



Article Application of Wavelet Transform for the Detection of Cetacean Acoustic Signals

Ruilin He ^{1,2,3,†}, Yang Dai ^{1,2,4,*,†}, Siyi Liu ⁵, Yuhao Yang ^{1,2,3}, Yingdong Wang ⁵, Wei Fan ^{1,2} and Shengmao Zhang ^{1,2}

- ¹ East China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Shanghai 200090, China
- ² Key Laboratory of Fisheries Remote Sensing, Ministry of Agriculture and Rural Affairs, Shanghai 200090, China
- ³ School of Navigation and Naval Architecture, Dalian University of Ocean, Dalian 116023, China
- ⁴ Laoshan Laboratory, Qingdao 266237, China
- ⁵ BeiDou Application & Research Institute Co., Ltd. of Norinco Group, Shanghai 200438, China
- * Correspondence: 13162787852@163.com
- + These authors contributed equally to this work.

Featured Application: We believe that this research will be helpful for the detection of Cetacean acoustic signals for research purposes or dataset building for the purpose of more accurate artificial neural network training.

Abstract: *Cetaceans* are an important part of the ocean ecosystem and are widely distributed in seas across the world. *Cetaceans* are heavily reliant on acoustic signals for communication. Some *Odontoceti* can perceive their environments using their sonar system, including the detection, localization, discrimination, and recognition of objects. Acoustic signals are one of the most commonly used types of data for *Cetacean* research, and it is necessary to develop *Cetacean* acoustic signal detection methods. This study compared the performance of a manual method, short-time Fourier transform (STFT), and wavelet transform (WT) in *Cetacean* acoustic signal detection. The results showed that WT performs better in click detection. According to this research, we propose using STFT for whistle and burst-pulse marking and WT for click marking in dataset building. This research will be helpful in facilitating research on the habits and behaviors of groups and individuals, thus providing information to develop methods for protecting species and developing biological resources.

Keywords: wavelet transform; short-time Fourier transform; hydrophone; *Cetacean*; acoustic signal; signal detection; click; whistle; burst pulse

1. Introduction

Cetaceans belong to *Whippomorpha*, including *Mysticeti* and *Odontoceti*. Acoustic signals are the most important way in which *Cetaceans* perceive their environment and communicate, especially active acoustic signals generated by *Cetaceans* [1,2]. *Odontoceti* can use their sonar systems for echolocation through wide-frequency band signals [3,4]. Since the 20th century, due to excessive whaling, these animals' habitats have been reduced in size, and some species are now endangered [5–8]. To protect *Cetaceans*, the International Whaling Commission (IWC) devised the International Convention for the Regulation of Whaling [9]. In recent years, studies have shown that in addition to human hunting, the noise generated by human activities also has a greatly negative impact on *Cetaceans* [1,10–15]. *Odontoceti*, especially *Delphinidae* and *Phocaenidae*, are highly dependent on their sonar systems for environmental perception and predation [1–3,16,17]. Human activities are frequent in rivers, estuaries, and coastal areas, and the noise produced results in serious interference with, and even damage to the auditory systems of *Odontoceti* [10–13,15]. Taking *Sousa chinensis* as an example, we can observe that their click signal is a broadband signal with a short



Citation: He, R.; Dai, Y.; Liu, S.; Yang, Y.; Wang, Y.; Fan, W.; Zhang, S. Application of Wavelet Transform for the Detection of Cetacean Acoustic Signals. *Appl. Sci.* **2023**, *13*, 4521. https://doi.org/10.3390/ app13074521

Academic Editor: Emanuel Guariglia

Received: 21 January 2023 Revised: 22 March 2023 Accepted: 28 March 2023 Published: 2 April 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). duration [18], which is easily masked by impact or knocking noises [13]. Young individuals can distinguish artificial knocking sounds and click signals from their own. However, the hearing capacity of elder individuals gradually becomes worse with increased age [19,20], resulting in the confusion of knocking sounds and clicking signals and causing them to enter rivers by mistake. Some aged *Sousa chinensis* have been killed due to grounding or fungal infection based on these mistakes. The study of *Cetaceans'* acoustic signals is important for species protection measures and biological resource development.

Hydrophone data are the most widely used type of information in *Cetacean* research because sound travels much further than light in water and is more easily recorded. The acoustic research on hydrophone data of *Cetaceans* began in the 1940s, given that hydrophone technology was developed in World War II. William E. Schevill et al. researched the acoustic signals of the *White porpoise* (*Delphinapterus leucas*) by a hydrophone working at 0.5 kHz~10 kHz in 1949 [21]. Since then, many researchers have used hydrophones to study the characteristics of the acoustic signals of *Cetaceans*. In 1993, Whitlow W.L. Au summarized the results of acoustic research on a variety of Cetaceans and compiled a book entitled The Sonar of Dolphins [3]. After determining the characteristics of Cetaceans' acoustic signals, researchers began to study the relationship between *Cetaceans'* acoustic signals, environment, and behaviors based on these characteristics. Therefore, the detection and analysis of acoustics became increasingly important. Tursiops truncatus are known to avoid obstacles in their paths while swimming and to locate fishes for food by sound reflection or by echolocation [4,22]. Johnson et al. found that the upper limit of hearing of *Tursiops* truncatus can reach 120~140 kHz. Liang Fang et al. studied the high-frequency echolocation signals of Sousa chinensis in Sanniang Bay, Guangxi Province, China, and found that the mean peak frequency was 109 kHz [18]. Liang Fang et al. researched the echolocation signals of captive and free-ranging *Neophocaena asiaeorientalis* and found that the main center frequency of clicks from individuals in the Baiji aquarium was 133 kHz, while that of individuals at the Shishou Tian-e-zhou Reserve was 128 kHz, and that of individuals at Tianxingzhou was 129 kHz [23]. In most studies, the researchers released the hydrophone into the water to collect acoustic signals. Some pressed the hydrophone against the skin of dolphins, as in T.H. Bullock et al.'s research [24]. Cetacean acoustic signals can be separated into whistle, burst-pulse, and click signals based on their time–frequency characteristics. The click signal is a broadband signal with an upper limit of frequency up to 150 kHz, with some reaching even more than 200 kHz. It is still difficult to identify the click signals of *Odontoceti* in large quantities for research purposes. On the one hand, marine environment noise leads to some interference with the sound signals. On the other hand, knocks and sounds of non-target marine animals may be misidentified [25]. Abbas et al. designed an FChOA-MLPNN for the automatic detection of marine mammal sounds [26]. However, the manual labeling of the dataset consumed a great deal of labor and could lead to the mislabeling of knock signals as click signals. Yang et al. transformed hydrophone data from the time domain into the time-frequency domain using a short-time Fourier transform (STFT). The acoustic signals of dolphins can be marked according to their duration, short energy, and spectral centroid. Due to the uncertainty principle, the time-frequency spectrum calculated using STFT cannot maintain a high temporal resolution and high-frequency resolution at the same time. Yang's method can mark whistle and burst pulses accurately, but there are some mismarks of clicks [27].

Since the development of the digital signal processing (DSP) method, a new signal analysis method has been produced, i.e., the wavelet transform (WT), where the resolution can be dynamically changed in accordance with the frequency. Although WT still cannot ensure both a high temporal resolution and high-frequency resolution in a window at the same time based on the uncertainty principle, the dynamic resolution captures greater details of wide-frequency band signals [28]. Thus far, WT has been applied to seismic signal recognition, part flaw detection, image processing, and other fields [29–31].

In this research, we studied the usage of WT for *Cetacean* acoustic signal detection; recognized whistle, burst-pulse, and click signals with WT; and compared these data with

the results obtained using traditional STFT and manual methods. This research lays the groundwork for the detection of many *Cetaceans'* acoustic signals and can support species protection efforts and the study of *Cetaceans'* habits.

2. Materials and Methods

Single-scalar hydrophones or scalar hydrophone arrays are the most widely used equipment for Cetacean acoustic research. Therefore, the currently known characteristics of Cetaceans' acoustic signals are based on the physical quantities that can be output using scalar hydrophones or gained by processing the output using scalar hydrophones, i.e., the duration, strength, and frequency, derived after time-frequency analysis. We designed a target signal detection procedure considering these three quantities (Figure 1). The procedure can be divided into three steps. Step 1, signal analysis, involves the collection and transfer of the time domain analog signal data into digital data using the hydrophone. Then, the background noise is flitted out as much as possible, and the time domain data are transformed into time-frequency domain data using a mathematical analysis method. Here, we chose two different methods: STFT and WT. The details of step 1 are shown in Section 2.2. Step 2, target signal detection, mainly serves to identify signals of a certain strength within the specified frequency range. These signals are strongly suspected to be target signals. Step 3, target signal marking, aims to determine the duration of each signal and mark the signals within the specified duration range. It should be noted that the strength, frequency range, and duration range should be set according to the target species. The details of step 2 and step 3 are shown in Section 2.3.



Figure 1. Flow chart of the target signal detection procedure.

2.1. Hydrophone Data

The dataset for the experiment was recorded by the South China Sea Fisheries Research Institute, Chinese Academy of Fisheries Science, at Guishan Offshore Wind Farm, Zhuhai, Guangdong Province, China (N22.142817, E133.7238333). Located at the site of an estuary of the Pearl River, Guishan is one of the habitats of the *Indo-Pacific humpback dolphin (Sousa* *chinensis*). A hydrophone was fixed on the pile of a wind power generator for data collection. The parameters and settings of the hydrophone are shown in Table 1.

Parameters	Value	
Туре	Sound Trap ST300 HF, Ocean Instruments	
Memory	256 GB	
Frequency band	20~144 kHz	
Sampling rate	288 kHz	
Resolution	16-bit	
Minimum self-noise	37 dB	

Table 1. Parameters of the hydrophone for data collection.

2.2. Methods of Signal Analysis

The original hydrophone data include time information and signal strength information, but the most notable feature of *Cetaceans'* acoustic signals is their frequency information. Therefore, we must transform the original time domain data into time–frequency domain data. In this research, we chose three different methods to process the data, including one manual method and two digital methods, STFT and WT. The manual method and STFT are commonly used for *Cetacean* acoustic signal detection. WT is a new method that has been applied for the processing of non-stationary signals or filters. We aimed to determine whether WT was valuable for *Cetacean* acoustic research by comparing these three methods.

2.2.1. Manual Method

Human sense is widely used in the field investigation of marine mammals. Both whistle and burst pulses are on the threshold of audibility, being easily identified by experienced fishers, trainers, and researchers. The frequency can also be roughly estimated using human sense, as a higher sound means a higher frequency. However, most of the frequency bands of clicks are higher than the human hearing range. Although the low-frequency part of the click range can be heard by trained persons, we believe that the characteristics of this part are not sufficient to judge whether the signal is a click. It is necessary to enable hearing of the higher-frequency band of clicks, which includes a much greater number of characteristics. In this research, the sampling rate of the original signal was 288 kHz, and we played the audio 20 times slower at 14.4 kHz. According to sampling theory, the upper limit of the frequency band was reduced to 7.2 kHz, enabling the inaudible high-frequency band of signals to be heard by humans. Clicks sound similar to firecracker explosions, making them more easily to distinguish.

2.2.2. STFT

STFT is a time–frequency transform method based on the Fourier transform (FT). At present, STFT is the most widely used method of marine mammal acoustic signal time–frequency analysis. Much commercial audio software and hydrophone software use STFT for visualization and data analysis, e.g., the commercial audio edit software Adobe Audition and the professional hydrophone data software Marco, Lucy, and Sound Trap Host. The principle of STFT is to divide the signal according to a certain window length and overlap. Then, one can analyze the signals in each window with FT and sort the results into chronological order. The function of STFT can be written as

$$X_{\text{STFT}}(t,f) = \int_{-\infty}^{\infty} x(\tau)h(\tau-t)e^{-j2\pi f\tau}d\tau,$$
(1)

where $x(\tau)$ is the input signal at the moment of τ , $h(\tau - t)$ is the window function, and t is the position of the window on the time axis. The settings of STFT in this research can be seen in Table 2.

Parameters	Setting
Software	SciPy 1.7.1 ¹
Window	Hamming
Window length	500 sampling points
Overlap	250 sampling points

Table 2. Parameters of STFT.

¹ A collection of mathematical algorithms (Software SciPy v1.7.1) [32]. Here, we used Python 3.9.7.

2.2.3. WT

Since the development of signal processing technology, many new methods have been developed for single- or two-dimension signal analysis [33–36]. Among these methods, WT is a widely used one [37]. WT appeared in the 1970s, and in recent years, it has been used for various types of signal analysis, such as signal and image processing [38,39]. The principle of WT is to use a scalable and displaceable finite-length base wavelet inner-product time domain data. The function of WT can be written as

$$X_{\rm WT}(a,\tau) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \varphi^*\left(\frac{n-\tau}{a}\right) dn,$$
(2)

where x(t) is the value of the time domain data at the moment of t; $\varphi(t)$ is a scalable and displaceable finite-length base wavelet; the positive coefficient a controls the scaling of $\varphi(t)$; and coefficient τ controls the displacement of $\varphi(t)$. The settings of WT in this research can be seen in Table 3.

Table 3. Parameters of WT.

Parameters	Setting	
Software	PyWavelets v1.1.1 ¹	
Base wavelet	Cmor100-100	
Total scale	1000	
Sampling period	Decided by sample rate of original signal	
Scales	Calculated by Function (3)	

¹ An open-source wavelet transform software (Software PyWavelets v1.1.1) [40]. Here, we used Python 3.9.7.

The scales controlling the window sizes can be calculated as

$$S = \frac{2fT}{A}, A = [T, T - 1, T - 2, \dots, 2, 1],$$
(3)

where *S* is the scale, *f* is the sample rate of the original signal, and *T* is the total scale.

Vicente J. Bolós et al. introduced a wavelet tool and windowed scale index to solve a particular case of Haar wavelet [41]. Magdalena Lepicka et al. also introduced a scale index in their research [42]. The scale index is necessary for orthogonal decomposition. The Cmor wavelet is a Gaussian envelope function and is related to non-orthogonal decomposition without a scale index. Both WT and STFT use the infinitesimal element method, in which the original data are divided into successive windows, while each window is transformed into the time–frequency domain, and the results are sorted in chronological order. Obviously, the shorter the window length is, the higher the temporal resolution will be. This is the uncertainty principle. We chose an original whistle signal without a filter and processed it using different window lengths to demonstrate the uncertainty principle (Figure 2). It can be seen that when the window length was short, the temporal resolution was high, with a low-frequency resolution. When the window length was longer, the temporal resolution became lower, with a higher frequency resolution. In essence, temporal resolution and frequency resolution are negatively correlated.



Figure 2. Uncertainty principle in time-frequency analysis.

As Figure 3 shows, the window length and overlap of STFT are constant values, i.e., they uniformly divide the entire signal into several parts and can be employed to calculate each part using FT. This data division method determines that STFT cannot increase the temporal resolution and frequency resolution at the same time, being affected by the uncertainty principle. In WT, each window is obtained by scaling and displacing the base wavelet. This data division method determines that WT has a higher temporal resolution with a lower frequency resolution for the high-frequency band and a lower temporal resolution with a higher frequency resolution for the low-frequency band.

The click signals have a short duration, high upper-frequency limit, and wide frequency band. Therefore, the detection of clicks requires a relatively high temporal resolution. The frequency resolutions of STFT are exactly the same in any frequency band. It is difficult for STFT to detect catastrophe signals (e.g., clicks). The temporal resolutions of WT can be dynamically adjusted according to the frequency band, thus realizing a higher temporal resolution at the high-frequency band and a higher frequency resolution at the low-frequency band. Furthermore, the base wavelet of WT can be modified, whereas the STFT wavelet type can only be a sinusoidal wave. Therefore, the modified WT, using the appropriate wavelet, performs better than STFT in aperiodic catastrophe signal detection. Capture audio randomly, and transform it into the time-frequency spectrum using STFT and WT, respectively. Figure 4a shows the time-frequency spectrum provided by STFT, and Figure 4b shows that provided by WT. It can be seen that WT retains more details of the signal, and the spectrum is clearer. For example, in Figure 4, the signals at around 50 ms, 200 ms, and 250 ms are so weak that they cannot be heard by humans and are not clear enough for STFT to detect them. However, in the case of WT, it is clear that there were some short-duration signals.



Figure 3. Resolutions of STFT and WT.

2.3. Target Signal Detection and Marking Experiment

Marine background noise is generated by geological activities, engineering machinery activities, sea surface winds and waves, biological activities, etc. After long-distance propagation through the seawater, these components overlap with each other, forming a complex high-intensity noise with a wide frequency band. It is necessary to denoise the original data before detection and marking. The loss function of sound propagation in the sea can be written as

$$PL = 20 \lg r + ar \times 10^{-3}, \tag{4}$$

where *PL* is the loss of the sound; *r* is the propagation distance; and *a* is the sound loss coefficient in the sea, which has a positive correlation with the frequency. In accordance with this function, the loss of sound has a positive correlation with the frequency and propagation distance. Although marine background noise is distributed throughout the whole frequency band, the energy is concentrated in the low-frequency band.



Figure 4. Time–frequency spectra provided by STFT and WT. (**a**) Spectrum calculated by STFT. (**b**) Spectrum calculated by WT.

Low-frequency noise can be filtered out using a high-pass filter or band-pass filter. A high-pass filter can filter out signals lower than the cut-off frequency. A band-pass filter can filter out non-specified frequency-band signals. Since the high-frequency characteristic is sufficient for detection, it is not necessary to retain the low-frequency band. Therefore, the Butterworth high-pass filter was selected to filter out all the signals lower than the cut-off frequency. This filter can ensure that the frequency response curve for the passband is as flat as possible so that the frequency response curve for the stop band gradually drops to zero. The function of the Butterworth high-pass filter can be written as

$$\left|H(\omega)\right|^{2} = \frac{1}{1 + \left(\frac{\omega_{\rm C}}{\omega}\right)^{2n'}} \tag{5}$$

where *n* is the order of the filter, and ω_C is the cut-off frequency. The filter $H(\omega)$ is approximately equal to 1, preserving the signal when the frequency ω is higher than ω_C . The cut-off frequency was set to 1 kHz for click detection and 0.1 kHz for whistle and burst pulse detection. The hydrophone, after filter application, was input into the procedure shown in Figure 1.

The time–frequency data obtained with STFT or WT represent a three-dimensional matrix $E^{t \times f}$, where *t* represents the position on the time axis, and *f* represents the position on the frequency axis. The set signal strength parameter M_e , M_e can be calculated with

$$M_{\rm e} = \sum_{f_{\rm min}}^{f_{\rm max}} f e_{(t,f)},\tag{6}$$

Next, we set a lower threshold T_e . The signal at *t* can be considered as part of the target signal when $M_e > T_e$. As shown in Figure 1, we obtain M_e at the moment *t*. It is considered that the data at moment *t* do not belong to the target signal if $M_e < T_e$, setting the non-signal mark $a_t = t$, while the signal mark b_t remains unchanged. It is considered that the data at moment *t* belong to the target signal if $M_e > T_e$, setting that the data at moment *t* belong to the target signal if $M_e > T_e$, setting that the data at moment *t* belong to the target signal if $M_e > T_e$, setting the signal mark $b_t = t$. If $a_t = t - 1$, the current moment is the starting point of the suspected signal, setting the signal starting mark u = t. If $a_t \neq t - 1$, $b_t = t - 1$. The current moment is the end point of the suspected signal, setting the signal ending mark as v = t. Setting the duration mark M_t , M_t can be calculated as:

$$M_{\rm t} = v - u, \tag{7}$$

where *v* is the end moment of a signal, and *u* is the starting moment of a signal. Then, we set an upper limit of duration M_{tmax} and lower limit of M_{tmin} . If $M_{\text{t}} \in (M_{\text{tmin}}, M_{\text{tmax}})$, it is determined that the signal belongs to the target signal.

3. Results

We analyzed the hydrophone data and marked the whistle, burst pulse, and click using the manual, STFT, and WT methods, respectively. Since the target species was the *Indo-Pacific humpback dolphin (Sousa chinensis)*, the parameters for detection were set, as shown in Table 4. The results are shown in Table 5.

Target Signal	Parameters	Value
	T _e	50,000
	f_{\min}	15 kHz
Whistle	f_{\max}	40 kHz
	$M_{ m tmin}$	0.5 s
	$M_{ m tmax}$	3 s
	T _e	50,000
	f_{\min}	10 Hz
Burst pulse	f_{max}	15 kHz
	$M_{ m tmin}$	0.2 s
	$M_{ m tmax}$	1 s
	T _e	40,000
	f_{\min}	40 kHz
Click	f_{\max}	144 kHz
	$M_{ m tmin}$	15 µs
	$M_{ m tmax}$	35 µs

Table 4. Parameters for detection.

Type of Signal	Manual	STFT	WT
Whistle	1062	1062	1059
Burst pulse	1361	1361	1352
Click	2501	2382	3057

Table 5. Results of target signal detection obtained using different methods.

Whistle and burst pulses are audible to humans. Therefore, the results of the manual method can be used as the standard. It can be seen that SFTF performed relatively well; it marked all the signals, whereas WT left several signals unmarked. As for clicks, it was not easy for us to judge directly whether the marked signals were target signals. Although WT marked more signals than the other methods in click detection, we could not directly assess its performance, as discussed in Section 4.

4. Discussion

Whistle and burst pulses are within the human hearing range, and the duration of the signal is long enough for our sense of hearing to recognize them. However, it is not easy for humans to recognize clicks, as they are too short for us to hear their features clearly. It is also difficult to distinguish clicks from other impulse signals, such as the knocking sounds of crustacean activity and offshore piling. Therefore, we believe that the results of manual methods for whistle and burst pulse are reliable and that these methods can be used as a standard, but we do not have a reliable standard for clicks.

It should be noticed that the frequency band of click is far beyond the upper limit of the human auditory threshold, and there is no clear standard that can be used to judge the recognition results. Although we increased the sampling time by 20 to render the click audible, the duration was still too short (approximately 440 μs). It is difficult to distinguish a click signal from other broadband pulse signals. For example, knocking sounds of objects impacting each other are so similar to clicks that some aged dolphins with hearing loss may misjudge them [19]. Therefore, it is reasonable to assume that there were several misjudgments in the experiment. WT reduced the influence of the uncertainty principle and had advantages in terms of non-stationary signal detection. Although WT detected more click signals in the experiment, it is uncertain whether WT performed better or, instead, marked interference signals. Is it possible for WT to distinguish clicks from other interference knocks by analyzing the duration of the signal? To determine this, some knocking sound samples were collected for an experiment. In this experiment, the data were pure interference knocks without Cetacean acoustic signals. We used the manual, STFT, and WT methods mentioned in Section 2 to process these data separately. Subsequently, all the detected signals were misdetections. These results indirectly reflect the anti-interference ability of the methods. The fewer misdetections there were, the better the anti-interference performance of the method in question was.

As shown in Figure 5, a hydrophone was tied to an unmovable cantilever with soft nylon string and released into a barrel filled with water. The parameters and settings of the hydrophone are shown in Table 6. Knocking sounds were made in the air (point A), on the barrel wall (point B), and in the water (point C), respectively (50 times each, totaling 150 times). The time–frequency spectrum of the clicks and knocking sounds at points A, B, and C are shown in Figure 6. Being affected by the air–water interface, the knocking sounds at point A were greatly weakened and could easily be eliminated using STFT and WT. The knocking sounds at points B and C were extremely similar to clicks and were difficult for the manual method and STFT to distinguish. WT had a better time resolution and could distinguish the knocking sounds through their duration. Based on analysis of the data obtained by the manual, STFT, and WT methods, respectively, the results are shown in Table 7. It can be seen that WT had fewer misdetections of knocking sounds than the other methods.



Figure 5. Sketch map of the knocking test.

Table 6. Parameters of the hydrophone for data collection.



Figure 6. The time–frequency spectrum of clicks and knocking sounds. (**a**) A knock sound at point A. (**b**) A knock sound at point B. (**c**) A knock sound at point C. (**d**) A typical click signal.

Type of Signal	Knock Number	Manual	STFT	WT
А	50	32	15	11
В	50	50	50	4
С	50	50	43	7

Table 7. The misdetection numbers of the different methods.

5. Conclusions

In this study, we compared WT with two other widely used methods for the detection of *Cetacean* acoustic signals. The results of the experiments showed that WT has certain advantages in the detection of *Cetacean* click signals. It seems that WT can detect more clicks for further research on aspects such as the regularity of click signals over a long timescale, which reflects the biological clock of *Odontoceti* or their territories. It is still difficult for us to verify the accuracy of signal detection. Therefore, we conducted a misdetection test to indirectly verify the accuracy. WT had the fewest misdetections, indicating that its anti-interference ability is better than that of the manual method and STFT in the detection of clicks. As for whistle and burst pulse, STFT performed better. Theoretically, the accuracy of time–frequency analysis depends on how similar the signal and the base wavelet are. Therefore, if we were to choose a suitable wavelet for WT, it might perform better. However, since STFT is sufficient for whistle and burst-pulse detection, it is not necessary to modify WT. In conclusion, in *Cetacean* acoustic signal detection, traditional STFT is better for whistles and burst pulses, and WT is better for clicks.

We believe that *Odontoceti* can distinguish the click signals emitted by themselves, avoiding interference from other individuals in the group. More precise click-detection methods will aid in the study of how these animals identify individuals. The existing studies have shown that the duration and interval of click signals are related to the distance and purpose of *Odontoceti*'s detection targets. The high time resolution of WT is very useful in this research. By combining hydrophone data and video monitoring data, we can expect that we will be able to determine the relationship between *Odontoceti* acoustic signals and some of their behaviors in the future. This study also offers suggestions on how to build more accurate datasets for different acoustic signals. It will be useful for the training of more accurate artificial neural networks for edge-computing signal-detecting devices.

Author Contributions: R.H. and Y.D. contributed equally to this paper. Conceptualization and methodology, R.H. and Y.D.; software, R.H. and Y.Y.; validation, R.H., Y.D. and Y.Y.; formal analysis, Y.D.; investigation, R.H., Y.D. and S.L.; resources, Y.D., S.L. and W.F.; data curation, R.H.; writing—original draft preparation, R.H.; writing—review and editing, Y.D., S.L. and Y.W.; visualization, R.H.; supervision, W.F. and S.Z.; project administration and funding acquisition, Y.D. and W.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Laoshan Laboratory, grant No. LSKJ202201801, and the Shanghai Science and Technology Committee, grant number 20dz1206400.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The following software used in the experiment can be downloaded at: https://docs.scipy.org/doc/scipy-1.7.1/reference/ (accessed on 5 August 2022), Software: SciPy 1.7.1 [32]. The following software used in the experiment can be downloaded at: https://pywavelets.readthedocs.io/en/v1.1.1/ (accessed on 5 August 2022), Software: PyWavelets v1.1.1 [40].

Acknowledgments: This research was also supported by the South China Sea Fisheries Research Institute, Chinese Academy of Fisheries Science, including the hydrophone dataset.

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. Würsig, B. Ethology and Behavioral Ecology of Odontocetes, 1st ed.; Springer International Publishing: Cham, Switzerland, 2019.
- King, S.L.; Guarino, E.; Donegan, K.; McMullen, C.; Jaakkola, K. Evidence that bottlenose dolphins can communicate with vocal signals to solve a cooperative task. *R. Soc. Open Sci.* 2021, *8*, 202073. [CrossRef] [PubMed]
- 3. Au, W.W.L. *The Sonar of Dolphins*, 1st ed.; Springer New York, Inc.: New York, NY, USA, 1993.
- 4. Kellogg, W.N. Auditory Perception of Submerged Objects by Porpoises. J. Acoust. Soc. Am. 1959, 31, 1–6. [CrossRef]
- 5. MacMillan, D.C.; Han, J. Cetacean By-Catch in the Korean Peninsula—By Chance or by Design? *Hum. Ecol.* **2011**, *39*, 757–768. [CrossRef]
- 6. Trevithick, H.P.; Lee, A.P. Modern Whaling. Oil Soap 1932, 9, 145–148. [CrossRef]
- 7. Horan, R.D.; Shortle, J.S. Optimal Management of Multiple Renewable Resource Stocks: An Application to Minke Whales. *Environ. Resour. Econ.* **1999**, *13*, 435–458. [CrossRef]
- 8. Aguilar, A.; Borrell, A. Unreported catches, impact of whaling and current status of blue whales in the South European Atlantic Shelf. *Sci. Rep.* **2022**, *12*, 5491. [CrossRef] [PubMed]
- 9. di Sciara, G.N.; Würsig, B. Marine Mammals: The Evolving Human Factor, 1st ed.; Springer International Publishing: Cham, Switzerland, 2022.
- 10. Fandel, A.D. *Ambient Sound Affects Movement and Calls of Bottlenose Dolphins;* University of Maryland: College Park, MD, USA, 2022; Available online: https://drum.lib.umd.edu/handle/1903/28701 (accessed on 29 November 2022).
- 11. Pine, M.K.; Jeffs, A.G.; Wang, D.; Radford, C.A. The potential for vessel noise to mask biologically important sounds within ecologically significant embayments. *Ocean. Coast. Manag.* **2016**, 127, 63–73. [CrossRef]
- 12. Li, S.; Wu, H.; Xu, Y.; Peng, C.; Fang, L.; Lin, M.; Xing, L.; Zhang, P. Mid- to high-frequency noise from high-speed boats and its potential impacts on humpback dolphins. *J. Acoust. Soc. Am.* **2015**, *138*, 942–952. [CrossRef]
- Wang, Z.; Wu, Y.; Duan, G.; Cao, H.; Liu, J.; Wang, K.; Wang, D. Assessing the underwater acoustics of the world's largest vibration hammer (OCTA-KONG) and its potential effects on the Indo-Pacific Humpbacked Dolphin (*Sousa chinensis*). *PLoS ONE* 2014, 9, e110590. [CrossRef]
- 14. Carstensen, J.; Henriksen, O.D.; Teilmann, J. Impacts of offshore wind farm construction on harbour porpoises: Acoustic monitoring of echo-location activity using porpoise detectors (T-PODs). *Mar. Ecol. Prog. Ser.* **2006**, *321*, 295–308. [CrossRef]
- 15. Brandt, M.J.; Diederichs, A.; Betke, K.; Nehls, G. Responses of harbour porpoises to pile driving at the Horns Rev II offshore wind farm in the Danish North Sea. *Mar. Ecol. Prog. Ser.* **2011**, *421*, 205–216. [CrossRef]
- 16. Mishima, Y.; Morisaka, T.; Ishikawa, M.; Yoshida, Y. Pulsed call sequences as contact calls in Pacific white-sided dolphins (*Lagenorhynchus obliquidens*). J. Acoust. Soc. Am. 2019, 146, 409–424. [CrossRef]
- 17. Akamatsu, T.; Wang, D.; Wang, K.; Li, S.; Dong, S. Scanning sonar of rolling porpoises during prey capture dives. *J. Exp. Biol.* **2010**, *213*, 146–152. [CrossRef]
- 18. Fang, L.; Li, S.; Wang, K.; Wang, Z.; Shi, W.; Wang, D. Echolocation signals of free-ranging Indo-Pacific humpback dolphins (*Sousa chinensis*) in Sanniang Bay, China. J. Acoust. Soc. Am. 2015, 183, 1346–1352. [CrossRef]
- Li, S.; Wang, D.; Wang, K.; Hoffmann-Kuhnt, M.; Fernando, N.; Taylor, E.A.; Lin, W.; Chen, J.; Ng, T. Possible age-related hearing loss (presbycusis) and corresponding change in echolocation parameters in a stranded Indo-Pacific humpback dolphin. *J. Exp. Biol.* 2013, 216, 4144–4153. [CrossRef]
- Bailey, H.; Senior, B.; Simmons, D.; Rusin, J.; Picken, G.; Thompson, P.M. Assessing underwater noise levels during pile-driving at an offshore windfarm and its potential effects on marine mammals. *Mar. Pollut. Bull.* 2010, 60, 888–897. [CrossRef]
- 21. Schevill, W.E.; Lawrence, B. Underwater Listening to the White Porpoise (*Delphinapterus leucas*). Science **1949**, 109, 143–144. [CrossRef]
- 22. Kellogg, W.N. Echo ranging in the porpoise. *Science* **1958**, *128*, 982. [CrossRef]
- 23. Fang, L.; Wang, D.; Li, Y.; Cheng, Z.; Pine, M.K.; Wang, K.; Li, S. The source parameters of echolocation clicks from captive and free-ranging Yangtze finless porpoises (*Neophocaena asiaeorientalis asiaeorientalis*). *PLoS ONE* **2015**, *10*, e0129143. [CrossRef]
- 24. Bullock, T.H.; Grinnell, A.D.; Ikezono, E.; Kameda, K.; Katsuki, Y.; Nomoto, M.; Sato, O.; Suga, N.; Yanagisawa, K. Electrophysiological studies of central auditory mechanisms in cetaceans. *Z. Für Vgl. Physiol.* **1968**, *59*, 117–156. [CrossRef]
- 25. Tyagi, K.D.; Bahl, R.; Kumar, A.; Saxena, S.; Kumar, S. Study of Noise Interfering with Dolphin Clicks. *Lect. Notes Electr. Eng.* **2019**, *526*, 353–362.
- 26. Saffari, A.; Khishe, M.; Zahiri, S.-H. Fuzzy-ChOA: An improved chimp optimization algorithm for marine mammal classification using artificial neural network. *Analog. Integr. Circuits Signal Process.* **2022**, *111*, 403–417. [CrossRef]
- Yang, Y.; He, R.; Dai, Y.; Fang, L.; He, L. The detection method of dolphin vocal endpoint based on time–frequency characteristics. *J. Appl. Acoust.* 2022, *8*, 1–12. Available online: http://kns.cnki.net/kcms/detail/11.2121.o4.20220811.1518.004.html (accessed on 13 August 2022).
- 28. Nanavati, S.P.; Panigrahi, P.K. Wavelet transform. Resonance 2004, 9, 50-64. [CrossRef]
- 29. Dutt, R.; Balouria, A.; Acharyya, A. Discrete wavelet transform based methodology for radar pulse deinterleaving. *CSIT* **2019**, 7, 141–147. [CrossRef]
- 30. Yasin, A.S.; Pavlova, O.N.; Pavlov, A.N. Speech signal filtration using double-density dual-tree complex wavelet transform. *Tech. Phys. Lett.* **2016**, *42*, 865–867. [CrossRef]

- 31. Rachdi, L.T.; Meherzi, F. Continuous Wavelet Transform and Uncertainty Principle Related to the Spherical Mean Operator. *Mediterr. J. Math.* 2017, 14, 11. [CrossRef]
- 32. SciPy Documentation—SciPy v1.7.1 Manual. Available online: https://docs.scipy.org/doc/scipy-1.7.1/reference/ (accessed on 5 August 2022).
- 33. Berry, M.V.; Lewis, Z.V. On the Weierstrass-Mandelbrot Fractal Function. Proc. R. Soc. Lond. 1980, 370, 459–484.
- Guido, R.C.; Pedroso, F.; Contreras, R.C.; Rodrigues, L.C.; Guariglia, E.; Neto, J.S. Introducing the Discrete Path Transform (DPT) and its applications in signal analysis, artefact removal, and spoken word recognition. *Digit. Signal Process.* 2021, 117, 103158. [CrossRef]
- 35. Guariglia, E. Harmonic Sierpinski Gasket and applications. *Entropy* **2018**, *20*, 714. [CrossRef] [PubMed]
- 36. Guariglia, E. Primality, Fractality, and Image Analysis. *Entropy* **2019**, 21, 304. [CrossRef] [PubMed]
- Guariglia, E.; Silvestrov, S. Fractional-Wavelet Analysis of Positive definite Distributions and Wavelets on D'(C). In *Engineering Mathematics II*; Silvestrov, S., Rančić, M., Eds.; Springer International Publishing: Cham, Switzerland, 2016; Volume 179, pp. 337–353.
- Zheng, X.; Tang, Y.Y.; Zhou, J. A framework of adaptive multiscale wavelet decomposition for signals on undirected graphs. *IEEE Trans. Signal Process.* 2019, 67, 1696–1711. [CrossRef]
- Yang, L.; Su, H.; Zhong, C.; Meng, Z. Hyperspectral image classification using wavelet transform-based smooth ordering. *Int. J. Wavelets Multiresolution Inf. Process.* 2019, 17, 1950050. [CrossRef]
- 40. Lee, G.R.; Gommers, R.; Wasilewski, F.; Wohlfahrt, K.; O'Leary, A. PyWavelets: A Python package for wavelet analysis. *J. Open Source Softw.* **2019**, *4*, 1237. [CrossRef]
- 41. Bolós, V.J.; Benítez, R.; Ferrer, R. A new wavelet tool to quantify non-periodicity of non-stationary economic time series. *Mathematics* **2020**, *8*, 844. [CrossRef]
- 42. Łępicka, M.; Górski, G.; Grądzka-Dahlke, M.; Litak, G.; Ambrożkiewicz, B. Analysis of tribological behaviour of titanium nitride-coated stainless steel with the use of wavelet-based methods. *Arch. Appl. Mech.* **2021**, *91*, 4475–4483. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.