	USA	Horizon						No detail	5.1–7.9	No detail			

Table 2. *Cont.*

Author	Year	Test Method	Target Object	Test Location (Outdoor Test or Field Survey)		Study Factors											IR Camera
				Environmental Conditions							Delamination Properties			Target Object			
				Region	Direction	Time window	Irradiation	Ambient Temperature	Wind	Others	Size * (cm)	Depth (cm)	Thickness (cm)	Material	Thermal Property	Others	
Watase et al. [152]	2015	Outdoor test	Concrete	USA	Horizon Soffit	✓		✓	✓	Relative humidity, Pressure	10	1–3	0.1				
Bauer et al. [153]	2016	Laboratory test, Field survey	Concrete + Tile	Brazil	East	✓					4	0.4–0.8	0.2				
Bauer et al. [154]	2016	Laboratory test	Concrete, Concrete + Tile				✓				4	0.4–0.8	0.2				
Ellenberg et al. [155]	2016	Outdoor test	Concrete	USA	Horizon						30–61	6.4–15.2	No detail				✓
Farrag et al. [102]	2016	Outdoor test	Concrete	UAE	Horizon		✓	✓	✓	Season	1.2–12.5	2.5–12.5	1.2–5.0	✓	✓	Rebar	
Hiasa et al. [156]	2016	Laboratory test	Concrete								10	1–3	0.1				✓
Huh et al. [157]	2016	Laboratory test	Concrete				✓				3–10	1–3	1				
Chiang & Guo [158]	2017	Field survey	Concrete + Tile	Taiwan	East, West, South, North	✓					No detail						
Hiasa et al. [88]	2017	Outdoor test, Field survey, Numerical simulation	Concrete	USA	Horizon						5–90	1.3–10.2	0.1–10				✓
Hiasa et al. [88]	2017	Outdoor test, Field survey, Numerical simulation	Concrete	USA	Horizon						5–90	1.3–10.2	0.1–10				✓
Hiasa et al. [159]	2017	Outdoor test, Numerical simulation	Concrete	USA	Horizon					Season	10–30	1.3–7.6	0.1–10				✓
Hiasa et al. [160]	2017	Field survey	Concrete	USA	Horizon	✓					No detail						✓
Hiasa et al. [161]	2017	Outdoor test	Concrete	USA	Horizon	✓					10.2	1.3–7.6	0.32				✓
Janků et al. [101]	2017	Outdoor test, Field survey	Concrete	Czech	Southwe area	✓	✓	✓		Weather	No detail	1–4	No detail				

Table 2. *Cont.*

Author	Year	Test Method	Target Object	Test Location (Outdoor Test or Field Survey)		Study Factors											
						Environmental Conditions					Delamination Properties			Target Object		IR Camera
				Region	Direction	Time window	Irradiation	Ambient Temperature	Wind	Others	Size * (cm)	Depth (cm)	Thickness (cm)	Thermal Material Property	Others	
Milovanović et al. [112]	2017	Laboratory test	Concrete								3–15	1–7	1–4	✓	Concrete age, Rebar	
Lourenço et al. [162]	2017	Outdoor test	IEICS / Brick + Tile	Portugal	West	✓				Weather	30	0.82	0.3	✓	Color, Water penetration	✓
Sultan & Washer [163]	2017	Outdoor test, Field survey	Concrete	USA	Horizon						15.2–60.9	5	2.54			
Tran et al. [164]	2017	Laboratory test	Concrete				✓	✓		Relative humidity	3–10	1–3	1			
Escobar-Wolf et al. [165]	2018	Laboratory test, Field survey	Concrete	USA	Horizon						2.5–10.2	2.5–5	1			✓
Güray & Birgül et al. [166]	2018	Numerical simulation	Concrete		Horizon	✓					10	1.1–4.1	0.2		Water penetration	
Hiasa et al. [90]	2018	Outdoor test, Numerical simulation	Concrete	USA	Horizon	✓				Weather	10	1.3–2.5	0.3		Surface obstacle	
Huh et al. [85]	2018	Laboratory test	Concrete				✓				10	1–8	1		Rebar	
Moropoulou et al. [43]	2018	Laboratory test, Numerical simulation	Stone				✓				1–3	2.5–3.5	No detail	✓		
Rocha et al. [103]	2018	Outdoor test	Concrete	Brazil	Horizon Shaded area	✓		✓		Relative humidity, Weather	10	2.5–7.5	0.3–1.2	✓		
Tran et al. [167]	2018	Laboratory test	Concrete				✓				7–15	2–8	1		Rebar	
Al Gharawi et al. [116]	2019	Outdoor test	Concrete	USA	South, North		✓			Month	30	2.5–12.7	1.3			✓
Cheng et al. [94]	2019	Laboratory test, Outdoor test, Numerical simulation	Concrete	USA	Horizon	✓	✓				5.1–15.2	3.8–8.9	0.4			
Mac et al. [56]	2019	Outdoor test	Concrete	Korea	Horizon	✓				Weather	5–15.8	2–7	1			✓
Vyas et al. [168]	2019	Outdoor test	Asphalt	India	Horizon	✓					60	5–10	No detail	✓		

Table 2. *Cont.*

Author	Year	Test Method	Target Object	Test Location (Outdoor Test or Field Survey)		Study Factors										
						Environmental Conditions					Delamination Properties			Target Object		IR Camera
				Region	Direction	Time window	Irradiation	Ambient Temperature	Wind	Others	Size * (cm)	Depth (cm)	Thickness (cm)	Material	Thermal Property / Others	
Cheng & Shen [110]	2019	Outdoor test, Field test	Concrete	USA	Horizon	✓					25	4.4–9.5	0.4			
Milovanovic et al. [169]	2020	Laboratory test	Concrete				✓				3–10	1–5	1–4			
Pozzer et al. [24]	2020	Outdoor test	Concrete	Brazil	Horizon	✓	✓	✓	✓	Relative humidity, Pressure	5–15	1–5	3			
Raja et al. [170]	2020	Laboratory test, Numerical simulation	Concrete				✓		✓		7–17	2.5–6.3	0.5			
Cheng & Shen [171]	2021	Laboratory test, Outdoor test	Concrete	USA	Horizon	✓	✓	✓			3–6	2.5–10	1–2			
Mac et al. [172]	2021	Outdoor test	Concrete	Korea	Soffit	✓		✓		Relative humidity	35–40	4–19.5	1			
Pozzer et al. [173]	2021	Outdoor test, Numerical simulation	Concrete	Brazil	Horizon	✓				Season	5–15	2–5	3			
Zheng et al. [174]	2021	Laboratory testOutdoor test	Concrete	China	Horizon						4–10	1.8–5	2.4–6.2			

Green shaded cells indicate factors studied by literature. * Size indicates the short side or the diameter of delamination.

5.1. Test Method

Test methods used in the previous studies are classified into four categories: laboratory test, outdoor test, field survey, and numerical simulation. Figure 4a indicates the frequencies of test methods employed by 66 studies. Laboratory tests and outdoor tests were mainly used by 41% and 48% of the literature, respectively. In contrast, the frequencies of field surveys and numerical simulations were low at approximately 30%. Thus, laboratory tests and outdoor tests predominated in previous studies.

In one laboratory test, specimens are prepared with polystyrene foam plates or other Insulation materials embedded to imitate delamination. Figure 5 shows a typical thermal contrast transition in a laboratory test. Artificial lamps heat the surface of a specimen during a heating period of 5–120 min [82,154]; thus, thermal contrast rises. After the lamps are turned off, thermal contrast continues to rise and reaches a peak. Then, thermal contrast decreases. The advantage of laboratory tests is that study factors can be controlled. The tests can investigate each factor independently and IRT detectability in ideal conditions with less noise. Many studies examined the impacts of delamination size and depth on detectability under laboratory conditions [54,82,167]. However, as it is not easy to simulate complex and dynamic outdoor conditions in a laboratory, the test is not appropriate to examine suitable time windows for the survey.

An outdoor test places specimens with simulated delamination in an outdoor location and observes the specimens for several days [159]. Thermal contrast generally behaves the curve shown in Figure 2c. The tests can examine detectability considering the combined effects of environmental factors [103,145,146]. Hence, the tests can investigate suitable time windows for the survey. However, environmental conditions around the specimens greatly depend on test region, climate, surface direction, etc. Thus, the results of outdoor tests are limited to a specific region and are not easy to be generalized.

A field survey is a method of inspecting existing infrastructures or buildings. The survey is often used to verify the results of laboratory tests and outdoor tests [101,148]. The difference from outdoor tests is that a field survey cannot control delamination properties; thus, some studies have compared the results using other NDTs [20,131,160]. Another disadvantage is the influence of noise, for example, reflections from surroundings [175,176], emissivity variation on the surface [27], subsurface material differences [97], and uneven solar heat gain [97]. In addition, thermal contrast can be caused by other subsurface defects, such as water penetration and high moisture content [20,176,177].

Numerical simulation or modeling may provide useful information on the impact of factors such as irradiation [20,170], defect size, and depth [82]. The accuracy of simulation results greatly relies on boundary settings; thus, validation according to laboratory tests or outdoor tests is essential. When modeling outdoor conditions, there are two types of input environmental data: meteorological observatory data [49] and standard environment data [20]. Software packages used in previous studies are general-purpose FEM software (e.g., COMSOL) [88] and transient thermal and humidity movement analysis programs for building envelopes (e.g., WUFI) [148].

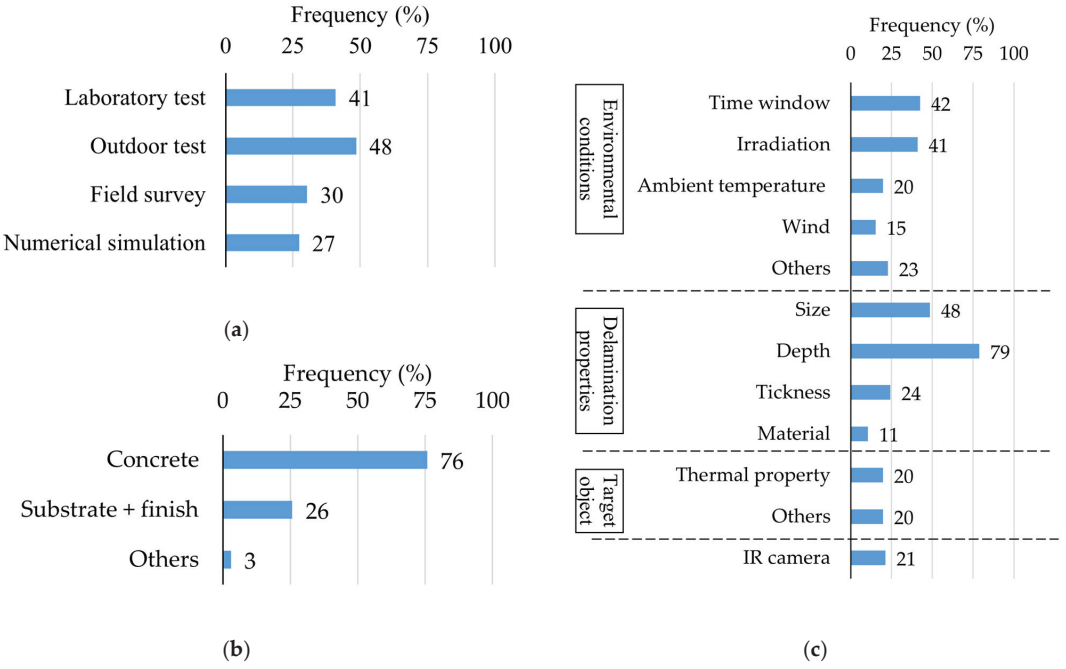


Figure 4. Statistics about methodologies in previous studies: (**a**) distribution of test methods; (**b**) distribution of target objects; (**c**) distribution of study factors.

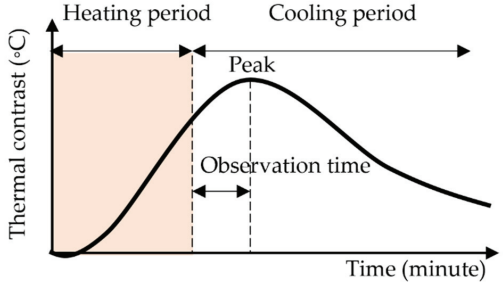


Figure 5. Typical thermal contrast transition in a laboratory test.

5.2. Target Object

Figure 4b shows that concrete was used as the target object in 76% of the previous studies because concrete is a fundamental and prevailing material in infrastructures and building structures. On the other hand, the substrate with finishes, the main materials in building facades, was at a low frequency of 26%. Substrates were mainly concrete, but few studies have examined the effects of bricks [162] or stones [54]. Finishes were tiles and mortar renders attached to substrates [109].

5.3. Test Location

The results of outdoor tests and field surveys may rely on the test region and surface direction. Most research was conducted under temperate climates, with high daily temperature changes and stable weather, for example, in the US [88] and Europe [20]. In contrast, there are few studies in the tropics [137].

Surface direction is also an important test condition because it relates to the magnitude and time of solar irradiation. In outdoor tests and field surveys, surface directions were mainly horizon or south elevation in the Northern Hemisphere. The horizontal direction assumed bridge decks, and the south elevation is considered ideal conditions with solar irradiation in buildings.

5.4. Metric and Criterion

Although metrics and criteria of detectable delamination are critical to identifying delamination and evaluating the impact of factors, there are no unified metrics and criteria. The metrics commonly used in previous studies are thermal contrast and signal-to-noise ratio (SNR).

Thermal contrast or ΔT, the temperature difference between sound and delamination area, is the most commonly used matric because it is simple and easy to analyze. However, the criterion of ΔT to be judged as delamination significantly differ depending on the studies, ranging from 0.2 to 1.2 °C. For example, Hiasa et al. [88,159] and Watase et al. [152] have defined a probable range for detectability as ±0.2 °C or larger and a certain range for detectability as ±0.4 °C or larger in outdoor tests for the concrete specimens. The reason was that Clark et al. [95] reported delamination on concrete bridges and masonry bridges was recognized when ΔT was more than ±0.2–0.3 °C. Additionally, Hiasa et al. [159] stated that ΔT of at least 10–20 times camera's NETD allowed inspectors to distinguish delamination from thermal noise. On the other hand, several studies [24,56,89,157,168] adopted ΔT of 0.5 °C as the criterion according to ASTM [47] for bridge deck inspection. Moreover, higher ΔT was used as the criterion. Farrag et al. [102] used ΔT of ±0.8 °C due to a more confident assessment. Another value of ΔT was ±1 °C. Washer et al. [100,145,146] mentioned that 1 °C was an order of magnitude larger than the thermal sensitivity of general IR cameras and was twice of ASTM [47]. Similarly, Raja et al. [170] employed ΔT of 1 °C in numerical simulations because the wind effect reduced ΔT to half. Chiang & Guo [158] also suggested ΔT of 1 °C as the criterion according to field surveys for tiled façades.

Meanwhile, some studies proposed multiple criteria of ΔT depending on test methods. For example, Zheng et al. [174] mentioned that it was difficult to identify temperature anomalies correctly by naked eyes when ΔT was less than 0.3 °C in the laboratory test and 1.2 °C in the outdoor test. Moreover, Sultan & Washer [163] examined the criteria quantitatively using receiver operating characteristics (ROC) analysis. As a result, 0.8 °C in the outdoor test and 0.6 °C in the field survey were optimum ΔT to balance true-positive and false-positive rates of delamination areas. As described above, the problem of thermal contrast is that the criterion is not established adequately. The reason may be that environmental conditions change thermal contrast and background noise.

The SNR is utilized as the metric to evaluate the detectability of delamination [10,67, 83,85,161] objectively. The *SNR*, which is used in engineering, compares single levels of a target area to signal levels of background noise, calculated by the following equation [85]:

$$SNR\ (\mathrm{dB}) = 20 \log_{10}(|S_{area} - N_{area}|/\sigma_{noise}), \tag{4}$$

where S_{area} is the average temperature value in the delamination area, N_{area} is the average value in the surrounding area, and σ_{noise} is the standard deviation in the surrounding area. Positive *SNR* means detectable delamination, and negative *SNR* means undetectable. The advantage of the *SNR* is that because of signal level evaluation, the metric and criterion can be applied not only to raw IR images but also processed images, such as PPT or PCT [10,169].

6. Affecting Factors of Detectability

This section compares and synthesizes previous studies on factors affecting delamination detectability. Figure 4c shows the frequency of study factors in the previous studies. Respectively, 43% and 40% of the studies examined the effects of time windows and irradiation. Meanwhile, only 20% and 15% of the studies dealt with ambient temperature and wind, respectively. This is probably because radiant heat transfer by sunlight is considered larger than convection heat transfer by the air. Regarding delamination properties, 78% and 48% of the studies investigated the effect of delamination depth and size, respectively. On the other hand, the effects of target objects and IR cameras were studied by approximately 20% of the literature. Therefore, time windows, irradiation, size, and depth are the main factors that attract attention among researchers.

6.1. Environmental Conditions

6.1.1. Time Window

Suitable time windows to conduct passive IRT are critical information for getting proper IR images to analyze. Multiple environmental factors, such as irradiation and ambient temperature change, can affect thermal contrast intricately. Thus, time windows are generally examined by outdoor tests and numerical simulations. Table 3 shows suitable time windows and interchange times in each direction under fine weather proposed by the literature.

Table 3. Suitable time windows and interchange times proposed by previous studies.

Direction	Author	Year	Time Windows
Horizontal surface	Yehia et al. [132]	2007	Defects of up to 3.8 cm deep can be detected between 10 a.m. and 3 p.m. Any defects cannot be detected during cooling cycle.
	Gucunski et al. [19]	2012	Defects at 40 min after sunrise are more apparent than at noon.
	Kee et al. [89]	2012	IR images obtained during cooling cycle are more evident than those obtained during heating cycle. Defects cannot be detected 3:45 h after sunrise. Shallow defects of 6.4 cm can be detected 7 h after sunrise.
	Watase et al. [152]	2015	Any time of day is suitable for 1 cm deep delamination, and 6 a.m. is best time.
	Hiasa et al. [90]	2018	Defects can be detected between 10 a.m. and 3 p.m. Defects can be detected between 5 p.m. and 8 am, and maximum contrast appears at 7 p.m. Cooling cycle is more suitable than the heating cycle for the inspection.
	Güray et al. [166]	2018	Favorable time window is between 3 p.m. and 7 p.m.
	Mac et al. [56]	2019	Optimal time windows for up to 4 cm deep defects are between 10 a.m. and 3 p.m. and between 7:30 p.m. and 2:00 a.m.
	Vyas et al. [168]	2019	Interchange times for asphalt unbonded by sand are between 8 a.m. and 10 a.m. and between 2:30 p.m. and 3:30 p.m.
	Pozzer et al. [24]	2020	Ideal time window is between 12 p.m. and 3 p.m.
South elevation (in the Northern Hemisphere)	Washer et al. [144]	2009	Optimum time is from 5–9 h after sunrise.
	Washer et al. [145]	2010	Optimum time is after 5:40 h after sunrise for 2.5 cm deep delamination and 9 h after for 12.7 cm.
	Scott et al. [147]	2012	Recommended time window is between 12 a.m. and 3 p.m. for under 6.5 cm deep delamination.
	Scott & Kruger [149]	2014	Optimum time window is between 11 a.m. and 1 p.m. for under 5 cm deep defects.

Table 3. *Cont.*

Direction	Author	Year	Time Windows
	Edis et al. [20]	2015	Interchange times occur between 5:30 a.m. and 6:50 a.m. and between 4:30 p.m. and 5:50 p.m.
	Chiang & Guo [158]	2017	Available time window is between 10 a.m. and 12 p.m.
	Janků et al. [101]	2017	Best time is around noon. Interchange time occurs at 4 p.m.
	Freitas et al. [148]	2018	Best time window is during hours of exposure to sunlight. Defects are less evident during cooling cycle than heating cycle.
East elevation	Bauer et al. [153]	2016	Defects are better visualized in early morning and late afternoon. Interchange time is around 12:30 p.m.
	Chiang & Guo [158]	2017	Available time window is between 9 a.m. and 11 a.m.
West elevation	Chiang & Guo [158]	2017	Available time window is between 12 p.m. and 2 p.m.
	Lourenço et al. [162]	2017	Desirable time during heating cycle is first 1:30 h after beginning of irradiation exposure. Desirable time during cooling cycle is beginning of cycle or 1 h after beginning of shadowing.
Shaded area/Soffit/North elevation (in the Northern Hemisphere)	Watase et al. [152]	2015	Favorable time window is midnight.
	Chiang & Guo [158]	2017	Available time window time is between 11 a.m. and 1 p.m.
	Janků et al. [101]	2017	Best conditions occur around noon.
	Rocha et al. [103]	2018	Best time window is between 10 a.m. 2 pm, specifically at noon. Interchange times are around 7 a.m. and 5 p.m.
	Mac et al. [172]	2021	First optimal time window is 7 h after decks are exposed to sunlight until 0.5 h after decks are not exposed. Second one is from 1.5 h to 3.5 h after decks are exposed to sunlight. Third one is 8 h after decks are not exposed to sunlight until 1 h after decks are exposed.

Regarding horizontal surface and south elevation, available time windows proposed by the literature are generally around 10 a.m. to 3 p.m. due to the presence of solar irradiation [56,90,101,132,147]. However, suitable or optimum time windows vary. Chiang & Guo [158] mentioned recommended time window of 10 a.m. to 12 p.m. according to the field survey for tile façades in Taiwan. Meanwhile, Scott et al. [147] suggested that the recommended time window was 12 p.m. to 3 p.m. for up to 6.5 cm deep delamination because of a time lag between the maximum solar loading at noon and thermal contrast responses. Pozzer et al. [24] statistically analyzed meteorological data and thermal contrast. They predicted favorable time windows from 12 p.m. to 3 p.m. due to high solar radiation, high ambient temperature, and low pressure.

Furthermore, several studies proposed that suitable time windows relied on delamination depth. The reason is that the deeper delamination is, the longer it takes for heat flow to reach delamination. Washer et al. [145] showed that the optimum time for 2.5 cm deep delamination was 5:40 h after sunrise and that for deep delamination of 12.7 cm was 9 h after sunrise. Similarly, Kee et al. [89] reported that 6.4 cm deep delamination could not be detected 3:45 h after sunrise even though it satisfied 3 h of solar irradiation required by ASTM [47]. In contrast, Watase et al. [152] argued that any time was suitable for shallow delamination of 1 cm. Additionally, delamination size can affect time windows. For example, Scott & Kruger [149] stated that the small delamination of 25 cm diameter generated the maximum contrast 4:30 h after sunrise, whereas the large delamination of 50 cm did 6:30 h after sunrise.

Meanwhile, several studies have focused on interchange times, which can not detect delamination due to low thermal contrast. Edis et al. [20] calculated that the interchange times happened on tile façades from 5:30 a.m. to 6:50 a.m. and from 4:30 p.m. to 5:50 p.m. Similarly, Janků et al. [101] confirmed that the times occurred at 8 a.m. and 4 p.m. in the

outdoor test. Hiasa et al. [90] also reported that the interchange time windows were 1 h in the morning and 2 h in the evening.

Overall, many studies examined suitable time windows for horizontal surface and south elevation during daytime. Although it is affected by delamination properties, the suitable time window is generally noon to early afternoon when delamination depth is under 6 cm.

Regarding east and west elevation, time windows with solar irradiation on the elevation should be optimal. It means that the suitable time for east elevation is in the morning and that for the west elevation is in the afternoon. For example, Buare et al. [153] observed in the field survey that the maximum contrast appeared at 8:30 a.m. on the east elevation, and thermal contrast declined toward 12:30 p.m. Thus, they proposed that the beginning of sun exposure was the optimal time window. Similarly, Lourenço et al. [162] pointed out that the desirable time was the first 1:30 h after the beginning of solar irradiation on the west elevation. Chiang & Guo [158] also mentioned that the recommended time windows were 9 a.m. to 11 a.m. on the east elevation and 12 p.m. to 2 p.m. on the west elevation. Therefore, suitable time windows for east or west elevation can be after direct sunlight exposure.

Nighttime or the cooling period is a candidate for suitable time windows; however, this is still being debated. One opinion is that nighttime is not appropriate or impossible to conduct IRT. Yehia et al. [132] failed to detect delamination in the outdoor test at night. Additionally, Freitas et al. [148] argued that nighttime inspection was available, while delamination during nighttime was less evident than those during daytime. The opposite opinion is that nighttime is more optimum than daytime because of less noise on IR images [19,90,161] or a long measurable duration [56,90]. Hiasa et al. [90] observed that IR images captured during daytime had much noise caused by nonuniform heating and shadows from surroundings. Moreover, Mac et al. [56] stated that the available time window during nighttime was from 7:30 p.m. to 2 a.m., which was longer than the window during daytime from 10 a.m. to 3 p.m. The difference in the literature on detectability during nighttime is considered due to environmental conditions.

Furthermore, there are still two opinions about suitable time windows during nighttime: early night or early morning. Hiasa et al. [153] mentioned that the maximum negative thermal contrast of 2.5 cm deep delamination occurred at around 7 pm, and the delamination was well recognized. Lourenço et al. [162] also insisted the optimum time was 1 h after the surface was covered in shades for tile facades. On the other hand, Kee et al. [89] suggested that even deep delamination, which was undetectable during daytime, could be detected 45 min after sunrise because of a long cooling duration until early morning. Hence, these studies indicate that optimum time windows during nighttime are dependent on delamination depth.

Shaded areas, soffit, or north elevation, which has no solar irradiation on the inspected surface, may exist on infrastructures and buildings. In these areas, suitable time windows during daytime are generally around noon due to the peak of ambient temperature; however, these time windows are shorter than those of sunny areas [101,103,158]. Regarding daytime and nighttime, previous studies do not agree with which time window is suitable. Rocha et al. [103] argued that thermal contrast during nighttime was smaller than that during daytime. In contrast, Watase et al. [152] proposed that midnight was the favorable time window for deck soffit rather than noon because of a high probability of days when thermal contrast exceeded the criterion of detectability. Thus, further studies are needed on suitable time windows for shaded areas.

As explained above, suitable time windows for IRT proposed by previous studies are not consistent. The reason is that the windows are affected not only by surface direction but also by environmental conditions and delamination properties. Therefore, investigating suitable time windows for each region and the target object is required to conduct IRT properly.

6.1.2. Irradiation

Solar irradiation is a primary stimulus producing heat flow [20,145]. It reaches 700 W/m^2 on a south elevation and 1300 W/m^2 on a horizontal surface at noon [88,148,159]. Previous studies have demonstrated that the larger the energy input is, the higher thermal contrast and SNR are generated in laboratory tests [10,85,139,157,169,170]. In contrast, delamination is difficult to be detected under low or no solar irradiation conditions [20,148,162]. In addition, detectable delamination depth is proportional to the heating time in the laboratory test [85,167]. Meanwhile, excessive energy input could decline the thermal contrast of shallow delamination [82,91]. Overall, a large amount of irradiation is generally a preferable condition for IRT.

Few studies quantitatively investigated the relationship between irradiation and thermal contrast. Washer et al. [145] conducted the outdoor tests for three months and argued that the daily total solar loading, not the maximum solar loading, had a high correlation with the maximum thermal contrast. The authors suggested that the total daily solar roading of at least 700 Wh/m^2 was required for 5.1 cm deep delamination to generate the detectable thermal contrast of 1 °C based on statistical analysis. Likewise, Raja et al. [170] proposed that the total irradiation of 680 Wh/m^2 produced the thermal contrast of 1 °C for 6.3 cm deep delamination based on the numerical simulations. In addition, the authors stated that a heat flux rate greatly influenced thermal contrast, especially for shallow and small delamination. These studies indicate that total irradiation of approximately 700 Wh/m^2 could be required to conduct passive IRT.

6.1.3. Ambient Temperature

Daily ambient temperature change is one of the drivers to generate thermal contrast due to convection heat transfer. The daily change is a primary heat source in shaded areas or under cloudy weather [20]. Multiple studies concluded that significant daily ambient temperature change increases thermal contrast and is preferred for IRT based on outdoor test results [100,101,103,137,146,152]. However, the amount of daily change required in shaded areas is not consistent among previous studies. For example, Janků et al. [101] confirmed that the daily change of more than 10 °C was necessary, while Rocha et al. [103] also insisted at least 5.4 °C. Likewise, Washer et al. [100] suggested a daily change of at least 8 °C for 5.1 cm deep delamination. Additionally, the authors proposed that the rate of ambient temperature change of at least 1.5 and −1.7 °C/h was favorable for daytime and nighttime inspection, respectively. Overall, the high daily ambient temperature change is favorable for passive IRT in shaded areas, although the requirement is still debated.

Ambient temperature values might influence thermal contrast. Tran et al. [164] mentioned that high ambient temperature increased thermal contrast, especially for large and shallow delamination, although the effect of temperature was significantly smaller than irradiation.

For buildings, an ambient temperature difference between indoor and outdoor can also affect thermal contrast. Edis et al. [20] conducted parametric studies on the effect of the difference using numerical simulation. The difference enhanced thermal contrast on the surface during daytime when the outdoor temperature was hotter than the indoor temperature. Thus, the effect of the ambient temperature difference should be considered when the difference is more than 10 °C.

6.1.4. Wind

Wind velocity is an environmental factor to be considered when performing passive IRT, as it relates to convection heat transfer [49]. High wind velocity increases heat transfer between the surface and the air [178]. Thus, the wind has different effects on thermal contrast depending on the presence of solar irradiation.

Under the condition of solar irradiation or during the heating cycle, high wind velocity decreases thermal contrast. The reason is that the surface temperature of a target object is generally higher than ambient temperature, so that the wind removes heat energy from the

surface. For example, Washer et al. [144] statistically analyzed the relationship between the maximum thermal contrast and average wind velocities in the outdoor tests. As a result, average wind velocity tended to be low when thermal contrast was high. Moreover, Raja et al. [170] quantitatively investigated the effect in the laboratory tests and stated that thermal contrast decreased as the wind velocity increased, especially for deep delamination. For example, the wind velocity of 7 m/s reduced thermal contrast by half for 6.3 cm deep delamination. Furthermore, the authors stated that the slight wind velocity of 1.4 m/s also decreased thermal contrast by 20%. Therefore, low wind velocity is preferable in sunny areas when solar irradiation is used as thermal stimulation.

In contrast, in shaded areas, high wind velocity could increase thermal contrast. The reason is that the surface temperature is generally lower than ambient temperature, and high wind velocity increases energy input from the air to the object's surface. Washer et al. [100] pointed out that high wind velocity improved thermal contrast based on the outdoor tests. Although high wind velocity is preferable in shaded areas, Washer et al. [146] suggested a guideline that average velocity during 6 h is limited to 4.4 m/s (16 km/h) because high wind velocity might indicate unstable weather conditions. Overall, wind positively affects thermal contrast in shaded areas, as opposed to sunny areas.

6.1.5. Relative Humidity

Relative humidity (RH) is considered to affect thermal contrast due to two theories. One theory is that high RH increases convection heat transfer between the object surface and atmosphere [179]. Thus, in shaded areas or soffit, high RH increases the effect of ambient temperature change on thermal contrast during the heating and cooling cycle [87,172]. The other theory is that high RH increases water adsorption on the surface. Rocha et al. [103] suggested that high RH during nighttime enhanced negative thermal contrast because water adsorption increases moisture content and thermal conductivity near surfaces. Therefore, high RH is typically a preferable condition for IRT.

However, the effect of RH may be limited and not significant. For example, Tran et al. [164] argued that thermal contrast under high RH was more evident than that under low RH for shallow delamination of 1 cm in the laboratory test. In comparison, there was no difference in thermal contrast for 2–3 cm deep delamination. Additionally, Washer et al. [87] mentioned that the effect of RH was not significant in sunny areas because the effect of solar irradiation is dominant. These studies indicate that the positive effects of high RH are less significant than other factors.

6.1.6. Others

Weather is closely related to other environmental factors. A sunny day is optimal for IRT regardless of sunny or shaded areas due to high solar radiation and high daily ambient temperature change [56,101,147,162]. A cloudy day is not recommended because of the small energy input from irradiation and ambient temperature change [148,162]. A partially cloudy day should also be avoided as rapid irradiation changes might make delamination identification difficult [162]. Regarding nighttime, a clear sky is also optimum because radiative cooling removes heat energy from the surface and enhances thermal contrast [90,103]. Overall, fine weather is desirable at all times. However, IRT cannot always be performed under fine weather, so that identifying acceptable weather conditions for IRT is necessary in practice.

A method to predict thermal contrast from meteorological data was proposed. Watase et al. [152] proposed multilinear regression formulas to estimate thermal contrast on bridge deck and soffit under Florida climate conditions. The variables of the formulas were ambient temperature at a bridge and ambient temperature and atmospheric pressure at a nearby meteorological observatory. Furthermore, Washer & Fuchs [180] developed a web-based application to predict whether passive IRT can be carried out based on meteorology records and weather forecasts. Likewise, Pozzer et al. [24] performed multivariate regression analysis under Brazilian climate conditions, considering interactions of meteorological

variables. They mentioned that significant dependent variables were ambient temperature, atmospheric pressure, solar radiation, and survey time. In contrast, the size and depth of delamination and wind velocity were not significantly related to thermal contrast. Although these predictions are useful in practice, these formulas are limited to specific regions and are not general.

Seasons are related to the amount of solar irradiation and daily ambient temperature change. Hence, the effects of seasons were examined using numerical simulation, but the results are not consistent. Hiasa et al. [159] concluded that seasonal effects were minor on the horizontal plane in Florida. In contrast, Pozzer [173] mentioned that spring and summer were desirable for IRT in Brazil because of high daily ambient temperature change. Therefore, preferred seasons for IRT depend on the region.

6.2. Delamination Properties

6.2.1. Size

Detecting small delamination at the early stage of deterioration leads to ensuring public safety. As shown in Figure 4c, half of the studies have examined the effect of the delamination size. Regarding the relationship between size and thermal contrast, Hiasa et al. [159] showed that size had a much more substantial effect on thermal contrast than thickness and volume of delamination by numerical simulation about outdoor tests. Moreover, Raja et al. [170] argued that the total heat input to create the contrast of 1 °C was inversely proportional to the area; thus, large delamination needed less input heat to be detected. However, Hiasa et al. [88] stated the size effects converged at approximately 40 cm. Additionally, the authors examined the impact of an aspect ratio of the delamination area. The thermal contrast of delamination with an aspect ratio of 25% or more was comparable to the contrast equal to the area of square or circle. In general, large-size delamination with a high aspect ratio has significant thermal contrast and is easily detected.

Delamination size is also related to the response time of the maximum thermal contrast. Maierhofer et al. [139] confirmed that observation time, shown in Figure 5, became longer as the area increased. Similarly, Scott & Kruger [149] mentioned that the delay of the maximum thermal contrast from peak irradiation increased as the size was large in the outdoor test. Thus, delamination size may change optimum time windows for inspection.

6.2.2. Depth

Depth from delamination to the surface significantly affects thermal contrast. Thus, detectable depth is an essential indicator of IRT abilities. Approximately 78% of the studies include depth as study parameters, as shown in Figure 4c. The range of depth examined is wide and depends on target objects assumed in the literature. For buildings, delamination was generally set to a depth of 0.5–3 cm [109,133,137,141,154,162]. For concrete civil infrastructures, the delamination depth was set to a depth of approximately 2–8 cm [82,85,101–103,139,157,159,163,166], which are standard concrete cover thickness [181]. Moreover, some studies examined 10 cm or more deep delamination to evaluate IRT limitations [19,82,88,89,102,132,139,145].

It is not easy to detect deep delamination as deep delamination has low thermal contrast. Table 4 lists the maximum detectable depth in previous studies by one-time data analysis. An overall trend is that maximum detectable depth depends on conditions. The detectable depth in (b) and (c) outdoor tests with solar irradiation tends to be deeper than that in (a) laboratory tests. The reason can be the difference in the total amount and time of energy input to test objects.

Table 4. Maximum detectable depth in literature.

Conditions	Maximum Detectable Depth in Literature
(a) Laboratory test	6 cm [143], 7 cm [10,85,167], 7.5 cm [82], 10 cm [138]
(b) Outdoor test with solar irradiation measured during heating cycle (daytime)	3 cm [159], 3.2 cm [174], 4 cm [56,101], 5.1 cm [88,132], 6.5 cm [147], 7.5cm [150],10 cm [102], 12.7 cm [87,144]
(c) Outdoor test with solar irradiation measured during cooling cycle (nighttime)	3 cm [159], 4 cm [56,101] 10.2 cm [88], 12.5 cm [102], 12.7 cm [87], 15.2 cm [89]
(d) Outdoor test in shaded areas	4 cm [101] 5 cm [103], 7.6 cm [87], 19.5 cm [172]

Furthermore, detectable depth was a controversial and much-disputed subject even under the same test condition. In (b) and (c) outdoor tests with solar irradiation, Washer et al. [87] mentioned that 12.7 cm and 7.6 cm deep delamination were detectable during the heating and cooling cycles, respectively. In contrast, Kee et al. [89] argued that 6.4 cm and 15.2 cm were the maximum detectable depths during the heating and cooling cycles, respectively. Besides, Hiasa et al. [159] reported that 5.1 cm deep delamination was not detectable at any time, and approximately 3 cm was the maximum depth in Florida. Similarly, Mac et al. [56] stated that delamination of up to 4 cm depth could be detected in South Korea. These differences in the detectable depth could be due to differences in environmental conditions, delamination properties, target objects, and metrics.

Depth estimation was also of great interest to researchers because depth is essential information to evaluate severity. For example, AASHTO Guide Manual for Bridge Element Inspection [182] assesses the severity of delamination based on its size and depth. Currently, two approaches to estimate depth were proposed: response time and thermal contrast magnitude.

The estimation method based on response time utilizes that delamination depth correlates with the time from energy input to the generation of thermal distribution on the surface [183]. In laboratory tests, this response time is defined as observation time, a difference from the end of the heating period to the peak [167], as shown in Figure 5. Many studies estimated delamination depth accurately using the observation time [10,82,85,139,157,164,167]. However, the coefficient of estimation formulas changes depending on environmental conditions and the thermal diffusivity of target objects [157,167]. Moreover, delamination size also influences response time and the observation time [139,149]. Thus, the estimation method based on response time is possible only under a specific controlled environmental condition, such as laboratory tests.

The estimation method based on thermal contrast magnitude uses the correlation between thermal contrast and depth. Tran et al. [164] showed in the laboratory test that the inverse of the cube of depth was proportional to the loss of contrast with relatively high accuracy. The authors insisted that this method was practical because it can quickly estimate depth without time-consuming analysis of observation time. Similarly, Raja et al. [170] demonstrated a linear correlation between the square of the depth and the total energy input to generate thermal contrast of 1 °C. However, these methods are difficult to be applied to outdoor tests because environmental conditions are not constant and change dynamically. Hence, Hiasa et al. [88] proposed a method of comparing actual thermal contrast to calculated thermal contrast at each depth by numerical simulation. Although it can estimate depth in outdoor tests, the method requires obtaining time-series data of irradiation and ambient temperature and the thermal properties of the target object. In addition, numerical simulation must be conducted for each depth based on those data.

As described above, depth estimation methods using response time or thermal contrast magnitude are possible under constant or controlled conditions such as laboratory tests. However, since environmental conditions fluctuate, further research is needed to estimate depth under outdoor conditions.

6.2.3. Width to Depth Ratio

Delamination width and depth are closely related to detectability while interacting. It is generally considered that the minimum detectable width is at least 1–2 times the depth or more [184]. Thus, many studies have investigated the width-to-depth ratio (WTDR) criterion of detectable delamination in laboratory tests [10,82,85,157,167] and outdoor tests [56,102,174]. Figure 6 shows the syntheses of the previous results of detectability with respect to the width and depth of delamination in concrete specimens. The data were categorized according to test conditions. The WTDR criteria proposed by the literature are also displayed in Figure 6. A WTDR corresponds to the slope of the straight line through the origin of figures. The overall tendency is that the upper left region of each graph, high WTDR, clearly has a high probability of delamination detection. The reason can be that the larger WTDR delamination is, the higher the thermal contrast is and the easier it is to detect by IRT [56,174,185]. Furthermore, the results of the same width and depth delamination are not consistent enough, especially for delamination near the proposed WTDR criteria. This inconsistency can be due to the difference in environmental conditions, delamination properties, and detection metrics.

More specifically, each condition has a different tendency for detectable delamination distribution and WTDR criteria. In (a) laboratory tests, the distribution results are almost consistent among the literature compared to outdoor tests. This is probably because such laboratory tests can optimize energy input and remove unintended noise from the surroundings. Additionally, the detectable and undetectable regions are relatively separated by a straight line. Thus, the WTDR criteria proposed by the literature are relatively low values of 1.11–1.43 [82,85]. This means that laboratory tests can detect small and deep delamination. In (b) outdoor tests with solar irradiation measured during the heating cycle, the WTDR criteria of 1.8–2.25 were proposed [56,174], which are higher than those in (a) laboratory tests. In (c) outdoor tests with solar irradiation measured during the cooling cycle, the distribution of detectable delamination and the proposed WTDR criteria differ significantly depending on the studies. Mac et al. [56] suggested that the WTDR criterion was 2.5 in Korea, whereas Farrag et al. [102] proposed the that of 0.4–0.5 in the UAE. This difference can be attributed to intense solar irradiation during daytime in the UAE. Figure 6c indicates that the proposed WTDR of 2.5 [56] relatively agrees with the results of other studies. In (d) outdoor tests in shaded areas, WTDR criteria were not proposed by previous studies to our knowledge. Although the number of results is not adequate, Figure 6d suggests that the distribution is not significantly different from (c) outdoor tests with solar irradiation measured during the cooling cycle.

As described above, the WTDR criteria of detectable delamination are influenced by test methods, the presence of solar irradiation, measurement cycle, and test regions. As a result of integrating previous studies, WTDR criteria are approximately 1.25 in (a) laboratory test, 2.0 in (b) outdoor test with irradiation during the heating cycle, and 2.5 in (c) outdoor test with irradiation during the cooling cycle and (d) outdoor test in shaded areas.

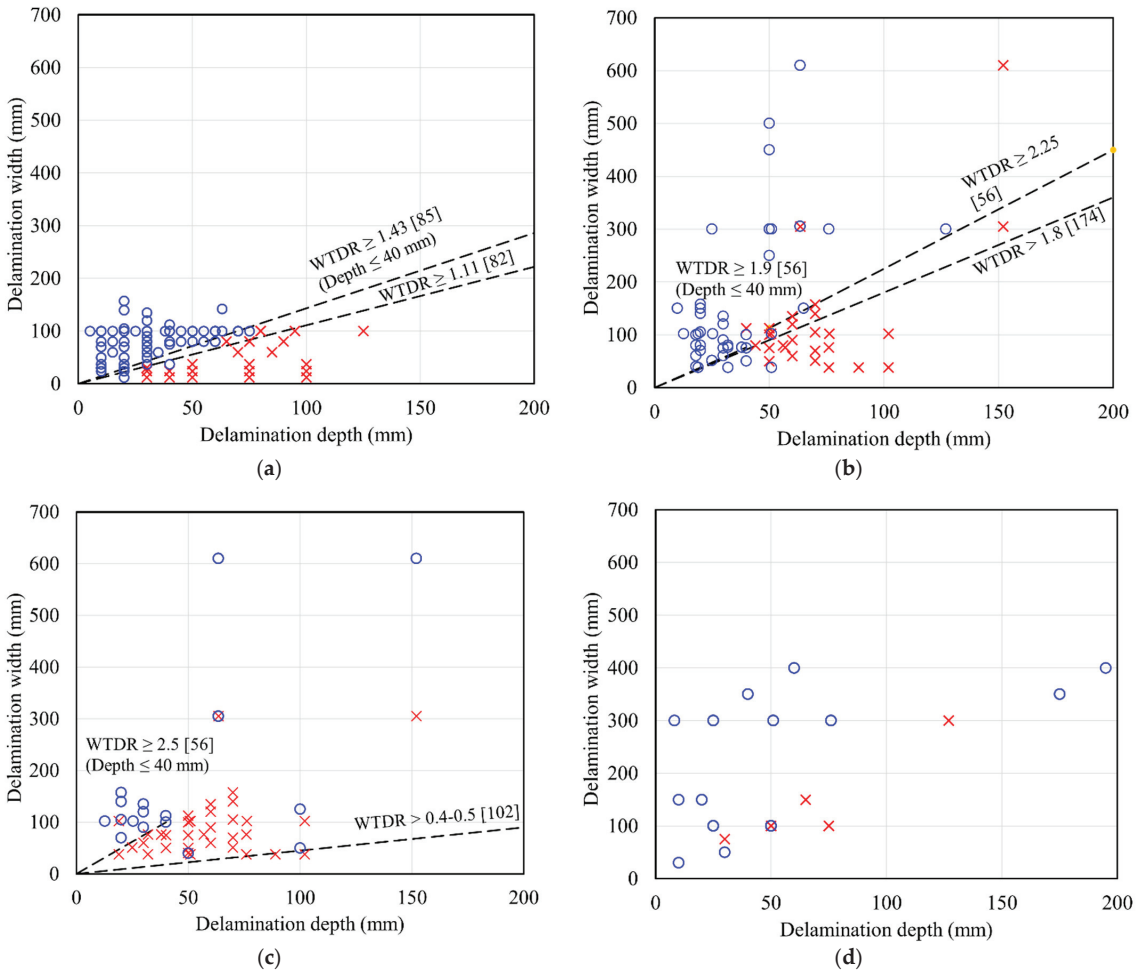


Figure 6. Synthesises of literature about detectability with respect to depth and size of delamination in concrete: (**a**) laboratory test; (**b**) outdoor test with solar irradiation measured during heating cycle (daytime); (**c**) outdoor test with solar irradiation measured during cooling cycle; (**d**) outdoor test in shaded areas. Legends are that blue circles indicate detectable delamination, while red crosses indicate undetectable delamination. Synthesized data have following terms: target object is ordinary concrete; analysis method is one-time data analysis; width represents diameter or shortest side of delamination; detectability is determined according to each study. Sources are [10,56,82,85,87,89,100–103,132,133,143,145–147,149,157,159,164,167,170,174,186].

6.2.4. Thickness

Delamination thickness is also a factor to consider for its impact on detectability. Previous studies have generally set the thickness of 0.1–2 cm by adjusting the thickness of embedded materials. Thick delamination has a low overall heat transfer coefficient; thus, it generates significant thermal contrast regardless of environmental conditions or measurement cycles [20,88,89,102,103,141,146,159]. For example, Kee et al. [89] reported that delamination of 0.1–0.2 cm thickness at 6.4 cm depth was detectable, while thin delamination of 0.03 cm thickness was undetectable. However, the effect of thickness may

converge at a certain value. Hiasa et al. [88] showed convergence at 1 cm thickness by the numerical simulation.

Meanwhile, thickness is considered the minor effect on thermal contrast among the geometric factors of delamination [91]. Hiasa et al. [159] demonstrated that the most influential factor was the area of delamination, followed by thickness. Similarly, Farrag et al. [102] showed that thickness was the geometric aspect with the least effect on thermal contrast. These results indicate that IRT is relatively robust to the effect of delamination thickness.

6.2.5. Material

Delamination is usually filled with air; thus, the thermal properties of delamination are considered to resemble the air. However, making air-filled delamination in a concrete specimen with a predetermined size and depth is not easy, except for delamination beneath tiles. To simulate delamination, materials with low thermal conductivity are embedded in specimens. Thus, several studies have examined the effect of embedded materials [54,82,102,132,141,168]. For example, Yehia et al. [132] maintained that air-filled delamination was more visible than delamination simulated with polyethylene foam. Contrary to this, Cotič et al. [82] mentioned no significant difference between thermal contrasts above polystyrene foam and air-filled void. Although the results of these studies are not consistent enough, polyethylene foam is generally used as the material to simulate delamination. The reason may be that the difference between the thermal conductivity of polyethylene foam (0.033–0.045 W/mK [187]) and air (0.022 W/mK) is negligible for that of concrete (1.6–2.1 W/mK [139,157]). Therefore, the results of IRT by polyethylene foam could be applied to the actual delamination.

6.3. Target Object

6.3.1. Thermal Property

The materials of the target object affect thermal contrast because heat flow is determined by thermophysical properties of the materials: thermal conductivity, specific heat capacity, and density. The properties of concrete change depending on compression strength and mix proportions. For example, Rocha et al. [103] and Farrag et al. [102] stated that concrete with a low water-to-cement ratio or high strength concrete generated high thermal contrast in outdoor tests because of high thermal conductivity and high density. On the other hand, Maierhofer et al. [143] mentioned that thermal contrast decreased slightly along with the concrete strength increase. Additionally, the authors showed that density significantly affected thermal contrast, while thermal conductivity had minor effects. As mentioned above, there are some debates about the effects of materials on thermal contrast.

Building walls are generally layered with different materials rather than the single material of concrete. When finish materials are the same, substrate materials can also affect thermal contrast. Lourenço et al. [162] examined an external thermal insulating composite system (ETICS) and brick masonry with tile finish. In addition, Meola [141] investigated marble, brick, and tuff with render finish. These studies indicate that subsurface materials with high thermal conductivity generate high thermal contrast. The reason can be that substrate material with high conductivity increase the ratio of the difference in thermal transmission coefficients between sound area and delamination area. This means that delamination becomes difficult to be detected in the order of concrete, bricks and insulation in substrates.

6.3.2. Others

Other factors related to target objects investigated by the previous studies include rebars, water penetration, and surface conditions. The effects of these factors might be inevitable when inspecting existing infrastructures and buildings.

Rebars are usually embedded in concrete parallel to the surface to reinforce concrete structures. Rebars have a high thermal conductivity of 12.5 W/mK, much higher than concrete of 1.8 W/mK. Therefore, rebars may diffuse heat flow parallel to the surface,

resulting in low thermal contrast. According to laboratory tests and outdoor tests, the effect of rebars is different depending on the relative position of rebars and delamination [85,102,143,147,167]. When delamination occurs between rebars and the surface, rebars have little impact on thermal contrast and detectability [85,167]. On the other hand, when delamination occurs deeper than rebars, the effect is not consistent enough between previous studies. Scott et al. [147] stated no differences in thermal contrast between the presence and absence of rebars. In contrast, Huh et al. [85] argued that delamination indicated significantly lower SNR than delamination above rebars; thus, the delamination under rebars was not easy to be detected. Moreover, rebar density also affects thermal contrast. Maierhofer et al. [143] pointed out that high rebar density slightly decreased thermal contrast. In addition, rebars can influence response time. Tran et al. [167] revealed that rebars above delamination shortened observation time; thus, the depth of delamination may be estimated to be shallower than the actual depth. Overall, the effects of rebars on thermal contrast rely on the relative position between delamination and rebars.

Water penetration or high moisture content in a target object generates nonuniform temperature distribution on a surface due to three physical phenomena: evaporative cooling [25,162,177], the increase in specific heat capacity of the object [175,188], and the increase in thermal conductivity of the object [7,26]. Water penetration may occur at the same time as delamination in target objects. Edis et al. [188] surveyed glazed tile façade buildings and stated that both delamination and high moisture content areas had positive thermal contrast in midafternoon (e.g., 4:30 pm) under sunlight exposure conditions. However, water penetration into a delamination cavity may cause negative effects on thermal contrast and detectability. Lourenço et al. [162] conducted the outdoor tests in which water was poured into the back of nonadhesive tiles. Water penetration created opposite thermal behavior to delamination and decreased thermal contrast. Similarly, Güray et al. [166] stated that water-filled delamination could not be detected at any time. To address the issue caused by water penetration, Lourenço et al. [162] proposed inspecting target objects in different conditions: after rainy days and under dry conditions. Therefore, since water penetration could generate thermal contrast or reduce detectability, IRT surveys after rain or under wet conditions should be avoided.

Surface conditions, such as color and obstacles on the surface, affect IR images. Building facades are generally colored with paint or colorful materials. Lourenço et al. [162] studied the effects of surface color using white and black tiles. Black color, which absorbs a large amount of solar irradiation, contributed to high thermal contrast during the heating and cooling cycle. Thus, surface color affects detectability in sunny areas, and dark colors are advantageous for IRT.

The surfaces of infrastructures and buildings are not always clean and may have small obstacles. Hiasa et al. [90] stated that the obstacles could be discriminated on IR images because obstacles were smaller than a deck surface and quickly heated up and cooled down. The authors also suggested that visual images could help to distinguish obstacles certainly. To complement the information of IR images with visual images, simultaneously capturing IR and visual images is recommended.

6.4. IR Camera

6.4.1. IR Camera Type

Two types of IR cameras are generally used for IRT: a short-wavelength (SW) camera and a long-wavelength (LW) camera [79]. Table 5 indicates the characteristics of types of IR cameras. SW and LW cameras can detect infrared rays in the high atmospheric transmission band of 3–5 μm and 8–14 μm, respectively, known as atmospheric windows [189]. This difference in the band creates the characteristics of these cameras.

Table 5. Characteristics of types of IR camera.

Items	Short-Wavelength (SW) Camera	Long-Wavelength (LW) Camera
Spectral range	3–5 μm	8–14 μm
Detector type	InSb, Quantum detector	Microbolometer, Thermal detector
Cooling	Cooling	Uncooling
Thermal sensitivity, NETD	Fine	Middle
Shutter speed	Fast (e.g., 10 μs–10 ms)	Slow (e.g., 10 ms)
Camera cost	High	Low–middle

SW cameras use a cooled quantum detector sensitive to high-energy emissions from hot objects [79]. Thus, the quality of IR images is high when a target object is at a high temperature. In contrast, SW cameras are not suitable for measurements at a low temperature below 10 °C [96]. Additionally, the cameras are less affected by reflections of surrounding buildings or the sky on glazed facades [96]. The disadvantage is that the cameras are susceptible to solar reflections on the surface. Therefore, SW cameras tend to be used at night [156].

LW cameras use an uncooled microbolometer detector sensitive to low-energy emissions. Thus, the quality of IR images is relatively high when a target is at a low temperature. In addition, LW cameras are less subject to solar reflections on surfaces. In contrast, the cameras are susceptible to reflections of surrounding buildings and the sky on glazed tiles or smooth surfaces [14,96,97,160]. Hence, LW cameras are often used for daytime measurements [156]. Currently, many LW cameras are being developed, including affordable models [160] and lightweight models for unmanned aerial vehicles [56,190].

Regarding the influence of IR camera type and model, Hiasa et al. [156,159,160] compared two LW cameras and an SW camera, and Bauer et al. [14] examined two LW cameras of different manufacturers. Although different IR cameras output different temperature values even for the same object, there were no significant differences in thermal contrast and detectability. Therefore, selecting the type of IR camera is advisable according to the type of surrounding noise.

6.4.2. Distance and Spatial Resolution

A short distance from the IR camera to the target object is considered ideal [27,137,160,162]; however, surveys at short distances are not always possible due to the limited accessibility of existing infrastructures. The distance can affect detectability in three aspects: atmospheric attenuation, captured area, and spatial resolution.

Atmospheric attenuation is a phenomenon in which water vapor and carbon dioxide in the atmosphere absorb IR [191]. Due to the low impact of atmospheric attenuation, short distance measurements can provide accurate temperature values with few errors [27]. Furthermore, the effect of distance on detectability depends on the camera types because the atmospheric attenuation relies on spectral ranges [191]. For example, Hiasa et al. [160] mentioned that LW cameras were relatively affected by distance, while SW cameras were less affected. However, both cameras appropriately captured thermal contrasts, which are important to detect delamination. Overall, it is considered that distances within 10 m have little impact on detectability [137,177,192].

The size of a captured area may influence the efficiency of IRT surveys and detectability. The area captured is determined by an IR camera's field of view (FOV) and distance. FOV indicates the largest area that an IR camera can capture, described in horizontal and vertical degrees, and is determined by the focal length and the detector size of the IR camera. A long-distance measurement can capture a large area at once and improve efficiency. However, this IR image tends to include surroundings or nontarget objects with high or low temperatures. IR cameras automatically adjust the span of the temperature color scale based on the maximum and minimum temperature in an IR image. Therefore, Lourenço et al. [162] stated that the surroundings and nontarget objects widened the scale of the

image, making it difficult to emphasize the slight thermal contrast of delamination in visual analysis. Thus, short-distance measurement is recommended.

To measure the temperature value of a small area accurately, at least a smaller spatial resolution than the area is required [79]. Spatial resolution refers to the physical size of a target object per pixel and is determined by the multiplication of instantaneous field of view (IFOV) and distance. IFOV is determined by FOV and sensor resolution (the number of pixels). Therefore, spatial resolution becomes large as the distance increases and the sensor resolution decreases. Hiasa et al. [160] mentioned that the IR camera with a small spatial resolution (sensor resolution is 640 $\times$ 512 pixels) had higher sensitivity for detecting delamination than that with a large spatial resolution (sensor resolution is 320 $\times$ 240 pixels) at the same distance. Thus, using the IR camera with high sensor resolution is one way to keep detectability for long-distance measurements. However, the sensor resolution of IR cameras is lower than that of visual cameras and is commonly limited to 640 $\times$ 512 pixels [55]. Hence, Scott et al. [147] suggested using a telescope lens of small IFOV for long-distance measurements to keep the spatial resolution. Selecting an appropriate distance, FOV, and sensor resolution is important for detecting small delamination.

As described above, distance is related to detectability from the aspects of atmospheric attenuation, captured area, and spatial resolution. It is desirable to capture IR images as close as possible while balancing productivity and limitation of accessibility.

6.4.3. Angle

An observation angle could affect temperature values measured by IR cameras and detectability. This is because the emissivity of objects relies on the angle with respect to the surface. In general, the emissivity of nonmetallic materials is stable from the angles of 0° to 45° and decreases at higher angles [193,194]. Several studies suggested that thermal contrast is stable when angles are within 45°, and delamination can be detected although measured temperature values might change [147,156,190]. Additionally, Ortiz et al. [190] noted that the angle of 0° should be avoided for glazed surfaces because an IR camera may capture the reflection of the inspector or the IR camera on the surface.

At angles above 45°, the detectability of IRT may decline because of thermal contrast reduction or reflection noise. Scott et al. [147] reported that only shallow delamination, which was high thermal contrast, could be detected at the angle of 80° in the outdoor test. Moreover, Ortiz et al. [190] argued that measurement errors increased sharply due to reflections from the sky and the sun. Although the survey with angles above 45° may detect delamination, the angle within 45° is desirable to keep the reliability of IRT.

6.4.4. Platform

When surveying a wide area, mounting an IR camera on a platform can enhance the IRT survey's efficiency compared to by hand. For example, in the bridge deck inspection, an IR camera fixed to the top of a car continuously captures a road lane [195]. However, IR images captured on vehicles may be blurry or low quality due to the effects of moving speed or vibration. Thus, ASTM [47] limits the speed to 16 km/h or less. To survey with normal car speed without closing road lanes, Hiasa et al. [160,161] have examined the effects of speed on IR images using the two types of IR cameras. As a result, the SW camera with fast shutter speed could acquire IR images with high quality at 48 or 64 km/h, whereas the LW cameras with slow shutter speed captured blur IR images. Hence, measurement at high-speed movement requires SW cameras.

Recent developments in robotics allow inspectors to use unmanned aerial vehicles (UAVs) as a platform to inspect infrastructures and buildings [64,196,197]. UAVs with IR cameras can access any location without scaffolds and efficiently capture IR images at appropriate distances and angles [165,198,199]. Some studies have compared UAVs with traditional platforms, a tripod or cart, in outdoor tests using LW cameras [56,155]. As a

result, mounting the camera on UAVs has little effect on the quality of IR images at a resting state or slow speeds.

7. Conclusions

Capturing latent defects at the early stage of delamination even before delaminated objects falling is essential for integral components of infrastructures and buildings. With this in mind, a comprehensive review on the use of IRT to detect delamination on infrastructures and buildings was presented.

Three classifications of IRT for assessing defects were explained to clarify the methodologies used in delamination detection. Regarding delamination detection, the principle, evaluation protocols with one-time and time-series data analysis, and standards and guidelines were consolidated. Additionally, the performance of IRT in detecting delamination was compared with that of other NDTs.

Experimental methodologies employed by studies over the last 20 years on factors affecting delamination detection were discussed. Furthermore, the impact of factors on detectability was also investigated. Factors studied include environmental conditions, delamination properties, target objects, and IR cameras. Although the results of the studies were not always consistent due to the differences in experimental conditions, general desirable conditions for IRT are summarized below:

- Suitable time windows for the inspection depend on the direction of the inspection surface and delamination depth. For shallow delamination on a horizontal surface or south elevation, the windows are noon to early afternoon and late evening to early night.
- A large amount of total solar irradiation is desirable because irradiation is the primary heat source to generate thermal contrast.
- High daily ambient temperature change allows IRT even in shaded areas.
- A low wind velocity is preferable in sunny areas.
- Fine weather is optimum for the heating and cooling cycles because of solar irradiation, high daily ambient temperature changes, and radiative cooling.
- Delamination of large size has high thermal contrast and is easy to detect.
- The detectable depth of delamination is greatly affected by environmental conditions. Delamination of at least 3–5 cm or less could be detected in outdoor conditions.
- The width to depth ratio (WDTR) of delamination also affects detectability. The WTDR criteria of detectable delamination are 1.25 under laboratory conditions and 2–2.5 under outdoor conditions.
- The target object with high thermal conductivity has high thermal contrast, and the detectability is low on the insulation walls or low-strength concrete.
- Water penetration into delamination causes the opposite behavior of the thermal contrast of delamination.
- Dark color surfaces in sunny areas are advantageous for inspection.
- The influence of obstacles on the surface can be removed by complementing IR images with visual images.
- Both types of SW cameras and LW cameras can be used for inspection. An appropriate type should be selected according to the noise of the surrounding environment.
- The close distance from an IR camera to a target object is desirable in terms of atmospheric attenuation, captured area, and spatial resolution while balancing productivity and limitation of accessibility.
- When IR camera platforms, such as vehicles or UAVs, move quickly, SW cameras can collect clear IR images compared with LW cameras.

The results of this study could be used as the benchmarks for setting standardized testing criteria, as well as for comparison of results for future works on the use of infrared thermography for detection of delamination on infrastructures and buildings.

Author Contributions: Conceptualization, K.T. and M.Y.L.C.; methodology, K.T.; investigation K.T.; formal analysis, K.T.; resources, K.T.; data curation, K.T.; writing—original draft preparation, K.T.; writing—review and editing, K.T. and M.Y.L.C.; supervision M.Y.L.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. American Road and Transportation Builders Association. Bridge Conditions Report 2021. Available online: https://artbabridgereport.org/ (accessed on 8 October 2021).
2. Chew, M.Y.L. Façade Inspection for Falling Objects from Tall Buildings in Singapore. *Int. J. Build. Pathol. Adapt.* **2021**. ahead-of-print. [CrossRef]
3. U.S. Department of Transportation. Highway Bridge Inspections. Available online: https://www.transportation.gov/testimony/highway-bridge-inspections (accessed on 8 October 2021).
4. Kimura, T.; Zhang, T.; Fukuda, H. A Proposal for the Development of a Building Management System for Extending the Lifespan of Housing Complexes in Japan. *Sustainability* **2019**, *11*, 5622. [CrossRef]
5. Hong Kong Government Buildings Department. Mandatory Building Inspection Scheme. Available online: https://www.bd.gov.hk/en/safety-inspection/mbis/index.html (accessed on 8 October 2021).
6. Mohammadi, J. Building Facade Inspection Process: Administration and Management Matters. *Pract. Period. Struct. Des. Constr.* **2021**, *26*, 02521001. [CrossRef]
7. *BS EN 13187*; Thermal Performance of Buildings. Qualitative Detection of Thermal Irregularities in Building Envelopes. Infrared Method. British Standards Institute: London, UK, 1999.
8. Chew, M.Y.L.; De Silva, N. Factorial Method for Performance Assessment of Building Facades. *J. Constr. Eng. Manag.* **2004**, *130*, 525–533. [CrossRef]
9. Lourenço, T.; Matias, L.; Faria, P. Anomaly Diagnosis in Ceramic Claddings by Thermography—A Review. In Proceedings of the 7th International Conference on Safety and Durability of Structures (ICOSADOS 2016), Vila Real, Portugal, 10–12 May 2016; pp. 1–8.
10. Tran, Q.H.; Huh, J.; Kang, C.; Lee, B.Y.; Kim, I.-T.; Ahn, J.-H. Detectability of Subsurface Defects with Different Width-to-Depth Ratios in Concrete Structures Using Pulsed Thermography. *J. Nondestruct. Eval.* **2018**, *37*, 32. [CrossRef]
11. Chew, M.Y.L. The Study of Adhesion Failure of Wall Tiles. *Build. Environ.* **1992**, *27*, 493–499. [CrossRef]
12. Chew, M.Y.L.; Tan, S.S.; Kang, K.H. Contribution Analysis of Maintainability Factors for Cladding Facades. *Archit. Sci. Rev.* **2005**, *48*, 215–227. [CrossRef]
13. Chew, M.Y.L.; Tan, S.S.; Kang, K.H. A Technical Evaluation Index for Curtain Wall and Cladding Facades. *Struct. Surv.* **2004**, *22*, 210–227. [CrossRef]
14. Bauer, E.; de Freitas, V.P.; Mustelier, N.; Barreira, E.; de Freitas, S.S. Infrared Thermography—Evaluation of the Results Reproducibility. *Struct. Surv.* **2015**, *33*, 20–35. [CrossRef]
15. El Masri, Y.; Rakha, T. A Scoping Review of Non-Destructive Testing (NDT) Techniques in Building Performance Diagnostic Inspections. *Constr. Build. Mater.* **2020**, *265*, 120542. [CrossRef]
16. Dwivedi, S.K.; Vishwakarma, M.; Soni, P.A. Advances and Researches on Non Destructive Testing: A Review. *Mater. Today Proc.* **2018**, *5*, 3690–3698. [CrossRef]
17. McCann, D.M.; Forde, M.C. Review of NDT Methods in the Assessment of Concrete and Masonry Structures. *NDT E Int.* **2001**, *34*, 71–84. [CrossRef]
18. Omar, T.; Nehdi, M. Condition Assessment of Reinforced Concrete Bridges: Current Practice and Research Challenges. *Infrastructures* **2018**, *3*, 36. [CrossRef]
19. National Academies of Sciences, Engineering, and Medicine. *Nondestructive Testing to Identify Concrete Bridge Deck Deterioration*; The National Academies Press: Washington, DC, USA, 2012. [CrossRef]
20. Edis, E.; Flores-Colen, I.; De Brito, J. Building Thermography: Detection of Delamination of Adhered Ceramic Claddings Using the Passive Approach. *J. Nondestruct. Eval.* **2015**, *34*, 1–13. [CrossRef]
21. Kylili, A.; Fokaides, P.A.; Christou, P.; Kalogirou, S.A. Infrared Thermography (IRT) Applications for Building Diagnostics: A Review. *Appl. Energy* **2014**, *134*, 531–549. [CrossRef]
22. Garrido, I.; Lagüela, S.; Arias, P. Infrared Thermography's Application to Infrastructure Inspections. *Infrastructures* **2018**, *3*, 35. [CrossRef]

23. Usamentiaga, R.; Venegas, P.; Guerediaga, J.; Vega, L.; Molleda, J.; Bulnes, F.G. Infrared Thermography for Temperature Measurement and Non-Destructive Testing. *Sensors* **2014**, *14*, 12305–12348. [CrossRef] [PubMed]
24. Pozzer, S.; Pravia, Z.M.C.; Rezazadeh Azar, E.; Dalla Rosa, F. Statistical Analysis of Favorable Conditions for Thermographic Inspection of Concrete Slabs. *J. Civ. Struct. Health Monit.* **2020**, *10*, 609–626. [CrossRef]
25. Grinzato, E.; Ludwig, N.; Cadelano, G.; Bertucci, M.; Gargano, M.; Bison, P. Infrared Thermography for Moisture Detection: A Laboratory Study and in-Situ Test. *Mater. Eval.* **2011**, *69*, 97–104.
26. *ASTM C1153-10*; Standard Practice for Location of Wet Insulation in Roofing Systems Using Infrared Imaging. ASTM Internatinal: West Conshohocken, PA, USA, 2015. [CrossRef]
27. Barreira, E.; Almeida, R.M. *Infrared Thermography for Building Moisture Inspection*; Springer: Cham, Switzerland, 2019. [CrossRef]
28. Garrido, I.; Solla, M.; Lagüela, S.; Fernández, N. IRT and GPR Techniques for Moisture Detection and Characterisation in Buildings. *Sensors* **2020**, *20*, 6421. [CrossRef]
29. Ruiz Valero, L.; Flores Sasso, V.; Prieto Vicioso, E. In Situ Assessment of Superficial Moisture Condition in Façades of Historic Building Using Non-Destructive Techniques. *Case Stud. Constr. Mater.* **2019**, *10*, e00228. [CrossRef]
30. Avdelidis, N.P.; Moropoulou, A.; Theoulakis, P. Detection of Water Deposits and Movement in Porous Materials by Infrared Imaging. *Infrared Phys. Technol.* **2003**, *44*, 183–190. [CrossRef]
31. Barreira, E.; Almeida, R.M.S.F. Drying Evaluation Using Infrared Thermography. *Energy Procedia* **2015**, *78*, 170–175. [CrossRef]
32. Fox, M.; Goodhew, S.; De Wilde, P. Building Defect Detection: External Versus Internal Thermography. *Build. Environ.* **2016**, *105*, 317–331. [CrossRef]
33. *ASTM C1060-11a*; Standard Practice for Thermographic Inspection of Insulation Installations in Envelope Cavities of Frame Buildings. ASTM Internatinal: West Conshohocken, PA, USA, 2015. [CrossRef]
34. Tejedor, B.; Barreira, E.; Almeida, R.M.S.F.; Casals, M. Thermographic 2d U-Value Map for Quantifying Thermal Bridges in Building Façades. *Energy Build.* **2020**, *224*, 110176. [CrossRef]
35. Taylor, T.; Counsell, J.; Gill, S. Energy Efficiency Is More Than Skin Deep: Improving Construction Quality Control in New-Build Housing Using Thermography. *Energy Build.* **2013**, *66*, 222–231. [CrossRef]
36. Ibarra-Castanedo, C.; Sfarra, S.; Klein, M.; Maldague, X. Solar Loading Thermography: Time-Lapsed Thermographic Survey and Advanced Thermographic Signal Processing for the Inspection of Civil Engineering and Cultural Heritage Structures. *Infrared Phys. Technol.* **2017**, *82*, 56–74. [CrossRef]
37. Bauer, E.; Milhomem, P.M.; Aidar, L.A.G. Evaluating the Damage Degree of Cracking in Facades Using Infrared Thermography. *J. Civ. Struct. Health Monit.* **2018**, *8*, 517–528. [CrossRef]
38. Jo, J.; Jadidi, Z. A High Precision Crack Classification System Using Multi-Layered Image Processing and Deep Belief Learning. *Struct. Infrastruct. Eng.* **2020**, *16*, 297–305. [CrossRef]
39. Aggelis, D.G.; Kordatos, E.Z.; Strantza, M.; Soulioti, D.V.; Matikas, T.E. NDT Approach for Characterization of Subsurface Cracks in Concrete. *Constr. Build. Mater.* **2011**, *25*, 3089–3097. [CrossRef]
40. Dufour, M.B.; Derome, D.; Zmeureanu, R. Analysis of Thermograms for the Estimation of Dimensions of Cracks in Building Envelope. *Infrared Phys. Technol.* **2009**, *52*, 70–78. [CrossRef]
41. Barreira, E.; Almeida, R.M.S.F.; Moreira, M. An Infrared Thermography Passive Approach to Assess the Effect of Leakage Points in Buildings. *Energy Build.* **2017**, *140*, 224–235. [CrossRef]
42. Lerma, J.L.; Cabrelles, M.; Portalés, C. Multitemporal Thermal Analysis to Detect Moisture on a Building Façade. *Constr. Build. Mater.* **2011**, *25*, 2190–2197. [CrossRef]
43. Moropoulou, A.; Avdelidis, N.; Karoglou, M.; Delegou, E.; Alexakis, E.; Keramidas, V. Multispectral Applications of Infrared Thermography in the Diagnosis and Protection of Built Cultural Heritage. *Appl. Sci.* **2018**, *8*, 284. [CrossRef]
44. Theodorakeas, P.; Avdelidis, N.P.; Cheilakou, E.; Koui, M. Quantitative Analysis of Plastered Mosaics by Means of Active Infrared Thermography. *Constr. Build. Mater.* **2014**, *73*, 417–425. [CrossRef]
45. Yousefi, B.; Sfarra, S.; Ibarra-Castanedo, C.; Avdelidis, N.P.; Maldague, X.P.V. Thermography Data Fusion and Nonnegative Matrix Factorization for the Evaluation of Cultural Heritage Objects and Buildings. *J. Therm. Anal. Calorim.* **2018**, *136*, 943–955. [CrossRef]
46. Sfarra, S.; Yao, Y.; Zhang, H.; Perilli, S.; Scozzafava, M.; Avdelidis, N.P.; Maldague, X.P.V. Precious Walls Built in Indoor Environments Inspected Numerically and Experimentally within Long-Wave Infrared (Lwir) and Radio Regions. *J. Therm. Anal. Calorim.* **2019**, *137*, 1083–1111. [CrossRef]
47. *ASTM D4788-03*; Standard Test Method for Detecting Delaminations in Bridge Decks Using Infrared Thermography. ASTM Internatinal: West Conshohocken, PA, USA, 2013. [CrossRef]
48. Residential Energy Services Network. *Resnet Interim Guidelines for Thermographic Inspections of Buildings*; Residential Energy Services Network: Oceanside, CA, USA, 2010.
49. Rumbayan, R.; Washer, G.A. Modeling of Environmental Effects on Thermal Detection of Subsurface Damage in Concrete. *Res. Nondestruct. Eval.* **2014**, *25*, 235–252. [CrossRef]
50. Kirimtat, A.; Krejcar, O. A Review of Infrared Thermography for the Investigation of Building Envelopes: Advances and Prospects. *Energy Build.* **2018**, *176*, 390–406. [CrossRef]
51. Milovanović, B.; Banjad Pečur, I. Review of Active IR Thermography for Detection and Characterization of Defects in Reinforced Concrete. *J. Imaging* **2016**, *2*, 11. [CrossRef]

52. Du, B.; He, Y.; He, Y.; Zhang, C. Progress and Trends in Fault Diagnosis for Renewable and Sustainable Energy System Based on Infrared Thermography: A Review. *Infrared Phys. Technol.* **2020**, *109*, 103383. [CrossRef]
53. Lehmann, B.; Ghazi Wakili, K.; Frank, T.; Vera Collado, B.; Tanner, C. Effects of Individual Climatic Parameters on the Infrared Thermography of Buildings. *Appl. Energy* **2013**, *110*, 29–43. [CrossRef]
54. Meola, C. Infrared Thermography of Masonry Structures. *Infrared Phys. Technol.* **2007**, *49*, 228–233. [CrossRef]
55. Garrido, I.; Lagüela, S.; Otero, R.; Arias, P. Thermographic Methodologies Used in Infrastructure Inspection: A Review—Data Acquisition Procedures. *Infrared Phys. Technol.* **2020**, *111*, 103481. [CrossRef]
56. Mac, V.H.; Tran, Q.H.; Huh, J.; Doan, N.S.; Kang, C.; Han, D. Detection of Delamination with Various Width-to-Depth Ratios in Concrete Bridge Deck Using Passive IRT: Limits and Applicability. *Materials* **2019**, *12*, 3996. [CrossRef]
57. Lahiri, B.B.; Bagavathiappan, S.; Jayakumar, T.; Philip, J. Medical Applications of Infrared Thermography: A Review. *Infrared Phys. Technol.* **2012**, *55*, 221–235. [CrossRef] [PubMed]
58. Ciampa, F.; Mahmoodi, P.; Pinto, F.; Meo, M. Recent Advances in Active Infrared Thermography for Non-Destructive Testing of Aerospace Components. *Sensors* **2018**, *18*, 609. [CrossRef] [PubMed]
59. Li, L.; Zhang, Q.; Huang, D. A Review of Imaging Techniques for Plant Phenotyping. *Sensors* **2014**, *14*, 20078–20111. [CrossRef] [PubMed]
60. Bagavathiappan, S.; Lahiri, B.B.; Saravanan, T.; Philip, J.; Jayakumar, T. Infrared Thermography for Condition Monitoring—A Review. *Infrared Phys. Technol.* **2013**, *60*, 35–55. [CrossRef]
61. Sarawade, A.A.; Charniya, N.N. Infrared Thermography and Its Applications: A Review. In Proceedings of the International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 15–16 October 2018; pp. 280–285. [CrossRef]
62. Alfredo Osornio-Rios, R.; Antonino-Daviu, J.A.; de Jesus Romero-Troncoso, R. Recent Industrial Applications of Infrared Thermography: A Review. *IEEE Trans. Ind. Inform.* **2019**, *15*, 615–625. [CrossRef]
63. Doshvarpassand, S.; Wu, C.; Wang, X. An Overview of Corrosion Defect Characterization Using Active Infrared Thermography. *Infrared Phys. Technol.* **2019**, *96*, 366–389. [CrossRef]
64. Gallardo-Saavedra, S.; Hernández-Callejo, L.; Duque-Perez, O. Technological Review of the Instrumentation Used in Aerial Thermographic Inspection of Photovoltaic Plants. *Renew. Sustain. Energy Rev.* **2018**, *93*, 566–579. [CrossRef]
65. Yumnam, M.; Gupta, H.; Ghosh, D.; Jaganathan, J. Inspection of Concrete Structures Externally Reinforced with Frp Composites Using Active Infrared Thermography: A Review. *Constr. Build. Mater.* **2021**, *310*, 125265. [CrossRef]
66. Yang, R.; He, Y. Optically and Non-Optically Excited Thermography for Composites: A Review. *Infrared Phys. Technol.* **2016**, *75*, 26–50. [CrossRef]
67. Vergani, L.; Colombo, C.; Libonati, F. A Review of Thermographic Techniques for Damage Investigation in Composites. *Frat. Ed Integritá Strutt.* **2013**, *8*, 1–12. [CrossRef]
68. Mercuri, F.; Orazi, N.; Paoloni, S.; Cicero, C.; Zammit, U. Pulsed Thermography Applied to the Study of Cultural Heritage. *Appl. Sci.* **2017**, *7*, 1010. [CrossRef]
69. Orazi, N. The Study of Artistic Bronzes by Infrared Thermography: A Review. *J. Cult. Herit.* **2020**, *42*, 280–289. [CrossRef]
70. Lucchi, E. Applications of the Infrared Thermography in the Energy Audit of Buildings: A Review. *Renew. Sustain. Energy Rev.* **2018**, *82*, 3077–3090. [CrossRef]
71. Nardi, I.; Lucchi, E.; de Rubeis, T.; Ambrosini, D. Quantification of Heat Energy Losses through the Building Envelope: A State-of-the-art Analysis with Critical and Comprehensive Review on Infrared Thermography. *Build. Environ.* **2018**, *146*, 190–205. [CrossRef]
72. Bienvenido-Huertas, D.; Moyano, J.; Marín, D.; Fresco-Contreras, R. Review of in Situ Methods for Assessing the Thermal Transmittance of Walls. *Renew. Sustain. Energy Rev.* **2019**, *102*, 356–371. [CrossRef]
73. Sirca, G.F., Jr.; Adeli, H. Infrared Thermography for Detecting Defects in Concrete Structures. *J. Civ. Eng. Manag.* **2018**, *24*, 508–515. [CrossRef]
74. Garrido, I.; Lagüela, S.; Otero, R.; Arias, P. Thermographic Methodologies Used in Infrastructure Inspection: A Review—Post-Processing Procedures. *Appl. Energy* **2020**, *266*, 114857. [CrossRef]
75. Fox, M.; Coley, D.; Goodhew, S.; de Wilde, P. Thermography Methodologies for Detecting Energy Related Building Defects. *Renew. Sustain. Energy Rev.* **2014**, *40*, 296–310. [CrossRef]
76. FLIR. User's Manual, Flir Ex Series. Available online: https://www.flir-direct.com/pdfs/cache/www.flir-direct.com/e5/manual/e5-manual.pdf (accessed on 8 October 2021).
77. Fokaides, P.A.; Kalogirou, S.A. Application of Infrared Thermography for the Determination of the Overall Heat Transfer Coefficient (U-Value) in Building Envelopes. *Appl. Energy* **2011**, *88*, 4358–4365. [CrossRef]
78. Avdelidis, N.P.; Moropoulou, A. Emissivity Considerations in Building Thermography. *Energy Build.* **2003**, *35*, 663–667. [CrossRef]
79. Walker, N.J. *Infrared Thermography Handbook—Volume 1. Principles and Practice*; The British Institute of Non-Destructive Testing: Northampton, UK, 2004.
80. Lo, T.Y.; Choi, K.T.W. Building Defects Diagnosis by Infrared Thermography. *Struct. Surv.* **2004**, *22*, 259–263. [CrossRef]
81. Zhang, H.; Yang, R.; He, Y.; Foudazi, A.; Cheng, L.; Tian, G. A Review of Microwave Thermography Nondestructive Testing and Evaluation. *Sensors* **2017**, *17*, 1123. [CrossRef] [PubMed]
82. Cotič, P.; Kolarič, D.; Bosiljkov, V.; Bosiljkov, V.B.; Jagličić, Z. Determination of the Applicability and Limits of Void and Delamination Detection in Concrete Structures Using Infrared Thermography. *NDT E Int.* **2015**, *74*, 87–93. [CrossRef]

83. *BS ISO 9869-2*; Thermal Insulation. Building Elements. In-Situ Measurement of Thermal Resistance and Thermal Transmittance, Part 2: Infrared Method for Frame Structure Dwelling. British Standards Institute: London, UK, 2018.
84. Biddulph, P.; Gori, V.; Elwell, C.A.; Scott, C.; Rye, C.; Lowe, R.; Oreszczyn, T. Inferring the Thermal Resistance and Effective Thermal Mass of a Wall Using Frequent Temperature and Heat Flux Measurements. *Energy Build.* **2014**, *78*, 10–16. [CrossRef]
85. Huh, J.; Mac, V.; Tran, Q.; Lee, K.-Y.; Lee, J.-I.; Kang, C. Detectability of Delamination in Concrete Structure Using Active Infrared Thermography in Terms of Signal-to-Noise Ratio. *Appl. Sci.* **2018**, *8*, 1986. [CrossRef]
86. Barreira, E.; Almeida, R.M.S.F.; Simões, M.L.; Rebelo, D. Quantitative Infrared Thermography to Evaluate the Humidification of Lightweight Concrete. *Sensors* **2020**, *20*, 1664. [CrossRef]
87. Washer, G.; Fenwick, R.; Bolleni, N. *Development of Hand-Held Thermographic Inspection Technologies*; OR10-007; Department of Transportation: Jefferson City, MO, USA, 2009.
88. Hiasa, S.; Birgul, R.; Catbas, F.N. Effect of Defect Size on Subsurface Defect Detectability and Defect Depth Estimation for Concrete Structures by Infrared Thermography. *J. Nondestruct. Eval.* **2017**, *36*, 1–21. [CrossRef]
89. Kee, S.-H.; Oh, T.; Popovics, J.S.; Arndt, R.W.; Zhu, J. Nondestructive Bridge Deck Testing with Air-Coupled Impact-Echo and Infrared Thermography. *J. Bridge Eng.* **2012**, *17*, 928–939. [CrossRef]
90. Hiasa, S.; Birgul, R.; Matsumoto, M.; Necati Catbas, F. Experimental and Numerical Studies for Suitable Infrared Thermography Implementation on Concrete Bridge Decks. *Measurement* **2018**, *121*, 144–159. [CrossRef]
91. Qu, Z.; Jiang, P.; Zhang, W. Development and Application of Infrared Thermography Non-Destructive Testing Techniques. *Sensors* **2020**, *20*, 3851. [CrossRef] [PubMed]
92. Omar, T.; Nehdi, M.L.; Zayed, T. Infrared Thermography Model for Automated Detection of Delamination in RC Bridge Decks. *Constr. Build. Mater.* **2018**, *168*, 313–327. [CrossRef]
93. Balaras, C.A.; Argiriou, A.A. Infrared Thermography for Building Diagnostics. *Energy Build.* **2002**, *34*, 171–183. [CrossRef]
94. Cheng, C.; Na, R.; Shen, Z. Thermographic Laplacian-Pyramid Filtering to Enhance Delamination Detection in Concrete Structure. *Infrared Phys. Technol.* **2019**, *97*, 162–176. [CrossRef]
95. Clark, M.R.; McCann, D.M.; Forde, M.C. Application of Infrared Thermography to the Non-Destructive Testing of Concrete and Masonry Bridges. *NDT E Int.* **2003**, *36*, 265–275. [CrossRef]
96. Japan Building and Equipment Long-Life Cycle Association. *Periodic Inspection Manual for Exterior Tile Wall and Plaster Wall (Engl. Transl.)*, 3rd ed.; BELCA: Tokyo, Japan, 2009.
97. Edis, E.; Flores-Colen, I.; de Brito, J. Passive Thermographic Inspection of Adhered Ceramic Claddings: Limitation and Conditioning Factors. *J. Perform. Constr. Facil.* **2013**, *27*, 737–747. [CrossRef]
98. Rao, D.P. Infrared Thermography and Its Applications in Civil Engineering. *Indian Concr. J.* **2008**, *82*, 41–50.
99. *BS EN ISO 6781-3*; Performance of Buildings—Detection of Heat, Air and Moisture Irregularities in Buildings by Infrared Methods—Part 3: Qualifications of Equipment Operators, Data Analysts and Report Writers. British Standards Institute: London, UK, 2015.
100. Washer, G.; Fenwick, R.; Nelson, S.; Rumbayan, R. Guidelines for Thermographic Inspection of Concrete Bridge Components in Shaded Conditions. *Transp. Res. Rec.* **2013**, *2360*, 13–20. [CrossRef]
101. Janků, M.; Březina, I.; Grošek, J. Use of Infrared Thermography to Detect Defects on Concrete Bridges. *Procedia Eng.* **2017**, *190*, 62–69. [CrossRef]
102. Farrag, S.; Yehia, S.; Qaddoumi, N. Investigation of Mix-Variation Effect on Defect-Detection Ability Using Infrared Thermography as a Nondestructive Evaluation Technique. *J. Bridge Eng.* **2016**, *21*, 4015055. [CrossRef]
103. Rocha, J.H.A.; Póvoas, Y.V.; Santos, C.F. Detection of Delaminations in Sunlight-Unexposed Concrete Elements of Bridges Using Infrared Thermography. *J. Nondestruct. Eval.* **2018**, *38*, 8. [CrossRef]
104. Oh, T.; Kee, S.-H.; Arndt, R.W.; Popovics, J.S.; Zhu, J. Comparison of NDT Methods for Assessment of a Concrete Bridge Deck. *J. Eng. Mech.* **2013**, *139*, 305–314. [CrossRef]
105. Garrido, I.; Lagüela, S.; Sfarra, S.; Madruga, F.J.; Arias, P. Automatic Detection of Moistures in Different Construction Materials from Thermographic Images. *J. Therm. Anal. Calorim.* **2019**, *138*, 1649–1668. [CrossRef]
106. Omar, T.; Nehdi, M.L. Remote Sensing of Concrete Bridge Decks Using Unmanned Aerial Vehicle Infrared Thermography. *Autom. Constr.* **2017**, *83*, 360–371. [CrossRef]
107. Park, G.; Lee, M.; Jang, H.; Kim, C. Thermal Anomaly Detection in Walls via CNN-Based Segmentation. *Autom. Constr.* **2021**, *125*, 103627. [CrossRef]
108. Cheng, C.; Shang, Z.; Shen, Z. Bridge Deck Delamination Segmentation Based on Aerial Thermography through Regularized Grayscale Morphological Reconstruction and Gradient Statistics. *Infrared Phys. Technol.* **2019**, *98*, 240–249. [CrossRef]
109. Lai, W.W.-L.; Lee, K.-K.; Poon, C.-S. Validation of Size Estimation of Debonds in External Wall's Composite Finishes Via Passive Infrared Thermography and a Gradient Algorithm. *Constr. Build. Mater.* **2015**, *87*, 113–124. [CrossRef]
110. Cheng, C.; Shen, Z. The Application of Gray-Scale Level-Set Method in Segmentation of Concrete Deck Delamination Using Infrared Images. *Constr. Build. Mater.* **2020**, *240*, 117974. [CrossRef]
111. Edis, E.; Flores-Colen, I.; De Brito, J. Time-Dependent Passive Building Thermography for Detecting Delamination of Adhered Ceramic Cladding. *J. Nondestruct. Eval.* **2015**, *34*, 1–16. [CrossRef]
112. Milovanović, B.; Banjad Pečur, I.; Štirmer, N. The Methodology for Defect Quantification in Concrete Using Ir Thermography. *J. Civ. Eng. Manag.* **2017**, *23*, 573–582. [CrossRef]

113. Marinetti, S.; Grinzato, E.; Bison, P.G.; Bozzi, E.; Chimenti, M.; Pieri, G.; Salvetti, O. Statistical Analysis of IR Thermographic Sequences by PCA. *Infrared Phys. Technol.* **2004**, *46*, 85–91. [CrossRef]
114. Krankenhagen, R.; Maierhofer, C. Pulse Phase Thermography for Characterising Large Historical Building Façades after Solar Heating and Shadow Cast—A Case Study. *Quant. InfraRed Thermogr. J.* **2014**, *11*, 10–28. [CrossRef]
115. Marinetti, S.; Finesso, L.; Marsilio, E. Matrix Factorization Methods: Application to Thermal NDT/E. *NDT E Int.* **2006**, *39*, 611–616. [CrossRef]
116. Al Gharawi, M.; Adu-Gyamfi, Y.; Washer, G. A Framework for Automated Time-Lapse Thermography Data Processing. *Constr. Build. Mater.* **2019**, *227*, 116507. [CrossRef]
117. Japan Public Works Research Institute. *Inspection Manual for Prevention of Delamination and Spalling on Concrete Infrastructure by Infrared Thermography (Engl. Transl.)*; Japan Public Works Research Institute: Tokyo, Japan, 2005.
118. *NDIS 3428*; Non-Destructive Testing Method for Evaluation of Defects and Deteriorations in Surface Layer of Buildings and Structures by Infrared Thermography. Standard of the Japanese Society for Non-Destructive Inspection (NDIS): Tokyo, Japan, 2009.
119. Singleton, F. The Beaufort Scale of Winds—Its Relevance, and Its Use by Sailors. *Weather* **2008**, *63*, 37–41. [CrossRef]
120. *ASTM E1543-14*; Standard Practice for Noise Equivalent Temperature Difference of Thermal Imaging Systems. ASTM Internatinal: West Conshohocken, PA, USA, 2018. [CrossRef]
121. Xie, C.; Hazzard, E. Infrared Imaging for Inquiry-Based Learning. *Phys. Teach.* **2011**, *49*, 368–372. [CrossRef]
122. Kashif Ur Rehman, S.; Ibrahim, Z.; Memon, S.A.; Jameel, M. Nondestructive Test Methods for Concrete Bridges: A Review. *Constr. Build. Mater.* **2016**, *107*, 58–86. [CrossRef]
123. ACI Committee 228. *ACI PRC-228.2-13 Report on Nondestructive Test Methods for Evaluation of Concrete in Structures*; American Concrete Institute: Farmington Hills, MI, USA, 2013.
124. *ASTM D4580/D4580M-12(2018)*; Standard Practice for Measuring Delaminations in Concrete Bridge Decks by Sounding. ASTM Internatinal: West Conshohocken, PA, USA, 2018. [CrossRef]
125. Skłodowski, R.; Drdácký, M.; Skłodowski, M. Identifying Subsurface Detachment Defects by Acoustic Tracing. *NDT E Int.* **2013**, *56*, 56–64. [CrossRef]
126. Lin, S.; Meng, D.; Choi, H.; Shams, S.; Azari, H. Laboratory Assessment of Nine Methods for Nondestructive Evaluation of Concrete Bridge Decks with Overlays. *Constr. Build. Mater.* **2018**, *188*, 966–982. [CrossRef]
127. Dorafshan, S.; Azari, H. Evaluation of Bridge Decks with Overlays Using Impact Echo, a Deep Learning Approach. *Autom. Constr.* **2020**, *113*, 103133. [CrossRef]
128. Lagüela, S.; Solla, M.; Puente, I.; Prego, F.J. Joint Use of GPR, IRT and TLS Techniques for the Integral Damage Detection in Paving. *Constr. Build. Mater.* **2018**, *174*, 749–760. [CrossRef]
129. Sultan, A.A.; Washer, G.A. Comparison of Two Nondestructive Evaluation Technologies for the Condition Assessment of Bridge Decks. *Transp. Res. Rec.* **2018**, *2672*, 113–122. [CrossRef]
130. Dinh, K.; Gucunski, N. Factors Affecting the Detectability of Concrete Delamination in GPR Images. *Constr. Build. Mater.* **2021**, *274*, 121837. [CrossRef]
131. Vaghefi, K.; Ahlborn, T.M.; Harris, D.K.; Brooks, C.N. Combined Imaging Technologies for Concrete Bridge Deck Condition Assessment. *J. Perform. Constr. Facil.* **2015**, *29*, 04014102. [CrossRef]
132. Yehia, S.; Abudayyeh, O.; Nabulsi, S.; Abdelqader, I. Detection of Common Defects in Concrete Bridge Decks Using Nondestructive Evaluation Techniques. *J. Bridge Eng.* **2007**, *12*, 215–225. [CrossRef]
133. Cheng, C.-C.; Cheng, T.-M.; Chiang, C.-H. Defect Detection of Concrete Structures Using Both Infrared Thermography and Elastic Waves. *Autom. Constr.* **2008**, *18*, 87–92. [CrossRef]
134. Khan, F.; Bartoli, I. Detection of Delamination in Concrete Slabs Combining Infrared Thermography and Impact Echo Techniques: A Comparative Experimental Study. In Proceedings of the Structural Health Monitoring and Inspection of Advanced Materials, Aerospace, and Civil Infrastructure, San Diego, CA, USA, 8–12 March 2015; pp. 94370I-1–94370I-11. [CrossRef]
135. Janků, M.; Cikrle, P.; Grošek, J.; Anton, O.; Stryk, J. Comparison of Infrared Thermography, Ground-Penetrating Radar and Ultrasonic Pulse Echo for Detecting Delaminations in Concrete Bridges. *Constr. Build. Mater.* **2019**, *225*, 1098–1111. [CrossRef]
136. Cotič, P.; Jagličić, Z.; Niederleithinger, E.; Stoppel, M.; Bosiljkov, V. Image Fusion for Improved Detection of near-Surface Defects in NDT-Ce Using Unsupervised Clustering Methods. *J. Nondestruct. Eval.* **2014**, *33*, 384–397. [CrossRef]
137. Chew, M.Y.L. Assessing Building Façades Using Infra-Red Thermography. *Struct. Surv.* **1998**, *16*, 81–86. [CrossRef]
138. Maierhofer, C.; Brink, A.; Röllig, M.; Wiggenhauser, H. Transient Thermography for Structural Investigation of Concrete and Composites in the near Surface Region. *Infrared Phys. Technol.* **2002**, *43*, 271–278. [CrossRef]
139. Maierhofer, C.; Wiggenhauser, H.; Brink, A.; Röllig, M. Quantitative Numerical Analysis of Transient IR-Experiments on Buildings. *Infrared Phys. Technol.* **2004**, *46*, 173–180. [CrossRef]
140. Maierhofer, C.; Brink, A.; Röllig, M.; Wiggenhauser, H. Quantitative Impulse-Thermography as Non-Destructive Testing Method in Civil Engineering—Experimental Results and Numerical Simulations. *Constr. Build. Mater.* **2005**, *19*, 731–737. [CrossRef]
141. Meola, C.; Di Maio, R.; Roberti, N.; Carlomagno, G.M. Application of Infrared Thermography and Geophysical Methods for Defect Detection in Architectural Structures. *Eng. Fail. Anal.* **2005**, *12*, 875–892. [CrossRef]
142. Maierhofer, C.; Arndt, R.; Röllig, M.; Rieck, C.; Walther, A.; Scheel, H.; Hillemeier, B. Application of Impulse-Thermography for Non-Destructive Assessment of Concrete Structures. *Cem. Concr. Compos.* **2006**, *28*, 393–401. [CrossRef]

143. Maierhofer, C.; Arndt, R.; Röllig, M. Influence of Concrete Properties on the Detection of Voids with Impulse-Thermography. *Infrared Phys. Technol.* **2007**, *49*, 213–217. [CrossRef]
144. Washer, G.; Fenwick, R.; Bolleni, N.; Harper, J. Effects of Environmental Variables on Infrared Imaging of Subsurface Features of Concrete Bridges. *Transp. Res. Rec.* **2009**, *2108*, 107–114. [CrossRef]
145. Washer, G.; Fenwick, R.; Bolleni, N. Effects of Solar Loading on Infrared Imaging of Subsurface Features in Concrete. *J. Bridge Eng.* **2010**, *15*, 384–390. [CrossRef]
146. Washer, G.; Bolleni, N.; Fenwick, R. Thermographic Imaging of Subsurface Deterioration in Concrete Bridges. *Transp. Res. Rec.* **2010**, *2201*, 27–33. [CrossRef]
147. Scott, M.; Luttig, H.; Strydom, M.; Gonelli, M.; Kruger, D.; Rankine, R.; Broodryk, T. Passive Infrared Thermography as a Diagnostic Tool in Civil Engineering Structural Material Health Monitoring. In Proceedings of the Concrete Repair, Rehabilitation and Retrofitting III: 3rd International Conference on Concrete Repair, Rehabilitation and Retrofitting, Cape Town, South Africa, 3–5 September 2012; p. 274.
148. de Freitas, S.S.; de Freitas, V.P.; Barreira, E. Detection of Façade Plaster Detachments Using Infrared Thermography—A Nondestructive Technique. *Constr. Build. Mater.* **2014**, *70*, 80–87. [CrossRef]
149. Scott, M.; Kruger, D. Infrared Thermography as a Diagnostic Tool for Subsurface Assessments of Concrete Structures. In *Construction Materials and Structures*; IOS Press: Amsterdam, The Netherlands, 2014; pp. 904–909. [CrossRef]
150. Alfredo-Cruz, R.; Quintero-Ortiz, L.A.; Galán-Pinilla, C.A.; Espinosa-García, E.J. Non-Destructive Techniques in Concrete Elements for Bridges' Evalutation. *Rev. Fac. De Ing.* **2015**, *24*, 81. [CrossRef]
151. Khan, F.; Bolhassani, M.; Kontsos, A.; Hamid, A.; Bartoli, I. Modeling and Experimental Implementation of Infrared Thermography on Concrete Masonry Structures. *Infrared Phys. Technol.* **2015**, *69*, 228–237. [CrossRef]
152. Watase, A.; Birgul, R.; Hiasa, S.; Matsumoto, M.; Mitani, K.; Catbas, F.N. Practical Identification of Favorable Time Windows for Infrared Thermography for Concrete Bridge Evaluation. *Constr. Build. Mater.* **2015**, *101*, 1016–1030. [CrossRef]
153. Bauer, E.; Pavón, E.; Pereira, C.H.F.; Nascimento, M.L.M. Criteria for Identification of Ceramic Detachments in Building Facades with Infrared Thermography. In *Recent Developments in Building Diagnosis Techniques*; Delgado, J.M.P.Q., Ed.; Springer: Singapore, 2016; pp. 51–68. [CrossRef]
154. Bauer, E.; Pavón, E.; Barreira, E.; Kraus De Castro, E. Analysis of Building Facade Defects Using Infrared Thermography: Laboratory Studies. *J. Build. Eng.* **2016**, *6*, 93–104. [CrossRef]
155. Ellenberg, A.; Kontsos, A.; Moon, F.; Bartoli, I. Bridge Deck Delamination Identification from Unmanned Aerial Vehicle Infrared Imagery. *Autom. Constr.* **2016**, *72*, 155–165. [CrossRef]
156. Hiasa, S.; Birgul, R.; Catbas, F.N. Infrared Thermography for Civil Structural Assessment: Demonstrations with Laboratory and Field Studies. *J. Civ. Struct. Health Monit.* **2016**, *6*, 619–636. [CrossRef]
157. Huh, J.; Tran, Q.H.; Lee, J.-H.; Han, D.; Ahn, J.-H.; Yim, S. Experimental Study on Detection of Deterioration in Concrete Using Infrared Thermography Technique. *Adv. Mater. Sci. Eng.* **2016**, *2016*, 1053856. [CrossRef]
158. Chiang, L.-W.; Guo, S.-J. The Research of Infrared Thermal Image Diagnostic Model for Building External Wall Tiles. *Adv. Technol. Innov.* **2017**, *2*, 56–60.
159. Hiasa, S.; Birgul, R.; Catbas, F.N. Investigation of Effective Utilization of Infrared Thermography (IRT) through Advanced Finite Element Modeling. *Constr. Build. Mater.* **2017**, *150*, 295–309. [CrossRef]
160. Hiasa, S.; Necati Catbas, F.; Matsumoto, M.; Mitani, K. Considerations and Issues in the Utilization of Infrared Thermography for Concrete Bridge Inspection at Normal Driving Speeds. *J. Bridge Eng.* **2017**, *22*, 4017101. [CrossRef]
161. Hiasa, S.; Catbas, F.N.; Matsumoto, M.; Mitani, K. Monitoring Concrete Bridge Decks Using Infrared Thermography with High Speed Vehicles. *Struct. Monit. Maint.* **2016**, *3*, 277–296. [CrossRef]
162. Lourenço, T.; Matias, L.; Faria, P. Anomalies Detection in Adhesive Wall Tiling Systems by Infrared Thermography. *Constr. Build. Mater.* **2017**, *148*, 419–428. [CrossRef]
163. Sultan, A.A.; Washer, G. A Pixel-by-Pixel Reliability Analysis of Infrared Thermography (IRT) for the Detection of Subsurface Delamination. *NDT E Int.* **2017**, *92*, 177–186. [CrossRef]
164. Tran, Q.H.; Han, D.; Kang, C.; Haldar, A.; Huh, J. Effects of Ambient Temperature and Relative Humidity on Subsurface Defect Detection in Concrete Structures by Active Thermal Imaging. *Sensors* **2017**, *17*, 1718. [CrossRef]
165. Escobar-Wolf, R.; Oommen, T.; Brooks, C.N.; Dobson, R.J.; Ahlborn, T.M. Unmanned Aerial Vehicle (UAV)-Based Assessment of Concrete Bridge Deck Delamination Using Thermal and Visible Camera Sensors: A Preliminary Analysis. *Res. Nondestruct. Eval.* **2018**, *29*, 183–198. [CrossRef]
166. Güray, E.; Birgül, R. Determination of Favorable Time Window for Infrared Inspection by Numerical Simulation of Heat Propagation in Concrete. In Proceedings of the 3rd International Sustainable Buildings Symposium, Dubai, United Arab Emirates, 15–17 March 2018; pp. 577–591. [CrossRef]
167. Tran, Q.H.; Huh, J.; Mac, V.H.; Kang, C.; Han, D. Effects of Rebars on the Detectability of Subsurface Defects in Concrete Bridges Using Square Pulse Thermography. *NDT E Int.* **2018**, *100*, 92–100. [CrossRef]
168. Vyas, V.; Patil, V.J.; Singh, A.P.; Srivastava, A. Application of Infrared Thermography for Debonding Detection in Asphalt Pavements. *J. Civ. Struct. Health Monit.* **2019**, *9*, 325–337. [CrossRef]
169. Milovanovic, B.; Gasi, M.; Gumbarevic, S. Principal Component Thermography for Defect Detection in Concrete. *Sensors* **2020**, *20*, 3891. [CrossRef]

170. Raja, B.N.K.; Miramini, S.; Duffield, C.; Sofi, M.; Mendis, P.; Zhang, L. The Influence of Ambient Environmental Conditions in Detecting Bridge Concrete Deck Delamination Using Infrared Thermography (IRT). *Struct. Control. Health Monit.* **2020**, *27*, e2506. [CrossRef]
171. Cheng, C.; Shen, Z. Semi Real-Time Detection of Subsurface Consolidation Defects During Concrete Curing Stage. *Constr. Build. Mater.* **2021**, *270*, 121489. [CrossRef]
172. Mac, V.H.; Huh, J.; Doan, N.S.; Lee, B.Y.; Haldar, A. A Comprehensive Study on Identification of Both Deck and Soffit Defects in Concrete Bridge Decks through Thermographic Investigation of Shaded Side under Natural Conditions. *Constr. Build. Mater.* **2021**, *303*, 124452. [CrossRef]
173. Pozzer, S.; Dalla Rosa, F.; Pravia, Z.M.C.; Rezazadeh Azar, E.; Maldague, X. Long-Term Numerical Analysis of Subsurface Delamination Detection in Concrete Slabs Via Infrared Thermography. *Appl. Sci.* **2021**, *11*, 4323. [CrossRef]
174. Zheng, D.; Tan, S.; Li, X.; Cai, H. Research on the Infrared Thermographic Detection of Concrete under Solar Heating. *Adv. Civ. Eng.* **2021**, *2021*, 6692729. [CrossRef]
175. Edis, E.; Flores-Colen, I.; de Brito, J. Quasi-Quantitative Infrared Thermographic Detection of Moisture Variation in Facades with Adhered Ceramic Cladding Using Principal Component Analysis. *Build. Environ.* **2015**, *94*, 97–108. [CrossRef]
176. Li, Z.; Yao, W.; Lee, S.; Lee, C.; Yang, Z. Application of Infrared Thermography Technique in Building Finish Evaluation. *J. Nondestruct. Eval.* **2000**, *19*, 11–19. [CrossRef]
177. Barreira, E.; Almeida, R.M.S.F.; Delgado, J.M.P.Q. Infrared Thermography for Assessing Moisture Related Phenomena in Building Components. *Constr. Build. Mater.* **2016**, *110*, 251–269. [CrossRef]
178. Kumar, S.; Mullick, S.C. Wind Heat Transfer Coefficient in Solar Collectors in Outdoor Conditions. *Sol. Energy* **2010**, *84*, 956–963. [CrossRef]
179. Zhang, J.; Gupta, A.; Baker, J. Effect of Relative Humidity on the Prediction of Natural Convection Heat Transfer Coefficients. *Heat Transf. Eng.* **2007**, *28*, 335–342. [CrossRef]
180. Washer, G.A.; Fuchs, P.A. Developments in the Use of Infrared Thermography for the Condition Assessment of Concrete. In Proceedings of the International Symposium Non-Destructive Testing in Civil Engineering, Berlin, Germany, 15–17 September 2015.
181. *SS EN 1992-1-1*; Eurocode 2: Design of Concrete Structures, Part 1-1 General Rules and Rules for Buildings. Singapore Standards Council: Singapore, 2008.
182. American Association of State Highway and Transportation Officials (AASHTO). *Aashto Guide Manual for Bridge Element Inspection*; AASHTO: Washington, DC, USA, 2011.
183. Maldague, X.P.V. *Nondestructive Evaluation of Materials by Infrared Thermography*; Springer: Berlin/Heidelberg, Germany, 2012. [CrossRef]
184. Maldague, X.P.V. *Theory and Practice of Infrared Technology for Nondestructive Testing*; Wiley: New York, NY, USA, 2001.
185. Meola, C. Infrared Thermography in the Architectural Field. *Sci. World* **2013**, *2013*, 323948. [CrossRef]
186. Pozzer, S.; Rezazadeh Azar, E.; Dalla Rosa, F.; Chamberlain Pravia, Z.M. Semantic Segmentation of Defects in Infrared Thermographic Images of Highly Damaged Concrete Structures. *J. Perform. Constr. Facil.* **2021**, *35*, 04020131. [CrossRef]
187. Deshmukh, G.; Birwal, P.; Datir, R.; Patel, S. Thermal Insulation Materials: A Tool for Energy Conservation. *J. Food Process. Technol.* **2017**, *8*, 1–4. [CrossRef]
188. Edis, E.; Flores-Colen, I.; de Brito, J. Passive Thermographic Detection of Moisture Problems in Façades with Adhered Ceramic Cladding. *Constr. Build. Mater.* **2014**, *51*, 187–197. [CrossRef]
189. Minkina, W.; Dudzik, S. *Infrared Thermography: Errors and Uncertainties*, 3rd ed.; J. Wiley: Chichester, UK, 2009. [CrossRef]
190. Ortiz-Sanz, J.; Gil-Docampo, M.; Arza-García, M.; Cañas-Guerrero, I. IR Thermography from UAVs to Monitor Thermal Anomalies in the Envelopes of Traditional Wine Cellars: Field Test. *Remote Sens.* **2019**, *11*, 1424. [CrossRef]
191. Minkina, W.; Klecha, D. Atmospheric Transmission Coefficient Modelling in the Infrared for Thermovision Measurements. *J. Sens. Sens. Syst.* **2016**, *5*, 17–23. [CrossRef]
192. Tashan, J.; Al-Mahaidi, R.; Mamkak, A. Defect Size Measurement and Far Distance Infrared Detection in CFRP-Concrete and CFRP-Steel Systems. *Aust. J. Struct. Eng.* **2015**, *17*, 2–13. [CrossRef]
193. Vollmer, M.; Möllmann, K.-P. *Infrared Thermal Imaging: Fundamentals, Research and Applications*, 2nd ed.; Wiley-VCH: Weinheim, Germany, 2018. [CrossRef]
194. Muniz, P.R.; Cani, S.P.N.; Magalhaes, R.d.S. Influence of Field of View of Thermal Imagers and Angle of View on Temperature Measurements by Infrared Thermovision. *IEEE Sens. J.* **2014**, *14*, 729–733. [CrossRef]
195. Omar, T.; Nehdi, M.L.; Zayed, T. Rational Condition Assessment of RC Bridge Decks Subjected to Corrosion-Induced Delamination. *J. Mater. Civ. Eng.* **2018**, *30*, 04017259. [CrossRef]
196. Ahmed, H.; La, H.M.; Gucunski, N. Review of Non-Destructive Civil Infrastructure Evaluation for Bridges: State-of-the-Art Robotic Platforms, Sensors and Algorithms. *Sensors* **2020**, *20*, 3954. [CrossRef] [PubMed]
197. Lattanzi, D.; Miller, G. Review of Robotic Infrastructure Inspection Systems. *J. Infrastruct. Syst.* **2017**, *23*, 04017004. [CrossRef]

198. Carrio, A.; Pestana, J.; Sanchez-Lopez, J.-L.; Suarez-Fernandez, R.; Campoy, P.; Tendero, R.; García-De-Viedma, M.; González-Rodrigo, B.; Bonatti, J.; Rejas-Ayuga, J.G. Ubristes: UAV-Based Building Rehabilitation with Visible and Thermal Infrared Remote Sensing. In Proceedings of the Robot 2015: Second Iberian Robotics Conference, Lisbon, Portugal, 19–21 November 2015; pp. 245–256. [CrossRef]
199. Pan, N.-H.; Tsai, C.-H.; Chen, K.-Y.; Sung, J. Enhancement of External Wall Decoration Material for the Building in Safety Inspection Method. *J. Civ. Eng. Manag.* **2020**, *26*, 216–226. [CrossRef]

applied sciences

MDPI

Article

The Dependence of Ultrasonic Velocity in Ultra-Low Expansion Glass on Temperature

Wenqing Wei [1,2], Yongfeng Zhang [1,3], Zongzheng Du [1], Minwei Song [1], Yuanyuan Zhang [1,*] and Hong Liu [1,2,*]

[1] Institute of Optics and Electronics, Chinese Academy of Sciences, Chengdu 610209, China; weiwenqing15@mails.ucas.ac.cn (W.W.); zhangyongfeng15@mails.ucas.edu.cn (Y.Z.); 13739450124@163.com (Z.D.); 18780175291@139.com (M.S.)
[2] University of Chinese Academy of Sciences, Beijing 100049, China
[3] Key Laboratory of Adaptive Optics, Chinese Academy of Sciences, Chengdu 610209, China
* Correspondence: zhangyy@ioe.ac.cn (Y.Z.); liuh972@sina.com (H.L.)

Abstract: The coefficient of thermal expansion (CTE) is an important property of ultra-low expansion (ULE) glass, and the ultrasonic velocity method has shown excellent performance for the nondestructive measurement of CTE in large ULE glass. In this method, the accurate acquisition of the ultrasonic velocity in ULE glass is necessary. Herein, we present a correlation method to determine the ultrasonic TOF in ULE glass and to further obtain the ultrasonic longitudinal wave velocity (c_L) indirectly. The performance of this method was verified by simulations. Considering the dependence of c_L on temperature (T), we carried out the derivation of the analytical model between c_L and T. Based on reasonable constant assumptions in the physical sense, a c_L–T exponential model was produced, and some experimental results support this model. Additional experiments were carried out to validate the accuracy of the c_L–T exponential model. The studies we conducted indicate that the c_L–T exponential model can reliably predict the ultrasonic velocity in ULE glass at different temperatures, providing a means for the nondestructive CTE measurement of large ULE glass at a specified temperature.

Keywords: ultra-low expansion glass; ultrasonic velocity; correlation method; temperature coefficient; exponential model

Citation: Wei, W.; Zhang, Y.; Du, Z.; Song, M.; Zhang, Y.; Liu, H. The Dependence of Ultrasonic Velocity in Ultra-Low Expansion Glass on Temperature. *Appl. Sci.* **2022**, *12*, 577. https://doi.org/10.3390/app12020577

Academic Editors: Phong B. Dao, Lei Qiu, Liang Yu and Zahra Sharif Khodaei

Received: 6 December 2021
Accepted: 31 December 2021
Published: 7 January 2022

1. Introduction

Ultra-low expansion (ULE) glass is a SiO_2–TiO_2 glass formed by flame hydrolysis and vapor deposition (nominal composition: 93 wt% SiO_2 and 7 wt% TiO_2) that has found applications in large telescope mirror blanks because of its near-zero coefficient of thermal expansion (CTE) over the 5~35 °C temperature range [1,2]. However, the uniformity of glass material has a significant impact on the CTE homogeneity of the final ULE glass products, which results in figure distortion and in the degradation of the optics. It is therefore necessary to know the CTE characteristics of ULE glass to better understand—and thus better control—the fabrication process for manufacturing the highest quality ULE boules.

Commonly used methods for measuring the CTE of ULE glass include the pushing-rod dilatometer [3], interferometric [4,5], and photoelastic analysis [6–8] methods. All involve destructive measurements, and the procedures are time-consuming and expensive, so they are not suitable for the detection of the CTE uniformity of large ULE glass. The ultrasonic velocity in ULE glass is proportional to its CTE, a fact that can be utilized to nondestructively assess the absolute and relative CTE of large ULE glass, and its feasibility has been demonstrated by researchers [9,10]. In the process of using ultrasonic velocity to nondestructively detedmine CTE, there are two key points to focus on. One is how to guarantee high accuracy in the ultrasonic velocity measurements, and the second is the

correspondence between CTE and ultrasonic velocity at certain given temperatures, mainly because both ultrasonic velocity and CTE are temperature-dependent characteristics.

For the first point, the ultrasonic velocity can be measured by various effective methods, including the threshold method [11,12], the zero-crossing method [13], the peak method [14,15], and the correlation approach [16,17]. The first three methods are based on using the local signal characteristics to measure physical quantities, and the measurement findings are strongly reliant on local signal characteristics and are easily influenced by human factors (such as subjectively selected thresholds) or noise. By contrast, the correlation method considers global signals, which contribute to its excellent noise robustness, and it is free of subjective effects. In this study, we use the correlation method to accurately measure the ultrasonic velocity in ULE glass.

We now turn to the dependence of ultrasonic velocity on temperature. The use of ultrasonic velocity measurements to characterize the CTE of ULE glass was first discovered by researchers at Corning. In their research, some troublesome operations and correction models were used to achieve nondestructive CTE testing [10]. On the one hand, the internal physical mechanism is not very clear. On the other hand, the various cumbersome practical steps (including ensuring the constant temperature of the sample to reach thermal equilibrium, removal of the water bath, quick-drying, and ultrasonic velocity measurement [9]) can introduce uncertainty regarding the sample temperature, which, in turn, affects the accuracy of the CTE measurement results. In later research, the tested sample is not forcefully separated from the constant temperature water bath during ultrasonic CTE measurement [18], which reduces the CTE measurement error, but this also increases the time cost of the entire measurement. Herein, we have analytically modeled the influencing mechanisms, which will help to simplify the practical procedures.

The paper is organized as follows. The principle and method of ultrasonic velocity measurement are mathematically stated in the following section. Section 3 summarizes the derivation of the dependence of ultrasonic velocity on temperature. Section 4 describes the composition of an ultrasonic velocity measurement system and the preparation of the tested samples in detail. In Section 5, the uncertainty and stability of the ultrasonic velocity measurement are discussed, the experimental establishment of the ultrasonic velocity–temperature exponential model is presented, and a comparison of the predicted ultrasonic velocity with the actual measured ultrasonic velocity is listed. Finally, the conclusions are presented in Section 6.

2. Principle and Method of Ultrasonic Velocity Measurement

2.1. Principle of the Ultrasonic Velocity Measurement

Many instruments are now employed to measure ultrasonic velocity, most of which use the pulse reflection method [19–21]. This method usually involves time-domain analysis based on the ultrasonic time of flight (TOF). In this analysis, the ultrasonic velocity is obtained by the ratio of the material thickness to the ultrasonic TOF [13,22]. On this basis, we built a fully integrated high-precision system to measure the ultrasonic velocity in ULE glass using the immersion pulse reflection method.

Figure 1 depicts the schematic diagram of the immersion pulse reflection method. The time of S_F relative to the time base "0" point is called t_1. The time interval between B_1 and S_F is recorded as Δt_1, and the time interval between B_2 and B_1 is recorded as Δt_2, which have the following relationship:

$$\begin{aligned} t_1 &= 2H/c_{\mathrm{L}}^{\mathrm{water}}, \\ \Delta t_1 = \Delta t_2 &= 2d/c_{\mathrm{L}}^{\mathrm{sample}}. \end{aligned} \tag{1}$$

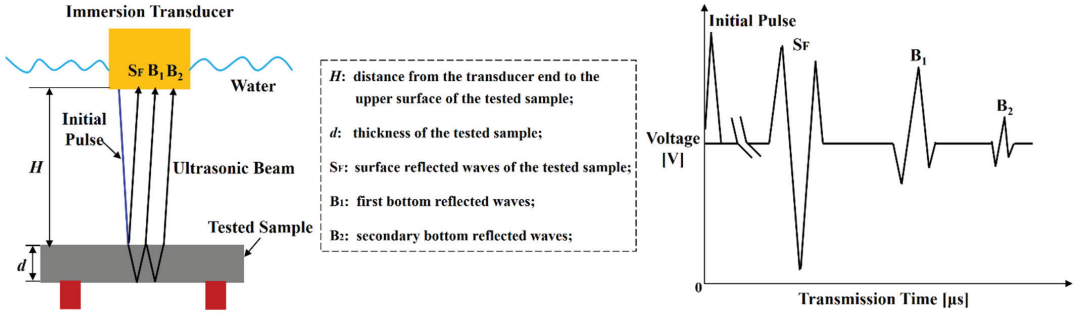


Figure 1. Schematic diagram of the immersion pulse reflection method: (**a**) the transmission process of ultrasonic waves; (**b**) the transmission time of ultrasonic waves.

The ultrasonic longitudinal wave velocity (c_L) of a sample can be calculated from Equation (1):

$$c_L^{\text{sample}} = 2d/\Delta t_1 = 2d/\Delta t_2. \quad (2)$$

Using Equation (2), we can determine the c_L of the tested sample by first determining Δt_1 or Δt_2 of the ultrasonic waveform and the known thickness d.

2.2. Correlation Method for Measuring the TOF

2.2.1. The Correlation Calculation Principle

As the B_1 and B_2 signals of the ULE samples were readily available and were highly similar, the correlation method was used to determine the time interval between B_1 and B_2 as the ultrasonic TOF in the samples. In detail, the ultrasonic signal that was sampled and saved in the PC is designated as $x(t)$. By separating the B_1 and B_2 signals and noting them as $x_1(t)$ and $x_2(t)$, respectively, the correlation coefficient (R) can be represented as

$$R = \frac{\int_{-\infty}^{+\infty} x_1(t)x_2(t)\mathrm{d}t}{\sqrt{\left[\int_{-\infty}^{+\infty} x_1^2(t)\mathrm{d}t \int_{-\infty}^{+\infty} x_2^2(t)\mathrm{d}t\right]}} (|R| \le 1). \quad (3)$$

The ultrasonic signals that were collected by the data acquisition card and the PC, on the other hand, were two discrete signal arrays. As a result, for the discrete signals, the normalized R can be given as

$$R = \frac{\sum x_1(i)x_2(i) - \sum x_1(i)\sum x_2(i)/n}{\sqrt{\left(\left[\sum x_1^2(i) - (\sum x_1(i))^2/n\right]\left[\sum x_2^2(i) - (\sum x_2(i))^2/n\right]\right)}}, i = 1, 2, \ldots, n. \quad (4)$$

where n is the computed length of the signal array, and i is the location inside the signal array.

In the time domain of signal $x(t)$, there is a time interval between the signals $x_1(t)$ and $x_2(t)$. The time interval between $x_1(t)$ and $x_2(t)$ is the round trip time of the ultrasonic wave propagated in the thickness direction, i.e., the TOF. Equation (4) is used to generate the correlation array, with the position having the largest correlation coefficient corresponding to the temporal position of m. The corresponding TOF (Δt) is equal to m/f_S if the sampling frequency is noted as f_S. Finally, the following equation can be used to calculate the ultrasonic velocity:

$$c_L = \frac{2d \cdot f_S}{m}. \quad (5)$$

2.2.2. Simulation of the Correlation Method

First, the reliability of the proposed algorithm needed to be validated by simulations in which the simulated ultrasonic signal is a declining sinusoidal function.

$$x(t) = \beta \cdot \exp(-\alpha_s t) \cdot \sin(2\pi f_c t) \tag{6}$$

where α_s is the declining coefficient of the sinusoidal function set as 9×10^6 Np/m; f_c is the center frequency of the ultrasonic transducer set as 5 MHz; β is the amplitude of the signal; and the amplitudes of the initial and echo signals are set to 1 and 0.3 V, respectively.

The sampling frequency of the signal was set as 2.5 GHz, and the TOF between the initial and echo signals was set as 17.36 µs. To come closer to the actual measurement signal, Gaussian noise with +15 dB SNR (Signal-to-Noise Ratio) was added to the ideal declining sinusoidal signal. The simulated signal waveform and the correlation distribution calculation results are shown in Figure 2a,b, respectively.

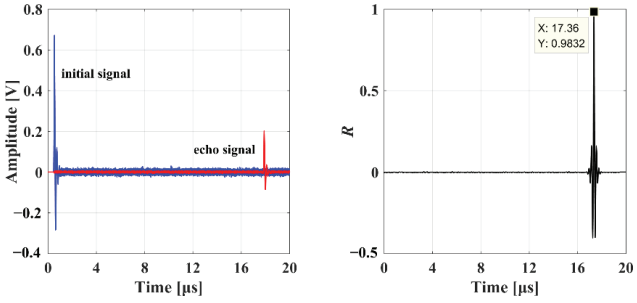


Figure 2. The correlation calculation result of a simulated signal: (**a**) the waveform of the simulated signal; (**b**) the correlation coefficient distribution.

Where the X value is the time corresponding to the largest correlation coefficient, and the Y value is the largest correlation coefficient.

In Figure 2, the calculated correlation coefficient has a clear peak on the time axis, with a TOF of 17.36 µs corresponding to the peak. The calculated results were in good agreement with the theoretical settings, so the algorithm is suitable for ultrasonic signal processing in ultrasonic velocity measurements.

3. Theoretical Model between Ultrasonic Velocity and Temperature

Ultrasound is defined as an elastic wave of high frequency that propagates in a medium. Therefore, when an elastic wave propagates in an isotropic medium without being affected by volume stress, the ultrasonic longitudinal wave velocity, denoted as c_L, can be found from the following equation:

$$c_L = \sqrt{\frac{E(1-v)}{\rho(1+v)(1-2v)}}, \tag{7}$$

where ρ is density, v is Poisson's ratio, and E is Young's modulus.

Equation (7) indicates that the c_L in the material is mainly related to Young's modulus, density, and Poisson's ratio. For ULE glass, within the upper limit of 11.5 wt%, slight changes in the TiO_2 content will not change the density or Poisson's ratio but will cause changes to Young's modulus [23]. Therefore, the difference in c_L is mainly determined by the difference in Young's modulus. Then, the effect of temperature on the ultrasonic velocity essentially reflects its effect on Young's modulus. The microscopic analysis of Young's modulus shows that the reaction of a solid to all forces comes from the potential

energy of the interaction between atoms. The interaction potential $U(r)$ of a pair of two atoms separated by r can be expressed as:

$$U(r) = -\frac{A}{r^n} + \frac{B}{r^m}, \tag{8}$$

where A, B, n, and m are all positive constants. The first term represents the energy of attraction, and the second term represents the energy of repulsion.

Assuming that the solid is stretched along the crystal axis when the tensile force changes df, the interatomic distance r changes to dr. At this time, the cross-sectional area r^2 of a unit cell is regarded as inconvenient. Then, Young's elastic modulus can be expressed as:

$$E = \frac{\sigma}{\varepsilon} = \frac{\frac{\mathrm{d}f}{r^2}}{\frac{\mathrm{d}r}{r}} = \frac{\mathrm{d}f}{r\mathrm{d}r}, \tag{9}$$

where σ and ε represent the stress and strain, respectively.

Since the binding force f of the tensioned solid is only related to the first term of Equation (8), its magnitude is:

$$f(r) = -\frac{\mathrm{d}U(r)}{\mathrm{d}r} = -\frac{nA}{r^{n+1}}. \tag{10}$$

Equation (10) is then substituted into Equation (9) by deriving the derivative for r to obtain

$$E = \frac{n(n+1)A}{r^{n+3}}. \tag{11}$$

Assume $K = (n + 1)\ A$ and $Q = n + 3$. Equation (11) can be changed to:

$$E = \frac{nK}{r^Q}. \tag{12}$$

Equation (12) takes the derivative of T and divides both sides by Er^Q at the same time, shifting the term to obtain:

$$\frac{\mathrm{d}E}{E\mathrm{d}T} = Q\frac{\mathrm{d}r}{r\mathrm{d}T}. \tag{13}$$

It is assumed that the distance between atoms still obeys the following rules when a solid is heated and expanded:

$$r = r_0(1 + \alpha T), \tag{14}$$

where r_0 is the atomic distance when the absolute temperature $T_0 = 0$; α is the linear expansion coefficient of the solid, and its differential definition is:

$$\begin{cases} \alpha = \frac{1}{r}\frac{dr}{dT}, \eta = \frac{1}{E}\frac{dE}{dT}. \end{cases} \tag{15}$$

The η in Equation (15) is the temperature coefficient of the elastic modulus. Relevant studies have demonstrated that the elastic modulus of ULE glass increases with the increase in temperature, and the increment dE of E has a positive value [23]. When both sides of Equation (13) are multiplied by dT, Equation (14) is then substituted in and integrated, and we obtain:

$$\int_{E_0}^{E} \frac{\mathrm{d}E}{E} = Q\int_{T_0}^{T} \frac{\mathrm{d}(1+\alpha T)}{(1+\alpha T)}. \tag{16}$$

From Equation (16) and considering $T_0 = 0$, we obtain:

$$E = E_0\left(\frac{1+\alpha T}{1+\alpha T_0}\right)^Q \approx E_0(1 + Q\alpha T). \tag{17}$$

From Equations (13) and (15), we see that:

$$Q{=}\eta/\alpha. \tag{18}$$

Therefore, Equation (17) is changed to:

$$E = E_0(1 + \eta T). \tag{19}$$

Equation (19) shows that the elastic modulus of ULE glass increases as the temperature increases, where η is the temperature coefficient of the elastic modulus. Therefore, it can be inferred that the c_L in ULE glass also increases as T increases. A review of the data reveals that when considering the effects of thermal expansion, there is a specific equation relating the c_L and T, which is as follows [24]:

$$c_L^T = c_L^{T_0}[1 + \alpha_L(T - T_0)], \tag{20}$$

where α_L is the temperature coefficient of the ultrasonic longitudinal wave velocity, a positive value of about 10^{-4} orders of magnitude, and c_L^T, $c_L^{T_0}$ indicate the ultrasonic longitudinal wave velocity of the material at temperatures T and T_0, respectively.

Although there are certain differences in α_L at different temperature points, we can use the c_L–T data in a small temperature range to fit and solve the average α_L in this temperature range. The method of a differential equation is introduced here to obtain the mathematical relationship between c_L and T in a small temperature range. Writing Equation (20) in a differential form, we obtain

$$\alpha_L = \frac{dc_L}{c_L dT},\ T_1 < T < T_2. \tag{21}$$

Solving this equation yields

$$c_L = c' \cdot e^{\alpha_L T}. \tag{22}$$

Equation (22) shows that the relationship between c_L and T is theoretically exponential within a certain temperature range.

4. Tested Material and Experimental Setup

4.1. Materials and Sample Preparation

The selected ULE glass was Corning Code 7972 glass. Considering the boundary effect of ultrasonic wave propagation, the experimental samples were prepared in a cylindrical shape with a cross-sectional area that was larger than that of the ultrasonic transducer. To avoid the adverse effects of the scattering attenuation of ultrasound at the interface of the sample, the two end faces of the tested sample should be flat and parallel to each other. As shown in Figure 3, six cylindrical glass samples were cut with an equal thickness (50–0.1 mm) along the radial position of the ULE 7972 boule (No. 82714) using an abrasive water jet; then, these samples were finely ground and polished to achieve a flatness of 0.5λ and parallelism of 20 μm. The six samples were numbered 1#~6# from the inside to the outside of the boule according to the increasing CTE. The details of the ULE samples are summarized in Table 1.

4.2. Ultrasonic Measurement System

The schematic diagram of the c_L. measurement experimental system is illustrated in Figure 4. A 75 MHz ultrasonic pulser/receiver (Olympus, Waltham, MA, USA, model 5073PR) was used, which was connected to a 400 MHz data acquisition card (Spectrum Instrumentation GmbH, Grosshansdorf, Germany, model M4i.2220-x8) that transmits the ultrasonic signals to a computer to be processed. All signals were captured with 2,500,000 points at a sampling rate of 2.5 GHz. After the acquisition, the data were properly processed to determine the involved ultrasonic velocities. The temperature of the tested

sample was controlled by a thermostatic water tank (Hangzhou Qiwei, Hangzhou, China, model DHC-05-B) with a temperature control precision of ±0.05 °C.

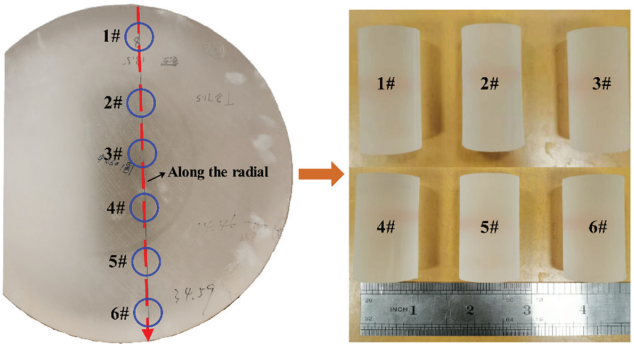


Figure 3. The prepared ULE glass samples: (**a**) source of samples; (**b**) photograph of samples.

Table 1. The important parameters of the ULE samples.

No.	Thickness (mm)	CTE (5~35 °C) (ppb/°C)
1#	49.936	−1
2#	49.932	0
3#	49.937	1
4#	49.940	2
5#	49.941	2
6#	49.927	3

Note: No.—sample number; thickness—measured thickness of the ULE samples; CTE (5~35 °C)—average CTE over the temperature range 5~35 °C; ppb/°C—unit of CTE (1 ppb/°C = 1×10^{-9}/°C).

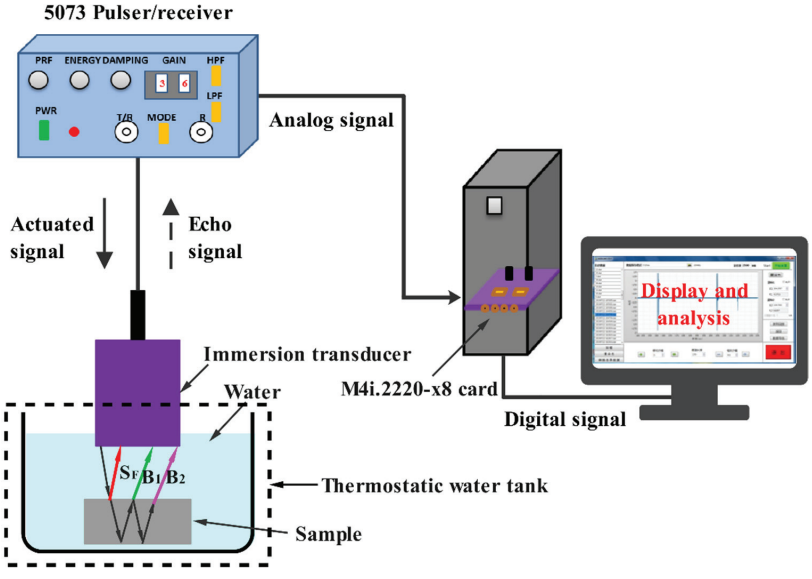


Figure 4. Schematic diagram of the experimental system for measuring ultrasonic velocity.

The self-generating and self-receiving transducer model was utilized to transmit and receive the ultrasonic waves. Considering the diameter and thickness of the tested sample, wideband focusing 19.05 mm diameter longitudinal wave immersion transducers (Olympus, Waltham MA, USA) with 3.5 and 5 MHz nominal center frequencies were employed in this investigation. Figure 5 illustrates the ultrasonic velocity measurement results for the samples 1#~6# at 20 °C using two different ultrasonic frequencies: 3.5 and 5 MHz. Each sample was subjected to five replicate measurements by each transducer. The ultrasonic frequency had little effect on the ultrasonic velocity of the sample, as the c_L in the sample only depends on its inherent physical parameters, including its bulk modulus and density [25]. In addition, the repeatability of the c_L measurement was essentially the same for both frequency transducers, both within 0.1 m/s. This suggests that any frequency could be used to characterize the CTE of ULE glass in engineering applications when using ultrasonic velocity measurement methods.

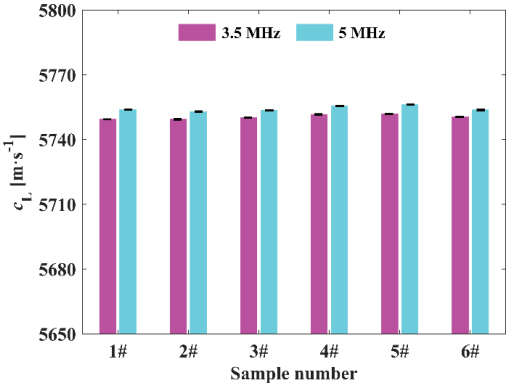


Figure 5. c_L measurement results of the ULE samples 1#~6# at frequencies of 3.5 and 5 MHz.

5. Results and Discussion

5.1. c_L Measurement at a Single Temperature

5.1.1. Uncertainty Analysis

In this work, an ultrasonic transducer with a 5 MHz center frequency was employed to measure the ultrasonic velocities in all of the prepared ULE samples. Figure 6 gives an example of the received ultrasonic signals, which formed a series of ultrasonic longitudinal waves with a sampling rate of 2.5 GHz. High-frequency noise was found to be present in the actual ultrasonic signals. However, since the correlation of the noisy signal was very small, the effect of such noise was removed via the correlation calculation. The results of the TOF calculations obtained using this method are depicted in Figure 6. Again, there is a peak in the correlation coefficient distribution plot corresponding to a TOF of Δt = 17.3383 μs.

The uncertainty in measuring ultrasonic velocity was also investigated, and the expression is shown in Equation (23).

$$u_{c_L} = \sqrt{\left(\frac{2}{\Delta t}u_d\right)^2 + \left(\frac{2d}{(\Delta t)^2}u_{\Delta t}\right)^2} \tag{23}$$

where u_{c_L}, u_d, and $u_{\Delta t}$ are the uncertainties regarding the c_L, d, and TOF in the ULE glass, respectively.

As illustrated in the first item of Equation (23), the uncertainty related to the d was $u_{c_L}(d) = (2/\Delta t)\Delta u_d$. A typical value of Δt is 17.3562 μs. When measuring the d of the ULE sample, u_d was evaluated by the measurement precision of a micrometer (0.001 mm), thus $u_{c_L}(d) = 0.12$ m/s. After substituting the measured values into the second item of Equation (23), the uncertainty introduced by the TOF (Δt) was obtained: $u_{c_L}(\Delta t) = 0.332 \times 10^9 u_{\Delta t}$.

Since $u_{\Delta t}$ was evaluated by the sampling period of the data acquisition card (0.4 ns), the typical $u_{c_L}(\Delta t)$ was 0.13 m/s. Therefore, when measuring the ULE glass using the proposed experimental setup, the calculated total uncertainty was $u_{c_L} = 0.2$ m/s, ensuring that the ultrasonic velocity measurements at a single temperature had high reliability.

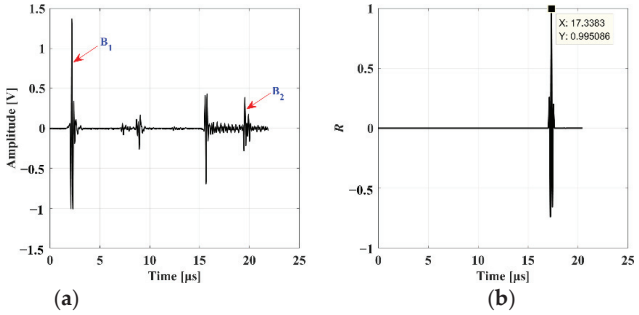


Figure 6. The correlation calculation result of an actual received ultrasonic signal: (**a**) the ultrasonic echo signal; (**b**) the correlation coefficient distribution.

5.1.2. Stability of c_L Measurement

To determine whether the proposed measurement system could provide reliable c_L measurements over long periods of time, the stability of the measured c_L value also needs to be considered when the tested samples reach thermal equilibrium. Long-term measurement of the ultrasonic velocity in samples 1#~6# with a high CTE was taken at the same temperature (20 °C), and the c_L data were recorded at 1 h intervals, for a total of seven measurements in one day. Figure 7 depicts the variation in c_L with the measurement time.

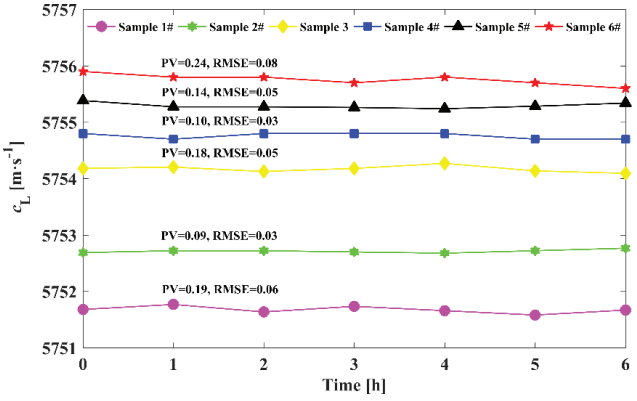


Figure 7. The change in c_L with time for samples 1#~6#.

The c_L measured in the same sample was almost constant at the different times while maintaining a constant temperature, and the RMSE (Root Mean Squared Error) of the c_L changes were all within 0.10 m/s, which indicate that the measurement system that was built in this paper has stable performance. This also provides a strong guarantee for the c_L measurement of a large batch of ULE glass samples.

5.2. Measurement and Analysis of c_L–T Data

5.2.1. Acquisition of c_L–T Data

The tested sample and an ultrasonic transducer fixed above the sample were placed into a thermostatic water tank, and the temperature in the tank was steadily increased

in the range of 10~30 °C. Here, the temperature range of 10~30 °C was chosen for two reasons: (1) this corresponds to the real temperature range that ULE glass is likely to experience when undergoing ultrasonic measurements, i.e., the range of room temperature throughout the year, and (2) it corresponds to the temperature range of the ultrasonic transducer in use. We were utilizing a standard immersion transducer, whose normal operating temperature range is 10~60 °C, beyond which the piezoelectric action of the transducer would be weakened, making high-amplitude data acquisition difficult. The ultrasonic echo signals were manually sampled and stored by the PC at a temperature interval of 1 °C, which was chosen to account for the maximum number of temperature points to be sampled and the required modeling time. The temperature was held constant within ±0.05 °C, as measured using a digital thermometer probe in the bath.

To ensure that the glass sample was in thermal equilibrium during the TOF measurement, the sample was immersed in a controlled water bath for at least 94 min, which was calculated according to the time t (hours) to reach equilibrium for a given glass thickness d (cm), which could be expressed as $t = d^2/16$ in [26]. To reduce measurement errors, the echo signals were acquired three times at each temperature point. By using the described correlation algorithm method, the ultrasonic TOF was obtained, and the average values were used to calculate the c_L of the tested ULE samples. Nevertheless, the experiment did not consider the change in the d and ρ of the tested samples.

The obtained ultrasonic velocities of samples 1#~6# are presented in Figure 8 in terms of temperature. The ultrasonic velocities are observed to increase as the temperature increases. Changes in c_L are correlated with temperature changes, regardless of the absolute value of the CTE. Comparing Figure 8a–f, we can also observe that the ultrasonic velocity differs at the same temperature and that the c_L increases with the CTE of ULE glass.

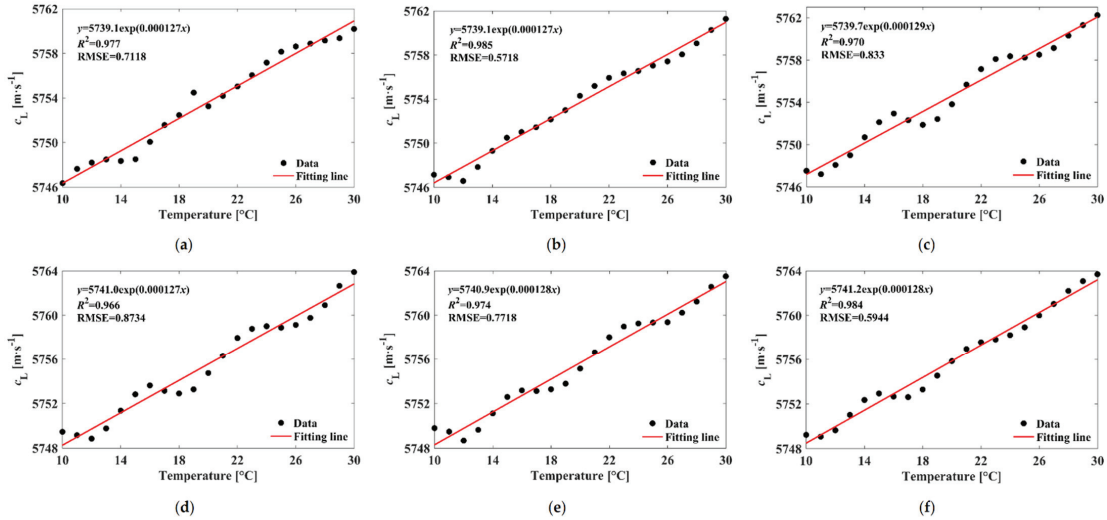


Figure 8. Changes in c_L(m·s^{-1}) with temperature for all measured samples: (**a**) sample 1#; (**b**) sample 2#; (**c**) sample 3#; (**d**) sample 4#; (**e**) sample 5#; (**f**) sample 6#.

5.2.2. Analysis of the Change in the c_L with T

The data from Figure 8 were used to directly plot the changes in c_L as a function of T according to the exponential fitting in Equation (22). Table 2 displays the results of all six samples, with excellent exponential fits given for each case and with α_L mostly ranging from 0.000127 to 0.000129, which shows that for ULE materials with very small CTE differences, their α_L is almost constant. This means that the temperature effect pattern on ultrasonic velocity was essentially the same in ULE glasses with different CTE values.

Therefore, the average value of these α_L coefficients could be taken as the temperature correction coefficient of ultrasonic velocity in ULE glass. Of course, because of the small difference in the CTE, the c' in Equation (22) showed relatively large fluctuations. In the physical sense, the c' indicates that c_L correpsonds to 0 °C. It was found that as the CTE of the ULE sample increased, c' increased correspondingly, which is consistent with the mechanism of the linear positive correlation between c_L and the CTE.

Table 2. Data from the six samples shown in Figure 8 were taken to exponentially fit the relationship between c_L and T.

No.	$\alpha_L/\times 10^{-6}$	R^2	RMSE
1#	127	0.977	0.7118
2#	127	0.985	0.5718
3#	129	0.970	0.8330
4#	127	0.966	0.8734
5#	128	0.974	0.7718
6#	128	0.984	0.5944

Note: No.—sample number; α_L—temperature coefficient of ultrasonic longitudinal wave velocity; R^2—*R*-square (a measure of goodness of fit).

To represent the more pronounced difference in the ultrasonic velocity with increasing temperature for the ULE samples with different CTE values, a new coefficient is defined as follows:

$$\beta_L = \frac{\Delta c_L}{c_L \Delta T}. \tag{24}$$

Here, we take the c_L corresponding to 10 °C of the samples as the basis c_L, and calculate the β_L via Equation (24) with temperature increments of 5 °C, i.e., at 15 °C, 20 °C, 25 °C, and 30 °C. The β_L for samples 1#, 2#, and 3# is shown in Figure 9. We noted that for the different temperature ranges, the speed of the ultrasonic velocity changes with the temperature was inconsistent. This is possibly because the elastic modulus also has a temperature coefficient β_E, which varies in different temperature ranges. The PV value of β_L for the three samples was 0.000098, 0.000056, and 0.000101. Considering that the difference was minor, we conclude that the multi-data point fitting method that was used to obtain α_L in a given temperature range in this paper is a more feasible equivalent to an average function and has wider temperature applicability.

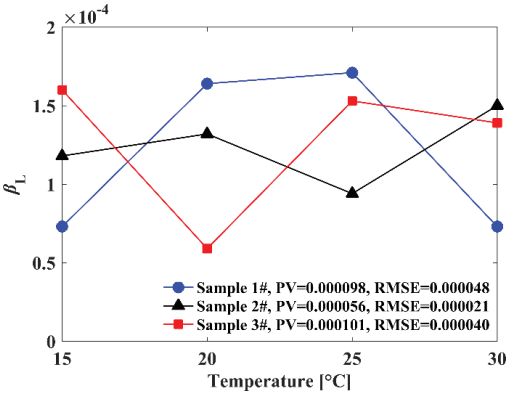


Figure 9. The relationship between β_L and temperature for ULE samples with different CTEs.

Samples 4# and 5#, which had the same CTE, were also analyzed for differences in the c_L as the temperature increased, and the results are shown in Figure 10. We found that the two samples had similar patterns of variation in β_L. This implies that the changing trends in the ultrasonic velocity with temperature was, to some extent, consistent with the

changing trend of the CTE with temperature; however, this assumes that the thickness of the ULE glass sample did not change over a wide temperature range.

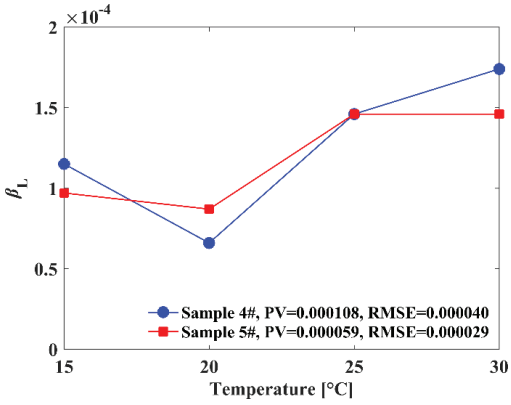


Figure 10. The relationship between β_L and temperature for ULE samples with the same CTE.

The correlation between c_L and T in ULE glass over a wide range of CTE values from −1 ppb/°C to +3 ppb/°C was investigated and discussed. We will next validate the accuracy of the c_L–T exponential model for ULE samples with different CTE values through a series of experiments.

5.3. Accuracy Validation of c_L–T Model

The accuracy of the c_L–T model was decisive for its application in engineering. Samples 1#, 3#, and 5# were randomly selected, and their ultrasonic velocities were measured at ten random temperature points in the range of 10~30 °C. The measured values were compared with the model predictions in this paper, and the results are shown in Table 3.

Table 3. Experimental data from the model validation and their error analysis.

No.	Temperature (°C)	c_m(m/s)	c_p(m/s)	$c_m - c_p$(m/s)	$\lvert\delta\rvert$ %
1#	10.2	5745.5	5746.3	−0.8	0.014
	12.7	5748.1	5748.2	−0.1	0.002
	14.3	5749.6	5749.3	0.3	0.005
	15.2	5749.4	5750.0	−0.6	0.010
	17.6	5750.3	5751.7	−1.4	0.024
	21.8	5754.4	5754.8	−0.4	0.007
	23.5	5754.9	5756.1	−1.2	0.021
	25.6	5457.1	5757.6	−0.5	0.009
	26.0	5757.5	5757.9	−0.4	0.007
	27.9	5759.4	5759.3	0.1	0.002
				σ_c = 0.75	
3#	10.8	5747.5	5746.8	0.7	0.012
	12.4	5747.9	5748.0	−0.1	0.002
	15.2	5751.7	5750.1	1.6	0.028
	16.5	5751.5	5751.0	0.5	0.009
	16.9	5751.4	5751.3	0.1	0.002
	18.4	5751.5	5752.4	−0.9	0.016
	23.3	5757.3	5756.1	1.2	0.021
	24.3	5757.2	5756.8	0.4	0.007
	25.6	5757.5	5757.8	−0.3	0.005
	29.0	5761.1	5760.3	0.8	0.014
				σ_c = 0.85	

Table 3. Cont.

No.	Temperature (°C)	c_m(m/s)	c_p(m/s)	$c_m - c_p$(m/s)	$\|\delta\|$ %
5#	10.6	5750.4	5749.7	0.7	0.012
	11.3	5750.2	5750.2	0.0	0.000
	13.5	5752.1	5751.8	0.3	0.005
	14.1	5753.1	5752.3	0.8	0.014
	15.1	5754.4	5753.0	1.4	0.024
	17.5	5754.2	5754.8	−0.6	0.010
	20.3	5756.8	5756.8	0.0	0.000
	25.2	5760.3	5760.5	−0.2	0.003
	25.9	5760.4	5761.0	−0.6	0.010
	29.9	5764.8	5763.9	0.9	0.016
				σ_c = 0.73	

Note: c_m is the actual measured ultrasonic velocity, c_p is the ultrasonic velocity predicted by the fitting model, and δ denotes the relative error. The deviation σ_c is a measure of the inaccuracy.

Considering that the measurement locations of the three samples in the model validation procedure may vary, it was reasonable to use a recalculation of c' based on the standard ultrasonic velocity at 20 °C to determine the c_L–T exponential model, which may differ slightly from the c' obtained by modeling the above c_L–T data. The c_L values in the tested glasses at an average room temperature of 20 °C were substituted into the derived exponential model based on the average temperature coefficient from the above fitting analysis to determine a c_L–T exponential model for each tested glass. As seen in Table 3, the standard deviation of the predicted values for three samples were all within 0.90 m/s, and the relative errors between measured and model-predicted values were mostly within 0.020%, which suggests that the models fitted in this work exhibit high precision.

Of course, it should be noted that increasing the sample thickness may further reduce the errors in c_L measurement, but it also imposes more stringent requirements on the frequency of the transducer and time required for the sample to reach thermal equilibrium.

6. Conclusions

In this investigation, the aim was to theoretically analyze and validate the dependence of the ultrasonic velocity in ULE glass on temperature. Based on the simulation and experimental investigations, we can draw the following conclusions:

1. The proposed pulse reflection immersion method provides reliable and stable measurements of the ultrasonic echo signal of ULE glass, and the calculation of the signal based on the correlation method can be used to conveniently and accurately extract the ultrasonic TOF of the tested sample and then obtain the ultrasonic velocity.
2. The application of the proposed method for six ULE samples with different CTE values indicates that ultrasonic velocity increases as the experimental temperature is increased. Furthermore, the c_L–T exponential model was theoretically analyzed and experimentally fitted. The predicted c_L using the exponential model at ten random temperature points shows good agreement with the actual measured ultrasonic velocities at the same temperature.

These findings indicate the promising potential of the c_L–T exponential model to determine the ultrasonic velocity at a specified temperature for ultrasonic nondestructive CTE measurement in large ULE glass, which will be critical for reliably evaluating the CTE homogeneity of large ULE glass at a specified temperature.

There are some comments that can be made about the general applicability of our method. In this paper, the derivation and verification of the c_L–T relationship are only for isotropic ULE glass, but the method for deriving exponential model can also be applied to anisotropic materials, such as anisotropic crystals, and for obtaining the longitudinal wave velocity, transverse wave velocity, and then the Poisson's ratio. The only distinction is α_L. As for the experimental verification method for the exponential model, the correlation

method, it is also applicable for the measurement of the TOF, regardless of whether the material is isotropic.

Author Contributions: Conceptualization, W.W., Y.Z. (Yuanyuan Zhang) and H.L.; methodology, Y.Z. (Yongfeng Zhang) and H.L.; software, W.W. and Y.Z. (Yongfeng Zhang); validation, W.W. and H.L.; investigation, W.W. and M.S.; resources, Y.Z. (Yuanyuan Zhang), M.S. and H.L.; data curation, W.W.; writing—original draft preparation, W.W.; writing—review and editing, Y.Z. (Yongfeng Zhang), Z.D. and H.L.; visualization, W.W.; project administration, Y.Z. (Yuanyuan Zhang). All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. VanBrocklin, R.R.; Edwards, M.J.; Wells, B. Review of Corning's capabilities for ULE mirror blank manufacturing for an extremely large telescope. In Proceedings of the SPIE—The International Society for Optical Engineering, Orlando, FL, USA, 6 July 2006; Volume 6273, p. 627301.
2. Kendrick, S.E.; Stahl, H.P. Large aperture space telescope mirror fabrication trades. In Proceedings of the SPIE—The International Society for Optical Engineering, Marseille, France, 12 July 2008; Volume 7010, p. 70102G.
3. Plummer, W.A.; Hagy, H.E. Precision thermal expansion measurements on low expansion optical materials. *Appl. Opt.* **1968**, *7*, 825–831. [CrossRef]
4. Imai, H.; Okaji, M.; Kishii, T.; Sagara, H.; Aikawa, H.; Kato, R. Measurement of thermal expansivity of low-expansion glasses by interferometric methods: Results of an interlaboratory comparison. *Int. J. Thermophys.* **1990**, *11*, 937–947. [CrossRef]
5. Kato, R.; Azumi, T.; Maesono, A. Measurement of thermal expansion of low-expansion glasses by a laser interferometric thermal expansion meter. *High Temp.-High Press.* **1991**, *23*, 615–620.
6. Hagy, H.E.; Smith, A.F. Sandwich seal in the development and control of sealing glasses. *J. Can. Ceram. Soc.* **1969**, *38*, 63–68.
7. Hagy, H.E.; Best, M.E. Comparison of two high-precision nondestructive measurement methods for evaluating thermal expansion differences in the 8.3-m ultralow-expansion Subaru primary mirror blank. *Appl. Opt.* **1996**, *35*, 1126–1128. [CrossRef]
8. Gulati, S.T.; Hagy, H.E. Theory of the narrow sandwich seal. *J. Am. Ceram. Soc.* **2006**, *61*, 260–263. [CrossRef]
9. Hagy, H.E. High precision photoelastic and ultrasonic techniques for determining absolute and differential thermal expansion of titania–silica glasses. *Appl. Opt.* **1973**, *12*, 1440–1446. [CrossRef] [PubMed]
10. Hagy, H.E.; Shirkey, W. Determining absolute thermal expansion of titania–silica glasses: A refined ultrasonic method. *Appl. Opt.* **1975**, *14*, 2099–2103. [CrossRef] [PubMed]
11. Barshan, B. Fast processing techniques for accurate ultrasonic range measurements. *Meas. Sci. Technol.* **2000**, *11*, 45–50. [CrossRef]
12. Xia, N.; Zhao, P.; Zhang, J.; Xie, J.; Fu, J. Investigation of ultrasound velocity measurements of polymeric parts with different surface roughness. *Polym. Test.* **2020**, *81*, 106231. [CrossRef]
13. Zhao, P.; Peng, Y.; Yang, W. Crystallization measurements via ultrasonic velocity: Study of poly (lactic acid) parts. *J. Polym. Sci. Part B Polym. Phys.* **2015**, *53*, 700–708. [CrossRef]
14. Jenot, F.; Ouaftouh, M.; Duquennoy, M.; Ourak, M. Corrosion thickness gauging in plates using Lamb wave group velocity measurements. *Meas. Sci. Technol.* **2001**, *12*, 1287–1293. [CrossRef]
15. Kim, Y.H.; Song, S.J.; Lee, J.K. Technique for measurements of elastic wave velocities and thickness of solid plate from access on only one side. *Jpn. J. Appl. Phys.* **2005**, *44*, 5240–5243. [CrossRef]
16. Hu, E.; Wang, W. The elastic constants measurement of metal alloy by using ultrasonic nondestructive method at different temperature. *Math. Probl. Eng.* **2016**, *2016*, 6762076. [CrossRef]
17. Zhao, P.; Xia, N.; Zhang, J. Measurement of molecular orientation using longitudinal ultrasound and its first application in in-situ characterization. *Polymer* **2020**, *187*, 122092. [CrossRef]
18. Edwards, M.J.; Bullock, E.H.; Morton, D.E. Improved precision of absolute thermal expansion measurements for ULE glass. In Proceedings of the SPIE—The International Society for Optical Engineering, Denver, CO, USA, 11 November 1996; Volume 2857, pp. 58–63.
19. Leydier, A.; Mathieu, J.; Despaux, G. The two coupling fluids method for ultrasonic velocity measurement. Application to biological tissues. *Meas. Sci. Technol.* **2009**, *20*, 095801. [CrossRef]
20. Souri, D. Ultrasonic velocities, elastic modulus and hardness of ternary Sb-V_2O_5-TeO_2 glasses. *J. Non-Cryst. Solids* **2017**, *470*, 112–121. [CrossRef]

21. Minh, H.N.; Du, J.; Raum, K. Estimation of thickness and speed of sound in cortical bone using multifocus pulse-echo ultrasound. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2019**, *67*, 568–579. [CrossRef] [PubMed]
22. Daugschies, M.; Rohde, K.; Glüer, C.C.; Barkmann, R. The preliminary evaluation of a 1 MHz ultrasound probe for measuring the elastic anisotropy of human cortical bone. *Ultrasonics* **2014**, *54*, 4–10. [CrossRef] [PubMed]
23. Gulati, S.T. Mechanical properties of SiO_2 vs. SiO_2-TiO_2 bulk glasses and fibers. *MRS Online Proc. Libr.* **1991**, *244*, 67–84. [CrossRef]
24. Jiang, Z. Effect of temperature on ultrasonic velocity and stress measurement (in Chinese). *Nondestruct. Test.* **1999**, *21*, 245–248.
25. Perepechko, I.; Leib, G. *Acoustic Methods of Investigating Polymers*; Mir Publishers: Moscow, Russia, 1975.
26. Lillie, H.R.; Ritland, H.N. Fine annealing of optical glass. *J. Am. Ceram. Soc.* **1954**, *37*, 466–473. [CrossRef]

applied sciences

MDPI

Article

A Characterization Method for Pavement Structural Condition Assessment Based on the Distribution Parameter of the Vehicle Vibration Signal

Weiguo Wang [1,*], Shishi Zhou [2] and Qun Yang [2]

[1] School of Civil Engineering, Southeast University, Nanjing 211189, China
[2] The Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, Shanghai 201804, China; 1710932@tongji.edu.cn (S.Z.); qunyang.w@tongji.edu.cn (Q.Y.)
* Correspondence: 230188209@seu.edu.cn

Featured Application: This article has potential applications in rapid pavement structure evaluation based on vehicle vibration.

Abstract: A pavement structural survey plays a vital role in road maintenance and management. This study was intended to explore the feasibility of a non-stop pavement structure assessment method by analyzing the vibration data from a vehicle sensor. In this study, three falling weight deflectometer (FWD) tests and four vehicle vibration tests were conducted on five pavement structures. The FWD test results show that the continuously reinforced composite pavement has a higher structural stiffness than the semi-rigid base asphalt pavement. According to the statistical distribution of vehicle acceleration, a distribution parameter, the peak probability density (PPD), was proposed. The correlation coefficient (−0.722) of the center deflection (D1) and PPD indicates a strong correlation between the two variables. Therefore, PPD is strongly correlated with pavement structural stiffness. This study proposed a novel characterization method for pavement structural conditions based on the distribution parameter of the vehicle vibration signal.

Keywords: pavement; structural condition; vibration signal; probability density distribution; correlation coefficient; falling weight deflectometer test

Citation: Wang, W.; Zhou, S.; Yang, Q. A Characterization Method for Pavement Structural Condition Assessment Based on the Distribution Parameter of the Vehicle Vibration Signal. *Appl. Sci.* **2022**, *12*, 683. https://doi.org/10.3390/app12020683

Academic Editors: Liang Yu, Phong B. Dao, Lei Qiu and Zahra Sharif Khodaei

Received: 15 December 2021
Accepted: 9 January 2022
Published: 11 January 2022

1. Introduction

The good structural condition of infrastructure is the basis for ensuring its safe operation and providing the corresponding services. Therefore, structural health monitoring (SHM) has always been a hot spot in the engineering industry. Among SHM methods, vibration-based SHM uses the dynamic response of structures such as acceleration to reflect the structural condition [1]. This method has the merit of remote testing, while the drawbacks are that it requires an installation cost, and the power should be more accessible.

During the operation of a pavement, its structural stiffness is weakening year by year. Obtaining the structural condition of the pavement is an essential part of a pavement condition survey and evaluation, which is the basis for the transport agencies to make maintenance decisions, ensure the pavement's structural security, and maintain its level of service.

The assessment methods of pavement structure conditions mainly include the core-drilling method, dynamic deflection method, and some innovative methods using distributed fiber and ground-penetrating radar, etc. [2,3]. The core-drilling sampling method is inconvenient to use because it destroys the pavement structure and interrupts traffic. Therefore, it is often adopted for verification rather than testing. In contrast, the dynamic deflection method has a broader application. Researchers have mainly studied the performance evaluation of pavement structure based on the modulus back-calculation and deflection basin parameters [4–6]. The former has received more attention.

In 1977, Hou [7] proposed a method in his doctoral dissertation to compare the error between the two deflections was minimal. Later, many researchers adopted a similar idea to invert the structural layer modulus based on the falling weight deflectometer (FWD) deflection basin data. As time went by, some scholars employed advanced analytical techniques such as neural networks, genetic algorithms to solve the non-uniqueness of the solution in modulus inversion. Li and Wang [8] developed an artificial neural network and genetic algorithm (ANN-GA) method to back-calculate the layer moduli of flexible pavement from the falling weight deflectometer (FWD) test. The back-calculated parameters can be directly applied to the mechanistic-empirical design for pavement overlays.

Deflection data analysis can effectively evaluate the performance of pavement structure, but this kind of method requires parking before testing. Closing roads will affect traffic flow; therefore, it is difficult to apply on roads with heavy traffic. This study explores the feasibility of a non-stop pavement structure assessment method by analyzing vehicle vibration data.

Pavement condition includes structural condition and surface condition. Pavement condition surveys based on vehicle vibration primarily focus on surface performance [9,10], including pavement roughness [11], surface stresses such as potholes and transverse cracks, etc. [12–14]. Still, there are very few studies on pavement structures. In fact, the vibration of a vehicle running on the pavement can reflect structural conditions to a certain extent, and only this relationship may be implicit.

Inspired by vibration-based structural health monitoring, Yang and co-workers demonstrated for the first time the feasibility of analyzing a bridge structure through the vibration signals of the driven vehicle [15]. Over the past years, they still focused on this issue and have conducted a series of analyses, from the initial extraction of the first-order frequency to the determination of modal parameters such as multi-order frequencies, damping, and mode shapes, then to damage identification and location, and expansion joint detection, etc. [16–19]. Unlike traditional SHM, these methods are called indirect SHM (ISHM) because the sensors are installed on a moving vehicle rather than on the structure itself [20]. ISHM has the merits of being mobile, economic, and efficient [19]. In response to the problem that the manually extracted features are not very sensitive to damage, Liu, et al. [21] proposed a physics-guided algorithm to extract an effective feature to determine the damage in a bridge.

With the rapid development of smartphones, in recent years, there has been some research on ISHM, based on the data obtained by the built-in sensors of smartphones [22,23]. These data are mainly derived from a triaxial accelerometer, a global positioning system (GPS), and an inertial measurement unit (IMU). This kind of method has an even lower cost and can collect more data. By using a smartphone, Quqa, et al. [1] explored the feasibility of using bicycles to extract common modal parameters of structures, namely the natural frequency modal shape and the operating deflection shape (ODS), which has been demonstrated on a footbridge in Italy.

From the above analysis, it can be seen that the research objects of the vehicle-vibration-based method mainly focus on the bridge structure or the pavement surface condition. However, the analysis of the pavement structure condition through indirect vehicle vibration needs more research. Therefore, this study intends to analyze this aspect, aiming to provide a method of characterizing pavement structural conditions based on characteristic parameters of vibration signal distribution. The result may provide a basic methodological framework for studying the pavement structural performance evolution in the long-term operation process.

This paper is organized as follows. Section 2 introduces the test conditions, including the pavement conditions and experiment conditions; Section 3 analyzes the results of the dynamic deflection test and the vehicle vibration test and then establishes a link between the two; Section 4 presents some discussion about the proposed method and Section 5 states the conclusion of this study.

2. Test Introduction

2.1. Pavement Condition

The authors conducted four vehicle vibration tests and three dynamic deflection tests on five pavement structures in Shanghai from 2018 to 2021. Five test sections were selected, including three continuously reinforced composite pavements with different structures, one semi-rigid base asphalt pavement, and one pavement that was the junction of a composite section and a semi-rigid base section. The basic parameters of the test section are shown in Table 1.

Table 1. Basic parameters of the test section.

Section	Length	Structure	Cross-Section
1	200 m	Continuously reinforced composite pavement A	Two-way 4-lane
2	200 m	Continuously reinforced composite pavement B	Two-way 6-lane
3	200 m	Continuously reinforced composite pavement B + semi-rigid base pavement	Two-way 6-lane
4	200 m	Continuously reinforced composite pavement C	Two-way 6-lane
5	200 m	Semi-rigid base pavement	Two-way 6-lane

It can be seen from Table 2 that the surface layer of the five test sections had the same structure, that is, 4 cm SMA-13 + 6 cm AC-20C. The base, subbase, and subgrade structures were quite different.

Table 2. Schematic diagram of the pavement structures.

Section 1	Section 2	Section 3		Section 4	Section 5
4 cm SMA-13 (SBS modified)	4 cm SMA-13 (SBS modified)	4 cm SMA-13 (SBS modified)	4 cm SMA-13 (SBS modified)	4 cm SMA-13 (SBS modified)	4 cm SMA-13 (SBS modified)
6 cm AC-20C (anti-rutting agent)	6 cm AC-20C (anti-rutting agent)	6 cm AC-20C (anti-rutting agent)	6 cm AC-20C (anti-rutting agent)	6 cm AC-20C (anti-rutting agent)	6 cm AC-20C (anti-rutting agent)
26 cm continuously reinforced concrete	26 cm continuously reinforced concrete	8 cm AC-25C (rock asphalt modification)	26 cm continuously reinforced concrete	26 cm continuously reinforced concrete	8 cm AC-25C (rock asphalt modification)
–	–	0.6 cm slurry seal	–	0.6 cm slurry seal	0.6 cm slurry seal
20 cm cold regeneration (three-slag base regeneration)	20 cm plain concrete	40 cm cement stabilized macadam	20 cm plain concrete	40 cm cement stabilized crushed stone leveling layer	40 cm cement stabilized macadam
	20 cm graded gravel				

2.2. Test Condition

This study carried out three FWD tests and four vibration tests on the test sections in 2018, 2019, 2020, and 2021. To avoid the interference of temperature, the temperature of these tests were close, as shown in Table 3.

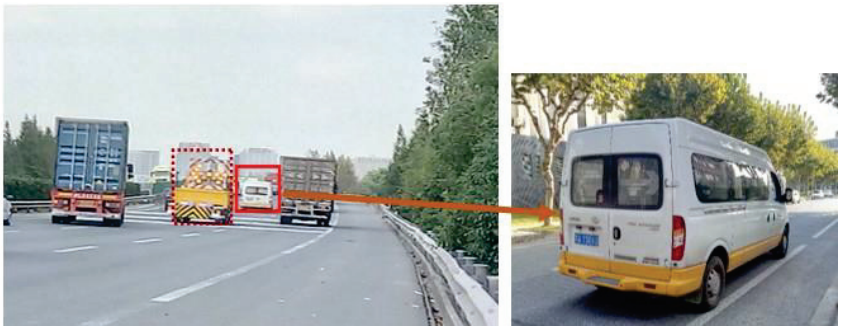
Table 3. Date and temperature.

Test	1st	2nd	3rd	4th
Date	23 September 2018	9 September 2019	20 September 2020	21 May 2021
Temperature	22~27 °C	23~28 °C	21~27 °C	22~27 °C

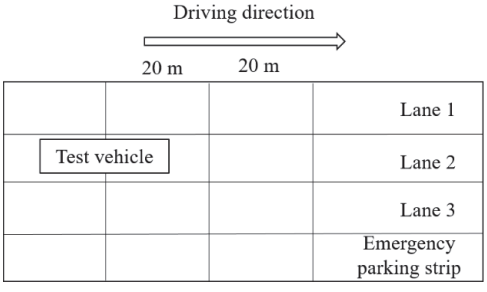
2.2.1. Deflection Test

Evaluating the pavement structure based on the collected pavement surface deflection derived from the falling weight deflectometer (FWD) is common. The multipoint vehicle-mounted falling weight deflectometer (SHN-FWD-MV) was used in this study, as displayed

in Figure 1a. The test was conducted every 20 m along the longitudinal direction. The schematic diagram of the field test is shown in Figure 1b.



(**a**)



(**b**)

Figure 1. FWD field test and the schematic diagram. (**a**) FWD field test. (**b**) The schematic diagram of FWD test.

The deflection basin data were obtained through the test results of nine displacement sensors. The distances between the nine sensors and the center point of the load were 0, 200, 300, 450, 600, 900, 1200, 1500, and 1800 mm, numbered d1~d9, and the corresponding sensor values were D1~D9. The load size was 50 kN, and the load plate radius was 15 cm.

2.2.2. Vehicle Vibration Test

The test equipment included a test car, a vibration sensor, a constant current adapter, a data acquisition card, and a laptop computer. For the placement position, some studies simply place the sensors on the vehicle floor, seat, or center console for ease of installation. Still, these locations are far away from where the tires and the pavement are in contact and have passed the vibration damping system; therefore, the vibration is relatively small and less sensitive. In this study, the vibration sensor was installed on the knuckle of the right front wheel, which is not subjected to the vibration damping system and thus is more sensitive to pavement conditions. The sensor position is shown in Figure 2. A motion camera was used to record the surface condition of the testing pavement, which was installed on the hood, as shown in Figure 3.

Increasing the sampling frequency generally can obtain more information, which is conducive to identifying pavement conditions. However, meanwhile, high sampling frequency leads to a large amount of data storage and lower computation efficiency. The sampling frequency of the vehicle vibration test was 1280 Hz.

Figure 2. The vibration sensor.

Figure 3. The motion camera.

After installing the vibration test equipment and adjusting the test parameters, we drove the test vehicle at a uniform speed on the five test sections. When a vehicle is traveling at a high speed, the tire may jump off the ground at certain positions, which is contrary to the assumption that the distance between the tire and the pavement surface is zero [24]. Therefore, the driving speed should not be too high, and the vehicle speed was maintained at 70 km/h during the test. The traffic volume of the test sections is large, and there may be congestion during the weekday, therefore, the tests were implemented on weekend afternoons.

3. Test Result Analysis

3.1. Deflection Test

Each measuring point was hammered three times during the pavement surface deflection test. The first stroke is for trial and the average deflection value of the second and third strokes was used for analysis. There were 11 measuring points in each section when the length of the test section was 200 m, and the interval was 20 m. The average value of the 11 measuring points was taken to draw the deflection basin diagram, which is shown in Figure 4. The central deflection of the five sections in the three tests is shown in Figure 5.

The central deflection (D1) is generally used to evaluate the overall structural stiffness of the pavement. The larger the D1 is, the lower the structural stiffness is. According to Figures 4 and 5:

- The structural stiffness of the continuously reinforced composite pavement (Sections 2 and 4) is better than the semi-rigid base asphalt pavement (Section 5);
- The structural stiffness of the junction (Section 3) of the two pavement structures is in the middle;

- Although Section 1 is a composite pavement, its structural stiffness is worse than that of Sections 2 and 4 because the subbase of Section 1 is weak among the five pavements (see Table 2);
- With the extension of the operating time, D1 has increased, indicating that the pavement structural stiffness has decreased.

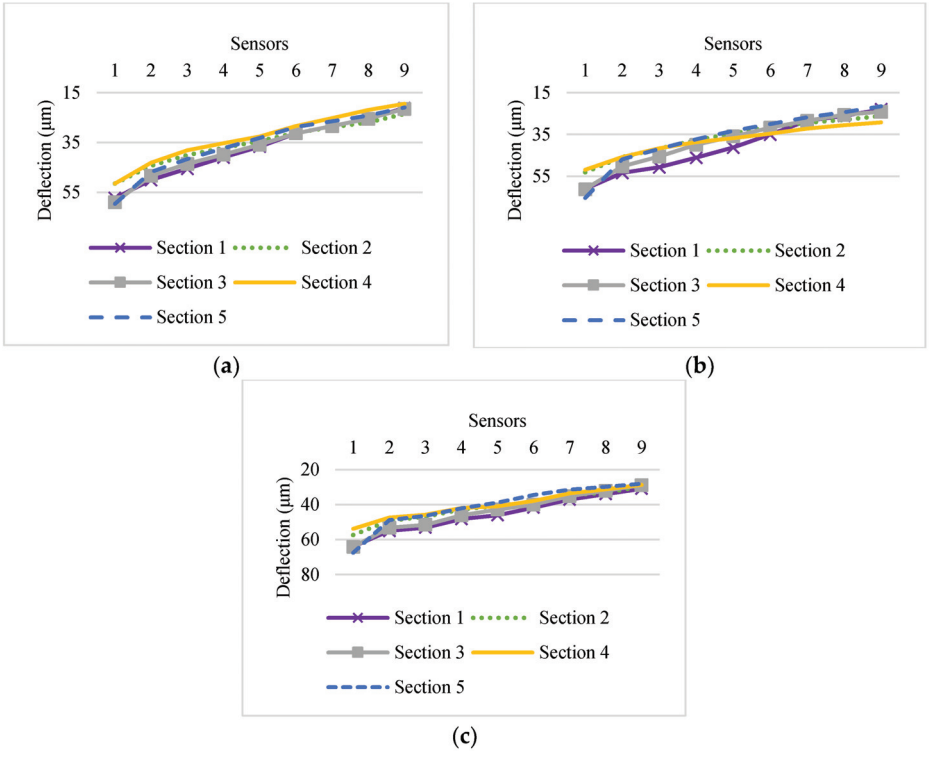


Figure 4. Deflection basin of each test. (**a**) The first test. (**b**) The second test. (**c**) The third test.

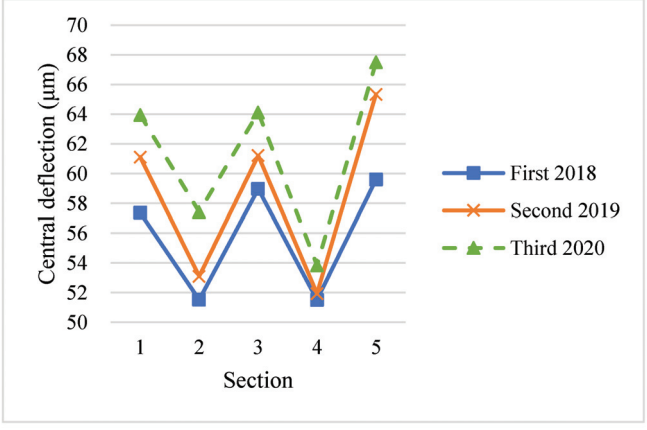


Figure 5. D1 of five sections in three tests.

3.2. Vehicle Acceleration Analysis

3.2.1. Accelerogram Analysis

The vehicle acceleration mainly comes from the road excitation while driving, and the inertia force of the engine. The vehicle acceleration when only the engine was running and that when the vehicle was driving on a road pavement is shown in Figure 6. Figure 6 shows that the acceleration induced by the engine was quite small compared with that induced by the road. It also means that the placement of the vibration sensor was rational.

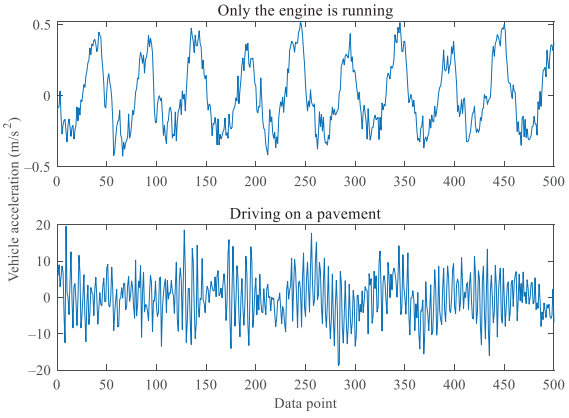


Figure 6. The vehicle acceleration when only the engine is running and when driving on a pavement.

Noises were inevitably introduced during the vibration test, including the engine noise and the instrument noise, etc. Wavelet denoising is considered to be one of the best tools in engineering signal analysis [25]. Therefore, we used the wavelet transform to denoise the raw data. The wavelet basis we used was Symlet 3 and the decomposition level was 3. These parameters are not fixed and should be determined according to the signal itself (such as signal sampling frequency, etc.) and the actual situation of engineering applications. The vibration data below were all processed after the denoising.

The time-domain waveforms of the first vehicle vibration test on the five sections are shown in Figure 7. It appears that the vehicle acceleration of Sections 3 and 5 are relatively large. Apart from that, the difference in the vehicle acceleration in each section is not obvious, and further analysis is needed.

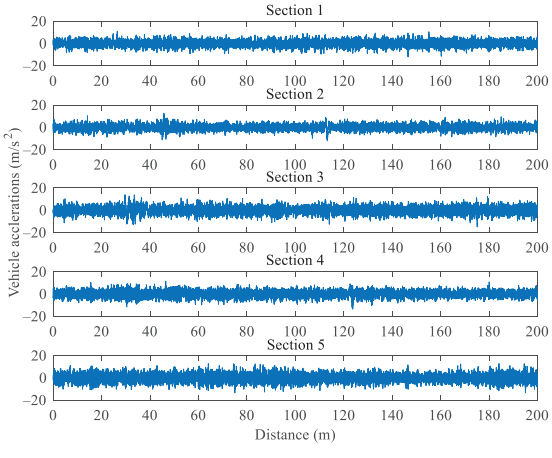


Figure 7. The time-domain diagram of the acceleration at the first test.

3.2.2. Acceleration Statistical Analysis

In order to make a statistical analysis of the vehicle vibration acceleration of 5 sections, the probability density distribution histogram of the acceleration at the first test is drawn in Figure 8. Through the Kolmogorov–Smirnov test, the vehicle acceleration data are found to be normally distributed since the *p*-value is larger than 0.05. The acceleration probability density fitting graphs of the four tests are drawn in Figure 9.

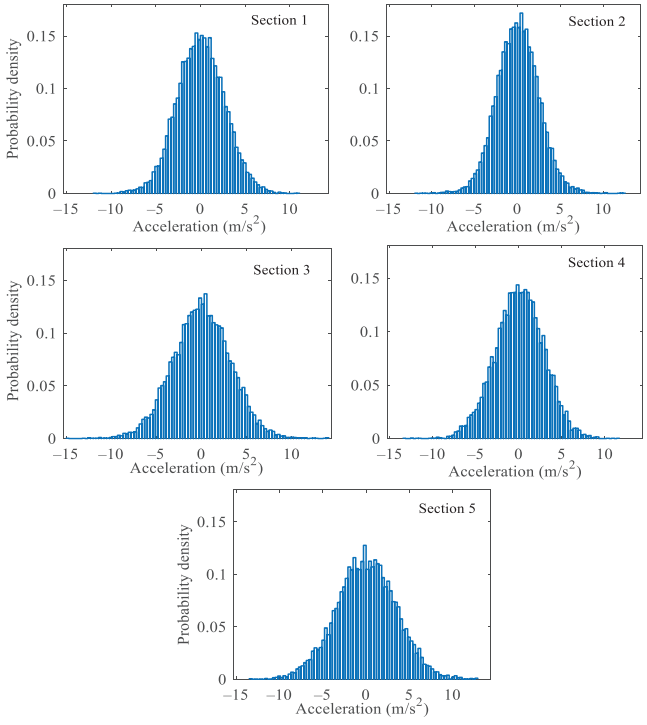
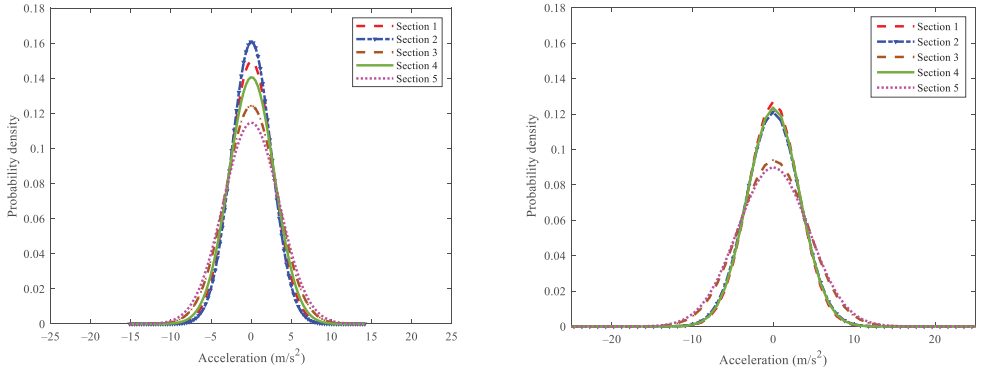


Figure 8. The probability density histogram of vehicle acceleration.



(**a**) (**b**)

Figure 9. *Cont.*

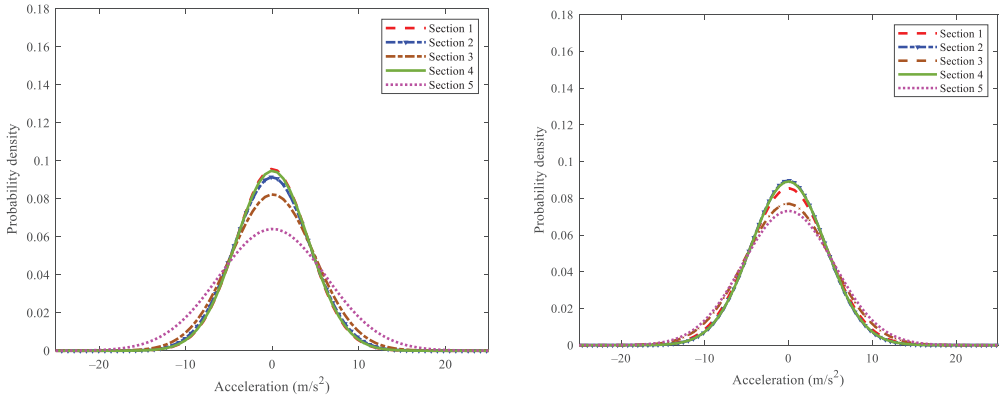


Figure 9. The normal fitting diagram of the probability density of vehicle acceleration. (**a**) The first test. (**b**) The second test. (**c**) The third test. (**d**) The fourth test.

According to Figure 9, for the continuously reinforced composite pavement (Sections 1, 2 and 4), the probability density distribution is relatively concentrated, and the curve is steeper. The probability density distribution is more scattered, and the curve is gentler for the semi-rigid base asphalt pavement (Section 5). For the junction of the two pavements (Section 3), the probability density distribution map is located between the two. From the perspective of the longitudinal time series, compared with the previous test, the probability density distribution curve of the latter test has a lower height and a larger width.

To quantify the above description, we calculated the width and height of the probability density distribution curve. The height is easy to understand, that is, the maximum probability density. Suppose the width is defined as the width of the abscissa axis, that is, the difference between the maximum acceleration and the minimum acceleration. In that case, it is easily affected by some abnormal values. Therefore, the width here is defined as the width of two inflection points on both sides of the probability density distribution curve. The point of inflection is the point where the second derivative of the function is zero.

Acceleration datum x obeys normal distribution; μ is the mean, and σ is the standard deviation. The probability density distribution function is as follows:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{1}$$

Find the second derivative of $f(x)$ as follows:

$$f''(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)\left[\left(\frac{x-\mu}{\sigma^2}\right)^2 - \frac{1}{\sigma^2}\right] \tag{2}$$

Let $f''(x) = 0$, and we can find that the function has two solutions:

$$x_1 = \mu + \sigma,\ x_2 = \mu - \sigma \tag{3}$$

In this study, the mean value of the vibration acceleration is calculated as 0; therefore, the two solutions of the function are as follows:

$$x_1 = \sigma,\ x_2 = -\sigma \tag{4}$$

Therefore, on both sides of the probability density distribution curve, there are two symmetrical inflection points, that is, the width of the curve is 2σ.

By substituting $x = \mu$ into Equation (1), the peak probability density (PPD) can be computed as follows:

$$\mathrm{PPD} = \frac{1}{\sqrt{2\pi\sigma}} \tag{5}$$

It can be seen that the height and width of the probability density distribution curve are only related to σ; hence we only need one parameter. The height is more evident in the figure than the width, and the final parameter selected is the peak height of the probability density distribution curve, that is, PPD.

3.3. Correlation Analysis of Vehicle Acceleration and Pavement Surface Deflection

Table 2 shows that the surface layer of the five test sections has the same structure, namely 4 cm SMA-13 and 6 cm AC-20C. Moreover, the surface condition is quite similar according to the video recorded by the motion camera. Therefore, it is reasonable to compare the vehicle acceleration and pavement surface deflection.

Pearson's correlation coefficient is commonly used to measure the correlation between two variables (the peak probability density of acceleration and central deflection). In statistics [26], Pearson's correlation coefficient (r) can measure the correlation between two variables, $-1 \leq r \leq 1$. If r is positive, there is a positive correlation between the two variables, and the larger the value of r is, the stronger the correlation is. A negative r value means the two variables have a negative correlation; the smaller the value, the stronger the correlation. There is no correlation between the two variables if r is 0.

The relationship between the correlation degree of the variable and $|r|$ is shown in Table 4.

Table 4. Correlation degree of the variable.

$\|r\|$	0.8~1.0	0.6~0.8	0.4~0.6	0.2~0.4	0~0.2
Correlation degree	Very strong	Strong	Middle	Weak	Very weak or no correlation

For the peak probability density (PPD) and the center deflection (D1), their scatter diagram is shown in Figure 10. The greater the acceleration probability density is, the smaller the center deflection is. There is a linear correlation between the two variables so that the Pearson correlation coefficient can be used for evaluation.

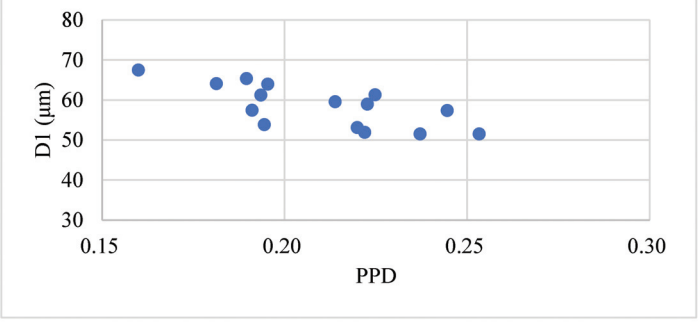


Figure 10. Relationship between PPD and D1.

The Pearson correlation coefficient (r) between PPD and D1 is calculated to be -0.722. According to Table 4, there is a strong negative correlation between them. D1 is negatively correlated with pavement structural stiffness. Therefore, it can be concluded that a strong positive correlation exists between PPD and the pavement structural stiffness. The greater

the PPD is, the higher the stiffness of the pavement structure is. The reason may be that the stronger pavement causes the vehicle to vibrate more frequently. Frequent vibrations will cause the vehicle acceleration to cross zero more frequently, which means that the probability of a vehicle acceleration value being zero is greater. When the abscissas are equally spaced, greater probability means higher probability density. This result preliminarily proves that the vehicle acceleration can reflect the pavement's structural stiffness.

4. Discussions

A pavement structural survey is a complex issue that requires either laborious work or advanced equipment. Assessing a pavement's structural stiffness through vehicle vibration signals has the merits of convenience, low cost, and being non-destructive. However, it is a relatively novel approach that lacks sufficient research. A vibration-based pavement structural survey is challenging, and this study analyzed its feasibility.

According to the result of this study, the structural difference between the continuously reinforced composite pavement and semi-rigid base asphalt pavement is easy to distinguish. However, classifying three types of continuously reinforced composite pavement (A, B and C) is still arduous because their pavement structure is similar. In our future research, we need to improve the proposed method by adjusting the testing vehicle (such as increasing the weight) and developing thorough data analysis methods.

The road surface is paved with a temperature-influenced material, the asphalt mixture. Temperature changes would lead to changes in the pavement structural stiffness. Therefore, the surface deflection obtained by the FWD test usually requires temperature correction. The vehicle vibration test obtains the vehicle acceleration through the contact between the tire and the road, so the data would also be affected by the temperature. In this study, each vehicle vibration test and the corresponding deflection test were carried out on the same day. The temperatures were relatively similar, avoiding the influence of ambient temperature. However, in future research, the impact of temperature changes on the test data and results should be considered for a more detailed analysis.

5. Conclusions

In this study, a preliminary exploration was made on the evaluation of pavement structure conditions based on vehicle vibration. For the five pavement structures, four vehicle vibration tests and three pavement surface deflection tests were conductedfrom 2018 to 2021. According to the results of the FWD test, the pavement structure stiffness is relatively high for the continuously reinforced composite pavement. The structural stiffness of the semi-rigid base asphalt pavement is poor. For the junction of the composite pavement and the semi-rigid base pavement, the structural stiffness is middling. With the increase in the operating time, the pavement structural stiffness decreased.

According to the vehicle vibration test result, the peak probability density (PPD) of vehicle vibration acceleration was relatively large for the pavement section with high structural stiffness; the value was relatively small for the pavement with poor structural stiffness. However, for the three types of continuously reinforced composite pavements with slight differences in structural stiffness, it is difficult to describe the structural differences using this indicator.

The correlation analysis between the PPD of vehicle vibration acceleration and the center deflection (D1) shows that the correlation coefficient is -0.722, which is a strong positive correlation. Therefore, it can be inferred that the vehicle vibration acceleration can reflect the stiffness of the pavement structure to a certain extent, but the exact relationship between the two needs further research and analysis. This may be due to the lightweight of the test car and the insufficient sensitivity of the structure. Subsequent tests can be carried out with heavier vehicles in order to have a higher sensitivity to structural performance.

Author Contributions: Conceptualization, W.W.; funding acquisition, Q.Y.; investigation, W.W. and S.Z.; methodology, W.W.; software, Q.Y.; writing—original draft, S.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China, grant number 51778482.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author, Wang Weiguo, 230188209@seu.edu.cn.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Quqa, S.; Giordano, P.F.; Limongelli, M.P. Shared micromobility-driven modal identification of urban bridges. *Autom. Constr.* **2021**, *134*, 104048. [CrossRef]
2. Zhao, H.D.; Wu, D.F.; Zeng, M.Y.; Zhong, S. A Vibration-Based Vehicle Classification System using Distributed Optical Sensing Technology. *Transp. Res. Rec.* **2018**, *2672*, 12–23. [CrossRef]
3. Guattari, C.; D'Amico, F.; Benedetto, A. Integrated road pavement survey using GPR and LFWD. In Proceedings of the XIII Internarional Conference on Ground Penetrating Radar, Lecce, Italy, 21–25 June 2010.
4. Mehta, Y.; Roque, R. Evaluation of FWD data for determination of layer moduli of pavements. *J. Mater. Civ. Eng.* **2003**, *15*, 25–31. [CrossRef]
5. Lai, J.C.; Liu, J.; Huang, C.W. The Application of Frequency-Temperature Superposition Principle for Back-Calculation of Falling Weight Deflectometer. *Appl. Sci.* **2020**, *10*, 132. [CrossRef]
6. Xie, Z.; Cong, L.; Guo, Z. Evaluation of Subsoil Modulus for Asphalt Pavement Based on FWD Deflection Basin Parameter. *J. Highw. Transp. Res. Dev.* **2009**, *26*, 28–31.
7. Hou, Y. Evaluation of Layered Material Properties from Measured Surface Deflection. Ph.D. Thesis, University of Utah, Salt Lake, UT, USA, 1977.
8. Li, M.Y.; Wang, H. Development of ANN-GA program for backcalculation of pavement moduli under FWD testing with viscoelastic and nonlinear parameters. *Int. J. Pavement Eng.* **2019**, *20*, 490–498. [CrossRef]
9. Seraj, F.; Van Der Zwaag, B.J.; Dilo, A.; Luarasi, T.; Havinga, P. Roads: A road pavement monitoring system for anomaly detection using smart phones. In *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Cham, Switzerland, 2016; pp. 128–146.
10. Chuang, Y.; Perng, N.H.; Han, J.Y. Pavement performance monitoring and anomaly recognition based on crowdsourcing spatiotemporal data. *Autom. Constr.* **2019**, *160*, 102882. [CrossRef]
11. Liu, C.L.; Wu, D.F.; Li, Y.S.; Jiang, S.C.; Du, Y.C. Mathematical insights into the relationship between pavement roughness and vehicle vibration. *Int. J. Pavement Eng.* **2020**. [CrossRef]
12. Harikrishnan, P.M.; Gopi, V.P. Vehicle Vibration Signal Processing for Road Surface Monitoring. *IEEE Sens. J.* **2017**, *17*, 5192–5197. [CrossRef]
13. Yang, Q.; Zhou, S.S. Identification of asphalt pavement transverse cracking based on vehicle vibration signal analysis. *Road Mater. Pavement Des.* **2021**, *20*, 1780–1798. [CrossRef]
14. Mednis, A.; Strazdins, G.; Zviedris, R.; Kanonirs, G.; Selavo, L. Real time pothole detection using Android smartphones with accelerometers. In Proceedings of the 2011 International Conference on Distributed Computing in Sensor Systems (DCOSS 2011), Barcelona, Spain, 27–29 June 2011.
15. Yang, Y.B.; Lin, C.W.; Yau, J.D. Extracting bridge frequencies from the dynamic response of a passing vehicle. *J. Sound. Vib.* **2004**, *272*, 471–493. [CrossRef]
16. Lin, C.W.; Yang, Y.B. Use of a passing vehicle to scan the fundamental bridge frequencies: An experimental verification. *Eng. Struct.* **2005**, *27*, 1865–1878. [CrossRef]
17. Yang, Y.B.; Yau, J.D.; Urushadze, S. Scanning the modal coupling of slender suspension footbridges by a virtual moving vehicle. *Eng. Struct.* **2019**, *180*, 574–585. [CrossRef]
18. Yang, Y.; Jia, B.Y.L. Structural damage identification of a bridge from a passing test vehicle using frequency domain method. In *Mechanics of Structures and Materials: Advancements and Challenges, Proceedings of the 24th Australasian Conference on the Mechanics of Structures and Materials, ACMSM24 2016, Perth, Australia, 6–9 December 2016*; CRC Press/Balkem: Perth, Australia, 2017; pp. 1125–1130.
19. Yang, Y.B.; Yang, J.P. State-of-the-Art Review on Modal Identification and Damage Detection of Bridges by Moving Test Vehicles. *Int. J. Struct. Stab. Dyn.* **2018**, *18*, 1850025. [CrossRef]
20. Cerda, F.; Garrett, J.; Bielak, J.; Barrera, J.; Zhuang, Z.; Chen, S.; McCann, M.; Kovaevi, J.; Rizzo, P. Indirect structural health monitoring in bridges: Scale experiments. In Proceedings of the Sixth International Conference on Bridge Maintenance, Safety and Management, Stresa, Italy, 8–12 July 2012; pp. 346–353.
21. Liu, J.; Chen, B.; Chen, S.; Berges, M.; Bielak, J.; Noh, H. Damage-Sensitive and Domain-Invariant Feature Extraction for Vehicle-Vibration-Based Bridge Health Monitoring. In Proceedings of the ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing, Barcelona, Spain, 4–8 May 2020; pp. 3007–3011.

22. Feng, M.; Fukuda, Y.; Mizuta, M.; Ozer, E. Citizen sensors for SHM: Use of accelerometer data from smartphones. *Sensors* **2015**, *15*, 2980–2998. [CrossRef] [PubMed]
23. McGetrick, P.J.; Hester, D.; Taylor, S.E. Implementation of a drive-by monitoring system for transport infrastructure utilising smartphone technology and GNSS. *J. Civ. Struct. Health Monit.* **2017**, *7*, 175–189. [CrossRef]
24. Du, Y.C.; Liu, C.L.; Wu, D.F.; Jiang, S.C. Measurement of International Roughness Index by Using Z-Axis Accelerometers and GPS. *Math. Probl. Eng.* **2014**, *2014*, 928980. [CrossRef] [PubMed]
25. Wang, J.; Tang, X. Wavelet Denoise Method Applied in Load Spectrum Analysis of Engineering Vehicles. *Adv. Mater. Res.* **2010**, *108*, 1320–1325. [CrossRef]
26. David, S.M.; William, N. Methods of describing correlations: Scatter diagrams and correlation coefficients. In *Statistics: Concepts and Controversies*, 8th ed.; CMC Publishing House: Golden, CO, USA, 2017; pp. 280–385.

sensors

MDPI

Article

Crack Detection in Frozen Soils Using Infrared Thermographic Camera

Yang Zhao [1], Yufan Han [2], Cheng Chen [1] and Hyungjoon Seo [3,*]

[1] Department of Civil Engineering, Xi'an Jiaotong Liverpool University, Suzhou 215000, China; yang.zhao@xjtlu.edu.cn (Y.Z.); cheng.chen19@student.xjtlu.edu.cn (C.C.)
[2] Department of Computer Science, Xi'an Jiaotong Liverpool University, Suzhou 215000, China; Yufan.Han20@student.xjtlu.edu.cn
[3] Department of Civil Engineering and Industrial Design, University of Liverpool, Liverpool L69 7ZX, UK
* Correspondence: hyungjoon.seo@liverpool.ac.uk

Abstract: Frozen soils are encountered on construction sites in the polar regions or regions where artificial frozen ground (AFG) methods are used. Thus, efficient ways to monitor the behavior and potential failure of frozen soils are currently in demand. The advancement of thermographic technology presents an alternative solution as deformation occurring in frozen soils generate heat via inter-particle friction, and thus a subsequent increase in temperature. In this research, uniaxial compression tests were conducted on cylindrical frozen soil specimens of three types, namely clay, sand, and gravel. During the tests, surface temperature profiles of the specimens were recorded through an infrared video camera. The thermographic videos were analyzed, and subsequent results showed that temperature increases caused by frictional heat could be observed in all three frozen soil specimens. Therefore, increases in temperature can be deemed as an indicator for the potential failure of frozen soils and this method is applicable for monitoring purposes.

Keywords: frozen soil; infrared camera; crack

Citation: Zhao, Y.; Han, Y.; Chen, C.; Seo, H. Crack Detection in Frozen Soils Using Infrared Thermographic Camera. *Sensors* **2022**, *22*, 885. https://doi.org/10.3390/s22030885

Academic Editors: Phong B. Dao, Lei Qiu, Liang Yu and Zahra Sharif Khodaei

Received: 8 December 2021
Accepted: 21 January 2022
Published: 24 January 2022

1. Introduction

Historically, humans lived and built structures on frozen ground among Arctic areas where annual average temperatures were below the freezing point of water. Artificial frozen ground (AFG) methods have recently gained popularity as a technique to stabilize soil during excavation. Due to the bonding effects of ice, frozen soil is stronger and more rigid than regular soil and may demonstrate mechanical behavior similar to concrete. The crack initiation and propagation of concrete has been studied by numerous researchers, but few studies have focused on cracks in frozen soil.

As cracking occurs, there is a relative movement of soil particles around the crack, and hence frictional heat. The frictional heat may cause temperature increases, which can be observed using a thermographic device such as an infrared camera. Researchers successfully used an infrared camera to detect various phenomena with temperature changes, such as oil products spreading on water surfaces [1], wild fires [2], and wind flow [3]. Specifically, the thermographic technique was applied to structural defects. The behavior of rock and soil was determined through monitoring with infrared thermal cameras [4,5]. Liu, et al. [6] applied infrared monitoring in an experimental study of a tunnel. Seo, et al. [7], Seo [8] detected crack formation in pillars using an infrared camera. Moreover, an infrared thermographic camera was applied in ice detection on wind turbine blades [9] and aircraft air foils [10].

Though in some cases, the defects fail to generate temperature changes when the structure is at rest. However, when heated by an external source, defects will display different temperature compared with the rest of the structure. Broberg [11] used an infrared camera to detect welding defects based on the temperature difference between defects and

surrounding surfaces while the weld was heated by a flash lamp as an external IR source. Štarman and Matz [12] observed the propagation of artificially generated thermal pulses in steel bars using an infrared camera to locate the presence of cracks. Afshani, et al. [13] conducted a study to detect defects in the lining of a tunnel with an infrared thermal camera. Recently, deep learning and machine learning analyses were applied to find a crack in infrastructures [14,15]. Moreover, the three-dimensional monitoring system was applied for monitoring the displacement and tilt of infrastructure using laser scanning [16]. In this paper, an experimental study was conducted by simulating frozen soils. The cracks of different frozen soils were identified through infrared topography.

2. Methods

2.1. Specimen Preparation

Three cases of soil category were considered in this study, namely clay, sand, and gravel. For the clay specimen, an undisturbed clay sample extracted from a depth of 3 m was cut into a cylinder of 100 mm (diameter) by 150 mm (height). Next, the sample was submerged in a water tank for 24 h until reaching saturation. Afterward, it was placed in a freezer for at least 24 h. For the sand and gravel specimens, soil particles and water were poured into molds iteratively to ensure the specimen was fully saturated and composition was relatively uniform across the height. Then, they were frozen in a freezer before being removed from the mold. Subsequently, the specimen was again stored in the freezer for 24 h before the test. Figure 1 shows the process of forming the frozen sand specimen. Due to mechanical disturbances during the removal of the specimen, the height of the sand and gravel specimen was not strictly controlled. The properties of the specimens is shown in Table 1 and photos of specimens were taken before uniaxial compression tests as shown in Figure 2.

Figure 1. Preparation of sand specimen.

Table 1. Properties of specimens (N.M. represents no measurement).

Material	Height (mm)	Diameter (mm)	Weight (kg)	Density (kg/m^3)	Specific Gravity	Porosity
clay	150	100	N.M.	N.M.	N.M.	N.M.
sand	171	100	2.62	1948.76	2.25	0.24
gravel	179	100	2.75	1952.96	2.51	0.37

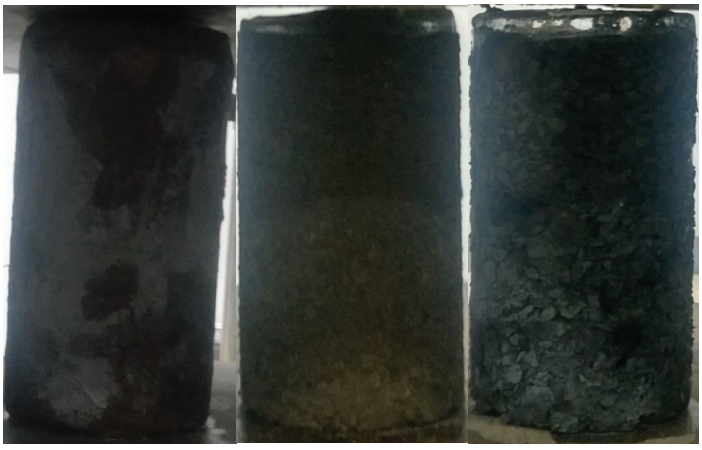


Figure 2. Frozen soil specimens before uniaxial compression tests (clay, sand, and gravel from left to right).

2.2. Thermal Graphic Imaging

Temperature changes can be detected remotely and non-intrusively. Additionally, the thermograph provides 2D geometric information. Therefore, an infrared camera was used in this research to detect crack formation in frozen soils. The infrared camera used was the FLIR E60. The exported results file consisted of sequences of thermographic photos taken at a frequency of 30 Hz and resolution of 320 * 240 pixels. Each pixel in a frame had its temperature measured. The exported thermographic files were analyzed using the software FLIR Research IR. The thermal sensitivity was 0.05 °C and the accuracy was ±2 °C.

All bodies above 0 K (Kelvins) emit electromagnetic radiation. The characteristic of the radiation depends on the temperature of the body, thus the temperature of the body can be determined by the measured radiance. Infrared is a section of electromagnetic radiation with wave length ranging from 780 nm to 1 mm [17]. At room temperature, the majority of radiation energy lies within the range of infrared, which makes an infrared sensor the ideal selection for the purpose of temperature measurement. When an object is photographed using an infrared camera, the reflected infrared radiation emitted by the object is received by the infrared camera's sensor and simultaneously converted to an electronic signal. Then, the electronic signal is processed by the controlling software to generate a thermal graphic image or, in short, a thermograph.

The governing equation for the relationship between thermal radiation intensity and temperature is shown by Stefan–Boltzmann's law (1)

$$W = \epsilon \sigma T^4 \tag{1}$$

where W is radiance intensity (W/m^2) measured by the infrared sensor, Stefan–Boltzmann constant $\sigma = 5.67 \times 10^{-8} \frac{W}{m^2 \cdot K^4}$, T is the temperature of the object (K), ϵ is emissivity. In this research, software FLIR Research and Design [18] was used to process the infrared measurement and the back-calculation from the measured radiance intensity to temperature based on Equation (1).

Ideally, to obtain the precise relationship of values W and T, it is necessary that we quantify the values of ϵ precisely at any given time and location, which requires robust and sophisticated calibration and strict control over atmospheric conditions. Such requirements are practically difficult and unnecessary if the sole purpose concerns the occurrence of abnormal changes in temperature rather than measuring the exact value of the temperature. In this research, $\epsilon = 0.95$ is assumed.

2.3. Test Procedures

In this research, uniaxial compression tests were conducted on frozen soil specimens while the specimens were subjected to an infrared camera. The arrangement of the experimental apparatus is shown in Figure 3.

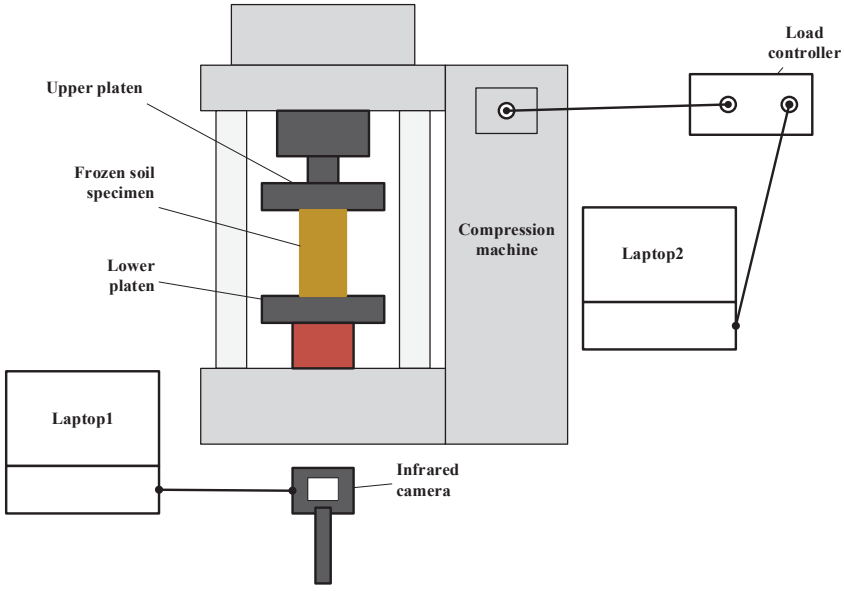


Figure 3. Experimental apparatus setup.

The infrared camera was placed approximately 1 m in front of the compression machine. The load was recorded at a frequency of 10 Hz. The T-load curve was visible simultaneously on laptop 2 controlling the compression machine. The uniaxial compression test was terminated when the specimen failed or the lift of the piston reached its capacity.

3. Result and Analysis

Tests were conducted on 6 samples for each soil type where multiple samples for each soil type can be seen in Figure 1. However, we were unable to observe the cracking behaviors on all of the tests using an infrared camera due to the following three conditions: (a) cracks occurred on the backsides of the samples, (b) samples disintegrated without cracking due to melting (see Figure 4a,b), and (c) samples yielded without cracking (see Figure 4c). Cracking occurred on either one or both sides of the specimen by chance, which cannot be practically controlled. Yielding and disintegrating of specimens occurred as the specimens were partially melted. The authors indeed attempted to repeat the tests under a lower room temperature (around 15 °C) in the winter, however, satisfactory improvements on avoiding the melting of specimens were not achieved. To avoid the melting of specimens, the compression tests on specimens should be conducted under temperatures below the freezing point of water [19–21]. In the previous research [19–21], the low temperature of specimens during testing was maintained by immersing the specimens in liquid cool materials. Unfortunately, such cooling systems were not available to the authors and restricted the direct observation of specimens using the infrared thermographic camera during compression tests. Therefore, specimens were not guaranteed to be completely frozen during our tests. One sample of each soil type was analyzed as they demonstrated evident cracking behavior.

(**a**) (**b**) (**c**)

Figure 4. Failure of specimens without cracking behavior: (**a**) gravel; (**b**) sand; and (**c**) clay.

Due to a lack of pre-existing knowledge of conditions under which a crack would appear in thermograph of frozen soils, the thermographs were carefully examined frame by frame. The qualification of a crack can be qualitatively described as follows:

(a) Significant temporal temperature variation, compared to the immediate surrounding area (baseline), occurring at certain spots of a relatively small area.
(b) The temperature variation initiates within the specimen surface and is not transmitted from the interface between the specimen and the ambient environment.
(c) The temperature variation is sustained for more than 1 frame, which disqualifies false positives caused by random flocculation in the measured value of temperature.
(d) The temperature variation is not caused by mass transportation i.e., movement of disintegrated soil particles or water flow thawed from ice.

According to the above description, two types of curves are plotted to demonstrate the temperature change at a crack point. The T-t curve shows how the temperature at the crack point varies temporally. The T-d curve shows how the temperature at the crack point varies spatially. Moreover, the load-time curve is used to qualitatively determine the strain-stress status of the specimen. It was assumed that the frozen soil specimens displayed elastic-plastic behavior. The period before the load peaks is defined as the elastic stage and the period after as the plastic stage. To convert the unit of distance measured in the thermograph from pixel to mm, the thickness of the lower platen of the compression machine was used to calculate the conversion ratio from pixel to mm. The calculated conversion ratios for frozen clay, sand and gravel were 0.60, 0.62, and 0.65 mm/pixel, respectively.

3.1. Crack in Frozen Clay

As shown in the load-time curve in Figure 5, the load peaked at 96.30 s after the test commenced. After 96.30 s, the temperature of certain points on the specimen's surface appears to increase at a slightly higher rate than the rest. One example is indicated in Figure 6 as the crack point. A baseline point about 4 mm downwards of the crack point was selected to mitigate the effect of measurement errors and the effect of heating from the ambient atmosphere on the variation of temperature. The ΔT (temperature change)-t (time) curve for the crack point and baseline points are plotted in Figure 7. The temperature profile was set relative to temperature changes, with 0 chosen as the initial temperature at the reference point. The temperature changes were filtered by implementation of a Savitzky–Golay filter.

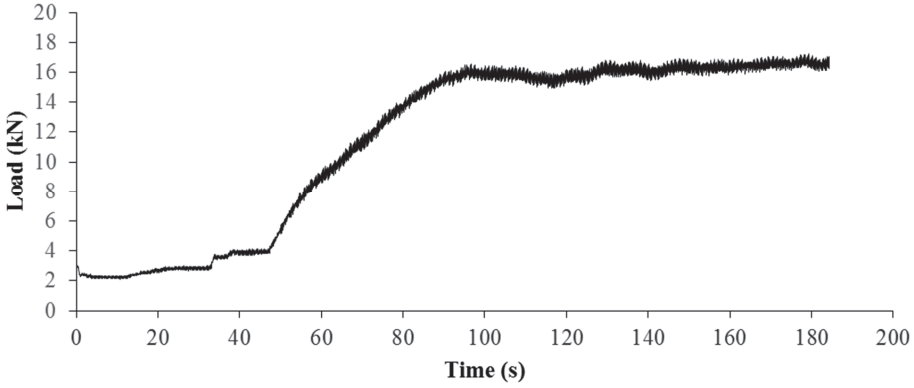


Figure 5. Uniaxial load-time curve of clay specimen.

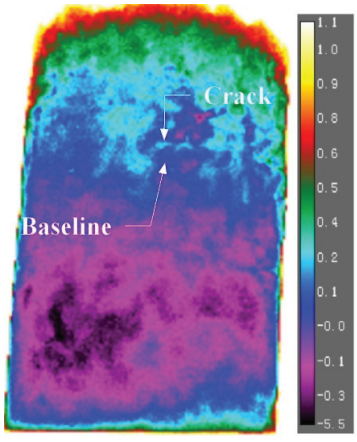


Figure 6. Temperature profile of clay specimen (unit of scale bar is in °C).

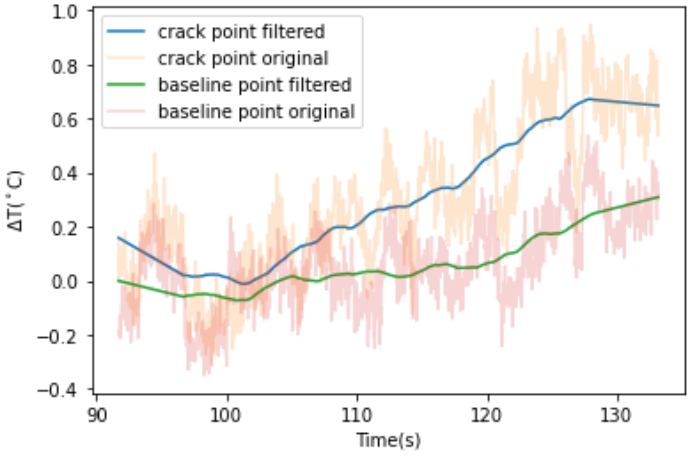


Figure 7. Temperature–time curve of crack point and baseline for clay specimen.

Disregarding the random flocculation, from 105 s to 123 s, the temperature increases from −3.02 °C to −2.41 °C.

The temperature increase at the crack point was further verified by the temperature profile along a line approximately perpendicular to the crack as shown in Figure 8. The temperature profile at the start of the test (0 s), immediately before the crack occurs (105 s), and after the crack forms (125 s), are plotted. Before the crack occurs, there is no significant difference in temperature between the potential crack point and the rest. After the crack is formed, the temperature at the crack point is 0.57 °C higher than those at the points not influenced by the crack.

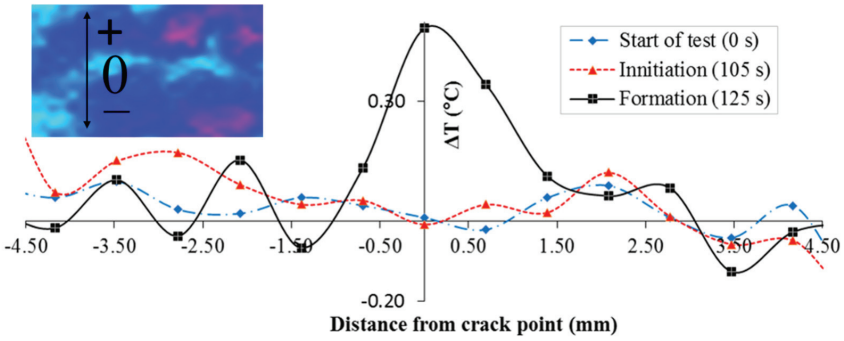


Figure 8. Temperature profile along a measurement line approximately perpendicular to the crack (the position of the measurement line is marked on the temperature profile attached on the top left).

3.2. Bulge Effect in Frozen Sand

Although there was no individual crack observed in the sand case, as the frozen sand specimen was compressed, the expansion in the radial direction, which we termed bulge, became significantly visible after yield and the observation can also be correlated to temperature variations. The temperature of three representative points located at the upper part (U.), lower left part (L.L), lower right part (L.R), and the average temperature of a square area (S.) were selected to demonstrate the effect of the bulge on temperature. The locations are indicated in Figure 9. In Figure 9, temperatures are mostly above 0 °C as the surface of the specimen was covered by water instead of ice. The ambient air temperature was around 30 °C. The ice on the top of the specimen melted immediately after contact between the specimen and upper platen of the compression machine. Due to gravitational effects, the water subsequently ran off downwards onto the specimen surface. Although the specimen was completely frozen initially, certain parts melted during the compression test. The same observation was made for the clay specimen and gravel specimen.

The temperature and load curves are plotted together in Figure 10. As the bulge of the specimen becomes evident after yield (64.9 s), the temperature of each single point of the specimen surface demonstrated 4 simultaneous impulsive increases initiating at 66.80 s, 88.30 s, 99.53 s, and 110.30 s, respectively. Impulses were not observed at points outside of the specimen surface, which signifies these impulses were not measurement errors. The magnitude of ΔT for the impulses ranged from 0.37 to 0.95 °C.

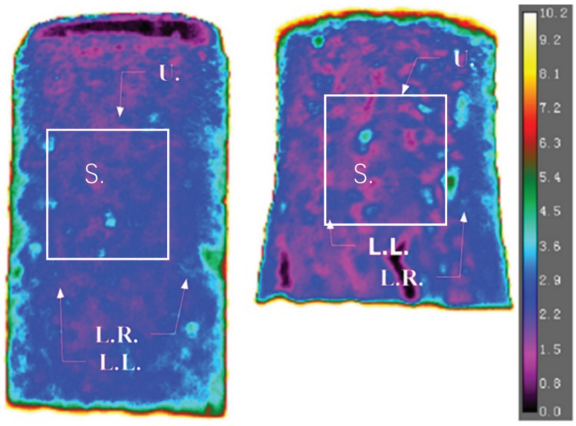


Figure 9. Temperature profile of frozen sand specimen at the start (**right**) and end (**left**) of the compression test (unit of scale bar is in °C).

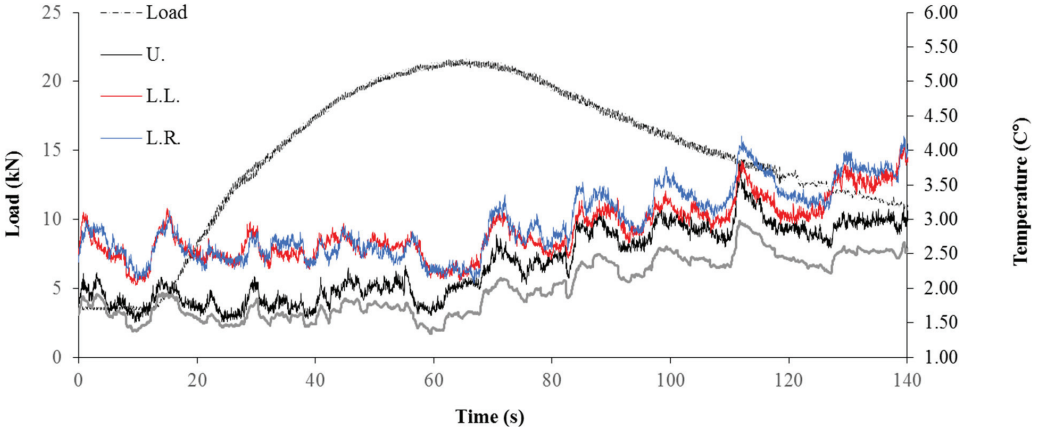


Figure 10. Load curve and temperature–time curve of sand specimen.

3.3. Cracks in Frozen Gravel

Two types of cracks were observed during the compression test of the gravel specimen. The cracks, which occurred in-between gravel particles, caused increases of temperature at the crack point. Such cracks were denoted as I1, I2 I3, and I4 for 'increase'. Temperature decreases at the crack point were observed for cracks which occurred within the ice block. Such cracks were denoted as D1, D2, D3, D4, D5, and D6 following the capitalized initial of the word 'decrease'. The ideal temperature variation for cracks of types I (increase) and D (decrease) are illustrated in Figure 11. When cracks occur between gravel particles, the inter-particle friction generates heat. However, those frictional forces were relatively negligible between ice surfaces due to their smoothness. Although there was heat generated around ice particles, the heat would probably be consumed by the melting of ice as opposed to an increase in temperature. The temperature decrease is due to a temperature gradient travelling from the surface to the inner core of the specimen. As the crack widens, the inner surface of the specimen, the temperature of which is lower compared to the outside, is exposed to the camera. The location of these two types of cracks is indicated in Figure 12.

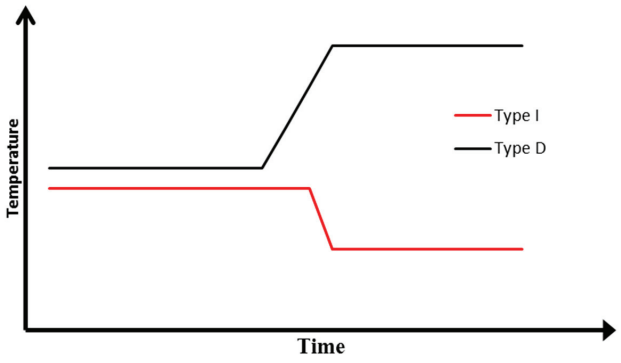
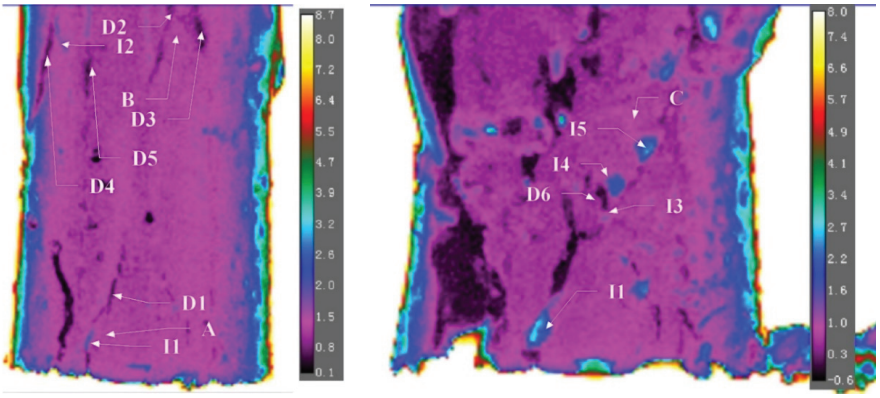


Figure 11. Ideal temperature–time curves of crack type I and type D.



(**a**) (**b**)

Figure 12. Locations of cracks before plastic deformation (**a**) and after plastic deformation (**b**).

Temporal variations for cracks I1, D2 and D5 are shown in Figures 13a, 13b and 13c, respectively. Crack I1 occurs at the time of yield. Before the crack initiates, the temperature of the crack point is relatively constant and approximately equal to the baseline temperature. The ΔT-t curve of the crack point deviates from that of the baseline since t = 126.33 s. From t = 126.33 to t = 128.80 s, the temperature at the crack point increases by 0.6 °C while that at the baseline remains relatively stable. After the crack is formed, the temperature at the crack remains relatively steady from 128.90 to 135.33 s but subsequently decreases due to the decrease in load, as shown in Figure 14.

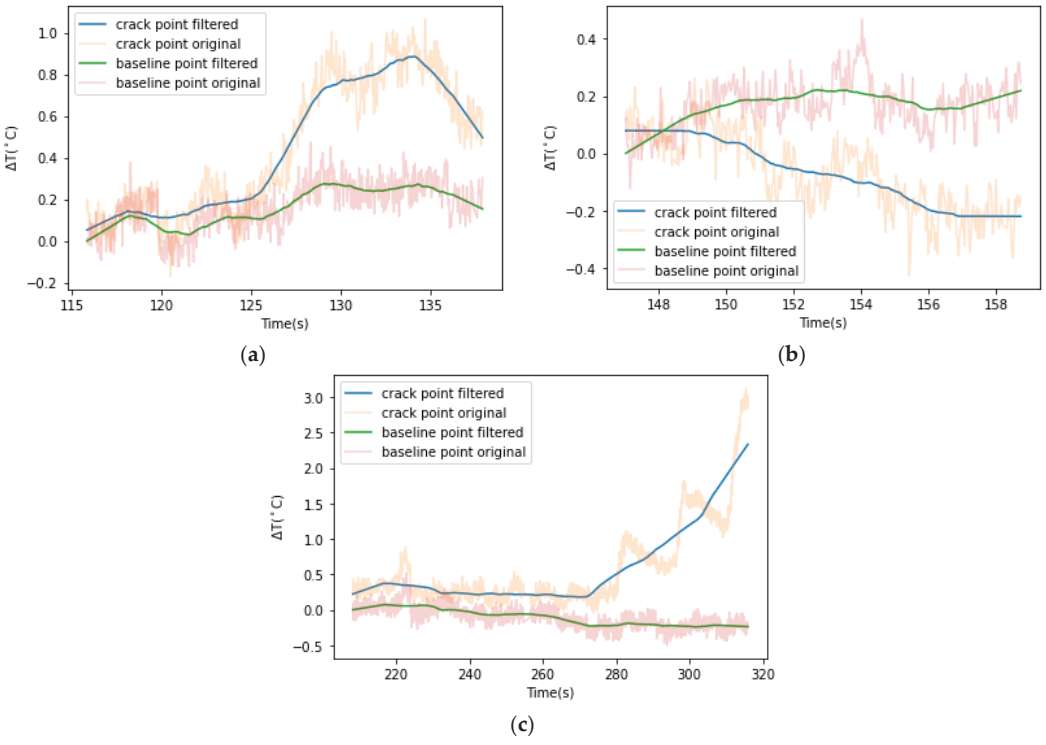


Figure 13. Temporal temperature variation at crack I1 D2 I5 and their immediate baseline A, B, C: (**a**) crack I1; (**b**) crack D2; and (**c**) crack D5.

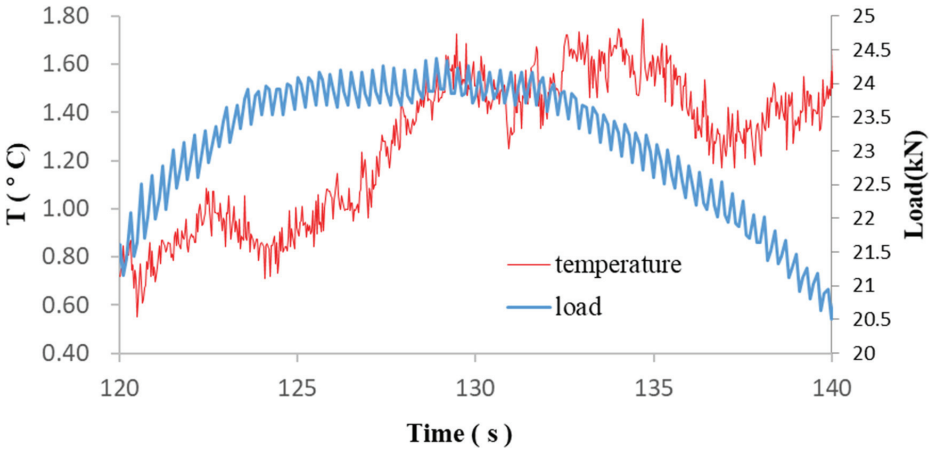


Figure 14. Variation of load and temperature with time at crack point I1.

Crack D2 forms at the transitional point between stress softening and plastic deformation, which takes less time than the formation of crack I1. This is compatible with the explanation that there is minimal friction and ice is extremely brittle The magnitude of the temperature decrease is less important than the temperature increase for the crack between

gravel particles as the decrease only demonstrates the depth of the crack that develops into the specimen. Crack I5 occurs along the global failure surface during plastic deformation. The crack occurred at 265 s with a temperature increase of 0.51 °C. Following the relative movement between two parts of the specimen on each side of the failure, the surface starts at a 280 s to display a temperature increase that takes 3 steps, which corresponds to the steps of relative movement along the failure surface observed in the video. The load, and hence the stress within the specimen, is less than those observed previously during the elastic stage when I1 occurs. However, the relative movement along the global failure surface is more intense. Thus, both the rate of increase and the total amount of increase in temperature is larger. From the 280 s to 315 s, the temperature increase due to friction along the global failure surface is 2.66 °C. The spatial temperature distribution across I1, D2, and I5 are also plotted to demonstrate the increase during the formation of the crack in Figure 15. ΔT_{I1} is defined as

$$\Delta T_{I1} = T_{I1} - T_{average} \quad (2)$$

where T_{I1} is the temperature at crack I1 at formation and $T_{average}$ is the average at the immediate surrounding area of crack I1.

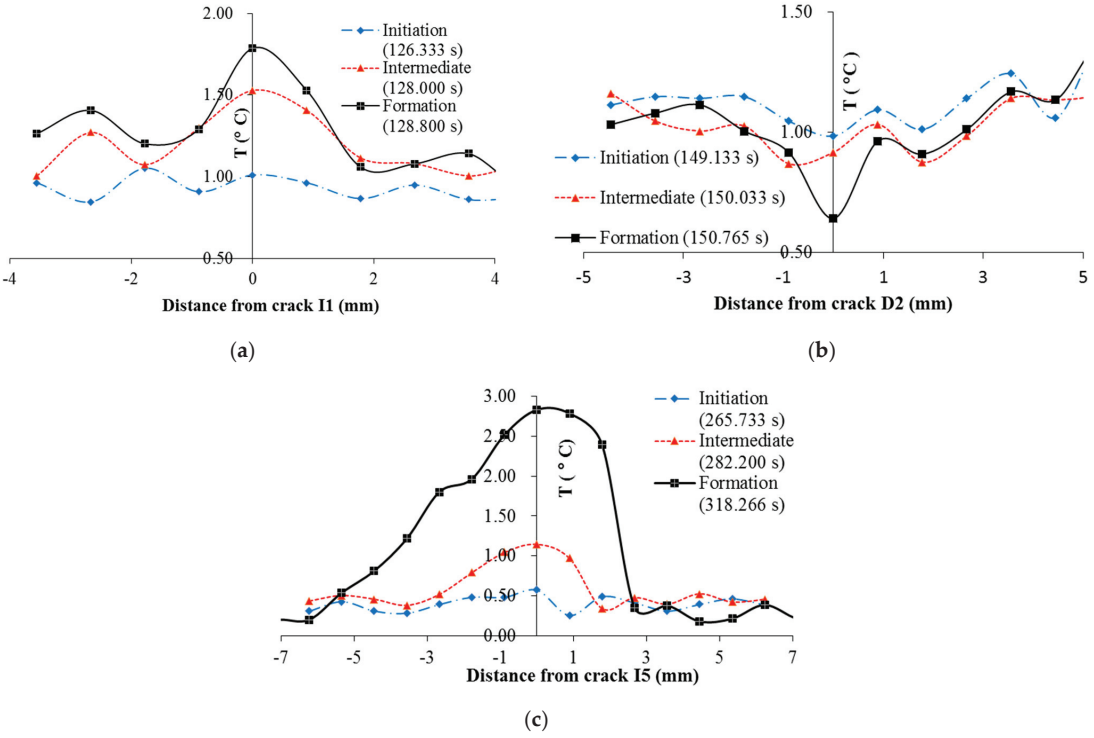


Figure 15. Temperature profile during crack formation: (**a**) crack I1; (**b**) crack D2, and (**c**) crack I5.

In the same manner, the time, t_i, represents when the crack is fully developed and ΔT_i is recorded for cracks I2, I3, I5, D1, D3, D4, and D6. After a crack has formed, the temperature may still vary, subject to the change in stress-strain conditions. Thus, the ΔT for crack I1 at the stress-softening stage (denoted by the initial 's' as subscript in $I1_s$) and plastic stage (denoted by the initial 'p' as subscript in $I1_p$) and ΔT for Crack I2 at the plastic stage were also calculated.

The series of (t_i, ΔT_i) are plotted together with the load-time curve in Figure 16. During the elastic stage, only two cracks, I1 and D1, occur slightly before yield (128.9 s). As the load

decreases during the stress-softening period, the ΔT at crack I1 also decreases. A series of cracks consisting of 4 types D and 1 type I occur immediately before the transition to plastic deformation at 162.0 s. Afterward, the load remains relatively constant as the compression reaches its plastic stage and no crack occurs until t = 265 s when the cracks later form the initiating global failure line. From t = 265 s to t 300 s, the global failure surface gradually grows by connecting the individual cracks and the temperature along with it generally increases. The total increase in temperature at these series of cracks are all above 2.1 °C.

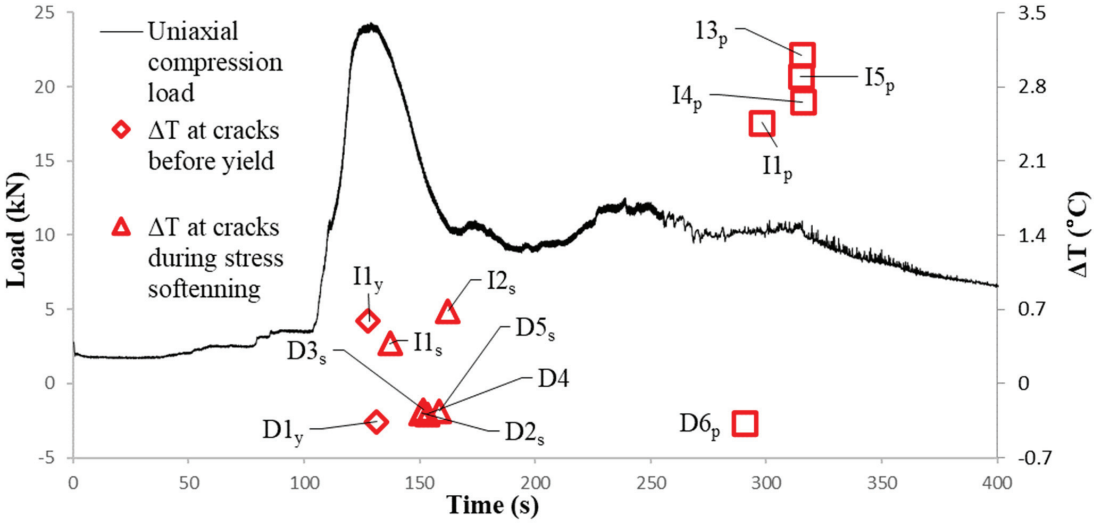


Figure 16. Load–time curve with magnitudes of temperature change at cracks.

Figure 17 shows the temperature profile and normalized temperature profile around I1 at yield, stress softening and failure. ΔT is a result of frictional heat but its magnitude also depends on the thermal properties. It is assumed that the thermal properties and conditions of crack I1 remain constant during the test.

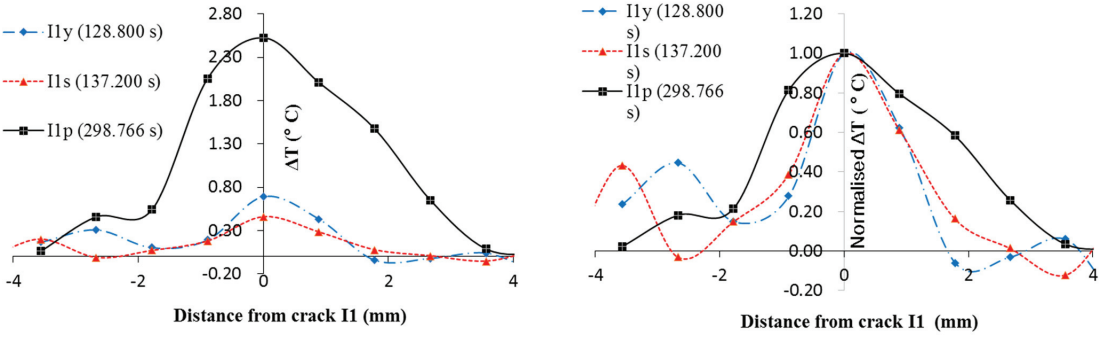


Figure 17. Temperature profile and normalized ΔT at crack I1 at yield, stress softening, and plastic stage.

For the crack I1, values μ and C can be assumed constant during the test. Although the stress at the crack was not measured, it can be inferred that the stress is proportional to the load. Before the plastic stage, the deformation of the entire specimen was relatively uniform and there was little inter-particle movement. There, the ΔT is mostly determined by the

load. The load is the highest at yield and subsequently decreases during the stress-softening stage while the rate of displacement shows no significant change throughout the period. Thus, ΔT_y is higher than ΔT_s. After the plastic stage was reached, although the stress is lower than before, the rate of displacement is significantly higher. The ΔT_p continues increasing and, when the specimen eventually fails, ΔT_p is much higher than ΔT_y and ΔT_s.

With the normalized profile, the propagation of frictional heat can be examined. The dimension affected by the frictional heat expands during the test but is ultimately limited by the dimensions of the gravel particle directly subjected to frictional forces. It seems the temperature of ice within the void of particles remains unaffected by the frictional heat. There are two possible explanations:

1. The ice has a much higher heat capacity than the gravel.
2. The heat transferred to the ice will be consumed by the melting of ice rather than causing an increase in temperature.

3.4. Comparison of Behavior of Plastic Stage

In Figure 18 the spatial temperature profile of cracks from all 3 cases are plotted together. The maximum value of ΔT is observed in Figure 17 also plotted together with the load curve in Figure 19a. The strain is estimated from the deformation of specimen in the videos. However, the magnitude of ΔT cannot provide an undistorted view of the magnitude of friction as the ΔT also depends on thermal parameters such as the specific heat capacity besides the frictional heat. Thus, the frictional heat is back-calculated from ΔT.

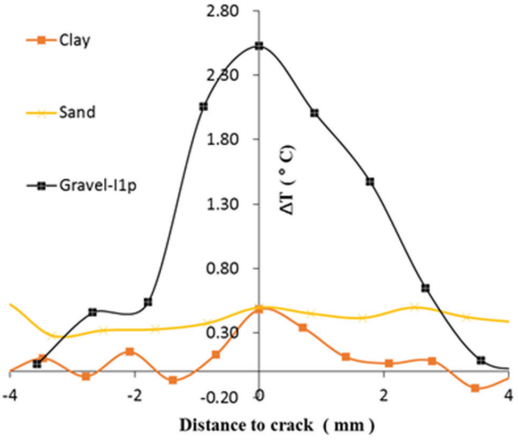


Figure 18. Comparison of temperature profiles across cracks during plastic deformation.

A body of unit amount of mass is considered for each specimen and the body is assumed as an isolated system. Regarding the clay and gravel specimens, the body is located in the pixel referred to as the crack point. For the sand specimen, the body is located in an arbitrary position within rectangle S. Thus, the amount of frictional heat that occurred in those bodies of unit weight can be calculated as:

$$\Delta H = C_m \times \Delta T \tag{3}$$

where C_m is the specific heat capacity of unit mass, ΔT is the temporal temperature change resulting from crack or bulge. The value of C_m is obtained from the literature as shown in Table 2.

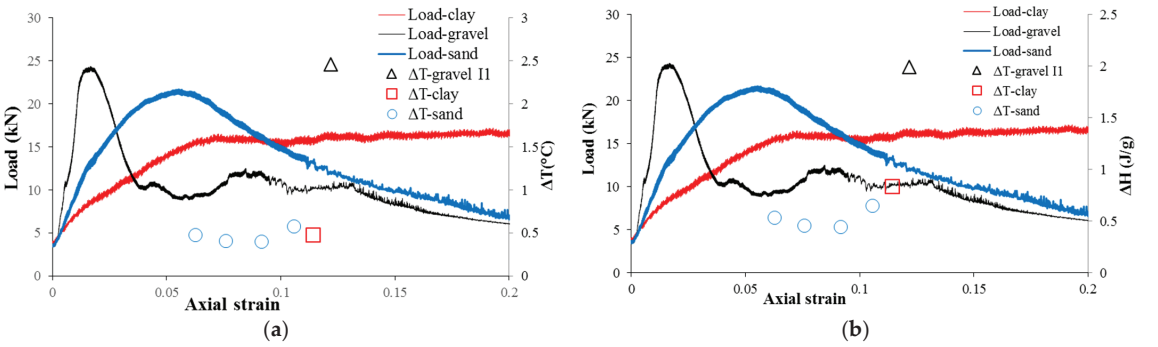


Figure 19. Inter-particle friction in frozen soils: (**a**) comparison of ΔT with stain and load; and (**b**) comparison of ΔH with stain and load.

Table 2. Values of C_m from the literature.

Material	C_m(J · gK^{-1} · °C^{-1})	Source	Description
Clay	1.750	[22]	Saturated and frozen, measured at 0 °C
Sand	0.770	[23]	Dry, measured at 0 °C
Gravel	0.810		

As the spatial resolution was about 0.4mm/pixel in this study, for the calculation of the gravel specimen, pixels of the crack point observed were within the gravel particle, thus the c_m for dry gravel was directly used. For the clay and sand specimens, the c_m for the soil–water mixture should be used. The c_m for sand is very much dependent on the void ratio of the sand specimen. The c_m for the sand–ice mixture is assumed to be the weighted average of specific heat capacity for ice and dry sand and calculated as

$$C_m = \frac{C_s \times m_s + C_i \times m_i}{m_s + m_i} \tag{4}$$

where C_s is the specific heat capacity per unit mass of sand, C_w is the specific heat capacity per unit of ice, m_s and m_w are the mass of sand and ice(water) constituting the specimen, $C_s = 0.770$ J/(gK) is cited in Table 2, $m_s = 2257$ g and $m_w = 370$ g are measured when the specimens are prepared. C_i is estimated according to an empirical equation proposed by Dickinson and Osborne [24] as follows:

$$C_i = 2.114 + 0.007789T \tag{5}$$

The value of T (temperature) should be assigned to the unit of K. To be consistent with the heat capacity for gravel and clay measured at 0 °C, T = 273.15 K is used in Equation (4), C_i =2.116 J/(gK). With all the values on the right-hand side in Equation (3) obtained, C_m can be calculated. The values of C_m adopted for all three cases are listed in Table 3 and these values were used in Equation (4) to calculate the ΔH for each case.

Table 3. Values C_m of adopted in this research.

Material	C_m (J/(gK))
Frozen Clay	1.750
Frozen Sand	1.116
Frozen Gravel	0.810

Figure 19b shows the ΔH in unit mass resulting from inter-particle friction and load–strain curve for all 3 cases altogether. As shown in Table 3, the equivalent unit weight specific heat capacity C_m of frozen gravel is the lowest among the three different specimens. It is assumed the frictional heat generated is proportional to the work done by the compression. The work done by compression is approximately proportional to the product of load and axial strain in Figure 19. Before the maximum load points of the specimens, the gravel specimen experienced the most frictional heat ΔH. Therefore, according to Equation (4), $\Delta T = \Delta H / C_m$, the temperature change of gravel is higher than that of sand and clay.

For the clay specimen, during the plastic deform, there was significant radial deformation at the top and bottom of the specimen, which occurred immediately contact of the platen of the compression machine. Moreover, the rest of specimen deformed similar to an extremely viscous fluid. The friction results from the differential displacement rates between neighboring clay particles. For the sand specimen, as the specimen bulges, the sand particle layers must rearrange. Particles are squeezed into neighboring layers in an axial direction and hence friction occurs. As the friction is distributed among the entire specimen, thus ΔH in unit mass is the least among the 3 cases during plastic deformation. For the gravel specimen, the ΔH is more than 2 times that of the other two cases, because the friction concentrates on the global failure surface where two parts of the specimen slide against each other. The same set of curves except for that of the sand case are also replotted in Figure 20 with normalized ΔT. This shows that the affected distance of the crack in the clay case is less than that of the gravel case. This might be related to the particle size of the soil. As can be seen in the case of gravel, the surface of the temperature rise propagates through the ice in the voids of the soil particle skeleton. Moreover, the estimated specific heat of gravel was less than half of that of frozen clay. This reveals that it took less amount of heat for the frozen gravel to increase a unit of temperature than for the frozen clay. Thus, the temperature increase due to frictional heat occurred in a smaller spatial dimension in frozen clay than in frozen gravel.

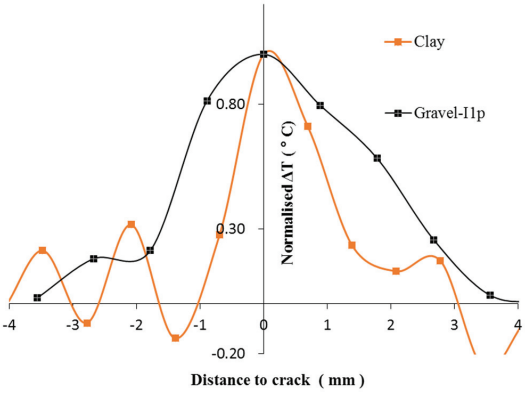


Figure 20. Comparison of ΔT normalised ΔT between clay and gravel cases.

4. Discussion

The identification of the change in thermographic profile was labor intensive and difficult to manage in real time. Thus, the application of this method for monitoring purpose on-site requires significant improvements in efficiency in the future. For example, deep-learning image pattern recognition methods could be applied to automate the process of identification of cracking patterns in thermography.

In this research, tests were conducted at summer room temperature conditions (between 25 °C and 32 °C) and thermodynamic processes occurred, such as heat transmission between specimens and platens of compression machine occurred. It was assumed that temperature changes were not affected by thermodynamic processes. Emissivity of speci-

mens were assumed constant and uniform among all samples. Both assumptions may not agree with real-life conditions upon application of this method in the field.

5. Conclusions

In this research, the thermographic profile of frozen soil specimens under uniaxial compression test was studied. For all 3 cases, abnormal temperature variations on specimen surfaces due to inter-particle friction occurred were observed.

a. In frozen clay specimens, the temperature increases at cracks were identified only after plastic deformation occurred.
b. For frozen sand, simultaneous temperature increases were observed along the entire specimen as it bulged at the plastic stage.
c. In frozen gravel, temperature changes were observed for before cracks appeared in addition to a yield. For cracks in ice particles, there were temperature decreases due to changes in geometry. For cracks in gravel particles, there were temperature increases due to inter-particle friction.
d. Compared in terms of ΔH in unit mass it was shown that the friction in gravel was the strongest and that in sand was the least.
e. The propagation of temperature increases from cracks were also examined in frozen clay and gravel cases. Subsequently, it was shown that the temperature rise propagates further in gravel than in sand.

Author Contributions: Conceptualization, H.S.; methodology, H.S. and Y.Z.; software, Y.Z.; validation, H.S. and Y.Z.; formal analysis, Y.Z.; investigation, Y.Z. and C.C.; resources, H.S.; data curation, Y.Z. and C.C.; writing—original draft preparation, Y.Z.; writing—review and editing, H.S. and Y.H.; visualization, Y.Z.; supervision, H.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pilžis, K.; Vaišis, V. Detection of Oil Product on the Water Surface with Thermal Infrared Camera. *Naft. Aptikimas Vandens Paviršiuje Naudojant Infraraudonųjų Spindulių Kamerą* **2017**, *9*, 357–362. [CrossRef]
2. Fukuhara, T.; Kouyama, T.; Kato, S.; Nakamura, R.; Takahashi, Y.; Akiyama, H. Detection of Small Wildfire by Thermal Infrared Camera With the Uncooled Microbolometer Array for 50-kg Class Satellite. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 4314–4324. [CrossRef]
3. Hagen, N. Passive imaging of wind surface flow using an infrared camera. *Infrared Phys. Technol.* **2017**, *87*, 47–54. [CrossRef]
4. Freund, F.T.; Takeuchi, A.; Lau, B.W.S.; Al-Manaseer, A.; Fu, C.C.; Bryant, N.A.; Ouzounov, D. Stimulated infrared emission from rocks: Assessing a stress indicator. *eEarth (eE) Discuss. (eED)* **2007**, *2*, 7–16. [CrossRef]
5. Huang, J.; Liu, S.; Li, T.; Song, L.; Ni, Q. Experimental study on thermal infrared radiation variation of loaded sandstone in the cold sky background. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016; pp. 4222–4225.
6. Liu, X.; Liang, Z.; Zhang, Y.; Liang, P.; Tian, B. Experimental study on the monitoring of rockburst in tunnels under dry and saturated conditions using AE and infrared monitoring. *Tunn. Undergr. Space Technol.* **2018**, *82*, 517–528. [CrossRef]
7. Seo, H.; Choi, H.; Park, J.; Lee, I.-M. Crack Detection in Pillars Using Infrared Thermographic Imaging. *Geotech. Test. J.* **2017**, *40*, 20150245. [CrossRef]
8. Seo, H. Infrared thermography for detecting cracks in pillar models with different reinforcing systems. *Tunn. Undergr. Space Technol.* **2021**, *116*, 104118. [CrossRef]
9. Gómez Muñoz, C.Q.; García Márquez, F.P.; Sánchez Tomás, J.M. Ice detection using thermal infrared radiometry on wind turbine blades. *Measurement* **2016**, *93*, 157–163. [CrossRef]

10. Lin, Z.Z.G.J.Y.; Chengrui, Z. Active infrared icing detection using neural networks. *J. Huazhong Univ. Sci. Technol. (Nat. Sci. Ed.)* **2010**.
11. Broberg, P. Surface crack detection in welds using thermography. *NDT E Int.* **2013**, *57*, 69–73. [CrossRef]
12. Štarman, S.; Matz, V. Automated System for Crack Detection Using Infrared Thermographic Testing. In Proceedings of the 4th International CANDU In-service Inspection Workshop and NDT in Canada 2012 Conference, Toronto, ON, USA, 18–21 June 2012.
13. Afshani, A.; Kawakami, K.; Konishi, S.; Akagi, H. Study of infrared thermal application for detecting defects within tunnel lining. *Tunn. Undergr. Space Technol.* **2019**, *86*, 186–197. [CrossRef]
14. Chen, C.; Seo, H.; Jun, C.H.; Zhao, Y. Pavement crack detection and classification based on fusion feature of LBP and PCA with SVM. *Int. J. Pavement Eng.* **2021**, 1–10. [CrossRef]
15. Chen, C.; Seo, H.; Zhao, Y. A novel pavement transverse cracks detection model using WT-CNN and STFT-CNN for smartphone data analysis. *Int. J. Pavement Eng.* **2021**, 1–13. [CrossRef]
16. Zhao, Y.; Seo, H.; Chen, C. Displacement mapping of point clouds: Application of retaining structures composed of sheet piles. *J. Civ. Struct. Health Monit.* **2021**, *11*, 915–930. [CrossRef]
17. Budzier, H.; Gerlach, G. *Thermal infrared Sensors: Theory, Optimization, and Practice*; Wiley: Chichester, UK; Hoboken, NJ, USA, 2010; Volume 2010.
18. Flir System, I. *Flir R&D Software 3.3*; Flir System, Inc.: Wilsonville, OR, USA, 2020.
19. Bragg, R.A.; Andersland, O.B. Strain rate, temperature, and sample size effects on compression and tensile properties of frozen sand. *Eng. Geol.* **1981**, *18*, 35–46. [CrossRef]
20. Baker, T.; Jones, S.; Parameswaran, V. *Confined and Unconfined Compression Tests on Frozen Sand*; Laboratory Testing of Frozen Soil, 2010; pp. 387–393.
21. Liu, X.-z.; Liu, P. Experimental research on the compressive fracture toughness of wing fracture of frozen soil. *Cold Reg. Sci. Technol.* **2011**, *65*, 421–428. [CrossRef]
22. Kozlowski, T. Modulated Differential Scanning Calorimetry (MDSC) studies on low-temperature freezing of water adsorbed on clays, apparent specific heat of soil water and specific heat of dry soil. *Cold Reg. Sci. Technol.* **2012**, *78*, 89–96. [CrossRef]
23. Goranson, R.W. Heat capacity of rocks. In *Handbook of Physical Constants*; Birch, F., Ed.; Geological Society of America: Boulder, CO, USA, 1942; p. 36.
24. Dickinson, H.C.; Osborne, N.S. *Specific Heat of Fusion of Ice*; US Government Printing Office: Boston, MA, USA, 1915.

sensors

MDPI

Article

Verification of Tensile Force Estimation Method for Temporary Steel Rods of FCM Bridges Based on Area of Magnetic Hysteresis Curve Using Embedded Elasto-Magnetic Sensor

Won-Kyu Kim [1,2], Junkyeong Kim [3,*], Jooyoung Park [4], Ju-Won Kim [5] and Seunghee Park [6,7,*]

[1] Equipment Group, Industrial Materials Team 2, Materials Division, Samsung C&T Corporation, Seoul 05510, Korea; kwk1104@gmail.com
[2] Department of Convergence Engineering for Future City, Sungkyunkwan University, Suwon 16419, Korea
[3] Safety Inspection for Infrastructure Laboratory (SIIL), Advanced Institute of Convergence Technology, Suwon 16229, Korea
[4] Department of the Civil, Architectural and Environmental System Engineering, Sungkyunkwan University, Suwon 16419, Korea; mitjy26@gmail.com
[5] Department of Safety Engineering, Dongguk University-Gyeongju, Gyeongju 38066, Korea; juwon@dongguk.ac.kr
[6] School of Civil, Architectural Engineering & Landscape Architecture, Sungkyunkwan University, Suwon 16419, Korea
[7] Technical Research Center, Smart Inside AI Co., Ltd., Suwon 16419, Korea
* Correspondence: junkyeong@snu.ac.kr (J.K.); shparkpc@skku.edu (S.P.)

Abstract: The free cantilever method (FCM) is a bridge construction method in which the left and right segments are joined in sequence from a pier without using a bottom strut. To support the imbalance of the left and right moments during construction, temporary steel rods, upon which tensile force is applied that cannot be managed after construction, are embedded in the pier. If there is an excessive loss of tensile force applied to the steel rods, the segments can collapse owing to the unbalanced moment, which may cause personal and property damage. Therefore, it is essential to monitor the tensile force in the temporary steel rods to prevent such accidents. In this study, a tensile force estimation method for the temporary steel rods of an FCM bridge using embedded Elasto-Magnetic (EM) sensors was proposed. After the tensile force was applied to the steel rods, the change in tensile force was monitored according to the changing area of a magnetic hysteresis curve, as measured by the embedded EM sensors. To verify the field applicability of the proposed method, the EM sensors were installed in an FCM bridge pier under construction. The three sensors were installed in conjunction with a sheath tube, and the magnetic hysteresis curve was measured over nine months. Temperature data from the measurement period were used to compensate for the error due to daily temperature fluctuations. The estimated tensile force was consistent with an error range of ±4% when compared with the reference value measured by the load cell. Based on the results of this experiment, the applicability of the proposed method was demonstrated.

Keywords: free cantilever method (FCM); temporary steel rod; tensile force; embedded Elasto-Magnetic (EM) sensor; magnetic hysteresis curve; temperature compensation

Citation: Kim, W.-K.; Kim, J.; Park, J.; Kim, J.-W.; Park, S. Verification of Tensile Force Estimation Method for Temporary Steel Rods of FCM Bridges Based on Area of Magnetic Hysteresis Curve Using Embedded Elasto-Magnetic Sensor. *Sensors* **2022**, *22*, 1005. https://doi.org/10.3390/s22031005

Academic Editor: Mohammad Noori

Received: 19 December 2021
Accepted: 26 January 2022
Published: 27 January 2022

1. Introduction

As the construction industry develops, the importance of effective and efficient maintenance techniques for the structures is increasingly emphasized. In the case of bridges, the need for more advanced maintenance technology is gradually increasing as construction methods progress [1,2]. Prestressed Concrete (PSC) has been used in numerous bridges since the late 1960s [3] when the post-tension method was first used in bridge construction. PSC is characterized by a higher tensile strength than plain concrete because the initial

stress is pre-applied to the concrete. In addition, PSC offers the advantage that the construction can be performed at a low cost, which is increasingly demanded in the bridge construction industry [4].

The free cantilever method (FCM) is a construction method for PSC bridges that does not install scaffolding systems under the bridge and completes the superstructure of the bridge by sequentially joining the segments to form a span by post-tensioning and balancing them left and right from each pier using special erection equipment. The FCM has advantages in terms of construction conditions and its period. As it does not use a shore or a scaffold, the bridge can be constructed under extreme conditions, such as in deep valleys or at sea. Furthermore, the construction period can be shortened in the case of using the precast concrete method, where manufactured girders are assembled at the construction site, rather than the in situ concrete method. Thus, the FCM is consistently used for bridge construction. However, considering the characteristics of the FCM, which joins segments from side to side, an unbalanced moment may occur owing to the load generated during construction, as shown in Figure 1.

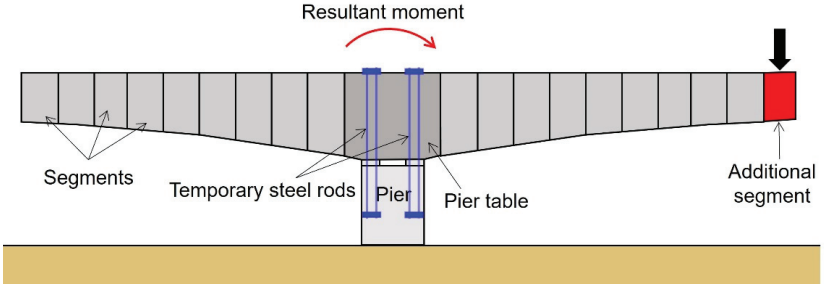


Figure 1. Illustration of unbalanced moment occurrence during construction of FCM bridges.

To prevent such a problem, temporary steel rods are embedded in the pier table. Then, calculated tensile forces are introduced to withstand the loads and their resulting moments during the construction stage. However, because the temporary steel rods are embedded in the pier table, it is almost impossible to visually observe and determine whether the tensile forces introduced to the rods are being maintained properly. If the signs of destruction or loss of tension are not recognized at the early stages, and consequential actions are not taken in a timely manner, accidents, such as sudden collapses and falls, may occur during construction. Therefore, a Non-Destructive Test (NDT) technique, which can be applied to the temporary steel rods during construction, is necessary to reduce the potential accidents.

Recently, a great deal of research has been conducted on measuring steel tension, such as the tensile force estimation method of a PS tendon using a Fiber Bragg Grating (FBG) sensor [5,6] or an Elasto-Magnetic (EM) sensor [7–9] and a tensile force measurement method of steel wire using a natural frequency measurement [10]. In addition, tensile force measurements of unbonded steel wires using longitudinal guided ultrasonic measurement techniques [11,12] and magnetic flux transmission monitoring techniques using magnetic circuits [13] have been studied. A method that uses an acoustoelastic theory was proposed to evaluate the prestress levels in post-tensioned steel strands employing changes in longitudinal stress wave velocity [14,15]. The aforementioned studies showed the estimated tensile force from the sensor response according to the introduced tensile force. However, the changes in the measured value due to environmental changes, such as temperature, were not taken into consideration with a sufficiently long observation period. A study that considered the temperature condition of the target structure was also proposed using various sensors. An SHM system for a long-span, cable-membrane structure was also proposed by Tang et al. [16]. It was applied to monitor the structural static responses, structural vibration, and environmental effects of the structure using various types of sensors, including a magnetic flux sensor, an FBG strain gauge, an FGB thermometer,

an accelerometer, and so on. A study to estimate the girder deflection under thermal actions was conducted for a cable-stayed bridge [17]. Although those studies showed the correlation between the temperature distribution on the members of the bridge and the deflection of the girder by thermal actions, this study focuses on the change in the signal of the sensor, which measures the tension force as the change in temperature of the target member leads to the change in the obtained signals. Therefore, in this study, a technique was proposed for estimating the tensile force during the construction of an FCM bridge using embedded EM sensors to monitor the magnetic hysteresis of the temporary steel rods, considering the temperature effect on the sensors continually. The embedded EM sensor consists of two coils and a bobbin. The primary coil generates a magnetic field after magnetizing the steel rods, and the secondary coil serves to measure the magnetic flux generated in the magnetized steel rods. The magnetic hysteresis curve generated from the measured magnetic flux varies depending on the change in the magnetic field according to the change in the tension of the steel rods [18,19]. It leads to the change of the area of the hysteresis curve. The tensile force was estimated by calculating the changing area. However, the sensor coil, being made of copper, is very sensitive to the change of temperature [20] and, thus, affects the sensor response. The signals from the sensors may not be interpreted accurately without reflecting on this phenomenon. Therefore, in this study, a temperature compensation method was applied for the correction of the errors caused by the temperature change.

2. Theoretical Background and Methods

2.1. EM Sensor

The EM sensor used in this paper is composed of a part for inducing voltage to generate magnetic fields and a part for measuring it, as shown in Figure 2. The coils are wound on the outer surface of the bobbin, and insulation covers are wrapped between each coil. The coils pulled out of the bobbin are connected to the data-acquisition device through connectors. A protection cover is on the outermost surface of the EM sensor, and it protects the sensors from the concrete pouring. For the attachment of the EM sensor to the sheath around the temporary steel rod, conchoids are on the inner surface of both ends of the bobbin.

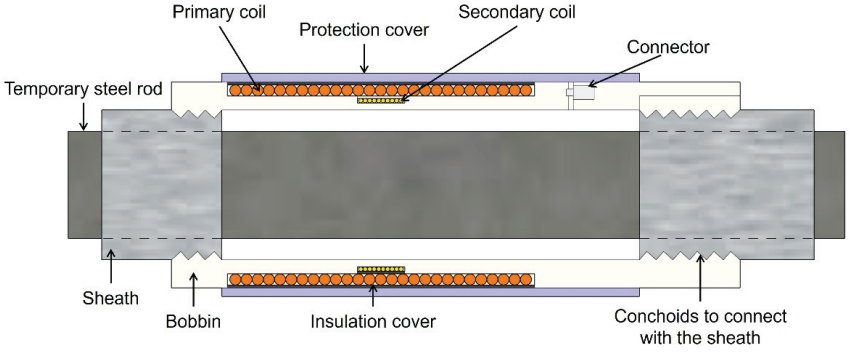


Figure 2. Schematic diagram of embedded EM sensors.

2.2. Prestress Loss in PSC Bridge

Prestress is introduced to the PSC bridge at each major construction stage of the bridge superstructure. The loss of prestress introduced to the structure has various causes, and the type of the loss is usually classified into immediate loss and long-term loss [21]. Immediate loss occurs when the prestress is introduced, and it includes the loss due to friction between the PS steel and the sheath pipe, the anchorage slip, and the elastic deformation of the concrete. Long-term loss refers to the loss that occurs over time after the introduction of

the prestress, and it is related to the creep and drying shrinkage of the concrete and the relaxation of the PS steel.

2.3. Tensile Force Estimation through Measuring Area of Magnetic Hysteresis Curve

In this study, the magnetic hysteresis curves were measured using embedded EM sensors to monitor the change in the tensile force of the temporary steel rods during the construction of an FCM bridge. Three embedded EM sensors were installed with a sheath pipe outside the steel rod used for the construction. As shown in Figure 3, when a voltage is applied to the primary coil of the installed EM sensor, the magnetic field of the steel rod becomes saturated. On the other hand, when a reverse voltage is applied, the direction of magnetic field in the steel rod is reversed, creating a magnetic hysteresis curve (B–H Loop) [22]. However, the shape of the hysteresis curve begins to change when some amount of force is applied to induce stresses inside the steel rod [23]. The letter A in Figure 3 shows a state of magnetization generated in the absence of stress, while the letter B indicates a magnetization state after stress is applied [24].

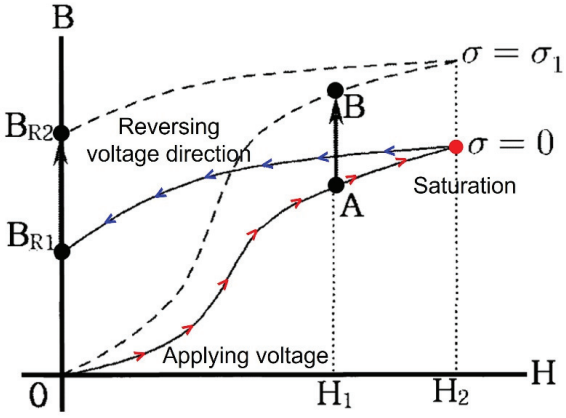


Figure 3. Change of magnetic hysteresis curve according to the effect of tensile force.

The magnetic hysteresis curve measured by this principle represents the relationship between the strength of the magnetic field and the magnetic flux density of a ferromagnetic material, which is used to indicate the magnetic property of the ferromagnetic material. In addition, when the tensile force introduced to the steel rod changes, the magnetic properties change by the inverse magnetostriction effect, leading to the consequential change of the magnetic hysteresis curve. When the tensile force increases, the magnetic flux density increases, leading to an increase in the magnetic flux leakage. As the magnetic flux leakage increases, the area of the magnetic hysteresis curve increases. Therefore, it is possible to estimate the state of the introduced tensile force by measuring the area increase and decrease of the magnetic hysteresis curve for the temporary steel rods.

2.4. Temperature Compensation Method

A temperature compensation technique was adopted to correct the measurement errors due to the temperature changes during the measurement. Temperature compensation is essential because steel materials, including the temporary steel rods used in this study, are greatly affected by external forces and temperature changes [25–27]. Such compensation techniques are broadly divided into software-based methods and hardware design methods. However, the methods through hardware compensation have a limitation in their use in the field because they suffer from poor reliability due to their inability to cope with design errors that may occur during production [28]. Therefore, in this study, a software compensation technique using a quadratic polynomial was applied. The polynomial compensation technique is widely used as a temperature compensation technique for measurement data.

The quadratic polynomial employed in this study is represented by the following regression equation [27]:

$$F_p = a_{00} + a_{10}V_T + a_{01}T + a_{20}V_T^2 + a_{11}V_T \cdot T + a_{02}T^2, \quad (1)$$

where F_p is the tensile force predicted by the quadratic polynomial fitting; V_T is the applied voltage; T is the on-site temperature; and a_{00}, a_{10}, a_{01}, a_{20}, a_{11}, and a_{02} are the second-order fitting coefficients. The change in the area of the graph of the measured magnetic hysteresis curve, using embedded EM sensors, was approximated to the second order through regression analysis. The temperature dependence was compensated for by substituting the temperature data measured in the field into the equation with their respective measurement times.

3. Field Experimental Results and Discussion

3.1. Experimental Setup

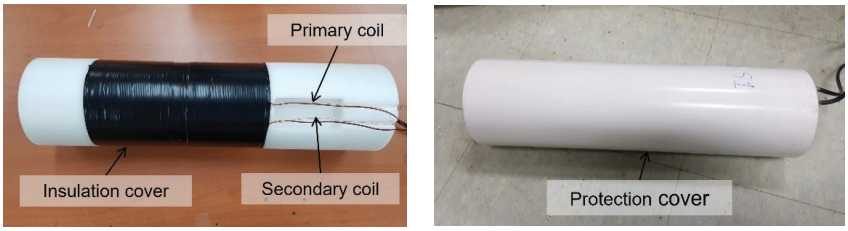
To apply the embedded EM sensors to the site, a tensile force monitoring experiment was conducted by measuring the magnetic hysteresis curve at the construction site of a PSC bridge where the FCM was applied. The test bridge was a box-type PSC girder bridge with a span of 640 m and a width of 24.51 m at a construction site in Asan city, Chungcheongnam-do, Republic of Korea. The steel rods applied to the site were circular rods with a diameter of 47 mm, an ultimate strength (F_u) of 1820 kN, and a yield strength (F_y) of 1650 kN. A load cell was installed with the application of three embedded EM sensors to measure the accurate prestressed force. The load cell used in the experiment is a VW type (SJ-3000), made by Sungjin Geotec in South Korea. Its specifications are shown in Table 1. The VW type load cell uses a principle by which the vibration wire generates the resonant frequency, and the frequency is transmitted to the output device to display the necessary engineering unit when it is magnetized by the magnetic coil mounted due to the load. The embedded EM sensor and its specifications are shown in Figure 4 and Table 2, respectively. The bobbin was made of a primary coil part with a diameter of 117 mm and a secondary coil part with a diameter of 107 mm. The primary and the secondary coils were wound 300 times and 120 times, respectively.

Table 1. Specifications of load cell (SJ-3000).

Classification		Values and Description
Capacity		1177 kN
Ultimate overload		150% of Capacity
Resolution		0.025% F.S.
Accuracy		±0.1~±1% F.S.
Linearity error		±0.5% F.S.
Material		SCM alloy steel
Gauge		3 VW Strain gauge (4 Strain gauge)
Thermal expansion coefficient		10.8×10^{-6}/°C
Operating temp. range		−40 °C~80 °C
Temp. sensor	Type	NTC Thermistor (3KD-ATF)
	operating range	−40 °C~80 °C
	Accuracy	Thermistor: ±1 °C
Waterproof		Fluoride O-ring, High-density vacuum grease coating
Weight		4.95 kg

Table 2. Specifications of the embedded EM sensor.

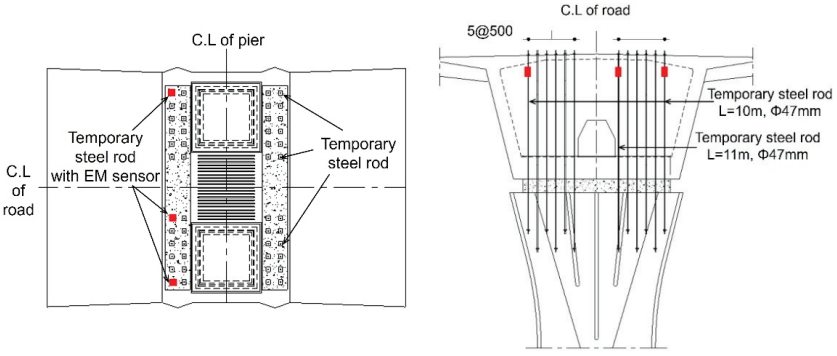
Classification	Primary Coil	Secondary Coil
Diameter of bobbin (mm)	117	107
Diameter of coil (mm)	1.2	0.3
Number of turns	300	120



(**a**) (**b**)

Figure 4. Fabricated EM sensor: (**a**) parts of EM sensor with insulation cover; (**b**) EM sensor equipped with protection cover.

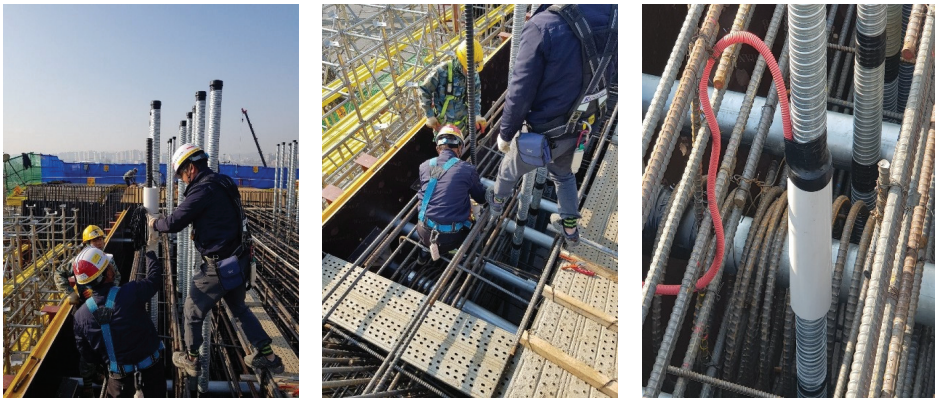
Figure 5 shows the location of the temporary steel rods and the EM sensors. Forty-eight steel rods were installed in the pier at intervals of 500 mm from each other, and they were connected from the pier to the pier table. The EM sensors were installed on three steel rods, two of which were connected to the steel rods with the length of 10 m; the other one was connected to the steel rod with the length of 11 m.



(**a**) (**b**)

Figure 5. Layout of pier and pier table: (**a**) locations of EM sensors installed on pier table. (**b**) cross-section of pier and pier table.

Figure 6 shows the installation process of the EM sensors. It was decided that the locations of the EM sensors would be at three spots on the pier head, where two of them were eccentric sections, and the other was a midsection. The sensors were installed with an external sheath for the temporary steel rods. The cables connected to the sensor were pulled out of the bridge using a cable tube to prevent damage during the concrete pouring. After the concrete was hardened, a tensile force of 900 kN was introduced to all the temporary steel rods. The magnetic hysteresis curve was measured by installing a container for the measurement under the pier. The measurement was continuously conducted in 30 min intervals for approximately 10 months, from 17 March 2019 to 6 January 2020.



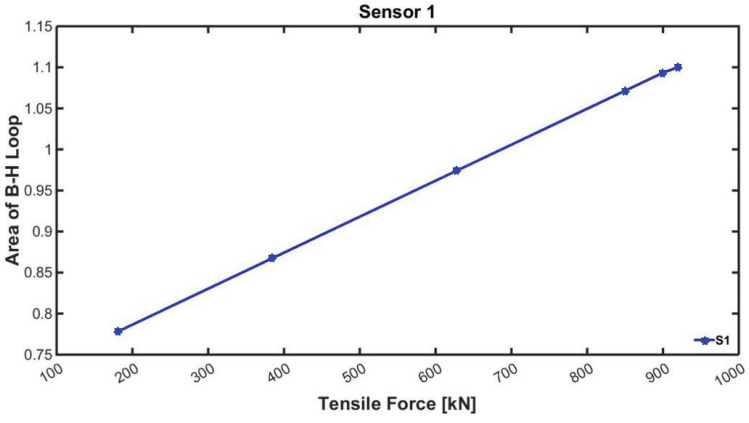
(**a**) (**b**) (**c**)

Figure 6. Installation process of embedded EM sensors: (**a**) inserting the sensor after cutting the sheath; (**b**) sheath tube and sensor combination; (**c**) electric wire protection with cable tube.

A module incorporating a voltage amplifier, a data acquisition device, and a desktop computer equipped with the NI (National Instrument) LabVIEW software package (Version 19.0) was used for the measurement. The measurements were performed five times to reduce the measurement error from each sensor.

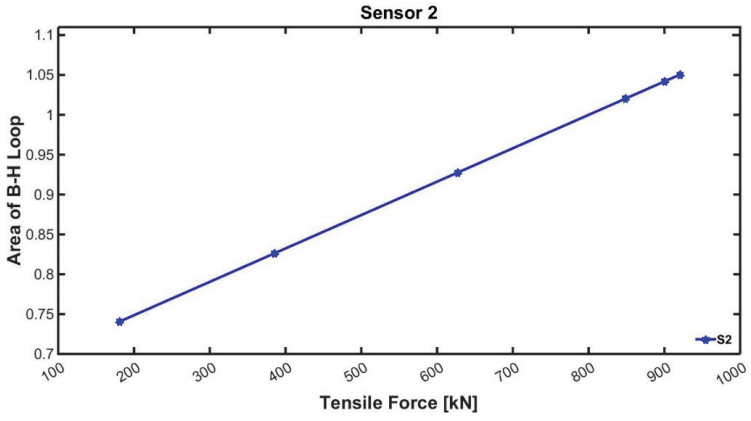
3.2. Initial Value Calibration of Tensile Force

The initial value calibration results using the EM sensors are shown in Figure 7. The tensile force was measured six times by progressively increasing the prestressing force to 180 kN, 383 kN, 628 kN, 849 kN, 900 kN, and 915 kN. The area of the magnetic hysteresis curve was measured simultaneously with the prestressing process.

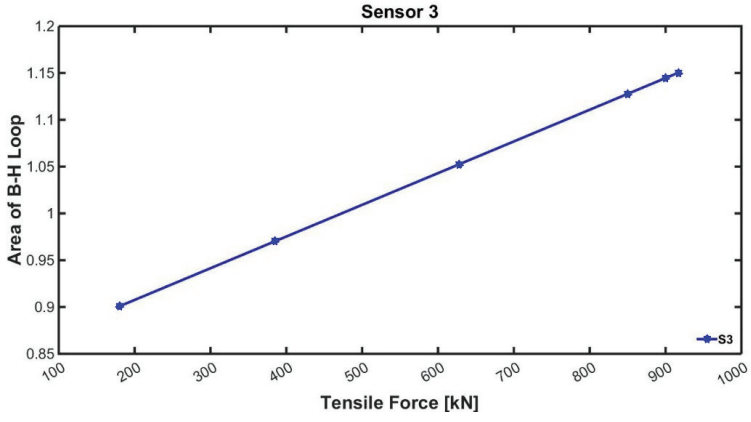


(**a**)

Figure 7. *Cont.*



(**b**)



(**c**)

Figure 7. Calibration results using EM sensors: (**a**) Sensor 1; (**b**) Sensor 2; (**c**) Sensor 3.

The area of the B-H curve corresponding to the value of the tension force introduced to each steel rod could be obtained through the initial value calibration. The area of the curve increases as the tension force gets developed to the designed tensile force. Therefore, it was confirmed that the area of the B-H curve and the introduced tensile force are in positive correlation.

3.3. Measurement Results of the Field Experiment

The results of measuring the tensile force using the load cell are shown in Figure 8. The load cell data were used as the absolute values of the force introduced to the temporary steel rods. In addition, the results of the change in area of the magnetic hysteresis curve measured using the embedded EM sensors and the change in field temperature are shown in Figure 9.

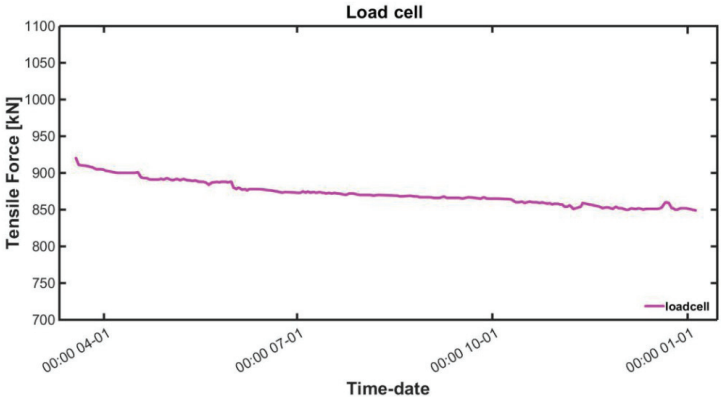


Figure 8. Tensile force measurement result using a load cell.

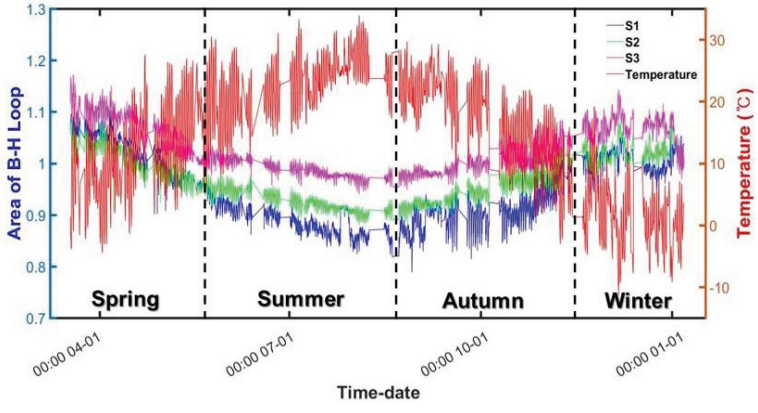
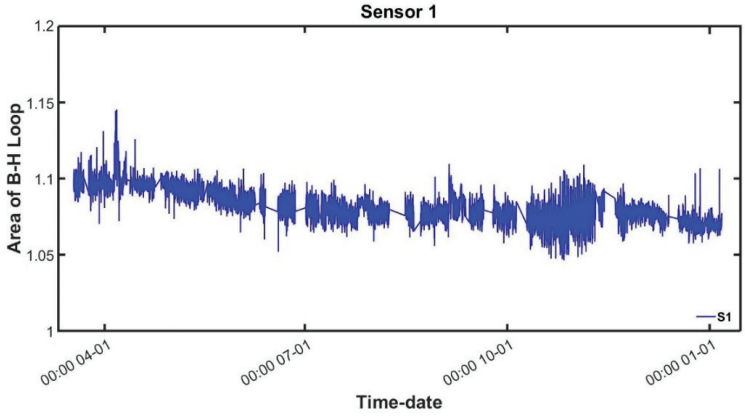


Figure 9. Changes in the area of magnetic hysteresis curves and field temperature during the measurement.

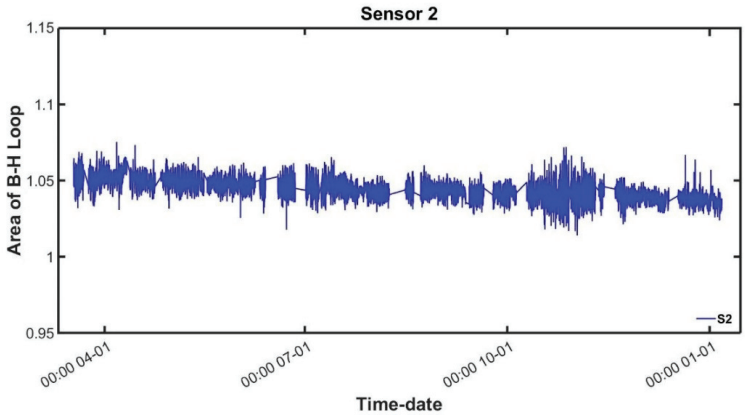
During the measurement, the input voltage remained constant to a 0.02-Hz triangular wave of ±3 V. Although the voltage input value to the sensor remained constant, the area decreased as summer approached and then increased again as winter approached. The results of Figure 9 indicate that the area of the magnetic hysteresis curve measured by the embedded EM sensor exhibits an opposite trend to that of the temperature change.

3.4. Tensile Force Estimation with Temperature-Compensated Data

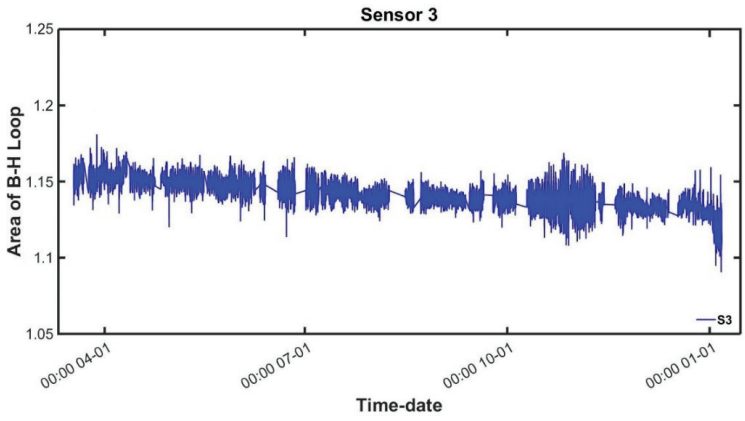
The temperature compensation method was described in the former section, and the technique was employed in the data-compensation process. The area of the measured hysteresis curve, which showed an inverse relation to the temperature data, was corrected through temperature compensation, and the results are shown in Figure 10.



(**a**)



(**b**)



(**c**)

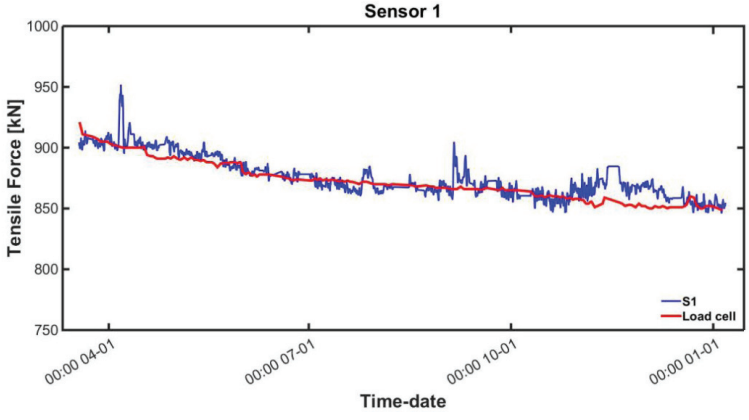
Figure 10. B–H loop area changes of sensors after temperature compensation: (**a**) Sensor 1; (**b**) Sensor 2; (**c**) Sensor 3.

The trend of the magnetic hysteresis curve due to the temperature change was modified from each sensor. It could be identified that the area decreased consistently in accordance with prestress loss during the measurement. This result agreed with the descending tendency of the tension force introduced to the temporary rods.

The tension force was estimated using the area of the magnetic hysteresis curve from each sensor after correcting by temperature compensation. The results of comparing the measured force with the load cell measurements are shown in Table 3. In addition, the sensor measurements and the load cell results are plotted in Figure 11.

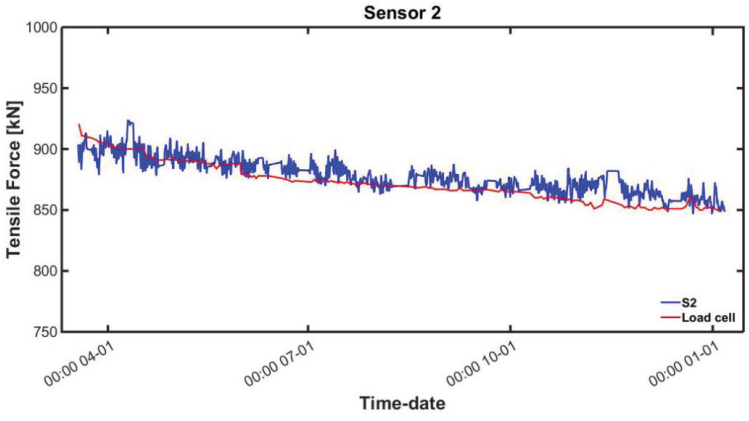
Table 3. Comparison of the estimated tensile force obtained using embedded EM sensors and the tensile force measured using a load cell.

Date		Load Cell (kN)	Sensor 1		Sensor 2		Sensor 3		Temperature (°C)
			Estimated Tension (kN)	Error Rate (%)	Estimated Tension (kN)	Error Rate (%)	Estimated Tension (kN)	Error Rate (%)	
1 April 2019	7:00	891	893.89	0.32	896.06	0.57	890.64	0.04	1.1
12 April 2019	14:00	880	884.04	0.46	884.10	0.47	880.23	0.03	13.0
13 May 2019	10:00	875	874.15	0.10	882.55	0.86	880.84	0.67	23.1
30 May 2019	9:00	872	866.71	0.61	890.79	2.15	881.98	1.14	18.6
14 June 2019	8:00	869	861.24	0.89	871.08	0.24	872.91	0.45	22.8
17 July 2019	10:00	865	860.11	0.57	864.91	0.01	871.28	0.73	25.5
14 August 2019	11:00	861	875.60	1.70	870.31	1.08	864.30	0.38	30.0
6 September 2019	9:00	858	858.15	0.02	863.12	0.60	864.65	0.78	23.2
23 October 2019	15:00	850	877.66	3.25	877.31	3.21	867.11	2.01	18.5
6 December 2019	14:00	852	847.68	0.51	844.30	0.90	828.37	2.77	4.3

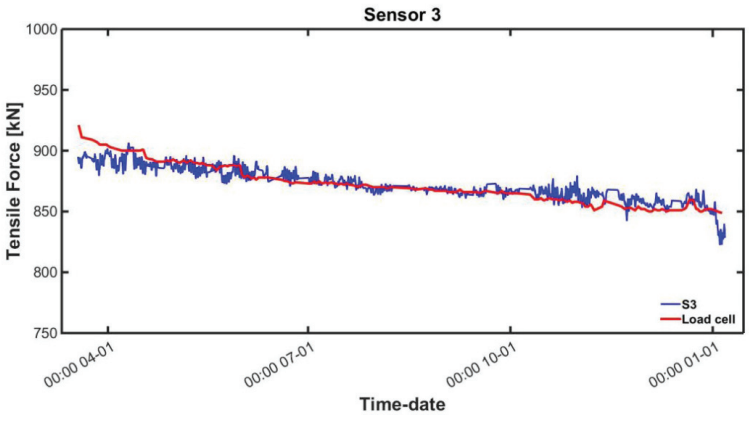


(**a**)

Figure 11. *Cont.*



(**b**)



(**c**)

Figure 11. Estimated tension after temperature compensation and actual tension: (**a**) Sensor 1; (**b**) Sensor 2; (**c**) Sensor 3.

4. Conclusions

The tensile force management of the temporary steel rods plays a key role at the phase of attaching segments during the construction of an FCM bridge to prevent any possible collapse and to meet the construction schedule. Therefore, a tensile force estimation method to measure the area of the magnetic hysteresis curve of the temporary steel installed in an FCM bridge was proposed in this study. For the measurement, the embedded EM sensors were installed in combination with a sheath tube outside the temporary steel rod. The magnetic field strength and the induced magnetic flux density were obtained to form the magnetic hysteresis curve. It was possible to estimate the tension by tracking the changes in the magnetic hysteresis of the steel rods. As a noticeable trend in the curve due to the temperature was observed, a temperature compensation technique using quadratic polynomial fitting was applied to offset the measurement error. Three EM sensors were employed to increase the reliability of the measurement data. To verify the applicability of the study, the experiment using the EM sensors was conducted at an actual FCM bridge construction site. Three EM sensors were installed at different spots on the pier head. An initial value compensation method was exploited as a reference by measuring the tensile

force and the area of the magnetic hysteresis curve at the same time during the prestressing process. In the field experiment, steel rods with a diameter of 47 mm, an ultimate strength (F_u) of 1820 kN, and a yield strength (F_y) of 1650 kN were used; the bridge was a PSC box girder bridge with a total length of 640 m and a width of 24.51 m. The results of the tension estimation were compared to the tension measured using a load cell installed on a steel rod with sensor 1. The estimated results of sensors 1, 2, and 3 showed average error rates of 0.68%, 0.95%, and 0.82%, respectively. The field applicability of the long-term tensile force monitoring method using embedded EM sensors was verified through experiment and analysis. This study's reliability would be further increased by acquiring more data through the long-term monitoring of additional bridge construction sites and by adopting more accurate methods to reduce the error rate. Based on this, it is expected to develop into a method for estimating the tensile force that can be applied to all PSC bridges to be constructed using FCM.

Author Contributions: Conceptualization, W.-K.K. and J.K.; methodology, W.-K.K. and J.K.; software, J.K. and W.-K.K.; validation, W.-K.K., J.P., and J.-W.K.; formal analysis, W.-K.K.; investigation, W.-K.K.; resources, S.P.; data curation, W.-K.K.; writing—original draft preparation, W.-K.K.; writing—review and editing, J.P.; visualization, W.-K.K.; supervision, J.K.; project administration, W.-K.K.; funding acquisition, S.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Korea Ministry of Land, Infrastructure and Transport [MOLIT] as an [Innovative Talent Education Program for Smart City], a grant (20CTAP-C157163-01) from the Infrastructure and Transportation Technology Promotion Research Program funded by the Ministry of Land, Infrastructure and Transport of the Korean government, the National Research Foundation of Korea (NRF) grant funded by the Korea government(MSIT) (NRF-2021R1A4A3033128), and the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2017-R1A6A3A04011933).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lee, C.; Park, S.; Bolander, J.E.; Pyo, S. Monitoring the hardening process of ultra high performance concrete using decomposed modes of guided waves. *Constr. Build. Mater.* **2018**, *163*, 267–276. [CrossRef]
2. Kim, J.-W.; Park, M.; Kim, J.; Park, S. Improvement of MFL sensing–based damage detection and quantification for steel bar NDE. *Smart Struct. Syst.* **2018**, *22*, 239–247.
3. Hewson, N.R. Prestressed concrete in bridgeworks. In *Prestressed Concrete Bridges: Design and Construction*; Thomas Telford: London, UK, 2003; pp. 12–17.
4. Rens, K.L.; Wipf, T.J.; Klaiber, F.W. Review of nondestructive evaluation techniques of civil infrastructure. *J. Perform. Constr. Fac.* **1997**, *11*, 152–160. [CrossRef]
5. Kim, J.M.; Kim, H.W.; Park, Y.H.; Yang, I.H.; Kim, Y.S. FBG sensors encapsulated into 7-wire steel strand for tension monitoring of a prestressing tendon. *Adv. Struct. Eng.* **2012**, *15*, 907–917. [CrossRef]
6. Lan, C.; Zhou, Z.; Ou, J. Monitoring of structural prestress loss in RC beams by inner distributed Brillouin and fiber Bragg grating sensors on a single optical fiber. *Struct. Control Health Monit.* **2014**, *21*, 317–330. [CrossRef]
7. Wang, M.L.; Wang, G.; Zhao, A. Application of EM Stress Sensors in Large Steel Cables. In *Sensing Issues in Civil Structural Health Monitoring*; Springer: New York City, NY, USA, 2005; pp. 145–154.
8. Kim, J.; Kim, J.-W.; Lee, C.; Park, S. Development of Embedded EM Sensors for Estimating Tensile Forces of PSC Girder Bridge. *Sensors* **2017**, *17*, 1989. [CrossRef] [PubMed]
9. Kim, J.; Kim, J.-W.; Park, S. Investigation of Applicability of an Embedded EM Sensor to Measure the Tension of a PSC Girder. *J. Sens.* **2019**, *2019*, 2469647. [CrossRef]
10. Saiidi, M.; Douglas, B.; Feng, S. Prestress Force Effect on Vibration Frequency of Concrete Bridges. *J. Struct. Eng.* **1994**, *120*, 2233–2241. [CrossRef]
11. Chen, H.L.; Wissawapaisal, K. Measurement of Tensile Forces in a Seven-wire Prestressing Strands using Stress Waves. *J. Eng. Mech.* **2001**, *127*, 599–606. [CrossRef]
12. Salamone, S.; Bartoli, I.; Nucera, C. Health Monitoring of Prestressing Tendons in Post-tensioned Concrete Structures. In Proceedings of the SPIE Smart Structures and Materials + Nondestructive Evaluation and Health Monitoring, San Diego, CA, USA, 14 April 2011.

13. Joh, C.; Lee, J.W.; Kwahk, I. Feasibility study of stress measurement in prestressing tendons using Villari effect and induced magnetic field. *Int. J. Distrib. Sens. Netw.* **2013**, *9*, 249829. [CrossRef]
14. Chen, H.L.; Wissawapaisal, K. Application of Wigner-Ville Transform to Evaluate Tensile Forces in Seven-Wire Prestressing Strands. *J. Eng. Mech.* **2002**, *128*, 1206–1214. [CrossRef]
15. Chaki, S.; Bourse, G. Stress Level Measurement in Prestressed Steel Strands Using Acoustoelastic Effect. *Exp. Mech.* **2009**, *49*, 673–681. [CrossRef]
16. Tang, T.; Yang, D.H.; Wang, L.; Zhang, J.R.; Yi, T.H. Design and application of structural health monitoring system in long-span cable-membrane structure. *Earthq. Eng. Eng. Vib.* **2019**, *18*, 461–474. [CrossRef]
17. Yang, D.H.; Yi, T.H.; Li, H.N.; Zhang, Y.F. Correlation-Based Estimation Method for Cable-Stayed Bridge Girder Deflection Variability under Thermal Action. *J. Perform. Constr. Facil.* **2018**, *32*, 04018070. [CrossRef]
18. Zhukov, A.; Vazquez, M.; Garcia-Beneytez, J.M. Magnetoelastic sensor for signature identification based on mechanomagnetic effect in amorphous wires. *J. Phys. IV* **1998**, *8*, Pr2-763–Pr2-766. [CrossRef]
19. Zhukov, A.; Zhukova, V.; Larin, V.; Gonzalez, J. Tailoring of magnetic anisotropy of Fe-rich microwires by stress induced anisotropy. *Phys. B Condens. Matter* **2006**, *384*, 1–4. [CrossRef]
20. Dellinger, J.H. *The Temperature Coefficient of Resistance of Copper*; U.S. Government Printing Office: Washington, DC, USA, 1911; pp. 74–76.
21. Hewson, N.R. Prestress design. In *Prestressed Concrete Bridges: Design and Construction*; Thomas Telford: London, UK, 2003; pp. 67–72.
22. Jiles, D. *Introduction to Magnetism and Magnetic Materials*, 3rd ed.; CRC Press: Boca Raton, FL, USA, 2015; pp. 109–130.
23. Szpunar, B.; Szpunar, J. Influence of stresses on the hysteresis curve in constructional steel. *IEEE Trans. Magn.* **1984**, *20*, 1882–1884. [CrossRef]
24. Kim, J. Tensile Force Measurement Technique for Prestressed Tendon of PSC Structures using Embedded EM Sensor and Machine Learning Method. Ph.D. Thesis, Sungkyunkwan University, Suwon, Korea, 2018.
25. Lloyd, G.M.; Singh, V.; Wang, M.L. Experimental evaluation of differential thermal errors in magnetoelastic stress sensors for Re < 180. In Proceedings of the Sensors 2002 IEEE, Orlando, FL, USA, 12–14 June 2002.
26. Lloyd, G.M.; Singh, V.; Wang, M.L.; Hovorka, O. Temperature Compensation and Scalability of Hysteretic/Anhysteretic Magnetic-Property Sensors. *IEEE Sens. J.* **2003**, *3*, 708–716. [CrossRef]
27. Zhang, R.; Duan, Y.; Zhao, Y.; He, X. Temperature Compensation of Elasto-Magneto-Electric (EME) Sensors in Cable Force Monitoring Using BP Neural Network. *Sensors* **2018**, *18*, 2176. [CrossRef]
28. Duan, Y.; Zhang, R.; Zhao, Y.; Or, S.; Fan, K.; Tang, Z. Smart elasto-magneto-electric (EME) sensors for stress monitoring of steel structures in railway infrastructures. *J. Zhejiang. Univ. Sci. A* **2011**, *12*, 895–901. [CrossRef]

applied sciences

MDPI

Article

A Filtering Method for Suppressing the Lift-Off Interference in Magnetic Flux Leakage Detection of Rail Head Surface Defect

Yinliang Jia [1], Yichen Lu [1,*], Longhui Xiong [2], Yuhua Zhang [2], Ping Wang [1] and Huangjian Zhou [1]

[1] College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China; jyl@nuaa.edu.cn (Y.J.); zeitping@nuaa.edu.cn (P.W.); Samuel@nuaa.edu.cn (H.Z.)
[2] Infrastructure Inspection Research Institute, China Academy of Railway Sciences Corporation Limited, Beijing 100081, China; xionglonghui@rails.cn (L.X.); zhangyuhua@rails.cn (Y.Z.)
* Correspondence: yichenlu@nuaa.edu.cn

Abstract: Magnetic flux leakage (MFL) detection is a common nondestructive detection method which is usually used to detect the surface defects of steel pipes and rails. To suppress the interference of lift-off on the detection signal of the defects in rail head surfaces, a filtering method is proposed according to the distribution characteristics of the defect leakage magnetic field (LMF) in different directions. The sensor array is used to confirm the reference signal according to the difference between the signals in x and z directions. The installation mode of the sensors is deduced according to the distribution of the defect LMF. The experimental results show that this method can effectively suppress the lift-off interference in the MFL signal of the defects in the rail head surfaces.

Keywords: rail defects; MFL detection; filtering method; array sensor

Citation: Jia, Y.; Lu, Y.; Xiong, L.; Zhang, Y.; Wang, P.; Zhou, H. A Filtering Method for Suppressing the Lift-Off Interference in Magnetic Flux Leakage Detection of Rail Head Surface Defect. *Appl. Sci.* **2022**, *12*, 1740. https://doi.org/10.3390/app12031740

Academic Editor: Phong B. Dao

Received: 17 January 2022
Accepted: 3 February 2022
Published: 8 February 2022

1. Introduction

Railway is an important mode of public transportation in contemporary society. With the increase in railway running mileage and speed, the risk of the rail defects affected by repeated extrusion, friction, and impact of train wheel sets becomes higher [1].

To keep safe, some nondestructive testing techniques, such as ultrasonic, eddy current, and MFL, have been used in the detection of rail defects [2]. Ultrasonic detection is a widely used and mature rail detection technology [3,4]. It is a convenient method for detection of the internal defects of a rail, but due to the need for a coupling agent, ultrasonic detection cannot complete noncontact detection, so the detection speed cannot be further improved, and it is difficult to detect the surface defects by ultrasound.

The eddy current detection needs no exchange agent. It is cheap, simple, and reliable; however, traditional eddy current testing has obvious skin effect, which makes it difficult to detect subsurface defects and is susceptible to the specimen surface state [5].

MFL is an electromagnetic detection technique developed from magnetic particle detection [6]. The advantages of MFL are high sensitivity, high speed, and simple operation. It can detect the surface or subsurface defects [7], but the detection signal is easily affected by the lift-off change, the magnetic field excitation, and others.

The principle of MFL detection is shown in Figure 1. When magnetized ferromagnetic material has defects in its surface, such as cracks, laminations, or magnetic nonuniformities in the manufacturing process, the refraction of the magnetic induction line changes the magnetic circuit, which leads to the leakage of part of the magnetic flux. By detecting the magnetic field above the rail surface with sensors, a defect in the surface can be found and further information about it can be obtained. Multiple sensors are usually arranged along the width of a rail to detect the defect in the entire surface.

The vertical distance from the sensor or yoke to the rail is called lift-off. The LMF of a defect near the rail surface changes quickly in the vertical direction. During detection, the lift-off changes affected by vibration and other factors, thus interfering with the detection

of the defects. With the development of MFL detection [8–10], various requirements are proposed for railway detection equipment, such as high speed, high efficiency, and high resolution, so interferences such as lift-off interference must be suppressed.

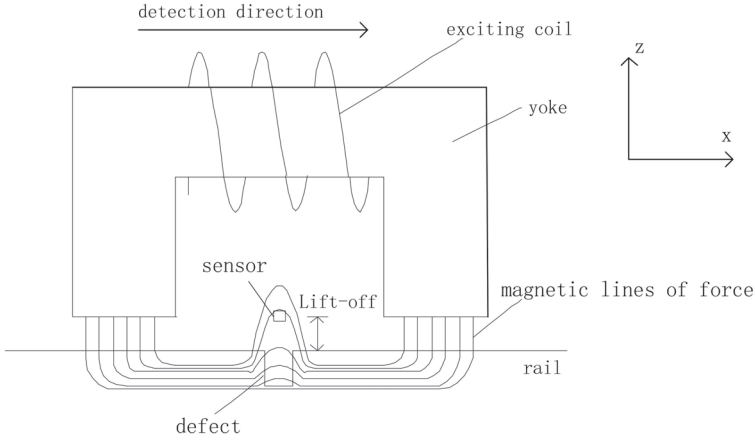


Figure 1. MFL detection.

The interference can be suppressed by hardware such as differential circuit [11], but as the interference is affected by detection speed, the state of the rail surface, and other factors, a circuit is difficult to apply to various situations. Therefore, in addition to hardware filtering, digital filtering is needed. The rail head upper surface defect is a common early defect, and it is usually defected by MFL. This paper presents a method to suppress the lift-off interference in MFL detection.

2. Related Works

Researchers have worked a lot on MFL detection for defects, such as the influence of speed on MFL signal [12], magnetization time of high-speed MFL detection [13], and excitation and vibration interference on the signal. In view of the interference, several methods are proposed to suppress the interference.

Karuppasamy optimized the MFL technique by finite element model to detect outer surface defects in ferromagnetic steam generator tubes. In this method, the iron core, coil, coil current, and position of the sensor are optimized, and the MFL signal is predicted. The influence of length, width, and depth of defects on the MFL signal is analyzed. The reliability of detection is improved [14].

A probe consisting of a quantum well Hall-effect sensor, an illuminating electromagnet and sensor circuitry, is designed. This probe is used to apply magnetic fields of various frequencies and field strengths to ascertain a frequency and field range best suited to detecting longitudinal surface-breaking toe cracks in ground mild steel welds [15].

As for the denoising process of the MFL signals, a multilevel filtering approach based on wavelet denoising combined with median filtering is proposed. By analyzing and comparing the denoising properties of three wavelet families, two wavelet bases with the best denoising performance are recognized and selected. Then, the median filtering method is cascaded [16].

To cover the entire workpiece or obtain sufficient data, sensor arrays are used [17]. The researchers carefully set up the position of the sensors in the probe and the relative position of the sensors [18,19]. Because the distance between the sensors is close and their vibration is similar, the lift-off interference can be suppressed through the relationship between the signals.

Ding Shunyi proposes a noise reconstruction algorithm of transverse sensor array and constructed an adaptive interference canceller to filter the multichannel MFL signal of rail treads [20].

Ji Kailun proposed a way of building a virtual channel signal which does not contain any defect. Multiple sensors are arranged, the sensor with the minimum absolute value of signal is considered as having not detected any defect, and this signal is used as a reference signal for filtering [21]. This method suppresses interference to a certain extent, but the lift-off interference and the leakage magnetic field of a defect may be in the reversed phase. It is unreasonable to take the minimum as the reference signal, which will cause certain misjudgment.

A defect detection method of longitudinal array sensors is proposed according to the characteristics of the signals detected by the magnetic sensor in x and z directions [22]. The method reduces the misjudgment caused by lift-off interference but is only suitable for judging whether there is a defect and cannot be used for filtering.

The sensor signal changes caused by vibration are found to suppress the lift-off interference, but their methods to find the signal changes are unreasonable [22]. Inspired by these references, this paper presents a method to determine the signal change caused by vibration through the difference between the signals in x and z directions of two sensors, to suppress the lift-off noise.

3. Filtering Method

3.1. MFL Analysis

The detection range of a single magnetic sensor is small because the LMF of a defect is always weak and small. Two rows of magnetic sensors are arranged along the y direction to cover the entire surface of the rail head, as shown in Figure 2. The defects and varied lift-off both change the output of the sensors.

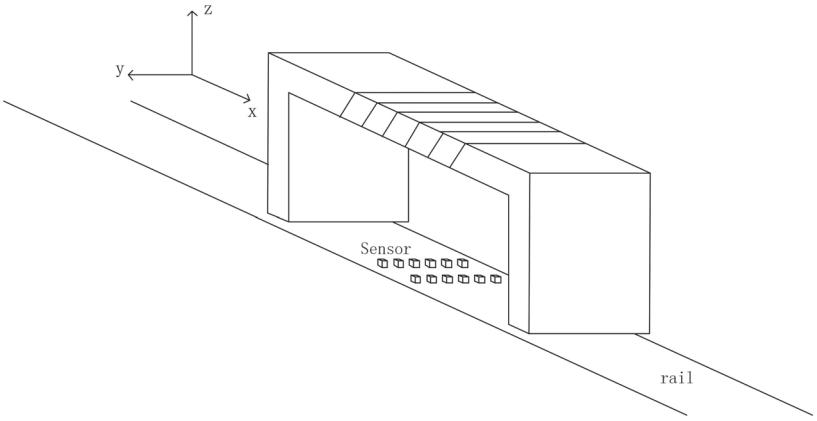


Figure 2. Array sensors layout.

The surface defects are generally distributed near the edge or the center of the rail head, and usually not across the entire surface or perpendicular to the x direction. Therefore, when the sensors in a row pass through a defect, some magnetic sensors may detect the defect and their outputs change while other sensors do not detect the defect and their outputs do not change by the defect. The output of a sensor that detects the defect is affected by both the defect and the change of the lift-off, while the output of a sensor that does not detect the defect is affected only by the change of the lift-off.

Figure 3 shows the x and z plane near a defect. The defect width and depth are denoted as 2a and b, respectively. The distance between the front and rear sensors rows is denoted as l, and the lift-off is denoted as d. The magnetic field intensity of the point P(x,z)

above the defect is $H(x,z)$ and its x and y component are $H_x(x,z)$ and $H_z(x,z)$, respectively. $H_x(x,z)$ and $H_z(x,z)$ can be obtained by Equations (1) and (2), respectively [19].

$$H_x = H_x(x,z) = \frac{\sigma_{ms}}{2\pi}\left[\tan^{-1}\frac{b+z}{a-x} + \tan^{-1}\frac{b+z}{a+x} - \tan^{-1}\frac{z}{a-x} - \tan^{-1}\frac{z}{a+x}\right] \quad (1)$$

$$H_z(x,z) = \frac{\sigma_{ms}}{4\pi}\ln\frac{\left[(x+a)^2+(z+b)^2\right]\left[(x-a)^2+z^2\right]}{\left[(x-a)^2+(z+b)^2\right]\left[(x+a)^2+z^2\right]} \quad (2)$$

where σ_{ms} is the magnetic charge density of the defect side, which can be calculated by Equation (3):

$$\sigma_{ms} = 5.3\left(\frac{b/a+1}{b/(a\mu)+1}\right)H \quad (3)$$

where μ is the magnetic permeability of the material, and H is the applied magnetic field strength.

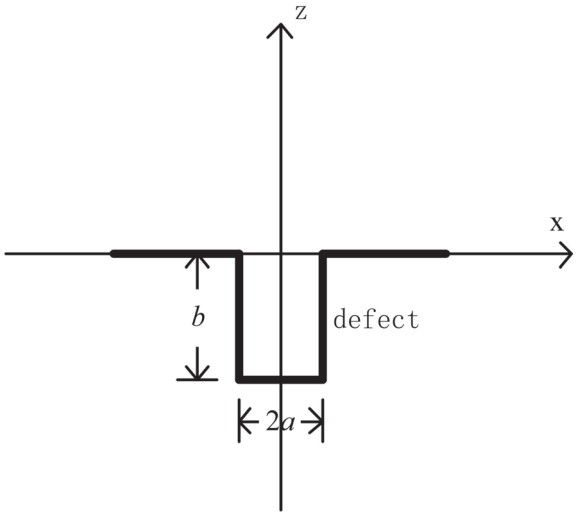


Figure 3. The x and z plane near a defect.

According to Equations (1) and (2), if z = 1 mm, a = 1 mm, b = 1 mm, and σ_{ms} = 1, the magnetic field distribution in x and z directions of the defect is shown in Figure 4.

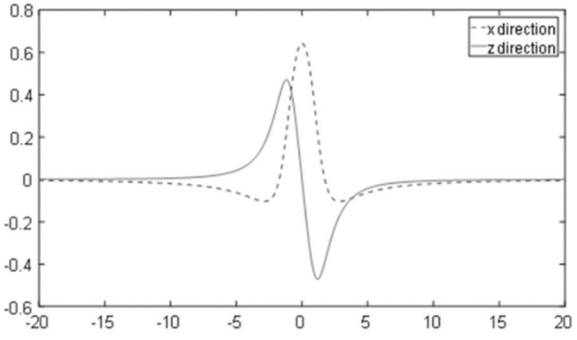


Figure 4. Magnetic field distribution in the x and z directions.

3.2. Principle of Filtering Method

According to Figure 4, the maximum of an MFL in the x direction of a defect is directly above the defect, while the maximum and minimum of the MFL in the z direction is on both sides of it. Two sensors are installed in the front and rear: a sensor which is used to detect the LMF in the z direction is in the front, and the other sensor, which is used to detect the LMF in the x direction, is installed in the rear. When the rear sensor is directly above a defect, its output is the maximum. At this time, the output of the front sensor is negative. The absolute value of the difference between the outputs is larger than that of each output.

Without any defect, the lift-off interference in x and z directions are in the same direction. The smaller the lift-off, the greater the output in both x and z directions. In a no-defect area, if one sensor measures the LMF in the x direction and the other sensor measures the LMF in the z direction, the absolute value of the difference between their output signals is smaller than that of each output.

Two sensors which detect the x and z directions MFL, respectively, are a pair, and several pairs of sensors are arranged in y the direction, as shown in Figure 2. If the distance between a pair of sensors is reasonable, the smaller the absolute value, the more likely it is that no defect is measured. As described in Section 3.1, there are always several sensors in a row that do not detect the LMF of a defect. It can be considered that the pair of sensors with the smallest absolute value of the difference of their outputs does not detect the LMF of a defect, and their outputs are taken as a reference signal that does not contain the defect signal but only the lift-off interference.

The absolute value of the difference between the outputs of the two sensors should be as large as possible to find the reference signal accurately. If a sensor detects the maximum in the x direction and the other sensor detects the minimum in the z direction, the absolute value of the difference between the outputs of the two sensors is the maximum.

According to Equation (1), if the lift-off is z_0, which remains unchanged, $H_x(x, z_0)$ is the maximum when x = 0. According to Figure 4, the minimum of $H_z(x, z_0)$ occurs on the positive half-axis of x. Due to the small defect, the magnetic dipole model can be introduced to simplify Equation (2) [19].

$$H_z(x, z) = \frac{\sigma_{ms}}{4\pi} \ln \frac{\left[(x-a)^2 + z^2\right]}{\left[(x+a)^2 + z^2\right]} \tag{4}$$

As the sensors are close to each other and their lift-offs are similar, $z = z_0$. The derivative of Equation (4) is taken to x and set as equal to 0.

$$H'_z(x, z_0) = 0 \tag{5}$$

Equation (5) is then solved, that is, $H_z(x, z_0)$ is the maximum when $x = -\sqrt{a^2 + z_0^2}$, and $H_z(x, z_0)$ is the minimum when $x = \sqrt{a^2 + z_0^2}$. Obviously, if the distance between two sensors is $\sqrt{a^2 + z_0^2}$, when the rear sensor detects the maximum in the x direction, the front sensor detects the minimum in the z direction.

In actual detection, the spacing, which is denoted as l between two sensors, is determined as the minimum width and depth of the defect that should be detected. Figure 5 shows the distribution of LMFs of the defects with different sizes. The distance of 10 is calculated according to the minimum defect. For a large defect, its LMF maximum in the x direction is larger than that of a small defect. When a sensor is direct above a larger defect and its output is the maximum in the x direction, the other sensor with a distance of l0 from this sensor does not measure the minimum in the z direction. Nevertheless, its output is smaller than the minimum of the minimal defect. In other words, the absolute value of the difference between the outputs of a pair of sensors is greater than that of a small defect.

Therefore, the spacing of the two sensors set to a small defect is also suitable for detecting large defects.

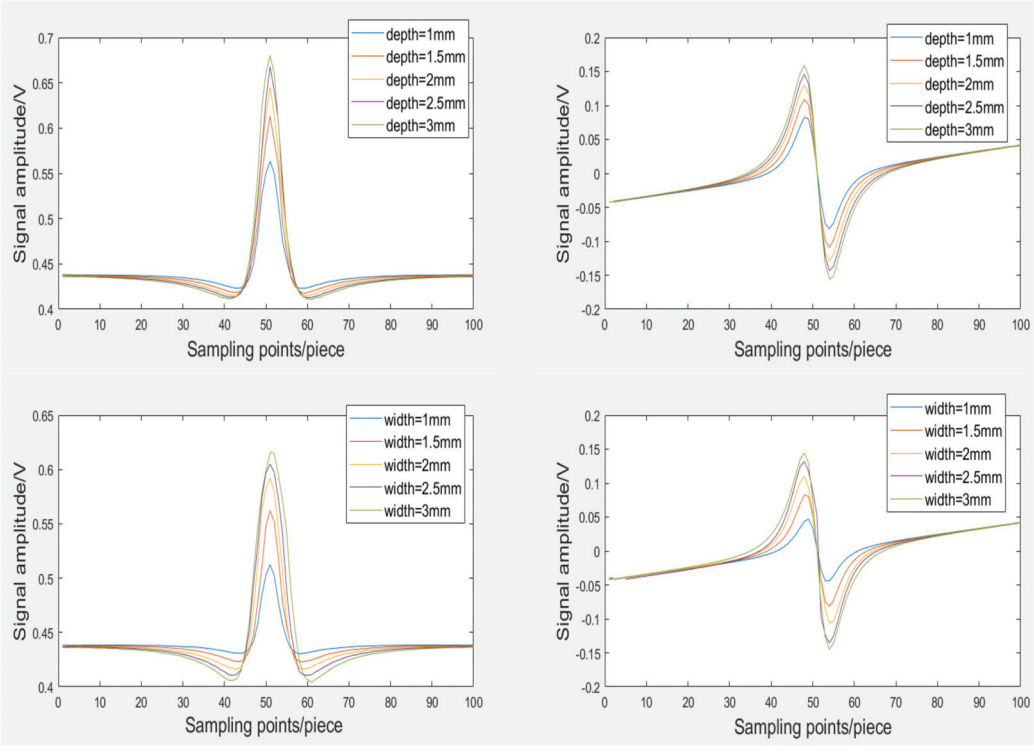


Figure 5. Magnetic field distribution of defects in x and z directions of different sizes.

In Figure 5, the LMF of a defect is within a certain range. To accurately identify and reconstruct a defect, the LMF should be measured as completely as possible. However, at a distance from the defect, the LMF is usually weak and difficult to distinguish from interference. The boundary of LMF in this paper is that the absolute value of LMF reduces to 10% of the maximum in the z direction. Beyond this range, the LMF is relatively small and negligible.

According to Formula (1), if $z = z_0$ unchanged, $H_z(x, z_0)$ is the maximum when $x = -\sqrt{a^2 + z_0^2}$. The leakage magnetic field distributes in 2L centered on the defect. According to Equation (4), L is calculated by Equation (6).

$$\ln \frac{\left[(-L-a)^2 + z_0^2\right]}{\left[(-L+a)^2 + z_0^2\right]} = \frac{1}{10} \ln \frac{\left[\left(-\sqrt{a^2 + z_0^2} - a\right)^2 + z_0^2\right]}{\left[\left(-\sqrt{a^2 + z_0^2} + a\right)^2 + z_0^2\right]} \tag{6}$$

The detection speed is v m/s, and the sampling speed is s point/s. J_0 is a sampling point in the z direction. If the distance between j_0 and the center of a defect is less than $L + \sqrt{a^2 + z_0^2}$, there will be a maximum point in the x direction and a minimum point in the z direction from sampling point $j_0 - B$ to $j_0 + B$.

$$B = \frac{\left(L + \sqrt{a^2 + z_0^2}\right)s}{v} \tag{7}$$

The absolute value of the difference between the output signals of the two sensors is large at the maximum point, so it is beneficial to find the reference signal. The absolute value of the difference between the outputs of two sensors in sampling point j is denoted by aj. When sampling point j_0 is processed, $aj_0 - B$ to $aj_0 + B$ are calculated first. Secondly, the maximum in $aj_0 - B$ to $aj_0 + B$ is found. The maximum of ith pair of sensors is denoted by MAXi. At last, the pair of sensors with the smallest MAXi (i = 1,2, ... ,N) is found and their output is taken as the reference signal.

3.3. Filtering Algorithm

As shown in Figure 2, two rows of sensors are arranged at intervals along the y direction below a yoke to detect magnetic fields in the x and z directions, respectively. The distance between the two rows of magnetic sensors is $\sqrt{a^2 + z_0^2}$. The two magnetic sensors Sx[i] and Sz[i] (i = 1, 2, ... , N) are a pair, and there are N pairs of magnetic sensors. The filtering steps are as follows.

1. Each sensor output is sampled, and the number of sampling points is denoted by M. The sampling results of Sx[i] and Sz[i] are array Sx[i,j] and Sz[i,j] (j = 1, 2, ... , M).
2. The array $R_x[i,k]$ and $R_z[i,k]$ is calculated:

$$R_x[i,k] = \begin{cases} S_x[i,1] & 1 \le k \le B \\ S_x[i,j] & B < k \le M + B \\ S_x[i,M] & M + B < k \le M + 2B \end{cases}$$

$$R_z[i,k] = \begin{cases} S_z[i,1] & 1 \le k \le B \\ S_z[i,j] & B < k \le M + B \\ S_z[i,M] & M + B < k \le M + 2B \end{cases}$$

   q is a loop variable with an initial value of B + 1.
3. $MAX[i,q] = MAX|R_x[i,p] - R_z[i,p]|$ $(p = q - B,\ q - B + 1, \ldots, q + B)$.
4. The minimum in MAX[..., q] (i = 1, 2, ... , N) is found. The number of the pair that has the minimum is denoted by i0. The sampling point of the pair sensors are taken as the reference signal. $R_x[i,q] - R_x[i_0,q]$, $R_z[i,q] - R_z[i_0,q]$ (i = 1, 2, ... , N): The differences are taken as the filtering results of each sensor at this sampling point.
5. q = q + 1. If $q \le M + B$, the process returns to step 3, otherwise the filtering ends.

4. Experimental Results and Analysis

4.1. Finite Element Simulation Results and Analysis

The finite element analysis software used in this paper is Ansys Maxwell. An MFL detection model is built as shown in Figure 6. The yoke is made of ferromagnetic material. The excitation coil adopts 4000 turns of copper wire. The DC excitation voltage is 60 V, and the lift-off between the rail and the yoke is 2.0 mm. Two detection points with a lift-off of 2.0 mm are set up above the rail.

The width and depth of the defect are 2.0 mm and 1.0 mm, respectively and the lift-off is 2.0 mm, which means that a = 1.0 mm, b = 1.0 mm, and z = 2.0 mm. According to Formula (4), when the distance between a pair of sensors is $\sqrt{a^2 + z^2}$, which is 2.2 mm, the absolute value of the difference between the outputs of the two sensors is the maximum. The distance between two sensors is set to 1.8 mm, 2.2 mm, and 2.6 mm, respectively. The simulation results are shown in Figure 7a–c.

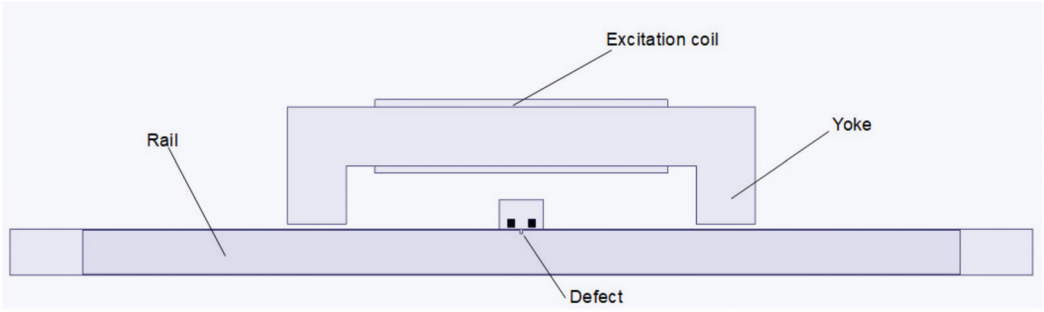
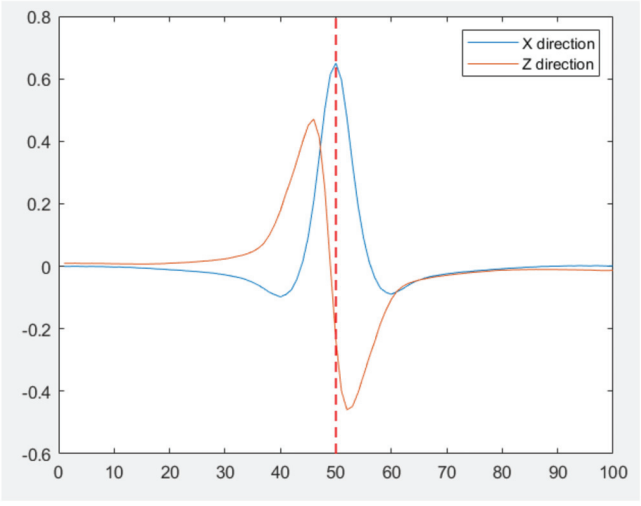
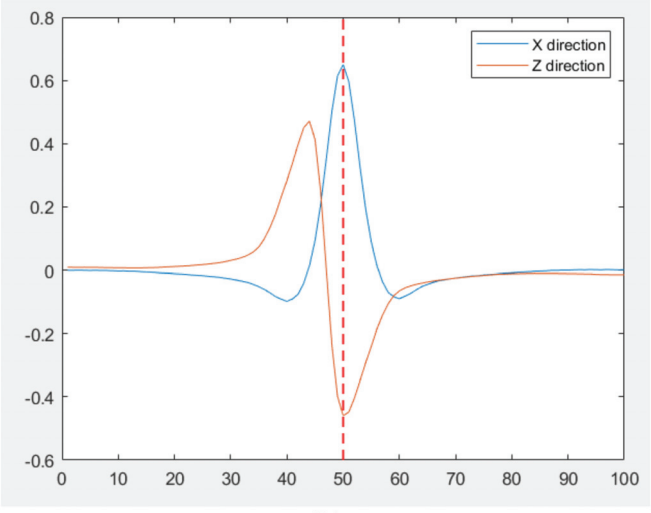


Figure 6. Simulation model.

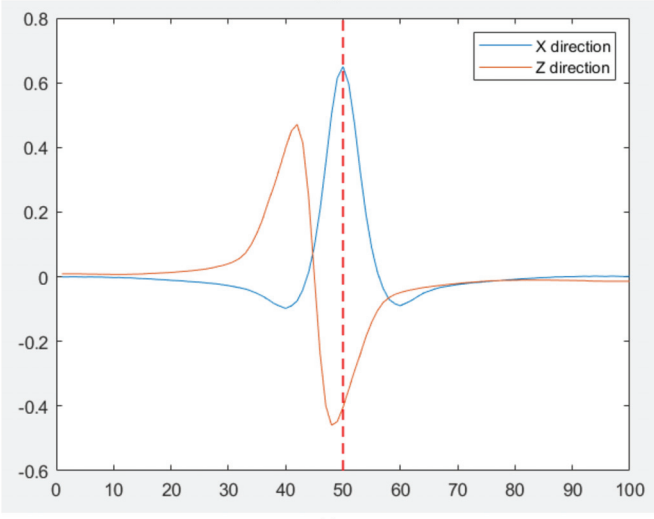


(**a**)



(**b**)

Figure 7. *Cont.*



(**c**)

Figure 7. The x and z direction simulation results with different distance. (**a**) 1.8 mm; (**b**) 2.2 mm; (**c**) 2.6 mm.

According to the simulation results in Figure 7a–c, the phase difference of the pair of sensors becomes increasingly significant with the increase in the distance between them. When the distance is 2.2 mm, one sensor detects the maximum in the x direction and the other sensor detects the minimum in the z direction, which is indicated by the red line in Figure 7. The absolute value of the difference between them is the maximum which is in accordance with the above conclusion.

4.2. Physical Experiment Results and Analysis

4.2.1. Experiment System

An experimental system was built and is shown in Figure 8. The system consists of a computer, a signal conditioning circuit, a DAQ card, and a probe. The sensor is an SL-106C, whose length and width are both 1.5 mm. The amplifier is AD620, and the magnification is 100. The DAQ card is ADLINK DAQ2208, and the sampling frequency of each channel is set to 4 kHz. The probe, which consists of two rows, eight sensors per row, and an excitation device, is mounted on a detection car that is placed on the rail. The detection speed was 2.0 m/s.

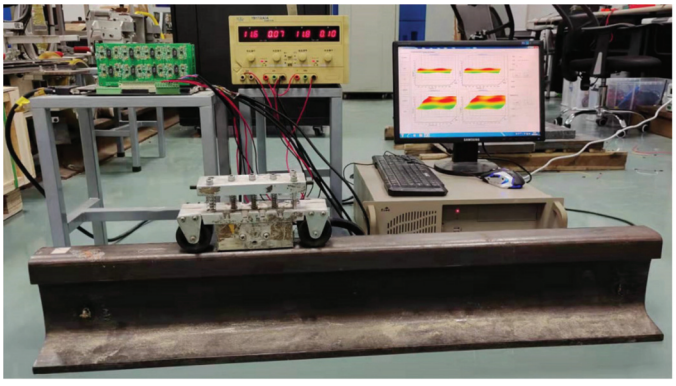


Figure 8. MFL detection system.

4.2.2. Single Defect Experiment

The detection object is an artificial defect in the surface of a used rail, as shown in Figure 9. The direction of detection is from left to right. A wire is tied to the rail, which will change the lift-off of the handcart to simulate the vibration. If no vibration is encountered, the lift-off of the sensor is 2.0 mm, and when the handcart wheel rolls over the wire, the lift-off will fluctuate within 2.0~3.5 mm. The defect is 2.0 mm in width and 1.0 mm in depth. Therefore, a = 1.0 mm, z_0 = 2.0 mm. The distance between a pair of sensors is $\sqrt{a^2 + z_0^2}$, that is 2.2 mm. L is 41.4 mm and B is 87, according to Formulas (5) and (6), respectively.

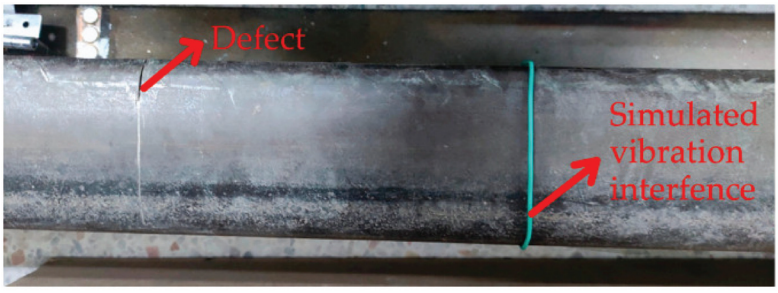
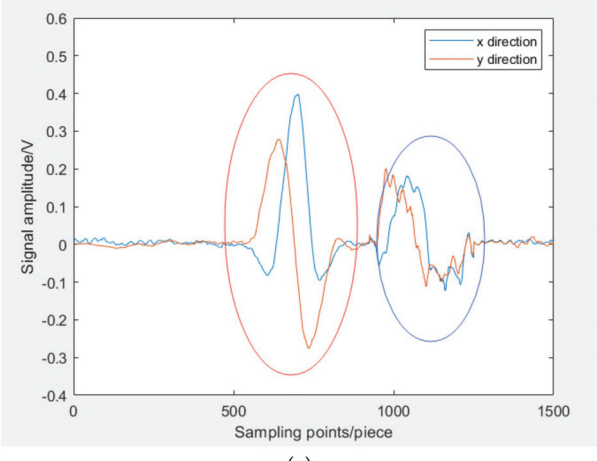


Figure 9. A rail with artificial defect.

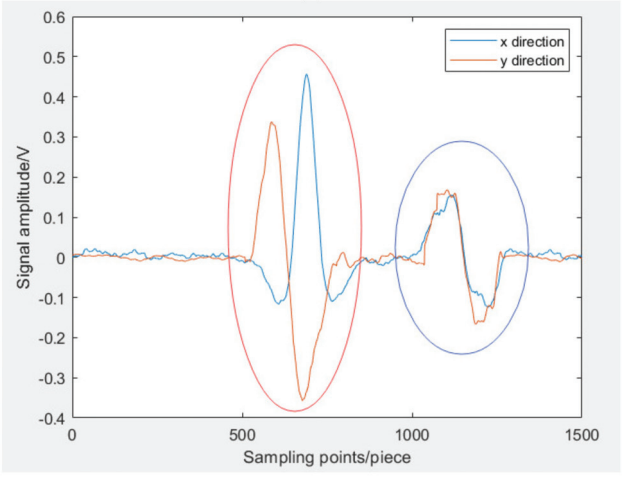
A pair of sensors which detect the defect in x and z directions, respectively, are arranged in the probe. The distance between them is set to 1.8 mm, 2.2 mm, and 2.6 mm respectively. The detection results are shown in Figure 10.

According to Figure 10a–c, the phase of the signal of two sensors is different near the defect and when the sensors pass through the defect, the difference in the phase increases with increases in the sensor distance, which is shown in the red region. As shown in the blue region, if vibration has occurred, the phases of the two signals are basically similar. Therefore, the phase relationship of the two signals can be used to determine whether there is a defect. In addition, when the distance between a pair of sensors is $\sqrt{a^2 + z_0^2}$, which is 2.2 mm, a sensor detects the maximum in the x direction and the other detects the minimum in the z direction, and the absolute value of the difference between the outputs of the two sensors is the maximum.

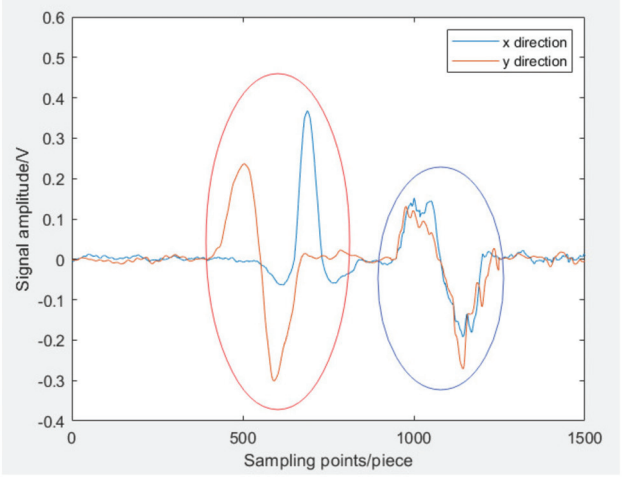


(a)

Figure 10. *Cont.*



(**b**)



(**c**)

Figure 10. Detection results with different distances. (**a**) 1.8 mm; (**b**) 2.2 mm; (**c**) 2.6 mm.

4.2.3. Multiple Defects Experiment

The detection results in the x and z directions of four different defects are shown in Figure 11a–d, respectively. The lift-off is 2.0 mm. The distance between two sensors is 2.2 mm.

As shown in Figure 11a–d, the amplitude of the signal will increase when the depth of the defect increases. When the distance between a pair of sensors is 2.2 mm in Figure 11a,b, one sensor detects the maximum in the x direction and the other detects the minimum in the z direction, and the absolute value of the difference between the outputs of the two sensors is the maximum. As shown in Figure 11c,d, the width becomes larger when a sensor is directly above the larger defect and its output is the maximum in the x direction; the other does not measure the minimum in the z direction, but the absolute value of the difference between them is greater than that of a small defect. Therefore, the distance calculated according to the minimum defect is also applicable to the detection of large defects.

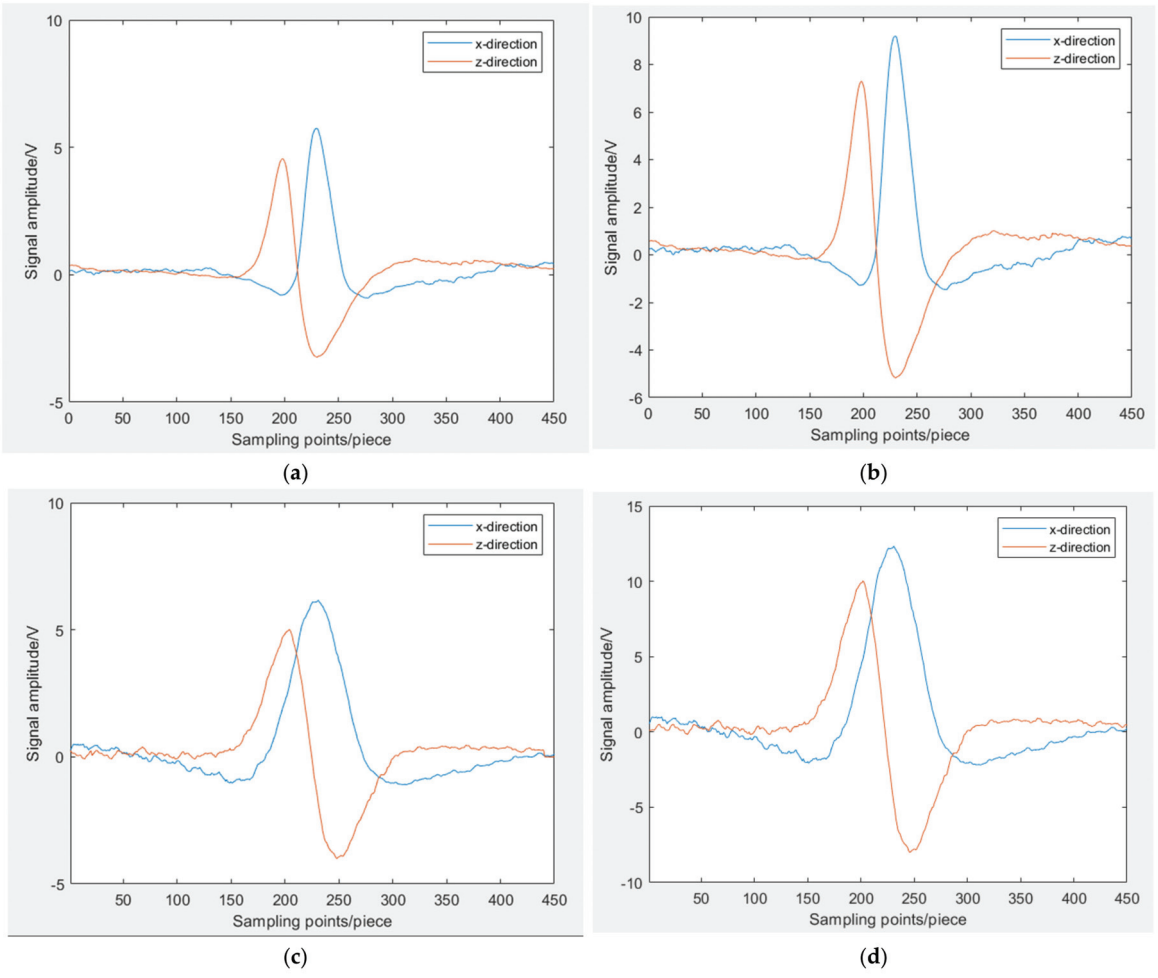


Figure 11. Detection results with different defects. (**a**) width = 2 mm, depth = 1 mm; (**b**) width = 2 mm, depth = 2 mm; (**c**) width = 3 mm, depth = 1 mm; (**d**) width = 3 mm, depth = 2 mm.

A used rail with eight artificial defects is detected, and a wire is tied to the rail also, as shown in Figure 12. The width and depth of each defect are shown in Table 1. The lift-off of the sensor is also fixed at 2.0 mm. The distance between the two sensors is 2.2 mm.

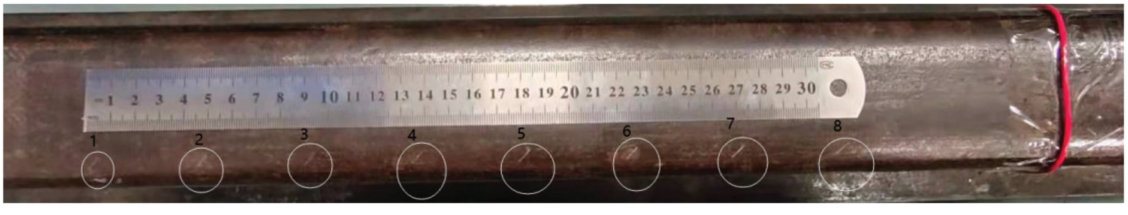


Figure 12. A rail with multiple artificial defects.

Table 1. The width and depth of each defect.

	Defect 1	Defect 2	Defect 3	Defect 4	Defect 5	Defect 6	Defect 7	Defect 8
Width	2.0 mm	2.2 mm	2.2 mm	2.5 mm	2.5 mm	2.5 mm	2.5 mm	2.5 mm
Depth	1.0 mm	1.2 mm	1.2 mm	2.0 mm	2.5 mm	2.5 mm	2.8 mm	3.0 mm

Figure 13 shows a multichannel three-dimensional diagram of the sensors. The sensors near the outer edge all detected defect signals, and all the sensors were affected by lift-off interference and the vibration signals changed synchronously. Figure 14 shows the original signal of the seventh sensor in blue and the reference signal in red. The vibration signals are marked with rectangular boxes and the defect signals are marked with elliptical boxes. Obviously, the amplitude of the signal caused by vibration is similar to that caused by a small defect, so the detection is disturbed by vibration. The reference signal retains the changes caused by vibration and contains no defect signal.

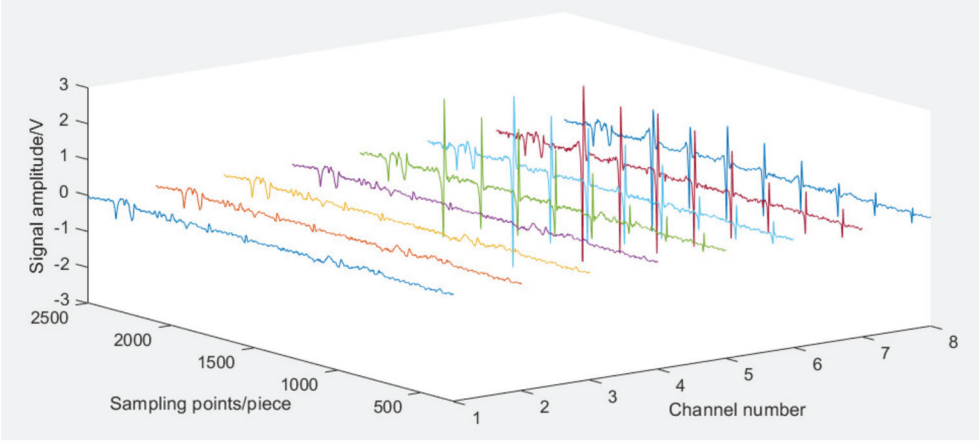


Figure 13. Multichannel original MFL signal.

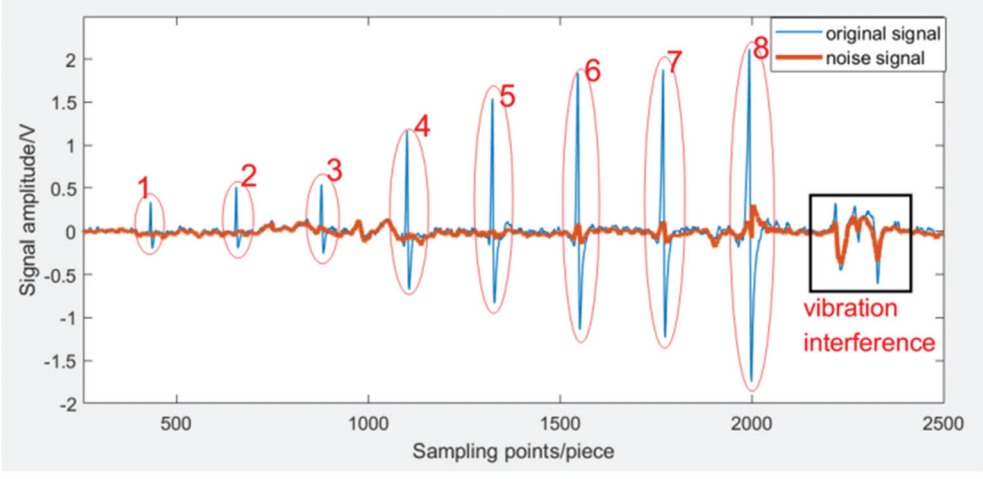


Figure 14. Original signal and vibration signal of the seventh sensor.

The filtered signal is shown in Figure 15. Compared with the original signal in Figure 14, the defect signal is basically unchanged, while the vibration signal decreases significantly. Figure 16 shows the three-dimensional image of the multichannel filtered signal. After filtering, the vibration interference is obviously reduced.

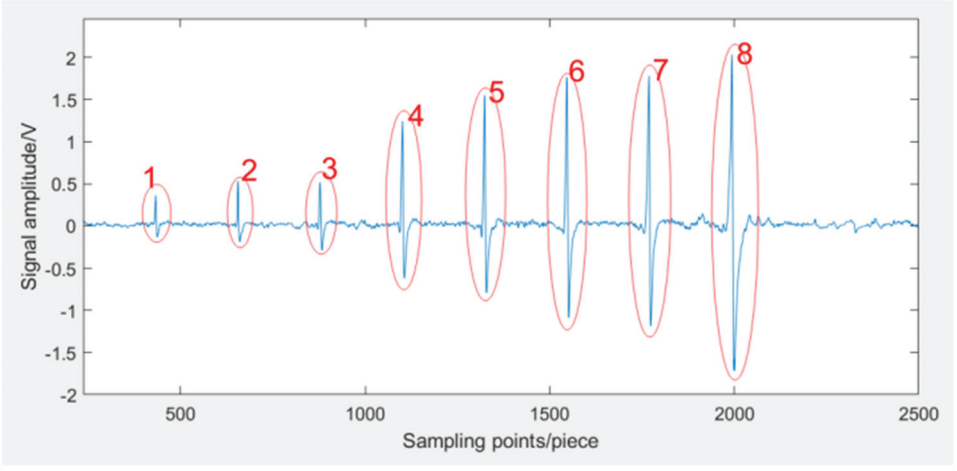


Figure 15. Filtered signal of the seventh sensor.

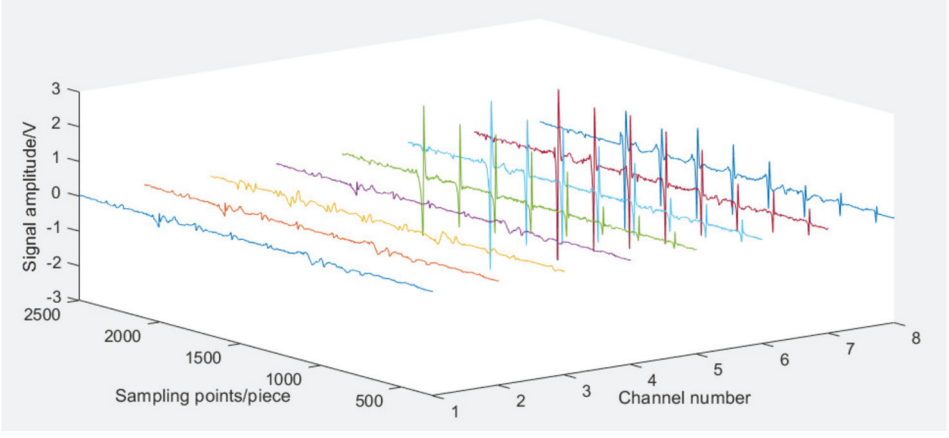


Figure 16. Multichannel filtered signal.

5. Conclusions

In this paper, a filtering method is proposed to suppress the lift-off interference of rail defect MFL detection. According to the different signal characteristics of sensors in x and z directions, the structure of the array sensor is adopted, which can effectively extract the reference signal that contains only interference signals such as lift-off. Thus, the lift-off interference can be suppressed, and a relatively complete waveform of defect leakage magnetic field detection will be obtained. In the future, the natural defects detection test will be carried out, and the identification and reconstruction methods of defects will be studied.

Author Contributions: Conceptualization, Y.J. and Y.L.; methodology, Y.J.; software, Y.L.; validation, Y.J., Y.L. and L.X.; formal analysis, Y.J.; investigation, Y.Z.; resources, Y.L.; data curation, Y.L.; writing—

original draft preparation, Y.L.; writing—review and editing, Y.J.; visualization, H.Z.; supervision, P.W.; project administration, P.W.; funding acquisition, P.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the National Key R&D Program of China (2018YFB2100903) and Scientific Research Projects of China Academy of Railway Sciences (2020YJ153).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhang, H.; Song, Y.N.; Wang, Y.N. Review of rail defect non-destructive testing and evaluation. *Chin. J. Sci. Instrum.* **2019**, *40*, 11–25.
2. Alahakoon, S.; Sun, Y.Q.; Spiryagin, M.; Cole, C. Rail Flaw Detection Technologies for Safer, Reliable Transportation: A Review. *J. Dyn. Syst. Meas. Control* **2018**, *140*, 020801. [CrossRef]
3. Fu, L.Z.; Wu, J.J.; Lang, X.W. Experiment research on switch rail defect monitoring using ultrasonic guided waves. *Railway Engineering* **2018**, *31*, 129–133.
4. Miki, M.; Ogata, M. Phased array ultrasonic testing methods for welds in bogie frames of railway vehicles. *Insight* **2015**, *57*, 382–388. [CrossRef]
5. Hu, Q.W.; Wang, X.B.; Wu, W. Eddy current testing and evaluation method for fine cracks on rail surface. *Equip. Manag. Maint.* **2021**, *42*, 34–35.
6. Wilson, J.W.; Tian, G.Y. 3D magnetic field sensing for magnetic flux leakage defect characterisation. *Insight* **2006**, *48*, 357–359. [CrossRef]
7. Gao, L. Research on Image Detection and Recognition Method of Steel Plate Surface Defect Based on Magnetic Flux Leakage Signal. Master's Thesis, Northeastern University, Shen Yang, China, 2017.
8. Azari, H.; Al Ghorbanpoor; Shams, S. Development of Robotic Nondestructive Testing of Steel Corrosion of Prestressed Concrete Bridge Girders using Magnetic Flux Leakage System. *Transp. Res. Record.* **2020**, *2674*, 466–476. [CrossRef]
9. Peng, X.; Anyaoha, U.; Tsukada, K. Analysis of Magnetic-Flux Leakage(MFL) Data for Pipeline Corrosion Assessment. *IEEE Trans. Magn.* **2020**, *57*, 6200410. [CrossRef]
10. Piao, G.Y.; Guo, J.B.; Hu, T.H.; Leung, H.; Deng, Y.M. Fast reconstruction of 3-D defect profile from MFL signals using key physics-based parameters and SVM. *NDT E Int.* **2019**, *103*, 26–38. [CrossRef]
11. Wu, D.H.; Su, L.X.; Wang, X.H. A Novel Non-destructive Testing Method by Measuring the Change Rate of Magnetic Flux Leakage. *J. Nondestruct. Eval.* **2017**, *36*, 24.
12. Usarek, Z.; Chmielewski, M.; Piotrowski, L. Reduction of the Velocity Impact on the Magnetic Flux Leakage Signal. *J. Nondestruct. Eval.* **2019**, *38*, 28. [CrossRef]
13. Feng, B.; Kang, Y.H.; Sun, Y.H. Magnetization Time Lag Caused by Eddy Currents and Its Influence on High-Speed Magnetic Flux Leakage Testing. *Res. Nondestruct. Eval.* **2019**, *30*, 189–204. [CrossRef]
14. Karuppasamy, P.; Abudhahir, A.; Jayakumar, T. Model-Based Optimization of MFL Testing of Ferromagnetic Steam Generator Tubes. *J. Nondestruct. Eval.* **2016**, *35*, 5. [CrossRef]
15. Watson, J.M.; Liang, C.W.; Sexton, J.; Missous, M. Magnetic field frequency optimisation for MFL imaging using QWHE sensors. *Insight* **2020**, *62*, 396–401. [CrossRef]
16. Zhang, O.; Wei, X.Y. De-noising of Magnetic Flux Leakage Signals Based on Wavelet Filtering Method. *Res. Nondestruct. Eval.* **2019**, *30*, 269–286. [CrossRef]
17. Liu, S.W.; Sun, Y.H.; Deng, Z.Y.; He, L.S. A novel method of omnidirectional defects detection by MFL testing under single axial magnetization at the production stage of lathy ferromagnetic materials. *Sens. Actuator A-Phys.* **2017**, *262*, 35–45. [CrossRef]
18. Liu, X.C.; Wang, Y.J.; Wu, B.; He, C.F. Design of Tunnel Magnetoresistive-Based Circular MFL Sensor Array for the Detection of Flaws in Steel Wire Rope. *J. Sens.* **2016**, *6*, 198065.
19. Wu, D.H.; Liu, Z.T.; Su, L.X. New MFL detection method based on differential peak extraction using dual sensors. *Chin. J. Sci. Instrum.* **2016**, *37*, 1218–1225.
20. Ding, S.Y. Research on De-Noising and Recognition of Rail Magnetic Flux Leakage Signal Based on Adaptive Filtering and Random Forest. Master's Thesis, Nanjing University of Aeronautics and Astronautics, Nan Jing, China, 2020.
21. Ji, K.L.; Wang, P.; Jia, Y.L.; Ye, Y.F.; Ding, S.Y. Adaptive Filtering Method of MFL Signal on Rail Top Surface Defect Detection. *IEEE Access* **2021**, *9*, 87351–87359. [CrossRef]
22. Jia, Y.L.; Zhang, S.C.; Wang, P.; Ji, K.L. A Method for Detecting Surface Defects in Railhead by Magnetic Flux Leakage. *Appl. Sci.* **2021**, *11*, 9489. [CrossRef]

sensors

MDPI

Article

A New Design of the Dual-Mode and Pure Longitudinal EMAT by Using a Radial-Flux-Focusing Magnet

Xu Zhang [1], Weiwen Li [1], Bo Li [1], Jun Tu [1], Chunhui Liao [1], Qiao Wu [1], Sheng Feng [2] and Xiaochun Song [1,*]

[1] Hubei Key Laboratory of Modern Manufacturing Quantity Engineering, School of Mechanical Engineering, Hubei University of Technology, Wuhan 430068, China; zhangxu@mail.hbut.edu.cn (X.Z.); 102010061@hbut.edu.cn (W.L.); 102110151@hbut.edu.cn (B.L.); juntu@hbut.edu.cn (J.T.); 20161010@hbut.edu.cn (C.L.); qiao_wu@hbut.edu.cn (Q.W.)

[2] Xianning Institute of Information and Standardization, Xianning 437000, China; 101810129@hbut.edu.cn

* Correspondence: songxc@mail.hbut.edu.cn

Abstract: An electromagnetic acoustic transducer (EMAT) is suitable for measuring the propagation time more accurately without causing abrasion to the transducer during testing due to the principle of its excitation. This work designs a flux-concentrating EMAT with a radial-flux-focusing permanent magnet to significantly enhance static magnetic field strength. Through theoretical analysis and finite element simulation, two kinds of coils are designed according to the concentration areas of the horizontal and vertical components of the magnetic field. One is used to generate pure longitudinal waves, and the other is used to generate both longitudinal waves and shear waves. The experimental comparison shows that the amplitudes of the pure longitudinal wave and the dual-mode wave excited by the two kinds of coils with the radial-flux-focusing magnet are more than two times higher than those with the ordinary magnet. Therefore, the flux-concentrating EMAT with the appropriate coil provides an insight into realizing more accurate detection where longitudinal wave detection is required.

Keywords: electromagnetic acoustic transducer; longitudinal wave; flux-concentrating; transducer design

Citation: Zhang, X.; Li, W.; Li, B.; Tu, J.; Liao, C.; Wu, Q.; Feng, S.; Song, X. A New Design of the Dual-Mode and Pure Longitudinal EMAT by Using a Radial-Flux-Focusing Magnet. *Sensors* **2022**, *22*, 1316. https://doi.org/10.3390/s22041316

Academic Editors: Zahra Sharif Khodaei, Phong B. Dao, Lei Qiu and Liang Yu

Received: 3 January 2022
Accepted: 4 February 2022
Published: 9 February 2022

1. Introduction

Electromagnetic acoustic transducers (EMATs) are widely used in the field of non-destructive testing (NDT), and have the superiority of non-contact and no couplant with the specimen [1]. Unlike EMATs, the couplant is needed between the piezoelectric ultrasonic transducer and the specimen surface, so the elastic waves propagate not only in the specimen but also in the couplant and the transducer. Thus, the reflection echoes received by piezoelectric transducers are accompanied by interference echoes from the inside of the transducer and the interface of the couplant, which affects the calculation of the ultrasonic propagation time interval in the received signal [2]. However, the ultrasonic wave source excited by EMATs is inside the specimen, which is more suitable for measuring the propagation time accurately without causing abrasion to the transducer during testing [3]. At present, shear wave excitation is mainly used in EMATs for thickness measurement, flaw detection and defect detection in many industrial fields [4–7]. Due to the structural characteristics of shear wave EMATs [8–11], shear waves are easily generated and received. Longitudinal-wave EMATs have inefficient energy conversion because of the low horizontal magnetic field intensity required to generate longitudinal waves and the sizeable parasitic inductance of the coil [12,13]. However, the longitudinal-wave EMATs have a broad application prospect. Since the velocity of longitudinal waves is nearly twice that of shear waves, the longitudinal waves excited by electromagnetic ultrasound are more efficient in the detection of the thickness of large aluminum plates and other similar metals. Furthermore, the longitudinal waves combined with the shear waves excited by the EMAT

probe could take advantage of contactless detection in measurements of stress and elastic constants [14–18].

At present, there are few structures of longitudinal-wave EMATs. There are three typical designs: Hirao and Ogi [19] pointed out that a bulk wave EMAT that consists of a single-cylindrical permanent magnet and a spiral coil can generate longitudinal waves and radially polarized shear waves. Da Cunha and Jordan [20] proposed a longitudinal-wave EMAT consisting of a cylindrical permanent magnet and external magnetic rings, with an iron coupling between the internal magnetic rod and outer magnetic rings to provide a stronger horizontal magnetic field. Wu et al. [21] used a combination of a large-diameter center magnet and a ring magnet to enhance the strength of the horizontal magnetic field, and a sheet of copper was placed between the coil and the specimen to control the eddy current distribution, which aims to suppress shear waves. Considering the Lorentz forces as the dominant transduction mechanism, increasing the strength of the static magnetic field plays an important role in increasing ultrasonic wave amplitude. The optimization of the magnet arrangement has a significant effect on improving the static magnetic field [8,9,22,23]. In particular, the Halbach structure is better than the traditional soft iron backing for enhancing the static magnetic field and achieved good results in dual-mode excitation [24].

In this paper, a radial-flux-focusing magnet is proposed, inspired by Halbach's concept, which can provide a stronger magnetic field, thus improving the efficiency of the energy transfer of the longitudinal-wave EMAT and making the measurement signal easier to identify. This work also utilizes two kinds of spiral coils [25] by changing their size parameters according to the location of the area where the horizontal and vertical components of the magnetic field are concentrated. When dual modes are needed, the amplitude of shear waves and longitudinal waves simultaneously excited by the flux-concentrating EMAT can be increased using the large-diameter coil. When a pure longitudinal wave is needed, the amplitude of shear waves excited by the flux-concentrating EMAT with the small-diameter coil decreases, while the amplitude of longitudinal waves increases.

2. Configuration and Operating Principle of the Proposed Flux-Concentrating EMAT

The structure diagram of the ring-type EMAT is shown in Figure 1a. The permanent magnets of the transducer include a cylinder magnet in the middle and a surrounding circular magnet, and the gap between the two magnets is filled with epoxy resin. The poles of the circular-ring permanent magnet and the cylindrical permanent magnet are opposite at the same ends. Therefore, a horizontal radial magnetic field is distributed between the two magnets. Figure 1a shows a typical longitudinal-wave EMAT used as a comparison in this article since its structure is similar to the designed EMAT. Figure 1b shows the configuration of the flux-concentrating EMAT. A radial-magnetized annular magnet is filled between the cylinder magnet and the surrounding circular magnet, and another radial-magnetized annular magnet is also covered outside the surrounding circular magnet. The magnetizing directions of the magnets are all clearly illustrated in Figure 1.

The tested material used in this paper is aluminum, and in the non-ferromagnetic material, the EMAT excitation mechanism is the Lorentz force principle [26]. Figure 2 illustrates the generation of the longitudinal waves and shear wave in the non-ferromagnetic specimen. The white arrows in Figure 2 indicate the directions of Lorentz forces; the dotted lines indicate the direction of the magnetic field.

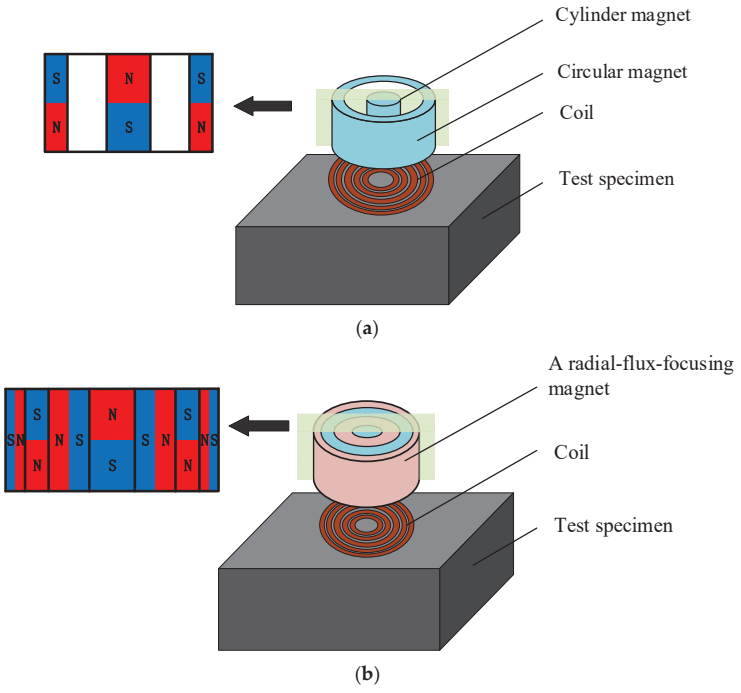


Figure 1. Configuration of (**a**) the ring-type EMAT and (**b**) the flux-concentrating EMAT.

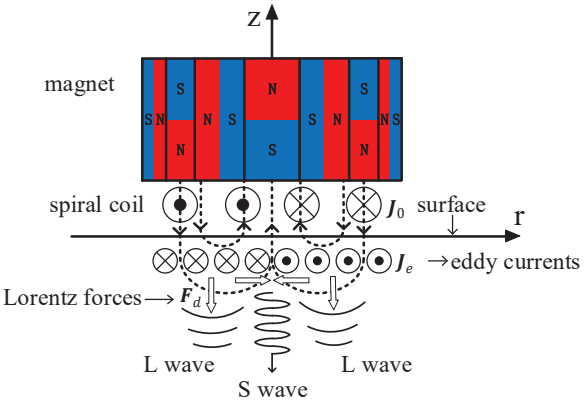


Figure 2. Schematic diagram of longitudinal wave and shear wave generation in EMAT.

A spiral coil is placed on the surface of the test piece, and an excitation current J_0 is pulsed through the spiral coil. The high-frequency dynamic magnetic field B_d will be induced in the test piece, and an eddy current J_e with the same frequency as the current in the coil is also induced in the surface skin depth. Under the action of the static magnetic field B_s and high-frequency dynamic magnetic field B_d, the particles on the specimen surface are subjected to the Lorentz forces F_s and F_d. F_s is the Lorentz force generated by the static magnetic field, and F_d is the Lorentz force generated by the dynamic magnetic field. As a result, high-frequency periodic vibration occurs on the surface of the specimen, and elastic deformation is generated. Thus, an ultrasonic wave is produced, propagating in

the specimen. The Lorentz force generated by the dynamic magnetic field B_d and the static magnetic field B_s on the specimen surface is [27]:

$$F_s = J_e \times B_s \tag{1}$$

$$F_d = J_e \times B_d \tag{2}$$

The total Lorentz force is:

$$F_L = F_s + F_d \tag{3}$$

In the process of ultrasonic excitation, the in-plane dynamic magnetic field of EMAT is relatively small compared with the out-of-plane dynamic magnetic field [26]. Therefore, an approximate calculation of F_d is made, and only the Lorentz force generated by the dynamic magnetic field in the *Z*-axis direction is considered. Considering that the static magnetic field from the magnet of the flux-concentrating EMAT has radial component B_{sr} and axial component B_{sz} into account, Equation (3) can be rewritten as:

$$F_s = F_r^{(s)} + F_z^{(s)} \tag{4}$$

$$F_s = J_e \times B_{sz} + J_e \times B_{sr} \tag{5}$$

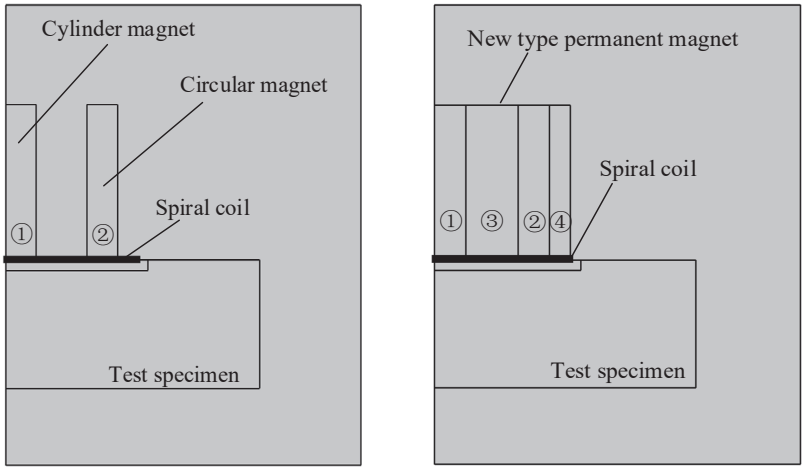
where $F_r^{(s)}$ and $F_z^{(s)}$ are Lorentz forces in the radial direction and axial direction, respectively.

Under the action of the static radial magnetic field, the Lorentz force on the charged particles of the specimen surface is perpendicular to the specimen surface in the *Z*-axis direction and parallel to the propagation direction of the ultrasonic wave, thus generating a longitudinal wave. Under the action of the static axial magnetic field, the vibration direction of particles on the specimen surface is in the *R*-axis direction. In other words, the direction of the Lorentz force is perpendicular to the propagation direction of the ultrasonic wave, thus generating shear waves. Lorentz forces concentrate on the specimen surface and generate time-dependent elastic stress waves in the specimen. Consequently, the Lorentz forces $F_r^{(s)}$ and $F_z^{(s)}$ generate longitudinal waves and shear waves, respectively, both propagating in the thickness direction at the same time [19]. The process of the EMATs receiving the signal is the inverse process of their transmission. The reflected echo reaches the specimen surface, making the surface particles vibrate and changing the current in the spiral coil under the action of the static magnetic field.

3. Simulation Analysis

3.1. Dynamic Magnetic Field in Specimen

Using the finite element software, COMSOL Multiphysics, two-dimensional axisymmetric solid simulation models are developed for the ring-type EMAT in Figure 1a and the flux-concentrating EMAT in Figure 1b, since they are all axisymmetric. Figure 3a,b shows the simulation model structures of the ring-type EMAT and the flux-concentrating EMAT with the left line as the symmetry axis, respectively. The parameters of magnets of the ring-type EMAT and flux-concentrating EMAT in finite element models are shown in Table 1. In models, the lift-off distance between the coil and the specimen is 0.2 mm, and the distance between the coil and the magnet is 0.3 mm. The right boundary of the 40 mm high aluminum specimen is set as a low reflection boundary to simulate the actual plate. The residual magnetic flux density of each permanent magnet is set as 1 T. By setting the frequency of the alternating current as 1 MHz, according to Equation (5), the Lorentz force is loaded into the Solid Mechanics Module to achieve the coupling of the electric, magnetic and elastic acoustic fields, so that the transmission of ultrasonic waves in the specimen can be observed in the time domain. Figure 4a,b shows the meshes of the two models. It should be noted that the maximum unit of the specimen mesh is set to 1/10 of the shortest wavelength to ensure the accuracy of the simulation.



(**a**) (**b**)

Figure 3. Model structure diagrams of (**a**) the ring-type longitudinal-wave EMAT and (**b**) the flux-concentrating longitudinal-wave EMAT. The No. 1 magnet is a cylinder magnet. The No. 2, No. 3 and No.4 magnets are circular magnets of different sizes.

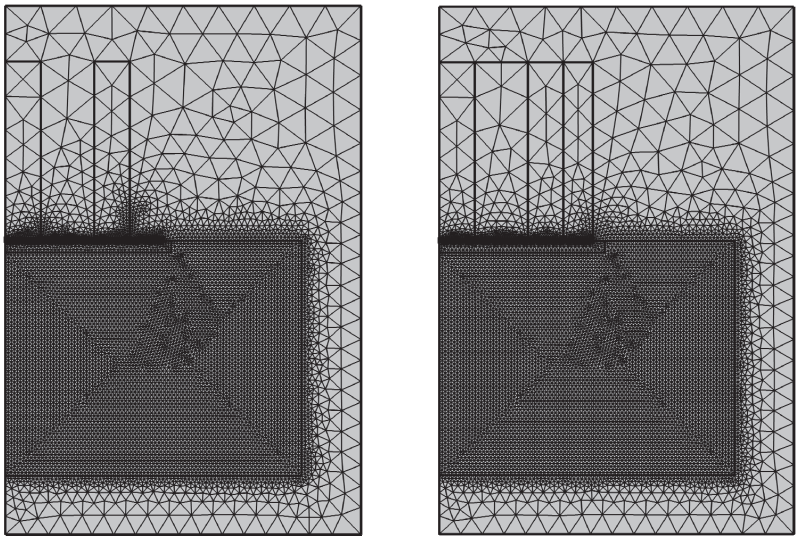
Table 1. Parameters of EMAT magnet used in this paper.

Magnet Number	Magnet Parameters	Symbol	Value (mm)
1	Diameter	D_1	12
	Height	H	30
2	Inner diameter	D_2	32
	Outer diameter	D_3	44
	Height	H	30
3	Inner diameter	D_1	12
	Outer diameter	D_2	32
	Height	H	30
4	Inner diameter	D_1	44
	Outer diameter	D_2	52
	Height	H	30

The magnetic flux density distributions in the ring-type EMAT and the flux-concentrating EMAT are shown in Figure 5, and Figure 6 shows the partially enlarged views of the bottom of the magnets shown in Figure 5. The white arrow in Figure 6 shows the magnetic field direction of the permanent magnet. The magnetic field direction on the specimen surface under the No. 3 and No. 4 magnets is primarily horizontal, while the magnetic field direction under the No. 1 and No. 2 permanent magnets is primarily vertical.

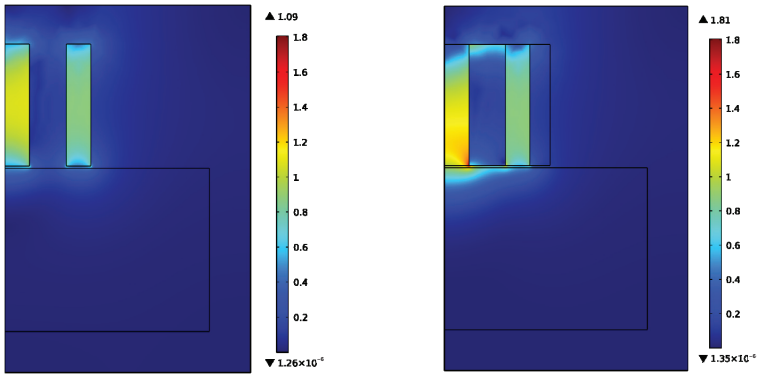
It can be seen from Figure 7a,b that the axial component of the magnetic flux density of the ring-type EMAT and the flux-concentrating EMAT is mainly distributed in the area below the central cylinder magnet and No. 2 circular magnet. At these locations, the vertical flux density is greater than the horizontal flux density. The shear wave signals received by the coils in these locations have a higher amplitude. However, the vertical axial magnetic flux density of the flux-concentrating EMAT is higher than that of the ring-type EMAT so that the flux-concentrating EMAT can generate stronger shear waves. It can be seen from Figure 7c,d that the radial component of the magnetic flux density, which has a local maximum near the outer edge of the ring magnet, is mainly distributed in the area between the central cylinder magnet and No. 2 circular magnet. The horizontal flux

density is greater than the vertical flux density in this area, where the longitudinal wave signals received by the coils have a higher amplitude. The horizontal radial magnetic flux density provided by the flux-concentrating EMAT occupies a larger part of the specimen surface, so that the flux-concentrating EMAT can generate stronger longitudinal waves. The comparison diagram of horizontal radial magnetic flux density B_{sr} and perpendicular axial magnetic flux density B_{sz} on the specimen surface is shown in Figure 8. As can be seen from the figure, the horizontal radial magnetic flux density of the flux-concentrating EMAT in the area from 6 mm to 16 mm from the center point on the specimen surface is significantly higher than that of the ring-type EMAT, which is about twice that of the ring-type EMAT. In addition, a strong vertical magnetic field is distributed in the circular region with a diameter of 12 mm in the center of the specimen surface to generate shear waves, which is about twice the vertical magnetic flux density of the ring-type EMAT. Therefore, the coils can be designed in different positions to generate shear waves or longitudinal waves.



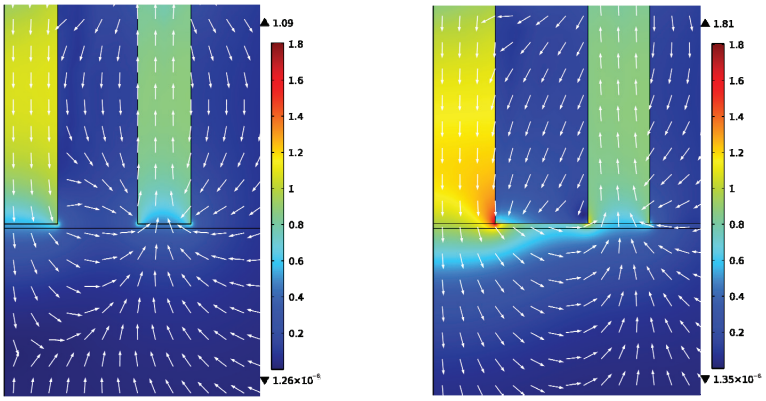
(**a**) (**b**)

Figure 4. Meshes of (**a**) the ring-type longitudinal-wave EMAT model and (**b**) the flux-concentrating longitudinal-wave EMAT model.



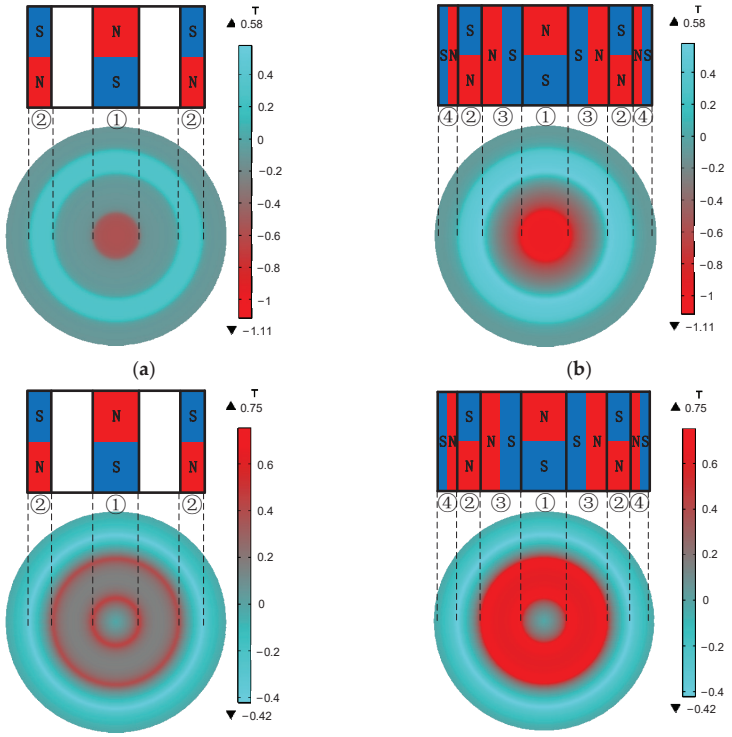
(**a**) (**b**)

Figure 5. The magnetic flux density distribution in (**a**) the ring-type EMAT and (**b**) the flux-concentrating EMAT.



(**a**) (**b**)

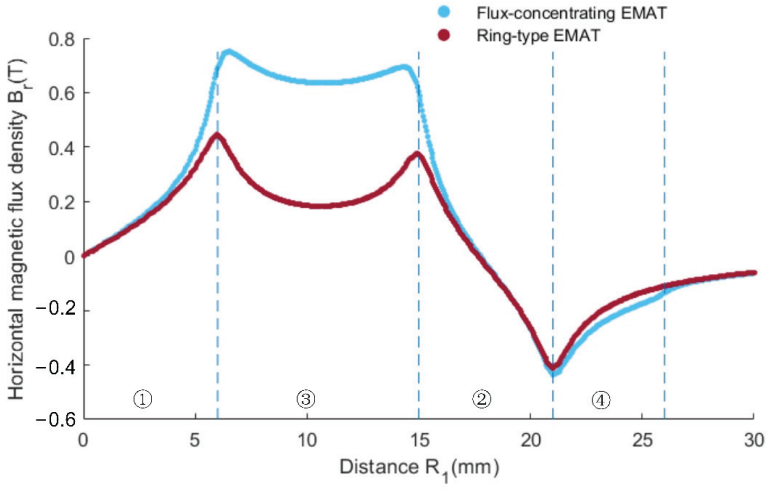
Figure 6. Local enlargement maps of (**a**) the ring-type EMAT and (**b**) the flux-concentrating EMAT.



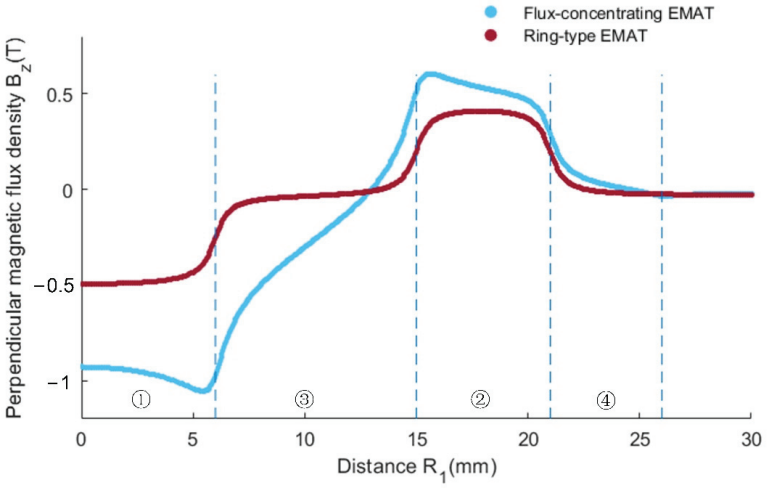
(**a**) (**b**)

(**c**) (**d**)

Figure 7. Magnet density distribution in a circular area with a radius of 30 mm from the center of the specimen surface. (**a**) Axial flux density component and (**c**) radial flux density component of the magnet in ring-type EMAT, (**b**) Axial flux density component and (**d**) radial flux density component of the magnet in flux-concentrating EMAT. The No. 1 magnet is a cylinder magnet. The No. 2, No. 3 and No.4 magnets are circular magnets of different sizes. The magnetizing directions of the magnets are shown.



(**a**)

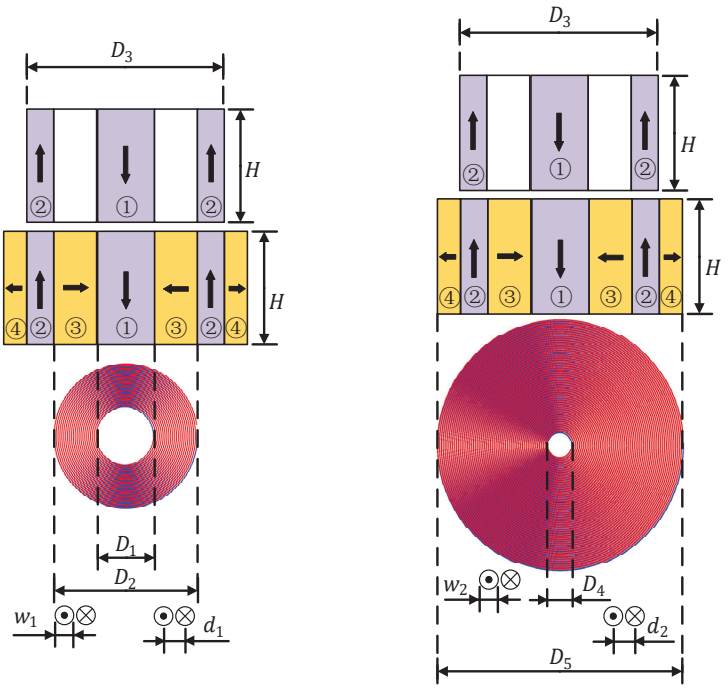


(**b**)

Figure 8. Magnetic flux profile of the ring-type EMAT and the flux-concentrating EMAT on specimen surface: (**a**) horizontal magnetic flux density, (**b**) perpendicular magnetic flux density.

3.2. Coil Design and Simulation of EMAT Signals

According to Figures 7 and 8, we designed two kinds of coils with different sizes, shown in Figure 9. One is the L-mode coil that generates pure longitudinal waves, and the other is the dual-mode coil that can excite both longitudinal waves and shear waves. The coil size parameters are shown in Table 2. A high-frequency pulse current of 1 MHz is passed through the coil. The specimen below the coil induced eddy currents on the surface. Figure 10 shows the eddy current distribution on the specimen surface. The L-mode coil shown in Figure 9a, which can excite and receive relatively pure longitudinal waves, is in the radial magnetic flux density region. The diameter of the dual-mode coil, which is capable of exciting and receiving both longitudinal waves and shear waves, is equivalent to the overall diameter of the magnet.

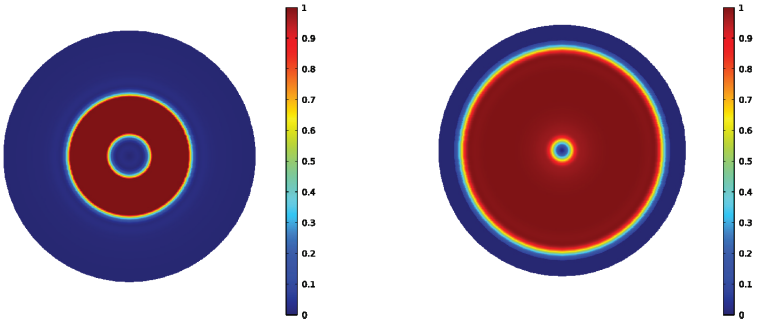


(**a**) (**b**)

Figure 9. Top view of (**a**) L-mode coil and (**b**) dual-mode coil. The No. 1 magnet is a cylinder magnet. The No. 2, No. 3 and No.4 magnets are circular magnets of different sizes. The magnetizing directions of the magnets are shown.

Table 2. Size parameters of L-mode coil and dual-mode coil.

Coil Type	Size Parameters	Symbol	Value (mm)
L-mode coil	Inner diameter	D_1	12
	Outer diameter	D_2	32
	Wire width	w_1	0.2
	Wire spacing	d_1	0.2
Dual-mode coil	Inner diameter	D_4	5
	Outer diameter	D_5	52
	Wire width	w_2	0.2
	Wire spacing	d_2	0.4

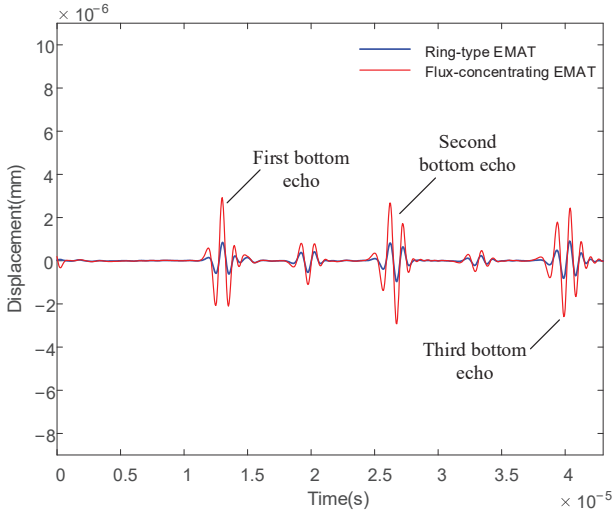


(**a**) (**b**)

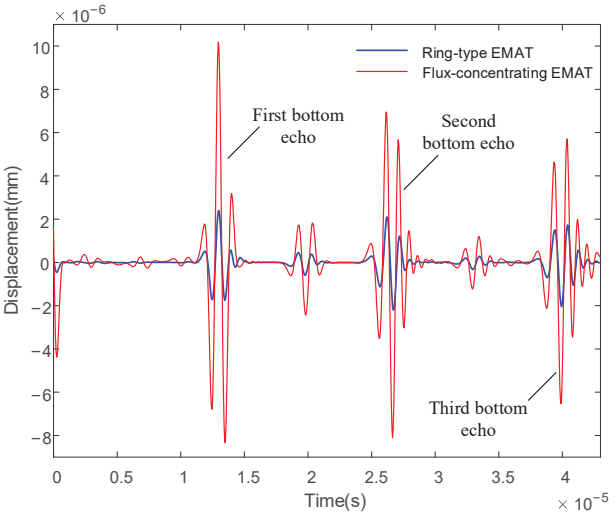
Figure 10. Eddy current distribution in a circular region with a radius of 30 mm from the center of the specimen surface: (**a**) L-mode coil, (**b**) dual-mode coil.

Figure 11 shows the simulated average displacement of ring-type EMAT and flux-concentrating EMAT on the specimen surface as it changes over time. From the figure, we can see the simulated first bottom echo at 13.1 μs, the simulated second bottom echo at 26.2 μs and the simulated third bottom echo at 39.3 μs. The time interval between the two adjacent echoes is about 13 μs. The specimen in the model is aluminum with a thickness of 40 mm, in which the longitudinal wave propagation speed is 6100 m/s, and the shear wave propagation speed is 3050 m/s. Therefore, the theoretical peak time of the first longitudinal-wave echo is 13.1 μs. The theoretical peak time of the first shear wave echo is 26.2 μs and the theoretical peak time of the third bottom echo is 39.3 μs. The peak position of the simulated waveform is completely consistent with the theory. In Figure 11a,b, the displacement amplitudes of the flux-concentrating EMAT are all significantly higher than those of the ring-type EMAT no matter which coil is used. The first and third bottom echoes are mainly longitudinal waves, while the second bottom echo contains shear and longitudinal waves. Therefore, the flux-concentrating EMAT can generate stronger longitudinal waves than the ring-type EMAT. In addition, it is found from the simulation that there is an echo signal between the two adjacent echoes, which is generated by the mode conversion of longitudinal waves reflected on the surface or underside of the specimen.

To better show the mode conversion mechanism, the snapshots of the displacement with the same color scale are shown in Figure 12. Figure 12a–d refer to the flux-concentrating EMAT and ring-type EMAT working in dual and longitudinal mode, respectively. It is known that the velocity of a longitudinal wave is faster than that of a shear wave and the wavelength of the longitudinal wave is longer. Thus, according to the relative position and wavelength, all of the important waves related to Figure 11 in the snapshots are annotated. The middle echo, which appears after the first longitudinal wave reaches the bottom at 7 μs, is the shear wave according to the snapshots and propagating time.



(**a**)



(**b**)

Figure 11. Average displacements of ring-type EMAT and flux-concentrating EMAT on the specimen surface (**a**) using dual-mode coil and (**b**) using L-mode coil.

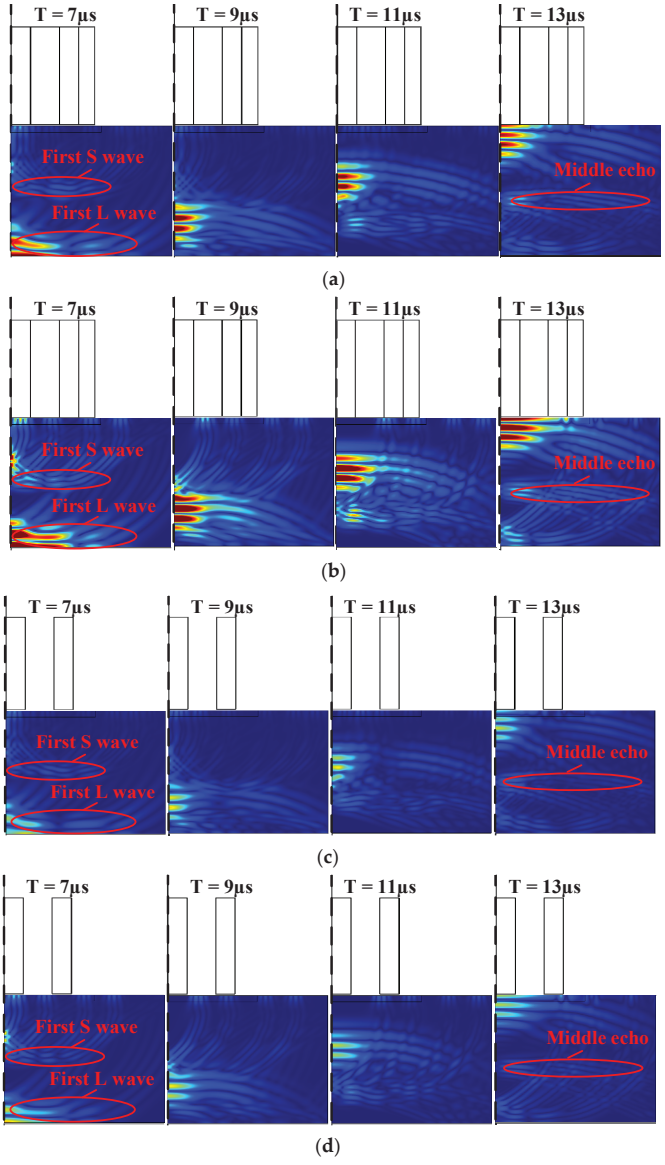


Figure 12. Total displacement propagation snapshots for (**a**) the flux-concentrating EMAT with dual-mode coil, (**b**) the flux-concentrating EMAT with L-mode coil, (**c**) the ring-type EMAT with dual-mode coil and (**d**) the ring-type EMAT with L-mode coil. The symmetry axis is the left dotted line.

4. Experiment

To verify the optimized performance of the flux-concentrating EMAT, a comparative experiment was conducted between the ring-type EMAT and the flux-concentrating EMAT. The schematic diagram of the experimental setup is shown in Figure 13. The magnets, combined with the spiral coil to simultaneously transmit and receive signals, were placed on top of the aluminum plate. The RPR-4000 pulse generator/receiver was chosen to excite a high-frequency pulse through the spiral coil in the experiment, whose duplex protected the preamplifier from the influence of the excitation pulse. The relevant parameters of the

permanent magnet of the ring-type EMAT and the flux-concentrating EMAT are shown in Table 1, and the fabrication of the magnet and coil are shown in Figures 14 and 15, respectively. What is worth mentioning is that the angular sector magnets of the flux-concentrating EMAT are hard to assemble because of the repulsive forces between adjacent magnets. The magnets need to be glued together, and the outside is secured with a metal ring. In addition, the center frequency of the excitation pulse was set as 1 MHz and the peak voltage of the excitation signal was up to 730 V. The test specimen was an aluminum plate with a thickness of 40 mm.

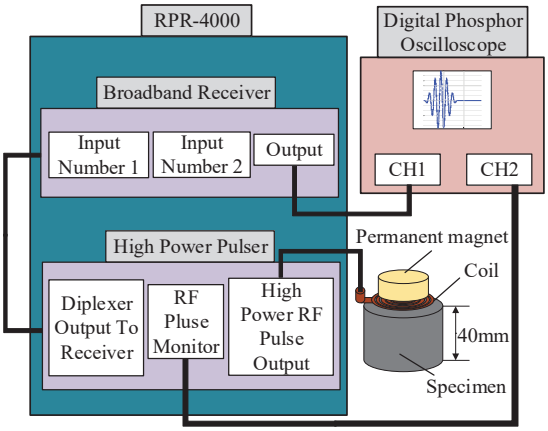
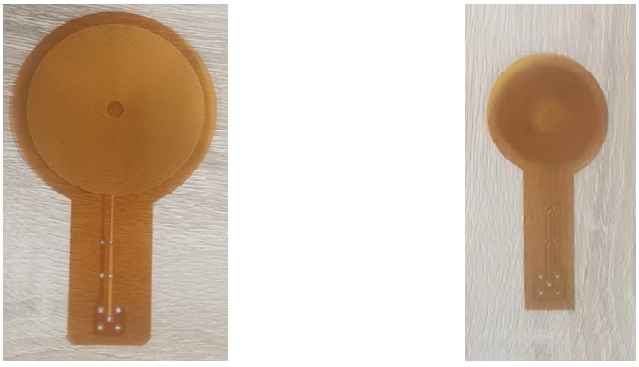


Figure 13. Experimental setup for measurement with EMATs.

(**a**) (**b**)

Figure 14. The fabrication of permanent magnets of (**a**) the ring-type EMAT and (**b**) the flux-concentrating EMAT.

The signal diagram, shown in Figure 16, of the dual-mode coil, which is shown in Figure 15a, demonstrates how a 40 mm thick aluminum plate was obtained. It can be seen that the time interval between the two bottom echoes measured by the ring-type EMAT and the flux-concentrating EMAT is about 13 μs, but the amplitude of the bottom echoes significantly increases. Compared with the amplitude of the signal received by the ring-type EMAT, the first bottom echo is increased by 276%, and the second bottom echo is increased by 391%.



(**a**) (**b**)

Figure 15. The fabrication of (**a**) dual-mode coil and (**b**) L-mode coil.

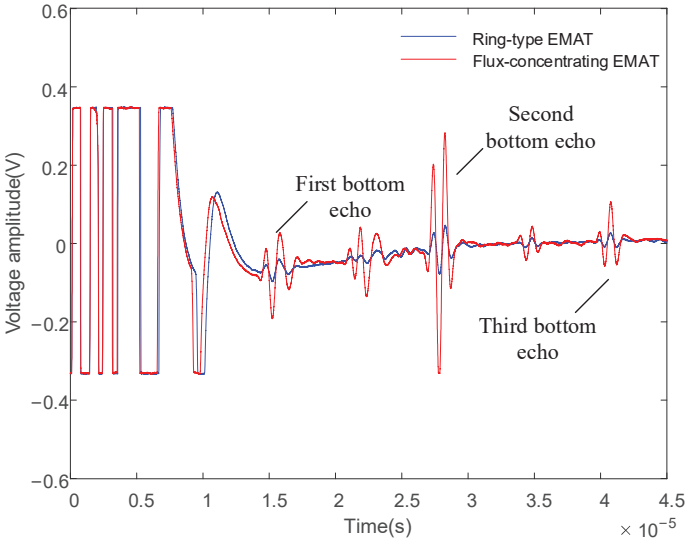


Figure 16. Signal diagram comparison of the ring-type EMAT and the flux-concentrating EMAT using the dual-mode coil.

The experimental results show that the flux-concentrating EMAT with the radial-flux-focusing magnet can simultaneously generate stronger longitudinal waves and shear waves. Figure 16 shows that the back echo signal of the flux-concentrating EMAT is easy to identify and has a high SNR. The first bottom echo is the longitudinal wave; the second bottom echo is a mixture of longitudinal waves and shear waves; and the third bottom echo is the longitudinal wave. In the signal diagram, we found a wave peak between two adjacent bottom echoes, which was concluded from the simulation, as the shear wave signal was generated by the wave mode transformation when the longitudinal wave reflected through the surface or bottom.

To suppress the generation of shear waves and obtain pure longitudinal waves, the L-mode coil shown in Figure 15b is used to measure the aluminum plate with a thickness of 40 mm. The comparison diagram of measurement signals of the ring-type EMAT and the flux-concentrating EMAT is shown in Figure 17.

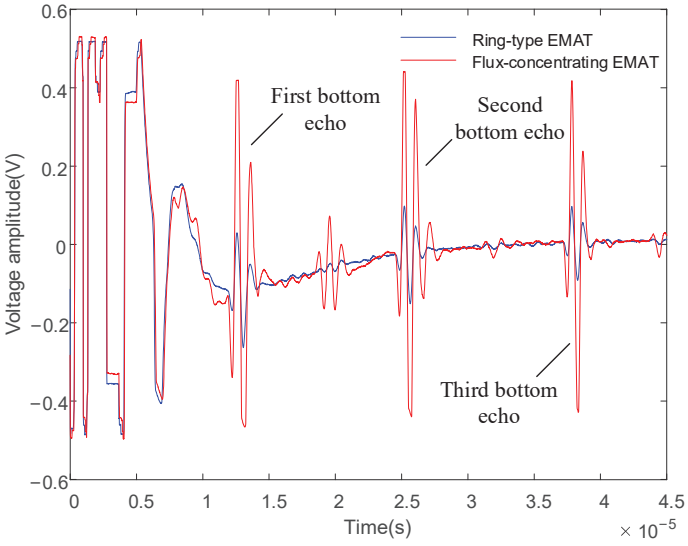


Figure 17. Signal diagram comparison of the ring-type EMAT and the flux-concentrating EMAT using the L-mode coil.

From the longitudinal wave signal diagram shown in Figure 17, the L-mode coil can suppress the generation of shear waves and stimulate purer longitudinal waves with better signal recognition. Compared with the ring-type EMAT, the voltage amplitude of the first bottom echo of the received signal increases by 202%.

5. Conclusions

In this paper, a flux-concentrating EMAT composed of a radial-flux-focusing magnet is proposed. Based on the magnet of the ring-type EMAT, the radial-flux-focusing magnet is filled with a radial-magnetized annular magnet, which can provide a strong horizontal-radial magnetic field and a vertical axial magnetic field on the premise that the overall volume of the magnet does not change significantly. The ring-type EMAT with a similar structure and no magnetization is used for comparison.

Two-dimensional axisymmetric solid simulation models of the flux-concentrating EMAT and the ring-type EMAT operating on the non-ferromagnetic aluminum specimen were established to simulate the distribution of static magnetic flux. According to the simulation results, the flux-concentrating EMAT can provide a stronger horizontal radial magnetic field below the region of the No. 3 magnet and a vertical axial magnetic field in the central area, compared to the ring-type EMAT.

Based on theoretical and finite element analysis, two kinds of spiral coils with different sizes are designed: one is a longitudinal-wave EMAT, and the other is a dual-mode EMAT that can excite both longitudinal waves and shear waves. According to the simulation results, the flux-concentrating EMAT can improve the amplitude of the proposed wave mode significantly.

Finally, the ring-type EMAT and the flux-concentrating EMAT were used to measure the thickness of the 40 mm aluminum plate, and their received signals were compared. When the dual-mode coil is used, the voltage amplitude of the signal received by the flux-concentrating EMAT increases by 276% compared with the ring-type EMAT. When the L-mode coil is used, the voltage amplitude increases by 202%. It is verified that under the same pulse excitation, the received signal of the flux-concentrating EMAT is more accessible to identify than that of the ring-type EMAT. The voltage amplitude increases significantly, and the SNR is higher.

The flux-concentrating EMAT is not easy to assemble but has a higher energy conversion efficiency than the ring-type EMAT. Although the volume of the former is 63.7 cm^3 and the volume of the latter is only 41.6 cm^3, provided that the same static magnetic field strength is needed, the flux-concentrating structure could be combined with more economical magnets of lower energy products, and different coils according to the distribution of magnet direction could be used to achieve dual-mode excitation or pure longitudinal wave excitation according to specific needs. For pure longitudinal wave excitation, the shear wave signal is suppressed while the longitudinal wave signal is strengthened in resonant ultrasound spectroscopy, thus the interference of shear waves to longitudinal waves is greatly reduced. The designed EMAT could be applied to increase the detection speed in the field of thickness measurement. For dual-mode excitation, the flux-concentrating EMAT could have similar applications in elastic-constant extraction, which is consistent with other similar studies [20,24]. Moreover, the cylindrical-design EMAT are fit for measuring the axial stress of non-ferromagnetic cylinder-like aluminum bolts, and we hope to conduct further research on this topic in the future.

Author Contributions: Conceptualization, X.Z.; Data curation, C.L.; Formal analysis, Q.W.; Investigation, B.L.; Methodology, W.L.; Project administration, X.S.; Supervision, J.T.; Writing—original draft, X.Z.; Writing—review and editing, S.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China, grant number 51807052, 51707058 and Hubei Natural Science Foundation Innovation Group Project, grant number 2019CFA021.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Sun, H.; Huang, S.; Wang, Q.; Wang, S.; Zhao, W. Orthogonal Optimal Design Method for Point-Focusing EMAT Considering Focal Area Dimensions. *Sens. Actuators A Phys.* **2020**, *312*, 112109. [CrossRef]
2. Hirao, M.; Ogi, H. Introduction. In *Electromagnetic Acoustic Transducers: Noncontacting Ultrasonic Measurements Using EMATs*; Hirao, M., Ogi, H., Eds.; Springer: Tokyo, Japan, 2017; pp. 1–11.
3. Pei, C.; Zhao, S.; Xiao, P.; Chen, Z. A modified meander-line-coil EMAT design for signal amplitude enhancement. *Sens. Actuators A Phys.* **2016**, *247*, 539–546. [CrossRef]
4. Baskaran, G.; Balasubramaniam, K.; Lakshmana Rao, C. Shear-wave time of flight diffraction (S-TOFD) technique. *NDT E Int.* **2006**, *39*, 458–467. [CrossRef]
5. Parra-Raad, J.; Khalili, P.; Cegla, F. Shear waves with orthogonal polarisations for thickness measurement and crack detection using EMATs. *NDT E Int.* **2020**, *111*, 102212. [CrossRef]
6. Park, J.; Lee, J.; Min, J.; Cho, Y. Defects Inspection in Wires by Nonlinear Ultrasonic-Guided Wave Generated by Electromagnetic Sensors. *Appl. Sci.* **2020**, *10*, 4479. [CrossRef]
7. Zhang, K.; Yi, P.; Li, Y.; Hui, B.; Zhang, X. A New Method to Evaluate Surface Defects with an Electromagnetic Acoustic Transducer. *Sensors* **2015**, *15*, 17420–17432. [CrossRef]
8. Sun, H.; Urayama, R.; Uchimoto, T.; Takagi, T.; Hashimoto, M. Small electromagnetic acoustic transducer with an enhanced unique magnet configuration. *NDT E Int.* **2020**, *110*, 102205. [CrossRef]
9. Tagawa, A.; Ueda, M.; Yamashita, T. Development of the ISI Device for Fast Breeder Reactor MONJU Reactor Vessel. *J. Power Energy Syst.* **2006**, *1*, 3–12. [CrossRef]
10. Zhang, Y.; Liu, W.; Li, N.; Qian, Z.; Wang, B.; Liu, D.; Li, X. Design of a new type of omnidirectional shear-horizontal EMAT by the use of half-ring magnets and PCB technology. *Ultrasonics* **2021**, *115*, 106465. [CrossRef]
11. Shankar, S.; Balasubramaniam, K. Characterising the beam formation of SH waves using double-row Staggered Halbach EMAT configurations. *NDT E Int.* **2021**, *121*, 102465. [CrossRef]
12. Shimizu, H.; Bahr, A.J. Improved Design for Non-Contacting Electromagnetic-Acoustic Transducers. In Proceedings of the 1977 Ultrasonics Symposium, Menlo Park, CA, USA, 26–28 October 1977; pp. 89–93.

13. Zhai, G.; Liang, B.; Li, X.; Ge, Y.; Wang, S. High-temperature EMAT with double-coil configuration generates shear and longitudinal wave modes in paramagnetic steel. *NDT E Int.* **2022**, *125*, 102572. [CrossRef]
14. Chaki, S.; Corneloup, G.; Lillamand, I.; Walaszek, H. Combination of Longitudinal and Transverse Ultrasonic Waves for In Situ Control of the Tightening of Bolts. *J. Press. Vessel. Technol.* **2006**, *129*, 383–390. [CrossRef]
15. Chen, P.; He, X.; Wang, X. Ultrasonic Measurement of Axial Stress Using High-Frequency Cylindrical Guided Wave. *IEEE Sens. J.* **2020**, *21*, 6691–6697. [CrossRef]
16. Chen, P.; He, X.; Song, W. Parameter Recognition of Mode-Converted Wave in Single-Source Ultrasound Using Gabor Transform for Bolt Axial Stress Evaluation. *J. Sens.* **2020**, *2020*, 8883845. [CrossRef]
17. Liu, H.; Liu, T.; Li, Y.; Liu, Y.; Zhang, X.; Wang, Y.; Gao, S. Uniaxial stress in-situ measurement using EMAT shear and longitudinal waves: Transducer design and experiments. *Appl. Acoust.* **2021**, *175*, 107781. [CrossRef]
18. Ding, X.; Wu, X.; Wang, Y. Bolt axial stress measurement based on a mode-converted ultrasound method using an electromagnetic acoustic transducer. *Ultrasonics* **2014**, *54*, 914–920. [CrossRef]
19. Hirao, M.; Ogi, H. Available EMATs. In *Electromagnetic Acoustic Transducers: Noncontacting Ultrasonic Measurements Using EMATs*; Hirao, M., Ogi, H., Eds.; Springer: Tokyo, Japan, 2017; pp. 39–67.
20. Cunha, M.P.D.; Jordan, J.W. Improved longitudinal EMAT transducer for elastic constant extraction. In Proceedings of the 2005 IEEE International Frequency Control Symposium and Exposition, Vancouver, BC, Canada, 29–31 August 2005; pp. 426–432.
21. Wu, X.; Ding, X.; Wang, L. An improved longitudinal wave EMAT based on the shielding effect. *Int. J. Appl. Electromagn. Mech.* **2014**, *45*, 227–233. [CrossRef]
22. Zhang, X.; Feng, S.; Tu, J.; Song, X. An improved design of shear horizontal guided wave electromagnetic acoustic transducer. *Insight-Non-Destr. Test. Cond. Monit.* **2020**, *62*, 494–497. [CrossRef]
23. Xu, Y.; Tagawa, A.; Ueda, M.; Yamashita, T.; Ohtsuka, Y.; Osafune, K.; Nishikawa, M. A New Structure of SH Wave Electromagnetic Acoustic Transducer (EMAT). In *Acoustical Imaging*; Springer: Dordrecht, The Netherlands, 2004; pp. 175–183.
24. Li, Y.; Liu, T.; Liu, Y.; Liu, H.; Wang, Y. Measurement of Elastic Constants Using Halbach Array Enhanced EMAT. In Proceedings of the 2019 IEEE International Ultrasonics Symposium (IUS), Glasgow, UK, 6–9 October 2019.
25. Jia, X.; Ouyang, Q.; Zhang, X. An Improved Design of the Spiral-Coil EMAT for Enhancing the Signal Amplitude. *Sensors* **2017**, *17*, 1106. [CrossRef]
26. Jian, X.; Dixon, S.; Grattan, K.T.V.; Edwards, R.S. A model for pulsed Rayleigh wave and optimal EMAT design. *Sens. Actuators A Phys.* **2006**, *128*, 296–304. [CrossRef]
27. Ren, W.; He, J.; Dixon, S.; Xu, K. Enhancement of EMAT's efficiency by using silicon steel laminations back-plate. *Sens. Actuators A Phys.* **2018**, *274*, 189–198. [CrossRef]

applied sciences

MDPI

Article

Percussion-Based Pipeline Ponding Detection Using a Convolutional Neural Network

Dan Yang [1,2,*], Mengzhou Xiong [1,3], Tao Wang [1] and Guangtao Lu [1,2]

[1] Key Laboratory for Metallurgical Equipment and Control of Ministry of Education, Wuhan University of Science and Technology, Wuhan 430081, China; wkdxmz195@wust.edu.cn (M.X.); wangtao77@wust.edu.cn (T.W.); luguangtao@wust.edu.cn (G.L.)
[2] Hubei Key Laboratory of Mechanical Transmission and Manufacturing Engineering, Wuhan University of Science and Technology, Wuhan 430081, China
[3] Precision Manufacturing Institute, Wuhan University of Science and Technology, Wuhan 430081, China
* Correspondence: yangdan@wust.edu.cn

Abstract: Pipeline transportation is the main method for long-distance gas transportation; however, ponding in the pipeline can affect transportation efficiency and even cause corrosion to the pipeline in some cases. A non-destructive method to detect pipeline ponding using percussion acoustic signals and a convolution neural network (CNN) is proposed in this paper. During the process of detection, a constant energy spring impact hammer is used to apply an impact on the pipeline, and the percussive acoustic signals are collected. A Mel spectrogram is used to extract the acoustic feature of the percussive acoustic signal with different ponding volumes in the pipeline. The Mel spectrogram is transferred to the input layer of the CNN and the convolutional kernel matrix of the CNN realizes the recognition of pipeline ponding volume. The recognition results show that the CNN can identify the amount of pipeline ponding with the percussive acoustic signals, which use the Mel spectrogram as the acoustic feature. Compared with the support vector machine (SVM) model and the decision tree model, the CNN model has better recognition performance. Therefore, the percussion-based pipeline ponding detection using the convolutional neural network method proposed in this paper has high application potential.

Keywords: pipeline ponding; percussion detection method; Mel spectrogram; convolutional neural network (CNN)

Citation: Yang, D.; Xiong, M.; Wang, T.; Lu, G. Percussion-Based Pipeline Ponding Detection Using a Convolutional Neural Network. *Appl. Sci.* **2022**, *12*, 2127. https://doi.org/10.3390/app12042127

Academic Editor: José A.F.O. Correia

Received: 20 January 2022
Accepted: 16 February 2022
Published: 18 February 2022

1. Introduction

As a main method of oil and gas transportation, pipelines play an important role in transporting supplies [1,2]. During their long-term service life, various types of pipeline damages are related to pipeline ponding; corrosion, perforation, and leakage are not uncommon, and they usually bring about serious safety hazards to pipeline transportation [3]. Therefore, to ensure the safe and stable operation of pipelines, pipeline ponding detection has become more important and urgent.

In pipeline ponding detection, changes in the ponding volume will cause changes in the structural characteristics of the pipeline system composed of pipeline and ponding. Therefore, some developed methods for the monitoring of structural characteristics may provide an approach as the reference for pipeline ponding detection. In recent decades, several common methods for pipeline structure characteristic detection have been introduced, including the CCTV (closed-circuit television) inspection method [4], the ultrasonic testing method [5] and the radiography method [6]. The CCTV inspection method presents very rich internal information of the pipeline in the form of photos or videos [7] by a robotic system with a camera [8]. However, the CCTV method is greatly affected by environmental factors, and its detection accuracy of the pipeline evaluation depends largely on the quality of the hardware system and the experience of the inspectors [9]. The ultrasonic

testing method can estimate the health state by analyzing reflection waves [10,11] in the pipelines. It is sensitive to changes in structural state and can be related to several structural characteristics [12]. However, the signals collected by the ultrasonic method are usually accompanied with noise, and effective noise reduction methods are required to obtain useful information [13,14]. The radiography method detects the pipeline by evaluating the attenuation of the rays [15] which pass through the pipeline. This method can be used for pipelines with complex geometric shapes [16]. However, its detection accuracy decreases when it is employed for vertical angle defect detection [17], and the rays are harmful to human health [18]. Therefore this method's practical application is very limited.

Compared with the aforementioned detection approaches, the percussive detection method [19–21] has the characteristics of deep detection and fast transmission speeds, and is user-friendly [22]. It is used to determine the pipeline structure characteristic by the sounds generated through impact on the pipeline under test [23]. Traditional percussive detection method still requires engineering experience, which can be subjective and inefficient [24]. This is solved by using the powerful computing power of computers or the automatic prediction and classification properties of machine learning. Furui Wang et al. proposed a new percussion-based method using analytical modeling and numerical simulation, whereby a percussion-induced sound pressure level (SPL) could be obtained via the acoustic radiation mode approach. The corresponding numerical simulation was developed with a focus on the acoustic–structure coupling, and the acoustic boundary conditions were satisfied through a perfectly matched layer (PML) [25]. Liqiong Zheng et al. used Mel-frequency cepstral coefficients (MFCCs) as the features of percussion-induced acoustics, and support vector machine (SVM)-based machine learning was utilized to classify results [26]. Dongdong Chen et al. used power spectrum density (PSD) to process percussive sound, and a decision tree machine (DTM) learning approach was used to classify results [27].

CNN, one of the representative algorithms of deep learning, which automatically predicts and classifies the data [28], can overcome the drawbacks of percussive detection methods that requires engineering experience, and can therefore obtain superior results in visual classification tasks [29]. In the classification of audio data, as CNN cannot process sound directly [30], the sound of digital signals is often converted into spectrogram images [31] by a Short-time Fourier transform (STFT) or a wavelet transform. In particular, the STFT is a low-complexity time–frequency method capable of analyzing non-stationary signals which has a low computational burden [32]. However, the dimension of the spectrogram after STFT is relatively high, resulting in a large amount of subsequent CNN calculation, which increases the complexity of CNN learning. Furthermore, a nonlinear transformation can be applied to the frequency axis after the STFT process, to obtain a Mel spectrogram with lower dimensions, by compressing the frequency range [33]. This makes it easier for the CNN to extract and process specific features.

This paper proposes a non-destructive detection method for pipeline ponding by referring to a pipeline structure characteristic detection method which combines the percussive detection method and a CNN. During detection, a constant energy spring impact hammer is first used to impact the pipeline under different ponding volumes to generate sound, and the collected acoustic signals are converted into the Mel spectrogram. Then, the CNN is used to perform a two-dimensional convolution operation on the Mel spectrogram and the convolution kernel matrix, and realize the identification of pipelines with different ponding volumes according to the output matrix. The rest of this paper is organized as follows: Section 2 introduces the principle of percussion-based pipeline ponding detection using CNN and network model evaluation metrics; Section 3 introduces the experimental equipment and experimental procedures; Section 4 presents the experimental results and comparative analysis with other recognition models; Section 5 summarizes the advantages and disadvantages of the method proposed in this paper.

2. Materials and Methods

2.1. Working Principle

The flowchart of the proposed method is presented in Figure 1. In general, it consists of three steps: percussion signal acquisition, signal processing, and automatic pattern recognition based on the CNN. In the first step, the acoustic signal generated by the percussion on the pipeline with different ponding volumes was recorded by a microphone, where six ponding volumes were considered. The signal processing step included three consecutive processing stages: preprocessing, STFT method, and Mel filtering. Pre-processing was applied to the percussion signal to delete any low-frequency interference components in the sound signal, and to increase the proportion of high-frequency components. Then, using both overlap and a Hamming window, the STFT was used to obtain the time–frequency plane of the current signal. Finally, the Mel filtering was applied to the frequency axis after the STFT to obtain the Mel spectrogram with lower dimensions by compressing the frequency range, which made the CNN less computationally intensive. In the pattern recognition step, a CNN is proposed to classify the ponding volume case in an automatic way. It is worth noting that the time–frequency plane obtained through the Mel spectrogram was treated as an image in order to implement a conventional two-dimensional (2D) CNN. In the 2D CNN design, learning rates, batch sizes, and dataset split ratios were analyzed.

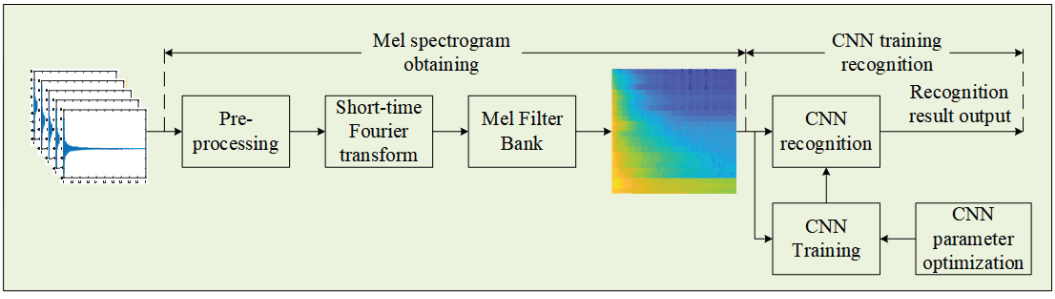


Figure 1. Schematic diagram of the working principles.

2.2. Mel Spectrogram

The Mel spectrogram is obtained with the following procedures:

I: Perform pre-processing of the selected signal including pre-emphasis, framing and windowing;

II: Perform short-time Fourier transform of the pre-processed data;

III: Perform Mel filtering of the data after step II to obtain the Mel spectrogram.

2.3. CNN

The recognition process of the convolutional model can be divided into two parts: CNN training and CNN recognition. In the training process of CNNs, the model parameters and training steps are preset; then, the model parameters are continuously corrected through the data forward propagation process, and error backward propagation process, until the convolutional model meets the requirements. In the CNN recognition, the high-dimensional features extracted by convolution and pooling operations are matched with the trained model to output recognition results.

The structure of the CNN model proposed in this paper is shown in Figure 2. It consists of four nonlinear trainable convolutional layers, four nonlinear fixed convolutional layers (Pooling Layer) and one fully connected layer.

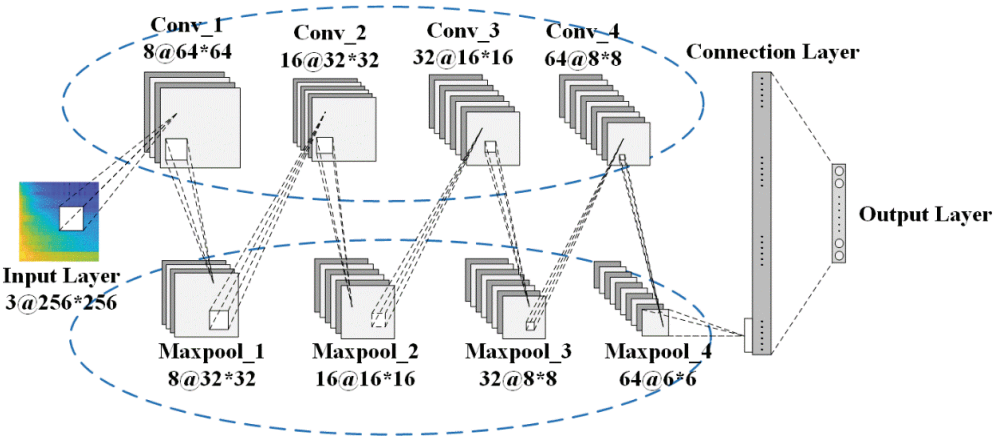


Figure 2. The CNN model.

Among them, the role of the convolutional layer was to perform adaptive feature extraction on the Mel spectrogram, which was achieved by convolutional operations of the convolutional kernel matrix [34]. The operation of the convolutional layer is as follows:

$$C_l = \sum_{x=1}^{m} \sum_{y=1}^{n} \sum_{z=1}^{p} a_{x,y,z} \omega_{x,y,z}^{l} + b^{l}, l = 1, 2, \ldots, q \quad (1)$$

where l is the convolutional kernel number, C_l is the lth layer feature map of CNN, a is the input of convolutional layer, ω is the weight matrix, b is the bias term of convolutional kernel, and x, y, z are the different dimensions of the input data.

Adding a pooling layer after the convolution layer allows downsampling of the input features while preserving the dominant features, which can reduce the model parameters at the same time as suppressing overfitting [35]. The CNN model proposed in this paper uses maximum value pooling, and its expression is:

$$G_l = downsamp(H_l) = \max H_l(v_1, v_2) \quad (2)$$

where H_l is the pooling layer input feature, G_l is the pooling layer output feature, and (v_1, v_2) is the classification element that is pooled for the previous layer.

After the Mel spectrum is propagated through several convolutional and pooling layers alternately, the fully connected layer network is relied upon to classify the extracted features, and its expression is:

$$h_l = f(W_l * h_{l-1} + b_l) \quad (3)$$

where h_{l-1} is the output of the previous network layer, h_l is the output of the current fully connected layer, W_l is the weight, b_l is the bias, and $f(*)$ is the activation function.

2.4. CNN Model Evaluation Metrics

The performance of the final trained CNN model needed to be evaluated by corresponding metrics [36]. Common evaluation metrics for classification tasks are Precision, Recall, and F1-Measure [37,38], which have the following equations:

$$P = \frac{TP}{TP + FP} \quad (4)$$

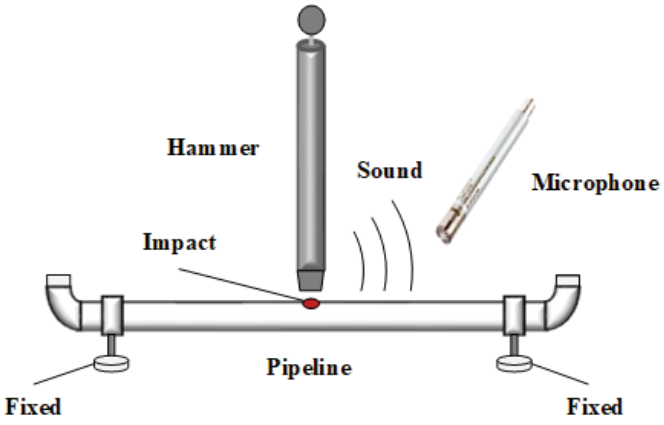
$$R = \frac{TP}{TP + FN} \quad (5)$$

$$F1 = \frac{2 * P * R}{P + R} \tag{6}$$

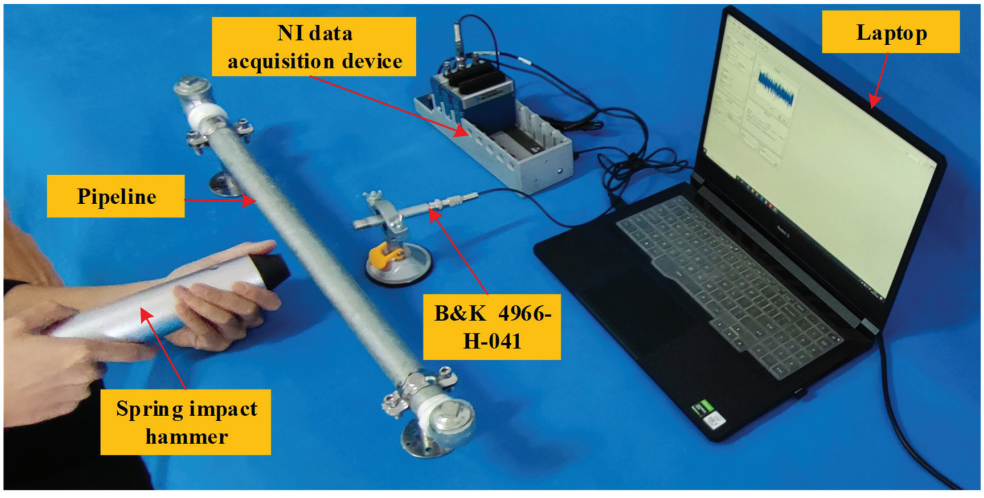
where *TP* indicates a positive sample is correctly identified as a positive sample, *TN* indicates a negative sample is correctly identified as a negative sample, *FP* indicates a false positive sample (which means a negative sample is incorrectly identified as a positive sample), and *FN* indicates a false negative sample (which means a positive sample is incorrectly identified as a negative sample).

3. Experimental Setup and Procedures

As shown in Figure 3, the pipeline was fixed by a holding device, a spring-loaded impact hammer applied an impact on the middle position of the pipeline, and a microphone with a frequency band of 10 Hz~20 kHz was placed about 5 cm away from the impact position to capture the percussive acoustic signal generated by the impact. During the experiments, the sampling rate of the data acquisition device was set to 100 kHz.



(**a**)



(**b**)

Figure 3. (**a**) Schematic of the experimental setup; (**b**) experimental setup.

In the tests, six pipelines specimens with different dimensions were fabricated; the dimensions of these specimens are listed in Table 1.

Table 1. Dimensions of the pipeline specimens.

Pipeline Number	Outer Diameter/mm	Inner Diameter/mm	Length/mm
1#	Φ32	Φ25	60
2#	Φ32	Φ25	100
3#	Φ42	Φ35	60
4#	Φ42	Φ35	100
5#	Φ48	Φ41	60
6#	Φ48	Φ41	100

During the test, to simulate different ponding states of the pipelines, the specimens were filled with different volume percentages of water. There were a total of six experimental cases, which are listed in Table 2. The energy of each impact of the spring-loaded hammer was constant at 1J. Only the selected signal was filtered with a band-pass filter matching the microphone frequency, and 100 experiments were performed for each case.

Table 2. Different experimental cases with different volume percentage of water.

Name	Value					
Case	0	1	2	3	4	5
Water as a percentage of pipeline volume (%)	0	10	20	30	40	50

4. Experimental Results

4.1. Mel-Feature Extraction

The typical percussive sound signals of the pipeline with experimental cases are shown in Figure 4.

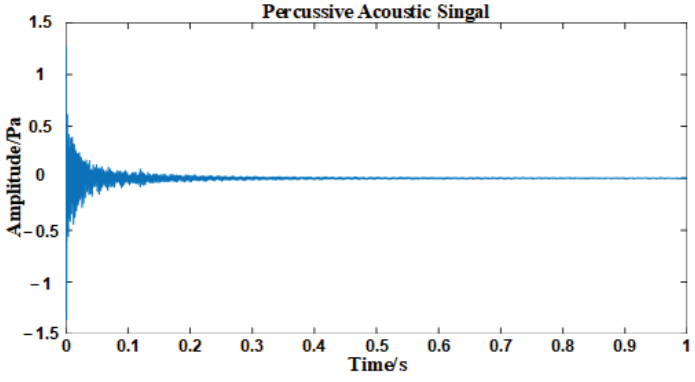


Figure 4. One of the sound signals recorded by the microphone.

The filtered signals were then converted into a Mel spectrogram and the parameters [39,40] were set, as shown in Table 3. The extracted Mel spectrogram features are shown in Figure 5. The results show that the differences in the Mel spectrogram of the six ponding volumes of the 1#pipeline are very small and difficult to distinguish with the naked eye.

4.2. Identification of the Amount of Ponding Volume in a Single Pipeline

Before the CNN is trained, a finer selection of other parameters such as learning rate and batch size can be carried out. The learning rate determines the step size of adjusting weights and error reduction in the training process. Figure 6 shows the obtained results for

different learning rates by considering only one epoch. One epoch is a complete pass over the entire dataset. The results demonstrate that extreme values have a negative impact on accuracy. Therefore, in this work, a learning rate value of 0.01 was used, as it presented a higher accuracy and accelerated the error convergence. Table 4 shows the results of accuracy and computation time obtained using different values of batch size. The batch size determines the size of the subset of the entire dataset used in each training iteration. As indicated in Table 4, a small batch size value generates high accuracy, but results in a high computation time. On the contrary, a high value of batch size reduces the computational time, but the accuracy is negatively affected. In this regard, we chose a batch size of 30 because it provided high accuracy and a suitable computational time. Additionally, SGDM was used as the optimizer and ReLU was used as the activation function.

Table 3. Mel spectrogram Parameters.

Name	Value
Fs/Hz	100,000
Window	Hamming
Window Length	2048
Overlap Length	1024
FFT Length	4096
NumBands	24

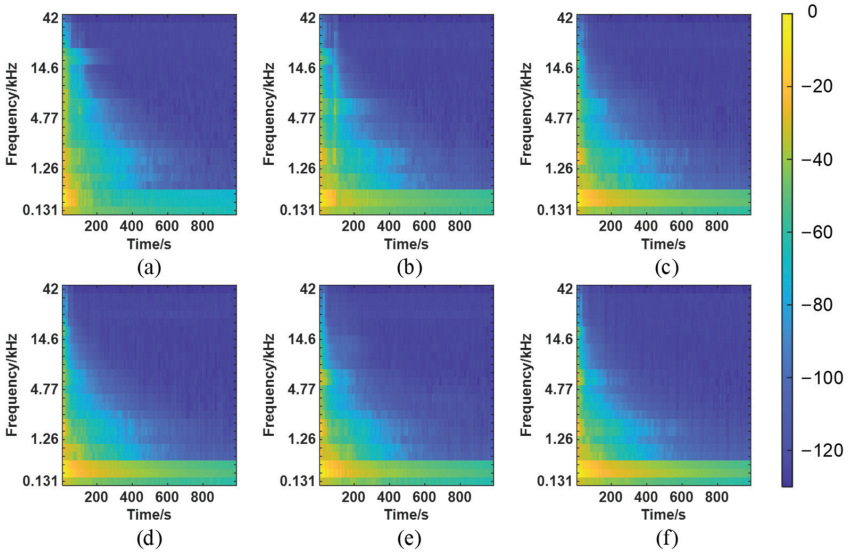


Figure 5. Mel spectrogram of 1#pipeline for 6 cases: (**a**) 0 case; (**b**) 1 case; (**c**) 2 case; (**d**) 3 case; (**e**) 4 case; (**f**) 5 case.

After we selected the above-mentioned parameters, the CNN could be completely trained and validated. However, before using the dataset to train the model, the whole dataset needed to be divided into a training set and a validation set. With the dataset well partitioned, the speed of model applications can be improved. If the partitioning is not good, it can greatly affect the deployment of the model applications. Table 5 shows the results of accuracy and computation time obtained using different ratios of dataset split. This table demonstrates that the CNN model has the highest accuracy and its application speed is optimal when the dataset splitting ratio is 7:3. Therefore, in the training process of the convolutional model, 70 sets of data obtained under each experimental case were randomly selected and converted into the Mel spectrogram, then input into the CNN as the

training set. The remaining 30 sets were input into the trained CNN model as validation sets to complete the recognition of the pipeline ponding volume.

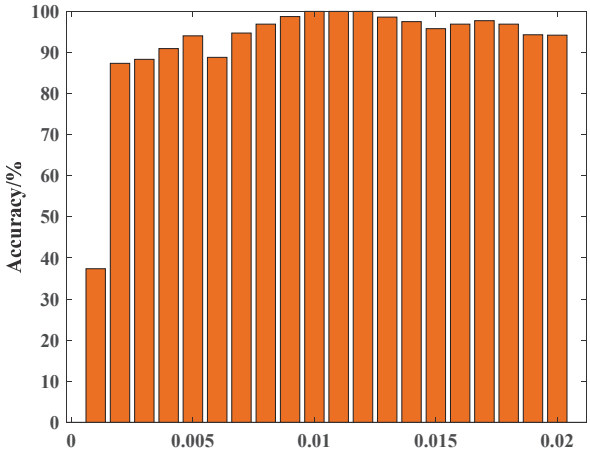


Figure 6. Obtained accuracy for different learning rate values.

Table 4. Results for different batch size values.

Name	Value					
Batch size	5	10	15	20	25	30
Accuracy (%)	98.57	97.14	98.32	98.57	99.32	100
Time/s	529	267	170	131	104	94
Batch size	35	40	45	50	55	60
Accuracy (%)	98.73	93.10	91.67	84.76	86.19	83.24
Time/s	84	72	67	65	59	61
Batch size	65	70	75	80	85	90
Accuracy (%)	81.36	92.38	82.14	87.14	87.62	90.00
Time/s	54	55	48	49	41	42
Batch size	95	100				
Accuracy (%)	87.62	91.43				
Time/s	40	41				

Table 5. Results of different splitting datasets.

Name	Case				
Dataset split ratio	1:1	3:2	7:3	4:1	9:1
Accuracy (%)	98.47	97.83	100	97.50	98.70
Time/s	86	92	81	97	129

The training process of the CNN model for six ponding volume cases in the 1$^{\#}$pipeline is shown in Figure 7, and the recognition results are shown in Table 6.

Figure 7a shows that, with the increase in training times, the accuracy rate increases alternately and its fluctuation is large; after the number of training times reaches 146, the accuracy rate reaches 98.34%. Figure 7b indicates that the value of the loss function decreases continuously with the increase in training times, and finally stabilizes at about 0.086. Table 6 shows the CNN predictions for different case validation sets, and it can be calculated that the accuracies are 96.67%, 100%, 100%, 96.67%, 100% and 96.67%, respectively. The results show that the proposed approach can classify different ponding volume cases with high accuracy.

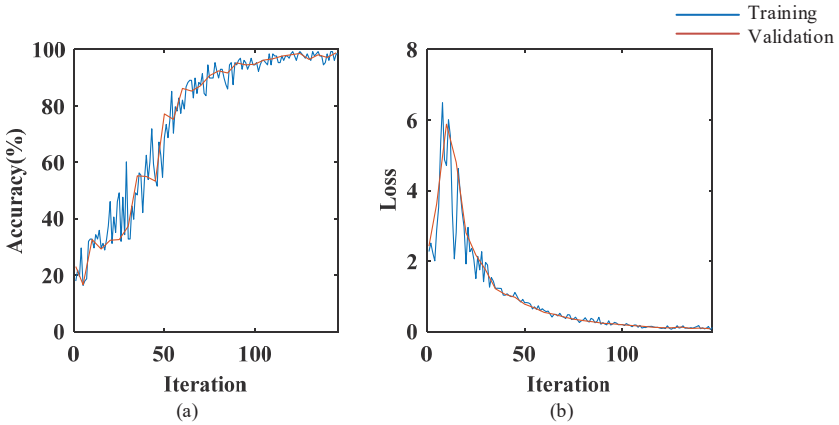


Figure 7. The CNN model training process. (**a**) Accuracy (%), (**b**) Loss.

Table 6. The CNN identification results of the 1#pipeline.

		Target Class					
		0	**1**	**2**	**3**	**4**	**5**
Predicted Class	0	29	0	0	0	0	0
	1	1	30	0	0	0	0
	2	0	0	30	1	0	0
	3	0	0	0	29	0	0
	4	0	0	0	0	30	1
	5	0	0	0	0	0	29
		Total accuracy (%)					98.34

4.3. The CNN Model Evaluation of Ponding Volume in Different Pipelines

Based on the proposed method, the recognition of ponding volume for different pipelines was also performed. The three common evaluation metrics of Precision (P), Recall (R), and F1-Measure (F1) in the classification task were chosen to evaluate the final trained CNN model, as shown in Table 7.

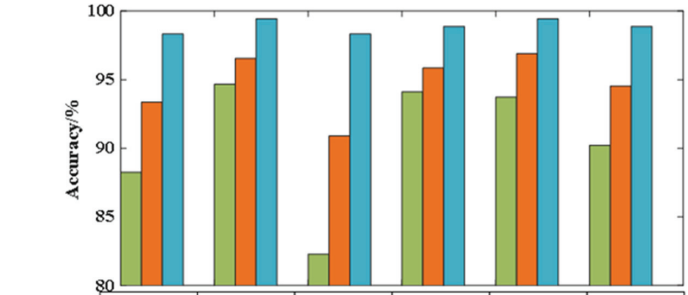
Table 7. Three common evaluation metrics results of six pipeline dimensions.

		1#Pipeline			**2#Pipeline**			**3#Pipeline**		
		R	**P**	**F1**	**R**	**P**	**F1**	**R**	**P**	**F1**
Case	0	96.7	100	98.3	100	96.8	98.4	100	100	100
	1	100	96.8	98.4	100	100	100	100	90.5	95.2
	2	100	96.8	98.4	100	100	100	90	100	94.7
	3	96.7	100	98.3	100	100	100	100	100	100
	4	100	96.8	98.4	96.7	100	98.3	100	100	100
	5	96.7	100	98.3	100	100	100	100	100	100
		4#Pipeline			**5#Pipeline**			**6#Pipeline**		
		R	**P**	**F1**	**R**	**P**	**F1**	**R**	**P**	**F1**
Case	0	100	96.8	98.4	100	100	100	100	100	100
	1	100	100	100	100	100	100	100	100	100
	2	96.7	100	98.3	100	100	100	100	100	100
	3	100	96.8	98.4	100	100	100	96.7	96.7	96.7
	4	100	100	100	96.7	100	98.3	96.7	100	98.3
	5	96.7	100	98.3	100	96.8	98.4	100	96.8	98.4

Table 7 demonstrates that the output performance of the six pipeline CNN models is an accuracy rate of 90.9–100%, a recall rate of 90–100%, and an F1-Measure of 94.7–100%. The results show the proposed approach is effective and the evaluation results can accurately classify the ponding volume in different pipelines.

4.4. Comparison of Proposed CNN Model with Other Models

To compare the proposed method with the current common methods, experiments of identical strategies but using DTM and SVM were conducted, with the Mel spectrogram as the input image. The SVM process was performed with the LIBSVM toolbox [41], with RBF as the kernel function and a kernel function with a parameter coefficient g of 2^{-27}, and a penalty factor coefficient c of 2^6 [22]. The DTM utilized the TreeBagger function, and NumTrees is set to 50 [42]. These recognition results are shown in Figure 8.



	1#	2#	3#	4#	5#	6#
Decision Tree	88.25	94.67	85.28	94.11	93.72	90.21
SVM	93.36	96.55	90.89	95.85	96.89	94.54
CNN	98.33	99.44	98.33	98.88	99.44	98.88

Figure 8. Comparison of recognition accuracy of three models.

In Figure 8, the symbols of 1#, 2#, 3#, 4#, 5#, 6# denote six different pipelines as shown in Table 1, respectively. This figure highlights that the recognition accuracies of the DTM with ponding volume of six pipelines are between 88.25% and 94.67%, the recognition accuracies of the SVM between 90.89% and 96.89%, and the recognition accuracies of the CNN between 98.33% and 99.44%. This proves that the CNN recognition model is more stable and has a higher accuracy than the other two models.

5. Conclusions

The paper has proposed a novel approach to identifying pipeline ponding volumes, by combining the percussive detection method and a CNN. The proposed approach is low-cost but user-friendly and effective. The experiment was performed based on the proposed method and the experimental results show the effectivity and high accuracy of the proposed recognition model. The major findings of the proposed approach can be summarized as follows:

- The way of processing percussion-caused audio signal by converting to Mel spectrogram can be considered as a novel and cost-effective approach in detecting pipeline ponding volume. It presents a simple but very effective acoustic signal processing method;
- The actual output of the CNN is basically consistent with the theoretical output during the proposed approach. The results demonstrate that the CNN recognition accuracy reaches 98.34% and can be effectively adopted to pipeline ponding detection;

- The proposed method is suitable for the detection of ponding volume in pipelines of different specifications, and the output performance of the six pipelines in the CNN models had an accuracy rate of 90.9–100%, a recall rate of 90–100%, and an F1-Measure of 94.7–100%;
- The recognition accuracy of CNN falls between 98.33% and 99.44%, which indicates that this recognition model has a more stable and superior performance than the DTM recognition model and the SVM recognition model. Therefore, it can be concluded that the method combining the percussive detection method and the CNN proposed in this paper has better application prospects in pipeline ponding detection.

The research in this paper demonstrates the feasibility and effectiveness of the proposed pipeline ponding detection method. The essence and mechanism of the proposed method is identifying underlying dynamical characteristics of percussion-caused audio signals of pipeline ponding. However, this paper also has its shortcomings: the length and diameter of the six different pipelines selected were too singular to determine the effective detection distance of the proposed percussion detection method. In follow-up research, designing corresponding experiments to detect the effective distance of the percussive detection method in pipeline health detection will become our research focus.

Author Contributions: Conceptualization, D.Y.; Data curation, M.X.; Funding acquisition, G.L.; Investigation, M.X.; Methodology, D.Y.; Project administration, M.X.; Resources, D.Y.; Software, M.X.; Supervision, D.Y. and T.W.; Validation, M.X.; Writing—original draft preparation, M.X.; Writing—review and editing, T.W. and G.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by National Natural Science Foundation of China (Grant No.: 51808417).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Due to the nature of this research, participants of this study did not agree for their data to be shared publicly and only available upon reasonable request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhang, J.; Wang, Z.; Liu, S.; Zhang, W.; Yu, J.; Sun, B. Prediction of hydrate deposition in pipelines to improve gas transportation efficiency and safety. *Appl. Energy* **2019**, *253*, 113521. [CrossRef]
2. Zhu, Y.; Wang, P.; Wang, Y.; Tong, R.; Yu, B.; Qu, Z. Assessment method for gas supply reliability of natural gas pipeline networks considering failure and repair. *J. Nat. Gas Sci. Eng.* **2021**, *88*, 103817. [CrossRef]
3. Huh, C.; Kang, S.G.; Cho, M.I.; Baek, J.H. Effect of Water and Nitrogen Impurities on CO_2 Pipeline Transport for Geological Storage. *Energy Procedia* **2011**, *4*, 2214–2221. [CrossRef]
4. Chae, M.; Jeong, H.D. Acceptance Sampling Plans for Pipeline Condition Assessment. *J. Pipeline Syst. Eng. Pract.* **2019**, *10*, 04019024. [CrossRef]
5. Zeng, W.; Dang, X.; Li, S.; Wang, H.; Wang, H.; Wang, B. Application of non-contact magnetic corresponding on the detection for natural gas pipeline. *E3S Web Conf.* **2020**, *185*, 01090. [CrossRef]
6. Licata, M.; Parker, H.M.O.; Aspinall, M.D.; Bandala, M.; Cave, F.; Conway, S.; Gerta, D.; Joyce, M.J. Fast neutron and γ-ray backscatter radiography for the characterization of corrosion-born defects in oil pipelines. *Eur. Phys. J. Conf.* **2020**, *225*, 06009. [CrossRef]
7. Soltysik, R.C. CCTV Pipeline Inspection System Data Management System and Computer-Based Monitoring/Action Application. U.S. Patent 7916170, 29 March 2011.
8. Khan, M.S. An acoustic based approach for mitigating sewer system overflows. In Proceedings of the Global Humanitarian Technology Conference, Seattle, DC, USA, 13–16 October 2016; pp. 782–789.
9. Hemavathi, R.; Pushpalatha, B.A. Crack and Object Detection in Pipeline using Inspection Robot. *J. Trend Sci. Res. Dev.* **2018**, *2*, 1072–1077.
10. Wang, T.; Wei, D.; Shao, J.; Li, Y.; Song, G. Structural Stress Monitoring Based on Piezoelectric Impedance Frequency Shift. *J. Aerosp. Eng.* **2018**, *31*, 04018092. [CrossRef]

11. Mustapha, S. Ultrasonic method for Measuring transport parameters using only the reflected waves at the first interface of porous materials having a rigid frame. *INTER—NOISE NOISE—CON Congr. Conf. Proc.* **2016**, *253*, 7258–7263.
12. Finger, C.; Saydak, L.; Vu, G.; Timothy, J.J.; Meschke, G.; Saenger, E.H. Sensitivity of Ultrasonic Coda Wave Interferometry to Material Damage—Observations from a Virtual Concrete Lab. *Materials* **2021**, *14*, 4033. [CrossRef]
13. Zheng, G.; Tian, Y.; Zhao, W.; Jia, S.; He, N. Band-Stop Filtering Method of Combining Functions of Butterworth and Hann Windows to Ultrasonic Guided Wave. *J. Pipeline Syst. Eng. Pract.* **2022**, *13*, 04021076. [CrossRef]
14. Yu, Y.; Safari, A.; Niu, X.; Drinkwater, B.; Horoshenkov, K.V. Acoustic and ultrasonic techniques for defect detection and condition monitoring in water and sewerage pipes: A review. *Appl. Acoust.* **2021**, *183*, 108282. [CrossRef]
15. Saracino, G.; Ambrosino, F.; Bonechi, L.; Cimmino, L.; D'Alessandro, R.; D'Errico, M.; Noli, P.; Scognamiglio, L.; Strolin, P. Applications of muon absorption radiography to the fields of archaeology and civil engineering. *Philos. Trans. Ser. A Math. Phys. Eng. Sci.* **2018**, *377*, 20180057. [CrossRef]
16. Yao, M.; Duvauchelle, P.; Kaftandjian, V.; Peterzol-Parmentier, A.; Schumm, A. Simulation of Computed Radiography X-ray Imaging Chain Dedicated to Complex Shape Objects. *Eur. Conf. Non Destr. Test.* **2014**, *10*, 6–10.
17. Schulze, R.; Krummenauer, F.; Schalldach, F.; d'Hoedt, B. Precision and accuracy of measurements in digital panoramic radiography. *Dento Maxillo Facial Radiol.* **2000**, *29*, 52–56. [CrossRef]
18. Ju, F.H.; Gong, X.B.; Jiang, L.B.; Hong, H.H.; Yang, J.C.; Xu, T.Z.; Chen, Y.; Wang, Z. Chronic myeloid leukaemia following repeated exposure to chest radiography and computed tomography in a patient with pneumothorax: A case report and literature review. *Oncol. Lett.* **2016**, *11*, 2398–2402. [CrossRef]
19. Adams, R.D.; Cawley, P.; Pye, C.J.; Stone, B.J. A vibration technique for non-destructively assessing the integrity of structures. *J. Mech. Eng. Sci.* **1978**, *20*, 93–100. [CrossRef]
20. Cawley, P.; Adams, R.D. The mechanics of the coin—tap method of nondestructive testing. *J. Sound Vib.* **1988**, *122*, 299–316. [CrossRef]
21. Cawley, P.; Adams, R.D. Sensitivity of the coin—tap method of nonde-structive testing. *Mater. Eval.* **1989**, *47*, 558–563.
22. Kong, Q.; Zhu, J.; Ho SC, M.; Song, G. Tapping and listening: A new approach to bolt looseness monitoring. *Smart Mater. Struct.* **2018**, *27*, 07LT02. [CrossRef]
23. Adams, R.D. Vibration measurements in nondestructive testing. In Proceedings of the 3rd International Conference on Emerging Technologies in Non Destructive Testing, Thessaloniki, Greece, 26–28 May 2003; pp. 27–35.
24. Wang, F.; Song, G. A novel percussion-based method for multi-bolt looseness detection using one-dimensional memory augmented convolutional long short-term memory networks. *Mech. Syst. Signal Process.* **2021**, *161*, 107955. [CrossRef]
25. Wang, F.; Ho, S.; Song, G. Modeling and analysis of an impact-acoustic method for bolt looseness identification. *Mech. Syst. Signal Processing* **2019**, *133*, 106249. [CrossRef]
26. Zheng, L.; Cheng, H.; Huo, L.; Song, G. Monitor concrete moisture level using percussion and machine learning. *Constr. Build. Mater.* **2019**, *229*, 117077. [CrossRef]
27. Chen, D.; Montano, V.; Huo, L.; Fan, S.; Song, G. Detection of subsurface voids in concrete-filled steel tubular (CFST) structure using percussion approach. *Constr. Build. Mater.* **2020**, *262*, 119761. [CrossRef]
28. Lall, A.; Scalzo, F.; Ullman, H.; Liebeskind, D.S.; Chien, A. Abstract P494: Automatically Predicting Modified Treatment in Cerebral Ischemia Scores From Patient Digital Subtraction Angiography Using Deep Learning. *Stroke* **2021**, *52* (Suppl. 1), AP494. [CrossRef]
29. Sharif Razavian, A.; Azizpour, H.; Sullivan, J.; Carlsson, S. CNN Features off-the-shelf: An Astounding Baseline for Recognition. In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops, Columbus, OH, USA, 23–28 June 2014; pp. 806–813.
30. Permana, S.D.H.; Saputra, G.; Arifitama, B.; Caesarendra, W.; Rahim, R. Classification of Bird Sounds as an Early Warning Method of Forest Fires using Convolutional Neural Network (CNN) Algorithm. *J. King Saud Univ.—Comput. Inf. Sci.* **2021**, *in press*. [CrossRef]
31. Hidayat, A.A.; Cenggoro, T.W.; Pardamean, B. Convolutional Neural Networks for Scops Owl Sound Classification. *Procedia Comput. Sci.* **2021**, *179*, 81–87. [CrossRef]
32. Valtierra-Rodriguez, M.; Rivera-Guillen, J.R.; Basurto-Hurtado, J.A.; De-Santiago-Perez, J.J.; Granados-Lieberman, D.; Amezquita-Sanchez, J.P. Convolutional Neural Network and Motor Current Signature Analysis during the Transient State for Detection of Broken Rotor Bars in Induction Motors. *Sensors* **2020**, *20*, 3721. [CrossRef]
33. Liu, F.; Shen, T.; Luo, Z.; Zhao, D.; Guo, S. Underwater target recognition using convolutional recurrent neural networks with 3-D Mel-spectrogram and data augmentation. *Appl. Acoust.* **2021**, *178*, 107989. [CrossRef]
34. Xie, J.; Hu, K.; Guo, Y.; Zhu, Q.; Yu, J. On loss functions and CNNs for improved bioacoustic signal classification. *Ecol. Inform.* **2021**, *64*, 101331. [CrossRef]
35. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al Dujaili, A.; Duan, Y.; Al Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *J. Big Data* **2021**, *8*, 53. [CrossRef]
36. Rodríguez-González, A.; Torres-Niño, J.; Valencia-Garcia, R.; Mayer, M.A.; Alor-Hernandez, G. Using experts feedback in clinical case resolution and arbitration as accuracy diagnosis methodology. *Comput. Biol. Med.* **2013**, *43*, 975–986. [CrossRef]

37. Manochandar, S.; Punniyamoorthy, M. A new user similarity measure in a new prediction model for collaborative filtering. *Appl. Intell.* **2020**, *5*, 586–615. [CrossRef]
38. Yan, L.; Zhong, B.; Ma, K.K. Confusion-Aware Convolutional Neural Network for Image Classification. In Proceedings of the International Conference on Neural Information Processing, Sydney, Australia, 12–15 December 2019.
39. Jung, S.Y.; Liao, C.H.; Wu, Y.S.; Yuan, S.M.; Sun, C.T. Efficiently Classifying Lung Sounds through Depthwise Separable CNN Models with Fused STFT and MFCC Features. *Diagnostics* **2021**, *11*, 732. [CrossRef]
40. Algermissen, S.; Hörnlein, M. Person Identification by Footstep Sound Using Convolutional Neural Networks. *Appl. Mech.* **2021**, *2*, 257–273. [CrossRef]
41. Chang, C.-C.; Lin, C.-J. LIBSVM: A library for support vector machines. *ACM Trans. Intel. Syst. Technol.* **2011**, *2*, 1–27. [CrossRef]
42. Cheng, H.; Wang, F.; Huo, L.; Song, G. Detection of sand deposition in pipeline using percussion, voice recognition, and support vector machine. *Struct. Health Monit.* **2020**, *19*, 2075–2090. [CrossRef]

materials

MDPI

Article

Guided Wave Phase Velocity Dispersion Reconstruction Based on Enhanced Phased Spectrum Method

Vykintas Samaitis * and Liudas Mažeika

Prof. K. Baršauskas Ultrasound Research Institute, Kaunas University of Technology, K. Baršausko St. 59, LT-51423 Kaunas, Lithuania; liudas.mazeika@ktu.lt
* Correspondence: vykintas.samaitis@ktu.lt

Abstract: Fibre-reinforced composite laminates are frequently used in various engineering structures, due to their increased weight-to-stiffness ratio, which allows to fulfil certain regulations of CO_2 emissions. Limited inter-laminar strength makes composites prone to formation of various defects, which leads to progressive degradation of residual strength and fatigue life of the structure. Using ultrasonic guided waves is a common technique for assessing the structural integrity of composite laminates. Phase velocity is one of the fundamental characteristics of guided waves and can be used for defect detection, material property estimation, and evaluation of dispersion. In this paper, a phase velocity reconstruction approach, based on the phase-shift method, was proposed, which uses frequency sweep excitation to estimate velocity at specific frequency harmonics. In contrast to the conventional phase spectrum technique, the proposed approach is applicable to the narrowband piezoelectric transducers and suitable for the reconstruction of dispersion curves for direct, converted, and multiple co-existing modes with high accuracy. The proposed technique was validated with finite element simulations and experiments, both on isotropic and anisotropic structures, analysing the direct, converted, and overlapped modes. The results demonstrated that, using the proposed technique, the phase velocity dispersion can be reconstructed at -20 dB level bandwidth of the transducer, with a relative error of $\pm 4\%$, compared to the theoretical velocity predictions.

Keywords: guided waves; mechanical properties; phase velocity; non-destructive testing; composites

Citation: Samaitis, V.; Mažeika, L. Guided Wave Phase Velocity Dispersion Reconstruction Based on Enhanced Phased Spectrum Method. *Materials* **2022**, *15*, 1614. https://doi.org/10.3390/ma15041614

Academic Editors: Phong B. Dao, Lei Qiu, Liang Yu and Zahra Sharif Khodaei

Received: 10 January 2022
Accepted: 18 February 2022
Published: 21 February 2022

1. Introduction

The composites market is one of the strategic development areas of the European Union, which aims to strengthen their competitiveness and extend the use of composites in the sectors of aerospace, automotive, and renewable energy [1]. The current EU demand of carbon fibre is estimated to be 35% of the global demand, and it will have an annual growth of 10–12% [1,2], while the UK market will grow from 2.5 billion up to 10 billion pounds a year by 2030 [3–5]. Fibre-reinforced composite laminates can offer increased strength- and stiffness-to-weight ratios, which allow for meeting the demanding requirements of CO_2 emissions. However, composites have limited interlaminar strength and are prone to formation of fibre breakage, matrix cracking, delaminations, porosity, and other structural defects. Such defects are usually hidden and progressively degrade the residual strength and fatigue life, eventually leading to sudden structural failure. Using ultrasonic guided waves is a common method for periodic inspection and monitoring of structural integrity of plate-like composite laminates, that offers large inspection areas and sensitivity to structural damage of various kinds [6–8]. To date, many studies are available that employ guided waves for the detection and quantification of impact damage [9–12], delaminations [13–17], and other defects in composite laminates. Guided wave propagation in composites is determined by many factors, including, but not limited to, multi-layered structure and anisotropy, object boundaries, dispersion, multiple co-existing modes, and mode conversion. Phase velocity is one of the fundamental properties of guided wave modes that depends on composition, structural integrity, elastic properties, and frequency-thickness product of composite.

Velocity measurements can be exploited both for material characterisation and damage detection, offering several benefits, such as validation of material properties, identification of wave-packets in complex guided wave signals, and sizing of defects [18–20].

However, reconstruction of phase velocity from overlapped, multimodal signals, and multi-layered anisotropic structures has been a long-standing problem. Initial phase velocity measurement approaches used threshold, zero crossing, or cross-correlation methods to evaluate the time-of-flight (ToF) of well-isolated guided wave modes [21]. The threshold method, in its simplest form, captures the time instance at which the signal crosses certain amplitude level. As these methods are based on signal amplitude, they are susceptible to noise and any other variation of the signal shape; hence, more advanced threshold-based ToF evaluation methods, such as variable ratios or similarity-based double threshold, were proposed [22,23]. The zero-crossing method seeks to obtain time instances at which the amplitude of the signal is equal to zero. To improve the accuracy of ToF estimation, using the zero-crossing technique, and avoid cycle skip problems, multiple zero-crossing points are being estimated within the same signal [24]. It is known that zero-crossing technique suffers from the phase uncertainty, especially at large propagation distances and under significant dispersion, as it become impossible to follow signal phase of the elongating wave packet and to avoid the cycle-skip. Recently, a technique based on zero-crossing and spectrum decomposition was proposed which exploits signals measured at sufficiently close distances and estimates the phase velocity, based on zero-crossing evaluation on signals filtered with different bandpass filters [25]. Cross-correlation technique is based on the measurement of correlation lag, between the received and reference signals. Such technique is considered suitable for low signal-to-noise ratio (SNR) signals, while the ToF accuracy mainly depends on the sampling ratio [26]. However, it is reported that cross-correlation-based ToF estimation may become significantly biased while analysing signals distorted due to scattering or dispersion [27].

The abovementioned ToF estimation methods can effectively be used for well-isolated and undistorted signals; however, they usually fail in analysing the overlapped, multimodal, scattered, and dispersed responses. Model-based approaches can partly deal with this problem by solving multi-dimensional and non-linear optimisation problems, while fitting synthetic signals to a segment of ultrasonic structural response. By using matching pursuit, chirplet transform, empirical mode decomposition or wavelet methods it is possible to decompose multimodal signals and to estimate their properties, such as frequency or ToF [28–32]. However, model-based methods are usually computationally expensive, as transformations are calculated in multi-dimensional space, while the selection of the mother wavelet or atoms is non-trivial task and may lead to unexpected results. It has been demonstrated that phase and group velocities can be reconstructed using phase-shift methods. First proposed by Sachse [33] and used by Schumacher [34], phase-shift methods are based on the estimation of the phase difference between transmitted and received signals, which is proportional to propagation distance. Initially, phase-shift methods were extensively used for bulk waves and later applied to laser-induced guided waves. In contrast to broadband laser-based excitation, piezoelectric sensors, that are more cost effective and commonly used in structural health monitoring applications, usually have quite narrow frequency band, due to the type of excitation, vibration mode, and size of the transducer; hence, the phase velocity reconstruction zone essentially becomes limited. Moreover, in order to avoid phase ambiguity, the distance between transmitted and recorded signals is required to be up to one wavelength, which limits spatial velocity distribution reconstruction capabilities.

In this paper, a phase velocity reconstruction approach is presented that uses phase-shift method and excitation frequency sweep to obtain phase velocity estimations in the entire band of transducer. Two sensors, positioned in close proximity, are used to record signals propagated through the structure and estimate the phase-shift between the signals. At each excitation frequency, the reconstruction of phase velocity is performed at specific frequency components only, which correspond to the peak values of the magnitude spectra. These peak frequencies depend on the frequency response of the excitation signal; hence,

phase velocity values can be collected at different frequencies, allowing us to achieve a wideband reconstruction. The validity of the approach is demonstrated through simulations and experiments by reconstructing the phase velocities of S_0 and converted A_0 modes, as well as identifying guided wave modes in complex multimodal signals.

In contrast to the classic phased spectrum method, the proposed approach allows to reduce the relative error of the phase velocity reconstruction from $\pm 11\%$ to $\pm 4\%$ and increase significantly the reconstruction bandwidth from -6 dB to -20 dB of the ultrasonic probe. As a result, using only two signals, measured in close proximity, the proposed phased spectrum method can achieve the reconstruction accuracy and bandwidth, which, to date, could be achieved only with techniques that include scanning of the sensor over a sufficiently large area.

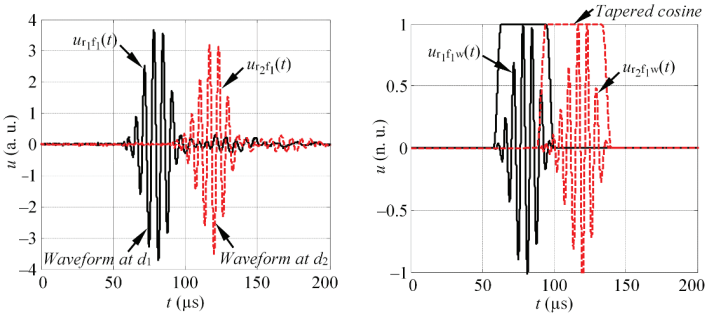
2. Description of Proposed Phase Velocity Estimation Method

The proposed phase velocity reconstruction approach employs a classic phase-shift method to estimate the velocity values at specific frequencies that correspond to peak values of the magnitude spectra of received signal. By repeating this procedure at different excitation frequencies, velocity values can be reconstructed at wide band, covering the entire bandwidth of the transducer. Variation of the excitation frequency allow different harmonics to be enhanced or suppressed, which is the key factor if reconstruction is performed at peak values of magnitude spectra only. The algorithm of the proposed method can be summarized with the following steps:

1. The transducer is driven by a burst at a central frequency of f_1, and the waveforms $u_{\mathrm{r1f1}}(t)$ and $u_{\mathrm{r2f1}}(t)$ are registered with receivers r_1 and r_2, each positioned at a distances d_1 and d_2 from the source (see Figure 1a for reference).
2. The waveforms $u_{\mathrm{r1f1}}(t)$ and $u_{\mathrm{r2f1}}(t)$ are windowed using the tapered cosine window $w(t)$ to isolate the wave packets of particular mode (see Figure 1b):

$$u_{\mathrm{r_1f_1w}}(t) = u_{\mathrm{r_1f_1}}(t) \cdot w(t - t_1), \;\; u_{\mathrm{r_2f_1w}}(t) = u_{\mathrm{r_2f_1}}(t) \cdot w(t - t_2) \tag{1}$$

where $u_{\mathrm{r1f1w}}(t)$ and $u_{\mathrm{r2f1w}}(t)$ represent the windowed versions of the waveforms $u_{\mathrm{r1f1}}(t)$ and $u_{\mathrm{r2f1}}(t)$, respectively; t_1 and t_2 correspond to the time instances of the maximum amplitude of the wave packet.



(**a**) (**b**)

Figure 1. (**a**) The example of the waveform, registered with receivers r_1 and r_2, at distances d_1 and d_2; (**b**) the illustration of waveform windowing to isolate the wave packet of single mode.

3. Each waveform, $u_{\mathrm{r1f1w}}(t)$ and $u_{\mathrm{r2f1w}}(t)$, is shifted in the time domain by $-t_{\mathrm{m1}}$ and $-t_{\mathrm{m2}}$, to avoid the uncertainties in the phase unwrapping procedure. The waveforms can be shifted according to the centroid of signal [35] or maximum value of the Hilbert envelope [36], in case of moderate dispersion:

$$u_{\mathrm{r_1f_1s}}(t) = u_{\mathrm{r_1f_1w}}(t + t_{\mathrm{m_1}}), \;\; u_{\mathrm{r_2f_1s}}(t) = u_{\mathrm{r_2f_1w}}(t + t_{\mathrm{m_2}}), \tag{2}$$

$$t_{m_1} = \arg\left(\max_t\left[\mathrm{HT}\left|u_{r_1f_1w}(t)\right|\right]\right),\ t_{m_2} = \arg\left(\max_t\left[\mathrm{HT}\left|u_{r_2f_1w}(t)\right|\right]\right),$$

where HT denotes the Hilbert transform; t_{m1} and t_{m2} are the time instances, which corresponds to the maximum of Hilbert envelope, in such a way that the influence of the signal delay due to phase velocity is compensated. The shift in time domain is illustrated in Figure 2a.

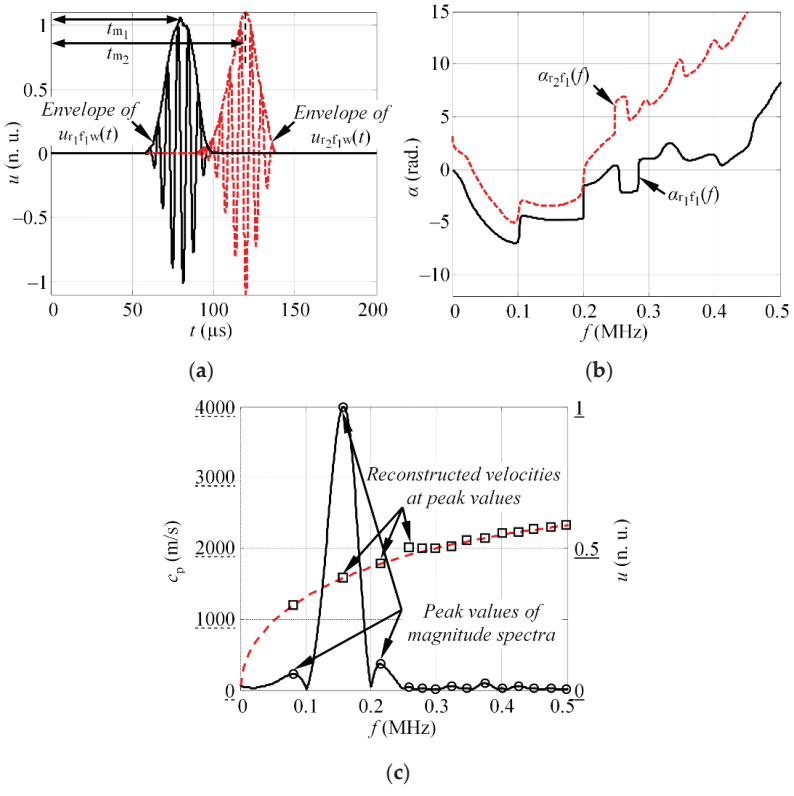


Figure 2. (**a**) The illustration of the shift of waveform in time domain to the maximum value of the Hilbert envelope; (**b**) the phase spectra of the waveforms, registered at distances d_1 and d_2; (**c**) the normalized magnitude spectra of the waveform, captured with receiver r_1 with the local maximum frequency values (circle markers), at which the phase velocity values are estimated (square markers) (dashed line represents the theoretical DC).

4. The complex frequency spectra of each time-shifted waveform, $u_{r1f1s}(t)$ and $u_{r2f1s}(t)$, is obtained employing the Fourier transform:

$$U_{r_1f_1}(jf) = \mathrm{FT}\left[u_{r_1f_1s}(t)\right],\ U_{r_2f_1}(jf) = \mathrm{FT}\left[u_{r_2f_1s}(t)\right] \quad (3)$$

where FT represents the Fourier transform.

5. The phase difference $\Delta\phi(f)$ between shifted signals $u_{r1f1s}(t)$ and $u_{r2f1s}(t)$, is estimated for a given frequency band f (see Figure 2b):

$$\Delta\varphi_{f_1}(f) = \left(\alpha_{r_1f_1}(f) - \alpha_{r_2f_1}(f)\right), \quad (4)$$

$$\alpha_{r_1f_1}(f) = \arctan\left[\frac{\mathrm{Im}\left[U_{r_1f_1}(jf)\right]}{\mathrm{Re}\left[U_{r_1f_1}(jf)\right]}\right],\ \alpha_{r_2f_1}(f) = \arctan\left[\frac{\mathrm{Im}\left[U_{r_2f_1}(jf)\right]}{\mathrm{Re}\left[U_{r_2f_1}(jf)\right]}\right],$$

where Im and Re represent the imaginary and real of the complex Fourier spectra.

Note that the phases $\alpha_{r1f1}(f)$ and $\alpha_{r2f1}(f)$ are calculated in a range of $[-\pi \dots \pi]$ radians. If the true phase of the particular frequency is less than $-\pi$ radians, it will be represented below the π radians. This means that some discontinuities will appear, in case the phase goes beyond the $\pm\pi$ radian limit. Therefore, the phases $\alpha_{r1f1}(f)$ and $\alpha_{r2f1}(f)$ have to be unwrapped.

6. The phase velocity, as a function of frequency, is calculated at particular frequencies, $f_{1,k1}$, using a modified version of the phase spectrum method:

$$c_p(f_{1,k_1}) = \frac{2\pi f_{1,k_1} d}{\Delta\varphi_{f_1}(f_{1,k_1}) - 2\pi f_{1,k_1}(t_{m_1} - t_{m_2})}, \quad (5)$$

where $f_{1,k1}$ are the frequencies that corresponds to the peak values of the magnitude spectra $|U_{r1f1}(jf)|$ at excitation frequency f_1; $k = 1 \div K_1$, K_1—is the total number of detected peaks at excitation frequency f_1, and d is the separation distance between the receivers r_1 and r_2 ($d = d_2 - d_1$). The frequency selection for phase velocity estimation is illustrated in Figure 2c.

7. The intermediate values of the phase velocities at other frequencies are obtained by changing the excitation frequency to f_2 and repeating the whole routine described above. The final result is obtained by combining the calculations at different excitation frequencies $f_1 \dots f_N$:

$$c_p(f) = \underset{f}{\mathrm{sort}}\{c_p(f_{1,k_1}), \dots, c_p(f_{n,k_n}), \dots, c_p(f_{N,k_N})\}, \quad (6)$$

where N is the number of excitation frequencies used to drive the emitter.

The method presented above is applicable to flat structures with uniform thickness, which can be multi-layered, anisotropic, or isotropic. In contrast to the conventional phase spectrum method, it provides better accuracy of velocity estimation, which will be demonstrated in the subsequent Chapter.

3. Experimental Validation on Isotropic Samples

In this section, the proposed phase velocity estimation approach is validated with the appropriate experiments. For this purpose, the phase velocity values, extracted with the proposed approach, are compared with the theoretical calculations, which were considered a reference. In this study, the velocities of the S_0 mode in the aluminium sample will be analysed.

The experiments were carried out on the aluminium alloy 2024 T6 plate, which was 2 mm thick, 650 mm wide, and 1250 mm long. The well-known isotropic material was deliberately selected for this study, in order to be able to compare the experimental results with the theoretically estimated values. The S_0 mode was launched into the structure by attaching the thickness mode transducer to the edge of the Al plate, as is shown in Figure 3. For the reception, two transducers, r_1 and r_2, possessing the same characteristics, were bonded perpendicularly to the upper surface of the specimen at distances d_1 = 450 mm and d_2 = 550 mm from the source (see Figure 3).

In this paper, transducers with a central frequency of 240 kHz and bandwidth of 340 kHz at -6 dB level were used. The frequency response of the probe can be seen in Figure 4a. To reconstruct the dispersion curve under the wide band, two different scenarios employing the square pulse excitation were used, as follows: n_1 = 3 cycles, f_1 = 150 kHz; and n_2 = 3 cycles, f_2 = 200 kHz. Such excitation frequencies were deliberately selected, according to the magnitude spectrum of excitation pulse, which can be seen in Figure 4b. The results, presented in the figure, demonstrate that a minor shift of excitation frequency from 150 to 200 kHz enables peak amplitudes of the magnitude spectra to be obtained at

different frequencies. Moreover, the local maximum values, in case of 200 kHz excitation, mostly correspond to the local minimum frequencies of 150 kHz excitation. Thus, excitation under the selected frequencies enables a large variety of reconstruction frequencies to be obtained. In this case, it was presumed that the selected excitation frequencies will provide a sufficient amount of velocity values. In other cases, more excitation frequencies may be used, exploiting the whole bandwidth of the transducer.

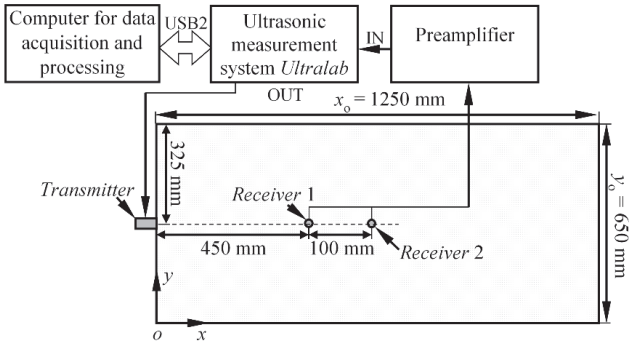


Figure 3. The schematic diagram of the experimental set-up for the validation of phase velocity estimation method.

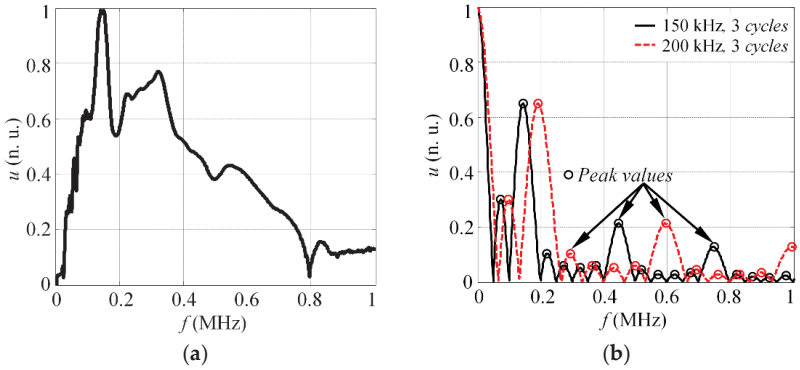


Figure 4. (**a**) The frequency response of the thickness mode transducer used for the experiments; (**b**) the magnitude spectra of three cycles (150 kHz) (solid line) and three cycles (200 kHz) (dashed line) square excitation pulse.

The experimental waveforms of the S_0 mode, at distances d_1 and d_2, under the f_1 = 150 and f_2 = 200 kHz excitation, are presented in Figure 5a,b, respectively. The magnitude spectra, $|U_{r1f1}(jf)|$ and $|U_{r2f2}(jf)|$, of the windowed S_0 mode wave packet can be seen in Figure 5c. The frequencies at which the phase velocity values were extracted are indicated with circle markers. Finally, the reconstructed dispersion curve of the phase velocity for the S_0 mode, along with theoretical estimation, is shown on Figure 5d. The theoretical dispersion curve was calculated by employing the SAFE method and material properties of aluminium 2024 T6 (the density: ρ = 2780 kg/m^3; Young's modulus: E = 72 GPa; Poisson's ratio: υ = 0.35).

The results in Figure 5d show that the phase velocities are reconstructed in the frequency band up to 0.8 MHz. According to the frequency response of the transducer used in this study (see Figure 4a), the technique enables the phase velocities in the −20 dB level bandwidth of the actuator to be reconstructed. In this study, a total of K = 52 velocity values were extracted at a band up to 1 MHz. This means that using two frequencies to drive the

transducer, 52 reconstruction points were observed that correspond to peak frequencies of the magnitude spectra. Such a number of reconstruction points is relative and depends on the total number of excitation frequencies, N, and obtained number of peak values of magnitude spectra, in case of each excitation frequency.

It is noteworthy that the general reliability of the phase spectrum method depends on the proper selection of the time window to crop the wave packet of the single mode for FFT. The proposed method implicitly assumes that only one mode is present at the selected time window.

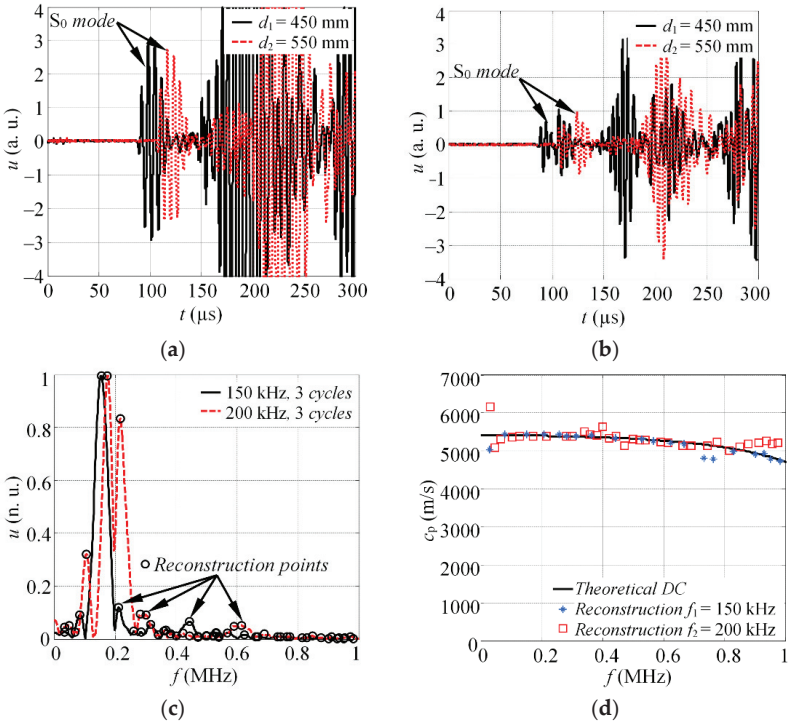


Figure 5. The experimental waveforms of the S_0 mode, at distances d_1 and d_2, in case of (**a**) 150 (**b**) and 200 kHz excitation; (**c**) the magnitude spectra of windowed S_0 mode at different excitation frequencies; (**d**) the combined reconstruction of dispersion relations.

In order to estimate the agreement of the results with theoretical phase velocities, the standard deviation (STD) was used as a measure of spread:

$$\sigma = \sqrt{\frac{1}{K_1 - 1} \sum_{i=1}^{K_1} \left| \left(c_\mathrm{p}(f_i) - c_\mathrm{t}(f_i) \right) - \mu \right|^2}, \tag{7}$$

$$\mu = \frac{1}{K_1} \sum_{i=1}^{K_1} \left(c_\mathrm{p}(f_i) - c_\mathrm{t}(f_i) \right), \tag{8}$$

where K_1 is number of points in reconstructed phase velocities, $c_\mathrm{p}(f_i)$ is a vector of reconstructed phase velocity values, and $c_\mathrm{t}(f_i)$ are the corresponding reference phase velocity values, calculated using the SAFE method. The estimated standard deviation of the calculated phase velocity values is σ = 161 m/s. This leads to the conclusion that 40 out of 52 velocity values (77%) are within the standard deviation range, as shown in Figure 6a.

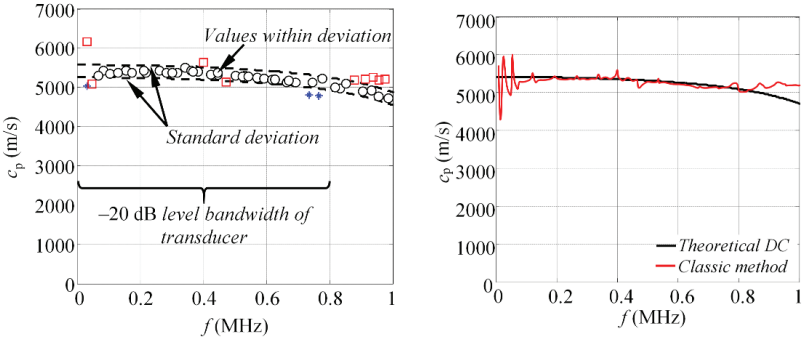


Figure 6. (**a**) The graphic representation of standard deviation, showing the spread of estimated phase velocity values and (**b**) reconstruction of phase velocity dispersion curves using the classic phased spectrum method.

The experimental results, presented in this section, demonstrate that proposed approach reconstructs the phase velocity values at frequencies up to 800 kHz for the selected probe. At frequencies above 800 kHz, the approach starts to fail at capturing the pattern of the dispersion curve. Hence, it can be said that the phase velocity values of the S_0 mode can be reconstructed at -20 dB bandwidth or 0.1 level of the transducer, according to its normalized magnitude spectra, presented at Figure 4a. The standard deviation of the reconstructed phase velocities, calculated according to Equation (7), is 161 m/s, which provides relative error of phase velocity estimation equal to $\pm 3\%$ for the S_0 mode, calculated according to:

$$\delta = \left(\frac{\sigma \times 100\%}{\mu_{c_t(f)}} \right), \tag{9}$$

where $\mu_{c_t(f)}$ is the mean theoretical phase velocity value in the selected frequency band under analysis.

In order to emphasize the achieved improvement, the signals of the S_0 mode, obtained at 200 kHz, were processed using classic phased spectrum method, described in [33,34]. The reconstructed phase velocity curve is presented at Figure 6b. The results indicate that highest velocity reconstruction accuracy can be obtained at frequency band 200–340 kHz, which corresponds to -6 dB bandwidth of the sensor. The standard deviation of the S_0 mode phase velocity reconstruction is estimated to be 592 m/s for the classic phase spectrum method, which gives $\pm 11\%$ relative phase velocity reconstruction error. In can be concluded that proposed approach allow to increase the reconstruction bandwidth, from 140 to 800 kHz, and reduce the relative velocity estimation error, from $\pm 11\%$ to $\pm 3\%$, for the S_0 mode.

4. Identification of Converted Modes

In this section, the numerical validation of the proposed phase velocity reconstruction method will be presented. The major focus will be given to the method performance, in case the analysed signal is surrounded by the wave packets of other co-existing modes. To achieve the purpose of this study, the phase velocities of the converted A_0 mode will be analysed, which convert from the S_0 mode, due to the presence of notch.

To fulfil the scope of this research, the 3D linear structural mechanics finite element model of isotropic aluminium alloy 2024 T6 plate ($600 \times 200 \times 2$ mm) is considered. The top view of the analysed structure is presented on Figure 7. The S_0 mode was initially launched into the structure by applying the in-plane force to the shortest edge of the Al plate. To generate the converted A_0 mode, the vertical 36 mm wide (along x axis) crack-type defect, with a depth of 66% of the plate thickness, was introduced by duplicating the

nodes of the mesh. In such way, a complete disbond was simulated, without changing the shape of finite element model. It was shown by the various researchers that, if a crack is not symmetrical to the middle plane of the plate, according to the thickness, the mode conversion takes place upon the wave interaction with the notch, and both the S_0 and A_0 modes are expected as the reflected and transmitted waves [37]. The defect was centred, with respect to the short edge of the sample, and situated at the distance of 200 mm from source of Lamb waves (see Figure 7).

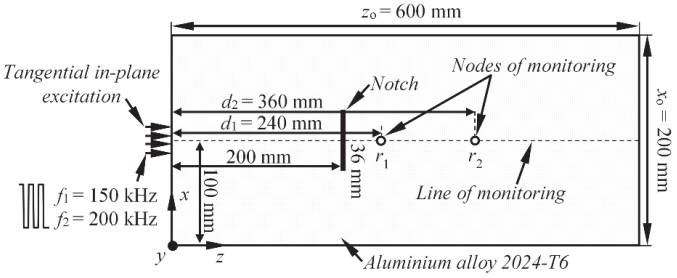


Figure 7. Illustration of the set-up of numerical aluminium plate FE model with the notch.

Throughout the simulations, the ANSYS 17.1 implicit solver and 3D structural solid solid64 finite elements were used, which are defined by eight nodes having three degrees of freedom at each node and 2 × 2 × 2 integration points. The finite elements were hexahedrons, meshed using structured grid. Once again, two different scenarios employing the square pulse excitation were used, as it was described in the previous section. At first, the excitation pulse consisted of n_1 = 3 cycles and a central frequency of f_1 = 150 kHz. Meanwhile, in the second case, the Lamb waves were excited with n_2 = 3 cycles at f_2 = 200 kHz. The average mesh size was equal to 0.5 mm, which corresponds to 21 nodes per wavelength for the slowest A_0 mode at f_1 and 17 nodes per wavelength at f_2. The integration steps in the time domain were 0.33 and 0.25 μs, respectively, which produces a 1/20 of the period, both at f_1 and at f_2. The variable monitored in this study was a vertical component of particle velocity (y) along the centreline of the sample. The waveforms for the phase velocity estimation were selected along the centreline of the sample at distances d_1 = 240 mm and d_2 = 360 mm. The B-scan images of the longitudinal (z) and vertical component (y) of the particle velocity, showing the S_0 and converted A_0 modes, are presented in Figure 8a,b.

The simulated waveforms of the converted A_0 modes, at distances d_1 and d_2, in case of f_1 = 150 kHz and f_2 = 200 kHz excitation, are presented in Figure 9a,b. The selected time windows to cut the wave packet of single mode are indicated with vertical dashed lines. The magnitude spectra of windowed A_0 mode, at frequencies f_1 and f_2, along with indicated reconstruction frequencies, can be seen on Figure 9c. Finally, the comparison of estimated DC, with the theoretical calculations, is shown on Figure 9d. The results demonstrate a good match between the estimated results and theoretical phase velocities, calculated with the SAFE method. The standard deviation of the reconstructed velocities is equal to σ = 47.3 m/s. Overall, the K = 32 velocity values were extracted, while 20 (63%) of them were within the range of standard deviation. Even though the number of reconstruction points is less than from the experiments present in previous section, Figure 9d suggests that its quite sufficient for the reconstruction of the segment of dispersion curve. The proposed approach is not limited with two excitation frequencies; hence, the number of reconstruction points can be increased if the segment of dispersion curve is not represented properly.

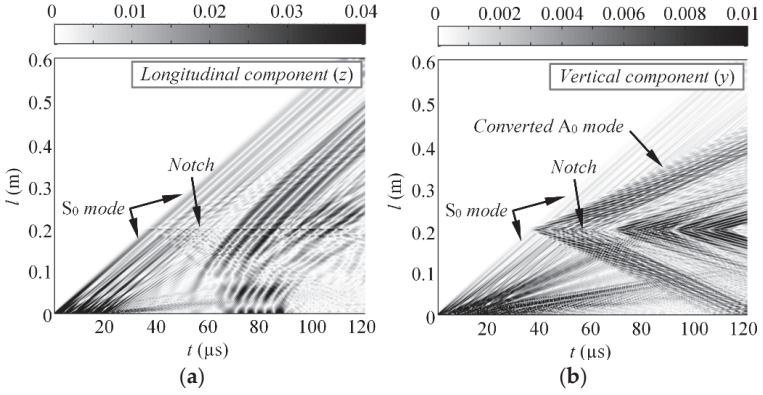


Figure 8. The (**a**) longitudinal and (**b**) vertical component of particle velocity along the centreline of the sample, in case of 150 kHz excitation.

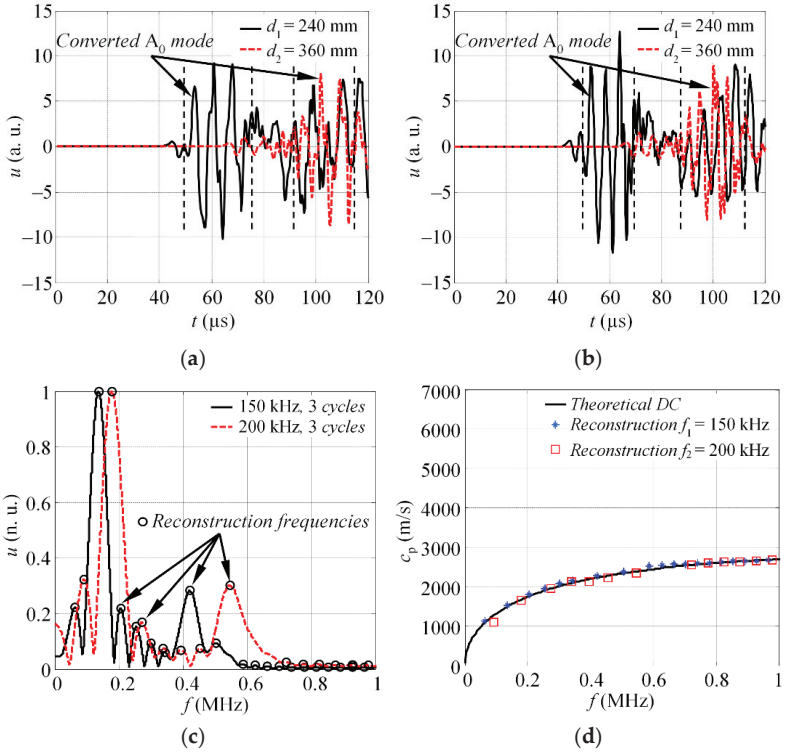


Figure 9. The simulated waveforms of the converted A_0 mode, at distances d_1 and d_2, in case of (**a**) 150 and (**b**) 200 kHz excitation; (**c**) the magnitude spectra of windowed A_0 mode at different excitation frequencies; (**d**) the combined reconstruction of phase velocity dispersion curve along with the theoretical estimation.

As it was mentioned previously, the time window selection ambiguity is quite essential in the success of phase velocity reconstruction using phase-shift method, especially for overlapped modes. Hence, the selected time window must hold the single mode only. For complex structures, where signals undergo many reflections, the reconstruction can be

quite uncertain. On the other hand, it will be demonstrated in the next chapter that phase velocities can be estimated using part of the signal only. In such a case, the position of the time window must be optimised, i.e., by solving a minimisation problem, to get reasonable velocity reconstruction results.

5. Analysis of Multimodal Signals in Anisotropic Structures

In this section, the performance of the proposed phase velocity reconstruction approach is validated qualitatively by analysing the experimental multimodal signals in an anisotropic structure. For this purpose, the experiments have been carried out in a pitch-catch configuration on the 6-ply GFRP plate (biaxial: 0° and 90°/bias: ±45°/biaxial: 0° and 90°), with dimensions x_o = 2000, y_o = 1000, and 4 mm thickness (see Figure 10).

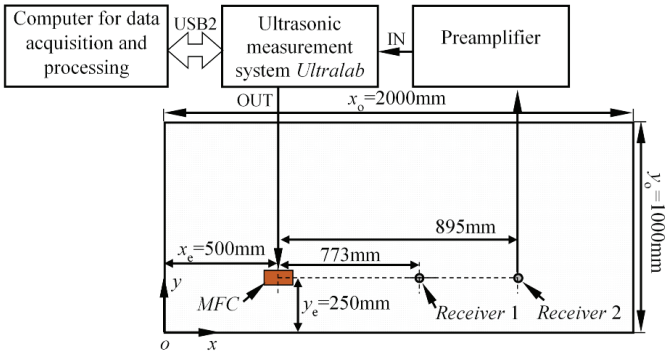


Figure 10. The schematic diagram of the experiments for validation of phase velocity reconstruction approach.

The Lamb waves were generated using the MFC transducer, centred at the coordinates x_e = 500 mm, y_e = 250 mm. It was bonded to the surface of the specimen using a thin layer of gasket maker. The emitter was excited by a three-cycle square pulse, with a central frequency of 100 kHz, where the fundamental A_0 and S_0 modes exist in the structure. In this case, the measurements were recorded at a single excitation frequency. Two waveforms were recorded along the wave path (0° propagation), at the distances d_1 = 773 mm and d_2 = 895 mm from the source of Lamb waves (see Figure 10). The proposed phase velocity estimation method was used to extract velocities of the four wave packets: direct S_0, bottom reflected S_0, left top edge reflected S_0, and direct A_0 mode. The experimentally obtained waveforms, at the distances d_1 and d_2, are presented in Figure 11a,b. The start and stop points of the time windows used to crop the wave packets are indicated by dashed squares.

The reconstructed phase velocities of different reflections can be seen in Figure 12a–d. The standard deviations for each case of reconstruction are summarized in Table 1. Note that the reconstructed velocity values below 40 kHz were not considered in the calculations of STD.

Table 1. The standard deviation of the estimated phase velocities for different GW mode packets.

Type of Mode	Velocity Standard Deviation, σ (m/s)
S_0 direct	97.4
S_0 bottom edge	202.7
S_0 left and top edge	224.5
A_0 direct	51.5

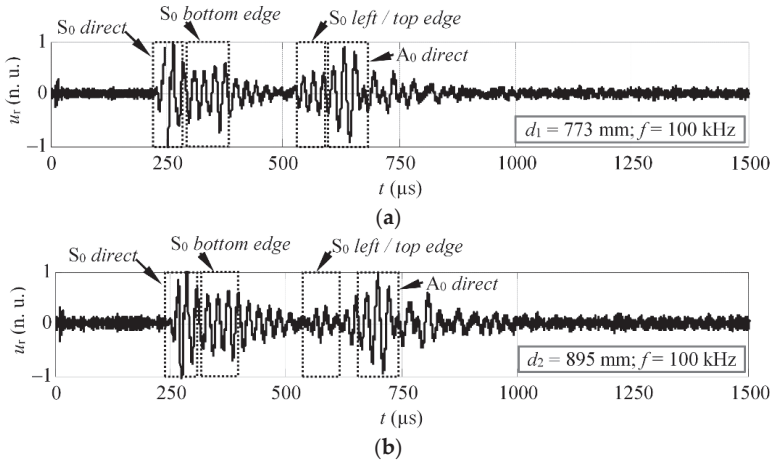


Figure 11. The experimental waveforms obtained on the GFRP sample along the propagation path, at distances (**a**) d_1 and (**b**) d_2.

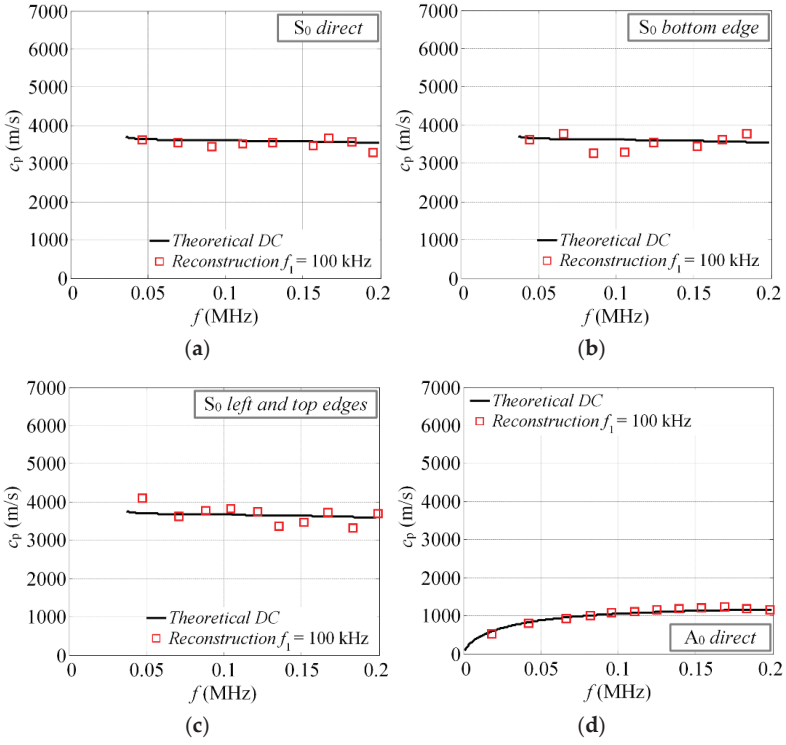


Figure 12. The reconstructed phase velocities of the S_0 and A_0 modes: (**a**) S_0 direct; (**b**) S_0 bottom edge reflected; (**c**) S_0 left and top edge reflected; (**d**) A_0 direct.

The results presented above (Figure 12) were found to be in quite good agreement with the theoretical calculations. It suggests that the proposed technique can be used with a certain reliability to extract the phase velocities of GW and identify modes in complex signals. The results show that the velocities of the direct modes are closer to the theoretical

values, in comparison to the reflected ones. The average deviation for the direct modes (A_0 and S_0) is approximately 75 m/s, while for the reflected S_0 modes, it is 213 m/s. Several factors may influence the reliability of the results, though. First of all, the selected time windows in Figure 11a,b (dashed squares) may give an idea that this procedure is not very straightforward, especially for the reflected modes. As it turns out, in some cases, part of the wave packet has to be cropped to get better velocity estimation. Another important factor is the propagation distance, which varies for modes arriving at different directions. It means that the distance (d) has to be predefined for each wave packet separately. If the propagation distance is not known in advance, an additional error will be obtained. The study revealed that the proposed velocity estimation technique gives an approximate relative experimental error of $\pm 4\%$, in comparison to theoretical predictions. Meanwhile, for the incident modes, the relative error is always less than a $\pm 2.5\%$. For example, the 2D FFT method gives an error of approximately of 1% [38]. However, in the study above, the authors used a set of 64 time series, spatially sampled at 1 mm, to achieve such accuracy.

6. Conclusions

In this paper, a phase velocity reconstruction approach, based on the phased spectrum method, was developed, which exploits several excitation frequencies of the ultrasonic probe and estimates phase velocity values at peak frequencies of the magnitude spectra. The proposed approach allows us to reconstruct phase velocities with high accuracy in wide frequency bandwidth using only two waveforms measured at close proximity. In contrast to the classic phase spectrum method, the proposed technique offers an increased reconstruction bandwidth (from -6 to -20 dB) and reduced relative error of phase velocity reconstruction (from $\pm 11\%$ to $\pm 4\%$). The main outcomes of the research can be summarized as follows:

- It was found that the accuracy of the classic phase spectrum method can be improved if several frequencies are used to drive the transducer, while the phase velocities are reconstructed at peak values of Fourier spectra only. Such an approach allows us to avoid low energy frequency components, where the velocity estimation error is likely to increase.
- The initial experiments demonstrated that the proposed phased spectrum method can increase the reconstruction bandwidth, from -6 to -20 dB, of the sensor and improve the standard deviation of velocity reconstruction, from 592 to 161 m/s. For the experimental S_0 mode, this results in a velocity estimation relative error improvement, from $\pm 11\%$ to $\pm 3\%$.
- The finite element simulations demonstrated the applicability of the proposed approach in detecting converted guided wave modes. It was demonstrated that the phase velocities of converted modes can be reconstructed with a standard deviation of 47.3 m/s, even if the modes are partly overlapped with direct waves.
- Finally, the proposed method was demonstrated to be appropriate for the analysis of complex guided wave signals, with multiple co-existing modes. It was estimated that average deviation for the direct modes (A_0 and S_0) is approximately 75 m/s, while for the reflected S_0 modes, it is 213 m/s. While analysing the overlapped complex guided wave signals, the proper selection of time gate is the most important parameter for the accuracy of reconstruction.
- It was estimated that the average phase velocity reconstruction error of the proposed method, including both symmetrical and asymmetrical modes, is up to $\pm 4\%$. The classic phase spectrum method provides an approximate reconstruction error of $\pm 11\%$. Other techniques, reported in the literature, can achieve velocity reconstruction with an average error of $\pm 1\%$; however, at least 64 signals need to be acquired to achieve such accuracy.

Author Contributions: Conceptualization, V.S. and L.M.; methodology, L.M.; validation, V.S. and L.M.; formal analysis, L.M.; investigation, V.S. and L.M.; data curation, V.S.; writing—original draft

preparation, V.S.; writing—review and editing, L.M.; visualization, V.S.; supervision, L.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Research Foundation of the Research Council of Lithuania, under the project UNITE " Ultrasonic tomography for spatial reconstruction of material properties from limited number of positions for NDT and monitoring applications ", no. MIP-2048.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Pavel, C.C.; Blagoeva, D.T. Materials Impact on the EU's Competitiveness of the Renewable Energy, Storage and E-Mobility Sectors. 2017. Available online: http://publications.jrc.ec.europa.eu/repository/bitstream/JRC108356/jrc108356_report_competitiveness_online.pdf (accessed on 10 January 2022).
2. Kühnel, M.; Kraus, T. The Global CFRP Market 2016. 2016. Available online: https://elib.dlr.de/109625/1/CFRP%20market%20report%20ICC%202016%20K%C3%BChnel.pdf (accessed on 10 January 2022).
3. Composite Materials Market Forecast for the United Kingdom (UK) from 2015 to 2030, by Industrial Sector. 2016. Available online: https://www.statista.com/statistics/624539/composite-market-industry-sector-uk/ (accessed on 10 January 2022).
4. Hinton, M. The Composites Age Is Arriving. 2017. Available online: https://www.bristol.ac.uk/media-library/sites/composites/documents/Bristol%20Composites%20Inst%20Launch%20151117v4.pdf (accessed on 10 January 2022).
5. The UK Composites Strategy. 2009. Available online: https://compositesuk.co.uk/system/files/UK%20Composites%20Strategy%202009.pdf (accessed on 10 January 2022).
6. Zhao, X.; Gao, H.; Zhang, G.; Ayhan, B.; Yan, F.; Kwan, C.; Rose, J.L. Active Health Monitoring of an Aircraft Wing with Embedded Piezoelectric Sensor/Actuator Network: I. Defect Detection, Localization and Growth Monitoring. *Smart Mater. Struct.* **2007**, *16*, 1208–1217. [CrossRef]
7. Mei, H.; Haider, M.F.; Joseph, R.; Migot, A.; Giurgiutiu, V. Recent Advances in Piezoelectric Wafer Active Sensors for Structural Health Monitoring Applications. *Sensors* **2019**, *19*, 383. [CrossRef] [PubMed]
8. Sikdar, S.; Banerjee, S. *Structural Health Monitoring of Advanced Composites Using Guided Waves*; LAP LAMBERT Academic Publishing: Saarbrucken, Germany, 2017.
9. Hervin, F.; Maio, L.; Fromme, P. Guided Wave Scattering at a Delamination in a Quasi-Isotropic Composite Laminate: Experiment and Simulation. *Compos. Struct.* **2021**, *275*, 114406. [CrossRef]
10. Castaings, M.; Singh, D.; Viot, P. Sizing of Impact Damages in Composite Materials using Ultrasonic Guided Waves. *NDT E Int.* **2012**, *46*, 22–31. [CrossRef]
11. De Luca, A.; Caputo, F.; Sharif Khodaei, Z.; Aliabadi, M.H. Damage Characterization of Composite Plates Under Low Velocity Impact using Ultrasonic Guided Waves. *Compos. Part B Eng.* **2018**, *138*, 168–180. [CrossRef]
12. Zhang, B.; Sun, X.C.; Eaton, M.J.; Marks, R.; Clarke, A.; Featherston, C.A.; Kawashita, L.F.; Hallett, S.R. An Integrated Numerical Model for Investigating Guided Waves in Impact-Damaged Composite Laminates. *Compos. Struct.* **2017**, *176*, 945–960. [CrossRef]
13. Ramadas, C.; Balasubramaniam, K.; Joshi, M.; Krishnamurthy, C.V. Interaction of Guided Lamb Waves with an Asymmetrically Located Delamination in a Laminated Composite Plate. *Smart Mater. Struct.* **2010**, *19*, 065009. [CrossRef]
14. Samaitis, V.; Mažeika, L.; Rekuvienė, R. Assessment of the Length and Depth of Delamination-Type Defects using Ultrasonic Guided Waves. *Appl. Sci.* **2020**, *10*, 5236. [CrossRef]
15. Leckey, C.A.C.; Seebo, J.P. Guided Wave Energy Trapping to Detect Hidden Multilayer Delamination Damage. *AIP Conf. Proc.* **2015**, *1650*, 1162–1169.
16. Munian, R.K.; Mahapatra, D.R.; Gopalakrishnan, S. Lamb Wave Interaction with Composite Delamination. *Compos. Struct.* **2018**, *206*, 484–498. [CrossRef]
17. Munian, R.K.; Roy Mahapatra, D.; Gopalakrishnan, S. Ultrasonic Guided Wave Scattering due to Delamination in Curved Composite Structures. *Compos. Struct.* **2020**, *239*, 111987. [CrossRef]
18. White, A.; Hong, J.; Hong, S.; Choi, J. Parameter Estimation for Wavelet Transformed Ultrasonic Signals. *NDT E Int.* **2011**, *44*, 32–40. [CrossRef]
19. Hoseini, M.R.; Wang, X.; Zuo, M.J. Estimating Ultrasonic Time of Flight using Envelope and Quasi Maximum Likelihood Method for Damage Detection and Assessment. *Measurement* **2012**, *45*, 2072–2080. [CrossRef]
20. Raišutis, R.; Tiwari, K.A.; Žukauskas, E.; Tumšys, O.; Draudvilienė, L. A Novel Defect Estimation Approach in Wind Turbine Blades Based on Phase Velocity Variation of Ultrasonic Guided Waves. *Sensors* **2021**, *21*, 4879. [CrossRef] [PubMed]
21. Andria, G.; Attivissimo, F.; Giaquinto, N. Digital Signal Processing Techniques for Accurate Ultrasonic Sensor Measurement. *Measurement* **2001**, *30*, 105–114. [CrossRef]
22. Zhu, W.; Xu, K.; Fang, M.; Shen, Z.; Tian, L. Variable Ratio Threshold and Zero-Crossing Detection Based Signal Processing Method for Ultrasonic Gas Flow Meter. *Measurement* **2017**, *103*, 343–352. [CrossRef]

23. Fang, Z.; Hu, L.; Mao, K.; Chen, W.; Fu, X. Similarity Judgment-Based Double-Threshold Method for Time-of-Flight Determination in an Ultrasonic Gas Flowmeter. *IEEE Trans. Instrum. Meas.* **2018**, *67*, 24–32. [CrossRef]
24. Fang, Z.; Su, R.; Hu, L.; Fu, X. A Simple and Easy-Implemented Time-of-Flight Determination Method for Liquid Ultrasonic Flow Meters Based on Ultrasonic Signal Onset Detection and Multiple-Zero-Crossing Technique. *Measurement* **2021**, *168*, 108398. [CrossRef]
25. Draudviliene, L.; Tumsys, O.; Mazeika, L.; Zukauskas, E. Estimation of the Lamb Wave Phase Velocity Dispersion Curves using Only Two Adjacent Signals. *Compos. Struct.* **2021**, *258*, 113174. [CrossRef]
26. Queirós, R.; Martins, R.C.; Girão, P.S.; Cruz Serra, A. A New Method for High Resolution Ultrasonic Ranging in Air. In Proceedings of the XVIII IMEKO World Congress Metrology for a Sustainable Development, Rio de Janeiro, Brazil, 17–22 September 2006.
27. Angrisani, L.; Moriello, R.S.L. Estimating Ultrasonic Time-of-Flight through Quadrature Demodulation. *IEEE Trans. Instrum. Meas.* **2006**, *55*, 54–62. [CrossRef]
28. Feng, Z.; Liang, M.; Chu, F. Recent Advances in Time–frequency Analysis Methods for Machinery Fault Diagnosis: A Review with Application Examples. *Mech. Syst. Signal Processing* **2013**, *38*, 165–205. [CrossRef]
29. Kim, C.; Park, K. Mode Separation and Characterization of Torsional Guided Wave Signals Reflected from Defects using Chirplet Transform. *NDT E Int.* **2015**, *74*, 15–23. [CrossRef]
30. Le, T.; Argoul, P. Continuous Wavelet Transform for Modal Identification using Free Decay Response. *J. Sound Vibrat.* **2004**, *277*, 73–100. [CrossRef]
31. Xu, B.; Giurgiutiu, V.; Yu, L. Lamb Waves Decomposition and Mode Identification using Matching Pursuit Method. In *Proceedings Volume 7292, Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2009*; SPIE: San Diego, CA, USA, 2009.
32. Frank Pai, P. Time–frequency Analysis for Parametric and Non-Parametric Identification of Nonlinear Dynamical Systems. *Mech. Syst. Signal Processing* **2013**, *36*, 332–353. [CrossRef]
33. Sachse, W.; Pao, Y. On the Determination of Phase and Group Velocities of Dispersive Waves in Solids. *J. Appl. Phys.* **1978**, *49*, 4320–4327. [CrossRef]
34. Schumacher, N.A.; Burger, C.P.; Gien, P.H. A Laser-based Investigation of Higher-order Modes in Transient Lamb Waves. *J. Acoust. Soc. Am.* **1993**, *93*, 2981–2984. [CrossRef]
35. Rokhlin, S.I.; Lewis, D.K.; Graff, K.F.; Adler, L. Real-time Study of Frequency Dependence of Attenuation and Velocity of Ultrasonic Waves during the Curing Reaction of Epoxy Resin. *J. Acoust. Soc. Am.* **1986**, *79*, 1786–1793. [CrossRef]
36. Luo, H.; Fang, X.; Ertas, B. Hilbert Transform and its Engineering Applications. *AIAA J.* **2009**, *47*, 923–932. [CrossRef]
37. Castaings, M.; Le Clezio, E.; Hosten, B. Modal Decomposition Method for Modeling the Interaction of Lamb Waves with Cracks. *J. Acoust. Soc. Am.* **2002**, *112*, 2567–2582. [CrossRef]
38. Alleyne, D.; Cawley, P. A Two-Dimensional Fourier Transform Method for the Measurement of Propagating Multimode Signals. *J. Acoust. Soc. Am.* **1991**, *89*, 1159–1168. [CrossRef]

sensors

MDPI

Article

Gearbox Fault Diagnosis Based on Improved Variational Mode Extraction

Yuanjing Guo [1,*], Shaofei Jiang [2], Youdong Yang [1], Xiaohang Jin [2] and Yanding Wei [3]

[1] Zhijiang College, Zhejiang University of Technology, Shaoxing 312030, China; yydong@zjut.edu.cn
[2] College of Mechanical Engineering, Zhejiang University of Technology, Hangzhou 310023, China; jsf75@zjut.edu.cn (S.J.); xhjin@zjut.edu.cn (X.J.)
[3] Key Laboratory of Advanced Manufacturing Technology of Zhejiang Province, College of Mechanical Engineering, Zhejiang University, Hangzhou 310027, China; weiyd@zju.edu.cn
* Correspondence: guozi@zzjc.edu.cn

Abstract: Gearboxes are widely used in drive systems of rotating machinery. The health status of gearboxes considerably influences the normal and reliable operation of rotating machinery. When a gearbox experiences tooth failure, a vibration signal with impulse features is excited. However, these impulse features tend to be relatively weak and difficult to extract. To solve this problem, a novel approach for gearbox fault feature extraction and fault diagnosis based on improved variational mode extraction (VME) is proposed. Since the initial value of the desired mode center frequency and the value of the penalty parameter in VME must be assigned, a short-time Fourier transform (STFT) was performed, and a new index, the standard deviation of differential values of envelope maxima positions (SDE), is proposed. The feasibility and effectiveness of the proposed approach was verified by a simulation signal and two datasets associated with a gearbox test bench. The results demonstrate that the VME-based approach outperforms the variational mode decomposition (VMD) approach.

Keywords: gearbox; variational mode extraction; short-time Fourier transform; SDE index; fault feature extraction

Citation: Guo, Y.; Jiang, S.; Yang, Y.; Jin, X.; Wei, Y. Gearbox Fault Diagnosis Based on Improved Variational Mode Extraction. *Sensors* **2022**, *22*, 1779. https://doi.org/10.3390/s22051779

Academic Editors: Zahra Sharif Khodaei, Phong B. Dao, Lei Qiu and Liang Yu

Received: 29 January 2022
Accepted: 22 February 2022
Published: 24 February 2022

1. Introduction

Gearboxes are widely used in rotating machinery, such as wind turbines, nuclear power units, high-speed rail, and new energy vehicles, and their performance influences the normal and reliable operation of rotating machinery. Gearboxes usually operate under a dynamic load or in overload conditions and are prone to suffer from various kinds of defects, such as fatigue pitting, wear, tooth spalling, and tooth fracture [1–3]. To evaluate the operating status of rotating machinery and ensure safe production and effective maintenance, it is important to monitor the gearbox condition and diagnose faults. When a gearbox experiences tooth failure, the contact stiffness at the damage point changes, and the meshing movement is not smooth during operation, exciting impulsive vibration. When rotating machinery operates at a constant speed, impulsive vibration can be demonstrated by periodic impulse features in the sampled vibration signal [4,5]. The impulse feature frequency is closely related to the health status of the gearbox, which suggests that different frequencies indicate different fault states [6]. Therefore, extracting the fault feature frequency from the sampled vibration signal is a feasible solution for gearbox fault diagnosis [7]. However, gearboxes usually consist of multiple rotating parts. The vibrations excited by different parts are coupled to each other. In particular, the transient impulses caused by the damaged part attenuate along the transmission path from the source to the sensor, which is usually located on the casing. In addition, the sampled vibration signal often contains environmental noise [8]. Therefore, extracting the weak impulse features from vibration signals with heavy noise becomes a challenging task in practice [9–11].

Essentially, the process of fault feature extraction comprises the elimination of noise and interference components in vibration signals. An effective approach to solve this problem is signal decomposition, the variants of which include wavelet decomposition (WT) [12–14], empirical mode decomposition (EMD) [11,15,16], local mean decomposition (LMD) [9,10,17,18], and empirical wavelet transform [19–22]. However, WT is not a self-adaptive signal analysis method because it is restricted by the selection of the wavelet basis function and number of decomposition levels [17]. Although EMD can self-adaptively decompose a multimodulated signal into a series of intrinsic mode functions (IMFs), it lacks a theoretical basis and has some inherent defects, such as sensitivity to noise, end effects, and mode mixing [4,5,15,16]. Like EMD, LMD adaptively decomposes a multicomponent signal into several single-component AM–FM signals but encounters several technical problems, such as end effects and mode mixing [17,18]. EWT decomposes a signal into several AM–FM monocomponents that have compact support in the Fourier spectrum; however, the boundaries of the frequency partition are difficult to set, and this aspect may result in inaccurate and even invalid components [21,22].

In 2014, Dragomiretskiy and Zosso proposed a novel signal processing method known as variational mode decomposition (VMD) [23]. Based on complete mathematical principles, VMD decomposes a signal into an ensemble of band-limited subsignals, i.e., modes. The mode compact around a center frequency is estimated online in a nonrecursive manner. VMD is highly effective when applied to machinery fault diagnosis [5,24,25]. However, the actual performance of VMD is critically affected by the mode number and quadratic penalty parameter. Presetting the two parameters usually requires experience and experimentation. Moreover, it is difficult to determine optimal preset values. To address these limitations, many optimization algorithms, such as genetic optimization [26], particle swarm optimization [25,27], grasshopper optimization [24,28], gray wolf optimization [29,30], cuckoo search algorithm [31], artificial bee colony algorithm [32], and chaotic pigeon-inspired optimization [33], have been studied and applied to achieve the optimal values for the two parameters. Many parameter-optimized or -enhanced VMD methods have been proposed. However, these optimization methods require a large number of iterative operations and are time-consuming, resulting in a low computational efficiency [1], especially given the quadratic penalty parameter usually takes a large value, and its optimal value needs to be searched in a large range. Therefore, the enhanced VMD method combined with several optimization algorithms must be further studied and discussed to enhance its practicality.

In gearbox fault-diagnosis applications, the fault feature information is usually included in a certain frequency band of the sampled vibration signal. Therefore, among the obtained modes of VMD, only one invaluable target mode exists, and the other modes may not be as important. Selection of the target mode is challenging because some experience and knowledge of fault diagnosis is required, or several indicators, such as kurtosis [26], weighted kurtosis index [24,29,32], Shannon entropy [34], and sample entropy [35], must be considered.

In 2017, Nazari and Sakhaei proposed a new efficient method known as variational mode extraction (VME) to derive respiratory signals from ECG [36]. VME, which is based on a similar principle as that of VMD, directly decomposes a signal into two modes, i.e., the desired mode and residual signal. Therefore, VME resolves the difficulty of determining the mode number and selecting the target mode associated with VMD. Nevertheless, the application of VME necessitates the consideration of two key aspects. First, the initial center frequency of the desired mode must be correctly preset according to the analyzed signal. Otherwise, the desired mode may not contain valuable fault-related information. Second, the penalty parameter must be set to an appropriate value. If the penalty parameter value is excessively small, considerable noise and interference may remain in the desired mode. Moreover, the penalty parameter value must not be excessively large; otherwise, the desired mode may lose useful information.

Considering these aspects, this paper proposes an improved VME method, in which the initial value of the desired mode center frequency is selected by the STFT spectrum, and

the optimal value of the penalty parameter is determined using a novel index: the standard deviation of differential values of envelope maxima positions (SDE). The subsequent sections of this paper are organized as follows. The VME method is briefly introduced in Section 2. Section 3 describes the effect of the initial value of the desired mode center frequency and the penalty parameter value on the performance of VME, as well as the methods to determine these parameters. The improved VME method for gearbox fault diagnosis is summarized in Section 4. Two cases of gearbox fault vibration signal analysis based on the improved VME method are described in Section 5. Concluding remarks are presented in Section 6.

2. Variational Mode Extraction

2.1. Basic Theory

The basic principle of VME is similar to that of VMD; however, VME explicitly decomposes an original signal, $x(t)$, into two modes: the desired mode, $u_\mathrm{d}(t)$, and residual signal, $x_\mathrm{r}(t)$, as indicated in Equation (1):

$$x(t) = u_\mathrm{d}(t) + x_\mathrm{r}(t) \tag{1}$$

where $u_\mathrm{d}(t)$ is a compact function around the center frequency, ω_d. In the application of rotating machinery fault diagnosis, $u_\mathrm{d}(t)$ denotes the fault feature component. To extract $u_\mathrm{d}(t)$ from $x(t)$, the spectral overlap between the two components must be minimized. In other words, $x_\mathrm{r}(t)$ has the minimal energy in the frequency band of $u_\mathrm{d}(t)$, particularly zero energy in the center frequency, ω_d of $u_\mathrm{d}(t)$. Moreover, the sum of $u_\mathrm{d}(t)$ and $x_\mathrm{r}(t)$ must completely reconstruct the original signal, $x(t)$. Therefore, the mathematical model of the VME method can be expressed by the constrained minimization problem in Equation (2):

$$\begin{aligned} \min_{u_\mathrm{d},\omega_\mathrm{d},f_\mathrm{r}} \quad & \left\{ \alpha \left\| \partial_t \left[\left(\delta(t) + \tfrac{\mathrm{j}}{\pi t} \right) * u_\mathrm{d}(t) \right] \mathrm{e}^{-\mathrm{j}\omega_\mathrm{d} t} \right\|_2^2 + \| \beta(t) * x_\mathrm{r}(t) \|_2^2 \right\} \\ \text{s.t.} \quad & u_\mathrm{d}(t) + x_\mathrm{r}(t) = x(t) \end{aligned} \tag{2}$$

where ∂_t denotes the partial derivative with time, t; $\delta(t)$ is the Dirac delta distribution; and * is the convolution operator. $\beta(t)$ represents the impulse response of the filter, which is used to filter the frequencies in $x_\mathrm{r}(t)$ overlapping with $u_\mathrm{d}(t)$ and can be described as in Equation (3):

$$\hat{\beta}(\omega) = \frac{1}{\alpha(\omega - \omega_\mathrm{d})^2} \tag{3}$$

Combining Equations (2) and (3), it can be observed that the penalty parameter, α, plays a critical role in controlling the balance between the compactness of $u_\mathrm{d}(t)$ and spectral overlap between $u_\mathrm{d}(t)$ and $x_\mathrm{r}(t)$.

To solve the constrained minimization problem, the Lagrangian multiplier, $\lambda(t)$, is introduced. Hence, the constrained minimization problem defined in Equation (2) can be converted into an unconstrained optimization problem, as expressed in Equation (4):

$$\begin{aligned} L(u_\mathrm{d}, \omega_\mathrm{d}, \lambda) \quad & = \alpha \left\| \partial_t \left[\left(\delta(t) + \tfrac{\mathrm{j}}{\pi t} \right) * u_\mathrm{d}(t) \right] \mathrm{e}^{-\mathrm{j}\omega_\mathrm{d} t} \right\|_2^2 + \| \beta(t) * x_\mathrm{r}(t) \|_2^2 \\ & + \| x(t) - (u_\mathrm{d}(t) + x_\mathrm{r}(t)) \|_2^2 + \langle \lambda(t), x(t) - (u_\mathrm{d}(t) + x_\mathrm{r}(t)) \rangle \end{aligned} \tag{4}$$

According to Parseval's theorem, the optimization problem presented in Equation (4) can be rewritten in the frequency domain as in Equation (5):

$$\begin{aligned} L(u_\mathrm{d}, \omega_\mathrm{d}, \lambda) \quad & = \alpha \| \mathrm{j}(\omega - \omega_\mathrm{d}) [(1 + \mathrm{sgn}(\omega)) \cdot \hat{u}_\mathrm{d}(\omega)] \|_2^2 + \left\| \hat{\beta}(\omega) \cdot \hat{x}_\mathrm{r}(\omega) \right\|_2^2 \\ & + \| \hat{x}(\omega) - (\hat{u}_\mathrm{d}(\omega) + \hat{x}_\mathrm{r}(\omega)) \|_2^2 + \left\langle \hat{\lambda}(\omega), \hat{x}(\omega) - (\hat{u}_\mathrm{d}(\omega) + \hat{x}_\mathrm{r}(\omega)) \right\rangle \end{aligned} \tag{5}$$

The alternating direction method of multipliers (ADMM) is adopted to solve Equation (5) to identify the saddle point of the augmented Lagrangian expression. As the

optimization problem is converted into a series of iterative suboptimization problems solved by ADMM, the updating equations, in terms of $\hat{u}_d$, ω_d, and $\hat{\lambda}$ are expressed as Equations (6)–(8), respectively:

$$\hat{u}_d^{n+1}(\omega) = \frac{\hat{x}(\omega) + \alpha^2(\omega - \omega_d^{n+1})^4 \cdot \hat{u}_d^n(\omega) + \hat{\lambda}(\omega)/2}{\left[1 + \alpha^2(\omega - \omega_d^{n+1})^4\right]\left[1 + 2\alpha(\omega - \omega_d^n)^2\right]} \tag{6}$$

$$\omega_d^{n+1} = \frac{\int_0^\infty \omega \left|\hat{u}_d^{n+1}(\omega)\right|^2 d\omega}{\int_0^\infty \left|\hat{u}_d^{n+1}(\omega)\right|^2 d\omega} \tag{7}$$

$$\hat{\lambda}^{n+1} = \hat{\lambda}^n + \tau \cdot \left[\frac{\hat{x}(\omega) - \hat{u}_d^{n+1}(\omega)}{1 + \alpha^2(\omega - \omega_d^{n+1})^4}\right] \tag{8}$$

where $\hat{u}_d^n(\omega)$ is the obtained desired mode in the nth iteration with the center frequency ω_d^n, $\hat{x}(\omega)$ represents the original signal Fourier transforms, n is the number of iterations, and τ is the iteration step length. The steps involved in VME are as follows [36].

Step 1: initialize $\hat{u}_d^1$, ω_d^1, $\hat{\lambda}^1$, and n = 1.

Step 2: update $\hat{u}_d$ and ω_d according to Equations (6) and (7), respectively.

Step 3: update $\hat{\lambda}$ according to Equation (8).

Step 4: $n = n + 1$; repeat Steps 2–3 until the iteration termination condition,

$$\frac{\left\|\hat{u}_d^{n+1} - \hat{u}_d^n\right\|_2^2}{\left\|\hat{u}_d^n\right\|_2^2} < \varepsilon, \tag{9}$$

is satisfied.

Step 5: obtain the desire mode, $u_d(t)$.

2.2. Analysis of Parameter Influence

In the implementation of VME, the update step size, τ, is often set as zero to ensure that the algorithm converges effectively [36]. The value of the convergence tolerance, ε, determines the reconstruction accuracy of the VME decomposition; therefore, this parameter is usually set as an extremely small positive value (e.g., 1×10^{-6} in our experiments). The performance of VME is significantly dependent on the decomposition parameters, such as the initial center frequency, ω_d, of the desired mode and penalty parameter α, which should be preset in a restrained manner. The effects of these parameters on VME performance are analyzed below.

(1) Initial center frequency, ω_d, of the desired mode. The frequency band position of the desired mode extracted from the original signal is determined by the initial value of ω_d. If the initial value of ω_d is inappropriately initialized, the desired mode likely does not contain valuable information. Although VME is not highly sensitive to the initial ω_d value, which can be preset in a wide range [36], the approximate range of the initial ω_d value must be reasonably determined.

(2) Penalty parameter α. As this parameter controls the compactness of the desired mode, it also determines the degree of spectral overlap between the desired mode and residual signal. When α is extremely small, the bandwidth of the desired mode is extremely large; therefore, lots of interference components or additional noise may be included in the desired mode, which can hinder the identification of valuable information. Normally, parameter α is necessarily set to a large value to ensure that the detected center frequency is closely related to the desired mode [36]. However, when α is set to an extremely large value, the bandwidth of the desired mode is extremely small, and part of the useful information may be lost, especially when the

center frequency, ω_d, is not appropriately set. Consequently, the penalty parameter, α, must be preset to an appropriate value.

3. Simulation Analysis and Discussion

3.1. Simulation Signal Construction

This section describes the construction of a simulation vibration signal to evaluate the effects of the initial center frequency, ω_d, and penalty parameter, α, on the fault feature extraction performance of VME. Considering these effects, methods to determine the two parameters are established.

Consider a pair of engaged gears, in which one of the gears is assumed to have a local damage fault. When the gears are operating smoothly, the fault vibration signal, $s(t)$, can be considered to be composed of three main components, including the multiple AM–FM components, $s_m(t)$, induced by gear meshing and shaft rotation; a periodic impulsive component, $s_p(t)$, excited by gear local damage; and additive noise, $s_n(t)$. The mathematical model of the fault vibration signal, $s(t)$, can be presented as:

$$s(t) = s_m(t) + s_p(t) + s_n(t) \tag{10}$$

$$s_m(t) = \sum_{n=1}^{N} X(n)[1 + a_n(t)]\cos[2\pi n Z f_r t + \varphi_n + b_n(t)] \tag{11}$$

$$s_p(t) = \sum_{k=1}^{K} P(k)e^{-\zeta(t-kT)}\cos[2\pi Z f_r(t-kT)] \tag{12}$$

where Z and f_r are the number of teeth and rotation frequency of the damaged gear, respectively; $Z{\cdot}f_r$ denotes the meshing frequency of the gears, represented by f_m, i.e., $f_m = Z{\cdot}f_r$; N is the number of AM–FM components; $X(n)$, φ_n, $a_n(t)$, and $b_n(t)$ denote the amplitude, initial phase, amplitude-modulation function, and phase-modulation function of the nth AM-FM component, respectively; and K, ζ, and T denote the number, amplitude-damping coefficient, and period of impulses, respectively. When the damage gear rotates once, the fault impulse feature is triggered once; thus, $T = 1/f_r$, and the fault feature frequency, $f_{3f} = f_r$. $P(k)$ is the amplitude of the kth impulse. Furthermore, the modulation functions, $a_n(t)$ and $b_n(t)$, can be described as:

$$\begin{cases} a_n(t) = A_n\cos(2\pi n f_r t + \alpha_n) \\ b_n(t) = B_n\cos(2\pi n f_r t + \beta_n) \end{cases} \tag{13}$$

where A_n and α_n denote the amplitude and phase of the amplitude-modulation function, respectively, and B_n and β_n denote the amplitude and phase of the phase-modulation function, respectively. The fault vibration signal of this pair of gears can be simulated by setting the appropriate values for the relevant parameters. The values of all parameters in the simulated signal are listed in Table 1.

Table 1. Parameters of the simulation signal model.

Z	f_r	f_m	N	$X(n)$	φ_n	A_n	α_n	B_n	β_n	ς	T
25	20 Hz	500 Hz	1	2	π	0.25	π	0.5	π	60π	0.05 s

The sampling frequency, f_s, is set as 10 k Hz, and the sampling time, t_s, is set as 0.4 s. Therefore, the number of sampling points is 4000. In the sampling period, the first impulse feature is assumed to appear at $t_1 = 0.025$ s, and the number of impulses is $K = \lfloor(t_s - t_1){\cdot}f_r\rfloor$, where $\lfloor\cdot\rfloor$ denotes the floor function. The amplitude sequences, $P(k)$ (k=1,2, … , 8), for the impulses are obtained from a Gaussian distribution with a mean of 1.3 and standard deviation of 0.2. White Gaussian noise is introduced such that the signal-to-noise ratio (SNR) of the simulated signal is −8 dB. The time-domain waveform of the simulation signal, $s(t)$, is shown in Figure 1a. As the impulsive component is almost buried in the

strong background noise, it is difficult to detect the fault by observing the amplitude of the simulation signal. The square envelope spectrum (SES) [37] of the simulation signal, $s(t)$, is shown in Figure 1b, in which the Y-axis is the normalized amplitude obtained by dividing each amplitude by the maximum amplitude. From the SES, the fault feature frequency can be extracted, although it is easily disturbed by other frequency contents with large amplitudes. VME is applied to extract the weak fault features from the simulation signal, and further analysis is performed.

(**a**) (**b**)

Figure 1. The simulation signal: (**a**) time–domain waveform and (**b**) SES.

3.2. Initial Value Estimation of ω_d

As mentioned, the fault feature extraction effect of VME is highly dependent on the initial value of the desired mode center frequency, ω_{d}. An appropriate initial value of ω_{d} can help ensure that the desired mode contains complete and relatively pure fault features. In this study, the initial value of ω_{d} is selected using a short-time Fourier transform (STFT).

STFT is a classical method for joint time–frequency analysis that is suitable for not only stationary signals but also time-varying, nonstationary signals. The basic idea of STFT is to multiply the original signal by a fixed-length window function, assuming that the signal section in the short time interval of the analysis window is stationary, and to implement a Fourier transform (FT) for the signal section to obtain a local spectrum. By translating the analysis window along the entire time horizon, the original signal is analyzed by FT in a section-by-section manner, and a set of local spectra is obtained during each time interval. Therefore, the basic calculation formula of STFT can be expressed as:

$$STFT\{x(t)\}(t,f) = \int_{-\infty}^{+\infty} x(\tau)g(\tau - t)\mathrm{e}^{-\mathrm{j}2\pi f\tau}\mathrm{d}\tau \tag{14}$$

where $x(t)$ is the original signal, t and τ denote the time, f denotes the frequency, and $g(\tau - t)$ is a window function with the center at time t. For a given time, t, $STFT(t,f)$ can be considered to be the FT spectrum. In particular, when the window function is set as 1, the STFT degenerates into the traditional FT.

To obtain relatively optimal localization performance, the type and width of the window function must be selected appropriately according to the signal characteristics. In this study, considering the exponential attenuation tendency of the transient impulse feature, the Gaussian window function is applied, and its length is set as 65 for a high time resolution. The time–frequency spectrum of the simulated signal by STFT is shown in Figure 2. Because of the heavy noise, the morphology and periodicity of the impulses undergo considerable interference. Nevertheless, it can be seen that the impulses repeatedly appear in a narrow frequency range centered at 500 Hz along the time axis. Therefore, the initial center frequency, ω_{d}, is set as $2\pi{\cdot}500$ rad/s, which can also be considered an optimal value.

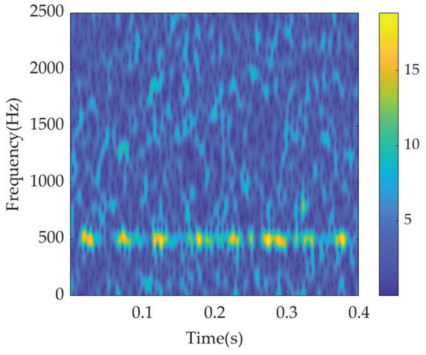


Figure 2. STFT spectrum of the simulation signal.

Because a reliable basis for the selection of the penalty parameter, α, is not available, the α value is set as 10,000, 80,000 and 150,000. Experimental results indicate that it is reasonable to set α as 80,000 because the desired mode obtained using this value has the fewest interference components, and the fault feature frequency can be successfully extracted. When α is extremely small, the spectral overlap between the desired mode and residual signal is large; therefore, the desired mode may contain many noise components. When α is extremely large, the desired mode exhibits a high compactness. Thus, useful components may be lost, and the signal may degenerate into a harmonic signal with a frequency of ω_{d}, from which the fault feature frequency cannot be extracted. Thus, an optimal value must be set for the penalty parameter, α, which requires further study.

3.3. SDE Index

To investigate the effect of parameter settings on VME performance, it is critical to construct an accurate objective function. The desired mode, $u_{\mathrm{d}}(t)$, theoretically obtained by VME contains a series of fault impulse features with as complete a periodicity as possible and as little interference or noise as possible. To assess the periodicity and purity of the desired mode, this paper proposes a new index called the standard deviation of differential values of envelope maxima positions (SDE), which can be determined using the following steps.

Step 1: Demodulate the desired mode, $u_{\mathrm{d}}(t)$, using the Hilbert transform to obtain the upper envelope, $ue_{\mathrm{d}}(t)$.

Step 2: Find the positions of the local maximum points in $ue_{\mathrm{d}}(t)$ that are greater than a certain threshold, m_{th}, expressed as:

$$m_{\mathrm{th}} = \eta \cdot \{\max[ue_{\mathrm{d}}(t)] - \min[ue_{\mathrm{d}}(t)]\} \tag{15}$$

where η is the threshold coefficient and, in this paper, all the η values are set as 0.25. These positions of the target points are denoted by p_i, $i = 1, 2, \ldots, N$, where N is the number of local maximum points.

Step 3: Calculate the differential values, d_j, of series p_n,

$$d_{\mathrm{j}} = p_{j+1} - p_j, j = 1, 2, \cdots, N-1. \tag{16}$$

Step 4: Calculate the standard deviation of series d_j, i.e., SDE, and use it as the indicator of the periodic fault impulse features.

$$SDE = \sqrt{\frac{1}{N-1}\sum_{j=1}^{N-1}\left(d_j - \overline{d}\right)^2} \tag{17}$$

where $\overline{d}$ is the mean value of d_j.

Ideally, regardless of the amplitude differences between the fault impulses, the positions of the maximum amplitudes are uniform due to periodicity, and the calculated SDE index is zero. In practice, the fault impulse features in vibration signals may not have strict periodicity, causing the position intervals of the fault impulses to fluctuate slightly and randomly. Nevertheless, the calculated SDE index is expected to be extremely small due to the randomness of fault impulse occurrence positions. According to the calculation process described above, in an ideal situation, the value of the SDE index is mainly determined by the maximum value of each fault impulse feature. Therefore, the SDE index is adaptive to the periodicity of fault impulse features and can clarify the purity of the fault-related information in the desired mode obtained by VME.

3.4. Effect of α

To study the effect of the penalty parameter, α, on VME performance and to identify a way to select an optimal value for α, the initial center frequency, ω_d, is set as a fixed value of $2\pi \cdot 500$ rad/s, and the value of α increases gradually from α_{min} to α_{max} at a step size of s_α. At the nth step, the value, α_n, of the penalty parameter is calculated as:

$$\alpha_n = \alpha_{min} + (n-1) \cdot s_\alpha, n = 1, 2, \cdots, (\alpha_{max} - \alpha_{min})/s_\alpha \quad (18)$$

For the simulation signal, α_{min}, α_{min}, and s_α are set as 500, 150,000, and 500, respectively. VME is implemented with a fixed ω_d value of $2\pi \cdot 500$ rad/s and different α values of α_n, n=1, 2, $\cdots$, 299. The SDE index value for each desired mode is calculated. The curve of the SDE index with different penalty parameter, α, values is shown in Figure 3.

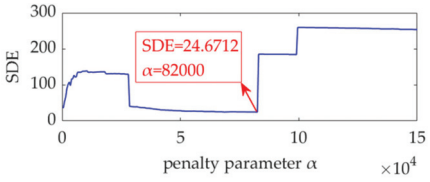


Figure 3. Relationship between the SDE index and penalty parameter α.

In the initial stage of the curve, the SDE index acquires a small value. The main reason for this is that as the value of α is extremely small, the obtained desired mode still contains considerable noise, and the position distribution of the desired mode peaks has strong randomness and uniformity. As the value of α continues to increase, the noise in the obtained desired mode is gradually suppressed, the randomness and uniformity of its peak distribution are broken, and the resulted values of the SDE index also increase. When the noise in the desired mode is eliminated, the SDE index reaches its minimal value due to the periodicity of fault impulse features. However, if the value of α is further enhanced, the fault impulse features in the obtained desired mode will be suppressed, especially those with small amplitudes, and the periodicity of the fault impulse features will be disturbed, leading to nonuniformity of the peaks in the desired mode and a large value of the SDE index. Such an analysis shows that when the SDE index reaches the minimum value of 24.6712, the optimal value of α is obtained, expressed as α_{opt} = 82,000. VME achieves the optimal effect for fault impulse feature extraction. In other words, the obtained desired mode exhibits minimal interference with the noise and exhibits the most significant periodicity. The desired mode, along with its STFT spectrum, and the extracted fault feature frequency using SES are shown in Figure 4.

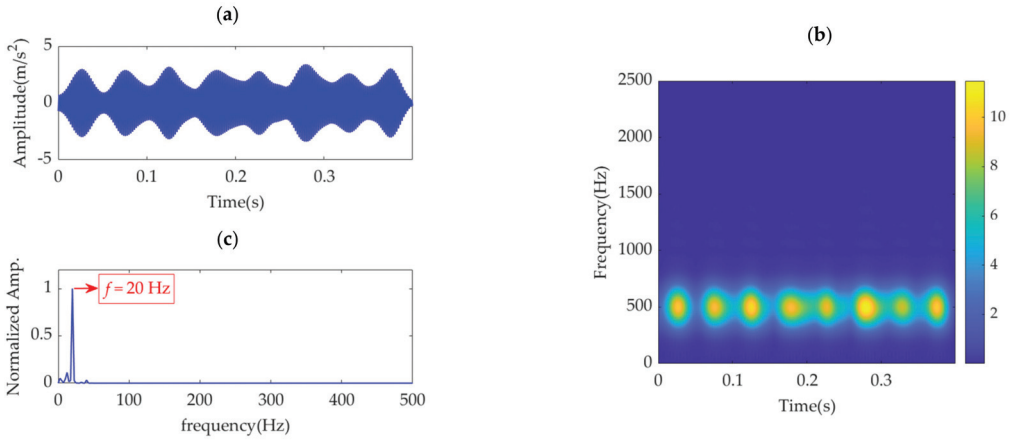


Figure 4. Desired mode obtained by VME with α_{opt} = 82,000: (**a**) time−domain waveform, (**b**) STFT spectrum, and (**c**) SES.

In addition, Figure 3 shows that the values of the SDE index vary slightly around α_{opt}. Within this varying interval, some other values of the SDE index are selected, e.g., *SDE* = 27.3522 and *SDE* = 24.8577, corresponding to α = 52,000 and α = 68,500, respectively. The numerical experimental results show that the effects of α = 52,000 and α = 68,500 on the VME performance are basically identical to those of α_{opt} = 82,000. Therefore, although there exists an optimal value for α, the performance of VME is not highly sensitive to α as long as the α value is selected within a reasonable range, which can be easily estimated using the α–SDE index curve.

3.5. Effect of Initial ω_d

The influence of the penalty parameter, α, on VME performance is examined with the initial center frequency ω_d set as 2π·500 rad/s, as obtained using the STFT. Moreover, the optimal value of α is α_{opt} = 82,000. Thus, α is set as a fixed value of 82,000, and ω_d increases from ω_{dmin} to ω_{dmax}, with a step size of s_ω. At the nth step, the value of ω_{dn} is calculated as:

$$\omega_{dn} = \omega_{dmin} + (n-1) \cdot s_\omega, n = 1, 2, \cdots, (\omega_{dmax} - \omega_{dmin})/s_\alpha \quad (19)$$

As shown in Figure 2, periodic fault impulse features appear in a frequency band with the center frequency, f_d = 500 Hz. To include this frequency band, ω_{dmin}, ω_{dmax}, and s_ω are set as 2π·300 rad/s, 2π·700 rad/s, and 2π rad/s, respectively. VME is implemented with initial center frequency, ω_{dn} (n = 1, 2, . . . , 401), to achieve the desired mode, and the corresponding SDE index value is calculated. The curve of the SDE index with ω_{dn} is shown in Figure 5.

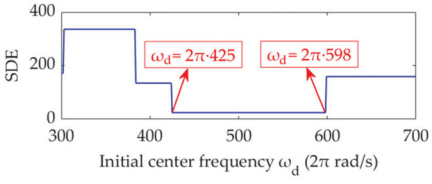


Figure 5. Relationship between the initial center frequency, ω_d, and SDE index.

The SDE index is minimized in the ω_d range of $2\pi \cdot 425$ to $2\pi \cdot 598$ rad/s. In this range of ω_d, VME exhibits the same performance for periodic fault impulse feature extraction. In other words, as long as the ω_d value is estimated within an appropriate range, the performance of VME is not sensitive to ω_d. In practice, the targeted value range of ω_d can be clarified using the STFT spectrum, and a reasonable or optimal ω_d value can be easily obtained.

3.6. Comparative Study between the SDE Index and Other Indices

Many indicators have been proposed for impulse detection and fault feature extraction from rotating machinery vibration signals. Examples of widely used indicators are kurtosis [1]; the weighted kurtosis index (also known as KCI) [24,29,32,38]; envelope kurtosis (EK) [39]; the Gini index (GI) [40]; and entropy-class indices, such as information entropy (IE, also known as Shannon entropy) [26,34], sample entropy (SE) [35,41], permutation entropy (PE) [15], power spectral entropy (PSE) [42], envelope entropy (EE) [30], and envelope spectrum entropy (ESE) [43]. To verify the superiority of the proposed SDE index, the values of these 10 indices are calculated for the desired modes obtained by VME with different penalty parameters, α_n ($n = 1, 2, \ldots, 300$). The relationship between these 10 indices and penalty parameter α are shown in Figure 6.

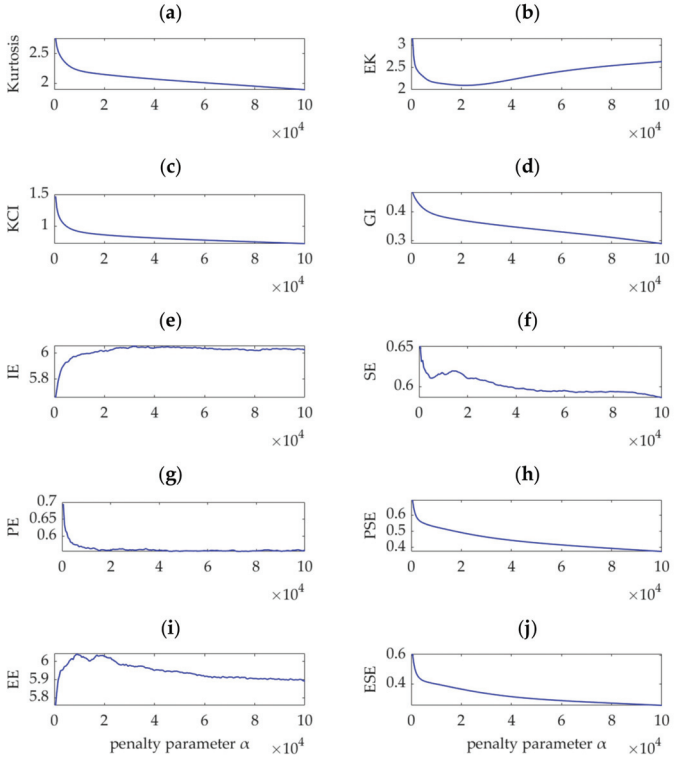


Figure 6. Relationship between different indices and penalty parameter α.

VME is equivalent to the implementation of a bandpass filter around the center frequency, ω_d, and the penalty parameter, α, determines the passband width of the filter. Theoretically, when α is small, considerable noise and other interference components remain in the desired mode. As α increases and the compactness of the desired mode intensifies, the noise and interference components are gradually depressed. When α increases to a

certain value, the optimal fault impulse features are obtained. As α further increases, the fault-related features are damped accordingly, and the desired mode tends to be a single-component signal in which the fault-related information may be completely lost. Therefore, an effective index must achieve a significant extremum or sudden change in the search range of α to ensure that it can indicate the possible optimal value of α. Considering this aspect, the indices of kurtosis, KCI, GI, PE, PSE, and ESE cannot indicate the possible optimal α value due to their monotonic changes, as shown in Figure 6a,c,d,g,h,j. Nevertheless, in the search range of α, the EK index is minimized, as shown in Figure 6b, and the corresponding value of α is 21,500. In contrast, the IE and EE indices are maximized, as shown in Figure 6e,i, and the corresponding values of α are 32,500 and 9000, respectively. In addition, Figure 6f shows that the SE index obtains a local maximum value, which indicates α to be 14,000. Therefore, the optimal values of α are respectively assumed to be 9000 and 32,500, and the obtained desired modes, together with their corresponding square envelope spectra, are shown in Figures 7 and 8, respectively. Comparing Figures 8 and 9 with Figure 5, although the differences between the performances of VME with different optimal α values are not significant, the desired mode obtained through VME with the proposed SDE index has fewer interference components. This outcome verifies the robustness of VME in response to penalty parameter α and the superiority of the SDE index compared to other indices in this application.

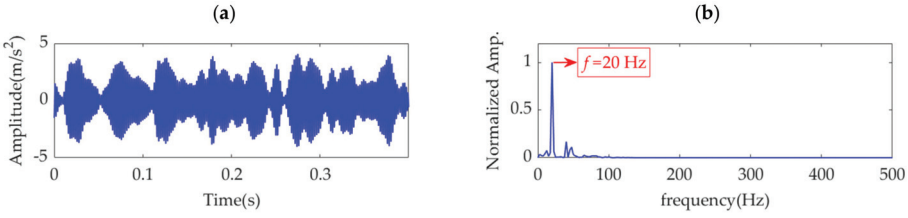


Figure 7. Desired mode obtained by VME with α = 9000: (**a**) time–domain waveform and (**b**) SES.

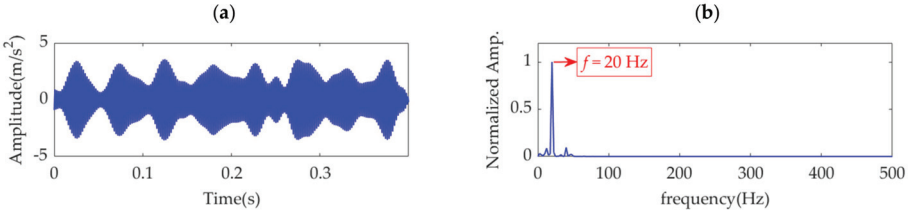


Figure 8. Desired mode obtained by VME with α = 32,500: (**a**) time–domain waveform and (**b**) SES.

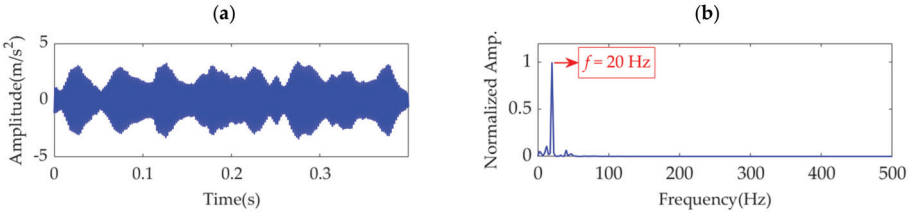


Figure 9. Target mode obtained by VMD: (**a**) time–domain waveform and (**b**) SES.

3.7. Comparative Study between VME and VMD

To further evaluate the superiority of VME, the simulation vibration signal is decomposed using VMD. To ensure a fair comparison, the value of penalty parameter α_d in VMD

is the same as the previously obtained optimal value of penalty parameter α in VME, i.e., α_d = 82,000. After several trials, the number of modes in VMD is determinately set as 8. In the VMD results, the target mode that contains fault impulse features is the second mode, as shown in Figure 9a. The SES of the target mode is shown in Figure 9b. Comparison of Figures 4c and 9b shows that the desired mode obtained by VME is similar to the target mode obtained by VMD, although the fault feature frequency can be successfully extracted from either of the two modes. However, compared with VME, the VMD method exhibits several deficiencies:

(1) The exact number of modes is not easy to determine, and the selection of the target mode with valuable fault-related information is challenging to perform simultaneously. These two issues can be solved through trials or an optimization method; however, the execution efficiency of VMD may decrease.
(2) Considerable noise is present in the target mode of VMD, and less noise is present in the desired mode of VME. This phenomenon likely occurs because there exist certain overlaps between the target mode and its previous and latter modes. Therefore, a small amount of noise associated with the previous and latter modes remains in the target mode.

As VME decomposes a signal into two layers, i.e., the desired mode and residual signal, the problems of decomposed mode number determination and target mode selection associated with VMD can be avoided. In addition, the penalty parameter, α, in VME only adjusts the spectrum overlap degree between the desired mode and residual signal. In contrast, the penalty parameter, α_d, in VMD must balance the spectrum overlap between all adjacent modes. Therefore, typically, the desired mode obtained by VME is purer than the target mode obtained by VMD.

4. Improved VME Method for Gearbox Fault Diagnosis

Based on the previous analysis, this paper proposes an improved VME method to extract the fault-related features from vibration signals for gearbox fault diagnosis. The fault diagnosis process is shown in Figure 10. The main implementation steps are described as follows.

Step 1: The time–frequency spectrum of the vibration signal sampled from the gearbox is obtained using STFT to detect the approximate frequency band location of fault-related features. Accordingly, the initial value of the center frequency, ω_d, is selected and considered the optimal value.

Step 2: The value range of penalty parameter α is set as [α_{min}, α_{max}], and the increasing step size is set as s_α. Subsequently, α is increased from α_{min}. In the nth step, the α value is expressed as α_n.

Step 3: VME is implemented with parameter group [ω_d, α_n] in each step for the original vibration signal to obtain the desired mode. The corresponding SDE index value is calculated.

Step 4: Once α increases to α_{max} and the last SDE value is calculated, the α–SDE curve is plotted. Using this curve, the optimal α value is selected as α_{opt}, corresponding to the minimal SDE index value.

Step 5: VME is implemented with the optimal parameter group [ω_d, α_{opt}] for the original vibration signal to obtain the optimal desired mode.

Step 6: SES analysis is performed for the optimal desired mode to extract the fault feature frequency and diagnose gearbox faults.

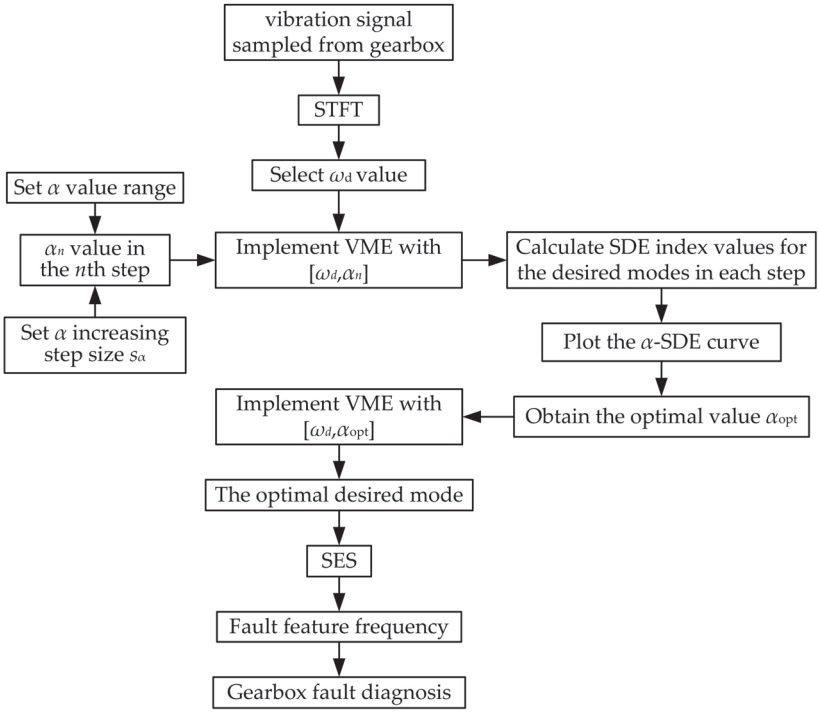


Figure 10. Process flow of gearbox fault diagnosis based on improved VME.

5. Experimental Evaluation

As described in this section, two vibration datasets of gearboxes obtained using a wind turbine gearbox simulation test bench are sampled and considered to verify the effectiveness and feasibility of the improved VME method for gearbox fault diagnosis.

5.1. Gearbox Test Bench

The gearbox test bench is constructed to simulate the function of the wind turbine gearbox and evaluate the gearbox fault diagnosis methods, shown in Figure 11. The mechanical components in the test bench include the driving motor, cycloidal-pin planetary gear speed reducer, double-row conic roller bearing, two-stage speed-increase planetary gearbox, one-stage speed-increase parallel shaft gearbox, torque and speed sensors ($\pm$30 Nm torque measurement range, 1024 line incremental encoder, and $\pm$6000 rpm standard speed measurement range), and load motor. The combination of the driving motor and cycloidal-pin planetary gear speed reducer produces a low speed to simulate the wind turbine main shaft speed. The combination of a two-stage speed-increase planetary gearbox and a one-stage speed-increase parallel shaft gearbox is used to imitate the wind turbine gearbox with the widely used configuration involving two-stage planetary and one-stage parallel shaft gear transmission. In this study, the test object is the parallel shaft gearbox, which simulates the high-speed stage prone to failure in a wind turbine gearbox. The gear ratio, gear modulus, number of gear teeth, and number of pinion teeth of the parallel shaft gearbox are 2.45, 2, 54, and 22, respectively.

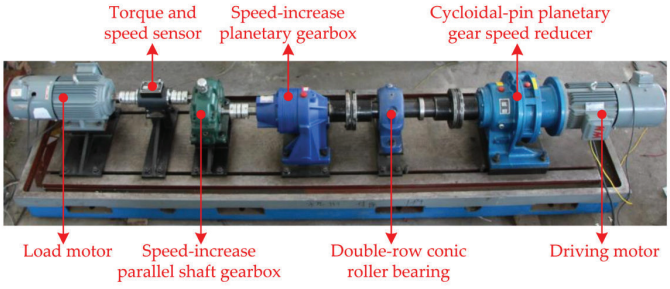


Figure 11. Gearbox test bench.

Two types of gear damages, i.e., pinion-tooth partial fracture and gear-tooth fracture, are manually induced in the experimental gearbox, as shown in Figure 12. Subsequently, operational tests are conducted over the test bench for each fault state. In each test, the load motor is set to operate in the power-generation state and output a torque of 3 Nm. An IEPE (integrated-electronics piezoelectric) accelerometer (±50 g full scale range, 100 mV/g sensitivity, and 0.2–5000 Hz frequency range) is mounted on the experimental gearbox top cover using a suitable magnetic mount to acquire the vibration signal. The sampling frequency of vibration signal is set as 5.12 kHz.

Figure 12. (**a**) Pinion-tooth partial fracture fault and (**b**) gear-tooth fracture fault.

5.2. Pinion Fault Vibration Dataset Analysis

In the fault state of the pinion-tooth partial fracture, the driving motor rotation speed is set as 700 rpm. The experimental gearbox output rotation speed is 1347 rpm, as measured by the sensor. Therefore, the theoretical fault feature frequency of the pinion in the experimental gearbox is calculated as f_p = 22.45 Hz. The vibration dataset with 4000 points sampled from the experimental gearbox is shown in Figure 13.

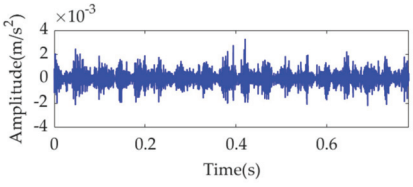


Figure 13. Sampled vibration dataset in the pinion fault state.

Because the fault impulse features exhibit significant periodicity in the time-domain waveform, analysis of this dataset may not be fruitful. To more effectively reflect the superiority of the proposed method, white Gaussian noise is added to the dataset to ensure that the signal-to-noise ratio (SNR) is −7 dB. The obtained dataset with heavy noise is shown in Figure 14a. Due to the low SNR, it is challenging to identify the impulse features in the time-domain waveform. Figure 14b shows the SES of the obtained dataset. Although a feature frequency of 21.76 Hz can be extracted using the SES, many interference components and noise are included in the SES. The STFT spectrum of the noisy dataset is

shown in Figure 14c. It is challenging to identify the impulse features in the waveform. In the STFT spectrum, a few impulse features can be observed, but these features do not exhibit any regularity. Nevertheless, the strongest impulse feature appears at a position of approximately [0.4715 s, 1490 Hz], as shown in the red box of Figure 14c. Hence, the initial center frequency, ω_d, of the desired mode is accordingly selected as $2\pi \cdot 1490$ rad/s. The penalty parameter, α, is set to increase from 100 to 30,000, with a step size of 100. Subsequently, in each step, VME is implemented with a fixed ω_d value of $2\pi \cdot 1490$ rad/s and an increased α value pertaining to the current step. Moreover, the corresponding SDE index value of the desired mode is calculated. The relationship between the penalty parameter, α, and SDE index is represented as a curve in Figure 15.

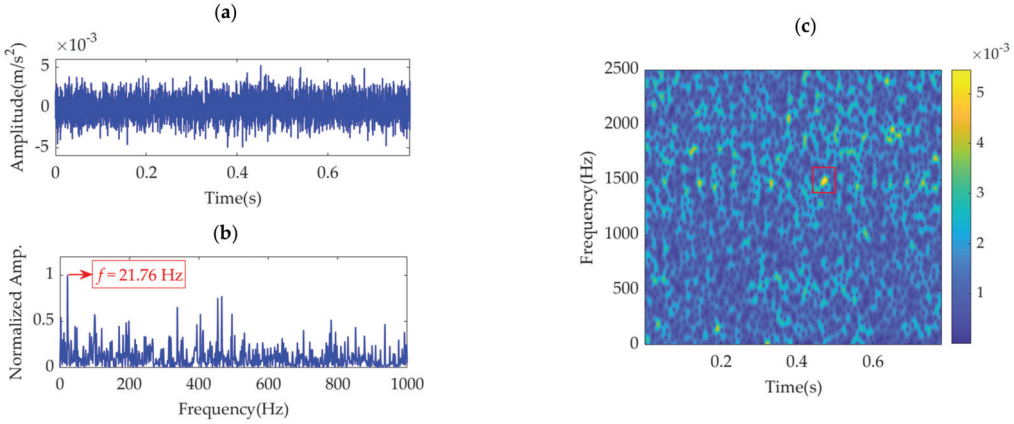


Figure 14. Vibration dataset with additional noise: (**a**) time–domain waveform, (**b**) SES, and (**c**) STFT spectrum.

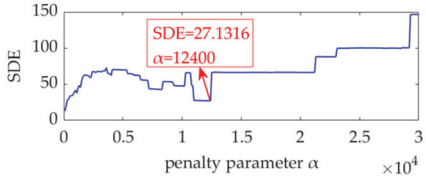


Figure 15. Relationship between the penalty parameter, α, and SDE index.

Notably, the α value of 100 corresponding to the SDE index minimum value of 13.22 is not the desired value for VME, as such a small SDE value is caused by strong random noise in the obtained desired mode. In this scenario, the optimal α value must be considered the value associated with the minimum SDE index value of 27.1316; hence, α_{opt} = 12,400. Accordingly, VME is implemented with an initial ω_d value of $2\pi \cdot 1490$ Hz and an optimal α value of 12,400. The obtained desired mode and its STFT spectrum are shown in Figure 16a,b, respectively. The noise in the desired mode is substantially eliminated, and regular impulse features can be clearly observed. The SES of the desired mode is shown in Figure 16c. The feature frequency of 21.76 Hz is successfully extracted, consistent with the pinion fault feature frequency, f_p.

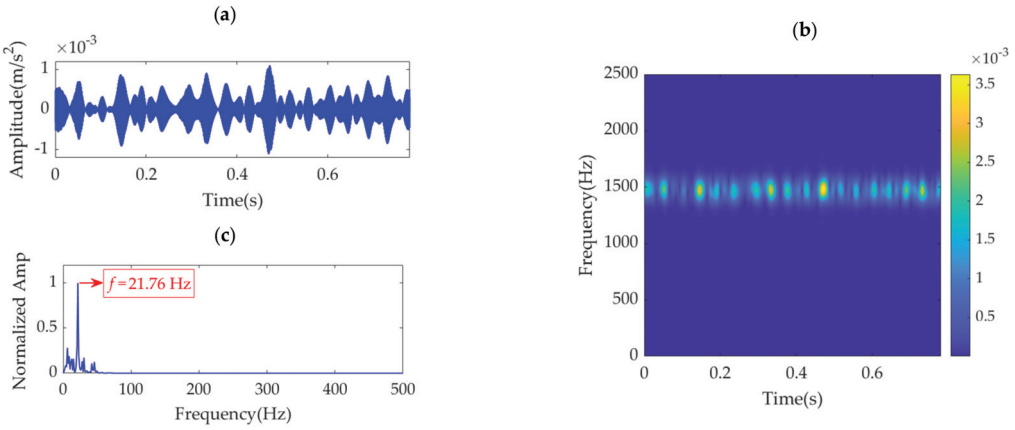


Figure 16. Desired mode obtained by VME with the optimal parameters: (**a**) time–domain waveform, (**b**) STFT spectrum, and (**c**) SES.

Furthermore, the penalty parameter, α, is set as the fixed value of α_{opt} = 12,400. Subsequently, the initial center frequency, ω_d, is set to gradually increase from $2\pi\cdot1200$ rad/s to $2\pi\cdot1600$ rad/s in step sizes of 2π rad/s. The SDE index value of the desired mode in each step is calculated. Based on these values, the relationship curve between ω_d and the SDE index is obtained, as shown in Figure 17. The findings indicate that the initial ω_d value, $2\pi\cdot1490$ rad/s, selected from the STFT spectrum is the optimal value that results in a minimum SDE index value 27.1316. Notably, there exist more than one optimal initial ω_d value that minimizes the SDE index in the case of α_{opt} = 12,400. Moreover, the ω_d values that minimize the SDE index are not continuous. Between the ω_d values that minimize the SDE index, there are some other ω_d values that lead to large SDE index values. These values are also the optimal initial ω_d values, although their corresponding optimal α values must be adjusted. For instance, we set the ω_d value as $2\pi\cdot1493$ rad/s, leading to an SDE index value 65.7457, and the obtained optimal α value is an arbitrary value in the range of 12,600 to 13,000, rather than 12,400. In fact, as long as the selected ω_d value lies in the range of $2\pi\cdot1474$ to $2\pi\cdot1521$ rad/s that can be estimated by the STFT spectrum, it can be considered the optimal ω_d value. Nevertheless, its corresponding optimal α values may be adjusted. Therefore, there exist numerous optimal values for the parameter combination $[\omega_d, \alpha]$, and the proposed VME method can easily obtain a group of optimal parameters. This outcome verifies the effectiveness and robustness of the SDE index and the practicality of the proposed VME method.

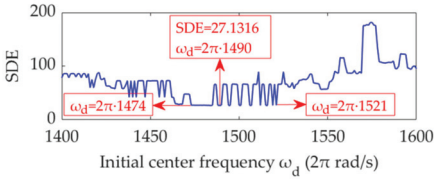


Figure 17. Relationship between the initial center frequency, ω_d, and SDE index.

To verify the effectiveness of the SDE index, the 10 indices mentioned above, i.e., kurtosis, KCI, EK, GI, IE, SE, PE, PSE, EE, and ESE, for the vibration dataset in this case are calculated. As α increases, noise and interference components are first suppressed. When α is up to a certain value, i.e., the optimal value, fault features will get the best presentation. However, if α continues to increase, the fault features will be also suppressed. The relationship curves between α and the 10 indices are shown in Figure 18. Due to monotonous

attenuation shown in Figure 18f,g, the SE and PE indices cannot indicate the optimal α values. However, as shown in Figure 18a–e,h–j, all the other indices exhibit sudden changes and achieve the maximum, minimum, or local extremum in the α range 10,200–10,800. In this case, the optimal α values indicated by these indices are not considerably different from that pertaining to the SDE index. Similarly, for the desired mode and fault feature frequency extraction, VME with the optimal α value indicated by the SDE index achieves the same performance as VME with the optimal α values indicated by the above indices. This finding demonstrates that the performance of VME is robust to the penalty parameter α. Moreover, this finding shows the effectiveness of the SDE index, although the SDE index is different from the abovementioned existing indices.

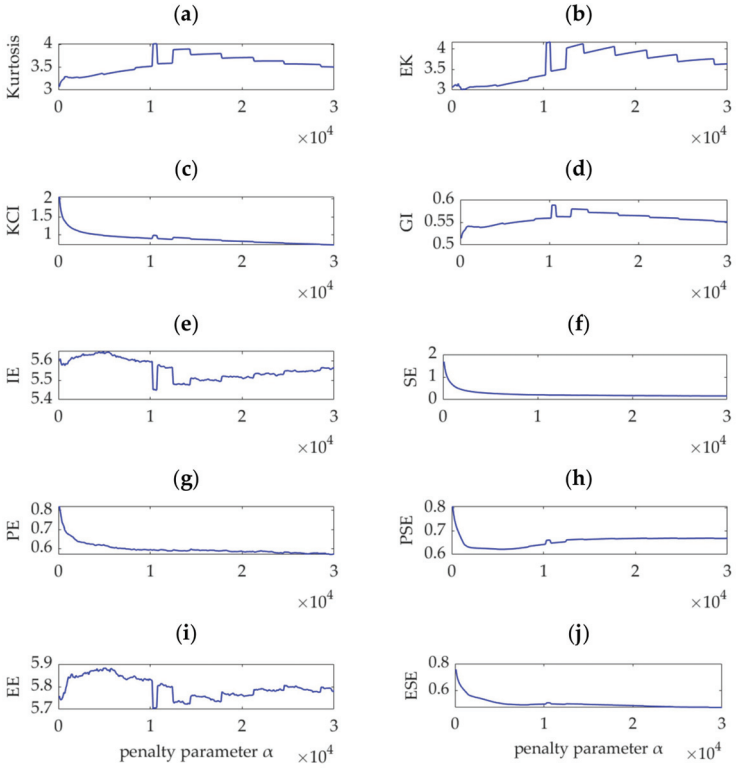


Figure 18. Relationship between different indices and the penalty parameter, α.

To verify the superiority of the improved VME method, the classical VMD method is applied to analyze the pinion fault vibration dataset. In the implementation of VMD, the penalty parameter, α, is set as 12,400, similar to the value in the VME method, and the mode number is set as 8, determined after several trials. By experience, the target mode is selected as the fourth mode. The time-domain waveform and SES are shown in Figure 19. Comparison of Figures 16 and 19 shows that, both the improved VME method and classical VMD method can extract the pinion fault feature frequency from the vibration dataset with heavy noise. Nevertheless, the implementation of VMD may require several trials and experience, whereas the application of the improved VME method is more targeted and thus simpler and more practicable.

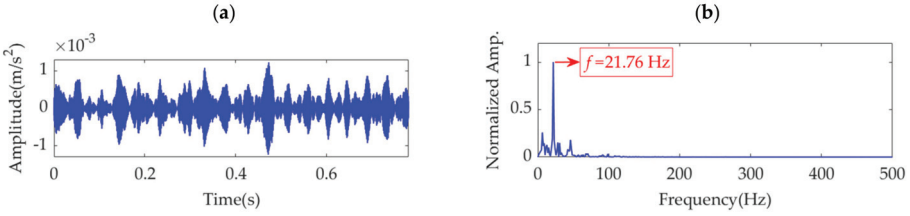


Figure 19. Target mode obtained by VMD: (**a**) time–domain waveform and (**b**) SES.

5.3. Gear Fault Vibration Dataset Analysis

The driving motor rotation speed is set as 900 rpm, and the experimental gearbox output rotation speed is 1730 rpm, as measured by the sensor. Hence, the gear fault feature frequency, f_g, in the experimental gearbox is calculated as 11.77 Hz. The vibration dataset with 4000 points sampled from the experimental gearbox is shown in Figure 20a. Because of the heavy noise, the regular impulse features cannot be recognized in the time-domain waveform. In the SES shown in Figure 20b, several fault-related spectrum lines can be observed. However, the spectrum line at the fault feature frequency cannot be identified. The STFT spectrum shown in Figure 20c indicates that intermittent but not significant fault features occur at approximately 1400 Hz along the time axis.

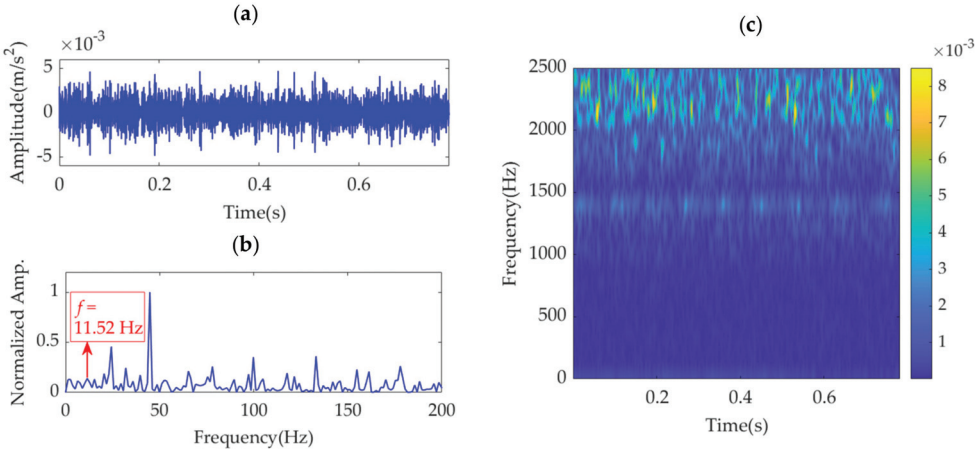


Figure 20. Gear fault vibration dataset: (**a**) time–domain waveform, (**b**) SES, and (**c**) STFT spectrum.

In the implementation of the improved VME method for the gear fault vibration dataset, the initial value of the center frequency, ω_d, is selected as $2\pi \cdot 1400$ rad/s. The optimal value of penalty parameter α is searched in the range from 500 to 100,000, with an increasing step of 100. The relationship between α and the SDE index is obtained and represented by the curve shown in Figure 21. The optimal α value indicated by the minimum SDE index value is α_{opt} = 61,300. VME with the optimal parameter group $[\omega_d, \alpha_{opt}]$ is implemented for the vibration dataset. The obtained optimal desired mode and its STFT spectrum are shown in Figure 22a,b, respectively. The periodic fault impulse features are completely extracted with little noise. Figure 22c shows the SES of the optimal desired mode. The extracted feature frequency of 11.52 Hz is the same as the gear fault feature frequency, f_g.

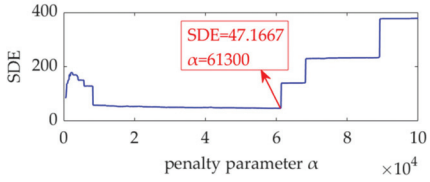


Figure 21. Relationship between penalty parameter α and the SDE index.

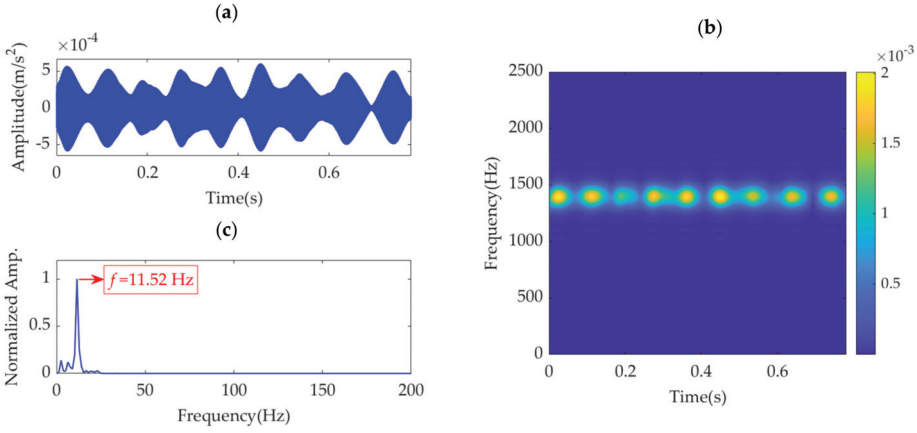


Figure 22. Desired mode obtained by VME with the optimal parameters: (**a**) time–domain waveform, (**b**) STFT spectrum, and (**c**) SES.

We set the penalty parameter, α, as a fixed value of 61300, and the initial value of the center frequency, ω_d, increases from $2\pi \cdot 1200$ rad/s to $2\pi \cdot 1600$ rad/s, with a step size of 2π rad/s. Using the same method as mentioned previously, the relationship curves between ω_d and the SDE index are obtained, as shown in Figure 23. The value of $2\pi \cdot 1400$ rad/s, which minimizes the SDE index, is indeed the optimal value of ω_d. In addition, any value within the range of $2\pi \cdot 1368$ rad/s to $2\pi \cdot 1434$ rad/s can be considered the optimal initial value of ω_d. This aspect suggests that the performance of VME is highly robust to the initial ω_d value. Moreover, the appropriate initial ω_d value can be selected using the STFT spectrum in gearbox fault feature extraction.

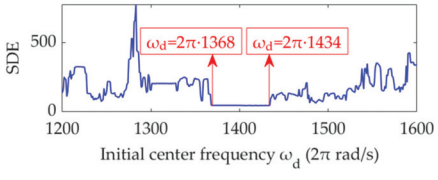


Figure 23. Relationship between the initial center frequency, ω_d, and SDE index.

Like in the previous case, the 10 indices, i.e., kurtosis, KCI, EK, GI, IE, SE, PE, PSE, EE, and ESE for the gear vibration dataset are calculated, and the relationship curves between α and the 10 indices are obtained. Except for the EK and EE indices which are shown in Figure 24b,i, all the other indices vary monotonously with increasing α, as shown in Figure 24a,c–h,j, and no extrema or sudden change values are observed. Therefore, these indices cannot indicate the optimal values of α in the search range. Figure 24b shows that the EK index is minimized, and its corresponding α value is 44,100. Figure 24i shows that a

maximum value exists for the EE index, and the corresponding α value is 31,400. Results of experiments show that the VME methods with different optimal α values indicated by the SDE, EK, and EE indices exhibit similar performance for gear fault feature extraction. However, considering the large differences between the three optimal α values, VME is highly robust to α, and the optimal α value can be selected within a large range. This aspect can be verified by the SDE index variation curve shown in Figure 21, in which the SDE index presents a nearly horizontal varying trend in the α range of 8,300–61,300. As the other indices do not have such an indicative function, the SDE index is considered superior.

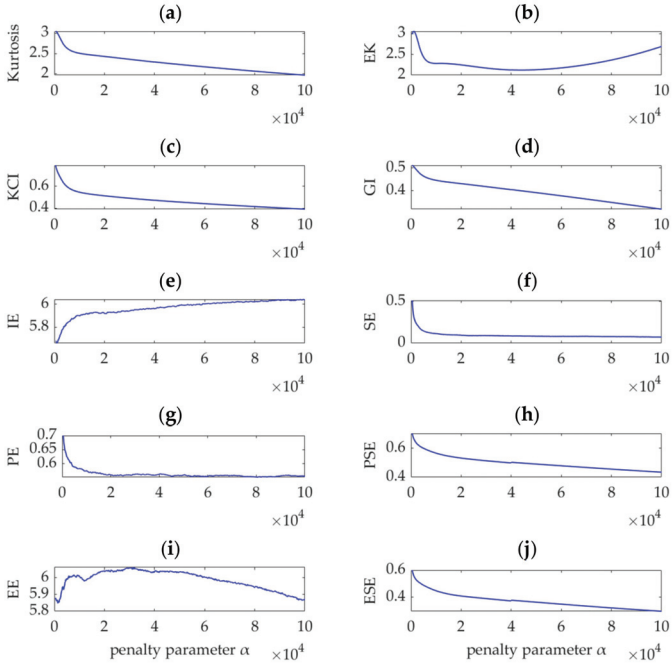


Figure 24. Relationship between different indices and the penalty parameter, α.

Furthermore, VMD is used to analyze the gear fault vibration dataset. In the implementation of VMD, its penalty parameter is set to the same value as that in the VME method, i.e., 61,300, and the number of decomposed modes is set as 8, which is an appropriate value determined by several trials. The target mode containing valuable fault feature information is the fourth mode. The corresponding time–domain waveform and SES are shown in Figure 25. Although the gear fault feature frequency can be extracted from the target mode, comparison of Figure 25 with Figure 22 shows that the desired mode obtained by VME exhibits a smaller amount of noise and fewer interference components than that in the target mode obtained by VMD.

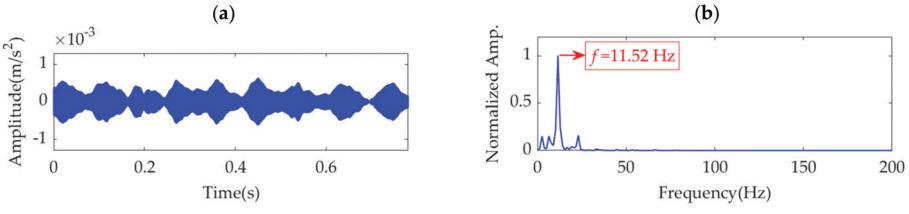


Figure 25. Target mode obtained using VMD: (**a**) time–domain waveform and (**b**) SES.

6. Conclusions

VME directly decomposes a signal into the desired mode and residual signal. Compared with VMD, VME is more feasible to implement. The performance of VME is highly robust to the initial value of the desired mode center frequency and the penalty parameter. In the application of fault feature extraction and fault diagnosis for gearboxes, a reasonable initial center frequency value can be easily selected using the STFT spectrum. Moreover, the proposed SDE index can effectively indicate the optimal value of the penalty parameter and is superior to the currently widely used kurtosis-class and entropy-class indices. Accordingly, the proposed improved VME approach exhibits powerful anti-noise performance and can thus extract clear and complete fault impulse features from gearbox vibration signals with low SNRs. In addition, compared with the target mode extracted by VMD, the desired mode extracted by the improved VME approach contains fewer noise and interference components and thus achieves a superior effect for gearbox fault feature extraction. In the future, we may use the improved VME approach in feature extraction of large-scale data for gearboxes and develop intelligent fault diagnosis approaches based on deep learning.

Author Contributions: Writing—original draft, methodology, Y.G.; supervision, S.J.; formal analysis, Y.Y.; validation, X.J.; resources, Y.W. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (Grant No. 51975517, 52105282), the Key R&D Program of Zhejiang Province (Grant No. 2021C01086), the Zhejiang Provincial Natural Science Foundation of China (Grant No. LY20E050020), the Open Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China (Grant No.ICT2021B43), and the Ningbo Natural Science Foundation of China (Grant No. 2021J038).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Li, J.; Yao, X.; Wang, H.; Zhang, J. Periodic impulses extraction based on improved adaptive VMD and sparse code shrinkage denoising and its application in rotating machinery fault diagnosis. *Mech. Syst. Signal Process.* **2019**, *126*, 568–589. [CrossRef]
2. Li, Q.; Ji, X.; Liang, S.Y. Incipient fault feature extraction for rotating machinery based on improved AR-minimum entropy deconvolution combined with variational mode decomposition approach. *Entropy* **2017**, *19*, 317. [CrossRef]
3. Shao, H.; Lin, J.; Zhang, L.; Galar, D.; Kumar, U. A novel approach of multisensory fusion to collaborative fault diagnosis in maintenance. *Inf. Fusion* **2021**, *74*, 65–76. [CrossRef]
4. Jiang, X.; Wang, J.; Shi, J.; Shen, C.; Huang, W.; Zhu, Z. A coarse-to-fine decomposing strategy of VMD for extraction of weak repetitive transients in fault diagnosis of rotating machines. *Mech. Syst. Signal Process.* **2019**, *116*, 668–692. [CrossRef]
5. Zhou, X.; Li, Y.; Jiang, L.; Zhou, L. Fault feature extraction for rolling bearings based on parameter-adaptive variational mode decomposition and multi-point optimal minimum entropy deconvolution. *Measurement* **2021**, *173*, 108469. [CrossRef]
6. Ding, X.; Li, Q.; Lin, L.; He, Q.; Shao, Y. Fast time-frequency manifold learning and its reconstruction for transient feature extraction in rotating machinery fault diagnosis. *Measurement* **2019**, *141*, 380–395. [CrossRef]
7. He, Z.; Shao, H.; Zhong, X.; Zhao, X. Ensemble transfer CNNs driven by multi-channel signals for fault diagnosis of rotating machinery cross working conditions. *Knowl.-Based Syst.* **2020**, *207*, 106396. [CrossRef]
8. Wang, L.; Shao, Y. Fault feature extraction of rotating machinery using a reweighted complete ensemble empirical mode decomposition with adaptive noise and demodulation analysis. *Mech. Syst. Signal Process.* **2020**, *138*, 106545. [CrossRef]
9. Yao, J.; Zhao, J.; Deng, Y.; Langari, R. Weak fault feature extraction of rotating machinery based on double-window spectrum fusion enhancement. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 1029–1040. [CrossRef]
10. Yu, J.; Lv, J. Weak fault feature extraction of rolling bearings using local mean decomposition-based multilayer hybrid denoising. *IEEE Trans. Instrum. Meas.* **2017**, *66*, 3148–3159. [CrossRef]
11. Li, H.; Liu, T.; Wu, X.; Chen, Q. Application of EEMD and improved frequency band entropy in bearing fault feature extraction. *ISA Trans.* **2019**, *88*, 170–185. [CrossRef]
12. Islam, M.M.; Kim, J.M. Automated bearing fault diagnosis scheme using 2D representation of wavelet packet transform and deep convolutional neural network. *Comput. Ind.* **2019**, *106*, 142–153. [CrossRef]

13. Chen, R.; Huang, X.; Yang, L.; Xu, X.; Zhang, X.; Zhang, Y. Intelligent fault diagnosis method of planetary gearboxes based on convolution neural network and discrete wavelet transform. *Comput. Ind.* **2019**, *106*, 48–59. [CrossRef]
14. Belai, K.; Miloudi, A.; Bournine, H. The processing of resonances excited by gear faults using continuous wavelet transform with adaptive complex Morlet wavelet and sparsity measurement. *Measurement* **2021**, *180*, 109576. [CrossRef]
15. Ying, W.; Zheng, J.; Pan, H.; Liu, Q. Permutation entropy-based improved uniform phase empirical mode decomposition for mechanical fault diagnosis. *Digit. Signal Process.* **2021**, *117*, 103167. [CrossRef]
16. Zheng, J.; Su, M.; Ying, W.; Tong, J.; Pan, Z. Improved uniform phase empirical mode decomposition and its application in machinery fault diagnosis. *Measurement* **2021**, *179*, 109425. [CrossRef]
17. Wang, L.; Liu, Z.; Miao, Q.; Zhang, X. Time–frequency analysis based on ensemble local mean decomposition and fast kurtogram for rotating machinery fault diagnosis. *Mech. Syst. Signal Process.* **2018**, *103*, 60–75. [CrossRef]
18. Zhao, H.; Wang, J.; Jay, L.; Li, Y. A compound interpolation envelope local mean decomposition and its application for fault diagnosis of reciprocating compressors. *Mech. Syst. Signal Process.* **2018**, *110*, 273–295.
19. Teng, W.; Ding, X.; Cheng, H.; Han, C.; Liu, Y.; Mu, H. Compound faults diagnosis and analysis for a wind turbine gearbox via a novel vibration model and empirical wavelet transform. *Renew. Energy* **2019**, *136*, 393–402. [CrossRef]
20. Xu, Y.; Deng, Y.; Zhao, J.; Tian, W.; Ma, C. A novel rolling bearing fault diagnosis method based on empirical wavelet transform and spectral trend. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 2891–2904. [CrossRef]
21. Kim, Y.; Ha, J.M.; Na, K.; Park, J.; Youn, B.D. Cepstrum-assisted empirical wavelet transform (CEWT)-based improved demodulation analysis for fault diagnostics of planetary gearboxes. *Measurement* **2021**, *183*, 109796. [CrossRef]
22. Zhang, K.; Ma, C.; Xu, Y.; Chen, P.; Du, J. Feature extraction method based on adaptive and concise empirical wavelet transform and its applications in bearing fault diagnosis. *Measurement* **2021**, *172*, 108976. [CrossRef]
23. Dragomiretskiy, K.; Zosso, D. Variational mode decomposition. *IEEE Trans. Signal Process.* **2014**, *62*, 531–544. [CrossRef]
24. Zhang, X.; Miao, Q.; Zhang, H.; Wang, L. A parameter-adaptive VMD method based on grasshopper optimization algorithm to analyze vibration signals from rotating machinery. *Mech. Syst. Signal Process.* **2018**, *108*, 58–72. [CrossRef]
25. Ding, J.; Huang, L.; Xiao, D.; Li, X. GMPSO-VMD algorithm and its application to rolling bearing fault feature extraction. *Sensors* **2020**, *20*, 1946. [CrossRef] [PubMed]
26. Zhang, C.; Wang, Y.; Deng, W. Fault diagnosis for rolling bearings using optimized variational mode decomposition and resonance demodulation. *Entropy* **2020**, *22*, 739. [CrossRef] [PubMed]
27. Wang, X.; Yang, Z. Novel particle swarm optimization-based variational mode decomposition method for the fault diagnosis of complex rotating machinery. *IEEE/ASME Trans. Mech.* **2018**, *23*, 68–79. [CrossRef]
28. Miao, Y.; Zhao, M.; Lin, J. Identification of mechanical compound-fault based on the improved parameter-adaptive variational mode decomposition. *ISA Trans.* **2019**, *84*, 82–95. [CrossRef] [PubMed]
29. Gu, R.; Chen, J.; Hong, R.; Wang, H.; Wu, W. Incipient fault diagnosis of rolling bearings based on adaptive variational mode decomposition and Teager energy operator. *Measurement* **2020**, *149*, 106941. [CrossRef]
30. Gai, J.; Shen, J.; Hu, Y.; Wang, H. An integrated method based on hybrid grey wolf optimizer improved variational mode decomposition and deep neural network for fault diagnosis of rolling bearing. *Measurement* **2020**, *162*, 107901. [CrossRef]
31. Yan, X.; Jia, M. Application of CSA-VMD and optimal scale morphological slice bispectrum in enhancing outer race fault detection of rolling element bearings. *Mech. Syst. Signal Process.* **2019**, *122*, 56–86. [CrossRef]
32. Zhu, S.; Xia, H.; Peng, B.; Zio, E.; Wang, Z.; Jiang, Y. Feature extraction for early fault detection in rotating machinery of nuclear power plants based on adaptive VMD and Teager energy operator. *Ann. Nucl. Energy* **2021**, *160*, 108392. [CrossRef]
33. Xu, B.; Zhou, F.; Li, H.; Yan, B.; Liu, Y. Early fault feature extraction of bearings based on Teager energy operator and optimal VMD. *ISA Trans.* **2019**, *86*, 249–265. [CrossRef]
34. Wan, S.; Zhang, X.; Dou, L. Shannon entropy of binary wavelet packet subbands and its application in bearing fault extraction. *Entropy* **2018**, *20*, 260. [CrossRef]
35. Guan, Z.; Liao, Z.; Li, K.; Chen, P. A precise diagnosis method of structural faults of rotating machinery based on combination of empirical mode decomposition, sample entropy, and deep belief network. *Sensors* **2019**, *19*, 591. [CrossRef]
36. Nazari, M.; Sakhaei, S.M. Variational mode extraction: A new efficient method to derive respiratory signals from ECG. *IEEE J. Biomed. Health* **2017**, *22*, 1059–1067. [CrossRef]
37. Antoni, J. The infogram: Entropic evidence of the signature of repetitive transients. *Mech. Syst. Signal Process.* **2016**, *74*, 73–94. [CrossRef]
38. Dibaj, A.; Hassannejad, R.; Ettefagh, M.M.; Ehghaghi, M.B. Incipient fault diagnosis of bearings based on parameter-optimized VMD and envelope spectrum weighted kurtosis index with a new sensitivity assessment threshold. *ISA Trans.* **2021**, *114*, 413–433. [CrossRef]
39. Wang, D.; Peng, Z.; Xi, L. The sum of weighted normalized square envelope: A unified framework for kurtosis, negative entropy, Gini index and smoothness index for machine health monitoring. *Mech. Syst. Signal Process.* **2020**, *140*, 106725. [CrossRef]
40. Li, H.; Liu, T.; Wu, X.; Chen, Q. An optimized VMD method and its applications in bearing fault diagnosis. *Measurement* **2020**, *166*, 108185. [CrossRef]
41. Gao, S.; Ren, Y.; Zhang, Y.; Li, T. Fault diagnosis of rolling bearings based on improved energy entropy and fault location of triangulation of amplitude attenuation outer raceway. *Measurement* **2021**, *185*, 109974. [CrossRef]

42. Li, Y.; Cheng, G.; Liu, C.; Chen, X. Study on planetary gear fault diagnosis based on variational mode decomposition and deep neural networks. *Measurement* **2018**, *130*, 94–104. [CrossRef]
43. Yan, X.; Jia, M.; Xiang, L. Compound fault diagnosis of rotating machinery based on OVMD and a 1.5-dimension envelope spectrum. *Meas. Sci. Technol.* **2016**, *27*, 75002. [CrossRef]

energies

MDPI

Article

Numerical Simulation of Aging by Water-Trees of XPLE Insulator Used in a Single Hi-Voltage Phase of Smart Composite Power Cables for Offshore Farms

Drissi-Habti Monssef [1,*,†], Manepalli Sriharsha [1], Neginhal Abhijit [1], Carvelli Valter [2,†] and Bonamy Pierre-Jean [3]

[1] COSYS Department, Université Gustave Eiffel, F-77447 Marne-la-Valleée, France; sriharsha.manepalli@gmail.com (M.S.); abhijit.neghinal@gmail.com (N.A.)
[2] Department of Architecture, Polytechnic University of Milan, 20133 Milan, Italy; valter.carvelli@politecnico.it
[3] MEDYSYS, 91400 Orsay, France; pj.bonamy@medysys.fr
* Correspondence: monssef.drissi-habti@univ-eiffel.fr; Tel.: +33-6-0927-1370
† Directors—International Associated Lab. SenSIN-CT, Université Gustave Eiffel—Politecnico Milan, F-77447 Marne-la-Valleée, France.

Abstract: Submarine power cables are expected to last 20 years without maintenance to be considered technologically reliable enough and economically beneficial. One of the main issues facing this target is the development of what is called commonly water-trees (nanometer-sized flaws filled with residual humidity), that form within XLPE (cross-linked Polyethylene) insulators and then migrate towards copper, thus leading to its corrosion and further to possible shut-down. Water trees are resulting from the coalescence of nanovoids filled with residual humidity that migrate towards copper under the combined effects of electrical forces and plastic deformation. The nanovoids are originated during manufacturing, shipping, handling and embedding in deep seas. The formation of these nanovoids leads to the degradation of the service lifetime of submarine power cables. Current research is intended to come up with a way to go a little further towards the generalization of coalescence of n nanovoids. In the perspective of multi-physics modeling, a preliminary 3D finite element model was built. Although water voids are distributed randomly inside XLPE, in this study, two extreme cases where the voids are present parallel and perpendicular to the copper surface, were considered for simplification. This will enable checking the electric field effect on neighbouring voids, in both cases as well as the influence of the proximity of the conductor on the plasticity of voids, that further leads to their coalescence. It is worthwhile to note that assessing water-trees formation and propagation through an experimental campaign of ageing tests may extend over decades. It would therefore be an exceptional opportunity to be able to get insight into this mechanism through numerical modeling that needs a much shorter time. The premilinary model suggested is expected to be extended in the future so that to include more variables (distribution and shapes of nano-voids, water pressure, molecular modeling, electric discharge.

Keywords: wind energy; Hi-Voltage Power cables; offshore; phase; water-trees; aging; numerical modeling; XPLE insulator

Citation: Monssef, D.-H.; Sriharsha, M.; Abhijit, N.; Valter, C.; Pierre-Jean, B. Numerical Simulation of Aging by Water-Trees of XPLE Insulator Used in a Single Hi-Voltage Phase of Smart Composite Power Cables for Offshore Farms. *Energies* **2022**, *15*, 1844. https://doi.org/10.3390/en15051844

Academic Editors: Phong B. Dao, Lei Qiu, Liang Yu and Zahra Sharif Khodaei

Received: 10 January 2022
Accepted: 23 February 2022
Published: 2 March 2022

1. Introduction

Offshore energy is one of the most prominent renewable energy resources which is currently under development, globally. This energy generation relies on some critical structures, among which are wind-blades and Hi-Voltage electrical transport cables that connect offshore farms to onshore [1–6]. A lot of research is carried out to suggest the best ways to maintain the reliability of offshore structures. The transport of wind-generated energy is achieved by submarine Hi-Voltage Electric Power cables. Modern cables are expected to go beyond simple embedded fiber-optics for telecommunications. In fact, cables should absolutely carry optical fibers for strain and temperature monitoring, at

various internal locations, over its total length [1]. These power cables once deployed in deep-sea will work under demanding-environmental conditions for over 20 to 25 years and efficient methods are necessary to predict their lifetimes [6]. The most preferred insulator material is cross-linked polyethylene (XLPE) due to its high performance. It can be used up to 420 kV system voltage. Figure 1 shows the complex structure of submarine power cable. The deterioration phenomena that degrades the service lifespan of these XLPE power cables is know as water treeing. Water-trees are very small voids of a few nano-meters as depicted in Figure 2. Cable manufacturers and power suppliers need to understand the influence of water-trees, their coalescence and their migration towards copper, that may lead to corrosion with the fear of shutdown that will result in an economical catastrophe.

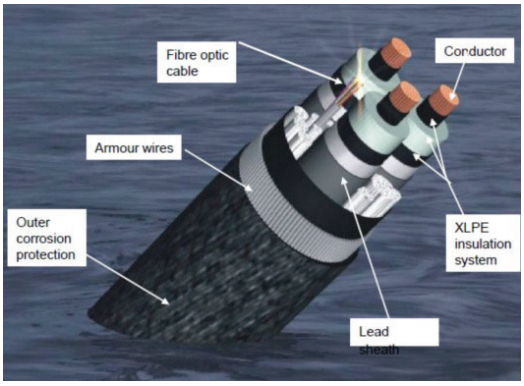


Figure 1. The components of a submarine power cable, Copyright ESCA [7].

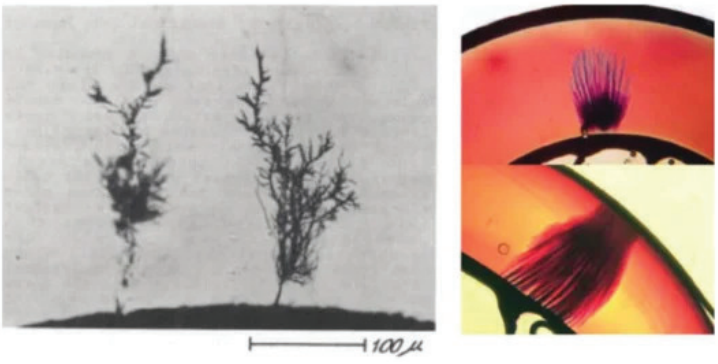


Figure 2. The Electrical and water treeing shown in the polyethylene insulated cable in the left. Electrical trees are grown from field-enhancing defects at the insulation interface, in the right water trees growing from the conductor screen. Taken from [8].

A brief summary of bibliography defines water trees as nanovoids that need a residual humidity within the insulation to propagate (not free water) and do not grow up under low electrical voltage. Water trees are located in the insulation materials (Bow tie tress, the case study of current article) as well as in the interface between the insulator and the screen (Vented trees), from defects that exist at these both areas. During the elaboration process of the insulator, many nanovoids are introduced in the dielectric (stream curing), that absorb humidity and start to coalesce to form water-trees. Water-trees are partially conductive tridimensional structures that look like plumes and are related to polymers crystalline structures. Under the combined effects of high electrical voltage and residual water, water-trees propagate and increase size by a channeling effects that took them generally all the way to the conductor, which starts corroding. The bibliography pertaining

to modeling of degradation by water-trees is numerous. Few examples just for reminder, the growth of water tress are heavily influenced by external and internal factors. The external factors (Operational) are when water trees arise during the process of manufacturing, construction and installation [9,10]. Internal factors are the material factors of cables. A lot of publications are available that cover the effects due to external factors on water tree growth, but there is little literature available that addresses internal factors. There is an increase in temperature of the insulator due to rising loads and electrical current, which also triggers the formation of water-tress inside XLPE [11]. Thermal aging influence during operational service was identified to be significant on several properties of the insulation material (electrical, physico-chemical...). Thermal degradation was shown to ease the initiation as well as the propagation of water-trees [9,10]. The same authors used a modeling based on finite element along with Taguchi method to identify and study key-factors that influence the electric field at the tip of water-trees initiated from the outer surface of XLPE cable insulation. Dielectric properties of water-trees as well as the properties of the insulation material were shown to have a heavy influence of the increase of tree tip's electric field intensity and this influence is higher than tree's geometrical parameters. The permittivity distribution within the water-tree has also been identified as the key-factor over all others. Accelerated aging by water-trees based on increasing voltage, frequency as well as various techniques of producing artificial water-trees have been used to get insight into the influence of this aging on dielectric properties of the XPLE insulator [12]. It was shown that the electric field at the tip of the water tree is increased due to the increase of both the size and permittivity of the water-tree. However, there remains the question whether accelerated aging is or is not representative of what is going on inside the cable. Lin et al. [13] have modeled the effects of mechanical behavior on the size increase of water-trees, at various temperatures. It was shown that the increase of size of water trees is depending mainly on the mechanical behavior of the insulation material. It was also shown that water tree propagation is slowed by the raise of temperature while reverse mechanism is shown when yield strength is lowered. This adds to the fact that the development of water trees is function of the electric field and residual humidity [14].

The pioneering article [2] that supports current work has suggested a model of the coalescence of two neighbouring nano-voids. It is assumed that plastic deformation of the nanovoids that arise from the combined effects of electrical field and mechanical loading is responsible for two adjacent voids to coalesce, thus promoting the growth of water trees. In this article, the growth of water trees due to internal factors was investigated by considering some special cases.

Using computer modelling and numerical analysis a better understanding of electric field distribution caused by water-treeing mechanism can be gained. Heuristically, a water tree is a nanovoid that propagates under the influence of an electrical field and experimental findings state that the study of water growth rates as a function of polymer morphology, electric field strength and frequency would need very long time [8]. Water voids can be of any shape but for the sake of simplicity, they are considered as ellipsoids [15].

Under the influence of alternating electric field, the XLPE material suffers from continuous Maxwell stress which leads to the fracture of molecular chains [16]. Water voids connect through channels and that advancing in the direction of electric field, which causes degradation of the dielectric strength of insulation [17] and also results in the migration of water-trees to conductor followed by its corrosion with probable shutdown. The range of electric fields considered for this analysis was 10 kV/mm to 100 kV/mm. This was then applied to evaluate the effective plastic strain when water voids are placed in both parallel and perpendicular to the conducting surface. Obviously these two extreme distributions are not the only ones that can exist and there are a plethora of other cases. Nevertheless, our current objective is to set-up a preliminary model that study the effects of the proximity of nano-voids to copper conductor and of the Maxwell forces, as well as these same effects when one moves away gradually from the conductor. This would allow us to better specify the mechanisms when we will consider the random distributions of defects within the

insulator. Except its concept, the article is not yet to be applied to industry. In fact, one should keep in mind that experimental aging is extremely tedious, given the very high aging resistance of XPLE. As a matter of fact, many researchers tried to extract results from fast experimental tests although those tests are not representative of "natural" aging. So the scope of current research is precisely to go towards simulation of natural aging of a single phase (not the cable) using numerical modeling. For all work, the depth at the bed sea where the cable is placed is considered to be 800 m (mean depth that was suggested by a cable manufacturer). The effects of water pressure, the external temperature are not studied, at this stage.

The electromechanical modeling is set-up to simulate an electromechanical system before actual system is built to examine physical parameters of the system. The present research was carried out in the electromechanics perspective, using COMSOL Multiphysics. COMSOL Multiphysics is a finite element analysis solver that helps to conduct multiphysics simulations. Each module in the physics interfaces that are available in this platform is fully multiphysics enabled. There are various study capabilities that include stationery, eigen frequency, time-dependent, frequency response, buckling, and parametric studies.

2. Previous Model

In a previous article [2], numerical models were developed that include a 3D Finite Element (FE) analysis of a water void formed inside an insulator, using COMSOL Multiphysics software. In the perspective of this modeling, ellipsoids were considered [2] and the displacement field of these water voids was proven to be function of the applied electric field. At the tips of these water voids (modeled as ellipsoids), the strain is strongly dependent on the relative distance between two neighboring nanovoids and the magnitude of the applied electric field. The model that was proposed was a sort of introduction towards the assessment of ageing of XPLE insulator, which is extremely tough to simulate experimentally. Indeed, XPLE is a very well mastered material and experimental ageing may take decades before delivering meaningful inputs. Such a study is a typical case where appropriate numerical models can bring a strong support to researchers to figure out ageing under coalescence of nanovoids (water-treing).

The initiation and development of water trees is function of the electric field and leftover water substance [2,14]. The varied electric field coupled to the polarization effect shapes water nanovoids into ellipsoids [15]. The application of an electric field actuates high Maxwell stress close the tips of ellipsoidal water voids, resulting in their distortion. Agreeing to [18], moderate electric fields of 10–50 KV/mm don't cause permanent deformation of water voids, whereas [19] demonstrated that electric fields over 100 KV/mm cause plastic deformation. Hence, in our study an electric field of 50–100 KV/mm was applied to assess the effective plastic strain generated at the tips of water voids. In relation to a numerical study of water treeing in power cable insulation from a mechanical point of view, refs. [20,21] detailed that weariness of insulation caused by dielectro-phoretic stresses around the voids might lead to the development of water trees.

It was appeared [8] that water-trees growth proceeds through coalescence of ellipsoidal-shaped nanovoids. Nanovoids were originating from extrusion process and also generated when cables are under service, time water molecules reach the nanovoids, because excess of humidity. After initiation, water-trees start their migration towards copper which cause corrosion to propagate and damage cables. Figure 3 represents the steps involved in the development of water trees. Ref. [2] proved that Maxwell stress increases as the relative inter-voids distance decreases. That leads to the increase of the Von Mises stress. Figure 4 shows the close relationship between the effective plastic strain at the tips of the nanovoids electric field. Typically, it demonstrated that the tips of water nanovoids reaching plasticity with an increase in electric.

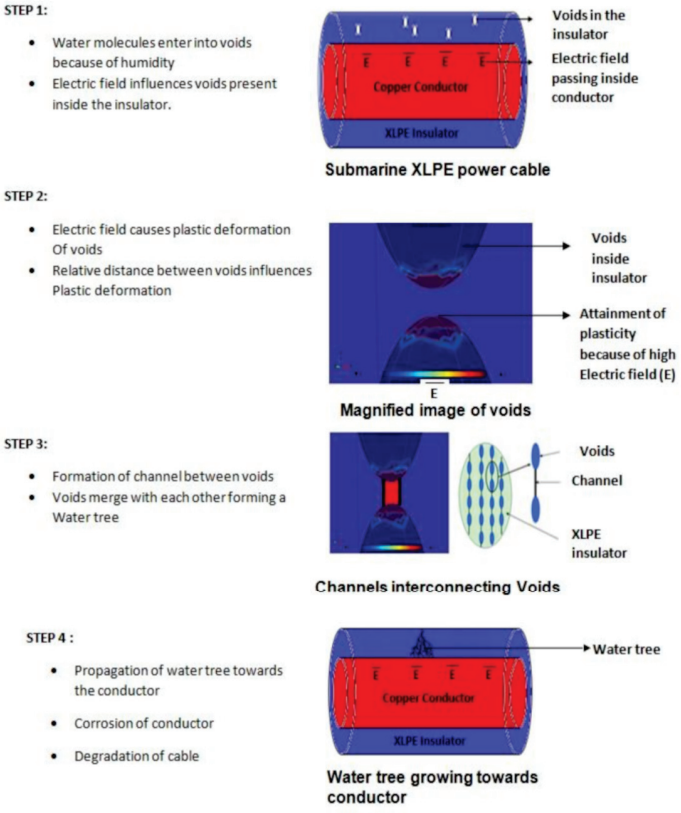


Figure 3. Schematic representation of steps involved in the growth of water trees [2].

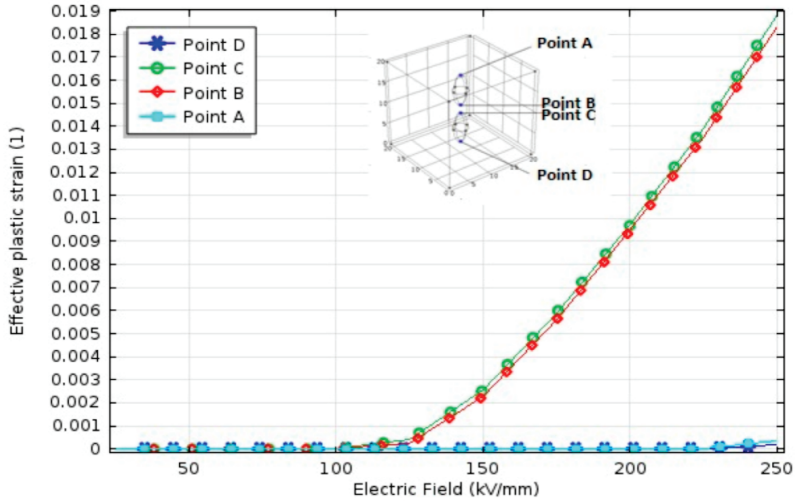


Figure 4. Effective plastic strain at the tips of ellipsoids with increasing electric field (50–250 kV/mm) and decrease in distance between points B and C at distance 0.5 μm.

The results of [2] shed some light on this juncture, in spite of the fact that significant inquire yet to be done. We noted the significantly affects on the coalescence and formation of water voids with respect to applied electric field and relative distance between two adjacent nanovoids. In the current article, we went a little further to enlarge the research, which may enable generalizing of previous model to coalescence of n voids at the step level of a single phase (not the cable) and examining how these defects initiate corrosion of copper. The continuation of this work is the aim of the current research.

3. Objective of Current Research

The void coalescence in XLPE power cable has a significance with respect to applied electric field and the spacing between the nano-voids [2]. The present work is intended to come up with a way towards the generalization of the coalescence of n nanovoids at the step level of a single phase. For that, a preliminary 3D finite element model was developed. A multi-physics model is presented and simulated the combined effects of mechanical stress and electric field. In an insulator, water voids can occur in random locations but for simplicity, two extreme scenarios were considered. In the first extreme scenario, water voids were placed parallel to the copper surface and close to it, knowing that the distance to copper has influence on plastic deformation of voids. In the second scenario, the water voids were placed perpendicular to copper Surface. For both cases, the idea upstream is to check the effects of Maxwell forces on nanovoids when their distance from copper conductor is varied either when they are aligned horizontally and/or vertically. Figure 5 shows the illustration of this project. The electric field varied from 10 kV/mm to 100 kV/mm.In all the work, he cable is assumed to be at a depth of 800 m and is considered submitted to hydrostatic loading

In the Figure 5, Case 1 represents water voids placed parallel, Case 2 represents water voids placed perpendicular and Case 3 represents random placement of water voids to the electrical conductor. Figure 6 illustrates the one of the extreme cases where nanovoids are placed parallel to the conductor of the single phase considered. The cuboid represents XLPE insulator, which is attached to copper conductor. This model replicates submarine power cables considered at a nano-scale. This article covers the different aspects of the coalescence of nanovoids by preforming the following cases:

- calculating the displacement and von Mises stress of water voids placed perpendicular and parallel to the copper surface with varied voltage;
- studying the plastic deformation of nanovoids placed in both configurations knowing that plastic deformation of void tips is leading to their coalescence, then compare the results.

The Section 4 of this article describes the modeling tools, the Section 5 describes the multiphyics or coupled problem approach to understand the electro-mechanical problem in this juncture and the Section 6 describes the results obtained.

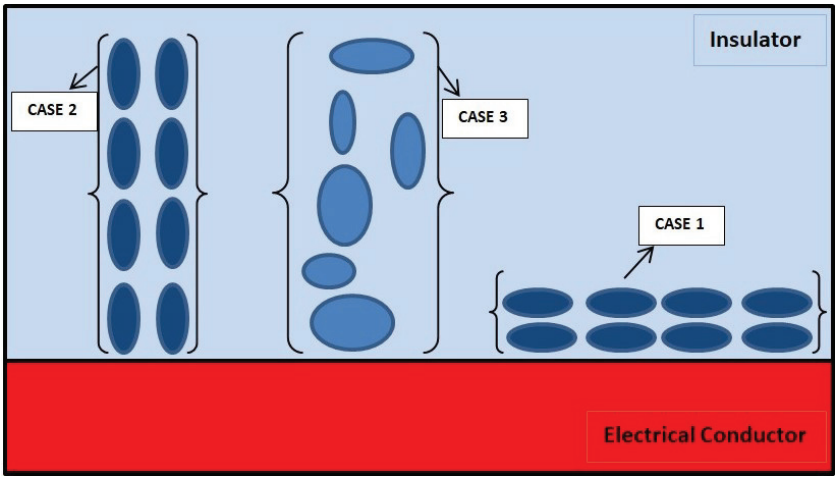


Figure 5. Illustration of the possible cases where voids can be placed in an insulator for the analysis.

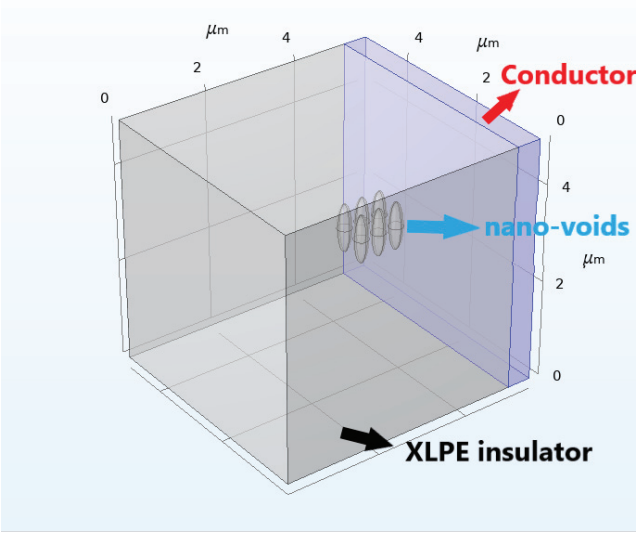


Figure 6. CASE 1: Illustration of the case where nanovoids are put parallel to copper.

4. Numerical Modeling

4.1. Geometric Parameters

The amount of water content that can be present inside insulator is around 1% to 6% of the total insulation volume [8]. The density is approximately 106 mm^{-3} [14]. In this work, water void is considered as 0.1 μm. The number of voids present in each scenario was limited to four as it was difficult to run the tests for a large number of voids in the perspective of multi-physics modeling. As shown in the previous section, copper can be represented as an empty cubic cell with conducting properties. The insulator can be represented as a cubic cell in which water voids are presents in the form of ellipsoids. For this study, the geometric parameters for each domain are considered as shown in Table 1.

Table 1. Modeling parameters.

Name of the Domain	Value	Unit
XPLE	l = 10, b = 10, h = 10	μm
Copper	l = 10, b = 10, h = 10	μm
Water void (ellipsoid)	a = 0.6, b = 0.6, c = 1.8	μm

FE Software

As mentioned earlier, modeling and simulation were carried-out using COMSOL Multiphysics. This solver is most preferred in multiphysics perspective as it provides a wide range of possible operating conditons and physical effects. The current case study was carried-out using electro-mechanics module that is found under the structural mechanics physics tree. For the sake of simplicity for the modeling and analysis, water voids were considered as ellipsoids provided with the suitable materials [2].

4.2. Material Parameters

The material parameters that were used for each domain in this study are shown in Tables 2–4.

Table 2. XPLE Properties [2].

Name	Value	Unit
Young's Modulus	3.5	GPa
Relative permittivity	2.3	1
Poisson's Ratio	0.3	kg/m^3
Density	930	1
Electrical conductivity	1×10^{-15}	S/m
Initial yield stress	18	MPa

Table 3. Properties of water-voids [2].

Name	Value	Unit
Electrical conductivity	-	-
Relative permittivity	80	1

Table 4. Properties of Copper [6].

Name	Value	Unit
Young's Modulus	112	GPa
Relative permittivity	1	1
Poisson's Ratio	0.30	1
Density of Copper	9400	kg/m^3
Electrical conductivity	58.7×10^6	S/m

4.3. Electromechanics Model

A 3D electro-mechanics model is used to describe the deformation of water void at the interface of the conductor and insulator of the single phase considered (not the cable). Initially, a single water void is placed in between the two cubic cells and its deformation was studied with varying electric potential. The equations that governs the electro-mechanics module in COMSOL Multiphysics are:

$$-\nabla.\sigma = F \quad (1)$$

$$-\nabla.D = \rho \quad (2)$$

Equation (1) is the equilibrium equation of continuum mechanics and Equation (2) is the Gauss law of electricity where σ is the Cauchy stress tensor, F is the body force, D is the electric displacement and ρ is the free charge density. The study was performed to understand the distortion of water voids at their tips. Further, the study was extended to understand the plastic deformation between them. The electric field intensity is the negative of potential gradient as shown in Equation (3). The constitutive equations that relates the electric field intensity and displacement (D) in an insulator is shown in Equation (4).

$$E = -\nabla V \tag{3}$$

$$D = \epsilon_0 \epsilon_r E \tag{4}$$

As mentioned earlier plastic deformation of water voids was studied using von Mises yielding criterion. This study was achieved by using plasticity theory. By default, the problem was first set to Linear elastic material and plasticity was added as a sub node(attribute) to it. As we know, in the elastic region, the stress and strain are proportional to each other. At some point, there will be yielding and that is the end of the elasticity. So, we have to specify the initial yield stress. After this behaviour, the model meets perfect plasticity (Figure 7).

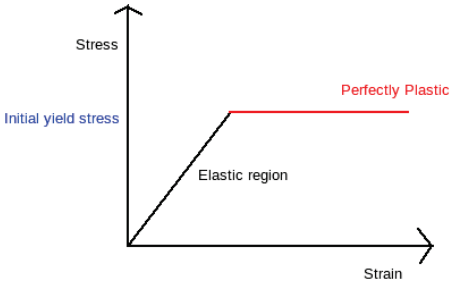


Figure 7. Stress-Strain relationship.

In our case, as the electric field increases, von Mises stress also increases, especially at the tips of ellipsoids (i.e., nanovoids). So, to study plasticity of water voids we found a value during simulations where von Mises stress surpasses the yield strength of XLPE. Figure 8 shows the meshing which was used for case 1. The coarse mesh was used for both XLPE and Copper model, a normal mesh was used for ellipsoids.

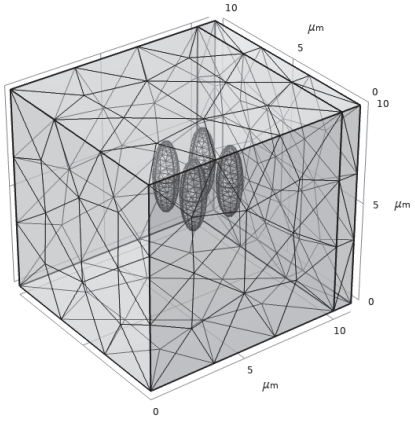


Figure 8. Meshing in Case 1.

5. Results and Discussion

To understand the deformation caused due to applied electric field, various simulations were executed to gzt some interesting results. With the increase of electric field up to 250 kV/mm, von Mises stress also gets higher at the tips of water void void. At the tips of two neighboring nanovoids the displacement is a function of their relative distance and on the magnitude of the electric field [2].

5.1. Case 1: Water Voids Placed Parallel to the Copper Surface

As specified above, the first part of this study involves the placement of water voids parallel to copper surface. Since the analysis was carried-out on several 3D voids present in an insulator, to compensate with the computational cost and resources, the study was limited to 100 kV/mm. Initially, the electric potential was provided to observe the displacement of water voids. Then, the electric field was varied between 10 kV/mm to 100 kV/mm. Figures 9 and 10 show the displacement of water voids and electric field distribution inside XLPE of the single phase considered.

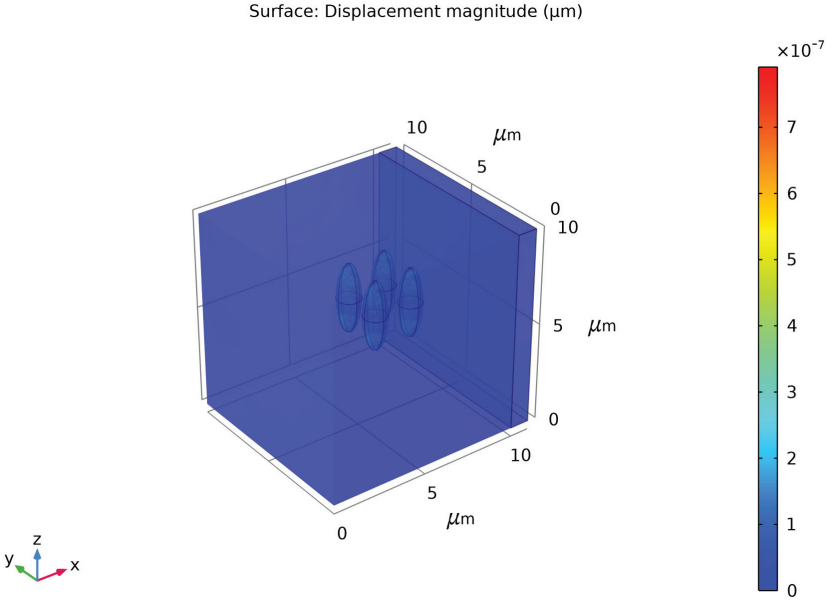


Figure 9. Deformation of voids at their tips. Note the colors of the nanovoids, especially red that shows plastic deformation stage.

When the nano-voids are placed in the direction of electric field they suffer a maximum deformation [2]. The deformation on the voids can also vary concerning the location inside the insulator. Figure 11 shows the deformation as a form of displacement concerning the electric field. As electric field increases the deformation also increases. Points A and C are most likely to be deformed because they were placed nearer to the edge where the influence will be higher.

As shown in the illustration, the points in the figures are the tips of the water voids. The deformation of the voids also depends on their position and their alignment [2]. Generally, for two voids to merge, it involves plastic deformation of the localized at the tips [2]. Now, to study the merging of the material von Mises criterion was used. In a material, plastic deformation is attained due to an external load and that is or electric field. This electric field causes Maxwell stress to act along the voids surfaces.

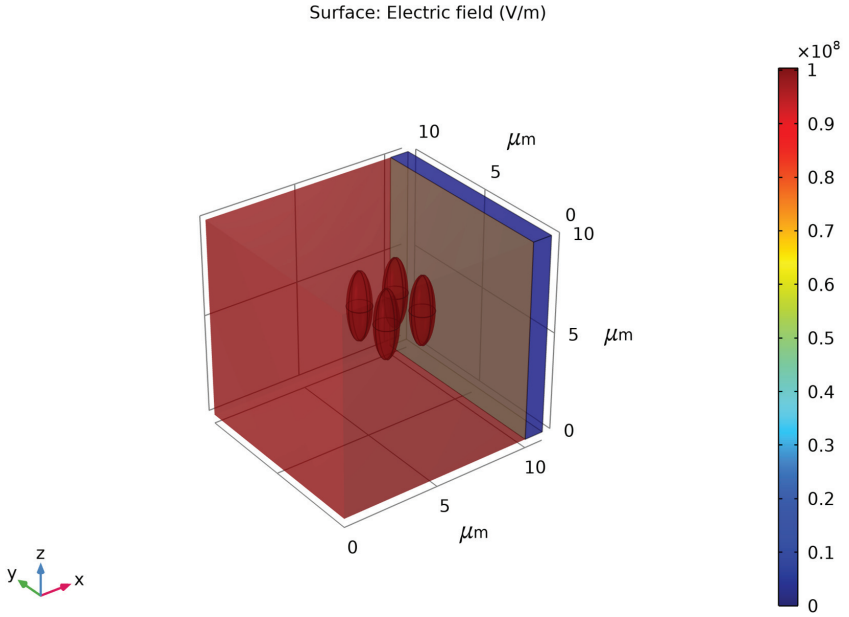


Figure 10. Electric field distribution inside XPLE.

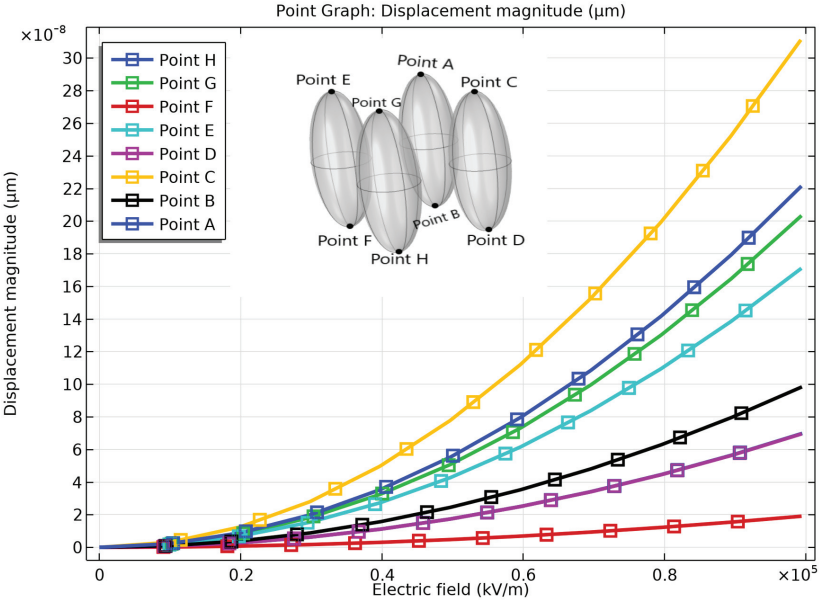


Figure 11. Deformation vs electric field.

Figure 12 shows the increase of von Mises stress especially at tips. This is due to the induced Maxwell stress. In this Figure 12 shows that at 80 kV/mm von Mises stress surpasses the yield strength of XLPE.

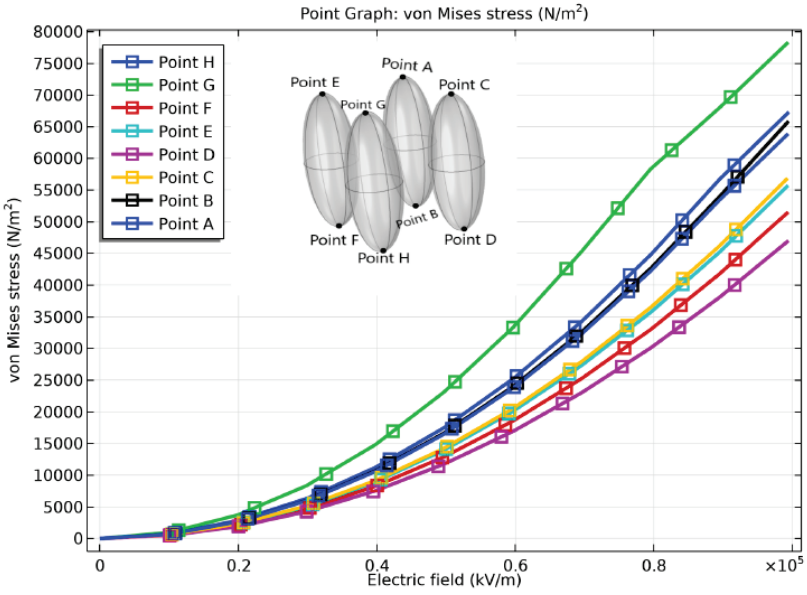


Figure 12. Von Mises stress vs. electrical field.

When Von Mises stress reaches a critical value which can be seen from the Figure 12 water voids experience permanent deformation, 80×10^3 Pa. That is, at around 0.8×10^5 kV/m (i.e., 80 kV/mm) the plastic deformation is taking place at point G. In the Figure 13 around 0.7×10^5 kV/m (i.e., 70 kV/mm) there is a huge shift of the slope in the displacement as the electric filed increases and it shows that initial yield stress required for studying plasticity is 1.4×10^{10} N/m^2 . That means, at this critical value, voids experience permanent deformation. And, at electric field 0.5×10^5 kV/m the equivalent plastic strain was observed.

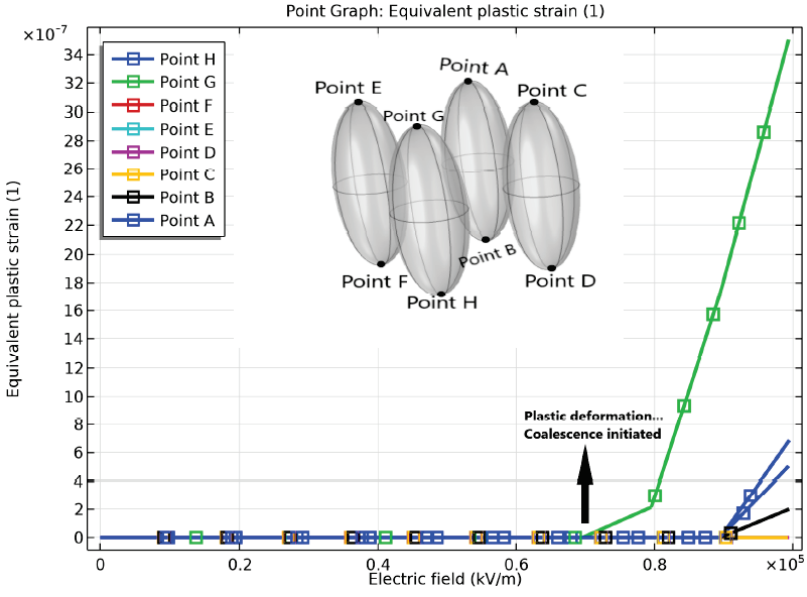


Figure 13. Plastic strain vs. electric field.

5.2. Case 2 : Water Voids Placed Perpendicular to the Copper Surface

In this case, water voids were placed perpendicular to Copper. The entire procedure is the same as case 1. As mentioned earlier, the deformation of voids is expressed as the displacement at the tips of water voids, Figure 14.

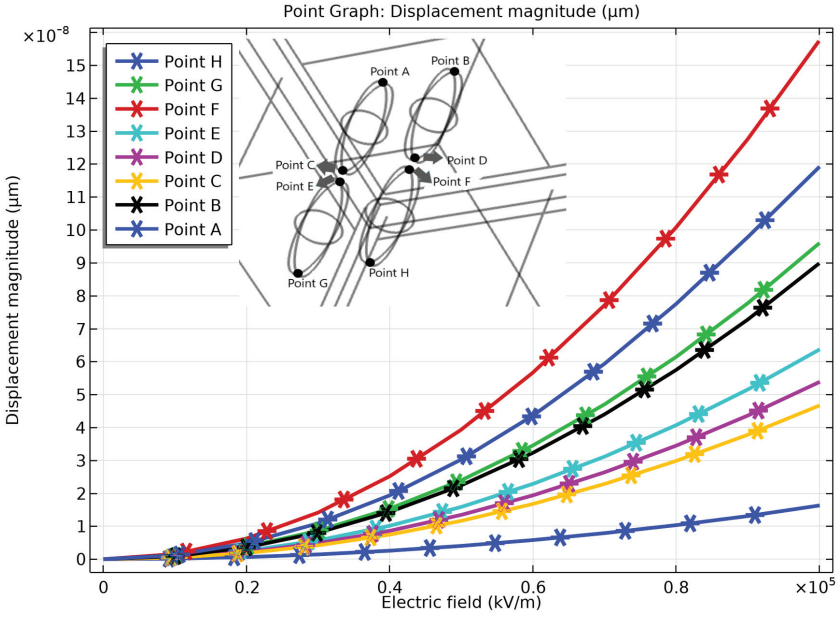


Figure 14. Deformation as a function of Electric field for case 2.

The deformation of nanovoids at their tips was observed as electric field increased. The affect on water voids due to electric field strictly depends on the placement of nanovoids inside the insulator. As specified earlier, for two voids to merge, it involves plastic deformation of of the tips to take place [2]. Now, to study the failure of the material, von Mises criterion was used. As shown in Figure 15, the nanovoids in this juncture most likely execute permanent deformation and at 19×10^3 Pa. Around 0.7×10^5 kV/m (i.e., 70 kV/mm) the plastic deformation is taking place at point D and followed by other void tips in XLPE as shown in Figure 16.

The aim of the current research is to observe the plastic deformation at tips in both cases and as the electric field increases the merging of localized zones are most likely expand forming water trees. The spacing considered in the first scenario was different from the second as the modeled configuration is different. The spacing considered was 2 μm for the first scenario and whereas 1.8 μm for the second. In this analysis spacing was not varied in any of the cases.

As specified in the Section 2, the void coalescence was studied at the tips of water voids while setting the critical distance between the two adjacent water voids(i.e., nano voids) as 2 μm and at an applied electric field of 225 kV/mm. In both cases, it we can observe the cause of plastic deformation by high Von Mises stress and this Von Mises stress was influenced by an amplified electric field. The localized damage will be achieved if the order of plasticity increases. The plastic deformation caused by high von Mises stress obtained from the influence of an amplified electric field at the tips of voids. From both the cases, the results obtained helped us to visualize plastic deformation of nanovoids placed at two extreme cases inside an insulating material.

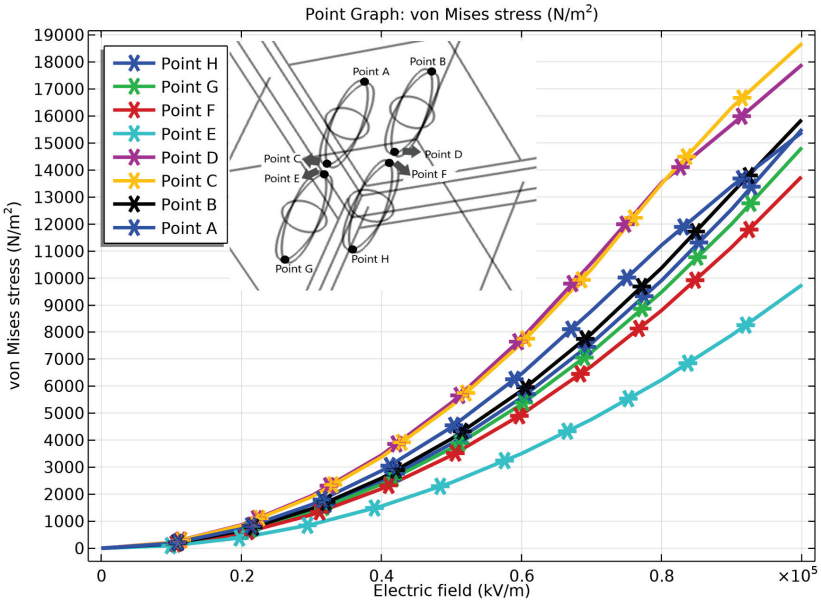


Figure 15. Von Mises stress vs. Electric field for case 2.

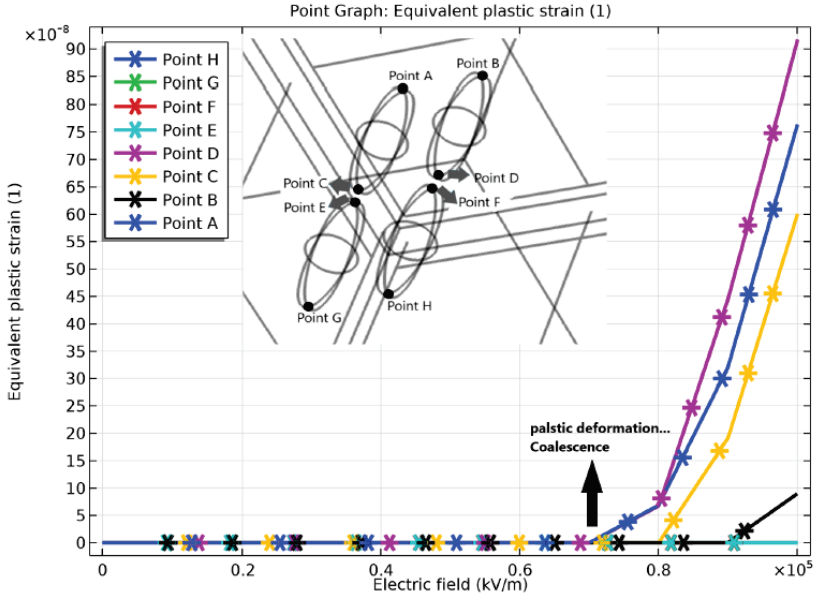


Figure 16. Equivalent plastic strain vs Electric field for case 2.

The plastically-deformed zones expand with increase of the electric field intensity and the merging of voids will be initiated. This kind of model is interesting because it is difficult to perform experimental analysis. As a top view of the modeling, this research tends to indicate that numerical simulation of ageing by water trees can be approximated correctly. The parameters that are critical for this mechanism to be activated can be accurately identified and varied so that enabling insightful conclusions to be out-comed, which can help to figure out how ageing proceeds in such critical structures. Of course, we are still

far from a model that can be used by engineers. This will need to upgrade the model by introducing many other variables such that the random distribution of nanovoids, the variation of external temperature, the external pressure, the movement of cables... This effort is currently under way.

6. Conclusions

Submarine power cables are expected to last 20 years without maintenance to be considered technologically and economically reliable, enough. One of the main issues facing this target is the development of what is called commonly water-trees (nanometer-sized flaws) that form within XLPE insulators and then migrate towards copper, thus leading to its corrosion and further to shut-down. Therefore, cable manufacturers are keen to understand the possible parameters that influence the growth of water-trees. The results achieved in this article were based on generalizing coalescence to nanovoids placed parallel and perpendicular to the conductor of the single phase considered and gave a preliminary understanding on the issues regarding the electrical breakdown. Two extreme cases, where water voids placed parallel and perpendicular to the copper surface, were considered. The simulations were run up to 100 kV/mm to understand the plastic deformation especially at the water void's tips, which leads to coalescence. The spacing between water voids was kept minimum. As a future work, more variables that were not considered in current article will be introduced (shape and distribution of voids, chemical aspects of voids merging, effects of electric discharge, effect of water pressure...). Also, finer meshing can be implemented to get more precise results. The coalescence at tips of water voids was observed at 70 kV/mm for the first scenario, whereas 80 kV/mm for the second scenario which is matching with the actual functional values. Based on this, one can conclude that with the presence of n nanovoids, coalescence can happen at a much faster rate that suspected, with a possible shutdown. As mentioned earlier, technically nano-voids can be present in any configuration inside an insulator but two extreme cases were considered for the sake of preliminary investigation. Future work can consider randomly distributed nanovoids and check for the conditions of their merging. Significant research remains to be done on the investigation of coalescence of n nanovoids with different densities and temperature involved.

Author Contributions: D.-H.M. is in charge of supervision, conceptualization, methodology, resources and review editing, M.S. and N.A. are in charge of software, validation and formal analysis, C.V. and B.P.-J. are in charge of investigation. All authors have read and agreed to the published version of the manuscript.

Funding: Funding was provided by The French Research Agency (ANR) under the Contract Flow-Cam Project.

Acknowledgments: This research has been conducted within Flow-Cam FLOW CAM (FLoating Offshore Wind turbine CAble Monitoring) Project in the framework of Martera Era-Net Cofund. The French National Research Organizeation, ANR, provided funding to this work. I-Site Future provided a master-grant to SM. MDH would like to thank the funders for their support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Drissi-Habti, M.; Raman, V.; Khadour, A.; Timorian, S. Fiber Optic Sensor Embedment Study for Multi-Parameter Strain Sensing. *Sensors* **2017**, *17*, 667. [CrossRef] [PubMed]
2. Drissi-Habti, M.; Raj-Jiyoti, D.; Vijayaraghavan, S.; Fouad, E. Numerical Simulation for Void Coalescence (Water Treeing) in XLPE Insulation of Submarine Composite Power Cables. *Energies* **2020**, *13*, 5472. [CrossRef]
3. Raman, V.; Drissi-Habti, M.; Guillaumat, L.; Khadhour, A. Numerical simulation analysis as a tool to identify areas of weakness in a turbine wind-blade and solutions for their reinforcement. *Compos. Part B Eng.* **2016**, *103*, 5472. [CrossRef]
4. Famakin, S.; Kim, C. Modeling for underground cable water tree growth dynamics. *J. Power Energy Eng.* **2019**, *7*, 51. [CrossRef]
5. Hinrichsen, P.F.; Houdayer, A.; Belhadfa, A.; Crine, J.P.; Pelissou, S.; Cholewa, M. A localized trace element analysis of water trees in XLPE cable insulation by micro-PIXE and EDX. *IEEE Trans. Electr. Insul.* **1988**, *23*, 971–978. [CrossRef]

6. Matine, A.; Drissi-Habti, M. On-Coupling Mechanical, Electrical and Thermal Behavior of Submarine Power Phases. *Energies* **2019**, *12*, 1009. [CrossRef]
7. The European Subsea Cables Association. Available online: http://escaeu.org (accessed on 26 January 2022).
8. Teyssedre, G.; Laurent, C. Advances in high-field insulating polymeric materials over the past 50 years. *IEEE Electr. Insul. Mag.* **2013**, *29*, 26–36. [CrossRef]
9. Kim, C.; Jin, Z.; Huang, X.; Jiang, P.; Ke, Q. Investigation on water treeing behaviors of thermally aged XLPE cable insulation. *Polym. Degrad. Stab.* **2007**, *13*, 92. [CrossRef]
10. Kim, C.; Duan, J.; Huang, X.; Kim, S.; Jiang, P.; Kim, H.; Hyon, S. Numerical analysis on water treeing deterioration of XLPE cable insulation using combination of FEM and Taguchi method. *Eur. Trans. Electr. Power* **2010**, *13*, 749–759. [CrossRef]
11. Rajagopala, K. Electric and Thermal Properties of Wet Cable 10 kV and 15 kV. *TELKOMNIKA Indones. J. Electr. Eng.* **2012**, *10*, 1904–1916.
12. Radu, I.; Acedo, M.; Filippini, J.C.; Notingher, P.; Ftutos, F. The effect of water treeing on the electric field distribution of XLPE. Consequences for the dielectric strength. *IEEE Trans. Dielectr. Electr. Insul.* **2000**, *7*, 860–868. [CrossRef]
13. Lin, S.; Zhou, K.; Li, Y.; Meng, P. Water-tree propagation in a wide temperature range insight into the role of mechanical behavior of cross-linked poly-ethlylene (XPLE) Material. *Polymers* **2021**, *13*, 40. [CrossRef] [PubMed]
14. Ross, R. Inception and propagation mechanisms of water treeing. *IEEE Trans. Dielectr. Electr. Insul.* **1998**, *5*, 660–680. [CrossRef]
15. Capaccioli, S.; Lucchesi, M.; Casalini, R.; Rolla, P.A.; Bona, N. Effect of water inclusions on charge transport and polarization in porous media. *IEEE Trans. Dielectr. Electr. Insul.* **2001**, *8*, 454–460. [CrossRef]
16. Tao, W.; Ma, Z.; Wang, W.; Liu, J.; Zhou, K.; Li, T.; Huang, M. The mechanism of water tree growth in XLPE cables based on the finite element method. In Proceedings of the 2016 IEEE International Conference on High Voltage Engineering and Application (ICHVE), Chengdu, China, 19–22 September 2016; pp. 1–4.
17. Boggs, S.A.; Mashikian, M.S. Role of semiconducting compounds in water treeing of XLPE cable insulation. *IEEE Electr. Insul. Mag.* **1994**, *10*, 23–27. [CrossRef]
18. Crine, J.-P. Influence of electro-mechanical stress on electrical properties of dielectric polymers. *IEEE Trans. Dielectr. Electr. Insul.* **2001**, *12*, 791–800. [CrossRef]
19. Sakamoto, M.; Yahagi, K. Influence of High Electric Fields on Capacitance Measurements in PE. *Jpn. J. Appl. Phys.* **1980**, *19*, 253–259. [CrossRef]
20. Wang, Z.; Marcolongo, P.; Lemberg, J.A.; Panganiban, B.; Evans, J.W.; Ritchie, R.O.; Wright, P.K. Mechanical fatigue as a mechanism of water tree propagation in TR-XLPE. *IEEE Trans. Dielectr. Electr. Insul.* **2012**, *19*, 321–330. [CrossRef]
21. Meziani, M.; Mekhaldi, A.; Teguar, M. Impact of presence of water tree and microcavities on the electric field distribution in XLPE insulation (ICEE). In Proceedings of the 2015 4th International Conference on Electrical Engineering, Boumerdes, Algeria, 13–15 December 2015; pp. 1–5.

sensors

MDPI

Article

SDFormer: A Novel Transformer Neural Network for Structural Damage Identification by Segmenting the Strain Field Map

Zhaoyang Li [1,2,3], Ping Xu [1,2,3], Jie Xing [1,2,3] and Chengxing Yang [1,2,3,*]

[1] School of Traffic and Transportation Engineering, Central South University, Changsha 410075, China; zhaoyangli@csu.edu.cn (Z.L.); xuping@csu.edu.cn (P.X.); xingjie@csu.edu.cn (J.X.)
[2] Key Laboratory for Track Traffic Safety of Ministry of Education, Central South University, Changsha 410075, China
[3] Joint International Research Laboratory of Key Technology for Rail Traffic Safety, Central South University, Changsha 410075, China
* Correspondence: chengxing_yang_hn@163.com

Abstract: Damage identification is a key problem in the field of structural health monitoring, which is of great significance to improve the reliability and safety of engineering structures. In the past, the structural strain damage identification method based on specific damage index needs the designer to have rich experience and background knowledge, and the designed damage index is hard to apply to different structures. In this paper, a U-shaped efficient structural strain damage identification network SDFormer (structural damage transformer) based on self-attention feature is proposed. SDFormer regards the problem of structural strain damage identification as an image segmentation problem, and introduces advanced image segmentation technology for structural damage identification. This network takes the strain field map of the structure as the input, and then outputs the predicted damage location and level. In the SDFormer, the low-level and high-level features are smoothly fused by skip connection, and the self-attention module is used to obtain damage feature information, to effectively improve the performance of the model. SDFormer can directly construct the mapping between strain field map and damage distribution without complex damage index design. While ensuring the accuracy, it improves the identification efficiency. The effectiveness and accuracy of the model are verified by numerical experiments, and the performance of an advanced convolutional neural network is compared. The results show that SDFormer has better performance than the advanced convolutional neural network. Further, an anti-noise experiment is designed to verify the anti-noise and robustness of the model. The anti-noise performance of SDFormer is better than that of the comparison model in the anti-noise experimental results, which proves that the model has good anti-noise and robustness.

Keywords: structural damage identification; strain field; transformer neural network; deep learning; segmentation

Citation: Li, Z.; Xu, P.; Xing, J.; Yang, C. SDFormer: A Novel Transformer Neural Network for Structural Damage Identification by Segmenting the Strain Field Map. *Sensors* **2022**, *22*, 2358. https://doi.org/10.3390/s22062358

Academic Editors: Zahra Sharif Khodaei, Phong B. Dao, Liang Yu and Lei Qiu

Received: 18 February 2022
Accepted: 16 March 2022
Published: 18 March 2022

1. Introduction

The long-term use of engineering structures often results in structural aging and fatigue due to long-term external force load and the influence of various physical and chemical factors, resulting in changes in size, shape, and their own material properties. The resulting structural damage threatens the safe use of engineering structures. Therefore, structural damage identification has always been a key issue in the field of structural health monitoring. For large engineering structures, especially the key components that were hard to disassemble for separate maintenance, non-destructive testing (NDT) technology was mainly used to carry out regular maintenance and testing on engineering structural parts. Its technical means include magnetic particle flaw [1], ultrasonic [2], X-ray [3], eddy current [4], etc. However, these methods had some limitations. Magnetic particle testing and eddy current testing were affected by the material and structure of the tested parts, and

it was hard to detect the deep internal cracks in the structure. Ultrasonic testing was easy to operate and has realized automatic testing, but it was not suitable for the identification of complex structures, and it could not achieve real-time identification because of its low efficiency. X-ray testing had a high detection efficiency, but it had radiation impact on long-term workers. Moreover, these methods require inspectors to have rich experience and there is a need to confirm and process the test data manually. These shortcomings virtually increase the test cost.

With the progress of detection sensor technology, especially the emergence of fiber Bragg grating sensing technology [5] and digital speckle measurement technology [6], the acquisition of strain field data becomes convenient, and the damage identification method based on strain modal data is widely used because of the high sensitivity of strain to local damage. In the field of structural damage identification based on strain mode information, the traditional method was to design a sensitive and accurate damage index to identify the damage in the tested structure. Kaewunruen et al. [7] proposed a damage identification index based on curvature analysis based on the strain measured by the fiber Bragg grating strain sensor, which was used to identify the damaged cavity under railway sleeper. WL Bayissa et al. [8] proposed spectral or mean strain energy (SSE) for structural damage identification of plate-like structures, and a large number of numerical simulation studies show that SSE index had high sensitivity to damage and good identification performance. Fariba et al. [9] proposed a strain frequency response function (SFRF) for strain damage identification of structures, and studies showed that their method had low error rate and strong anti-noise ability.

However, the method of designing strain damage index needs a deep understanding and background knowledge of the detected structure, and it was hard to migrate to different application scenarios and different structures. Due to the popularity of sensors, structural health monitoring has higher and higher requirements for real-time monitoring, and a large number of sensor data need to be processed. The method based on strain damage index struggled to overcome the problem of low data processing efficiency caused by data explosion, and the environmental noise and unpredictable load in the actual scene would lead to large deviation of strain damage index.

In recent years, with the rapid development of deep learning technology and big data technology, the data-driven structural health monitoring had become a research hotspot. Osama Abelijaber et al. [10] achieved damage location of truss steel structure by using a one-dimensional convolutional neural network. Seokgoo Kim [11] proposed a convolutional neural network based on signal segmentation to identify the damage of different gears and achieved good results. Bao [12] imagized one-dimensional time series data and trained the deep neural network on this basis, which was applied to abnormal acceleration data detection of a long-span cable-stayed bridge, and achieved an accuracy of 87% with satisfactory results. Hyo et al. [13] used convolutional neural network to predict the strain changes of key structural points of concrete, to evaluate the life of concrete structures. Jinhua Lin et al. [14] used X-ray detection images of forged parts to train convolutional neural network for damage identification of forged parts, and obtained an accuracy rate of more than 96% with high identification efficiency. Zhang et al. [15] proposed an improved deep learning algorithm for crack identification on asphalt pavement. Their improvement gave the algorithm higher accuracy and significantly reduced the probability of misjudgment. Kumar et al. [16] proposed a two-stage fault detection and classification method designed for a rotating motor, which combines numerical features and shallow neural network. Jin and Chen [17] proposed a novel end-to-end intelligent vibration signal classification framework based on modified transformer neural network, and proved that the proposed method is better than the signal classification method based on convolution neural network and recurrent neural network through two cases.

The structural damage location and identification method based on deep learning could adaptively extract the structural feature information from the original response data for damage identification. However, the current research is still limited to images or

strain mode data of several key nodes for structural damage identification. Those methods were restricted by external conditions such as insufficient illumination and detection points, which limits its application scope. There was still a lack of a mapping method between strain field data and structural damage information to realize real-time damage identification. In recent years, the rapid development of deep learning in the field of image semantic segmentation provides a new way to realize this idea. Fully convolutional network (FCN) [18] was the first end-to-end deep convolution neural network (CNN) for image semantic segmentation. It brings image semantic segmentation into the era of deep learning. After FCN, the DeepLab [19] series and the PSPnet [20] network had brought far-reaching hole convolution and feature pyramid structure, respectively, and their design ideas have influenced others so far. UNet is the first deep convolution neural network using symmetrical encoder-decoder structure for medical image segmentation. After combining with residual network (ResNet) [21], the effect has been greatly improved, and its simple symmetrical structure has been borrowed by most later deep neural networks. Since 2021, the vision transformer network, represented by ViT [22], has defeated the traditional CNN network in many tasks. In particular, the emergence of the swin transformer [23] has brought the whole field of image segmentation into a new era. The accuracy of image semantic segmentation is improving with the development of deep learning. At present, there is still no relevant research trying to use image segmentation technology to segment the strain field for structural damage identification. However, the research on damage assessment of composite laminated structures by Tran-Ngoc et al. [24] and Samir Khatir et al. [25] gives enlightenment to this study. In their research, the composite structure to be detected is first meshed, and then damage identification is carried out based on specific damage indicators combined with artificial neural network. Their research is still limited to designing specific damage indicators for damage identification, but the idea of meshing the detection area and then identifying each grid is worth learning from. Therefore, when the mesh of the structure is dense enough, the problem of structural damage identification can be regarded as an image segmentation problem, and then the current advanced depth image segmentation algorithm can be used for structural damage identification.

In this paper, the structural damage identification problem based on strain field data is regarded as an image semantic segmentation problem. With the help of deep learning technology, a structural strain damage identification network combining encoder–decoder configuration and advanced transformer structure, SDFormer, is proposed. This neural network was trained by the simulation results, and the predicted damage map of the structure was predicted from the strain field map. Compared with the real damage map, the feasibility of the model was verified.

To present our method, this paper is organized as follows: Section 2 introduces the process of the damage identification method and the proposed network architecture, Section 3 describes the setup of the numerical experiment and results of the method, Section 4 discusses the advantages and disadvantages of the method, and Section 5 contains the conclusion of this paper.

2. Proposed Method

2.1. Overall Architecture

In this paper, a U-shaped network, structural damage transformer (SDFormer), which is based on swin transformer [23], is proposed for structural strain damage identification. The network takes the strain field as input and outputs the corresponding damage position and level. The structural strain damage identification framework in this paper includes three stages: generate strain damage dataset stage, training stage, and online testing stage, as shown in Figure 1.

In the generate dataset stage, there are two ways to build the dataset: using the strain field map detected by sensors or using the strain field map simulated by finite element model. The dataset should include two parts: 1. strain field map; 2. the real damage map, which should correspond to the strain field map one-to-one. The cost of tagging the strain

field map dataset detected by sensors is prohibitive. Therefore, in this paper, the strain field map simulated by finite element model is used to build the dataset. We preset the structural damage location and level, and built the corresponding finite element model, then obtained the simulated strain field map as the input of the neural network, and the preset structural damage map as the real damage map label of the neural network. The setting method of simulated damage is introduced in Section 3.

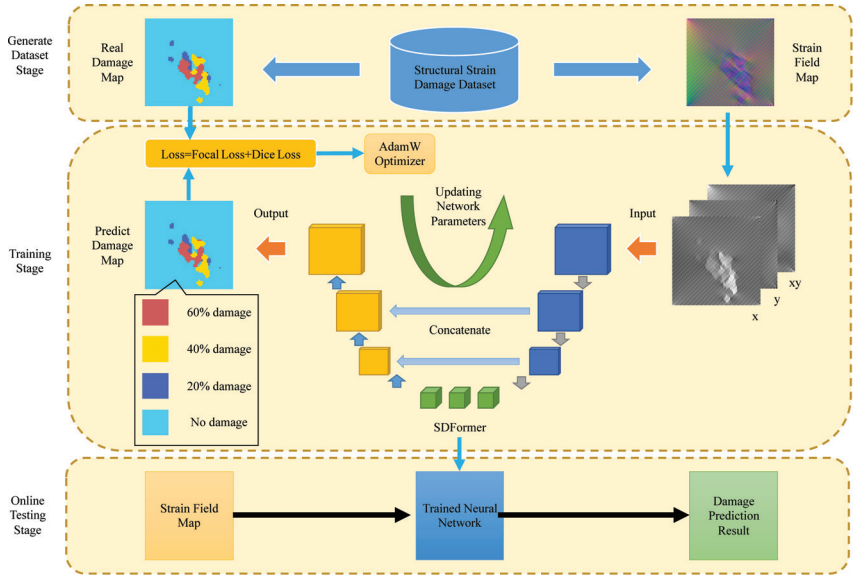


Figure 1. The structural strain damage identification framework.

In the training stage, the strain field data are processed into tensors with the size of $H \times W \times N$, where H and W represent the length and width of the strain field, and N is the number of channels of the strain field data. It corresponds to the strain data in x, y, and xy directions respectively, $N = 3$. Then, the processed strain field data is used as the input of the SDFormer. The SDFormer outputs the prediction probability tensor P with the size of $H \times W \times Cls$, where H and W correspond to the size of the strain field, Cls is the number of predicted damage categories, $Cls = d + 1$, where d is the number of preset strain damage levels, and 1 refers to the normal category without damage. Three strain damage levels of 20%, 40%, and 60% are preset in this paper. Therefore, Cls set in this paper is 4. $P_{i,j,k}$ is the probability of k-th damage of the element with coordinates (i, j). When $k = 0$, it represents the probability of no strain damage of the element. The real damage map T is processed into a [0,1] tensor with the size of $H \times W \times Cls$, too. $T_{i,j,k} = 1$ means that the element with coordinate (i, j) produces the k-th damage. Then, P and T will be used to calculate the gradient of the loss function to the SDFormer parameters, and use the AdamW [26] algorithm to update the SDFormer parameters. This process is cycled until loss converges to obtain the trained SDFormer.

In the online testing stage, the real-time detected strain field data is processed into the format which is required by the network and input into SDFormer, then the output of SDFormer is the predicted damage map.

2.2. SDFormer

The architecture of the proposed SDFormer in this paper is presented in Figure 2. The SDFormer consists of patch block, encoder block, bottleneck, decoder block, and seghead, in which the symmetrical encoder and decoder blocks are connected by skip connection.

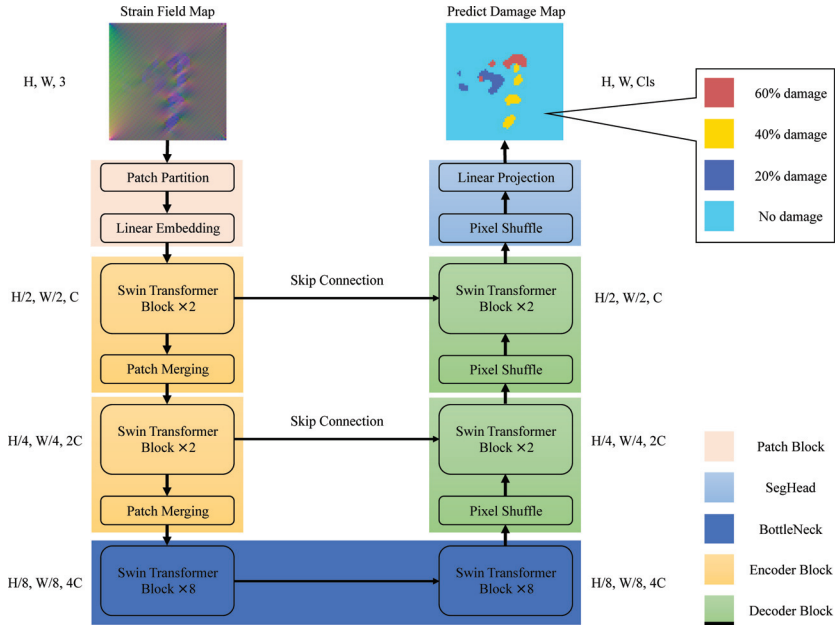


Figure 2. Overview of SDFormer network architecture.

In the patch block, the strain field map S would be split into several non-overlapping patches with a size of 2×2, and then flattened. These flattened patches would be transformed into vectors with length C by an linear embedding layer. The encoder block consists of two swin transformer blocks and a patch merging layer. The feature map from the previous layer would be processed by the swin transformer block first, and the processed feature map is retained. The processed feature map would be sent to the patch merging layer and the decoder block symmetrical to the current encoder block at the same time. Then, patch merging downsamples the feature map to obtain the global feature information and send it to the next layer. This design is very common in the field of computer vision. In convolutional neural networks, a similar operation is the pooling layer. In fact, patch merging is similar to patch block, except that it further merges the split patches, so we call it patch merging. The bottleneck consists of two groups of symmetrical swin transformer blocks. We could control the complexity of the model to meet the actual needs by adjusting the number of swin transformer blocks in bottleneck. The decoder block performs the opposite operation to the encoder block. The input feature map first goes through a pixel shuffle [27] layer for upsampling, and then it would be concatenated with the feature map from the symmetrical encoder block in the channel dimension and sent to the swin transformer blocks, and the processed feature map would be sent to the next layer.

Finally, in SegHead, a pixel shuffle layer restores the feature map to the same size as the input strain field map, and then obtains the final damage prediction probability tensor P through the linear projection layer, which is a fully connected layer.

2.2.1. Patch Block and Patch Merging

In the patch block, the strain field map S is split into non-overlapping patches with a size of 2×2, and then flattened, as shown in Figure 3a. In this way, the size of S will change from $H \times W \times 3$ to $\frac{H}{2} \times \frac{W}{2} \times 12$. Then, a fully connected layer is used to embed each patch, and the feature dimension of each patch is transformed into C. In this paper, $C = 64$. The size of the output feature map F is $\frac{H}{2} \times \frac{W}{2} \times C$.

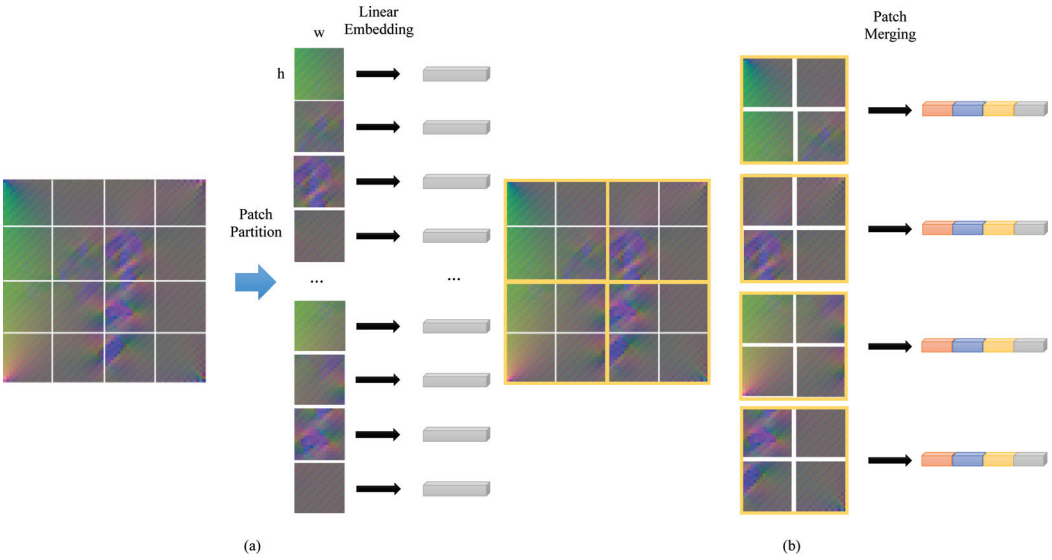


Figure 3. Patch block and patch merging: (**a**) Patch block: splits the strain field map and embeds them. (**b**) Patch merging: concatenates the patch and re-embeds.

Patch merging merges the patches of the previous layer feature map, as shown in Figure 3b. Similar to patch block, patch merging uses a 2×2 sliding window to split the feature map and merge the patches in the window, so that its size changes from $H \times W \times C$ to $\frac{H}{2} \times \frac{W}{2} \times 4C$. Then, a fully connected layer is used to reduce the dimension of the embedded feature from $4C$ to $2C$. Thus, the size of the feature map output by patch merging is $\frac{H}{2} \times \frac{W}{2} \times 2C$. The calculation formula of patch merging is shown in Formulas (1) and (2).

$$x^l = MLP(PatchMerging(x^{l-1})) \tag{1}$$

$$MLP(x) = wx + b \tag{2}$$

where $w \in R^{4C \times 2C}$ is the weight of the full connection layer and $b \in R^{2C}$ is the bias of the full connection layer. It is worth noting that $w \in R^{4C \times 2C}$ only affects the embedded feature dimension.

2.2.2. Pixel Shuffle

Pixel shuffle [27] is a parameterless upper sampling layer. Its principle is to reorder an input tensor to complete the upper sampling of the input tensor. Its mathematical expression can be written as Formula (3)

$$PixelShuffle(T)_{x,y,c} = T_{\lfloor x/r \rfloor, \lfloor y/r \rfloor, C \cdot mod(y,r) + C \cdot mod(x,r) + c} \tag{3}$$

where $T \in R^{H \times W \times C \cdot r^2}$ is the input tensor, r is the magnification of upsampling, and C is the number of channels of the output tensor. When the size of input tensor is $H \times W \times C \cdot r^2$, the size of output tensor is $r \cdot H \times r \cdot W \times C$.

In addition, a fully connected layer is added in front of the pixel shuffle layer to adjust the embedded feature dimension of the input feature map to meet the requirements of the pixel shuffle layer. In the SDFormer, all feature maps passing through the pixel shuffle layer will be upsampled to twice the original size.

2.2.3. Swin Transformer Block

The basic unit of SDFormer is the swin transformer block, which is an important part in the encoder block, decoder block, or bottleneck. Different to the traditional multi-head self-attention module, the swin transformer block is constructed by the multi-head attention module based on fixed window and shifted window. The structure of the swin transformer block is shown in Figure 4a, which includes four layernorm (LN) layers, a window-based multi-head self-attention module (W-MSA), a shifted window-based multi-head self-attention module (SW-MSA), and two multilayer perceptron (MLP) layers with nonlinear activation function GeLU.

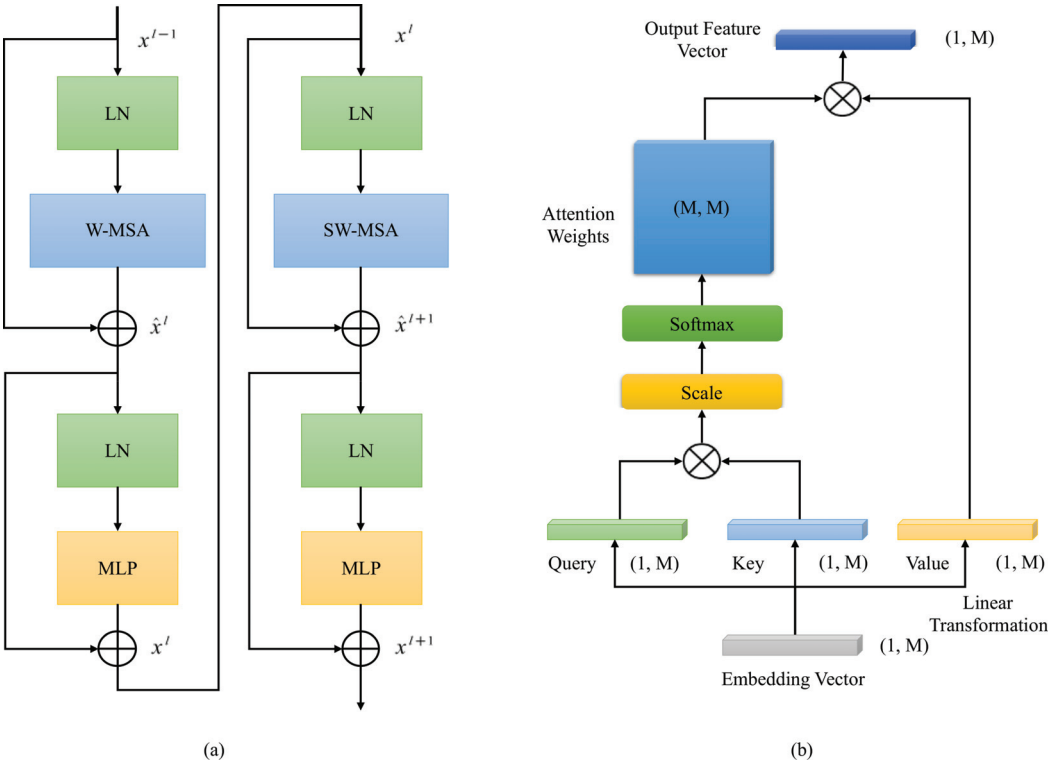


Figure 4. Overview of swin transformer block and the multi-head self-attention (MSA) module. (**a**) Architecture of swin transformer block. (**b**) Calculation flow of multi-head self-attention (MSA) module.

W-MSA splits the feature map into several non-overlapping windows, and then calculates the local self-attention in each window, which can reduce the computational complexity significantly while maintaining the global attention. At the same time, SW-MSA is proposed to solve the disadvantage of lack of cross-window connection in W-MSA and enhance the capability of the whole model, and cross-window connection is introduced on the basis of SW-MSA and maintains its efficient computing power. The difference between SW-MSA and W-MSA is that a window shift operation is added before W-MSA. This small change makes the whole network model better able to deal with the global information. In practical implementation, shifting the feature map achieves the same effect. W-MSA and SW-MSA are used alternately to form a swin transformer block, which can be expressed by the following formula:

$$\hat{x}^l = \text{W-MSA}\left(LN\left(x^{l-1}\right)\right) + x^{l-1} \tag{4}$$

$$x^l = MLP\left(LN\left(\hat{x}^l\right)\right) + \hat{x}^l \tag{5}$$

$$\hat{x}^{l+1} = \text{SW-MSA}\left(LN\left(x^l\right)\right) + x^l \tag{6}$$

$$x^{l+1} = MLP\left(LN\left(\hat{x}^{l+1}\right)\right) + \hat{x}^{l+1} \tag{7}$$

$$LN(x) = \frac{x - E[x]}{\sqrt{Var(x) + \epsilon}} \cdot \gamma + \beta \tag{8}$$

where $\hat{x}^l$ and x^l are the feature maps output by the ith (S)W-MSA module and MLP module, respectively, $E[x]$ is the mean value of x, $Var(x)$ is the variance of x, ϵ is usually 1×10^{-5}, and γ and β are learnable affine transformation parameters. When calculating self-attention, similar to previous studies [28,29], we first use an MLP layer to map the input embedding feature vector into three vectors of the same size: query, key, and value. Then we calculate the similarity between query and key, obtain an attention weight matrix after passing through a softmax layer, and then multiply the attention weight matrix and the value to obtain the final output. The formula of self-attention is as follows:

$$Attention(Q, K, V) = SoftMax\left(\frac{QK^T}{\sqrt{d}} + B\right)V \tag{9}$$

where $Q, K, V \in R^{M^2 \times d}$ are query, key, and value respectively, M is the window size for calculating local self-attention, d is the dimension size of query or key, and $B \in R^{M^2 \times M^2}$ represents an offset matrix. Since storing an $M^2 \times M^2$ matrix requires a lot of memory, and the relative position of the patch in the window is often between $[-M+1, M-1]$, a $\hat{B} \in R^{(2M-1)\times(2M-1)}$ is usually used to save the value of B.

2.3. Loss and Optimizer

Subtle strain damage is very common in engineering structures. This makes the area of strain damage account for a small part in the strain field map. Then, there is a serious sample imbalance in the training data, which makes the model hard to fit. In order to solve this problem, the following loss function is designed for model training:

$$L(y, p) = L_{dice}(y, p) + L_{focal}(y, p) \tag{10}$$

$$L_{dice}(y, p) = 1 - \frac{2|y \cap p|}{|y| + |p|} \tag{11}$$

$$L_{focal}(y, p) = -\alpha y(1-p)^{\gamma} log(p) - (1-\alpha)(1-y)p^{\gamma} log(1-p) \tag{12}$$

where y is the real damage map, p is the predicted damage map, $|y \cap p|$ is the area of the intersection area between the real damage map and the predicted damage map, $|y|$ is the area of the real damage map, $|p|$ is the area of the predicted damage map, α and γ are adjustable super parameters, α defaults to 0.5, and γ defaults to 2. It can be seen that when the predicted damage distribution is more similar to the real damage distribution, the dice loss (L_{dice}) [30] is smaller. However, when dice loss encounters small damage that is difficult to distinguish, there will be a problem that the overall loss value is very small but the prediction effect is poor, resulting in difficulty in model convergence. Therefore, we add a focal loss (L_{focal}) [31] item to increase the loss of these difficult samples, making the network more sensitive to such damage. In the training process, AdamW [26] algorithm was used to train the model. AdamW algorithm is an improved version of Adam [32], and its mathematical formula is as follows:

$$g_t = \nabla f_t(\theta_{t-1}) \tag{13}$$

$$m_t = \beta_1 m_{t-1} + (1-\beta_1) g_t \tag{14}$$

$$v_t = \beta_2 v_{t-1} + (1-\beta_2) g_t^2 \tag{15}$$

$$\hat{m}_t = m_t / (1-\beta_1^t) \tag{16}$$

$$\hat{v}_t = v_t / (1-\beta_2^t) \tag{17}$$

$$\theta_t = \theta_{t-1} - \eta_t \left(\alpha \hat{m}_t / (\sqrt{\hat{v}_t} + \epsilon) + \lambda \theta_{t-1} \right) \tag{18}$$

where θ_t is the network model parameter at time t; g_t is the gradient of model parameters at time $t-1$; η_t is the learning rate at time t; m_t is the first moment vector, $m_{t=0} = \mathbf{0}$; v_t is the second moment vector, $v_{t=0} = \mathbf{0}$; $\hat{m}_t$ is the first moment vector after correcting the offset; $\hat{v}_t$ is the second moment vector after correcting the offset; β_1 and β_2 default to 0.9 and 0.999 respectively; α defaults to 0.001; ϵ defaults to 1×10^{-8}; λ is the weight attenuation coefficient, and it defaults to 0.01.

3. Numerical Experiments and Results

In this section, two finite element simulation datasets were constructed to verify the effectiveness of the proposed SDFormer. Moreover, this study compares SDFormer with three popular CNN methods, including UNet [33], PSPnet [20], and DeepLabV3 [19], where ResNet-50 [21] was taken as the backbone network, to evaluate the performance of the SDFormer. The loss function and optimization algorithm used in these network training are consistent with SDFormer, as mentioned in Section 2.3.

3.1. Numerical Experiments Setup

3.1.1. Damage Simulation

It is a simple and practical method to simulate material damage by replacing the elastic modulus of the original material with that of the damaged material. This method, based on strain equivalence hypothesis, has been widely used because of its simple implementation. In the strain equivalence hypothesis, the deformation of damaged material due to stress is equivalent to the deformation of a hypothetical non-destructive material under stress, and the actual effective bearing area of damaged material is equal to the virtual bearing area of non-destructive material. Its mathematical formula is as follows:

$$\frac{\sigma'}{E} = \frac{\sigma}{E'} \tag{19}$$

$$\sigma' = \frac{\sigma}{1-D} \tag{20}$$

where E is the elastic modulus of undamaged material, E' is the elastic modulus of damaged material, σ is the nominal stress, and σ' is the effective stress; D is the damage factor, and its value is between 0 and 1. Combining Formulas (19) and (20):

$$D = 1 - \frac{E'}{E} \tag{21}$$

It can be seen that when there is no damage, $D = 0$, and the elastic modulus of the material does not change. When complete damage occurs, $D = 1$, and the elastic modulus of the material becomes 0. Based on Formula (21), this study sets up three kinds of damage in the case of $D = 0.2$, $D = 0.4$, and $D = 0.6$ to construct the datasets.

In Case 1 and Case 2, the damage simulation is carried out according to the following steps:

1. For the structural component S, the material Mat of S and the corresponding elastic modulus E are known. The preset damage level $D \in (0.2, 0.4, 0.6)$ defines four

materials (three damaged materials and one undamaged material), in which the undamaged material corresponds to *Mat*, the elastic modulus of the three damaged materials is set as $E' = E(1 - D)$, according to Formula (21), and other material properties are consistent with *Mat*.

2. Mesh *S*, and randomly select an area not exceeding 5% of the total area of the structure for each damaged material. If the selected area of two damaged materials overlaps, the one set later is the material of the overlapping area. In addition, other conditions are consistent, and a damage sample finite element model is obtained.
3. The damage map and corresponding strain field data in X, Y, and XY directions are obtained by solving the finite element model obtained in step 2.
4. Repeat steps 2 and 3 until enough data samples are obtained to construct the structural strain damage dataset.

3.1.2. Case 1

In this case, a simply supported plate model was constructed, as shown in Figure 5. The simply supported plate is a flat plate with a size of 640 mm × 640 mm composed of shell elements. One end of the plate was fixed, and the other end was applied with a uniformly distributed load of 1000 N/cm in the positive direction of the *X*-axis.

The default material of the plate is aluminum, the elastic modulus is 70 Gpa, the Poisson's ratio is 0.33, and the density is 2700 kg/m^3. In this case, three damage levels were set as mentioned in Section 3.1.1, i.e., 20%, 40%, and 60% material damage. Three hypothetical materials were defined to simulate this damage. The elastic moduli of these three hypothetical materials were 80%, 60%, and 40% of aluminum, respectively. Other material properties are consistent with aluminum.

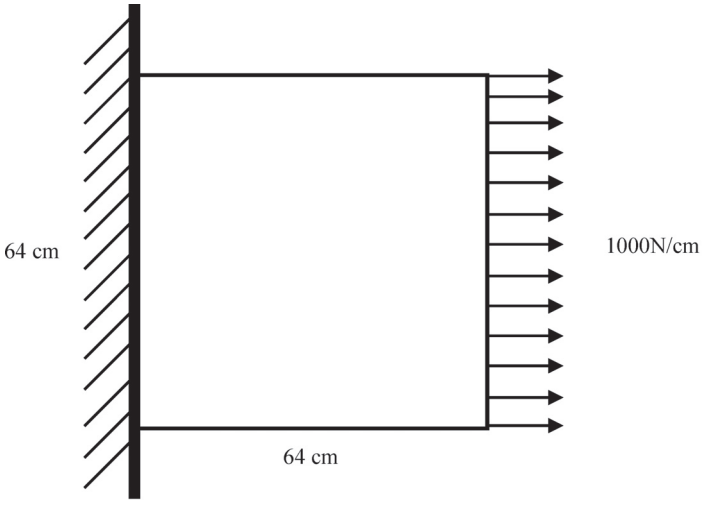


Figure 5. Schematic diagram of plate finite element model.

Following the four steps of damage simulation in Section 3.1.1, the plate model was meshed according to the size of 10 mm × 10 mm, and 4096 finite elements were obtained. The strain field data obtained by solving the finite element model would be saved as a tensor with the size of 64 × 64 × 3 as the strain field map in the dataset, and the damage selection area would be saved as the real damage map $T^{(64\times64)} = (t_{ij})$, where $t_{ij} \in (0, 1, 2, 3)$ refers to the damage level $D = (0, 0.2, 0.4, 0.6)$. This process is repeated to obtain 2000 groups of data samples under different damage conditions. In this paper, a Python script was used to call ABAQUS to build this dataset. And an example of the plate dataset is shown in Figure 6.

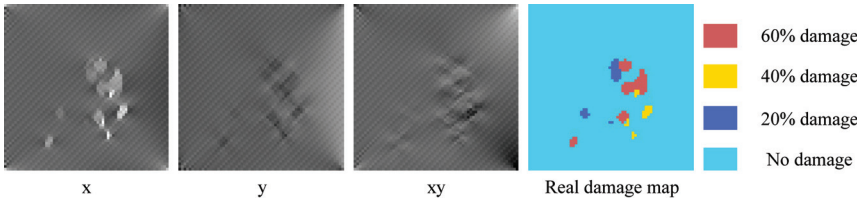


Figure 6. Example of plate dataset.

3.1.3. Case 2

In this case, a strain damage dataset of a sleeper beam of a metro train is constructed. The sleeper beam of the metro train is shown in Figure 7a, and is composed of a cover plate, a bottom plate, and reinforcing structures in the middle.

This case takes the strain damage identification of the cover plate of a sleeper beam as the research task, as shown in Figure 7b. The cover plate of the sleeper beam is a rectangular non-perforated plate. As an important component of bearing and connecting the vehicle body and traveling part, the state of the cover plate is related to the driving safety of the train. However, the damage of the sleeper cover plate is hard to detect because of the particularity of the sleeper position. Thus, this case expects to use the method proposed in this paper to conduct the real-time strain damage identification of the sleeper beam cover plate.

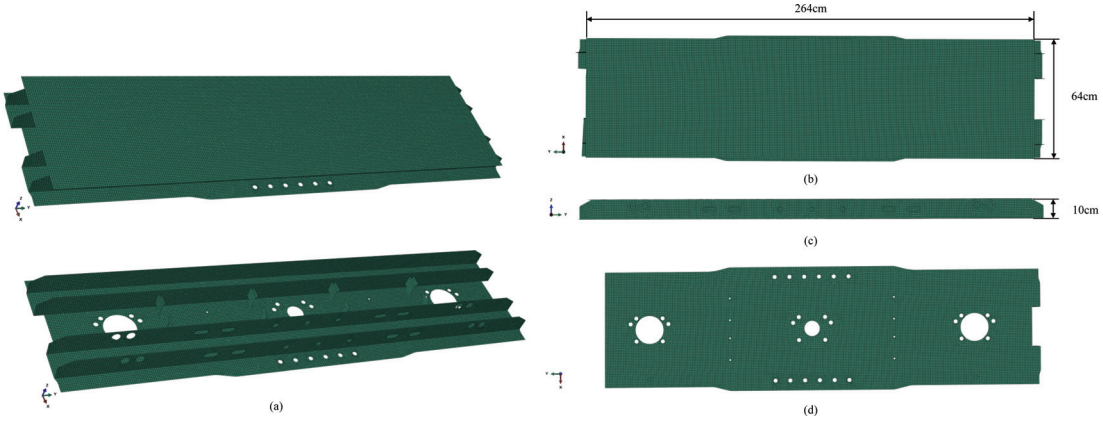


Figure 7. Schematic diagram of sleeper beam. (**a**) Overall structure of the sleeper beam. (**b**) Top view of the sleeper beam, the length of the cover plate is 264 cm and the width is 64cm. (**c**) Front view of the sleeper beam, the height of the sleeper beam is 10 cm. (**d**) The lower cover plate of the sleeper beam.

In this case, the size of the cover plate of the sleeper beam was 2640 mm $\times$ 640 mm $\times$ 10 mm. The bottom plate of the sleeper beam is fixed, and a uniform load of 0.3 Mpa in the $-z$ direction is applied to the cover plate. The default material of the sleeper beam is steel, the elastic modulus is 206 GPa, the Poisson's ratio is 0.3, and the density is 7850 kg/m^3. Similar to Case 1, this case also sets the same three damage levels. Three hypothetical materials were defined to simulate this damage. The elastic moduli of the three hypothetical materials are 80%, 60%, and 40% of steel, respectively. Other material properties were consistent with steel.

Following the four steps of damage simulation in Section 3.1.1, the cover plate was meshed according to the size of 10 mm $\times$ 10 mm, and 16,896 finite elements were obtained. The strain field data obtained by solving the finite element model would be saved as a tensor with the size of 264 $\times$ 64 $\times$ 3 as the strain field map in the dataset, and the damage selection area would be saved as the real damage map $T^{(264\times64)} = (t_{ij})$, where $t_{ij} \in (0, 1, 2, 3)$ refers

to the damage level $D = (0, 0.2, 0.4, 0.6)$. This process is repeated to obtain 2000 groups of data samples under different damage conditions. The dataset of this case is also built by calling ABAQUS with Python script. And an example of the sleeper beam dataset is shown in Figure 8.

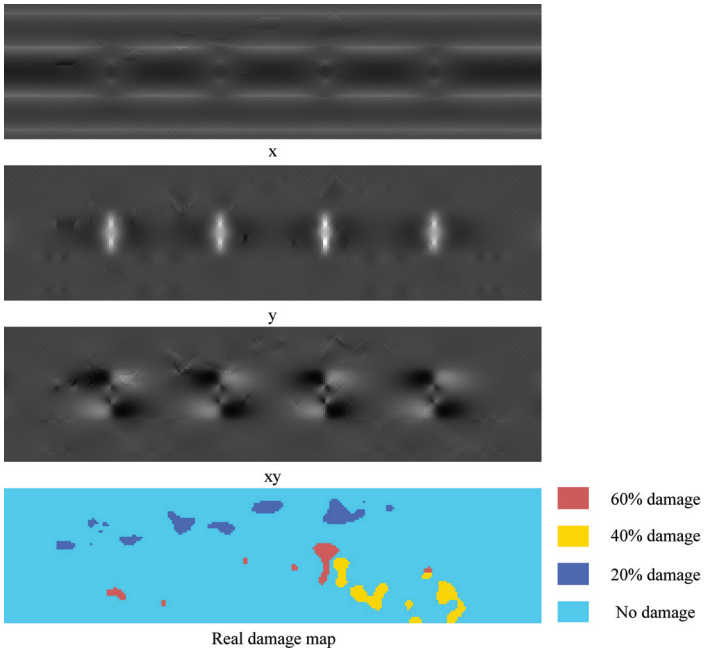


Figure 8. Example of sleeper beam dataset.

3.1.4. Evaluation Metrics

In this paper, mean intersection over union (MIoU) is used to evaluate the performance of each model. MIoU is calculated according to the following formula:

$$MIoU = \frac{1}{k+1} \sum_{i=0}^{k} \frac{p_{ii}}{\sum_{j=0}^{k} p_{ij} + \sum_{j=0}^{k} p_{ji} - p_{ii}} \tag{22}$$

where p_{ij} is the number of class i damage units predicted as class j damage in the network output results, and k is the number of damage categories.

3.2. Result

In this section, the datasets of Case 1 and Case 2 were divided into training set and testing set according to the ratio of 8:2, respectively, to test the performance of PSPnet [20], DeepLabV3 [19], UNet [33], and SDFormer. The maximum training epochs of all models including comparison models and the proposed model are 50; the batch size is 4; the initial learning rate is set to 10^{-5}; and AdamW algorithm is selected as the model optimization algorithm.

Figure 9 shows the training loss curve and MIoU curve of DeepLabV3, PSPnet, UNet, and SDFormer on the plate dataset and the sleeper beam dataset. It could be found that under the same training conditions, SDFormer can make the loss function converge faster and obtain better MIoU performance than the other three models. It means that SDFormer has stronger capacity to extract features from the strain field map and a higher performance upper limit. This is very important for the accurate detection of structural damage.

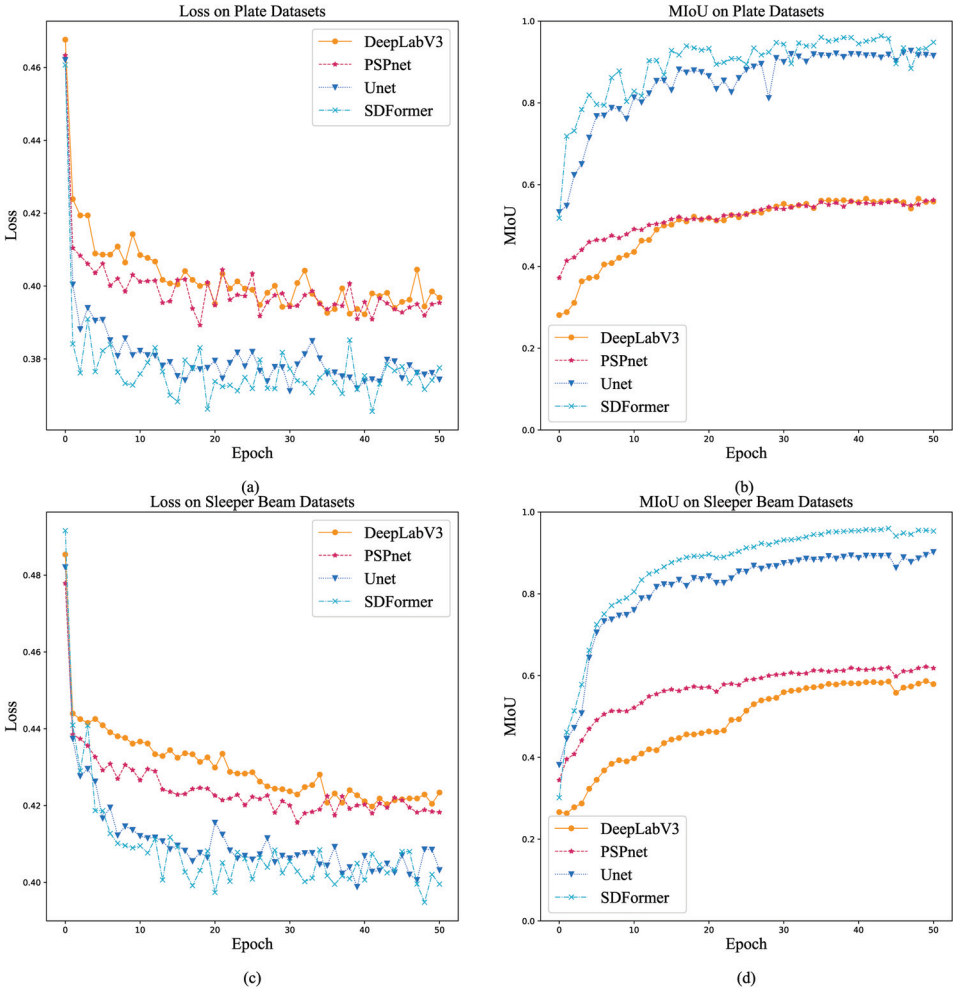


Figure 9. Training loss curve and MIoU curve. (**a**) Training loss curve on plate dataset. (**b**) Training MIoU curve on plate dataset. (**c**) Training loss curve on sleeper beam dataset. (**d**) Training MIoU curve on sleeper beam dataset.

Figure 10 compares the test results of PSPnet, DeepLabV3, UNet, and SDFormer on the plate dataset, and Figure 11 compares the test results on the sleeper beam dataset. Comparing the test results of each model on the two datasets, it could be found that for structural damage, DeepLabV3 and PSPnet could only give the approximate damage location, but could not predict the detailed parts of the damage. In contrast, UNet and SDFormer could give a prediction map very close to the real damage map. Further observing the white dashed circle in Figures 10 and 11, it can be seen that SDFormer predicts the edge of the structural damage area more smoothly and accurately than UNet. This is due to the powerful self-attention module of the swin transformer block.

Tables 1 and 2 summarize the MIoU performance, floating-point operations per second (FLOPs), and number of parameters (params) of DeepLabV3, PSPnet, UNet, and SDFormer on the plate dataset and the sleeper beam dataset. It can be seen that the MIoU performances of SDFormer and UNet far exceed those of PSPnet and DeepLabV3 in both datasets, and they were in an absolute leading position. Compared with UNet, SDFormer has

higher MIoU performance and less parameters. Higher FLOPs means that the calculation requirements of SDFormer would be higher, but it is still within the acceptable range. Moreover, because the structure of the sleeper beam is more complex than that of the plate, the MIoU performance of each model was reduced on the sleeper beam dataset.

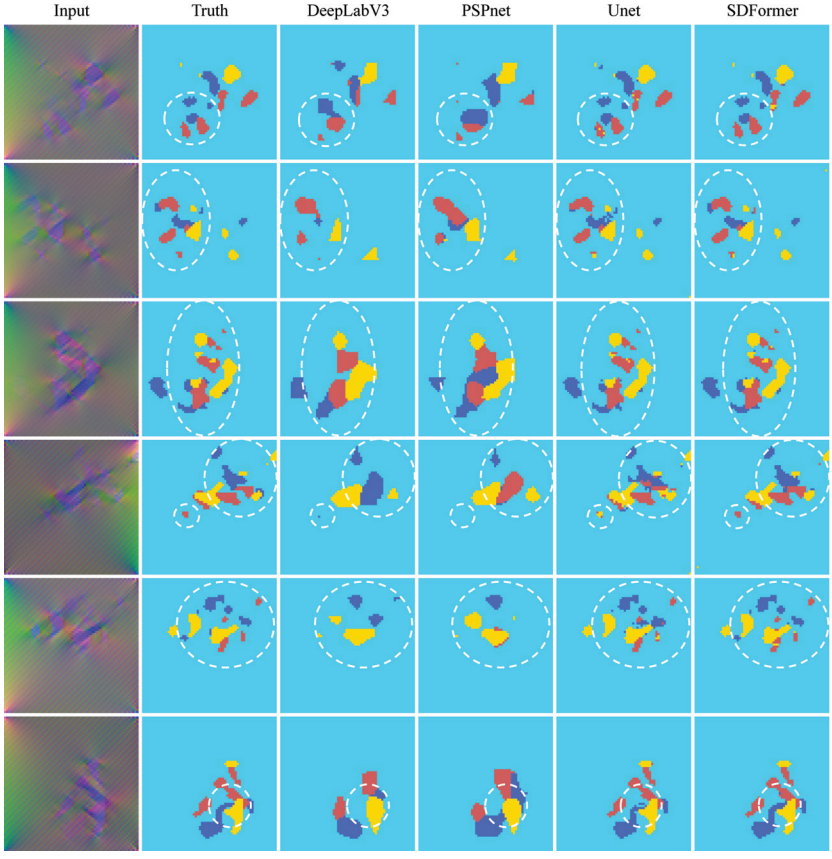


Figure 10. Test result of DeepLabV3, PSPnet, UNet, and SDFormer on the plate dataset.

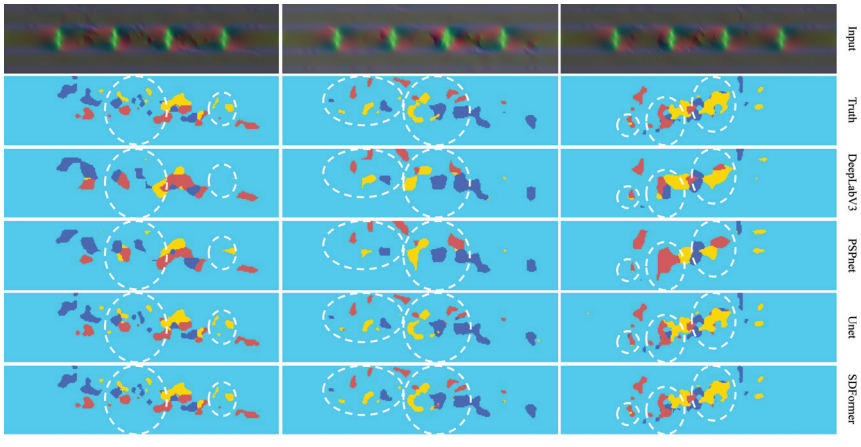


Figure 11. Test result of DeepLabV3, PSPnet, UNet, and SDFormer on the sleeper beam dataset.

Table 1. Result on plate dataset.

Methods	Backbone	Input Size	MIoU. (×100%)	FLOPs	Params. (M)
DeepLabV3	ResNet-50	64 × 64	55.97	2561 M	39.63
PSPnet	ResNet-50	64 × 64	56.93	161 M	1.75
UNet	ResNet-50	64 × 64	95.21	669 M	32.5
SDFormer(ours)	-	64 × 64	97.07	1807 M	18.8

Table 2. Result on Sleeper beam dataset.

Methods	Backbone	Input Size	MIoU. (×100%)	FLOPs	Params. (M)
DeepLabV3	ResNet-50	64 × 264	58.65	102.41 G	39.63
PSPnet	ResNet-50	64 × 264	62.05	6.51 G	1.75
UNet	ResNet-50	64 × 264	90.21	26.74 G	32.5
SDFormer(ours)	-	64 × 264	95.93	72.29 G	18.8

3.3. Anti-Noise Experiment Setup

Anti-noise performance is also a key capability for industrial applications. In order to test the anti-noise performance of SDFormer proposed in this paper, anti-noise experiments based on plate dataset and sleeper beam dataset are set up in this section. In this section, an anti-noise experiment is set up to observe the MIoU performance transformation of the model trained with noiseless data when the noise level increases. In the anti-noise experiment, the normal distribution noise was directly added to the strain field map to simulate the noise interference in the actual detection, and the model trained in Section 3.2 would be used for testing to observe the change of their MIoU performance with the increase of noise level. The noise setting formula is as follows:

$$\hat{x} = (1 + \mu \cdot N)x \tag{23}$$

where x is the original strain field map, $\hat{x}$ is the strain field map with noise, N is the normal distribution random number tensor, its size is consistent with x, and μ is the noise level.

In this section, five noise levels of 4%, 8%, 12%, 16%, and 20% are set for testing. The evaluation metric of anti-noise experiment was MIoU. Similarly, this section compares the anti-noise performance of four pretrained models, PSPnet, DeepLabV3, UNet, and SDFormer, on the plate dataset and the sleeper beam dataset.

3.4. Anti-Noise Result

Figure 12 shows the trend of MIoU performance of DeepLabV3, PSPnet, UNet, and SDFormer models in plate dataset and sleeper beam dataset with the improvement of noise level, and Tables 3 and 4 summarize the anti-noise performance of the proposed model and comparison of models in these two datasets. With the improvement of noise level, the MIoU performance of each model decreases gradually. Compared with the other three models, SDFormer could always maintain the best MIoU performance. This shows that SDFormer has good anti-noise ability and can maintain a good identification accuracy under certain noise conditions.

Table 3. MIoU of anti-noise test on plate dataset.

Methods	$\mu = 0\%$	$\mu = 4\%$	$\mu = 8\%$	$\mu = 12\%$	$\mu = 16\%$	$\mu = 20\%$
DeepLabV3	55.97	55.63	51.74	46.40	42.51	39.48
PSPnet	56.93	55.86	53.43	50.23	46.31	42.51
UNet	95.21	94.13	82.61	69.38	58.37	52.11
SDFormer(ours)	97.07	95.78	85.37	73.56	62.71	54.79

Table 4. MIoU of anti-noise test on sleeper beam dataset.

Methods	$\mu = 0\%$	$\mu = 4\%$	$\mu = 8\%$	$\mu = 12\%$	$\mu = 16\%$	$\mu = 20\%$
DeepLabV3	58.65	58.31	57.89	57.08	55.01	51.46
PSPnet	62.05	61.61	61.54	59.82	57.11	52.98
UNet	90.21	89.09	83.60	75.25	67.29	60.17
SDFormer(ours)	95.93	94.96	88.93	82.91	73.62	64.17

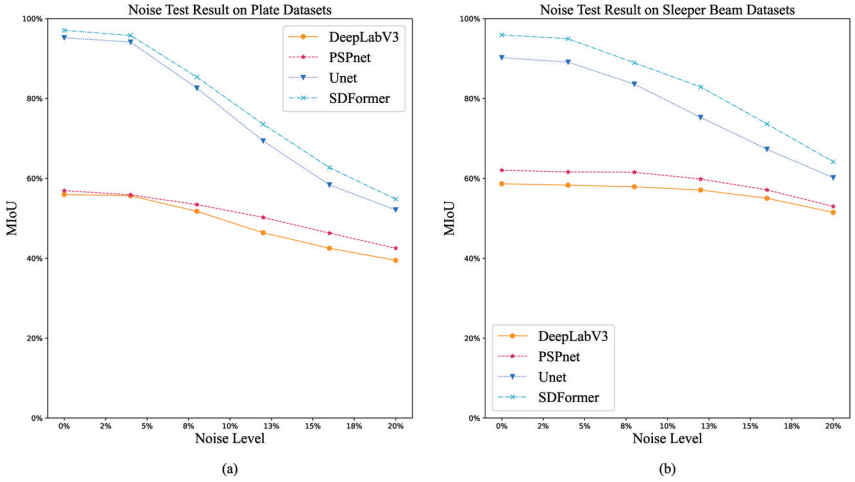


Figure 12. Anti-noise test result. (**a**) Anti-noise test result on the plate dataset. (**b**) Anti-noise test result on the sleeper beam dataset.

4. Discussion

Through the above comparison between numerical results and anti-noise experiment, it was found that the effects of SDFormer and UNet on damage segmentation are much better than DeepLabV3 and PSPnet because they both use the structure of skip connection. The skip connection structure makes the fusion of low-level features and high-level features of the model smoother. More low-level features could be effectively retained, which improves the prediction performance of the whole model. Further comparing the test results in Figures 10 and 11, it could be found that the prediction result of SDFormer for the edge of the damaged area is closer to the real label than that of UNet. This is due to the self-attention module in SDFormer's swin transformer block, which effectively enhances the weight of structural damage parts in the feature map, and then improves the overall performance of the model. With the increase of strain field map noise, SDFormer still obtains the effect better than other models, which fully shows the superiority of the model proposed in this paper. The framework of using strain field data to train a neural network to predict structural strain damage proposed in this paper is feasible, and can realize real-time damage identification.

However, it is undeniable that due to the lack of measured data in practical application scenarios, the SDFormer network proposed in this paper can only be trained and verified with the data simulated by finite element model. Labeling the strain field data from sensors is costly work, which may be a major obstacle to the promotion of the method, but it does not affect the effectiveness of this method. The problem of high data cost can be replaced by simulation data close to the real situation. In addition, although the number of parameters in SDFormer is less than UNet, its time-consumption is three times that of UNet, which will reduce the identification speed of SDFormer when it takes strain damage identification of large structures. However, due to the long time consumed by the generation of strain damage, the requirement of real-time damage identification would not

be too high. Therefore, the current identification speed of SDFormer is capable of meeting the requirements.

In future work, we will explore the feasibility of applying the simulation data training model to the real structure scene, and find a method to obtain the real sensor data at a low cost.

5. Conclusions

In this paper, a novel U-shaped high-efficiency structural strain damage identification neural network, i.e, SDFormer, is proposed. In this method, the strain field data is converted into the form of a strain field map and input into SDFormer. The numerical simulation and experimental study were conducted on a simply supported plate and a sleeper beam of a rail train. The damage identification performance and anti-noise performance of SDFormer were studied through numerical analysis, and compared with the advanced convolutional neural network. Based on theoretical analysis, numerical experiments, and anti-noise experiments, the following conclusions can be obtained:

1. A novel structural strain damage identification strategy is proposed in this paper. This strategy takes the strain field map of the structure as the input, and uses the image segmentation algorithm to identify the damage location and level. This damage identification process is simple and there is no need for complex damage index design. On the premise of ensuring the accuracy, it can greatly simplify the process of damage identification and improve the efficiency.
2. According to the results of numerical experiments, compared with the advanced convolutional neural network, the SDFormer can achieve better damage identification performance with fewer parameters. The damage identification results of SDFormer are closer to the real damage map than those of the comparison model, which shows that SDFormer has excellent structural strain damage identification performance.
3. The results of anti-noise experiment show that SDFormer can still maintain better identification performance than that of the comparison models, although the identification performance of SDFormer decreases under the influence of different noise levels, which illustrates that the SDFormer has good noise resistance and robustness.

Author Contributions: Conceptualization, Z.L. and P.X.; methodology, Z.L.; software, Z.L.; validation, Z.L.; investigation, Z.L.; resources, P.X.; data curation, Z.L.; writing—original draft preparation, Z.L.; writing—review and editing, P.X., J.X., and C.Y.; visualization, Z.L.; supervision, P.X. and C.Y.; project administration, P.X.; funding acquisition, P.X. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China [No. 51675537] and the Leading Talents in Science and Technology of Hunan Province in 2019 [No. 2019RS3018].

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Eisenmann, D.J.; Enyart, D.; Lo, C.; Brasche, L. Review of Progress in Magnetic Particle Inspection. In *40th Annual Review of Progress in Quantitative Nondestructive Evaluation: Incorporating the 10th International Conference on Barkhausen Noise and Micromagnetic Testing, VOLS 33A & 33B*; American Institute of Physics: College Park, MD, USA, 2014.
2. Marcantonio, V.; Monarca, D.; Colantoni, A.; Cecchini, M. Ultrasonic waves for materials evaluation in fatigue, thermal and corrosion damage: A review. *Mech. Syst. Signal Process.* **2019**, *120*, 32–42. [CrossRef]
3. Withers, P.J.; Preuss, M. Fatigue and Damage in Structural Materials Studied by X-ray Tomography. *Annu. Rev. Mater. Res.* **2012**, *42*, 81–103. [CrossRef]
4. Sophian, A.; Tian, G.; Fan, M. Pulsed Eddy Current Non-destructive Testing and Evaluation: A Review. *Chin. J. Mech. Eng.* **2017**, *30*, 500–514. [CrossRef]

5. Campanella, C.E.; Cuccovillo, A.; Campanella, C.; Yurt, A.; Passaro, V.M.N. Fibre Bragg Grating Based Strain Sensors: Review of Technology and Applications. *Sensors* **2018**, *18*, 3115. [CrossRef] [PubMed]
6. Yang, L.; Xie, X.; Zhu, L.; Wu, S.; Wang, Y. Review of Electronic Speckle Pattern Interferometry (ESPI) for Three Dimensional Displacement Measurement. *Chin. J. Mech. Eng.* **2014**, *27*, 1–13. [CrossRef]
7. Kaewunruen, S.; Janeliukstis, R.; Freimanis, A.; Goto, K. Normalised curvature square ratio for detection of ballast voids and pockets under rail track sleepers. *J. Phys. Conf. Ser.* **2018**, *1106*, 012002. [CrossRef]
8. Bayissa, W.L.; Haritos, N. Structural damage identification in plates using spectral strain energy analysis. *J. Sound Vib.* **2007**, *307*, 226–249. [CrossRef]
9. Shadan, F.; Khoshnoudian, F.; Esfandiari, A. Structural Damage Identification Based on Strain Frequency Response Functions. *Int. J. Struct. Stab. Dyn.* **2018**, *18*, 1850159. [CrossRef]
10. Abdeljaber, O.; Avci, O.; Kiranyaz, S.; Gabbouj, M.; Inman, D.J. Real-time vibration-based structural damage detection using one-dimensional convolutional neural networks. *J. Sound Vib.* **2017**, *388*, 154–170. [CrossRef]
11. Kim, S.; Choi, J.H. Convolutional neural network for gear fault diagnosis based on signal segmentation approach. *Struct. Health Monit. Int. J.* **2019**, *18*, 1401–1415. [CrossRef]
12. Bao, Y.; Tang, Z.; Li, H.; Zhang, Y. Computer vision and deep learning-based data anomaly detection method for structural health monitoring. *Struct. Health Monit. Int. J.* **2019**, *18*, 401–421. [CrossRef]
13. Park, H.S.; An, J.H.; Park, Y.J.; Oh, B.K. Convolutional neural network-based safety evaluation method for structures with dynamic responses. *Expert Syst. Appl.* **2020**, *158*, 113634. [CrossRef]
14. Lin, J.; Yu, Y.; Lin, M.; Wang, Y. Detection of a casting defect tracked by deep convolution neural network. *Int. J. Adv. Manuf. Technol.* **2018**, *97*, 573–581. [CrossRef]
15. Zhang, K.; Cheng, H.D.; Zhang, B. Unified Approach to Pavement Crack and Sealed Crack Detection Using Preclassification Based on Transfer Learning. *J. Comput. Civ. Eng.* **2018**, *32*, 04018001. [CrossRef]
16. Kumar, R.; Cirrincione, G.; Cirrincione, M.; Tortella, A.; Andriollo, M. Induction Machine Fault Detection and Classification Using Non-Parametric, Statistical-Frequency Features and Shallow Neural Networks. *IEEE Trans. Energy Convers.* **2020**, *36*, 1070–1080. [CrossRef]
17. Jin, C.; Chen, X. An end-to-end framework combining time-frequency expert knowledge and modified transformer networks for vibration signal classification. *Expert Syst. Appl.* **2021**, *171*, 114570. [CrossRef]
18. Shelhamer, E.; Long, J.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *39*, 640–651. [CrossRef]
19. Chen, L.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking Atrous Convolution for Semantic Image Segmentation. *arXiv* **2017**, arXiv:1706.05587.
20. Zhao, H.; Shi, J.; Qi, X.; Wang, X.; Jia, J. Pyramid Scene Parsing Network. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 6230–6239. [CrossRef]
21. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778. [CrossRef]
22. Dosovitskiy, A.; Beyer, L.; Kolesnikov, A.; Weissenborn, D.; Zhai, X.; Unterthiner, T.; Dehghani, M.; Minderer, M.; Heigold, G.; Gelly, S.; et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *arXiv* **2021**, arXiv:2010.11929.
23. Liu, Z.; Lin, Y.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin transformer: Hierarchical Vision Transformer using Shifted Windows. *arXiv* **2021**, arXiv:2103.14030.
24. Tran-Ngoc, H.; Khatir, S.; Ho-Khac, H.; De Roeck, G.; Bui-Tien, T.; Abdel Wahab, M. Efficient Artificial neural networks based on a hybrid metaheuristic optimization algorithm for damage detection in laminated composite structures. *Compos. Struct.* **2021**, *262*, 113339. [CrossRef]
25. Khatir, S.; Tiachacht, S.; Le Thanh, C.; Ghandourah, E.; Mirjalili, S.; Abdel Wahab, M. An improved Artificial Neural Network using Arithmetic Optimization Algorithm for damage assessment in FGM composite plates. *Compos. Struct.* **2021**, *273*, 114287. [CrossRef]
26. Loshchilov, I.; Hutter, F. Decoupled Weight Decay Regularization. In Proceedings of the International Conference on Learning Representations, New Orleans, LA, USA, 6–9 May 2019.
27. Shi, W.; Caballero, J.; Huszár, F.; Totz, J.; Aitken, A.P.; Bishop, R.; Rueckert, D.; Wang, Z. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network. *arXiv* **2016**, arXiv:1609.05158.
28. Han, H.; Gu, J.; Zheng, Z.; Dai, J.; Wei, Y. Relation Networks for Object Detection. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Salt Lake City, UT, USA, 18–23 June 2018; pp. 3588–3597.
29. Hu, H.; Zhang, Z.; Xie, Z.; Lin, S. Local Relation Networks for Image Recognition. In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 3464–3473.
30. Milletari, F.; Navab, N.; Ahmadi, S.A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. In Proceedings of the 2016 Fourth International Conference on 3D Vision (3DV), Stanford, CA, USA, 25–28 October 2016; pp. 565–571. [CrossRef]
31. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal Loss for Dense Object Detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 318–327. [CrossRef] [PubMed]

32. Kingma, D.; Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* **2014**, arXiv:1412.6980.
33. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*; Springer International Publishing: Cham, Switzerland, 2015; pp. 234–241.

sensors

MDPI

Perspective

Concept of Placement of Fiber-Optic Sensor in Smart Energy Transport Cable under Tensile Loading

Monssef Drissi-Habti [1,*,†], Neginhal Abhijit [1], Manepalli Sriharsha [1], Valter Carvelli [2,†] and Pierre-Jean Bonamy [3]

[1] COSYS Department, Université Gustave Eiffel, Champs/Marne, F-77447 Marne-la-Valleée, France; abhijit.neghinal@gmail.com (N.A.); sriharsha.manepalli@gmail.com (M.S.)
[2] Department A.B.C., Politecnico di Milano, Piazza Leonardo Da Vinci 32, 20133 Milan, Italy; valter.carvelli@politecnico.it
[3] MEDYSYS, 91400 Orsay, France; pj.bonamy@medysys.fr
* Correspondence: monssef.drissi-habti@univ-eiffel.fr; Tel.: +33-6-0927-1370
† Directors—International Associated Lab. SenSIN-CT, Université G. Eiffel-Politecnico di Milano, 20133 Milan, Italy.

Abstract: Due to the exponential growth in offshore renewable energies and structures such as floating offshore wind turbines and wave power converters, the research and engineering in this field is experiencing exceptional development. This emergence of offshore renewable energy requires power cables which are usually made up of copper to transport this energy ashore. These power cables are critical structures that must withstand harsh environmental conditions, handling, and shipping, at high seas which can cause copper wires to deform well above the limit of proportionality and consequently break. Copper, being an excellent electric conductor, has, however, very weak mechanical properties. If plasticity propagates inside copper not only will the mechanical properties be affected, but the electrical properties are also disrupted. Constantly monitoring such large-scale structures can be carried out by providing continuous strain using fiber-optic sensors (FOSs). The embedding of optical fibers within the cables (not within the phase) is practiced. Nevertheless, these optical fibers are first introduced into a cylinder of larger diameter than the optical fiber before this same fiber is embedded within the insulator surrounding the phases. Therefore, this type of embedding can in no way give a precise idea of the true deformation of the copper wires inside the phase. In this article, a set of numerical simulations are carried-out on a single phase (we are not yet working on the whole cable) with the aim of conceptualizing the placement of FOSs that will monitor strain and temperature within the conductor. It is well known that copper wire must never exceed temperatures above 90 °C, as this will result in shutdown of the whole system and therefore result in heavy maintenance, which would be a real catastrophe, economically speaking. This research explores the option of embedding sensors in several areas of the phase and how this can enable obtaining strain values that are representative of what really is happening in the conductor. It is, therefore, the primary objective of the current preliminary model to try to prove that the principle of embedding sensors in between copper wires can be envisaged, in particular to obtain an accurate idea about strain tensor of helical ones (multi-parameter strain sensing). The challenge is to ensure that they are not plastically deformed and hence able to transport electricity without exceeding or even becoming closer to 90 °C (fear of shutdown). The research solely focuses on mechanical aspects of the sensors. There are certainly some others, pertaining to sensors physics, instrumentation, and engineering, that are of prime importance, too. The upstream strategy of this research is to come up with a general concept that can be refined later by including, step by step, all the aspects listed above.

Keywords: smart composite; strain; fiber optic sensor (FOS); cross-linked polyethylene (XLPE); multi-axial strains; smart energy transport cable; thermomechanical coupling

Citation: Drissi-Habti, M.; Abhijit, N.; Sriharsha, M.; Carvelli, V.; Bonamy, P.-J. Concept of Placement of Fiber-Optic Sensor in Smart Energy Transport Cable under Tensile Loading. *Sensors* **2022**, *22*, 2444. https://doi.org/10.3390/s22072444

Academic Editors: Phong B. Dao, Lei Qiu, Liang Yu and Zahra Sharif Khodaei

Received: 16 February 2022
Accepted: 14 March 2022
Published: 22 March 2022

1. Introduction

Extensive research work has been devoted to offshore wind generation in recent years in aspects of both large-dimension wind blades and high-voltage electrical transport cables [1–7]. The key point of structural health monitoring (SHM) is the adequate technology selection and the most accurate placement of the necessary sensors for measurement of parameters that are directly responsible for the performance of the system. The health of the system is influenced by type of loading, magnitude of loading, the area where the cable is placed (for example, area where the cable is heavily deformed under bending), pre-existing anomalies in the installation surfaces, anchor drops, and shark attacks. The influences of individual events are manifested in terms of perturbations such as strains, damage, and/or cracks, which can lead to thermal gradients and high heat. Therefore, selection of appropriate sensors requires knowledge of the relationship between the events and their effects [8]. According to current market survey, the energy produced by offshore farms is financially costlier than the energy produced by fossil fuels. The difference becomes much smaller if we consider the side effects caused by the fossil fuels, such as global warming and environmental pollution. This difference could be improved by reducing the maintenance cost [9]. This can be made possible by considering several techniques, among which structural health monitoring (SHM) is one. In this decade, the offshore cable failures account for 80% of total financial losses and insurance claims. In the past 7 years, about 90 offshore cable failures have been reported with over EUR 350 million in insurance claims. According to [10], the repair costs of an offshore cable can be between EUR 0.7 million and EUR 1.5 million, which is very expensive.

The technique of embedding FOS can help monitor the cable (Figures 1 and 2) [11]. There are a wide variety of applications of embedding fiber optic sensors in composite materials which include vibration measurements, cure process monitoring, temperature measurements, thermal expansion measurements, detection of delamination/debonding, three-dimensional strain measurements, thermal strain measurements, relative humidity measurements, and detection of cracking. For all these applications, the key requirement is to measure either strain or temperature or both parameters [12]. The different types of FOS reported for strain/temperature measurements in composite materials are fiber Bragg grating (FBG) sensors, interferometric fiber optic sensors, polarimetric sensors, fiber optic micro-bend sensors, distributed sensors (using techniques such as Rayleigh scattering, Raman scattering, and Brillouin scattering), and hybrid sensors.

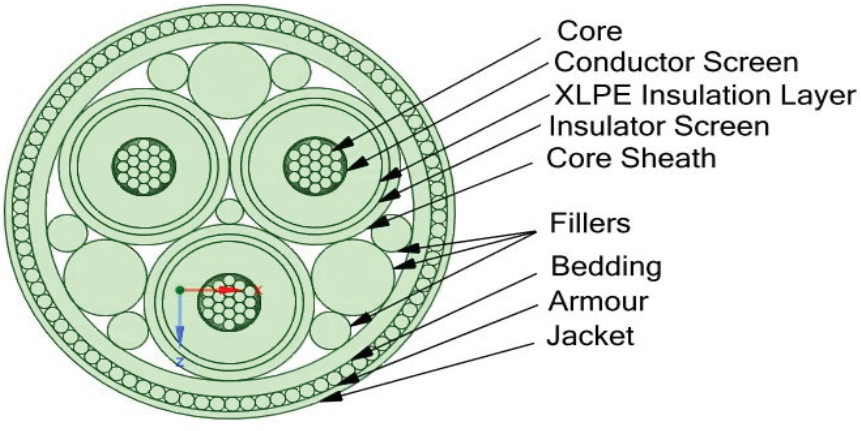


Figure 1. Cross-section of the cable.

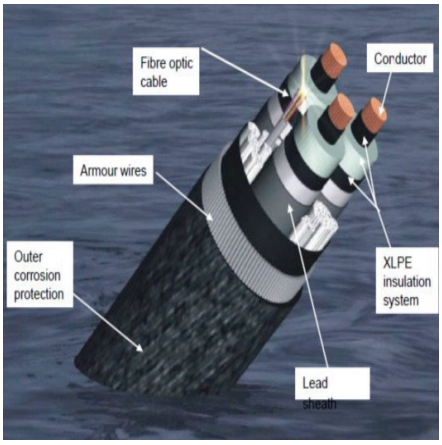


Figure 2. Electric transport cable for offshore farms [11].

The core of the power phase is made up of copper, which is a weak metal. If plasticity propagates inside copper wires, both their individual mechanical properties and their individual electrical properties are disrupted. Plasticity is the consequence of massive defects proliferation into crystal lattices of the metal that lower electrical and thermal conduction [13]. Looking at the generic structure of submarine power cables, there is no potential room for inserting a separate cable with FOS inside it for monitoring of the core. Nowadays, optical fibers which are used for communications are placed inside the cable in place of the fillers with their separate individual casing. Placing the FOS, to monitor the health of the core inside the filler, means placing the FOS far from the copper core, and this may result in inaccurate data due to major offset and presence of other materials, of which some might be in contact and some may not. Along with this, placing another cable for FOS adds up to more space inside the submarine cable. The increase in small space accounts for millions of funds when considered for thousands of kilometers of submarine power cables. Thus, this article suggests a solution of embedding FOS in between copper wires or within the insulator XLPE (cross-linked polyethylene). The placements strategy is balanced between two aspects; first, when placing fiber-optics within copper wires, sensors will be facing extensive sliding friction, which may be detrimental and leads to their failure. Second, embedding sensors within the insulator seems easy, but technically may lead to extensive and variable offset. This work is intended to numerically study these two strategies and extract the most feasible cases. It is worthwhile to note that this article is tackling the issues from a mechanical prospective, knowing that there are other aspects related to sensors physics and instrumentation that are not addressed.

Presently, XLPE is widely used in submarine cables due to its good electrical and thermomechanical behaviors and relatively low cost [14]. The maximum heat-resistant operating temperature of XLPE insulation is 90 °C and the overload temperature should be less than 105 °C [15] in order to avoid premature aging of the dielectric [16]. If one or more copper wires are broken and/or permanently deformed in the core, it leads to upshoot of the temperature of the core [7], and once the core rises above 90 °C, the whole system shuts down. In addition, high temperature will accelerate XLPE insulation aging and shorten cable service life [17]. On the contrary, if the operating temperature of the cable conductor is much lower than 90 °C, the operating efficiency of the cable will be reduced [18]. Therefore, conductor temperature will directly affect the aging of XLPE insulation and the operation of cable; thus, real-time cable temperature measurement is critical to the safe and steady operation of the XLPE power cable [19]. Monitoring of stress and temperature through embedded sensors is therefore mandatory. This is the main reason for this work.

Embedding of optical sensors within XPLE is quite easy to handle technically, whereas embedding between copper wires is technologically challenging. Indeed, either embedding

along the central copper wire and/or helically-wound the same way as other copper wires can lead to extensive sliding/friction and eventually fiber-optic sensor failure. To avoid this, FOSs can be covered by friction-resistant fiber-optic coatings, whereas embedding within XPLE can underestimate the magnitude of thermal parameters inside the phases. The question raised in this article is whether the underestimate is constant or not, after the embedding location is shifted from copper to XPLE (at several distances). The objective of the current article is to check these points numerically, which will help further engineers to target the right cost-effective solutions. Until today, optical fibers are not embedded within the phase, but in the cables and, more precisely, in the insulator that surrounds the phases. Moreover, such optical fibers are not embedded to monitor wire plasticity (fine damage) as we wish to do. Indeed, these optical fibers are embedded into a cylindrical tube, with a diameter much greater than that of the optical fiber, and this difference in diameter in no way makes it possible to follow the true deformation of the copper wires. What we bring back as innovation consists of (i) the embedding of the optical fiber in the phase (and not in the cable), (ii) the embedding of the optical fibers between copper wires to monitor true wire strains, along the central wire in order to measure its deformation, (iii) around the helical wires in order to measure their 3D deformation as well, and (iv) in the insulator at different distances from the center of the phase, to check for the magnitude of the strain offset vs. strain measured at the central wire location. It is worthwhile to note that the article is concentrating on mechanical aspects pertaining to sensors embedding. The aspects linked to the type of sensors physics and associated engineering are too premature to address at this stage.

2. Geometric Model

We will be working on a single phase of the 35kV Nexans submarine power XLPE cable, 2XS2YRAA 18/30(36) kV. The conductor screen is insulated by a thick layer of XLPE (cross-linked polyethylene). The XLPE insulation is mostly recommended for usage in submarine cables as they perform well in both high and low temperatures [20]. At high temperature, the elastic modulus of XLPE decreases while the thermal expansion increases. XLPE has a high thermal expansion as its temperature increases from 25 °C to 105 °C, XLPE expands by 5% while the copper expands by less than 3% in the same temperature range [21]. The insulator is surrounded by a thin insulator screen made up of semi-conducting material (Table 1).

Table 1. Material table assigned to single phase cable.

Materials	Density (kg/m^3)	Young's Modulus (MPa)	Poisson's Ratio	Thermal Conductivity (W/mK)
Copper	8300	1.1×10^5	0.3	370
Semi-Conducting Polymer	1000	1500	0.4	0.1
XLPE	955	1250	0.4	0.28
Polyethelene Black	958	1050	0.4	0.2
Acrylate	950	2700	0.35	0.2
Polymide	1100	3000	0.42	0.8

3. Mechanical Boundary Conditions

Submarine cables are designed in such a way that they can withstand all mechanical stress during manufacturing, transport, handling, installation, and operation. These stresses imposed on submarine cables are completely different from the factors which are affecting the underground cables. In case of inappropriate design and layouts, the cable is more prone to damages and results in high repair cost. Thus, such inappropriate designs are usually abandoned and replaced rather than being rectified. From the previous studies, the tensile and torsional loading cases were considered on a cable armor by using the energy method to create a balance equation [22]. The cable will generate a torsional force

when the cable is subjected to axial tensile loading, and vice versa will be in the case of torsional loading, that is, it will generate an axial tensile force [23]. With the introduction to finite element method, we can effectively analyze tensile properties of the cable. There was an observation made that the tensile stiffness of the cable increases with increase in the radial stiffness up to a typical value and then drops sharply [24]. In this article, a pure tensile loading of a single phase of the cable was assumed, which is not reflecting the reality. The idea here is to have the basic outputs for further research. In previous research, such as [24,25], the primary loading considered is a pure tensile loading case. The tensile load usually occurs while the cables are being laid from the ship. There are three major components which contribute to the tension forces at the laying wheel: [26]

(i) Static weight of the cable between the ship and sea floor.
(ii) Residual bottom tension, which translates to an extra tension force at the laying wheel.
(iii) Dynamic forces when the laying wheel is moving up and down.

Tensile loading on one face of the phase was considered, as shown in Figure 3. The tensile loading is applied along the axial/longitudinal axis of the submarine phase with magnitude of 80 MPa. This stress value is considered according to the hydrostatic law assuming the phase is at a depth of 8000 m. The opposite end of the phase is constrained with all degrees of freedom fixed. Numerical simulations have been conducted assuming that copper wires are arranged in a concentric pattern in a three-layered configuration (1 + 8 + 14). These wires are helically wrapped around the central wire of the core. For the copper helical wires and the helical FOS (mentioned in another test case), the length:pitch ratio has been considered as 1:1 and the twist angle of the helical wires are modeled at 0 deg. The various parts of the model have been explained elsewhere (cf. references of the team, for example, Matine and Drissi-Habti). The contact assumed is linear. All simulations were conducted with a fiber-optic made of silica glass (density 2.4, Young's modulus 72 GPa, and Poisson's ratio 0.17) and acrylate coating (density 0.9, Young's modulus 2.7 GPa, and Poisson's ratio 0.35).

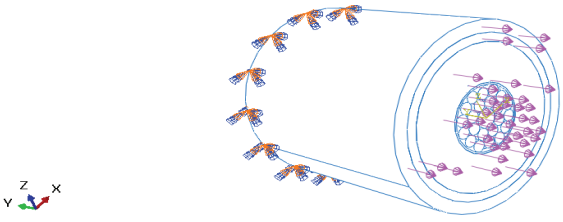


Figure 3. Pure tensile loading condition.

4. Thermal Boundary Condition

There are two common methods for calculating temperature and the ampacity of submarine cables. One is the equivalent thermal resistance method that is based on International Standard IEC 60287. This method has high efficiency and accuracy in calculating the temperature and the ampacity of the direct buried cables. It can satisfy the calculation of the load in the simple scenario. For instance, the IEC calculation cannot address the physical problems coupling with air convection, radiation, and heat transfer, and it will cause large errors in calculation of the multi-loop and complex environment. Another method is the numerical method, including the boundary element method [27], difference method, and finite element method [28]. The finite element method that is based on COMSOL can simulate actual working conditions and perform the coupling calculation of multiple physical fields [14]. The conventions of the IEC 60287 and working of cable at normal operation of the cable will be considered from [7]. In this paper [7], it has been considered that sea water is the surrounding media with an ambient temperature of 15 °C. The thermal analysis of cables aims at computing the temperature rise inside the cables due to the heat generated inside the conductor. During the working of the submarine power

cable, the heat is generated in the cable core (heat source) and is transported to the outside of the cable and disappears into the ambient water(heat sink). The transportation from the heat source to heat sink involves various different layers. The heat transfer is assumed to take place in a steady-state phase, i.e., the analysis will be independent of time (Figure 4). The transfer of heat is assumed to take place by conduction only, and the radiation effects are considered to be negligible. In addition, the material is considered isotropic, which means material properties are the same in all directions.

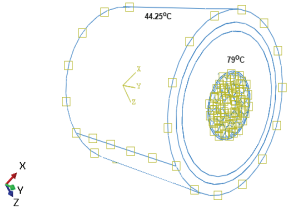


Figure 4. Thermal loading condition.

5. Placement of FOS

In previous articles [1–4], simulations have been conducted to assess the placement of FOS linearly parallel to bonded joints versus a placement based on a sinusoidal fiber covering the largest part of bond joint surfaces and/or composite plies in the case of large-sized wind-blades for offshore farms Figure 5. It has been shown that a single sinusoidally-placed FOS can cover the largest part of the surface to be monitored, in comparison to a linearly-placed FOS concept that needs many parallel FOSs, which is costly. The monitoring that uses a sinusoidal-placed FOS comes up with multi-parameter strain recording, which is a significant advantage over linearly-placed FOS. The simulation conducted took care to refine the placement of FOS to the edge to avoid FOS slippage and sensibility reduction.

More recent works [1–4] have emphasized the use of dual sinusoidal FOSs placement Figures 6 and 7, which was shown to screen up some drawback of monitoring by a single sinusoidal FOS. Indeed, dual-sinusoidal placement offers maximum surface monitoring. More broadly, the most important point to be kept in mind is the wide potential of applications of distributed FOS in many structures and/or infrastructures, whatever the geometries, sizes, and shapes, provided that embedding is performed in a smart way that preserves the integrity and physical properties of sensors and parent structures as well. To allow these sensors to offer the full measure of their performances, without turning out to have "flaws" passing through their miniaturizing as much as possible, their placement far enough from the edges and implementing appropriate design and efficient technologies (right choice of fiber-optic cores, coatings, etc.) are the conditions that may guarantee the full and reliable use of the so-made smart structures.

In the following, the same reasoning conducted in the above for the case of a single high-voltage transmission phase is used. In the above, the monitoring has been optimized in order to ensure the maximum coverage of the area of the composite structure under surveillance, while in the case of the phase, the problem of helical copper wires is more complex, due to the multiplicity of constraints to which they are subjected and the fear of plasticity which threatens them.

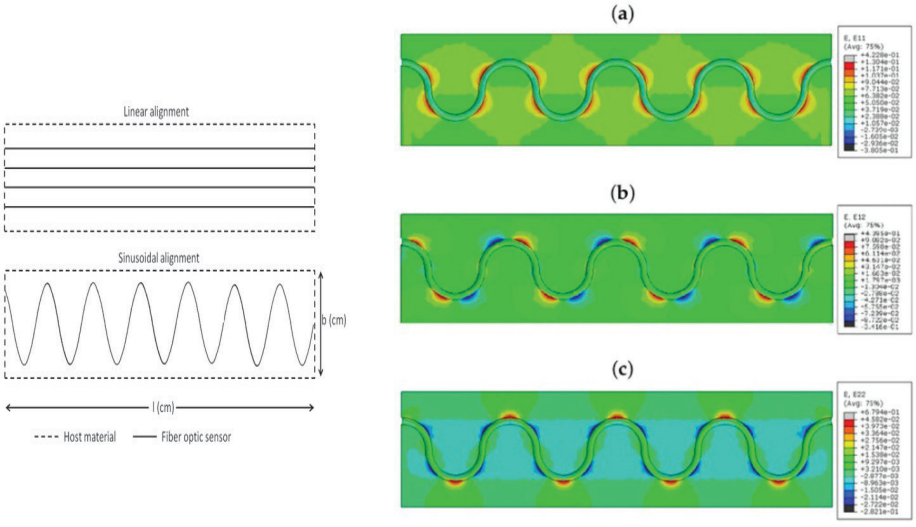


Figure 5. Illustration of linear and sinusoidal fiber FOS placement strategies: (**a**) E11 strain component; (**b**) E12 strain component; (**c**) E22 strain component (acrylate-coated FOS numerical model; the value observed in the host epoxy), sinusoidal placement. Multi-parameter strainsin linear, shear, and lateral directions) can be recorded on sinusoidal placement, while this is not shown when using straight-placed FOS.

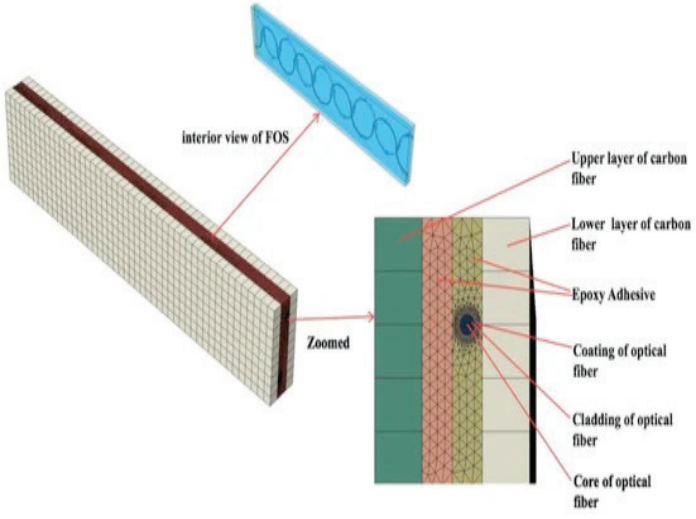


Figure 6. Orientation of dual-sinusoidal FOS.

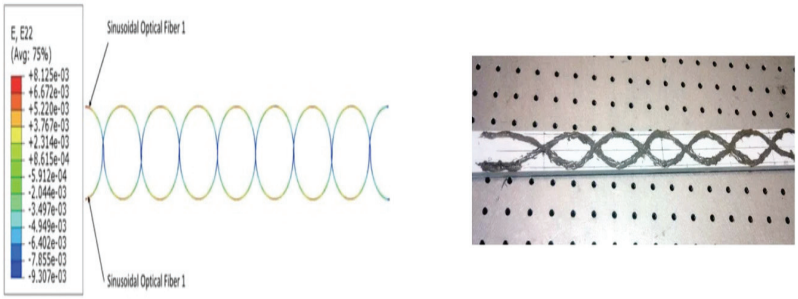


Figure 7. Dual-sinusoidal FOS placement embedded on a glass composite specimen along with strain measurements showing dual-sinusoidal FOS placement (placement in dual-sinusoidal mode, in phase opposition, provides coverage complementary to that of a single sinusoidal fiber) and a view of the specimen that was used for experimental validation.

6. Results and Analysis

A set of test cases were considered, trying to determine the best optimal location for embedding FOSs inside the core and/or as close as possible to the core. This is necessary because the closer the FOSs are to the core, the more accurate the monitoring results will be. Thus, FOSs were embedded close to the core but without compromising their safety. It is worthwhile to note that, here, the article is discussing the conceptualization of FOS embedding inside a single electric phase only, and it is still a little far from engineering applications, which would impose extensive experimental validation. Thus, the current article has to be considered mainly as a proof of concept.

(i) Test case A—Placement of linear FOS in parallel arrangement with varying length from the core.
(ii) Test case B—Placement of FOS helically wound around the complete core

6.1. Test Case A—Placement of Linear FOS in Parallel Arrangement with Varying Length from the Core

The longitudinal length of the linearly-placed FOS is considered in all simulations as 40 mm, which corresponds to the helical step. Figure 8 is one of the examples of the test case where linear FOS is embedded 4 mm inside the XLPE insulator. The strain values in various directions are represented by Exx in the figures.

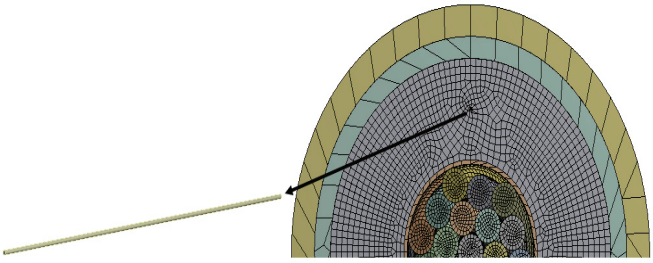


Figure 8. Placement of linear FOS 4 mm inside XLPE.

From Figure 9, different color schemes of linear-placed FOS can be observed. The blue color scheme refers to FOS placed parallel to the central wire, whereas the other color schemes of green, pink, and orange refer to the FOS embedded inside the XLPE insulator over the thickness. The length considered in all simulations is 40 mm, which corresponds to the helical pitch.

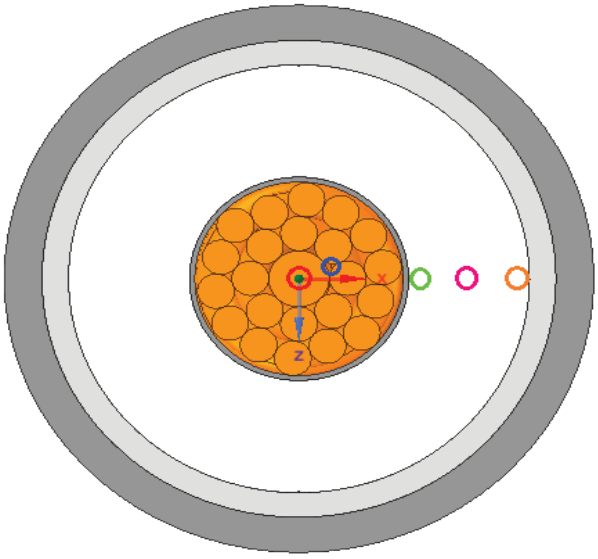


Figure 9. Color scheme for placement of FOS.

The FOS being in direct contact with copper wire is expected to record an averaged value of copper deformation. The red curve represent strain values of the central wire of the copper core. The strain values in the FOS which is placed parallel and in contact with the core (in blue) provides almost the same value of all directional strain as that of the copper wire (in red). Statistically, this seems to be the most convenient case which could be adopted, but arrangement of FOS parallel to the central wire of the core is risky and accounts for a high probability of failure, unless proper coating is used. All simulations were conducted using fiber-optic made of silica glass (density 2.4, Young's modulus 72 GPa, and Poisson's ratio 0.17) and acrylate coating (density 0.9, Young's modulus 2.7 GPa, and Poisson's ratio 0.35).

From Figures 10–12, linear FOS placed 1 mm inside the XLPE insulator (in green) provides us with similar strain values to those of copper but it experiences significant divergence in strain values after a given FOS length. The most probable reason for this divergence is due to the boundary condition which is being applied at the free end. Another reason is that this placement corresponds to the case where the FOS sensor is embedded within two different media (copper and XPLE) that have different mechanical properties and, consequently, depending where mesh nodes are located, calculations are oscillating between two different values of the Young's modulus, which may introduce some errors. This may come from numerical simulation.There are fluctuations in the calculation of the Young's modulus that will randomly change depending on where the mesh noodles are located. Embedding FOS much further into the insulator, 4 mm inside the XLPE insulator (shown in pink) and away from the central core, was also tested. There is slight difference in the strain magnitude and there still exists divergence in the strain values after certain length of the FOS, which needs to be considered. This can also be considered as a potential case for the optimal placement of FOS. In the case of embedding FOS 8 mm inside the insulator, XLPE (shown in orange) nearly edges the boundary of the insulator screen.This lies approximately 15 mm away from the central core, which involves some offset in the values of the strain. From the figure, in spite of placing FOS at a comparatively far distance from copper core, similar trends and divergences exist when compared with the previous test case.

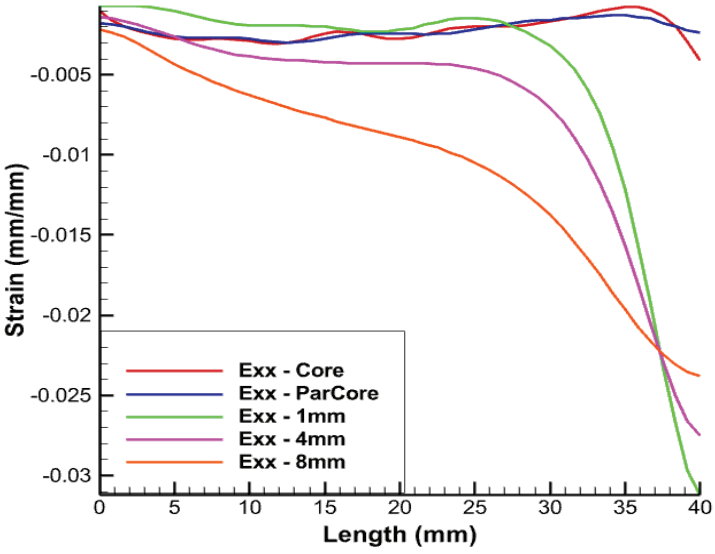


Figure 10. E_{xx} strain for pure tension test case.

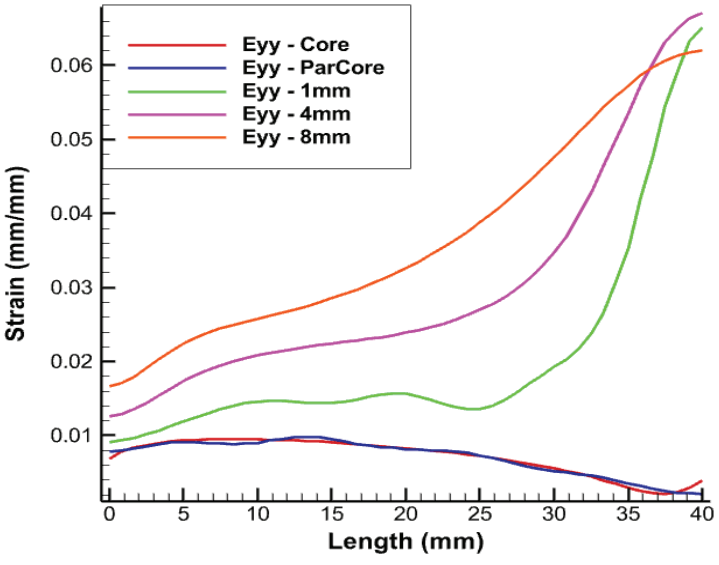


Figure 11. E_{yy} strain for pure tension test case.

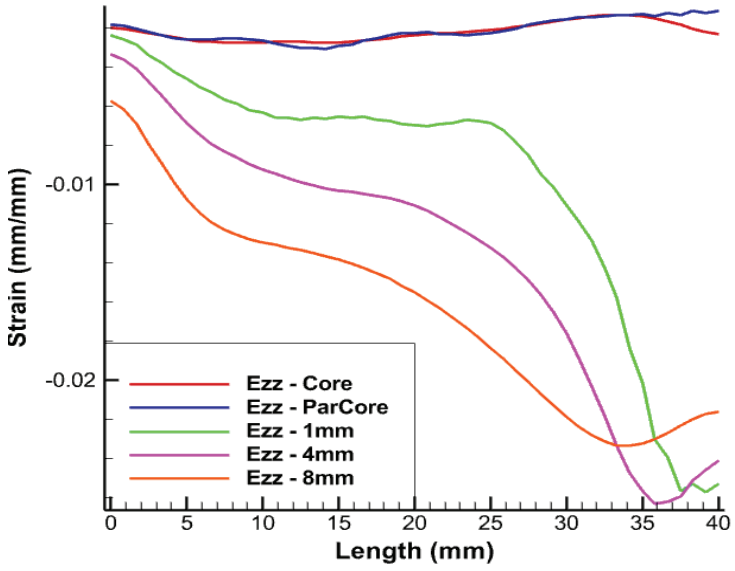


Figure 12. E_{zz} strain for pure tension test case.

From the above results, a better perspective about the placement of FOS for accurate SHM can be observed. Placement of an FOS inside the copper core provides exactly the same results as that of copper, but technically implementing this case is somewhat risky and must be considered along with appropriate fiber-optics coating technologies. The same risk is still to be considered when FOS is helically wound the same way as copper wires, although this would bring the most valuable outputs, i.e., multi-parameter strains. Thus, this case can be considered if full assurance is given that FOS will not suffer failure following helical winding coupled to sliding/friction with copper wires. The second case with placement of FOS 1 mm inside the XLPE provides almost the same lateral strain (E_{xx}). There is not a major difference either in the longitudinal strain. With the other two cases, there exist a difference in longitudinal strain and lateral strain up to magnitude e^{-2}. The case with embedding FOS 1 mm inside the XLPE insulator (in green) is the most appropriate and reliable in terms of FOS safety and monitoring.

6.2. Test Case B: Placement of FOS Helically Wound around the Complete Core

From previous research [3], using sinusoidal placement over the linear one provides multi-parametric strain values, i.e., strains in linear, shear, and lateral directions, whereas the linear FOS arrangement just provides the strain values in the lateral direction and is not sensible on longitudinal and shear loads [1]. Thus, FOS was also tested numerically, helically wound around a helical copper wire. The pitch size of the helical FOS is considered the same as that of the copper coil, i.e., pitch of the helical wire used is half of the total axial length. Along with this, the helical wound FOS covers a much greater area compared to parallel FOS arrangement [1]. This should tend to give much more complete results compared to the parallel FOS placement. In other words, one can hopefully access the complete information about the average status of partial and/or total plasticity of copper in the three dimensions, which is considerable, knowing that all the wires, except the central one, are helically-wound (Figures 13–16).

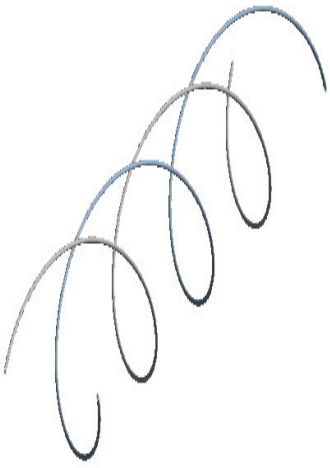


Figure 13. Helical arrangement FOS.

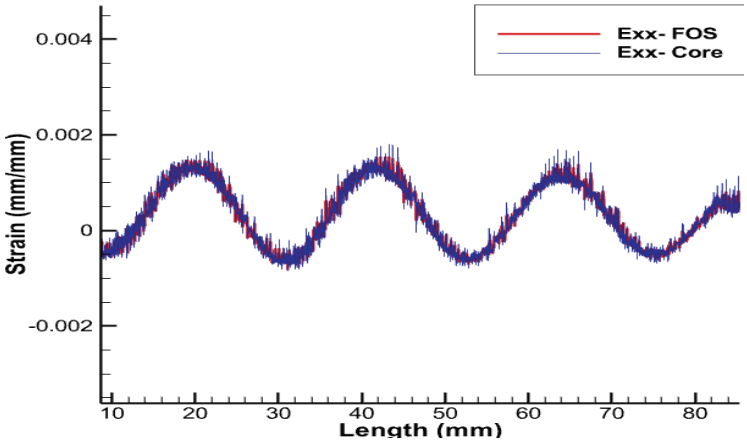


Figure 14. E_{xx} strain for helical arrangement of FOS.

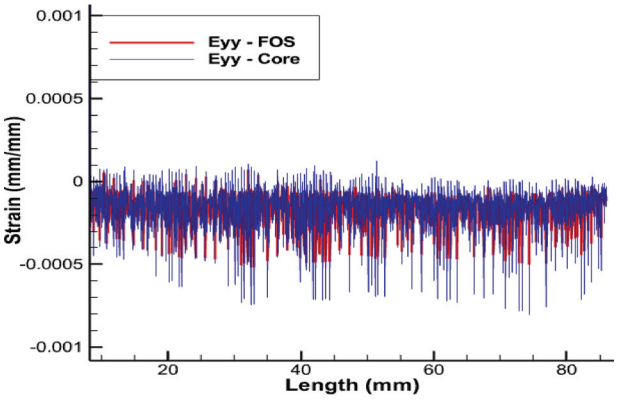


Figure 15. E_{yy} strain for helical arrangement of FOS.

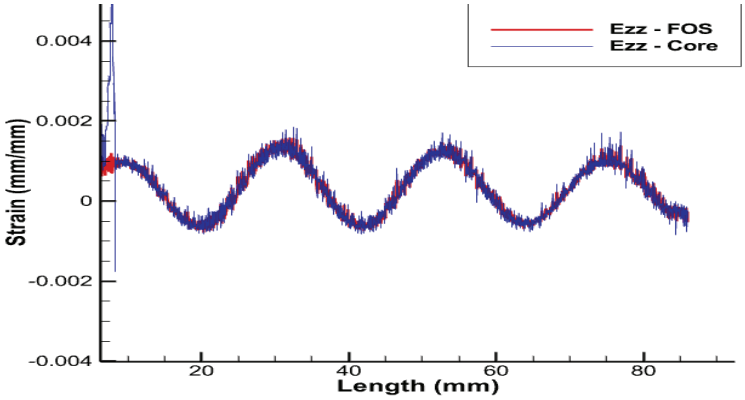


Figure 16. E_{zz} strain for helical arrangement of FOS.

Strain values of the outermost helical copper wire and the helical FOS were compared (Figures 13–15). FOS is directly in contact with copper in this case, but the trend of both values of strains is almost the same. The issue with this case is the technical difficulty in placing the helically wound FOS around a single copper wires, knowing that helically-wound strategy is the most informative one since it wil come up with multiparameter strain monitoring of copper wires. Note that crumbling and twisting pressure force acting on the FOS from the core and from the conductor screen leaves a window for cause of failure. Also, during the process of installation and retrieval, there is also a high possibility of damaging the FOS when it brushes over the copper-wound wires.

This solution comes with a strong technology challenge, namely how to embed fiber-optic sensors between copper wires while ensuring they can withstand the stress–strain state of helically-wound copper wires. It is worthwhile to note that this is out of the scope of the current article for the time being. Our research should be viewed as no more than a preliminary numerical model that is trying to tackle a complex and challenging problem.

7. Conclusions and Discussion

The concept of using embedding FOSs inside high-voltage electric transport phases (not cables) for offshore farms was numerically introduced as a preliminary study to directly monitor true strain of copper wires. The idea is to be informed well before extensive plasticity develops within the wires, thus avoiding electric shutdown. The strain values of FOSs were compared with the strain values of copper cores. The test case where FOS was placed parallel to the central wire produces very good results, although integrity of FOS must be carefully checked by using appropriate coating technologies. The other test cases were favorable, of which the case of embedding the FOS 1 mm inside the XLPE insulator was promising, too. This could be further implemented and tested to perform the most reliable SHM for smart energy transport phases. The other test case, where the FOS is helically wound around and is directly in contact with copper core, provides the most information with the lateral, shear, and longitudinal strain all known at the same time, although it raises a strong technical challenge, which could be dealt with in the future upcoming developments. Finally, through this article, the option of embedding optical fiber within the phase for finer monitoring of plastic deformation of copper wires and their individual failure seems interesting, although extensive work pertaining to refining sensors placements, physics of sensors, and engineering still needs to be performed, to better emphasize the concept which is suggested. Current results, although not totally enough to validate the concept at the engineering step, do however light the way towards obtaining real, accurate, and in situ 3D strain of copper wires inside electric transport cables for offshore farms.

Author Contributions: M.D.-H. is in charge of supervision, conceptualization, methodology, resources and review editing, N.A. and M.S. are in charge of software, validation and formal analysis, V.C. and P.-J.B. are in charge of investigation. All authors have read and agreed to the published version of the manuscript.

Funding: The research received no external funding.

Acknowledgments: This research has been conducted within Flow-Cam (FLoating Offshore Wind turbine CAble Monitoring) project in the framework of Martera Era-Net Cofund. The French National Research Organizeation, ANR, provided funding to this work. I-Site Future provided a master grant to AN. MDH would like to thank the funders for their support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Drissi-Habti, M.; Raman, V.; Khadour, A.; Timorian, S. Fiber Optic Sensor Embedment Study for Multi-Parameter Strain Sensing. *Sensors* **2017**, *17*, 667. [CrossRef] [PubMed]
2. Raman, V.; Monssef Drissi-Habti, L.G.; Khadhour, A. Numerical simulation analysis as a tool to identify areas of weakness in a turbine wind-blade and solutions for their reinforcement. *Compos. Part B Eng.* **2016**, *103*, 23–39. [CrossRef]
3. Raman, V.; Monssef Drissi-Habti, P.L.; Khadour, A. Finer SHM-coverage of inter-plies and bondings in smart composite by dual sinusoidal placed distributed optical fiber sensors. *Sensors* **2019**, *19*, 742. [CrossRef] [PubMed]
4. Raman, V.; Drissi-Habti, M. Numerical simulation of a resistant structural bonding in wind-turbine blade through the use of composite cord stitching. *Compos. Part B Eng.* **2019**, *176*, 107094. [CrossRef]
5. Drissi-Habti, M.; Raj-Jiyoti, D.; Vijayaraghavan, S.; Fouad, E.C. Numerical Simulation for Void Coalescence (Water Treeing) in XLPE Insulation of Submarine Composite Power Cables. *Energies* **2020**, *13*, 5472. [CrossRef]
6. Raman, V.; Drissi-Habti, M.; Kadhour, A. Fiber Optic Sensor Embedment Study for Multi-Parameter Strain Sensing in Smart Composites. In Proceedings of the American Society for Composites—Thirty-Second Technical Conference, West Lafayette, IN, USA, 23–25 October 2017.
7. Matine, A.; Drissi-Habti, M. On-Coupling Mechanical, Electrical and Thermal Behavior of Submarine Power Phases. *Energies* **2019**, *12*, 1009. [CrossRef]
8. Ansari, F. Structural health monitoring with fiber optic sensors. *Front. Mech. Eng. China* **2009**, *4*, 103–110. [CrossRef]
9. Karimirad, M. *Offshore Energy Structures: For Wind Power, Wave Energy and Hybrid Marine Platforms*; Springer: New York, NY, USA, 2014.
10. Gulski, E.; Anders, G.; Jongen, R.; Parciak, J.; Siemiński, J.; Piesowicz, E.; Paszkiewicz, S.; Irska, I. Discussion of electrical and thermal aspects of offshore wind farms' power cables reliability. *Renew. Sustain. Energy Rev.* **2021**, *151*, 111580. [CrossRef]
11. Available online: https://www.nexans.com/en/ (accessed on 15 February 2022).
12. Ramakrishnan, M.; Rajan, G.; Semenova, Y.; Farrell, G. Overview of fiber optic sensor technologies for strain/temperature sensing applications in composite materials. *Sensors* **2016**, *16*, 99. [CrossRef]
13. Napieralski, P.; Juszczak, E.N.; Zeroukhi, Y. Nonuniform distribution of conductivity resulting from the stress exerted on a stranded cable during the manufacturing process. *IEEE Trans. Ind. Appl.* **2016**, *52*, 3886–3892. [CrossRef]
14. Zhang, Y.; Chen, X.; Zhang, H.; Liu, J.; Zhang, C.; Jiao, J. Analysis on the temperature field and the ampacity of XLPE submarine HV cable based on electro-thermal-flow multiphysics coupling simulation. *Polymers* **2020**, *12*, 952. [CrossRef] [PubMed]
15. Aras, F.; Biçen, Y. Thermal modelling and analysis of high-voltage insulated power cables under transient loads. *Comput. Appl. Eng. Educ.* **2013**, *21*, 516–529. [CrossRef]
16. Pilgrim, J.; Lewin, P.; Gorwadia, A.; Waite, F.; Payne, D. Quantifying possible transmission network benefits from higher cable conductor temperatures. *IET Gener. Transm. Distrib.* **2013**, *7*, 636–644. [CrossRef]
17. Robinson, G. Ageing characteristics of paper-insulated power cables. *Power Eng. J.* **1990**, *4*, 95–100. [CrossRef]
18. Yang, B.; Wu, K.; Zheng, Z.; Sun, W.; Wang, Y.; Yang, S.; Liang, R.; Feng, L. Study on improving the current carrying capacity of power cable in real-time operating environment. In Proceedings of the 2016 China International Conference on Electricity Distribution (CICED), Xi'an, China, 10–13 August 2016; pp. 1–4.
19. Yang, L.; Hu, Z.; Hao, Y.; Qiu, W.; Li, L. Internal temperature measurement and conductor temperature calculation of XLPE power cable based on optical fiber at different radial positions. *Eng. Fail. Anal.* **2021**, *125*, 105407. [CrossRef]
20. Cao, L.; Zhong, L.; Li, Y.; Gao, J.; Chen, G. Crosslinking Dependence of Direct Current Breakdown Performance for XLPE-PS Composites at Different Temperatures. *Polymers* **2021**, *13*, 219. [CrossRef]
21. Robaidi, A.; Khateeb, A.M.I.N.S.A.; Dweiri, R.; Gaghbeer, A.T. The Effect of radiation on the thermal expansion of kaolin nano filled-xlpe materials. *Int. J. Metall. Materials Sci. Eng.* **2012**, *2*, 74–84.
22. Knapp, R. Nonlinear analysis of a helically armored cable with nonuniform mechanical properties in tension and torsion. In Proceedings of the OCEAN 75 Conference, San Diego, CA, USA, 22–25 September 1975; pp. 155–164.
23. Knapp, R.H. Derivation of a new stiffness matrix for helically armoured cables considering tension and torsion. *Int. J. Numer. Methods Eng.* **1979**, *14*, 515–529. [CrossRef]

24. Yue, Q.; Lu, Q.; Yan, J.; Zheng, J.; Palmer, A. Tension behavior prediction of flexible pipelines in shallow water. *Ocean. Eng.* **2013**, *58*, 201–207. [CrossRef]
25. Lu, Q.; Chen, J.; Yang, Z.; Chao, Y.Y.; Yue, Q. Numerical and experimental analysis of umbilical cables under tension. *Adv. Compos. Lett.* **2017**, *26*, 096369351702600205.
26. Worzyk, T. *Submarine Power Cables: Design, Installation, Repair, Environmental Aspects*; Springer Science & Business Media: New York, NY, USA, 2009.
27. Hiranandani, A. Calculation of conductor temperatures and ampacities of cable systems using a generalized finite difference model. *IEEE Trans. Power Deliv.* **1991**, *6*, 15–24. [CrossRef]
28. Flatabo, N. Transient heat conduction problems in power cables solved by the finite element method. *IEEE Trans. Power Appar. Syst.* **1973**, *1*, 56–63. [CrossRef]

sensors

MDPI

Article

Experimental Research on Vibration-Damping Effect of Combined Shear Hinge Prefabricated Steel Spring Floating Slab Track

Zhiping Zeng [1,2], Xudong Huang [1,*], Zhuang Li [1], Weidong Wang [1,2], Zixiao Shi [1], Yu Yuan [3] and Abdulmumin Ahmed Shuaibu [1,4]

[1] School of Civil Engineering, Central South University, Changsha 410075, China; 203160@csu.edu.cn (Z.Z.); zhuang_li2@yeah.net (Z.L.); wd1997@csu.edu.cn (W.W.); 8210200726@csu.edu.cn (Z.S.); abdulshub4u@csu.edu.cn (A.A.S.)
[2] MOE Key Laboratory of Engineering Structures of Heavy Haul Railway, Central South University, Changsha 410075, China
[3] Information Engineering School, Nanchang University, Nanchang 330031, China; yuyuan@email.ncu.edu.cn
[4] Department of Civil Engineering, Faculty of Engineering, Ahmadu Bello University Zaria, Kaduna 800242, Nigeria
* Correspondence: samhuang@csu.edu.cn

Abstract: Objective: The cast-in-place steel spring floating slab track (SSFST) is difficult to maintain and repair, while the mechanical strength of the end of the traditional prefabricated SSFST is poor. In order to overcome the above shortcomings, a shear-hinge-combined prefabricated SSFST was developed, and an indoor test was carried out to analyze its vibration-damping effect. Methods: A combined shear hinge SSFST connection model with two length sizes was established. The dynamic response amplitude and frequency response characteristics of the foundation (ground) under different isolator installations and fatigue loads were studied, and the vibration-damping performance of two sizes of combined shear hinge SSFST was evaluated. Results: The vibration-damping effect of the steel spring vibration isolator mainly acts in the middle and low-frequency bands of 16–400 Hz, and the vibration near 10 Hz will be aggravated after the vibration isolator is installed. The vibration index and variation law of the two sizes of SSFST are similar, and the vibration response of 4.8 m SSFST is slightly less than 3.6 m SSFST. There is almost no change in each index when the load is 5 million times, and there is a certain range of change when the load is 10 million times, but the overall change is small. Conclusions: The combined shear hinge prefabricated SSFST can have an excellent isolation effect on vibration and can still maintain good vibration-damping ability within 10 million fatigue loads (about 5 years); 4.8 m SSFST should be laid in straight sections with higher train speeds, while 3.6 m SSFST should be applied in curved sections to ensure smooth lines.

Keywords: subway; combined shear hinge; steel spring floating slab track; prefabricated; vibration isolator

Citation: Zeng, Z.; Huang, X.; Li, Z.; Wang, W.; Shi, Z.; Yuan, Y.; Shuaibu, A.A. Experimental Research on Vibration-Damping Effect of Combined Shear Hinge Prefabricated Steel Spring Floating Slab Track. *Sensors* **2022**, *22*, 2567. https://doi.org/10.3390/s22072567

Academic Editors: Phong B. Dao, Lei Qiu, Liang Yu, Zahra Sharif Khodaei and Francesc Pozo

Received: 4 February 2022
Accepted: 25 March 2022
Published: 27 March 2022

1. Introduction

With the rapid development of the rail transit industry worldwide and the further acceleration of urbanization, most economically developed cities need an urban railway system between the main railway and the general low-speed urban rail transit to further provide linkage between urban and suburban areas [1–3]. Therefore, urban express rail transits have been rapidly developed to meet the medium- and long-distance travel requirements of urban residents [4,5]. However, due to the normally dense land use in urban areas, urban express rail transits often pass very close residential areas [6–8]. For this reason, urban express rail transit is likely to have a greater impact on the environment, particularly in terms of vibration and noise caused by train operation. Therefore, there is a need for an efficient vibration damping measure to solve this increasingly serious problem [9].

Among all urban rail transit vibration isolation measures, the floating slab track (FST) is one of the most efficient vibration-damping structures and has some outstanding advantages [10,11]. In comparison with an integrated slab track system, the FST more effectively reduces the transmission of force [12,13]. However, steel spring FST (SSFST) represents one of the best methods of urban rail transit damping currently recognized and has been widely used around the world [14]. For example, the Japan Tsukuba Express Line, built in 2005, has an SSFST section that has facilitated a maximum speed of 160 km/h. However, South Korea Busan and Cheonan Stations also adopted SSFST for vibration damping and subsequently recorded speeds of 350 km/h, making them the world's fastest and most heavily loaded FST application case [15].

As the speed of the train increases, the unevenness when the train passes the end of the slab is reduced, the force on the end of the track slab of the steel spring floating slab of different lengths is improved, and the shear hinge device needs to be arranged between the track slabs [16]. However, whether the shear hinge will affect the vibration-damping effect of the floating slab track under the train impact, and whether the shear hinge will affect the vibration of the two types of floating slab track under different fatigue loads, are questions that need to be investigated. Therefore, it is necessary to analyze the vibration-damping performance of the floating slab track under the combination of shear hinges, and to establish a full-scale model of the prefabricated steel spring floating slab track with two lengths of shear hinge combination, which could be used to analyze the dynamic response amplitude and frequency response characteristics of the track foundation under different fatigue loads.

In the process of engineering practice, the vibration-damping effect of SSFST is particularly valued by the owner and design unit, and there are few studies on the vibration-damping performance of the prefabricated steel spring track system under fatigue load. Therefore, in order to analyze the vibration-damping performance of SSFST after fatigue load, the ground (foundation) measuring point 150 cm from the longitudinal centerline of SSFST (that is, 15 cm from the edge of the slab) is taken as the research object. Through indoor fatigue tests on SSFST with two lengths of 3.6 m and 4.8 m, the time domain, frequency domain, and vibration level characteristics of the foundation vibration when SSFST is subjected to different times of fatigue loads are analyzed, which could provide a theoretical reference for related engineering applications.

Therefore, two full-scale combined shear hinge prefabricated SSFST models with different lengths are established to analyze the dynamic response amplitude and frequency response characteristics at the track foundation under different fatigue loads and different impact positions. The vibration law and frequency response distribution characteristics at the foundation of two kinds of combined shear hinge SSFST structures are explored, and the vibration-damping performance of two kinds of combined shear hinge SSFST structures is evaluated, which accumulates effective test data for the study of vibration-damping performance of combined shear hinge SSFST.

2. Model Building and Test Plan

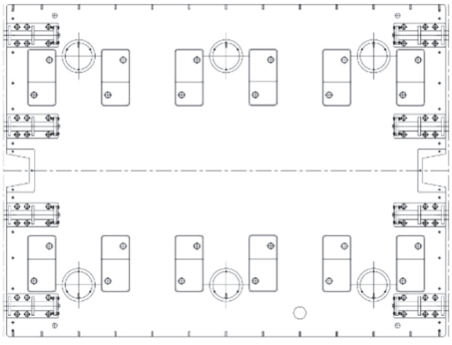
2.1. Full-Scale Model of SSFST

In order to ensure the safe and stable operation of trains and save the number of fasteners, the distance between fasteners in Chinese subways is generally set to 0.6 m [17]. If the steel spring vibration isolator is to be set in the design, at least 2 rows of fasteners are required, then the track slab needs to be set to an integer multiple of 1.2 m, that is, the lengths of 1.2 m, 2.4 m, 3.6 m, 4.8 m, 6 m, etc. When the length of the track slab is too short, the structural vibration is unstable and is not conducive to the safety of the train. If the length of the track slab is too long, it is not conducive to construction and hoisting and it is difficult to lay in the curved section. According to the theoretical calculation of relevant research results, it is found that the dynamic indicators decrease with the increase of the track slab length. When the track slab length exceeds 3.6 m, the reduction degree of the

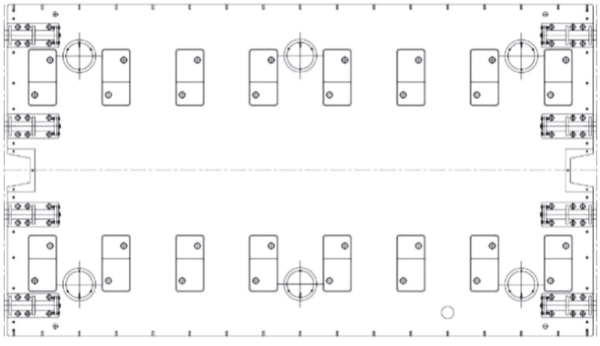
dynamic effect will no longer be obvious. Therefore, the lengths of 3.6 m and 4.8 m are selected in the design.

According to the prefabricated SSFST type laid in the subway main line tunnel, two track bed slabs (3.6 m and 4.8 m) were selected for the drop hammer impact test, and the vibration acceleration of the ground (foundation) measuring point 15 cm away from the edge of the SSFST slab was tested. The SSFST structure is mainly composed of steel rails, spring bars VII type subway fasteners, track bed slabs, vibration isolators, and shear hinges.

The prefabricated SSFST adopts a 3-3 vibration isolator arrangement and consists of a complete assembly structure consisting of steel springs, damping, leveling washers, locking slabs, locking bolts, upper shear hinges, and prefabricated slabs, as shown in Figure 1. SSFST test component parameters are shown in Table 1.



(**a**)



(**b**)

Figure 1. Schematic diagram of floating slab track. (**a**) Design drawing of 3.6 m SSFST. (**b**) Design drawing of 4.8 m SSFST.

Table 1. SSFST test component parameters.

Length	Component	Amount	Parameter
3.6 m	Prefabricated slab	1	3600 mm(length) × 2700 mm(width) × 340 mm(thickness), C50
	Vibration isolator	6	6.66 kN/mm
	Shear hinge	2	SDT-55 × 870
4.8 m	Prefabricated slab	1	4800 mm(length) × 2700 mm(width) × 340 mm(thickness), C50
	Vibration isolator	6	7.50 kN/mm
	Shear hinge	2	SDT-55 × 870

The design schematic of the combined shear hinge and the site are shown in Figure 2. When the shear hinge device is in use, the first shear hinge clamping plate 105 and the second shear hinge clamping plate 111 are moved to the top of the first floating slab 100 by the first adjusting bolt 103 and the second adjusting bolt 109. The shear hinge 201 is placed in front of the first floating slab 100 and the second floating slab 106, and the first shear hinge base 200 is threadedly connected to the first floating slab 100 through fixing bolts 202. The first adjusting bolt 103 is adjusted to make the first shear hinge clamping plate 105 contact and clamp the first shear hinge base 200, which effectively prevents the shear hinge 201 from breaking due to loosening of the fixing bolts 202 in front of the first shear hinge base 200. The movable baffle 205 is taken out through the baffle connecting block 206, and the first connecting plate 204 and the second connecting plate 209 are put into the interior of the first shear hinge support 203 and the second shear hinge support 210, where the damping sleeve 207 and the spring 208 can achieve a certain buffering effect. After the placement is completed, the movable baffle 205 is reset, so that the staff can replace the damping sleeve 207 and the shear hinge 201 at any time. The second shear hinge base 211 is fixed to the front of the second floating slab 106 by the fixing bolts 202, the second adjusting bolts 109 are adjusted; when the second shear hinge clamping plate 111 is brought into contact with the second shear hinge base 211 and clamped, the installation work of the device can be completed [18].

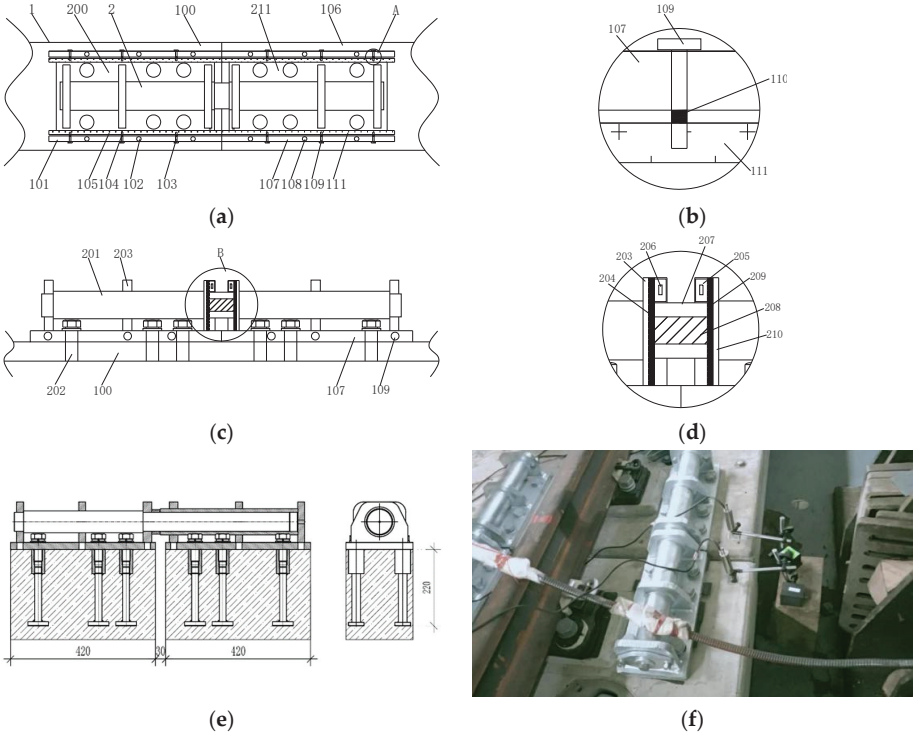


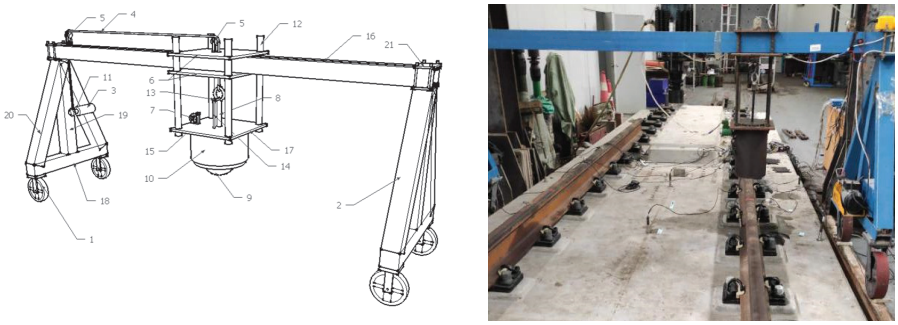
Figure 2. Floating slab track combined shear hinge structure design and physical site drawing. (**a**) Top view schematic. (**b**) The enlarged schematic diagram at A. (**c**) Side view schematic. (**d**) The enlarged schematic diagram at B. (**e**) Design drawing of shear hinge structure. (**f**) On-site installation of shear hinge.

2.2. Drop Hammer Impact Test Instrument

The drop hammer impact test is to test the vibration acceleration of the SSFST foundation under the drop hammer impact before and after the installation of the spring isolator, after 5 million times and 10 million times of fatigue tests. It is possible to understand the

vertical vibration acceleration, frequency spectrum characteristics, and Z vibration level of the SSFST and its foundation under the impact of the drop weight before the installation of the spring isolator, after the installation and after the fatigue test, and to grasp the vibration-damping effect of the spring SSFST vibration isolator after the installation of the spring SSFST vibration isolator under the impact of the falling weight, and after the fatigue test of 5 million times and 10 million times, relative to the vibration isolator before the installation of the vibration isolator.

The drop hammer impact device used in the test can accurately position the drop hammer at the required height and accurately control the position of the hammer body impacting the surface of the rail. The drop weight frame can move freely on the beam and can also provide different drop weight heights, which is suitable for the drop hammer impact test of different sizes and different types of tracks, as shown in Figure 3.



(**a**) (**b**)

Figure 3. Drop hammer impact test device. (**a**) Design drawing. (**b**) Site installation.

Among them, the weight of the drop hammer is 50 kg, the vertical drop height is 100 mm, and the hammering points at the two rails of the SSFST (the hammering points are located at 1/2 of each of the two sizes of SSFST) and no fewer than 10 effective impacts are performed.

2.3. Fatigue Load Test System

The type of fatigue machine used in this test is PMW800-500 electro-hydraulic type, and its hydraulic pulse displacement is 800 mL/min. The whole device includes the host and controller. The hydraulic head can apply a maximum pressure of 500 kN. Combined with the gantry loading device and special railway auxiliary appliances, up to four hydraulic actuators can be configured for loading at the same time. Therefore, the test system is mainly composed of SSFST system, loading device, data acquisition device, and so on.

The schematic diagram of the test system is shown in Figure 4, and the field of the loading device is shown in Figure 5. During the fatigue load test, the rail is pressed by the force adding frame. Load value 80 kN < P < 308 kN, number of cycles N1 = 5×10^6, N2 = 1×10^7, and the load frequency is 3 Hz. The reason is that the axle load of Chinese subway type B car is 14 t [19], and one bogie has two axles (weight is 28 t), and then multiplied by the safety factor of 1.1, so the upper limit of the fatigue load is taken as $28 \times 10 \times 1.1 = 308$ kN. The lower limit of fatigue load = upper limit $\times$ 0.25, so on the basis of 308 kN, $308 \times 0.25 = 77$ kN, taking a multiple of 10 kN is 80 kN. The length of Chinese subway type B car is 12.6 m, each car has 2 bogies, and the actual running speed of the train is generally 70 km/h (19.44 m/s), so in the average time interval of $12.6/(2 \times 19.44) = 0.324$ s, the track slab will pass a bogie load, which is about 3 Hz. Therefore, the fatigue load value in this study can simulate the real train load well and has reliability.

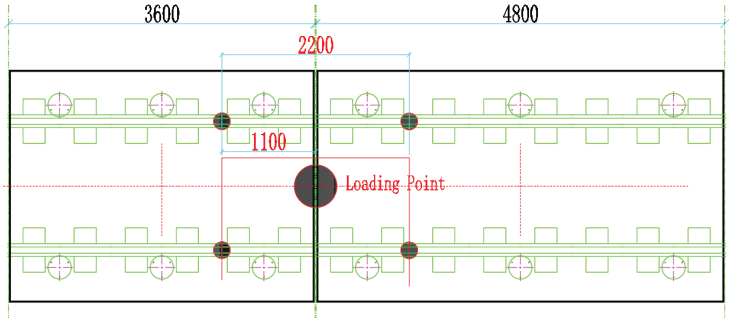


Figure 4. Schematic diagram of fatigue test system of SSFST system model.

Figure 5. Loading device site.

2.4. Introduction to the Test Plan

2.4.1. Test Point Layout

Since the function of the shear hinge is the same under different settings, it provides the establishment of longitudinal shear force constraints for the end of the track slab, and the end conditions of different track slabs are similar. In order to study the vibration-damping performance of the track slabs with two lengths at the same time, under the same experimental cost and conditions, 3.6 m + 4.8 m is the best experimental plan. To fully explore the damping performance of SSFST, the acceleration at the ground foundation is mainly collected. Therefore, the acceleration sensors are arranged on the foundation of the track slab, and the frequency response characteristics of the foundation under different working conditions are explored through the comparative analysis of the acceleration at the foundation of each track. In the test, the drop hammer impact test is mainly performed on two types of track slabs. When arranging the acceleration measuring points, the 0.5 g acceleration sensor is mainly arranged on the ground in the middle of the SSFST, as shown in Figure 6.

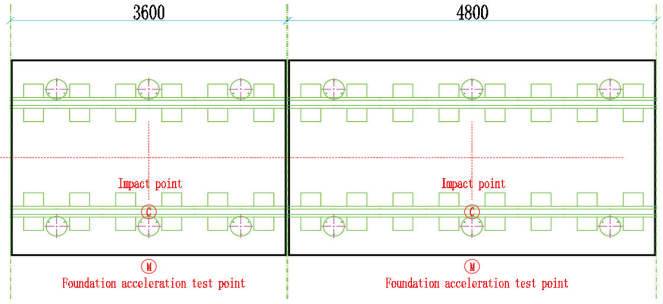


Figure 6. Layout of the accelerometer measuring points.

2.4.2. Test Sampling Parameters and Data Processing Principles

In order to collect more accurate data, the collection frequency of each parameter is set to 20,000 Hz. When organizing test data, system errors should be eliminated, and suspicious data generated by negligence errors should be discarded. The time-domain waveform should be pre-checked to remove singular items, correct zero-line drift, remove trend items, and remove other errors to ensure the accuracy and authenticity of data analysis [20].

According to the setting of the test conditions, in order to explore the influence of the falling weight impact on the vibration damping characteristics of SSFST, the analysis is carried out from the time domain and frequency domain, respectively. A comparative study is undertaken of the basic vibration characteristics of SSFST without vibration isolator, with vibration isolator, after 5 million fatigue cycles, and after 10 million fatigue cycles, combining the impact and vibration attenuation analysis results of the two SSFST structures and the vibration conditions of the foundation under different fatigue times, and then the vibration-damping performance of the two subway SSFST structures is evaluated.

3. Time Domain and Frequency Domain Analysis

3.1. Time Domain Analysis

The purpose of the test research is mainly to study the vibration-damping performance of the two SSFST structures. Since the main purpose of SSFST is to reduce the vibration of the foundation, it is mainly concerned with the vertical acceleration response of the vibration at the foundation under the impact load of the falling weight.

Due to the environmental vibration around the laboratory during the test, a certain amount of clutter will be generated, and the accelerometer at the foundation has high accuracy and is very susceptible to the influence of small external vibrations. Therefore, in order to eliminate the influence of clutter on the test analysis, it is necessary to reduce the noise of the collected signal.

Among the current denoising methods, wavelet threshold denoising has a better application [20–22]. The basic method is to use a family of functions to approximate the objective function hierarchically. Among them, the generated function $\Psi(t)$ will to meet the laws (1) and (2):

Time domain:

$$\int_{-\infty}^{\infty} \Psi(t)dt = 0 \tag{1}$$

Frequency domain:

$$C_{\psi} = \int_{-\infty}^{\infty} \frac{|\varphi^*(w)|}{|w|}dw < \infty \tag{2}$$

where $\varphi(w)$ is the Fourier transform of $\Psi(t)$ and $\varphi^*(w)$ is the complex conjugate function of $\varphi(w)$.

When selecting the basis function and decomposition series of the vibration acceleration signal at the basis of this research, the following strategies are adopted:

(1) Intercept the original signal for 0.4 s, use wavelet soft threshold denoising decomposition to denoise, and use the default threshold for the threshold. A three-level wavelet decomposition is performed on the basic vibration signal, and an approximation signal is obtained.
(2) Reconstruct the decomposed signal to obtain the noise-reduced basic vibration signal for subsequent frequency domain and vibration level analysis.

The dbN wavelet denoising method is adopted in this study (db is the abbreviation of Daubechies, and N is the wavelet order). The number of wavelet decomposition layers is one of the factors affecting the effect of wavelet noise reduction. The larger the number of decomposition layers, the more obvious the difference between noise characteristics and

signal characteristics will be, which is more favorable for signal and noise separation. As far as reconstruction is concerned, the larger the number of decomposition layers, the larger the error in reconstruction and the greater the distortion; when the number of decomposition scales is too small, the maximum value of the modulus corresponding to the noise cannot be sufficiently attenuated, which will cause the signal to be indistinguishable from the noise. It can be seen that if the decomposition scale is too large, some important local features in the signal will be lost. Therefore, it is often decomposed into 3~6 levels when performing wavelet noise reduction. In this study, the signal-to-noise ratio has reached 70.7013 when it is decomposed into three levels, indicating that the noise signal is greatly reduced. In order to ensure that important local features in the signal are not lost, level 1, 2, and 3 wavelet decomposition are selected [23].

The mathematical characteristics and smoothness of the signal after soft threshold processing are also better, and the processing result is also more reliable. Therefore, we choose the dbN wavelet denoising method with soft threshold. In order to ensure that the signal is not distorted after noise reduction, we used the corresponding multi-scale one-dimensional wavelet decomposition function "wavedec" when using MATLAB to calculate, and set up a three-level decomposition function to achieve the removal of high-frequency white noise [24].

Take the acceleration data of the measuring point at the foundation after the vibration isolator is installed at 3.6 m SSFST as an example, as shown in Figure 7. After reconstructing the decomposed signal, it can be seen that the overall waveform of the signal after denoising has not changed much, but the clutter of the signal has been significantly reduced; as shown in Figure 8, 0~0.03 s and 0.17~0.30 s are the most significant, indicating that the wavelet soft threshold denoising is suitable for this type of signal processing. After clutter processing, it will be more conducive to subsequent analysis.

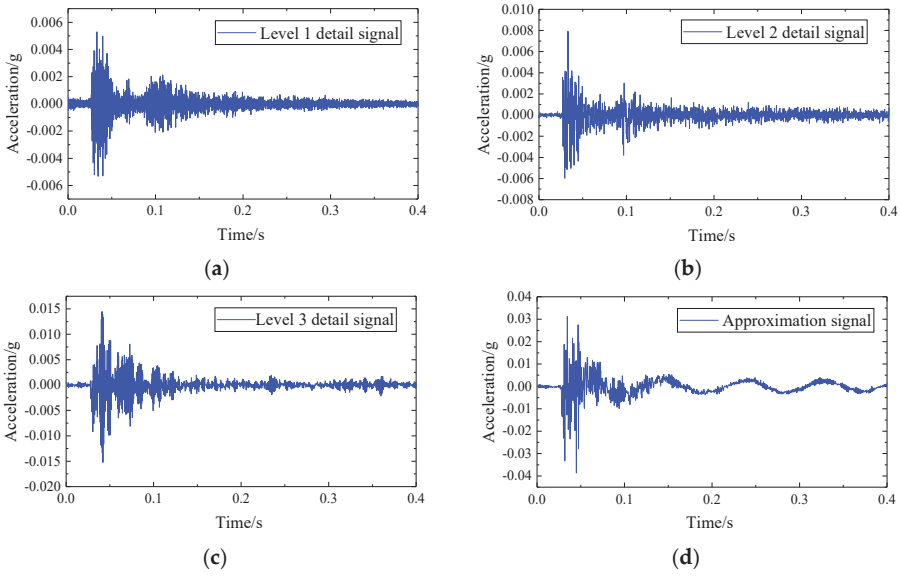


Figure 7. Wavelet soft threshold de-noising decomposes the detailed signal. (**a**) Level 1 detail signal. (**b**) Level 2 detail signal. (**c**) Level 3 detail signal. (**d**) Approximation signal.

The change law of the acceleration time history waveform of the track structure components at the basic acceleration measurement point is relatively similar. Therefore, for different parts of the basic measuring point, the measuring point closest to the impact point is selected for analysis, which is 15 cm away from the edge of the slab. The time-history

waveforms of vibration acceleration of two SSFST structural components are shown in Figures 9 and 10.

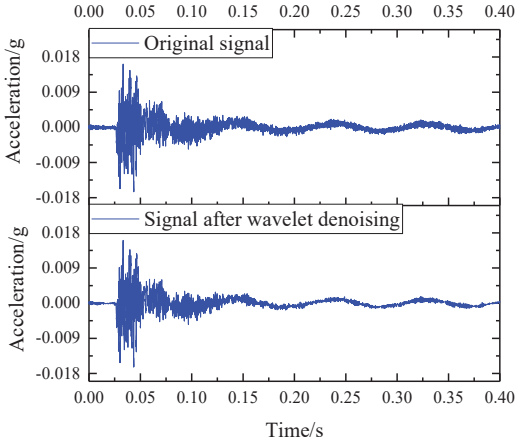
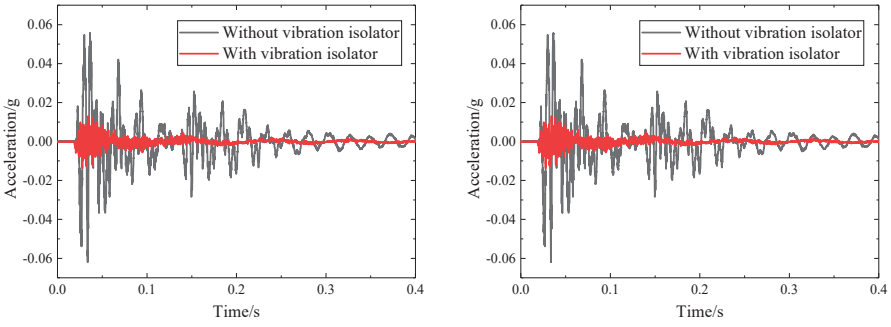
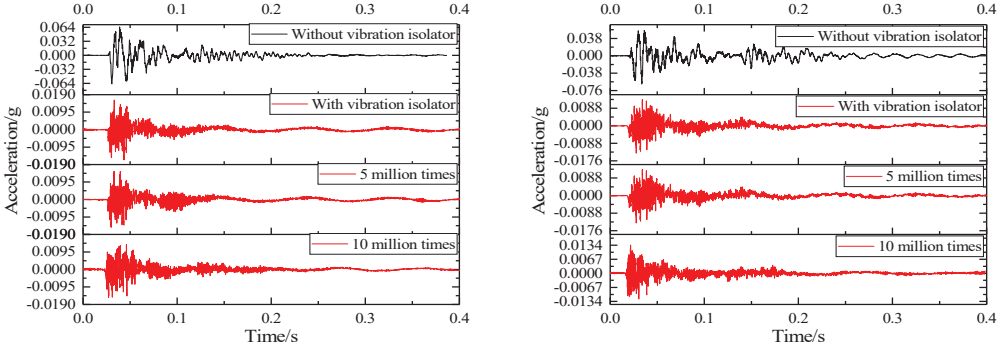


Figure 8. Comparison of the original signal and the signal after wavelet denoising.



(**a**) 3.6 m (**b**) 4.8 m

Figure 9. Acceleration time history waveform with or without vibration isolation measures.



(**a**) 3.6 m (**b**) 4.8 m

Figure 10. Basic acceleration time history waveform for different fatigue times.

According to Figures 9 and 10:

(1) During the SSFST drop hammer impact test, it can be seen that the positive value of acceleration at the foundation is slightly larger than its negative value. After the vibration isolator is installed, its peak acceleration is greatly reduced, and its drop is about 80%, which shows that the vibration isolator has a key effect on reducing the dynamic response of the SSFST system to the foundation.
(2) After the prefabricated SSFST is installed with the vibration isolator, the period of shock and vibration level is continuously shortened. After the installation of the vibration isolator, the total vibration duration of the SSFST is less than the prefabricated SSFST without the vibration isolator. It can be seen that the SSFST with the vibration isolator can achieve excellent vibration energy absorption.
(3) SSFST of two lengths has similar basic vibration laws, and both acceleration amplitude and attenuation waveform are similar. Among them, the acceleration of 4.8 m SSFST is slightly smaller than 3.6 m SSFST, which indicates that the length of the track slab has a small effect on the attenuation of the foundation vibration.
(4) With the application of fatigue load, the waveform of the acceleration waveform at the foundation has slight changes, but the overall change is small.

The time domain analysis mainly focuses on the peak value, peak-to-peak value, and effective value of the track structure impact vibration. The specific meaning is as follows:

(1) The peak value mainly represents the maximum value of the acceleration at the foundation when the SSFST is subjected to an impact, which can directly reflect the acceleration of the foundation when the impact is received.
(2) The peak-to-peak value is the difference between the maximum value and the minimum value of the acceleration at the foundation, and its magnitude has a greater correlation with the shock received at the foundation, so it can be used as an important indicator for evaluating the shock response received at the foundation.
(3) The effective value is also called root-mean-square value, which can represent the vibration energy generated when the foundation is impacted.

Table 2 shows the peak and effective values of vibration acceleration of the two SSFST foundations under different assembly conditions and fatigue loading times of the vibration isolator.

Table 2. Time domain analysis comparison of vibration acceleration (unit: g).

Fatigue Load	Without Vibration Isolator (WOVI)		With Vibration Isolator (WVI)		5 Million Times (5 MT)		10 Million Times (10 MT)	
Type	3.6 m	4.8 m	3.6 m	4.8 m	3.6 m	4.8 m	3.6 m	4.8 m
Max	0.0657	0.0563	0.0162	0.0143	0.0154	0.0141	0.0138	0.0127
Minimum	−0.0660	−0.0627	−0.0165	−0.0145	−0.0160	−0.0143	−0.0153	−0.0139
Peak	−0.0660	−0.0627	−0.0165	−0.0145	−0.0160	−0.0143	−0.0155	−0.0139
Peak-to-peak	0.1317	0.1190	0.0327	0.0284	0.0314	0.0280	0.0291	0.0226
Effective value	0.0105	0.0099	0.0021	0.0018	0.0020	0.0017	0.0018	0.0016

Among them, the effective value is calculated by selecting the entire vibration period, and the duration is set to 0.40 s and is shown in Figure 11; it can be seen that:

(1) After the installation of the vibration isolator, the foundation acceleration peak, peak-to-peak value, and effective value are significantly reduced, and the drop rate is about 80%. Therefore, SSFST is beneficial to reduce the vibration response level of the track structure.
(2) When the number of fatigue loads increases, the maximum, minimum, peak, and effective values of the two SSFST foundations have little change, and the overall decrease slightly. It can be seen that the vibration effect of this type of SSFST is less affected by the fatigue load.
(3) Comparing the peak and effective value of vibration acceleration at the foundation under different assembly conditions of vibration isolators and the number of fatigue loads, it can be seen that the vibration response amplitude of 3.6 m SSFST is slightly

larger than 4.8 m SSFST, and the acceleration level of both at the foundation is generally low.

(4) When laying SSFST tracks in subway tunnels, when wheel-rail impact occurs, it is beneficial to reduce the vibration response of the tunnel base and lining structure and reduce the possibility of damage to the track and tunnel structure.

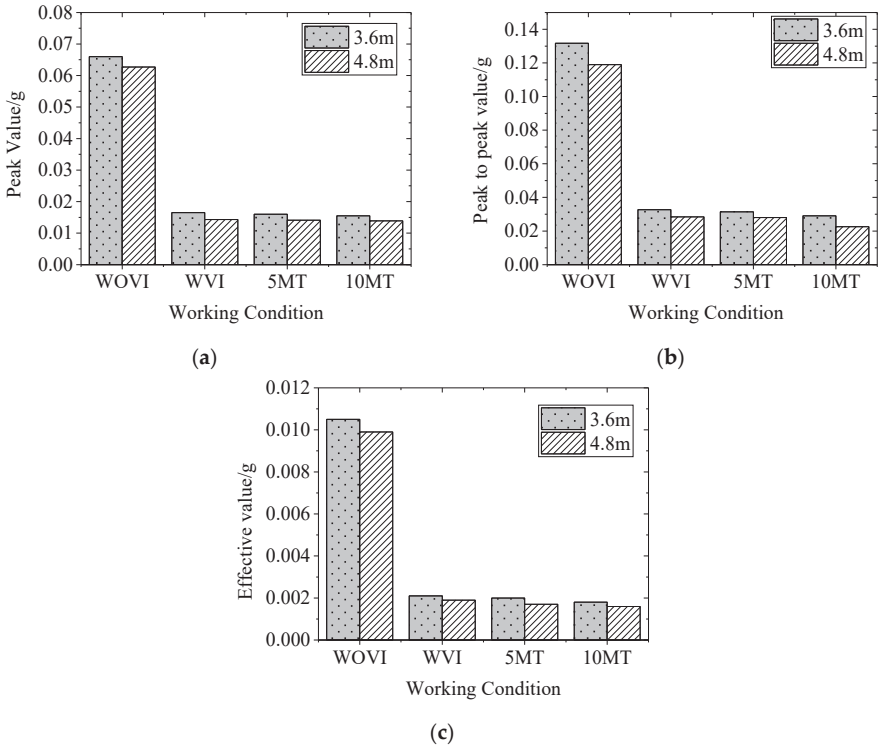


Figure 11. Comparison diagram of time-domain analysis of vibration acceleration. (**a**) Peak value. (**b**) Peak to peak value. (**c**) Effective value.

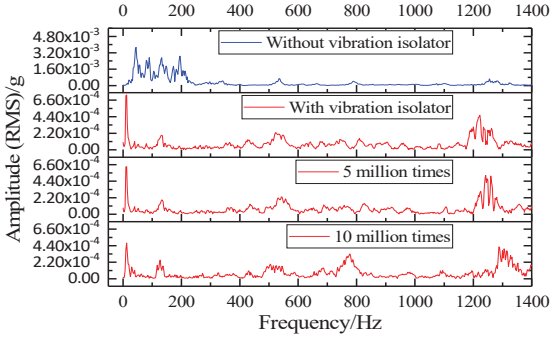
3.2. Frequency Domain Analysis

In the frequency domain analysis of vibration signals, frequency spectrum is often used as an important research method. Through the transformation of parameters such as amplitude in different frequency ranges, the amplitude and phase contained in the vibration signal can be displayed. In the analysis, the Fourier transform is the most used. When in use, the frequency and time domain of the vibration signal can be related to each other, and then the signal characteristics contained in the time domain signal can be displayed and used for analysis. Therefore, the Fourier transform has been widely used [25,26].

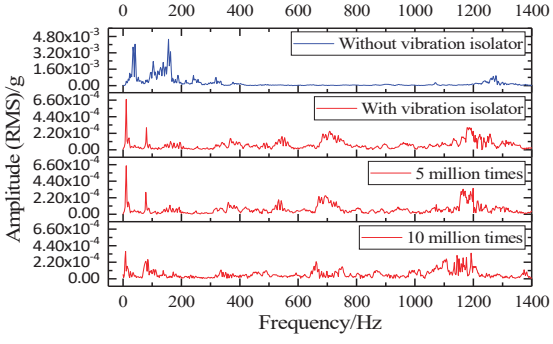
In this study, based on the FFT method, the vibration signals at the foundation of the prefabricated SSFST under different fatigue times were transformed. Since the amplitude spectrum (RMS) mainly reflects the effective value of the vibration acceleration signal, it has a good reflection on the distribution of the vibration signal in each frequency band. Therefore, a comparative study was made on the amplitude spectrum (RMS) of the foundation shock vibration.

In order to study the vibration distribution of the foundation in different frequency domains under different working conditions, the amplitude spectrum (RMS) of the foundation is obtained through fast Fourier transform (FFT), as shown in Figure 12. The

amplitude spectrum (RMS) can well show the changing law of the amplitude waveform under different working conditions, that is, it is easier to compare and analyze.



(**a**) 3.6 m



(**b**) 4.8 m

Figure 12. SSFST basic amplitude (RMS) comparison waveform.

According to Figure 12a, when the vibration isolator is not installed, the amplitude spectrum (RMS) of 3.6 m SSFST has 7 obvious peak points, located at 44.99 Hz, 89.98 Hz, 132.48 Hz, 194.98 Hz, 534.93 Hz, 789.91 Hz, and 1254.84 Hz, and their RMS is 37.28×10^{-4} g, 27.27×10^{-4} g, 27.12×10^{-4} g, 28.73×10^{-4} g, 6.79×10^{-4} g, 4.19×10^{-4} g, and 6.54×10^{-4} g, respectively. After installing the vibration isolator, the amplitude spectrum (RMS) has four obvious peak points, located at 10.50 Hz, 134.98 Hz, 524.93 Hz, and 1222.35 Hz, its RMS is 7.31×10^{-4} g, 1.91×10^{-4} g, 2.23×10^{-4} g, and 6.49×10^{-4} g, respectively. After 5 million fatigue loads, the positions and amplitudes of the above four peak points have relatively small changes, which are located at 10.25 Hz, 132.48 Hz, 542.43 Hz, and 1243.34 Hz, respectively, and its RMS is 6.34×10^{-4} g, 1.91×10^{-4} g, 2.30×10^{-4} g, and 5.16×10^{-4} g. The peak point and amplitude of 10 million times of fatigue change greatly, and another peak point appears at 777.41 Hz; all peak points are located at 10.498 Hz, 127.42 Hz, 504.43 Hz, 777.41 Hz, and 1287.34 Hz, respectively, and its RMS is 4.74×10^{-4} g, 2.49×10^{-4} g, 1.94×10^{-4} g, and 4.25×10^{-4} g, respectively.

The basic vibration acceleration of 4.8 m SSFST shown in Figure 12b also obeys a similar law, that is, the spectrogram changes greatly when the fatigue load is 10 million times. When the vibration isolator is not installed, the amplitude spectrum (RMS) of 4.8 m SSFST has 3 obvious peak points, located at 41.355 Hz, 156.23 Hz, and 1277.43 Hz, and its RMS is 40.46×10^{-4} g, 45.26×10^{-4} g, and 9.61×10^{-4} g, respectively. After installing the vibration isolator, there are 7 obvious peak points, located at 9.51 Hz, 80.41 Hz, 163.12 Hz, 367.61 Hz, 544.51 Hz, 707.63 Hz, and 1185.53 Hz, and the RMS is 3.67×10^{-4} g, 2.93×10^{-4} g, 1.11×10^{-4} g, 1.51×10^{-4} g, 1.76×10^{-4} g, 2.46×10^{-4} g, and

2.95×10^{-4} g, respectively. After fatigue loading, the positions of the above peak points have changed. After 5 million fatigue loads, the positions and amplitudes of the above seven peak points have relatively small changes, located at 9.49 Hz, 78.12 Hz, 160.82 Hz, 360.71 Hz, 533.02 Hz, 693.85 Hz, and 1164.85 Hz, and its RMS is 6.47×10^{-4} g, 2.92×10^{-4} g, 1.21×10^{-4} g, 1.60×10^{-4} g, 1.93×10^{-4} g, 2.47×10^{-4} g, and 3.34×10^{-4} g, respectively. The peak point and amplitude change greatly when fatigue is 10 million times. The number of obvious peak points is 6, located at 9.19 Hz, 85.01 Hz, 335.42 Hz, 661.68 Hz, 751.29 Hz, and 1192.42 Hz, and its RMS is 3.63×10^{-4} g, 2.56×10^{-4} g, 1.18×10^{-4} g, 2.27×10^{-4} g, 1.59×10^{-4} g, and 3.41×10^{-4} g, respectively.

4. Hilbert–Huang Transform Analysis

The shock component in the vibration signal of the SSFST system accounts for a large proportion, the frequency component is very rich, and the vibration information contained in different frequency bands is different [27]. When analyzing the vibration signal of the SSFST system, the Intrinsic Mode Function (IMF) components of each order after the Empirical mode decomposition (EMD) will contain the natural vibration components of different frequency bands caused by different vibration isolation methods [28–31]. According to the theory of kurtosis coefficient, IMF components with larger kurtosis values have more periodic impact components [32].

Due to the effect of the vibration isolator of the SSFST system, the dynamic response of the SSFST structure will have a huge impact, and the amplitude and frequency of the basic acceleration response will change. In order to analyze the impact of the SSFST system vibration isolator on the track structure response, the foundation is the key test site to test the dynamic impact of the SSFST, and the kurtosis value improved HHT method is used to perform time-frequency analysis and energy analysis on the acceleration response of the track structure.

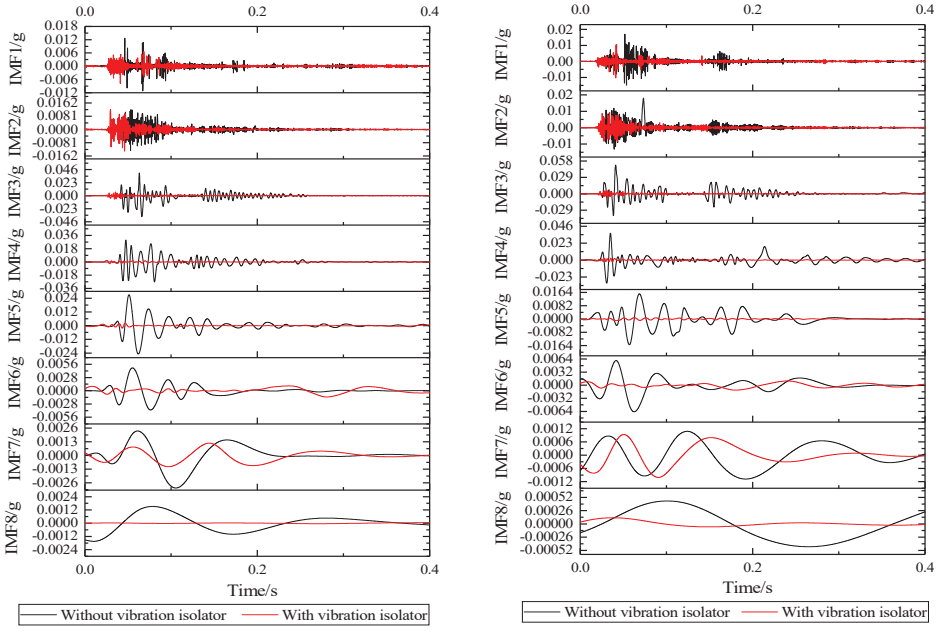
4.1. EMD Analysis

According to the frequency domain analysis in Section 3, it can be seen that the vibration signal has multiple peak points in the frequency domain, and there is a superposition of multiple modal waveforms at the same time. Therefore, we can consider decomposing the vibration signal at the foundation to analyze the vibration waveform of different frequency bands.

According to the characteristics of IMF, for the vibration acceleration at the foundation of SSFST of non-stationary signal, EMD can be used for analysis and processing to obtain the IMF of each frequency band. Through the energy intensity under the characteristic scale reflected by IMF, as shown in the impact intensity of the vibration acceleration signal in each frequency band, it is used to compare the vibration acceleration signal of the foundation under different vibration isolators and fatigue conditions [33].

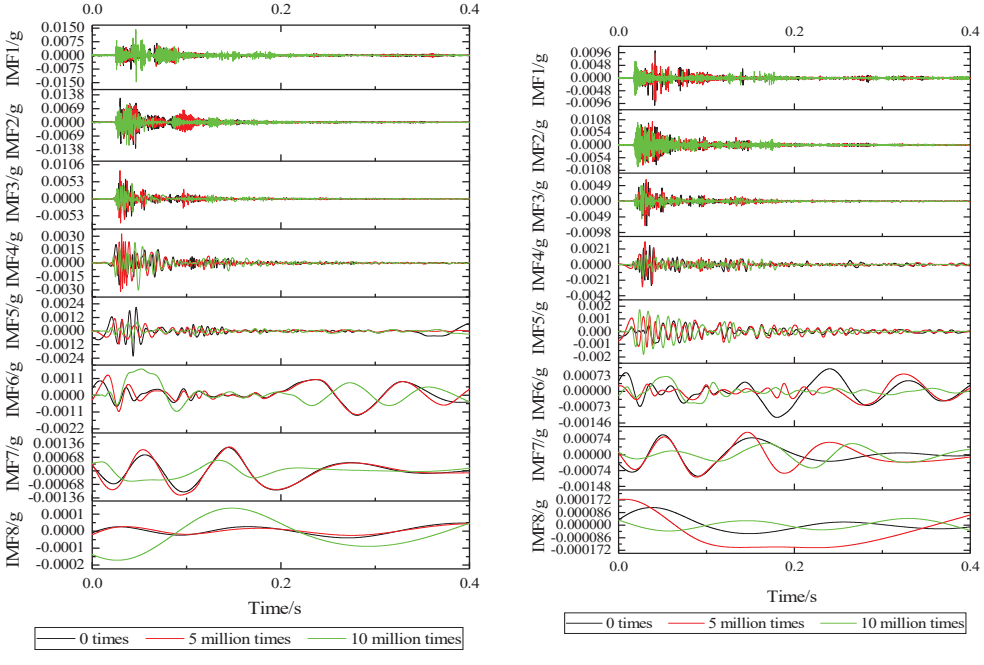
Since the vibration acceleration signal at the foundation of SSFST presents nonlinear and non-stationary characteristics, it is very suitable to use the EMD method to analyze its modal components. At present, there are few studies in SSFST vibration analysis. Therefore, the use of EMD to obtain the IMF components of the vibration signal at different frequency bands can provide a more comprehensive understanding of the change law of the vibration signal.

Using the above steps, EMD transformation is performed on the vibration acceleration time-domain waveform in Figure 10 to obtain the basic acceleration EMD waveform with or without vibration isolators and the basic acceleration EMD waveform under different fatigue load times, as shown in Figures 13 and 14. It can be seen that the vibration signal can be decomposed into 8 orders by using the EMD method, and the decomposed vibration signal presents a relatively regular pure oscillation characteristic. As the decomposition order increases, the amplitude of the IMF decreases, the period becomes longer, and the frequency decreases. It can be seen that EMD has better adaptability to non-stationary fundamental vibration acceleration signals.



(**a**) 3.6 m (**b**) 4.8 m

Figure 13. Comparison of basic acceleration EMD waveforms with and without vibration isolators.



(**a**) 3.6 m (**b**) 4.8 m

Figure 14. EMD waveform of foundation acceleration under different fatigue load times.

According to Figure 13: IMF of various orders with or without vibration isolators, it can be seen that the vibration amplitudes of IMF1, IMF2, IMF7, and IMF8 of the two are maintained at an order of magnitude, which is relatively closer, in which it can be seen that the vibration isolator has limited damping effect on the ultra-high-frequency band and the ultra-low-frequency band. For IMF3~IMF6, it can be seen that after the installation of the vibration isolator, the vibration amplitude has been significantly attenuated, especially in the frequency range of IMF4, and the vibration IMF at the foundation of the two sizes of SSFST is basically the same.

According to Figure 14, it can be seen that as the number of fatigue loads increases, the IMF component of the vibration acceleration at the foundation changes to a certain extent. At 5 million times of fatigue, the IMF component at the foundation of the two sizes of SSFST has a small change range, and the IMF changes of each order of vibration of the two sizes are consistent. When the number of fatigue reaches 10 million times, the IMF component changes greatly, especially in the IMF1~IMF2 frequency band; the IMF7~IMF8 frequency band drops significantly; and the IMF3~IMF6 frequency band shows irregular changes.

According to the IMF waveform decomposed by the EMD method, the IMF maximum value under different working conditions is summarized and analyzed, as shown in the Tables 3 and 4. With the increase of IMF order, the frequency and amplitude of vibration show a decreasing trend; after installing the vibration isolator, the maximum value of the IMF waveform of IMF3~IMF8 decreased significantly; comparing the maximum positive value and the maximum negative value of the IMF waveform, it can be seen that the absolute value of the maximum negative value of the IMF without vibration isolator is generally smaller than the maximum positive value, while the absolute value of the maximum negative value of IMF with vibration isolator is generally greater than the maximum positive value; and with the increase of the number of fatigue loads, the maximum value of IMF gradually showed a downward trend.

Table 3. Summary of the maximum negative values of each IMF waveform (unit: g).

Size	Working Condition	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8
3.6 m	Without vibration isolator	−0.0112	−0.0128	−0.0370	−0.0271	−0.0248	−0.0041	−0.0031	−0.0016
	With vibration isolator	−0.0077	−0.0131	−0.0062	−0.0025	−0.0022	−0.0013	−0.0011	0.0000
	5 million times	−0.0061	−0.0108	−0.0076	−0.0032	−0.0012	−0.0013	−0.0012	0.0000
	10 million times	−0.0145	−0.0116	−0.0044	−0.0030	−0.0009	−0.0011	−0.0005	−0.0002
4.8 m	Without vibration isolator	−0.0149	−0.0132	−0.0385	−0.0314	−0.0159	−0.0065	−0.0011	−0.0004
	With vibration isolator	−0.0103	−0.0084	−0.0076	−0.0028	−0.0009	−0.0012	−0.0010	−0.0001
	5 million times	−0.0089	−0.0086	−0.0077	−0.0030	−0.0012	−0.0006	−0.0010	−0.0002
	10 million times	−0.0066	−0.0093	−0.0055	−0.0017	−0.0018	−0.0006	−0.0006	0.0000

Table 4. Summary of the maximum positive value of each IMF waveform (unit: g).

Size	Working Condition	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8
3.6 m	Without vibration isolator	0.0127	0.0122	0.0398	0.0297	0.0270	0.0048	0.0023	0.0015
	With vibration isolator	0.0066	0.0121	0.0061	0.0021	0.0021	0.0010	0.0012	0.0000
	5 million times	0.0051	0.0096	0.0088	0.0033	0.0011	0.0013	0.0012	0.0000
	10 million times	0.0142	0.0089	0.0047	0.0023	0.0008	0.0017	0.0005	0.0001
4.8 m	Without vibration isolator	0.0169	0.0180	0.0507	0.0366	0.0155	0.0060	0.0011	0.0005
	With vibration isolator	0.0103	0.0094	0.0058	0.0026	0.0008	0.0010	0.0009	0.0001
	5 million times	0.0093	0.0100	0.0068	0.0031	0.0012	0.0008	0.0011	0.0002
	10 million times	0.0064	0.0096	0.0049	0.0013	0.0017	0.0007	0.0005	0.0000

4.2. Kurtosis Analysis

In order to analyze the vibration severity of each IMF in Section 4.1, the kurtosis value is used for research. The kurtosis value is also called the fourth-order central moment, which is often used for statistical estimation of the signal. It is used to express the convexity of the peak of the analyzed vibration signal, has no unit, and is a dimensionless value [34]. The larger the kurtosis value, the sharper the peak of the vibration signal, which means that

the vibration at the foundation is more intense in this study. The mathematical expression of the kurtosis value is:

$$K = \frac{E(x-\mu)^4}{\sigma^4} = \frac{1}{N}\sum_{i=1}^{N}[\frac{x_i-\mu}{\sigma}]^4 \quad (3)$$

In the formula, K—the kurtosis value of the vibration signal at the foundation; N—the number of sampling points in the vibration signal; μ—the mean value of the vibration signal; and σ—the standard deviation of the vibration signal.

In the field of mechanical engineering diagnosis, it is generally considered that when the kurtosis value is equal to 3, it is the normal kurtosis value, also called zero kurtosis. At this time, the concavity and convexity at the peak of the analyzed waveform signal is consistent with the normal distribution waveform. When σ decreases, it indicates that the vibration signal is more concentrated, and the kurtosis value increases. On the contrary, when σ increases, this indicates that the peak value of the vibration signal is more dispersed, and the kurtosis value decreases. According to the above rule, when the kurtosis value is greater than 3, it is recorded as positive, and when the kurtosis value is less than 3, it is recorded as negative. By using the kurtosis index to analyze the dramatic magnitude of IMF changes, it can also be used to understand the degree of peak fundamental vibration acceleration in each frequency band. Therefore, the IMF kurtosis values of each order in Section 4.1 are calculated, as shown in Table 5.

Table 5. IMF kurtosis value of vibration acceleration of SSFST system components.

Size	Working Condition	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8
3.6 m	Without vibration isolator	55.651	15.281	17.178	10.485	12.149	7.734	4.863	2.739
	With vibration isolator	33.294	33.107	40.849	18.419	10.579	2.906	2.704	2.116
	5 million times	26.238	30.885	29.416	18.065	10.149	2.851	2.531	2.063
	10 million times	23.472	22.742	25.266	14.071	8.862	3.725	2.484	1.967
4.8 m	Without vibration isolator	48.728	16.509	15.663	15.672	5.911	7.301	2.058	1.633
	With vibration isolator	28.936	25.593	40.848	23.911	7.629	5.288	3.099	3.227
	5 million times	26.062	24.315	31.146	15.54	5.404	3.571	2.549	2.13
	10 million times	24.424	20.564	28.221	9.048	4.728	2.949	2.288	1.677

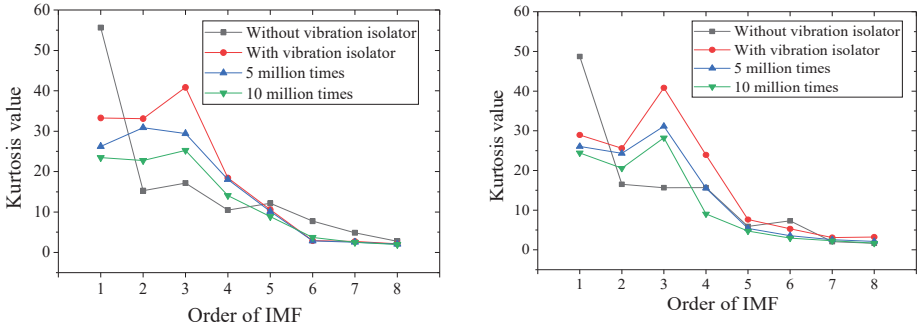
Drawing the kurtosis values of each order into the change waveform shown in Figure 15, we can see that:

(1) When there is no vibration isolator, the IMF1 kurtosis value of the vibration signal at the foundation is larger, and the kurtosis value decreases sharply at IMF2 and is smaller than the working condition of the vibration isolator. After IMF3, the kurtosis value changes steadily, gradually decreases, and gradually becomes negative kurtosis at IMF7~IMF8;

(2) After installing the vibration isolator, the IMF kurtosis values of each order are reduced, the IMF2~IMF4 is higher than without vibration isolators, and the peak signal density of the IMF3 frequency band is the highest. As the number of fatigue increases, the kurtosis value of this frequency range first increases and then decreases, indicating that the fatigue load will affect the distribution of the vibration acceleration IMF peak point at the foundation.

In order to analyze the variation range of the kurtosis value under different installation conditions of vibration isolators and the number of fatigue loads, the kurtosis value change ratios after installing the vibration isolator, after 5 million times of fatigue and after 10 million times of fatigue are summarized, as shown in the Table 6. It can be seen that IMF2~IMF4 of 3.6 m SSFST ground vibration has a relatively large increase after installing the vibration isolator, while IMF2~IMF4 and IMF7~IMF8 of 4.8 m SSFST ground vibration have a significant increase; under the action of fatigue load, all orders of IMF showed a downward trend. The decrease range of IMF5~IMF8 of 3.6 m SSFST ground vibration was

greater than 4.8 m SSFST, while the reduction ratio of IMF1~IMF4 was less than 4.8 m SSFST.



(**a**) 3.6 m (**b**) 4.8 m

Figure 15. IMF kurtosis trend graph of basic acceleration.

Table 6. The ratio of changes in kurtosis values.

Size	Working Condition	IMF1	IMF2	IMF3	IMF4	IMF5	IMF6	IMF7	IMF8
3.6 m	With vibration isolator	0.598	2.167	2.378	1.757	0.871	0.376	0.556	0.773
	5 million times	0.788	0.933	0.720	0.981	0.959	0.981	0.936	0.975
	10 million times	0.705	0.687	0.619	0.764	0.838	1.282	0.919	0.930
4.8 m	With vibration isolator	0.594	1.550	2.608	1.526	1.291	0.724	1.506	1.976
	5 million times	0.901	0.950	0.762	0.650	0.708	0.675	0.823	0.660
	10 million times	0.844	0.804	0.691	0.378	0.620	0.558	0.738	0.520

4.3. Hilbert Envelope Spectrum Analysis

In the field of mechanical engineering, Hilbert envelope spectrum is often used for fault diagnosis. This method is based on the traditional spectrum analysis technology, combined with the envelope detection of the vibration signal, which can remove the high-frequency components of the vibration signal and amplify the defect signal to facilitate diagnosis and analysis. The function after the Hilbert transformation can be used to construct the analytic function, and then the Hilbert envelope spectrum can be obtained by performing FFT on this basis [35,36].

The Hilbert envelope spectrum analysis method is as follows.

The original vibration signal undergoes Hilbert transform, which makes the vibration signal produce a phase shift of 90°. The transformed analytical signal can be used to construct an envelope signal. The Hilbert transform is defined as:

$$\hat{x}(t) = H[x(t)] = \int_{-\infty}^{+\infty} \frac{x(\tau)}{\pi(t-\tau)} d\tau = x(t) \cdot \frac{1}{\pi t} \tag{4}$$

$\hat{x}(t)$ is the signal after the original vibration signal x(t) is transformed by Hilbert, and then the analytical signal can be obtained:

$$Z(t) = x(t) + j\hat{x}(t) \tag{5}$$

On this basis, the envelope function $A(t)$ is constructed, which is defined as follows:

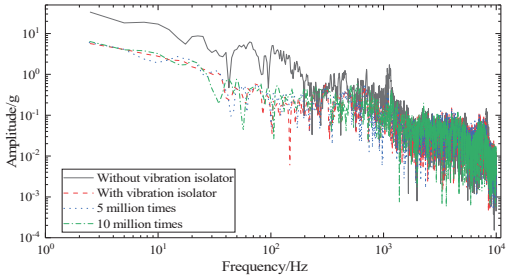
$$A(t) = \frac{\hat{x}(t)}{x(t)} \tag{6}$$

where: $A(t)$—the envelope signal of the original vibration signal $x(t)$.

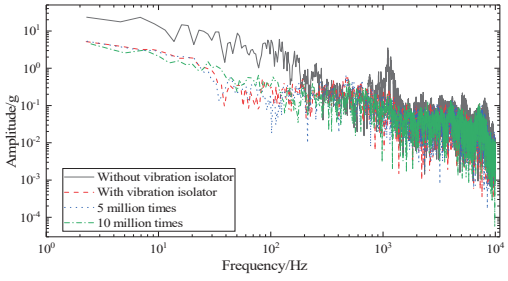
The instantaneous phase $\varphi(t)$ is defined as:

$$\varphi(t) = arctan[\frac{\hat{x}(t)}{x(t)}] \tag{7}$$

Since the vibration of SSFST foundation is most concerned about the middle and low frequencies, according to the above steps, the original vibration signal is detected and the high-frequency components are removed, so that the Hilbert envelope spectrum is drawn. The low-frequency defect signals in the spectrum are easier to identify through the method. Therefore, the Hilbert envelope spectrum analysis is performed on the SSFST basic vibration acceleration of two lengths, as shown in Figure 16.



(**a**) 3.6 m



(**b**) 4.8 m

Figure 16. Hilbert envelope spectrum of basic acceleration.

It can be seen that after Hilbert detection, the amplitude of frequencies above 1000 Hz is greatly reduced. The vibration acceleration envelope signal when there is no vibration isolator can be regarded as a fault signal. It can be seen that the frequency amplitude below 110 Hz is larger than that of the vibration isolator. After the fatigue load, the Hilbert envelope waveform has certain fluctuations, especially in the frequency range between 10 Hz and 30~50 Hz. The laws reflected under different fatigue loading times are consistent with the previous analysis results, that is, 5 million times of fatigue has a small effect, and 10 million times of fatigue has a greater effect.

According to the Hilbert envelope spectrum, as the frequency increases, the Hilbert amplitude gradually decreases, indicating that the foundation structure receives mainly the drop hammer impact transmitted from the track structure. The maximum Hilbert amplitudes of the 3.6 m SSFST ground vibration under the four working conditions are 33.413 g, 6.408 g, 6.000 g, and 5.670 g, respectively. In the range of 0~80 Hz, the prominent protruding points without vibration isolators are located at 7.50 Hz, 24.98 Hz, and 64.52 Hz, and there is a significant depression point at 42.49 Hz. There are many fluctuations in the basic Hilbert amplitude under different fatigue loading times. In the range of 0~80 Hz, before fatigue, 5 million times of fatigue, and 10 million times of fatigue, the first large fluctuation occurs at about 10 Hz, where the Hilbert amplitudes are 17.061 g, 2.788 g,

1.935 g, and 3.245 g. The maximum Hilbert amplitudes of the 4.8 m SSFST ground vibration under the four working conditions are 23.431 g, 5.316 g, 5.173 g, and 4.490 g, respectively. In the range of 0~80 Hz, the prominent protruding points without vibration isolators are located at 6.89 Hz, 16.08 Hz, 22.98 Hz, 32.16 Hz, and 50.54 Hz, and there is a significant depression point at 39.06 Hz. The first large fluctuation appeared at 5 Hz, and the Hilbert amplitudes were 17.622 g, 3.919 g, 3.731 g, and 2.577 g, respectively.

5. Vibration Level Analysis

5.1. Frequency Division Vibration Level Analysis

According to the "Technical Specification for Floating Slab Track" (CJJ/T 191-2012) [37], when evaluating the vibration-damping performance of the SSFST in this study, the test frequency range that needs to be used is 1~200 Hz. The vertical acceleration should be used in the test, and the vibration acceleration with or without vibration isolation measures should be tested for comparison by calculating the root mean square difference ΔLa of the divided frequency vibration level when the SSFST with or without vibration isolator is used as the evaluation parameter. The difference ΔL_{max} of the frequency division vibration level is used as the evaluation quantity of the SSFST vibration-damping effect, and the calculation formula is as follows:

$$\Delta L_a = 10\lg\left(\sum_{i=1}^{n} 10^{\frac{VL_q(i)}{10}}\right) - 10\lg\left(\sum_{i=1}^{n} 10^{\frac{VL_h(i)}{10}}\right) \tag{8}$$

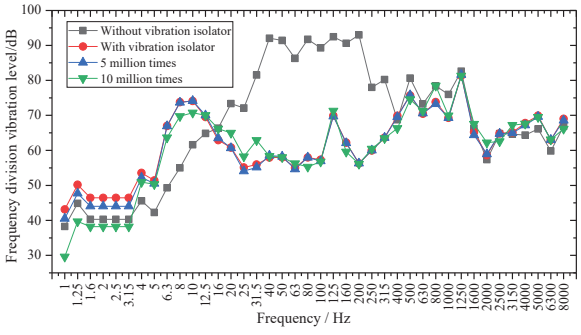
$$\Delta L_{\max} = \max_{i=1\to n}\left[VL_q(i) - VL_h(i)\right] \tag{9}$$

$$\Delta L_{\min} = \min_{i=1\to n}\left[VL_q(i) - VL_h(i)\right] \tag{10}$$

where $V_{Lq(i)}$—in this study, the measured vertical acceleration on the ground without vibration isolators is the divided-frequency vibration level when the vertical acceleration is in the 1/3 octave band;

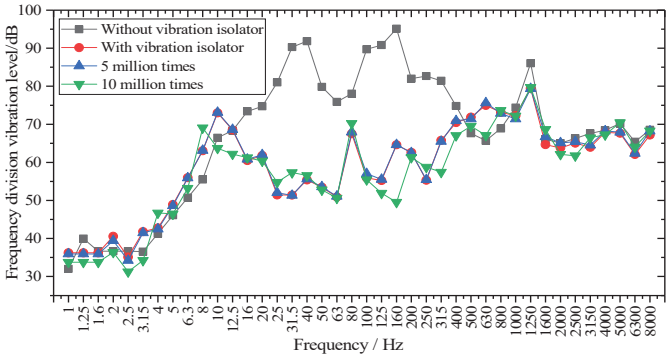
$V_{Lh(i)}$—in this study, the tested vertical acceleration is the divided-frequency vibration level when the vertical acceleration is in the 1/3 octave band.

The 1/3 octave frequency at the basis of the two lengths of SSFST is shown in Figure 17, and the insertion loss analysis is shown in Figure 18.



(**a**) 3.6 m.

Figure 17. *Cont.*

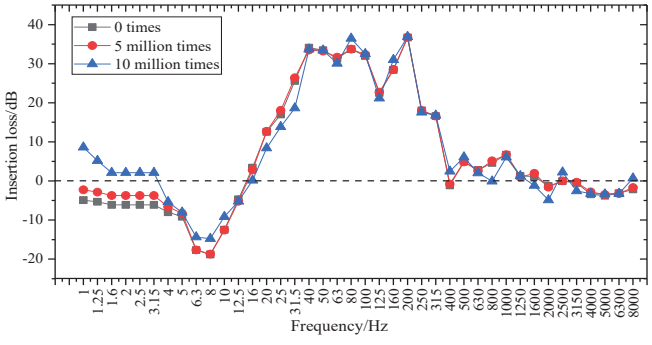


(**b**) 4.8 m.

Figure 17. 1/3 octave waveform at the base of SSFST.

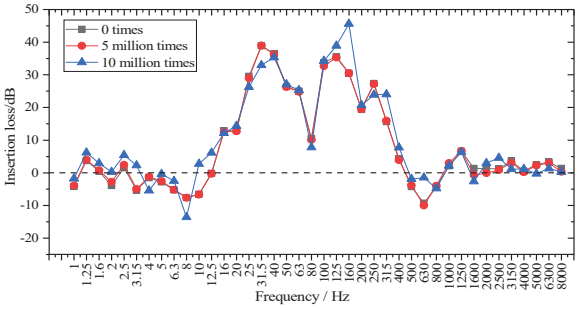
According to Figure 17, in the range of 1~200 Hz, the 1/3 octave at the foundation is distributed at 20~95 dB, and the peak value is at 100~200 Hz when the vibration isolator is not installed, located at 40 Hz, 80 Hz, 125 Hz, 200 Hz, 500 Hz, 800 Hz, and 1250 Hz, respectively, and their frequency positions correspond to the seven peak points analyzed in Section 3.2 (44.99 Hz, 89.98 Hz, 132.48 Hz, 194.98 Hz, 534.93 Hz, 789.91 Hz, and 1254.84 Hz) and are consistent. After the vibration isolator is installed, the vibration level at the base is significantly reduced in the frequency range of 16~400 Hz, and the peak value is at 8~12.5 Hz. The peak points correspond to 10.50 Hz, 134.98 Hz, 524.93 Hz, and 1222.35 Hz. The peak point of the vibration level curve under the fatigue load of 5 million times changes little, while the vibration level curve fluctuates greatly after the fatigue load of 10 million times.

From the comparison of Figure 17a,b, it can be seen that with the increase of fatigue load, the 1/3 octave of the foundation at different frequency bands fluctuates to a certain extent, but the variation range is small. The two-length SSFSTs have very similar basic frequency division vibration level curves under various working conditions, and each amplitude frequency division vibration level of the 4.8 m SSFST foundation is slightly smaller than that of the 3.6 m SSFST foundation.



(**a**) 3.6 m.

Figure 18. *Cont.*



(**b**) 4.8 m.

Figure 18. Insertion loss curve at the foundation of SSFST.

In order to study the effect of fatigue on the insertion loss of steel spring isolators, the fatigue loss of two SSFSTs was plotted as a function of fatigue load, as shown in Figure 18. It can be seen that in the range of 1~200 Hz, the vibration isolation effect of the vibration isolator on the foundation vibration is mainly reflected in the 16~400 Hz range. This phenomenon is consistent with the results in Section 3.2, that is, the vibration damping of the steel spring vibration isolator is mainly reflected in low frequency. At the same time, the insertion loss of the base frequency in the 5~16 Hz range is negative, indicating that the vibration in this frequency range is intensified, which is consistent with the conclusion that the maximum amplitude appears around 10 Hz after the isolator is installed in the analysis results of the amplitude spectrum (RMS). With the increase of fatigue load, the vibration isolator has a certain fluctuation in the insertion loss of the foundation at different frequency bands, but the variation amplitude is limited.

5.2. Z Vibration Level Analysis

The Z vibration levels (VLz) at different track structure components of the two SSFSTs are shown in Table 7 and Figure 19. It can be seen that during the drop hammer impact, when weighing according to ISO2631-1:1985 [38], the spring vibration isolator's foundation VLz gradually decreases before and after installation, after 5 million fatigue tests, and after 10 million fatigue tests. The Z vibration level results of the two sizes of SSFST are similar, and it can be seen that the increase in fatigue load will not weaken the vibration-damping effect of SSFST. The relevant literature [39] shows that the load caused by the operation of the subway line in one year is equivalent to 2 million fatigue loads. In this experiment, the vibration-damping effect changes little after 5 million and 10 million times, which just proves that the floating slab track can still guarantee a good vibration-damping effect after about 2.5 years and 5 years, which shows that the track structure has high reliability and has a shelf life of vibration-damping effect about 5 years.

Table 7. The Z vibration level of the foundation weighting (VLz, unit: dB).

Track Size	Data	Track Structure Components			
		Without Vibration Isolator	With Vibration Isolator	After 5 Million Fatigue	After 10 Million Fatigue
3.6 m	Z vibration level	96.0	73.9	72.8	72.2
	Insertion loss		22.1	23.2	23.8
4.8 m	Z vibration level	94.0	73.7	72.9	72.4
	Insertion loss		20.3	21.1	21.6

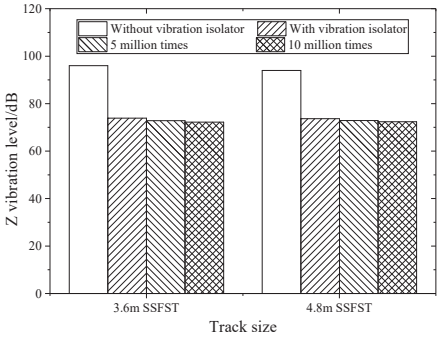


Figure 19. The Z vibration level of the foundation.

5.3. Comparison with Cast-in-Place SSFST Railway Field Test

The combined shear hinge prefabricated SSFST has not been widely laid on the railway site at present. In order to reflect the advanced nature of the track slab, we compared the field test data of the cast-in-place SSFST previously tested with it. The results show that this new type of combined shear hinge SSFST is not only beneficial to the rapid construction on site but also has good engineering application value, and its vibration-damping effect can also reach the level of cast-in-place SSFST. At the same time, the frequency response of the vibration isolator reflected by the prefabricated SSFST indoor test and the cast-in-place SSFST field test is consistent, and it can be seen that the drop-weight method has high feasibility for evaluating the vibration damping characteristics of the SSFST that has not been constructed on site.

The wall vibration of a cast-in-place SSFST tunnel of a subway line is selected to be measured by an acceleration sensor, and the length of the cast-in-place SSFST is 25 m. At the same time, an ordinary monolithic track (OMT) section was selected for comparison with the tunnel wall vibration, and each measurement section used an acceleration sensor to measure the vertical acceleration of the tunnel wall (Figure 20). The site layout is shown in the Figure 21.

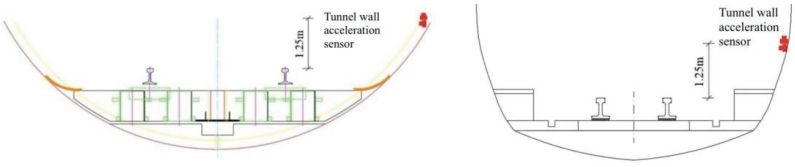


Figure 20. Layout of tunnel wall acceleration test points. (**a**) Cast-in-place SSFST section. (**b**) OMT section.

Figure 21. Field test diagram. (**a**) Acceleration measuring point of tunnel wall. (**b**) Overall view.

Through the test, the comparison chart of tunnel wall vibration in Figure 22 is obtained. For the tunnel wall acceleration, the SSFST section is significantly smaller than the OMT section, and the amplitude is about 0.2~0.3 of the OMT section. It can be seen that SSFST has excellent vibration isolation effect, which can filter and weaken the vibration very efficiently, which is of great significance to ensure the vibration damping requirements along the subway line. In the frequency band below 12.5~16 Hz, the insertion loss of the frequency division vibration level of SSFST is negative, indicating that the setting of SSFST will aggravate the vibration response of this frequency band. The reason is that the natural vibration frequency of SSFST is mainly located in this section. Excitation in the frequency band results in increased vibration in this frequency band. The insertion loss curve of the frequency band near 3150 Hz shows a valley value, and the vibration-damping effect of the high-frequency band also decreases, indicating that the vibration-damping effect of SSFST is mainly concentrated in the 16~3150 Hz frequency band. Through the analysis of the insertion loss curve, it can be seen that the vibration-damping effect reaches its peak value near 200~250 Hz, which indicates that the vibration isolation and filtering of the steel spring vibration isolator for medium and high-frequency signals are mainly concentrated in this frequency band. Therefore, the vibration damping characteristics of SSFST reflected by the field test are consistent with the results of the laboratory test.

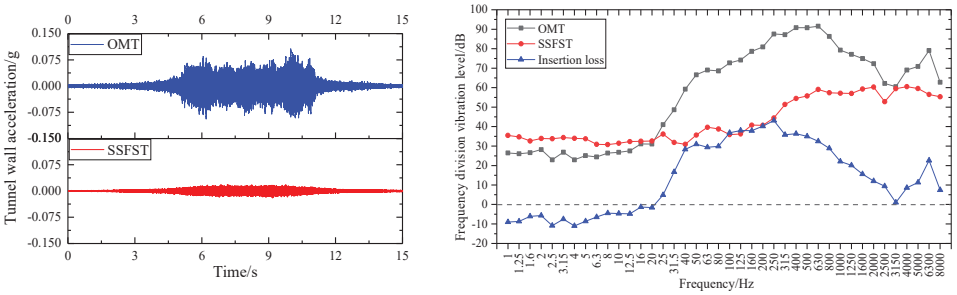


Figure 22. Comparison of vertical vibration of tunnel wall. (**a**) Acceleration time domain curve. (**b**) Frequency division vibration level and insertion loss.

According to "Technical Guidelines for Environmental Impact Assessment" HJ453-2008 [40], for the environmental vibration caused by train operation, the train speed correction formula at different speeds is:

$$C_v = 20\lg\frac{v}{v_0} \tag{11}$$

where:

v_0—the reference speed of the source intensity, in km/h;
v—the speed of the train, in km/h.

Taking the data measured when 20 trains pass through the test section, the VL_Z value and the vibration difference value ΔVL_Z of the tunnel wall measurement point of each test section under the two track bed forms are analyzed and statistically obtained, see Table 8. After the vehicle speed is corrected to the same speed, the Z vibration level of the tunnel wall in the SSFST section is significantly smaller than that in the OMT section, and the insertion loss ΔVL_Z is 24.42 dB. Compared with the combined shear hinge prefabricated SSFST in this test, the vibration-damping effect of the two types of SSFST is above 20 dB. Due to the different working conditions without vibration isolators, the vibration level and the value of the vibration-damping effect are different, but the overall law and the magnitude of the vibration-damping effect are consistent and the reliability is high.

Table 8. Tunnel wall VL_Z and its vibration difference ΔVL_Z.

Track Bed Form	OMT	SSFST
VL_Z (dB)	70.83	46.61
Speed (km/h)	71.36	73.00
Corrected value after the speed is unified (dB)	0.00	−0.20
VL_Z after speed correction (dB)	70.83	46.41
Insertion loss ΔVL_Z (dB)	24.42	

In summary, the prefabricated SSFST indoor test and the cast-in-place SSFST field test reflect the same vibration damping frequency response law of the isolator. It can be seen that the drop-weight method has high feasibility for evaluating the vibration damping characteristics of the SSFST that is not constructed on site.

6. Conclusions

By establishing a full-scale model of the combined shear hinge SSFST, the drop hammer impact test of the combined shear hinge SSFST under different isolator assembly conditions and fatigue times was carried out. The dynamic response of the foundation (ground) of two lengths of SSFST structures under the impact of a drop weight was investigated. According to the time domain, frequency domain, vibration level, and other indicators, the vibration-damping performance of the combined shear hinge SSFST structure is compared, and the vibration excitation response of the SSFST foundation under different working conditions is evaluated. The main conclusions are as follows:

(1) Through amplitude (RMS) analysis, EMD analysis, kurtosis value analysis, Hilbert envelope spectrum analysis, and vibration level analysis, it can be seen that this novel type of combined shear hinge SSFST can have excellent isolation effect on vibration and can still maintain good vibration damping ability within 10 million fatigue loads (about 5 years).

(2) By comparing the magnitude (RMS) spectrum, the IMF waveform of the EMD method, the frequency division vibration level curve and the insertion loss curve, it can be seen that the frequency of the peak point of the frequency division vibration level curve corresponds to the peak point frequency of the amplitude (RMS) spectrum. The vibration damping results of each frequency band reflected by the insertion loss curve are consistent with the IMF display results after EMD transformation, that is, the vibration-damping effect of the steel spring vibration isolator mainly acts on the middle and low-frequency bands of 16–400 Hz.

(3) According to the amplitude (RMS) spectrum, frequency division vibration level curve, and insertion loss curve, it can be seen that the vibration near 10 Hz will be intensified after the vibration isolator is installed, while the human body is more sensitive to the vibration response in the frequency range of 0~80 Hz. Therefore, this problem should be improved in the subsequent SSFST design to weaken the vibration response of the floating slab to the foundation in this frequency band.

(4) The vibration indicators and variation laws of the two sizes of SSFST are similar, and the acceleration amplitude and attenuation laws are consistent. The vibration response of 4.8 m SSFST is slightly smaller than that of 3.6 m SSFST. Since 3.6 m SSFST is easier to lay a smooth line, in practical applications, 3.6 m SSFST should be arranged in the curved section, and 4.8 m SSFST should be arranged in the straight-line section with higher train speed.

(5) Compared with the field experiment method, the drop-weight method can also systematically study the vibration characteristics of this novel combined shear hinge SSFST before laying, and the displayed SSFST vibration damping characteristics are consistent with the field test and can conduct a comprehensive evaluation of the vibration-damping effect of SSFST at a lower cost.

Author Contributions: Conceptualization, Z.Z. and W.W.; methodology, Z.L.; software, X.H.; validation, A.A.S. and X.H.; formal analysis, X.H.; investigation, X.H.; resources, Z.Z.; data curation, Z.L.; writing—original draft preparation, X.H. and Z.S.; writing—review and editing, Y.Y. and X.H.; visualization, X.H.; supervision, Z.Z.; project administration, W.W.; funding acquisition, Z.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the High-speed Railway Joint Fund of National Natural Science Foundation of China grant Number U1734208, the Major Program of National Natural Science Foundation of China grant Number 11790283, and the Hunan Provincial Natural Science Foundation of China grant Number 2019JJ40384.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Grootenhuis, P. Floating track slab isolation for railways. *J. Sound Vib.* **1977**, *51*, 443–448. [CrossRef]
2. Ding, D.; Liu, W.; Li, K.; Sun, X.; Liu, W. Low frequency vibration tests on a floating slab track in an underground laboratory. *J. Zhejiang Univ. Sci. A Appl. Phys. Eng.* **2011**, *12*, 345–359. [CrossRef]
3. Zeng, Z.; Huang, X.; Wang, W.; Zhu, Z.; Wu, Z. Comparative Research on the Vibration Transmission Law of Different Subway Track Types. *J. Railw. Eng. Soc.* **2021**, *38*, 74–79. (In Chinese)
4. Ji, T.; Qiu, L.; Zhou, B. Theoretical analysis of long floating slab track(part I)—Dispersion graph and critical velocity. *Chin. J. Comput. Mech.* **2009**, *26*, 919–923.
5. Cui, F.; Chew, C. The effectiveness of floating slab track system—Part I. Receptance methods. *Appl. Acoust.* **2000**, *61*, 441–453. [CrossRef]
6. Zeng, Z.; Huang, X.; Wang, J.; Liu, F.; Wang, W.; Shuaibu, A. Wheel–rail stochastic dynamics and rail wear analysis of small radius waveformd sections of a tram line based on generalized probability density evolution. *Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit* **2021**, *235*, 543–588. [CrossRef]
7. Dere, Y. Effectiveness of the floating slab track system constructed at Konya Light Rail. *Measurement* **2016**, *89*, 48–54. [CrossRef]
8. Geng, C.; Tian, M. Vibration Frequency Analysis of Floating Slab Track System. *Urban Mass Transit* **2007**, *53*, 155–164.
9. Wang, L.; Xuan, Y.; Wan, J.; Jiang, J. Simulation research on the vibration insulation performance of the floating slab track structure. *China Railw. Sci.* **2005**, *26*, 48–52.
10. Zhou, B.; Xie, X.; Yang, Y. Simulation of wave propagation of floating slab track-tunnel-soil system by 2d theoretical model. *Int. J. Struct. Stab. Dyn.* **2013**, *14*, 1350051. [CrossRef]
11. Sun, C.; Gao, L.; Ma, J. Environmental vibration isolation tests and analysis on steel spring floating slab track used in subway. *Int. Symp. Environ. Vib.* **2011**, *1*, 840–842.
12. Ding, D.; Liu, W.; Zhang, B.; Sun, X. Modal Analysis on the Floating Slab Track. *J. China Railw. Soc.* **2008**, *30*, 61–64. (In Chinese)
13. Liu, W.; Ding, D.; Li, K.; Zhang, H. Experimental study of the low-frequency vibration characteristics of steel spring floating slab track. *China Civ. Eng. J.* **2011**, *44*, 118–125. (In Chinese)
14. Wei, K.; Dou, Y.; Yang, Q. Random vibration analysis and parameter optimization of steel-spring floating-slab track. *J. Huazhong Univ. Sci. Technol. (Nat. Sci. Ed.)* **2017**, *45*, 115–119.
15. Kang, L.; Lei, Z. Shield Tunnel Structure Modal Analysis with Floating Slab Track on Elastic Foundation. *Adv. Mater. Res.* **2013**, *838–841*, 1393–1398. [CrossRef]
16. Zhu, S.; Cai, C. Interface damage and its effect on vibrations of slab track under temperature and vehicle dynamic loads. *Int. J. Non-Linear Mech.* **2014**, *58*, 222–232. [CrossRef]
17. Zhu, S.; Luo, J.; Wang, M.; Cai, C. Mechanical characteristic variation of ballastless track in high-speed railway: Effect of train–track interaction and environment loads. *Railw. Eng. Sci.* **2020**, *28*, 94–109. [CrossRef]
18. Zeng, Z.; Huang, X.; Ye, M.; Zhu, Z.; Xiao, Y.; Peng, G.; Hu, G.; Hu, J.; Xu, R. Combined shear hinge device of floating slab track. *Chin. Invent. Pat.* **2021**, 1–8. (In Chinese)
19. Luo, J.; Zhu, S.; Zhai, W. An advanced train-slab track spatially coupled dynamics model: Theoretical methodologies and numerical applications. *J. Sound Vib.* **2021**, *501*, 116059. [CrossRef]
20. Zeng, Z.; Huang, X.; Yan, B.; Wang, W.; Shuaibu, A.; He, X. Research on the fatigue performance of self-compacting concrete structure in CRTSIII slab ballastless track under the action of heavy haul train. *Constr. Build. Mater.* **2021**, *303*, 124465. [CrossRef]
21. *GB 10071-88*; Urban Area Environmental Vibration Measurement Method. National Environmental Protection Agency: Beijing, China, 1989.

22. Huang, N.; Shen, Z.; Long, S.; Wu, M.; Shih, H.; Zheng, Q.; Yen, N.; Tung, C.; Liu, H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **1998**, *454*, 903–995. [CrossRef]
23. Gu, J.; Peng, Y.; Lu, H.; Cao, B.; Chen, G. Compound fault diagnosis and identification of hoist spindle device based on hilbert huang and energy entropy. *J. Mech. Sci. Technol.* **2021**, *35*, 4281–4290. [CrossRef]
24. Gao, C.; Zhou, S. Wavelet transform threshold noise reduction methods in the oil pipeline leakage monitoring and positioning system. *J. Elec.* **2010**, *27*, 405–411. [CrossRef]
25. Li, Y.; Lin, J.; Niu, G.; Wu, M.; Wei, X. A Hilbert–Huang Transform-Based Adaptive Fault Detection and Classification Method for Microgrids. *Energies* **2021**, *14*, 5040. [CrossRef]
26. Li, X.; Chen, S.; Liu, S.; Li, Z. AE waveform characteristics of rock mass under uniaxial loading based on Hilbert-Huang transform. *J. Cent. South Univ.* **2021**, *28*, 1843–1856. [CrossRef]
27. Diao, Y.; Jia, D.; Liu, G.; Sun, Z.; Xu, J. Structural damage identification using modified Hilbert–Huang transform and support vector machine. *J. Civ. Struct. Health Monit.* **2021**, *11*, 1155–1174. [CrossRef]
28. Murzagulov, D.; Zamyatin, A.; Romanovich, O. Approach to the Detection of Anomalies in Process Signals by Using the Hilbert–Huang Transform. *Optoelectron. Instrum. Data Process.* **2021**, *57*, 27–36. [CrossRef]
29. Neukirch, M.; García-Jerez, A.; Villaseñor, A.; Luzón, F.; Ruiz, M.; Molina, L. Horizontal-to-Vertical Spectral Ratio of Ambient Vibration Obtained with Hilbert–Huang Transform. *Sensors* **2021**, *21*, 3292. [CrossRef] [PubMed]
30. Alberti, T.; Milillo, A.; Laurenza, M.; Massetti, S.; Ivanovski, S.; Ippolito, A.; Plainaki, C.; De Angelis, E.; Mangano, V.; Mura, A.; et al. Multiscale Features of the Near-Hermean Environment as Derived Through the Hilbert-Huang Transform. *Front. Phys.* **2021**, *9*, 668098. [CrossRef]
31. Das, I.; Arif, M.; Oo, A.; Subhani, M. An Improved Hilbert–Huang Transform for Vibration-Based Damage Detection of Utility Timber Poles. *Appl. Sci.* **2021**, *11*, 2974. [CrossRef]
32. Zhao, Y.; Shan, R.; Wang, H. Research on vibration effect of tunnel blasting based on an improved Hilbert–Huang transform. *Environ. Earth Sci.* **2021**, *80*, 206. [CrossRef]
33. Ezzeiri, S.; Hamdi, E. Ultrasonic rock microcracking characterization and classification using Hilbert–Huang transform. *Innov. Infrastruct. Solut.* **2020**, *5*, 100. [CrossRef]
34. AlSaleh, R.; Ni, Y. Condition assessment of a supertall structure by an improved Hilbert-Huang transform and empirical mode decomposition method. *Int. J. Sustain. Mater. Struct. Syst.* **2020**, *4*, 244–267.
35. Harsuko, M.; Zulfakriza, Z.; Nugraha, A.; Sarjan, A.; Widiyantoro, S.; Rosalia, S.; Puspito, N.; Sahara, D. Investigation of Hilbert–Huang Transform and Fourier Transform for Horizontal-to-Vertical Spectral Ratio Analysis: Understanding the Shallow Structure in Mataram City, Lombok, Indonesia. *Front. Earth Sci.* **2020**, *8*, 334. [CrossRef]
36. Zhou, H.; Jia, M. Analysis of rolling bearing fault diagnosis based on EMD and kurtosis Hilbert envelope demodulation. *J. Mech. Electr. Eng.* **2014**, *31*, 1136–1167. (In Chinese)
37. *CJJ/T 191-2012*; Technical Code for Floating Slab Track. Ministry of Housing and Urban-Rural Development of the People's Republic of China: Beijing, China, 2012. (In Chinese)
38. *ISO 2631/1-1985*; Evaluation of Human Exposure to Whole-Body Vibration. The International Organization for Standardization: Delft, The Netherlands, 1985.
39. Lian, S.; Yin, X. *Theoretical Research and Practice of Floating Slab Track*; China Railway Press: Beijing, China, 2021. (In Chinese)
40. *HJ453-2018*; Technical Guidelines for Environmental Impact Assessment Urban Rail Transit. Ministry of Ecology and Environment: Beijing, China, 2018. (In Chinese)

applied sciences

MDPI

Article

Damage Detection of Continuous Beam Bridge Based on Maximum Successful Approximation Approach of Wavelet Coefficients of Vehicle Response

Kai Liu, Haopeng Qi and Zengshou Sun *

Civil Engineering, Zhengzhou University, Zhengzhou 450001, China; 202012212014069@gs.zzu.edu.cn (K.L.); 202022212014173@gs.zzu.edu.cn (H.Q.)
* Correspondence: zengshou@zzu.edu.cn

Abstract: In view of problems such as closed traffic, the large number of sensors required, and the labor-intensive and time-consuming nature of previous bridge detection, this paper analyzes the dynamic response of the vehicle body of the continuous girder bridge under the action of vehicle load. Based on theoretical analysis and formula derivation, a new method of bridge damage detection based on coupled vehicle–bridge vibration is conceived. This method can accurately identify the location of bridge damage and approximately estimate the degree of bridge damage. The method is as follows: Taking the continuous beam bridge as an example, first, use the tractor inspection vehicle model to drive over the continuous beam bridge before and after the damage, and collect the acceleration response of the vehicle body. Then, the acceleration response difference is transformed by wavelet transform. Furthermore, perform the innovative use of the maximum successive approximation approach to process wavelet transform coefficients, which can identify the location of the bridge damage. Additionally, study the impact of vehicle speed, vehicle weight, road surface roughness, and noise on this damage detection method. In addition, a method for judging bridge damage degree based on wavelet transform coefficients is proposed, and the judgment error basically meets the requirements.

Keywords: vehicle scanning method (VSM); damage identification; degree of damage; continuous beam bridge

Citation: Liu, K.; Qi, H.; Sun, Z. Damage Detection of Continuous Beam Bridge Based on Maximum Successful Approximation Approach of Wavelet Coefficients of Vehicle Response. *Appl. Sci.* **2022**, *12*, 3743. https://doi.org/10.3390/app12083743

Academic Editors: Phong B. Dao, Liang Yu and Lei Qiu

Received: 12 February 2022
Accepted: 6 April 2022
Published: 8 April 2022

1. Introduction

With the rapid development of the number of bridges and the speed of bridge construction, bridge destruction or collapse accidents occur from time to time, causing serious economic losses and negative social impacts. At 8:45 on 14 July 2011, the deck of Fujian Wuyishan Mansion Bridge collapsed, causing a bus to fall into the riverbank about 8.8 m below the bridge. The bus driver was killed on the spot and 22 people on board were injured [1]. If bridges are regularly inspected and repaired, disasters can be avoided. The use of renewable materials such as tire rubber [2] in bridge damage can effectively improve its performance. Therefore, bridge damage detection and assessment have gradually become a popular research direction. Damage is the change in the geometrical and material properties of a structure, resulting in structural deformation, increased vibration, or reduced load-bearing capacity, and failure or partial failure of the structure [3]. The overall detection method of bridge structures can be divided into the static detection method and the dynamic detection method [4]. The dynamic detection method uses the structural dynamic response as a measure of the overall state of the structure. When a structural component is damaged, the stiffness, mass, or damping characteristics of the structural member will change, and the structural dynamic response will also change accordingly. The collected structure dynamic response data is processed and analyzed, which can effectively identify structural damage information [5].

In 2004, Yang proposed that the dynamic response of vehicles crossing the bridge contains the modal information of the bridge, which was verified by numerical simulation [6]. In 2005, Yang conducted a field test and successfully extracted the first-order vibration frequency of the bridge from the corresponding vehicle by using a tractor to tow a single-axle detection vehicle [7]. When the bridge is damaged, the stiffness will change, so the identified modal parameters can be used to determine the damage. In addition, when the vehicle passes through the damage location, its dynamic response signal will suddenly change, and the time–frequency domain signal analysis method can also be used to locate the damage. Under the influence of Yang et al., Bu et al. proposed to use the dynamic response of a vehicle body for bridge damage identification in 2006, use the stiffness reduction to simulate the damage element, and use the acceleration of the vehicle body as the index for damage identification [8]. In 1974, wavelet transform was first proposed by J. Morlet and applied to petroleum signal processing; as a time–frequency domain analysis method, wavelet transform can effectively identify the singular points in the signal, and the damage location can be effectively identified by using the amplification characteristics of wavelet transform. In 2010, Nguyen and tran carried out wavelet transform on the displacement response signal of mobile vehicles, which can identify multiple damages of bridges [9]. In 2013, sun Zengshou summarized the damage detection methods based on vehicle–bridge coupling vibration response [10]. In 2014, Yang's team and OBrien successfully identified the modes of the bridge using two different methods simultaneously [11,12]. In 2017, Obrien decomposed the intrinsic mode functions (IMFs) of each order from the vehicle corresponding by empirical modal decomposition and identified the damage location by the IMF before and after the damage [13]. Qian Yao introduced the theoretical solution of the contact point response in 2018, using empirical mode decomposition and Hilbert transformation to extract instantaneous amplitude squares from the drive component response, which can identify the damage location [14]. In 2019, Marashi et al. divided the bridge into multiple sections, measured the short-term transmissibility of the axle acceleration time–history response, and then used the amplitude modulation program to connect the local mode shapes of each section in series to obtain the entire mode shape of the bridge [15]. In 2020, Locke et al. established a finite element model for driving recognition of simply supported girder bridges. Through numerical simulation, they studied the effects of temperature changes, vehicle body speed, accompanying traffic flow, bridge deck roughness, and different damage on the driving test results of simply supported girder bridges. It further confirms the validity and application potential of the bridge indirect measurement method based on vehicle response [16].

In the past two decades, scholars have proposed a large number of structural damage identification methods based on the overall parameters of the bridge and the response of the bridge, and have been applied to a certain extent; Clemente P., Bongiovanni G. et al. obtained bridge parameters through a vibration test of Indiano cable-stayed bridge in Florence, and used these parameters to establish a finite element model and analyze the influence of this model on static and seismic loads [17]. Although the results of traditional detection methods are accurate, there are shortcomings such as inconvenient detection and a lot of preparation. Bridge damage detection methods based on vehicle response can overcome the above shortcomings, so this paper combines the existing research results of predecessors, further explores the bridge damage detection method based on vehicle response, uses wavelet analysis as the main tool, analyzes and studies the dynamic response of vehicles when passing through the bridge, and proposes corresponding structural damage indicators. The method proposed in this paper is applicable regardless of whether the pile foundation adopts full-scale reinforced concrete piles or reinforced micropiles [18]. Although the detection method proposed in this paper can quickly and conveniently identify the bridge damage, its detection results are greatly affected by the accuracy of the sensor. In addition, the experiment using this method has not been carried out yet, and further research is needed.

2. Theoretical Research on Indirect Measurement Method Based on Vehicle–Bridge Coupling

The damage detection method based on vehicle response is based on the vehicle–bridge coupling vibration system model, as shown in Figure 1, which analyzes the interaction mechanism between vehicle and bridge and obtains the bridge damage information. This method is a continuation of the damage detection method of multi-point input single-point output, in the hope that the damage detection can be carried out through the dynamic response of bridge vehicles, which can greatly reduce the number of sensors and does not affect the traffic.

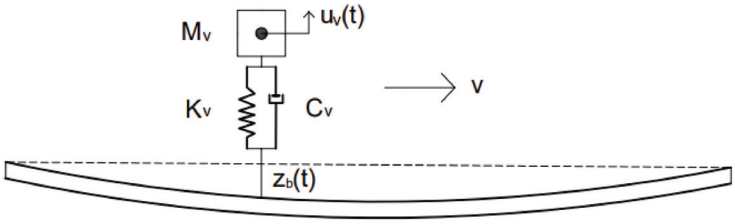


Figure 1. Mathematical model for vehicle scanning method.

2.1. Theoretical Research on Vehicle–Axle Coupling Vibration

The whole mathematical model can be regarded as two subsystems; one is the vehicle subsystem and the other is the bridge subsystem. In the solution of vehicle–bridge coupling vibration equation [19], bridge mass, damping and stiffness matrices are generally constant matrices, bridge damping is selected as Rayleigh damping, and differential equations for bridge and vehicle vibration can be expressed as:

$$M_b\ddot{Z}_b(t) + C_b\dot{Z}_b(t) + K_bZ_b(t) = P_b(t) \tag{1}$$

where: M_b is the bridge quality matrix, C_b is the bridge damping matrix, K_b is the bridge stiffness matrix, Z_b is the bridge displacement matrix, $\dot{Z}_b$ is the bridge speed matrix, and $\ddot{Z}_b$ is the bridge acceleration matrix.

$$[M_v]\{\ddot{u}_v(t)\} + [C_v]\{\dot{u}_v(t)\} + [K_v]\{u_v(t)\} = \{P_v(t)\} \tag{2}$$

where $[M_v]$, $[C_v]$, and $[K_v]$ are vehicle generalized mass matrix, generalized damping matrix, and generalized stiffness matrix, respectively, $\{\ddot{u}_v(t)\}$, $\{\dot{u}_v(t)\}$, and $\{u_v(t)\}$ are vehicle acceleration vector, velocity vector, and displacement vector, respectively, and $\{P_v(t)\}$ is load vector.

In the vehicle–bridge coupling vibration system, the solution process of differential equation of vehicle–bridge vibration is based on the geometric compatibility and mutual constraint of force balance at the contact point between wheel and bridge surface, the force acting on the wheel is obtained by using the relative displacement between the axle, and the vehicle matrix operation is used to obtain the force of the vehicle on the bridge. The displacement of the wheel relative to the road surface can be expressed as:

$$\Delta_i = Z_{wi} - Z_i - r_i(x) \tag{3}$$

where Z_{wi} is the vertical displacement of the wheel, Z_i is the vertical displacement of the bridge at the contact point between the bridge and the wheel, and $r_i(x)$ is the pavement roughness at the bridge and point.

There is a pair of forces with equal size and opposite direction at the contact point between the wheel and the bridge. The gravity of the vehicle on the bridge should be considered in calculation and programming. The force balance can be expressed as:

$$F_{wi} = -F_{bi} = c_{wi}(\Delta_i') + k_{wi}(\Delta_i) \tag{4}$$

where c_{wi} is wheel damping, k_{wi} is the wheel stiffness, and Δ_i is the vertical relative displacement between wheel and pavement.

The equation selects the Newmark-β method for solving the vehicle–bridge coupling system. The displacement and velocity of vehicles calculated by the Newmark-β method are used as the initial conditions for the next time interval to solve the vehicle–bridge coupling system. The differential equation of motion of the vehicle is:

$$[M]\{\ddot{u}(t)\} + [C]\{\dot{u}(t)\} + [K]\{u(t)\} = \{P(t)\} \tag{5}$$

where $\{\ddot{u}\}$ can be represented by linear acceleration interpolation, and $\{\ddot{u}_i\}$ can be represented by α control parameters, that is:

$$\{\ddot{u}\} = (1-\delta)\{\ddot{u}_i\} + \delta\{\ddot{u}_{i+1}\}\ (0 \le \delta \le 1) \tag{6}$$

$$\{\ddot{u}\} = (1-2\alpha)\{\ddot{u}_i\} + 2\alpha\{\ddot{u}_{i+1}\}\ \left(0 \le \alpha \le \frac{1}{2}\right) \tag{7}$$

In turn, the structural acceleration vector and velocity vector at the t_{i+1} moment can be derived:

$$\{\ddot{u}_{i+1}\} = a_0(\{u_{i+1}\} - \{u_i\}) - a_2\{\dot{u}_i\} - a_3\{\ddot{u}_i\} \tag{8}$$

$$\{\dot{u}_{i+1}\} = a_1(\{u_{i+1}\} - \{u_i\}) - a_4\{\dot{u}_i\} - a_5\{\ddot{u}_i\} \tag{9}$$

where $a_0 = \frac{1}{\alpha\Delta t^2}$; $a_1 = \frac{\delta}{\alpha\Delta t}$; $a_2 = \frac{1}{\alpha\Delta t}$; $a_3 = \frac{1}{2\alpha} - 1$; $a_4 = \frac{\delta}{\alpha} - 1$; $a_5 = \frac{\Delta t}{2}\left(\frac{\delta}{\alpha} - 2\right)$; $a_6 = \Delta t(1-\delta)$; $a_7 = \delta\Delta t$.

Then, the equation of motion of the structure at time t_{i+1} can be written according to Equation (5):

$$[M]\{\ddot{u}_{i+1}(t)\} + [C]\{\dot{u}_{i+1}(t)\} + [K]\{u_{i+1}(t)\} = \{P_{i+1}(t)\} \tag{10}$$

Replace Equations (8) and (9) into Equation (10):

$$(\alpha_0[M] + \alpha_1[C] + [K])\{u_{i+1}\} = \{P_{i+1}\} + [M](a_0\{u_i\} + a_2\{\dot{u}_i\} + a_3\{\ddot{u}_i\}) + [C](a_1\{u_i\} + a_4\{\dot{u}_i\} + a_5\{\ddot{u}_i\}) \tag{11}$$

According to Equation (10), the structural displacement vector $\{u_{i+1}\}$ at t_{i+1} time can be obtained by substituting the structural displacement vector $\{u_{i+1}\}$ into Equations (7) and (8) to obtain the structural acceleration vector $\{\ddot{u}_{i+1}\}$ and the structural velocity vector $\{\dot{u}_{i+1}\}$ at the t_{i+1} time.

2.2. Theoretical Analysis of Detection Method Based on Vehicle Response

Indirect measurement of bridge damage using vehicle response begins with the bridge frequency indirect measurement method proposed by Yang and Chang et al., and the first-order frequency of the bridge can be identified by performing an FFT transformation on the acceleration response of the vehicle. Yang et al. derived the vehicle acceleration response [20]:

$$\ddot{q}_v(t) = \sum_{i=1}^{n} \left\{ \begin{array}{c} P_i cos(\omega_v t) + Q_i cos(2\omega_{ci} t) + \\ X_i cos[(\omega_{ci} - \omega_i)t] + Y_i cos[(\omega_{ci} + \omega_i)t] \end{array} \right\} \tag{12}$$

where P_i, Q_i, X_i, and Y_i is a time-independent coefficient, ω_i is the $i-$th natural frequency of the bridge, ω_v is the natural vibration frequency of the vehicle, and $\omega_{ci} = i\pi v/L$ is the vehicle loading frequency.

The load vector $\{P_{i+1}\}$ at the right end of the vehicle vibration differential equation can be expressed by the bridge displacement and speed, and the load vector $\{P_{i+1}\}$ of 1/4 vehicle model is:

$$\{P_{i+1}\} = \begin{pmatrix} 0 \\ c_{w1}f'_z(x) + k_{w1}f_z(x) \end{pmatrix} \tag{13}$$

where c_{w1} is wheel damping, k_{w1} is the wheel stiffness, $f_z(x)$ is the bridge displacement, and $f'_z(x)$ is the bridge speed.

With the tractor test truck model, both the tractor and the inspection vehicle can be reduced to a concentrated force compared to the bridge span, as shown in Figure 2.

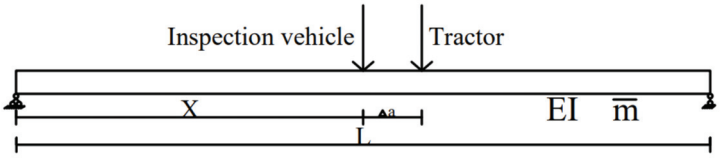


Figure 2. Schematic diagram of simplified calculation for tractor inspection vehicle.

The deflection at the position of the detection vehicle can be calculated using the superposition principle:

$$f_z(x) = \frac{F_c x^2 (x-l)^2}{3EIl} + \frac{F_q}{EI} \cdot \left[-\frac{(l-(x+\Delta a))x^3}{6l} + \left(\frac{(x+\Delta a)^3}{6l} + \frac{(x+\Delta a)l}{3} - \frac{(x+\Delta a)^2}{2} \right) x \right] \tag{14}$$

where Δa is the distance between the tractor and the center of mass of the test vehicle.

Replace Equations (13) and (14) into Formula (11) to detect the displacement vector of the vehicle at the driving position $\{u_{i+1}\}$:

$$\{u_{i+1}\} = \frac{1}{b_2 b_1 - b_3^2} \begin{bmatrix} b_1(b_4+b_6) + b_3(c_{w1}f'_z(x) + k_{w1}f_z(x) + b_5 + b_7) \\ b_3(b_4+b_6) + b_2(c_{w1}f'_z(x) + k_{w1}f_z(x) + b_5 + b_7) \end{bmatrix} \tag{15}$$

where:

$$\begin{gathered} b_1 = a_0 m_{w1} + a_1(c_{u1} + c_{w1}) + (k_{u1} + k_{w1}) \\ b_2 = a_0 m_v + a_1 c_{u1} + k_{u1} \\ b_3 = a_1 c_{u1} + k_{u1} \\ b_4 = m_v\left(a_0 u_{i1} + a_2 \dot{u}_{i1} + a_3 \ddot{u}_{i1}\right) \\ b_5 = m_{w1}\left(a_0 u_{i2} + a_2 \dot{u}_{i2} + a_3 \ddot{u}_{i2}\right) \\ b_6 = c_{u1}\left(a_1 u_{i1} + a_4 \dot{u}_{i1} + a_5 \ddot{u}_{i1}\right) - c_{u1}\left(a_1 u_{i2} + a_4 \dot{u}_{i2} + a_5 \ddot{u}_{i2}\right) \\ b_7 = (c_{u1} + c_{w1})\left(a_1 u_{i2} + a_4 \dot{u}_{i2} + a_5 \ddot{u}_{i2}\right) - c_{u1}\left(a_1 u_{i1} + a_4 \dot{u}_{i1} + a_5 \ddot{u}_{i1}\right) \end{gathered} \tag{16}$$

Replace Equation (15) with Equation (8) to detect the acceleration vector of the car at the driving position $\{\ddot{u}_{i+1}\}$

$$\{\ddot{u}_{i+1}\} = \begin{bmatrix} \frac{a_0}{b_2 b_1 - b_3^2}\left(b_1(b_4+b_6) + b_3\begin{pmatrix} c_{w1}f'_z(x) \\ +k_{w1}f_z(x) + b_5 + b_7 \end{pmatrix}\right) - a_0 u_{i1} - a_2 \dot{u}_{i1} - a_3 \ddot{u}_{i1} \\ \frac{a_0}{b_2 b_1 - b_3^2}\left(b_3(b_4+b_6) + b_2\begin{pmatrix} c_{w1}f'_z(x) \\ +k_{w1}f_z(x) + b_5 + b_7 \end{pmatrix}\right) - a_0 u_{i2} - a_2 \dot{u}_{i2} - a_3 \ddot{u}_{i2} \end{bmatrix} \tag{17}$$

The acceleration response of the car body is:

$$a_{v1} = \frac{a_0}{b_2 b_1 - b_3^2}\begin{pmatrix} b_1(b_4+b_6) + b_3(c_{w1}f'_z(x) \\ +k_{w1}f_z(x) + b_5 + b_7) \end{pmatrix} - a_0 u_{i1} - a_2 \dot{u}_{i1} - a_3 \ddot{u}_{i1} \tag{18}$$

When there is damage somewhere in the bridge and the bridge stiffness decreases, the deflection of the bridge at the damage location is:

$$f_{z\theta_d} = \frac{1}{1-\theta_d}\left\{\frac{F_q}{EI}\left[\begin{array}{c}\left(\frac{(x+\Delta a)^3}{6l}+\frac{(x+\Delta a)l}{3}-\frac{(x+\Delta a)^2}{2}\right)x \\ -\frac{\left(l-(x+\Delta a)x^3\right)}{6l}\end{array}\right]+\frac{F_c x^2(x-l)^2}{6l}\right\}=\frac{1}{1-\theta_d}f_z(x) \tag{19}$$

where θ_d is the degree of damage to the bridge.

The vehicle's body acceleration response at the point of injury is:

$$a_{v1\theta}=\frac{a_0}{b_2b_1-b_3^2}\left(\begin{array}{c}b_1(b_4+b_6)+b_3(\frac{1}{1-\theta_d}c_{w1}f_z'(x) \\ +\frac{1}{1-\theta_d}k_{w1}f_z(x)+b_5+b_7)\end{array}\right)-a_0u_{i1}-a_2\dot{u}_{i1}-a_3\ddot{u}_{i1} \tag{20}$$

From Equations (18) and (20), it can be seen that the acceleration response of the car body contains bridge deflection and speed response information, and the acceleration response of the car body through the bridge is processed and analyzed, and the bridge information can be identified. When a bridge is damaged, the deflection of the bridge changes accordingly at the location of the damage, and this change will be reflected in the acceleration response of the car body, so the damage location can be identified by using the acceleration response of the car body.

2.3. Wavelet Analysis

At the initial stage of structural damage or when the damage is relatively small, the changes of structural response caused by the damage are also very small. In order to highlight the damage information in vehicle response, wavelet analysis is applied in dynamic response signal analysis. As a new method in the field of signal processing, wavelet analysis can effectively identify the subtle differences in signals. It carries out multi-scale analysis on the signal in time domain and frequency domain, so that the signal can be time-refined at high frequency and frequency-refined at low frequency. Therefore, it is widely used in the field of signal detection.

Wavelet transforms are divided into continuous wavelet transforms and discrete wavelet transforms, continuous wavelet transform (CWT) can decompose the original signal into a two-dimensional signal; by adjusting the scale factor m and translation factor n, the original signal $f(s)$ is analyzed to obtain the signal local detection information. The original signal $f(s)$ is continuously transformed at the scale factor m and the translation factor n, because the adjustment factor has continuity, so the wavelet transformation coefficient will generate a large amount of redundant data, which is not conducive to signal processing, so a discrete wavelet transform (DWT) is introduced. The scale factor m and translation factor n are generally calculated as follows:

$$m = m_0^j \tag{21}$$

$$n = km_0^j n_0 \tag{22}$$

where k and j are integers and $m_0 \neq 1$.

The discrete wavelet transform function is:

$$WT_f(m,n) = |m_0|^{-\frac{j}{2}}\int_{-\infty}^{+\infty} f(s)\phi^*{}_{m,n}\left(m_0^{-j}s - kn_0\right)ds \tag{23}$$

Discrete wavelet transform effectively reduces the redundant data generated by continuous wavelet transform and improves the calculation efficiency and accuracy. In this paper, discrete wavelet transform is used to analyze the vehicle dynamic response data and accurately find the subtle changes in the data, which is conducive to bridge damage identification.

2.4. Example Verification

In order to verify the accuracy of the program, the vehicle response spectrum obtained in this paper is compared with reference [7]. When the verification program is correct in this section, the same bridge and vehicle parameters as those in the literature are selected. The schematic diagram of the model is shown in Figure 1. The span diameter of the simple branch girder bridge was 25 m; cross-sectional area A = 2 m^2; cross-sectional moment of inertia I = 0.12 m^2; bridge elastic modulus E = 2.75 × 10^{10} N/m^2; bridge unit length mass of 4800 kg/m; vehicle body mass m_v = 1200 kg; wheel stiffness k_w = 500,000 N/m; vehicle speed v = 10 m/s; other parameters of the vehicle were set to 0. The extracted acceleration response was used as a fast Fourier transform. The spectrum of acceleration response obtained is shown in in Figure 3.

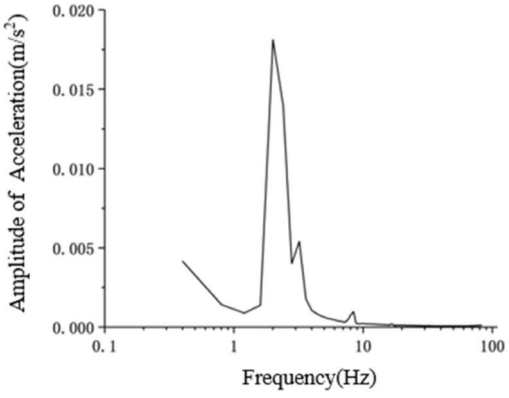


Figure 3. The response spectrum of this paper.

It can be seen from Figure 3 that the extracted bridge frequency and vehicle frequency are 2.01 Hz and 3.22 Hz, respectively, which are very close to the bridge frequency and vehicle frequency of 2.08 Hz and 3.25 Hz in the literature. Input and output are the same, which proves the correctness of the program. This shows that the dynamic performance of the bridge can be identified and determined from the vehicle response.

2.5. Influence of Road Roughness

In practical measurement, vehicle response is not only related to vehicle performance and bridge performance, but also related to road roughness. Additionally, uneven bridge deck is the main source of motivation that affects vehicle vibration, especially vertical vibration of vehicles [21], and naturally affects indirect detection methods based on vehicle response.

According to the mathematical definition, road surface irregularity is a random function, and the harmonic superposition method can simulate a stationary random process. The main idea is to express road roughness as the sum of a large number of sine and cosine functions with random phases. It has the characteristics of fast calculation speed and high accuracy. The road surface irregularity value at each position is calculated by the harmonic superposition method, and the calculation formula is as follows

$$r(x) = \sum_{k=1}^{N} \alpha_k cos(\omega_k x + \varphi_k) \tag{24}$$

$$\alpha_k = \sqrt{4G_d(n_k)\Delta n} \tag{25}$$

$$n_k = n_1 + \left(k - \frac{1}{2}\right)\Delta n \; k = 1, \; 2, \; \ldots N \tag{26}$$

$$\Delta n = \frac{n_2 - n_1}{N} \tag{27}$$

In the formula: $r(x)$ is the uneven sample waveform; ω_k is the circular frequency, with a value of $2\pi n_k$; φ_k is a random phase angle, with a value between $[0, 2\pi]$; N is a sufficiently large integer, which is an equal fraction of the spatial frequency; and $G_d(n)$ is the pavement power spectrum function.

The harmonic superposition method is used to simulate road irregularity, and MATLAB software is used to simulate the irregularity sample functions of three pavement grades A–C, as shown in Figure 4.

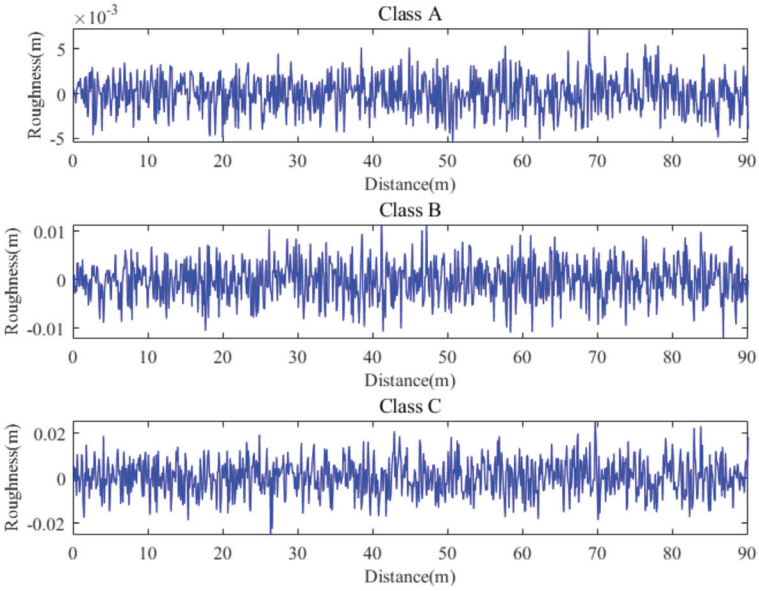


Figure 4. Three classes of road roughness.

It can be seen from Figure 4 that the pavement irregularity value increases with the decrease in pavement grade. The peak irregularity values of three pavement grades A–C are 5.78 mm, 11.49 mm, and 27.86 mm, respectively.

3. Identification of Bridge Damage Location under the Action of Mobile Vehicles

3.1. Theoretical Basis

Before the bridge is put into use and in normal operation, many experiments are carried out to observe its performance. In this chapter, combined with the vehicle–bridge coupling vibration analysis algorithm mentioned above, considering the influence of road roughness, an indirect measurement method based on health data is proposed. The vehicle response when the detection vehicle passes through the damaged bridge is compared with the vehicle response data when the bridge is in a healthy state. The acceleration responses of the two vehicles are subtracted; then, the acceleration response difference is transformed by wavelet transform, and the bridge damage location is located through the peak value of wavelet transform.

By subtracting Equations (18) and (20), the acceleration response difference in the vehicle body can be obtained:

$$\Delta a_{v\theta}(x) = \begin{cases} H(x) & Undamaged\ location \\ \frac{a_0 b_3}{b_2 b_1 - b_3^2}(c_{w1} f_z'(x) + k_{w1} f_z(x)) \cdot \frac{\theta_d}{1-\theta_d} + H(x) & Damage\ location \end{cases} \tag{28}$$

where:

$$H(x) = a_0 u_{i11} + a_2 \dot{u}_{i11} + a_3 \ddot{u}_{i11} - a_0 u_{i12} - a_2 \dot{u}_{i12} - a_3 \ddot{u}_{i12} + \frac{a_0}{b_2 b_1 - b_3^2} \begin{pmatrix} b_1(b_{42} + b_{62} - b_{41} - b_{61}) \\ +b_{52} + b_{72} - b_{51} - b_{71} \end{pmatrix} \quad (29)$$

It can be seen from Equation (28) that the acceleration response difference in the vehicle body has a large sudden change at the damage position, but the sudden change point cannot be identified on the acceleration response difference curve of the vehicle body. The damage position of the continuous beam bridge can be identified by discrete wavelet transform of the acceleration response difference in the vehicle body.

3.2. Model Establishment and Simplification

Taking the three-span continuous beam bridge with equal sections as an example, the total length of the bridge is 90 m, the span combination is 30 m + 30 m + 30 m, the bridge section height is 0.5 m, the bridge width is 3 m, the bridge elastic modulus E = 3×10^{10} N/m^2, and the mass per unit length of the bridge is 3300 kg/m. The beam3 element is used to simulate the bridge element, the division accuracy is 0.1 m, and the bridge damping adopts Rayleigh damping. The vehicle model adopts the tractor test vehicle model in the literature [20]; the acceleration sensor is placed at the centroid of the detection vehicle body, and there is no torque transmission between the two parts connecting the tractor and the detection vehicle. The tractor is only used as a provider of forward power of the detection vehicle, therefore there is no torque transmission between the tractor and the detection vehicle, so the tractor is simplified as a concentrated force acting on the center of mass of the tractor, as shown in Figure 5. The vehicle model is symmetrical left and right, so the detection vehicle model uses a 1/2 model for simplified calculation, and the tractor concentration force is 5880 N.

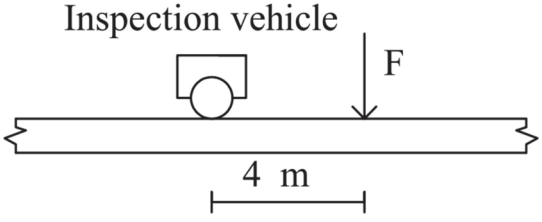


Figure 5. Simplified schematic diagram of tractor-testing vehicle (unit: m).

Vehicle parameters are set according to the actual vehicle parameters and documents [22,23], as shown in Table 1.

Table 1. Vehicle parameters of the test vehicle.

Properties	Symbol	Unit	Value
Body mass	m_v	kg	2000
Car suspension and wheel mass	m_{wi} (i = 1.2)	kg	100
Stiffness coefficient of vehicle suspension	k_{ui} (i = 1.2)	N/m	4.07×10^5
Vehicle suspension damping coefficient	c_{ui} (i = 1.2)	N·s/m	7509
Wheel stiffness coefficient	k_{wi} (i = 1.2)	N/m	1.56×10^5
Wheel damping coefficient	c_{wi} (i = 1.2)	N·s/m	1000
The lateral distance between the wheel and the center of mass of the car body	d_i (i = 1.2)	m	1

3.3. Bridge Damage Conditions

In this paper, the bridge damage unit is simulated by using the stiffness reduction in the bridge damage area, and the damage degree is expressed by θ_d. Six working conditions

are preliminarily assumed, and the damage position x is the distance between the damage center point and the left end of the continuous girder bridge, as shown in the bridge damage condition Table 2.

Table 2. Bridge damage conditions.

Working Condition Description	Serial Number	Damage Distance (x/m)	Degree of Damage (θ_d/%)
No damage	Working condition 1	\	0
Single damage	Working condition 2	5	20
	Working condition 3	20	30
	Working condition 4	45	20
Multiple damage	Working condition 5	20	30
		45	20
	Working condition 6	15	10
		45	20
		75	40

3.4. Identification of the Damage Location

For some new bridges, there are generally baseline data of the bridge's intact condition. The bridge damage detection method based on bridge health data mainly collects the vertical acceleration response of the car body before and after the damage, and performs discrete wavelet transformation on the acceleration response difference in the car body before and after the bridge damage to identify the damage location. A simplified model is used for calculation in this section. The vehicle driving speed is 72 km/h, the road surface is uneven using Class A pavement, and the above damage conditions are simulated and calculated by using finite element software.

The wavelet basis function selects the bior4.4 wavelet and db20 wavelet for discrete wavelet transformation, and the d1 coefficient in the bior4.4 wavelet transform and the d5 coefficient in the db20 wavelet is selected to determine the damage position, and the damage position recognition of the working conditions 2–6 is shown in Figures 6 and 7.

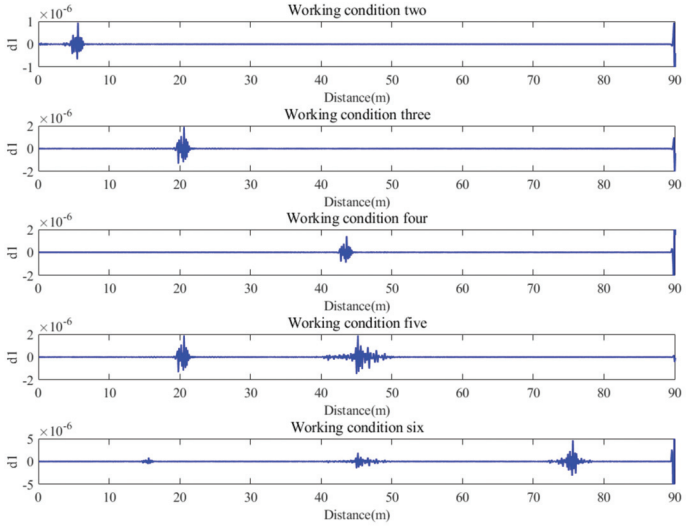


Figure 6. Wavelet transform coefficient (bior4.4–d1) diagram.

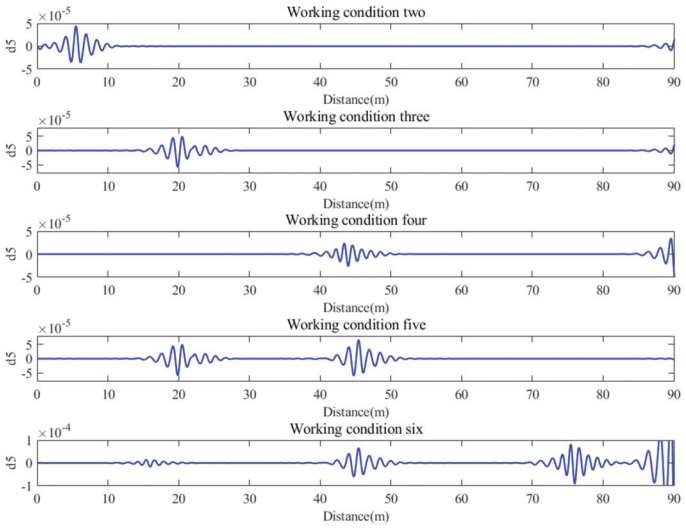


Figure 7. Wavelet transform coefficient (db20–d5) diagram.

It can be seen from the wavelet variation coefficients in Figures 8 and 9 that both wavelet coefficients have obvious sudden changes near the damage location. However, there is some interference near the damage location, which is not conducive to the accurate identification of the damage location and is easy to cause misjudgment. However, through the maximum successful approximation approach (MSAA) processing on the wavelet transform coefficients, the damage location can be accurately identified and misjudgment can be reduced. Figures 8 and 9 shows the wavelet coefficients after the maximum successive approximation processing.

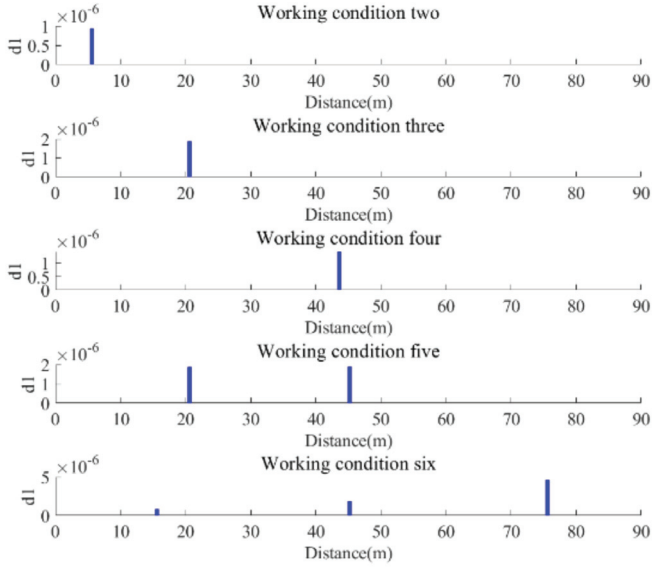


Figure 8. Bior4.4–d1 coefficient after MASS treatment.

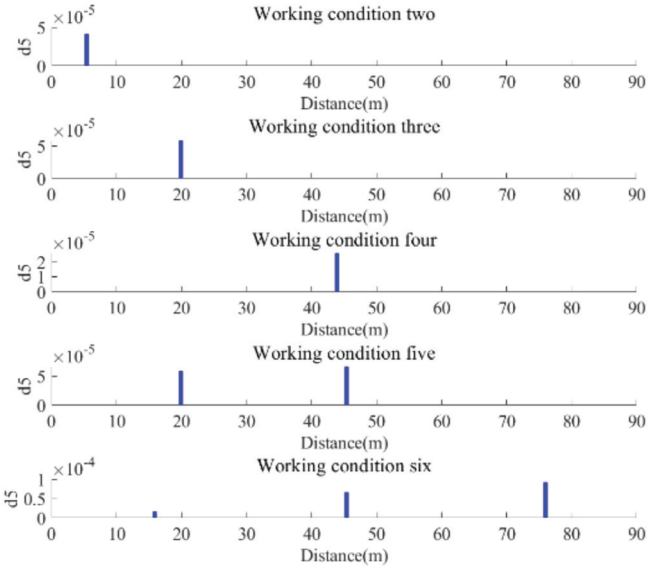


Figure 9. db20–d5 coefficient after MASS treatment.

It can be seen from Figures 8 and 9 that the wavelet transform coefficient of the acceleration response difference in the detection vehicle has a large mutation at the damage position, and the damage position can be determined. Using MSAA method to process wavelet coefficients can accurately locate the damage location and reduce misjudgment. Under the finite element simulation calculation, the bridge damage detection method based on health data can accurately locate the single damage and multiple damage of a continuous beam bridge. In condition 6, it can be seen that the peak value of wavelet transform coefficient increases with the increase in damage degree.

4. Influence Factors on Damage Location Identification

Bridge damage identification based on vehicle response involves vehicle parameters, bridge parameters, noise, and other factors, which will have a certain impact on bridge damage identification. Therefore, studying the influence of various factors on bridge damage identification will help to accurately locate bridge damage and select parameters. The following mainly studies the influence of vehicle weight, vehicle speed, road roughness, noise, and other factors on the damage location identification of a continuous beam bridge.

4.1. Influence of Vehicle Weight on Damage Location Identification

The weight of the inspection vehicle is an important parameter affecting vehicle response and bridge response. Considering the actual vehicle situation, the weight of the inspection vehicle was calculated by taking 1000 kg, 1500 kg, 2000 kg, and 2500 kg, respectively. Other vehicle parameters and bridge parameters were set according to Section 3.2. In order to study the influence of vehicle weight on damage recognition, we selected condition 3 in Section 3.3 (single damage; damage location: 20 m; damage degree: 30%) for calculation, and conducted discrete wavelet transform on the acceleration response difference in the vehicle body for bridge damage identification. The wavelet transform coefficients under different vehicle weights are shown in Figure 10.

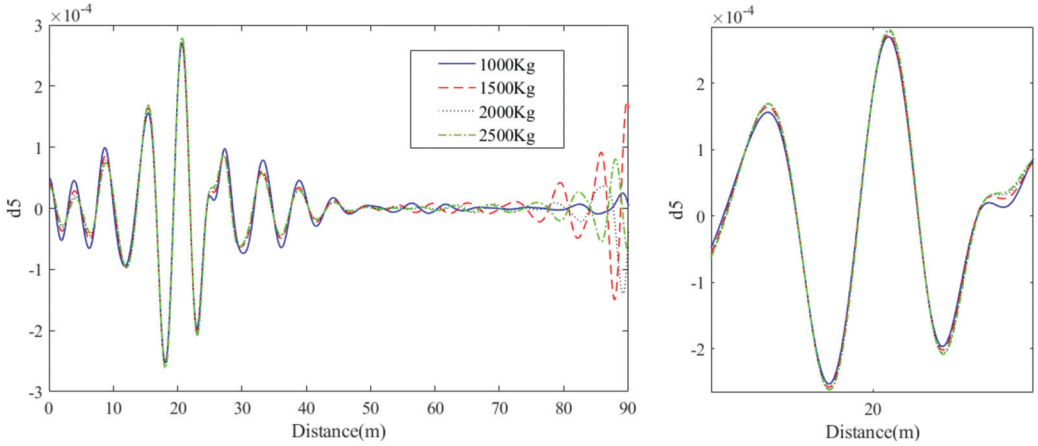


Figure 10. Wavelet transform coefficients and partial enlarged diagrams under different vehicle weights.

As can be seen from Figure 10, with the increase in vehicle weight, the bridge damage can still be accurately identified by using the wavelet transform coefficient of vehicle acceleration difference, but the coefficient peaks of different vehicle weights are slightly different at the damage location. At the same time, when the vehicle weight exceeds 1000 kg, there is more interference in the wavelet transform coefficient near 90 m, which is not conducive to the location of bridge damage. According to the finite element simulation calculation results, the vehicle weight of the test vehicle shall be 1000 kg as far as possible.

4.2. Influence of Vehicle Speed on Damage Location Identification

The change of the vehicle excitation frequency will affect the dynamic response of the vehicle–bridge coupling system. Naturally, it will have a certain impact on the bridge damage identification. When the vehicle speed is within a specific range, the vehicle acceleration signal changes drastically, and the wavelet transform coefficient peak value is obvious, which is beneficial to the location of bridge damage. The inspection vehicle weighed 1000 kg. The parameters of other vehicles and bridges remain unchanged, and the vehicle speeds were calculated as 10 m/s, 20 m/s, and 30 m/s, respectively, to study the influence of the speed change of the detected vehicle on the bridge damage identification. Working condition 3 in Section 3.3 (single damage; damage location: 20 m; damage degree 30%) was selected for calculation, and discrete wavelet transform was performed on the difference in acceleration response of the vehicle body to perform bridge damage identification, as shown in Figure 11.

When the vehicle speed is 10 m/s, 20 m/s, and 30 m/s, the damage can be accurately identified, and as the vehicle speed increases, the wavelet transform peak value first decreases and then increases. In theory, the test vehicle can travel at a speed of 10–30 m/s. However, if the sampling rate of the sensor is considered, it is required to collect enough sampling signals in a single test run; therefore, it is suggested that the speed should not be too high; when the speed is at 20 m/s and 30 m/s, the wavelet coefficients near the support will affect the damage identification. In addition, if the speed is too low it will lead to insufficient bridge excitation. Considering detection speed and detection effect, a vehicle speed of 10 m/s is recommended in this article.

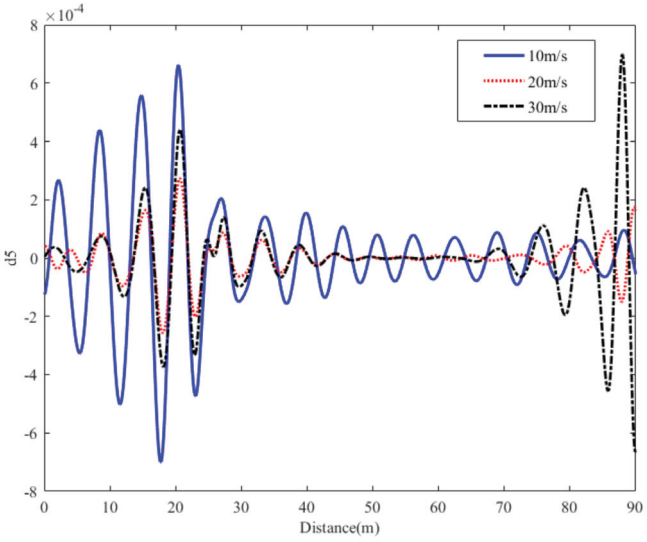


Figure 11. Wavelet transform coefficients at different vehicle speeds.

4.3. Influence of Road Surface Roughness on Damage Identification

During on-site inspection, the uneven road surface will inevitably have a greater impact on the response of the vehicle, thereby affecting the identification of the bridge damage's location. Using random simulation in Section 2.5 to generate grade A–C road roughness class, the impact of road irregularities on bridge damage identification was studied. The working condition 3 in Section 3.3 (single damage; damage location: 20 m; damage degree 30%) was selected for calculation, and discrete wavelet transform was performed on the difference in acceleration response of the vehicle body to perform bridge damage identification, as shown in Figure 12.

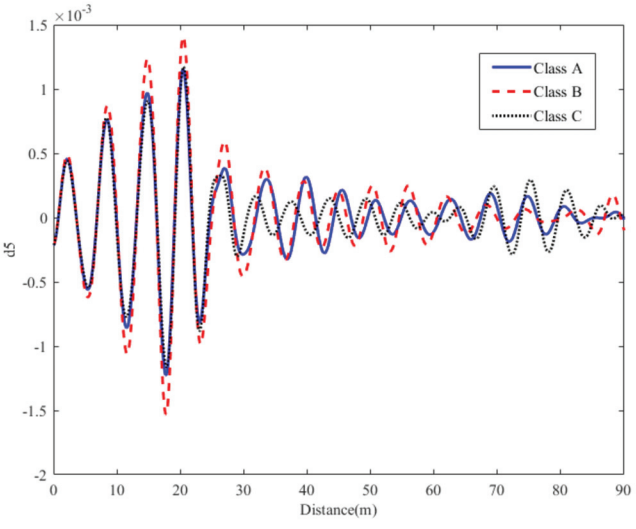


Figure 12. Wavelet transform coefficients under different road smoothness grades.

Analyzing Figure 12, it can be obtained that the damage location of bridge can be accurately identified under different road surface roughness, which shows that the road surface roughness has little influence on the bridge damage identification method. Under ideal road conditions, the use of acceleration response difference can effectively reduce the impact of roughness on the damage results, but it cannot be completely eliminated.

4.4. Influence of Noise on Damage Identification

During the on-site measurement and inspection process, due to the influence of the measurement environment and the measurement system, noise is inevitable, and the noise will "contaminate" the measurement data, which will have a greater impact on the later data processing and analysis. In order to study the effect of noise on the damage recognition effect of the method, white noise was superimposed on the acceleration response of the vehicle body, and the expression is as follows:

$$a_n = a + E_n N_{oise} \sigma(a) \tag{30}$$

where a_n is the acceleration response with noise, a is the original acceleration response, E_n is the noise level, N_{oise} is the noise obeying the standard normal distribution N (0, 1), and $\sigma(a)$ is the standard deviation of the original acceleration response.

The noise level was 3%, 5%, and 7% in sequence, and the above Section 3.3 working condition 3 (single damage; damage location: 20 m; damage degree 30%) was selected as the damage condition, and the damage identification of the bridge by noise was studied. The influence of this is shown in Figures 13 and 14.

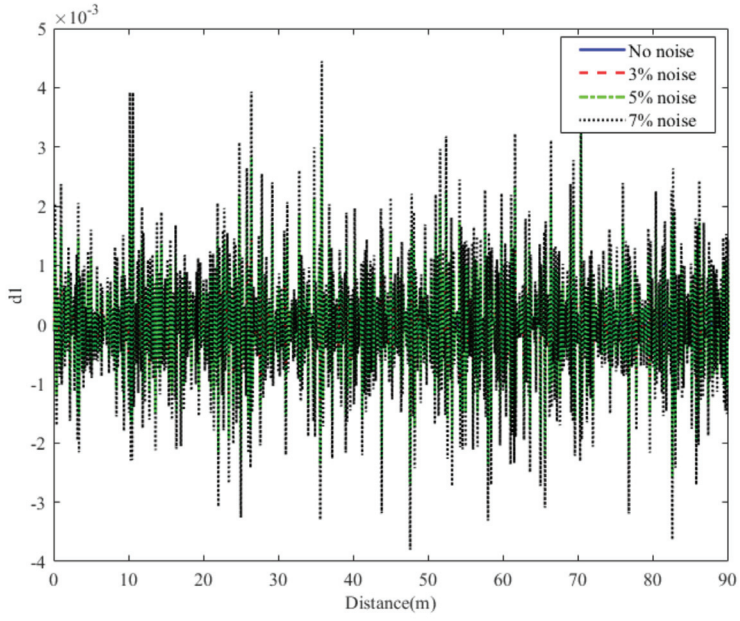


Figure 13. Wavelet transform coefficient (d1) under different noise levels.

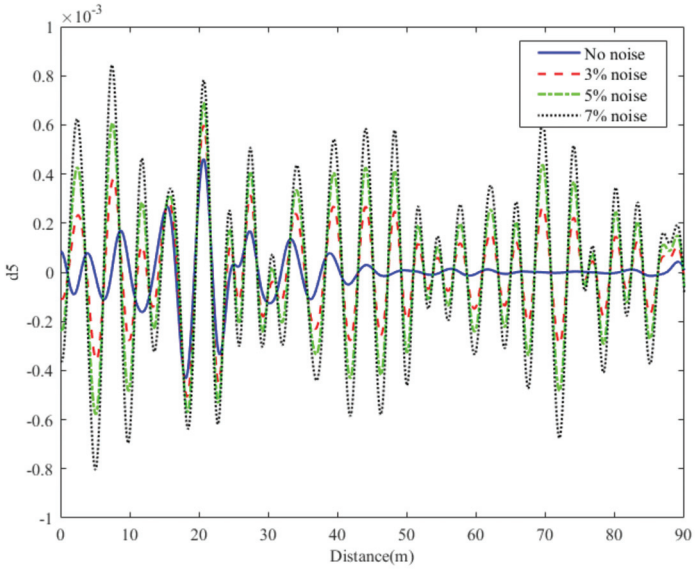


Figure 14. Wavelet transform coefficient (d5) under different noise levels.

When the sensor records the acceleration of the vehicle body, it will inevitably be affected by environmental noise. The frequency and energy of these noises will affect the wavelet transform, which may cause the results of the wavelet transform to be irregular, and the damage location cannot be identified. It can be seen from Figure 14 that when selecting the low-frequency area (d5) for damage identification, the wavelet transform coefficients under different levels of noise are basically the same as those under noise-free conditions, indicating that selecting the d5 coefficient for damage location identification has good anti-noise qualities. It can be seen from Figure 13 that when the high-frequency area (d1) is selected for damage identification, it cannot be identified at all in the presence of noise, and the amplitude of the wavelet coefficient increases with the increase in the noise level. This is mainly because the high-frequency energy of noise is higher than the low-frequency energy, which affects the damage identification of wavelet transform coefficients in the high-frequency region. However, the original signal can be denoised. The original signal is first decomposed in multiple layers, high-frequency coefficients are thresholded, and low-frequency coefficients are retained. Then, the signal is reconstructed through wavelet inverse transform to reduce the influence of noise on damage identification.

5. Judgment of Damage Degree of Continuous Beam Bridge

In bridge damage identification, it is not only necessary to identify the location of the bridge damage, but also to determine the degree of damage. This section will be based on the bridge deflection and deflection difference to judge the damage degree of the bridge.

5.1. Damage Index

The MSAA coefficient is used as the coefficient peak value of the wavelet transform coefficient at the damage position. By processing the MSAA coefficient, the damage degree index can be obtained. Compared with the low frequency coefficient (d5), the high frequency coefficient (d1) has more concentrated energy, and the MSAA coefficient peak value changes more regularly. Therefore, the bridge detection method based on bridge health data selects the bior4.4-d1 coefficient to determine the degree of damage.

The MSAA coefficient is the peak coefficient of the damage location obtained after the discrete wavelet transformation of the vehicle body acceleration response difference.

Therefore, the MSAA coefficient is closely related to the vehicle body acceleration response difference. It can be seen from Equation (28) that the sudden change point of the vehicle body acceleration response difference at the damage location is mainly related to the healthy bridge deflection $f_z(x)$, so the MSAA coefficient at the damage location is divided by the healthy bridge deflection $f_z(x_{\theta_d})$ at the same location. The D_1 coefficient of continuous beam bridge damage detection method based on bridge health data is:

$$D_1 = \frac{w_m}{f_z(x_{\theta_d})} \tag{31}$$

$$f_z(x_{\theta_d}) = \frac{F_q}{EI} \cdot \left[\begin{array}{c} -\frac{\left(l-\left(x_{\theta_d}+\Delta a\right)\right){x_{\theta_d}}^3}{6l} \\ +\left(\frac{(x_{\theta_d}+\Delta a)^3}{6l} + \frac{\left(x_{\theta_d}+\Delta a\right)l}{3} - \frac{(x_{\theta_d}+\Delta a)^2}{2}\right)x_{\theta_d} \end{array} \right] + \frac{F_c {x_{\theta_d}}^2 (x_{\theta_d}-l)^2}{3EIl} \tag{32}$$

where w_m is the MSAA coefficient at the damage location, $f_z(x_{\theta_d})$ is the healthy bridge deflection value at the damage location, and x_{θ_d} is the distance from the damage location to the left end of the span.

The MSAA coefficients of the same degree of damage vary greatly in different positions. The change trend in MSAA coefficient is similar to that of healthy bridge deflection. Under the same damage degree, the MASS coefficient of the mid span is the largest, and it gradually decreases from the middle of the span to the support, approximately in the shape of a parabola. Therefore, damage location adjustment factor D_2 should be introduced, so that the damage degree index of each position of the same damage degree is kept within a certain range as in Formula (33)

$$D_2 = {x_d}^{0.5} \tag{33}$$

where x_d is the distance between the damage location and the nearest support of the span, x_{e1} is the distance between the damage location and the left end of the span, x_{e2} is the distance between the damage location of the side span and the nearest bridge end, and l_b is the length of the side span.

From Formulas (27) and (29), the damage degree index D of the continuous beam bridge damage detection method based on bridge health data is:

$$D = D_1 \cdot D_2 = \frac{w_m}{f_z(x_{\theta_d})} \cdot {x_d}^{0.5} \tag{34}$$

The MSAA coefficient of the same damage position increases as the damage degree increases, and the change trend is in the form of a power function. Multiple damage degree indicators at a certain point of the bridge are used to fit the damage degree function $S(x)$, and the damage degree function $S_{d1}(x)$. Multiple damage degree indexes at 20 m from the damage position are selected for fitting, and the damage degree is 10%, 20%, 30%, 40%, 50%, and Matlab is selected for function fitting. Substituting the damage degree index D into the damage degree function $S(x)$, the damage degree function $S_{d1}(D)$ is expressed as:

$$S_{d1}(D) = 10.49 \times D^{0.4519} - 0.1707 \tag{35}$$

where D is an index of the degree of damage.

5.2. Judgment of Bridge Damage Degree Based on Health Data

Corresponding data processing was performed on the MSAA coefficients of each bridge damage condition in Table 2 to obtain the damage degree index D. Substituting the damage degree index D into the damage degree function $S_{d1}(D)$, the bridge damage degree can be determined, as shown in Table 3.

Table 3. Judgment of Bridge Damage Degree.

Damage Conditions	Damage Position (x/m)	Degree of Damage (θ_d/%)	MSAA Coefficient	Damage Index D	Degree of Damage ($S_{d1}(D)$/%)	Judgment Error (R_d/%)
Working condition 2	5	20	9.44×10^{-7}	4.57×10^{-4}	18.17	−1.83
Working condition 3	20	30	1.89×10^{-6}	9.06×10^{-4}	29.95	−0.05
Working condition 4	43	20	1.42×10^{-6}	4.86×10^{-4}	19.08	−0.92
Working condition 5	20	30	1.80×10^{-6}	9.08×10^{-4}	29.99	−0.01
	45	20	1.89×10^{-6}	4.90×10^{-4}	19.21	−0.79
Working condition 6	15	10	8.55×10^{-7}	2.56×10^{-4}	10.71	0.71
	45	20	1.90×10^{-6}	4.85×10^{-4}	19.06	−0.94
	75	40	4.64×10^{-6}	1.43×10^{-3}	40.07	0.07

It can be seen from Table 3 that the above-mentioned data processing method can be used to determine the degree of damage at each position of the bridge, with a maximum error of 2.81%, which is relatively small. The bridge damage degree determination is based on the MSAA coefficient through data fitting, and the MSAA coefficient is derived from the wavelet transform coefficient, so the accuracy of the damage degree determination is also affected by environmental noise, sensor accuracy, and road surface roughness.

6. Conclusions

(1) Using the maximum value successive approximation method to process the wavelet transform coefficient can accurately identify the damage location, and the identification accuracy is affected by vehicle speed, vehicle weight, road surface roughness, and noise. The recommended speed is 10 m/s, and the recommended vehicle weight is 1000 kg.

(2) This damage detection method has good noise immunity. The vehicle response signal needs to be processed by wavelet noise reduction when the noise is greater. When the noise level is not higher than 5%, the location of the bridge damage can be identified successfully.

(3) The corresponding damage degree index D and the damage degree function $S_{d1}(D)$ are put forward. According to this, the damage degree of the bridge can be judged, and the judgment error of the damage degree basically meets the requirements, which is suitable for continuous beam bridges with different span combinations.

Author Contributions: Methodology, Z.S. and K.L.; software, K.L.; writing—original draft preparation, Z.S. and K.L.; writing—review and editing, Z.S. and K.L.; supervision, H.Q. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China(NSFC), grant No. 50878198.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank Zhengzhou University for supporting the lab, and Yirui Tang for suggesting improvements to the method in this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Chen, Q. Wuyishan Gongguan Bridge with a "Tragic Fate". *China Safety Production News N*, 19 July 2011.
2. Zadehmohamad, M.; Bazaz, J.B.; Riahipour, R.; Farhangi, V. Physical modeling of the long-term behavior of integral abutment bridge backfill reinforced with tire-rubber. *Int. J. Geo-Eng.* **2021**, *12*, 36. [CrossRef]
3. Bandara, R.P.; Chan, T.; Thambiratnam, D. Structural damage detection method using frequency response functions. *Struct. Health Monit.* **2014**, *13*, 418–429. [CrossRef]
4. Zhan, J.; Xia, H.; Zhang, N. Bridge damage identification method based on online vibration response. *China Railw. Sci.* **2011**, *32*, 58–62.

5. Zhao, H. Performance Monitoring and Evaluation of Long-Span Cable-Stayed Bridges during Operation Period. Ph.D. Thesis, Southwest Jiaotong University, Chengdu, China, 2015.
6. Yang, Y.-B.; Lin, C.; Yau, J. Extracting bridge frequencies from the dynamic response of a passing vehicle. *J. Sound Vib.* **2004**, *272*, 471–493. [CrossRef]
7. Yang, Y.-B.; Lin, C. Vehicle–bridge interaction dynamics and potential applications. *J. Sound Vib.* **2005**, *284*, 205–226. [CrossRef]
8. Bu, J.Q.; Law, S.S.; Zhu, X.Q. Innovative Bridge Condition Assessment from Dynamic Response of a Passing Vehicle. *J. Eng. Mech.* **2006**, *132*, 1372–1379. [CrossRef]
9. Nguyen, K.V.; Tran, H.T. Multi-cracks detection of a beam-like structure based on the on-vehicle vibration signal and wavelet analysis. *J. Sound Vib.* **2010**, *329*, 4455–4465. [CrossRef]
10. Sun, Z.; Zhang, Y.; Tong, S. Summary of Bridge Damage Detection Methods Based on Vehicle-Bridge Coupling Vibration. *World Earthq. Eng.* **2013**, *29*, 1–8.
11. Yang, Y.-B.; Li, Y.; Chang, K. Constructing the mode shapes of a bridge from a passing vehicle: A theoretical study. *Smart Struct. Syst.* **2014**, *13*, 797–819. [CrossRef]
12. Malekjafarian, A.; Obrien, E. Identification of bridge mode shapes using Short Time Frequency Domain Decomposition of the responses measured in a passing vehicle. *Eng. Struct.* **2014**, *81*, 386–397. [CrossRef]
13. Obrien, E.J.; Malekjafarian, A.; González, A. Application of empirical mode decomposition to drive-by bridge damage detection. *Eur. J. Mech. A/Solids* **2017**, *61*, 151–163. [CrossRef]
14. Qian, Y. Bridge Frequency Identification and Damage Detection Based on Indirect Measurement Method. Master's Thesis, Chongqing University, Chongqing, China, 2018.
15. Marashi, S.M.; Pashaei, M.H.; Khatibi, M.M. Estimating the Mode Shapes of a Bridge Using Short Time Transmissibility Measurement from a Passing Vehicle. *Appl. Comput. Mech.* **2019**, *5*, 735–748.
16. Locke, W.; Sybrand, J.; Redmond, L.; Safro, I.; Atamturktur, S. Using Drive-by health monitoring to detect bridge damage considering environmental and operational effects. *J. Sound Vib.* **2020**, *468*, 115088. [CrossRef]
17. Clemente, P.; Bongiovanni, G.; Buffarini, G.; Saitta, F. Structural health status assessment of a cable-stayed bridge by means of experimental vibration analysi. *J. Civ. Struct. Health Monit.* **2019**, *9*, 655–669. [CrossRef]
18. Farhangi, V.; Karakouzian, M. Design of Bridge Foundations Using Reinforced Micropiles. In Proceedings of the International Road Federation Global R2T Conference & Expo, Las Vegas, NV, USA, 19–22 November 2019.
19. Zhao, L. Vehicle-Bridge Coupling Vibration Analysis of Long-span Highway Concrete Filled Steel Tube Arch Bridge. Master's Thesis, Zhengzhou University, Zhengzhou, China, 2017.
20. Lin, C.; Yang, Y.-B. Use of a passing vehicle to scan the fundamental bridge frequencies: An experimental verification. *Eng. Struct.* **2005**, *27*, 1865–1878. [CrossRef]
21. Han, W.; Ma, L.; Yuan, S.; Zhao, S. Analysis of the influence of non-uniform excitation of road surface roughness on the response of vehicle-bridge coupled vibration system. *China Civ. Eng. J.* **2011**, *44*, 81–90.
22. Guo, X.; Peng, M.; Zou, J.; Zhang, J.; Zhang, H. Research on Influencing Factors of Energy Feedback Potential of Commercial Vehicle Suspension. *Chin. J. Highw. Transp.* **2016**, *29*, 151–158.
23. Yao, C. Analysis and Research on Road Vehicle-Bridge Coupling Vibration Considering the Influence of Bridge Surface Irregularities. Master's Thesis, Central South University, Changsha, China, 2011.

applied sciences

MDPI

Article

Research on Damage Localization of Steel Truss–Concrete Composite Beam Based on Digital Orthoimage

Rui Luo [1], Zhixiang Zhou [2,*], Xi Chu [2], Xiaoliang Liao [2] and Junhao Meng [1]

[1] State Key Laboratory of Mountain Bridge and Tunnel Engineering, Chongqing Jiaotong University, Chongqing 400074, China; luorui_cq@163.com (R.L.); mengjunhao19980806@163.com (J.M.)
[2] College of Civil and Transportation Engineering, Shenzhen University, Shenzhen 518060, China; chuxi@szu.edu.cn (X.C.); liaoxiaoliang2020@email.szu.edu.cn (X.L.)
* Correspondence: zhixiangzhou@szu.edu.cn

Featured Application: Holographic deformation data of a bridge structure can be obtained from the image data, and areas with excessive structural stiffness loss can be monitored to provide key monitoring locations for routine bridge inspections.

Abstract: Most structural health monitoring is carried out for a limited number of key measurement points of a bridge, and incomplete measurement data lead to incomplete mechanical equation inversion results, which is a key problem faced in bridge damage identification. The ability of digital images to holographically describe structural morphology can effectively alleviate the problem of damage identification due to incomplete test data. Based on digital image processing technology, a matrix similarity damage identification method based on a structural digital orthoimage was proposed. Firstly, a steel truss–concrete composite beam specimen with a complex support bar system was designed and fabricated in the laboratory, and the digital orthoimage of the test beam was obtained by the perspective transformation of the original image of the test beam. The body contour of the structure was extracted from the digital orthoimage of the test beam, and wavelet threshold denoising was performed on the lower edge profile to obtain the deflection curves of the structure under different working conditions. The verification results show that the maximum error of the deflection curve is 3.42%, which proves that the digital orthoimage can accurately and completely reflect the deformation of the structure. Finally, based on the digital orthophoto of the test beam, a matrix similarity test before and after the damage was carried out, and the results show that the singularities of the similarity distribution are consistent with the location of the damage; furthermore, the accurate positioning of the damage in different working conditions is achieved.

Keywords: bridge structures; damage identification; digital orthoimage

Citation: Luo, L.; Zhou, Z.; Chu, X.; Liao, X.; Meng, J. Research on Damage Localization of Steel Truss–concrete Composite Beam Based on Digital Orthoimage. *Appl. Sci.* **2022**, *12*, 3883. https://doi.org/10.3390/app12083883

Academic Editors: Liang Yu, Phong B. Dao and Lei Qiu

Received: 25 January 2022
Accepted: 1 April 2022
Published: 12 April 2022

1. Introduction

With the rapid development of Computer Science and Technology, the difficulty of identifying structural damage caused by incomplete test data can be effectively alleviated by the ability of images to completely describe the structure morphology. Non-contact data collection through digital images has the advantages of being holographic, convenient, and economical compared to traditional point sensor-based monitoring methods. In recent years, with the development of hardware conditions such as pre-installed camera, UAVs, and wearable virtual reality equipment, image archives of structural damage can be created based on accumulated monitoring data and previous inspection results. Therefore, vision-based damage detection and identification techniques are more easily applied to real structures, offering the possibility of identifying structural damage from images. Computer vision and image processing techniques have been widely used for damage identification on image datasets such as concrete cracks [1], concrete spalling [2], pavement cracks [3],

underground concrete pipe cracks [4], and asphalt pavement potholes [5]. For the detection of local cracks in structures, image processing techniques meet the need for intuitive and fast detection [6,7]. Prasanna et al. [8] conducted a study on the automated detection of cracks in concrete bridges. The images are first filtered and smoothed and noise is removed using increasing structural elements with alternating opening and closing; then, the edges of the bridge cracks are accurately extracted using a multi-scale morphological edge detector, and the development of the cracks is tracked and localized. Sarvestani et al. [9] have developed a vision-based image acquisition robot and subsequently proposed a more advanced automatic vision monitoring system. The system uses a vision-based remote-controlled robot for image acquisition and digital image processing software to identify the crack size in the captured images, making the inspection process faster, safer, more reliable, and less costly. Dyke et al. [10] proposed a vision-based bridge crack detection technique by automatically processing target detection and grouping. Yang and Nagarajaiah et al. [11,12] combine a low-rank approach with a sparse representation to detect local structural damage in real time using video. The monitoring of various complex, hidden, and high-altitude parts of bridges has been achieved by using intelligent robots and drones instead of the traditional manual safety monitoring methods [13]. The use of camera-equipped unmanned aerial vehicles (UAVs) for bridge safety condition monitoring is growing exponentially [14]. At present, with the improvements of monitoring equipment, lightweight, miniaturized, and accurate monitoring devices of various types can be mounted on UAVs. Therefore, new intelligent bridge monitoring technology based on the UAV monitoring platform has become a hot field of current research and technical application. Xu et al. [15] developed a novel system framework for bridge inspection and management: i.e., images are collected by camera-mounted UAS, inspection data are collected and processed based on computer vision algorithms, and a bridge information model (BrIM) is used to store and manage all relevant information. Morgenthal et al. [16] use camera-equipped UAVs to collect high-definition image data of bridge structures; the flight paths are automatically calculated by 3D models, and the intelligent safety assessment of large bridges is achieved by using the machine learning-based identification of typical damage patterns. Zhong et al. [17] use a UAV and three-points laser rangefinder to collect images of bridge structures and construct a training model of intelligent crack morphology extraction based on the Support Vector Machine (SVM) to realize the intelligent recognition of bridge crack width. Liang et al. [18] designed a bridge monitoring scheme using an unmanned aircraft with a high-definition gimbal camera to collect images of bridge cables in an intensive batch according to the structural characteristics of the bridge and the distribution form of the cables. The effective information is extracted through image processing, and the health condition of the bridge cables is evaluated comprehensively according to the relevant specifications. Lin et al. [19] designed an automatic bridge crack detection system combined with a real-time integrated image processing method, which can be assembled on an unmanned aircraft for real-time data acquisition and processing and which can effectively detect bridge cracks with higher accuracy and speed compared with other detection methods.

All of these methods use classical image processing algorithms for the damage identification of structural cracked areas, and the idea is to directly process the underlying pixels and local regions in the image and then output the target region of interest. These classical treatments have certain limitations, such as the need to manually design filters in advance to detect damage and the need to make assumptions about the crack geometry [20–22], which leads to a serious dependence of the damage identification results on manually hand-picked parameters. If the understanding of cracks and the establishment of crack models are not in place, the extraction of cracks will be invalid, and damage identification will not be possible. The most important point is that the biggest problem of image processing techniques applied in structural damage identification is that the technique requires a priori knowledge of the cracking position. In other words, the location of the damage cracking on the surface of the structure needs to be known in advance in order to track the development of cracks, meaning that the first key issue of damage identification—the

location of the damage—is difficult to discriminate by using existing image processing techniques. Under the condition that the location of the damage is not known in advance, the existing image processing methods are limited by the monitoring resolution, which makes it difficult to locate the early cracking of the structure. Therefore, the research and application of image processing methods in structural damage identification are currently focused on the algorithmic capability of the images, but there is no research on the correlation between the deep information changes in the images due to structural damage and the structural mechanical behavior. In particular, structural types with many truss rods, such as steel truss–concrete composite beams, are prone to damage and require an efficient monitoring method.

Any structure can be considered as a mechanical system consisting of stiffness, mass, and a damping matrix. Once structural damage occurs, the structural parameters are changed, resulting in a change in the response of the system. Therefore, changes in the morphological characteristics of the structure can be considered as a sign of the occurrence of early structural damage. The rapid rise of computer image processing technology in recent years has provided the technical support to solve the defects of incomplete measurement data in the parameter damage identification method. By taking the advantages of parametric damage recognition theory and digital image processing technology, respectively, and correlating digital images with the mechanical behavior of structures, a new method of structural damage recognition based on the holographic morphological monitoring of bridges can be formed through cross research.

2. Structural Damage Identification Method Based on Digital Image Processing

2.1. Damage Recognition Principle

The structure consists of several spatial units, and the characteristics of each unit can be reflected by the spatial stiffness matrix. Structural damage is essentially a change in the local stiffness of the structure: i.e., a change in the substructure stiffness matrix, and a change in the substructure position of the stiffness matrix, which can be interpreted as the appearance of bridge damage. The rod end force vector of an element can be expressed as

$$F^{(e)} = \left\{ \begin{array}{cccccccccccc} F_{Ni} & F_{Qij} & F_{Qiz} & M_{ix} & M_{ij} & M_{iz} & F_{Nj} & F_{Qjy} & F_{Qjz} & M_{jx} & M_{jy} & M_{jz} \end{array} \right\}^T \quad (1)$$

The rod end displacement vector of an element can be expressed as

$$\delta^{(e)} = \left\{ \begin{array}{cccccccccccc} u_i & v_i & w_i & \theta_{ix} & \theta_{iy} & \theta_i & u_j & v_j & w_j & \theta_{jx} & \theta_{jy} & \theta_{jz} \end{array} \right\}^T \quad (2)$$

The stiffness equation of an element can be expressed as

$$F^{(e)} = K^{(e)}\delta^{(e)} \quad (3)$$

Several elements can be superimposed together to form a bridge stiffness matrix equation, as shown in Equation (4).

$$\begin{bmatrix} k_{11} & k_{12} & k_{13} & \cdots & k_{1,n-2} & k_{1,n-1} & k_{1,n} \\ k_{21} & k_{22} & k_{23} & \cdots & k_{2,n-2} & k_{2,n-1} & k_{2,n} \\ k_{31} & k_{32} & k_{33} & \cdots & k_{2,n-2} & k_{3,n-1} & k_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ k_{n-2,1} & k_{n-2,2} & k_{n-2,3} & \cdots & k_{n-2,n-2} & k_{n-2,n-1} & k_{n-2,n} \\ k_{n-1,1} & k_{n-1,2} & k_{n-1,3} & \cdots & k_{n-1,n-2} & k_{n-1,n-1} & k_{n-1,n} \\ k_{n,1} & k_{n,2} & k_{n,3} & \cdots & k_{n,n-2} & k_{n,n-1} & k_{n,n} \end{bmatrix} \begin{Bmatrix} u_1 \\ v_1 \\ w_1 \\ \vdots \\ u_n \\ v_n \\ w_n \end{Bmatrix} = \begin{Bmatrix} X_1 \\ Y_1 \\ Z_1 \\ \vdots \\ X_n \\ Y_n \\ Z_n \end{Bmatrix} \quad (4)$$

The bridge structural system necessarily follows the mechanical matrix equation as

$$\{d\} = [K]^{-1}\{F\} \quad (5)$$

Equation (5) shows that there is an inevitable intrinsic connection between the displacement state $\{d\}$, which reflects the deformation characteristics of the bridge, and the

stiffness matrix $[K]$, which characterizes the safety state of the bridge structure, and the various load effects $\{F\}$ acting on the bridge structure. If essential damage occurs in a part of the bridge structure, the stiffness of the corresponding member or linkage in the stiffness matrix $[K]$ is degraded, and there is a change in the displacement state $\{d\}$ of the bridge corresponding to it, meaning that the structural morphology will be different from the previous state. In summary, the changes in the morphology of a structure are bound to have a gradual development and evolution process due to the accumulation of internal damage and external loading effects, which is the landing point of various technologies in the field of structural health monitoring: trying to detect and capture the abnormal deformation of the structure as early as possible and then invert the safety state of the structure based on the monitoring data.

2.2. Image Matrix Similarity Damages Identification Method

The structural image matrix can be considered as a matrix consisting of pixel coordinates containing structural morphological information, denoted by the symbol a, as shown in Figure 1.

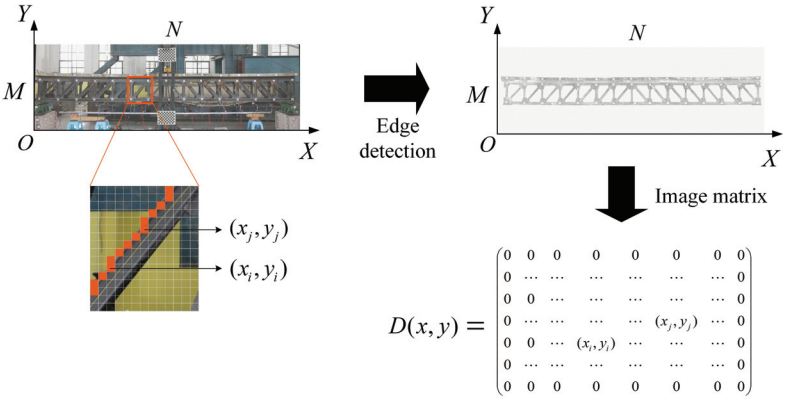


Figure 1. Image matrix.

If a structure is damaged at a place, the degradation of its stiffness matrix will destroy the original deformation continuity, and the abnormal changes at the damage site will be more obvious compared with other parts, which will be characterized as discontinuous pixel distribution at the edge of the structure on the image, so the similarity test of the structure image matrix can be carried out for damage identification. Similarity analysis is an analytical method to evaluate the degree of similarity between two things, and the commonly used similarity analysis methods are the Euclidean metric, Pearson correlation coefficient, and cosine similarity [23]; in this paper, the Euclidean distance metric is used as the index of image matrix similarity analysis, and the similarity function can be expressed by Equation (6).

$$\mathrm{sim}(D_0, D_1) = \rho\left(D_0\left(x^{(0)}, y^{(0)}\right), D_1\left(x^{(1)}, y^{(1)}\right)\right) = \sqrt{\left(x^{(0)} - x^{(1)}\right)^2 + \left(y^{(0)} - y^{(1)}\right)^2} \quad (6)$$

where $\mathrm{sim}(D_0, D_1)$ is the similarity between two image matrices D_0 and D_1.

$\rho\left(D_0\left(x^0, y^0\right), D_1\left(x^1, y^1\right)\right)$ represent the Euclidean distance between the corresponding elements of the two image matrices D_0 and D_1. x^0, y^0, x^1, and y^1, respectively, represent the horizontal and vertical coordinates of the elements in D_0 and D_1. From the relationship between structural damage and morphology mentioned earlier, it can be seen that when the structure is not damaged, the similarity distribution of the structure under different working conditions is a straight line or a continuous smooth curve. After the damage

occurs, the matrix similarity curve will show an abnormal peak response at the damage site, and the principle is shown in Figure 2.

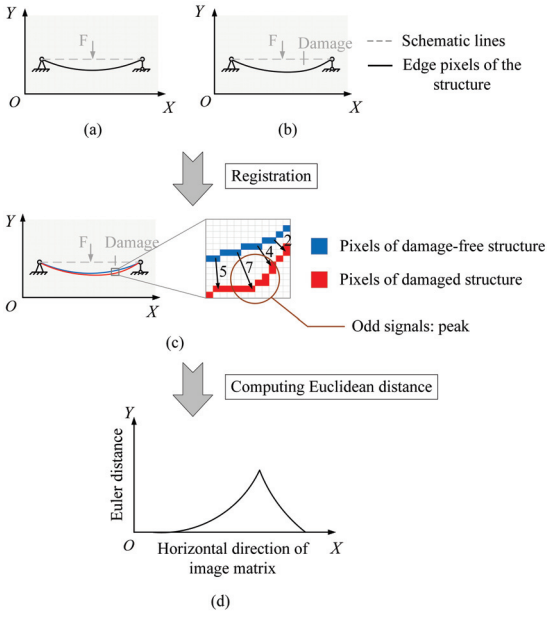


Figure 2. Image matrix similarity analysis process; (**a**) Image matrix of damage-free structure; (**b**) Image matrix of damaged structure; (**c**) Matching tie point; (**d**) Euclidean distance of the homologous points.

2.3. Numerical Validation of Image Matrix Similarity Damage Localization Method

A numerical model of the simply supported beam is established to verify the above damage localization method. The model span is 500 mm, and the cross-section is a rectangular section of size 20 mm × 6 mm. The beam is of uniform mass, the material is structural steel, the modulus of elasticity is 2×10^5 MPa, and the concentrated force of 1 kN acts in the middle of the span, as shown in Figure 3.

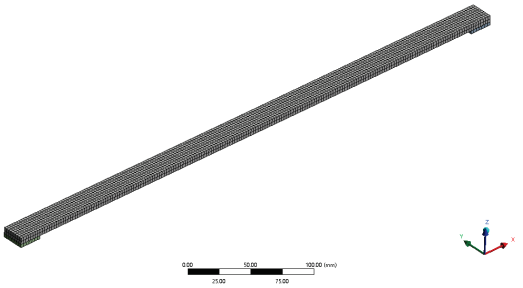


Figure 3. Finite element model of simply supported beam.

The mesh nodes on the $X - Z$ surface of the numerical model are simulated as image matrices. The deflection diagram and the corresponding image matrix of the undamaged model under the action of 1 kN concentrated force are shown in Figure 4.

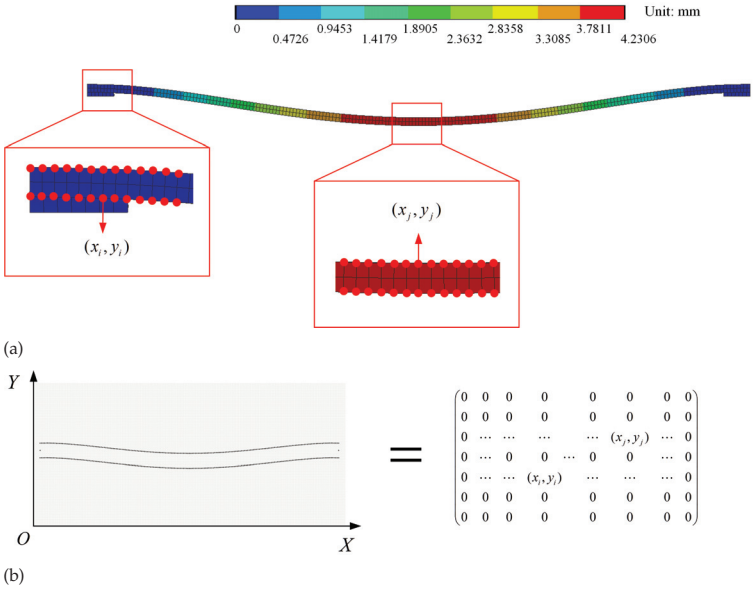


Figure 4. Acquisition of finite element model image matrix; (**a**) the model nodes (red dots in the diagram) represent the non-zero elements of the image matrix; (**b**) image matrix for finite element models.

(1) Single damage identification verification.

A damage crack is set at 200 mm from the right support with a width of 1 mm and a depth of 3 mm. Figure 5 shows the deflection of the damage model.

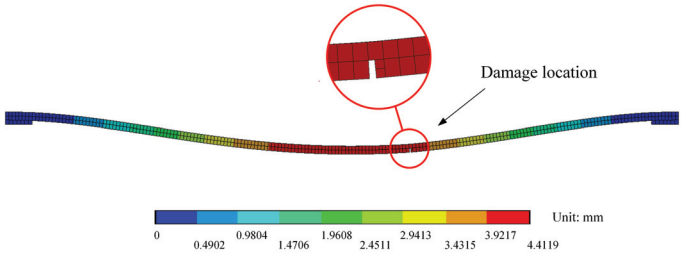


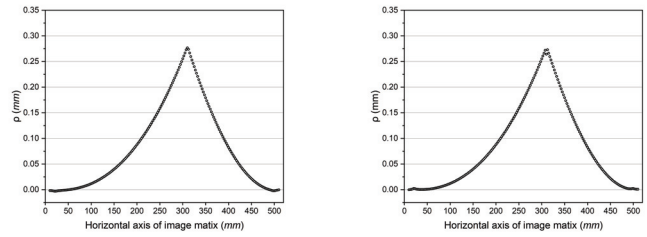
Figure 5. Deflection diagram of $X - Z$ section for a single crack in a simply supported beam.

The Euclidean distances of the corresponding points of the structure with no damage and post-damage are calculated separately under the same load. Figure 6 shows the Euclidean distances of the upper and lower edges of the model.

The image matrix similarity curves in Figure 6 have obvious peaks at the damage locations, indicating that the image matrix similarity analysis method can effectively identify the location of structural damage in the single damage condition.

(2) Eccentric multi-damage identification verification.

Set two cracks 100 mm and 200 mm from the right support, with a crack width of 1 mm and depth of 3 mm. The multi-damage model deformation is shown in Figure 7.



(a) (b)

Figure 6. Results of similarity analysis of single damage image matrix for simply supported beams; (**a**) Euclidean distance of the same name point on the upper edge of the model; (**b**) Euclidean distance of the same name point on the lower edge of the model.

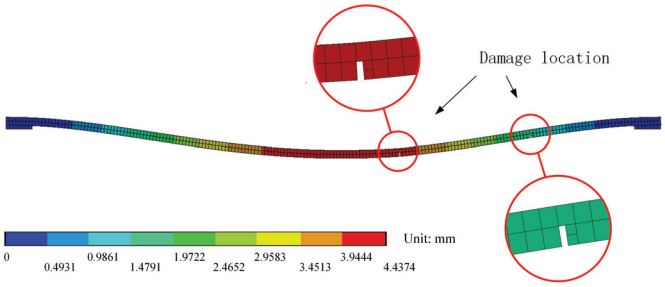
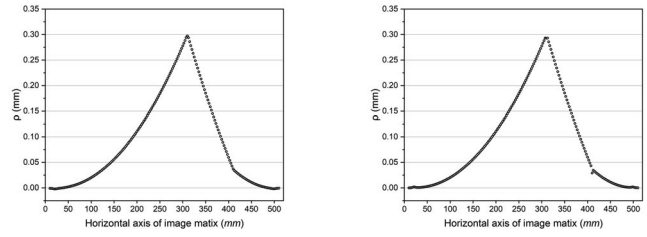


Figure 7. Location of the two cracks (eccentric situation).

Calculate the Euclidean distances of the homologous point in the image matrix of undamaged and post-damaged structures. The results of the analysis of the upper and lower edges of the simply supported beam are shown in Figure 8.



(a) (b)

Figure 8. Results of similarity analysis of two eccentric damage image matrices for simply supported beams; (**a**) Euclidean distance of the homologous point on the upper edge of the model; (**b**) Euclidean distance of the homologous point on the lower edge.

(3) Symmetric multi-damage identification verification.

Similarly, set up two cracks 200 mm from each side of the support with a width of 1 mm and a depth of 3 mm. The symmetric multi-damage model is shown in Figure 9.

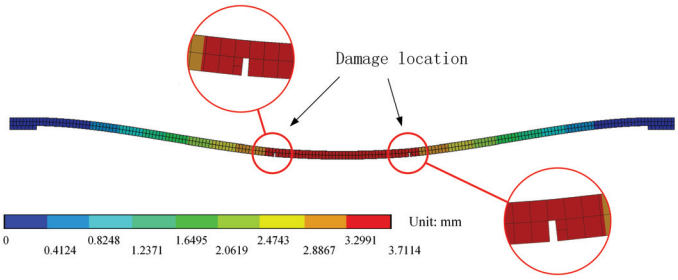
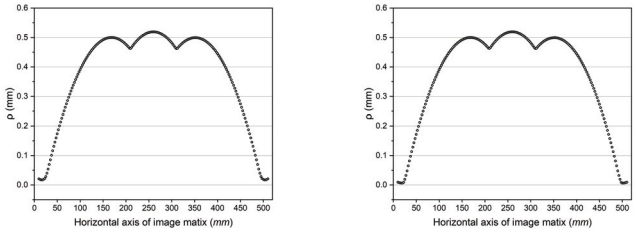


Figure 9. Location of the two cracks (Symmetrical situation).

The Euclidean distances of the homologous points in the image matrix of the undamaged and post-damaged structures are calculated. The results of the upper and lower edge analysis are shown in Figure 10.



(a) (b)

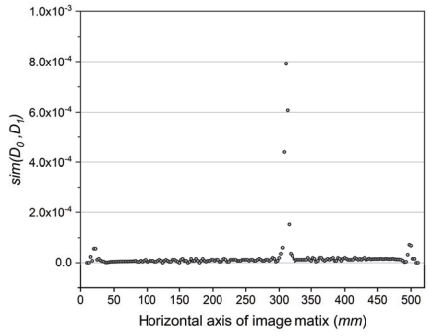
Figure 10. Results of matrix similarity analysis of two symmetrical damage images of simply supported beams; (**a**) Euclidean distance of the same name point on the upper edge of the model; (**b**) Euclidean distance of the same name point on the lower edge.

According to the above analysis, it is found that the extreme points of the similarity curve may not be the damage locations, and the discontinuity points are the damage locations, which is consistent with the deformation coordination of the structure. Therefore, the second-order derivative of the Euclidean distance curve is used to amplify this damage signal, so the similarity is calculated by Equation (7).

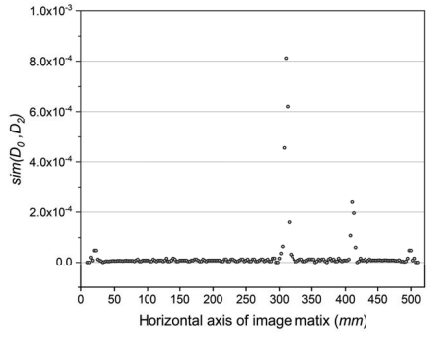
$$\mathrm{sim}(D_0, D_1) = \left| \frac{d^2\rho}{dx^2} \right| \tag{7}$$

The image matrix similarity analysis of the finite element model for single-damage and multi-damage conditions is re-performed using Equation (7), and the results can be obtained as in Figure 11.

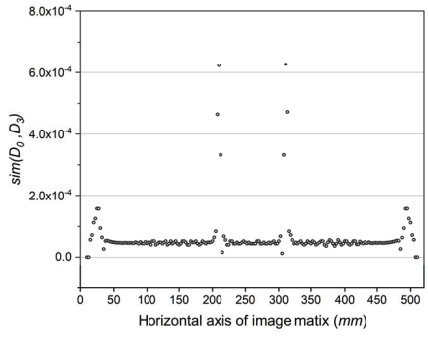
Figure 11 illustrates that the results of image matrix similarity analysis under multi-damage conditions are similar to those of single-damage identification, and the image matrix similarity has a clear peak at the damage location. In summary, the image matrix similarity analysis method can accurately locate the damage location on a multi-damaged simply supported beam. It is verified that the difference in the location of the damage does not affect its regularity.



(a)



(b)



(c)

Figure 11. Results of image matrix similarity analysis for various damage conditions; (**a**) single damage condition (position: $x = 300$ mm); (**b**) eccentric two damage condition (position: $x = 300$ mm & $x = 400$ mm); (**c**) symmetrical two damage condition (position: $x = 200$ mm & $x = 300$ mm).

3. Static Test of a Steel Truss–Concrete Composite Beam

To test the effectiveness of the image matrix similarity method proposed in this paper for damage identification on real structures, steel truss–concrete composite beam specimens have been designed and fabricated. The image matrix was obtained by extracting the pixel coordinates of the structure edges in the photos. The displacements extracted from the images were compared with the data from the dial gauges and the 3D laser scanner to ensure the accuracy of the images in describing the structural deformation.

3.1. Specimen Preparation

Figure 12a gives the dimensions and construction of the steel truss–concrete composite beam for the test. The bridge deck slab is made of C50 precast concrete with T-shaped cross section, total height of 140 mm, width of 500 mm and flange plate thickness of 100 mm. The bridge deck slab is assembled from 5 standard sections and 2 end sections, the length of the middle 5 standard sections is 1000 mm and the length of the end 2 sections is 1080 mm. The steel trusses are fabricated and welded in the factory. The completed steel truss–concrete composite beam specimen is shown in Figure 12b.

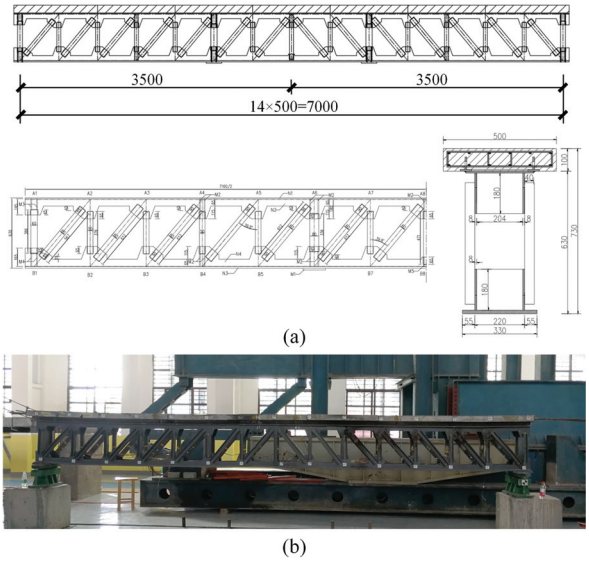


Figure 12. Steel truss–concrete composite beam for damage identification tests; (**a**) design dimensions of the test beam; (**b**) specimen entity.

3.2. Test Loading Protype and Damage

The loading point of the test beam is located in the middle of the span. Hydraulic jack (Maximum load 50 tons) and counterforce frame are used as loading devices, as shown in Figure 13, and pressure sensors are installed on the jack to ensure accurate loading. The specimen support type is the hinged support. The test beams in each damage condition were loaded with 150 kN and 250 kN as load class. After each class of loading, hold the load for two minutes to ensure the full deformation of the test beam, and then take pictures of the specimen and collect the displacement data. The loading protype is shown in Figure 14.

The damage to the rod was simulated by cutting the truss vertical rod, and the damage working condition of the specimen is shown in Table 1. The vertical rod number and the location of the damaged vertical rod are shown in Figure 15.

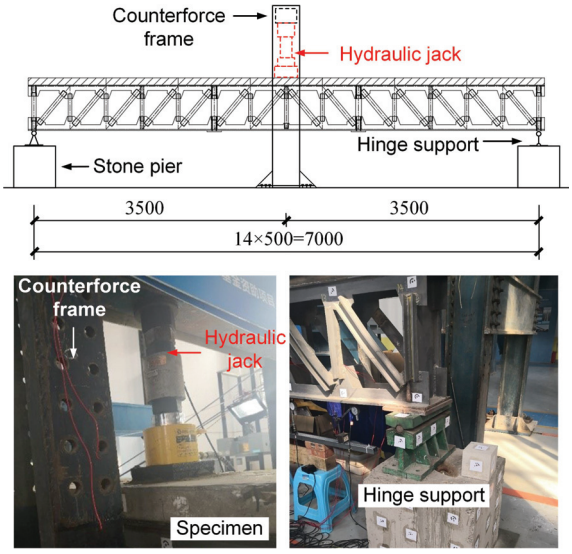


Figure 13. Arrangement of loading devices.

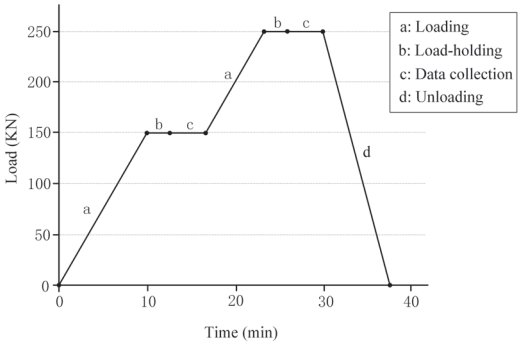


Figure 14. Loading system of static load test.

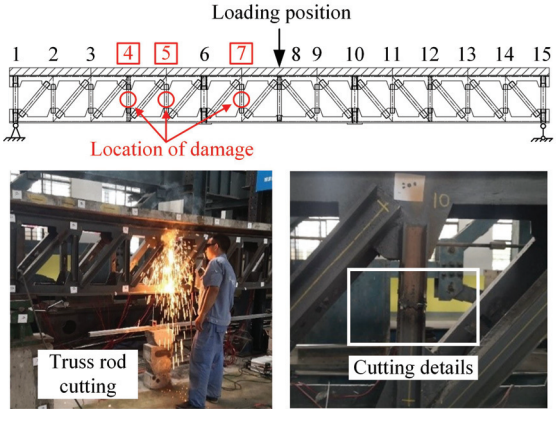


Figure 15. Location of loading and truss rod cutting.

Table 1. Summary of damage and load conditions.

Damage Conditions	Location of Damage	Loading Situation	Load (KN)
No damage		O_1 O_2 O_3	0 150 250
Damage to a rod	Rod No. 7	A_1 A_1 A_1	0 150 250
Damage to two rods	Rod No. 5	B_1 B_1 B_1	0 150 250
Damage to three rods	Rod No. 4	C_1 C_1 C_1	0 150 250

3.3. Image Data Acquisition

Using Canon 5DSR camera to take photos of the experiment specimen, the camera and lens parameters are shown in Table 2, and the camera calibration [24] results are shown in Table 3. The directly acquired specimen images are affected by the natural perspective and show obvious near-large and far-small features (Figure 16), resulting in the images not correctly reflecting the deformation of the structure, so a perspective transformation of the specimen images is required.

Table 2. Camera and lens parameters.

Number of Pixels	Sensor Size	Image Size	Aspect Ratio	Pixel Size	Lens Models	Lens Relative Aperture	Focal Length
50.6 million	36 × 24 mm	8688 × 5792	3:2	4.14 µm	EF 24–70 mm f/2.8LII	F2.8–F22	24–70 mm

Table 3. Calibration results.

Actual Image Size (Pixel)		Actual Image Centre Size (Pixel)		Actual Focal Length (mm)		Radial Distortion Parameters			Tangential Distortion Parameters	
X	Y	x_0	y_0	f_x	f_y	k_1	k_2	k_3	p_1	p_2
8688	5792	4319.74	2885.85	24.3	24.3	0.122	0.108	0.024	0	0

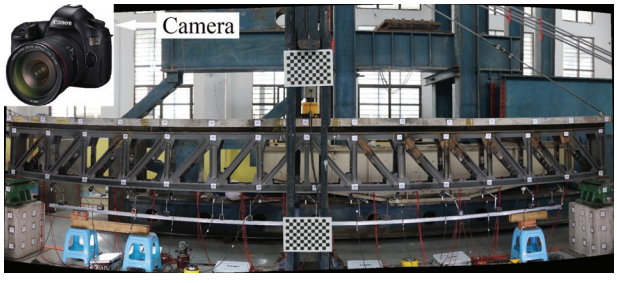


Figure 16. Test beam image without perspective transformation processing.

Image geometric transformation refers to the geometric transformation of image pixel positions without changing the original image content, mainly including translation, rotation, scaling, reflection, and misalignment. The perspective transformation is a combination of basic geometric transformations. The image of the measured structure will be distorted, when photographed under the condition of non-orthogonal projection [25,26]. If the image is mapped to the plane where the target structure is measured, called the measuring plane, in other words, the camera shoots perpendicular to the measurement plane, the real shape of the target structure can be obtained, and the perspective transformation model is shown in Figure 17.

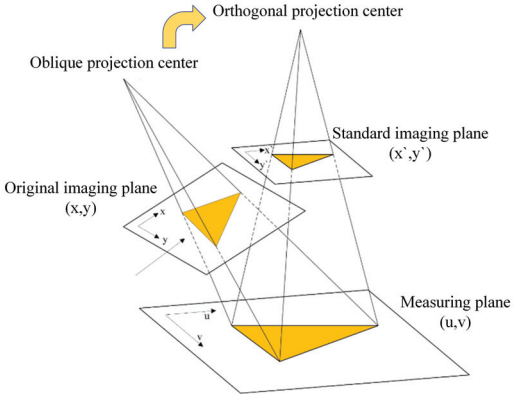


Figure 17. Image perspective transformation model.

Using the projection center of the camera as the origin to establish a three-dimensional Cartesian coordinate system, called the camera 3D coordinate system, let the original imaging plane be the xy plane, with the focal point at $[0, 0, f]^T$, $(f > 0)$. A two-dimensional coordinate system is established in the measuring plane, and the origin of this coordinate system is $[x_0, y_0, z_0]^T$ in the camera 3D coordinate system, the unit vector in the X-axis direction is $[u_1, u_2, u_3]^T$, and the unit vector in the Y-axis direction is $[v_1, v_2, v_3]^T$. The vectors adhere to the rules as follows:

$$\begin{cases} u_1v_1 + u_2v_2 + u_3v_3 = 0 \\ u_1^2 + u_2^2 + u_3^2 = v_1^2 + v_2^2 + v_3^2 = 1 \end{cases} \tag{8}$$

Then, the point with coordinates $[u, v]^T$ in the target object plane (measuring plane) can be expressed in the camera 3D coordinate system as

$$u \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + v \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} \tag{9}$$

Assuming that the point in the imaging plane has point coordinates $[x, y, 0]^T$, then $\exists k \in R$ must satisfy the following equation.

$$u \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + v \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ f \end{bmatrix} = k \left(\begin{bmatrix} 0 \\ 0 \\ f \end{bmatrix} - \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \right) \tag{10}$$

The following equations can be obtained.

$$-k \begin{bmatrix} x \\ y \end{bmatrix} = u \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + v \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} u_1 & v_1 & x_0 \\ u_2 & v_2 & y_0 \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{11}$$

$$kf = uu_3 + vv_3 + z_0 - f = \begin{bmatrix} u_3 & v_3 & z_0 - f \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{12}$$

From Equation (12), we then obtain the following:

$$-k = \begin{bmatrix} -\frac{u_3}{f} & -\frac{v_3}{f} & -\frac{z_0 - f}{f} \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{13}$$

Combining this with Equation (11) gives

$$-k \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} u_1 & v_1 & x_0 \\ u_2 & v_2 & y_0 \\ -\frac{u_3}{f} & -\frac{v_3}{f} & -\frac{z_0 - f}{f} \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{14}$$

Introducing the parameter matrix M, we obtain

$$M = \begin{bmatrix} u_1 & v_1 & x_0 \\ u_2 & v_2 & y_0 \\ -\frac{u_3}{f} & -\frac{v_3}{f} & -\frac{z_0 - f}{f} \end{bmatrix} \tag{15}$$

If the measuring plane does not pass through the focal point $[0, 0, f]^T$, then the matrix M must be an invertible matrix. Under normal operation, the focal point does not lie on the measuring plane, so the matrix M is always an invertible matrix. When the camera moves to a new position to photograph the target structure, the relative spatial position of the camera and the target structure changes. The change can be equated to the condition whereby the camera imaging plane does not move, while the focal length and the actual spatial position of the target structure change accordingly. Let the coordinates of the camera focus become $[0, 0, f']^T$ and the origin of the measuring plane become $[x'_0, y'_0, z'_0]^{\mathrm{T}}$. The unit vectors in the x, y axis of the scenic plane become $\left[u'_1, u'_2, u'_3\right]^{\mathrm{T}}$ and $\left[v'_1, v'_2, v'_3\right]^{\mathrm{T}}$, respectively. Similarly, $\exists k' \in R$ makes the coordinate point $[u, v]^{\mathrm{T}}$ on the measuring plane and the point $[x', y', 0]^{\mathrm{T}}$ of the imaging plane corresponding to the point $[u, v]^{\mathrm{T}}$ satisfy the following equation.

$$-k' \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} u'_1 & v'_1 & x'_0 \\ u'_2 & v'_2 & y'_0 \\ -\frac{u'_3}{f'} & -\frac{v'_3}{f'} & -\frac{z'_0 - f'}{f'} \end{bmatrix} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \tag{16}$$

Then, the parameter matrix M' is shown as

$$M' = \begin{bmatrix} u'_1 & v'_1 & x'_0 \\ u'_2 & v'_2 & y'_0 \\ -\frac{u'_3}{f'} & -\frac{v'_3}{f'} & -\frac{z'_0 - f'}{f'} \end{bmatrix} \tag{17}$$

Combining Equations (14) and (16) yields

$$-k' \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = M' \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = -kM'M^{-1} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \tag{18}$$

Assuming

$$M' \cdot M^{-1} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \tag{19}$$

Then, we can obtain

$$\begin{cases} k'x' = k(m_{11}x + m_{12}y + m_{13}) \\ k'y' = k(m_{21}x + m_{22}y + m_{23}) \\ k' = k(m_{31}x + m_{32}y + m_{33}) \end{cases} \tag{20}$$

Therefore, there is

$$\begin{cases} x' = \frac{m_{11}x + m_{12}y + m_{13}}{m_{31}x + m_{32}y + m_{33}} \\ y' = \frac{m_{21}x + m_{22}y + m_{23}}{m_{31}x + m_{32}y + m_{33}} \end{cases} \tag{21}$$

The point (x, y) on the original imaging plane can be transformed into the new imaging point (x', y') by perspective transformation, and (x', y') is the point of the orthorectified image that is not affected by perspective. The orthographic projection of the test beam after perspective transformation is shown in Figure 18.

Figure 18. Test beam image after perspective transformation process.

As can be seen from Figure 18, the specimen image after perspective transformation is no longer influenced by perspective and is free from the imaging characteristic of "large near and small far". Thus, the specimen image has the characteristics of orthographic projection, and the deformation information of the structure can be analyzed on this basis.

3.4. Accuracy-Verified Data Acquisition

3.4.1. Deflection Gauge Measurement

The deflection gauge is a traditional deformation measurement instrument that is often used as a basis for verifying the accuracy of experimental data because of its high accuracy [27].The conventional displacement measurement method of this test uses the deflection electric measurement system, which includes deflection gauge (range 0–30 mm, accuracy 0.01–mm) and DH5902N test and analysis system. The DH5902 acquisition frequency is set to 1 time in 2 seconds. Seven deflection gauges are installed directly below the vertical rod of the test beam at $L/8$, $L/4$, $3L/8$, $L/2$, $5L/8$, $3L/4$ and at the two supports, and the data acquisition system is arranged as shown in Figure 19.

The deflections of different load classes for each damage condition measured by the deflection gauge are summarized in Table 4. The process of collecting deflection data under no-damage conditions is shown in Figure 20.

From Table 4, it is found that in some positions, the damaged bars increase but the deflections decrease. The reason is that the steel joist–concrete combination beam has a complex internal support bar system, and the difference of structural deflection before and after the damage is not significant, which indicates that the deflection is not sensitive to the damage of the complex structure. Conventional single-point measurement damage identification methods are not ideal for this type of structure. Compared with the conventional damage identification methods, using this type of specimen as the test object can fully demonstrate the innovation and efficiency of structural holographic morphological data in the damage identification problem.

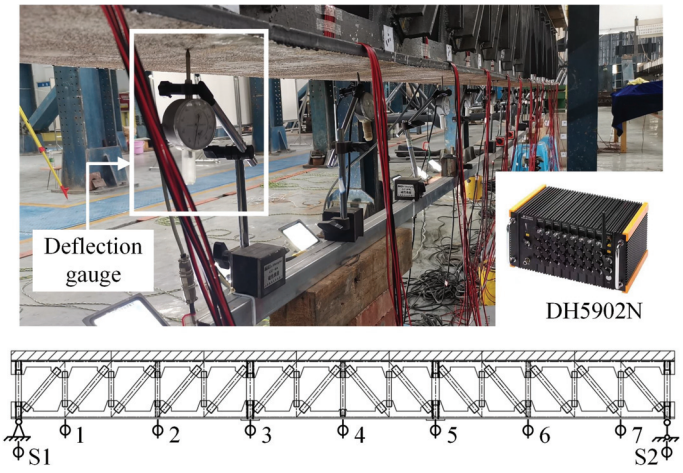


Figure 19. Deflection gauge layout.

Table 4. Deflection measurement results.

Number of Damaged Rods	Working Conditions	Load (kN)	Deflection Gauge Values (mm)								
			S1	L/8	L/4	3L/8	L/2	5L/8	3L/4	7L/8	S2
0	O_1	0	0	0	0	0	0	0	0	0	0
	O_2	150	0.717	3.884	6.101	8.696	9.514	8.725	6.33	2.696	0.732
	O_3	250	0.844	5.751	9.875	14.194	15.539	14.32	10.15	4.101	0.913
1	A_1	0	0	0	0	0	0	0	0	0	0
	A_2	150	0.689	2.401	5.949	8.541	10.693	8.921	6.198	2.001	0.694
	A_3	250	0.828	4.213	9.668	13.871	16.815	14.103	10.104	3.319	0.851
2	B_1	0	0	0	0	0	0	0	0	0	0
	B_2	150	0.704	3.356	5.959	8.764	10.61	8.923	6.358	2.603	0.688
	B_3	250	0.813	5.336	9.62	14.107	16.67	14.521	10.14	4.288	0.843
3	C_1	0	0	0	0	0	0	0	0	0	0
	C_2	150	0.655	4.155	6.479	9.246	10.583	9.524	6.384	3.139	0.664
	C_3	250	0.805	6.102	10.234	14.589	16.631	14.782	10.134	4.558	0.822

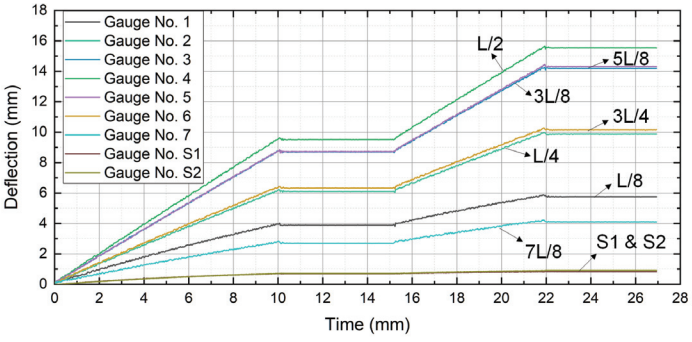
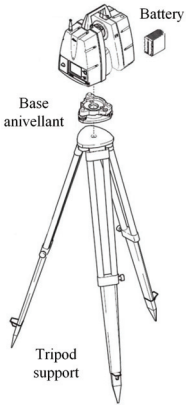


Figure 20. Deflection meter measurement process under no-damage conditions.

3.4.2. Three-Dimensional Laser Point Cloud Data Acquisition

Three-dimensional laser scanning technology is a non-contact, fast, and accurate measurement technique, and its application to the inspection of engineering structures has become a new trend of development [28,29]. It is widely used in the field of accurate deformation monitoring for bridges, buildings, tunnels, pipe racks, and other projects [30]. Ling Xiaochun [31] studied the influence of distance, incident angle, and target color on the accuracy of the scanner during measurement. Using the plane fitting method to analyze the accuracy, it was verified that the accuracy of a ground-based 3D laser is 1 to 2 mm, which meets the nominal accuracy; the angle measurement accuracy is about 15″; and the point position accuracy can reach the millimeter level. Moreover, this technology is based on the principle of laser ranging and can acquire a large amount of morphological data of the target object, so it is feasible to use 3D laser scanning data for the comparison and verification of images.

The Leica ScanStation P50 3D laser scanner is used to verify the extracted structural holographic deformation data, the basic parameters of which are shown in Figure 21. The scanning resolution is 0.8 mm/10 m, the point accuracy is 30 mm/50 m, the target acquisition accuracy is 2 mm/50 m, and the noise accuracy is 0.4 mm/10 m. According to the analysis results of the scanning angle effect on the point cloud density obtained from the previous study [32], the following scanning measurement verification test scheme was used: the scanning pattern was three-station joint scanning, the scanning radius was set to 10 m, and the scanning angle was $[-30°, 30°]$. The layout of the site is shown in Figure 22.



Leica ScanStation P50

Range accuracy	1.2 mm + 10ppm over full range (120 m / 270m mode) 3 mm + 10ppm over full range (570 m / >1 km mode)
Angular accuracy	8" horizontal; 8" vertical
Target acquisition	2mm standard deviation at 50m
Scan rate	Up to 1'000'000 points per second
Field-of-View	
Horizontal	360°
Vertical	290°

Figure 21. Basic parameters of Leica ScanStation P50 3D laser scanner.

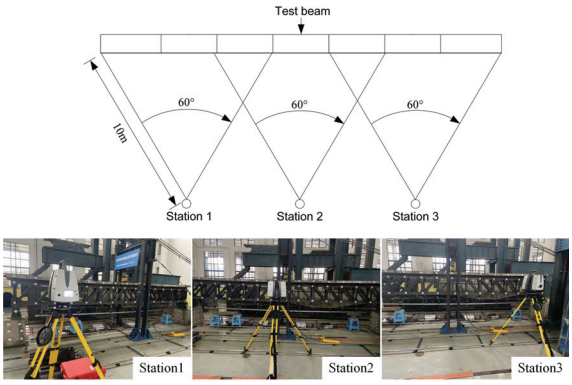


Figure 22. Layout of 3D laser scanning station.

Under this scanning scheme, the theoretical value of high-density point cloud coverage on the side of the test beam is >91.5% with high accuracy, which makes the 3D laser scanning data valuable for verifying the morphological deviations extracted from the images. The verification process of 3D laser scanning on the holographic morphology of the structural images is detailed in Section 4.3. In order to match the 3D laser scanning accuracy verification test, 20 mm diameter coded marker points were arranged on the upper and lower chord node plates of the test beam, and the location and number of the marker points are shown in Figure 23. The results of the 3D laser scan are shown in Figure 24.

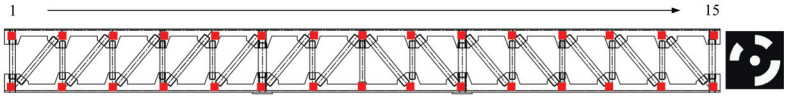


Figure 23. Layout of global sign points.

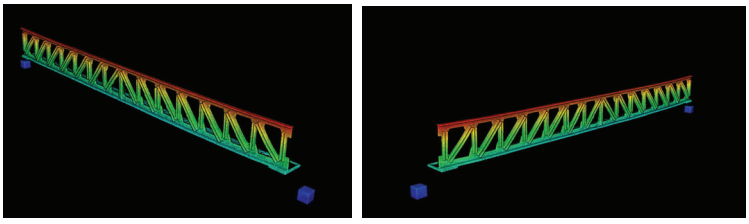


Figure 24. Test beam scanning results.

4. Bridge Structure Morphology Extraction

Based on the equivalent orthographic projection image, the features of the test beam under the O_1 working condition obtained by using the SIFT feature point extraction method [33] are shown in Figure 25.

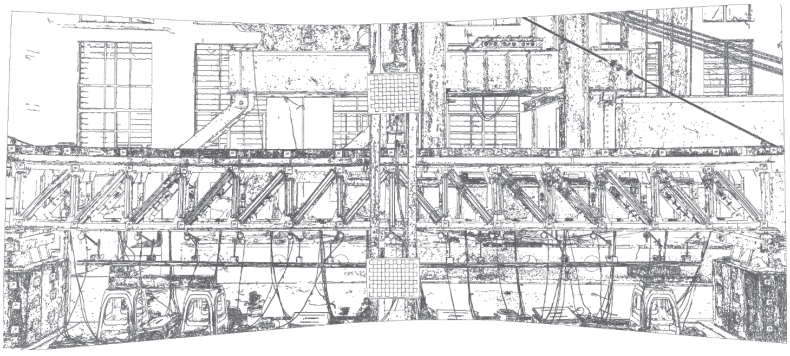


Figure 25. Test beam image feature points.

The main body edge contains the morphological information of the structure, which is the main area of structural feature point distribution and constitutes an important carrier of structural holographic deformation data. After simplifying the feature extraction results of the O_1 working condition, unnecessary environmental feature points are removed, and the main features of the structure are highlighted as shown in Figure 26.

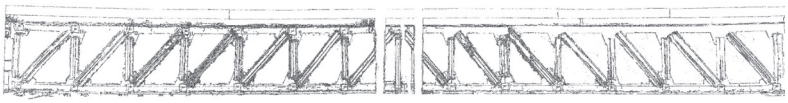


Figure 26. Main features of test beam.

4.1. Regression of Discontinuous Edges of Test Beam Images

The intrinsic nature of image features is points with discontinuous grayscale variations, and the dramatic grayscale variations imply the presence of high-frequency components of the signal of interest in the vicinity of the features. The previously extracted structural image features are not enough for structural holographic deformation monitoring. The reason for this is that it is difficult to clearly delineate the high-frequency components of the signal of interest on the bridge structure images from the environmental noise. For example, in a segment of the bridge edge image signal shown in Figure 27, point A is the ideal signal identified in the feature extraction process, and there is a step in the feature gray change at this point. Since the differential operation in the feature extraction process leads to the amplification of the noise signal, whether the step signal identified at points B and C is the true edge of the structure needs to be treated with caution. In fact, points B and C are most likely the synthesis of the characteristic signal with some noise.

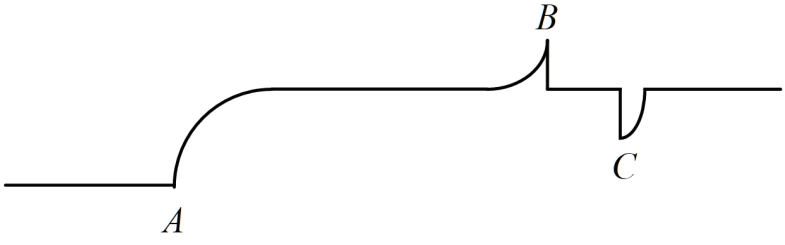


Figure 27. Schematic diagram of edge point and noise point.

Due to the variable environment of bridge structure monitoring, points with natural grayscale excess and continuous gradient changes, such as point A in Figure 27, are rarer on actual bridge images. The geographical and lighting environment in which the bridge is located would make most of the structural features accompanied by environmental noise, forming a large number of feature points with complex components, such as B and C. This leads to the real feature signal being smoothed out by the Gaussian spatial filter after the noise is mixed with the structural feature signal, and the actual performance of this problem on the bridge image is edge discontinuity and missing edge information (as shown in Figure 28). Structural edges are an important source of feature point generation, the missing edges cannot provide stable and rich point source data for structural holographic morphology analysis, and the missing edges need to be regressed.

The dilation and erosion operations are the two most important image boundary processing methods in morphology [34–38], and they are also the key means to achieve the concentration of structure-extracted edges toward the actual edges and the regression of edge breakpoints in this section. We use $f(x, y)$ to describe grayscale images, $b(i, j)$ for structural elements, and D_f, and D_b to define the domains of f and b. The dilation and erosion operations are described in Equations (22) and (23).

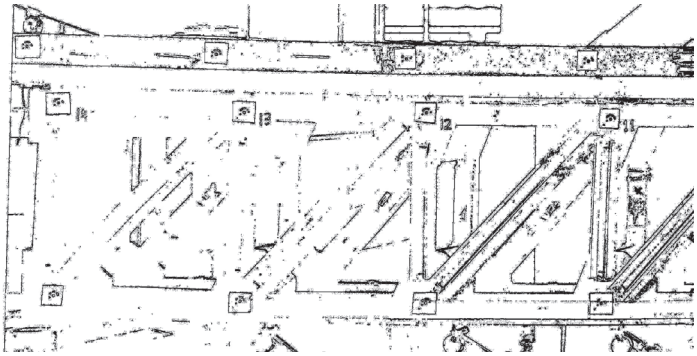


Figure 28. Edge discontinuity of bridge structure image.

(1) Definition of erosion:

$$(f\Theta b)(x,y) = \min\left\{f(x+i,y+i) - b(i,j) \mid (x+i,y+j) \in D_f; (i,j) \in D_b\right\} \quad (22)$$

(2) Definition of dilation:

$$(f \oplus b)(x,y) = \max\left\{f(x-i,y-i) - b(i,j) \mid (x-i,y-j) \in D_f; (i,j) \in D_b\right\} \quad (23)$$

The edges of the images obtained by the erosion or dilation operations alone are rough, and the edge noise composition is complicated in the actual image processing, so it is necessary to combine the erosion and dilation operations. In order to eliminate the external discrete points of the structure edge and smooth the structure edge, the process of first erosion and then dilation is used for edge treatment. For the problem of discontinuous breaks in the edge, the process of first dilation and then erosion is used to naturally connect the edge breaks while ensuring that the edge distribution does not expand outward. The morphological process of first erosion and then dilation is called the open operation (Equation (24)), and the process of first dilation and then erosion is called the closed operation (Equation (25)), and the above basic morphological operations are shown schematically in Figure 29.

(3) Open operation definition.

$$f(x,y) \circ b(x,y) = [(f\Theta b) \oplus b](x,y) \quad (24)$$

(4) Closed operation definition.

$$f(x,y) \bullet b(x,y) = [(f \oplus b)\Theta b](x,y) \quad (25)$$

From Figure 29, it can be seen that morphological operations can improve the continuity of edges and achieve the approximation of discontinuous edges to continuous edges when there are discontinuities in the edges of bridge images. The before-and-after comparison of the regression of the discontinuous edge of the test beam image is shown in Figure 30. It can be seen that after edge regression, the discrete edge points of the image become continuous curves, while smoothing out most of the noise. This image can be initially used to quantitatively identify the target deformation.

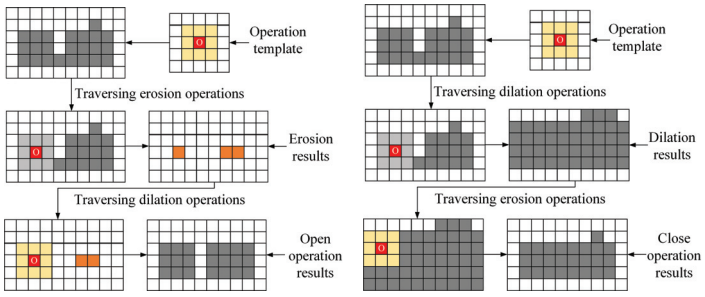


Figure 29. Basic operation diagram of morphology; (**a**) open operation; (**b**) closed operation.

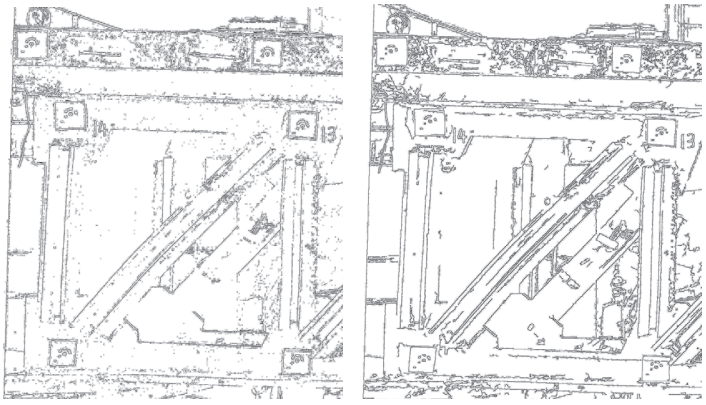


Figure 30. Comparison before and after regression treatment of discontinuous edges of the test beam; (**a**) before; (**b**) after.

4.2. Calibration of Bridge Image Resolution

The deformation of the structure is reflected in the image as a change in the pixel position of the deformed area. In the same shooting environment, how much deformation can cause the pixel position to change becomes a resolution issue for bridge image deformation monitoring. Theoretically, the deformation value has to be larger than the pixel size to be recognized by the bridge structure deformation monitoring method. The yellow calibration line is drawn in the truss vertical rod (Figure 31), which is a series of regular pixel matrix arrangements on the image, the physical size of these pixels in relation to the actual specimen is the monitoring resolution, and the calibration model is shown in Figure 32. The resolution calibration is calculated as $R = L/n$ (mm/pixel). The test beam has a total of 15 vertical rods, and yellow calibration lines are drawn in the middle of each vertical rod during the test. The exact length of each calibration line is actually measured, the number of pixels of the calibration line is counted on the image, and then the monitoring resolution is obtained as shown in Table 5.

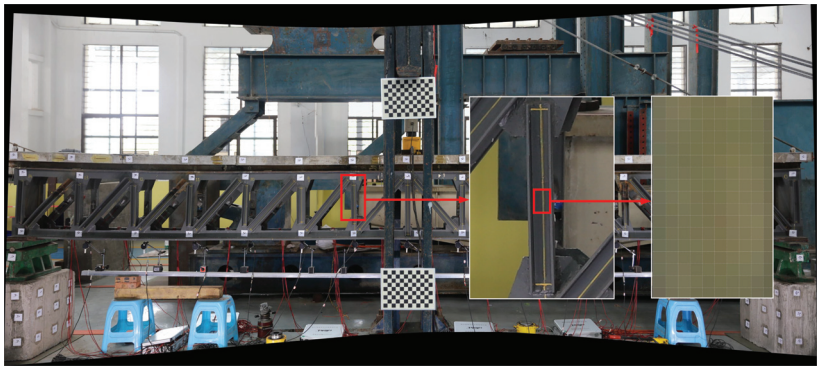


Figure 31. Array arrangement of pixels in image.

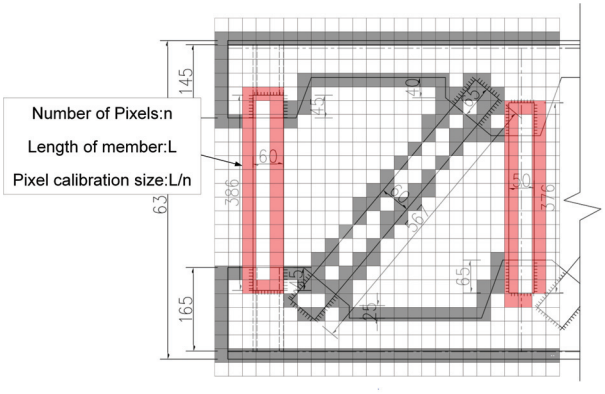


Figure 32. Pixel size calibration model.

Table 5. Resolution calibration table.

Vertical Truss Rod Number	Number of Pixels	Calibration Line Length (mm)	Calibrated Value (mm/px)	Calibration Average (mm/px)
1	1986	367.17	0.1849	
2	1999	359.21	0.1797	
3	1993	360.1	0.1807	
4	1991	357.3	0.1795	
5	1998	356.32	0.1783	
6	1994	312.39	0.1567	
7	1989	355.91	0.1789	
8	2654	453.94	0.171	0.1771
9	1982	356.83	0.18	
10	1993	321.44	0.1613	
11	1988	357.77	0.18	
12	1987	356.75	0.1795	
13	1994	361.37	0.1812	
14	1989	363.1	0.1826	
15	1983	359.47	0.1813	

The monitoring resolution of the bridge image is obtained by calibrating the pixel size. Using this resolution, the pixel dimensions of bridge images can be converted to

deformation monitoring dimensions, and then the measured deformation value of the structure can be obtained from the pixel bit difference.

4.3. Structure Body Morphology Extraction Results Validation

In this section, the morphology of the main body of the structural image extracted in Section 3.4.2 is verified quantitatively using 3D laser scanning measurements that have the ability to describe the holographic morphology of the structure and proven experience in engineering application.

After the 3D point cloud of the test beam is eliminated from perspective, a digital orthophoto map (DOM) of the point cloud of the test beam is made as shown in Figure 33. The calibrated morphology of the test beam obtained in Section 4.2 is superimposed on Figure 33, and the comparison of the main morphology of the structural image and the actual morphology of the DOM is shown in Figure 34.

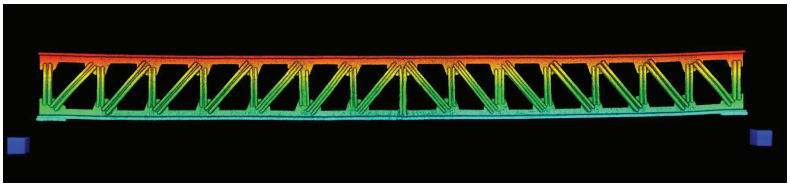


Figure 33. DOM obtained by 3D laser scanning of test beam.

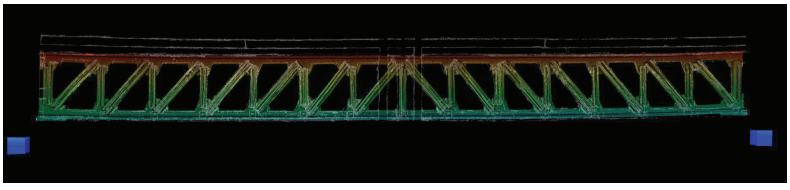


Figure 34. Direct contrast effect between image main shape and actual shape of test beam.

As can be seen from Figure 34, the main morphology of the extracted structure image matches well with the actual morphology of the structure obtained by 3D machine light scanning. The quantitative verification of the structural morphology extracted from the images follows the method used in the previous study [32]: the point cloud data are constituted as the NURBS surface of the test beam, the coordinates of the top left and bottom left vertices of the vertical rods are extracted, and all the vertical rod heights L_i and vertical rod spacing D_i of the test beam are calculated according to the patterns of Figure 35; at the same time, L_i' and D_i' of the same positions of the vertical rods in the structural morphology of the images are calculated, and the comparison results are shown in Table 6.

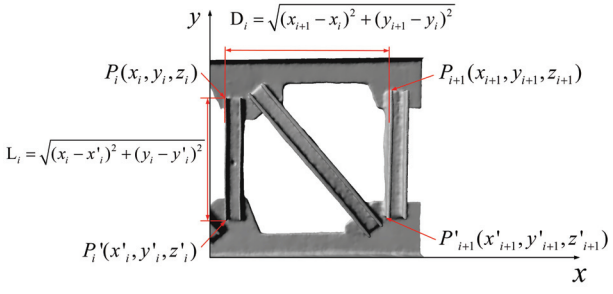


Figure 35. Calculation method of height and spacing of vertical bar by 3D laser scanning.

Table 6. Test beam image vertical bar height L_i', spacing D_i' extraction result verification table.

Vertical Rod Number	Rod Height Extracted from 3D Scan L_i/mm	Rod Height Extracted from Image L_i'/mm	Spacing Extracted from 3D Scan D_i/mm	Spacing Extracted from Image D_i'/mm
1	387.96	387.31 (−0.65)	509.87	509.43 (−0.44)
2	379.41	380.24 (0.83)	503.01	503.74 (0.73)
3	380.67	380.31 (−0.36)	501.74	502.21 (0.47)
4	377.67	378.06 (0.39)	500.58	499.76 (−0.82)
5	376.83	376.26 (−0.57)	491.36	491.05 (−0.31)
6	332.09	332.59 (0.5)	517.49	517.79 (0.30)
7	375.21	375.59 (0.38)	503.05	503.22 (0.17)
8	473.48	474.46 (0.98)	507.04	507.28 (0.24)
9	376.98	376.76 (−0.22)	523.49	524.12 (0.63)
10	341.28	342.06 (0.78)	488.16	488.74 (0.58)
11	377.17	377.25 (0.08)	508.43	507.96 (−0.47)
12	376.15	375.74 (−0.41)	509.25	509.44 (0.19)
13	381.90	382.41 (0.51)	509.56	509.86 (0.30)
14	383.17	383.45 (0.28)	513.39	512.81 (−0.58)
15	379.50	379.21 (−0.29)		

The values in parentheses are the absolute errors of the image morphology extraction values of the vertical rods compared with the 3D laser scanning results.

From Table 6, the length and spacing of vertical rods in the image-extracted structural morphology are consistent with the actual morphology of the structure with a maximum absolute error of 0.98 mm, indicating that the extracted image morphology of the test beam can correctly reflect the actual morphology of the structure. Therefore, the holographic morphological changes of any characteristic position of the structure in the image can be quantitatively analyzed.

5. Damage Identification Based on Image Matrix Similarity

5.1. Bridge Structure Edge Deformation Analysis

Since the deflection gauge data used as accuracy verification in Section 3.4.2 came from the bottom of the specimen beam, the lowermost edge of the test beam was selected as the extraction location of the characteristic edge for the overall structural deformation analysis, as shown in Figure 36. Figure 37 shows the contour line of the lower edge of the experimental beam under nondestructive conditions, and the part obscured by the reaction frame is taken to fill in the edge distribution of its neighborhood.

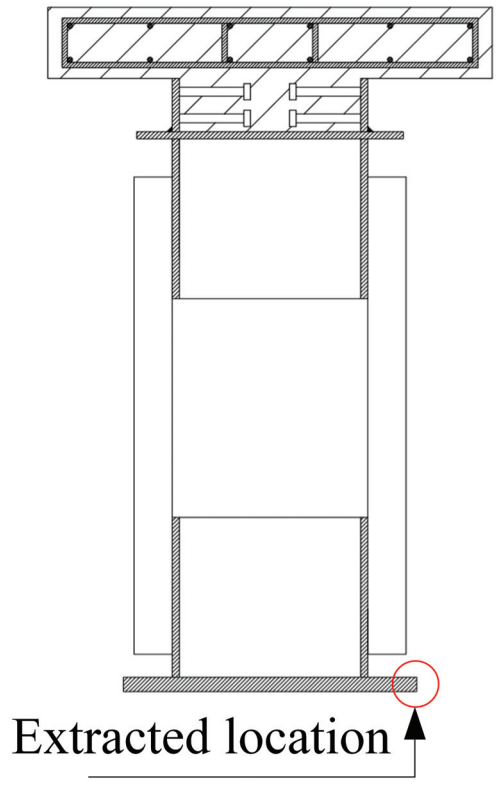


Figure 36. Edge extraction position.

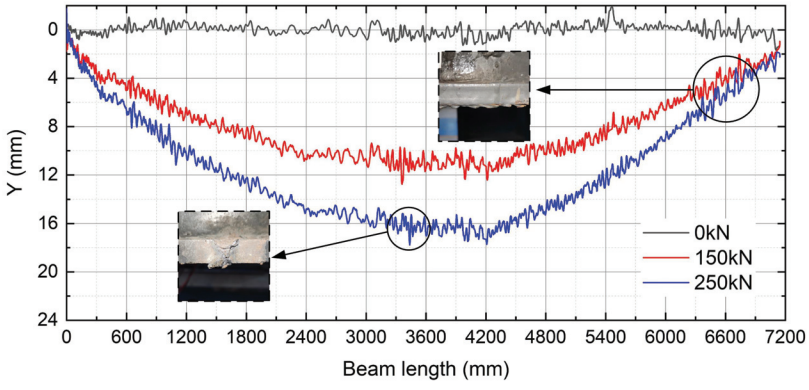


Figure 37. Original edge of the beam.

The extracted structure edge lineshapes for each loading case illustrated in Figure 37 show a distinct sawtooth effect, and the peaks and valleys of the oscillations are located in the same cross-section of the structure. There are three reasons for this regular discrepancy between the image extracted edge contours and the actual structure edges. 1, the edges of the actual structure have some discontinuous parts (As in the square in Figure 37) due to manufacturing errors and wear, which makes the contour pixel distribution at those parts show oscillations characteristics. 2, as shown in Figure 38 for the red pixel, the extracted

structural edges consist of one or more pixels, which causes some of the pixels to be lined up side by side hence forming a pixel bandwidth. This is due to the poor lighting of the shot scene causing image noise to be mixed with the edges, and thus the edges cannot be accurately located within a pixel range. The final edge position depends on the gradient of the grayscale values of all pixels in the pixel bandwidth. 3, as shown in Figure 38 for the pixel step location, there is a step at the point of continuous edges of the structure, and the step does not match the deformation trend of the structure, resulting in an unsmooth contour curve.

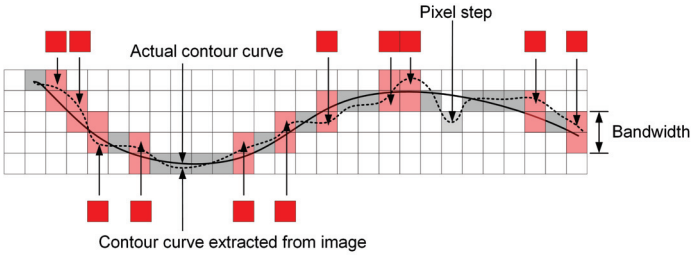


Figure 38. The reasons for the oscillation effect of structural image edges.

In order to eliminate this curve oscillation pattern, a structured edge line filtering method based on improved wavelet threshold denoising function [39–41] is adopted to perform signal decomposition, threshold filtering and signal reconstruction on the edge profile signal of Figure 37, so as to achieve noise reduction [42,43], and the processed edge profile is shown in Figure 39.

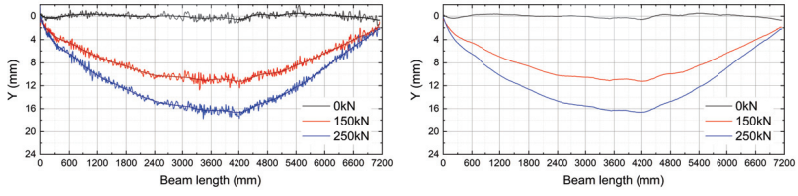


Figure 39. Lower edge curve of the test beam; (**a**) Original edge of the beam; (**b**) Edge curve after noise reduction.

To verify the accuracy of the extracted edge curves, an error analysis is performed on this curve. The data measured by the deflection gauge at each working condition were compared with the extracted beam edge curves for the corresponding working condition. Due to the large amount of data, only the comparison results for the damaged three rods are shown in Table 7.

As demonstrated in Table 7, the structural edge curves are consistent with the deformation values obtained from conventional measurements, with a maximum absolute error of 3.42%.

Table 7. Error analysis of beam deformation extracted from images.

Load Conditions	Load	Deflection Gauge Location	Deflection Gauge Values R_1/mm	Deformation Curve Values R_2/mm	Error $\frac{\lvert R_1-R_2\rvert}{R_1}$%
C_2	150kN	S1	0.655	0.634	3.11
		L/8	4.155	4.184	0.70
		L/4	6.479	6.609	2.02
		3L/8	9.246	8.929	3.42
		L/2	10.583	10.464	1.12
		5L/8	9.524	9.807	2.98
		3L/4	6.384	6.559	2.75
		7L/8	3.139	3.222	2.65
		S2	0.664	0.647	2.56
C_3	250kN	S1	0.805	0.779	3.19
		L/8	6.102	6.258	2.56
		L/4	10.234	10.229	0.04
		3L/8	14.589	14.842	1.74
		L/2	16.631	16.536	0.57
		5L/8	14.782	14.434	2.35
		3L/4	10.134	10.441	3.03
		7L/8	4.558	4.43	2.80
		S2	0.822	0.806	1.88

5.2. Damage Identification

5.2.1. Single Damage Test Beam Image Matrix Similarity Analysis

Under 250 kN load condition, the lower edge curve of the test beam without damage and one rod damage (No. 7 rod) is extracted according to the previous method, and after noise reduction, the pixel coordinates of the lower edge curve form the image coordinate matrices D_0 and D_1. Then, the Euclidean distance of the homologous image points of the two image matrices is calculated using the Equation (6), whose calculation process is shown in Figure 40, and further the similarity $sim(D_0, D_1)$ distribution (Figure 41) is obtained by amplifying the damaged signal according to the Equation (7). To facilitate the view of the damage recognition effect, only the image matrix similarity results for the truss vertical rod regions are shown in Figure 41.

From Figure 41, the similarity peaks at vertical bar #7, indicating that damage is more likely to occur here than at other bars. By connecting the maximum similarity at each vertical bar, the similarity envelope of the test beam in the single damage condition can be obtained (Figure 42). The similarity envelope gives a more concise and obvious indication of the damage location.

Figure 42 shows that the image matrix similarity analysis method works well for single damage identification of the test beam and can accurately locate the damage location.

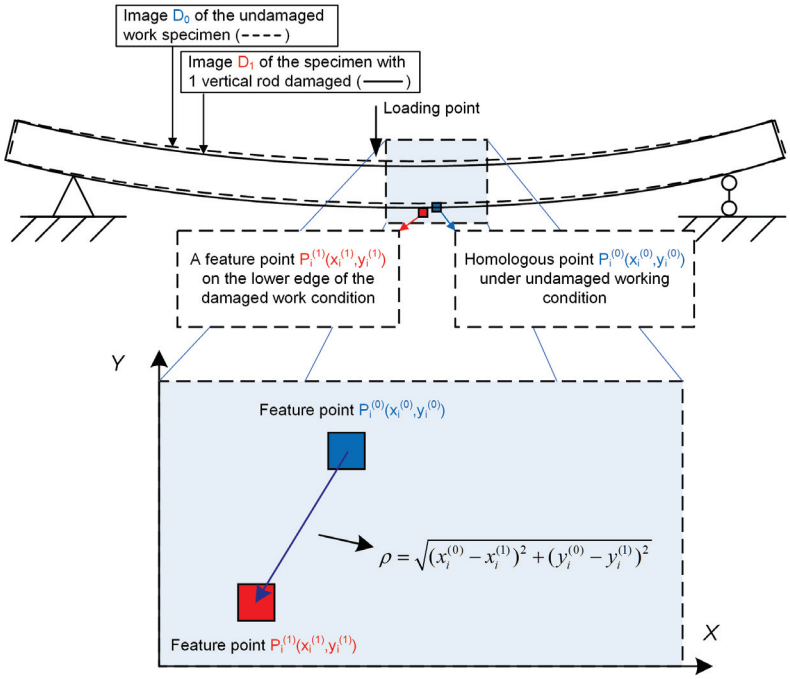


Figure 40. The calculation process of Euclidean distance in Equation (6).

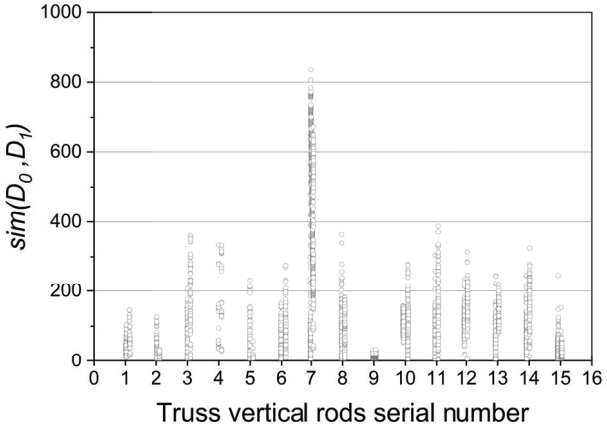


Figure 41. Single damage identification results of test beam.

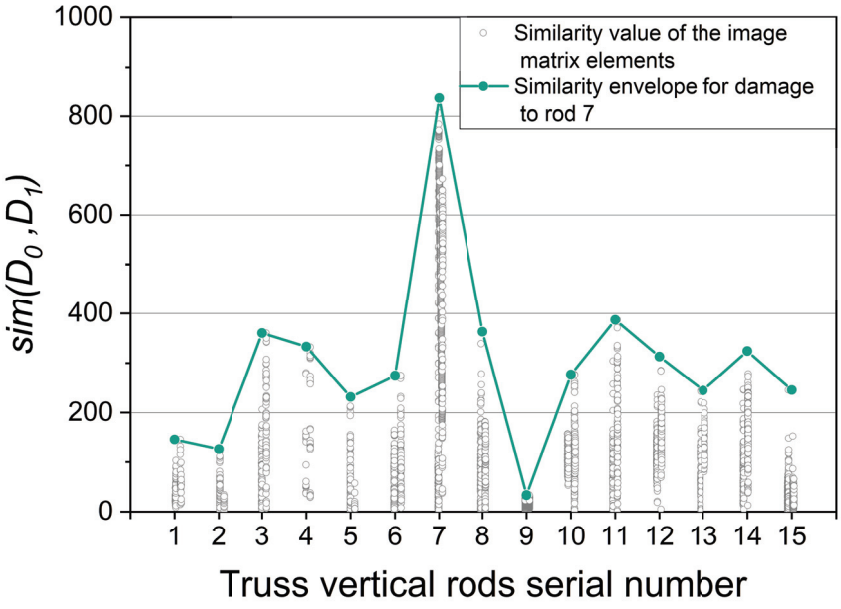


Figure 42. Single damage condition similarity envelope.

5.2.2. Multi-Damage Test Beam Image Matrix Similarity Analysis

The structural edge curves of two-bar damage (bars #5 and #7) and three-bar damage (bars #4, #5 and #7) under 250 kN load condition are selected, and the steps in Section 5.2.1 are repeated to obtain the image coordinate matrices D_2 and D_3. The similarity between them and the image matrices in the no damage condition is calculated respectively as shown in Figure 43.

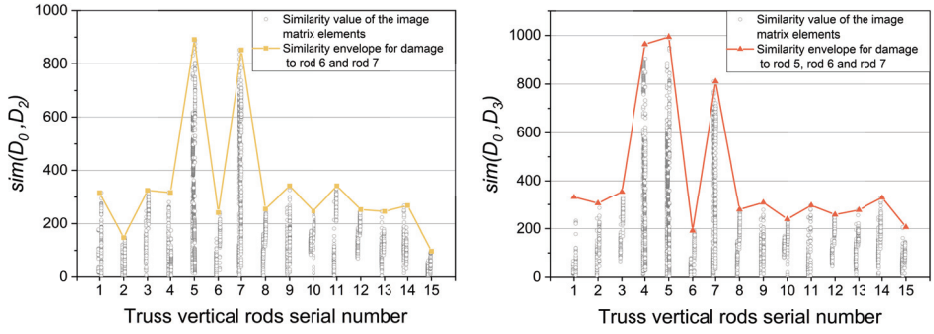


Figure 43. Multi-damage condition similarity envelope; (**a**) two-damage conditions; (**b**) three-damage conditions.

To compare the image matrix similarity damage identification effects for different damage conditions, the similarity envelopes are integrated into the same coordinate system, as in Figure 44.

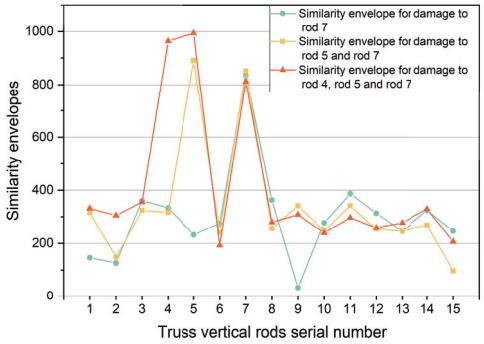


Figure 44. Similarity envelopes for all damage conditions.

In summary, the analysis found that the test results are consistent with the results of the numerical simulations in Section 2.3, with the similarity (*sim*) envelope having a significant peak response at the damage location of the test specimen. This result indicates that the image matrix similarity analysis method proposed in this paper can accurately identify the damage locations on the test beams under each working condition and, at the same time, demonstrates that the overall deformation curve of the structure contains anomalous signals arising from local stiffness degradation caused by the damage.

6. Conclusions

Using digital image processing techniques, an image matrix similarity damage identification method for bridge structures is proposed based on the ability of images to describe the holographic deformation of structures. The test results show that the method can accurately identify the singular signal of local stiffness degradation caused by structural damage from the overall structural deformation and achieve the localization of damage. This solves the difficulty of damage identification due to incomplete test data and provides a new idea for beam damage identification. The research in this paper has engineering application prospects for small and medium-span bridges and for fast, economic, and efficient long-term health monitoring. The main conclusions of this paper are as follows.

(1) A digital orthophoto-based image matrix similarity damage identification method is proposed. Numerical simulation studies show that the damage will break the structural system deformation coordination, the abnormal distribution of pixels at the damage site will manifest odd signals in the matrix similarity test before and after the damage, the signal characteristics are shown as similarity curve discontinuity, and the discontinuity location is consistent with the damage site.

(2) A damage loading test of steel truss–concrete composite beam has been conducted, and the original test beam image is corrected by perspective transformation to obtain an equivalent digital orthophoto. Further, the holographic morphology of the test beams extracted by the SIFT feature extraction algorithm is verified, and the results show that the extracted morphology of the test beams is consistent with the real morphology of the structure.

(3) The lower edge curve of the experimental beam is affected by noise with regular oscillation, and the noise-containing curve is processed by the wavelet denoising function to obtain a continuous smooth lower edge curve of the structure. The validation results show that the maximum error of the noise-reduced curve is 3.42%.

(4) The image matrices of the structure before and after the damage are obtained from the coordinates of the lower edge curve. The similarity envelopes for each damage condition were derived by calculating the similarity of the image matrices under single damage condition and multiple damage conditions. The peak of the envelope is consistent

with the position of the damaged rod, which verifies the accuracy of the damage position identification in practical applications.

Author Contributions: Conceptualization, R.L. and X.C.; methodology, formal analysis, and investigation, R.L., X.L. and J.M.; writing—original draft preparation, R.L.; writing—review and editing, R.L. and X.C.; visualization, R.L.; supervision, Z.Z. and X.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (Grant No. 51778094).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Schlune, H.; Plos, M.; Gylltoft, K. Improved bridge evaluation through finite element model updating using static and dynamic measurements. *Eng. Struct.* **2009**, *31*, 1477–1485. [CrossRef]
2. Marinone, T.; Dardeno, T.; Avitabile, P. Reduced model approximation approach using model updating methodologies. *J. Eng. Mech.* **2018**, *144*, 04018005. [CrossRef]
3. Zhang, J.; Guo, S.; Wu, Z.; Zhang, Q. Structural identification and damage detection through long-gauge strain measurements. *Eng. Struct.* **2015**, *99*, 173–183. [CrossRef]
4. Cui, F.; Yuan, W.; Shi, J. Damage Detection of Structures Based on Static Response. *J. Tongji Univ.* **2000**, *1*, 8–11.
5. Xu, Z.D.; Li, S.; Zeng, X. Distributed strain damage identification technique for long-span bridges under ambient excitation. *Int. J. Struct. Stab. Dyn.* **2018**, *18*, 1850133. [CrossRef]
6. Banan, M.R.; Banan, M.R.; Hjelmstad, K. Parameter estimation of structures from static response. I. Computational aspects. *J. Struct. Eng.* **1994**, *120*, 3243–3258. [CrossRef]
7. Kourehli, S.S.; Bagheri, A.; Amiri, G.G.; Ghafory-Ashtiany, M. Structural damage detection using incomplete modal data and incomplete static response. *KSCE J. Civ. Eng.* **2013**, *17*, 216–223. [CrossRef]
8. Li, S.; Ma, L.; Tian, H. Structural damage detection method using incomplete measured modal data. *J. Vib. Shock* **2015**, *34*, 196–203.
9. He, R.S.; Hwang, S.F. Damage detection by a hybrid real-parameter genetic algorithm under the assistance of grey relation analysis. *Eng. Appl. Artif. Intell.* **2007**, *20*, 980–992. [CrossRef]
10. Savadkoohi, A.T.; Molinari, M.; Bursi, O.S.; Friswell, M.I. Finite element model updating of a semi-rigid moment resisting structure. *Struct. Control Health Monit.* **2011**, *18*, 149–168. [CrossRef]
11. Hua, J. *Research on Bridge'S Damage Detection and Evaluation Based on Static Test Data*; Southwest Jiaotong University: Chengdu, China, 2005.
12. Liu, G.; Mao, Z. Structural damage diagnosis with uncertainties quantified using interval analysis. *Struct. Control Health Monit.* **2017**, *24*, e1989. [CrossRef]
13. Yan, L. *The Research of Suspension Bridge Damage Detection Based on the Grey Relation Theory and Genetic Algorithm*; Lanzhou Jiaotong University: Lanzhou, China, 2015.
14. WenLong, G.; Jun, C.; DaShan, L. Structural damage identification based on quantum particle swarm optimization algorithm. *J. Dyn. Control* **2015**, *13*, 388–393.
15. Fang, Z.; Zhang, G.G.; Tang, S.H.; Chen, S.J. Finite element modeling and model updating of concrete cable-stayed bridge. *China J. Highw. Transp.* **2013**, *26*, 77–85.
16. Wang, Q.; Wu, N. Detecting the delamination location of a beam with a wavelet transform: An experimental study. *Smart Mater. Struct.* **2010**, *20*, 012002. [CrossRef]
17. Song, Y.Z.; Bowen, C.R.; Kim, H.A.; Nassehi, A.; Padget, J.; Gathercole, N.; Dent, A. Non-invasive damage detection in beams using marker extraction and wavelets. *Mech. Syst. Signal Process.* **2014**, *49*, 13–23. [CrossRef]
18. Ma, H.; Zhang, W.; Wang, Z. Application of wavelet analysis in cantilever beam crack identification. *Chin. J. Comput. Mech.* **2008**, *05*, 148–154.
19. XiaoWei, Y.; Tan, L.; ChuanZhi, D. Structural damage detection based on Kalman filter and neutral axis location. *China J. Zhejiang Univ. (Eng. Sci.)* **2017**, *10*, 137–143.
20. Pan, H.; Azimi, M.; Yan, F.; Lin, Z. Time-frequency-based data-driven structural diagnosis and damage detection for cable-stayed bridges. *J. Bridge Eng.* **2018**, *23*, 04018033. [CrossRef]
21. Fan, W.; Qiao, P. Vibration-based damage identification methods: A review and comparative study. *Struct. Health Monit.* **2011**, *10*, 83–111. [CrossRef]

22. Brownjohn, J.M.; De Stefano, A.; Xu, Y.L.; Wenzel, H.; Aktan, A.E. Vibration-based monitoring of civil infrastructure: Challenges and successes. *J. Civ. Struct. Health Monit.* **2011**, *1*, 79–95. [CrossRef]
23. Zheng, J.; Chen, D.; Hu, H. Boundary Adjusted Network Based on Cosine Similarity for Temporal Action Proposal Generation. *Neural Process. Lett.* **2021**, *53*, 2813–2828. [CrossRef]
24. Chu, X.; Zhou, Z.; Deng, G.; Duan, X.; Jiang, X. An overall deformation monitoring method of structure based on tracking deformation contour. *Appl. Sci.* **2019**, *9*, 4532. [CrossRef]
25. Wang, J.; Mao, Y.; Yan, T.; Liu, Y. Perspective Transformation Algorithm for Light Field Image. *Laser Optoelectron. Prog.* **2019**, *56*, 151003. [CrossRef]
26. Yuan, R.; Liu, M.; Hui, M.; Zhao, Y.; Dong, L. Depth map stitching based on binocular vision. *Laser Optoelectron. Prog.* **2018**, *55*, 282–287.
27. Jiang, J.; Zou, Y.; Yang, J.; Zhou, J.; Zhang, Z.; Huang, Z. Study on Bending Performance of Epoxy Adhesive Prefabricated UHPC-Steel Composite Bridge Deck. *Adv. Civ. Eng.* **2021**, *2021*, 6658451. [CrossRef]
28. Pinpin, L.; Wenge, Q.; Yunjian, C.; Feng, L. Application of 3D laser scanning in underground station cavity clusters. *Adv. Civ. Eng.* **2021**, *2021*, 8896363. [CrossRef]
29. Zhang, H.; Xia, J. Research on convergence analysis method of metro tunnel section–Based on mobile 3D laser scanning technology. *IOP Conf. Ser. Earth Environ. Sci.* **2021**, *669*, 012008. [CrossRef]
30. Wu, C.; Yuan, Y.; Tang, Y.; Tian, B. Application of Terrestrial Laser Scanning (TLS) in the Architecture, Engineering and Construction (AEC) Industry. *Sensors* **2022**, *22*, 265. [CrossRef]
31. Ling, X. Research on Building Measurement Accuracy Verification Based on Terrestrial 3D Laser Scanner. *IOP Conf. Ser. Earth Environ. Sci.* **2021**, *632*, 052086. [CrossRef]
32. Xi, C.X.; xiang Zhou, Z.; Xiang, X.; He, S.; Hou, X. Monitoring of long-span bridge deformation based on 3D laser scanning. *Ingénierie Des Systèmes D'information* **2018**, *18*, 113–130. [CrossRef]
33. Zhou, S.; Wu, X.; Qi, Y.; Luo, S.; Xie, X. Video shot boundary detection based on multi-level features collaboration. *Signal Image Video Process.* **2021**, *15*, 627–635. [CrossRef]
34. Hashimoto, R.F.; Barrera, J.; Ferreira, C.E. A Combinatorial Optimization Technique for the Sequential Decomposition of Erosions and Dilations. *J. Math. Imaging Vis.* **2004**, *13*, 17–33. [CrossRef]
35. Pillon, P.E.; Pedrino, E.C.; Roda, V.O.; do Carmo Nicoletti, M. A hardware oriented ad-hoc computer-based method for binary structuring element decomposition based on genetic algorithms. *Integr. Comput. Aided Eng.* **2016**, *23*, 369–383. [CrossRef]
36. Sussner, P.; Pardalos, P.M.; Ritter, G.X. On integer programming approaches for morphological template decomposition problems in computer vision. *J. Comb. Optim.* **1997**, *1*, 165–178. [CrossRef]
37. Deng, S.; Huang, Y. Fast algorithm of dilation and erosion for binary image. *Comput. Eng. Appl.* **2017**, *53*, 207–211.
38. Dokladalova, E. Algorithmes et Architectures Efficaces Pour Vision Embarquée. Ph.D. Thesis, Université Paris Est, Paris, France, 2019.
39. Chu, X.; Zhou, Z.; Deng, G.; Jiang, T.; Lei, Y. Study on Damage Identification of Beam Bridge Based on Characteristic Curvature and Improved Wavelet Threshold De-Noising Algorithm. *Adv. Model. Anal.* **2017**, *60*, 505–524. [CrossRef]
40. Tengjiao, J. *Experimental Study on Damage Conditions of the Steel-Concrete Composite Beam Based on the Bridge Surface*; Chongqing Jiaotong University: Chongqing, China, 2018.
41. Xi, C. *Holographic Shape Monitoring and Damage Identification of Bridge Structure Based on Fixed Axis Rotation Photography*; Chongqing Jiaotong University: Chongqing, China, 2020.
42. Andrade, L.C.M.D.; Oleskovicz, M.; Fernandes, R.A.S. Adaptive threshold based on wavelet transform applied to the segmentation of single and combined power quality disturbances. *Appl. Soft Comput.* **2016**, *38*, 967–977. [CrossRef]
43. Hu, Z.; Liu, L. Applications of wavelet analysis in differential propagation phase shift data de-noising. *Adv. Atmos. Sci.* **2014**, *31*, 825–835. [CrossRef]

aerospace

MDPI

Article

Fault Detection of Aero-Engine Sensor Based on Inception-CNN

Xiao Du [1], Jiajie Chen [1], Haibo Zhang [1] and Jiqiang Wang [2,*]

[1] College of Energy and Power Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China; duxiao@nuaa.edu.cn (X.D.); cjj9654@nuaa.edu.cn (J.C.); zhhbason@nuaa.edu.cn (H.Z.)
[2] Ningbo Institute of Materials Technology & Engineering, Chinese Academy of Sciences, Ningbo 315201, China
* Correspondence: wangjiqiang@nimte.ac.cn

Abstract: The aero-engine system is complex, and the working environment is harsh. As the fundamental component of the aero-engine control system, the sensor must monitor its health status. Traditional sensor fault detection algorithms often have many parameters, complex architecture, and low detection accuracy. Aiming at this problem, a convolutional neural network (CNN) whose basic unit is an inception block composed of convolution kernels of different sizes in parallel is proposed. The network fully extracts redundant analytical information between sensors through different size convolution kernels and uses it for aero-engine sensor fault detection. On the sensor failure dataset generated by the Monte Carlo simulation method, the detection accuracy of Inception-CNN is 95.41%, which improves the prediction accuracy by 17.27% and 12.69% compared with the best-performing non-neural network algorithm and simple BP neural networks tested in the paper, respectively. In addition, the method simplifies the traditional fault detection unit composed of multiple fusion algorithms into one detection algorithm, which reduces the complexity of the algorithm. Finally, the effectiveness and feasibility of the method are verified in two aspects of the typical sensor fault detection effect and fault detection and isolation process.

Keywords: CNN; sensor fault detection; aircraft engine; Monte Carlo simulation method; inception block

Citation: Du, X.; Chen, J.; Zhang, H.; Wang, J. Fault Detection of Aero-Engine Sensor Based on Inception-CNN. *Aerospace* **2022**, *9*, 236. https://doi.org/10.3390/aerospace9050236

Academic Editors: Phong B. Dao, Lei Qiu and Liang Yu

Received: 1 March 2022
Accepted: 21 April 2022
Published: 25 April 2022

1. Introduction

As the measuring element of the aero-engine control system, the function of the sensor is fundamental and essential for the system. However, with the increasing control demand and the increasing complexity of the control system, new requirements are put forward for the number and reliability of sensors, and in the poor working environment of aero-engine sensors, faults are difficult to avoid [1,2]. According to statistics, sensor faults cover more than 80% of faults in the aero-engine control system [3], so the fault detection and isolation (FDI) of aero-engine sensors are vital to improving the reliability of aero-engine control systems.

At present, the FDI methods of aero-engine sensors can be divided into two types: model-based and signal-based. The main idea of the model-based method is to establish the mapping relationship between the actual input and output of the controlled object by analyzing the physical characteristics of the controlled object and then analyzing the residuals between the output of the actual system and the output of the model for diagnosis and isolation [4]. The prominent representatives are the Kalman filter [5–10], particle filter [11], etc. Based on the model method, sensor fault detection becomes very simple under the premise of accurate model mapping. However, the complexity of actual controlled objects is the difficulty of its practical application.

Signal-based methods directly analyze signals through reliability analysis or machine learning, among which machine learning has made significant progress in the field of sensor FDI. Regarding fault identification based on one-dimensional sensor signals, literature [12] expands the one-dimensional signals of a single sensor into two-dimensional data. It

inputs them into a CNN for fault detection through time series misalignment. However, this approach has poor randomness and reliability for aero-engines with complex control systems and diverse working conditions. Therefore, the most common solution is to use the analytically redundant space information between multiple sensors to detect faults of single or multiple sensors. In literature [13–15], the support vector machine (SVM) algorithm is used to build a nonlinear mapping relationship between faulty sensors and healthy sensors. Finally, an appropriate threshold is selected for the predicted value of the algorithm to determine whether the sensor is faulty. The literature [16–18] used the Levenberg–Marquardt algorithm, Artificial Bee Colony algorithm, and adaptive method to optimize a BP neural network to construct the mapping relationship between sensors for banning SVM. All of the above methods have high accuracy for single-sensor fault diagnosis. However, for multi-sensor diagnosis, it is necessary to build multiple algorithm units and match them with a complex logical judgment method or fusion judgment algorithm to make its accuracy usable. As mentioned in reference [19], assuming that no more than three sensors fail simultaneously among the nine sensors, 84 multi-input multi-output generalized regression neural networks (MIMOGRNNs) should be constructed for fusion diagnosis. In literature [20,21], the structure of the self-associative neural network and the extreme learning machine algorithm are optimized and changed by the genetic algorithm, respectively, to achieve multi-sensor detection, but its detection accuracy is still far from the 95% specified in the literature [22].

In recent years, CNN, as an essential part of deep learning algorithms, has had many applications in different engineering fields. Especially in data-based prediction, CNN shows good performance due to its robust feature extraction ability. Literature [23] proposed a prediction model based on CNN and lion algorithm fusion; the model was improved by niche immunity and finally used in the short-term load prediction of electric vehicles; the optimized prediction model can allow the existence of deformed data and has good prediction performance. Literature [24] proposed an innovative hyperspectral remote sensing image classification method based on CNN. The parameter update of the model adopts an extreme learning machine; finally, the model's accuracy has been verified on the remote sensing image Jiuzhaigou. Literature [25] proposed a prediction model for detecting power theft by combining the bidirectional gated recurrent unit and CNN. Compared with the multilayer perceptron, the long short-term memory network (LSTM), and the single gated recurrent unit, the model performs better in this application scenario. Moreover, in the field of power protection, the author proposes a deep hybrid neural network model that combines CNN and particle swarm optimization (PSO) in literature [26]. The model uses CNN as the feature extractor of the PSO algorithm, which aims to condense useful feature information in the original time series. Literature [27] proposed a one-dimensional CNN prediction model for the real-time detection of motor faults. The model has a simple structure, low hardware requirements, and a fast detection speed. Literature [28] uses CNN to predict the sound source of plate-like structures. Compared with stacked autoencoders, CNN can accept more information input and has a flexible information input method. In the literature [29], the author converts the vibration sensor signal of the motor into a three-dimensional feature matrix as the input of the CNN, and different sensors occupy different channel dimensions in the feature matrix. This paper provides ideas for the situation in which there are different kinds of feature data as CNN input. In the literature [30], the author used the output of the virtual sensor of the aero-engine physical model and the Kalman filter as the input of a one-dimensional CNN and realized the life prediction of the aero-engine. Compared with feed-forward neural networks (FNN) and LSTM, this method has smaller variance and higher accuracy in multiple predictions. In literature [31], the object detection deep learning algorithm YOLOv3 based on CNN is used. The prediction model built by this algorithm realizes the detection of the number of people in an air-conditioned room and reduces the energy consumption of the air conditioning. Literature [32] built a predictive model for power transformer and cyberattack fault diagnosis.

This paper proposes a convolutional neural network based on the inception block; the Inception-CNN uses the convolution layer and pooling layer to extract the analytical redundancy information between each sensor. Next, this information is used to adjust network parameters and predict the probability of individual sensor failures.

The innovations of this paper are as follows.

(1) The construction of the forecasting model

Compared to the various signal-based methods mentioned above, Inception-CNN can take advantage of the characteristics of the inception block to incorporate more interphase information of sensors of different scales into the fault detection model, which is undoubtedly beneficial to the signal-based method. In addition, the method is based on the self-iterative characteristics of neural networks; thus, it does not require a tedious parameter optimization process, and can focus more on the construction of the network model itself.

(2) Generation of fault datasets

Actual aero-engine sensor failure data are very sparse, which contradicts the large amount of training data required for deep learning algorithms. This paper generates sufficient data for training by the Monte Carlo simulation method.

(3) The optimization of the forecasting model input

In this paper, the one-dimensional data of the sensor are converted into a two-dimensional feature matrix by the pan and regroup method. This method makes the extraction of sensor phase information more sufficient, and the prediction model's accuracy is improved.

(4) The optimization of the forecasting model details

The activation function and output mode of the predictive model are optimized so that the model can perform fault detection on all target sensors simultaneously, and the accuracy of the model is improved.

The rest of the paper is organized as follows: Section 2 introduces the research object of the prediction model in this paper and the construction of training data and validation data. Section 3 shows the various parts of the CNN and how they work and then shows the specific architecture of the Inception-CNN prediction model. Section 4 verifies the feasibility and effectiveness of Inception-CNN in terms of fault detection effect and FDI process, respectively. Section 5 concludes this paper.

2. Data Preparation

This paper takes a Geared Turbofan Engine (GTF) as the research object, which can be expressed as Equation (1):

$$\left[x_s^{(t)}, x_v^{(t)}\right] = F\left(w^{(t)}, \theta^{(t)}\right) \tag{1}$$

Among them, w is the operating condition, including height (alt), flight speed (Mach), power lever angle (PLA), etc. θ is a health parameter. x_s is a measurable physical property, including the values of each sensor, and x_v is an unobservable physical quantity.

Due to the limitation of calculation conditions, this paper only studies the ground operation of an aero-engine, namely height = 0 and speed = 0. Ten sensors are selected as the research object of this paper, as shown in Equation (2), and the cross-section position of the aero-engine is shown in Figure 1.

$$x_{sensor} = \left[\ \text{NL, NH, Pt25, Ps3, Pt5, Pt7, Tt25, Tt36, Tt45, Tt5}\ \right] \tag{2}$$

In this paper, the method is tested and verified by the GTF model of NASA/T-MATS master [33] in the simulation environment of Matlab/Simulink2021a.

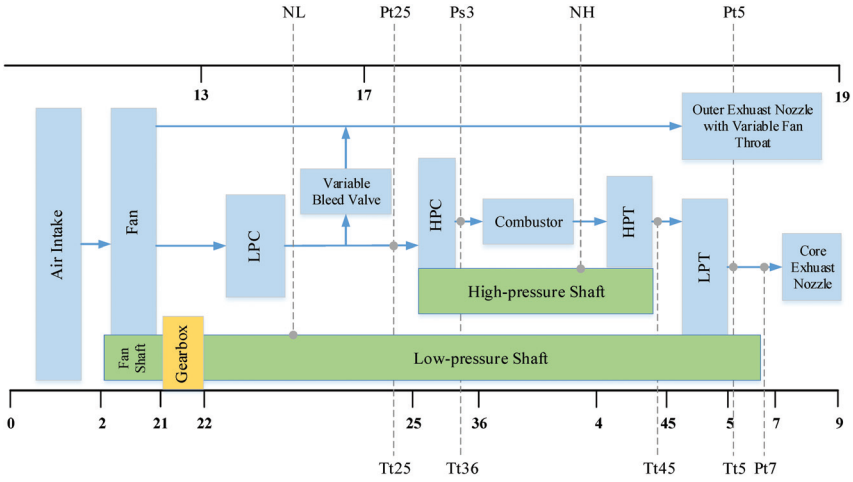


Figure 1. Schematic diagram of the engine structure and the position of the sensors used for detection. Abbreviations: low-pressure compressor (LPC), high-pressure compressor (HPC), high-pressure turbine (HPT), low-pressure turbine (LPT).

By referring to the verification method of the civil turbofan engine fault diagnosis system proposed by Donald et al. [34,35], the Monte Carlo simulation method is used to generate the training and verification data set of the neural network. Firstly, the power lever angle (PLA) of the healthy running engine model is given, as shown in Figure 2, and sensor data x_{sensor} are derived. Then, ten sensor data x_{sensor} are randomly disturbed by the normal distribution, as shown in Equation (3).

$$y^i_{sensor} = x^i_{sensor} + \alpha^i X \qquad X \sim N(0,1) \tag{3}$$

The coefficient α^i is determined by the average $Mean\left(x^i_{sensor}\right)$ of the input sensor value.

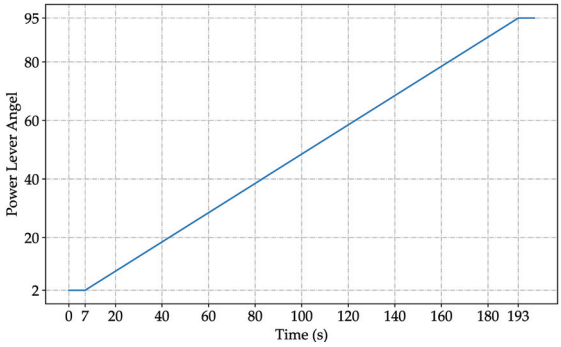


Figure 2. The PLA used by the model to generate the data.

In the training data set and validation data set, the data point that exceeds 5% of the original value of the sensor is set as the fault point. For example, if $\left|\left(y^i_{sensor}\right)_j - \left(x^i_{sensor}\right)_j\right| > 0.05 \times \left(x^i_{sensor}\right)_j$, sensor i is considered a failure at j data points, whereas sensor i is considered to have no failure. Since class imbalance can adversely affect network training, this paper adjusts α^i to reduce the gap between faulty and healthy data points and loops to generate data multiple times. Since the probability of failure of more than three sensors at the same time is extremely small, and the redundant information between the sensors decreases with the increase in the number of simultaneously failed sensors,

the data points where more than three sensors fail at the same time $\sum_{i=1}^{10} \varphi\left[\left(x_{sensor}^{i}\right)_{j}\right] > 3$ are deleted. The scatter plot of a small part of the value of NL, NH, Pt25, and Tt25 sensors after adding disturbance is shown in Figure 3, where the abscissa represents the label of the data points.

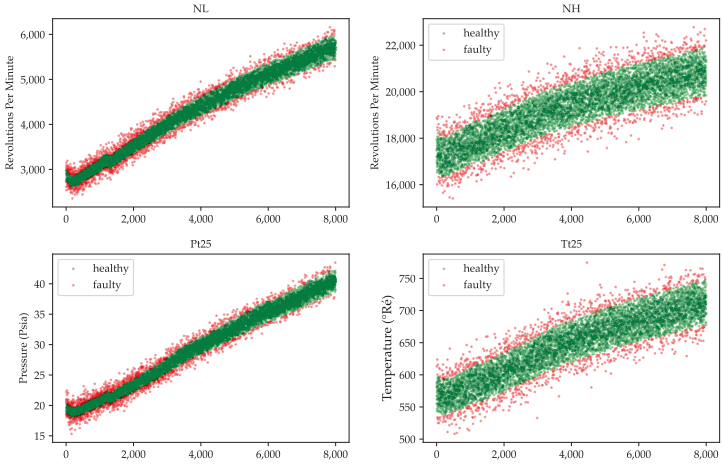


Figure 3. Scatter plot of some sensor data after adding perturbation.

In the final method verification link, we adjust the coefficient α^{i} to ensure that the sensor data y_{sensor} are equivalent to the actual engine data collected and filtered, add the typical faults shown in Table 1, and change the PLA to verify the neural network's generalization performance.

Table 1. Typical sensor failure.

Type of Failure	Cause of Failure
Impulse	Random disturbances, surges, sparks, etc.
Drift	Sensor component deterioration, etc.
Bias	Bias current, bias voltage, etc.

3. Introduction to Inception-CNN

3.1. Introduction to CNN

As a transformed form of multilayer perceptron, the convolutional neural network (CNN) was developed based on studies on the visual cortex of cats [36]. It was initially applied in the field of image recognition. Now, CNN has become a hot spot in many research fields. A typical CNN consists of convolutional layers, pooling layers, and fully connected layers, as shown in Figure 4.

Convolutional layers extract local features of the input data by a perceptual structure and reduce the number of parameters of the CNN by sharing weights. The pooling layer merges adjacent data into a single datum, reduces the dimensionality of the data, speeds up the calculation, and prevents the overfitting of the parameters. The fully connected layer is the basic unit of the BP neural network, which generates output based on the feature data extracted by the previous layer.

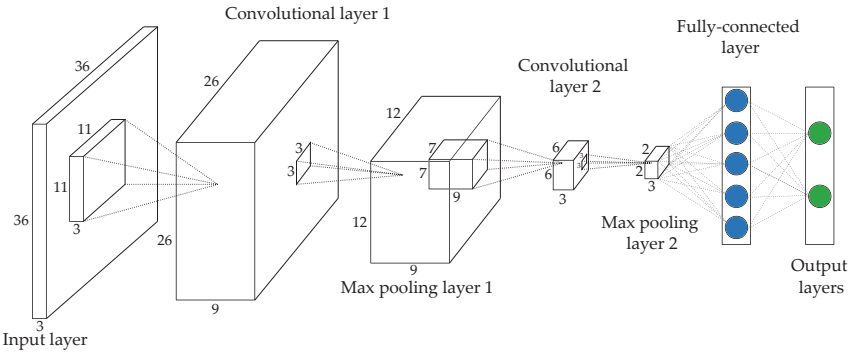


Figure 4. Schematic diagram of a typical convolutional neural network.

3.2. The Basic Module of CNN

The input of each node of the convolutional layer is only the local features of the previous layer, and the convolution kernel is used to convert the subnode matrix of the current layer into a unit node matrix with unlimited channel dimensions in the next layer. Under normal circumstances, the convolution layer generally only converts the channel dimension of the input data and adopts the method of edge zero-padding to ensure that the length and width of the input data remain unchanged. For example, the convolution kernel transforms the matrix of $m_1 \times n_1 \times k_1$ into $1 \times 1 \times k_2$ as Equation (4).

$$g(i) = f\left(\Sigma_{x=1}^{m_1}\Sigma_{y=1}^{n_1}\Sigma_{z=1}^{k_1} a_{x,y,z} \times w_{x,y,z}^{i} + b^{i}\right), 0 < i \leqslant k_2 \tag{4}$$

Among them, $g(i)$ represents the value of the ith identity matrix of this node, and $a_{x,y,z}$ represents a certain length and width of the convolution kernel sampling in the input matrix of this node. The interception matrix with the channel dimension (x, y, z). $w_{x,y,z}^{i}$ represents the convolution kernel weight value matrix corresponding to the ith unit matrix, b^{i} represents the bias value corresponding to the ith unit matrix, and f represents the activation function. In this paper, Scaled Exponential Linear Units (SELU) is selected as the activation function, which normalizes the data distribution and ensures that the gradient will not explode or disappear during the training process. The SELU activation function can be expressed as Equation (5):

$$selu(z) = \lambda \begin{cases} z \; z > 0 \\ \alpha e^{z} - \alpha \; z \leqslant 0 \end{cases} \tag{5}$$

where z represents the output value of the convolution operation, and λ and α are constants, $\lambda \cong 1.051$, $\alpha \cong 1.673$.

Pooling layers sample the data by sliding the pooling kernel. Unlike convolutional layers, pooling layers generally only change the length and width of the input matrix. The most common types of pooling layers are max pooling and average pooling. The process of converting the $m_1 \times n_1 \times k_1$ matrix to $1 \times 1 \times k_1$ by max layer pooling is as Equation (6):

$$g(i) = Subsampling\left(a_{x,y}^{i}\right), \quad 0 < x \leqslant m_1, 0 < y \leqslant n_1, 0 < i \leqslant k_1 \tag{6}$$

Among them, $g(i)$ represents the value of the ith unit matrix of this node, and $a_{x,y}^{i}$ represents the interception matrix of a length and width (x, y) sampled by the pooling kernel when the input matrix of this node corresponds to the ith unit matrix. The pooling layer reduces the data size, speeds up computation, and prevents parameter overfitting without losing data features.

Finally, the network expands the feature matrix extracted by the convolutional and pooling layers into a one-dimensional vector, which is input to the fully connected layer and produces an output based on these features.

3.3. The Calculation Process of CNN

Minimizing the loss function $L(W, b)$ as much as possible is the goal of neural network training, where W and b represent the weights and bias parameters in the neural network, respectively. The loss function consists of two parts: the first part is the residual between the output value and the expected value, and the second part is the regularization loss caused by overfitting, which is regulated by the parameter θ. The loss function can be expressed as Equation (7):

$$L(W, b) = E(W, b) + \frac{\theta}{2} W^T W \tag{7}$$

This paper uses the Stochastic Gradient Descent plus Momentum (SGDm) as the solver to update the CNN parameters. Through the back-propagation of the loss function, the trainable parameters of each layer in the CNN can be updated layer by layer. The mathematical representations are as Equations (8) and (9):

$$W_i = W_{i-1} + m_i, m_i = \left(\eta \frac{\partial L(W_i, b_i)}{\partial W_i} + \beta m_{i-1} \right) \tag{8}$$

$$b_i = b_{i-1} + m_i, m_i = \left(\eta \frac{\partial L(W_i, b_i)}{\partial b_i} + \beta m_{i-1} \right) \tag{9}$$

where η is the learning rate, and β is the momentum coefficient.

3.4. Inception Block

Although a deep neural network can be abstractly considered the superposition of several neural networks, different network architectures and choices of hyperparameters will make a huge difference in the performance of deep neural networks. Furthermore, with the deepening of the network, the interpretability of the deep neural network becomes worse. A deep neural network with good performance often needs to explore the design intuition formed in this field for many years, so it is vital to master and use the previous exploration results.

The inception block is the basic module of the GoogLeNet network architecture proposed in 2014 [37], and GoogLeNet prevailed in the ImageNet competition that year.

Earlier convolutional neural networks are often series architectures. The size of the convolution kernel is also very different, such as from 1×1 to 11×11, etc. However, the inception block parallels convolutional layers with convolution kernels of different sizes in order to have the ability to extract correlated features at different scales. It has been shown that using convolution kernels of different sizes is beneficial for feature extraction.

In addition, deeper neural networks are often helpful to improve prediction accuracy. However, deep networks also bring the risk of gradient disappearance, which will cause the model to fail to converge eventually. Furthermore, due to the exponential increase in computing parameters in deeper networks, the demands on computing equipment are further increased. The inception block can reduce the network computing parameters effectively.

As shown in Figure 5, the inception block consists of four parallel paths. The first three paths use convolutional layers with kernel sizes 1×1, 3×3, and 5×5 to extract information from different spatial sizes. The two paths in the middle first perform 1×1-convolution on the input to reduce the number of channels and reduce the complexity of the model. The fourth path uses a pooling layer with a pooling kernel size 3×3 and then uses a 1×1 convolution kernel to vary the number of channels. All four paths use appropriate padding to match the height and width of the input and output and, finally, join the output of each line in the channel dimension and form the output of the inception block. In the inception block, the adjusted hyperparameters are mainly the number of output channels per layer.

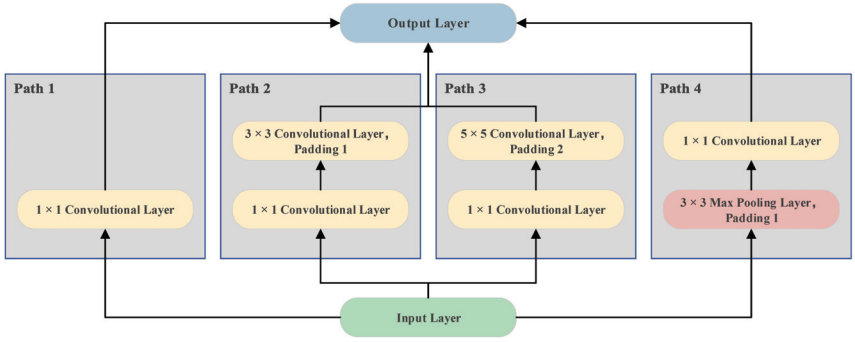


Figure 5. Schematic diagram of the inception block.

The specific representation of the first inception block used in this paper is shown in Figure 6. The input is an 10×10 matrix with a channel number of 48, which passes through four paths, respectively. After the different convolution pooling operations, which are shown in Figure 5, the number of channels is changed and merged into the output matrix in the channel dimension, as the following input to one layer of the neural network.

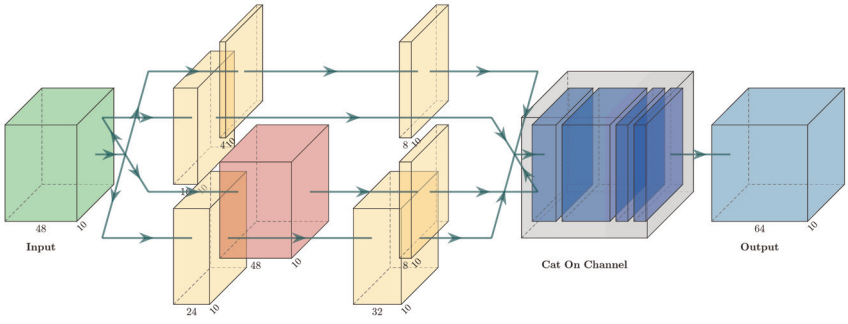


Figure 6. Concrete representation of an inception block.

3.5. Architecture of Inception-CNN

Considering that the convolution kernel has the characteristics of the local receptive field, for one-dimensional data, it is difficult for the convolution kernel to extract the interphase features of the data with a large separation distance. In order to fully extract the relevant features between sensor signals, this paper performs two-dimensional expansion and recombination of the original data before the data are input into the neural network. That is, the size of the data $1 \times n$ is expanded to $n \times n$.

As shown in Figure 7, first, through the method described in Section 1, a one-dimensional sensor data set is obtained by collecting data from the engine model. Different colored and numbered squares represent data from different sensors. Then, the one-dimensional data group is expanded into a two-dimensional data group while rearranging the data through the data translation operation. For example, the second row of the two-dimensional data group in the figure is obtained by shifting the original one-dimensional data group by two units.

Suppose that a convolution kernel of size three is used to extract the interphase features of the sensor data. In this case, it is impossible to extract the interphase features of the sensor data with an interval greater than 3 in the original one-dimensional data set. However, it can be done in the two-dimensional data set after translation expansion.

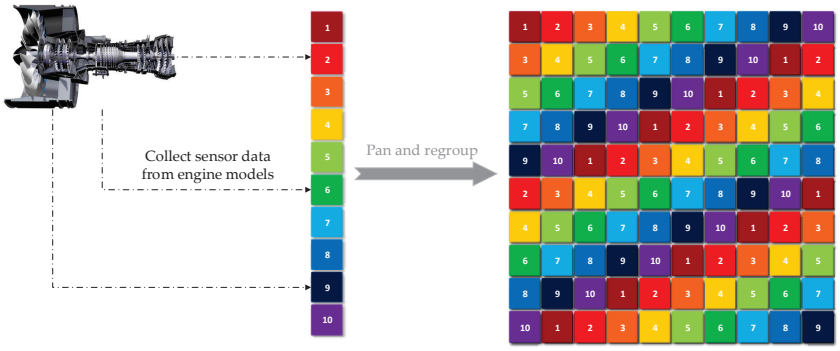


Figure 7. Expansion and recombination of sensor data.

Since the neural network model requires a large amount of data for parameter updating, this paper adopts the operating mode of offline training and online detection. The Inception-CNN architecture is shown in Figure 8. A large amount of perturbed and labeled sensor data is used as the network input during offline training, and the network parameters are updated. During online diagnosis, the data format is kept unchanged, the sensor data that match the actual situation are input to the network, and the prediction is made based on the updated parameters during offline training.

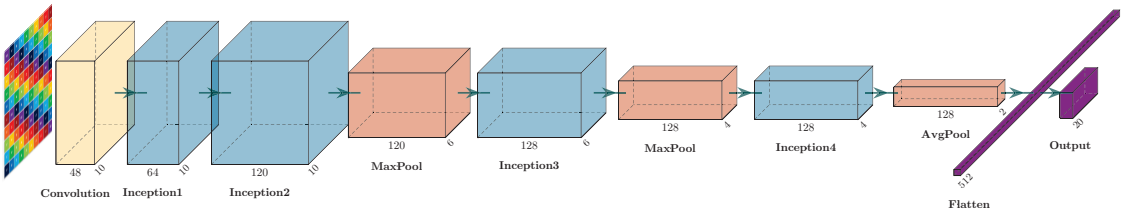


Figure 8. The overall architecture of Inception-CNN.

As shown in Figure 8, the two-dimensional sensor data set expanded by translation is input into the network and passes through one convolution layer, four inception blocks, two max pooling layers, and one average pooling layer. The first convolutional layer has 48 convolution kernels of size 3×3, which converts the original data of size $1 \times 10 \times 10$ into a feature response matrix of size $48 \times 10 \times 10$. The pooling kernel size of all pooling layers in the network is 2×2. After the last average pooling, the three-dimensional matrix is converted to a one-dimensional vector with no information loss. Finally, the extracted feature response vector is input to the fully connected layer and produces an output with 20 nodes.

The parameters of the four inception block convolution kernels and pooling kernels are shown in Figure 5. The changed hyperparameters are only the number of channels output by each layer. The channel parameters are shown in Table 2.

Table 2. Channel parameters for each inception block.

Serial Number	Path 1	Path 2	Path 3	Path 4	Number of Output Channels
1	16	24.32	4.8	48.8	64
2	32	32.48	8.24	64.16	120
3	48	24.52	4.12	120.16	128
4	40	28.56	6.16	128.16	128

3.6. Loss Function

In the traditional multiclassification problem using neural networks, if the classification category is n, it is necessary to construct an n-dimensional one-hot encoding vector to represent the category, and network output nodes are required. For example, in the three-classification problem, the three one-hot vectors $[1,0,0]$ $[0,1,0]$ $[0,0,1]$ represent three different categories. It can be seen that one element in the vector is 1, and the others are 0.

$$y_i = \frac{e^{a_i}}{\sum_{k=1}^{C} e^{a_k}} \quad \forall i \in 1 \cdots C \tag{10}$$

Equation (10) is the softmax function expression, where a is the output value of the last fully connected layer. By calculating the natural logarithm, the softmax function converts the network's output into the probability of each category on the one-hot encoding, i.e., $[p1, p2, p3]$. Among them, $p1 + p2 + p3 = 1$. The softmax function is widely used in multi-classification problems due to its fit with one-hot encoding and its ability to widen the difference between categories in most cases.

However, in the sensor fault detection problem in this paper, due to the assumption that multiple sensors fail simultaneously, the detection vector output by the network will present multiple 1s at the same time. For example, $[1,0,1,0,0,0,0,0,0,0]$ means that the NL sensor and the Pt25 sensor fail simultaneously, so the softmax function is no longer applicable in the network, and the sensor failure probability is only calculated during model validation.

For multi-classification problems, a combination of the softmax function and cross-entropy loss function is often used. The expression of the cross-entropy loss function is as Equation (11).

$$E(W,b) = -\frac{1}{N}\sum_{i}\sum_{c=1}^{M} y_{ic}\log(p_{ic}) \tag{11}$$

Among them, M is the number of categories, y is the symbolic function, i is the batch data size, and p is the predicted probability of the category, which is the output of softmax. Since softmax is no longer applicable, the output of the fully connected layer a is replaced with p.

There are two types of individual sensors, faulty and healthy, and one-hot encoding is constructed for each sensor separately, so the number of final output nodes of Inception-CNN is 10×2, a total of 20 output nodes. Every two nodes constitute a prediction unit of the sensor. The output value of the cross-entropy loss function of each prediction unit is accumulated as a total loss value for the back-propagation of the network.

In summary, the loss function expression used in this paper is as Equation (12):

$$E(W,b) = \sum_{j=1}^{10}\left(-\frac{1}{N}\sum_{i}\sum_{c=1}^{M} y_{ic}\log(a_{ic})\right) \tag{12}$$

where j is the number of the prediction unit, and a is the output of the last fully connected layer of the network.

4. Validation and Analysis

4.1. Training the Networks

The Inception-CNN and other comparative neural networks in this paper are trained in the simulation environment of Python3.9/Pytorch1.10.1. The rest of the non-neural network machine learning algorithms are trained in the simulation environment of Matlab2021a. For the data set constructed in Section 1, the accuracy of some neural networks during the training process is shown in Figure 9, and the accuracy of each algorithm when it converges is shown in Table 3:

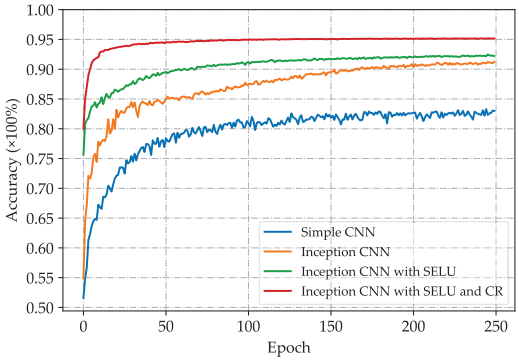


Figure 9. Accuracy changes during some neural networks' training.

Table 3. Performance comparison of Inception-CNN and other algorithms.

Algorithm	Activation Function	Loss Function	Accuracy at Convergence
Gaussian Naive Bayes	\	\	67.81%
Cubic SVM	\	\	73.35%
Cosine KNN	\	\	75.54%
Gaussian SVM	\	\	78.14%
BP Neural Networks	ReLU (Rectified Linear Unit)	MSE	82.72%
Simple 4-Tier CNN	ReLU	MSE	85.41%
Inception-CNN	ReLU	MSE	91.82%
Inception-CNN	SELU	MSE	92.13%
Inception-CNN	SELU	Cross-Entropy	95.41%

Each neural network sets the same training hyperparameters. The learning rate is 0.0001, and the batch size is 256.

It can be seen from Table 3 that the accuracy of the neural network is significantly improved compared to other non-neural network algorithms. For example, a simple BP neural network improves the accuracy by 4.58% compared to the best-performing Gaussian SVM in a non-neural network. Compared with the simple BP neural network, the accuracy of the simple four-layer CNN network is increased by 2.69%, so the CNN fits this data set better than other algorithms; thus, the improved CNN was chosen as the fault detection algorithm in this paper.

Figure 9 takes Epoch as the abscissa and the ordinate as the accuracy of each network on the test data set. The first three CNN networks use the mean square error (MSE) loss function, and the fourth network uses the improved cross-entropy loss function in the paper. It can be seen from Figure 9 and Table 3 that the CNN based on the inception block has a rapid increase in accuracy and a higher upper limit than the simple four-layer CNN. After replacing the traditional ReLU activation function with the SELU activation function, the convergence time of the network is reduced, and the training time is reduced by approximately 60%. After replacing MSE with the improved cross-entropy loss function in this paper, the network's sensitivity to wrong predictions is improved. The accuracy rate rises faster, and the accuracy during convergence is 3.28% higher than that of Inception-CNN using the MSE loss function. The final accuracy of Inception-CNN used in this paper is 95.41%.

4.2. Typical Sensor Failure Validation

The typical sensor fault verification link is divided into two parts. On the one hand, it tests the detection effect of Inception-CNN itself for typical faults, and on the other hand, it verifies the effectiveness of the network in the FDI process.

Excellent generalization performance is an important indicator to measure the performance of neural network models. For the research content of this paper, even if the data set for training Inception-CNN is only composed of a simple slope response after adding disturbance, it is still expected to have good prediction results for all ground operating conditions of the engine model.

Therefore, in a typical sensor failure verification process, the steady-state or transition-state conditions used for each verification are different.

4.2.1. Validation of Fault Detection Effect

First, we verify the case of a single fault and simultaneous multiple faults in the steady state of the engine.

A single sensor failure occurs at a steady state, as shown in Figure 10. Under different steady-state conditions, the impulse, drift, and bias fault are added to the NL, Pt25, and Tt36 sensor, respectively. It can be seen from the figure that when the fault probability is 80% as the threshold, the injected impulse fault, drift fault, and bias fault are identified by the fault detection system after delays of around 0.1 s, 4.2 s, and 0 s, respectively. The drift fault has a relatively long detection delay due to the small offset in the early stage when the fault occurs. In addition, after injecting faults, the failure probabilities of other sensors without faults are all below 6% and fluctuate much less than the threshold.

As described in Section 1, when constructing the neural network training data, this paper removes data points where more than three sensors fail simultaneously. Thus, the simultaneous failure of up to three sensors is the assumption of this paper.

The simultaneous failure of multiple sensors in a steady state is shown in Figure 11. Under a specific steady-state condition, the bias, impulse, and drift fault are injected into the NH, Ps3, and Tt25 sensor at approximately 5 s, 6 s, and 9 s, respectively. The detection delay of each fault is around 0 s, 1.6 s, and 2.5 s, respectively. Among them, the delay of drift fault detection is still caused by insufficient offset in the early stage of fault occurrence. As can be seen from the figure on the right, although multiple faults occur simultaneously, it does not affect the accurate identification of a single fault by the fault detection system, and the detection delay is hardly affected by multi-sensor faults. Among the non-faulty sensors, the failure probability of the Pt25 sensor fluctuates up to approximately 19%, but it is still far below the failure probability threshold.

Next, we verify the case of a single fault and simultaneous multiple faults in the transition state of the engine.

A single sensor failure occurs at a transition state, as shown in Figure 12. Under different transition state conditions, the impulse, drift, and bias fault were injected into the NH, Pt5, and Tt45 sensor, respectively. The situation is similar to a single sensor failure in a steady state. The detection system identifies each fault after delaying by 0.1 s, 2.6 s, and 0 s, respectively. It can be seen that the performance of Inception-CNN in the transition state and steady state is comparable.

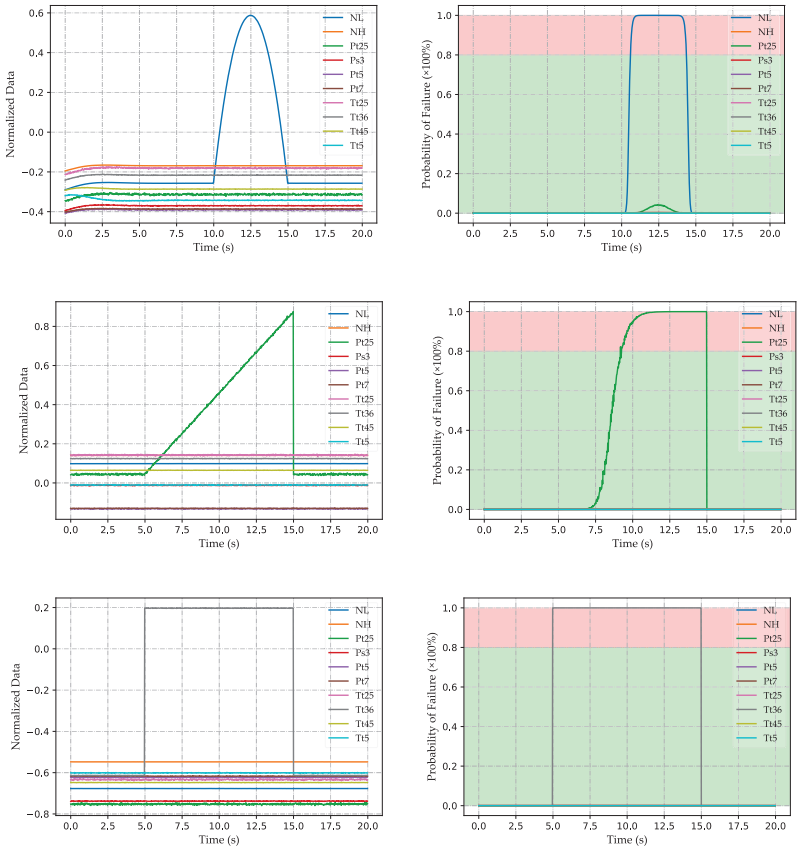


Figure 10. Detection of single sensor failure in steady state.

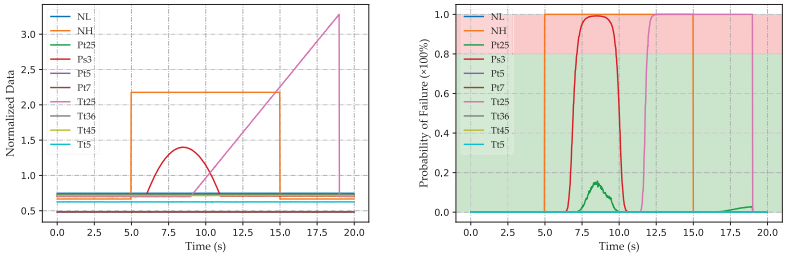


Figure 11. Detection of simultaneous failure of multiple sensors in steady state.

The simultaneous failure of multiple sensors in a transition state is shown in Figure 13. Under a specific steady-state condition, the drift, impulse, and bias fault are injected into the NL, Pt7, and Tt5 sensor at approximately 5 s, 3 s, and 9 s, respectively. The detection delay of each fault is approximately 5.9 s, 0.2 s, and 0 s, respectively. The delay in detecting drift faults is due to the fact that the injected drift fault offset is small, and the drift fault itself has the characteristic of slowly increasing the offset.

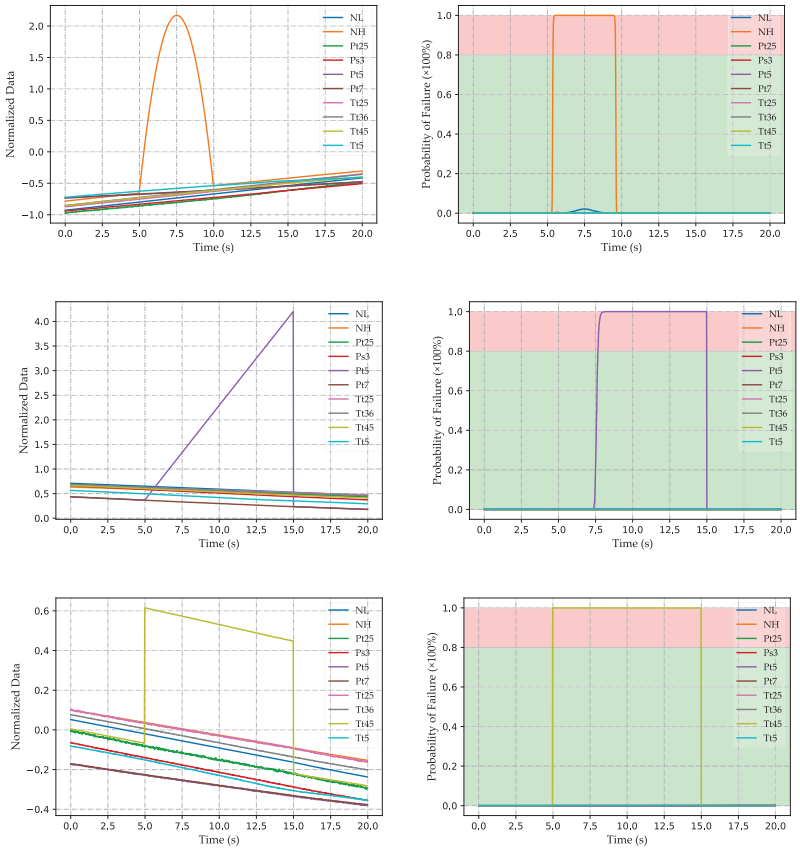


Figure 12. Detection of single sensor failure in transition state.

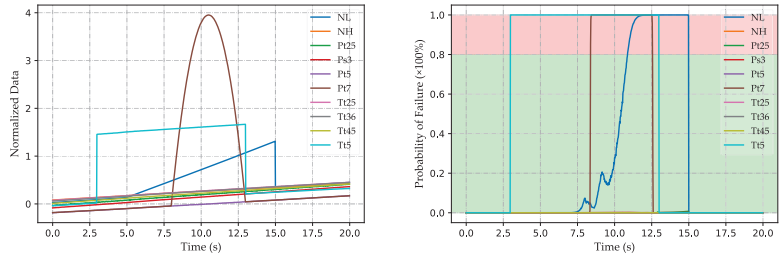


Figure 13. Detection of simultaneous failure of multiple sensors in transition state

4.2.2. Validation in the FDI Process

As shown in Figure 14, the FDI process consists of a fault detection module (FDM) and a fault isolate module (FIM). The FDM contains fault detection estimators (FDE) running in real time, and the FIM contains fault isolate estimators (FIE) that require the output of the FDM to activate. The essence of each estimator is a method of fault detection or isolation. The Inception-CNN proposed in this paper is an FDE in the FDI process.

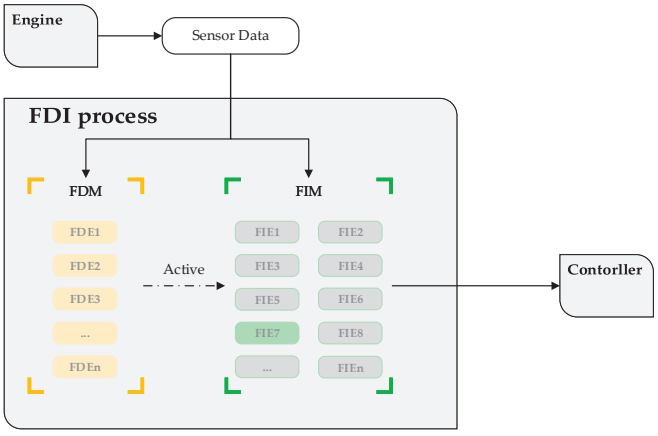


Figure 14. Schematic diagram of FDI process.

The Gaussian process regression algorithm is used to build the FIE module.

Due to space limitations, the FDI process verification is only carried out for the simultaneous failure of multiple sensors. First, we verify the FDI flow of the engine at a steady state.

As shown in Figure 15, when multiple sensors fail simultaneously in a steady state, FDM effectively activates FIM through the built-in Inception-CNN. After the FDI process, the injected fault offsets of drift, impulse, and bias faults are reduced by around 81%, 61%, and 100%, respectively, which correspond to different faults' detection delays. The lower impulse fault offset reduction percentage is the lower injected impulse fault offset.

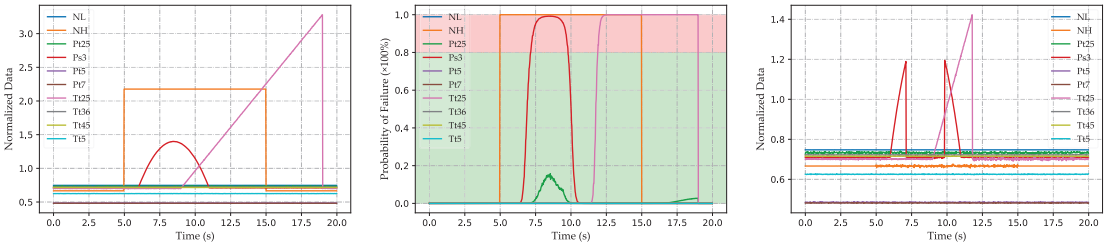


Figure 15. Validation of the FDI process with simultaneous multi-sensor failures in steady state.

When multiple sensors fail simultaneously in the transition state, it is similar to the steady state, as shown in Figure 16. After the FDI process, the drift, impulse, and bias fault offsets were reduced by 76%, 56%, and 100%, respectively. The effectiveness of Inception-CNN in the FDI process is verified.

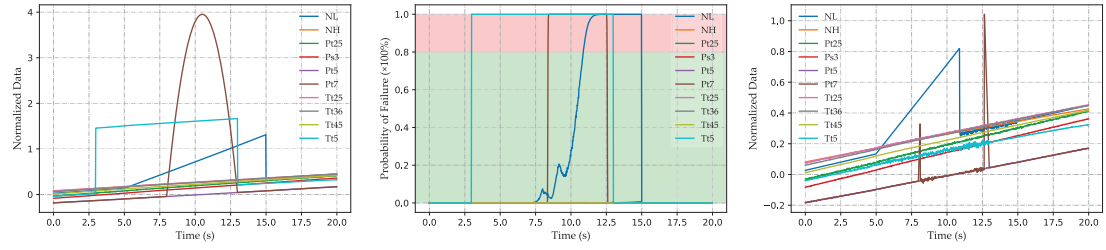


Figure 16. Validation of the FDI process with simultaneous multi-sensor failures in transition state.

5. Conclusions

In this paper, a convolutional neural network based on an inception block is proposed and utilized for aero-engine sensor fault detection. The traditional detection unit composed of multiple fusion algorithms is simplified into one detection algorithm, which mainly solves the problems of traditional sensor FDI methods with many parameters and complex systems.

The effectiveness and feasibility of the method are verified by the detection effect and the FDI process. On the data set of this paper, the detection accuracy of Inception-CNN is 95.41%, which improves the prediction accuracy by 17.27% and 12.69% compared with the best-performing non-neural network algorithm and simple BP neural networks tested in the paper, respectively.

In addition, this paper constructs the training data set and validation data set of the signal-based sensor FDI method through the Monte Carlo simulation method, which solves the problem wherein the experiment cannot be carried out due to insufficient fault data.

The sensor fault detection method based on Inception-CNN proposed in this paper is a data-driven algorithm. Its accuracy and applicability are positively related to the quality of the data. Thus, in the future, this method can be combined with the mechanism study of the engine to improve the algorithm's performance through higher-quality data. In addition, the research content of this paper can be combined with the research on the safety control strategy of aero-engines based on sensor value judgment. Taking the research [38] of Cao et al. as an example, if FDI is used as the front module of the safety protection control module in the control loop, the possibility of control strategy failure due to sensor failure can be reduced, and the robustness of the system can be improved.

Author Contributions: Conceptualization, X.D. and J.W.; Data curation, X.D.; Investigation, X.D.; Methodology, X.D., J.C. and J.W.; Project administration, X.D.; Resources, X.D., H.Z. and J.W.; Software, X.D.; Supervision, J.W.; Validation, X.D.; Visualization, X.D.; Writing—original draft, X.D. and J.C.; Writing—review and editing, X.D., J.C., H.Z. and J.W. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the Central Military Commission Foundation to Strengthen Program Technology Fund (No. 2019-JCJQ-JJ-347); Central Military Commission Special Fund for Defense Science, Technology and Innovation (20-163-00-TS-009-096-01); Aviation Science Fund of China-Xi'an 631 Research Institute (No. 201919052001); Kunpeng Talent Plan of Zhejiang Province (E10938DL03); and Startup Foundation of Ningbo Institute of Materials Technology and Engineering, CAS (E10921RA06).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available on request from the first author or corresponding author.

Acknowledgments: The authors would like to acknowledge the open-source developers of the deep learning framework and graphing tools used in this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Garg, S. Controls and Health Management Technologies for Intelligent Aerospace Propulsion Systems. In Proceedings of the 42nd AIAA Aerospace Sciences Meeting and Exhibit, Reno, Nevada, 5–8 January 2004; American Institute of Aeronautics and Astronautics: Reno, Nevada, 2004. [CrossRef]
2. Jaw, L.; Mattingly, J. *Aircraft Engine Controls*; American Institute of Aeronautics and Astronautics, Inc.: Washington, DC, USA, 2009. [CrossRef]
3. He, B.; Yu, D.; Shi, X. Application of simulation model of sensors to analysis fault of control system of turbojet engine. *J. Propuls. Technol.* **2001**, *5*, 364–367. [CrossRef]

4. Giantomassi, A.; Ferracuti, F.; Iarlori, S.; Ippoliti, G.; Longhi, S. Electric Motor Fault Detection and Diagnosis by Kernel Density Estimation and Kullback-Leibler Divergence Based on Stator Current Measurements. *IEEE Trans. Ind. Electron*. **2015**, 62, 1770–1780. [CrossRef]
5. Dai, X.; Gao, Z. From Model, Signal to Knowledge: A Data-Driven Perspective of Fault Detection and Diagnosis. *IEEE Trans. Ind. Inf.* **2013**, 9, 2226–2238. [CrossRef]
6. Yuan, C.; Yao, H. Aero-engine adaptive model re-construction under sensor failure. *J. Aerosp. Power* **2006**, *21*, 195–199. [CrossRef]
7. Wei, X.; Yingqing, G. Multiple Sensors Soft Failure Diagnosis for Aircraft Engine Control System. *Comput. Meas. Control* **2007**, *15*, 585–587. [CrossRef]
8. Zhao, C.; Ye, Z.; Wang, J.; Yin, B. Rapid prototype real-time simulation of turbo-fan engine sensor fault diagnosis. *J. Aerosp. Power* **2014**, *29*, 451–457. [CrossRef]
9. Han, B.; Gou, L.; Mao, N.; Wang, X. A method of multiple fault diagnosis based on fault matching. *Aeronaut. Comput. Tech.* **2015**, *45*, 96–99.
10. Lu, F.; Chen, Y.; Huang, J.; Zhang, D.; Liu, N. An Integrated Nonlinear Model-Based Approach to Gas Turbine Engine Sensor Fault Diagnostics. *Proc. Inst. Mech. Eng. Part G J. Aerosp. Eng.* **2014**, *228*, 2007–2021. [CrossRef]
11. Tadić, P.; Đurović, Ž. Particle Filtering for Sensor Fault Diagnosis and Identification in Nonlinear Plants. *J. Process Control* **2014**, *24*, 401–409. [CrossRef]
12. Liu, R.; Meng, G.; Yang, B.; Sun, C.; Chen, X. Dislocated Time Series Convolutional Neural Architecture: An Intelligent Fault Diagnosis Approach for Electric Machine. *IEEE Trans. Ind. Inf.* **2017**, *13*, 1310–1320. [CrossRef]
13. Cai, K.; Sun, Y.; Yao, W. Fault diagnosis and adaptive reconfiguration control for sensors in aeroengine. *Electron. Opt. Control* **2009**, *23*, 1118–1126.
14. Duan, S.; Li, Q.; Zhao, Y. Fault diagnosis for sensors of aero-engine based on improved least squares support vector regression. In Proceedings of the 2011 Eighth International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), Shanghai, China, 26–28 July 2011; Volume 3, pp. 1962–1966. [CrossRef]
15. Lu, F.; Huang, J.; Chen, Y.; Song, Y. Research on sensor fault diagnosis of aero-engine based on data fusion of spso-svr. *J. Aerosp. Power* **2009**, *24*, 1856–1865. [CrossRef]
16. Liu, Y.; Yang, D.; Liu, Y.; Kang, J. Failure diagnose research for the plane engine of basic neural nework. *Electron. Des. Eng.* **2012**, *20*, 89–92. [CrossRef]
17. Xian, N. Aero-engine fault diagnosis method based on abc-bp neural network. *Equip. Manuf. Technol.* **2018**, *5*, 173–175.
18. Jianliang, A.; Yang, X. Fault diagnosis of aero-engine based on self-adaptive neural network. *Sci. Sin. Technol.* **2018**, *48*, 326–335.
19. ChangZheng, L.; Zhang, Y. Sensor fault detection based on general regression neural network. *J. Propuls. Technol.* **2017**, *38*, 2130–2137. [CrossRef]
20. Lv, S.; Guo, Y.; Sun, H. Aero-engine sensor data preprocessing based on sdq algorithm of ga-aann neural network. *J. Propuls. Technol.* **2018**, *39*, 1142–1150. [CrossRef]
21. Cui, J.; Liu, H.; Tao, S.; Yu, M.; Gao, Y.; Automation, S. Aeroengine fault diagnosis method based on elm. *Fire Control Command Control* **2018**, *43*, 113–121.
22. Yao, H. *Full Authority Digital Electronic Control System for Aero-Engine*; Aviation Industry Press: Beijing, China, 2014.
23. Li, Y.; Huang, Y.; Zhang, M. Short-Term Load Forecasting for Electric Vehicle Charging Station Based on Niche Immunity Lion Algorithm and Convolutional Neural Network. *Energies* **2018**, *11*, 1253. [CrossRef]
24. Ye, A.; Zhou, X.; Miao, F. Innovative Hyperspectral Image Classification Approach Using Optimized CNN and ELM. *Electronics* **2022**, *11*, 775. [CrossRef]
25. Duarte Soares, L.; de Souza Queiroz, A.; López, G.P.; Carreño-Franco, E.M.; López-Lezama, J.M.; Muñoz-Galeano, N. BiGRU-CNN Neural Network Applied to Electric Energy Theft Detection. *Electronics* **2022**, *11*, 693. [CrossRef]
26. Ullah, A.; Javaid, N.; Samuel, O.; Imran, M.; Shoaib, M. CNN and GRU Based Deep Neural Network for Electricity Theft Detection to Secure Smart Grid. In Proceedings of the 2020 International Wireless Communications and Mobile Computing (IWCMC), Limassol, Cyprus, 15–19 June 2020; IEEE: Limassol, Cyprus, 2020; pp. 1598–1602. [CrossRef]
27. Ince, T.; Kiranyaz, S.; Eren, L.; Askar, M.; Gabbouj, M. Real-Time Motor Fault Detection by 1-D Convolutional Neural Networks. *IEEE Trans. Ind. Electron*. **2016**, *63*, 7067–7075. [CrossRef]
28. Ebrahimkhanlou, A.; Salamone, S. Single-Sensor Acoustic Emission Source Localization in Plate-Like Structures Using Deep Learning. *Aerospace* **2018**, *5*, 50. [CrossRef]
29. Wang, H.; Li, S.; Song, L.; Cui, L.; Wang, P. An Enhanced Intelligent Diagnosis Method Based on Multi-Sensor Image Fusion via Improved Deep Learning Network. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 2648–2657. [CrossRef]
30. Arias Chao, M.; Kulkarni, C.; Goebel, K.; Fink, O. Fusing Physics-Based and Deep Learning Models for Prognostics. *Reliab. Eng. Syst. Saf.* **2022**, *217*, 107961. [CrossRef]
31. Elsisi, M.; Tran, M.-Q.; Mahmoud, K.; Lehtonen, M.; Darwish, M.M.F. Deep Learning-Based Industry 4.0 and Internet of Things towards Effective Energy Management for Smart Buildings. *Sensors* **2021**, *21*, 1038. [CrossRef]
32. Elsisi, M.; Tran, M.; Mahmoud, K.; Mansour, D.-E.A.; Lehtonen, M.; Darwish, M.M.F. Effective IoT-Based Deep Learning Platform for Online Fault Diagnosis of Power Transformers against Cyberattacks and Data Uncertainties. *Measurement* **2022**, 190, 110686. [CrossRef]

33. Chapman, J.W.; Lavelle, T.M. *Toolbox for the Modeling and Analysis of Thermodynamic Systems (T-MATS) User's Guide*; NASA: Washington, DC, USA, 2014; Volume 50. Available online: https://ntrs.nasa.gov/api/citations/20140012486/downloads/20140012486.pdf (accessed on 10 September 2021).
34. Simon, D.L.; Bird, J.; Davison, C.; Volponi, A.; Iverson, R.E. Benchmarking Gas Path Diagnostic Methods: A Public Approach. In *Volume 2: Controls, Diagnostics and Instrumentation; Cycle Innovations; Electric Power*; ASMEDC: Berlin, Germany, 2008; pp. 325–336. [CrossRef]
35. Simon, D.L. *Propulsion Diagnostic Method Evaluation Strategy (ProDiMES) User's Guide*; NASA: Washington, DC, USA, 2010; Volume 47. Available online: https://ntrs.nasa.gov/api/citations/20100005639/downloads/20100005639.pdf (accessed on 15 September 2021).
36. Hubel, D.H.; Wiesel, T.N. Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex. *J. Physiol.* **1962**, *160*, 106–154. [CrossRef] [PubMed]
37. Szegedy, C.; Liu, W.; Ji, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going Deeper with Convolutions. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; IEEE: Boston, MA, USA, 2015; pp. 1–9. [CrossRef]
38. Chen, C.; Wang, M.; Dimirovski, G.M.; Zhao, J. Safety Protection Control for Aeroengines Based on Finite Times of Controller Switches. In Proceedings of the 31st Chinese Control Conference, Hefei, China, 25–27 July 2012; IEEE: Beijing, China, 2012; pp. 2071–2076.

sensors

MDPI

Article

Design of a Chipless RFID Tag to Monitor the Performance of Organic Coatings on Architectural Cladding

Tim Savill * and Eifion Jewell

Materials Research Center, College of Engineering, Swansea University Bay Campus, Swansea SA1 8EN, UK; e.jewell@swansea.ac.uk
* Correspondence: 863047@swansea.ac.uk

Abstract: Coating degradation is a critical issue when steel surfaces are subject to weathering. This paper presents a chipless, passive antenna tag, which can be applied onto organically coated steel. Simulations indicated that changes associated with organic coating degradation, such as the formation of defects and electrolyte uptake, produced changes in the backscattered radar cross section tag response. This may be used to determine the condition of the organic coating. Simulating multiple aging effects simultaneously produced a linear reduction in tag resonant frequency, suggesting coating monitoring and lifetime estimation may be possible via this method. For coatings thinner than calculations would suggest to be optimum, it was found that the simulated response could be improved by the use of a thin substrate between the coated sample and the antenna without vastly affecting results. Experimental results showed that changes to the dielectric properties of the coating through both the uptake of water and chemical degradation were detected through changes in the resonant frequency.

Keywords: chipless RFID; corrosion; organic coating; degradation; sensors

Citation: Savill, T.; Jewell, E. Design of a Chipless RFID Tag to Monitor the Performance of Organic Coatings on Architectural Cladding. *Sensors* **2022**, 22, 3312. https://doi.org/10.3390/s22093312

Academic Editor: Antonio Lázaro

Received: 11 March 2022
Accepted: 23 April 2022
Published: 26 April 2022

1. Introduction

Organic coatings are widely used in the construction industry for aesthetic, corrosion protection, and weather resistance reasons. In fact, in 2020, the EU produced 130 million metric tonnes of hot rolled steel, of which 5 million metric tonnes was organically coated [1]. Despite this, it has been suggested by some studies that the reason for the high cost of corrosion is due, in part, to the poor selection of protective measures [2]. Hence, one widely suggested strategy for corrosion protection is to 'develop advanced life prediction and performance assessment methods and to move to a greater degree of corrosion monitoring [3].

Currently, coatings used on construction panels are rarely monitored in-situ and the expected performance of the overall building envelope is often only estimated from lab-based accelerated corrosion testing. Real time monitoring could allow more accurate estimates of building cladding lifespan, as well as required maintenance schedules, providing the customer with active performance data [3]. A significant amount of emerging research in this field shows the appetite for this technology [4].

An ideal solution would be a wireless, internet of things style sensor system which allows remote, live monitoring of the organic coatings. The oil and gas industry have long been aware of the benefits of corrosion monitoring of pipelines [5]. However, they are currently the only significant industry with commonly used, commercially viable monitoring methods. The construction industry is starting to catch up, although current research suggests monitoring is focused on load critical components, especially reinforced concrete rebar.

Some research has been carried out into coated panel sensors, such as strain gauges [6], corrosion indicating paint [7], electrical resistance probes [8], and numerous EIS (electrical

impedance spectroscopy) modifications [9–17]. These sensors work in small scale testing. However, they suffer from the same drawbacks of practical use. For example, they often either modify the paint system, require connection to the underlying metal substrate, or require modification between the layers of the coating. This makes these systems difficult to implement and replace if necessary, and, in some instances, reduce the effectiveness of the coating itself.

Sensors for monitoring coated steel products have to be capable of detecting the kind of changes in coatings that are indicative of degradation or failure. These include the formation of defects and the associated uptake of oxygen and ions from the environment, coating adhesion loss, electrolyte penetration, or corrosion initiation [18]. Furthermore, degradation can also produce increased porosity of the coating through the action of UV, blistering due to osmotic or electrochemical effects, and decreased coating thickness through erosion and coating mineralization and oxidation [18–21]. Tracking the spread and size of corrosion effects would also be required to allow an indication of the severity and overall condition of the product. Cut edge corrosion, when corrosion begins and propagates from a metal edge exposed during manufacture, is one of the most commonly seen defects for organically coated steel [22].

This paper outlines the basics of a new passive RFID (resonant frequency identification) technique that aims to allow easy performance monitoring of organically coated steel cladding products without some of these drawbacks. A passive RFID sensor presents a number of promising features for organic coating monitoring. As a cheap and 'semi-remote' method, it would allow for a medium range, non-destructive monitoring method of a number of assets. In fact, a significant amount of work has been carried out on developing RFID based sensors [23–25]. The approach taken by the vast majority of the sensors currently being developed is to use a so called 'chipless' RFID design.

Chipless RFID tags were first considered as a concept in order to allow RFID tagging to compete financially with low-cost barcode tagging [25,26]. By removing the integrated circuit, chipless RFID is greatly reduced in cost and allows consideration of fully printable tags, increasing ease and rate of production [26]. Chipless sensors offer a much simpler approach in which some of the complications of impedance matching as well as other complications are negated compared to chipped devices.

As a sensor, they are stated to have better robustness, lower radiated power, and longer life than traditional sensors [25]. Three types of chipless RFID sensors exist, namely time-domain reflectometry (TDR), frequency modulation, and phase encoded sensors [25], and these are summarized in Figure 1. For each of these types, different tag technologies exist, such as surface acoustic wave (SAW) and radar cross section backscatter (RCS).

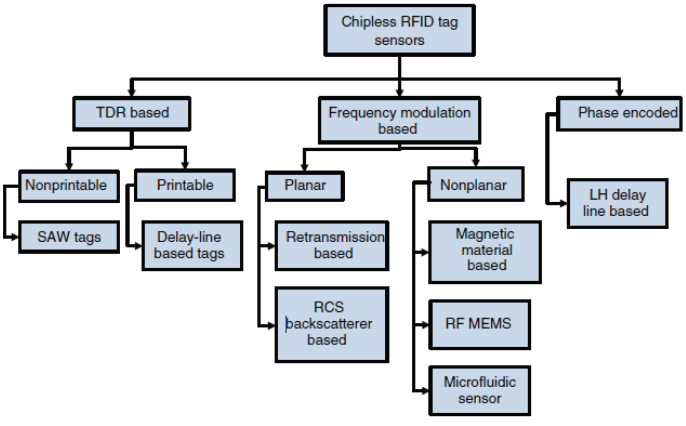


Figure 1. Types of chipless RFID tags (Reprinted with permission from Ref. [25] 2016, John Wiley and Sons).

A recent study involved using a chipless RFID circular microstrip patch antenna (CMPA) to monitor crack growth in aluminium [27]. Microstrip antennas are commonly used in wireless communication, including in high performance situations, such as in satellites and aircraft [28]. They have advantages of low price and profile, simplicity, and versatility [28]. A basic microstrip antenna is composed of a ground plane, which is covered by a dielectric layer with a metal antenna strip on it. A microstrip patch antenna is differentiated because it consists of relatively large sections, or patches, of metal [29]. A CMPA is simply an MPA in which the patch is circular in shape, as shown in Figure 2.

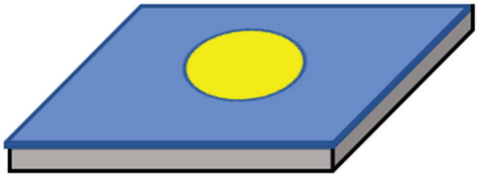


Figure 2. A basic CMPA.

In this paper, a similar approach was used to that in [28], which utilizes radar cross section (RCS) backscatter-based technology, which is based on frequency modulation. The principle is described visually in Figure 3 and involves the use of a microstrip resonator. An interrogation signal is aimed at the chipless tag which will resonate at a unique frequency and hence produce a signature backscattered signal which is received by a further antenna for analysis [25,27].

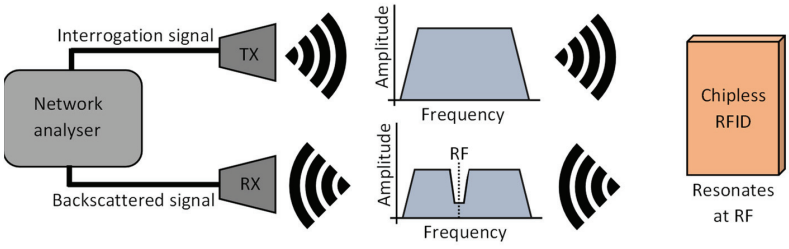


Figure 3. Basic working principle of an RCS backscattered RFID tag.

The monitoring system developed in this paper has parallels to that designed by Marindra [27], which utilised the fact that the resonant frequency of a CMP antenna is affected by changes in the ground plane. This study aimed to use a similar antenna system to monitor changes in the dielectric layer of such a system with the dielectric layer, in this case, being an organic coating layer on a steel product.

As described in [28], the dominant TM^{z}_{mn0} mode for a circular microstrip patch antenna is the TM_{110} mode for which the resonant frequency is defined by:

$$(f_r)_{110} = \frac{1.8412}{2\pi a_e \sqrt{\mu\epsilon}} = \frac{1.8412 v_0}{2\pi a_e \sqrt{\epsilon_r}} \quad (1)$$

where:

$$a_e = a\left\{1 + \frac{2h}{\pi a \epsilon_r}\left[ln\left(\frac{\pi a}{2h}\right) + 1.7726\right]\right\}^{\frac{1}{2}} \quad (2)$$

where a is the diameter of the circular patch, h is the thickness of the dielectric layer, ϵ_r is the relative permittivity of the dielectric layer, v_0 is the speed of light, μ is the permeability of the dielectric, and ϵ is the permittivity of the dielectric layer.

Hence, it can be concluded that:

$$(f_r)_{110} \propto f(\epsilon_r) \tag{3}$$

$$(f_r)_{110} \propto f(h) \tag{4}$$

Therefore, the resonant frequency of a patch antenna is related to a function of both the relative permittivity (dielectric constant) and the height of the dielectric. As a result, any changes to the dielectric properties of an organically coated product could be detected via this approach. This would include degradation due to water ingress, defect formation, or paint formation failures such as chalking. This approach was utilized by [28] to monitor the response of an epoxy coated resonator and was able to detect changes in the dielectric properties of the epoxy coating produced by the absorption and desorption of water. The proposed method differs from this method as the antenna can be placed onto, rather than under, the coating of interest which may interfere with the adhesion and/or protective properties of the coating. Furthermore, as the proposed device is solely reliant on backscatter RCS measurement, a physical wired connection to the antenna, as in [30], is not required simplifying the procedure.

The aim of the investigation was to develop the concept outlined, in order to prove its applicability to architecturally painted steel. The strategy employed used a combination of simulation and laboratory experimental techniques. The simulation provided a means of estimating the response change when variations in material properties which could be expected during degradation occurred within a purely simulation space. The simulated response could then be compared and validated using the laboratory derived experimental results. This parallel approach ensured that a greater understanding of underlying physics could be established while also proving the applicability in the real world.

2. Methodology

2.1. Simulation

The tag was designed and tested in CST (Computer Simulation Technology) Design Environment 2019, a 3D electromagnetic analysis software commonly used for antenna design and analysis [31]. Two approaches, the basic design of each shown in Figure 4, were considered. The first approach (NS) solely used the organic coating as the dielectric layer, and the sensor patch is attached to this, while the second approach (S) used a dielectric substrate between the sensor patch and the coated steel panel.

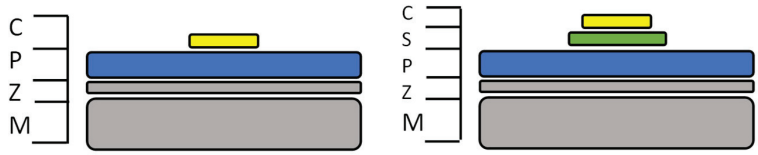


Figure 4. Two tag systems NS, no substrate, (**left**) and S, substrate, (**right**) showing the copper sensor (C), the substrate layer (S), the paint system (P) the zinc galvanized layer (Z) and the steel metal substrate layer (M).

Two coating systems were also considered and designed in the software based on two commonly used coated steel products composed of a steel substrate, a zinc metal coating layer, and a dielectric organic coating layer. This is referred to as coated steel in the following work. A polyvinyl chloride (PVC)-based coating with a total nominal coating thickness of 211 microns and a polyurethane (PU)-based coating with a total nominal coating thickness of 41 microns were designed. These are two commonly used coating systems for architectural steel cladding and are referred to in the following work as PVC and PU. The dimensions of the tag are displayed in Figure 5 and listed in Table 1.

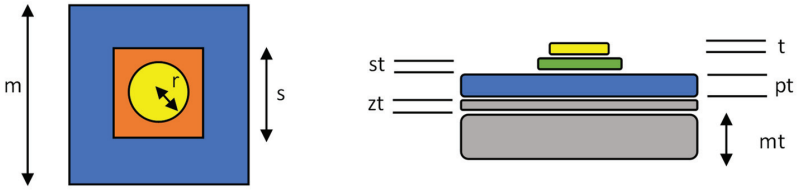


Figure 5. Tag systems showing key dimensions which are defined in Table 1.

Table 1. Simulation parameters.

Component	Symbol	Dimension/mm
Antenna Radius	r	10
Antenna Thickness	t	0.035
Dielectric Thickness	pt	0.211 (PVC) 0.040 (PU)
Zinc Thickness	zt	0.04
Metal Thickness	mt	0.6
Dielectric Constant (PVC & PU)	e	3.5
Sample Height/Width	m	60
Substrate Thickness	st	0.00 (NS) 0.20 (S)
Substrate Height/Width	s	0.00 (NS) 22.0 (S)

In the modelling of the cladding panels, several assumptions were made for the initial testing. Firstly, the coating layers of pre-treatment, primer, and topcoats were combined into one homogenous layer with a cumulative thickness and an arbitrary dielectric constant of 3.5. The metal substrate was defined as 1010 steel and the metallic coating layer as pure zinc. The size of the cladding piece was set to 6 cm^2, and the substrate layer used was FR-4 (lossy). The properties for each of the materials used are shown in Table 2.

Table 2. Simulation material properties.

Component	Value/Properties	Reference
Steel Conductivity (σ_S)	6.99×10^6 S/m	[32,33]
Zinc Conductivity (σ_z)	1.69×10^7 S/m	[32,33]
Fr-4 Dielectric Constant (ε_r)	4.3	[33]
Fr-4 Loss Tangent (δ)	0.025	[33]

Although the actual steel substrate and zinc layer are composed of either different or more complex alloys, the conductivity values used were expected to be a fair representative value. Furthermore, additional simulations, not provided in this work, showed that small deviations in these values of conductivity did not noticeably affect the resonant frequency and only produced small changes in the measured RCS.

Perhaps the largest source of uncertainty is the accuracy in modelling the coating layers as one homogenous layer, whereas in reality the pre-treatment, primer, and topcoat(s) have different requirements and hence compositions. This was mainly done for ease of modelling, and it is believed to be representative of a weighted average of the multiple dielectric layers. It is possible that experimental confirmation of this assumption may be further required. A dielectric constant of 3.5 was considered an accurate representation of both systems as the coatings are based on polyvinyl chloride chemistry, which has a stated dielectric constant of 3.4 [34,35], and polyurethane chemistry, which has a stated dielectric constant of 3.2–3.6 [36,37].

The simulation, shown in Figure 6, was set up with a frequency range of 2–6 GHz of plane wave excitation in the z direction. The E-plane of the excitation plane wave was oriented parallel with the y axis and the H-plane was oriented parallel with the x axis.

The plane wave excitation source and an RCS Probe were set −200 mm away from the sample in z direction to provide the simulated interrogation signal and response signal measurement of the tag, respectively, at 200 mm distance. Finally, all boundaries were set to open (add space), which is used to simulate free space surrounding the set up and is recommended for antenna calculations [38].

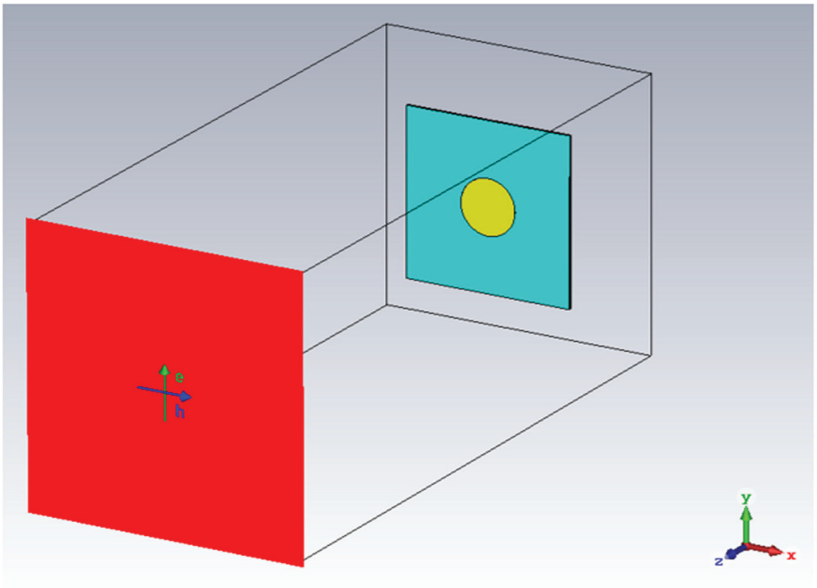


Figure 6. Simulation set up in CST Studio Suite.

The resonant frequency of the tag was determined by measuring the frequency at the point of maximum drop of RCS amplitude. The RCS change at resonant frequency was calculated by taking the measured RCS (RCS_{RF}) away from the average of the RCS at the beginning of the valley (RCS_L) and the RCS at the end of the valley (RCS_H):

$$\Delta RCS = \frac{RCS_L + RCS_H}{2} - RCS_{RF} \quad (5)$$

2.2. Production of the Antenna

In order to validate the simulations, several tags were produced, as shown in Figure 7. The antennae were manufactured from adhesive copper tape (RS Pro foil, thickness 0.035 mm, from RS Components, Ireland) using a 20 mm paper punch. This method was initially used for simplicity and because it gave reasonable cut quality and uniformity. As copper rapidly tarnishes antennae were also produced from adhesive aluminium tape (RS Pro foil thickness 0.04 mm), this is a far more environmentally resilient substrate and hence it was considered as an alternative. The 20-mm circles were then either attached directly to the sample, or to the Fr-4 substrate, which was attached to the sample via thin double sided adhesive tape. An example of the two produced antenna systems is shown in Figure 7. A number of coated steel samples were used in this study with the colours used including white, anthracite, grey (PVC only), and silver (PU only). This was done to understand the effect the different coating colours may or may not have on the measured resonant frequencies. The samples used are shown in Figure 8.

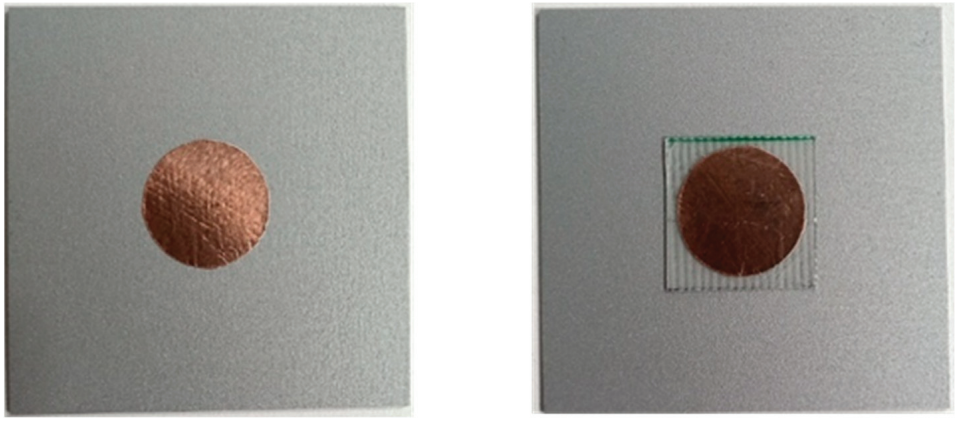


Figure 7. The manufactured Chipless RFID Tag deployed onto the same without (**left**) and with (**right**) the substrate layer.

Figure 8. The samples used in this study.

2.3. Antenna Measurement

To measure the antenna response, a vector network analyser (VNA) (model) was used, connected to two horn antennas (TX and RX) (model). A signal sweep from 4 to 6 GHz was emitted from the TX horn antenna and the resulting S_{21} from the RX antenna was recorded. To ensure similar positioning of the antenna in each test, a small acrylic stand was produced, and the position of each horn antenna was marked on the test workspace to ensure the same range was used. The distance between horn antennas was 5 cm and the distance to the sample was 20 cm as with the simulations. The test set up is shown in Figure 9.

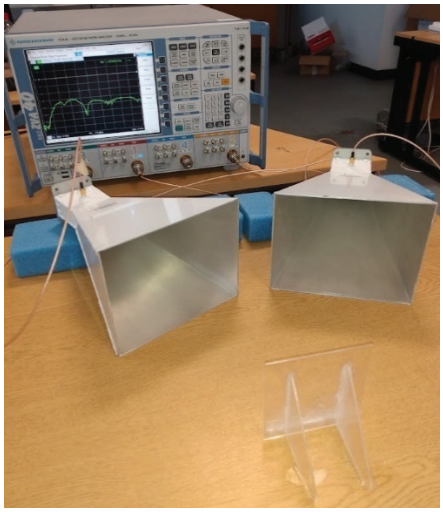


Figure 9. VNA and horn antennas used to interrogate the sample.

3. Results

3.1. Simulations

3.1.1. Initial Simulations

The simulated RCS response for each system is shown in Figure 10 and the main results that can be drawn from each RCS response are summarized in Table 3. It can be seen that the PVC system gives a similar response when the antenna is placed both directly on the coated steel and on a substrate. The PU coating system, comparatively, does not show a very significant response when the antenna is placed directly on the cladding. However, a much more prominent, albeit smaller than the PVC system, response is recorded when a substrate layer is used. Both systems show a small shift to a smaller resonant frequency when the substrate layer is used. This is expected as the substrate has a larger dielectric constant of 4.3. Adding this larger value dielectric layer has the effect of decreasing the resonant frequency of the system according to Equation (1).

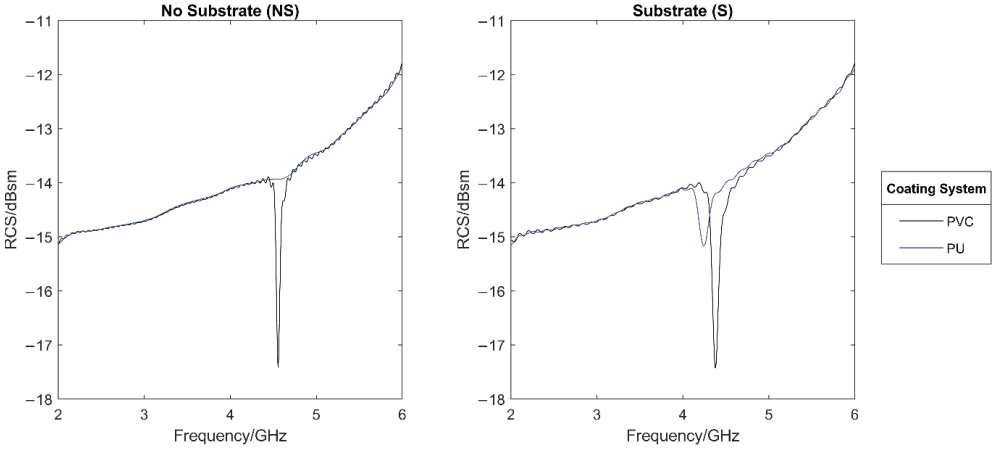


Figure 10. Simulated RCS response of sensors mounted directly on coated system (NS, **left**) and sensors mounted on an additional substrate (S, **right**).

Table 3. Results of initial simulations.

System	Resonant Frequency/GHz	RCS at Resonant Frequency/dBsm	RCS Change/dBsm
PVC-NS	4.556	−17.41	3.48
PU-NS	4.672	−13.87	0.03
PVC-S	4.380	−17.43	3.33
PU-S	4.244	−15.17	1.02

As the only difference in simulations between system PVC-NS and PU-NS is the paint thickness, the poor response from the PU system (PU-NS) must be caused by the change in dielectric layer thickness. It is stated in [28] that the usual dielectric substrate height h for a microstrip patch antenna with a dielectric constant between approximately 2 and 12 is:

$$0.003\lambda_0 \leq h \leq 0.05\lambda_0 \quad (6)$$

The calculated corresponding height range for the observed resonant frequencies seen is shown in Table 4. It can be concluded that the optimum range is not satisfied by the PU system paint thickness unless it is further increased with the substrate used in PU-S. As stated in [39], a decreased thickness of dielectric can lead to greater losses and a decreased efficiency, thus explaining the lack of a significant response by the PU model.

Table 4. Calculated optimum range of dielectric height.

System	Resonant Frequency/GHz	System h/mm	h Min/mm	h Max/mm
PVC-NS	4.556	0.211	0.20	3.29
PU-NS	4.672	0.040	0.19	3.21
PVC-S	4.380	0.411	0.21	3.42
PU-S	4.244	0.240	0.21	3.53

This conclusion is reinforced by comparing the surface current distribution maps of each system at resonant frequency, as shown in Figure 11. It can be seen that the PU sample with no substrate has a significantly lower current distribution surrounding the antenna patch than any of the other systems. This is thought to be down to the poor efficiency of the antenna because of the thin dielectric layer, which causes far less resonance and hence electrical current. As a result, the backscattered RCS records far less of a change in the signal as little power has been transferred.

3.1.2. Simulated Changes in Dielectric Constant

To monitor the effect of changes in the dielectric constant of the paint layer, a parametric sweep was performed in which the dielectric constant of the paint was changed from 2 to 5 in increments of 0.2.

As can be seen in Figure 12, for all systems, a clear trend was seen in that as the dielectric constant was increased, the resonant frequency of the system decreased. This is expected from Equations (1) and (3). However, it was also seen that the use of the substrate layer decreased the magnitude of the resonant frequency change. In Figure 13, it can be seen that the total change in resonant frequency for systems PVC-NS and PU-NS is far greater than that for PVC-S and PU-S. This is because the substrate has an unchanging dielectric constant and so is thought to average out some of the dielectric changes introduced.

What is perhaps more unexpected is that, by using a substrate layer, the RCS at resonant frequency showed a far clearer trend of reducing with increased value for the dielectric constant. This can be seen by comparing the four systems against each other, as shown in Figure 13.

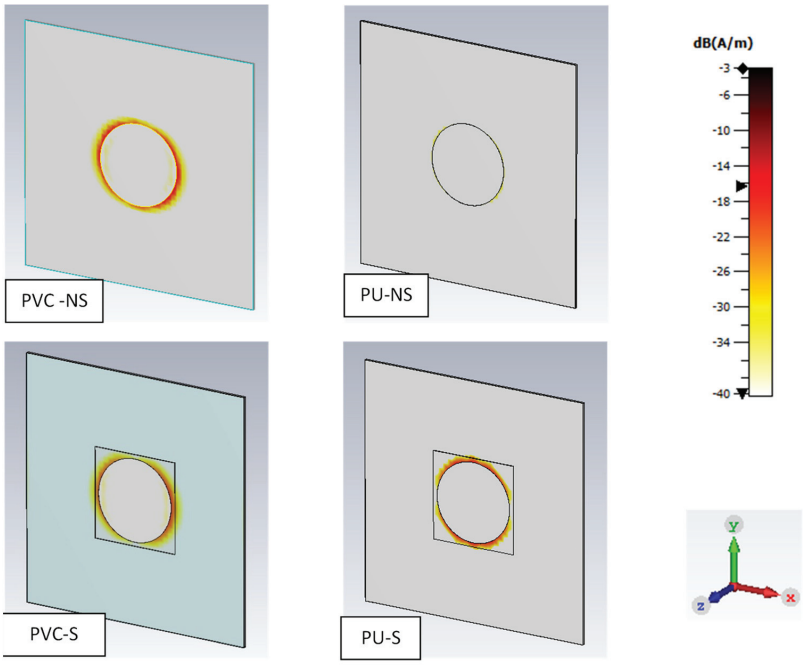


Figure 11. Simulated surface current distribution on each system at resonant frequency.

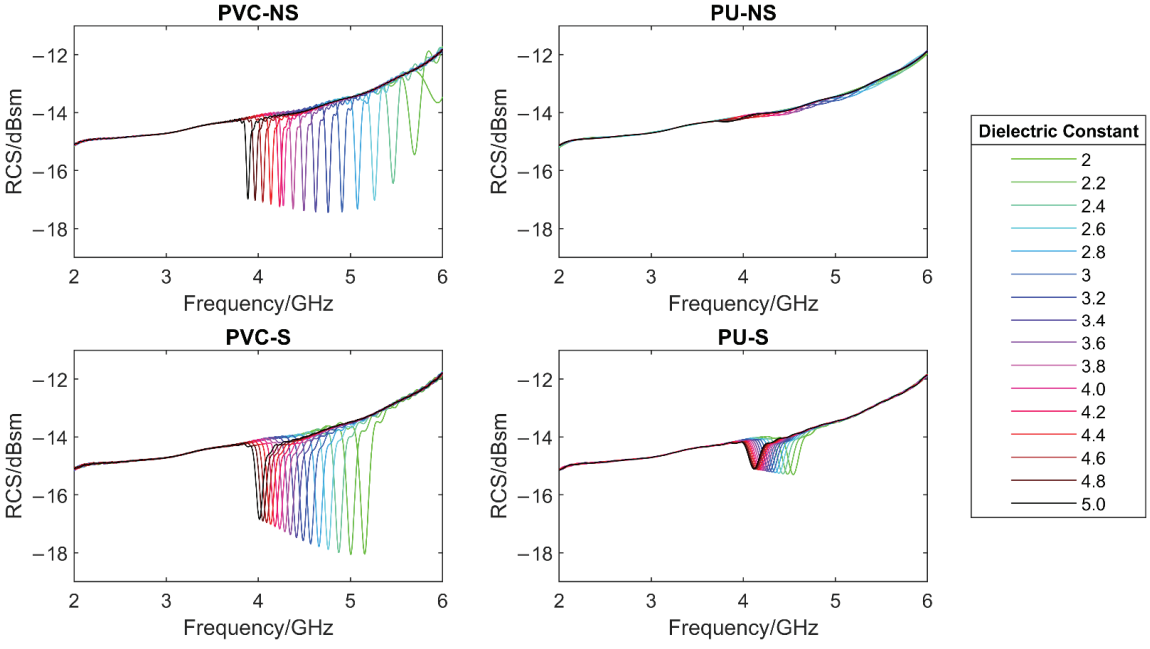


Figure 12. The effect of changes in the dielectric constant of the paint layer on the simulated RCS response.

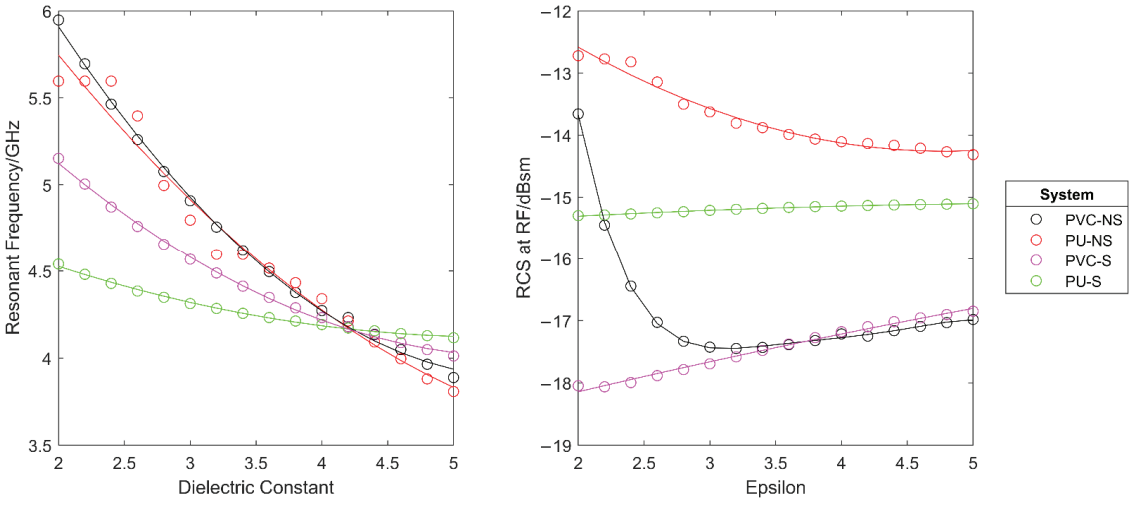


Figure 13. The effect of changes in the dielectric constant of the paint layer on the resonant frequency (**left**) and the RCS at resonant frequency (**right**) of each system.

This set of results helps to validate the model from the assumption that a dielectric constant of 3.5 was used and hence provides more confidence in the simulation process. If the coating system used has a different dielectric constant, a result will still be expected, however it will show a shift to a lower frequency, as shown by this simulation.

3.1.3. Simulated Aging/Degradation

In reality, as a coated steel sample ages, a number of the effects simulated above occur simultaneously. Hence, it is important to determine the effect of multiple expected changes on the response. It is also important to determine if the multiple effects cause the overall change in response to reduce or even cancel out. To do this, a test was performed in which aging and degradation of the coating was simulated eight times and with each iteration. As displayed in Table 5:

- The dielectric constant of the paint layer was increased by 0.025 to simulate water ingress. Previous work such as [30] has shown that water uptake by a coating leads to an increased dielectric constant of up to at least 8%.
- The diameter of small defect holes in the paint were increased by 0.01 mm to simulate defect growth. These were designed to be placed 'randomly' with no specific pattern to attempt to mimic as close as possible that which would occur in reality. The defects were placed centered on $(-4, 0)$ $(3, 3)$ $(-3, -5)$ $(-2, 7)$, with $(0, 0)$ being the centre of the CMPA. This sort of defect is a known failure method in organic coatings as a result of mechanical shock and/or aging [40].
- The paint thickness was decreased by 0.001 mm to simulated UV degradation. A decrease in organic coating thickness is known to occur through chain scission as a result of organic coating exposure to UV, oxygen, and other atmospheric contaminants [41].

This offered a reasonable level of replication of the real-life ageing process of a coated specimen which undergoes the changes simulated in this exercise. The simulated RCS response for each system is shown in Figure 14 and the effect of the aging severity on the resonant frequency and the RCS at resonant frequency are shown in Figure 15. It can be seen that the resonant frequency shifts to lower frequencies as the simulated aging severity increases. Furthermore, the RCS at resonant frequency is seen to decrease in magnitude with simulated aging. Using a substrate had little effect on the change in frequency with

aging. However, as before, this makes identification of resonant frequencies in the thinner paint system far easier, by increasing the depth of the trough.

Table 5. Parameters used for each aging severity level.

Aging Severity	Dielectric Constant	Defect Diameter/mm	Paint Thickness/mm
1	3.500	0.00	0.211 (PVC) 0.040 (PU)
2	3.525	0.01	0.210 (PVC) 0.039 (PU)
3	3.550	0.02	0.209 (PVC) 0.038 (PU)
4	3.575	0.03	0.208 (PVC) 0.037 (PU)
5	3.600	0.04	0.207 (PVC) 0.036 (PU)
6	3.625	0.05	0.206 (PVC) 0.035 (PU)
7	3.650	0.06	0.205 (PVC) 0.034 (PU)
8	3.675	0.07	0.204 (PVC) 0.033 (PU)

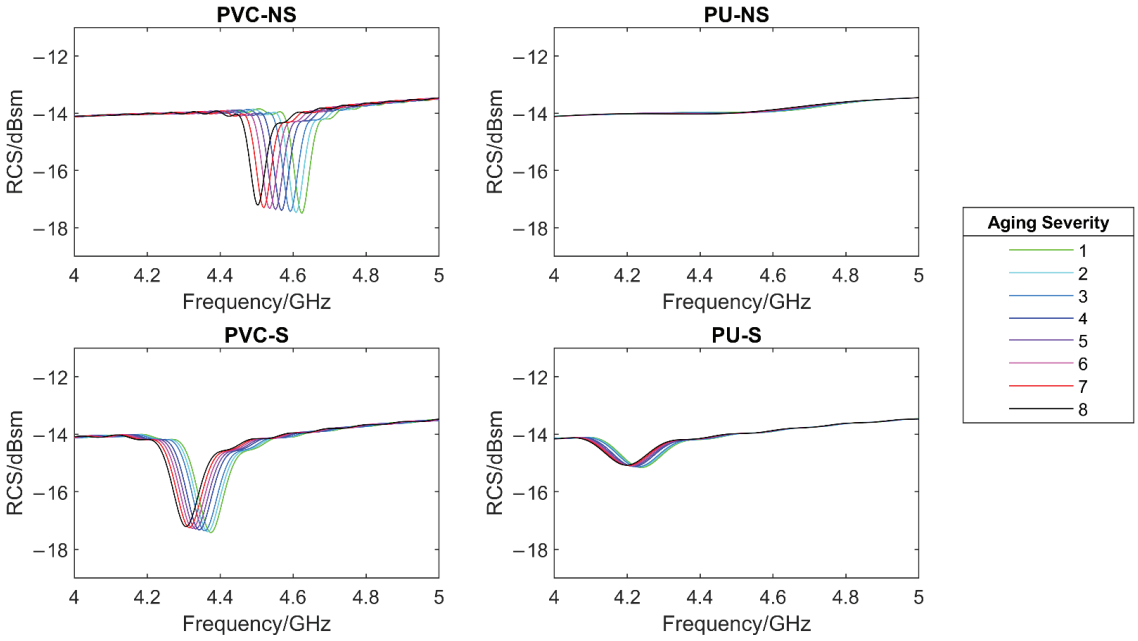


Figure 14. The effect of simulated aging on the simulated RCS response of each system.

3.2. Experimental

3.2.1. Initial Test of the System

Figure 16 shows the comparison between the simulated and measured RCS for both paint and mounting systems. This test was carried out on virgin, white coloured samples. It can be seen that interference was captured compared to the ideal simulated scenario. This made the detection of the resonant frequencies for the NS samples (no additional substrate) impossible to determine. However, the resonant frequencies of system two (marked A and B) were detectable. These were observed to be at a higher frequency and provided a smaller RCS drop that was predicted by the simulations. However, this is not unexpected due small variations between the values and assumptions used in the simulations and the actual real values as mentioned previously in Section 3.1. The resonant frequency troughs did however appear in a similar position and size relative to each other when compared to the simulated results giving confidence in their correct identification. While it is true that noise would be greatly reduced through the use of an anechoic chamber, this would somewhat

defeat the desire to use this technique in the field. While, due to the low simulated RCS trough, it was expected that the NS-PU sample was undetectable, it was surprising that the NS-PVC sample also produced no clear resonant frequency trough, especially as in simulations this showed a clear decrease in RCS of around 3.5 dBsm at resonant frequency. It was theorized that this discrepancy between simulation and experimental may be due to the difference between modelling the coating perfectly as a layer of set dielectric properties and measuring on what is an inhomogeneous, multi-layer, multi component, complex paint system. While in a simulation, the paint system alone can produce a resonant antenna, in real life, the addition of a small FR-4 substrate with homogenous dielectric properties is required to support the resonance of the antenna. Hence, it was concluded that the concept of attaching the circle directly to the paint without an additional substrate was insufficiently responsive to use for further testing.

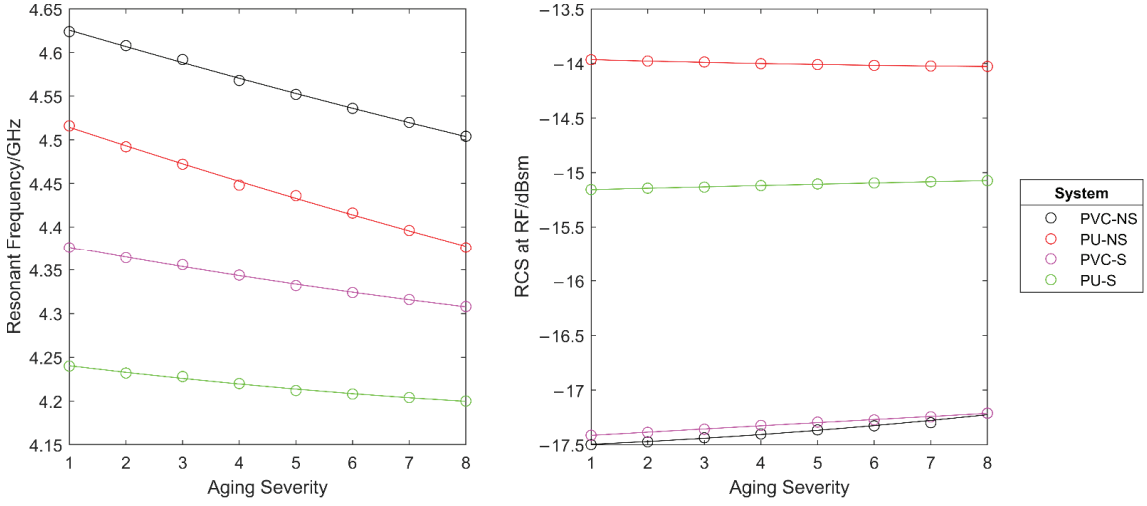


Figure 15. The effect of simulated aging on the resonant frequency (**left**) and the RCS at resonant frequency (**right**) of each system.

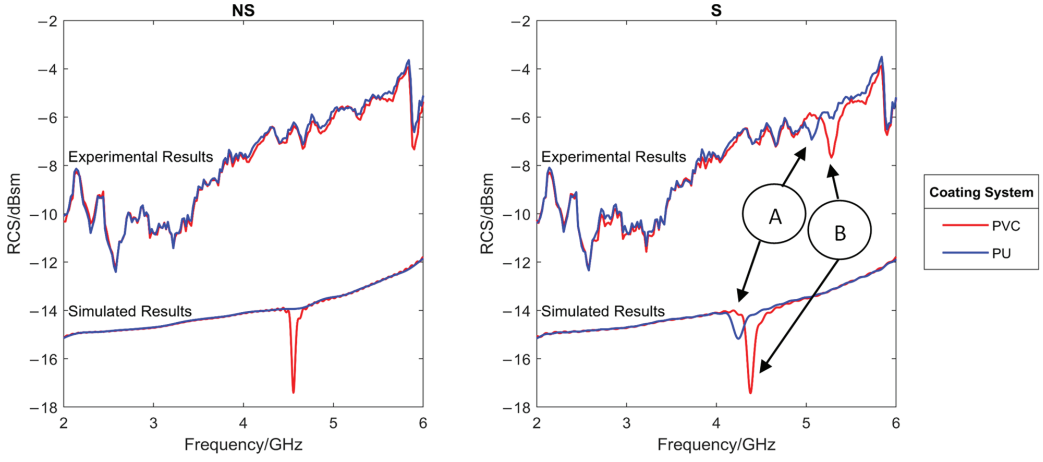


Figure 16. Comparison of the simulated and measured RCS response of sensors mounted directly on coated system (NS, **right**) and sensors mounted on an additional substrate (S, **left**). Identifiable peaks are matched between simulated and experimental results with designators A and B.

However, to reduce the impact of noise, the following results were calculated by subtracting the sample background RCS. This was the measured RCS of the sample without the antenna attached. As shown later, this allowed far easier trough detection and noise removal.

3.2.2. Detection of Artificial Weathering

Following the initial testing of the chipless RFID system, samples were produced with different artificial aging techniques. The set of samples (Salt) was produced via 10 weeks of ASTM B117 salt fog/spray exposure in a 100% humidity (5% NaCl solution, pH 7), 35 °C environment. A second set of samples (Chemical) was exposed to 50 MEK (Methyl Ethyl Ketone) double rubs, a 30 min acetone soak, and a two-hour boil. This was done as a proof of the concept of the technique to detect changes in organic coatings caused by degradation, and hence the use of common accelerated weathering techniques. The samples after exposure are shown in Figure 17.

It can be seen that the chemical samples do not appear, visually, that different to the control samples except for slight discoloration on the PVC grey and anthracite samples. The salt exposed samples show a far greater visual level of degradation with significant blistering and corrosion at the cut edge. However, it should be noted that the sensor tag is only expected to detect changes in the coating directly underneath where it is placed and that the middle of the samples showed little visual change to the control samples. Hence, the sensor is still monitoring an area in which degradation of the coating has not occurred significantly enough for a visual inspection to determine. After exposure to the environments, the samples were also examined using FTIR. Examples of the FTIR results are shown for a PVC and PU sample in Figure 18.

From these spectra, it is possible to see the effect the different treatments have had on the two paint systems. Perhaps most obvious is the significant emergence of the characteristic broad OH peak between approximately 3000 and 3500 cm^{-1} [42] in the samples exposed to the salt spray. This is indicative of moisture presence either in the coating itself or through the formation of micropores or blisters. While the chemically exposed samples do also perhaps show slight emergence of this peak, especially in the PVC samples, the effect is not as pronounced. However, from these results, it would be expected that the salt exposed samples would have a significantly different, higher, dielectric constant and that the chemically exposed samples may have a slightly higher dielectric constant. The respective obtained RCS results for a PVC and PU sample are shown in Figure 19.

It can be observed that the salt and chemical treatment has caused the resonant frequency to shift to the left and that this shift is greater in magnitude for the PVC sample. This shift is indicative of the types of changes expected under these tests. A comparison of the resonant frequencies for all the tested samples exposed to the salt and chemical degradation can be seen in Figure 20 where the resonant frequency was measured using three repeats. All the samples tested show a similar trend to that of the example above, even with the variation in measurements.

As expected, the shift in resonant frequency as a result of coating changes for the PU samples is far more reduced than for that of the PVC samples. As mentioned previously, the additional layer of FR-4 used to increase the signal for both paint systems effectively reduces the magnitude of the relative shift in resonant frequency due to its unchanging dielectric properties. The influence of this effect is far greater in the PU sample, where the FR-4 thickness (0.2 mm) constitutes a far larger proportion of the total antenna dielectric; approximately 83% compared to 50% of the total PVC antenna dielectric. However, despite this, it is still possible to distinguish the different treatments through resonant peak identification.

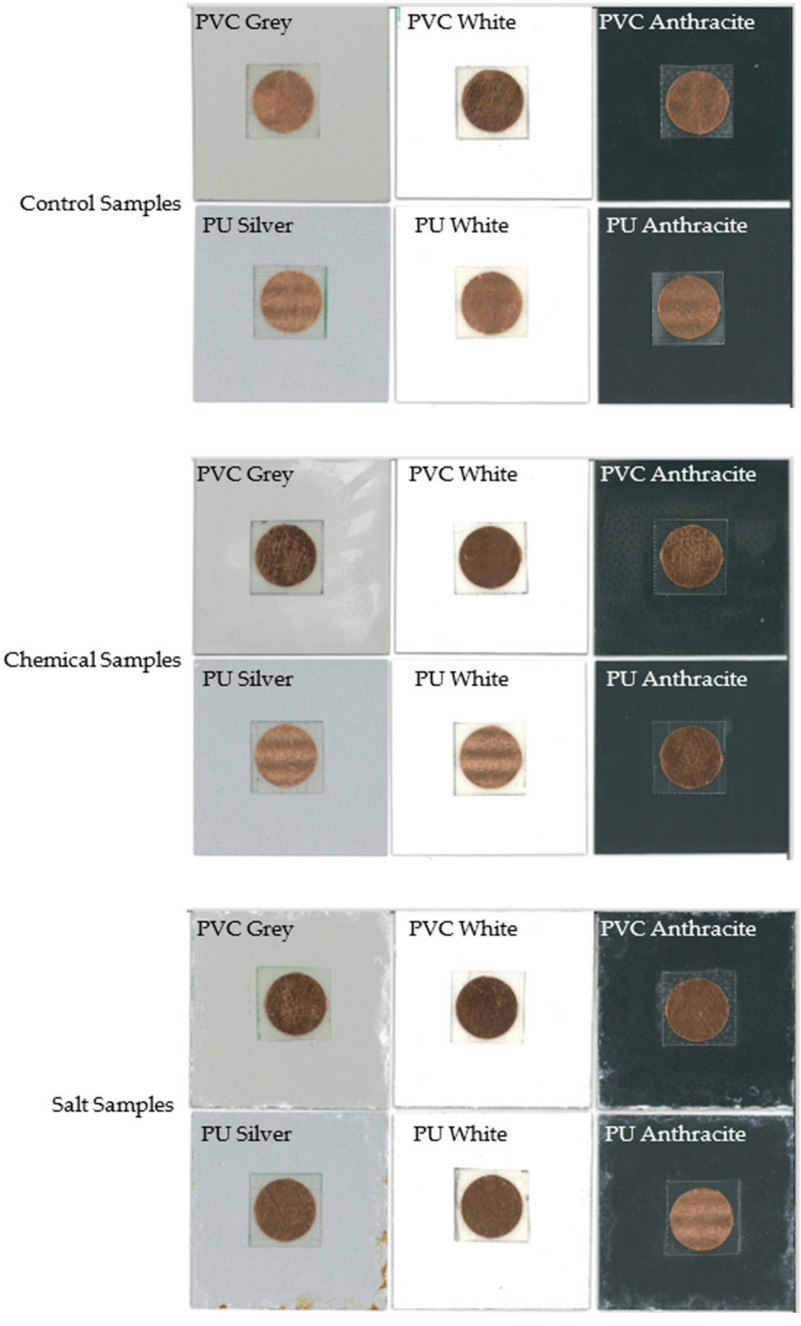


Figure 17. Sample appearance after no testing (control) (**top**), chemical (**middle**) and salt (**bottom**) exposure testing.

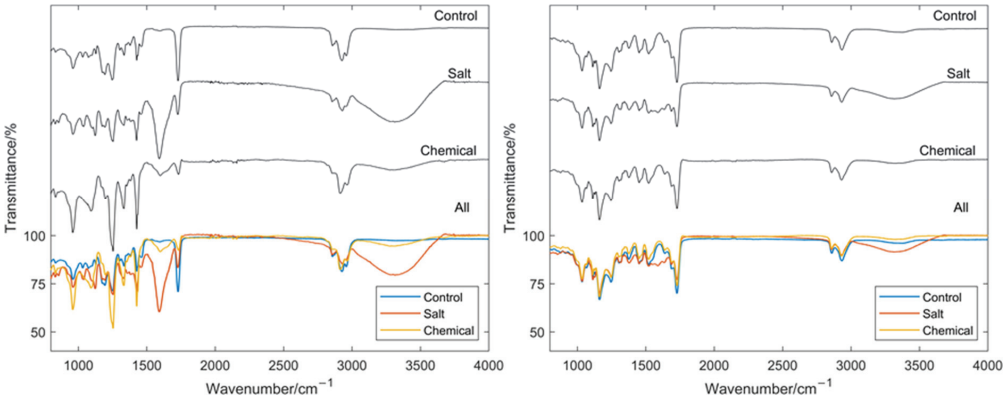


Figure 18. Resulting FTIR spectra for an example PVC sample (**left**) and example PU sample (**right**).

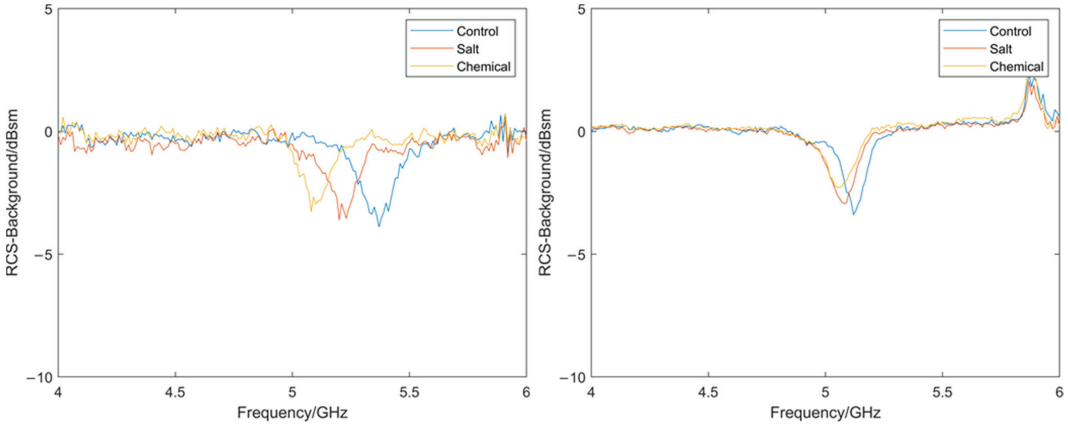


Figure 19. Measured RCS for a PVC coating system (**left**) and a PU coating system (**right**) samples exposed to salt and chemical degradation.

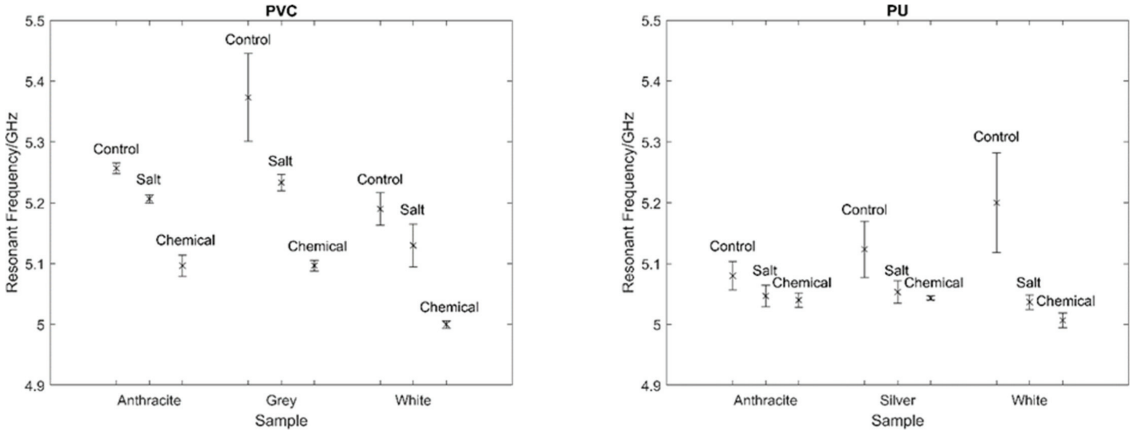


Figure 20. Measured resonant frequency for each PVC (**left**) and PU (**right**) sample tested and the associated standard error.

4. Discussion

There is no doubt that the monitoring of organic coatings applied to steel cladding on buildings is a difficult proposition. The current common monitoring practices for buildings are often simple and subjective manual inspections, and generally, a 'fix when failed' approach is taken. This means that degradation is only detected when significant coating damage and/or corrosion has taken place, increasing the cost of rectification. Recent techniques developed for monitoring early-stage coating deterioration suffer from requirements to modify the coating layers or connect to the substrate making results less representative of the bulk and more complex to carry out in-situ.

While not a perfect solution, the novel method developed shows considerable promise for the detection of pre-failure degradation in coated steels, which could be carried out quickly and reliably at low cost. It has shown its ability to detect both water uptake by the coating/coating–metal interface and degradation of the polymer itself before a significant visual change in the sample area has occurred. The degree of degradation is expected to be semi-quantitatively assessed by considering the size of the resonant frequency shift between the virgin sample and exposed sample with a greater degree of degradation producing a larger shift.

However, to further investigate and develop the technique, a number of additional studies are required. The sensitivity of the method needs to be established such that the change in resonant frequency can be more linked to the degree of degradation. For example, coating material integrity is often affected by UV, and this would be expected to provide a change in the dielectric properties of the coating. Typically, the change in structure that occurs when a coating is subjected to UV arises from the breakdown of the longer polymer chains, pigment bleaching, and cracking of the coating [41]. The precise impact of this and other failure modes on the dielectric properties of the coating should be addressed. This requires multiple coated substrate samples of varied levels of degradation, through multiple degradation mechanisms to be measured and analyzed.

Furthermore, the universality of the resonance value for multiple coating systems needs to be established. The absolute resonance is a function of many coating and substrate parameters. Hence, the coating systems of interest must first be tested in advance to determine the specific resonant frequency for the given coating thickness, pigmentation, polymer system and application involved. Thus, this technique would be used to monitor any resonance shift from this specific virgin sample datum, which will need to be considered as the prime indicator of coating degradation/substrate corrosion. It is therefore a relative, not an absolute measurement technique. Given the background noise observed within the laboratory, signal processing techniques may need to be refined in order to clearly identify the peaks in a consistent manner.

Finally, if the laboratory findings are positive then a simplified robust piece of hardware needs to be designed which can be used in the field. All of the hardware electronics for such a device exists and is relatively low cost, but it would need to be ruggedized for use in practice.

It is envisaged that this technique could be deployed via miniaturization into a portable handheld device, which is used to interrogate a tag that can be placed and then removed during routine inspections. The device would ensure that the interrogating antennas are fixed at a required angle and distance for measurement and a 'place and remove' testing procedure would ensure tarnishing of the tags does not impact the results and would reduce the aesthetic implications of constantly mounted tags. The resonant frequency measured could then be compared to the baseline initially measured resonant frequency and the shift related to the condition of the coating system. As each tag is only influenced by the coating directly beneath it several locations would have to be tested. However, this would allow a determination of the general condition of the coating across the building. Where the technique suggests early-stage coating degradation is occurring more rapidly than expected, maintenance and/or repair of the coating, via overpainting for example,

would reduce the likelihood of degradation continuing to the point where replacement is required.

5. Conclusions

This paper has introduced a new non-contact, non-destructive method for the monitoring of organically coated steel. The design of a CMPA was achieved via simulation and, by considering shifts in the resonant frequency of the CMPA, was shown to provide information as to the condition of an organic coating, even when the organic coating is relatively thin. It was observed that as a coating underwent simulated ageing, the expected response would be a decrease in the resonant frequency due to an increased dielectric constant and reduced thickness. For thinner coatings, the use of an additional piece of FR-4 substrate allowed a more substantial response, although it reduced the sensitivity of the method. Experimental results showed that background noise was the largest cause for concern, and hence resonant peaks could only be distinguished accurately if additional substrate thickness was used. Background removal via the measurement of the sample background RCS was also crucial to allow further ease of peak detection. However, the technique did respond as expected in the simulations when determining the resonant frequency of samples exposed to accelerated weathering conditions. These samples showed a clear decrease in resonant frequency as a result of water uptake and polymer degradation. This method shows promise as a rapid way to determine coating condition non-destructively and without contact with the product being tested.

Author Contributions: Conceptualization, T.S. and E.J.; methodology, T.S. and E.J.; software, N/A.; validation T.S.; formal analysis, T.S.; investigation, T.S.; resources, T.S. and E.J.; data curation, T.S.; writing—original draft preparation, T.S. and E.J.; writing—review and editing, T.S. and E.J.; visualization, T.S.; supervision, E.J.; project administration, E.J.; funding acquisition, E.J. All authors have read and agreed to the published version of the manuscript.

Funding: The authors would like to acknowledge the M2A funding from the European Social Fund via the Welsh Government (c80816), the Engineering and Physical Sciences Research Council (through UKRI) (Grant Ref: EP/L015099/1) and Tata Steel Colors that has made this research possible.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data supporting this study are provided in full in the 'Results' section of this paper.

Acknowledgments: The authors would like to thank the Swansea University Antenna and Smart Cities group specifically Hengyi Zhou and Benjamin Falkner for their help with VNA measurements.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Eurofer. *European Steel in Figures 2020*; Eurofer: Brussels, Belgium, 2021.
2. Biezma, M.; Cristobal, R.J. Methodology to Study Cost of Corrosion. *Int. J. Corros. Process. Corros. Control.* **2005**, *40*, 344–352. [CrossRef]
3. Koch, G.; Brongers, M.P.; Thompson, N.; Virmani, Y.; Payer, J. *Corrosion Costs and Preventative Strategies in the United States*; U.S. Department of Transport: Washington, DC, USA, 1998.
4. Savill, T.; Jewell, E. Techniques for in situ monitoring the performance of organic coatings and their applicability to the pre-finished steel industry: A review. *Sensors* **2021**, *21*, 6334. [CrossRef] [PubMed]
5. Singh, R. Corrosion and Corrosion Protection. In *Pipeline Integrity Handbook*, 2nd ed.; Elsevier: Amsterdam, The Netherlands, 2017; pp. 241–270.
6. Enser, H.; Kulha, P.; Sell, J.K.; Schatzl-Linder, M.; Strauß, B.; Hilber, W.; Jakoby, B. Printed strain gauges embedded in organic coatings—Analysis of gauge factor and temperature dependence. *Sens. Actuators A Phys.* **2018**, *276*, 137–143. [CrossRef]
7. Li, W.; Hintze, P.; Calle, L.M.; Buhrow, J.; Curran, J. *Smart Coating for Corrosion Indication and Prevention: Recent Progress*; Paper No: 09499; NACE: Houston, TX, USA, 2009; pp. 1–17.

8. Diler, E.; Lédan, F.; LeBozec, N.; Thierry, D. Real-time monitoring of the degradation of metallic and organic coatings using electrical resistance sensors. *Mater. Corros.* **2017**, *68*, 1365–1376. [CrossRef]
9. Cíhal, V.; Lasek, S.; Blahetová, M.; Kalabisová, E.; Krhutová, Z. Trends in the electrochemical polarization potentiodynamic reactivation method—EPR. *Chem. Biochem. Eng. Q.* **2007**, *21*, 47–54.
10. Cai, G.; Wang, H.; Jiang, D.; Dong, Z. Impedance sensor for the early failure diagnosis of organic coatings. *J. Coat. Technol. Res.* **2018**, *15*, 1259–1272. [CrossRef]
11. Hinton, A.J. Determination of Coating Adhesion Using Electrochemical Impedance Spectroscopy. *Solartron. Anal.* **2010**, *2*, 18–23. Available online: http://www.ameteksi.com/-/media/ameteksi/download_links/documentations/library/solartonanalytical/electrochemistry/technical%20note%20impedance%20spectroscopy%20for%20coatings.pdf?la=en (accessed on 11 March 2021).
12. Simpson, T.C.; Moran, P.J.; Hampel, H.; Davis, G.D.; Shaw, B.A.; Arah, C.O.; Fritz, T.L.; Zankel, K. Electrochemical monitoring of organic coating degradation during atmospheric or vapor phase exposure. *Corrosion* **1990**, *46*, 331–336. [CrossRef]
13. Davis, G.D.; Ross, R.A.; Dunn, R.C. Wireless, Battery-Powered Coating Health Monitor (CHM). In Proceedings of the CORROSION 2010, San Antonio, TX, USA, 14–18 March 2010; pp. 1–9.
14. Shi, Y.; You, C. A sensor system designed for remote coating degradation detection. In Proceedings of the 2010 IEEE Sensors Applications Symposium (SAS), Limerick, Ireland, 23–25 February 2010; pp. 257–260. Available online: http://ieeexplore.ieee.org/document/5439388/ (accessed on 11 March 2021).
15. Allahar, K.; Su, Q.; Bierwagen, G. Non-substrate EIS monitoring of organic coatings with embedded electrodes. *Prog. Org. Coat.* **2010**, *67*, 180–187. [CrossRef]
16. Van Westing, E.P.M.; Ferrari, G.M.; de Wit, J.H.W. The determination of coating performance with impedance measurements—II. *Water uptake of coatings. Corros Sci.* **1994**, *36*, 957–977. Available online: https://linkinghub.elsevier.com/retrieve/pii/0010938X9490197X (accessed on 11 March 2021). [CrossRef]
17. Kittel, J.; Celati, N.; Keddam, M.; Takenouti, H. New methods for the study of organic coatings by EIS. *Prog. Org. Coat.* **2001**, *41*, 93–98. Available online: http://linkinghub.elsevier.com/retrieve/pii/S0300944000001557 (accessed on 11 March 2021). [CrossRef]
18. Buchheit, R.G. Corrosion resistant coatings and paints. In *Handbook of Environmental Degradation of Materials*, 3rd ed.; Elsevier Inc.: Amsterdam, The Netherlands, 2005; pp. 367–385.
19. Bayer, G.T.; Zamanzadeh, M. Failure Analysis of Paints and Coatings. *Pigment Resin Technol.* **2014**, *35*, 1–37.
20. Tator, K.B. Coating Deterioration. In *Protective Organic Coatings*; ASM International: Almere, The Netherlands, 2015; pp. 462–473. Available online: https://dl.asminternational.org/books/book/13/chapter/143061/coating-deterioration (accessed on 11 March 2021).
21. Munger, C.G. Causes and Prevention of Paint Failure. In *Good Painting Practice, Steel Structures Paint Manual*; SSPC: Pittsburgh, PA, USA, 2014; pp. 1–21.
22. Ryan, P.A.; Wolstenholme, R.P.; Howell, D.M. *Durability of Cladding, A State of the Art Report*; Institution of Civil Engineers: London, UK, 1994.
23. Wiltshire, B.D.; Zarifi, T.; Zarifi, M.H. Passive Split Ring Resonator Tag Configuration for RFID-Based Wireless Permittivity Sensing. *IEEE Sens. J.* **2020**, *20*, 1904–1911. [CrossRef]
24. Amin, E.M.; Karmakar, N.C.; Jensen, B.W. Fully printable chipless RFID multi-parameter sensor. *Sens.Actuators A Phys.* **2016**, *248*, 223–232. [CrossRef]
25. Karmakar, N.C.; Amin, E.M.; Saha, J.K. Introduction. In *Chipless RFID Sensors*; Wiley: New York, NY, USA, 2016; pp. 1–245.
26. Preradovic, S. Chipless RFID System for Barcode Replacement. 2009. Available online: https://bridges.monash.edu/articles/thesis/Chipless_RFID_system_for_barcode_replacement/4546315/1 (accessed on 11 March 2021).
27. Marindra, A.M.J.; Tian, G.Y. Chipless RFID Sensor Tag for Metal Crack Detection and Characterization. *IEEE Trans. Microw. Theory Tech.* **2018**, *66*, 2452–2462. Available online: https://ieeexplore.ieee.org/document/8307772/ (accessed on 11 March 2021). [CrossRef]
28. Balanis, C.A. *Antenna Theory: Analysis and Design*, 3rd ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2005.
29. Milligan, T.A. *Modern Antenna Design*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2005; Available online: http://doi.wiley.com/10.1002/0471720615 (accessed on 11 March 2021).
30. Khalifeh, R.; Lescop, B.; Gallée, F.; Rioual, S. Development of a radio frequency resonator for monitoring water diffusion in organic coatings. *Sens. Actuators A Phys.* **2016**, *247*, 30–36. [CrossRef]
31. DassaultSystems. CST Studio Suite. *Electromagnetic Field Simulation Software.* 2022. Available online: https://www.3ds.com/products-services/simulia/products/cst-studio-suite/ (accessed on 7 April 2022).
32. ANSYSInc. Ansys Granta EduPack Software. Cambridge, UK. 2020. Available online: www.ansys.com/materials (accessed on 11 March 2021).
33. DassaultSystems. *CST Material Library*; CST Studio Suite: Providence, Rhode Island, 2021.
34. KAB_ELECTRO_ACOUSTICS. Dielectric Constants. pp. 1–52. Available online: https://www.kabusa.com/Dilectric-Constants.pdf (accessed on 4 June 2020).
35. Wroe, S.; Curtis-Thomas, C. Colorcoat HPS200 Ultra—Approval Inspection testing certification. *BBA Tech. Approv. Constr.* **2017**, 1–7.

36. MatWeb. Overview of materials for Thermoset Polyurethane. 2020. Available online: http://www.matweb.com/search/DataSheet.aspx?MatGUID=26606798bc9d4538a7c7eadf78ab082b (accessed on 4 June 2020).
37. Wroe, S.; Cooper, G. Colorcoat Prisma-Approval inspection testing certification. *BBA Tech. Approv. Constr.* **2010**, 1–8.
38. CST_Microwave_Studio. Boundary Conditions—Boundaries. 2013. Available online: http://www.mweda.com/cst/cst2013/mergedprojects/CST_MICROWAVE_STUDIO/special_solvopt/special_solvopt_boundary_conditions_boundaries.htm (accessed on 3 June 2020).
39. Pozar, D.M. Microstrip Antennas. *Electromagnetics* **1992**, *12*, 381–401. Available online: http://www.tandfonline.com/doi/abs/10.1080/02726349208908321 (accessed on 11 March 2021).
40. De Leon, A.; Advincula, R.C. Conducting Polymers with Superhydrophobic Effects as Anticorrosion Coating. In *Intelligent Coatings for Corrosion Control*; Elsevier Inc.: Amsterdam, The Netherlands, 2015; pp. 409–430.
41. Gerard Torney, N.; Harkin-Jones, E. *Investigating the Recyclability of Polypropylene*; Ulster University: Coleraine, UK, 2017. Available online: https://www.researchgate.net/publication/330797824 (accessed on 11 March 2021).
42. Merck. IR Spectrum Table. Technical Data Sheet. 2022. Available online: https://www.sigmaaldrich.com/GB/en/technical-documents/technical-article/analytical-chemistry/photometry-and-reflectometry/ir-spectrum-table (accessed on 21 February 2022).

energies

MDPI

Article

Horizontal Vibration Characteristics of a Tapered Pile in Arbitrarily Layered Soil

Xiaoyan Yang [1], Guosheng Jiang [1], Hao Liu [1,2,3,*], Wenbing Wu [1,2,3,*], Guoxiong Mei [1,2] and Zijian Yang [1]

[1] Faculty of Engineering, Zhejiang Institute, China University of Geosciences, Wuhan 430074, China; yxyxmu@163.com (X.Y.); jianggs@cug.edu.cn (G.J.); meiguox@163.com (G.M.); yzj1393648253@cug.edu.cn (Z.Y.)
[2] Guangxi Key Laboratory of Disaster Prevention and Engineering Safety, College of Civil Engineering and Architecture, Guangxi University, Nanning 530004, China
[3] Research Center of Coastal Urban Geotechnical Engineering, Zhejiang University, Hangzhou 310058, China
* Correspondence: liuhao370@cug.edu.cn (H.L.); wuwb@cug.edu.cn (W.W.)

Abstract: A tapered pile (TP) is a new type of pile with a good bearing capacity, and scholars have conducted in-depth research on its static bearing characteristics. However, there is relatively little research on its dynamic bearing characteristics. In this paper, the horizontal vibration behavior of a tapered pile in arbitrarily layered soil is studied. Utilizing the Winkler foundation model and Timoshenko beam model to simulate pile-surrounding soil (PSS) and a tapered pile, respectively, the horizontal vibration model of a tapered pile embedded in layered soil was built. The analytical solutions for the horizontal displacement (HD), bending moment (BM), and shear force (SF) of a tapered pile were derived, and then the solutions for the horizontal dynamic impedance (HDI), rocking dynamic impedance (RDI), and horizontal-rocking coupling dynamic impedance (HRDI) of pile head were obtained. Using the present solutions, the effects of soil and pile properties on the horizontal vibration characteristics of a tapered pile were systemically studied. The ability of a tapered pile–soil system to resist horizontal vibration can be improved by strengthening the upper soil, but this ability cannot be further improved by increasing the thickness of the strengthened upper soil if its thickness is greater than the critical influence thickness.

Keywords: tapered pile; horizontal vibration; arbitrarily layered soil; linear viscoelastic medium; dynamic impedance

Citation: Yang, X.; Jiang, G.; Liu, H.; Wu, W.; Mei, G.; Yang, Z. Horizontal Vibration Characteristics of a Tapered Pile in Arbitrarily Layered Soil. *Energies* **2022**, *15*, 3193. https://doi.org/10.3390/en15093193

Academic Editor: Phong B. Dao

Received: 6 February 2022
Accepted: 11 April 2022
Published: 27 April 2022

1. Introduction

A tapered pile is often utilized to replace the conventional cylindrical pile (CP) because of its good bearing capacity for wind farms, port engineering and other projects dominated by horizontal loads. Experiments conducted by El Naggar and Wei [1,2], as well as Ghazavi and Ahmadi [3], showed that the horizontal and axial bearing capacities of TPs increased by 77% and 80%, respectively, compared with those of a CP with the same length and volume. The static bearing characteristics of TPs have been thoroughly investigated by many researchers [4–8]. In addition to bearing static loads, the TP is also often subjected to dynamic loadings arising from traffic, wind, waves, machine vibrations and seismic loadings, etc. Therefore, the dynamic characteristics of TPs are experiencing an increasing research interest.

Over the past decades, many scholars have paid attention to the vertical and torsional dynamic analysis of TPs. Saha and Ghosh initially developed a finite difference model to study the vertical vibration behavior of TP [9]. Xie and Vaziri presented a numerical integration method for analyzing the response of TPs to vertical vibrations [10]. Ghazavi compared the vertical dynamic characteristics of TPs and CPs under different end-bearing conditions [11]. Wu et al. established an analytical model to investigate the construction effect on the vertical dynamic response of a TP [12], where the disturbed soil is simulated

by the radially inhomogeneous soil model [13–19]. Wang et al. extended Wu's study to the saturated soil case by introducing the fractional constitutive model [20]. Bryden et al. proposed a simple approximate method to quickly obtain the vertical stiffness and damping coefficients of TPs [21]. The study indicated that the resonant amplitude of a TP during vertical vibrations could be significantly decreased as its cone angle increased. As the material damping is often neglected in existing theoretical models for simplicity, Bryden et al. also proposed an analytical model to quantitatively study the effect of material damping on the vertical dynamic response of TPs [22]. The results indicated that the omission of the material damping would cause an overestimation of the resonant amplitude and the underestimation of the resonant frequency. Wu et al. first proposed the circumferential shear complex stiffness transfer model to investigate the influence of the construction disturbance effect on the torsional dynamic response of TPs [23]. Furthermore, Guan et al. theoretically studied the torsional vibration characteristics of TPs by taking both the compaction effect of PSS and the stress diffusion effect of pile end soil into account [24].

Compared with the significant achievements in the vertical and torsional dynamic analysis of TPs, the relevant literature on the horizontal dynamic characteristics of TPs is limited. Ghazavi used a two-phase analysis method to evaluate the horizontal dynamic behavior of a TP caused by kinematic seismic loading due to earthquakes [25]. Lee et al. assumed that a TP would behave as a vertical-loaded horizontal Euler–Bernoulli (EB) beam and investigated the influence of longitudinal loading on the transverse free vibration of TP [26]. It was found that the natural frequency significantly decreased with the increasing axial loading and tapered angle. The EB beam model is verified to be reasonably accurate for long or slender piles. However, for short stubby piles, the application of an EB beam model is doubtful because it does not consider the influence of shear displacement and rotatory inertia of pile body. Gupta and Basu compared the applicability of Timoshenko (TS), EB and rigid beam models in the analysis of piles that encountered horizontal static or dynamic loads [27,28]. They verified that the TS beam model could achieve a satisfactory accuracy in the cases of all horizontally loaded piles, while the EB beam model was most suitable for solid cross-section piles with a large slenderness ratio. Ding et al. [29] and Zheng et al. [30] studied the influence of longitudinal loading on the horizontal vibration characteristics of pipe plies with the EB and TS beam models, respectively. It was shown that the EB beam model overestimated the horizontal impedance of pile and the loading effect.

Reviews of the literature show that the existing research is mainly based on EB beam models for studying the horizontal vibration characteristics of a TP. However, a TP is a typical stubby pile with its radius decreasing linearly with depth; additionally, the tapered angle will also make pile–soil dynamic interaction more complicated. The application of an EB beam model for predicting the horizontal behavior of a TP is still questionable and needs to be calibrated [14,18]. In this study, analytical solutions for TP-encountered horizontal dynamic loads are derived with the TS beam model. The influences of soil and pile properties on the horizontal vibration characteristics of TPs are investigated based on the present solutions. A systematic parametric study is carried out to evaluate the performance and accuracy of EB and TS beam models for a TP subjected to horizontal dynamic loads. The results of this work may benefit other researchers working on the monitoring or evaluation of the TP–soil system used in construction when there are horizontal dynamic loads acting on the TP head.

2. Computational Model and Assumptions

The coupling horizontal vibration model of TP embedded in arbitrarily layered soil is depicted in Figure 1, where the pile is simulated by TS beam model and the PSS is simulated by the Winkler foundation model, respectively. The pile head bears a harmonic horizontal dynamic load $F(t) = Q_0 \cos \omega t$, whose amplitude is Q_0. ω is the circular frequency. The length and cone angle of TP are denoted by θ and L, separately. d_1 represents the diameter of pile head. The z-axis coincides with the downward pile axis. Considering the variable section of TP and the layered characteristics of PSS, the TP–soil system is divided into n

segments in the z-axis, which are numbered by 1, 2, … , j, … , n from pile head to pile bottom. l_j and h_j represent the thickness and top surface depth of the jth TP–soil system, respectively. When the number of segments n is large enough, each pile segment can be approximately regarded as a circular cross-section pile segment with equal diameter. In other words, if the number of segments n is large enough, the characteristics of the pile segment and surrounding soil layer are uniform with each pile segment, but may vary from segment to segment. k_{xj} and c_{xj} are the stiffness coefficient and damping coefficient of the soil acting on the shaft of the jth pile segment. According to the suggestion of Gazetas et al. [31], the values of k_{xj} and c_{xj} can be obtained as:

$$\begin{cases} k_{xj} = 1.2E_{\mathrm{S}j} \\ c_{xj} = 6.0{a_j}^{-\frac{1}{4}}\rho_{\mathrm{S}j}\, V_{\mathrm{S}j} d_j + 2k_{xj}\dfrac{\beta_{\mathrm{S}j}}{\omega} \end{cases} \tag{1}$$

where $V_{\mathrm{S}j} = \sqrt{G_{\mathrm{S}j}/2\rho_{\mathrm{S}j}(1+\nu_{\mathrm{S}j})}$, $G_{\mathrm{S}j}$, $\beta_{\mathrm{S}j}$, $\rho_{\mathrm{S}j}$ and $\nu_{\mathrm{S}j}$ represent the shear wave velocity, shear modulus, damping ratio, density and Poisson's ratio of the jth soil layer, respectively. $d_j = d_1 - 2z\tan\theta$ represents the diameter of the jth pile segment. $a_j = \omega \cdot d_j / V_{\mathrm{S}j}$ is the dimensionless frequency of the jth TP–soil segment.

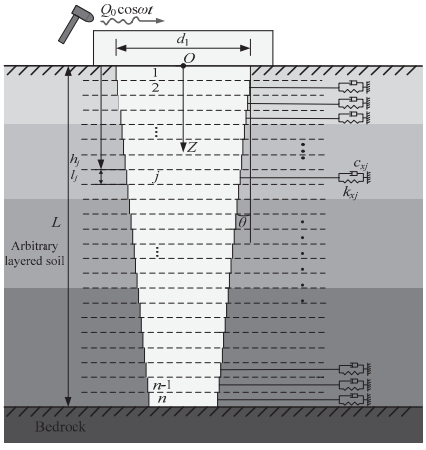


Figure 1. Transverse coupling vibration model of TP.

The analysis is conducted based on the following assumptions:

(1) The TP is a viscoelastic frustum, and its diameter reduces linearly along the length direction.
(2) To obtain the analytical solution for the established model, only the vertical load caused by the pile cap acting on the TP head is considered, and the interaction between the pile cap and its lower soil layer is not considered.
(3) The pile-surrounding soil is arbitrarily layered soil, and each soil layer is homogeneous and isotropic, which can be regarded as a linear viscoelastic medium.
(4) During horizontal vibration, the TP–soil system will undergo small deformations and strains.
(5) The soil and pile are in completely continuous contact, and the pile cap effect is not considered.
(6) During horizontal vibrations, the pile top does not rotate, and the pile bottom does not move because it is fixed by the bedrock.

3. Governing Equations and Their Solutions

The horizontal displacement (HD) and rotational angle (RA) of the jth pile segment are denoted by $u_j(z,t)$ and $\varphi_j(z,t)$, respectively. Based on the TS beam model, the governing equations of the jth pile segment are:

$$k' A_{\mathrm{P}j} G_{\mathrm{P}j}[\frac{\partial]\varphi_j(z,t)}{\partial z} - \frac{\partial^2 u_j(z,t)}{\partial z^2}] + k_{xj} \cdot u_j(z,t) + c_{xj} \cdot \frac{\partial \mathrm{u}_j(z,t)}{\partial t} + m_{\mathrm{P}j} \cdot \frac{\partial^2 u_j(z,t)}{\partial t^2} = 0 \quad (2)$$

$$E_{\mathrm{P}j} I_{\mathrm{P}j} \cdot \frac{\partial^2 \varphi_j(z,t)}{\partial z^2} + k' A_{\mathrm{P}j} G_{\mathrm{P}j}[\frac{\partial \mathrm{u}_j(z,t)}{\partial z} - \varphi_j(z,t)] = 0 \quad (3)$$

where $G_{\mathrm{P}j} = \mathrm{E}_{\mathrm{P}j}/(2(1+\nu_{\mathrm{P}j}))$, $E_{\mathrm{P}j}$, $\nu_{\mathrm{P}j}$, $m_{\mathrm{P}j}$, $A_{\mathrm{P}j}$ and $I_{\mathrm{P}j}$ are the shear modulus, elastic modulus, Poisson's ratio, mass, cross-section area and rotational moment of inertia of the jth pile segment. k' denotes the shear shape factor, which is taken as 0.75 when the pile section is circular.

For steady-state harmonic vibrations, the variable separation method is introduced to solve Equations (2) and (3). If $u_j(z,t) = U_j(z) \cdot \mathrm{e}^{\mathrm{i}\omega t}$ and $\varphi_j(z,t) = \psi_j(z) \cdot \mathrm{e}^{\mathrm{i}\omega t}$, these can be substituted into Equations (2) and (3) to yield:

$$k' A_{\mathrm{P}j} G_{\mathrm{P}j} \mathrm{e}^{\mathrm{i}\omega t} \cdot [\frac{d\psi_j(z)}{dz} - \frac{d^2 U_j(z)}{dz^2}] + U_j(z) \mathrm{e}^{\mathrm{i}\omega t} \cdot [k_{xj} - m_{\mathrm{P}j}\omega^2 + \mathrm{i}c_{xj}\omega] = 0 \quad (4)$$

$$E_{\mathrm{P}j} I_{\mathrm{P}j} \mathrm{e}^{\mathrm{i}\omega t} \cdot \frac{d^2 \psi_j(z)}{\partial z^2} + k' A_{\mathrm{P}j} G_{\mathrm{P}j} \mathrm{e}^{\mathrm{i}\omega t} \cdot [\frac{d\mathrm{U}_j(z)}{dz} - \psi_j(z)] = 0 \quad (5)$$

For simplicity, the following variables are introduced:

$$W_{\mathrm{P}j} = E_{\mathrm{P}j} I_{\mathrm{P}j} \quad (6)$$

$$J_{\mathrm{P}j} = k' A_{\mathrm{P}j} G_{\mathrm{P}j} \quad (7)$$

$$k_{\mathrm{S}j} = k_{xj} - m_{\mathrm{P}j}\omega^2 + \mathrm{i}c_{xj}\omega \quad (8)$$

Substituting the above variables into Equations (4) and (5) gives:

$$W_{\mathrm{P}j} \frac{d^4 U_j(z)}{dz^4} - \frac{k_{\mathrm{S}j} W_{\mathrm{P}j}}{J_{\mathrm{P}j}} \frac{d^2 U_j(z)}{dz^2} + k_{\mathrm{S}j} U_j(z) = 0 \quad (9)$$

$$\psi_j(z) = \frac{W_{\mathrm{P}j}}{J_{\mathrm{P}j}} \frac{d^3 U_j(z)}{dz^3} + (1 - \frac{k_{\mathrm{S}j} \cdot W_{\mathrm{P}j}}{J_{\mathrm{P}j}^2}) \frac{dU_j(z)}{dz} \quad (10)$$

Then, the general solutions of Equations (9) and (10) are obtained as:

$$U_j(z) = \mathrm{e}^{\alpha_j z}(A_{1j} \cos \beta_j z + B_{1j} \sin \beta_j z) + \mathrm{e}^{-\alpha_j z}(C_{1j} \cos \beta_j z + D_{1j} \sin \beta_j z) \quad (11)$$

$$\psi_j(z) = \mathrm{e}^{\alpha_j z}(A_{2j} \cos \beta_j z + B_{2j} \sin \beta_j z) + \mathrm{e}^{-\alpha_j z}(C_{2j} \cos \beta_j z + D_{2j} \sin \beta_j z) \quad (12)$$

where A_{1j}, A_{2j}, B_{1j}, B_{2j}, C_{1j}, C_{2j}, D_{1j} and D_{2j} are undetermined coefficients, and the coefficients α_j and β_j are written as:

$$\alpha_j = \sqrt{\sqrt{\frac{k_{\mathrm{S}j}}{4W_{\mathrm{P}j}}} + \frac{k_{\mathrm{S}j}}{4J_{\mathrm{S}j}}} \quad (13)$$

$$\beta_j = \sqrt{\sqrt{\frac{k_{\mathrm{S}j}}{4W_{\mathrm{P}j}}} - \frac{k_{\mathrm{S}j}}{4J_{\mathrm{S}j}}} \quad (14)$$

According to the theory of material mechanics, the bending moment (BM) amplitude and shear force (SF) amplitude are obtained as:

$$\begin{aligned} M_j(z) &= -W_{Pj}\frac{d\psi_j(z)}{dz} \\ &= -W_{Pj}[e^{\alpha_j z}(A_{3j}\cos\beta_j z + B_{3j}\sin\beta_j z) + e^{-\alpha_j z}(C_{3j}\cos\beta_j z + D_{3j}\sin\beta_j z)] \end{aligned} \tag{15}$$

$$\begin{aligned} Q_j(z) &= J_{Pj}[\frac{dU_j(z)}{dz} - \psi_j(z)] \\ &= W_{Pj}[e^{\alpha_j z}(A_{4j}\cos\beta_j z + B_{4j}\sin\beta_j z) + e^{-\alpha_j z}(C_{4j}\cos\beta_j z + D_{4j}\sin\beta_j z)] \end{aligned} \tag{16}$$

where A_{3j}, A_{4j}, B_{3j}, B_{4j}, C_{3j}, C_{4j}, D_{3j} and D_{4j} are undetermined coefficients.

The boundary conditions of the pile head bottom are derived as:

$$\psi_1(z)|_{z=0} = 0; \quad Q_1(z)|_{z=0} = Q_0 \tag{17}$$

$$U_n(z)|_{z=L} = 0; \quad \psi_n(z)|_{z=L} = 0 \tag{18}$$

The HD, RA, BM and SF at the interface of adjacent pile segments meet the continuous conditions, and their amplitudes are obtained as:

$$U_j(z)\big|_{z=h_j} = U_{j+1}(z)\big|_{z=h_j} \tag{19}$$

$$\psi_j(z)\big|_{z=h_j} = \psi_{j+1}(z)\big|_{z=h_j} \tag{20}$$

$$M_j(z)\big|_{z=h_j} = M_{j+1}(z)\big|_{z=h_j} \tag{21}$$

$$Q_j(z)\big|_{z=h_j} = Q_{j+1}(z)\big|_{z=h_j} \tag{22}$$

Transforming Equations (19)–(22) into a matrix relation gives:

$$\boldsymbol{T}_j(h_j)\boldsymbol{X}_j = \boldsymbol{T}_{j+1}(h_j)\boldsymbol{X}_{j+1} \tag{23}$$

where the expressions of each matrix are expressed as:

$$\boldsymbol{T}_j(h_j) = \begin{bmatrix} t_{1j} & t_{2j} & t_{3j} & t_{4j} \\ t_{5j}\cdot t_{1j} - t_{6j}\cdot t_{2j} & t_{6j}\cdot t_{1j} + t_{5j}\cdot t_{2j} & -t_{5j}\cdot t_{3j} - t_{6j}\cdot t_{4j} & t_{6j}\cdot t_{3j} - t_{5j}\cdot t_{4j} \\ t_{7j}\cdot t_{1j} - t_{8j}\cdot t_{2j} & t_{8j}\cdot t_{1j} + t_{7j}\cdot t_{2j} & t_{7j}\cdot t_{3j} + t_{8j}\cdot t_{4j} & -t_{8j}\cdot t_{3j} + t_{7j}\cdot t_{4j} \\ t_{9j}\cdot t_{1j} - t_{10j}\cdot t_{2j} & t_{10j}\cdot t_{1j} + t_{9j}\cdot t_{2j} & -t_{9j}\cdot t_{3j} - t_{10j}\cdot t_{4j} & t_{10j}\cdot t_{3j} - t_{9j}\cdot t_{4j} \end{bmatrix} \tag{24}$$

$$\boldsymbol{X}_j = \begin{bmatrix} A_{1j} & B_{1j} & C_{1j} & D_{1j} \end{bmatrix}^T \tag{25}$$

where $t_{1j} = e^{\alpha_j z}\cdot\cos\beta_j z$, $t_{2j} = e^{\alpha_j z}\cdot\sin\beta_j z$, $t_{3j} = e^{-\alpha_j z}\cdot\cos\beta_j z$, $t_{4j} = e^{-\alpha_j z}\cdot\sin\beta_j z$, $t_{5j} = \alpha_j(1 - \frac{k_{Sj}\cdot W_{Pj}}{J_{Pj}^2}) + \frac{W_{Pj}}{J_{Pj}}(\alpha_j^3 - 3\alpha_j\beta_j^2)$, $t_{6j} = \beta_j(1 - \frac{k_{Sj}\cdot W_{Pj}}{J_{Pj}^2}) + \frac{W_{Pj}}{J_{Pj}}(-\beta_j^3 + 3\alpha_j^2\beta_j)$, $t_{7j} = \alpha_j^2 - \beta_j^2 - \frac{k_{Sj}}{J_{Pj}}$, $t_{8j} = 2\alpha_j\beta_j$, $t_{9j} = \alpha_j\frac{k_{Sj}}{J_{Pj}} - \alpha_j^3 + 3\alpha_j\beta_j^2$, $t_{10j} = \beta_j\frac{k_{Sj}}{J_{Pj}} + \beta_j^3 - 3\alpha_j^2\beta_j$.

The matrix $\boldsymbol{X}_n$ can be obtained by the cumulative multiplication with Equation (23), that is:

$$\boldsymbol{X}_n = \left[\prod_{j=n}^{2}\boldsymbol{T}_j^{-1}(h_{j-1})\boldsymbol{T}_{j-1}(h_{j-1})\right]\boldsymbol{X}_1 \tag{26}$$

Substituting Equation (17) into Equations (12) and (16) gives:

$$\begin{bmatrix} t_{51} & t_{61} & -t_{51} & t_{61} \\ t_{91} & t_{10,1} & -t_{91} & t_{10,1} \end{bmatrix}\boldsymbol{X}_1 = \begin{Bmatrix} 0 \\ \frac{Q_0}{W_{Pj}} \end{Bmatrix} \tag{27}$$

Substituting Equation (18) into Equations (11) and (12) yields:

$$\begin{bmatrix} t_{1n} & t_{2n} & t_{3n} & t_{4n} \\ t_{5n} \cdot t_{1n} - t_{6n} \cdot t_{2n} & t_{6n} \cdot t_{1n} + t_{5n} \cdot t_{2n} & -t_{5n} \cdot t_{3n} - t_{6n} \cdot t_{4n} & t_{6n} \cdot t_{3n} - t_{5n} \cdot t_{4n} \end{bmatrix} \boldsymbol{X}_n = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \tag{28}$$

Substituting Equation (26) into Equation (28), and combining them with Equation (27), the equations group related to the matrix $\boldsymbol{X}_1$ can be obtained and then the matrix $\boldsymbol{X}_1$ can be solved. According to Equation (26), the matrix $\boldsymbol{X}_j$ of the jth pile segment can be successively deduced, then the HD function of the whole TP can be obtained. With the help of HD function, the RA, BM and SF are successively deduced. For example, the HD function of TP expressed by the piecewise function is derived as:

$$u(z,t)\big|_{z=z_0} = u_j(z,t)\big|_{z=z_0} (h_j \le z_0 \le h_{j+1}) \tag{29}$$

Then, the HDI K_h, RDI K_r, and HRDI K_hr of TP are derived as:

$$K_\mathrm{h} = \frac{Q_1(0)}{U_1(0)} \tag{30}$$

$$K_\mathrm{r} = \frac{M_1(0)}{\psi_1(0)} \tag{31}$$

$$K_\mathrm{hr} = \frac{Q_1(0)}{\psi_1(0)} \tag{32}$$

The dimensionless stiffness factor (k_h) and damping factor (c_h) of HDI are given as:

$$k_\mathrm{h} = \frac{d_1^3}{E_{\mathrm{P}1} I_{\mathrm{P}1}} \mathrm{Re}(K_\mathrm{h}) \tag{33}$$

$$c_\mathrm{h} = \frac{d_1^3}{E_{\mathrm{P}1} I_{\mathrm{P}1}} \mathrm{Im}(K_\mathrm{h}) \tag{34}$$

The dimensionless stiffness factor (k_r) and damping factor (c_r) of RDI are expressed as:

$$k_\mathrm{r} = \frac{d_1}{E_{\mathrm{P}1} I_{\mathrm{P}1}} \mathrm{Re}(K_\mathrm{r}) \tag{35}$$

$$c_\mathrm{r} = \frac{d_1}{E_{\mathrm{P}1} I_{\mathrm{P}1}} \mathrm{Im}(K_\mathrm{r}) \tag{36}$$

The dimensionless stiffness factor (k_hr) and damping factor (c_hr) of HRDI are obtained as:

$$k_\mathrm{hr} = \frac{d_1^2}{E_{\mathrm{P}1} I_{\mathrm{P}1}} \mathrm{Re}(K_\mathrm{hr}) \tag{37}$$

$$c_\mathrm{hr} = \frac{d_1^2}{E_{\mathrm{P}1} I_{\mathrm{P}1}} \mathrm{Im}(K_\mathrm{hr}) \tag{38}$$

4. Rationality Analysis of the Present Solutions

This work theoretically studies the horizontal vibration of TP in arbitrarily layered soil based on some assumptions. Therefore, the rationality of the present solutions should be analyzed before parametric study. For convenience, the PSS is taken as two layers with the same density of 2000 kg/m^3, a damping ratio of 0.05, and Poisson's ratio of 0.4. The property difference of these two soil layers can be reflected by changing the shear modulus. The parameters of TP are set as: pile length of 8 m, pile head diameter of 0.6 m, cone angle of 0.8°, density of 2500 kg/m^3, Poisson's ratio of 0.17, and elastic modulus of 20 GPa. The other calculation parameters are set as: the amplitude of harmonic horizontal dynamic load

Q_0 is 100 kN, the dimensionless frequency of pile head a_1 is 0.5, and the thickness of the upper soil layer is 1.6 m.

To analyze the rationality of the present solutions, the dimensionless forms of maximum HD, BM and SF of pile body are introduced as:

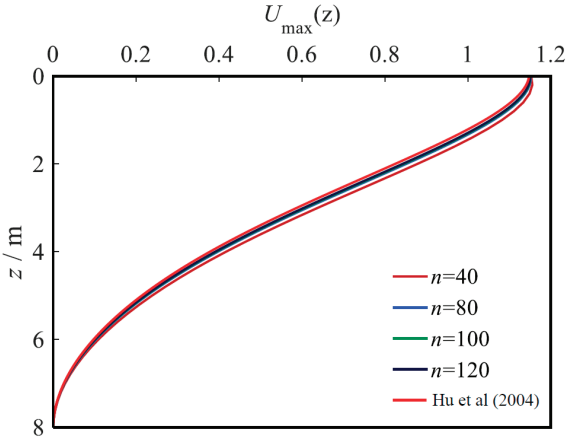
$$U_{j\max}(z) = E_{P1} d_1 \mathrm{u}_{j\max} / (500 Q_0) \tag{39}$$

$$M_{j\max}(z) = m_{j\max} / (2 Q_0 d_1) \tag{40}$$

$$Q_{j\max}(z) = q_{j\max} / Q_0 \tag{41}$$

To verify the reliability of the proposed solutions, the cone angle is set to 0°, so that the present solutions of TP can be reduced to the solutions of a circular cross-section pile with equal diameter and can then be compared with the solutions of Hu et al. [32]. The elastic moduli of the two PSS layers are given as 4 MPa. The number of TP segments is set as 40, 80, 100 and 120, respectively. Figure 2 shows the comparisons of HD, BM and SF of pile body calculated by the degenerate solution and Hu's solution [32], where the ordinate represents the soil depth. When the number of TP–soil segments n is greater than 80, the curves calculated by the two solutions tend to be stable and consistent. To ensure calculation accuracy, n is set to 100 in the following analyses, that is, the ratio of the thickness of the micro segment to the pile length is 1/100.

In the literature [14,18], it is found that the horizontal dynamic behaviors of long and thin piles gained by using the TS beam model and EB beam model are very close, but for TP, the difference between the two models is still unknown. Therefore, the difference between the TS beam model and EB beam model in analyzing the horizontal vibration characteristics of TP was studied to illustrate the rationality of the present solutions. Utilizing the same derivation process as this paper, the solutions of the horizontal vibration characteristics of TP based on the EB beam model were also obtained. However, due to content limitations, the derivation process based on the EB beam model was not given in this paper. For comparison, the length–diameter (diameter of pile head) ratio of the pile was set to 5, and the cone angles were set to 0°, 0.8° and 1.6°, respectively.



(**a**)

Figure 2. *Cont.*

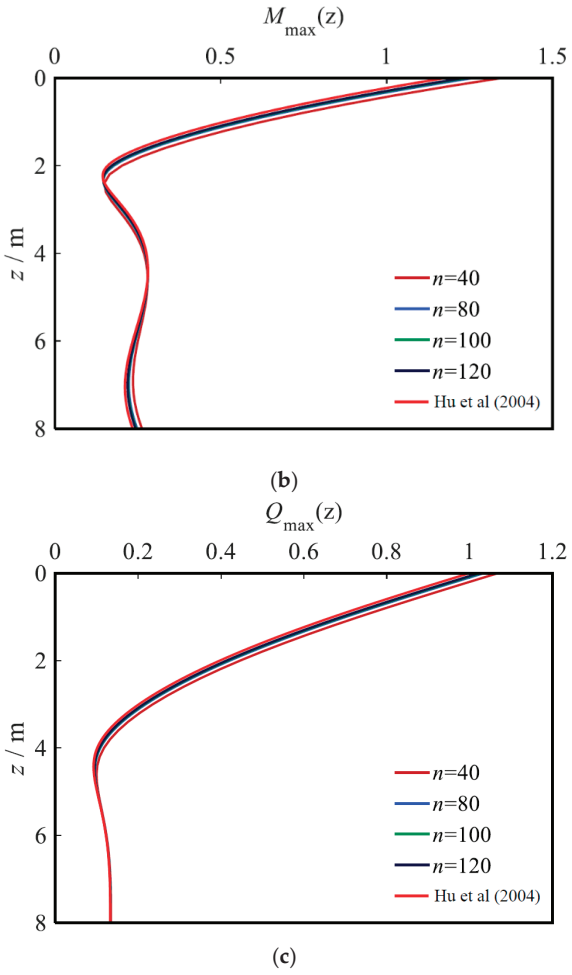


Figure 2. Comparison of the present solutions with the solution of Hu et al. [32]. (**a**) Horizontal displacement envelope, (**b**) bending moment envelope, (**c**) shear force envelope.

Figure 3 illustrates the comparison between the results obtained by the TS beam model and EB beam model. From Figure 3a, the HDs of TP obtained by both TS beam model and EB beam model decrease as the pile depth increases, and increase as the cone angle increases. The reason for this phenomenon is that, for the same diameter of pile head, the larger cone angle would lead to a smaller diameter of pile bottom, which results in an increase in HD of TP. The HD of TP obtained by the TS beam model is larger than that obtained by the EB beam model for the TS beam model and takes the shear deformation of TP into account. From Figure 3b, the change in the cone angle and calculation model has little effect on the BM of TP. From Figure 3c, the SFs of TP obtained by both the TS beam model and EB beam model decrease as the pile depth and cone angle increase. This is because, when the cone angle is smaller, the upper part of pile can bear more SF. In addition, the SF obtained by the TS beam model is smaller than that obtained by the EB beam model because the upper part of pile bears more SF when considering the shear deformation of TP. In general, since the TS beam model can consider the shear deformation of the TP, the deformation obtained by the TS beam model is relatively larger, which is conducive to the safe design of the pile foundation.

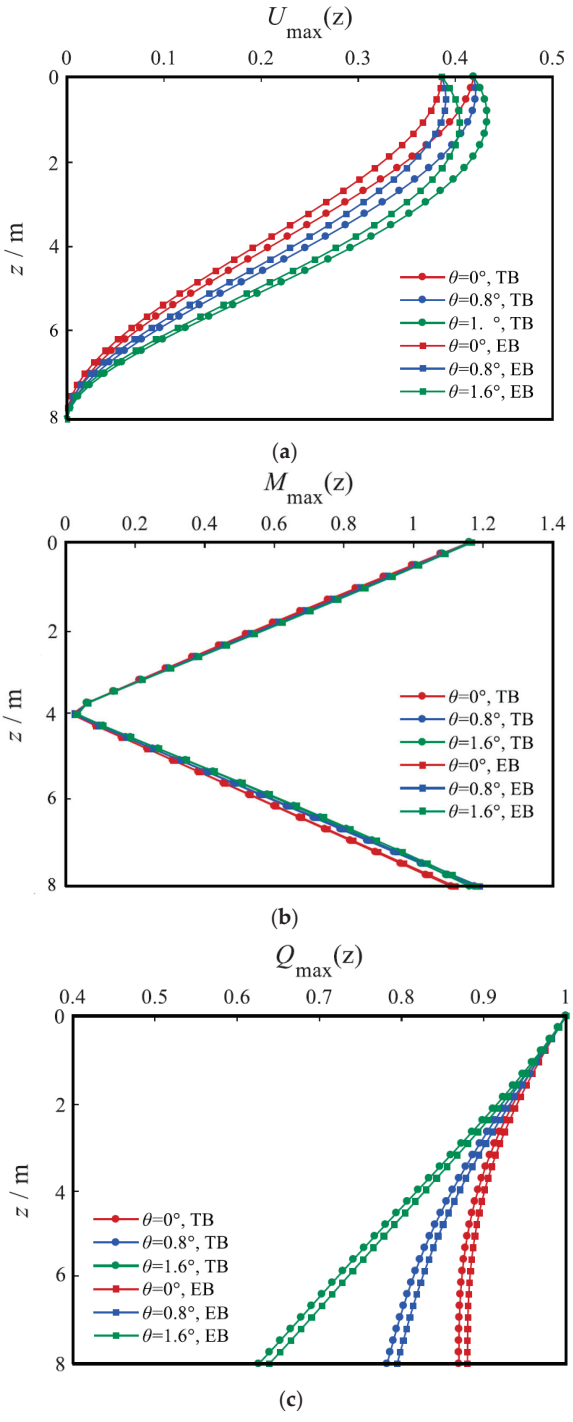


Figure 3. Comparison between the results obtained by TS beam model and EB beam model. (**a**) Horizontal displacement envelope, (**b**) bending moment envelope, (**c**) shear force envelope.

5. Parametric Study

In this section, the influences of soil and pile properties on the horizontal vibration characteristics of TPs are systemically researched. Unless otherwise specified, the parameters in this section are consistent with those in Section 4.

5.1. Influence of Shear Modulus of Soil

When analyzing the influence of the shear modulus of soil, there are two situations. The first situation is to keep the shear modulus of the lower soil unchanged at 4 MPa, and the shear moduli of upper soil are set as 2.4 MPa, 3.2 MPa, 4 MPa, 4.8 MPa and 5.6 MPa, respectively. The other situation is to keep the shear modulus of upper soil unchanged at 4 MPa, and the shear moduli of lower soil are set as 2.4 MPa, 3.2 MPa, 4 MPa, 4.8 MPa and 5.6 MPa, respectively.

Figure 4 depicts the HD curves of a pile body in the two situations. Whether the shear modulus of the upper soil or lower soil is changed, the HD of pile body decreases as the soil shear modulus increases. However, the influence of changing the shear modulus of upper soil on the HD of the TP is larger than that of lower soil because the upper soil directly bears a horizontal dynamic load transmitted from the TP. A further analysis of Figure 4 depicts that changing the shear modulus of the lower soil has a relatively uniform effect on the HD of the pile body above the pile bottom, while the HD of the pile bottom is always 0 because the pile bottom is fixed. The change in the shear modulus of the upper soil has little effect on the HD of the pile body at the position of 6~8 m. This is because the thickness of the upper soil is only 1.6 m, which makes it difficult to influence the horizontal vibration characteristics of the pile body near the bottom because its influence will be weakened by the increasing depth of the lower soil. In conclusion, the engineering properties of the upper soil are more important than those of the lower soil. From Figure 4a, when the shear moduli of the upper soil are 2.4 MPa and 5.6 MPa, the HD of pile top increases by 17.1% and decreases by 12.2%, respectively, compared with that of homogeneous foundation (i.e., $E_{S1}/E_{S2} = 1$). This phenomenon indicates that the soft upper soil has a greater effect on the horizontal vibration characteristics of TP than that of the hard upper soil. Therefore, in practical engineering, the upper soil can be treated and strengthened to improve the stability of the TP foundation.

5.2. Influence of Upper Soil Thickness

According to the analysis in Section 5.1, when the soil is soft in the upper part and hard in the lower part, the upper soil has a great effect on the horizontal vibration characteristics of the TP, but there is a critical-influence thickness of the upper soil. If the upper soil thickness is greater than the critical-influence thickness, the degree of influence does not increase. Therefore, this section intends to obtain the critical-influence thickness using a parametric study. The shear modulus of the upper soil is 2.4 MPa, and the shear modulus of the lower soil is 4 MPa. The thickness of upper soil is denoted by z_b.

Figure 5 illustrates the effect of the upper soil thickness on the HD of the pile head, where the abscissa is the upper soil thickness, and the ordinate is the HD of the pile head. The HD pile head increases as the upper soil thickness increases. However, if the upper soil thickness is larger than 5 m, the HD of the pile head reaches its maximum and remains effectively unchanged. Furthermore, the HD of pile body is calculated by setting the upper soil thickness as 4 m, 4.4 m, 4.8 m, 5.2 m and 5.6 m, respectively. As depicted in Figure 6, the two HD envelopes for the upper soil thicknesses of 5.2 m and 5.6 m completely coincide, which indicates that the upper soil thickness reaches its critical value. This phenomenon implies that, in the applications of soil improvement, if the upper soil thickness is higher than the critical-influence thickness, a further increase in the hard upper soil thickness will not further strengthen the ability of the TP–soil system to resist horizontal vibration.

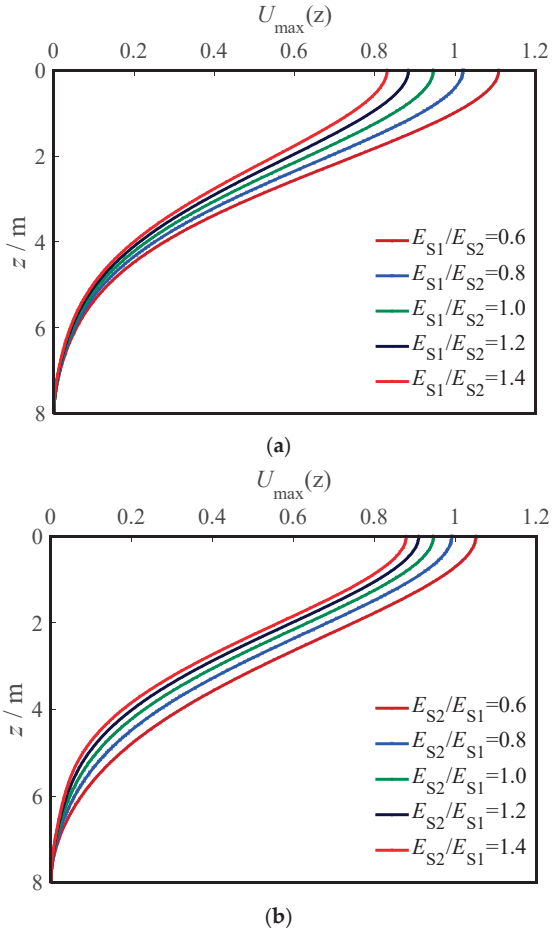


Figure 4. Influence of shear modulus of soil on the horizontal displacement of TP. (**a**) Variable shear modulus of upper soil and (**b**) variable shear modulus of lower soil.

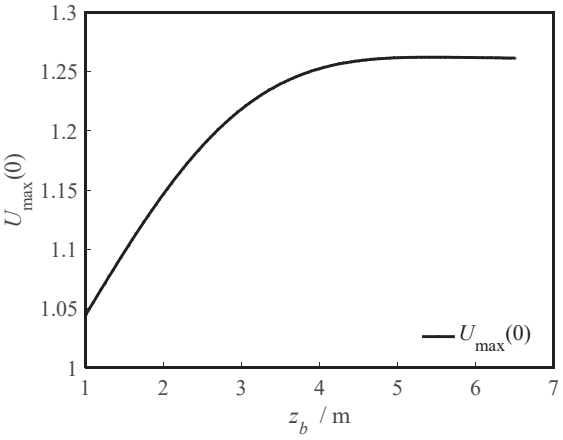


Figure 5. Influence of upper soil thickness on the horizontal displacement of pile head.

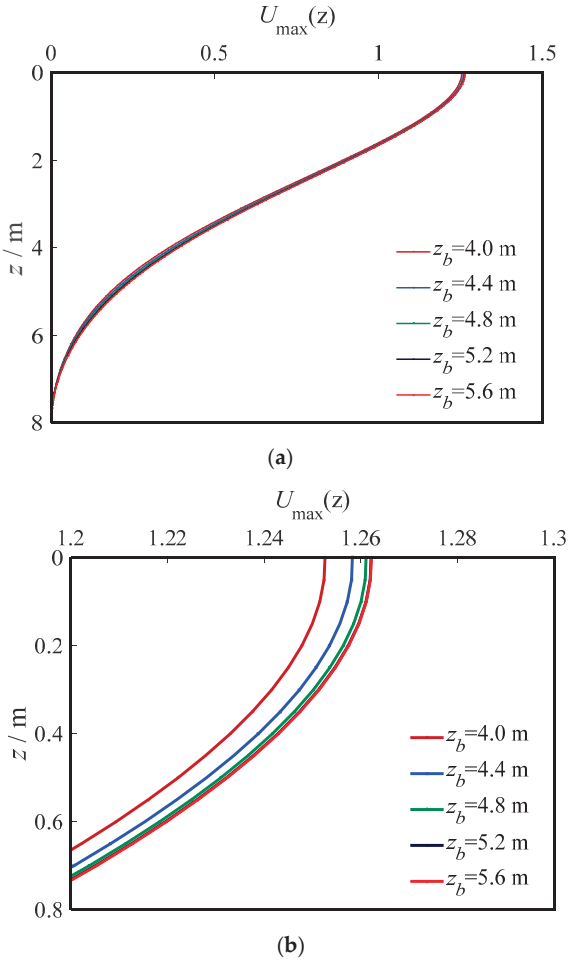


Figure 6. Influence of upper soil thickness on the horizontal displacement of pile body. (**a**) Horizontal displacement of TP (depth 0~8 m) and (**b**) horizontal displacement of TP (depth 0~0.8 m).

5.3. Influence of Weak Pile Segment Length

When pouring concrete in pile foundation engineering, the concrete is often segregated due to complex properties of soil layers, which lead to a weak pile segment in the pile body. The weak pile segment has a significant influence on its bearing capacity. Therefore, the influence of weak pile segment length is described in this section. The thickness of the upper soil is 5 m, shear modulus of upper soil is 2.4 MPa, and shear modulus of lower soil is 4 MPa; the shear modulus, mass density and Poisson's ratio of the weak pile segment are 2 GPa, 2000 kg/m^3 and 0.24, respectively. The top depth of the weak pile segment is set as 4 m (i.e., the middle part of the TP), and the weak pile segment length is denoted by l_e and set as 0.1 m, 0.2 m, 0.3 m, 0.4 m and 0.5 m in turn. The other parameters are the same as those in Section 4.

Figures 7–9 illustrate the influence of the weak pile segment on the HDI, RDI and HRDI of the pile head, respectively. The influence of the weak pile segment length on the three kinds of dynamic impedances of the pile head in the low-frequency range is less than that in the high-frequency range, and its degree of influence gradually increases as the frequency increases. Compared with the damping factor of the pile head, the stiffness factor of the pile head is more influenced by the weak pile segment length. When the concrete of

pile body is segregated, the absolute values of the three kinds of dynamic impedances of pile head gradually decrease, and its reduction rate is negatively correlated with weak pile segment length. Additionally, if the weak pile segment length is greater than 0.4 m, the three kinds of dynamic impedances of the pile head are essentially unaffected by the weak pile segment length.

5.4. Influence of the Weak Pile Segment Position

To investigate the influence of the weak pile segment position on the horizontal vibration characteristics of the TP, the weak pile segment length is set as 0.2 m and, in turn, can be set at the top, upper part, middle part, lower part and bottom of the TP. In other words, the depths of the top surface of the weak pile segment are set as 0.2 m, 2 m, 4 m, 6 m and 7.5 m, respectively.

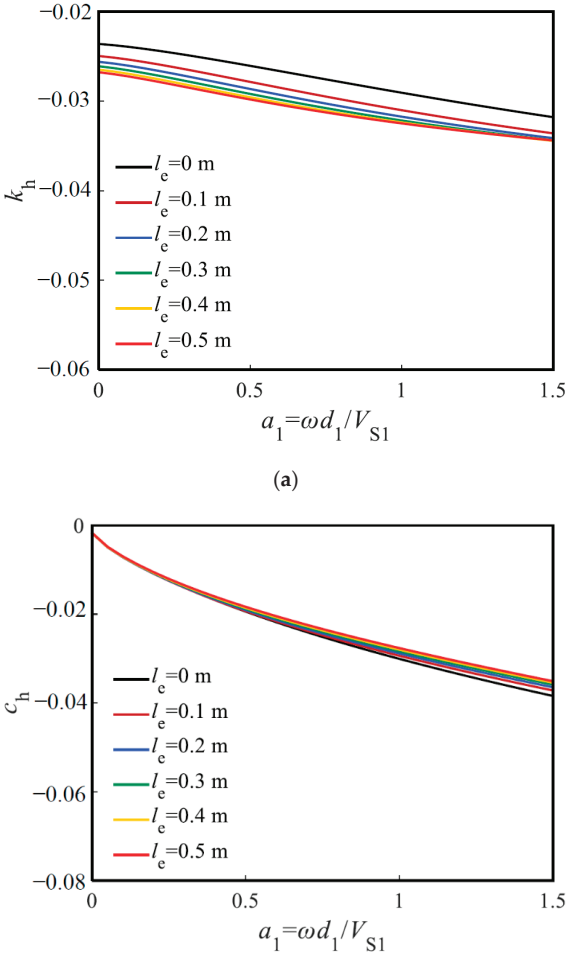


Figure 7. Influence of weak pile segment length on the HDI of pile head. (**a**) Stiffness factor of HDI and (**b**) damping factor of HDI.

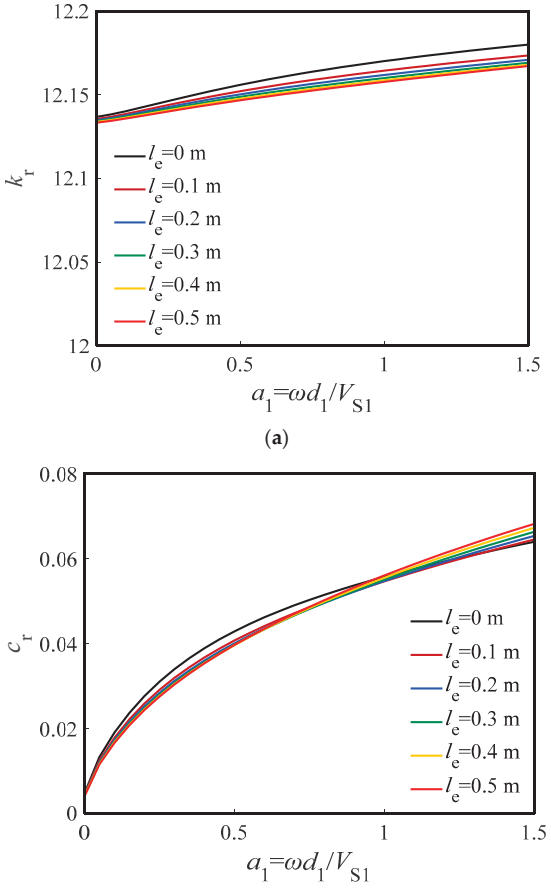


Figure 8. Influence of weak pile segment length on the RDI of pile head. (**a**) Stiffness factor of RDI and (**b**) damping factor of RDI.

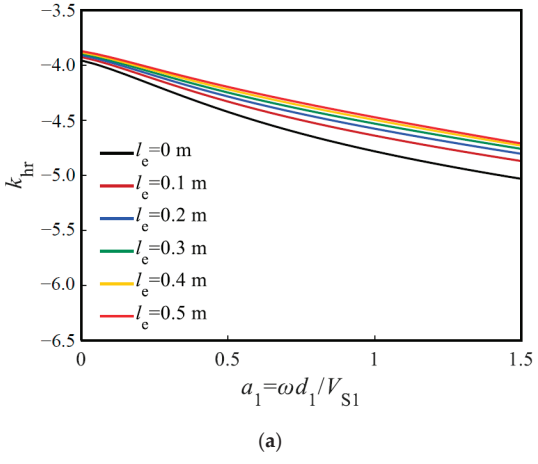
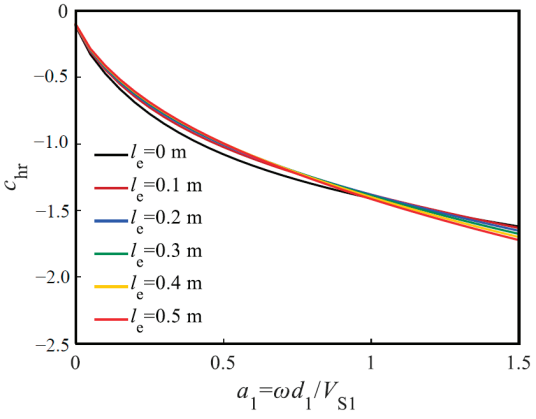


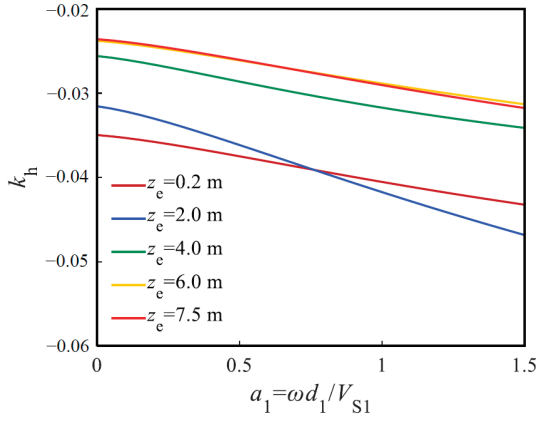
Figure 9. *Cont*.



(**b**)

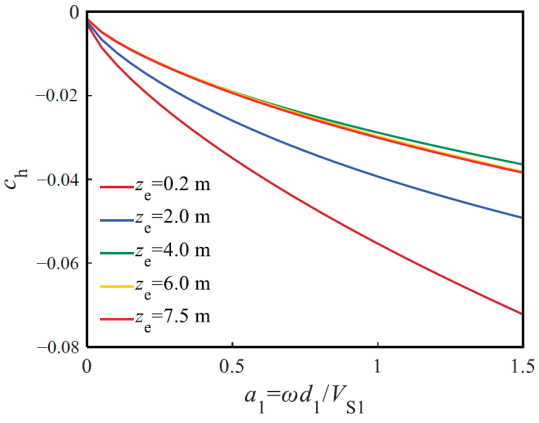
Figure 9. Influence of weak pile segment length on the HRDI of pile head. (**a**) Stiffness factor of HRDI and (**b**) damping factor of HRDI.

Figures 10–12 depict the influence of the weak pile segment position on the HDI, RDI and HRDI of the pile head, respectively. The closer the weak pile segment position to the pile head, the greater the absolute values of the three kinds of dynamic impedances of the pile head. When the weak pile segment depth increases, the absolute values of the three kinds of dynamic impedances decrease, and its reduction rate decreases with the increase in the depth of the weak pile segment. When the depth of weak pile segment is greater than 6 m, the three kinds of dynamic impedances of the pile head are essentially unaffected by the depth of the weak pile segment. This phenomenon indicates that, when the weak pile segment occurs at the lower part or the bottom of TP, its influence on the HDI of TP is smaller than that at other parts. Furthermore, in the whole frequency range, the weak pile segment position has relatively little influence on the three kinds of stiffness factors of pile head, but the three kinds of damping factors are greatly affected by the weak pile segment position in the high-frequency range. This phenomenon indicates that a high-frequency load will greatly harm the TP with the weak pile segment near the pile head and could lead to serious malfunctions in engineering quality. Therefore, attention should be paid to prevent the occurrence of a weak pile segment near the pile head in practical engineering.



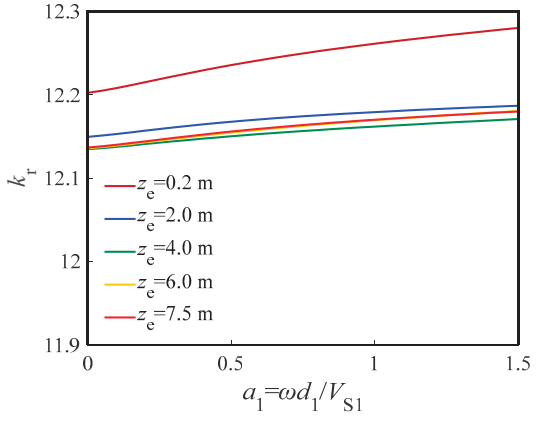
(**a**)

Figure 10. *Cont.*

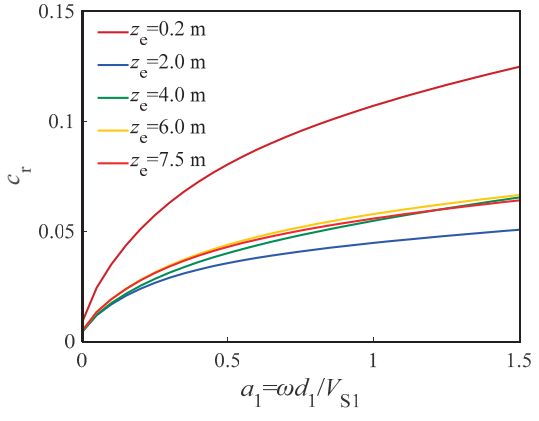


(**b**)

Figure 10. Influence of weak pile segment position on the HDI of pile head. (**a**) Stiffness factor of HDI and (**b**) damping factor of HDI.



(**a**)



(**b**)

Figure 11. Influence of weak pile segment position on the RDI of pile head. (**a**) Stiffness factor of RDI and (**b**) damping factor of RDI.

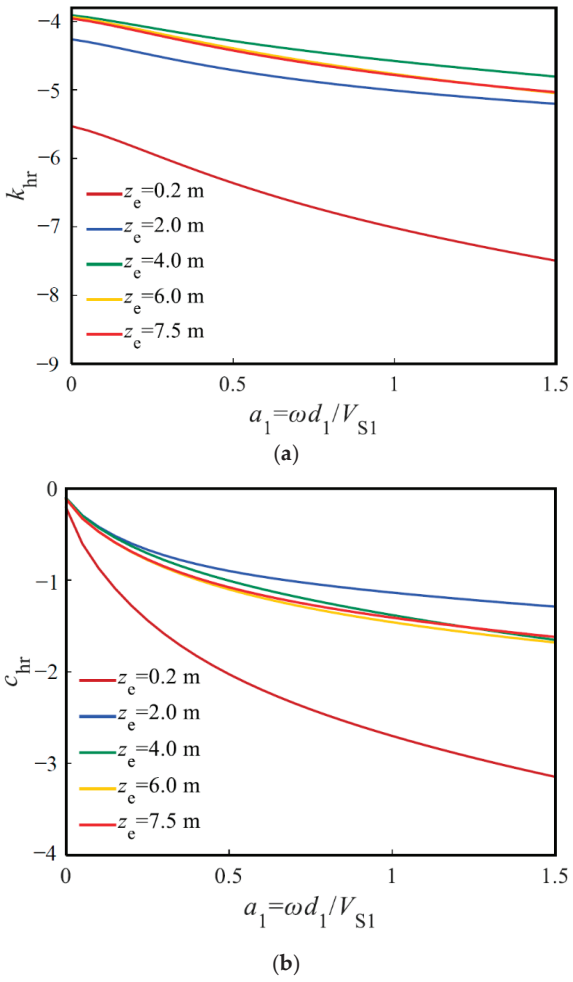


Figure 12. Influence of weak pile segment position on the HRDI of pile head. (**a**) Stiffness factor of HRDI and (**b**) damping factor of HRDI.

6. Conclusions

To investigate the horizontal vibration problems of a TP, this paper first proposes the analytical solutions for the HD, BM and SF of a TP undergoing a horizontally dynamic load. Then, the analytical solutions for the HDI, RDI and HRDI of the TP are also derived. Based on the present solutions, a parametric study is carried out to study the effect of soil and pile properties on the horizontal vibration characteristics of TPs. The main conclusions are summarized as follows:

(1) In practical engineering, the ability of the TP–soil system to resist horizontal vibrations can be improved by strengthening the upper soil. There is a critical-influence thickness for the influence of upper soil thickness on the horizontal vibration characteristics of TP. If the reinforcement depth exceeds the critical-influence thickness, increasing the reinforcement depth will not further enhance the ability of the TP–soil system to resist horizontal vibration.

(2) The effect of the weak pile segment length on the horizontal vibration characteristics of the TP in the low-frequency range is less than that in the high-frequency range.

Compared with the damping factor of dynamic impedance of the TP, the stiffness is more affected by the weak pile segment length.

(3) A closer position of the weak pile segment to the pile head can lead to greater absolute values for the three kinds of dynamic impedances of the pile head. When the depth of weak pile segment increases, the absolute values of the three kinds of dynamic impedance of the pile head decrease, and its reduction rate decreases with the increase in depth.

One limitation of this paper is that the interaction between the pile cap and soil is ignored, which may lead to a difference between the theoretical results of TP and its actual engineering characteristics. In the follow-up research, this team of authors will consider the influence of the dynamic interactions between pile cap and soil on the horizontal vibration characteristics of TP.

Author Contributions: X.Y.: analytical analysis, writing—original draft. G.J.: project administration, formal analysis. H.L.: analytical analysis, writing—review and editing. W.W.: project administration, funding acquisition, conceptualization. G.M.: formal analysis, review and editing. Z.Y.: Analytical analysis, writing—original draft. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the Outstanding Youth Project of Natural Science Foundation of Zhejiang Province (Grant No. LR21E080005), and the National Natural Science Foundation of China (Grant Nos. 52178371, 51878634, 52108347, 52108355, 52178321), and the Exploring Youth Project of Zhejiang Natural Science Foundation (Grant No. LQ22E080010).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the data used in this research can be easily accessible by downloading the various documents appropriately cited in the paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. El Naggar, M.H.; Wei, J.Q. Response of tapered piles subjected to lateral loading. *Can. Geotech. J.* **1999**, *36*, 52–71. [CrossRef]
2. El Naggar, M.H.; Wei, J.Q. Axial capacity of tapered piles established from model tests. *Can. Geotech. J.* **1999**, *36*, 1185–1194. [CrossRef]
3. Ghazavi, M.; Ahmadi, H.A. Long-term capacity of driven non-uniform piles in cohesive soil–field load tests. In Proceedings of the 8th International Conference on the Application of Stress Wave Theory to Piles, Lisbon, Portugal, 9 July–9 December 2008; pp. 132–139.
4. Zil'berberg, S.D.; Sherstnev, A.D. Construction of compaction tapered pile foundations. *Soil Mech. Found. Eng.* **1990**, *27*, 96–101. [CrossRef]
5. Wang, L.X.; Wu, W.B.; Zhang, Y.P.; Li, L.C.; Liu, H.; El Naggar, M.H. Nonlinear analysis of single pile settlement based on stress bubble fictitious soil pile model. *Int. J. Numer. Anal. Meth.* **2022**. [CrossRef]
6. El Naggar, M.H.; Mohammed, S. Evaluation of axial performance of tapered piles from centrifuge tests. *Can. Geotech. J.* **2000**, *37*, 1295–1308. [CrossRef]
7. Li, L.C.; Liu, H.; Wu, W.B.; Wen, M.J.; El Naggar, M.H.; Yang, Y.Z. Investigation on the behavior of hybrid pile foundation and its surrounding soil during cyclic lateral loading. *Ocean Eng.* **2021**, *240*, 110006. [CrossRef]
8. Li, L.C.; Zheng, M.Y.; Liu, X.; Wu, W.B.; Liu, H.; El Naggar, M.H.; Jiang, G.S. Numerical analysis of the cyclic loading behavior of monopile and hybrid pile foundation. *Comput. Geotech.* **2022**, *144*, 104635. [CrossRef]
9. Saha, S.; Ghosh, D. Vertical vibration of tapered piles. *J. Geotech. Eng.* **1986**, *112*, 290–302. [CrossRef]
10. Xie, J.; Vaziri, H.H. Vertical vibration of nonuniform piles. *J. Eng. Mech.* **1991**, *117*, 1105–1118. [CrossRef]
11. Ghazavi, M. Response of tapered piles to axial harmonic loading. *Can. Geotech. J.* **2008**, *45*, 1622–1628. [CrossRef]
12. Wu, W.B.; Jiang, G.S.; Dou, B.; Leo, C.J. Vertical dynamic impedance of tapered pile considering compacting effect. *Math. Probl. Eng.* **2013**, *2013*, 304856. [CrossRef]
13. El Naggar, M.H. Vertical and torsional soil reactions for radially inhomogeneous soil layer. *Struct. Eng. Mech.* **2000**, *10*, 299–312. [CrossRef]
14. Wu, W.B.; Yang, Z.J.; Liu, X.; Zhang, Y.P.; Liu, H.; El Naggar, M.H.; Xu, M.J.; Mei, G.X. Horizontal dynamic response of pile in unsaturated soil considering its construction disturbance effect. *Ocean Eng.* **2022**, *245*, 110483. [CrossRef]

15. Zhang, Y.P.; Jiang, G.S.; Wu, W.B.; El Naggar, M.H.; Liu, H.; Wen, M.J.; Wang, K.H. Analytical solution for distributed torsional low strain integrity test for pipe pile. *Int. J. Numer. Anal. Meth.* **2022**, *46*, 47–67. [CrossRef]
16. Zhang, Y.P.; Liu, H.; Wu, W.B.; Wang, S.; Wu, T.; Wen, M.J.; Jiang, G.S.; Mei, G.X. Interaction model for torsional dynamic response of thin-wall pipe piles embedded in both vertically and radially inhomogeneous soil. *Int. J. Geomech.* **2021**, *21*, 04021185. [CrossRef]
17. Wu, W.B.; Liu, H.; Yang, X.Y.; Jiang, G.S.; El Naggar, M.H.; Mei, G.X.; Liang, R.Z. New method to calculate apparent phase velocity of open-ended pipe pile. *Can. Geotech. J.* **2020**, *57*, 127–138. [CrossRef]
18. Huang, Y.M.; Wang, P.G.; Zhao, M.; Zhang, C.; Du, X.L. Dynamic responses of an end-bearing pile subjected to horizontal earthquakes considering water-pile-soil interactions. *Ocean Eng.* **2021**, *238*, 109726. [CrossRef]
19. Huang, K.; Sun, Y.W.; Zhou, D.Q.; Li, Y.J.; Jiang, M.; Huang, X.Q. Influence of water-rich tunnel by shield tunneling on existing bridge pile foundation in layered soils. *J. Cent. South Univ.* **2021**, *28*, 2574–2588. [CrossRef]
20. Wang, J.; Zhou, D.; Zhang, Y.Q.; Cai, W. Vertical impedance of a tapered pile in inhomogeneous saturated soil described by fractional viscoelastic model. *Appl. Math. Model* **2019**, *75*, 88–100. [CrossRef]
21. Bryden, C.; Arjomandi, K.; Valsangkar, A. Dynamic axial stiffness and damping parameters of tapered piles. *Int. J. Geomech.* **2018**, *18*, 06018014. [CrossRef]
22. Bryden, C.; Arjomandi, K.; Valsangkar, A. Dynamic axial response of tapered piles including material damping. *Pract. Period. Struct.* **2020**, *25*, 04020001. [CrossRef]
23. Wu, W.B.; Jiang, G.S.; Lü, S.H.; Huang, S.G.; Xie, B.H. Torsional dynamic impedance of a tapered pile considering its construction disturbance effect. *Mar. Georesour. Geotec.* **2016**, *34*, 321–330. [CrossRef]
24. Guan, W.J.; Wu, W.B.; Jiang, G.S.; Leo, C.J.; Deng, G.D. Torsional dynamic response of tapered pile considering compaction effect and stress diffusion effect. *J. Cent. South Univ.* **2020**, *27*, 3839–3851. [CrossRef]
25. Ghazavi, M. Analysis of kinematic seismic response of tapered piles. *Geotech. Geol. Eng.* **2007**, *25*, 37–44. [CrossRef]
26. Lee, J.K.; Park, S.H.; Kim, Y. Transverse free vibration of axially loaded tapered friction piles in heterogeneous soil. *Soil Dyn. Earthq. Eng.* **2019**, *117*, 116–121. [CrossRef]
27. Gupta, B.K.; Basu, D. Applicability of Timoshenko, Euler–Bernoulli and rigid beam theories in analysis of laterally loaded monopiles and piles. *Géotechnique* **2018**, *68*, 772–785. [CrossRef]
28. Gupta, B.K.; Basu, D. Timoshenko beam theory–based dynamic analysis of laterally loaded piles in multilayered viscoelastic soil. *J. Eng. Mech.* **2018**, *144*, 04018091. [CrossRef]
29. Ding, X.M.; Luan, L.B.; Zheng, C.J.; Zhou, W. Influence of the second-order effect of axial load on lateral dynamic response of a pipe pile in saturated soil layer. *Soil Dyn. Earthq. Eng.* **2017**, *103*, 86–94. [CrossRef]
30. Zheng, C.J.; Luan, L.B.; Qin, H.Y.; Zhou, H. Horizontal dynamic response of a combined loaded large-diameter pipe pile simulated by the Timoshenko bean theory. *Int. J. Struct. Stab. Dyn.* **2019**, *20*, 2071003. [CrossRef]
31. Gazetas, G.; Dobyr, R. Horizontal response of piles in layered soils. *J. Geotech. Eng.* **1984**, *110*, 20–40. [CrossRef]
32. Hu, A.F.; Xie, K.H.; Ying, H.W.; Qian, L. Analytical theory of lateral vibration of single pile in visco-elastic subgrade considering shear deformation. *Chin. J. Rock. Mech. Eng.* **2004**, *23*, 1515–1520. (In Chinese)

sensors

MDPI

Article

Analysis of the Effect of Velocity on the Eddy Current Effect of Metal Particles of Different Materials in Inductive Bridges

Wei Li [1], Shuang Yu [1], Hongpeng Zhang [1,*], Xingming Zhang [1,2], Chenzhao Bai [1], Haotian Shi [1], Yucai Xie [1], Chengjie Wang [1], Zhiwei Xu [1], Lin Zeng [3] and Yuqing Sun [1]

[1] School of Marine Engineering, Dalian Maritime University, Dalian 116026, China; dmuliwei@dlmu.edu.cn (W.L.); yush312@163.com (S.Y.); zhxm@hit.edu.cn (X.Z.); baichenz@163.com (C.B.); dmu6hao@163.com (H.S.); xyc86418332@163.com (Y.X.); wangcj@dlmu.edu.cn (C.W.); xuzhiwei201809@163.com (Z.X.); sunyq@dlmu.edu.cn (Y.S.)
[2] School of Naval Architecture and Ocean Engineering, Harbin Institute of Technology, Weihai 264209, China
[3] Laboratory of Biomedical Microsystems and Nano Devices, Bionic Sensing and Intelligence Center, Institute of Biomedical and Health Engineering, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China; lin.zeng@siat.ac.cn
* Correspondence: zhppeter@dlmu.edu.cn

Abstract: A method for analyzing the influence of velocity changes on metal signals of different materials in oil detection technology is proposed. The flow rate of metal contaminants in the oil will have a certain impact on the sensitivity of the output particle signal in terms of electromagnetic fields and circuits. The detection velocity is not only related to the sensitivity of the output particle signal, but also to the adaptability of high-speed and high-throughput in oil online monitoring. In this paper, based on a high-sensitivity inductive bridge, the eddy current effect of velocity in a time-harmonic magnetic field is theoretically analyzed and experimentally verified, the phenomenon of particle signal variation with velocity for different materials is analyzed and discussed, and finally the effect of velocity on the output signal of the processing circuit is also elaborated and experimentally verified. Experiments show that under the influence of the time-harmonic magnetic field, the increase of the velocity enhances the detection sensitivity of non-ferromagnetic metal particles and weakens the detection sensitivity of non-ferromagnetic particles. Under the influence of the processing circuit, different velocities will produce different signal gains, which will affect the stability of the signal at different velocities.

Keywords: inductance bridge; particle material; velocity; eddy current effect; oil detection

Citation: Li, W.; Yu, S.; Zhang, H.; Zhang, X.; Bai, C.; Shi, H.; Xie, Y.; Wang, C.; Xu, Z.; Zeng, L.; et al. Analysis of the Effect of Velocity on the Eddy Current Effect of Metal Particles of Different Materials in Inductive Bridges. *Sensors* **2022**, *22*, 3406. https://doi.org/10.3390/s22093406

Academic Editors: Phong B. Dao, Liang Yu and Lei Qiu

Received: 27 March 2022
Accepted: 27 April 2022
Published: 29 April 2022

1. Introduction

The Real Economy is the foundation of a country's economic development, involving industry, agriculture, construction, and other fields. The operation stability of various heavy machinery in the real economy determines the engineering efficiency and quality. Therefore, mechanical fault diagnosis is directly related to the economic cost of engineering and indirectly related to the development of the country's real economy. In the context of such engineering needs, the fault diagnosis of mechanical equipment shows its criticality. Fault diagnosis is a state monitoring technology for mechanical equipment during operation. By grasping the noise [1], vibration [2], temperature [3], oil quality [4,5], and other information of mechanical equipment during operation, it is possible to realize the type judgment of mechanical faults, fault location, and early prediction of mechanical faults [6]. There are four main means of fault diagnosis techniques [7]: noise detection, vibration detection, temperature detection, and oil detection. Among them, the oil detection technology [8,9] is widely used in engineering because it can detect various information in the mechanical system, so as to obtain more comprehensive fault information such as mechanical wear degree, fault type, and fault location. The detected information includes oil viscosity [10],

air bubbles [11], water content [12], and metal particle concentration [13,14], size [15], type [16] in the oil.

However, there are still problems of the low sensitivity of large-throughput detection and low-throughput detection which is not conducive to online detection [17]. Solving the contradiction has become the research direction of researchers in related fields. In 2009, Murali S et al. [18] designed a microfluidic oil detection chip based on the principle of capacitive Coulter counting, which combined the microfluidic chip with the oil detection technology, greatly reducing the size of the microfluidic oil detection chip and improving the sensitivity of metal particle detection in oil. In 2011, Du L et al. [19] designed a microfluidic oil detection chip based on the inductance Coulter principle, which can realize the differentiated identification of ferrous metals and non-ferrous metals. Zhang X et al. [20] studied the detection mechanism of the inductive oil detection chip, established a system model between the detection coil and metal particles, and obtained the mathematical model between the number of turns, density, excitation frequency, metal particle size, and output inductance of the detection coil. Respectively, Ma L et al. [21], Li Y et al. [22] used the mutual inductance between double coils and the optimization of triple coil structure to improve the detection sensitivity of the sensor. Davis J P et al. [23] realized the high-sensitivity identification of metal particles under the condition of the simplified circuit by using Maxwell bridge and subsequent filtering, amplifying, and rectifying circuits. As mentioned above, many researchers have been working to achieve high-sensitivity differential detection from various aspects, including structural design [24], circuit design [25], magnetic material addition [26]. To further realize online monitoring of mechanical faults and expand the detection throughput with high detection sensitivity, a series of studies have been carried out by the researchers concerned. Zhu X et al. [27] adopted the time-division multiplexing method to collect the sensing data of 9 channels in sequence, and added series diodes to eliminate the crosstalk between each channel. Jagtiani A V et al. [28] adopted the frequency division multiplexing method to modulate and demodulate multiple signals with different frequency carriers, increasing the detection throughput by 4 times. Bai C et al. [29] designed a high-throughput sensor with an annular flow channel structure, expanding the flow to 16 times that of the original microchannel.

Other related researchers started from the perspective of detection speed, explored the possibility of online monitoring of oil, and explored the impact of speed on detection sensitivity. Wang X et al. [30] explored the influence of particle velocity on the aliasing signal waveform, and put forward the conclusion that the aliasing induced voltage is proportional to the particle velocity. Wu Y et al. [31] deduced the analytical formula for the influence of the flow rate on the magnetic field, and proposed that the detection sensitivity of copper particles can be enhanced by increasing the oil detection flow rate. Liu E et al. [32] studied the influence of flow velocity on the detection signal, and proposed that the amplitude of the inductive signal of metal particles can be improved by appropriately reducing the flow velocity. The research results of the above researchers are valuable, but there is still controversy about the effect of flow rate on abrasive particles of different materials.

In this study, based on an inductive bridge, we explore the effect of particle velocities of different materials on the eddy current effect. The influence of particle motion velocity on the output signal of detected particles is analyzed by using the angle difference between the direction of particle motion and the direction of magnetic field lines on both sides of the coil. The characteristics of high-speed and high throughput regard online monitoring are fitted under the premise of high detection sensitivity. Compared with previous analysis of velocity, this method can visually verify the trend of particle velocity to signal change without interference from electromagnetic and detection circuits. In this paper, an inductive bridge and a processing circuit are used to replace the impedance analyzer, so the sensor still has high detection sensitivity under the premise of portability and low cost. And the content explored in this paper is in line with the actual engineering, which is conducive to the development of online monitoring. On this basis, it is committed to the realization

of online detection of oil, by exploring the influence of speed on the detection of metal particles of different materials.

2. Sensor Design and Theory Analysis

2.1. Inductance Detection Analysis

The design of the sensor chip is shown in Figure 1. The inside of the sensing chip includes the particle inlet and outlet, the detection channel, and the inductive bridge. In the inductive bridge, one of the spiral coils wrapped around the outside of the detection channel is the induction coil and the other spiral coil is the balance coil. Fixed resistance, variable resistance and two solenoid coils form an electric bridge, and the inductive pulses generated by particles passing through the induction coils are converted into differential voltage pulses by the inductive bridge. As shown in Figure 1, the two nodes on the bridge were connected to the AC excitation, and the voltage signals output by the other two nodes were connected to the subsequent detection circuit.

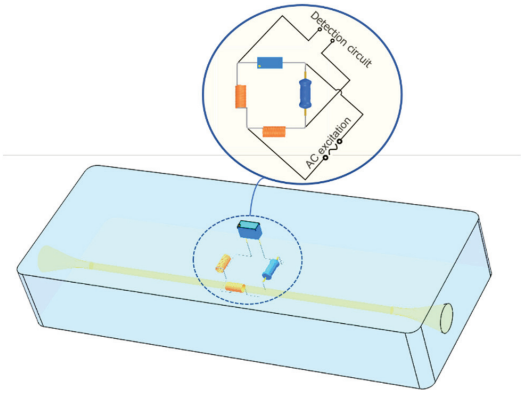


Figure 1. Schematic diagram of microfluidic chip sensing unit.

According to previous studies [33], the amount of inductance change produced by the particle at the central axis of the solenoid coil is:

$$\Delta L_x = \mathrm{Im}\left(\frac{\Delta z_{\max}}{\omega}\right) = \frac{4\pi\mu_0 N^2}{w^2 + d^2}\mathrm{Re}(k_p) \tag{1}$$

where $\Delta z_{\max}$ is the impedance output by the coil detecting the particle, ω is the angular frequency of the AC excitation, μ_0 is the vacuum permeability, N is the number of turns of the coil, w is the axial length of the coil, and d is the inner diameter of the coil.

The magnetization factor k_p of ferromagnetic metal particles is:

$$k_p = \frac{\left(-r^2k^2 + 2\mu_r + 1\right)\sin(rk) - rk(2\mu_r + 1)\cos(rk)}{\left(r^2k^2 + \mu_r - 1\right)\sin(rk) - rk(\mu_r - 1)\cos(rk)} \cdot \frac{r^3}{2} \tag{2}$$

The magnetization factor k_p of the non-ferromagnetic metal particles is:

$$k_p = \frac{1}{2} \cdot \left[r^3 + \frac{3r^2}{k}\cot(rk) - \frac{3r}{k^2}\right] \tag{3}$$

where r is the particle radius and μ_r is the relative magnetic permeability of the particle, which is assumed to be a real number in this paper and is not affected by frequency.

$$k = \sqrt{-j\omega\mu_r\mu_0\sigma} \tag{4}$$

where σ is the electrical conductivity.

When an AC excitation is applied to the coil, an alternating magnetic field is generated inside the coil. Since the relative permeability of ferromagnetic metal particles is much greater than 1, the magnetization effect generated when the ferromagnetic metal particles pass through the induction coil causes the magnetic flux in the coil to increase abruptly. However, the induced current is generated inside the ferromagnetic metal particles to obstruct the change of the original magnetic field, and the induced current generates magnetic flux in the opposite direction of the original magnetic field, which weakens the original magnetic field. That is, the particles produce eddy current effect in the time-harmonic magnetic field. For ferromagnetic particles, the magnetization effect is stronger than the eddy current effect, so the amount of change in coil inductance caused by ferromagnetic metal particles is positive.

Since the relative permeability of non-ferromagnetic metal particles is slightly less than 1, the eddy current effect dominates when the non-ferromagnetic metal particles pass through the induction coil, causing the original magnetic field to be weakened, so the change in coil inductance caused by the non-ferromagnetic metal particles is negative. The sensor can differentiate and detect ferromagnetic and non-ferromagnetic metal particles in the oil, based on the Coulter principle. The tribological information reflected by the detected metal particle information (including the number of metal particles obtained by the number of inductance signals, and the particle size of the metal particles obtained by the amplitude of the inductance signal) can be used to predict the degree of failure of the system.

2.2. Design and Principle Analysis of Inductor Bridge

Considering the bridge as an equivalent circuit, the equivalent circuit is shown in Figure 2 A, B, C, D are the endpoints on the bridge arms; L_x is the inductance of the induction coil; R_x is the internal resistance of the induction coil; L_n is the inductance of the reference coil; R_n is the internal resistance of the reference coil; R_a and R_b are the balance resistances on the two bridge arms, respectively.

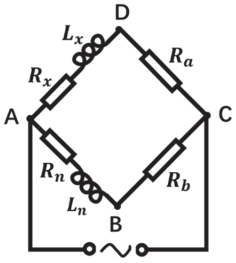


Figure 2. Inductive bridge equivalent circuit.

Simulate the equivalent circuit of the inductive bridge. The frequency of the AC excitation source was set to 1 MHz, and the voltage was set to 10 V. Explore the effect of resistance parameters on bridge sensitivity. First, change the resistance values of R_a and R_b at the same time; second, change the resistance values of the coil internal resistance R_x and R_n at the same time. Under the condition that the inductance base value of the two coils L_x, L_n was 1 μH and the inductance change ΔL_x was 0.1 μH, the change of the coil internal resistance had a great influence on the voltage difference U_{BD} between the two points B and D. As shown in Figure 3, the direction of the arrows shows the trend of bridge sensitivity as a function of internal resistance. As the coil internal resistance R_x decreased, the voltage value of U_{BD} increased under the condition of the same inductance variation ΔL_x, namely, the higher the quality factor of the coil, the higher the sensitivity of the bridge. Combined with the analysis of the simulation results, it is expected to select the coil with a smaller internal resistance R_x under the condition of similar base inductance L_x.

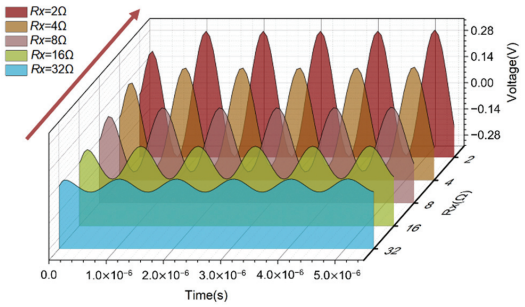


Figure 3. Effect of coil quality factor on bridge sensitivity.

As shown in Figure 4, w is the height of the coil; k is the number of layers of the coil; a is the wire diameter of the coil; d is the inner diameter of the coil.

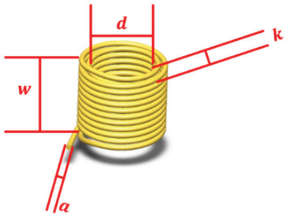


Figure 4. Schematic diagram of coil parameters.

The expression for the amount of change in coil inductance ΔL_x due to particles in Equation (1) shows that the amount of change in inductance ΔL_x is related to the coil turns N, axial length w, coil inner diameter d, particle radius r, relative permeability of particles μ_r, particle conductivity σ, and the angular frequency ω of the AC excitation applied to the coil. Among them, the factors affected by the coil itself are: the number of turns of the coil N, the axial length w, and the inner diameter of the coil d. Let $b = \frac{N^2}{w^2+d^2}$. The larger b is, the larger the change in coil inductance ΔL_x is, and the higher the sensitivity to detect particles as they pass through the coil. Since $\frac{\partial b}{\partial N} > 0$, $\frac{\partial b}{\partial w} < 0$, and $\frac{\partial b}{\partial d} < 0$, it means that the larger the number of coil turns N, the smaller the coil axis length w and the inner diameter d, the higher the coil inductance variation ΔL_x. In order to produce a large inductance variation ΔL_x between the metal particles and the coil, it is desired to increase the number of turns N of the coil, while being able to reduce the coil axis length w and the inner diameter d.

The expression for the estimated internal resistance of the spiral coil is:

$$R_x = \rho \cdot \frac{l}{S} = \rho \cdot \frac{2\pi \cdot \frac{N}{k} \cdot \sum_{i=0}^{k=i+1} \{[a(2i+1)+d]/2\}}{\pi \cdot \left(\frac{a}{2}\right)^2} \tag{5}$$

where ρ is the coil resistivity, l is the coil length, and S is the coil cross-sectional area.

It can be seen from Equation (5) that the larger the wire diameter a that determines the cross-sectional area of the coil S, the smaller the number of turns N, the axial length w, and the inner diameter d that collectively reflect the coil length l, the smaller the coil internal resistance R_x. Therefore, increasing the wire diameter of the coil a can reduce the internal resistance of the coil R_x without affecting the variation of the coil inductance ΔL_x.

Since the microfluidic chip adopts tiny coils, the wire diameter a which is smaller than the number of coil turns N, axial length w and inner diameter d is regarded as a fixed value, and the part as the numerator is ignored in the calculation of the partial derivative between the axial length w and the number of coil turns N. Then there is $\frac{\partial R}{\partial w} > 0$, $\frac{\partial R}{\partial d} > 0$, $\frac{\partial R}{\partial N} > 0$, that is, the coil resistance R_x decreases monotonically with the decrease of the

number of turns N, axial length w, and inner diameter d. Since the wire diameter a is much smaller than the number of coil layers k, the axial length w and the inner diameter d, namely, $\frac{\partial R}{\partial N} < \frac{\partial R}{\partial w}$ and $\frac{\partial R}{\partial N} < \frac{\partial R}{\partial d}$, which means the change rate of the coil resistance R_x with the number of turns N is smaller than the change rate of the resistance R_x with the axial length w and inner diameter d. Collectively, the coil resistance R_x can be reduced by decreasing the number of turns N, axial length w, and inner diameter d. The effect of a change in the number of turns N is less than the effect of a change in the axial length w and inner diameter d on the resistance R_x. Therefore, the number of turns N of the coil can be increased, while the resistance R_x can be appropriately sacrificed, and the coil inductance variation ΔL_x can be increased at the same time.

To summarize, in order to select a coil with smaller resistance R_x under the condition of larger inductance change ΔL_x, a larger coil wire diameter a, smaller axial length w, inner diameter d, and appropriately larger number of turns N can be selected.

Combined with the above theoretical analysis and debugging through experiments, the number of turns N = 230 turns, axial length w = 2.76 mm, inner diameter d = 0.54 mm, outer diameter D = 1.55 mm, wire diameter a = 0.07 mm, resistance R_x = 4 Ω of the solenoid coil were selected by the bridge sensing unit.

When the particles pass through the coil to generate an inductive signal, the two arms of the bridge generate a voltage difference [34], the voltage of the two arms are rectified into two DC pulsating voltage signals, and then filtered by the filter circuit to remove the high frequency noise signal [35], interspersed in the particle signal, and finally differentially amplified by an amplifier circuit to generate the differential voltage caused by the particles.

As shown in the block diagram of the detection circuit system in Figure 5, when the bridge is balanced:

$$Z_{AB} \cdot Z_{CD} = Z_{AD} \cdot Z_{BC} \tag{6}$$

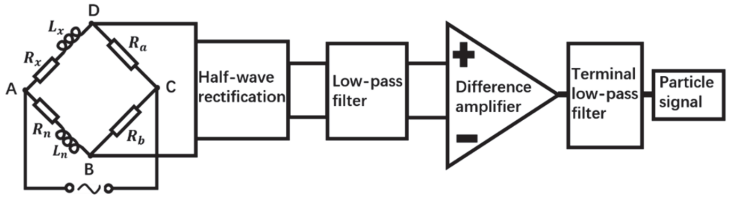


Figure 5. Schematic diagram of detection circuit system.

That is:

$$(j\omega L_x + R_x) \cdot R_b = (j\omega L_n + R_n) \cdot R_a \tag{7}$$

$$U_{BD} = U_{AD} - U_{AB} = U_i \cdot \left(\frac{Z_{AD}}{Z_{AD} + Z_{CD}} - \frac{Z_{AB}}{Z_{AB} + Z_{BC}} \right) \tag{8}$$

$$U_{BD} = U_i \{ \frac{\omega^2 \left[L_x^2 R_b (R_n + R_b) - L_n^2 R_a (R_x + R_a) \right] + (R_n + R_b)(R_x + R_a)(R_x R_b - R_n R_a)}{\left[\omega^2 L_x^2 + (R_x + R_a)^2 \right] \left[\omega^2 L_n^2 + (R_n + R_b)^2 \right]} + \frac{\omega L_x R_a \left[\omega^2 L_n^2 + (R_n + R_b)^2 \right] - \omega L_n R_b \left[\omega^2 L_x^2 + (R_x + R_a)^2 \right]}{\left[\omega^2 L_x^2 + (R_x + R_a)^2 \right] \left[\omega^2 L_n^2 + (R_n + R_b)^2 \right]} \cdot j \} \tag{9}$$

When ferromagnetic metal particles pass through an alternating time-harmonic magnetic field in the inductive bridge, the inductance value of the induction coil increases, resulting in an increase in U_{BD}, and the differential voltage is positive; when the non-ferromagnetic metal particles pass through the inductive bridge, the inductance value of the induction coil decreases, resulting in a decrease in U_{BD}, the differential voltage is negative.

In Figure 5, a half-wave rectifier circuit module using diodes was used to rectify two AC voltage signals into pulsating DC voltage signals respectively. Since the amplitude and direction of the DC signal do not change with time, the rectified pulsating DC voltage signal can realize the differential detection of ferromagnetic and non-ferromagnetic metal particles by an inductive bridge, and can be better collected and compared.

In the low-pass filter module using the two-stage wireless gain multiple feedback filter circuit, the cut-off frequency of the first-stage low-pass filter circuit was 16 kHz, and the cut-off frequency of the second-stage low-pass filter circuit was 1.6 kHz. The low-pass filter module can reduce the AC components in the two-channel pulsating DC voltage signals as much as possible, and filter out the high-frequency noise and harmonics mixed in the two-channel signals.

The differential amplifier module made the voltage difference between the two signals affected by the sensing coil and the reference coil, and the differential voltage signal combined with the reverse amplifier circuit could be amplified by 20 to 400 times.

In the previous experimental research in the laboratory, the DC signal generated by the AC voltage signal through the above-mentioned rectification, filtering and amplifying circuit still has relatively large noise. Therefore, the terminal filter circuit was adopted in the laboratory to further extract the particle signal with very low frequency. The terminal filter using the UAF42 active filter can adjust the cutoff frequency of the low-pass filter in the range of 0~5 kHz and the quality factor of the filter by changing the resistance value of the potentiometer.

The high-sensitivity detection of ferromagnetic and non-ferromagnetic metal particles in oil detection and the portability of the detection device are realized through the processing of the inductive bridge and the detection circuit.

2.3. The Effect of Velocity on the Magnetic Field of a Solenoid

The velocity of metal particles in the sensing coil is a non-negligible factor affecting the particle signal. The velocity of the metal particles passing through the sensing coil affects both the sensing signal [36] and the flux of the oil detection.

According to Faraday's law of electromagnetic induction:

$$\varepsilon = -\frac{d\Phi}{dt} \tag{10}$$

where ε is the induced electromotive force generated by the particle passing through the coil, and $\frac{d\Phi}{dt}$ is the rate of change of the magnetic flux with time.

According to earlier research in our lab [21]:

$$\varepsilon = \oint_C \vec{E}_{in} \cdot dl = -\frac{d\Phi}{dt} = -\frac{d}{dt}\iint_S \vec{B} \cdot d\vec{S} = -\iint_S \left(\frac{\partial \vec{B}}{\partial t} \cdot d\vec{S} + \vec{B} \cdot \frac{\partial}{\partial t} d\vec{S} \right) \tag{11}$$

where $\vec{E}_{in}$ is the electric field strength of the metal particle induced by the magnetic field, dl is the length differential of the vortex ring inside the metal particle, $\vec{B}$ is the magnetic induction between the metal particle and the coil, $\vec{S}$ is the area of the vortex ring, $-\iint_S \frac{\partial \vec{B}}{\partial t} \cdot d\vec{S}$ is the induced electromotive force ε_t due to the change in the magnetic field, and $-\iint_S \vec{B} \frac{\partial}{\partial t} \cdot d\vec{S}$ is the kinetic electromotive force ε_m due to the movement of the metal particle in the magnetic field.

When the metal particles pass through the alternating magnetic field in the coil, an eddy current ring is generated inside the metal, and the resulting motional electromotive force is:

$$\varepsilon_m = -\frac{d\phi_m}{dt} = \oint_C (\vec{v} \times \vec{B}) \cdot dl \tag{12}$$

where $\vec{v}$ is the moving velocity of the metal particles in the magnetic field.

According to the definition of inductance:

$$dL = \frac{d\Phi}{dI} = \frac{\oint_c (\vec{v} \times \vec{B}) \cdot dl\, dt}{I} \tag{13}$$

It can be seen from Equation (13) that in the alternating magnetic field, the coil inductance dL caused by the motional electromotive force ε_m increases with the increase of the velocity $\vec{v}$. That is, the greater the velocity, the greater the induced electromotive force containing the motional electromotive force, the stronger the obstruction of the original magnetic field by the induced electromotive force generated by the metal particles, and the more pronounced the eddy current effect inside the metal particles. This is reflected in the fact that, by increasing the velocity, the magnetization effect is weakened more by the eddy current effect in the detection of ferromagnetic metal particles, resulting in a weaker signal for ferromagnetic particles, while the eddy current effect is enhanced in the detection of non-ferromagnetic particles, resulting in a stronger signal for non-ferromagnetic particles.

3. Results and Discussion

Experiments to explore the velocity often use a plane coil, so that the metal particles move on the surface of the plane coil, that is, in the direction perpendicular to the magnetic field lines of the plane coil. This method amplifies the proportion of the motional electromotive force in the induced electromotive force, which is ideal for the study of velocity effect. However, during the actual experiments, many uncontrollable factors arise at the same time as the velocity increases, such as the small sampling rate of the equipment, particle vibration and increased noise. These factors also affect the output of the signal. Using the angle difference θ between the movement direction of the metal particles on both sides of the solenoid coil and the magnetic field line inside the solenoid, as shown in Figure 6, the influence trend of the dynamic electromotive force can be clearly seen by comparison.

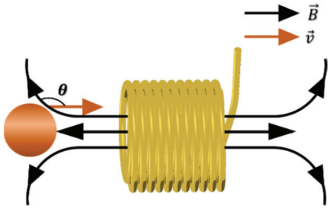


Figure 6. Schematic diagram of cutting magnetic field lines in the direction of movement of metal particles.

In this experiment, the experimental system was set up as shown in Figure 7, with the metal particles adhering to a nylon rope and a stepper motor controlling the velocity past the detection chip. The particle inductance signal sensed by the sensing coil in the detection chip is converted into two voltage signals by the inductance bridge in the detection chip. When the metal particles on the nylon rope pass through the sensing coil, a voltage difference is generated between the two voltage signals. The inductive bridge in the detection chip was given a 1.3 MHz, 10 V sinusoidal AC excitation source by a waveform generator. The subsequent processing circuit rectifies, filters, differentiates and amplifies the two voltage signals generated at points B and D of the inductive bridge, and the processing circuit was powered by a DC power supply of ±15 V DC. The voltage signal processed by the processing circuit is converted by the data acquisition card to the computer through analog-digital conversion. To avoid distortion of the sampled signal when the sampling rate was insufficient, the sampling rate was chosen to be much greater than the signal frequency. The sampling rate chosen was 2000 samples/s, much greater than the frequency of the signal required by Nyquist-Shannon sampling theorem for this experiment. The experiment proved that the sampling signal at this sampling rate was stable, with a set of signals for 350 μm iron particles at a velocity of 5 mm/s as shown in Figure 8:

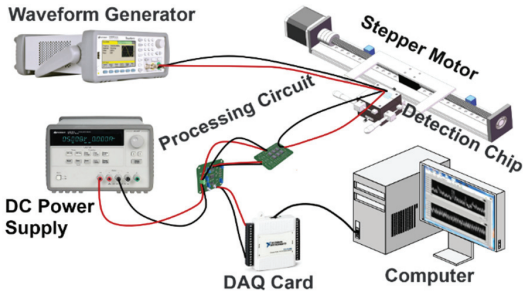


Figure 7. Experimental system diagram.

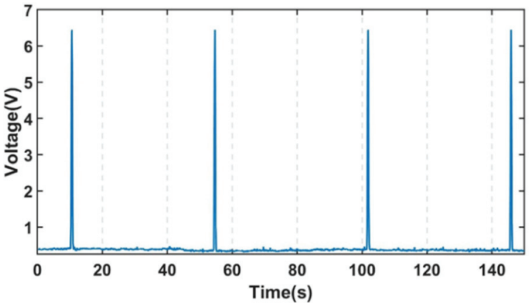


Figure 8. Steady sampling signal for 350 μm iron particles.

The experimental data of the acquired noise signal is shown in Figure 9. The noise signal is collected at each velocity for 10 s under the condition of no-load nylon rope. It can be seen in Figure 9 that the collected noise signal does not change much with the change of velocity.

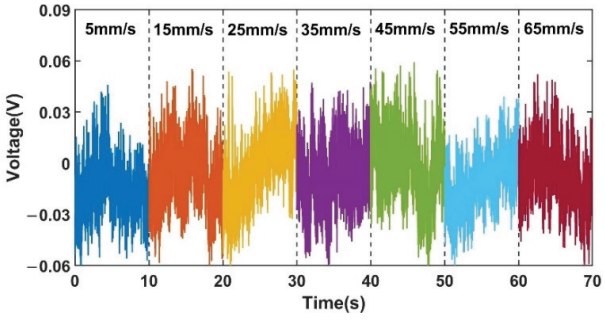


Figure 9. Noise signals at different velocities under no load.

Figure 10 shows a comparison of the signals of 350 μm ferromagnetic metal particles (350 μm Fe) at different velocities, capturing one signal from a set at each velocity, which was intercepted for 10 s. Figure 11 shows a comparison of the signals of non-ferromagnetic metal particles (350 μm Cu) at different velocities, capturing one signal from a set at each velocity, which was intercepted for 10 s. Figure 12 shows a comparison of the signals of non-ferromagnetic metal particles (350 μm Al) at different velocities, capturing one signal from a set at each velocity, which was intercepted for 10 s.

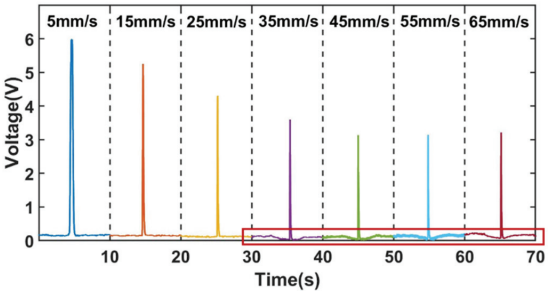


Figure 10. Comparison of signals at different velocities for 350 μm iron.

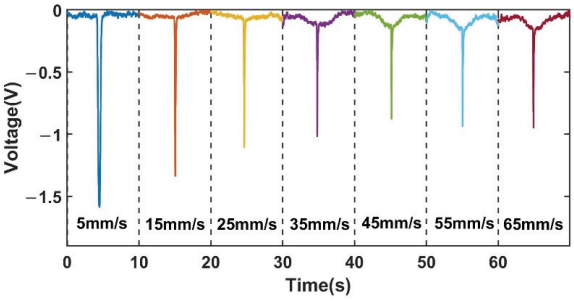


Figure 11. Comparison of signals at different velocities for 350 μm copper.

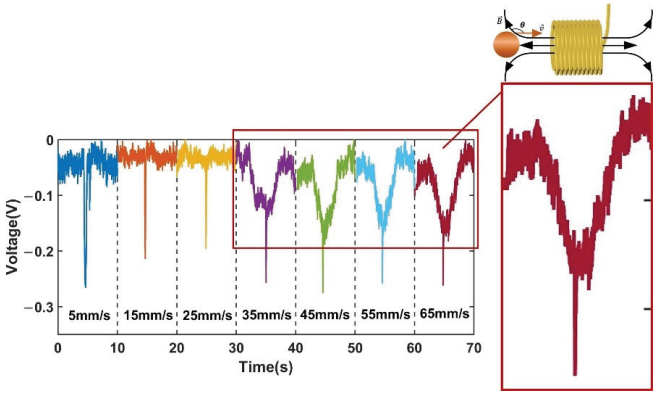


Figure 12. Comparison of signals at different velocities for 350 μm aluminum.

As can be seen in Figures 10–12, during the increase in velocity from 5 mm/s to 65 mm/s, the signal affected by the velocity of the cut magnetic induction lines before and after the spiral coil starts to increase significantly when the velocity increases to 35 mm/s. This phenomenon can be observed in the detection of iron particles, copper particles and aluminum particles. This phenomenon is caused by the angular difference θ between the moving direction of the particles and the direction of the alternating magnetic field. The change of velocity increases the induced electromotive force between the particles and the coil, and enhances the eddy current effect inside the particles. This phenomenon is consistent with the theoretical inference in 2.3.

Figure 13a shows the changing trend of the voltage signal amplitude affected by the cutting magnetic field velocity for the iron, copper, and aluminum metal particles on both sides of the coil.

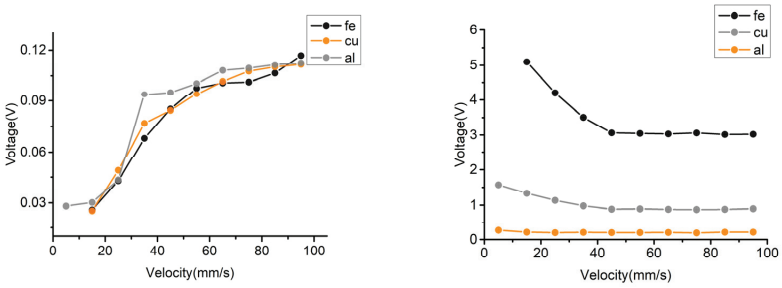


Figure 13. Diagram of signal variation with velocity: (**a**) Variation of metal particle voltage amplitude affected by velocity, (**b**) Variation of the total signal amplitude of the metal particles with varying velocity.

Although the eddy current effect caused by the enhanced induced electromotive force increases as the particle velocity increases, the amplitude of the voltage change caused by the velocity increases, resulting in a weakening of the signal of ferromagnetic metal particles and an enhancement of the signal of non-ferromagnetic metal particles. However, in general, the trend of the signal change of metal particles is still weakened, and the signal amplitude of non-ferromagnetic metal particles is still weakened. The signal weakening trend is shown in Figure 13b. The analysis of this phenomenon is presented as follows.

The increase of the detection velocity, on the one hand, for the time-harmonic magnetic field, changes the induced electromotive force between the metal particles and the coil in the magnetic field; on the other hand, for the detection circuit, it changes the frequency of the metal particle signal. That is, as the velocity of the metal particles increases, the signal frequency of the metal particles increases.

It can be seen by using the solenoid that in the time-harmonic magnetic field, the increase of the velocity leads to the enhancement of the induced electromotive force, which weakens the original magnetic induction intensity. This phenomenon is manifested as an increase in the negative detection signal caused by the induced electromotive force, which gradually increases with the increase of the detection velocity on both sides of the sensing area of the solenoid. This is reflected in the weakening of the detection signal of the ferromagnetic metal particles and the enhancement of the detection signal of the non-ferromagnetic metal particles. This experimental result is consistent with the theoretical inference in Section 2.3.

As for the non-magnetic factors during the experiment, that is, for the detection circuit, the change in the signal frequency of the metal particles causes a change in the signal output amplitude of the particles at the corresponding frequency components. According to the set stepper motor stroke, stepper motor acceleration time and the velocity of the particles passing through the sensing coil, it is known that when the particles pass through the sensing coil at a velocity of 5 mm/s, the particle signal frequency is 0.02 Hz; when the particle passes the sensing coil at a velocity of 15 mm/s, the frequency of the particle signal is 0.06 Hz; when the particle passes the sensing coil at a velocity of 25 mm/s, the frequency of the particle signal is 0.09 Hz; when the particle passes the sensing coil at a velocity of 35 mm/s, the frequency of the particle signal is 0.12 Hz; when the particle passes the sensing coil at a velocity of 45 mm/s, the frequency of the particle signal is 0.15 Hz; when the particle passes the sensing coil at a velocity of 55 mm/s, the frequency of the particle signal is 0.17 Hz; when the particle passes the sensing coil at a velocity of 65 mm/s, the frequency of the particle signal is 0.19 Hz. As can be seen, different detection velocities, will directly lead to changes in the frequency component of the acquired signal, and the change in frequency, which will further affect the effect of the processing of that frequency component in the detection circuit, namely, the change in the frequency component caused

by the velocity enhancement, will directly lead to changes in the output signal of the detection circuit.

As shown in Figure 14, the terminal filter selected in the detection circuit is taken as an example. For the selected terminal filter UAF42 low-pass filter, the cut-off frequency of the low-pass filter was adjusted to be very low at the beginning of the experiment, and the quality influence of the filter was not considered. The ripple attenuation in the passband of the low-pass filter was ignored, so the amplitude and frequency change of the particle signal when passing through the filtering and amplifying circuit in the detection circuit was not considered. As shown in Figure 15, the characteristic curves of Chebyshev-type low-pass filter vary in trend for different component parameters. The low-pass filter produces different amounts of equal ripple undulations d', d'' in the low frequency passband.

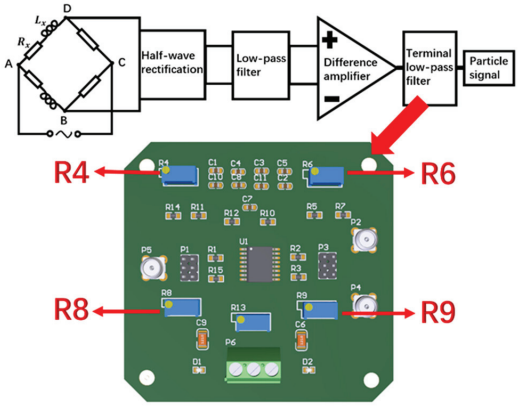


Figure 14. Terminal UAF42 low pass filter.

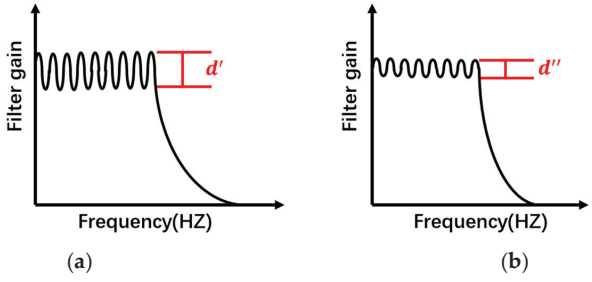


Figure 15. Characteristic curves of low-pass filter circuits with different design parameters: (**a**) Large passband ripple, (**b**) Small passband ripple.

Low-pass filters designed with different component parameters have different characteristic curves. When the ripple in the passband of the filter is large, the amplitude gain of the output signal fluctuates greatly due to the change of the signal frequency, as shown in Figure 15a; When the ripple in the passband of the filter is small, the amplitude gain of the output signal fluctuates less due to the change of the signal frequency, as shown in Figure 15b. Therefore, it is considered that as the detection velocity increases, the signal frequency component changes, resulting in a change in the output signal gain amplitude of the detection circuit.

Based on the above hypothesis, the experiment was further designed. By re-adjusting the parameters of the low-pass filter element, the amplitude-frequency characteristics [37] of the low-pass filter circuit were changed, and the response of the circuit to the signal frequency component that changes with the velocity was changed. The two potentiometers

R_4 and R_8 resistors of the low-pass filter were adjusted appropriately to change its quality factor; the two potentiometers R_6 and R_9 resistors of the low-pass filter were adjusted appropriately to change its low-pass cut-off frequency. The experimental results for the detection of 350 μm iron particles and 350 μm copper particles are shown in Figures 16 and 17, respectively. In Figure 16, A and B both represent 350 μm iron particles. Signal A is the voltage signal of 350 μm iron particles before adjusting the detection circuit, which is represented by the blue solid line; Trend A is the signal trend with increasing velocity, which is represented by the blue dashed line. Signal B is the voltage signal of 350 μm iron particles after adjusting the detection circuit, which is represented by the red solid line; Trend B is the signal trend with increasing velocity, which is represented by the red dashed line. In Figure 17, C and D both represent 350 μm copper particles. Signal C is the voltage signal of 350 μm copper particles before adjusting the detection circuit, which is represented by the blue solid line; Trend C is the signal trend with increasing velocity, which is represented by the blue dashed line; Signal D is the voltage signal of 350 μm copper particles after adjusting the detection circuit, which is represented by the red solid line; Trend D is the signal trend with increasing velocity, which is represented by the red dashed line.

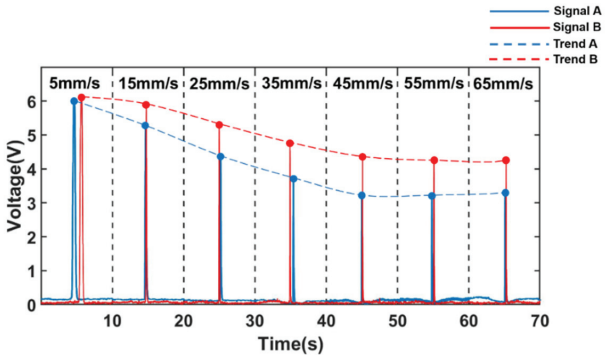


Figure 16. Signal change of ferromagnetic metal particles (350 μm Fe) with velocity change after adjusting the detection circuit.

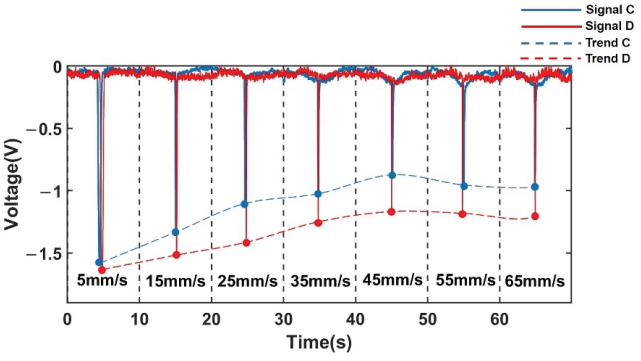


Figure 17. Signal change of non-ferromagnetic metal particles (350 μm Cu) with velocity change after adjusting the detection circuit.

It can be observed through two sets of experiments that after adjusting the detection circuit, the attenuation of the signal becomes slower.

Compared to the signal change amplitude of metal particles with velocity change before the adjustment of the detection circuit, the signal amplitude attenuation of iron particles was reduced from 2.6838 V to 1.8582 V, when the metal particle velocity was

increased from 5 mm/s to 65 mm/s after a slight adjustment to the detection circuit, as shown in Figure 16, and the signal amplitude attenuation was reduced by 0.8256 V. As shown in Figure 17, the signal amplitude attenuation of copper particles was reduced from 0.7469 V to 0.5097 V, and the signal amplitude attenuation was reduced by 0.2372 V. This shows that the attenuation of the total signal is indeed affected by the filter circuit. Similarly, the amplifying ability of the amplifying circuit to each frequency component is also different. It is inferred that the attenuation of the total signal amplitude is indeed influenced by the amplitude-frequency characteristics in the filtering and amplifying circuits in the detection circuit.

Another reason for the analysis of the total signal attenuation is that the experimental setup of the stepper motor itself has a high current, generating non-negligible electromagnetic interference, which leads to the enhancement of low-level noise and the weakening of high-level signals through radiation interference, ultimately leading to the attenuation of the detection signal amplitude, but the influencing factor is the effect on the overall frequency component, independent of the particle velocity.

The experimental results show that in a time-harmonic magnetic field, an increase in particle velocity increases the eddy current effect in the magnetic field, resulting in a decrease in signal amplitude for ferromagnetic metal particles and an increase in signal amplitude for non-ferromagnetic metal particles. Simultaneously, the quality of the detection circuit has an impact on the signals of particles with different velocities, and the design of the detection circuit should be optimized to improve the stability of the output signal in the passband.

4. Conclusions

We analyzed and summarized the theory and law of the electromagnetic field for inductance differential detection, deduced the formula that causes the voltage difference of the inductive bridge to change, and optimized the subsequent processing circuit. The designed sensor can achieve low cost and portability under the premise of ensuring higher sensitivity. Based on the inductive bridge, this paper draws the following conclusions by studying the effect of velocity on particles of different materials.

In terms of the sensitive influence of the coil quality factor on the inductive bridge: Through the simulation and analysis of the equivalent circuit of the inductive bridge, it is concluded that the higher the quality factor of the coil, the higher the detection sensitivity of the inductive bridge. The quality factor of the coil is mainly affected by N, w, d. To improve the detection sensitivity of the inductive bridge, it is possible to increase the coil diameter a and reduce w and d, and appropriately increase N. Based on this method, suitable sensing and reference coils are selected for the inductive bridge.

In terms of the effect of velocity on the electromagnetic field: In the alternating magnetic field, the general law that velocity affects the inductance signal was obtained, which was verified by theoretical analysis and experimental phenomena in the magnetic field. The increased velocity will enhance the eddy current effect of the particles in the magnetic field, which will further increase the signal of non-ferromagnetic metal particles and weaken the signal of ferromagnetic metal particles.

In terms of the effect of velocity on the processing circuit: We assumed that the velocity also affects the output signal of the processing circuit, due to the abnormal weakening of the detected signal. The conclusion is that the optimized processing circuit can suppress the instability of the output signal, which arises due to the variation of the particle velocity.

This research is helpful to the development of oil online monitoring technology, which starts from the characteristics of high flow rate and high flux of oil in the mechanical systems. Future work will concentrate on optimizing the design of the detection circuit to stabilize the particle output signal at different velocities in the circuit. In addition, combined with the effect of speed on magnetic fields and circuits, we will find a balance point in the contradiction between the effect of velocity on ferromagnetic and non-ferromagnetic metals.

Author Contributions: Conceptualization, W.L. and S.Y.; Data curation, S.Y., X.Z. and Y.X.; Formal analysis, C.B. and H.S.; Funding acquisition, W.L. and H.Z.; Investigation, Z.X.; Methodology, W.L., H.Z. and X.Z.; Project administration, W.L.; Resources, L.Z. and Y.S.; Software, S.Y. and Y.X.; Supervision, H.Z. and Y.S.; Validation, W.L. and S.Y.; Visualization, C.W. and Z.X.; Writing—original draft, W.L.; Writing—review & editing, W.L. and S.Y. All authors will be informed about each step of manuscript processing including submission, revision, revision reminder, etc. via emails from our system or assigned Assistant Editor. All authors have read and agreed to the published version of the manuscript.

Funding: This present project is supported by the National Natural Science Foundation of China (Grant No. 51679022), the Fundamental Research Funds for the Central Universities (Grant No. 3132022219), and the Liaoning Revitalization Talents Program (XLYC2002074).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Acknowledgments: The authors wish to thank the financial support. The authors would also like to thank Dalian Maritime University for their support of this research.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Towsyfyan, H.; Gu, F.; Ball, A.D.; Liang, B. Tribological behaviour diagnostic and fault detection of mechanical seals based on acoustic emission measurements. *Friction* **2019**, *7*, 572–586. [CrossRef]
2. Wang, T.; Han, Q.; Chu, F.; Feng, Z. Vibration based condition monitoring and fault diagnosis of wind turbine planetary gearbox: A review. *Mech. Syst. Signal Process.* **2019**, *126*, 662–685. [CrossRef]
3. Mohammed, A.; Djurovic, S. Electric Machine Bearing Health Monitoring and Ball Fault Detection by Simultaneous Thermo-Mechanical Fibre Optic Sensing. *IEEE Trans. Energy Convers.* **2021**, *36*, 71–80. [CrossRef]
4. Muthuvel, P.; George, B.; Ramadass, G.A. A Highly Sensitive In-Line Oil Wear Debris Sensor Based on Passive Wireless LC Sensing. *IEEE Sens. J.* **2021**, *21*, 6888–6896. [CrossRef]
5. Jia, R.; Ma, B.; Zheng, C.; Ba, X.; Wang, L.; Du, Q.; Wang, K. Comprehensive Improvement of the Sensitivity and Detectability of a Large-Aperture Electromagnetic Wear Particle Detector. *Sensors* **2019**, *19*, 3162. [CrossRef]
6. Wang, C.; Bai, C.; Yang, Z.; Zhang, H.; Li, W.; Wang, X.; Zheng, Y.; Ilerioluwa, L.; Sun, Y. Research on High Sensitivity Oil Debris Detection Sensor Using High Magnetic Permeability Material and Coil Mutual Inductance. *Sensors* **2022**, *22*, 1833. [CrossRef]
7. De Azevedo, H.D.M.; Araújo, A.M.; Bouchonneau, N. A review of wind turbine bearing condition monitoring: State of the art and challenges. *Renew. Sustain. Energy Rev.* **2016**, *56*, 368–379. [CrossRef]
8. Shi, H.; Zhang, H.; Huo, D.; Yu, S.; Su, J.; Xie, Y.; Li, W.; Ma, L.; Chen, H.; Sun, Y. An Ultrasensitive Microsensor Based on Impedance Analysis for Oil Condition Monitoring. *IEEE Trans. Ind. Electron.* **2022**, *69*, 7441–7450. [CrossRef]
9. Li, Y.; Yu, C.; Xue, B.; Zhang, H.; Zhang, X. A Double Lock-in Amplifier Circuit for Complex Domain Signal Detection of Particles in Oil. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–10. [CrossRef]
10. Jakoby, B.; Scherer, M.; Buskies, M.; Eisenschmid, H. An automotive engine oil viscosity sensor. *IEEE Sens. J.* **2003**, *3*, 562–568. [CrossRef]
11. Shi, H.; Zhang, H.; Ma, L.; Sun, Y.; Chen, H.; Zhao, X.; Bai, C.; Zhang, Y. Inductive-Capacitive Coulter Counting: Detection and Differentiation of Multi-Contaminants in Hydraulic Oil Using a Microfluidic Sensor. *IEEE Sens. J.* **2021**, *21*, 2067–2076. [CrossRef]
12. Birkin, P.R.; Youngs, J.J.; Truscott, T.T.; Martini, S. Development of an optical flow through detector for bubbles, crystals and particles in oils. *Phys. Chem. Chem. Phys.* **2022**, *24*, 1544–1552. [CrossRef] [PubMed]
13. Park, J.; Yoo, S.J.; Yoon, J.; Kim, Y.J. Inductive particle detection system for real-time monitoring of metals in airborne particles. *Sens. Actuators-Phys.* **2021**, *332*, 113153. [CrossRef]
14. Koo, H.; Ahn, H.S. Monitoring Inductance Change to Quantitatively Analyze Magnetic Wear Debris in Lubricating Oil. *Tribol. Lubr.* **2016**, *32*, 189–194. [CrossRef]
15. Haiden, C.; Wopelka, T.; Jech, M.; Keplinger, F.; Vellekoop, M.J. A Microfluidic Chip and Dark-Field Imaging System for Size Measurement of Metal Wear Particles in Oil. *IEEE Sens. J.* **2016**, *16*, 1182–1189. [CrossRef]
16. Zhang, H.; Chon, C.H.; Pan, X.; Li, D. Methods for counting particles in microfluidic applications. *Microfluid. Nanofluid.* **2009**, *7*, 739. [CrossRef]
17. Shi, H.; Bai, C.; Xie, Y.; Li, W.; Zhang, H.; Liu, Y.; Zheng, Y.; Zhang, S.; Zhang, Y.; Lu, H.; et al. Capacitive–Inductive Magnetic Plug Sensor with High Adaptability for Online Debris Monitoring. *IEEE Trans. Instrum. Meas.* **2022**, *71*, 1–8. [CrossRef]
18. Murali, S.; Jagtiani, A.V.; Xia, X.; Carletta, J.; Zhe, J. A microfluidic Coulter counting device for metal wear detection in lubrication oil. *Rev. Sci. Instrum.* **2009**, *80*, 016105. [CrossRef]

19. Du, L.; Zhe, J. A microfluidic inductive pulse sensor for real time detection of machine wear. In Proceedings of the 2011 IEEE 24th International Conference on Micro Electro Mechanical Systems, Cancun, Mexico, 23–27 January 2011; pp. 1079–1082. [CrossRef]
20. Zhang, X.; Zhang, H.; Sun, Y.; Chen, H.; Zhang, Y. Research on the Output Characteristics of Microfluidic Inductive Sensor. *J. Nanomater.* **2014**, *2014*, 725246. [CrossRef]
21. Ma, L.; Shi, H.; Zhang, H.; Li, G.; Shen, Y.; Zeng, N. High-sensitivity distinguishing and detection method for wear debris in oil of marine machinery. *Ocean Eng.* **2020**, *215*, 107452. [CrossRef]
22. Li, Y.; Wu, J.; Guo, Q. Design on Electromagnetic Detection Sensor on Wear Debris in Lubricating Oil. In Proceedings of the 2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Auckland, New Zealand, 20–23 May 2019; pp. 1–5. [CrossRef]
23. Davis, J.P.; Carletta, J.E.; Veillette, R.J.; Du, L.; Zhe, J. Instrumentation circuitry for an inductive wear debris sensor. In Proceedings of the 10th IEEE International NEWCAS Conference, Montreal, QC, Canada, 17–20 June 2012; pp. 501–504. [CrossRef]
24. Wu, X.; Zhang, Y.; Li, N.; Qian, Z.; Liu, D.; Qian, Z.; Zhang, C. A New Inductive Debris Sensor Based on Dual-Excitation Coils and Dual-Sensing Coils for Online Debris Monitoring. *Sensors* **2021**, *21*, 7556. [CrossRef] [PubMed]
25. Chady, T.; Enokizono, M.; Todaka, T.; Tsuchida, Y.; Yasutake, T. Identification of three-dimensional distribution of metal particles using electromagnetic tomography system. *J. Mater. Process. Technol.* **2007**, *181*, 177–181. [CrossRef]
26. Bai, C.; Zhang, H.; Wang, W.; Zhao, X.; Chen, H.; Zeng, N. Inductive-Capacitive Dual-Mode Oil Detection Sensor Based on Magnetic Nanoparticle Material. *IEEE Sens. J.* **2020**, *20*, 12274–12281. [CrossRef]
27. Zhu, X.; Du, L.; Zhe, J. A 3 × 3 wear debris sensor array for real time lubricant oil conditioning monitoring using synchronized sampling. *Mech. Syst. Signal Process.* **2017**, *83*, 296–304. [CrossRef]
28. Jagtiani, A.V.; Carletta, J.; Zhe, J. A microfluidic multichannel resistive pulse sensor using frequency division multiplexing for high throughput counting of micro particles. *J. Micromech. Microeng.* **2011**, *21*, 065004. [CrossRef]
29. Bai, C.; Zhang, H.; Zeng, L.; Zhao, X.; Yu, Z. High-Throughput Sensor to Detect Hydraulic Oil Contamination Based on Microfluidics. *IEEE Sens. J.* **2019**, *19*, 8590–8596. [CrossRef]
30. Wang, X.; Chen, P.; Luo, J.; Yang, L.; Feng, S. Characteristics and Superposition Regularity of Aliasing Signal of an Inductive Debris Sensor Based on a High-Gradient Magnetic Field. *IEEE Sens. J.* **2020**, *20*, 10071–10078. [CrossRef]
31. Wu, Y.; Zhang, H. Research on the effect of relative movement on the output characteristic of inductive sensors. *Sens. Actuators Phys.* **2017**, *267*, 485–490. [CrossRef]
32. Liu, E.; Zhang, H.; Wu, Y.; Fu, H.; Sun, Y.; Chen, H. Effect of oil velocity on sensitivity of micron metal particle detection by inductive sensor. *Opt. Precis Eng.* **2016**, *24*, 533–539. [CrossRef]
33. Zhang, X.; Zeng, L.; Zhang, H.; Huang, S. Magnetization Model and Detection Mechanism of a Microparticle in a Harmonic Magnetic Field. *IEEE ASME Trans. Mechatron.* **2019**, *24*, 1882–1892. [CrossRef]
34. Xie, Y.; Zhang, H.; Shi, H.; Zhang, Y. Frequency Research of Microfluidic Wear Debris Detection Chip Based on Inductive Wheatstone Bridge. In Proceedings of the 2021 IEEE 16th International Conference on Nano/Micro Engineered and Molecular Systems (NEMS), Xiamen, China, 25–29 April 2021; pp. 780–785. [CrossRef]
35. Xie, Y.; Shi, H.; Zhang, H.; Yu, S.; Llerioluwa, L.; Zheng, Y.; Li, G.; Sun, Y.; Chen, H. A Bridge-Type Inductance Sensor with a Two-Stage Filter Circuit for High-Precision Detection of Metal Debris in the Oil. *IEEE Sens. J.* **2021**, *21*, 17738–17748. [CrossRef]
36. Jia, R.; Ma, B.; Zheng, C.; Wang, L.; Ba, X.; Du, Q.; Wang, K. Magnetic Properties of Ferromagnetic Particles under Alternating Magnetic Fields: Focus on Particle Detection Sensor Applications. *Sensors* **2018**, *18*, 4144. [CrossRef]
37. Li, J.; Kong, M.; Xu, C.; Wang, S.; Fan, Y. An Integrated Instrumentation System for Velocity, Concentration and Mass Flow Rate Measurement of Solid Particles Based on Electrostatic and Capacitance Sensors. *Sensors* **2015**, *15*, 31023–31035. [CrossRef]

sensors

MDPI

Article

Improvement of Lightweight Convolutional Neural Network Model Based on YOLO Algorithm and Its Research in Pavement Defect Detection

Fu-Jun Du and Shuang-Jian Jiao *

Department of Civil Engineering, College of Engineering, Ocean University of China, Qingdao 266100, China; dufujun@stu.ouc.edu.cn

* Correspondence: oucjiaoshuangjian@163.com

Abstract: To ensure the safe operation of highway traffic lines, given the imperfect feature extraction of existing road pit defect detection models and the practicability of detection equipment, this paper proposes a lightweight target detection algorithm with enhanced feature extraction based on the YOLO (You Only Look Once) algorithm. The BIFPN (Bidirectional Feature Pyramid Network) network structure is used for multi-scale feature fusion to enhance the feature extraction ability, and Varifocal Loss is used to optimize the sample imbalance problem, which improves the accuracy of road defect target detection. In the evaluation test of the model in the constructed PCD1 (Pavement Check Dataset) dataset, the mAP@.5 (mean Average Precision when IoU = 0.5) of the BV-YOLOv5S (BiFPN Varifocal Loss-YOLOv5S) model increased by 4.1%, 3%, and 0.9%, respectively, compared with the YOLOv3-tiny, YOLOv5S, and B-YOLOv5S (BiFPN-YOLOv5S; BV-YOLOv5S does not use the Improved Focal Loss function) models. Through the analysis and comparison of experimental results, it is proved that the proposed BV-YOLOv5S network model performs better and is more reliable in the detection of pavement defects and can meet the needs of road safety detection projects with high real-time and flexibility requirements.

Keywords: pavement defects; deep learning; convolutional neural network; YOLOv5S; automated inspection; embedded equipment

Citation: Du, F.-J.; Jiao, S.-J. Improvement of Lightweight Convolutional Neural Network Model Based on YOLO Algorithm and Its Research in Pavement Defect Detection. *Sensors* **2022**, 22, 3537. https://doi.org/10.3390/s22093537

Academic Editors: Phong B. Dao, Liang Yu and Lei Qiu

Received: 29 March 2022
Accepted: 26 April 2022
Published: 6 May 2022

1. Introduction

It is understood that the United States has the largest highway network in the world. As of 2019, the total mileage of the U.S. highway network reached 6,853,024 km, of which about 63% has been paved. By 2020, the total mileage of China's highway network reached 5,198,000 km, of which 95% has been paved [1,2]. As more and more roads are paved, the maintenance of road pavements is a serious problem faced by the road maintenance departments of various countries. Pavement structural damage is the main cause of pavement defects. Once pavement cracks are formed, rainwater will accelerate the expansion of defects and create traps for moving vehicles, becoming one of the causes of car accidents. According to relevant data, from 2013 to 2016, a total of 11,386 people lost their lives due to road defects in India, which means that on average, about seven people die every day in India due to road surface defects [3]. In the United States, about half of the fatal car accidents that occur each year are caused by poor road maintenance. An 18-month study conducted by the Pacific Institute for Research and Evaluation examined information from the National Highway Traffic Safety Administration, because road defects, road icing, and other problems caused at least 42,000 Americans to lose their lives each year [4]. Pavement defects have introduced huge safety hazards to people's travel. Therefore, to protect people's lives and to provide for property safety, one of the important tasks of road safety and road maintenance is to discover and master road pavement defects in a timely manner.

In addition, highway engineering pavement defect detection technology is the most basic method of quality control in the construction stage of the life cycle of highway projects [5]. That technology is an important source of reference data for construction quality inspection, quality supervision, and quality control of highway engineering projects. It is also an important basis for engineering design parameters, construction quality control, construction completion acceptance, maintenance management, etc.

Pavement is the first part of the highway structure to be contacted by an external force. It bears the pressure of the passing vehicle loads and is also affected by other indirect factors, such as temperature change, corrosion, and human damage [6]. The damage to highway structures is often first manifested at the surface, so the quality of surface structure directly affects the quality of the entire life cycle of highway engineering. As an important component quality control in highway engineering construction, road engineering pavement defect detection technology has great significance for scheduling control and cost control during the life cycle of highway engineering. The detection of pavement defects does not essentially reduce the costs of highway engineering; however, through testing, a scientific and correct assessment of a highway project can be carried out reasonably to enhance maintenance, reduce the rate of accidents that are due to poor quality, reduce the costs of design, and prepare for preventive management and control to extend the service life of highway projects [7]. Therefore, research on the detection technology of pavement defects in highway engineering is of great importance.

The automatic detection of pavement defects is of great significance in the quality assessment of asphalt pavements [8]. In the assessment specifications of asphalt pavement in many countries, the assessment of pavement quality is determined according to the road defect condition. For example, in China, according to the MQI (Maintenance Quality Indicator) index of asphalt roads, the technical condition of roads is divided into five grades [9]. Pavement condition assessment is time-consuming and laborious repetitive work for humans, but it is simple and easy-to-operate work for computers. Therefore, evaluating road conditions by using computer vision technology can save significant labor costs and improve evaluation speed, while avoiding human subjective factors, to generate different evaluation criteria.

At present, there are three main methods in the detection of road surface defects: manual inspection, automatic detection, and image processing technology. In developing countries, the inspection of highway pavements is usually completed by manual inspection [10]. Traditional manual inspection has resulted in poor safety, low efficiency, and high costs, and is susceptible to subjective factors, resulting in different standards of judgment.

Defect detection of road pavement has been gradually replaced by automated equipment, such as inspection vehicles equipped with infrared or sensor equipment [11,12]. However, due to the complex characteristics of pavement defects and the road surface environment, automatic equipment brings certain difficulties to the detection of pavement defects, and presents difficulties in meeting the needs of actual engineering in terms of recognition accuracy and speed; in addition, the cost of automatic equipment is high and, accordingly, the cost of automatic detection is high.

Image processing technology has the advantages of low cost and high efficiency, and precision is gradually improved with the development of technology. Therefore, many researchers use image processing technology to detect road surface damage [13–17]. Traditional image processing techniques usually use manually selected features, such as color [18], texture [19], and geometric characteristics [20], to segment pavement defects, and then perform a classification and matching of machine learning algorithms to achieve the detection of pavement defects. However, given the complexity of road environments, the traditional image processing method cannot meet the requirements of model generalization ability and robustness in practical engineering through artificially designed feature extraction. Compared with traditional image processing technology, image processing technology based on deep learning theory has higher precision faster speed, and strong embeddability [21,22] in the detection of pavement defects, etc.

Computer vision technology based on deep learning has been widely used in the study of pavement defect detection. Hoang et al. [23] applied automatic detection technology to the investigation phase of asphalt pavement, using computer vision technology to detect repaired and unrepaired potholes, but its generalization ability was low. The technology needs further improvement if is to be applied to practical projects. Riid et al. [24] improved the detection performance of the model by manually selecting a deep convolutional neural network orthogonal frameset for training and digitizing, but the factors considered were too idealistic to be applied to actual highway engineering. Nguyen et al. [25] realized the automatic detection of cracks and potholes through the VGG16 network, and improved the robustness of the model through data enhancement processing technology, but the network structure of this model was large and it was difficult to meet the flexibility of embedded device applications. Maeda et al. [26] proposed the use of SSD Inception V2 and SSD MobileNet for road damage detection. The experimental results showed that the optimal model could achieve acceptable accuracy (77%) and recall (71%), outperforming conventional image processing techniques. Although the SSD network model can address inaccuracy in the detection of pavement defects, its speed is not as good as that of the YOLO series network in the detection of pavement defects.

Ping et al. [27] conducted experiments on the performance of YOLO, SSD, HOG with SVM, and Faster R-CNN network models for pavement defect detection; they constructed a dataset of pavement defects and then used different models for comparison The experimental results showed that the YOLOv3 model of the YOLO network algorithm series had the best performance in the detection of road surface defects, with fast speed and reliable detection results. Du et al. [28] used the YOLOv3 algorithm to construct a pavement defect detection model to achieve automatic feature extraction, and the detection speed was increased, but the flexibility was still poor, and it was difficult to meet the flexibility required by embedded systems. Liu et al. [29] used the combination of 3D ground-penetrating radar and the YOLO algorithm to detect damage to pavement structures. This method explored the damage in the deep structure of pavement through radar, which can detect potential structural damage in advance, but the cost was high. Park et al. [30] used YOLOv4 to carry out defect detection research for pavement potholes, but their research only focused on the ground penetration of pavement pothole defects, without more challenging crack detection studies. Baek et al. [31] improved the detection speed of the model by processing images of road pits in grayscale and then inputting them into the YOLO detection mode, and achieved good performance. To a certain extent, the processed images reduced the amount of information, resulting in the reduction of the generalization ability and robustness of the model. Pena-Caballero et al. [32] proposed deploying the YOLOv3 algorithm to an embeddable device to detect pavement defects, but the experimental results showed that due to the large network structure of the YOLOv3 network model, the performance of real-time detection in embedded systems needs to be improved. Ahmed [33] used YOLOv5, YOLOR, and Faster R-CNN deep learning network models in detecting pavement defects. The resulting analysis showed that the YOLOv5 model is extremely flexible and suitable for real-time detection scenarios of embedded devices. However, in terms of accuracy (mAP@0.5: 58.9%), further improvement was needed.

The prospects for the application of target detection technology based on deep learning theory in highway pavement defect detection are very considerable, but there are still some difficulties, including the following: (1) The existing pavement defect dataset is insufficient, and the field collection of data is dangerous, low in inefficiency, and high in cost. (2) The road environment is easily affected by passing vehicles and roadside greening facilities, which cause uneven light intensity, bringing difficulties to the target detection model. (3) Some pavement defects are small in size, large in number, and easily missed, and the target detection model has low detection accuracy for pavement defects [33]. To solve these problems, in this paper we adopt a new strategy to build a PCD1 dataset on pavement defects. Inspired by the YOLOv5 [34] model, we propose the BV-YOLOv5S deep learning network model, which improves the YOLOv5 model's performance. The feature extraction

network adopts the BiFPN network structure [35] to strengthen the feature extraction network, which aims to mine deep-level information about pavement defects in an image, improve the feature extraction performance of the model, and reduce missed detection caused by lighting and size. Further, based on the BV-YOLOv5S deep learning network model, Varifocal Loss [36] will be used to replace Focal Loss, so that the attention of model training is biased toward high-quality samples, making full use of the effective information in the dataset, and further improving the detection accuracy of the model.

The rest of this paper is organized as follows: (1) In Section 1, we introduce the current status of pavement defects and existing pavement defect detection methods. (2) The improved method of lightweight convolutional neural network for pavement defects is introduced in detail in Section 2. (3) In Section 3, we present the details of the experiment. (4) In Section 4, we discuss a comparison of the experimental results. (5) Finally, in Section 5, we summarize our work.

2. Methods

2.1. Introduction to Algorithm and Network Structure

YOLOv5S is a target detection method based on deep learning theory. It mines the data features in samples through a deep network structure [37], making the trained model more suitable for detection of targets with complex features, such as road defects, which can be detected by transfer learning [38]. Transfer learning quickly completes the training of network models and deploys them to target detection tasks in different backgrounds. Target detection work is an important branch in the field of computer vision [39]. The target detection network based on deep learning theory is mainly divided into two categories, according to the generation of candidate frames: the single-stage target detection network, without candidate frame generation, and the two-stage object detection network for box proposal generation. Typical single-stage detection networks include the YOLO [40] series networks and the SSD [41] series networks. The main operating principle of the single-stage target detection network is dividing the image into small squares, each of which has a preset fixed anchor. Then, the objects in the picture are divided into different small squares for classification, and the target detection work of the image to be tested is completed.

The two-stage target detection algorithm includes Fast RCNN [42], Faster RCNN [43], SPPNet [44], Mask R-CNN [45], and other network algorithms. The network principle is mainly that a large number of Windows are generated in the first stage. Windows uses a binary classification method to distinguish foreground and background. In the second stage, the region of interest (ROI) of target detection is used to deduct features from the feature map extracted by the convolutional neural network, and then the classification is performed again, which is different process compared to that of the first stage. The second stage of classification work is multi-target classification, which is used to distinguish the categories of different targets and to predict the position of an object by regression.

The accuracy of the single-stage detection network is limited, due to the imbalance of the samples. However, with the introduction of Focal Loss, the accuracy of the single-stage network target detection network has been greatly improved [46], and the single-stage network has an absolute advantage in terms of speed. Therefore, the single-stage detection network has been applied on a larger scale in the field of real-time detection, and the lightweight model based on the YOLO network has shown great flexibility in the deployment tests of embedded devices [47].

With the large-scale application of GPU and the rapid development of computer vision technology, the target detection technology based on deep learning theory has an absolute advantage in performance compared with traditional image processing technology [48,49]. At present, highway pavement defect detection has become an important topic in the field of computer vision technology. Single-stage object detection networks can be effectively transferred to highway pavement defect detection through transfer learning. Compared with other object detection algorithms, YOLOv5S has the advantages of fast detection speed, high accuracy, and flexibility [24]. However, when it is applied to the detection of

road pavement defects, there are still problems, such as difficulty in identification due to the small defect target, a high missed detection rate, and difficulty in mining the deep-level information of a defect image. To solve these difficulties, this paper proposes the BV-YOLOv5S network model, the network structure of which is shown in Figure 1, The network structure is mainly composed of backbone networks, feature extraction networks, and detection networks. To further improve accuracy, the BV-YOLOv5S deep learning network model uses Varifocal Loss to solve the sample imbalance problem in the single-stage target detection network model.

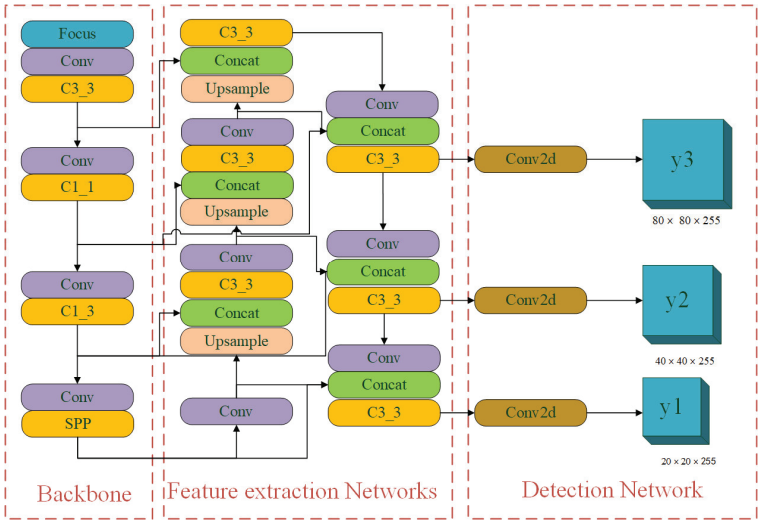


Figure 1. BV-YOLOv5S network structure. It is mainly composed of three parts: backbone networks, feature extraction networks, and detection networks.

To present the enhanced feature extraction network designed in this work in a more complete and detailed manner, we have drawn a detailed processing flow chart for the BV-YOLOv5S model to achieve road defect detection. Through the model's processing flow chart, we express the main work of this paper in detail, comprehensively and intuitively, as shown in Figure 2. In Figure 2, we divide the implementation process of the pavement defect detection algorithm into three parts: anchor, backbone, and head. We mainly redesigned the network structure for the head layer, in which the color deepened in the BifPN network structure, and an enhanced feature extraction network as shown in Figure 2 is constructed. Regarding the detailed network structure of BiFPN, we will introduce it in detail in Section 2.2.

2.2. Feature Extraction Networks of the BV-YOLOv5S Model

In the deep learning network model, with the continuous deepening of the number of network layers, each layer of a network will cause the loss of information to a certain extent and lead to the loss of features. By fusing multi-scale feature fusion networks, the detection accuracy of the model can be improved [50]. The feature extraction network of the YOLOv5S model adopts the PANet network structure. Focusing on the specificity of pavement defects, especially the characteristics of small pavement cracks, the model excavates deeper information about pavement defects shown in the picture. We improved the original feature extraction network, PANet, into a BiFPN network structure. It aims to enhance the depth of information mining and to further improve the feature extraction capability of the model. The network structures of BiFPN and PANet are shown in Figure 3. As shown in Figure 3b, the BiFPN structure diagram, the blue arrow is the top-down

path, which transmits the semantic information of high-level features, the red arrow is the bottom-up path, which transmits the location information of low-level features, and the purple part is an additional path when the input point and output point are located in the same layer, to fuse more features. Figure 4 is a visualization of the feature extraction of four types of road surfaces by PANet and BiFPN.

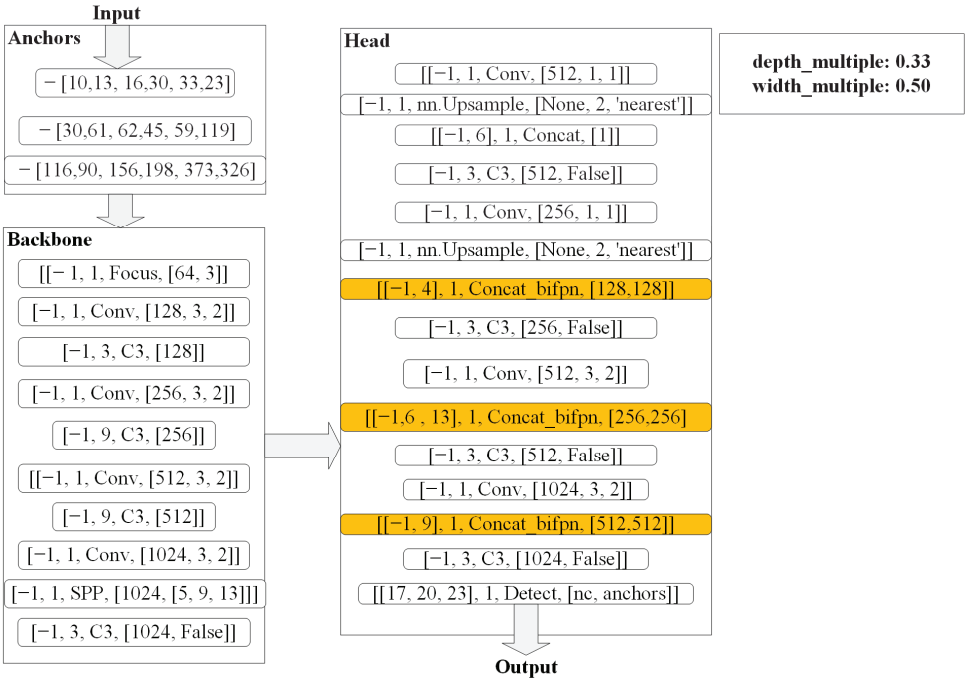


Figure 2. The detailed processing flow of the BV-YOLOv5S model to realize road defect detection. (Depth_multiple and width_multiple are the depth and width of the network, respectively. The anchor part is the size setting of the anchor. The content expression format of the backbone part [number, module, args] were from the input of the layer, where number is the number of the layer, module is the class name, and args is the initialization parameter of the class. The head part is the same as the backbone part in terms of content expression format).

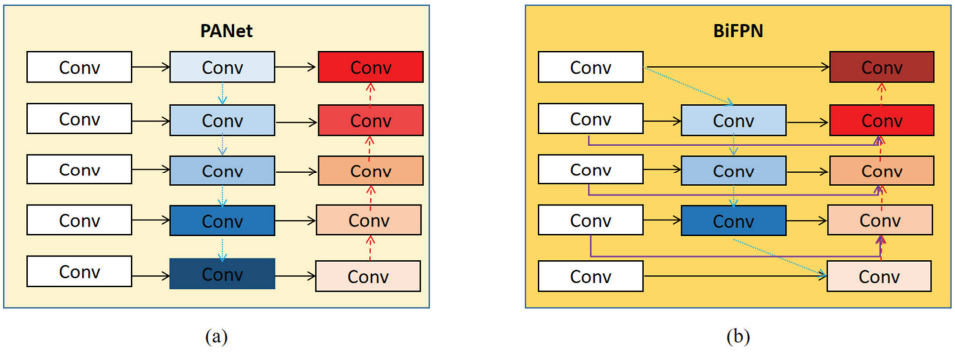


Figure 3. Comparison of BiFPN and PANet network structures [35,51]. (**a**) is the PANet network structure, and (**b**) is the BiFPN network structure, where conv is convolution node.

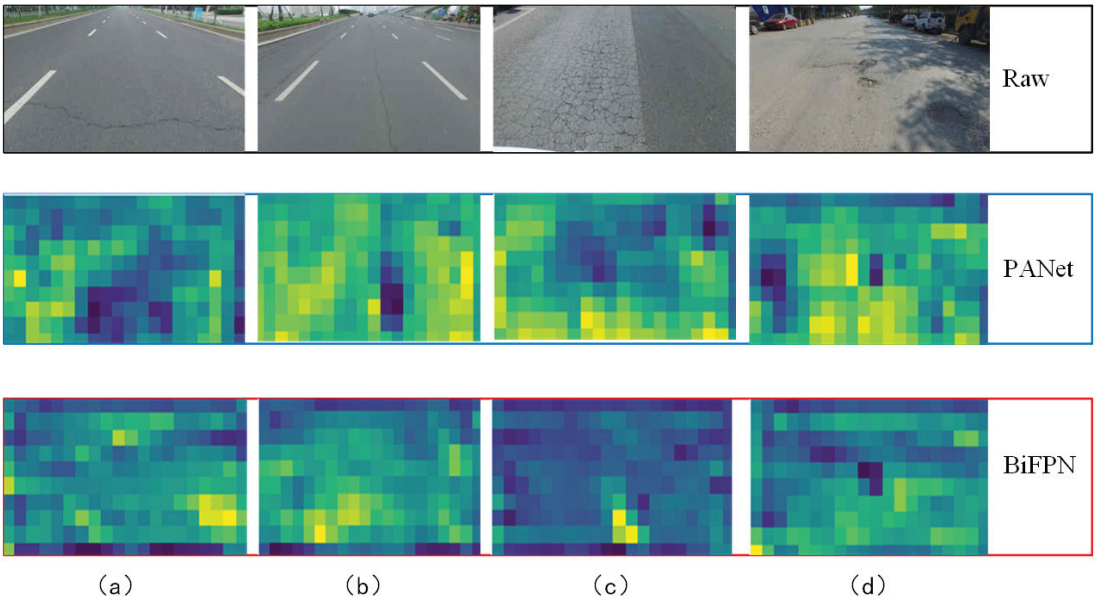


Figure 4. Feature extraction diagram of four types of pavement defects by BiFPN and PANet network structures. (**a**) lateral cracking, (**b**) longitudinal cracking, (**c**) alligator cracking, and (**d**) pothole.

To enhance the transmission of image features of the deep learning model, we adopted the BiFPN network structure to strengthen the feature extraction ability of the deep learning network, compared with the original feature fusion network, PANet, as shown in Figure 3a, eliminating single-sided input nodes with less contribution to fusion. As shown in Figure 3b, the BiFPN network strengthens higher-level feature fusion in the processing path, processing each bidirectional path (top-down and bottom-up) as a feature network layer, and repeats this process multiple times in the same layer. Through the fusion of weighted features, the importance of different input features is learned, and differentiated fusion is carried out for different features. BiFPN uses fast normalized fusion to fuse weighted features, which are defined as in Equation (1). The learning weight, wi, uses the ReLU activation function, with a value of $\varepsilon = 0.0001$, to strengthen the stability of the value. To further improve the detection efficiency of the deep network learning model, BiFPN uses separable convolutional fusion features and adds batch normalization and activation after each convolution. We took layer 6, as shown in Figure 4b, as an example and described the definitions of two fusion features as shown in Equation (2).

$$O = \sum_i \frac{w_i}{\varepsilon + \sum_j w_j} \cdot I_i \tag{1}$$

$$\begin{aligned} P_6^{td} &= Conv\left(\frac{w_1 \cdot P_6^{in} + w_2 \cdot Rwsize\left(P_7^{in}\right)}{w_1 + w_2 + \varepsilon}\right) \\ P_6^{out} &= Conv\left(\frac{w_1' \cdot P_6^{in} + w_2' \cdot P_6^{td} + w_3' \cdot Resize\left(P_5^{out}\right)}{w_1' + w_2' + w_3' + \varepsilon}\right) \end{aligned} \tag{2}$$

where P_6^{td} represents the middle feature of the sixth layer from top to bottom and P_6^{out} is the output feature of the sixth layer from the bottom up. Through the interconnection and fusion between different layers, BiFPN's bidirectional cross-scale connection and fast normalization fusion are finally realized.

2.3. Improved Focal Loss Function of BV-YOLOv5S Model

Existing detectors rank NMS detections by predicting an additional IoU score [52] or a center score [53] as an indicator incorporated into the prediction criteria, which can alleviate the error between a classification score and localization accuracy. However, the standard obtained by multiplying two imperfect scores may produce larger errors, and experiments have shown that this is not a perfect practice in the field of object detection [36]. In this paper, we determined an IoU-aware classification score as the target presence confidence by simultaneously predicting the perceptual classification score (IACS), representing the localization accuracy of a specific object class to be detected; we also determined the generated bounding box by incorporating additional predictions into the classification score, for a joint representation of the localization accuracy. The YOLOv5 deep learning network model uses Focal Loss to deal with the imbalance of positive and negative samples in the target detection of the YOLOv5 model. Its definition formula is shown in Equation (3), where α is the lost weight and $p\gamma$ is the weight of different samples; the sample weight increases for samples that are difficult to classify, reducing the impact of easy-to-classify samples on the loss function. The model pays more attention to the training of samples that are difficult to classify, but creates positive and negative samples equally; thus, it does not put the focus of training on high-quality samples. Given the complex background of pavement defects, the effective features in the sample images are difficult to highlight. In the BV-YOLOv5S deep network model, we used Varifocal Loss to deal with the problem of sample imbalance and to increase the weight of positive sample losses with high IoU values.

Focusing training on high-quality samples increases the robustness of the model. Its definition formula is as in Equation (4), where p is the predicted Iou-aware Cls_score (IACS) and q is the target IoU score. For the positive sample q in training, it is the IoU between the b box and the gt box, and for the negative sample q in the training, the value is 0.

$$\mathrm{FL}(\mathrm{p},\ \mathrm{y}) = \begin{cases} -\alpha(1-\mathrm{p})^{\gamma}\log(\mathrm{p}) & \text{if } \mathrm{y}=1 \\ -(1-\alpha)\mathrm{p}^{\gamma}\log(1-\mathrm{p}) & \text{otherwise} \end{cases} \tag{3}$$

$$\mathrm{VFL}(\mathrm{p},\ \mathrm{q}) = \begin{cases} -\mathrm{q}(\mathrm{q}\log(\mathrm{p}) + (1-\mathrm{q})\log(1-\mathrm{p})) & \mathrm{q} < 0 \\ -\alpha\mathrm{p}^{\gamma}\log(1-\mathrm{p}) & \mathrm{q} = 0 \end{cases} \tag{4}$$

Compared with other target detection objects, pavement defect images have the characteristics of complex backgrounds and diverse shapes; there is a high false detection rate during recognition, and the preparation of high-quality datasets raises problems of low efficiency and high cost. To solve these problems, we used Varifocal Loss instead of Focal Loss for calculation, predicting IACS and classification scores. At the same time, we focused on high-quality samples for training, and for efficiently we used the information in the dataset.

In this paper, our detailed improvement steps for the sample imbalance function, Focal Loss are as follows: first, according to the source code of Varifocal Loss [40], the applicability of the source code was converted to make the source code conform to the network structure of the YOLO algorithm; then, the original sample imbalance processing function was replaced to complete the construction of the BV-YOLOv5S road surface defect detection model.

3. Experiment

3.1. Experiment Environment and Evaluation Index

3.1.1. Experiment Environment

The model proposed in this paper was tested under laboratory conditions and the results were analyzed quantitatively and qualitatively. The deep learning model development tool used was Anaconda3, and the deep network learning model development hardware and environment in the laboratory were CPU: Intel(R) Xeon(R) Gold 5218; GPU: GeForce RTX 2080 Ti/11 GB; system: ubuntu18.04, CUDA10.2 with model acceleration training function, cuDNN7; pytorch1.7.0 was used as the training framework of the model.

3.1.2. Evaluation Metrics

To verify the performance of our proposed BV-YOLOv5S network model in pavement defect detection, a confusion matrix evaluation index was introduced to evaluate the model. The confusion matrix contains four types of definitions: TN (predict negative samples as negative samples), FN (predict positive samples as negative samples), TP (predict positive samples as positive samples), and FP (predict negative samples as positive samples).

(1) Precision, Recall, F1-score evaluation indicators

To comprehensively evaluate the performance of the model in the PCD1 dataset, we introduced quantitative analysis indicators: precision, recall, and F1-score as defined in Equations (5)–(7). Precision, recall, and F1-score have commonly used evaluation indicators for target detection algorithms, and they play an irreplaceable role in the evaluation of models.

$$\mathrm{P} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{5}$$

$$\mathrm{R} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{6}$$

$$\mathrm{F1} - \mathrm{score} = \frac{2 \times \mathrm{P} \times \mathrm{R}}{\mathrm{P} + \mathrm{R}} \tag{7}$$

(2) Detection rate

Detection speed is a critical requirement for engineering practical applications, and we used frame rate (FPS) to show detection speed, which is an important metric for model evaluation. Generally speaking, if FPS is $\geq$30, it can basically meet the demand, and the video detection function of FPS $\geq$ 60 is smooth.

(3) PR curve, mAP@.5 evaluation indicators

The PR curve takes recall as the abscissa and precision as the ordinate to draw the curve. If the PR curve of one model could completely wrap up the PR curve of the other model, then the performance of the former model could be considered to be better. If it could not be judged directly, the comparison could be made according to the area under each model curve.

mAP@.5 means that under the condition of lOU = 0.5, the average value of precision is calculated for the four types of defects. The average value of precision is an important indicator in the model evaluation process. AP is calculated by precision and recall, and AP is defined as:

$$AP = \frac{1}{11} \sum_{\gamma \in \{0, 0.1, 0.2 \cdots, 1\}} \rho_{\mathrm{interp}(r)} \tag{8}$$

When we calculated the AP of the four categories of defects under the condition of lOU = 0.5, we obtained the detected mAP@.5, which is defined as follows:

$$mAP = \frac{\sum_{i=1}^{4} AP_i}{4} \tag{9}$$

3.2. Data Collection and Processing

In this paper, we have constructed a pavement defect detection dataset—the PCD1 dataset. As far as we know, the current public datasets for defect recognition have different standards for such indicators as image shooting angle, light intensity, and clarity. The quality of such datasets is such that it is difficult to meet the requirements of use, so we decided to build our own defect dataset for road surfaces. Based on the existing public datasets, RDD2020 [54] and SDNET2018 [55], the quality of the PCD1 dataset was improved by using the Baidu Street View Map and web crawler technology, as well as field collection. The field acquisition device was the Huawei Mate30pro rear 40-megapixel camera. Some of the acquisition parameters were ISO = 50, F:1.6, S:1/1182 s, focal length: 27 mm, and shooting depression angle: 45°–60°. The PCD1 dataset was built based on diverse data,

clear pictures, and a shooting angle of 45°–60°. To make full use of the dataset and to improve the generalization ability of the model, the images containing multiple types of defects were preferentially used. After careful selection, the collected images were standardized before model training, and the images were reduced to 640 × 640 size so that the YOLO model could exert the best training performance. After standardization, according to different types of pavement defects, labeling was used to manually label data. The file format of the labeling was txt, and a total of 5600 pieces of data were formed. Figure 5 shows four typical defect types in PCD1. According to the needs of the experiment, it was randomly divided into a training set, a validation set, and a test set, according to 6:2:2. The details of the defect distribution of the dataset are shown in Table 1.

Figure 5. Partial image of four typical defect types in PCD1. (**a**) lateral cracking, (**b**) longitudinal cracking, (**c**) alligator cracking, and (**d**) pothole.

Table 1. Details of defect distribution in the PCD1 dataset.

Type of Dataset	Lateral Cracking	Longitudinal Cracking	Alligator Cracking	Pothole
Number	1350	1050	1400	1800

3.3. Data Collection and Processing

We trained YOLOv3-Tiny, YOLOv5s, B-YOLOv5s, and BV-YOLOv5S, as proposed in this paper, in the same training set that was independent of the validation and test sets. The YOLOv3-Tiny and YOLOv5S algorithms are widely used in the field of defect detection, due to their strong flexibility, high accuracy, and fast speed, and they are advanced in the application of lightweight convolutional neural networks [56]. Many researchers in other professional fields have applied the BiFPN network structure to the YOLO algorithm series and achieved good target detection results [57–59]. For this reason, we established the B-YOLOv5S network and applied it to the detection of pavement defects. We further compared and evaluated the performance of the BV-YOLOv5S model proposed in this paper in pavement defect applications.

We set the epoch to 1000 for the training of this study and used the cosine annealing method to adjust the learning rate. The initial learning rate was lr0 = 0.01, and the cyclic learning rate was lrf = 0.2, which was helpful for the model's convergence loss. To apply the YOLO network structure, the size of all input images was 640 × 640, and the training process of the whole model took about 6 h. Loss represented the gap between the predicted value and the actual value. As the gap gradually decreased and converged, it meant that the model approached the upper limit of performance determined by the dataset. A comparison of the training loss function curves of the four models is shown in Figure 6.

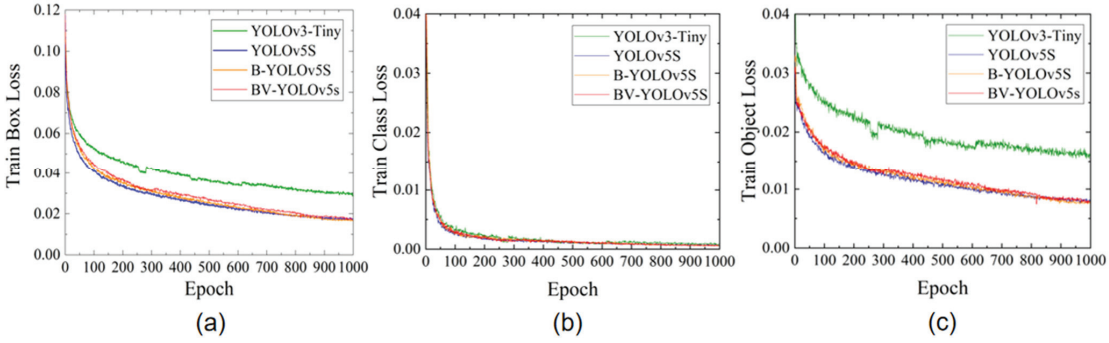


Figure 6. Comparison of training losses of four models. (**a**) train box loss. (**b**) train class loss. (**c**) train object loss.

As shown in Figure 6, the loss value fluctuates greatly at the beginning of training for each category, indicating that the initial hyperparameters we used were reasonable. After a certain number of iterations, the fluctuation of the loss curve gradually decreased, but as shown in Figure 6, we found that the convergence loss performance of YOLOv3-Tiny was poor. The loss function convergence performance of B-YOLOv5S and BV-YOLOv5S was similar, and better than the loss performance of YOLOv5S.

4. Results and Discussion

4.1. Evaluation Metrics

After the training of the four models, we used the test dataset independent of the training and validation sets to evaluate the models. During testing, we set IoU to 0.5 to divide the positive and negative samples, and we plotted the PR curves for the performance of different models. As shown in Figure 7, through the PR curve we could see that the performance of the improved YOLOv5S model was significantly better than that of YOLOv3-Tiny and YOLOv5S.

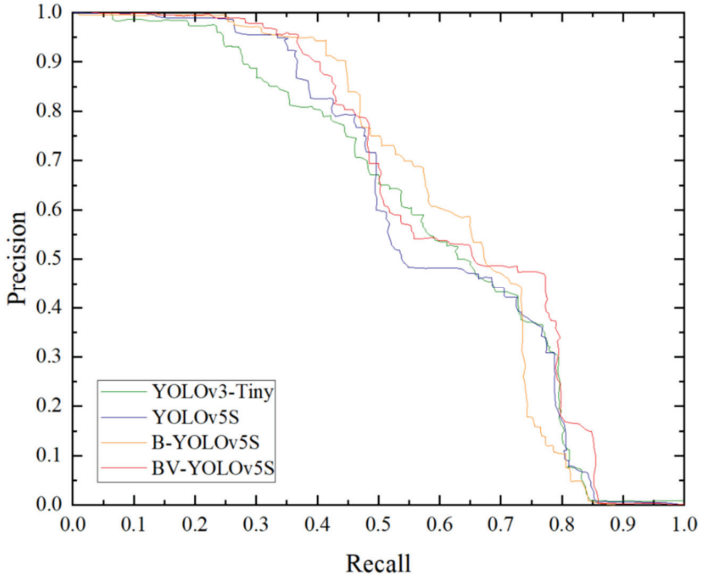


Figure 7. PR curves of four pavement defect detection models on the test set.

4.2. Discussion of Results

In Table 2, we summarize the performance of four deep learning network models in the PCD1 test dataset, and their confusion matrices are shown in Figure 8. From the confusion matrix, we found that BV-YOLOv5S was ahead or even far ahead of the other three models in the correct classification of pothole and alligator cracking, and was the same as B-YOLOv5 in the correct classification of longitudinal cracking and ahead of the other two models. In the correct classification of lateral cracking, although it was ahead of YOLOv3-Tiny and YOLOv5S, the correct rate was lower than that of B-YOLOv5S. Overall, the BV-YOLOv5S classification performance outperformed the other three models in the confusion matrix. However, for the defects of lateral cracking and longitudinal cracking, the classification effect of the four models was poor, resulting in a largely missed detection phenomenon. This phenomenon occurred because, for the deep network learning model, the small target of the crack, the variety of shape, and the uncertainty of the width caused certain difficulties in the feature information extraction of the target detection model, which was the difficulty in the field of small target detection.

Table 2. Test results of different defect detection models in the pavement defect dataset. (Through independent test set testing, the precision, recall, and F1-score values of each category were provided and averaged).

Model	mAP@.5	Precision	Recall	F1-Score	FPS
YOLOv3-Tiny	0.594	0.737	0.573	0.646	167
YOLOv5S	0.605	0.859	0.549	0.670	238
B-YOLOv5S	0.626	0.876	0.561	0.684	278
BV-YOLOv5S	0.635	0.864	0.590	0.701	263

This phenomenon will be the focus of our future work. In the next step, we will enhance the accuracy of the model in crack detection by further improving the feature extraction ability of the object detection model network.

Specifically, compared with the YOLOv3-Tiny model, the deep network learning model BV-YOLOv5S, the YOLOv5S model, and the B-YOLOv5S model mAP@.5 increased by 4.1%, 3%, and 0.9%, respectively. Precision increased by 12.7% and 0.5%, respectively, which was 1.2% lower than that of the B-YOLOv5S model; recall increased by 1.7%, 4.1%, and 2.9%, respectively. The F1-score increased by 5.5%, 3.1%, and 1.7%, respectively. As shown in Table 2, we found that the recall values of the four types of models were low, mainly due to the shallow network layers of the lightweight deep learning network model, The ability to extract features and learn was low, but this improved the running speed and flexibility of the model in meeting the needs of terminal deployment and practical engineering. We found that BV-YOLOv5S was significantly enhanced compared to YOLOv3-Tiny and YOLOv5S, in terms of speed, recall, and F1-score, which proved that our work is effective and can be used as a reference for other work in the defect detection field. According to the analysis of the experimental results of the detection speed, we found that the BiFPN network structure information processing speed was faster than that of the PANet network structure, which not only strengthened the feature extraction ability but also improved the detection speed of the model, and had better performance in real-time detection. However, the BV-YOLOv5S model adopted a more complex calculation method to deal with the problem of sample imbalance, which reduced the detection speed to a certain extent; nevertheless, it was sufficient for real-time detection requirements. In conclusion, our proposed BV-YOLOv5S deep network learning model comprehensively outperformed the YOLOv5S model in mAP@.5, precision, recall, F1-score, and its detection speed metrics went far beyond the current YOLOv3-Tiny network model, which is representative of lightweight target detection models.

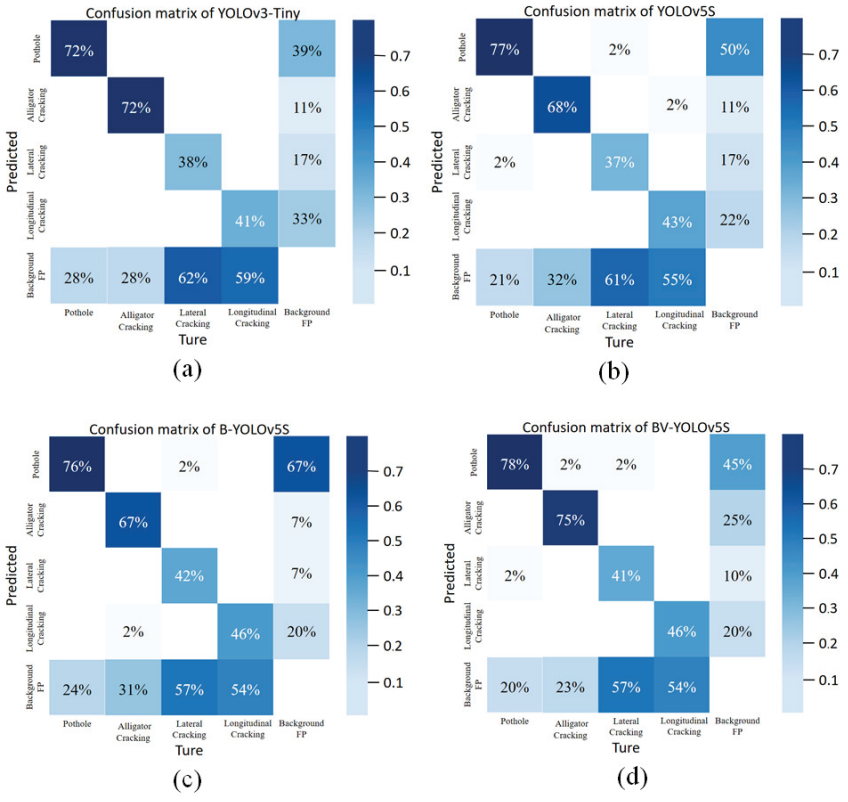


Figure 8. The confusion matrix of the comparison of the different defect detection algorithms. (**a**) YOLOv3-Tiny; (**b**) YOLOv5S; (**c**) B-YOLOv5S; (**d**) BV-YOLOv5S.

Through comprehensive comparative analysis of mAP@.5, precision, recall, F1-score, and FPS evaluation indicators in the experimental results, we believe that the BV-YOLOv5S model is more robust in performance than the YOLOv3-Tiny, YOLOv5S, and B-YOLOv5S network models. The BV-YOLOv5S model proposed in this paper has higher accuracy and flexibility, and has stronger advantages for the deployment and practical application of embedded devices.

Due to the complex road surface environment, especially on rural roads, there are many defects in road surfaces. To comprehensively measure the performance of the BV-YOLOv5S model, we used small targets, multi-targets, and shadow-occluded targets to conduct visual experiments, as shown in Figure 9.

As shown in Figure 9a, we found that the detection performance of the YOLOv3-Tiny model is obviously disturbed by rutting and lane boundaries, and the anti-interference ability was poor. The detection performance of YOLOv5S and B-YOLOv5S was also interfered with by rutting and roadway boundaries to a certain extent, and BV-YOLOv5S had strong anti-interference ability in this environment. However, Figure 9a shows that the model was too sensitive to the detection of pavement potholes; it is necessary to further improve the generalization ability of the model in the future.

Figure 9. The test results of different network models in a complex environment. (**a**) The model detection results of four types of pavement defects in a rutting interference environment. (**b**) Detection results of four types of pavement defect models under the conditions of many and small targets. (**c**) The detection results of the four types of pavement defect models under environments of uneven illumination intensity. (**d**) The model detection results of four types of pavement defects under partial shadow occlusion conditions. (**e**) Alligator crack detection in the presence of small target potholes.

As shown in Figure 9b, we tested the performance of the four defect models under the condition of multiple small targets, It can be seen from the figure that the YOLOv3-Tiny detection performance was average, and the small holes nearby were missed, which showed that the model was not sensitive to small targets, The performance of the YOLOv5S model was the worst; recent potholes were missed, and there were serious defects in performance. The B-YOLOv5S model had relatively good detection results. However, despite the missed detection of pits with inconspicuous defect characteristics in the vicinity and the missed detection of small targets at a distance, the BV-YOLOv5 model had the best performance in this detection. It could detect the pits in this range, showing an absolute advantage in performance.

As shown in Figure 9c, the pavement defect detection performance of the four models was tested under the condition of uneven illumination intensity. It can be seen that the YOLOv3-Tiny and YOLOv5S networks were greatly interfered with by light intensity, resulting in obvious losses in detection performance. The B-YOLOv5S model could detect defects but had low confidence. The BV-YOLOv5S could accurately detect defects with a high degree of confidence. It was not interfered with by uneven light intensity and had a strong anti-interference ability with respect to the external environment.

As shown in Figure 9d, the four models had similar detection capabilities under the condition of partial shadow occlusion, and a small part of the shadow did not have a great impact on the models. As shown in Figure 9e, we tested the model to detect alligator cracks under the interference of small target potholes. Through testing, we found that YOLOv3-Tiny had the lowest confidence, and the BV-YOLOv5S model had the highest

confidence in alligator crack detection, and only this model could detect potholes for small objects. This shows that, to a certain extent, the handling of sample imbalance by Varifocal Loss enhances the model's detection ability for small objects. It can be seen that by using BiFPN to enhance feature extraction, the detection accuracy of the model can be improved to a certain extent; in addition, by optimizing the sample imbalance processing method, the sensitivity of small target recognition can be improved, and the missed detection rate of targets with insignificant features can be reduced, with better performance in the detection of pavement defects. By further testing our improved model's performance, we confirmed its performance advantages. The BV-YOLOv5S showed stronger practical advantages, compared with the YOLOv3-Tiny, YOLOv5S, and B-YOLOv5S models. In the quantitative evaluation results and the qualitative analysis, the BV-YOLOv5S model proposed in this paper showed strong anti-interference ability, high sensitivity to small targets, a low multi-target missed detection rate, and little influence of external environment interference, with good robustness and generalization.

In this experiment, we used asphalt pavement as the background of the defect detection model. Compared with cement pavement, the feature structure of asphalt pavement is more complex. Through the experiments in this dataset, we confirmed our results. At the same time, we could quickly carry out the training and deployment of the pavement defect detection model for cement pavement and other targets through transfer learning.

5. Conclusions

In this paper, an improved lightweight deep network learning model, BV-YOLOv5S, was proposed for the detection of asphalt road surface defects by embedded devices. First, we established a high-standard, high-quality PCD1 dataset employing public datasets, Baidu Street View maps, web crawler, and field photography. The data types of the datasets were enriched, and new data collection strategies contributed to the target detection datasets. Second, to realize the applicability of the model in embedded systems and to perform accurate detection under complex road conditions, we proposed a BV-YOLOv5S deep network learning model, based on the YOLOv5S deep network learning model. In the feature extraction network, the BiFPN network was used to replace the PANet network, with the aim of mining deeper information in the pavement defect images. The BV-YOLOv5S network model improved the Focal Loss and used Varifocal Loss as the loss function to deal with the imbalance problem between samples. More of the model's attention was transferred to high-quality dataset samples, making full use of the effective information in the dataset. We trained and tested the proposed BV-YOLOv5S model and the YOLOv3-Tiny, YOLOv5S, and B-YOLOv5S network models under the same conditions. The experimental results showed that our proposed BV-YOLOv5S model in the mAP@.5 index improved by 4.1%, 3%, and 0.9%, respectively, compared with the deep network learning models YOLOv3-Tiny, YOLOv5S, and B-YOLOv5S. In addition, the recall and F1-score evaluation metrics were higher or even much higher than those of the B-YOLOv5S, YOLOv5S, and YOLOv3-Tiny models, O\outperforming YOLOv5S in detection speed and precision, and far exceeding YOLOv3-Tiny. The experimental results demonstrated that our proposed work is advanced.

Our proposed BV-YOLOv5S deep network learning model is most suitable for defect detection under actual road conditions and has good practicability and advancement. Compared with other object detection models, our model can be more flexibly deployed in mobile embedded devices, and compared with other lightweight object detection models, it has shown good performance in the tests. In subsequent work, the BV-YOLOv5S proposed in this paper will be applied to embedded devices in ordinary cars to further improve and realize the automatic detection of road surface defects.

Author Contributions: Conceptualization, F.-J.D. and S.-J.J.; methodology, software, validation, formal analysis, investigation, F.-J.D.; resources, F.-J.D.; data curation, writing—original draft preparation, F.-J.D.; writing—review and editing F.-J.D.; visualization, F.-J.D.; supervision, S.-J.J.; project administration, F.-J.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available, as they involve the subsequent applications for patents, software copyright, and the publication of project deliverables.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. CIA. Roadways—The World Factbook. Available online: www.cia.gov (accessed on 12 July 2021).
2. CIA. Public Road Length—2017 Miles by Functional System. Available online: www.cia.gov/the-world-factbook (accessed on 2 February 2019).
3. The Times of India. Deadly pits: Potholes Claimed 11,386 Lives during 2013–2016. Available online: https://timesofindia.indiatimes.com/india/deadly-pits-potholes-claimed-11386-lives-during-2013-16/articleshow/60774243.cms (accessed on 21 September 2017).
4. Greg Colemanlaw. Accidents and Injuries Caused by Bad Road Conditions. Available online: https://www.gregcolemanlaw.com/bad-road-damages-and-effects.html (accessed on 1 January 2022).
5. Pasha, A.; Mansourian, A.; Ravanshadnia, M. Evaluation of Work Zone Road User Cost of Pavements Based on Rehabilitation Strategy Approach. *J. Transp. Eng. Part B Pavements* **2021**, *147*, 4021015. [CrossRef]
6. Hosseini, A.; Faheem, A.; Titi, H.; Schwandt, S. Evaluation of the long-term performance of flexible pavements with respect to production and construction quality control indicators. *Constr. Build. Mater.* **2020**, *230*, 116998. [CrossRef]
7. Kumar, P.; Sharma, A.; Kota, S.R. Automatic Multiclass Instance Segmentation of Concrete Damage Using Deep Learning Model. *IEEE Access* **2021**, *9*, 90330–90345. [CrossRef]
8. Majidifard, H.; Adu-Gyamfi, Y.; Buttlar, W.G. Deep machine learning approach to develop a new asphalt pavement condition index. *Constr. Build. Mater.* **2020**, *247*, 118513. [CrossRef]
9. *JTG H20-2007*; Chinese Highway Technical Condition Evaluation Standard. Ministry of Transportation: Beijing, China, 2007.
10. Hoang, N.-D.; Nguyen, Q.-L. Automatic Recognition of Asphalt Pavement Cracks Based on Image Processing and Machine Learning Approaches: A Comparative Study on Classifier Performance. *Math. Probl. Eng.* **2018**, *2018*, 6290498. [CrossRef]
11. Radopoulou, S.C.; Brilakis, I. Automated Detection of Multiple Pavement Defects. *J. Comput. Civ. Eng.* **2017**, *31*, 4016057. [CrossRef]
12. Koch, C.; Georgieva, K.; Kasireddy, V.; Akinci, B.; Fieguth, P. A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure. *Adv. Eng. Inform.* **2015**, *29*, 196–210. [CrossRef]
13. Li, Z.; Wang, W.; Shui, P. Parameter Estimation and Two-Stage Segmentation Algorithm for the Chan-Vese Model. In Proceedings of the 2006 International Conference on Image Processing, Atlanta, GA, USA, 8–11 October 2006; pp. 201–204, ISBN 1-4244-0480-0.
14. Subirats, P.; Dumoulin, J.; Legeay, V.; Barba, D. Automation of Pavement Surface Crack Detection using the Continuous Wavelet Transform. In Proceedings of the 2006 International Conference on Image Processing, Atlanta, GA, USA, 8–11 October 2006; IEEE: Piscataway, NJ, USA, 2006; pp. 3037–3040, ISBN 1-4244-0480-0.
15. Li, B.; Wang, K.C.P.; Zhang, A.; Fei, Y.; Sollazzo, G. Automatic Segmentation and Enhancement of Pavement Cracks Based on 3D Pavement Images. *J. Adv. Transp.* **2019**, *2019*, 1813763. [CrossRef]
16. Zou, Q.; Cao, Y.; Li, Q.; Mao, Q.; Wang, S. CrackTree: Automatic crack detection from pavement images. *Pattern Recognit. Lett.* **2012**, *33*, 227–238. [CrossRef]
17. Hoang, N.-D.; Nguyen, Q.-L. A novel method for asphalt pavement crack classification based on image processing and machine learning. *Eng. Comput.* **2019**, *35*, 487–498. [CrossRef]
18. Salari, E.; Bao, G. Pavement Distress Detection and Severity Analysis. *Adv. Eng. Inform.* **2011**, *7877*, 25–27.
19. Moghadas Nejad, F.; Zakeri, H. A comparison of multi-resolution methods for detection and isolation of pavement distress. *Expert Syst. Appl.* **2011**, *38*, 2857–2872. [CrossRef]
20. Koch, C.; Brilakis, I. Pothole detection in asphalt pavement images. *Adv. Eng. Inform.* **2011**, *25*, 507–515. [CrossRef]
21. Zhang, L.; Yang, F.; Daniel Zhang, Y.; Zhu, Y.J. Road crack detection using deep convolutional neural network. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2006; IEEE: Piscataway, NJ, USA, 2006; pp. 3708–3712, ISBN 978-1-4673-9961-6.
22. Aswath, M.; Raj, S.J.; Mohanaprasad, K. Real-Time Pothole Detection with Onboard Sensors and Camera on Vehicles. In *Futuristic Communication and Network Technologies*; Sivasubramanian, A., Shastry, P.N., Hong, P.C., Eds.; Springer: Singapore, 2022; pp. 479–488, ISBN 978-981-16-4624-9.
23. Hoang, N.-D.; Huynh, T.-C.; Tran, V.-D. Computer Vision-Based Patched and Unpatched Pothole Classification Using Machine Learning Approach Optimized by Forensic-Based Investigation Metaheuristic. *Complexity* **2021**, *2021*, 3511375. [CrossRef]
24. Riid, A.; Lõuk, R.; Pihlak, R.; Tepljakov, A.; Vassiljeva, K. Pavement Distress Detection with Deep Learning Using the Orthoframes Acquired by a Mobile Mapping System. *Appl. Sci.* **2019**, *9*, 4829. [CrossRef]

25. Nguyen, H.T.; Nguyen, L.T.; Afanasiev, A.D.; Pham, L.T. Classification of Road Pavement Defects Based on Convolution Neural Network in Keras. *Aut. Control Comp. Sci.* **2022**, *56*, 17–25. [CrossRef]
26. Maeda, H.; Sekimoto, Y.; Seto, T.; Kashiyama, T.; Omata, H. Road Damage Detection and Classification Using Deep Neural Networks with Smartphone Images. *Comput.-Aided Civ. Infrastruct. Eng.* **2018**, *33*, 1127–1141. [CrossRef]
27. Ping, P.; Yang, X.; Gao, Z. A Deep Learning Approach for Street Pothole Detection. In Proceedings of the 2020 IEEE Sixth International Conference on Big Data Computing Service and Applications (BigDataService), Oxford, UK, 3–6 August 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 198–204, ISBN 978-1-7281-7022-0.
28. Du, Y.; Pan, N.; Xu, Z.; Deng, F.; Shen, Y.; Kang, H. Pavement distress detection and classification based on YOLO network. *Int. J. Pavement Eng.* **2021**, *22*, 1659–1672. [CrossRef]
29. Liu, C.; Wu, Y.; Liu, J.; Han, J. MTI-YOLO: A Light-Weight and Real-Time Deep Neural Network for Insulator Detection in Complex Aerial Images. *Energies* **2021**, *14*, 1426. [CrossRef]
30. Park, S.-S.; Tran, V.-T.; Lee, D.-E. Application of Various YOLO Models for Computer Vision-Based Real-Time Pothole Detection. *Appl. Sci.* **2021**, *11*, 11229. [CrossRef]
31. Baek, J.-W.; Chung, K. Pothole Classification Model Using Edge Detection in Road Image. *Appl. Sci.* **2020**, *10*, 6662. [CrossRef]
32. Pena-Caballero, C.; Kim, D.; Gonzalez, A.; Castellanos, O.; Cantu, A.; Ho, J. Real-Time Road Hazard Information System. *Infrastructures* **2020**, *5*, 75. [CrossRef]
33. Ahmed, K.R. Smart Pothole Detection Using Deep Learning Based on Dilated Convolution. *Sensors* **2021**, *21*, 8406. [CrossRef] [PubMed]
34. Ultralytics. Yolov5. Available online: https://github.com/ultralytics/yolov5 (accessed on 1 January 2021).
35. Tan, M.; Pang, R.; Le, V.Q. EfficientDet: Scalable and Efficient Object Detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 10781–10790.
36. Zhang, H.; Wang, Y.; Dayoub, F.; Sünderhauf, N. VarifocalNet: An IoU-aware Dense Object Detector. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 8514–8523.
37. Zhao, Z.-Q.; Zheng, P.; Xu, S.-T.; Wu, X. Object Detection with Deep Learning: A Review. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *30*, 3212–3232. [CrossRef]
38. Lu, J.; Behbood, V.; Hao, P.; Zuo, H.; Xue, S.; Zhang, G. Transfer learning using computational intelligence: A survey. *Knowl.-Based Syst.* **2015**, *80*, 14–23. [CrossRef]
39. Wu, X.; Sahoo, D.; Hoi, S.C.H. Recent advances in deep learning for object detection. *Neurocomputing* **2020**, *396*, 39–64. [CrossRef]
40. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
41. Cheng, Y.; Chen, C.; Gan, Z. Enhanced Single Shot MultiBox Detector for Pedestrian Detection. In Proceedings of the 3rd International Conference on Computer Science and Application Engineering, Sanya, China, 22–24 October 2019; pp. 1–7.
42. Girshick, R. Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015; pp. 1440–1448.
43. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *arXiv* **2015**, arXiv:1506.01497v3. [CrossRef]
44. He, K.; Zhang, X.; Ren, S.; Sun, J. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2014**, *37*, 1904–1916. [CrossRef]
45. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2961–2969.
46. Lin, T.-Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal Loss for Dense Object Detection. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2980–2988.
47. Jiang, Z.; Zhao, L.; Li, S.; Jia, Y. Real-time object detection method based on improved YOLOv4-tiny. *arXiv* **2020**, arXiv:2011.04244.
48. Ma, D.; Fang, H.; Wang, N.; Xue, B.; Dong, J.; Wang, F. A real-time crack detection algorithm for pavement based on CNN with multiple feature layers. *Road Mater. Pavement Des.* **2021**, *10338*, 1–17. [CrossRef]
49. Cha, Y.-J.; Choi, W.; Büyüköztürk, O. Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks. *Comput. -Aided Civ. Infrastruct. Eng.* **2017**, *32*, 361–378. [CrossRef]
50. Liu, C.; Wu, Y.; Liu, J.; Sun, Z.; Xu, H. Insulator Faults Detection in Aerial Images from High-Voltage Transmission Lines Based on Deep Learning Model. *Appl. Sci.* **2021**, *11*, 4647. [CrossRef]
51. Liu, S.; Qi, L.; Qin, H.; Shi, J.; Jia, J. Path Aggregation Network for Instance Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 8759–8768.
52. Wu, S.; Li, X.; Wang, X. IoU-Aware Single-Stage Object Detector for Accurate Localization. *Image Vis. Comput.* **2019**, *97*, 103911. Available online: http://arxiv.org/pdf/1912.05992v4 (accessed on 1 February 2022). [CrossRef]
53. Tian, Z.; Shen, C.; Chen, H.; He, T. FCOS: Fully Convolutional One-Stage Object Detection. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Seoul, Korea, 27 October–2 November 2019; pp. 9627–9636.
54. Arya, D.; Maeda, H.; Ghosh, S.K.; Toshniwal, D.; Sekimoto, Y. RDD2020: An annotated image dataset for automatic road damage detection using deep learning. *Data Brief* **2021**, *36*, 107133. [CrossRef]
55. Maguire, M.; Dorafshan, S.; Thomas, R.J. SDNET2018: A Concrete Crack Image Dataset for Machine Learning Applications. 2018. Available online: https://digitalcommons.usu.edu/all_datasets/48 (accessed on 1 February 2022).

56. Adibhatla, V.A.; Chih, H.-C.; Hsu, C.-C.; Cheng, J.; Abbod, M.F.; Shieh, J.-S. Applying deep learning to defect detection in printed circuit boards via a newest model of you-only-look-once. *Math. Biosci. Eng.* **2021**, *18*, 4411–4428. [CrossRef] [PubMed]
57. Jing, Y.; Ren, Y.; Liu, Y.; Wang, D.; Yu, L. Automatic Extraction of Damaged Houses by Earthquake Based on Improved YOLOv5: A Case Study in Yangbi. *Remote Sens.* **2022**, *14*, 382. [CrossRef]
58. Wang, Y.; Hua, C.; Ding, W.; Wu, R. Real-time detection of flame and smoke using an improved YOLOv4 network. *Signal Image Video Processing* **2022**, *288*, 30. [CrossRef]
59. Zhang, Z.; Lu, X.; Cao, G.; Yang, Y.; Jiao, L.; Liu, F. ViT-YOLO: Transformer-Based YOLO for Object Detection. In Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW), Montreal, BC, Canada, 11–17 October 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 2799–2808, ISBN 978-1-6654-0191-3.

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
www.mdpi.com

MDPI Books Editorial Office
E-mail: books@mdpi.com
www.mdpi.com/books

Academic Open
Access Publishing

mdpi.com

ISBN 978-3-0365-9932-8