

sensors

Wearable and Nearable Biosensors and Systems for Healthcare

Edited by

Marco Di Rienzo and Ramakrishna Mukkamala

Printed Edition of the Special Issue Published in *Sensors*

Wearable and Nearable Biosensors and Systems for Healthcare

Wearable and Nearable Biosensors and Systems for Healthcare

Editors

Marco Di Rienzo

Ramakrishna Mukkamala

MDPI • Basel • Beijing • Wuhan • Barcelona • Belgrade • Manchester • Tokyo • Cluj • Tianjin



Editors

Marco Di Rienzo
IRCCS Fondazione Don Carlo
Gnocchi
Italy

Ramakrishna Mukkamala
University of Pittsburgh
USA

Editorial Office

MDPI
St. Alban-Anlage 66
4052 Basel, Switzerland

This is a reprint of articles from the Special Issue published online in the open access journal *Sensors* (ISSN 1424-8220) (available at: https://www.mdpi.com/journal/sensors/special_issues/Nearable).

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

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

ISBN 978-3-0365-0974-7 (Hbk)

ISBN 978-3-0365-0975-4 (PDF)

© 2021 by the authors. Articles in this book are Open Access and distributed under the Creative Commons Attribution (CC BY) license, which allows users to download, copy and build upon published articles, as long as the author and publisher are properly credited, which ensures maximum dissemination and a wider impact of our publications.

The book as a whole is distributed by MDPI under the terms and conditions of the Creative Commons license CC BY-NC-ND.

Contents

About the Editors	vii
Marco Di Rienzo and Ramakrishna Mukkamala Wearable and Nearable Biosensors and Systems for Healthcare Reprinted from: <i>Sensors</i> 2021 , <i>21</i> , 1291, doi:10.3390/s21041291	1
Ikhtiyor Majidov and Taegkeun Whangbo Efficient Classification of Motor Imagery Electroencephalography Signals Using Deep Learning Methods Reprinted from: <i>Sensors</i> 2019 , <i>19</i> , 1736, doi:10.3390/s19071736	5
Dorien Huysmans, Pascal Borzée, Dries Testelmans, Bertien Buyse, Tim Willemen, Sabine Van Huffel and Carolina Varon Evaluation of a Commercial Ballistocardiography Sensor for Sleep Apnea Screening and Sleep Monitoring Reprinted from: <i>Sensors</i> 2019 , <i>19</i> , 2133, doi:10.3390/s19092133	19
Nicholas B. Bolus, Hyeon Ki Jeong, Daniel C. Whittingslow and Omer T. Inan A Glove-Based Form Factor for Collecting Joint Acoustic Emissions: Design and Validation Reprinted from: <i>Sensors</i> 2019 , <i>19</i> , 2683, doi:10.3390/s19122683	37
Michael Del Rosario, Nigel Lovell and Stephen Redmond Learning the Orientation of a Loosely-Fixed Wearable IMU Relative to the Body Improves the Recognition Rate of Human Postures and Activities Reprinted from: <i>Sensors</i> 2019 , <i>19</i> , 2845, doi:10.3390/s19132845	49
Yang Yao, Sungtae Shin, Azin Mousavi, Chang-Sei Kim, Lisheng Xu, Ramakrishna Mukkamala and Jin-Oh Hahn Unobtrusive Estimation of Cardiovascular Parameters with Limb Ballistocardiography Reprinted from: <i>Sensors</i> 2019 , <i>19</i> , 2922, doi:10.3390/s19132922	77
Stefanie Passler, Niklas Müller and Veit Senner In-Ear Pulse Rate Measurement: A Valid Alternative to Heart Rate Derived from Electrocardiography? Reprinted from: <i>Sensors</i> 2019 , <i>19</i> , 3641, doi:10.3390/s19173641	95
Federica Landreani, Andrea Faini, Alba Martin-Yebra, Mattia Morri, Gianfranco Parati and Enrico Gianluca Caiani Assessment of Ultra-Short Heart Variability Indices Derived by Smartphone Accelerometers for Stress Detection Reprinted from: <i>Sensors</i> 2019 , <i>19</i> , 3729, doi:10.3390/s19173729	111
Michel de Mathelin, Florent Nageotte, Philippe Zanne and Birgitta Dresch-Langley Sensors for Expert Grip Force Profiling: Towards Benchmarking Manual Control of a Robotic Device for Surgical Tool Movements † Reprinted from: <i>Sensors</i> 2019 , <i>19</i> , 4575, doi:10.3390/s19204575	127
Marco Di Rienzo, Giovannibattista Rizzo, Zeynep Melike Işilay and Prospero Lombardi SeisMote: A Multi-Sensor Wireless Platform for Cardiovascular Monitoring in Laboratory, Daily Life, and Telemedicine Reprinted from: <i>Sensors</i> 2019 , <i>19</i> , 680, doi:10.3390/s20030680	147

Shuai Yu and Sheng Liu

A Novel Adaptive Recursive Least Squares Filter to Remove the Motion Artifact in Seismocardiography

Reprinted from: *Sensors* **2020**, *20*, 1596, doi:10.3390/s20061596 163

Francesco Asci, Giovanni Costantini, Pietro Di Leo, Alessandro Zampogna, Giovanni Ruoppolo, Alfredo Berardelli, Giovanni Saggio and Antonio Suppa

Machine-Learning Analysis of Voice Samples Recorded through Smartphones: The Combined Effect of Ageing and Gender

Reprinted from: *Sensors* **2020**, *20*, 5022, doi:10.3390/s20185022 177

Muhammad Farooq, Talha Iqbal, Patricia Vazquez, Nazar Farid, Sudhin Thampi, William Wijns and Atif Shahzad

Thin-Film Flexible Wireless Pressure Sensor for Continuous Pressure Monitoring in Medical Applications

Reprinted from: *Sensors* **2020**, *20*, 6653, doi:10.3390/s20226653 195

About the Editors

Marco Di Rienzo received his MSc degree in Electronic Engineering from the Politecnico di Milano, Italy in 1980. He is the co-ordinator of the technological research in cardiovascular, wearable sensors and telemedicine areas at the IRCCS Fondazione Don Carlo Gnocchi. Adjunct Professor at the Faculty of Medicine, Università Statale, Milan. Research interests include signal processing, modelling of the cardiovascular control, cardiac mechanics, sleep, space physiology, seismocardiography and development of wearable systems for bio-signal monitoring. He is the author of more than 140 papers in peer-reviewed journals, co-inventor of four patents and serves as editor and referee for several international journals.

Ramakrishna Mukkamala is a Professor of Bioengineering and Anesthesiology and Perioperative Medicine at the University of Pittsburgh. He received graduate and post-doctoral training in bioelectrical engineering at MIT. He has been a dedicated cardiovascular researcher throughout his career. His interests include computational physiology, medical devices, mHealth, patient monitoring, physiologic sensors, and physiologic signal processing. He received an IEEE EMBS Most Impactful Paper Award and the Michigan State University Innovation of the Year Award in 2019 for his cuff-less blood pressure work.

Wearable and Nearable Biosensors and Systems for Healthcare

Marco Di Rienzo ^{1,*} and Ramakrishna Mukkamala ²¹ Polo Tecnologico, IRCCS Fondazione Don Carlo Gnocchi ONLUS, 20148 Milano, Italy² Department of Bioengineering and Department of Anesthesiology and Perioperative Medicine, University of Pittsburgh, Pittsburgh, PA 15261, USA; rmukkamala@pitt.edu

* Correspondence: mdirienzo@dongnocchi.it

Biosensors and systems in the form of wearables and “nearables” (i.e., everyday sensorized objects with transmitting capabilities such as smartphones) are rapidly evolving for use in healthcare. Unlike conventional approaches, these technologies can enable seamless or on-demand physiological monitoring anytime and anywhere. Such monitoring can be beneficial in various ways. Most notably, it can help transform healthcare from the current reactive, one-size-fits-all, hospital-centered, and volume-based system into a future proactive, personalized, decentralized, and valued-based system. This new system and other benefits of the technology hold great promise for longer and healthier living.

Wearable and nearable biosensors and systems have been made possible through integrated innovations in sensor design, electronics, data transmission, power management, and signal processing. Examples of measurements offered by these technologies include biopotentials, body motion, pressure, blood volume, temperature, and biochemical markers. Although much progress has been made in this field, many open challenges for the scientific community remain, especially for those applications requiring high accuracy.

The aim of this Special Issue of *Sensors* is to provide an open collection of state-of-the-art investigations on wearable and nearable biosensors and systems in order to foster further technological advances and the use of the technology to benefit healthcare. The 12 papers that constitute this Special Issue offer both depth and breadth pertaining to wearable and nearable technology [1–12]. Depth is afforded through a critical mass of studies on accelerometers [3–5,7,9,10], signal processing [1,2,4,5,7,10,11], and cardiovascular monitoring applications [2,5–7,9,10,12], whereas breadth is given through new biosensors [12] and data transmission [9], other clinical applications including surgical training [8] and brain-computer interfaces [1], and validation of commercial devices [6], which is crucial for adoption. We provide a flavor of each contribution below in order of appearance in this issue.

Majidov et al. [1] developed a machine learning technique to analyze EEG signals from a wearable electrode cap for brain–computer interface (BCI) applications. The technique was developed using a formal BCI competition dataset in which 18 subjects performed an imaginary movement of hands and feet, and comprised a number of analytical tools including online deep learning with data augmentation. They showed that the technique was able to increase the classification accuracy compared to earlier techniques.

Huysmans et al. [2] developed a machine learning technique to analyze a bed-based ballistocardiography (BCG) signal (i.e., a measure of the whole-body movement induced by the heartbeat) for sleep apnea screening and sleep monitoring. The fully automatic technique employed unsupervised, k-means clustering to reveal artifact (apnea) versus clean signals. They showed that the technique can detect subjects with significant apnea while also revealing how the bed pressure sensor should be used to enable future supervised learning.

Bolus et al. [3] developed a glove with a fingertip mounted accelerometer for monitoring the health of joints via acoustic emissions. The new form factor for measuring joint sounds eliminates the need for consumables like tape and associated interface noise. They

Citation: Rienzo, M.D.; Mukkamala, R. Wearable and Nearable Biosensors and Systems for Healthcare. *Sensors* **2021**, *21*, 1291. <https://doi.org/10.3390/s21041291>

Received: 7 February 2021

Accepted: 9 February 2021

Published: 11 February 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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/>).

showed that the device can yield reliable measurements under constant fingertip contact force in subjects during an intervention to alter the knee joint sound.

Del Rosario et al. [4] developed a machine learning technique to determine body position from a smartphone inertial measurement unit (IMU) placed at an arbitrary orientation for various applications including fall detection. The technique uses hand-crafted instead of deep-learning-based IMU features learned during walking periods as a reference for upright body posture. They showed that this technique can separate standing versus sedentary periods using only one smartphone IMU in the pocket of younger and older subjects.

Yao et al. [5] developed a wearable system for cuff-less tracking of blood pressure changes. The system measures a BCG signal via an armband accelerometer and a photoplethysmography (PPG) signal via a finger clip, extracts data-driven features including the time delay between the signals (pulse transit time), and performs cuff calibration to map the features to blood pressure. They showed that this system as well as a BCG-PPG weighing scale system could track blood pressure changes during interventions in healthy subjects.

Passier et al. [6] performed a validation study to compare two commercial in-ear PPG sensor devices for detecting heart rate during intense physical activity. The study is unique in terms of assessing external auditory canal PPG sensor devices and included 20 subjects during graded cycling. Both devices attained acceptable mean absolute heart rate errors compared to the reference ECG over a wide heart rate range but were not particularly precise.

Landreani et al. [7] developed a signal processing technique to quantify ultra-short-term heart rate variability via a smartphone accelerometer for stress detection. The technique involves placing the smartphone on the abdomen and detecting each heartbeat from the resulting BCG signal via cross correlation with a template. They showed the efficacy of the technique in detecting vagal withdrawal during mental arithmetic in healthy subjects.

de Mathelin et al. [8] developed a glove for establishing objective criteria of the expertise needed for surgeons to operate a transluminal robotic assistance system. The glove includes 12 wireless force sensitive resistors for measuring hand grip forces under visual feedback from the system. They revealed important differences in the handgrip forces of an expert versus a novice in performing an exemplary pick and drop task.

Di Rienzo et al. [9] developed a wearable acquisition platform for the monitoring of various cardiovascular features including pulse transit time and cardiac contractility. The platform is capable of measuring 36 signals from 12 wireless nodes, including ECG, seismocardiography (SCG, i.e., accelerometer-based measure of chest vibrations caused by the heartbeat), and PPG sensors. Field tests showed that the system can acquire good quality data in real life with a synchronization error between nodes lower than 1ms.

Yu et al. [10] developed a signal processing technique to remove the common motion artifact in the SCG signal. Since the artifact is typically mixed with the heartbeat in time and frequency, the technique is based on adaptive recursive least squares. They showed that the technique could extract a clear signal without further processing from only one accelerometer and detect the heart rate with high accuracy in healthy subjects.

Asci et al. [11] developed a machine learning technique to analyze smartphone voice signals for assessing physiological aging. The technique extracts thousands of signal features, performs feature reduction, and then applies a support vector machine to classify age and gender. They notably showed the efficacy of the technique in subjects in a free-living scenario to eliminate potential voice changes in supervised conditions.

Farooq et al. [12] developed a thin-film flexible wireless pressure sensor for interface pressure monitoring during leg compression treatment of venous insufficiency. The sensor is based on a pressure-dependent capacitance and inductive coil to allow passive and wireless measurement. Through analytical and experimental testing, they showed that the sensor offers sensitivity that is competitive with existing technology but with lower-cost fabrication.

We hope that this editorial serves as a useful guide to this collection of papers and that the Special Issue does end up inspiring future efforts to bring an array of wearable and nearable biosensors and systems to healthcare practices.

Funding: This work was supported in part by the Italian Ministry of Health and the US National Institutes of Health Grant EB027276.

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. Majidov, I.; Whangbo, T. Efficient Classification of Motor Imagery Electroencephalography Signals Using Deep Learning Methods. *Sensors* **2019**, *19*, 1736. [[CrossRef](#)] [[PubMed](#)]
2. Huysmans, D.; Borzée, P.; Testelmans, D.; Buyse, B.; Willems, T.; Van Huffel, S.; Varon, C. Evaluation of a Commercial Ballistocardiography Sensor for Sleep Apnea Screening and Sleep Monitoring. *Sensors* **2019**, *19*, 2133. [[CrossRef](#)] [[PubMed](#)]
3. Bolus, N.B.; Jeong, H.K.; Whittingslow, D.C.; Inan, O.T. A Glove-Based Form Factor for Collecting Joint Acoustic Emissions: Design and Validation. *Sensors* **2019**, *19*, 2683. [[CrossRef](#)] [[PubMed](#)]
4. Del Rosario, M.B.; Lovell, N.H.; Redmond, S.J. Learning the Orientation of a Loosely-Fixed Wearable IMU Relative to the Body Improves the Recognition Rate of Human Postures and Activities. *Sensors* **2019**, *19*, 2845. [[CrossRef](#)] [[PubMed](#)]
5. Yao, Y.; Shin, S.; Mousavi, A.; Kim, C.-S.; Xu, L.; Mukkamala, R.; Hahn, J.-O. Unobtrusive Estimation of Cardiovascular Parameters with Limb Ballistocardiography. *Sensors* **2019**, *19*, 2922. [[CrossRef](#)] [[PubMed](#)]
6. Passler, S.; Müller, N.; Senner, V. In-Ear Pulse Rate Measurement: A Valid Alternative to Heart Rate Derived from Electrocardiography? *Sensors* **2019**, *19*, 3641. [[CrossRef](#)] [[PubMed](#)]
7. Landreani, F.; Faini, A.; Martin-Yebra, A.; Morri, M.; Parati, G.; Caiani, E.G. Assessment of Ultra-Short Heart Variability Indices Derived by Smartphone Accelerometers for Stress Detection. *Sensors* **2019**, *19*, 3729. [[CrossRef](#)] [[PubMed](#)]
8. de Mathelin, M.; Nageotte, F.; Zanne, P.; Dresch-Langley, B. Sensors for Expert Grip Force Profiling: Towards Benchmarking Manual Control of a Robotic Device for Surgical Tool Movements. *Sensors* **2019**, *19*, 4575. [[CrossRef](#)] [[PubMed](#)]
9. Di Rienzo, M.; Rizzo, G.; Işilay, Z.M.; Lombardi, P. SeisMote: A Multi-Sensor Wireless Platform for Cardiovascular Monitoring in Laboratory, Daily Life, and Telemedicine. *Sensors* **2020**, *20*, 680. [[CrossRef](#)] [[PubMed](#)]
10. Yu, S.; Liu, S. A Novel Adaptive Recursive Least Squares Filter to Remove the Motion Artifact in Seismocardiography. *Sensors* **2020**, *20*, 1596. [[CrossRef](#)] [[PubMed](#)]
11. Asci, F.; Costantini, G.; Di Leo, P.; Zampogna, A.; Ruoppolo, G.; Berardelli, A.; Saggio, G.; Suppa, A. Machine-Learning Analysis of Voice Samples Recorded through Smartphones: The Combined Effect of Ageing and Gender. *Sensors* **2020**, *20*, 5022. [[CrossRef](#)] [[PubMed](#)]
12. Farooq, M.; Iqbal, T.; Vazquez, P.; Farid, N.; Thampi, S.; Wijns, W.; Shahzad, A. Thin-Film Flexible Wireless Pressure Sensor for Continuous Pressure Monitoring in Medical Applications. *Sensors* **2020**, *20*, 6653. [[CrossRef](#)] [[PubMed](#)]

Article

Efficient Classification of Motor Imagery Electroencephalography Signals Using Deep Learning Methods

Ikhtiyor Majidov and Taegkeun Whangbo *

Department of Computer Science Gachon University, Sujeong-Gu, Seongnam-Si, Gyeonggi-Do 13109, Korea; ixtiyoruz312@gmail.com

* Correspondence: twhangbo@gmail.com

Received: 24 February 2019; Accepted: 8 April 2019; Published: 11 April 2019

Abstract: Single-trial motor imagery classification is a crucial aspect of brain–computer applications. Therefore, it is necessary to extract and discriminate signal features involving motor imagery movements. Riemannian geometry-based feature extraction methods are effective when designing these types of motor-imagery-based brain–computer interface applications. In the field of information theory, Riemannian geometry is mainly used with covariance matrices. Accordingly, investigations showed that if the method is used after the execution of the filterbank approach, the covariance matrix preserves the frequency and spatial information of the signal. Deep-learning methods are superior when the data availability is abundant and while there is a large number of features. The purpose of this study is to a) show how to use a single deep-learning-based classifier in conjunction with BCI (brain–computer interface) applications with the CSP (common spatial features) and the Riemannian geometry feature extraction methods in BCI applications and to b) describe one of the wrapper feature-selection algorithms, referred to as the particle swarm optimization, in combination with a decision tree algorithm. In this work, the CSP method was used for a multiclass case by using only one classifier. Additionally, a combination of power spectrum density features with covariance matrices mapped onto the tangent space of a Riemannian manifold was used. Furthermore, the particle swarm optimization method was implied to ease the training by penalizing bad features, and the moving windows method was used for augmentation. After empirical study, the convolutional neural network was adopted to classify the pre-processed data. Our proposed method improved the classification accuracy for several subjects that comprised the well-known BCI competition IV 2a dataset.

Keywords: tangent space; Riemannian geometry; particle swarm optimization (PSO); BCI; EEG; electro-oscillography (EOG); CSP; FBCSP (filter bank common spatial pattern); online learning

1. Introduction

Creating brain–computer interface (BCI) applications based on electroencephalograms (EEG) is a challenging scientific task given that they translate the mental imagery to sets of commands without using any muscles. BCI applications are a valuable part of neuroscience, neural engineering, and medicine, in which robotics or mental issue detectors are used. To date, they have been used extensively in many areas of medicine to help people by connecting their minds to control devices or by detecting brain abnormalities [1–3]. This can be recognized by measuring the electric or magnetic fields generated by the central nervous system using electroencephalography (EEG) or magnetoencephalography (MEG) [4]. Electroencephalographic signals are usually recorded with the placement and use of EEG sensors with another kind of electrode placed onto the surface of the scalp using a 10–20 electrode placement system (Jasper 1958).

Frequency-based features have gained significant importance in this area, and various frequency-based feature extraction methods have been proposed [5]. The best-known approach is the power spectrum density approach [5]. It is computed using the classical signal processing algorithm known as the Fourier transform (FT) or the computationally efficient fast Fourier transform (FFT) form. The problem with this approach is that it lacks the capacity to preserve spatial information.

One of the algorithms used to retrieve spatial information from EEG signals is the common spatial pattern (CSP) algorithm. CSP first appeared in the area of EEG and MEG analyses as used by Koles [6]. It was initially proposed for two different classes, such as for left or right-hand movements [7]. It computes spatial filters which maximize the variance ratio of one of the label conditions with respect to another. The main disadvantage of this algorithm is that it works only with two classes. To solve this problem, several approaches have been proposed, such as the pairwise and the one-versus-rest approaches; however, the main problem of these algorithms is that they use multiple classifiers to obtain the final result. To use only one classifier, we employed the simultaneous diagonalization approach in conjunction with the information-theory-based approach, based on information theoretical feature extraction (ITFE) [4], which uses only one classifier. The detailed explanation of the ITFE method will be discussed in a subsequent subsection.

The presence of noise makes the classification of EEG signals difficult. EEG signals are prone to external and internal noise. Numerous methods have been proposed to reduce noise in EEG signals [8], such as independent component analysis (ICA) [9]. To obtain subject-dependent spatial and frequency-based features, combinations of these features are used. In the filterbank CSP (FBCSP) approach, filterbanks and CSPs were used [10].

Recently, Riemannian geometry-based feature extraction and classification methods have gained significant importance in BCI applications. These methods were initially used in BCI applications in [11] and won the first place in the BCI Challenge NER 2015. The same authors used the distance between the covariance matrix and Riemannian mean covariance matrix as a classification feature. Furthermore, Barachant et al. [11] introduced mapping covariance matrices on the tangent space of the Riemannian manifold. Mapped covariance matrices are represented as vectors in the tangent space.

The use of deep-learning-based classification methods in BCI applications is rare, owing to the complexity of the recording and the limited numbers of signals. In [12], Bashivan et al. used power spectrum densities based on three frequency ranges of EEG signals and generated images for each range by interpolating the topological features that preserved the surfaces of the brain. They used the VGG (visual geometry group) model and mixed 1D convolutions and long short-term memory (LSTM) layers. The results showed that the ConvNet and LSTM/1D-Conv yielded the best results compared to other architectures. In [13], the architecture was also based on convolutional neural networks, and the authors used the convolutional layer first and then the encoder part of the AutoEncoder. Additionally, they also used the power spectral densities of the fast Fourier transforms as a feature set.

In this work, initially, signals were augmented as represented in Section 3.1, filtered with the given filterbanks, and fed to the ITFE-based multiclass CSP algorithm to reject the unimportant parts. Thereafter, the results of the CSP method were concatenated and covariance matrices were computed. Next, the covariance matrices were mapped onto the tangent space of the Riemannian manifold, and the results were concatenated with flattened covariance matrices. Finally, power spectrum density (PSD) features were also concatenated with the last results, and training data was formed. Furthermore, the particle swarm optimization method was used to replace bad features with the linear average values of the neighboring features. Eventually, one layer of the convolutional neural network was implied to classify the pre-processed data.

The rest of this article is constructed as follows: Sections 2.1 and 2.2 explain the algorithm filterbank CSP (FBCSP), and Sections 2.3 and 2.4, respectively, explain the power spectrum density and the Riemannian geometry as well as its usage. Section 3 discusses our proposed method where the data augmentation method is presented in Section 3.1, and the filterbank particle swarm optimization

(FBPSO) feature-selection algorithm is represented in Section 3.2. Accordingly, datasets and generated results are presented in Section 4, followed by a brief summary of the study.

2. Related Studies

2.1. Filterbank Common Spatial Pattern

The FBCSP [10] is shown in Figure 1. The first stage is the bandpass filtering stage that uses ICA decomposition to reduce noise in the second stage. Moreover, the third stage fitted the CSP and formulated the spatial filter, followed by the application of the spatial filters and transformation of the results into CSP spaces.

2.1.1. Filtering at the Base Frequency

Initially, the signal was divided into filterbanks with the help of IIR filters. The cut-off frequencies of the filters were chosen in the range of 8–36 Hz because better results were obtained in this range.

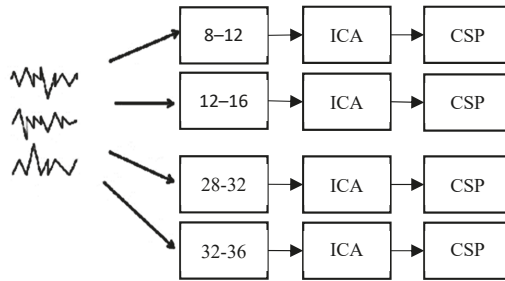


Figure 1. Division of signals into subsignals by filtering followed by the application of spatial filters, as described by [10]. CSP: common spatial pattern.

2.1.2. Common Spatial Patterns for Two Classes

The CSP algorithm was then applied after the blind ICA source separation algorithm was deployed because it constitutes one of the best ways to reduce noise in brain signals. The core idea of CSP is the maximization of one class of features and the minimization of another so that the resulting signals encode the most significant information [14].

If two conditions, a and b, and the respective matrices X^a and X^b exist, they define an $N \times T$ shape, where N is the number of electrodes, and T is the number of samples per electrode. Firstly, we have to find the normalized covariance of the matrices of the trials:

$$R_a^i = \frac{X_i^a X_i^{aT}}{\text{trace}(X_i^a X_i^{aT})}, \quad (1)$$

where R_a^i indicates the normalized covariance matrix of the i^{th} trial of a group or a condition “a” (similarly for R_b^i). Subsequently, the normalized covariance matrices are averaged according to

$$R^a = \frac{\sum_{i=1}^n R_a^i}{n}, \quad R^b = \frac{\sum_{i=1}^n R_b^i}{n}, \quad (2)$$

The summation of these allows the formulation of the composed matrix R^c and the estimation of the eigenvalues and vectors, whereby $R^c = B_c \lambda B_c^T$, and B_c is the matrix of eigenvectors. λ is a diagonal matrix of eigenvalues. The whitening transform is computed by $W_{\text{whitening}} = \lambda^{-1/2} B_c^T$. Let S_a and S_b be

$$S_a = W R_a W^T \text{ and } S_b = W R_b W^T \quad (3)$$

In the next step, we identify the aforementioned eigen decomposition of these matrices, which are expressed as

$$S_a = U\psi_a U^T \text{ and } S_b = U\psi_b U^T \quad (4)$$

The eigenvalues of the equations in (4) satisfy this equation $\psi_a + \psi_b = I$. Eventually, our spatial filter will be computed according to

$$P^T = U^T W, \quad (5)$$

Thus, this filter can be applied as

$$X_f^i = X^i P^T, \quad (6)$$

Our P matrix then satisfies these equations according to the representation listed in [15],

$$P^T R_a P = D_1, \quad (7)$$

$$P^T R_b P = D_2, \quad (8)$$

which means that the spatial filter P diagonalizes both covariance matrices.

After filtering the data, we can save only the important features and discard the less informative ones. To do this, $m/2$ rows above and $m/2$ rows below the matrix are selected so that information that represents both conditions is maintained. In the end, an $m \times T$ matrix is constructed, where m is the number of selected rows.

In this study, after the completion of the filtering process, we had seven spatially filtered filterbanks with sizes of $m \times T$, where $m = 4$ and $T = 500$ in this case. These filterbanks are concatenated, as represented in Figure 2. The axis of zero is used for concatenation so that the resulting matrix has the form of $(m \times 7) \times T$; that is, 42×500 . Accordingly, the computation of the covariance matrix based on this formulation yields a 36×36 matrix.

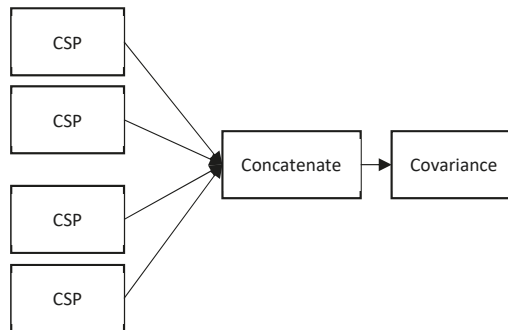


Figure 2. Graphical representation of the process of concatenation.

2.2. Multiclass Filterbank Common Spatial Pattern

The main objective associated with the use of spatial filters is to identify a signal's inner space with given conditions (classes). In this way, the CSP algorithm identifies one inner space whereby conditions are maximally represented. When multiple conditions exist (more than two), the problem is complex and difficult to handle. In [4], the information theoretic feature selection (ITFE) was proposed based on the joint approximate diagonalization (JAD) algorithm that solves minimal achievable classification and multiclass problems. It is based on the maximization of mutual information between data X and gives the labels $P^* = \text{argmax}\{I(c, P^T R)\}$ so that L rows from the data matrix X can be selected as a signal's inner space, which preserves most of the information. The implementation sequence of the algorithm is shown below.

First, the covariance matrices need to be computed as $R_{x|c_i}, i = 1, \dots, M$, where M is the number of classes.

Subsequently, the JAD algorithm has to be deployed. It is based on Equations (7) and (8), and as stated in [16], for every covariance matrix belonging to M classes, a W matrix has to be identified which diagonalizes all the covariance matrices as $W^T R_{x|c_i} W = D_{c_i}, i = 1, \dots, M$.

Next, every column of $w_j, j = 1, \dots, N$, of W where w_j is the j th column of matrix W is taken and changed so that it satisfies $w_j^T R_{x|c_i} w_j = 1$; then, the mutual information is computed according to the equation below (9) [4]:

$$I(c, w_j^T x) \approx - \sum_{i=1}^M P(c_i) \log \sqrt{w_j^T R_{x|c_i} w_j} - \frac{3}{16} \left(\sum_{i=1}^M P(c_i) \left(\left(w_j^T R_{x|c_i} w_j \right)^2 - 1 \right) \right)^2, \quad (9)$$

Eventually, L columns of W have to be selected and applied to the data X using the dot product.

2.3. Power Spectrum Density (PSD)

Simultaneously, the PSDs of all frequency bands were computed, including the mu (8–12 Hz), beta (13–25 Hz), and gamma (30–45 Hz) bands, with the measurement of the FBCSPs of signals and their concatenations, as for CSPs above. Some EEG headsets provide the PSDs automatically, and for convenience it is thus helpful to have a dataset that has this property.

2.4. Riemannian Geometry

The fundamental idea of this geometry is based on the mapping of the covariance matrix on the space which conveniently represents the data. The method learns the curve-shaped spaces that comprise Euclidean spaces. Therefore, as with the surface of the earth, the covariance matrix is located on a curve-based-space, and one should approach this accordingly. In the BCI field, it is assumed that the existence and use of the EEG signal are located in the specific curve-shaped space. Accordingly, if the signal is mapped to the defined space, it can be used more effectively. In this study, the concepts of Riemannian geometry were described, and the reader's attention is redirected to [11] and to the references listed therein for further instructions.

2.4.1. Spatial Covariance Matrices

Covariance matrices are computed using the equation given in (10). Suppose the original recording of the EEG signal is stored in matrix X , which has a shape of $N \times T$, where N is the number of channels and T is the number of samples. In the calibration mode, each EEG signal is divided into supervised segments called trials with sizes of $N \times T_s$, where T_s denotes the number of samples per trial. To apply it into Riemannian geometry-based algorithms, it should satisfy the equation $T_s \gg N$, and it should be symmetric positive-definite (SPD), which means it must be diagonalized with real positive eigenvalues [11].

$$C_i = \frac{X_i X_i^T}{N}, \quad (10)$$

2.4.2. Spatial Covariance Matrices

Let us consider that one has a set of matrices P which have $m \times m$ dimensions. As represented in [17], this set can be defined as $P(m) = \{P \in S(m) | u^T P u > 0, \forall u \in \mathbb{R}^m, u \neq 0\}$, where $S(m)$ indicates the space of all symmetric matrices. Moreover, this creates a manifold \mathcal{M} with the dimension $m(m+1)/2$.

The geodesic distance denotes the minimum length of the path between two points on the manifold \mathcal{M} . The geodesic distance between two SPD covariance matrices can be estimated using the equation below:

$$\delta_R(P_1, P_2) = \|\log(P_1^{-1}P_2)\|_F = \left[\sum_{i=1}^n \log^2 \lambda_i \right]^{1/2}, \quad (11)$$

where λ_i , $i = 1, \dots, n$ indicates the real eigenvalues of $P_1^{-1}P_2$.

2.4.3. Approximation of SPD Matrices

In Riemannian geometry, the mean of $n \geq 1$ SPD matrices, which is also referred to as the geometry mean, can be formulated based on the geodesic distance:

$$Q_{\text{mean}}(P_1, P_2, \dots, P_n) = \arg \min_{P \in P(m)} \sum_i^N \delta_R^2(P, P_i), \quad (12)$$

2.4.4. Tangent Space Mapping

Barachant et al. [11] also proposed a method which maps the covariance matrices in a tangent space of the Riemannian manifold. Each SPD matrix $P \in P(m)$, which is a point in Riemannian geometry, has a mapped version in the tangent space with the same dimension as that of the manifold $m(m+1)/2$. The tangent space mapping is computed using the following equation:

$$S_i = \text{upper}(\log(Q_{\text{mean}}^{-\frac{1}{2}} P_i Q_{\text{mean}}^{-\frac{1}{2}})), \quad (13)$$

As shown in Figure 3, each of the SPD matrices map onto the tangent space at point Q_{mean} , which is a point calculated with the use of Equation (12) and vectorized by obtaining only the triangular part of the matrix that corresponds to the dimensions of the tangent space.

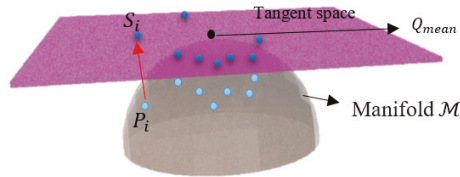


Figure 3. Graphic example of tangent space mapping, whereby the red arrow represents the exponential mapping of S_i .

2.5. Particle Swarm Optimization

Particle swarm optimization (PSO) is an optimization method that works like a genetic algorithm (GA) [18]. However, it is easier and can be implemented with a few lines of code. PSO originated from the traversing of flocks of birds and advances a problem by trying to improve the candidate solution in a continuous manner, based on a given quality ratio [19].

Algorithm 1. Simple particle swarm optimization (PSO) pseudocode.

1. **begin**
 2. initialize
 3. **for** i **in** $n_iterations = k$
 4. **for each** position p of particle compute fitness
 5. **if** fitness (p) > fitness (p_{best})
 6. $p_{\text{best}} = p$
 7. set best of p_{best} as g_{best}
 8. update particles velocity and position
 9. g_{best} is our result
-

The uniqueness of the algorithm compared to GA is based on the fact that it stores the variables for each particle's personal best position (pbest), global best position (gbest), velocity v , and current position x . As shown in the pseudocode of Algorithm 1, the individual best positions can be identified for each particle, and the best one is selected among all the pbest values of all the particles. In the long run, the velocity and the position of each particle can be updated using equation [19]:

$$v = wv + c_1 \text{rand}()(\text{pbest} - x) + c_2 \text{rand}()(\text{gbest} - x), \quad (14)$$

$$x = x + v, \quad (15)$$

where w indicates the inertia value that represents the percentage of the old velocity value which will be maintained, and $c_1 \in [0, 1]$ and $c_2 \in [0, 1]$ are known as the "acceleration coefficients" used for the selection of the pbest and gbest values, respectively. The function $\text{rand}()$ generates random (i.e., different) values between zero and one.

3. Proposed Method

In this study, a number of remarkable algorithms were used. After the data was read, data augmentation was applied. The augmentation process was completed by moving the sliced windows described in Section 3.1. The general architecture of the proposed method is depicted in Figure 4.

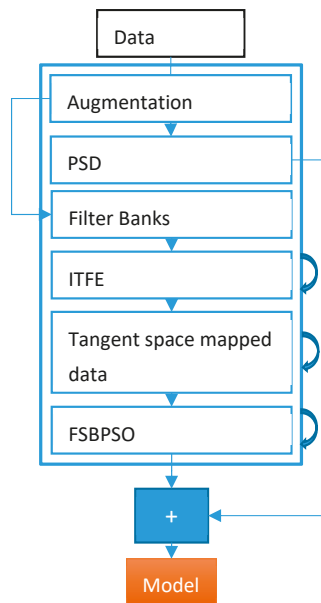


Figure 4. Overall architecture of the proposed method, where the curved arrows represent the algorithms which have two modes; that is, for training and testing.

3.1. Data Augmentation

For data augmentation, the sliced windows method with a window size of w and a moving time of t_{moving} was used. At each time instant, the window w was used such that there was a specific moving time t_{moving} , as represented in Figure 5.

3.2. FBCSP Algorithm

In this study, first, the filtering operation was used, as shown in Section 2.1, after the independent component analysis was adopted to reduce the noise of the data.

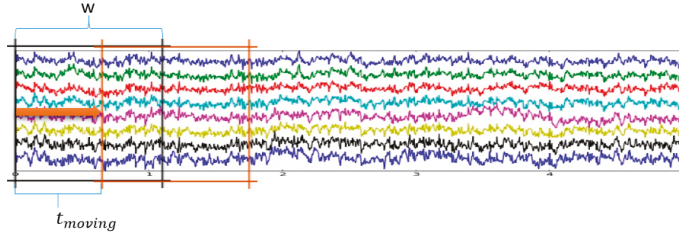


Figure 5. Graphical illustration of the augmentation process.

After the noise reduction, the CSP algorithm was deployed, as shown in Sections 2.1 and 2.2, and m rows of the spatially filtered data were selected. As mentioned earlier, they were concatenated. Suppose n_f frequency bands and n_t samples were selected in each time trial. Overall, $n \times (m \times n_f) \times n_t$ shaped data were taken, where n is the number of trials. The next step was the computation of the covariance matrices of the trials using the equation given in (10). Eventually, the data size was $n \times (m \times n_f) \times (m \times n_f)$, whereby the data comprises n square matrices.

3.3. PSD Algorithm

At the same time, the PSDs of the data were computed for the n_f^p frequency bands using three frequency bands; i.e., alpha, beta, and gamma. The dataset D_{psd} obtained from the use of this algorithm had a size of $n \times (N_{ch} n_f^p)$.

3.4. Tangent Space Mapping

After the completion of all the calculations indicated above, two types of data composed of covariance matrices and power spectral densities were retrieved.

The next step in this process was the deployment of the tangent space mapping explained in Section 2.4. The result was the vector D_{ts}^i , $i \in 1, \dots, n$, where n is the number of data samples (trials) given in our dataset with a dimension of $((m n_f)(m n_f + 1))/2$. Subsequently, for each trial, the tangent space D_{ts}^i and the flattened covariance matrix D_c^i were concatenated. Consequently, the D_{temp} matrix with the shape of $n \times ((m n_f)^2 + (m n_f)(m n_f + 1)/2)$ was obtained. Correspondingly, when the subsequent part was replaced with n_{new} , the shape of D_{temp} was $n \times n_{new}$.

3.5. FSBPSO Algorithm

In our study, the wrapper-based feature-selection algorithm called feature selection with the binary PSO algorithm FSBPSO was deployed. A similar approach was adopted in [20]. It is obvious from its name that FSBPSO is a feature-selection algorithm based on PSO optimization. As represented in Figure 6, the entire dataset was input at the beginning, and the FSBPSO initialized positions randomly by

$$X_{i,j} = \begin{cases} 1, & \text{if the feature selected} \\ 0, & \text{otherwise} \end{cases}, \quad (16)$$

where $j = j \in 1, \dots, N_{new}$, represents the number of features. Subsequently, based on the quality value, which determines the cost accuracy taken from the KNN algorithm, the pbest and gbest values were found, and the position and velocity values were updated based on Equations (14) and (15). At the

end, a vector with a shape of N_{new} that will consist of ones and zeros according to Equation (16) was achieved.

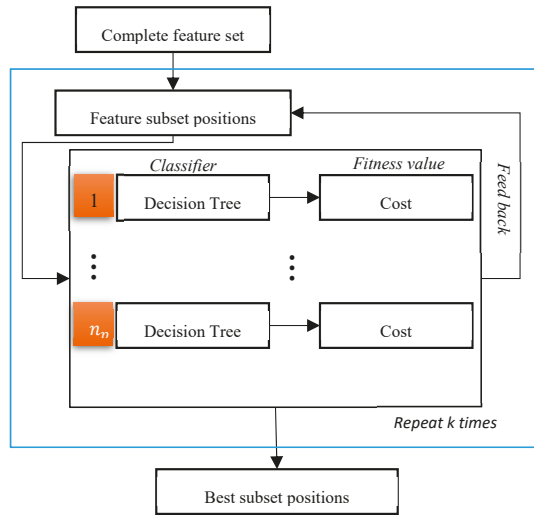


Figure 6. Graphical illustration of the filterbank particle swarm optimization (FBPSO) algorithm, where n_p is a number of particles for PSO.

At this stage, the positions of the features which best defined the data were retrieved. Herein, this feature selection was applied on the data D_{temp} , and a better accuracy than the concatenated case with D_{psd} was obtained. During the feature selection, the data with a linear interpolation scheme was interpolated to fill the missing gaps. This means that the non-selected features were effectively replaced with average values. Subsequently, to obtain the final data D , the feature selections of D_{temp} and D_{psd} were concatenated. Therefore, the final data D had shape dimensions equal to $n \times (n_{\text{new}} + N_{\text{ch}} n_f^p)$. The last part was set to $n_{\text{final}} = n_{\text{new}} + N_{\text{ch}} n_f^p$ so that in the end, D had shape dimensions equal to $n \times n_{\text{final}}$.

3.6. Architecture

The use of deep networks is very rare in the BCI field owing to the increased noise ratio and the lack of adequate data. In this study, the number of training samples or signal trials was improved using data augmentation, as mentioned earlier. The application of a number of large and deep networks has also been attempted. However, as the networks became deeper, overfitting increased. Thus, using very deep networks was stopped. Instead, to classify the data D , a 1D convolutional neural network (CNN) was used. It is not uncommon for CNNs to be used in the classification of EEG signals. In [21] and [12], 1D convolutional networks were used in combination with long short-term memory (LSTM). In the first case [21], the evoked results revealed that the combined architecture just outperformed the CNN. However, when they used only LSTM, the result was poor in comparison to the CNN. In [13], the authors used stacked CNNs as the encoder part of the AutoEncoder with the FFT features on two frequency bands: i.e., the mu (8–13Hz) and beta (13–30 Hz) bands.

In this study, two types of architectures were used. The first one was composed of CNN and output softmax layers only. In the second one, the CNN layer was followed by fully connected layers with 100 output units.

For training, the filter size was set to seven and the activation functions to rectified linear unit. To optimize our model, the Adam optimizer was used, and the learning rate applied varied based on

the dataset and subjects. The number of epochs for training was set to 20. Correspondingly, increasing this number led to overfitting.

4. Experiments and Results

4.1. Datasets

To show how well our algorithm works, two publicly available datasets, 2a and 2b from the BCI Competition IV, were selected. Both of them were mental imagery datasets, and both included left- and right-hand movements. These datasets were recorded during the imaginary movement of hands or feet and were sampled at a 250 Hz frequency rate. For detailed information, please refer to [22].

The dataset 2a consists of 22-channel data from nine subjects and was pre-split into training and testing sets. However, dataset 2b comprised nine subjects with three channels of EEG signals.

4.2. Results

In this study, experiments have been conducted with the use of the Python 3.6 environment with the MNE–Python EEG signal processing tool on an Intel (maximum 4.7 GHz) core i7 PC with 16 GB of RAM. The training and testing of each dataset was conducted separately for each subject. In the analysis process for dataset 2a, for example, the first subject’s training and testing process was completed on the A1T and A1E files, respectively. Herein, accuracies are listed only for the test set.

The results of dataset 2a are listed in Table 1. These values have been obtained using semi-supervised online learning methods, based on which our model was trained using predicted test data labels.

Table 1. Comparison of classification accuracy (in %) for the completion of the LvR task of the proposed method with published results of the dataset 2a obtained from the BCI competition IV that included left- and right-hand movement data.

Subject	Proposed Method	Raza et al. [23]	Gramound Gaur. [24]	SR-MDRM [25]	WOLA-CSP [26]
A1	90.14%	90.28%	91.49%	90.21%	86.81%
A2	62.35%	57.64%	61.27%	63.28%	63.89%
A3	89.17%	95.14%	94.89%	96.55%	94.44%
A4	79.23%	65.97%	76.72%	76.38%	68.75%
A5	76.45%	61.11%	58.52%	65.49%	56.25%
A6	69.35%	65.28%	68.52%	69.01%	69.44%
A7	84.08%	61.11%	78.57%	81.94%	78.47%
A8	86.10%	91.67%	97.01%	95.14%	97.91%
A9	87.06%	86.11%	93.85%	93.01%	93.75%
Average	80.44%	74.92%	79.93%	81.22%	78.86%

In the data augmentation process represented in Section 3.2, the number of iterations was four, the moving time was 0.3 s, and the window size w was 3 s. In the spatial filtering process, the number of selected rows m was six. The parameters of our FSBPSO algorithm during the stage of feature selection are listed in Table 2.

Table 2. Parameters for the FSBPSO algorithm.

c_1	c_2	w	n_p	K
0.9	0.8	0.7	8	200

For comparison, the results from several prior publications were distinguished based on state-of-the-art approaches. It is clear from the table that our method has comparable results with other methods (sample size = four subjects). If the average values of the results were taken into

consideration, one can see that our proposed method surpassed the majority of the other methods by reaching a classification accuracy of 80.44%, as shown in Table 1. The authors of SR-MDRM used a very similar approach to our method; however, in their work, a spatial filter regularized by coordinates of electrodes to ensure prior information from EEG channels and a minimum classification distance to the Riemannian mean was used. Furthermore, In WOLA-CSP [26], the WOLA algorithm, which consists of applying FFT to compute power spectrum density, shifting the spectrum of given frequencies measured by ERDSA, which is a specific-subject frequency selection algorithm, to the origin of the axis, and subsequently using IFFT to transform the spectrum back to the time domain, was used. Additionally, CSP to extract features and latent Dirichlet allocation (LDA) to classify the retrieved features were used; however, the data were still sparse.

In addition to this, the results obtained from the next dataset 2b are listed in Table 3. The dataset contained highly corrupted EEG signals, but the ICA algorithm managed to filter the signal. Data augmentation was conducted based on the same procedure as that used for the dataset 2a.

The features of the dataset have been obtained by the proposed algorithm, and the parameters of the FSBPSO algorithm are also the same as those for dataset 2a, as listed in Table 2.

Table 3. Comparison of the classification accuracy (in %) with other machine learning techniques based on the dataset 2b obtained from the BCI competition IV.

Subject	Proposed Method	LDA	SVM	Yousef and Ugur [13]
1	78.81%	64.58%	74.65%	76.00%
2	80.14%	58.08%	75.36%	65.80%
3	77.08%	61.11%	70.13%	75.30%
4	91.37%	67.34%	81.98%	95.30%
5	86.15%	66.55%	79.10%	83.00%
6	80.73%	59.72%	72.91%	79.50%
7	81.59%	69.79%	80.55%	74.50%
8	81.25%	67.10%	76.64%	75.30%
9	84.37%	60.06%	77.08%	73.30%
Average	82.39%	63.81%	76.49%	77.56%

To compare the results of dataset 2b, the feature set was evaluated with other well-known machine learning techniques, including the latent Dirichlet allocation (LDA) and support vector machines, and constructed similar encoder parts of the AutoEncoder with CNNs based on the formulation of the CNN-SAE model of Yousef and Ugur [13]. As shown, our method yields better results for all the tested subjects except for Subject 4, and the average value of the classification accuracy (82.39 %) also outperforms the corresponding accuracies of the other studies, as indicated in Table 3.

The comparisons listed above were for two-class cases only; that is, for the right and left hands. For multiclass cases, 10-fold cross-validation classification results for the 2a dataset from the BCI competition IV [27] are represented for all the classes, including the left hand, right hand, tongue, and both feet, as shown in Table 4.

For comparison, three related results for the dataset are represented, including the TSLDA [11], which represents the tangent space mapping feature extraction method alongside the LDA classification method, and the CSP* and the LDA method, in which the best features were selected for each subject according to the FDR criterion [11] in association with the LDA classification method. The results for the MDRM minimum distance to Riemannian manifold method [11] were also presented. Our method outperformed all the referenced examples, including the latter one.

Table 4. Comparison of cross-validation classification results (in %) with other related approaches for the 2a multiclass dataset obtained from the BCI competition IV.

Subject	Proposed Method	TSLDA	CSP*+LDA	MDRM
1	93.30%	80.50%	81.80%	77.80%
2	84.59%	51.30%	45.10%	44.10%
3	91.68%	87.59%	83.50%	76.80%
4	84.55%	59.30%	59.00%	54.90%
5	86.54%	45.00%	42.20%	43.80%
6	76.92%	55.30%	43.30%	47.10%
7	94.03%	82.10%	81.50%	72.00%
8	93.20%	84.80%	69.60%	75.20%
9	92.24%	86.10%	80.00%	76.60%
Average	87.94%	70.20%	65.11%	63.20%

5. Discussion and Conclusions

In this study, a number of well-known feature extraction methods were combined for EEG signal processing, and one of the deep-learning-based approaches was described. In addition to this, several deep learning techniques were also evaluated, such as online learning (unsupervised learning) and transfer learning. It was found that by using online learning, one could increase the classification accuracies of the EEG signals by almost 2%, while transfer learning led to bad results. It was observed that very deep networks can cause overfitting. However, if any one layer in the CNN network does not perform well, the parameters of the model are not strong enough to allow the learning of the problem. Therefore, in instances where this occurred, another fully connected layer was added immediately after the layer with poor performance. Deep-learning-based techniques often require an extensive number of resources. However, in this study, this problem was solved with the use of data augmentation.

Author Contributions: I.M. and T.W. designed the main idea of this study. I.M. programmed the Python code and performed all the experiments. T.W. supervised this study and contributed to the analysis and discussion of the algorithm and experimental results. The manuscript was prepared by I.M.

Funding: This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2019-2017-0-01630) supervised by the IITP (Institute for Information and Communications Technology Promotion).

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

References

- Vaughan, T.M.; Heetderks, W.J.; Trejo, L.J.; Rymer, W.Z.; Weinrich, M.; Moore, M.M.; Kübler, A.; Dobkin, B.H.; Birbaumer, N.; Donchin, E.; et al. Brain-computer interface technology: A review of the Second International Meeting. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2003**, *11*, 94–109. [[CrossRef](#)] [[PubMed](#)]
- Vidal, J.J. Toward direct brain-computer communication. *Annu. Rev. Biophys. Bioeng.* **1973**, *2*, 157–180. [[CrossRef](#)]
- Wolpaw, J.R.; Birbaumer, N.; McFarland, D.J.; Pfurtscheller, G.; Vaughan, T.M. Brain-computer interfaces for communication and control. *Clin. Neurophysiol.* **2002**, *113*, 767–791. [[CrossRef](#)]
- Grosse-Wentrup, M.; Buss, M. Multiclass common spatial patterns and information theoretic feature extraction. *IEEE Trans. Biomed. Eng.* **2008**, *55*, 1991–2000. [[CrossRef](#)]
- Zavaglia, M.; Astolfi, L.; Babiloni, F.; Ursino, M. The effect of connectivity on EEG rhythms, power spectral density and coherence among coupled neural populations: Analysis with a neural mass model. *IEEE Trans. Biomed. Eng.* **2008**, *55*, 69–77. [[CrossRef](#)] [[PubMed](#)]
- Koles, Z.J. The quantitative extraction and topographic mapping of the abnormal components in the clinical EEG. *Electroencephalogr. Clin. Neurophysiol.* **1991**, *79*, 440–447. [[CrossRef](#)]
- Ramoser, H.; Müller-Gerking, J.; Pfurtscheller, G. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Trans. Rehabil. Eng.* **2000**, *8*, 441–446. [[CrossRef](#)]

8. Nicolas-Alonso, L.F.; Gomez-Gil, J. Brain computer interfaces, a review. *Sensors* **2012**, *12*, 1211–1279. [[CrossRef](#)]
9. Comon, P. Independent component analysis, a new concept? *Signal Process.* **1994**, *36*, 287–314. [[CrossRef](#)]
10. Barachant, A.; Bonnet, S.; Congedo, M.; Jutten, C. Multiclass brain–computer interface classification by Riemannian geometry. *IEEE Trans. Biomed. Eng.* **2012**, *59*, 920–928. [[CrossRef](#)] [[PubMed](#)]
11. Forney, E. *Learning Deep Representations of EEG Signals in Mental-Task Brain-Computer Interfaces Using Convolutional and Recurrent Networks*; Colorado State University: Fort Collins, CO, USA, 2017.
12. Tabar, Y.R.; Halici, U. A novel deep learning approach for classification of EEG motor imagery signals. *J. Neural Eng.* **2016**, *14*, 016003. [[CrossRef](#)]
13. Ang, K.K.; Chin, Z.Y.; Zhang, H.; Guan, C. Filter bank common spatial pattern (FBCSP) in brain-computer interface. In Proceedings of the IEEE International Conference on Neural Networks (IJCNN 2008), Hong Kong, China, 1–8 June 2008.
14. Müller-Gerking, J.; Pfurtscheller, G.; Flyvbjerg, H. Designing optimal spatial filters for single-trial EEG classification in a movement task. *Clin. Neurophysiol.* **1999**, *110*, 787–798. [[CrossRef](#)]
15. Parra, L.C.; Spence, C.D.; Gerson, A.D.; Sajda, P. Recipes for the linear analysis of EEG. *Neuroimage* **2005**, *28*, 326–341. [[CrossRef](#)]
16. Ziehe, A.; Laskov, P.; Nolte, G.; MÃzler, K.-R. A fast algorithm for joint diagonalization with non-orthogonal transformations and its application to blind source separation. *J. Mach. Learn. Res.* **2004**, *5*, 777–800.
17. Islam, M.R.; Tanaka, T.; Molla, M.K.I. Multiband tangent space mapping and feature selection for classification of EEG during motor imagery. *J. Neural Eng.* **2018**, *15*, 046021. [[CrossRef](#)]
18. Boeringer, D.W.; Werner, D.H. Particle swarm optimization versus genetic algorithms for phased array synthesis. *IEEE Trans. Antennas Propag.* **2004**, *52*, 771–779. [[CrossRef](#)]
19. Blackwell, T.M.; Kennedy, J.; Poli, R. Particle swarm optimization. *Swarm Intell.* **2007**, *1*, 33–57.
20. Eberhart, R.; Kennedy, J. A new optimizer using particle swarm theory. In Proceedings of the Sixth International Symposium on Micro Machine and Human Science, Nagoya, Japan, 4–6 October 1995; pp. 39–43.
21. Zhang, Y.; Wang, S.; Phillips, P.; Ji, G. Binary PSO with mutation operator for feature selection using decision tree applied to spam detection. *Knowl. Based Syst.* **2014**, *64*, 22–31. [[CrossRef](#)]
22. Bashivan, P.; Rish, I.; Yeasin, M.; Codella, N. Learning representations from EEG with deep recurrent-convolutional neural networks. *arXiv*, 2015; arXiv:151106448.
23. Raza, H.; Cecotti, H.; Li, Y.; Prasad, G. Adaptive learning with covariate shift-detection for motor imagery-based brain–computer interface. *Soft Comput.* **2016**, *20*, 3085–3096. [[CrossRef](#)]
24. Gaur, P.; Pachori, R.B.; Wang, H.; Prasad, G. A multi-class EEG-based BCI classification using multivariate empirical mode decomposition based filtering and Riemannian geometry. *Expert Syst. Appl.* **2018**, *95*, 201–211. [[CrossRef](#)]
25. Singh, A.; Lal, S.; Guesgen, H.W. Reduce Calibration Time in Motor Imagery Using Spatially Regularized Symmetric Positives-Definite Matrices Based Classification. *Sensors* **2019**, *19*, 379. [[CrossRef](#)]
26. Belwafi, K.; Romain, O.; Gannouni, S.; Ghaffari, F.; Djemal, R.; Ouni, B. An embedded implementation based on adaptive filter bank for brain–computer interface systems. *J. Neurosci. Methods* **2018**, *305*, 1–16. [[CrossRef](#)] [[PubMed](#)]
27. Tangermann, M.; Müller, K.-R.; Aertsen, A.; Birbaumer, N.; Braun, C.; Brunner, C.; Leeb, R.; Mehring, C.; Miller, K.J.; Mueller-Putz, G. Review of the BCI competition IV. *Front. Neurosci.* **2012**, *6*, 55. [[CrossRef](#)]



Article

Evaluation of a Commercial Ballistocardiography Sensor for Sleep Apnea Screening and Sleep Monitoring

Dorien Huysmans ^{1,2,*}, Pascal Borzée ³, Dries Testelmans ³, Bertien Buyse ³, Tim Willemen ⁴, Sabine Van Huffel ^{1,2} and Carolina Varon ^{1,2}

¹ KU Leuven, Department of Electrical Engineering (ESAT), STADIUS Center for Dynamical Systems, Signal Processing, and Data Analytics, 3001 Leuven, Belgium; sabine.vanhuffel@esat.kuleuven.be (S.V.H.); carolina.varon@esat.kuleuven.be (C.V.)

² IMEC, 3001 Leuven, Belgium

³ UZ Leuven, Department of Pneumology, 3000 Leuven, Belgium; pascal.borzee@uzleuven.be (P.B.); dries.testelmans@uzleuven.be (D.T.); bertien.buyse@uzleuven.be (B.B.)

⁴ Equilli, 2800 Mechelen, Belgium; tim.willemen@equilli.com

* Correspondence: dorien.huysmans@esat.kuleuven.be; Tel.: +32-16-379-269

Received: 5 February 2019; Accepted: 4 May 2019; Published: 8 May 2019

Abstract: There exists a technological momentum towards the development of unobtrusive, simple, and reliable systems for long-term sleep monitoring. An off-the-shelf commercial pressure sensor meeting these requirements is the Emfit QS. First, the potential for sleep apnea screening was investigated by revealing clusters of contaminated and clean segments. A relationship between the irregularity of the data and the sleep apnea severity class was observed, which was valuable for screening (sensitivity 0.72, specificity 0.70), although the linear relation was limited (R^2 of 0.16). Secondly, the study explored the suitability of this commercial sensor to be merged with gold standard polysomnography data for future sleep monitoring. As polysomnography (PSG) and Emfit signals originate from different types of sensor modalities, they cannot be regarded as strictly coupled. Therefore, an automated synchronization procedure based on artefact patterns was developed. Additionally, the optimal position of the Emfit for capturing respiratory and cardiac information similar to the PSG was identified, resulting in a position as close as possible to the thorax. The proposed approach demonstrated the potential for unobtrusive screening of sleep apnea patients at home. Furthermore, the synchronization framework enabled supervised analysis of the commercial Emfit sensor for future sleep monitoring, which can be extended to other multi-modal systems that record movements during sleep.

Keywords: ballistocardiography; pressure sensor; Emfit; home monitoring; sleep recording; sleep apnea; unsupervised learning; synchronization

1. Introduction

Healthcare is evolving towards the application of automated systems for home-monitoring and pre-clinical screening to complement diagnostic routines. The current reference practice for diagnosis of sleep-related pathologies is a labor-intensive overnight stay in a specialized sleep center. There, a polysomnography (PSG) is performed, requiring the patient to wear encephalography electrodes, oronasal airflow sensors, thoracic and abdominal belts, electrocardiography (ECG) sensors, an oxygen saturation finger-clip sensor, a body position sensor, chin and leg electromyography and electrooculography sensors over a full night. This setup is highly obtrusive for the patient and impedes a normal night's sleep. Moreover, the PSG procedure requires well-trained staff for analysis and is

costly and burdensome. Sleep centers often have a limited capacity as well. Therefore, unobtrusive, cheap, and simple though reliable systems for monitoring at home are desired. These sensors could offer the ability to screen patients and prioritize them for hospital diagnostics, to increase healthcare accessibility or to enable long-term follow-up.

Among sleep disorders, obstructive sleep apnea (OSA) has the highest prevalence, from 13% to 33% in men and from 6% to 19% in women. However, these numbers are probably underestimated and are likely to grow as they are closely associated with obesity and advancing age [1]. OSA is characterized by events of breathing disturbance causing hypoxaemia, large chest motions, and arousals from sleep. These events fragment the patient's sleep and reduce phases of rapid eye movement and slow wave sleep. Consequently, OSA is an acknowledged risk factor for excessive daytime sleepiness, hypertension, and cardiovascular diseases [2]. The severity of sleep apnea is assessed by the Apnea–Hypopnea Index (AHI), which is the number of respiratory events (apneas and hypopneas) per hour. A patient is categorized as not suffering from sleep apnea ($0 \leq \text{AHI} < 5$), as having mild apnea ($5 \leq \text{AHI} < 15$), moderate apnea ($15 \leq \text{AHI} < 30$), or severe apnea ($30 \leq \text{AHI}$) [3].

In order to expand unobtrusive resources for home-based sleep apnea screening and sleep monitoring, a commercial off-the-shelf sensor was explored, the Emfit QS (referred to as Emfit, developed and manufactured by Emfit, Finland). The Emfit is a pressure sensor built from electromechanical film (EMFi), which is a polypropylene film including gas voids. The material is similar to piezoelectric materials as a displacement charge is produced when a force is being applied. However, the change of the internal electric field is caused by the movement of static charges that were injected during fabrication of the film [4]. From the pressure-modulated signal, a respiratory signal and ballistocardiography (BCG) signal can be derived. The latter is an unobtrusive measurement of the body's recoil caused by cardiovascular pulsation. As such, the sensor can provide information on sleep-disordered breathing as well as other origins of motion. A study by Koyama et al. [5], based on BCG, studied the feasibility of a piezoelectric sensor for apnea screening. They considered apneas during Cheyne–Stokes-like breathing to be correlated with AHI. This type of breathing is, however, only present in cardiac patients, thus targeting a subset of patients. Tenhunen et al. [6] evaluated a custom-made Emfit sheet and derived several parameters from breathing patterns to correlate these with AHI and assess sleep apnea severity. Despite the sensitivity of 0.95 in detecting subjects with AHI < 15 using a combined parameter, the method required annotators to score breathing patterns visually and made no contribution to the automatic detection of these patterns. The same authors derived heart rate variability (HRV) as well [7], which resembled known HRV results of sleep apnea patients during periodic apneic events. This revealed an increase in sympathetic activity and claimed a good reliability of detection of periodic sleep disordered breathing. However, periods with wakefulness, movements, and artefacts were manually omitted, which hinders the application of Emfit as a stand-alone device.

Currently, no fully automated sleep apnea screening method has been established based on the Emfit sensor. Moreover, no Emfit studies have been performed using the commercial off-the-shelf Emfit sensor, according to the knowledge of the authors of this study. Hence, the goal of the present study was twofold (see Figure 1). First, the potential of the Emfit sensor in a stand-alone setting for sleep apnea screening was investigated. Sleep apnea is characterized by breathing cessations, which are terminated by arousals often accompanied by large motions of the chest. These arousals and chest motions cause deviations in the signals, which were referred to as artefacts. Hence, the Emfit data was explored to reveal clusters of artefacts and clean segments in the signal. The characteristics of these clusters were linked to the AHI. This cluster analysis was performed unsupervised as the Emfit sensor was not automatically synchronized with the PSG and to avoid burdensome manual labeling of the data into clean and artefact segments. Secondly, the study explored the suitability of this commercial sensor to be merged with gold standard polysomnography data for future sleep monitoring. Therefore, an automated synchronization procedure based on the previously detected artefact patterns was developed, since PSG and Emfit signals originate from different types of sensor modalities and cannot be regarded as strictly coupled. After synchronization, two different positions

of the Emfit were investigated to find the optimal position for capturing respiratory and cardiac information similar to the PSG.

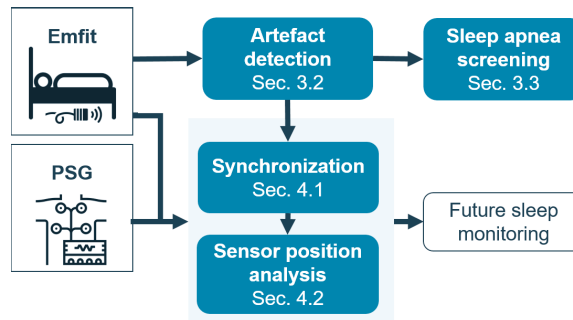


Figure 1. Overview of study objectives. First, the potential of the Emfit sensor for sleep apnea screening was investigated by searching for artefacts in the data caused by arousals and chest motions. Secondly, the study explored the suitability of this commercial sensor to be merged with gold standard polysomnography data for future sleep monitoring. Therefore, an automated synchronization procedure based on the previously detected artefact patterns was developed. After synchronization, the optimal position of the Emfit for capturing respiratory and cardiac information was identified.

2. Materials

The Emfit QS is a commercially available pressure sensor (542 mm × 70 mm × 1.4 mm). Both the raw data and prefiltered data was made available. The raw data was sampled at 100 Hz. The prefiltered data contained a bandpass filtered signal at [0.08, 3] Hz and a bandpass filtered signal at [6, 16] Hz to obtain the respiratory and BCG signals, respectively. Filtering techniques were not specified by the manufacturer. From the PSG system (B3IP, Medatec, Belgium) the thoracic belt and ECG signal were analyzed.

In this study, two setups of the sensor were investigated. The bed consisted of a mattress on top of which a mattress topper of approximately 4 cm thickness was added. One sensor was positioned underneath the thorax of the patient, separated by the mattress cover (position *Top*). A second sensor was placed beneath the topper (position *Bottom*) at a 2.5 cm horizontal distance to the top sensor (see Figure 2). The horizontal distance ensured the limiting of the influence of the top sensor and compensated the effect of patients moving down in the bed when lifting the head of the mattress upwards. This setup was applied simultaneously in two beds in the sleep laboratory.

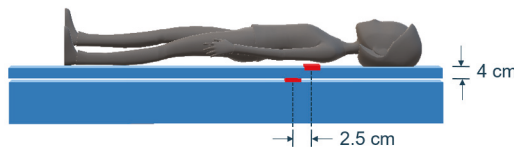


Figure 2. Setup of Emfit sensors. The bed consisted of a mattress with a mattress topper of approximately 4 cm thickness. One sensor was positioned underneath the thorax of the patient, separated by the mattress cover (position *Top*). A second sensor was placed beneath the topper (position *Bottom*) at a 2.5 cm horizontal distance to the top sensor.

The Emfit sensor and PSG simultaneously recorded data for patients referred for sleep diagnosis in the sleep laboratory of the University Hospitals Leuven (UZ Leuven). Overnight PSG signals were annotated by sleep specialists according to the American Academy of Sleep Medicine, 2012 scoring rules [8] to derive the AHI. The dataset was recorded in two phases with an interruption of 7.5 months.

The sensor setup remained the same; only the sensors were removed between phases and relocated as close as possible to the original location. Specifications of both datasets can be seen in Table 1. The last column, Top+Bottom, indicates the number of top sensor signals that have a corresponding bottom signal available. The reason for this was data loss due to technical problems, mostly with the bottom sensor.

Table 1. Datasets. The dataset was recorded in two phases with an interruption of 7.5 months. The sensor setup remained the same; only the sensors were removed between phases and relocated as close as possible to the original location. The datasets are characterized by age, Body Mass Index (BMI), Apnea-Hypopnea Index (AHI), male (M) or female (F) and number of available signals from the sensors at the top and/or bottom location.

	#Patients	Age (years)	BMI ($\frac{\text{kg}}{\text{m}^2}$)	AHI ($\frac{\text{events}}{\text{h}}$)	M/F	#Top	#Bottom	#Top + Bottom
Phase 1	31	49.7 ± 11.7	31.3 ± 7.9	30.4 ± 25.7	27/4	31	22	22
Phase 2	83	46.8 ± 12.7	31.6 ± 6.2	28.9 ± 26.3	50/33	83	33	29

All subjects gave their informed consent for inclusion before they participated in the study. The study was conducted in accordance with the Declaration of Helsinki, and the protocol with registration number B322201732928 was approved on November 8th 2018 by the UZ/KU Leuven Ethics Committee (Ethische Commissie Onderzoek UZ/KU Leuven).

3. Emfit-Based Sleep Apnea Screening

The Emfit sensor was evaluated in terms of its potential for sleep apnea screening in a stand-alone setting. As sleep apnea is characterized by breathing cessations that are often accompanied by large chest motions, these motions will induce deviations in the signal. These deviations will be referred to as artefacts, which, on the other hand, can also be induced by non-pathological body motions. It was hypothesized that the distortion of the data increased with AHI as more movement and arousals would be detected. Therefore, these artefacts were identified in the data by an unsupervised clustering method. First, the raw Emfit data was pre-processed. Thereafter, features were extracted that highlight irregularities in the signal. Features which optimally clustered the data were selected. Finally, the characteristics of the clustering were applied for sleep apnea screening.

3.1. Emfit Preprocessing

First, data quality was assessed by investigating the peak-to-peak amplitude (PP) distribution of the sensors after both measurement phases. Then, after subtraction of the mean value, the prefiltered respiratory signal of the Emfit sensor was further bandpass filtered to [0.08, 2] Hz. The respiratory signal was resampled at 4 Hz and the BCG signal at 50 Hz. As the signal amplitude was dependent on the weight and position of the patient, the signals were normalized. Normalization was based on the assumption that long-lasting periods of signal saturation corresponded to position changes by the patient. Segments between these periods were normalized by the median of the PP amplitude of this segment. If the median value was zero, the normalization of the previous segment was applied. This procedure was applied separately to the raw pressure, prefiltered respiratory, and prefiltered BCG signal. The periods of position changes and other saturated values were clipped to a value of 1, which was double the value of signals at the median amplitude.

Next, time–frequency domain information was extracted from the resulting signals by means of the discrete wavelet transform. To accentuate steep changes in the raw pressure signal indicating motion, a Daubechies 1 (i.e., db1 or Haar) wavelet was applied. Taking into account window size and sampling frequency, the signal was decomposed until level 8, i.e., [0.2, 0.4] Hz. The respiratory signal was approximated with a db4 wavelet (until level 3, [0.25, 0.5] Hz) and the BCG with db6 (until level 2, [6.25, 12.5] Hz). The respective wavelet shapes were chosen for their resemblance to the natural wave

shape. A total of 16 signals (original signals and decompositions) were used for the subsequent feature extraction step.

3.2. Artefact Detection

3.2.1. Feature Extraction

A feature window of 10 s was applied for sufficient time resolution and to include two to three breaths from the respiration signal. In total, 19 features were extracted in order to locate artefacts by inspecting outliers as well as irregularities (see Table 2). For features 9–19, the window was split into 3 equal subsegments over which PP was calculated, resulting in PP_3 [9].

Table 2. Features. Nineteen features were extracted in 10 s windows.

	Feature
1–3	Mean, Variance (Var), Standard Deviation (Std),
4–5	Kurtosis, Skewness
6–7	Kurtosis of Autocorrelation, Shannon Entropy
8	Peak-to-Peak Amplitude ($PP = \max(x) - \min(x)$)
	PP_3 : individual PP of 3 equal subsegments of window
9	Maximum (PP_3) / mean (PP_3)
10–11	Var (PP_3) (peakVar), Std (PP_3)
12–16	[10%, 25%, 50%, 75%, 90%] (PP_3)
17	Interquartile Range (PP_3)
18	Interdecile Range (PP_3)
19	Median Absolute Deviation (PP_3)

Time domain features were derived from both the untransformed signals and the three wavelet decomposed signals. These features were then normalized per subject using the z-score, and features with a Pearson correlation coefficient larger than 0.9 were removed. Lastly, feature values were transformed by means of the Euclidean norm normalization to decrease the effect of extreme values.

3.2.2. Unsupervised Feature Selection

The unsupervised feature selection framework was based on Robust Spectral learning (RSFS) [10] (see Figure 3). This method provides a ranking of features, depending on three parameters of the RSFS objective function, i.e., α , β , and γ . Input feature vectors were taken from a reduced training dataset selected using K-medoids clustering with $K = 2000$ and the Mahalanobis distance metric [11]. The K-medoids clustering was performed 100 times, such that the parameter optimization pipeline was run with 100 different training sets. Additionally, the Rényi entropy of every training set was calculated to verify the diversity within a training set and stability over training sets. Next, parameters α , β , and γ of the RSFS were taken from a 3D grid search over equispaced values in logarithmic scale from -3 to 3 . For every set α , β , and γ , a feature ranking was calculated and a number d of top-ranked features was selected. Subsequently, a k -means clustering in a d -dimensional space was performed 20 times using squared Euclidean distance and random initialization. The clustering performance was evaluated by the overall average silhouette score [12]. The pipeline was iterated for $d = [3, 5, 7]$ features and $k = 2$ clusters. After completion of these iterative steps, the pipeline optimized the parameters α , β , and γ , resulting in the feature ranking, as well as the optimal number of features d .

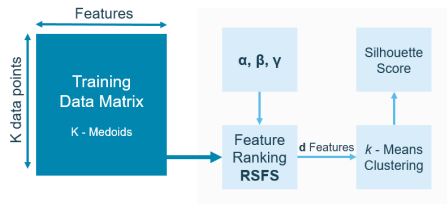


Figure 3. Pipeline for unsupervised feature selection. The input was a K-medoids clustering to reduce the dataset. This selection served as the input for unsupervised features selection. It was comprised of a parameter optimization that defined the feature ranking. The d top-ranked features were used for k -means clustering. The performance metric was the silhouette score. The pipeline was repeated for a different number of clusters k .

3.2.3. Clustering of Artefacts

With the optimized features, the training points were clustered using k -means with $k = 2$. From this clustered training set, the centroids of both clusters were identified. These centroids acted as target points for the test data to determine its associated cluster by mapping every test data point to the closest centroid. The characteristics of the clusters were analyzed based on their feature values and a pairwise Mann–Whitney U test. As the features were tailored to detect large deviations in the signal, it was assumed that one cluster contained clean and the other contaminated, or artefact, data segments.

3.3. Screening of Sleep Apnea

Artefacts present in the Emfit signal originated from different sources such as position changes and apneic arousals. It was hypothesized that more severe sleep apnea patients would have more artefacts present in their data compared to healthier subjects. Clustering of these artefacts was performed using k -means clustering. This method assumes globular data structures due to the use of the Voronoi diagram. However, artefacted segments exhibited a varying morphology, resulting in less globular clusters. Therefore, some artefacted segments might be assigned to the clean cluster. The cleanness of the clean segment cluster was inspected by taking into account the distances of segments in the clean cluster to the clean cluster centroid. Outlying values were discarded by only considering values below the 95th percentile of distances. This segment distance distribution was calculated for every subject. A larger 95th percentile would indicate larger distances within the clean cluster and thus more artefact-like segments, hence a larger AHI was expected for the subject.

Training of the cluster centroids was performed with the dataset from Phase 1 (see Table 1). The dataset from Phase 2 was applied for testing by mapping the data of individual subjects to the trained centroids and evaluating the cleanness of the cluster based on the 95th percentile.

4. Emfit Integration with Polysomnography for Sleep Monitoring

4.1. Artefact Pattern-Based Synchronization

The Emfit is a stand-alone device that was not connected to the PSG. Therefore, both sensors were not automatically synchronized. A synchronization of the Emfit with the PSG is necessary for further analysis of the Emfit signal in a supervised manner. Synchronization based on timestamps of both sensors was not sufficient as large delays were still present. Also, simultaneously tapping the mattress with built-in sensors and marking the PSG data with a synchronization button was not sufficient as it was difficult to discriminate normal movement behaviour in the Emfit data during wake from tapping. Therefore, an automated synchronization procedure was developed based on the signals' characteristics. To this end, the signal from the thoracic belt of the PSG was selected as a reference as its position was most proximate to the Emfit sensor. The Emfit respiratory signal and PSG respiratory effort signal, however, originate from different modalities. Therefore, a direct comparison of both

signals based on clean waveforms was not possible as wave shapes can differ. The synchronization was based on the observation that movement of the patient and large changes in ventilation due to apneic arousal were reflected in both the Emfit as well as in the PSG. For this reason, the synchronization made use of the occurrence and pattern of artefacts in the signals, which were derived in Section 3.2.

4.1.1. Polysomnography Preprocessing

The effect of movement caused by body posture changes was expected to be different in both modalities and more similar in the case of apneic breathing. Therefore, the central seven hours of sleep data were considered, as the patient was expected to be asleep. The PSG respiratory effort signal was bandpass filtered between [0.08–2] Hz using a Butterworth filter and downsampled from 500 Hz to 4 Hz. The data contained many small noisy peaks that were not necessarily present in both Emfit and PSG. To eliminate these, the top envelopes of the signals were derived using the secant method and a 1 s window.

4.1.2. Delay Detection

The Emfit and PSG signals could have large delays as well as a large variation in delay between patients. Moreover, synchronization becomes more difficult if a very high number of distortions are present, often observed in patients with a very large AHI, as shown in Figure 4. Therefore, the synchronization was performed in two steps: a coarse delay detection and a refined delay detection. The coarse delay detection step took into account large artefact patterns, thus a signal interval containing 18 artefact windows (Section 3.2.3) was defined in the Emfit respiration signal. This interval was compared with intervals of the PSG signal by correlation (see Figure 5). A large margin (35 min) was taken, as the initial delay between signals could be substantial. The shift for which the maximal cross-correlation occurred was defined as the delay for the considered Emfit artefact interval. After iteration over all artefact intervals, the final delay value for the coarse synchronization was selected as the maximum of the probability density estimation (PDE) of delays. The bandwidth of the kernel indicated the standard deviation of the PDE and thus the certainty of the estimated delay. After shifting the signal with the coarse delay, it is necessary to consider more confined artefact blocks and precisely locate these in the PSG signal. The refined delay detection meant meaning a reduction of the interval to six artefacts and the margin to 5 min.

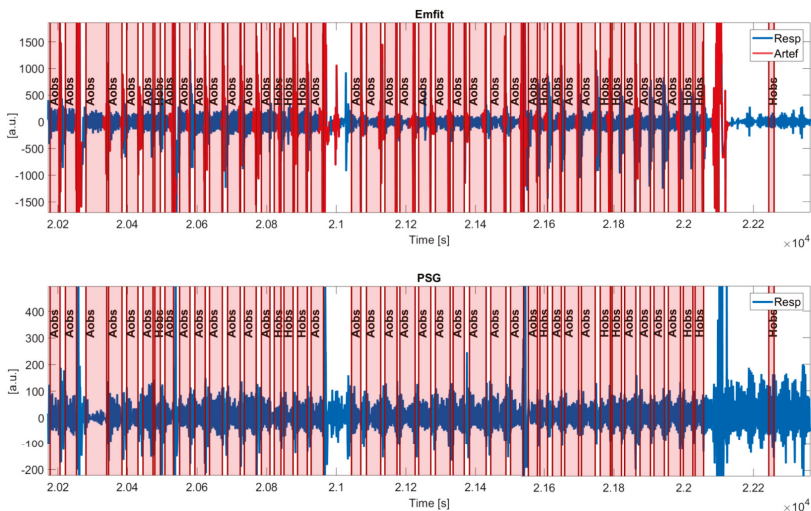


Figure 4. Signal of patient with large AHI. The signal contains consecutive apneic events, complicating synchronization.

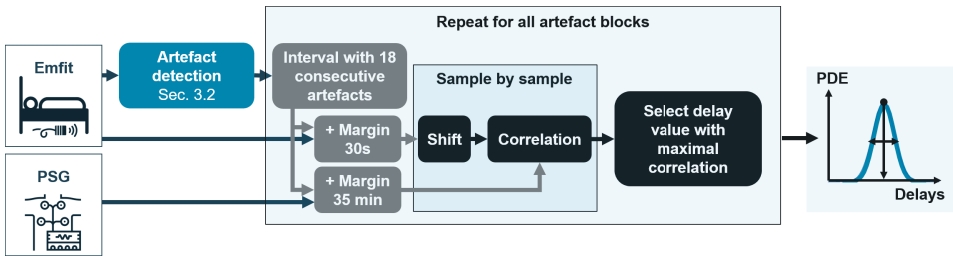


Figure 5. Procedure for coarse delay detection. The Emfit interval contained 18 artefact windows. The Emfit artefact block was shifted sample by sample along the PSG search interval. A probability density estimation (PDE) was derived over the series of optimal shifts.

4.2. Sensor Position Comparison

After synchronization of the Emfit with the PSG, the quality of the sensors was analyzed. The top sensor was expected to have a larger BCG signal quality, while the signal captured by the bottom sensor was attenuated by the mattress topper. The latter could lead to a better signal quality if many movement artefacts were present and for patients with an increased BMI. Without the attenuation, the signal would otherwise saturate. In a first phase, clean segments were extracted from the signals (see Figure 6). Based on the detected artefacts in the Emfit signal, segments of at least 1 min without an artefact were considered. These segments were compared to the corresponding segments from the PSG based on magnitude-squared (MS) coherence and correlation. Based on these statistics, the ability of the top and bottom sensors to capture heart rate and respiration information was assessed.

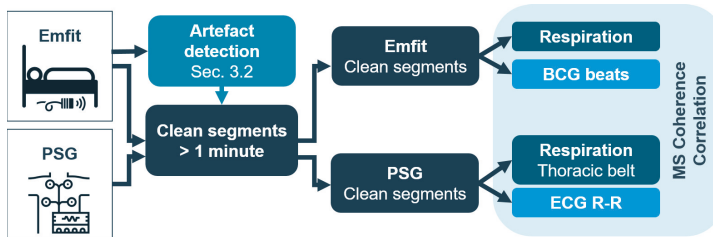


Figure 6. Procedure for sensor position comparison.

4.2.1. Tachogram Derivation from ECG and BCG

The comparison between the BCG and ECG was based on heart rate information. Therefore, the tachograms of both signals and their evenly sampled interpolation was derived. First, the ECG signal was cleaned and saturated segments were not considered. Next, the R-peaks were detected using the algorithm proposed in [13]. Beats of the BCG signal were detected by an adapted Pan–Tompkins algorithm described in [14]. The tachograms of both sensors were analyzed for outliers by an adaptive threshold. It was defined as the running standard deviation of the 20 most recent samples multiplied by a factor of 5. Thereafter, the tachograms were interpolated and resampled to 4 Hz.

4.2.2. Similarity Measures

Similarity was calculated between [0.1, 0.4] Hz for the respiratory signals of Emfit and PSG and between dynamic intervals for the interpolated tachograms of the BCG and ECG. For the latter, the maximum peak of the power spectral density of the ECG-derived tachogram in the low frequency (LF) band [0.03, 0.15] Hz and high frequency (HF) band [0.15, 0.4] Hz was determined. The frequency ranges covering the width at half maximum were considered. Additionally, the normalized cross-correlation was calculated between the HRV signals over lags in the interval [−15, 15] s.

In three cases, the clean segment was labeled as a segment containing no information, and no parameters were calculated: First, if the duration of one of the tachograms was smaller than 30 s; second, if the segment contained less than three detected heart beats; and finally, if the cross-correlation value was less than zero, as this indicates an erroneous tachogram of the BCG resulting from inferior data quality. The total length of clean segments over the total signal length was compared for the top and bottom sensors.

Since subjects have an unequal number of clean segments, some have a larger weight in the comparison as more of their segments are included. Therefore, a paired analysis was carried out as well. From every subject, the median values of the top and bottom parameter distribution were extracted and evaluated by a Wilcoxon signed rank test. The complete parameter distributions for coherence and correlation were compared for individual subjects as well. It was evaluated whether the top or bottom performed significantly better and whether there was a relation with the BMI of the patients.

5. Results

5.1. Emfit Data Usability Assessment

Generally, the amplitude of the top sensors was higher compared to the bottom sensors.

The top sensors had similar median PP amplitudes in both beds during Phase 1 as well as during Phase 2. When comparing both phases, the top sensors of Phase 1 had a higher median PP amplitude compared to Phase 2. The manufacturer claimed to not have performed upgrades, which could have affected the recordings. Alteration in amplitudes could be explained by slight changes in location when reinserting the sensors between two phases.

A similarity in the amplitude of bottom sensors in both beds was also observed during Phase 1. However, in Phase 2 the distribution of median PP amplitudes of bottom sensor 1 was significantly different, with a median of only 21% compared to bottom sensor 2. Bottom sensor 1 might have shifted location during recordings in Phase 2 and was left out of the analysis.

5.2. Unsupervised Feature Selection and Clustering

The pipeline was executed for $d = [3, 5, 7]$ and a cluster number $k = [2, 3, 4, 5]$ and repeated for 100 different training sets. The resulting silhouette score distribution is displayed in Figure 7 (borders indicating the 25th and 75th percentiles). It can be seen that a limited number of features as well as a lower number of clusters resulted in higher silhouette scores. The decrease of the average silhouette score with a higher number of clusters $k > 2$ suggested that the natural existing clusters might be split into multiple ones. Based on these results, the analysis was continued with feature number $d = 3$ and cluster number $k = 2$.

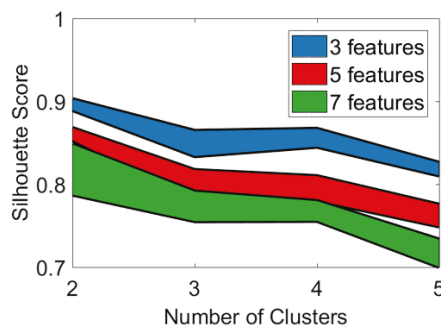
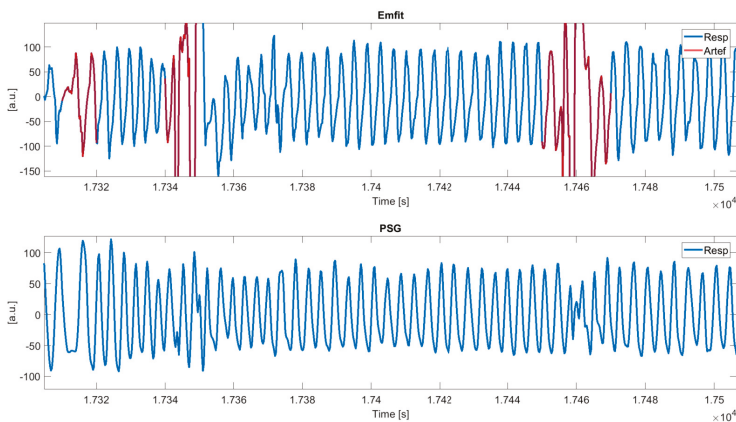


Figure 7. Silhouette score distribution. Borders indicate the 25th and 75th percentile of 100 iterations.

Optimal parameter sets $\{\alpha, \beta, \text{ and } \gamma\}$ varied slightly, hence feature ranking and the resulting silhouette score varied as well over K-medoids iterations. Within 100 iterations, 2 optimal feature subsets were put forward, each with a 15% occurrence. The feature subset resulting in the highest average silhouette score was finally selected, being the features *pressure peakVar* at wavelet decomposition level 2, 3, and 4 (see Table 2).

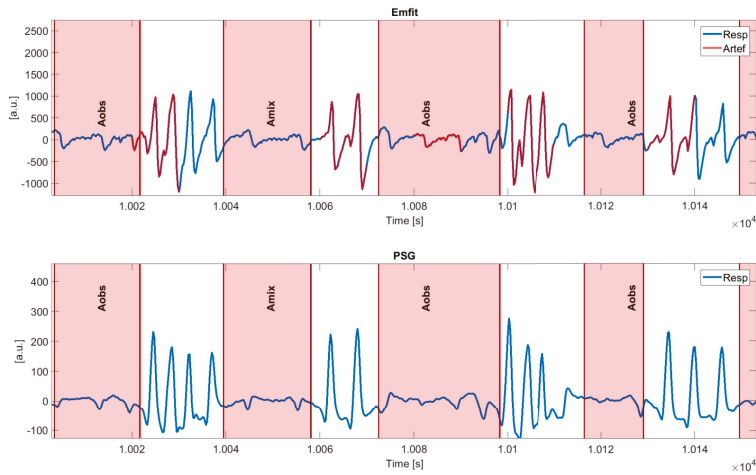
Evaluation of Rényi entropy values (mean of 1.41, standard deviation of 0.040) indicated a limited variability (see Section 3.2.2). Therefore, a random training set was chosen and clustered with the optimized features. This resulted in an overall average silhouette score of both clusters of 0.91. Cluster 1 contained training samples with a highly varying silhouette score. In contrast, cluster 2 was a very well-defined cluster and should contain samples with similar characteristics. One cluster containing higher values of features, corresponding to higher peak variations, was labeled as *artefact* cluster. The other cluster characterized stable segments without intermittent peaks and was labeled as *clean* cluster. The difference in data distribution between clusters was high (Mann–Whitney U test, $p < 0.001$), indicating that parameters were optimized to make a distinction between artefact and clean data. Mapping the test data to the trained centroids resulted in an overall average silhouette score for both clusters of 0.95.

Detailed examples of an Emfit signal with detected artefacts and (synchronized) PSG thoracic belt are displayed in Figure 8. Shaded intervals present apneic events and detected artefacts are indicated in red. Figure 8a illustrates that during normal breathing, both signals oscillate at the same frequency, although the Emfit signal is more heavily distorted during vibrations. Figure 8b shows artefacted segments following obstructive apneas (Aobs), suggesting the ability of the algorithm to capture apneic arousals and corresponding motions. Furthermore, Figure 8c displays artefact segments around 9450 s, which are not related to an apneic event and can be assigned to generic body movements. However, during [9500, 9700] s, obstructive hypopneas (Hobs) took place after which no artefacts were detected (except one). In this example, the reduction in ventilation is hardly captured in the Emfit signal.

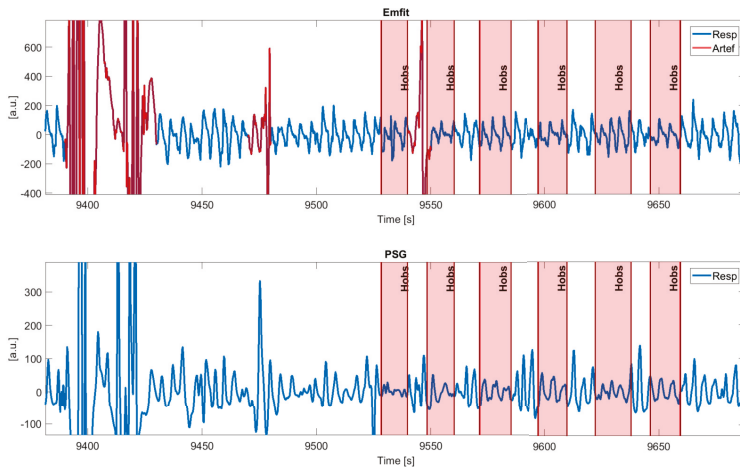


(a) During normal breathing both signals oscillate at the same frequency, although the Emfit signal is more heavily distorted during vibrations.

Figure 8. Cont.



(b) Artefactual segments follow obstructive apneas (Aobs), suggesting the ability of the algorithm to capture apneic arousals and corresponding motions.



(c) Artefact segments around 9450 s are not related to an apneic event and can be assigned to generic body movements. However, during [9500, 9700] s, obstructive hypopneas (Hobs) took place after which no artefacts were detected (except one). In this example, the reduction in ventilation is hardly captured in the Emfit signal.

Figure 8. Details of the synchronized Emfit and PSG signal with detected artefacts. Segments are shown during normal breathing, obstructive apneas (Aobs), and obstructive hypopneas (Hobs).

5.3. Screening of Sleep Apnea

As explained in Section 3.3, the cleanness of the clean segment cluster was inspected for every subject. For this, the 95th percentile of distance to the clean cluster centroid was derived. A linear regression of this metric with AHI is depicted in Figure 9. The regression displayed an upward trend; however, only a limited coefficient of determination R^2 of 0.16 was obtained.

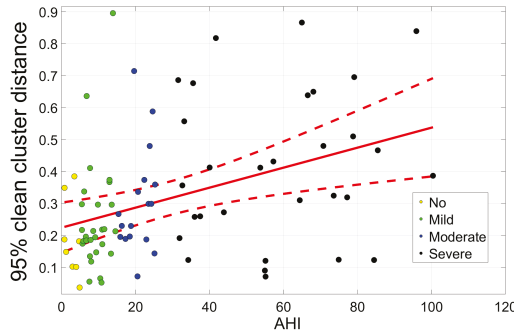


Figure 9. Linear regression of the 95th percentile of distance to the clean cluster centroid with AHI. The dashed lined is the 95% confidence interval with an R^2 value of 0.16.

The distance metric was also analyzed for standard sleep apnea classes of subjects, as shown in Figure 10, where the area under the curve (AUC) is defined of the receiver operating characteristic (ROC) curve. This also indicated a trend towards larger distances within the clean cluster and hence less regularity in the signal with increasing AHI. A Kruskal–Wallis test with Bonferroni correction between apnea classes indicated a significant difference ($p < 0.05$) between no and mild apnea versus severe apnea. Furthermore, a significant difference (Mann–Whitney U test, $p < 0.05$) was found between patients with $AHI < 15$ and $15 \leq AHI$. The ROC curve in Figure 10c displays the ability of screening of severe apnea patients ($AHI \geq 30$), where a sensitivity of 0.77 and specificity of 0.62 was reached. The ROC curve for more generally defined apnea patients ($AHI \geq 15$) reaches a sensitivity of 0.72 and specificity of 0.70. As a screening measure, a value of 0.229 for the 95th percentile of distance to the clean cluster centroid was taken.

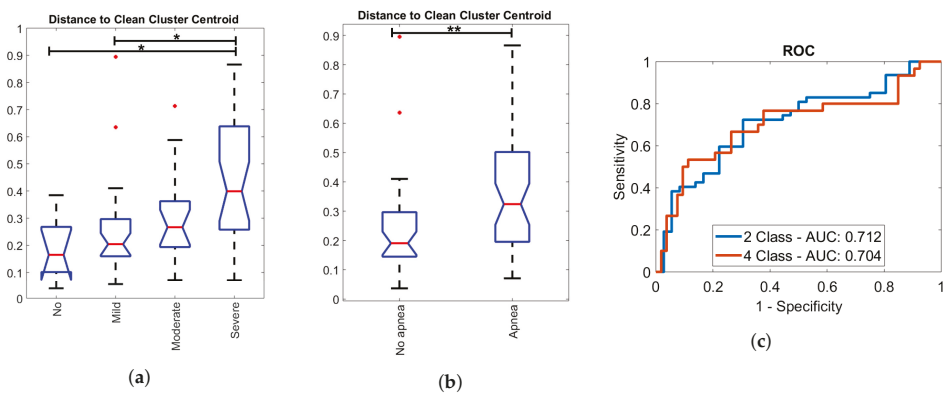


Figure 10. Screening of sleep apnea patients. The cleanness of the clean segment cluster was inspected for every subject by derivation of the 95th percentile of distance to the clean cluster centroid. These values were grouped according to the AHI of subjects. (a,b) A significant difference (Kruskal–Wallis test with Bonferroni correction, $p < 0.05$) was established between no and mild apnea versus severe apnea, as well as between patients with $AHI < 15$ and $15 \leq AHI$ (Mann–Whitney U test, $p < 0.05$). (c) The ROC curves display the ability for screening of severe apnea patients ($AHI \geq 30$) and more generally defined apnea patients ($AHI \geq 15$).

Since the resulting feature set consisted of relatively simple and similar features (see Section 5.2), the screening performance was compared to a threshold-based method as well. After normalization

of the data (see Section 3.1) and slicing into 10 s intervals, a window contained an artefact if any value exceeded the threshold. As such, the data of every patient was associated with a percentage of artefacts. Based on the artefact percentages and AHI of patients in the training data (Phase 1 of Table 1), an ROC analysis was performed. By analyzing the change in AUC with the selected signal amplitude threshold, an optimal threshold value was defined at 80% of the maximal amplitude. As such, a similar performance could be reached when screening patients from the test data (Phase 2). Here, the threshold was trained using the AHI labels. If the AHI is not available for training and an empirical threshold is taken at 50%, the results are close to random.

5.4. Artefact Pattern-Based Synchronization

The calculated delays of top and bottom Emfit sensors had a median value over all night recordings of $46.3 \text{ s} \pm 21.9 \text{ s}$. The accuracy of synchronization was verified by the bandwidth of the signal's delay distribution. Fifty percent of the data had a bandwidth value below 3.68, 75% below 7.90 and upper adjacent of 14.26. Signals with a delay distribution bandwidth above 7.90 were visually checked. Empirically, bandwidths between 7.90 and 14.26 resulted in a synchronization error lower than or equal to 10 s. The error margin of 10 s was considered manageable as this can be compensated by a correlation based on ECG. This procedure, explained in Section 4.2.2, searches over an interval of $[-15, 15] \text{ s}$ for the highest correlation. Bandwidths above 14.26 exhibited a varying range of synchronization errors, which comprised 13.7% of the data. Six subjects had to be removed from further analysis as the actual delay after synchronization was still more than 15 s.

5.5. Sensor Positioning Comparison

The parameters proposed in Section 4.2 were derived for all signals recorded by the top and bottom sensors (see Figure 11). Parameter distributions were similar for the top and bottom sensors; however, the median value of the top sensor was significantly higher. On an individual basis, in which the median value of distributions was taken into account, similar results were observed. The coherence parameters were significantly better for the top sensor, with $p < 0.05$ and a correlation with $p < 0.001$. On the other hand, the bottom sensor contained more clean segments ($p < 0.001$). Concerning the influence of BMI on the optimal sensor position, no correlation could be found between these measures. Furthermore, the shift during ECG–BCG correlation analysis was taken into account. The median optimal shift over all signals was -0.15 s with a bandwidth of 0.25.

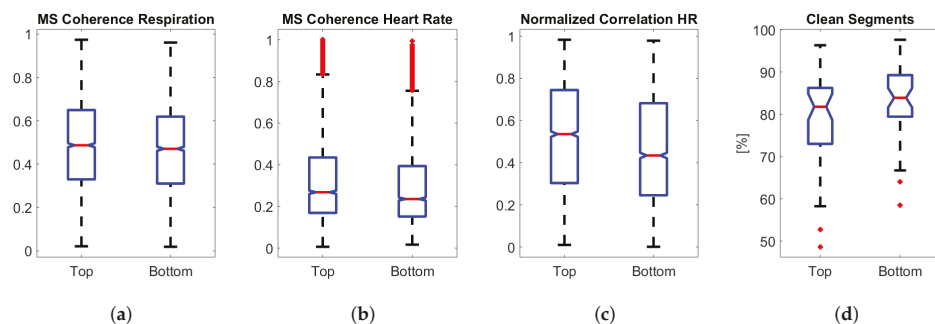


Figure 11. Parameter comparison of top and bottom sensors over the whole population. (a) Magnitude-squared coherence between Emfit and PSG respiration signals. (b) Magnitude-squared coherence between heart rate derived from BCG and ECG. (c) Cross-correlation between heart rate derived from BCG and ECG. (d) Percentage of clean segments that could be analyzed.

6. Discussion

The approach presented here demonstrated the potential for unobtrusive home-monitoring screening of patients at risk of sleep apnea with an off-the-shelf sensor intended for a home environment. Patients in which a large amount of artefacts are detected, due to position changes or apneic arousals, are considered as being at higher risk of suffering from sleep apnea. A trend was seen in the irregularity of the data with AHI (see Figure 10a), although the linear relation was limited (R^2 of 0.16). Moreover, a distinction was made between patients suffering from sleep apnea ($15 \leq \text{AHI}$) and patients considered healthy (see Figure 10b). A significant difference existed between both classes, which is a beneficial result for screening purposes. Doctors are most interested in the identification of these patients as they should be referred for further research in a sleep clinic and ideally prioritized on the waiting lists. The screening with ROC analysis resulted in a sensitivity of 0.72, specificity of 0.70, and diagnostic odds ratio ($\text{DOR} = \frac{\text{sensitivity} \times \text{specificity}}{(1 - \text{sensitivity}) \times (1 - \text{specificity})}$) of 6.00. Investigation of misclassification revealed a trend in the BMI towards higher values for false negatives and false positives, which can be attributed to saturation of the Emfit pressure signal with heavy weight. As patients with $35 \text{ kg/m}^2 \leq \text{BMI}$ are known to have an increased risk for sleep apnea, these were removed from the screening analysis. This increased the DOR of the EMFIT screening method for $15 \leq \text{AHI}$ to 8.96. Additionally, different body positions can have an influence on the signal and resulting misclassification, such as lying higher, lower, or sideways.

A similar screening procedure was performed in [15], in which a larger sensitivity (80%) and specificity (87%) for severe sleep apnea screening were obtained. The study was based on the dataset of Phase 1 but using a leave-one-subject-out approach for testing. In the current study, a separate test set (Phase 2) was applied for screening. The sensors of the test set were slightly relocated compared to the training set. This relocation could have changed the properties of the artefacts and of the signal itself, thereby deteriorating the results. Therefore, pre-processing was improved by a normalization of the input data as well as the interpretation of the clustering results. A more gradual increase in irregularity of the data with AHI was observed in this study, complicating the screening of specifically severe sleep apnea patients ($30 \leq \text{AHI}$).

In clinical practice, screening questionnaires for OSA are readily available. Chiu et al. [16] compared the screening performance of commonly used questionnaires such as the STOP-BANG questionnaire (SBQ), which was found to be a superior tool for detecting mild, moderate, and severe OSA. However, its sensitivity is high at the expense of low specificity ($15 \leq \text{AHI}$: sensitivity of 0.90, specificity of 0.36, and DOR of 5.05), and its DOR is inferior compared to the current Emfit-based method. Nonetheless, taking into account the different ratios of sensitivity and specificity for both screening methods, these could be applied simultaneously to reinforce each other. Nevertheless, as a screening sensitivity of 0.95 and specificity of 0.92 based on manual annotation of Emfit signals was reached by Tenhunen et al. [6], improvement in automated methods is possible.

On this matter, clustering of data in clean and artefact segments was performed using k -means clustering, which is a method assuming globular data structures. However, artefacted segments exhibited a varying morphology resulting in less globular clusters, causing artefacted segments to be assigned to the clean cluster. A more complex clustering algorithm such as kernel spectral clustering [17] may be able to capture the varying morphologies of artefacts in multiple clusters. On the other hand, the simplified threshold method for screening performed similarly to the unsupervised clustering-based method. However, to establish an optimized threshold, the AHI of patients is required. In contrast, the clustering method is purely data-driven and is trainable without prior knowledge. Furthermore, its application can be extended to capture different types of irregularities in the data.

In order to establish an integration of the Emfit sensor with the PSG, an automated synchronization approach was developed. Segments in Figure 8 show that wave shapes in both modalities are different. As such, signals cannot be compared as a whole based on cross-correlations and the procedure focused on detecting large artefact patterns first with a coarse synchronization step. In patients with very a high AHI, synchronization becomes more difficult as signal deviations are almost continuously present.

The synchronization approach was automated by the introduction of a performance indicator, namely the bandwidth of the delay distribution. A threshold of a bandwidth = 14.26 could be defined to ensure sufficient synchronization accuracy. Moreover, most of the data (86.3%) attained a value below the threshold. However, some signals exhibited a delay distribution bandwidth above 15 while synchronization was accurate enough. A reason was that some patients leave the bed overnight. Electrodes are detached and only noise is recorded, causing the synchronization between both sensors to be distorted. The optimal shift before and after detachment is different, causing the bandwidth of the shift distribution to increase. Leaving the bed is a typical event, hence future work for Emfit-PSG integration should include the detection of electrode detachment and separate synchronization on different segments of the night. Concerning other recordings, the delay was fixed over the night. The difference in delay among recordings was suspected in instabilities during recording of the Emfit data, transmission over the hospital's wifi network, or uploading to the Emfit server. Furthermore, synchronization in the signals of patients with a very large AHI (AHI > 90) was more difficult as artefactual segments were more similar due to almost continuous apneic events (see Figure 4). Different delays result in similar cross-correlation values. Additionally, signal quality tends to decrease, which causes the correlation value during synchronization to drop.

In a second stage, the sensor signals were precisely synchronized based on heart rate information instead of the respiratory signal. As the calculated delay between the tachograms of the ECG and BCG was small, a good synchronization was already reached during respiration-based synchronization. The presented framework for synchronization enabled a supervised analysis of the commercial Emfit sensor for future studies. Additionally, the framework can be applied to other multi-modal systems that record movements during sleep. This includes pressure-based signals of the thorax and respiratory-related signals, as simultaneous and similar artefacts can be expected in these signals.

Regarding the positioning of the Emfit sensor, it can be seen in Figure 11a–c that performance parameters exhibit similar distributions for the top and bottom. Parameters were only calculated for clean segments, therefore the percentage of (clean) segments included for analysis from every sensor was visualized in Figure 11d. From the bottom sensor, more clean segments of at least 1 min could be extracted as these signals were attenuated by the mattress topper and fewer artefacts were present in the signal. On the other hand, median values were significantly higher for the top sensor, indicating better sensor correspondence with the hospital's PSG. This is due to the fact that the recorded signal amplitude of the bottom sensor was lower, making it more difficult for the algorithm to detect heart beats in the BCG. In general, MS coherence and correlation values of Emfit compared to PSG were modest. The Emfit sensor has a different measuring mechanism to the PSG thoracic belt or the PSG ECG. Therefore, different frequency components can be expected in Emfit respiration signals compared to the PSG thoracic belt. Moreover, the sensor quality of Emfit is expected to be less consistent during the night due to the different body positions of the patient.

7. Conclusions

A commercial pressure sensor was explored in terms of its potential for sleep apnea screening. An unsupervised algorithmic pipeline based on clustering was developed to characterize artefacts. A parameter based on the cleanness of these clusters was extracted as an indicator for sleep apnea severity. To enable a supervised analysis of the sensor for sleep monitoring, an automated synchronization procedure was developed based on the occurrence of artefacts in the respiratory signal. The synchronization framework can be applied to other multi-modal systems that record movements during sleep. This includes pressure-based signals of the thorax and respiratory-related signals, as simultaneous and similar artefacts can be expected in these signals. Furthermore, two different Emfit setups were analyzed for optimal signal quality. Locating the sensor as close as possible to the thorax and placing the sensor on top of the mattress is preferred if both respiratory and cardiac information are required. However, the positioning of the sensor is less critical if only respiratory

information is required. Depending on the application, the signal-attenuating effect of a mattress topper could be advantageous.

Author Contributions: Conceptualization, D.H. and C.V.; Data curation, D.H., P.B., D.T., and B.B.; Formal analysis, D.H.; Funding acquisition, S.V.H. and C.V.; Investigation, D.H.; Methodology, D.H. and C.V.; Project administration, S.V.H. and C.V.; Resources, P.B., D.T., B.B., and T.W.; Software, D.H.; Supervision, S.V.H. and C.V.; Validation, D.H.; Visualization, D.H.; Writing—original draft, D.H.; Writing—review & editing, D.H., P.B., D.T., B.B., T.W., S.V.H., and C.V.

Funding: Agentschap Innoveren en Ondernemen (VLAIO): 150466: OSA+; Agentschap voor Innovatie door Wetenschap en Technologie (IWT): O&O HBC 2016 0184 eWatch; imec funds 2017; European Research Council: The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007–2013)/ERC Advanced Grant: BIOTENSORS (nr 339804). This paper reflects only the authors' views and the Union is not liable for any use that may be made of the contained information. Carolina Varon is a postdoctoral fellow of the Research Foundation-Flanders (FWO).

Conflicts of Interest: The authors declare no conflict of interest. The funding sponsors 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. Senaratna, C.V.; Perret, J.L.; Lodge, C.J.; Lowe, A.J.; Campbell, B.E.; Matheson, M.C.; Hamilton, G.S.; Dharmage, S.C. Prevalence of obstructive sleep apnea in the general population: A systematic review. *Sleep Med. Rev.* **2017**, *34*, 70–81.
2. Young, T.; Peppard, P.E.; Gottlieb, D.J. Epidemiology of obstructive sleep apnea: A population health perspective. *Am. J. Respir. Crit. Care Med.* **2002**, *165*, 1217–1239.
3. Kapur, V.K.; Auckley, D.H.; Chowdhuri, S.; Kuhlmann, D.C.; Mehra, R.; Harrod, C.G. Clinical practice guideline for diagnostic testing for adult obstructive sleep apnea: An American Academy of Sleep Medicine clinical practice guideline. *J. Clin. Sleep Med.* **2017**, *13*, 479–504.
4. Paaanen, M.; Leikkala, J.; Kirjavainen, K. ElectroMechanical Film (EMFi)—A new multipurpose electret material. *Sens. Actuators A Phys.* **2000**, *84*, 95–102.
5. Koyama, T.; Sato, S.; Kanbayashi, T.; Kondo, H.; Watanabe, H.; Nishino, S.; Shimizu, T.; Ito, H.; Ono, K. Apnea during Cheyne-Stokes-like breathing detected by a piezoelectric sensor for screening of sleep disordered breathing. *Sleep Biol. Rhythm.* **2015**, *13*, 57–67.
6. Tenhunen, M.; Elomaa, E.; Sistonen, H.; Rauhala, E.; Himanen, S.L. Emfit movement sensor in evaluating nocturnal breathing. *Respir. Physiol. Neurobiol.* **2013**, *187*, 2, 183–9.
7. Tenhunen, M.; Hyttinen, J.; Lipponen, J.A.; Virkkala, J.; Kuusimäki, S.; Tarvainen, M.P.; Karjalainen, P.A.; Himanen, S.L. Heart rate variability evaluation of Emfit sleep mattress breathing categories in NREM sleep. *Clin. Neurophysiol.* **2015**, *126*, 967–74.
8. Berry, R.; Budhiraja, R.; Gottlieb, D.; Gozal, D.; Iber, C.; Kapur, V.; Marcus, C.; Mehra, R.; Parthasarathy, S.; Quan, S.; et al. Rules for scoring respiratory events in sleep: Update of the 2007 AASM manual for the scoring of sleep and associated events. *J. Clin. Sleep Med.* **2012**, *8*, 597–619.
9. Bruser, C.; Diesel, J.; Zink, M.D.H.; Winter, S.; Schauerte, P.; Leonhardt, S. Automatic Detection of Atrial Fibrillation in Cardiac Vibration Signals. *IEEE J. Biomed. Health Inform.* **2013**, *17*, 162–171.
10. Shi, L.; Du, L.; Shen, Y.D. Robust Spectral Learning for Unsupervised Feature Selection. In Proceedings of the 2014 IEEE International Conference on Data Mining, Shenzhen, China, 14–17 December 2014; pp. 977–982.
11. Varon, C.; Alzate, C.; Suykens, J. Noise Level Estimation for Model Selection in Kernel PCA Denoising. *IEEE Trans. Neural Netw. Learn. Syst.* **2015**, *26*, 2650–2663.
12. Rousseeuw, P. Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65.
13. Varon, C.; Caicedo, A.; Testelmans, D.; Buyse, B.; Van Huffel, S. A Novel Algorithm for the Automatic Detection of Sleep Apnea From Single-Lead ECG. *IEEE Trans. Biomed. Eng.* **2015**, *62*, 2269–2278.
14. Willemen, T. Biomechanics Based Analysis of Sleep. Ph.D. Thesis, KU Leuven, Leuven, Belgium, 2015.
15. Huysmans, D.; Buyse, B.; Testelmans, D.; Van Huffel, S.; Varon, C. Unsupervised Artefact Detection and Screening Using Emfit Sensor in Patients with Sleep Apnea. In Proceedings of the 45th Annual Computing in Cardiology Conference, Maastricht, The Netherlands, 23–26 September 2018; pp. 1–4.

16. Chiu, H.Y.; Chen, P.Y.; Chuang, L.P.; Chen, N.H.; Tu, Y.K.; Hsieh, Y.J.; Wang, Y.C.; Guilleminault, C. Diagnostic accuracy of the Berlin questionnaire, STOP-BANG, STOP, and Epworth sleepiness scale in detecting obstructive sleep apnea: A bivariate meta-analysis. *Sleep Med. Rev.* **2017**, *36*, 57–70.
17. Suykens, J.; Alzate, C.; Leuven, K.U. Kernel Spectral Clustering: Model Representations, Sparsity and Out-of-Sample Extensions. In Proceedings of the 4th International Conference on Computational Harmonic Analysis, Hong Kong, China, 23–27 May 2011.



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).

Article

A Glove-Based Form Factor for Collecting Joint Acoustic Emissions: Design and Validation

Nicholas B. Bolus ¹, Hyeon Ki Jeong ², Daniel C. Whittingslow ^{3,4} and Omer T. Inan ^{2,3,*}

¹ Bioengineering Graduate Program, Georgia Institute of Technology, Atlanta, GA 30332, USA; nbolus@gatech.edu

² School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30313, USA; hjeong39@gatech.edu

³ Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA 30332, USA; d.c.whittingslow@emory.edu

⁴ School of Medicine, Emory University, Atlanta, GA 30322, USA

* Correspondence: inan@gatech.edu

Received: 13 May 2019; Accepted: 10 June 2019; Published: 13 June 2019

Abstract: Sounds produced by the articulation of joints have been shown to contain information characteristic of underlying joint health, morphology, and loading. In this work, we explore the use of a novel form factor for non-invasively acquiring acoustic/vibrational signals from the knee joint: an instrumented glove with a fingertip-mounted accelerometer. We validated the glove-based approach by comparing it to conventional mounting techniques (tape and foam microphone pads) in an experimental framework previously shown to reliably alter healthy knee joint sounds (vertical leg press). Measurements from healthy subjects (N = 11) in this proof-of-concept study demonstrated a highly consistent, monotonic, and significant ($p < 0.01$) increase in low-frequency signal root-mean-squared (RMS) amplitude—a straightforward metric relating to joint grinding loudness—with increasing vertical load across all three techniques. This finding suggests that a glove-based approach is a suitable alternative for collecting joint sounds that eliminates the need for consumables like tape and the interface noise associated with them.

Keywords: acoustic emissions; joint sounds; glove; wearable sensing; knee joint loading

1. Introduction

Injuries and chronic disorders affecting joints are pervasive and degrade quality of life for millions of individuals [1,2]. The knee joint, due to its anatomical complexity, role in weight bearing, and high, cyclical exposure to mechanical stress, is particularly susceptible to injury [3]. The current diagnostic standard for acute joint injury and chronic conditions such as osteoarthritis involves a combination of medical imaging, which can be costly and time-intensive, and physical examination, which often relies on subjective evaluations made on the part of either the clinician or the patient. Moreover, these methods are not ideally suited to longitudinal, comprehensive monitoring of joint health, which may benefit recovery.

Accordingly, recent work has demonstrated the viability of using the acoustic emissions produced by joints in motion—in particular, the knee—as an indicator of underlying joint health. McCoy et al. referred to the concept of sensing skin vibrations (i.e., their local accelerations) caused by joint articulation as “vibroarthrography” [4]. These vibrations produce an acoustic response in the surrounding media, which is why the signal is often termed a “joint sound” or “acoustic emission.” Arthro-acoustic techniques have been explored in both clinical [5–7] and ambulatory settings [8], using both benchtop and wearable equipment [9]. Results from these studies have demonstrated an ability to discriminate reliably between the acoustic signatures of healthy and impaired joints [6,10], and those

of joints under varying mechanical load [11]. The latter study validated the use of a vertical leg press as a reliable paradigm for modifying the acoustic output of a healthy knee, demonstrating a change in the heterogeneity of the joint sound as a function of percent body weight applied.

More recently, our group has begun to explore the use of alternative form factors for collecting joint sounds that would improve the quality and reliability of the measurements and eliminate the need for consumables like tape and adhesive microphone pads, which are the conventional means of mounting acoustic sensors on the skin. Drawing inspiration from manual auscultation, we designed a system in which contact microphones are embedded in a glove and placed at locations of interest around a joint to collect arthro-acoustic data. This approach offers several advantages, including the ability to finely regulate contact pressure at the sensor-to-skin interface (by leveraging the user's inherent motor control and tactile feedback mechanisms) while removing interface noise caused by adhesive, fabric, or other material interacting with the skin. Additionally, an adhesive-based solution is not ideally suited to applications involving repeated use, such as longitudinal tracking in a home setting. Conversely, a hand-worn system such as the one proposed in this work could be easily and repeatedly administered, and, furthermore, would provide an opportunity for an individual to actively engage in the management of one's own or a dependent's care—for example, a parent might use the glove to collect joint acoustic data on a child suffering from juvenile idiopathic arthritis.

In this study, we employed the healthy subjects vertical leg press paradigm (Figure 1) as a means to validate the glove-based approach, alongside two more conventional mounting techniques: fabric tape and adhesive foam mic pads. Achieving a similar result in terms of a quantity that reflects the internal state of the knee joint—i.e., the loudness of grinding—across these techniques would suggest that a glove-based system can provide clinical value comparable to more established techniques without the need for consumables. Multi-day repeatability testing was also conducted to assess the reliability of results derived from the glove-based system, as well as their agreement with results derived from conventional techniques.

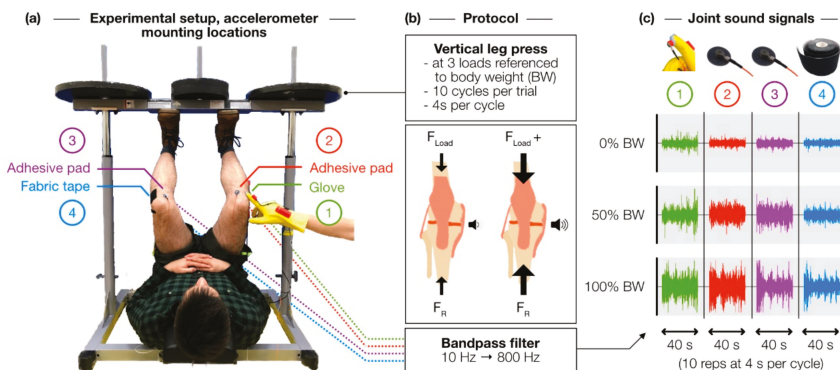


Figure 1. Experiment overview. (a) Four accelerometers were placed at regions of interest around the knee joint using different mounting techniques, and, in parallel, recorded vibrations produced by the joint during a vertical leg press exercise. (b) Increasing normal forces within the joint, we hypothesized, would increase the loudness of low-frequency grinding sounds within the knee. (c) Representative joint sound waveforms demonstrate how the amplitude of low-frequency vibrations increased as a function of percent body weight applied.

2. Methods

2.1. Design of a Glove-Based Form Factor

Our glove-based arthro-acoustic sensing system consists of (1) a glove to which various sensing and data acquisition components are mounted, (2) one or more fingertip modules in which the contact accelerometer and force sensor are integrated, and (3) a microcontroller for collecting data and driving feedback mechanisms for fingertip force regulation (Figure 2).

We used a latex/neoprene cleaning glove (Playtex, Dover, DE, USA), because it is easy to disinfect and because its elasticity enables a solid, contoured fit to the user's digits. Good coupling between the glove/sensors and the hand is critical for maintaining stable contact at the user–subject interface to minimize motion artifacts.

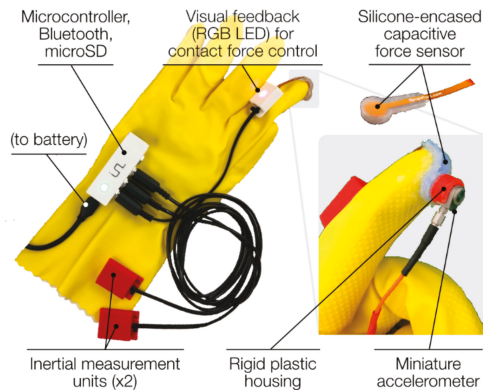


Figure 2. Design of a glove with embedded sensors for capturing joint sounds (accelerometer) and other contextual signals (inertial measurement units for limb motion, capacitive force sensor for sensor–skin contact pressure).

Sensing of the joint sound signal occurs at the fingertip, where a miniature, high-bandwidth, uniaxial accelerometer (sensitivity = 100 mV/g, frequency response $\pm 10\%$ = 2 to 10,000 Hz) (series 3225, Dytran Instruments, Inc., Chatsworth, CA, USA) is placed in a rigid plastic housing. The accelerometer is sensitive enough to resolve small vibrations caused by the articulation of the internal components of the knee joint that travel to the skin surface [12].

Sandwiched between the accelerometer and the fingertip is a capacitive force sensor (CS8-10N, SingleTact, Los Angeles, CA, USA) encased in silicone rubber (OOMOO 30, Smooth-On, Lower Macungie, PA, USA). The force sensor (full-scale range = 0–10 N) measures contact pressure between the accelerometer and the subject's skin. The utility of this measurement is twofold. First, it complements the acoustic signal captured by the accelerometer, providing context such as whether inconsistent contact is made, potentially a source of signal artifact; such context clues can help the researcher gauge the quality of the joint sound recording. Second, the contact force measurement, in conjunction with real-time sensory (e.g., visual, haptic) feedback, can be used as a mechanism for training users to apply consistent pressure at the sensor-to-skin interface, reducing inter-trial and inter-user variability of recordings. In our system, a multi-color LED on the dorsal surface of the index finger provides visual feedback of sensor contact force via an LED color scheme. A green light indicates that the user is pressing within a desired range of contact force for consistent signal acquisition, with light color changing from blue to red as force exits this range (Figure 3b). This feedback mechanism helped ensure that consistent contact pressure was maintained across trials and across subjects. Intermediate values of contact force (roughly between 4 and 7 N) were found to produce repeatable results in terms of root-mean-squared (RMS) amplitude in the frequency band of

interest, while pressing too hard (between 8 and 10 N) led to discomfort in some subjects. The current study did not directly assess the effects of contact force on signal properties, which is a limitation of the current approach that will be discussed further. The capacitive force sensor itself, though accurate and reliable, is delicate and prone to delamination, so the custom silicone rubber mold protects the sensor from damage while still allowing it to deflect and measure force.

Besides sensor–skin interface force, another potentially important variable to account for is the motion of the joint being assessed. To ensure consistent knee joint displacement and velocity—which can affect the acoustic output of the joint [9,13]—across repetitions of the leg press, we integrated two inertial measurement units (IMUs) (BNO055, Bosch Sensortec, Reutlingen, Germany)—one for the shank segment and one for the thigh—into the glove design. These particular sensors are able to perform onboard sensor fusion and thus output a quaternion estimate. These quaternions are used to estimate the knee joint angle across the leg press maneuver.

Data from the capacitive force sensor and both IMUs (Figure 3a,b) were collected by a Teensy 3.6 microcontroller (PJRC, Sherwood, OR, USA) at a sampling rate of 100 Hz and logged on a microSD card. The microcontroller was housed in a custom enclosure, along with a Bluetooth module (SPBT3.0DP1, STMicroelectronics, Geneva, Switzerland) for streaming data to a laptop and sending/receiving a start/stop signal from MATLAB (MathWorks, Natick, MA, USA). A National Instruments data acquisition unit (USB-4432, Austin, TX, USA) was used to collect the acoustic signals from the four accelerometers at 50 kHz per channel.

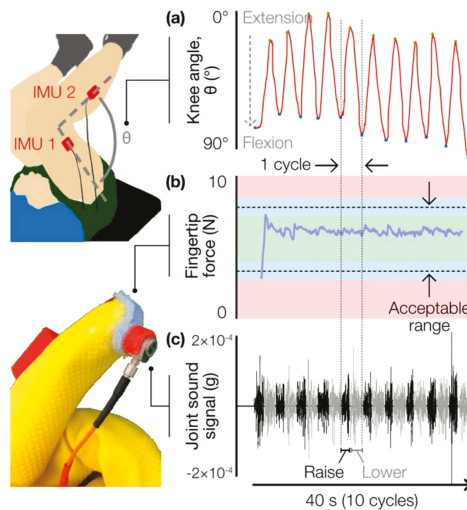


Figure 3. Sample time-series waveforms of signals collected by the glove system during a single experiment trial, consisting of 10 vertical leg press cycles. (a) IMUs were used to confirm that consistent knee range of motion (in degrees, °) was achieved at a constant cadence and to segment the joint sound signal into individual cycles. (b) Contact force (in N) at the fingertip was measured to confirm a consistent amount of pressure was applied. (c) The joint sound signal (local acceleration, in g) was captured by a fingertip-mounted vibration sensor and segmented into cycles consisting of extension (“raise”) and flexion (“lower”) phases.

2.2. Loading Experiment Protocol

All human subjects research was conducted under approval from the Georgia Institute of Technology Institutional Review Board. Eleven healthy subjects (seven male/four female, 25.1 ± 2.5 years, 71.4 ± 16.5 kg, 177.5 ± 11.4 cm) with no history of major knee injury were asked to perform a vertical leg press exercise at three loading conditions referenced to body weight (BW)—0% BW, 50% BW,

and 100% BW—while the joint sound signals from both knees were recorded simultaneously by four accelerometers—two on each knee. The accelerometer placement scheme is shown in Figure 1a. These locations have been shown to be effective for capturing the vibrations internal to the knee, and, importantly, they are anatomical landmarks that are easy to locate and provide relatively unimpeded (i.e., by muscle and fat) access to the internal joint space [14]. One accelerometer was affixed lateral to the patellar tendon of the left knee using fabric tape (Kinesio Tex, Kinesio, Albuquerque, NM, USA). Two accelerometers were affixed medial to the patellar tendon of both knees using double-sided adhesive foam microphone pads commonly used for skin-mounting lavalier microphones (23 mm Stickie, Rycote, Gloucestershire, UK). These sensors, attached to the corresponding locations on either knee using the same mounting technique, served as a matched comparison to indicate whether a subject’s left and right knee produced disparate results; in this case, a comparison across all four accelerometers would be invalid. Both the fabric tape and microphone pads have been used previously [11,12,14]. Finally, an accelerometer on the index fingertip of the glove was placed against the right knee lateral to the patellar tendon. For consistency, the glove-based acquisition was performed by the same individual for all subjects.

At each loading condition, the subject performed 10 repetitions of the leg press maneuver at a rate of one repetition every 4 seconds (i.e., raise for 2 s, lower for 2 s, repeat). Subjects were asked to traverse the same joint displacement each repetition, which was confirmed visually by marking off upper and lower positions on the leg press machine pylons. Consistent cadence was confirmed by ensuring that the time elapsed between each successive flexion-extension (FE) cycle—i.e., time between local minima of the joint angle estimation—deviated no more than 0.2 s from the ideal 4 s period. Loading conditions were randomized to minimize fatigue and learning effects. Two trials were conducted for each condition to confirm a consistent result.

2.3. Signal Processing and Data Analysis

We hypothesized that increasing vertical loading in healthy individuals would cause the articular surfaces in the knee to grind together more forcefully, thus increasing the loudness of the sounds associated with grinding (Figure 1b). By both visual and auditory assessment, we concluded that these grinding sounds were consistent with the lower-amplitude, lower-frequency component of the accelerometer signal. This hypothesis is supported by the fact that other joint sound sensing methods tend to focus on the low-frequency spectrum (<1 kHz) [15]. To that end, we posited that low-pass filtering the signal and then computing its RMS amplitude would give us a reasonable metric for grinding loudness.

The signals were digitally filtered using a Kaiser-window finite impulse response bandpass filter with bandwidth from 10 to 800 Hz. Frequency content below 10 Hz was removed to account for baseline wander of the accelerometer signal caused by coarse movement of the limb during the leg press task. This filtering approach is distinct from that of other work such as Reference [16], in which the primary goal was to capture large-amplitude, high-bandwidth peaks (“clicks” of the joint) in the acoustic signal. In those studies, air microphones offset from the skin surface were used to record the joint sounds instead of contact microphones placed against tissue; in such a scenario, the low-frequency, low-energy acoustic waves would be greatly attenuated at the skin–air interface, so their contribution was not considered.

IMU data were used to segment the filtered data into cycles (10 cycles per recording) (Figure 3c), and the signal RMS was computed on a cycle-to-cycle basis. Outlier RMS values (those that were more than three mean absolute deviations from the median) were rejected, and after confirming that the fingertip force and joint range of motion across each cycle were within acceptable ranges (Figure 3a,b), the remaining RMS values were averaged, yielding a single mean RMS value per loading condition for each subject. Each RMS value was normalized to that subject’s baseline RMS (i.e., the RMS value at 0% BW). This allowed for comparison of grinding loudness across subjects while minimizing inter-subject baseline RMS variability.

2.4. Repeatability Testing: Protocol and Analysis

2.4.1. Comparison of Repeatability between Mounting Techniques

To determine whether the glove-based joint sound sensing system can produce consistent, repeatable, and reliable measurements from cycle to cycle and from trial to trial, we analyzed the joint sounds from a single subject over three days using intraclass correlation coefficient (ICC). ICC is a widely used technique for assessing the degree of correlation and agreement between measurements [17,18].

The glove-based system and the two conventional mounting techniques (fabric tape and foam microphone pads) were used for comparison. Using each of these techniques, the accelerometer was placed on the medial side of the patella, and the subject was asked to perform five cycles of FE per trial, with three such trials conducted for each mounting technique. The signals were digitally bandpass-filtered (10–800 Hz, same as described above), and several key features commonly used in acoustic analysis were extracted for each FE cycle: acoustic energy, energy entropy, and median normalized frequency of the power spectrum. The data were organized into a matrix in which each row represented a single trial and each column represented a single FE cycle. A total of four datasets were used, three of which exclusively included trials of each of the three mounting techniques, with the fourth dataset containing all trials across all three mounting techniques. The three individual datasets were used to evaluate the internal consistency of each of the mounting techniques, while the combined dataset was used to assess the level of agreement among the three mounting techniques. Using the two-way random effects model [17], ICC values were calculated for each dataset to show the reliability of acoustic features calculated both across FE cycles and across trials.

2.4.2. Effect of Fingertip Contact Force Consistency on Repeatability

As mentioned previously, a capacitive force sensor embedded in the fingertip of the glove system provides information about sensor-to-skin contact—inconsistent or inadequate contact could produce artifacts in the joint sound signal, and, therefore, unreliable results. To assess the value of the force sensor experimentally, a single subject was asked to perform a seated, unloaded knee FE task while joint sound signals were acquired by a glove-mounted contact microphone at the same mounting position used in the loading experiment described above (i.e., lateral to the patellar tendon of the right knee), along with contact force and IMU data. Testing was performed on a single-subject basis to reduce the contribution of inter-subject variability on results. The subject performed eight trials of unloaded FE at the same cadence as the loading experiment (one repetition every 4 s), with each trial lasting 30 s in total. Across all eight trials, four were conducted under conditions of consistent contact in which the experimenter relied on visual feedback from the RGB LED to modulate fingertip force; the other four trials were conducted under conditions of inconsistent contact, in which the sensor occasionally lost contact with the skin due to imprecise fingertip force control. Repeatability of results was analyzed on a within-trial (i.e., between each FE cycle) and across-trial (i.e., average of all FE cycles from all four trials) basis. Specifically, consistency of the feature of interest, low-frequency RMS amplitude, was quantified using standard deviation (SD) and coefficient of variation (CV)—the ratio of sample standard deviation to sample mean—as metrics of reliability and repeatability.

3. Results and Discussion

3.1. Effect of Leg Press Load on Knee Grinding Loudness

The key result of this study is illustrated in Figure 4, which shows that relative grinding loudness (RMS of the low-pass-filtered joint sound signal, referenced to the no-load, or 0% BW, condition) within the knee increased significantly ($p < 0.01$, using paired sample t -test with Holm–Bonferroni correction) and monotonically with vertical loading for all three mounting techniques across subjects. Furthermore, comparison across techniques at each loading condition showed no significant ($p < 0.01$) differences in grinding loudness between the glove, mic pads, and tape. This finding—in particular, that the glove

achieved a comparable result (both in terms of the actual RMS quantity and its relationship to the test condition) to that of the other two conventional techniques—supports the idea that a glove-based form factor is an effective approach for capturing and extracting information from joint sounds.

3.2. Repeatability of Glove Versus Conventional Techniques

The central result of repeatability testing is shown in Table 1, which reports the ICC values calculated for each dataset described in Section 2.4.1. While there is no standard value for acceptable reliability using ICC, a general rule suggests that values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability, and values greater than 0.90 indicate excellent reliability [17]. For the glove-only dataset, repeatability analysis yielded ICC values of 0.984 (95% CI of 0.972–0.992) for the acoustic energy feature, 0.947 (95% CI of 0.905–0.975) for acoustic entropy, and 0.954 (95% CI of 0.916–0.977) for median-normalized frequency of the power spectrum (MDF). These results suggest that, in terms of three features commonly used to describe distinct characteristics of acoustic signals, a glove-based system can acquire consistent and repeatable joint sounds information between FE cycles and across trials. For the tape-only dataset, repeatability analysis yielded ICC values of 0.928 (95% CI of 0.877–0.962) for acoustic energy, 0.735 (95% CI of 0.608–0.859) for acoustic entropy, and 0.922 (95% CI of 0.867–0.958) for MDF. For the pads-only dataset, repeatability analysis yielded ICC values of 0.937 (95% CI of 0.893–0.967) for acoustic energy, 0.776 (95% CI of 0.608–0.859) for acoustic entropy, and 0.922 (95% CI of 0.867–0.958) for MDF. The results indicate that the level of reliability for pads and tape can be regarded as “good” to “excellent” for acoustic energy and MDF and “moderate” to “good” for acoustic entropy. For the dataset in which all three mounting techniques were included, repeatability analysis yielded ICC values of 0.982 (95% CI of 0.972–0.989) for acoustic energy, 0.836 (95% CI of 0.742–0.903) for acoustic entropy, and 0.976 (95% CI of 0.962–0.986) for MDF. These results demonstrate a high degree of agreement between features derived from each mounting technique, which further suggests that the glove-based system is a reliable alternative to the conventional methods of mounting the acoustic/vibration sensors to skin.

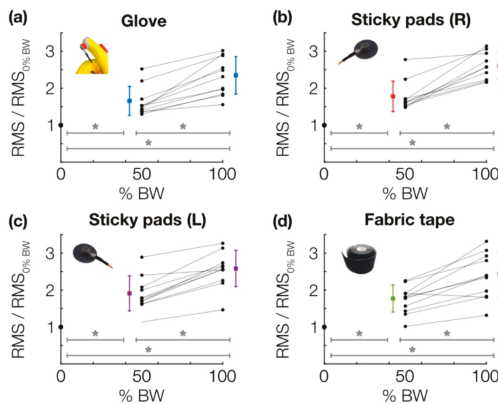


Figure 4. Relative grinding loudness vs. % body weight applied for each mounting technique, including (a) the instrumented glove, (b) adhesive microphone pads mounted on the right leg, (c) adhesive microphone pads mounted on the left leg (for determining comparability between left and right knees), and (d) fabric kinesiology tape. Across 11 subjects, each mounting technique demonstrates the same trend: a monotonic, significant increase in baseline-normalized RMS with increasing vertical load. (*) indicates significance ($p < 0.01$) as determined by paired Student’s *t*-test with Holm–Bonferroni correction.

Table 1. Results of repeatability testing using intraclass correlation coefficient (ICC) as an indicator of measurement repeatability and agreement among mounting techniques.

	Glove		Tape		Sticky Pads		Glove, Tape, and Sticky Pads	
	ICC	95% CI *	ICC	95% CI *	ICC	95% CI *	ICC	95% CI *
Acoustic Energy	0.984	0.972–0.992	0.928	0.877–0.962	0.937	0.893–0.967	0.982	0.972–0.989
Acoustic Entropy	0.947	0.905–0.975	0.735	0.648–0.859	0.776	0.672–0.871	0.836	0.742–0.903
Median Frequency Power	0.954	0.916–0.977	0.922	0.867–0.958	0.921	0.865–0.959	0.976	0.962–0.986

* CI = Confidence Interval.

3.3. Consistent Contact Force Improves Consistency of Results

Figure 5 shows a snippet of two representative trials comparing the effects of consistent versus inconsistent contact force on the joint sound signal captured by a fingertip-mounted accelerometer. Figure 5b depicts how loss of sensor contact (characterized by a rapid decrease in the contact force signal) coincides with regions of the joint sound signal corrupted by signal artifact. Importantly, these signals serve as an example of how the capacitive force sensor can be used to identify unreliable or low-SNR portions of a joint sound recording. Table 2 demonstrates the benefit of consistent contact on the acquired joint sound signal more quantitatively. In this table, values of mean, standard deviation, and coefficient of variation of “grinding loudness” (low-frequency RMS amplitude) are reported for each FE trial and across trials for both test conditions (i.e., consistent and inconsistent force applied at the fingertip). The key takeaway can be found in the last column of the table, in which the variation in grinding loudness across all FE cycles collected with consistent contact ($CV = 0.131$) can be seen to be, on average, less than 25% that of the trials with inconsistent contact ($CV = 0.550$). This finding highlights the fact that consistent contact is critical for obtaining reliable results.

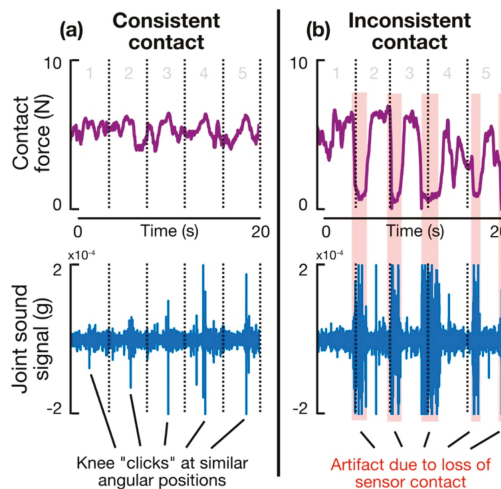


Figure 5. Fingertip contact force and joint sound signal waveforms for representative trials with (a) consistent and (b) inconsistent sensor-skin contact. Time duration of each waveform is 20 s, in which five seated knee flexion-extension cycles were completed at a rate of 4 s per cycle. Highlighted portions illustrate how a rapid decrease in contact force coincides with regions of the joint sound signal dominated by artifact. These data demonstrate that sensor-to-skin contact force can be used as a context clue for rejection of noisy, low-quality joint sound signals.

3.4. Considerations for a Hand-Worn Acoustic Sensing System

Using a glove-based system to measure joint sounds presents both benefits and challenges. This technique offers better sensor-to-skin contact, given the nervous system’s capacity for precise endpoint control, but introduces the possibility of user error. The mechanical sensitivity required to resolve vibrations on the surface of the skin caused by internal motion/friction of the joint makes the job of the glove wearer that much more difficult, for a slight change in fingertip contact with the subject’s skin can corrupt the underlying joint sound signal. Thus, contact force feedback and training are important to minimize human error. Other techniques such as tape do not suffer the same limitation, but they have their own. Namely, any approach that uses adhesive has the potential to couple interface sounds—e.g., tape lifting on/off of the skin—into the recording. Furthermore, these artifacts can be difficult to distinguish from the joint sound signal or can bury it entirely. In this way, a glove-based system, coupled with some feedback mechanism and adequate training, presents a major advantage over other, more established techniques. Additionally, a glove-based system eliminates the need for disposables like sticky pads or tape, the use of which may cause discomfort and of which the adhesive may degrade during use, leading to inconsistency in sensor-to-skin contact. Furthermore, the benefit of a glove form factor over a wearable brace with embedded sensors is its versatility of use across joints and across subjects of different sizes/shapes; a brace would require custom fitting—a potentially painstaking task to ensure optimal, consistent contact.

Table 2. Effects of contact force consistency on repeatability statistics of “grinding loudness” (low-frequency root-mean-square (RMS) amplitude) feature.

		Within Trial (Cycle-to-Cycle)			Across Trials		
		Mean	SD *	CV **	Mean	SD *	CV **
Consistent contact	trial 1	8.98×10^{-3}	7.10×10^{-4}	0.079	1.02×10^{-2}	1.34×10^{-3}	0.131
	trial 2	9.63×10^{-3}	1.56×10^{-3}	0.162			
	trial 3	1.09×10^{-2}	7.64×10^{-4}	0.070			
	trial 4	1.14×10^{-2}	4.13×10^{-4}	0.036			
Inconsistent contact	trial 1	1.93×10^{-2}	5.22×10^{-3}	0.270	3.24×10^{-2}	1.78×10^{-2}	0.550
	trial 2	2.32×10^{-2}	7.07×10^{-3}	0.305			
	trial 3	4.58×10^{-2}	2.15×10^{-2}	0.469			
	trial 4	4.11×10^{-2}	1.73×10^{-2}	0.422			

*SD = standard deviation, **CV = coefficient of variation.

4. Conclusions and Future Work

The work reported here—in which we used an experimental framework known to reliably alter healthy knee joint sounds, coupled with a simple metric (low-frequency RMS amplitude) that manifests some changing physical properties of the joint—serves chiefly as a proof-of-concept validation of our glove-based method of joint sound sensing. Validation was further conducted through repeatability testing, which indicated that our glove-based system was able to produce consistent results, particularly under conditions of consistent fingertip contact force, and a high level of agreement with conventional techniques used to couple vibration sensors to skin. Future efforts should focus on identifying additional metrics for comparing the utility/performance of various form factors and, more importantly, deploying the glove in affected populations (e.g., acute knee injury, arthritis). Inter-user variability should be studied to establish the repeatability of results when different users administer the glove. As mentioned previously, this variability may arise from a host of factors, including accuracy of sensor placement near the targeted anatomical landmark and amount of contact pressure applied. While these are limitations of the current approach, future efforts will focus on evaluating the effects of these variables experimentally. We believe that training (i.e., by a medical professional, user manual, or on-device sensory feedback) and experience will be critical for obtaining reliable results, as is the case with any self-administered medical exam or intervention. We envision that a wearable,

hand-worn system, when used in a home setting, could serve not only as an effective tool for capturing arthro-acoustic information but also as an opportunity for an individual to directly partake in the healthcare of a dependent, such as the parent of a child with juvenile idiopathic arthritis. Whatever the technique or application, as the value of a joint-sounds-based approach to joint health assessment is better developed, exploration of different techniques and improvements on existing ones will be critical for obtaining the best possible information and achieving the best possible clinical outcome.

Author Contributions: Conceptualization: N.B.B. and O.T.I.; Formal analysis: N.B.B. and H.K.J.; Funding acquisition: O.T.I.; Investigation: N.B.B.; Methodology: N.B.B. and H.K.J.; Project administration: O.T.I.; Resources: D.W. and O.T.I.; Validation: N.B.B. and H.K.J.; Writing—original draft: N.B.B. and H.K.J.; Writing—review & editing: D.W. and O.T.I.

Funding: This material is based upon work supported in part by the National Institutes of Health, National Institute of Biomedical Imaging and Bioengineering, Grant No. 1R01EB023808, as part of the NSF/NIH Smart and Connected Health Program, and the National Science Foundation CAREER Award, Grant No. 1749677.

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

References

1. Majewski, M.; Susanne, H.; Klaus, S. Epidemiology of athletic knee injuries: A 10-year study. *Knee* **2006**, *13*, 184–188. [[CrossRef](#)] [[PubMed](#)]
2. Hootman, J.M.; Helmick, C.G.; Barbour, K.E.; Theis, K.A.; Boring, M.A. Updated projected prevalence of self-reported doctor-diagnosed arthritis and arthritis-attributable activity limitation among US adults, 2015–2040. *Arthritis Rheumatol.* **2016**, *68*, 1582–1587. [[CrossRef](#)] [[PubMed](#)]
3. Woolf, A.D.; Åkesson, K. Understanding the burden of musculoskeletal conditions: The burden is huge and not reflected in national health priorities. *BMJ Br. Med. J.* **2001**, *322*, 1079. [[CrossRef](#)] [[PubMed](#)]
4. McCoy, G.F.; McCrea, J.D.; Beverland, D.E.; Kernohan, W.G.; Mollan, R. Vibration arthrography as a diagnostic aid in diseases of the knee. A preliminary report. *J. Bone Jt. Surg. Br. Vol.* **1987**, *69*, 288–293. [[CrossRef](#)]
5. Chu, M.L.; Gradisar, I.A.; Railey, M.R.; Bowling, G.F. Detection of knee joint diseases using acoustical pattern recognition technique. *J. Biomech.* **1976**, *9*, 113–114. [[CrossRef](#)]
6. Shark, L.-K.; Chen, H.; Goodacre, J. Discovering differences in acoustic emission between healthy and osteoarthritic knees using a four-phase model of sit-stand-sit movements. *Open Med. Inform. J.* **2010**, *4*, 116–125. [[CrossRef](#)] [[PubMed](#)]
7. Frank, C.B.; Rangayyan, R.M.; Bell, G.D. Analysis of knee joint sound signals for non-invasive diagnosis of cartilage pathology. *IEEE Eng. Med. Biol. Mag.* **1990**, *9*, 65–68. [[CrossRef](#)] [[PubMed](#)]
8. Töreyn, H.; Jeong, H.K.; Hersek, S.; Teague, C.N.; Inan, O.T. Quantifying the consistency of wearable knee acoustical emission measurements during complex motions. *IEEE J. Biomed. Health Inform.* **2016**, *20*, 1265–1272. [[CrossRef](#)] [[PubMed](#)]
9. Teague, C.N.; Hersek, S.; Töreyn, H.; Millard-Stafford, M.L.; Jones, M.L.; Kogler, G.F.; Sawka, M.N.; Inan, O.T. Novel methods for sensing acoustical emissions from the knee for wearable joint health assessment. *IEEE Trans. Biomed. Eng.* **2016**, *63*, 1581–1590. [[CrossRef](#)] [[PubMed](#)]
10. Hersek, S.; Pouyan, M.B.; Teague, C.N.; Sawka, M.N.; Millard-Stafford, M.L.; Kogler, G.F.; Wolkoff, P.; Inan, O.T. Acoustical emission analysis by unsupervised graph mining: A novel biomarker of knee health status. *IEEE Trans. Biomed. Eng.* **2018**, *65*, 1291–1300. [[CrossRef](#)] [[PubMed](#)]
11. Jeong, H.-K.; Pouyan, M.B.; Whittingslow, D.C.; Ganti, V.; Inan, O.T. Quantifying the Effects of Increasing Mechanical Stress on Knee Acoustical Emissions Using Unsupervised Graph Mining. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2018**, *26*, 594–601. [[CrossRef](#)] [[PubMed](#)]
12. Semiz, B.; Hersek, S.; Whittingslow, D.C.; Ponder, L.A.; Prahalad, S.; Inan, O.T. Using Knee Acoustical Emissions for Sensing Joint Health in Patients with Juvenile Idiopathic Arthritis: A Pilot Study. *IEEE Sens. J.* **2018**, *18*, 9128–9136. [[CrossRef](#)] [[PubMed](#)]
13. Teague, C.; Hersek, S.; Töreyn, H.; Millard-Stafford, M.L.; Jones, M.L.; Kogler, G.F.; Sawka, M.N.; Inan, O.T. Novel approaches to measure acoustic emissions as biomarkers for joint health assessment. In Proceedings of the 2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN), Cambridge, MA, USA, 9–12 June 2015; pp. 1–6.

14. Whittingslow, D.; Semiz, B.; Ponder, L.; Vega-Fernandez, P.; Inan, O.; Prahalad, S. Knee Joint Sounds: A Non-Invasive Modality for Classifying Knee Joint Health in Juvenile Idiopathic Arthritis. In Proceedings of the 2017 ACR/ARHP Annual Meeting, San Diego, CA, USA, 3–8 November 2017.
15. Tavathia, S.; Rangayyan, R.M.; Frank, C.B.; Bell, G.D.; Ladly, K.O.; Zhang, Y.-T. Analysis of knee vibration signals using linear prediction. *IEEE Trans. Biomed. Eng.* **1992**, *39*, 959–970. [[CrossRef](#)] [[PubMed](#)]
16. Jeong, H.K.; Whittingslow, D.; Inan, O.T. b-Value: A Potential Biomarker for Assessing Knee-Joint Health Using Acoustical Emission Sensing. *IEEE Sens. Lett.* **2018**, *2*, 1–4. [[CrossRef](#)] [[PubMed](#)]
17. Koo, T.K.; Li, M.Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* **2016**, *15*, 155–163. [[CrossRef](#)] [[PubMed](#)]
18. Koch, G.G. Intraclass Correlation Coefficient. Available online: <https://doi.org/10.1002/9781118445112.stat02461> (accessed on 1 May 2019).



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).

Article

Learning the Orientation of a Loosely-Fixed Wearable IMU Relative to the Body Improves the Recognition Rate of Human Postures and Activities

Michael B. Del Rosario ¹, Nigel H. Lovell ¹ and Stephen J. Redmond ^{1,2,3,*}

¹ Graduate School of Biomedical Engineering, UNSW, Sydney, NSW 2052, Australia; mikedelr@gmail.com (M.B.D.R.); n.lovell@unsw.edu.au (N.H.L.)

² UCD School of Electrical and Electronic Engineering, University College Dublin, Belfield, 4 Dublin, Ireland

³ UCD Centre for Biomedical Engineering, University College Dublin, Belfield, 4 Dublin, Ireland

* Correspondence: stephen.redmond@ucd.ie

Received: 14 May 2019; Accepted: 21 June 2019; Published: 26 June 2019

Abstract: Features were developed which accounted for the changing orientation of the inertial measurement unit (IMU) relative to the body, and demonstrably improved the performance of models for human activity recognition (HAR). The method is proficient at separating periods of standing and sedentary activity (i.e., sitting and/or lying) using only one IMU, even if it is arbitrarily oriented or subsequently re-oriented relative to the body; since the body is upright during walking, learning the IMU orientation during walking provides a reference orientation against which sitting and/or lying can be inferred. Thus, the two activities can be identified (irrespective of the cohort) by analyzing the magnitude of the angle of shortest rotation which would be required to bring the upright direction into coincidence with the average orientation from the most recent 2.5 s of IMU data. Models for HAR were trained using data obtained from a cohort of 37 older adults (83.9 ± 3.4 years) or 20 younger adults (21.9 ± 1.7 years). Test data were generated from the training data by virtually re-orienting the IMU so that it is representative of carrying the phone in five different orientations (relative to the thigh). The overall performance of the model for HAR was consistent whether the model was trained with the data from the younger cohort, and tested with the data from the older cohort after it had been virtually re-oriented (Cohen's Kappa 95% confidence interval [0.782, 0.793]; total class sensitivity 95% confidence interval [84.9%, 85.6%]), or the reciprocal scenario in which the model was trained with the data from the older cohort, and tested with the data from the younger cohort after it had been virtually re-oriented (Cohen's Kappa 95% confidence interval [0.765, 0.784]; total class sensitivity 95% confidence interval [82.3%, 83.7%]).

Keywords: quaternion; smartphone; feature engineering; human activity recognition; sensor fusion

1. Introduction

Wearable movement sensors, i.e., sensors that incorporate inertial measurement units (IMUs) and barometric altimeters, have been championed as tools that will positively impact health care [1]. These technologies have demonstrated their utility in the remote monitoring of patient rehabilitation [2], as well as the clinical analysis of gait [3], from which parameters can be extracted to predict falls in the elderly [4,5]. They have come to prominence in the management of Parkinson's disease [6], objectively quantifying patient tremor [7], and tracking the impact of the disease on their gait [8]. Moreover, wearables have been adopted for the longitudinal monitoring of physical activity, which can be used to identify those at risk of developing type-2 diabetes [9] and obesity [10].

1.1. Multiple Sensors or a Single Sensor?

The number of sensors that an individual needs to wear for adequate human activity recognition (HAR) is dependent on three factors: (i) the number of activities to be recognized, (ii) the location of the sensor(s) on the body, and (iii) the nature of the sensors (i.e., some arbitrary combination of accelerometers, gyroscopes, barometric pressure sensors, magnetometers, etc.). If the activities to be recognized involve the movement of each of the body's limbs (e.g., lunges, push ups, hand stands, etc.), multiple sensors may need to be worn on the body at specific locations to obtain measurements that allow the movement to be accurately identified. Wearing a single sensor on the body is sufficient if the model for HAR is only identifying gross body movements (e.g., standing, sitting, walking, running), placing the sensor near the body's center of mass is ideal (i.e., the thigh) [11].

In either case, wearing sensors at different locations on the body will increase the performance of a model for HAR [12], but at the expense of user compliance and burden [13], particularly if the sensor(s) must be placed somewhere uncomfortable or unsightly [14]. Consequently, wearable sensor systems for population-based studies are predominantly of the single-sensor variety [15,16], ideally integrating seamlessly into the daily lives of users (e.g., embedded within a watch, necklace, sock or belt) [17].

1.2. Smartphone-Based Human Activity Recognition

The dramatic recent increase in smartphone ownership [18,19] coupled with society's dependence on smartphones [20,21] has changed the paradigm. If individuals carry their device with them, the measurements from the smartphone's IMU and barometric altimeter can be analyzed to identify the users' gross body movements throughout the day. As a result, smartphones can be used as tools for the purposes of physical rehabilitation, weight loss, etc., in which the ability to recognize human activity is essential [22]. Finally, the greater penetration of smartphones amongst those of a lower socioeconomic status [23] would enable population-based interventions to be conducted at a reduced cost and with a wider reach. There are different models for HAR which can be adopted [24].

1.2.1. Fixed-to-the-Body

In this scenario, the IMUs embedded within smartphones are used in place of dedicated IMU devices to: (i) detect falls [25], (ii) monitor activities of daily living [26], (iii) monitor the performance of soccer and field hockey athletes [27], and (iv) swimmers [28]. These models assume that the smartphone will be worn at one location on the body and that the device's orientation relative to the body segment on which it is worn is known *a priori* and does not change during the monitoring period because it is held in place with a strap or similar apparatus.

1.2.2. Body-Position-Dependent

Conversely, models can be designed under the assumption that the smartphone is not strapped to the body and will be placed in either the user's pants/chest pocket [29], or hand or bag [30]. These models do not require the user's smartphone habits to change (as with those in Section 1.2.1) to accommodate the device being fixed to the body, however, this makes it challenging to infer the user's postural orientation due to the variability with which the sensor can be oriented in the pocket with respect to the body (i.e., the initial orientation of the device relative to the body segment on which it is worn cannot be controlled, and the orientation of the device can vary over time since it is not firmly fixed to the body).

1.2.3. Body-Position-Dependent

The final variant relaxes all constraints with respect to the smartphone's position and orientation on the body. Models are robust to device transitions from hand to pants/chest pocket [31], or bag, at any moment [32,33]. A trade-off for this robustness (compared to those discussed in Section 1.2.1

or 1.2.2) is that it can be difficult to determine body posture due to variability in both the device's location and its relative orientation to the body.

1.3. Extracting Information for Activity Recognition

There are two broad supervised learning approaches which have emerged to process sensor measurements for HAR: (i) feature engineering and classification, or (ii) deep learning.

1.3.1. Accounting for Variability in Device Orientation and Position

A number of pre-processing techniques have been proposed to reduce the variability in sensor measurements due to the inconsistency of the location and/or orientation with which the device is placed on the body. Khan et al. demonstrated that linear discriminant analysis (LDA) can be used to improve a classification algorithm's ability to distinguish between transitions from sitting to standing (and vice versa), and standing to lying (and vice versa) [34]. They also illustrated how kernel discriminant analysis (KDA) can estimate both the interclass and intraclass variance of features used to separate periods of walking, running, walking upstairs, and walking downstairs [35]. Henpraserttae et al. applied eigenvalue decomposition to tri-axial accelerometer measurements to infer the device's orientation with respect to the body by assuming that most of the acceleration due to body movement is in the forward direction, and that the vertical axis can be inferred from the low-pass filtered acceleration [36].

Yurtman and Barshan proposed another transformation based on singular value decomposition (SVD). They first pre-processed the data from a tri-axial accelerometer, tri-axial gyroscope, and tri-axial magnetometer so that it had unit variance, before SVD was applied to the entire time sequence to make the sensor measurements agnostic to the device's orientation [37]. Yurman et al. followed their seminal work with another method which combined the measurements from the accelerometer, gyroscope, and magnetometer, to estimate the sensor's orientation within the global frame of reference when it is firmly fixed to the body. The differential quaternions they generated, which estimated the relative change in the device's orientation between time intervals, enabled the raw sensor measurements to be expressed in a reference frame invariant to the sensor's orientation [38].

1.3.2. Feature Engineering and Classification

Feature engineering involves the application of domain knowledge to design hand-crafted features [39,40] which describe the changing characteristics of the data with respect to time. These features and labels (i.e., the human activities to be recognized which are temporally aligned with the feature values) are input to a classification algorithm (e.g., decision tree, support vector machine, Naïve Bayes, artificial neural networks, etc.) which tries to derive the best mathematical model that separates the labeled observations, based on the statistic distributions of those features.

1.3.3. Deep Learning/Deep Neural Networks

Alternatively, domain knowledge can be replaced with a standalone artificial intelligence solution that abstracts the entire feature extraction and classification process. Deep learning approaches are a natural extension to artificial neural networks, comprised of numerous neurons and layers, which attempt to learn both the Best Features and model for HAR by using the training labels to determine the value of the neurons' weights at each layer in the network [41]. The performance of these approaches are dependent on the network's architecture and the quality and quantity of the training data. While convolutional neural networks [42], short-time Fourier transforms combined with temporal convolution layers [43], long/short term memory (LSTM) networks [44], or a combination of convolution, recurrent, and LSTM network layers [45], have all been shown to perform exceptionally well, they incur a considerable energy cost when running on a smartphone due to the demands of real-time processing [46].

1.4. Contribution

This paper addresses the limitations associated with methods for inferring postural orientation that are dependent on the sensor's precise anatomical placement [47,48] by presenting a novel method (i.e., a hand-crafted feature) for identifying sedentary periods of activity that is robust to variability in the sensor's orientation. The sensor's orientation during walking periods is learned on-line and used as a reference for the upright body orientation (represented by the quaternion q_{upright}). Comparing the sensor's recent average orientation (over a sliding window) to q_{upright} enables standing and sedentary periods to be distinguished, regardless of the IMU's orientation. It is important to distinguish between standing and sedentary activities due to their differing energy expenditure profiles [49,50]. Furthermore, there are definitive relationships between total sedentary time per day and: type-2 diabetes [51,52]; cardiovascular mortality; all-cause mortality [53,54]; and even cancer [55].

2. Materials

Wearable sensor data from our previous work [56], in which a cohort of twenty younger adults (15 male and five female) of ages 21.9 ± 1.7 years (Human Research Ethics Advisory, reference number 08/2012/48) and 37 older adults (25 male and 12 female) of age 83.9 ± 3.4 years (Human Research Ethics Committee, reference number HC12316) performed nine human activities whilst a smartphone was placed in their pants pocket, was used to evaluate the method proposed herein. The younger adults were able-bodied university students recruited from the University of New South Wales, Sydney, Australia. The older adults were recruited from a cohort of participants enrolled in an existing study on memory and aging at Neuroscience Research Australia (NeuRA), Sydney, Australia. These participants were community-dwelling and retirement village residents living in inner and eastern Sydney; aged 65+ years; English-speaking; with a mini-mental state examination (MMSE) score of 24 or above; no acute psychiatric condition with psychosis or unstable medical condition; not currently participating in a fall prevention trial.

Sensor data from the IMU and barometric altimeter were originally sampled at $f_{\text{IMU}} = 100$ Hz and $f_{\text{bar}} = 16$ Hz, respectively. The measurements from the IMU and altimeter were also re-sampled at 40 Hz and 20 Hz, respectively, to demonstrate the method's ability to be adapted for a reduced sampling rate, thereby reducing the prospective power consumption of the algorithm. This is important because the usability of wearable sensors increases if they can operate continuously throughout the waking day [24,57].

Periods of human activity, originally labeled as elevator up/down, were relabeled as standing to focus on the clinical relevance of the activity rather than the wider context of the person being in an elevator; this naturally increased the classification performance by reducing the range of activities being classified. Additionally, sitting and lying were collectively re-labeled as sedentary. Consequently, the nine activities described in [56] were reduced to six: sedentary, standing, walking, walking upstairs (WU), walking downstairs (WD) and postural transitions (PT).

3. Methods

Note in the sections that follow, (i) quaternion multiplication (\otimes) and conjugation (*) are defined in [58]; (ii) vectors are bold-faced (i.e., \mathbf{b}); (iii) quaternions are bold-faced, italicized and normalized unless explicitly stated (i.e., $q = q/|q|$); (iv) vectors expressed in the sensor frame, or estimated in the global frame of reference, will be denoted with the superscripts $^s\mathbf{b}$, and $^g\mathbf{b}$, respectively; (v) a function will be denoted as $f(\dots)$ with arguments inside the brackets.

3.1. Generating Data Representative of Different Orientations

Each quaternion in Figures 1a–f was used to transform the accelerometer and gyroscope data (\mathbf{r}_{acc} and \mathbf{r}_{gyr} , respectively) collected in our previous work [56], into new accelerometer and gyroscope data (\mathbf{v}_{acc} and \mathbf{v}_{gyr} , respectively) that would have been obtained if the smartphone were re-oriented in the

pants pocket (Equation (1)). Note: (i) $\{\mathbf{r}_{acc}, \mathbf{r}_{gyr}, \mathbf{v}_{acc}, \mathbf{v}_{gyr}\} \in \mathbb{R}^3$; (ii) data from the barometric altimeter were not transformed as these scalar measurements are orientation invariant.

$$\begin{bmatrix} 0 & v_x & v_y & v_z \end{bmatrix} = \mathbf{q} \otimes \begin{bmatrix} 0 & r_x & r_y & r_z \end{bmatrix} \otimes \mathbf{q}^* \tag{1}$$

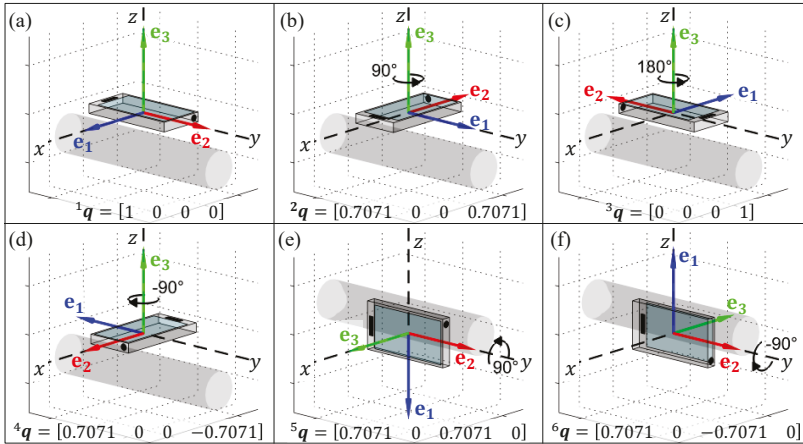


Figure 1. Six common ways that the inertial measurement units (IMU) might be placed in the pants pocket (assuming a seated position). The cylinder in each panel represents the orientation of the participant’s right thigh whilst seated (with the knee to the right-hand side of each image). The dashed lines labeled x , y , and z denote the original device reference frame, 1q , whilst the orthogonal basis defined by the vectors, \mathbf{e}_1 , \mathbf{e}_2 , and \mathbf{e}_3 illustrate the device orientation generated. In panels (a–d) the IMU is located on the anterior surface of the thigh (i.e., a pocket on the front of the pants), whilst in panels (e,f) the IMU is located on the lateral surface of the thigh (i.e., a pocket on the outer side of the pants).

3.2. Estimating the Orientation of the IMU

The data generated in Section 3.1 were fused using the adaptive error-state Kalman filter (AESKF) for orientation estimation, developed in our previous work [59], to estimate the device’s orientation. The tuning parameters of the AESKF algorithm are listed in Table 1. Note, whilst there are many algorithms that can be used to estimate the IMU’s orientation, the AESKF was chosen for its computational efficiency relative to other algorithms [59].

Table 1. Tuning parameters of the computationally-efficient adaptive error-state kalman filter.

Sampling Rate	${}^\dagger \sigma_G$	c_a	c_m	${}^\ddagger N_{short}$	${}^\ddagger N_{long}$	${}^* \zeta_a$	${}^\ddagger N_m$	${}^{**} \zeta_{xy}$
$f_s = 100$ Hz	0.01	0.1	0.99	7	49	1	7	5
$f_s = 40$ Hz	0.01	0.1	0.99	3	19	1	3	5

† rad/s; ‡ samples; * $m\ s^{-2}$; ** normalized units/s.

3.2.1. Removing the Heading from the Estimated Orientation

The estimated orientation, $\mathbf{q}_{AESKF,k'}$, had an arbitrary heading that did not contain any information about the orientation of the IMU on the individual’s body, since the person can face in any direction and perform the same activity. Consequently, this was removed by aligning the orientation, $\mathbf{q}_{AESKF,k'}$, with the north-facing component of the standard basis, $\mathbf{e}_x = [1\ 0\ 0]$. First, the x basis vector of

the quaternion, $q_{\text{AESKF},k}$, was identified and projected to the xy -plane (Equation (2)). Once $x_{xy,k}$ is determined, the quaternion, $q_{\text{north},k}$, that rotates the device orientation, $q_{\text{AESKF},k}$, northward can be calculated (Equations (4)–(6)). The resultant quaternion, q_k , had a fixed heading, i.e., the yaw angle, $\psi = 0$ (see Figure 2b), which ensured that the shortest rotation between two quaternions (Section 4.2.2), did not contain any information about changes in the device’s heading which normally occur due to turning the body.

$$x_{xy,k} = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 + q_0q_3) & 0 \end{bmatrix}_k \quad (2)$$

$$\theta_k = \cos^{-1} \left(\frac{(x_{xy,k} \cdot e_x) / \|x_{xy,k}\|}{\|x_{xy,k}\|} \right) \quad (3)$$

$$n_k = x_{xy,k} \times e_x \quad (4)$$

$$q_{\text{north},k} = \left[\cos\left(\frac{\theta}{2}\right) \ n \sin\left(\frac{\theta}{2}\right) \right]_k \quad (5)$$

$$q_k = q_{\text{north},k} \otimes q_{\text{AESKF},k} \quad (6)$$

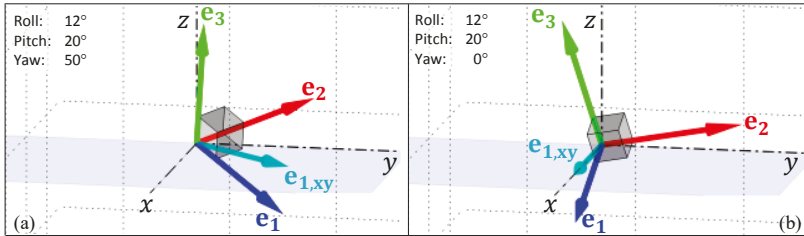


Figure 2. Effect of removing the yaw from an arbitrary orientation (the orthogonal basis defined by the vectors e_1 (in blue), e_2 (in red), e_3 (in green) by aligning it with the x -axis of the standard basis x, y, z): (a) orientation, with an arbitrary yaw angle; (b) the same orientation with the yaw component removed (see Equations (2)–(5)). Note: the light blue vector is $e_{1,xy}$, the e_1 basis vector of the orientation projected to the xy plane; the pitch and roll angles are preserved.

3.2.2. Smoothing the Estimated Orientation

The effects of the IMU shifting/re-orienting within the individual’s pants pocket as they move through the world were minimized by time-averaging the quaternion, q_k , using a computationally-efficient one-pass method [60], to create a moving average (window size N) of the device orientation (Equation (7)) from the last 2.5 s worth of data, \bar{q}_k .

$$\bar{q}_k = f_{q,\text{avg}}(q_{k-N+1}, \dots, q_k), \quad (7)$$

see Appendix A Equation (A1).

4. Feature Extraction

The features in Table 2 were aggregated (using a sliding window with 50% overlap) using sensor data from the most recent 2.5 s. Features (1)–(4) were obtained by processing the sensor measurements with finite impulse response (FIR) linear phase filters, as described in our previous work [56]. Whilst novel features (5)–(8) are described in Sections 4.1–4.3.

Table 2. Features extracted from the accelerometer, gyroscope, and barometric altimeter.

No.	Feature	Description
(1)	$\bar{\omega}_{\text{bpf},k}^2 = \frac{1}{N} \sum_{j=i}^k (\omega_{\text{bpf},x}^2 + \omega_{\text{bpf},y}^2 + \omega_{\text{bpf},z}^2)_j$	average squared band-pass-filtered angular velocity
(2)	$\bar{\Theta}_{\text{tilt},k} = \frac{1}{N} \sum_{j=i}^k \cos^{-1} \left(a_{\text{ipf},y} / \sqrt{(a_{\text{ipf},x}^2 + a_{\text{ipf},y}^2 + a_{\text{ipf},z}^2)} \right)_j^\dagger$	average inclination angle
(3)	$\bar{a}_{\text{ipfdif},k}^2 = \frac{1}{N} \sum_{j=i}^k (a_{\text{ipfdif},x}^2 + a_{\text{ipfdif},y}^2 + a_{\text{ipfdif},z}^2)_j$	average squared band-pass-filtered acceleration
(4)	$\Delta P_k = \frac{1}{N} \sum_{j=i}^k \partial p_j$	average differential pressure
(5)	${}^s\bar{\omega}_{xy,k}^2 = \frac{1}{N} \sum_{j=i}^k ({}^s\omega_x^2 + {}^s\omega_y^2)_j$	average squared pitch/roll angular velocity
(6)	$\bar{\theta}_{\text{tilt},k} = \frac{1}{N} \sum_{j=i}^k \theta_{\text{tilt},j}$	average of the shortest rotation between the upright and average orientations
(7)	$\Delta \bar{\theta}_{\text{tilt},k} = \left \bar{\theta}_{\text{tilt},k} - \bar{\theta}_{\text{tilt},k-N} \right $	change in the shortest rotation between the upward and average orientations
(8)	$\bar{v}_{z,k} = \frac{1}{N} \sum_{j=i}^k \dot{z}_j$	average velocity in the vertical direction of the estimated GFR

Note: $i = k - N + 1$; N —the number of samples in a 2.5 s analysis window ($[2.5 \times f_{\text{IMU}}]$);
The variables: $\omega_{\text{bpf},k}$, $a_{\text{ipfdif},k}$, $a_{\text{ipf},k}$ and ∂p correspond to the filtered signals described in [56];

[†] The tilt angle [48].

4.1. Squared Magnitude of Pitch/Roll Angular Velocity

In our previous work [56], the three orthogonal gyroscope measurements were each band-pass filtered (between 1 and 20 Hz) to isolate the frequency components predominantly due to walking [61]. The squared magnitude of these three band-pass filtered signals at each time sample, $\omega_{\text{bpf},k'}^2$ was used to distinguish between active/inactive periods of activity. Alternatively, the measurements can be expressed in the estimated global frame of reference (GFR) using the device orientation, \mathbf{q}_k . The squared magnitude of the pitch/roll rotation, ${}^s\omega_{xy,k'}^2$ can be isolated using Equations (8) and (9) since rotations about the vertical axis are primarily due to turning.

$$\begin{bmatrix} 0 & {}^s\omega_x & {}^s\omega_y & {}^s\omega_z \end{bmatrix}_k = \mathbf{q}_k \otimes \begin{bmatrix} 0 & {}^s\omega_x & {}^s\omega_y & {}^s\omega_z \end{bmatrix}_k \otimes (\mathbf{q}_k)^* \quad (8)$$

$${}^s\omega_{xy,k}^2 = {}^s\omega_{x,k}^2 + {}^s\omega_{y,k}^2 \quad (9)$$

4.2. Detecting Sedentary Periods

The tilt angle, $\Theta_{\text{tilt},k}$, was previously used [56] to identify the postural orientation of the body relative to the global frame of reference (GFR) [47,48]. In this previous formulation, the magnitude of $\Theta_{\text{tilt},k}$ is dependent on one of the axes of the IMU (the y -axis in Figure 3a) remaining in coincidence with the long axis of the thigh. This constraint is apparent when the IMU is oriented such that another axis is aligned with the long axis of the thigh (Figure 3b), but is also a problem if the orientation of the IMU shifts in the pocket. A new approach is proposed in Section 4.2.2 which compares the average recent orientation with the average orientation during walking periods (i.e., the sensor's orientation during walking periods is continuously learned and used to define the 'upright' orientation, against which all other orientations are compared).

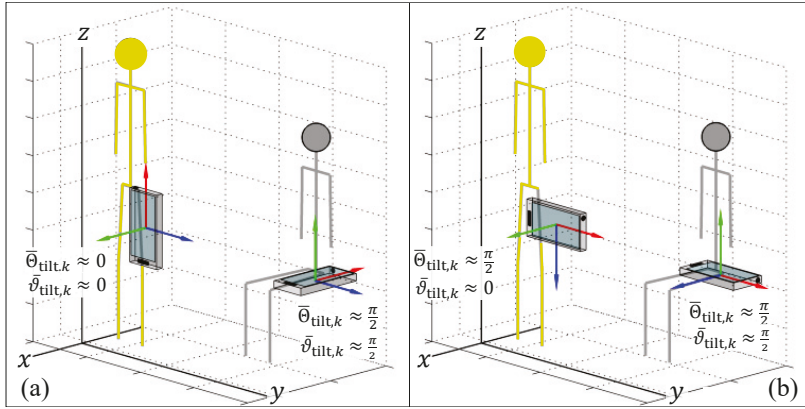


Figure 3. Two alternative methods for measuring the tilt. The limitations of the traditional tilt angle variable (i.e., the angle between the red basis vector and the z-axis of the global frame of reference (GFR)) are clear, making it impossible to separate standing (stick figure in gold) and sedentary periods (stick figure in grey). When the red basis vector: (a) runs along the length of the individual’s leg, the magnitude of the tilt angle changes by $\approx \frac{\pi}{2}$ radians (i.e., 90°) between standing and sedentary periods, making it easy to discriminate these postures; (b) red basis vector runs along the mediolateral axis of the individuals’s leg, so the magnitude of the tilt angle of the red vector remains relatively unchanged between standing and sedentary periods, resulting in confusion between the two postures.

4.2.1. Estimate the Upright Orientation using the Orientation during Walking Periods

Walking periods were identified using the method proposed by Jiménez et al. [62] which analyzed the squared magnitude of the raw tri-axial gyroscope signal, ${}^s\omega_k^2$ (see Equation (10)), and the magnitude of the unbiased sample variance, $\zeta_{\text{acc},k}^2$ (see Equation (12)), of the squared magnitude of the raw tri-axial accelerometer signal, ${}^s\mathbf{a}_k^2$ (see Equation (11)). When both signals are greater than pre-determined thresholds, the individual carrying the IMU was presumed to be walking (see Equation (13)). Note, $j = k - N + 1$ in Equation (12), and $\zeta_{\text{acc},k}^2$ was calculated using a computationally-efficient method [60] with a sliding window of 0.25 s; i.e., $N = \lfloor 0.25 \times f_{\text{IMU}} \rfloor$.

$${}^s\omega_k^2 = {}^s[\omega_x^2 + \omega_y^2 + \omega_z^2]_k \quad (10)$$

$${}^s\mathbf{a}_k^2 = {}^s[a_x^2 + a_y^2 + a_z^2]_k \quad (11)$$

$$\zeta_{\text{acc},k}^2 = \frac{1}{N-1} \left[\left(\sum_{i=j}^k ({}^s\mathbf{a}_i^2)^2 \right) - \frac{1}{N} \left(\sum_{i=j}^k {}^s\mathbf{a}_i^2 \right)^2 \right] \quad (12)$$

$$b_{\text{walk},k} = \begin{cases} 1, & (\omega_k^2 > 5 \text{ rad/s}) \cap (\zeta_{\text{acc},k}^2 > 10 \text{ m}^2/\text{s}^4) \\ 0, & \text{otherwise} \end{cases} \quad (13)$$

The scalar and vector components of each orientation, \mathbf{q}_k , which correspond to these walking periods are stored and used to calculate the ‘upright’ orientation, $\mathbf{q}_{\text{upright},k}$ (see Equation (14) and Appendix A). Note: * denotes the set of N indices corresponding to the most recent 2.5 s of data for which $b_{\text{walk},k} = 1$, and need not be a contiguous set of sample indices.

$$\mathbf{q}_{\text{upright},k} = f_{\mathbf{q},\text{avg}}(\mathbf{q}_{(k-N+1)^*}, \dots, \mathbf{q}_{k^*}), \quad (14)$$

see Appendix A Equation (A1).

The approach presented herein improves upon the method proposed by Elvira et al. because: (i) an orientation algorithm is utilized [63] whose estimated inclination angle is immune to magnetic interference [64,65]; (ii) an adaptive error-state Kalman filter (AESKF) is used which enables the estimated orientation to be quickly corrected during ‘quasi-static’ periods [59]; (iii) the estimated heading (i.e., the yaw component, ψ°) is removed from the orientation estimated (the importance of which is demonstrated in Section 3.2.1); (iv) most importantly this method is believed to be the first to demonstrate the utility of the shortest rotation between two quaternions as a feature for HAR.

4.2.2. Calculate the Shortest Rotation between the Upright Orientation and the Average Recent Orientation

Once the average recent orientation, \bar{q}_k , and the upright orientation, $q_{\text{upright},k}$, are known, the magnitude of the shortest rotation between them (see Equation (15)) can be calculated (see derivation in Appendix B) and used to distinguish standing and sedentary (seated/lying) periods, regardless of the IMU’s orientation relative to the thigh.

$$\vartheta_{\text{tilt},k} = f_{\text{angle}}(q_{\text{upright},k}, \bar{q}_k), \quad (15)$$

see Appendix B Equation (A2).

4.3. Estimating Velocity in the Vertical Direction of the GFR

The inertial acceleration in the sensor frame was obtained by measuring the magnitude of the Earth’s gravitational acceleration, $\|y_{a,0}\|$, (i.e., the accelerometer measurement during a quasi-static period, where the accelerometer is not moving) and expressing this measurement in the sensor frame of reference as ${}^s z_k$ using the accelerometer-corrected attitude, q_k (see Equations (16) and (17)). The acceleration due to gravity, as measured in the sensor frame of reference, ${}^s g_{\text{ref},k}$, can then be subtracted from the raw accelerometer measurement, ${}^s y_{a,k}$ (Equation (18)), to obtain the inertial acceleration in the sensor frame, ${}^s d_{a,k}$.

This acceleration can be expressed in the estimated GFR, ${}^s d_{a,k} = [{}^s d_{a,x} \quad {}^s d_{a,y} \quad {}^s d_{a,z}]_k$, using Equation (19). At this point, the sensor’s velocity in the vertical direction of the GFR can be estimated by fusing the vertical component of the acceleration, ${}^s d_{a,z,k}$, with the barometric pressure sensor measurements, p_k , using a complementary filter [66] or Kalman filter [67]. Assuming the external acceleration, $\dot{z}_k = {}^s d_{a,z,k}$, remains constant over the sampling interval, $T = \frac{1}{f_{\text{IMU}}}$, and the bandwidth of \dot{z}_k is less than $\frac{f_{\text{IMU}}}{2}$, the time-propagation of the altitude, z and velocity, \dot{z} , can be modeled [68] according to Equation (20):

$${}^s z_k = \begin{bmatrix} 2(q_1 q_3 - q_0 q_2) & 2(q_2 q_3 + q_0 q_1) & 2(q_0)^2 - 1 + 2(q_3)^2 \end{bmatrix}_k \quad (16)$$

$${}^s \check{g}_{\text{ref},k} = \|{}^s y_{a,0}\| \cdot {}^s z_k \quad (17)$$

$${}^s d_{a,k} = {}^s y_{a,k} - {}^s \check{g}_{\text{ref},k} \quad (18)$$

$$\begin{bmatrix} 0 & {}^s d_{a,x} & {}^s d_{a,y} & {}^s d_{a,z} \end{bmatrix}_k = q \otimes \begin{bmatrix} 0 & {}^s d_{a,x} & {}^s d_{a,y} & {}^s d_{a,z} \end{bmatrix}_k \otimes q^* \quad (19)$$

$$\begin{bmatrix} z \\ \dot{z} \end{bmatrix}_k = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} z \\ \dot{z} \end{bmatrix}_{k-1} + \begin{bmatrix} \frac{T^2}{2} \\ T \end{bmatrix} \dot{z}_k \quad (20)$$

$$x_k = A_k x_{k-1} + G_k u_k + w_k$$

4.3.1. Process Model

Imperfections in Equation (20), i.e., acceleration not being constant during the sampling interval, and noise in the acceleration input to the system, u_k , prevent the system’s true state, x , from being

observed. Consequently, the system's state can only be estimated as $\tilde{\mathbf{x}}_k$, by combining the process model with measurements obtained directly from the system. The 'prediction step' (i.e., Equations (22) and (23)) produces an *a priori* estimate of the system's state, $\tilde{\mathbf{x}}_k^-$, and covariance, \mathbf{P}_k^- . Note: (i) \mathbf{Q}_k is the process noise covariance matrix, (ii) $\frac{a_m}{2} \leq \sigma_{\text{acc}} \leq a_m$, where a_m is the magnitude of the maximum acceleration the system will experience [68].

$$\tilde{\mathbf{x}}_k^- = \mathbf{A}_k \tilde{\mathbf{x}}_{k-1}^+ + \mathbf{G}_k \mathbf{u}_k \quad (21)$$

$$\mathbf{P}_k^- = \mathbf{A}_k \mathbf{P}_{k-1}^+ \mathbf{A}_k^T + \mathbf{Q}_k \quad (22)$$

$$\mathbf{Q}_k = \mathbf{G}_k \mathbf{G}_k^T \sigma_{\text{acc}}^2 = \begin{bmatrix} \frac{1}{4} T^4 & \frac{1}{2} T^3 \\ \frac{1}{2} T^3 & T^2 \end{bmatrix} \sigma_{\text{acc}}^2 \quad (23)$$

4.3.2. Observation Model

The observation model (Equation (24)) transforms the state estimate, $\tilde{\mathbf{x}}_k$, to the domain of the barometric pressure sensor p_k (i.e., it converts altitude (in m) to air pressure (hPa) [69]), and enables the measurement residual, $\tilde{\mathbf{y}}_k$, to be calculated (Equation (26)). The measurement residual has a covariance \mathbf{S}_k that combines the covariance of the *a priori* state estimate, \mathbf{P}_k^- , and variance in the measurement from the barometric pressure sensor, $\mathbf{R}_k = \sigma_{\text{bar}}^2$ (Equation (27)); i.e., the variance in the barometric pressure when the device remains stationary. The gain in the filter, \mathbf{K}_k , can be determined by consolidating the covariances of the *a priori* state estimate and measurement residual (Equation (28)), thereby enabling the a posteriori state estimate, $\tilde{\mathbf{x}}_k^+$, and covariance, \mathbf{P}_k^+ , to be determined as described in Equations (29) and (30). Note: (i) \mathbf{H}_k is the Jacobian of $h(\mathbf{x}_k)$, that is, derivatives with respect to the elements of the state vector \mathbf{x}_k , evaluated at the estimate $\mathbf{x}_k = \tilde{\mathbf{x}}_k^-$; (ii) \mathbf{I}_2 a 2×2 identity matrix.

$$h(\mathbf{x}_k) = p_0 \left(1 - \frac{z_k}{44330.77} \right)^{5.26} \quad (24)$$

$$\mathbf{H}_k = \left[\begin{array}{cc} \frac{\partial h}{\partial z} & \frac{\partial h}{\partial z^2} \end{array} \right] \Bigg|_{h(\mathbf{x}=\tilde{\mathbf{x}}_k^-)} \quad (25)$$

$$\tilde{\mathbf{y}}_k = p_k - h(\tilde{\mathbf{x}}_k^-) \quad (26)$$

$$\mathbf{S}_k = \mathbf{H}_k \mathbf{P}_k^- \mathbf{H}_k^T + \mathbf{R}_k \quad (27)$$

$$\mathbf{K}_k = \mathbf{P}_k^- \mathbf{H}_k^T \mathbf{S}_k^{-1} \quad (28)$$

$$\tilde{\mathbf{x}}_k^+ = \tilde{\mathbf{x}}_k^- + \mathbf{K}_k \tilde{\mathbf{y}}_k \quad (29)$$

$$\mathbf{P}_k^+ = (\mathbf{I}_2 - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^- \quad (30)$$

It is hoped that the Kalman-filtered velocity estimate, \hat{z}_k , was able to distinguish between walking periods, upstairs ($\hat{z}_k \gg 0$), downstairs ($\hat{z}_k \ll 0$), and on a level surface ($\hat{z}_k \approx 0$). This would extend the utility of the Kalman-filtered velocity estimate, beyond applications in fall detection [70], for example.

5. Hierarchical Description of Human Activity

Rather than use one supervised machine learning algorithm to perform HAR, a hierarchical model of human activity (HMHA) [71,72] was devised and translated into a feature-based model (see Figure 4). A decision tree based on the classification and regression tree (CART) algorithm developed by Brieman [73] was trained for each node of the model and pruned so that there is only one leaf node for each activity class (see an example in Figure 5b). This approach minimized

over-fitting [74], ensured that the model was easily interpreted, and makes the process of HAR tractable in the event of misclassification [75]. In addition, the weights of each class were balanced when the decision tree was trained to ensure that the thresholds selected accounted for any class imbalances [76].

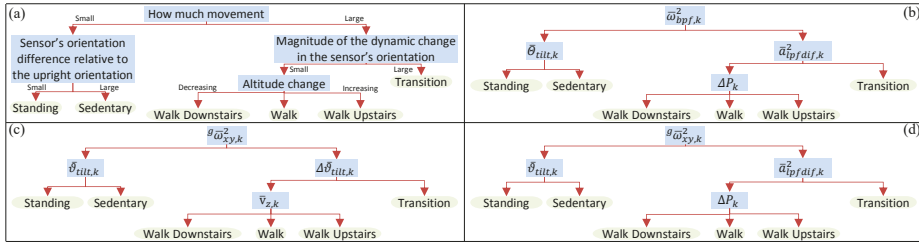


Figure 4. Illustration of how the activity classes can be separated using (a) a hierarchical description of human activity. A schematic for achieving the separation using the following features: (b) only original (i.e., features (1)–(4)), (c) only new (i.e., features (5)–(8)), and (d) best features from all eight old and New Features (i.e., features (3)–(6)) in Table 2. Each blue rectangle represents a classification and regression tree (CART) [73] implemented in MATLAB 2013b with ‘ClassificationTree.fit’. The CART algorithm used ‘uniform’ prior class probabilities to ensure that the thresholds selected accounted for any class imbalances.

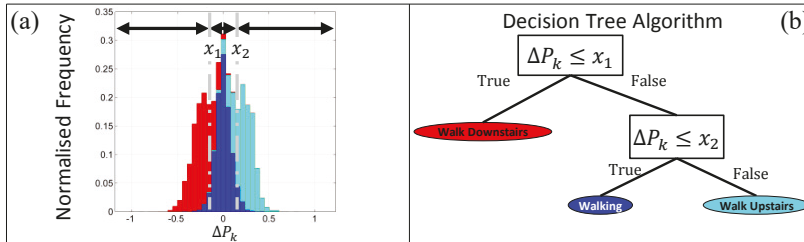


Figure 5. (a) Normalized frequency histogram for the ΔP_k feature and three activity classes (walk downstairs (in red), walking (in blue), and walk upstairs (in light blue), visualized as stacked bars); (b) An example of the decision tree, used at each node of the HMHA; i.e., the blue rectangles in Figure 4. x_1 and x_2 , are derived from Figure 5a according to the classification and regression tree [73].

6. Models and Performance Metrics

6.1. Performance at a High Sampling Rate

A number of models were developed in which a model for recognizing human activity was trained using all of the sensor data collected from the younger and/or older cohorts using either: (i) the Original Features, i.e., features (1)–(4) in Table 2; (ii) the New Features (i.e., features (5)–(8) in Table 2), and; (iii) four pairs of features (i.e., features 1 and 5; features 2 and 6; features 3 and 7; features 4 and 8 from Table 2) were provided to four separate instances of the CART algorithm to select the four Best Features to separate the human activities into distinct classes according to the structure of the HMHA described in Figure 4a; these pairings represent features which are similar in terms of the information they captures, e.g., features 1 and 5 capture angular velocity information in subtly different ways.

The robustness of each model for HAR was evaluated by virtually re-orienting the device (as described in Section 3.1) to obtain data from the younger and/or older cohorts that are representative of five different device orientations (see Figure 1b–f). Each model’s performance was evaluated by training the model with either of the younger and/or older cohorts data and testing the model with either of the younger and/or older cohorts data after it had been virtually re-oriented,

using 10-fold cross-validation. Ninety-five percent confidence intervals (95% CIs) were calculated for the: Cohen’s kappa (κ) and total classification sensitivity (%), as well as the sensitivity (%) and specificity (%) of each activity class. This process is repeated for the ‘Best Features’ (determined in the above search procedure).

6.2. Translating Performance to Different Sampling Rates

Finally, the HMHA are evaluated by training the model with data from the younger and/or older cohort at either (i) the original sampling rate (i.e., the IMU sampled at 100 Hz, and the barometric altimeter data sampled at 16 Hz), or (ii) a reduced sampling rate (i.e., the IMU resampled at 40 Hz, and the barometric altimeter data resampled at 20 Hz), and testing the model with the virtually rotated data from the younger and/or older cohort at the reciprocal sampling rate to determine if the performance and thresholds of the model are consistent. The metrics reported in Section 6.1 were also used to evaluate the model’s performance.

7. Results and Discussion

7.1. Comparing Features Using Shannon Entropy

When the Shannon entropy [77] of the training datasets (i.e., Figure 6a,c) or testing datasets (i.e., Figure 6b,d) were compared for the features ${}^s\bar{\omega}_{xy,k}^2$ and $\bar{\omega}_{bpf,k}^2$, two things become evident. Firstly, both features appear to be orientation invariant because the Shannon entropy is constant whether or not it is calculated from the training data or test data (i.e., here the test data was the virtually re-oriented training data). Secondly, the Shannon entropy was reduced by 0.085 bits when the quaternion-derived feature was used in place of the original feature proposed in our previous work, showing an improvement in the separation of the class distributions [56].

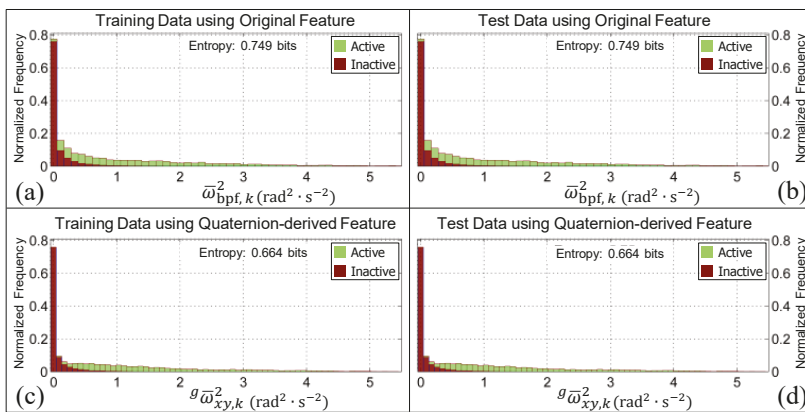


Figure 6. The normalized histograms of the active (i.e., the walking, walking upstairs, walking downstairs, and postural transition classes pooled together) and inactive (i.e., the standing and sedentary classes pooled together) classes for the features $\bar{\omega}_{bpf,k}^2$ and ${}^s\bar{\omega}_{xy,k}^2$. Panels (a,c) are generated from the training data (i.e., the pooled data from the younger and older cohorts, respectively). Panels (b,d) are generated from the test data (i.e., the pooled data from the younger and older cohort after they have been virtually re-oriented using the quaternions in Figure 1b–f). The bar charts in all panels are ‘stacked’. Note that the Shannon entropy for the quaternion-derived feature is both smaller and consistent, irrespective of the data it is calculated from, which suggests that it will be better at distinguishing the activity classes and is orientation invariant.

The features $\bar{\Theta}_{\text{tilt},k}$ and $\bar{\vartheta}_{\text{tilt},k}$ can be used to further separate the inactive class into standing and sedentary (i.e., sitting or lying) classes. When the Shannon entropy was calculated for both tilt angle features using the training data (see Figure 7a,c, respectively) the Shannon entropy [77] dropped by 0.142 bits when $\bar{\vartheta}_{\text{tilt},k}$ was used in place of $\bar{\Theta}_{\text{tilt},k}$. A more pronounced difference of 0.669 bits was observed between $\bar{\Theta}_{\text{tilt},k}$ and $\bar{\vartheta}_{\text{tilt},k}$ when the Shannon entropy was calculated using the test data (see Figure 7b,d, respectively). Whilst the Shannon entropy of $\bar{\vartheta}_{\text{tilt},k}$ increases by 0.291 bits when the test data are used in place of the training data, the shape of the normalized frequency distribution is more consistent for all device re-orientations when compared with $\bar{\Theta}_{\text{tilt},k}$ which increased by 0.818 bits for the re-oriented (test) data. This suggests that the quaternion-derived feature, $\bar{\vartheta}_{\text{tilt},k}$ is more robust to how a smartphone is initially placed in the pocket.

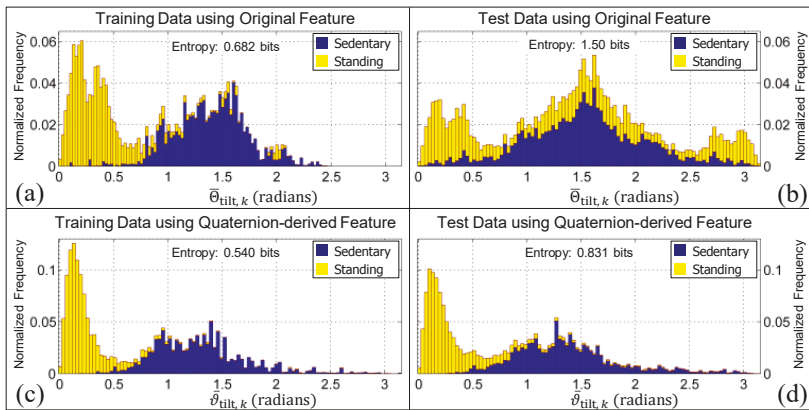


Figure 7. The normalized histograms of the sedentary and standing classes (illustrated in Figure 4) for the features $\bar{\Theta}_{\text{tilt},k}$ and $\bar{\vartheta}_{\text{tilt},k}$. Panels (a,c) are the histograms obtained when the training data are used (i.e., the pooled data from the younger and older cohorts, respectively). Panels (b,d) are the histograms obtained when the test data are used (i.e., the pooled data from the younger and older cohort after it had been virtually re-oriented using each of the quaternions in Figure 1b–f). The bar charts in all panels are ‘stacked’. Note that the entropy for the quaternion-derived feature is consistently smaller, which suggests that it will be better at distinguishing between activity classes.

Whilst the new feature, $\bar{\vartheta}_{\text{tilt},k}$, appears to improve the recognition rate of both standing and sedentary periods of activity, using the change in the shortest rotation between the upward and average orientations, $\Delta\bar{\vartheta}_{\text{tilt},k}$ to distinguish between postural transitions and periods of walking (i.e., walking upstairs, walking downstairs, or walking on a level surface) does not. This is evident by the increase in Shannon entropy (when $\bar{\vartheta}_{\text{tilt},k}$ is compared to $\Delta\bar{\vartheta}_{\text{tilt},k}$) whether or not the feature values are generated from the training or testing data, i.e., from 0.463 bits to 0.866 bits, or from 0.463 bits to 1.049 bits, respectively (see Figure 8). Additionally, since the Shannon entropy of $\bar{a}_{\text{pfdif},k}^2$ remains constant at 0.463 bits irrespective of the dataset used, it confirms that the feature previously described [56] is orientation invariant, as expected.

Conversely, both the average differential pressure, $\Delta P_{k'}$, and the velocity in the vertical direction of the estimated GFR, $\bar{v}_{z,k'}$ are orientation invariant as evident by the Shannon entropy which remains constant whether or not the training or testing data are used, for both the original feature (Figure 9a,b) and the quaternion-derived feature (Figure 9c,d). The Shannon entropy of ΔP_k (1.795 bits) is substantially smaller than $\bar{v}_{z,k}$ (4.070 bits) which suggests that the estimated velocity in the vertical direction (obtained by fusing vertical acceleration and barometric pressure using an extended Kalman filter) of the estimated GFR is not as useful in distinguishing between walking on flat or

inclined surfaces when compared to using the rate of change of pressure measured by the barometric altimeter alone.

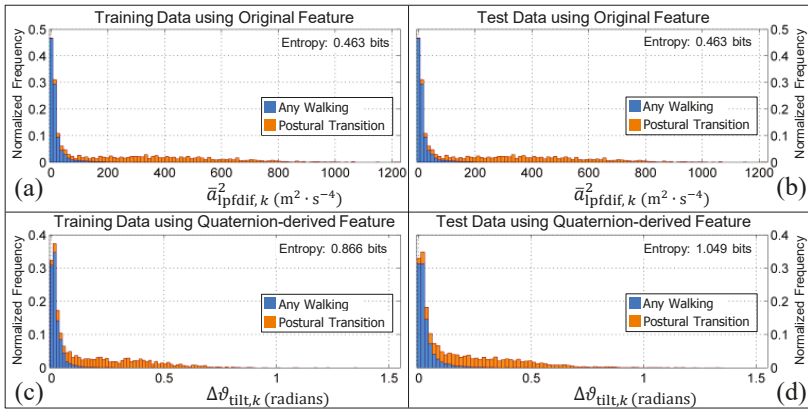


Figure 8. The normalized histograms of the walking (including walking up or down) and postural transition classes (illustrated in Figure 4) for the features $\bar{a}_{\text{pfdif},k}^2$ and $\Delta\theta_{\text{tilt},k}$. Panels (a,c) are the histograms obtained when the training data are used (i.e., the data from the younger and older cohorts, respectively). Panels (b,d) are the histograms obtained when the test data are used (i.e., the data from the younger and older cohort after it had been re-oriented with each of the quaternions in Figure 1b–f). The bar charts in all panels are ‘stacked’. Note that the Shannon entropy for the original feature is consistently lower than the quaternion-derived feature which suggests that it will be better at distinguishing the activity classes.

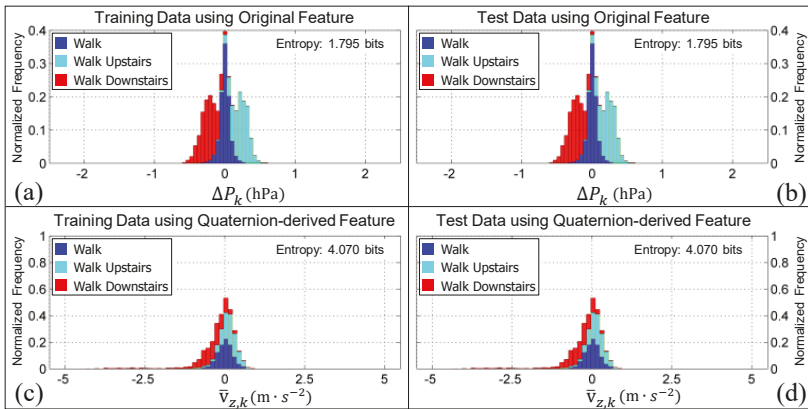


Figure 9. The normalized histograms of the walk, walking upstairs, and walking downstairs classes for the features ΔP_k and $\bar{v}_{z,k}$. Panels (a,c) are the histograms obtained when the training data are used (i.e., the data from the younger and older cohorts, respectively). Panels (b,d) are the histograms obtained when the test data are used (i.e., the data from the younger and older cohort after it had been re-oriented with each of the quaternions in Figure 1b–f). The bar charts in all panels are ‘stacked’. Note how the Shannon entropy of each feature remains constant whether it is computed from the training or test data (which, remember, is the training data re-oriented), confirming that both features are invariant to the initial orientation, as expected.

Although speculative, it is likely that large amplitude accelerations measured by the IMU in the pants pocket during walking is masking the subtle pattern changes in vertical acceleration associated with ascending/descending stairs. It is plausible that if the accelerometer had been placed in a chest pocket, an improved estimate of vertical acceleration may have been obtained by the Kalman filter.

7.2. Comparing the Overall Performance of Models for HAR

When the HMHA is trained and tested with data from the cohort of younger adults (at the original sampling rate: $f_{\text{IMU}} = 100$ Hz; $f_{\text{BAR}} = 16$ Hz), or trained with the older cohort, and tested on the data from the younger cohort after it has been re-oriented, the performance improvement of the models (i.e., the 95% confidence interval of the Cohen’s kappa, κ_{CI95} , and total class sensitivity, q_{CI95}) are negligible. For two out of the three remaining scenarios, there are substantial improvements in the model’s performance when the quaternion-derived features developed herein (i.e., ${}^s\bar{\omega}_{xy,k}^2$ and $\bar{\theta}_{\text{tilt},k}$) are incorporated into the process of human activity recognition. This can be observed in Table 3 when the model is trained with data from the older cohort and tested with the data from the older cohort after it has been re-oriented (i.e., κ_{CI95} increases from [0.685, 0.697] to [0.721, 0.733]; q_{CI95} increases from [77.6%, 78.5%] to [79.9%, 80.7%]), as well as when the model is trained with data from both cohorts and tested with the data from both cohorts after it has been re-oriented (i.e., κ_{CI95} increases from [0.702, 0.713] to [0.732, 0.742]; q_{CI95} increases from [78.4%, 79.2%] to [80.3%, 81.1%]).

Table 3. Ninety-five percent confidence intervals for the Cohen’s Kappa and total class sensitivity when hierarchical models for human activity recognition were developed with different features.

Cohen’s Kappa (κ_{CI95})					
Train	Test	Original Features	New Features	Best Features	Best Features [†]
Y	Y*	[0.820, 0.837]	[0.606, 0.629]	[0.827, 0.844]	[0.824, 0.841]
O	O*	[0.685, 0.697]	[0.478, 0.491]	[0.721, 0.733]	[0.730, 0.741]
Y	O*	[0.675, 0.688]	[0.468, 0.481]	[0.782, 0.793]	[0.779, 0.790]
O	Y*	[0.760, 0.780]	[0.590, 0.613]	[0.765, 0.784]	[0.761, 0.781]
Y&O	(Y&O)*	[0.702, 0.713]	[0.544, 0.556]	[0.732, 0.742]	[0.778, 0.787]
Total Class Sensitivity ($q_{\text{CI95} \%$)					
Train	Test	Original Features	New Features	Best Features	Best Features [†]
Y	Y*	[86.7, 88.0]	[69.0, 70.8]	[87.2, 88.5]	[87.0, 88.3]
O	O*	[77.6, 78.5]	[57.3, 58.4]	[79.9, 80.7]	[80.6, 81.5]
Y	O*	[77.7, 78.6]	[56.3, 57.3]	[84.9, 85.6]	[84.8, 85.5]
O	Y*	[82.0, 83.5]	[67.5, 69.3]	[82.3, 83.7]	[82.0, 83.5]
Y&O	(Y&O)*	[78.4, 79.2]	[63.9, 64.8]	[80.3, 81.1]	[84.1, 84.8]

* Test data were obtained by re-orienting the data from the younger (Y) and/or older (O) cohort (see Section 3.1); [†] IMU data were re-sampled at 40 Hz, barometer data were re-sampled at 20 Hz.

It is particularly noteworthy that the performance of the model trained with the ‘Best Features’ using the data collected from the younger cohort and tested with the ‘best features’ using the data collected from the older adults after it has been re-oriented (i.e., $\kappa_{\text{CI95}} = [0.782, 0.793]$; $q_{\text{CI95}} = [84.9\%, 85.6\%]$) is comparable to the performance of the model trained with the data collected from the older cohort and tested with the data from the younger cohort (i.e., $\kappa_{\text{CI95}} = [0.765, 0.784]$; $q_{\text{CI95}} = [82.3\%, 83.7\%]$). This contradicts the finding of our previous work [56] in which the performance of a model for HAR trained on younger cohorts degraded substantially when tested on older cohorts (due to the use of the tilt angle feature, $\bar{\Theta}_{\text{tilt},k}$, which was not orientation invariant (see Table 3, the column labeled ‘Original Features’)), compared to the opposite scenario in which the model is trained with the older cohort’s data and tested with the data from the younger cohort, which gives the better performance.

The improvements in total classification sensitivity and Cohen’s kappa gained by incorporating the quaternion-derived features (see Table 3 column labeled ‘Best Features’) persist when the data from the IMU are re-sampled at a reduced rate (see the column labeled ‘Best Features †’ in Table 3). This demonstrates the robustness of both the features and the HMHA to a decrease in the sample rate, which is an important design consideration given the limited battery life of wearable sensors (i.e., smartphones, smartwatches, etc.). Interestingly, there are marginal improvements in Cohen’s kappa (i.e., from $\kappa = [0.732, 0.742]$ when $f_{IMU} = 100$ Hz to $\kappa = [0.778, 0.787]$ when $f_{IMU} = 40$ Hz) and total class sensitivity (i.e., from $q_{CI95} = [80.3\%, 81.1\%]$ when $f_{IMU} = 100$ Hz to $q_{CI95} = [84.1\%, 84.8\%]$ when $f_{IMU} = 40$ Hz) when the model is trained with the data from the younger and older cohorts and tested with the data from the younger and older cohorts after it has been re-oriented. Upon analyzing the class sensitivity of these two hierarchical models of human activity, it is evident that this is primarily due to an increase in the sensitivity of detecting the walking class, from $\sim 72\%$ to $\sim 82\%$ (see Figure 10xv,xx). This improvement can be attributed to the use of the quaternion derived feature, ${}^s\tilde{\omega}_{xy,k}^2$ which measures the amount of pitch/roll rotation in the estimated GFR, a more consistent frame of reference than the local sensor frame.

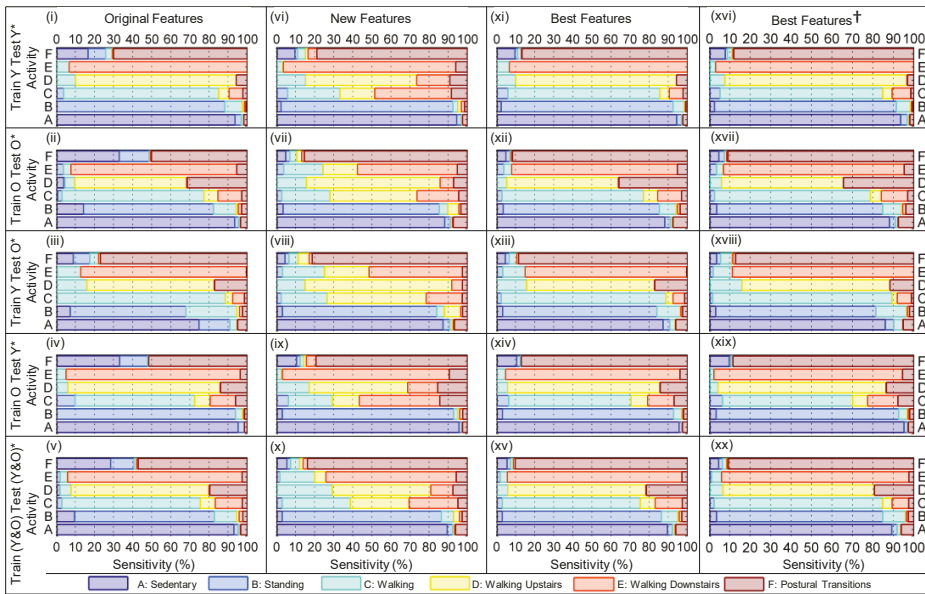


Figure 10. The column titled: ‘Original Features’ (i.e., panels (i–v)) correspond to hierarchical models of human activity that were trained and tested with the features developed in our previous work [56] using the hierarchical model of human activity (HMHA) illustrated in Figure 4b; ‘New Features’ (i.e., panels (vi–x)) correspond to hierarchical models of human activity that were trained and tested with the features developed herein using the HMHA illustrated in Figure 4c; ‘Best Features’ (i.e., panels (xi–xv)) correspond to hierarchical models of human activity that were trained and tested with a combination of the original and New Features using the HMHA illustrated in Figure 4d; ‘Best Features†’ (i.e., panels (xvi–xx)) equivalent to ‘Best Features’ with the IMU data re-sampled to 40 Hz, the barometer data to 20 Hz.

7.3. Identifying Which Features Drive Model Performance

When Figures 10i,xi are compared (i.e., a HMHA trained with the Original Features extracted from the younger cohort and a HMHA trained with the Best Features extracted from the younger cohort), the differences in the model’s performance become apparent. Most notably, the sensitivity for the

postural transition class increased from 70.05% to 86.64%. This improved performance is a by-product of modest increases in the model's ability to identify standing (i.e., the standing class sensitivity increased from 88.29% to 89.98%) and sedentary periods of activity (i.e., the class sensitivity increased from 93.67% to 94.42%). These pieces of evidence support the argument that the quaternion-derived feature, ${}^s\bar{\omega}_{xy,k}^2$, is better at distinguishing periods of activity (i.e., walking, or postural transitions) from periods of inactivity (i.e., standing or sedentary), a trend which is consistent across each of the five training and testing scenarios proposed in Section 6. However, it is likely that ${}^s\bar{\omega}_{xy,k}^2$ would not be as effective in this task if the smartphone is placed in the user's chest pocket, which pitches and rolls less when compared to the thigh (i.e., it rotates less about the x and y axis of the estimated GFR) during walking.

The underlying causes for the improvement in the HMHA become clear after analyzing the sensitivity of the activity classes listed in Figure 10. When the columns corresponding to the HMHAs built using the Original Features and New Features (e.g., when Figure 10i is compared to Figure 10vi, Figure 10ii to Figure 10vii, and so on) were compared, it is evident that the model's sensitivity to periods of walking upstairs decreased dramatically (e.g., in the case of (i) and (v), from 84.21% to 58.16%) when the differential pressure, ΔP_k , was replaced with the moving average velocity in the vertical direction of the estimated GFR, $\bar{v}_{z,k}$. This persisted whether the data from the younger or older cohort was used (i.e., when the columns entitled 'Original Features' and 'New Features' of Figure 10 are compared, the model's sensitivity to periods of walking either upstairs or downstairs is reduced). ΔP_k is superior to $\bar{v}_{z,k}$ in estimating vertical velocity and hence walking on stairs.

On the other hand, when $\bar{\Theta}_{\text{tilt},k}$ was substituted with $\bar{\delta}_{\text{tilt},k}$, the sensitivity of the model to standing classes increases (from 60–70% to >80%) when the HMHA was trained with: (a) the older cohort's data and tested with the older cohort's data after it had been re-oriented (the second row in Figure 10); (b) the younger cohort's data and tested with the older cohort's data after it had been re-oriented (the fourth row in Figure 10); (c) the data from both cohorts and tested with the data from both cohorts after it had been re-oriented. This improvement underscores the utility of learning the orientation of the device when the body is definitely upright (i.e., when walking), demonstrating how this method can intuitively account for the variability in sensor measurements which may arise due to inconsistent device orientation when the IMU is placed on the body.

In addition, the rate of misclassification of sedentary and stationary periods of activity as postural transitions decreases markedly. This phenomena is consistent across the five scenarios evaluated (recall Section 6). When the columns labeled 'Original Features' and 'Best Features' are compared row by row, periods of standing that were originally classified as postural transitions are all but eliminated (e.g., compare Figure 10i and Figure 10xi), whilst the misclassification rate of sedentary activity as postural transitions decreased from ~16% to ~9% (compare Figure 10i and Figure 10xi); ~32% to ~5% (compare Figure 10ii and Figure 10xii); ~9% to ~5% (compare Figure 10iii and Figure 10xiii); ~33% to ~10% (compare Figure 10iv and Figure 10xiv); ~29% to ~7% (compare Figure 10v and Figure 10xv).

7.4. Comparing Model Performance at Different Sampling Rates

Due to the limited battery life of smartphones, it is becoming increasingly important that algorithms for human activity recognition are able to operate at a reduced sampling rate without suffering a degradation in classification accuracy. Consequently, the robustness of the model's developed with the 'Best Features' were evaluated by training the model with the data collected from the younger and/or older cohort at 100 Hz, and testing the model's performance with data from the younger and/or older cohort at 40 Hz after it had been virtually re-oriented (and vice versa). From Table 4 it is evident that both the Cohen's kappa, and total class sensitivity of the HMHA proposed in Figure 4d remain consistent (i.e., the 95% confidence intervals overlap for almost all of the training and testing combinations evaluated) whether or not the HMHA is trained with the data at 100 Hz (i.e., the higher sampling rate) and tested with the re-oriented data at 40 Hz (i.e., the reduced

sampling rate), or the reciprocal scenario in which the HMHA is trained with the data at 40 Hz and tested with the re-oriented data at 100 Hz.

Table 4. Ninety-five percent confidence intervals for the Cohen’s Kappa and total class sensitivity (%) when a hierarchical model of human activity (HMHA) was developed with the Best Features at different sampling rates.

Train		Cohen’s Kappa (κ_{CI95})			Total Class Sensitivity (Q_{CI95} %)		
		100 Hz	100 Hz	40 Hz †	100 Hz	100 Hz	40 Hz †
	Test *	100 Hz	40 Hz †	100 Hz	100 Hz	40 Hz †	100 Hz
Y	Y *	[0.827, 0.844]	[0.819, 0.837]	[0.823, 0.841]	[87.2, 88.5]	[86.6, 87.9]	[87.0, 88.3]
O	O *	[0.721, 0.733]	[0.720, 0.732]	[0.728, 0.740]	[79.9, 80.7]	[79.8, 80.6]	[80.5, 81.3]
Y	O *	[0.782, 0.793]	[0.786, 0.796]	[0.777, 0.788]	[84.9, 85.6]	[85.2, 85.9]	[84.6, 85.4]
O	Y *	[0.765, 0.784]	[0.752, 0.772]	[0.769, 0.789]	[82.3, 83.7]	[81.3, 82.8]	[82.6, 84.1]
Y&O	(Y&O) *	[0.732, 0.742]	[0.726, 0.736]	[0.776, 0.786]	[80.3, 81.1]	[79.9, 80.6]	[84.0, 84.7]

* Test data were obtained by re-orienting the data from the younger (Y) and/or older (O) cohort (see Section 3.1); † IMU data were re-sampled at 40 Hz, barometer data were re-sampled at 20 Hz.

The sole exception to this trend is the scenario in which the data re-sampled at 40 Hz from both the younger and older cohorts are used to train the HMHA, whilst the data sampled at 100 Hz from the younger and older cohorts after it has been re-oriented are used to test the HMHA. In this particular scenario, the ninety-five percent confidence interval of the Cohen’s kappa, κ , increased by ~ 0.04 from $\kappa_{CI95} = [0.732, 0.742]$ to $\kappa_{CI95} = [0.776, 0.786]$. Similarly, the ninety-five percent confidence interval of the total class sensitivity increased by $\sim 3\%$ from $Q_{CI95} = [80.3\%, 81.1\%]$ to $Q_{CI95} = [84.0\%, 84.7\%]$ (see the bottom row of Table 4).

After analyzing Figure 11 it is evident that the model’s sensitivity to each activity remains relatively consistent as long as only one of the cohort’s data is used to train the model, and the other cohort’s data is used to test the model, irrespective of the sampling rate (i.e., when Figure 11iii,viii,xiii are compared; Figure 11iv,ix,xiv are compared, and so on). When both cohort’s data are used (i.e., when Figure 11v,x,xv are compared), the sensitivity of the model to the sedentary, standing, and postural transition classes is remarkably consistent whilst the sensitivity of the model to the three different walking, classes varies (whether or not the HMHA is trained with the data re-sampled at 40 Hz or the data sampled at 100 Hz). From Table 5 it is easy to see that this robustness in performance can be attributed to the relatively constant threshold values of ${}^s\hat{\omega}_{xy,k}^2$ and $\hat{\theta}_{\text{tilt},k}$ which change by < 0.1 ($\text{rad}^2 \cdot \text{s}^{-2}$ and radians, respectively), suggesting that these features are robust to both the variation in sampling rate and the cohort from which the threshold is extracted (i.e., the threshold changes little whether trained on the younger and/or older cohort’s data).

Interestingly, the recognition rate of the postural transition class also remains fairly consistent (i.e., between $\sim 89\text{--}91\%$), irrespective of the data which are used to train and test the HMHA. This suggests that $\hat{a}_{\text{pfdif},k}^2$ is also robust to variations in the sampling rate of the IMU and the cohort from which the threshold are determined (see Table 5).

In the case of the walking, walking upstairs, and walking downstairs classes, the differences were negligible when the HMHA was trained with the data sampled at 100 Hz and tested with the same data after it had been re-oriented; or trained with the data at 100 Hz and tested with the data re-sampled at 40 Hz after it had been re-oriented. However, when the model was trained with the data re-sampled at 40 Hz and tested with the data at 100 Hz there were slight changes in the class sensitivity when compared to either of the two previously mentioned scenarios.

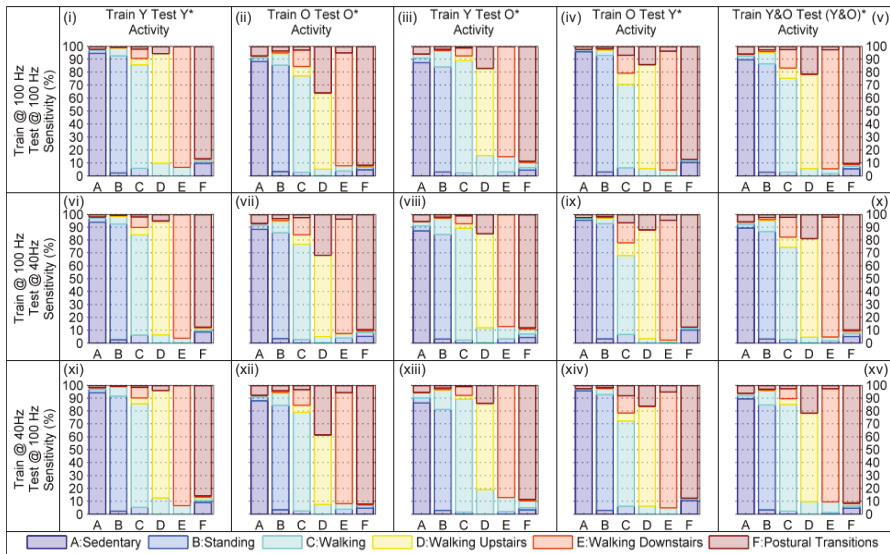


Figure 11. Class Sensitivity for a Hierarchical Model of Human Activity Recognition using the Best Features: (i–v) trained with the IMU data at 100 Hz and tested with the IMU data at 100 Hz after it has been re-oriented; (vi–x) trained with the IMU data at 100 Hz and tested with the IMU data at 40 Hz after it has been re-oriented; (xi–xv) trained with the IMU data at 40 Hz and tested with the IMU data at 100 Hz after it has been re-oriented.

In particular, the model’s sensitivity to the walking class increased from $\sim 71\text{--}72\%$ to $\sim 82\%$ (trained with data at 40 Hz, tested with data at 100 Hz) due to marked reductions in periods of walking upstairs and walking downstairs being incorrectly identified as walking on a level surface (i.e., from 7.95% to 4.5% and 15.32% to 7.82%, respectively). Similarly, the sensitivity of the walking upstairs class decreased from $\sim 76\%$ to $\sim 68\%$ (see Appendix C, the row labeled ‘Train Y&O Test (Y&O)*’ in Table A2) due to the increased misclassification of periods of walking upstairs as periods of walking on a level surface (i.e., from 4–5% to $\sim 10\%$; see Figure 11 and compare panels (x) and (xv)). This suggests that the smaller threshold of $\Delta P_k = 0.092 \text{ hPa}\cdot\text{s}^{-1}$ (see Table 5) is better (when compared to $\Delta P_k = 0.119 \text{ hPa}\cdot\text{s}^{-1}$) at distinguishing between periods of walking on a level surface versus walking upstairs.

This trend was mirrored in the reduction of the hierarchical model’s sensitivity to the walking downstairs class; i.e., decreasing from $\sim 91\text{--}92\%$ to $\sim 87\%$ (see Appendix C, the row labeled ‘Train Y&O Test (Y&O)*’ in Table A2) due to the increased misclassification of periods of walking downstairs as periods of walking on a level surface (i.e., from 3–4% to $\sim 8\%$; see Figure 11 and compare panels (x) and (xv)). Again, this suggests that the threshold of $\Delta P_k = -0.062 \text{ hPa}\cdot\text{s}^{-1}$ (see Table 5) is better than the threshold of $\Delta P_k = -0.094 \text{ hPa}\cdot\text{s}^{-1}$ (a change of $\sim 51\%$ in the threshold value) at distinguishing between periods of walking on a level surface versus walking downstairs.

Table 5. Comparison of thresholds when HMHA recognition were developed with the Best Features at different sampling rates.

Training Data	Feature	Threshold		Rule *
		100 Hz	40 Hz	
Y O Y&O	$s\bar{\omega}_{xy,k}^2$ ($\text{rad}^2 \cdot \text{s}^{-2}$)	0.232 0.249 0.230	0.195 0.237 0.202	Inactive [†] if $s\bar{\omega}_{xy,k}^2 \leq \text{threshold}$, else Active [‡]
Y O Y&O	$\bar{\theta}_{\text{tilt},k}$ (radians)	0.668 0.616 0.640	0.689 0.617 0.617	Standing if $\bar{\theta}_{\text{tilt},k} \leq \text{threshold}$, else Sedentary (i.e., sitting or lying)
Y O Y&O	$\bar{a}_{\text{ipfdif},k}^2$ ($\text{m}^2 \cdot \text{s}^{-4}$)	75.6 36.4 46.7	93.4 32.8 46.4	Any Walking [§] if $\bar{a}_{\text{ipfdif},k}^2 \leq \text{threshold}$, else Postural Transition
Y O Y&O	ΔP_k ($\text{hPa} \cdot \text{s}^{-1}$)	-0.107 -0.067 -0.062	-0.101 -0.068 -0.094	Walking Downstairs if $\Delta P_k \leq \text{threshold}$, else Walking
Y O Y&O	ΔP_k ($\text{hPa} \cdot \text{s}^{-1}$)	0.128 0.092 0.092	0.142 0.105 0.119	Walking if $\Delta P_k \leq \text{threshold}$, else Walking Upstairs

Inactive[†] — any of the standing or sedentary (sitting/lying) classes; active[‡] — any of the walking, walking upstairs, walking downstairs, or postural transition classes; any walking[§] — either of the walking, walking upstairs, or walking downstairs classes. * Each rule corresponds to a node of the HMHA illustrated in Figure 4d.

7.5. Comparison to the State-of-the-Art

In order to draw a fair comparison with other published work that is representative of state-of-the-art methods, the scope of these comparisons is limited to reports which only utilized the smartphone's internal sensing components to classify human activity. With this in mind, the state-of-the-art deep learning methods (recall Section 1.3.3) proposed by Ordoñez et al. [45] and Li et al. [78] are excluded because they utilize measurements from multiple IMUs that are placed at different anatomical locations on the body, whilst the works of Ravi et al. [43] and Ronao and Cho et al. [42] are included. Similarly, the 'feature engineering and classification'-based approaches (recall Section 1.3.2) developed by Bao and Intille et al. [79] are omitted, whilst the works of Anguita et al. [80] and Shoaib et al. [29] are included.

Anguita et al. developed a hardware-friendly multi-class support vector machine which processed the accelerometer and gyroscope data (at 50 Hz) from a waist-worn smartphone (i.e., attached to a belt worn about their waist) to identify activities of daily living in a cohort of thirty participants aged between nineteen and 48 years. From these six channels, they extracted 561 spatial or spectral features (every 1.25 seconds using 50% overlapping windows) to identify six activities with a sensitivity between 72% and 96%: walking (95.6%), walking upstairs (72.1%), walking downstairs (79.7%), standing (92.2%), sitting (96.4%), and lying (100%) [80].

Shoaib et al. evaluated the utility of a smartphone's internal sensors for the purposes of human activity recognition. They studied ten male participants, aged between 25 and 30 years, whilst a smartphone was firmly fixed to their body with a strap at one of five positions on their body (right and left front jeans pocket, on a belt near the right hip, right wrist, right upper arm). A smartphone application recorded the accelerometer, gyroscope, and magnetometer data at 50 Hz whilst each participant performed six activities of daily living (walking, jogging, sitting, standing, biking, walking upstairs, and walking downstairs) [29]. When features were extracted every two seconds (with 50% overlapping windows), the gyroscope-based features proved most effective in identifying periods of walking upstairs and walking downstairs (particularly when the sensor was placed in the jeans pocket or on the belt), whilst features from the magnetometer should only be used if they are independent of

heading. Moreover, they advocate against ‘blindly combining different sensors’, suggesting a more manual approach to system and feature design.

Deep learning approaches attempt to tease out more subtle differences, imperceptible by human observation, in wearable sensor data which can be used for the purposes of HAR. Ronao and Cho, recruited 30 participants (age range not disclosed) to evaluate the performance of a model for HAR, based on deep convolutional neural networks (convnet). The smartphone was placed in a pocket of the participants’ clothing (location on body not disclosed), whilst data from the accelerometer and gyroscope were recorded at 50 Hz [42]. When the data were segmented in 2.5 second intervals with 50% overlap, the convnet could identify six activities: walking (98.99%), walking upstairs (100.00%), walking downstairs (100.00%), standing (93.23%), sitting (88.88%), lying (87.71%); with an overall sensitivity of 94.79%. Before the accelerometer and gyroscope data could be processed, each channel (of six) needed to be normalized by subtracting the mean of each signal, and dividing each channel by the channel’s standard deviation. At this point, 2.5 second data segments are input to a five-layer convnet comprised of three convolutional/pooling layers (with 96, 192, 192 neurons in each layer, respectively), a fully connected layer comprised of 1000 neurons, and a softmax classification layer with six neurons.

Ravi et al. combined features extracted from the spectrogram (i.e., the short-time Fourier transform coefficients) of accelerometer and gyroscope signals (both of which sampled at either 50 Hz or 200 Hz, respectively) with a three layered network comprised of a temporal convolution layer (15 filters, 80 nodes), fully-connected layer, and soft-max classification layer for the purposes of HAR. Data was obtained from ten subjects (using five different smartphones) who were allowed to place the phone anywhere on their body (or in their hand/bag) whilst they performed six activities of daily living. The total class sensitivity of their model for HAR was 95.7% with class sensitivities of ~95% (running), ~95% (walking), ~96% (cycling), ~96% (casual movement), ~96% (public transport), ~98% (idle), and ~74% (standing). Whilst the features derived from the six channel spectrogram enabled highly-variable activities to be distinguished from repetitive activities, the absence of time-domain-based features limited the model’s ability to infer the user’s postural orientation, which was further limited by the fact that the phone could be placed at various parts of the body, in the hand, or in a bag [43].

The model for HAR constructed by Gu et al. [81] implemented denoising autoencoders (two layers, 1000 neurons per layer) combined with a softmax classification layer to automate the HAR process. Features were extracted from two-second intervals of data from the smartphone’s accelerometer, gyroscope, magnetometer, and barometer (all of which were sampled at 64 Hz, except for the barometer which was sampled at 32 Hz). Twelve participants (six male, six female) aged between twenty-five and thirty-five years were recruited to train the model to recognize eight activities of daily living. When the data from all four sensors were used by the denoising autoencoders (corrupting noise level = 0.5, learning rate = 1×10^{-3} , weight of sparsity penalty term = 1), the F-measure is 94.04% and the class sensitivity for the eight activities are: stationary (~98%), walking (~92%), stationary but using the phone (~96%), running (~97%), walking upstairs (~94%), walking downstairs (~93%), elevator up (~84%), elevator down (~87%).

A general pitfall of all of the above deep learning approaches is that this approach does not inherently allow the training of the neural network to be constrained by the domain knowledge that the smartphone could be placed anywhere on the body and with any orientation. For deep learning approaches, some safeguards against obtaining a classifier model which is not robust to such variability in smartphone placement and orientation involves collecting large datasets which capture this variability, or to perform preprocessing of the smartphone signals to generate features which are tolerant to such variability; the latter somewhat goes against the spirit of the deep learning approach.

8. Limitations

There are limitations with the study presented herein which need to be acknowledged. The model for HAR developed is dependent on the wearable sensor (i.e., device containing an IMU and barometer, such as a smartphone) remaining in the pants pocket throughout the day, which is not a realistic expectation since the individual's lower body garments may not always have a suitable pocket, or a pocket large enough to place the wearable sensor. If the wearable sensor is strapped to the thigh, the quaternion-derived feature, $\bar{\theta}_{\text{tilt},k}$, should always be able to separate standing and sedentary periods. If the device is sporadically removed from the pants pocket whilst the person is moving, it is conceivable that the walking detector (Equation (13)) could 'learn' an incorrect upright orientation, $q_{\text{upright},k'}$ thereby reducing the accuracy of the model for HAR until it relearns the correct upright orientation from the next 2.5 s of true walking data; robustness to this scenario will be evaluated in future work.

9. Conclusions and Future Work

This paper developed a model for HAR capable of recognizing six human activities (standing, sedentary, walk, walk upstairs, walk downstairs, as well as postural transitions between the standing and sedentary classes), regardless of the smartphone's orientation in the pants pocket by using a quaternion-based complementary filter [63] to estimate the device's orientation, thereby enabling sensor measurements to be expressed in a consistent frame of reference (the world/global frame). Four New Features were developed, and two were shown to be useful in the classification of human activities, namely ${}^S\hat{\omega}_{xy,k}^2$ which utilized an estimate of the IMU's orientation to determine the magnitude of the pitch/roll angular velocity, and $\bar{\theta}_{\text{tilt},k'}$ which measured the angle between the recent average orientation and the estimated upright orientation; upright orientation was estimated as the average orientation of the IMU when walking was detected. The success of these quaternion-derived features suggest that existing methods for recognizing human activities would benefit from converting all measurements to the global frame of reference where the feature values would be more consistent, especially if the orientation of the IMU with respect to the body is not fixed.

Author Contributions: Conceptualization, M.B.D.R. and S.J.R.; data curation, M.B.D.R.; formal analysis, M.B.D.R.; investigation, M.B.D.R. and S.J.R.; methodology, M.B.D.R., N.H.L. and S.J.R.; project administration, N.H.L. and S.J.R.; resources, N.H.L. and S.J.R.; supervision, N.H.L. and S.J.R.; validation, M.B.D.R.; visualization, M.B.D.R.; writing—original draft, M.B.D.R., N.H.L. and S.J.R.; eriting—review and editing, M.B.D.R., N.H.L. and S.J.R.

Funding: This research was funded by an Australian Research Council Discovery Projects grant (DP130102392).

Acknowledgments: We gratefully acknowledge our colleagues at UNSW, Jingjing Wang and Kejia Wang, and at Neuroscience Research Australia, Stephen Lord, Kim Delbaere, and Matthew Brodie, for their assistance in collecting data for the older cohort.

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

Appendix A. Average of Multiple Quaternions

Gramkow's method was used to calculate the average of N quaternions [82]. Note: the quaternion is normalized after each component of \bar{q} has been calculated (see Equation (A1)). If the scalar component of a quaternion, $q = [q_0 \ q_1 \ q_2 \ q_3]$, in the window was negative (i.e., if $q_0 < 0$) each component of the quaternion was negated (thereby preserving the rotational information since q and $-q$ represent the same rotation [83]) so that each quaternion lies in the same half plane.

$$\bar{q} = f_{q,\text{avg}}(q_1, \dots, q_N) = \left[\frac{\bar{q}_0}{\|\bar{q}\|} \quad \frac{\bar{q}_1}{\|\bar{q}\|} \quad \frac{\bar{q}_2}{\|\bar{q}\|} \quad \frac{\bar{q}_3}{\|\bar{q}\|} \right] \quad (\text{A1})$$

$$\bar{q}_j = \frac{1}{N} \sum_{k=1}^N q_{j,k} \text{ (where } j \in \{0, 1, 2, 3\}); \|\bar{q}\| = \sqrt{\bar{q}_0^2 + \bar{q}_1^2 + \bar{q}_2^2 + \bar{q}_3^2}$$

Appendix B. Shortest Rotation Between Two Quaternions

The rotation that brings two quaternions, q_A and q_B , into coincidence is $q_{AB} = q_A \otimes (q_B)^*$, i.e., $q_B \otimes q_{AB} = q_A$. The shortest angle between these quaternions, $0 \leq \theta \leq \pi$ is given by $q_{AB,0}$, the scalar component of q_{AB} , Equation (A2).

$$\theta = f_{\text{angle}}(q_A, q_B) = \begin{cases} 2 \cos^{-1}(q_{AB,0}), & q_{AB,0} \geq 0 \\ 2 \cos^{-1}(-q_{AB,0}), & q_{AB,0} < 0 \end{cases} \quad (\text{A2})$$

Appendix C. Ninety-five Percent Confidence Intervals for the Class Sensitivity and Class Specificity of the Hierarchical Models of Human Activity

Table A1. Ninety-five percent confidence intervals for the sensitivity and specificity of each activity class when HMHA were developed with different feature subsets ($f_{\text{IMU}} = 100$ Hz; $f_{\text{bar}} = 16$ Hz).

Activity	Sensitivity (%)			Specificity (%)			
	Original Features	New Features	Best Features	Original Features	New Features	Best Features	
Train Y Test Y*	Sedentary	[92.8, 94.5]	[93.6, 95.2]	[93.6, 95.2]	[99.3, 99.6]	[98.4, 98.9]	[98.4, 98.9]
	Standing	[87.1, 89.5]	[88.8, 91.1]	[88.8, 91.1]	[96.4, 97.2]	[96.1, 96.9]	[96.1, 96.9]
	Walking	[79.9, 82.7]	[25.8, 28.9]	[78.2, 81.0]	[94.5, 95.5]	[97.4, 98.1]	[96.0, 96.9]
	Walking Upstairs	[80.5, 87.9]	[53.2, 63.1]	[80.5, 87.9]	[97.7, 98.3]	[93.2, 94.2]	[98.0, 98.5]
	Walking Downstairs	[90.7, 95.9]	[87.2, 93.3]	[90.7, 95.9]	[97.1, 97.7]	[86.6, 87.9]	[97.2, 97.9]
	Postural Transitions	[64.0, 76.1]	[73.4, 84.2]	[82.1, 91.2]	[98.0, 98.6]	[95.3, 96.1]	[97.8, 98.3]
Train O Test O*	Sedentary	[92.9, 93.9]	[87.5, 88.8]	[87.5, 88.8]	[94.2, 94.8]	[98.6, 98.9]	[98.6, 98.9]
	Standing	[66.9, 69.2]	[80.9, 82.7]	[80.9, 82.7]	[97.2, 97.5]	[97.2, 97.5]	[97.2, 97.5]
	Walking	[73.6, 75.0]	[24.5, 25.8]	[73.4, 74.8]	[95.1, 95.7]	[96.6, 97.1]	[95.9, 96.5]
	Walking Upstairs	[54.4, 62.7]	[66.0, 73.8]	[54.2, 62.5]	[96.4, 96.7]	[81.4, 82.2]	[96.4, 96.8]
	Walking Downstairs	[83.9, 89.8]	[47.9, 56.7]	[83.9, 89.8]	[94.0, 94.5]	[90.3, 90.9]	[94.0, 94.5]
	Postural Transitions	[47.2, 53.2]	[83.3, 87.5]	[89.9, 93.3]	[96.3, 96.7]	[94.8, 95.2]	[94.9, 95.3]
Train Y Test O*	Sedentary	[73.9, 75.6]	[86.6, 87.9]	[86.6, 87.9]	[97.5, 97.9]	[98.7, 99.0]	[98.7, 99.0]
	Standing	[59.5, 61.9]	[79.8, 81.8]	[79.8, 81.8]	[93.6, 94.1]	[97.1, 97.5]	[97.1, 97.5]
	Walking	[87.7, 88.7]	[23.4, 24.8]	[85.7, 86.8]	[88.6, 89.5]	[96.5, 97.0]	[93.3, 94.0]
	Walking Upstairs	[62.7, 70.7]	[73.2, 80.3]	[62.9, 70.9]	[97.9, 98.2]	[79.0, 79.8]	[98.1, 98.4]
	Walking Downstairs	[83.9, 89.8]	[44.5, 53.3]	[81.5, 87.8]	[96.7, 97.0]	[91.3, 91.9]	[96.9, 97.2]
	Postural Transitions	[74.4, 79.5]	[78.8, 83.6]	[86.5, 90.4]	[96.9, 97.3]	[96.0, 96.4]	[96.6, 97.0]
Train O Test Y*	Sedentary	[94.5, 96.0]	[94.9, 96.3]	[94.9, 96.3]	[98.7, 99.2]	[98.1, 98.7]	[98.1, 98.7]
	Standing	[92.8, 94.7]	[88.5, 90.8]	[88.5, 90.8]	[93.7, 94.7]	[96.4, 97.2]	[96.4, 97.2]
	Walking	[61.1, 64.5]	[21.5, 24.4]	[62.5, 65.8]	[97.6, 98.3]	[97.6, 98.3]	[97.3, 98.0]
	Walking Upstairs	[75.7, 83.8]	[46.6, 56.6]	[75.7, 83.8]	[96.9, 97.6]	[94.7, 95.6]	[96.7, 97.4]
	Walking Downstairs	[88.5, 94.3]	[84.0, 90.9]	[88.5, 94.3]	[95.2, 96.0]	[86.2, 87.4]	[95.0, 95.8]
	Postural Transitions	[45.0, 58.3]	[73.9, 84.7]	[82.6, 91.6]	[96.2, 97.0]	[92.8, 93.8]	[95.4, 96.2]
Train Y&O Test (Y&O)*	Sedentary	[92.6, 93.4]	[88.9, 90.0]	[88.9, 90.0]	[95.6, 96.1]	[98.7, 98.9]	[98.7, 98.9]
	Standing	[72.2, 74.0]	[82.6, 84.1]	[82.6, 84.1]	[97.2, 97.5]	[97.2, 97.6]	[97.2, 97.6]
	Walking	[71.9, 73.2]	[35.0, 36.4]	[71.6, 72.9]	[94.8, 95.4]	[95.2, 95.7]	[95.8, 96.3]
	Walking Upstairs	[69.3, 75.1]	[48.3, 54.8]	[69.3, 75.1]	[96.3, 96.7]	[87.4, 88.0]	[96.4, 96.7]
	Walking Downstairs	[89.6, 93.3]	[65.1, 71.4]	[89.7, 93.4]	[93.6, 94.0]	[89.2, 89.8]	[93.7, 94.1]
	Postural Transitions	[54.4, 59.9]	[81.7, 85.8]	[88.6, 91.9]	[96.8, 97.1]	[95.1, 95.5]	[95.8, 96.1]

* The test data was obtained by virtually re-orienting the data from the younger (Y) and older (O) cohort as described in Section 3.1.

Table A2. Ninety-five percent confidence intervals for the sensitivity and specificity of each activity when hierarchical models for human activity recognition were developed with the best features.

Activity		Sensitivity (%)			Specificity (%)		
		Train Test	100 Hz 100 Hz	100 Hz 40 Hz [†]	40 Hz [†] 100 Hz	100 Hz 100 Hz	100 Hz 40 Hz [†]
Train Y Test Y [*]	Sedentary	[93.6, 95.2]	[93.0, 94.7]	[93.2, 94.9]	[98.4, 98.9]	[98.4, 98.9]	[98.5, 99.0]
	Standing	[88.8, 91.1]	[88.6, 90.9]	[87.8, 90.2]	[96.1, 96.9]	[95.7, 96.5]	[96.3, 97.1]
	Walking	[78.2, 81.0]	[76.2, 79.1]	[78.8, 81.6]	[96.0, 96.9]	[96.3, 97.2]	[95.3, 96.2]
	Walking Upstairs	[80.5, 87.9]	[84.6, 91.2]	[79.4, 86.9]	[98.0, 98.5]	[97.7, 98.3]	[98.2, 98.7]
	Walking Downstairs	[90.7, 95.9]	[94.1, 98.1]	[90.7, 95.9]	[97.2, 97.9]	[96.9, 97.5]	[96.8, 97.5]
	Postural Transition	[82.1, 91.2]	[83.2, 91.9]	[81.1, 90.4]	[97.8, 98.3]	[97.9, 98.5]	[98.1, 98.6]
Train O Test O [*]	Sedentary	[87.5, 88.8]	[87.7, 88.9]	[87.2, 88.5]	[98.6, 98.9]	[98.5, 98.8]	[98.6, 98.9]
	Standing	[80.9, 82.7]	[81.0, 82.8]	[79.9, 81.8]	[97.2, 97.5]	[97.1, 97.5]	[97.3, 97.7]
	Walking	[73.4, 74.8]	[73.1, 74.5]	[75.6, 77.0]	[95.9, 96.5]	[95.9, 96.4]	[95.7, 96.2]
	Walking Upstairs	[54.2, 62.5]	[58.8, 67.0]	[49.3, 57.7]	[96.4, 96.8]	[96.4, 96.8]	[97.3, 97.6]
	Walking Downstairs	[83.9, 89.8]	[85.8, 91.4]	[82.8, 88.9]	[94.0, 94.5]	[93.7, 94.2]	[94.2, 94.7]
	Postural Transition	[89.9, 93.3]	[87.6, 91.3]	[90.3, 93.6]	[94.9, 95.3]	[95.3, 95.7]	[94.4, 94.9]
Train Y Test O [*]	Sedentary	[86.6, 87.9]	[86.4, 87.7]	[85.5, 86.9]	[98.7, 99.0]	[98.7, 99.0]	[98.9, 99.1]
	Standing	[79.8, 81.8]	[80.1, 82.0]	[77.1, 79.1]	[97.1, 97.5]	[97.1, 97.4]	[97.5, 97.8]
	Walking	[85.7, 86.8]	[86.2, 87.2]	[86.9, 88.0]	[93.3, 94.0]	[93.3, 93.9]	[91.9, 92.6]
	Walking Upstairs	[62.9, 70.9]	[69.0, 76.5]	[62.6, 70.5]	[98.1, 98.4]	[98.2, 98.5]	[98.5, 98.7]
	Walking Downstairs	[81.5, 87.8]	[83.6, 89.6]	[84.1, 90.0]	[96.9, 97.2]	[96.9, 97.3]	[96.5, 96.9]
	Postural Transition	[86.5, 90.4]	[86.0, 90.0]	[86.5, 90.4]	[96.6, 97.0]	[96.8, 97.2]	[96.9, 97.3]
Train O Test Y [*]	Sedentary	[94.9, 96.3]	[94.5, 96.0]	[94.8, 96.3]	[98.1, 98.7]	[98.0, 98.6]	[98.1, 98.7]
	Standing	[88.5, 90.8]	[88.1, 90.4]	[88.4, 90.7]	[96.4, 97.2]	[96.0, 96.9]	[96.4, 97.2]
	Walking	[62.5, 65.8]	[59.4, 62.8]	[64.3, 67.6]	[97.3, 98.0]	[97.5, 98.2]	[97.2, 97.9]
	Walking Upstairs	[75.7, 83.8]	[80.3, 87.6]	[72.9, 81.3]	[96.7, 97.4]	[96.3, 97.0]	[97.6, 98.2]
	Walking Downstairs	[88.5, 94.3]	[90.2, 95.5]	[86.5, 92.8]	[95.0, 95.8]	[94.5, 95.3]	[95.2, 96.0]
	Postural Transition	[82.6, 91.6]	[83.2, 91.9]	[83.2, 91.9]	[95.4, 96.2]	[95.6, 96.3]	[94.8, 95.6]
Train Y&O Test (Y&O) [*]	Sedentary	[88.9, 90.0]	[88.7, 89.8]	[88.7, 89.8]	[98.7, 98.9]	[98.6, 98.9]	[98.6, 98.9]
	Standing	[82.6, 84.1]	[82.5, 84.0]	[80.5, 82.1]	[97.2, 97.6]	[97.1, 97.4]	[97.6, 97.9]
	Walking	[71.6, 72.9]	[70.5, 71.8]	[81.9, 83.0]	[95.8, 96.3]	[95.8, 96.3]	[94.5, 95.1]
	Walking Upstairs	[69.3, 75.1]	[73.4, 78.9]	[65.3, 71.3]	[96.4, 96.7]	[96.4, 96.7]	[97.9, 98.1]
	Walking Downstairs	[89.7, 93.4]	[91.1, 94.6]	[85.4, 89.8]	[93.7, 94.1]	[93.2, 93.7]	[96.4, 96.7]
	Postural Transition	[88.6, 91.9]	[88.0, 91.3]	[89.6, 92.7]	[95.8, 96.1]	[96.0, 96.4]	[95.6, 96.0]

* Test data were obtained by re-orienting the data from the younger (Y) and/or older (O) cohort (see Section 3.1); [†] IMU data were re-sampled at 40 Hz, barometer data were re-sampled at 20 Hz.

References

- Appelboom, G.; Camacho, E.; Abraham, M.E.; Bruce, S.S.; Dumont, E.L.; Zacharia, B.E.; D'Amico, R.; Slomian, J.; Reginster, J.Y.; Bruyère, O.; et al. Smart wearable body sensors for patient self-assessment and monitoring. *Arch. Public Health* **2014**, *72*, 28–36. [[CrossRef](#)] [[PubMed](#)]
- Rosario, M.D.; Lovell, N.H.; Fildes, J.; Holgate, K.; Yu, J.; Ferry, C.; Schreier, G.; Ooi, S.Y.; Redmond, S.J. Evaluation of an mHealth-based Adjunct to Outpatient Cardiac Rehabilitation. *IEEE J. Biomed. Health Inform.* **2018**, *22*, 1938–1948. [[CrossRef](#)] [[PubMed](#)]
- Seel, T.; Raisch, J.; Schauer, T. IMU-Based Joint Angle Measurement for Gait Analysis. *Sensors* **2014**, *14*, 6891–6909. [[CrossRef](#)] [[PubMed](#)]
- Dadashi, F.; Mariani, B.; Rochat, S.; Büla, C.J.; Santos-Eggimann, B.; Aminian, K. Gait and Foot Clearance Parameters Obtained Using Shoe-Worn Inertial Sensors in a Large-Population Sample of Older Adults. *Sensors* **2014**, *14*, 443–457. [[CrossRef](#)] [[PubMed](#)]
- Wang, K.; Delbaere, K.; Brodie, M.; Lovell, N.; Kark, L.; Lord, S.; Redmond, S. Differences between Gait on Stairs and Flat Surfaces in Relation to Fall Risk and Future Falls. *IEEE J. Biomed. Health Inform.* **2017**, *21*, 1479–1486. [[CrossRef](#)] [[PubMed](#)]

6. Casamassima, F.; Ferrari, A.; Milosevic, B.; Ginis, P.; Farella, E.; Rocchi, L. A Wearable System for Gait Training in Subjects with Parkinson's Disease. *Sensors* **2014**, *14*, 6229–6246. [CrossRef] [PubMed]
7. Patel, S.; Lorincz, K.; Hughes, R.; Huggins, N.; Growdon, J.; Standaert, D.; Akay, M.; Dy, J.; Welsh, M.; Bonato, P. Monitoring motor fluctuations in patients with Parkinson's disease using wearable sensors. *IEEE Trans. Inf. Technol. Biomed.* **2009**, *13*, 864–873. [CrossRef]
8. Hubble, R.P.; Naughton, G.A.; Silburn, P.A.; Cole, M.H. Wearable sensor use for assessing standing balance and walking stability in people with Parkinson's disease: A systematic review. *PLoS ONE* **2015**, *10*, 1–22. [CrossRef]
9. Healy, G.N.; Wijndaele, K.; Dunstan, D.W.; Shaw, J.E.; Salmon, J.; Zimmet, P.Z.; Owen, N. Objectively measured sedentary time, physical activity, and metabolic risk: The Australian diabetes, obesity and lifestyle study (AusDiab). *Diabetes Care* **2008**, *31*, 369–371. [CrossRef]
10. Tudor-Locke, C.; Brashear, M.; Johnson, W.; Katzmarzyk, P. Accelerometer profiles of physical activity and inactivity in normal weight, overweight, and obese U.S. men and women. *Int. J. Behav. Nutr. Phys. Act.* **2010**, *7*, 60–70. [CrossRef]
11. Westerterp, K. Physical activity assessment with accelerometers. *Int. J. Obes. Relat. Metab. Disord. J. Int. Assoc. Study Obes.* **1999**, *23* (Suppl. 3), S45–S49. [CrossRef]
12. Zhang, K.; Werner, P.; Sun, M.; Pi-Sunyer, F.X.; Boozer, C.N. Measurement of Human Daily Physical Activity. *Obes. Res.* **2003**, *11*, 33–40. [CrossRef] [PubMed]
13. Trost, S.G.; McIver, K.L.; Pate, R.R. Conducting accelerometer-based activity assessments in field-based research. *Med. Sci. Sports Exerc.* **2005**, *37*, S531–S543. [CrossRef] [PubMed]
14. Atallah, L.; Lo, B.; King, R.; Yang, G.Z. Sensor Positioning for Activity Recognition Using Wearable Accelerometers. *IEEE Trans. Biomed. Circuits Syst.* **2011**, *5*, 320–329. [CrossRef] [PubMed]
15. Ridgers, N.D.; Salmon, J.; Ridley, K.; O'Connell, E.; Arundell, L.; Timperio, A. Agreement between activPAL and ActiGraph for assessing children's sedentary time. *Int. J. Behav. Nutr. Phys. Act.* **2012**, *9*, 15. [CrossRef] [PubMed]
16. Steeves, J.A.; Bowles, H.R.; McClain, J.J.; Dodd, K.W.; Brychta, R.J.; Wang, J.; Chen, K.Y. Ability of thigh-worn ActiGraph and activPAL monitors to classify posture and motion. *Med. Sci. Sports Exerc.* **2015**, *47*, 952–959. [CrossRef] [PubMed]
17. Yang, C.C.; Hsu, Y.L. A Review of Accelerometry-Based Wearable Motion Detectors for Physical Activity Monitoring. *Sensors* **2010**, *10*, 7772–7788. [CrossRef]
18. Smith, A. Nearly Half of American Adults are Smartphone Owners. 2012. Available online: http://s4.goeshow.com/bricepac/pie/2013/PDF/Software_and_Apps_in_the_City_and_on_the_Campus_3-13-13.pdf (accessed on 24 May 2017).
19. Smith, A. Smartphone Ownership 2013. 2013. Available online: http://boletines.prisadigital.com/PIP_Smartphone_adoption_2013.pdf (accessed on 24 May 2017).
20. Ting, D.H.; Lim, S.F.; Patanmacia, T.S.; Low, C.G.; Ker, G.C. Dependency on Smartphone and the Impact on Purchase Behaviour. *Young Consum.* **2011**, *12*, 193–203. [CrossRef]
21. Sarwar, M.; Soomro, T.R. Impact of Smartphone's on Society. *Eur. J. Sci. Res.* **2013**, *98*, 216–226.
22. Fiordelli, M.; Diviani, N.; Schulz, J.P. Mapping mHealth Research: A Decade of Evolution. *J. Med. Internet Res.* **2013**, *15*, 1–14. [CrossRef]
23. Neubeck, L.; Lowres, N.; Benjamin, E.J.; Freedman, S.B.; Coorey, G.; Redfern, J. The mobile revolution—Using smartphone apps to prevent cardiovascular disease. *Nat. Rev. Cardiol.* **2015**, *12*, 350–360. [CrossRef] [PubMed]
24. del Rosario, M.B.; Redmond, S.J.; Lovell, N.H. Tracking the Evolution of Smartphone Sensing for Monitoring Human Movement. *Sensors* **2015**, *15*, 18901–18933. [CrossRef] [PubMed]
25. Mellone, S.; Tacconi, C.; Schwickert, L.; Klenk, J.; Becker, C.; Chiari, L. Smartphone-based solutions for fall detection and prevention: The FARSEEING approach. *Z. Gerontol. Geriatr.* **2012**, *45*, 722–727. [CrossRef] [PubMed]
26. Zhang, S.; McCullagh, P.; Nugent, C.; Zheng, H. Activity monitoring using a smart phone's accelerometer with hierarchical classification. In Proceedings of the Sixth International Conference on Intelligent Environments (IE), Kuala Lumpur, Malaysia, 19–21 July 2010; pp. 158–163.
27. Mitchell, E.; Monaghan, D.; Connor, N. Classification of sporting activities using smartphone accelerometers. *Sensors* **2013**, *13*, 5317–5337. [CrossRef] [PubMed]

28. Marshall, J. Smartphone sensing for distributed swim stroke coaching and research. In Proceedings of the 2013 ACM Conference on Pervasive and Ubiquitous Computing Adjunct Publication, 8–12 September 2013; ACM: New York, NY, USA, 2013; UbiComp '13 Adjunct, pp. 1413–1416.
29. Shoaib, M.; Bosch, S.; Incel, O.; Scholten, H.; Havinga, P. Fusion of smartphone motion sensors for physical activity recognition. *Sensors* **2014**, *14*, 10146–10176. [[CrossRef](#)] [[PubMed](#)]
30. Antos, S.A.; Albert, M.V.; Kording, K.P. Hand, belt, pocket or bag: Practical activity tracking with mobile phones. *J. Neurosci. Methods* **2014**, *231*, 22–30. [[CrossRef](#)] [[PubMed](#)]
31. Khan, A.M.; Tufail, A.; Khattak, A.M.; Laine, T.H. Activity recognition on smartphones via sensor-fusion and KDA-based SVMs. *Int. J. Distrib. Sens. Netw.* **2014**, *2014*, 1–14. [[CrossRef](#)]
32. Lu, H.; Yang, J.; Liu, Z.; Lane, N.D.; Choudhury, T.; Campbell, A.T. The jigsaw continuous sensing engine for mobile phone applications. In Proceedings of the 8th ACM Conference on Embedded Networked Sensor Systems, SenSys '10, Zurich, Switzerland, 3–5 November 2010; ACM: New York, NY, USA, 2010; pp. 71–84.
33. Anjum, A.; Ilyas, M. Activity recognition using smartphone sensors. In Proceedings of the IEEE Consumer Communications and Networking Conference (CCNC), Las Vegas, NV, USA, 11–14 January 2013; pp. 914–919.
34. Khan, A.M.; Lee, Y.K.; Lee, S.Y.; Kim, T.S. A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. *IEEE Trans. Inf. Technol. Biomed.* **2010**, *14*, 1166–1172. [[CrossRef](#)]
35. Khan, A.; Lee, Y.K.; Lee, S.; Kim, T.S. Human activity recognition via an accelerometer-enabled-smartphone using kernel discriminant analysis. In Proceedings of the 5th International Conference on Future Information Technology (FutureTech), Busan, Korea, 21–23 May 2010; pp. 1–6.
36. Henpraserttae, A.; Thiemjarus, S.; Marukat, S. Accurate activity recognition using a mobile phone regardless of device orientation and location. In Proceedings of the International Conference on Body Sensor Networks, Dallas, TX, USA, 23–25 May 2011; pp. 41–46.
37. Yurtman, A.; Barshan, B. Activity Recognition Invariant to Sensor Orientation with Wearable Motion Sensors. *Sensors* **2017**, *17*, 1838. [[CrossRef](#)]
38. Yurtman, A.; Barshan, B.; Fidan, B. Activity Recognition Invariant to Wearable Sensor Unit Orientation Using Differential Rotational Transformations Represented by Quaternions. *Sensors* **2018**, *18*, 2725. [[CrossRef](#)]
39. Preece, S.J.; Goulermas, J.Y.; Kenney, L.P.J.; Howard, D.; Meijer, K.; Crompton, R. Activity identification using body-mounted sensors—A review of classification techniques. *Physiol. Meas.* **2009**, *30*, 1–33. [[CrossRef](#)] [[PubMed](#)]
40. Figo, D.; Diniz, P.C.; Ferreira, D.R.; Cardoso, J.M.P. Preprocessing techniques for context recognition from accelerometer data. *Pers. Ubiquitous Comput.* **2010**, *14*, 645–662. [[CrossRef](#)]
41. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
42. Ronao, C.A.; Cho, S.B. Human activity recognition with smartphone sensors using deep learning neural networks. *Expert Syst. Appl.* **2016**, *59*, 235–244. [[CrossRef](#)]
43. Ravi, D.; Wong, C.; Lo, B.; Yang, G. A Deep Learning Approach to on-Node Sensor Data Analytics for Mobile or Wearable Devices. *IEEE J. Biomed. Health Inform.* **2017**, *21*, 56–64. [[CrossRef](#)] [[PubMed](#)]
44. Milenkoski, M.; Trivodaliev, K.; Kalajdziski, S.; Jovanov, M.; Stojkoska, B.R. Real time human activity recognition on smartphones using LSTM networks. In Proceedings of the International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 21–25 May 2018; pp. 1126–1131.
45. Ordóñez, F.J.; Roggen, D. Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition. *Sensors* **2016**, *16*, 115. [[CrossRef](#)] [[PubMed](#)]
46. Wang, J.; Chen, Y.; Hao, S.; Peng, X.; Hu, L. Deep learning for sensor-based activity recognition: A Survey. *Pattern Recognit. Lett.* **2018**, doi:10.1016/j.patrec.2018.02.010. [[CrossRef](#)]
47. Grant, P.M.; Ryan, C.G.; Tigbe, W.W.; Granat, M.H. The validation of a novel activity monitor in the measurement of posture and motion during everyday activities. *Br. J. Sports Med.* **2006**, *40*, 992–997. [[CrossRef](#)] [[PubMed](#)]
48. Fisher, C.J. *Using an Accelerometer for Inclination Sensing*; Application Note: AN-1057; Analog Devices: Norwood, MA, USA, 2010; pp. 1–8.
49. Reiff, C.; Marlatt, K.; Dengel, D.R. Difference in caloric expenditure in sitting versus standing desks. *J. Phys. Act. Health* **2012**, *9*, 1009–1011. [[CrossRef](#)]

50. Mansoubi, M.; Pearson, N.; Cledes, S.A.; Biddle, S.J.; Bodicoat, D.H.; Tolfrey, K.; Edwardson, C.L.; Yates, T. Energy expenditure during common sitting and standing tasks: examining the 1.5 MET definition of sedentary behaviour. *BMC Public Health* **2015**, *15*, 516–523. [[CrossRef](#)]
51. Petersen, C.B.; Bauman, A.; Tolstrup, J.S. Total sitting time and the risk of incident diabetes in Danish adults (the DANHES cohort) over 5 years: A prospective study. *Br. J. Sports Med.* **2016**, *50*, 1382–1387. [[CrossRef](#)] [[PubMed](#)]
52. Åsvold, B.O.; Midthjell, K.; Krokstad, S.; Rangul, V.; Bauman, A. Prolonged sitting may increase diabetes risk in physically inactive individuals: An 11 year follow-up of the HUNT Study, Norway. *Diabetologia* **2017**, *60*, 830–835. [[CrossRef](#)] [[PubMed](#)]
53. Wilmot, E.G.; Edwardson, C.L.; Achana, F.A.; Davies, M.J.; Gorely, T.; Gray, L.J.; Khunti, K.; Yates, T.; Biddle, S.J.H. Sedentary time in adults and the association with diabetes, cardiovascular disease and death: Systematic review and meta-analysis. *Diabetologia* **2012**, *55*, 2895–2905. [[CrossRef](#)] [[PubMed](#)]
54. Tran, B.; Falster, M.O.; Douglas, K.; Blyth, F.; Jorm, L.R. Health Behaviours and Potentially Preventable Hospitalisation: A Prospective Study of Older Australian Adults. *PLoS ONE* **2014**, *9*, e93111. [[CrossRef](#)] [[PubMed](#)]
55. Biswas, A.; Oh, P.I.; Faulkner, G.E.; Bajaj, R.R.; Silver, M.A.; Mitchell, M.S.; Alter, D.A. Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: A systematic review and meta-analysis. *Ann. Intern. Med.* **2015**, *162*, 123–132. [[CrossRef](#)]
56. Del Rosario, M.B.; Wang, K.; Wang, J.; Liu, Y.; Brodie, M.; Delbaere, K.; Lovell, N.H.; Lord, S.R.; Redmond, S.J. A comparison of activity classification in younger and older cohorts using a smartphone. *Physiol. Meas.* **2014**, *35*, 2269–2286. [[CrossRef](#)] [[PubMed](#)]
57. Khan, A.; Siddiqi, M.; Lee, S.W. Exploratory data analysis of acceleration signals to select light-weight and accurate features for real-time activity recognition on smartphones. *Sensors* **2013**, *13*, 13099–13122. [[CrossRef](#)]
58. Chou, J.C.K. Quaternion kinematic and dynamic differential equations. *IEEE Trans. Robot. Autom.* **1992**, *8*, 53–64. [[CrossRef](#)]
59. Del Rosario, M.B.; Khamis, H.; Ngo, P.; Lovell, N.H.; Redmond, S.J. Computationally-Efficient Adaptive Error-State Kalman Filter for Attitude Estimation. *IEEE Sens. J.* **2018**, *18*, 9332–9342. [[CrossRef](#)]
60. Welford, B.P. Note on a method for calculating corrected sums of squares and products. *Technometrics* **1962**, *4*, 419–420. [[CrossRef](#)]
61. Antonsson, E.K.; Mann, R.W. The frequency content of gait. *J. Biomech.* **1985**, *18*, 39–47. [[CrossRef](#)]
62. Jiménez, A.R.; Seco, F.; Prieto, J.C.; Guevara, J. Indoor pedestrian navigation using an INS/EKF framework for yaw drift reduction and a foot-mounted IMU. In Proceedings of the 2010 7th Workshop on Positioning Navigation and Communication, Dresden, Germany, 11–12 March 2010; pp. 135–143.
63. Del Rosario, M.B.; Lovell, N.H.; Redmond, S.J. Quaternion-Based Complementary Filter for Attitude Determination of a Smartphone. *IEEE Sens. J.* **2016**, *16*, 6008–6017. [[CrossRef](#)]
64. Madgwick, S.; Harrison, A.; Vaidyanathan, R. Estimation of IMU and MARG orientation using a gradient descent algorithm. In Proceedings of the IEEE International Conference on Rehabilitation Robotics, Zurich, Switzerland, 29 June–1 July 2011; pp. 1–7.
65. Elvira, V.; Nazabal-Renteria, A.; Artés-Rodríguez, A. A novel feature extraction technique for human activity recognition. In Proceedings of the 2014 IEEE Workshop on Statistical Signal Processing (SSP), Gold Coast, VIC, Australia, 29 June–2 July 2014; pp. 177–180.
66. Higgins, W.T. A comparison of complementary and Kalman filtering. *IEEE Trans. Aerosp. Electron. Syst.* **1975**, *AES-11*, 321–325. [[CrossRef](#)]
67. Sabatini, A.; Genovese, V. A sensor fusion method for tracking vertical velocity and height based on inertial and barometric altimeter measurements. *Sensors* **2014**, *14*, 13324–13347. [[CrossRef](#)]
68. Bar-Shalom, Y.; Li, X.R.; Kirubarajan, T. *Estimation with Applications to Tracking and Navigation*; John Wiley & Sons, Inc.: New York, NY, USA, 2002.
69. Guo, G. *Pressure Altimetry Using the MPL3115A2*; Application Note: AN4528; Freescale Semiconductor Ltd.: Austin, TX, USA, 2012; pp. 1–13.
70. Sabatini, A.M.; Ligorio, G.; Mannini, A.; Genovese, V.; Pinna, L. Prior-to- and Post-Impact Fall Detection Using Inertial and Barometric Altimeter Measurements. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2016**, *24*, 774–783. [[CrossRef](#)]

71. Capela, N.A.; Lemaire, E.D.; Baddour, N. Feature selection for wearable smartphone-based human activity recognition with able bodied, elderly, and stroke patients. *PLoS ONE* **2015**, *10*, e0124414. [[CrossRef](#)] [[PubMed](#)]
72. Capela, N.A.; Lemaire, E.D.; Baddour, N. Improving classification of sit, stand, and lie in a smartphone human activity recognition system. In Proceedings of the IEEE International Symposium on Medical Measurements and Applications, Turin, Italy, 7–9 May 2015; pp. 473–478.
73. Breiman, L.; Friedman, J.; Stone, C.J.; Olshen, R.A. *Classification and Regression Trees*; The Wadsworth and Brooks-Cole Statistics-Probability Series; Taylor & Francis: Abingdon, UK, 1984.
74. Esposito, F.; Malerba, D.; Semeraro, G.; Kay, J. A comparative analysis of methods for pruning decision trees. *IEEE Trans. Pattern Anal. Mach. Intell.* **1997**, *19*, 476–491. [[CrossRef](#)]
75. Aggarwal, J.; Ryoo, M. Human Activity Analysis: A Review. *ACM Comput. Surv.* **2011**, *43*, 1–43. [[CrossRef](#)]
76. Chawla, N.V., Data Mining for Imbalanced Datasets: An Overview. In *Data Mining and Knowledge Discovery Handbook*; Springer: Boston, MA, USA, 2010; pp. 875–886.
77. Shannon, C.; Weaver, W. *The Mathematical Theory of Communication*; Number v. 1 in The Mathematical Theory of Communication; University of Illinois Press: Champaign, IL, USA, 1949.
78. Li, F.; Shirahama, K.; Nisar, M.A.; Köping, L.; Grzegorzec, M. Comparison of Feature Learning Methods for Human Activity Recognition Using Wearable Sensors. *Sensors* **2018**, *18*, 679. [[CrossRef](#)]
79. Bao, L.; Intille, S.S. Activity Recognition from User-Annotated Acceleration Data. In *Pervasive Computing*; Ferscha, A., Mattern, F., Eds.; Springer: Berlin/Heidelberg, Germany, 2004; pp. 1–17.
80. Anguita, D.; Ghio, A.; Oneto, L.; Parra, X.; Reyes-Ortiz, J.L. Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. In Proceedings of the 4th International Conference on Ambient Assisted Living and Home Care, IWAAL'12, Vitoria-Gasteiz, Spain, 3–5 December 2012; Springer: Berlin/Heidelberg, Germany, 2012; pp. 216–223.
81. Gu, F.; Khoshelham, K.; Valaee, S.; Shang, J.; Zhang, R. Locomotion Activity Recognition Using Stacked Denoising Autoencoders. *IEEE Internet Things J.* **2018**, *5*, 2085–2093. [[CrossRef](#)]
82. Gramkow, C. On averaging rotations. *Int. J. Comput. Vis.* **2001**, *42*, 7–16. [[CrossRef](#)]
83. Dam, E.B.; Koch, M.; Lillholm, M. *Quaternions, Interpolation and Animation*; Technical Report; University of Copenhagen: Copenhagen, Denmark, 1998.



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).

Article

Unobtrusive Estimation of Cardiovascular Parameters with Limb Ballistocardiography

Yang Yao ^{1,2}, Sungtae Shin ², Azin Mousavi ², Chang-Sei Kim ³, Lisheng Xu ¹, Ramakrishna Mukkamala ⁴ and Jin-Oh Hahn ^{2,*}

¹ College of Medicine and Biological Information Engineering, Northeastern University, Shenyang 110169, China

² Department of Mechanical Engineering, University of Maryland, College Park, College Park, MD 20742, USA

³ School of Mechanical Engineering, Chonnam National University, Gwangju 61186, Korea

⁴ Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824, USA

* Correspondence: jhahn12@umd.edu; Tel.: +1-301-405-7864

Received: 23 May 2019; Accepted: 26 June 2019; Published: 1 July 2019

Abstract: This study investigates the potential of the limb ballistocardiogram (BCG) for unobtrusive estimation of cardiovascular (CV) parameters. In conjunction with the reference CV parameters (including diastolic, pulse, and systolic pressures, stroke volume, cardiac output, and total peripheral resistance), an upper-limb BCG based on an accelerometer embedded in a wearable armband and a lower-limb BCG based on a strain gauge embedded in a weighing scale were instrumented simultaneously with a finger photoplethysmogram (PPG). To standardize the analysis, the more convenient yet unconventional armband BCG was transformed into the more conventional weighing scale BCG (called the synthetic weighing scale BCG) using a signal processing procedure. The characteristic features were extracted from these BCG and PPG waveforms in the form of wave-to-wave time intervals, wave amplitudes, and wave-to-wave amplitudes. Then, the relationship between the characteristic features associated with (i) the weighing scale BCG-PPG pair and (ii) the synthetic weighing scale BCG-PPG pair versus the CV parameters, was analyzed using the multivariate linear regression analysis. The results indicated that each of the CV parameters of interest may be accurately estimated by a combination of as few as two characteristic features in the upper-limb or lower-limb BCG, and also that the characteristic features recruited for the CV parameters were to a large extent relevant according to the physiological mechanism underlying the BCG.

Keywords: ballistocardiography; ballistocardiogram; blood pressure; stroke volume; cardiac output; total peripheral resistance; photoplethysmography; photoplethysmogram

1. Introduction

The ballistocardiogram (BCG) is the recording of body movement (including displacement, velocity, and acceleration) in response to the ejection of the blood by the heart. In the absence of any external force acting on the body, the center of mass of the body must remain unchanged. Hence, as the blood circulates in the body, the rest of the body moves in the opposite direction to the circulating blood so that the center of mass of the entire body is maintained. This body movement may be recorded using a wide range of BCG instruments, such as a force plate [1–3], weighing scale [4–7], bed [8,9], chair [10–12], and wearables [13–15]. Being a response to the circulation of the blood, the BCG may be closely associated with the cardiovascular (CV) functions and thus possess clinical value. In fact, a recent study by us elucidated that the BCG is primarily attributed to the interaction of blood pressure (BP) at the aortic inlet and outlet as well as the apex of the aortic arch [16]. Hence, the BCG waveform

is largely shaped by the aortic BP waveforms and may thus serve as a window through which the shape of the aortic BP waveforms can be inferred (at least to a certain extent).

The measurement of clinically significant CV parameters often requires inconvenient instruments and even invasive procedures. For example, the gold standard arterial BP waveform is measured by invasive arterial catheterization [17]. There are non-invasive options such as volume-clamping techniques [18,19] and applanation tonometry [20], but these techniques require costly equipment and/or trained operators. The gold standard stroke volume (SV), cardiac output (CO), and total peripheral resistance (TPR) likewise require inconvenient and costly procedures such as dye injection [21], echocardiography [22], impedance cardiography [23], and electrical impedance tomography [24]. The CV parameters have also been derived indirectly using the so-called pulse contour methods [25–29]. These methods have been extensively investigated and demonstrated to be successful. Yet, the techniques still necessitate the measurement of (invasive or non-invasive) arterial BP waveforms.

Considering that the shape of the BCG may originate from the aortic BP waveforms, it is quite reasonable to conceive that the BCG (especially the characteristic features therein) may have a close relationship to the CV parameters. Combined with the unobtrusiveness of the BCG instrumentation, such a capability may open up unprecedented possibilities for ultra-convenient estimation of CV parameters in daily life. Regardless, the existing body of work on the use of the BCG for CV parameter estimation is quite sparse other than cuff-less BP. Indeed, it is only recently that a few pioneering studies to investigate the feasibility of the BCG for CV parameter monitoring appeared, including diastolic BP (DP) and systolic BP (SP) [1,2,14,30–32], SV and CO [30,33], peripheral blood oxygenation [30], and preload and afterload [34].

Motivated by this opportunity and limitations of the state-of-the-art techniques, the goal of this work was to investigate the potential of the limb BCG for unobtrusive estimation of CV parameters. The BCG in the head-to-foot direction was instrumented at the upper limb site using an accelerometer embedded in a wearable armband and at the lower limb site using a strain gauge embedded in a customized weighing scale, respectively, simultaneously with a finger photoplethysmogram (PPG). The weighing scale BCG is to a large extent analogous to the traditional whole-body BCG [16]. Thus, its physical implications may be drawn from earlier work on the BCG [35–37]. In contrast, there is relatively little prior work on the armband BCG, since these so-called wearable BCG has gained interest only recently by virtue of its convenience in instrumentation. Hence, the weighing scale BCG and the armband BCG may present contrasting trade-off between the accuracy of CV parameter estimation and compatibility to wearable implementation. To standardize the analysis of these distinct BCG signals, the armband BCG was transformed into the weighing scale BCG (called the synthetic weighing scale BCG) using a signal processing procedure. The characteristic features were extracted from these BCG and PPG waveforms in the form of wave-to-wave time intervals, wave amplitudes, and wave-to-wave amplitudes as well as BCG-PPG pulse transit time (PTT; the time interval between a major wave in the BCG and the diastolic foot of the PPG) [1,2,31]. Then, the relationship between the characteristic features associated with (i) the weighing scale BCG-PPG pair and (ii) the synthetic weighing scale BCG-PPG pair versus the CV parameters (including DP, pulse BP (PP), and SP, SV, CO, and TPR) was analyzed using the multivariate linear regression analysis.

2. Materials and Methods

To investigate the potential of the limb BCG for CV parameter estimation, the upper-limb and lower-limb BCG signals were analyzed with the following procedure: (i) experimental data acquisition; (ii) signal pre-conditioning; (iii) signal processing to transform the armband BCG into the weighing scale BCG; (iv) feature extraction; and (v) multivariate regression analysis (Figure 1a).

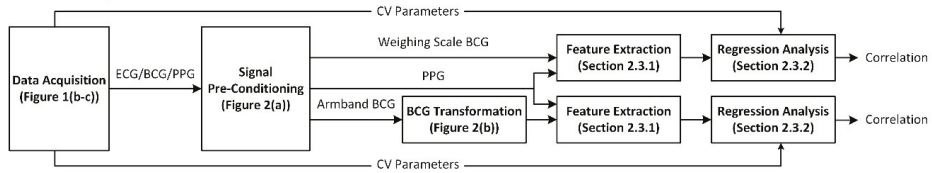
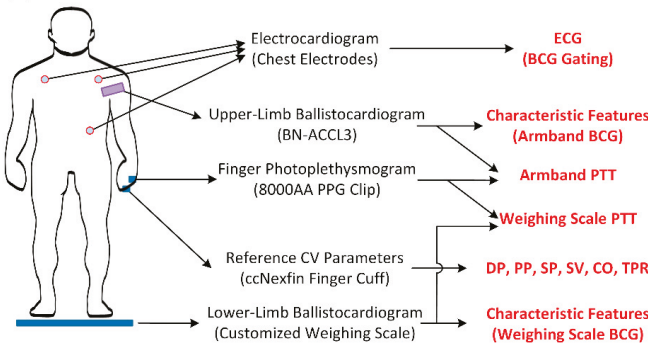
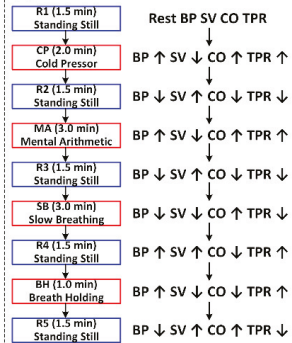
(a) BCG Analysis Procedure**(b) Measurements****(c) Interventions**

Figure 1. Study design and experimental protocol. (a) BCG signal analysis procedure to investigate the potential of limb BCG for cardiovascular (CV) parameter estimation. (b) Instrumented physiological signals. (c) Hemodynamic interventions. ECG: electrocardiogram. BCG: ballistocardiogram. PTT: pulse transit time. DP: diastolic pressure. PP: pulse pressure. SP: systolic pressure. SV: stroke volume. CO: cardiac output. TPR: total peripheral resistance.

2.1. Experimental Protocol

Under the approval obtained from the University of Maryland Institutional Review Board (IRB) and written informed consent, human subject study was conducted in 17 young healthy volunteers (age 25 ± 5 years old; gender 12 male and 5 female; height 174 ± 10 cm; weight 74 ± 17 kg), in strict accordance with the IRB guidelines.

From each subject, a wide variety of physiological signal waveforms required for investigating the relationship between the upper-limb (i.e., arm) and lower-limb (i.e., leg) BCG and the CV parameters was instrumented using off-the-shelf sensors as follows. First, the ECG was instrumented using three gel electrodes in a modified Lead II configuration interfaced to a wireless amplifier (BN-EL50, Biopac Systems, Goleta, CA, USA). Second, the reference CV parameters (including the BP waveform, SV, CO, and TPR) were instrumented using a fast servo-controlled finger cuff embedded with a blood volume waveform sensor on the ring finger of a hand to implement the volume clamping method [18,19] (ccNexfin, Edwards Lifesciences, Irvine, CA, USA). Third, the upper-limb BCG (called hereafter the armband BCG) was instrumented using a high-resolution accelerometer embedded in an armband equipped with a wireless amplifier (BN-ACCL3, Biopac Systems, Goleta, CA, USA). Fourth, the lower-limb BCG (called hereafter the weighing scale BCG) was instrumented using a strain gauge embedded in a customized weighing scale (BC534, Tanita, Tokyo, Japan). Fifth, the PPG signal was instrumented using a finger clip sensor (8000AA, Nonin Medical, Plymouth, MN, USA). All the devices were interfaced to a laptop computer by way of a data acquisition unit (MP150, Biopac Systems, Goleta, CA, USA) to synchronously instrument all the waveforms at 1 kHz sampling rate (Figure 1b).

The aforementioned physiological signal waveforms were acquired while the subjects underwent four hemodynamic interventions (Figure 1c). Each subject stood still for 1.5 min for an initial rest state (R1). Then, the subject underwent the cold pressor intervention (CP) for 2 min, in which the subject was asked to immerse a free hand in ice water. Followed by standing still for 1.5 min for a second

rest state (R2), the subject underwent the mental arithmetic intervention (MA) for 3 min, in which the subject was asked to repeatedly add the digits of a three-digit number and add the sum to the original number. Followed by standing still for 1.5 min for a third rest state (R3), the subject underwent the slow breathing intervention (SB) for 3 min, in which the subject was asked to take deep and slow breaths. Followed by standing still for 1.5 min for a fourth rest state (R4), the subject underwent the breath holding intervention (BH), in which the subject was asked to hold breath after normal exhalation. Lastly, the subject stood still for 1.5 min for a fifth rest state (R5). During the study, the subjects were asked to stand on the customized weighing scale with their arms placed at the side and still, and their movements minimized. Signal acquisition was continuously made throughout these states.

2.2. Signal Pre-Conditioning

In each subject, the acquired data were segmented into nine periods: R1, CP, R2, MA, R3, SB, R4, BH, and R5. Then, the physiological signal waveforms were pre-conditioned as follows on a period-by-period basis. First, the signals were smoothed via zero-phase filtering: the ECG and BP by a 1st-order Butterworth low-pass filter with a cut-off frequency of 20 Hz, and the BCG and PPG by a 2nd-order Butterworth band-pass filter with a pass band of 0.5~10 Hz. Second, the ECG R wave was extracted using the Pan Tompkins method. Third, the BCG and PPG beats were gated with the time instants corresponding to 10% of cardiac period before the R wave as gating locations. Fourth, beats associated with the low-quality armband and/or weighing scale BCG waveforms were discarded, by (i) calculating the amplitudes associated with all the armband and weighing scale BCG beats, and (ii) removing the beats associated with extraordinarily large or small BCG amplitude (i.e., outside of 3 scaled median absolute deviations (with the scaling factor of 1.4826) around the median amplitude) [38]. Fifth, the armband and weighing scale BCG signals were smoothed using a 10-beat exponential moving average filter to suppress the adverse impact of motion artifacts. The signal pre-conditioning procedure is depicted in Figure 2a.

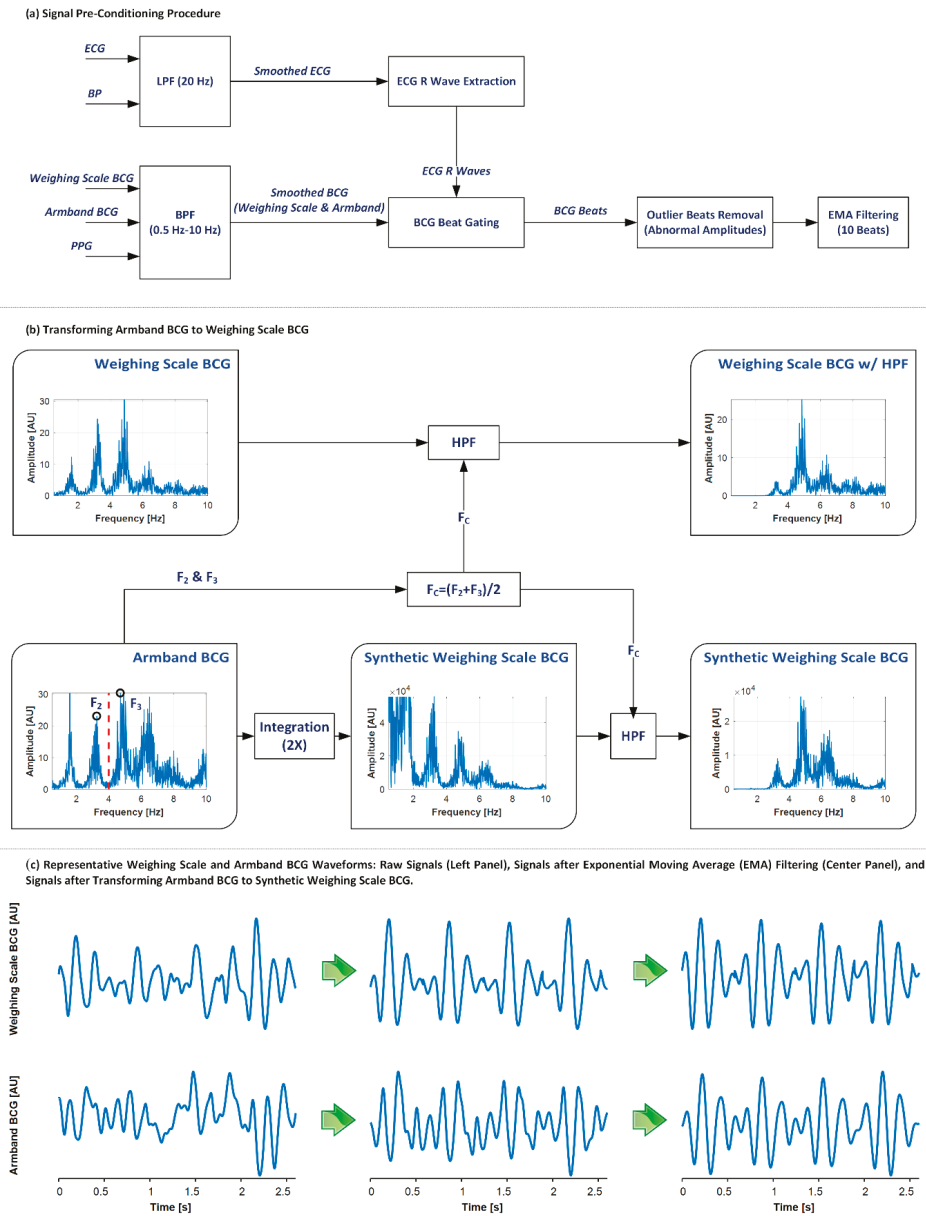


Figure 2. Procedure for signal pre-conditioning and transformation of armband BCG to weighing scale BCG. (a) Signal pre-conditioning procedure. LPF: low-pass filtering. BPF: band-pass filtering. EMA: exponential moving average. (b) Procedure for transforming armband BCG to weighing scale BCG. F_2 , F_3 , F_C : frequencies associated with 2nd (F_2) and 3rd (F_3) spectral peaks and the band (marked as red vertical line in the armband BCG signal) of high-pass filtering (HPF). (c) Representative weighing scale and armband BCG waveforms: raw signals (left), signals after EMA filtering (center), and signals after transforming the armband BCG to the synthetic weighing scale BCG.

2.3. Analysis of Weighing Scale Ballistocardiogram (BCG) for Cardiovascular (CV) Parameter Estimation

The weighing scale BCG was analyzed to investigate its association with the CV parameters in the following steps: (i) feature extraction and (ii) multivariate regression analysis.

2.3.1. Feature Extraction

The weighing scale BCG was labeled for the major I, J, and K waves as follows. The J wave was determined by finding the maximum peak in each BCG beat appearing after the ECG R wave. Then, the I and K waves were determined by finding the local minima right before and after the J wave, respectively. The foot of the PPG was determined using the intersecting tangent method [39]. By using these labels, a total of 16 characteristic features listed in Table 1 was constructed.

The reference CV parameters were computed as follows. In each cardiac beat, diastolic (DP) and systolic (SP) BP were computed as the minimum and maximum values in the BP waveform, while pulse pressure (PP) was computed as the difference between DP and SP. SV, CO, and TPR were computed as the mean values of the recorded SV, CO, and TPR values in each cardiac beat.

Table 1. Characteristic features extracted from the ballistocardiogram (BCG) in conjunction with the photoplethysmogram (PPG).

Symbol	Definition
PTT_I	Time interval between BCG I wave and PPG foot
PTT_J	Time interval between BCG J wave and PPG foot
PTT_K	Time interval between BCG K wave and PPG foot
T_{IJ}	Time interval between BCG I wave and J wave
T_{JK}	Time interval between BCG J wave and K wave
T_{IK}	Time interval between BCG I wave and K wave
T_{JJ}	Time interval between J waves of two consecutive BCG beats
A_I	Amplitude of BCG I wave
A_J	Amplitude of BCG J wave
A_K	Amplitude of BCG K wave
A_{IJ}	Amplitude difference between I wave and J wave
A_{JK}	Amplitude difference between J wave and K wave
$A_{IJ} \cdot PTT_I^2$	Surrogate of SV^*
$A_{JK} \cdot PTT_I^2$	Surrogate of SV^*
RMS	Root mean square of BCG waveform $s[i]$, $i = 1 \sim n$: $\sqrt{\sum_{i=1}^n s[i]^2 / n}$
E	Energy of BCG waveform $s[i]$, $i = 1 \sim n$: $\sum_{i=1}^n s[i]^2$

*: Considering that A_{IJ} and A_{JK} are approximately associated with PP and PTT^2 is proportional to arterial compliance, $A_{IJ} \cdot PTT_I^2$ and $A_{JK} \cdot PTT_I^2$ are approximately associated with stroke volume (SV).

2.3.2. Data Analysis

The data were analyzed in the following steps. First, the outliers in the extracted characteristic features were identified and removed. Second, the sample size of the characteristic features was increased. Third, the relationship between the characteristic features and the CV parameters were analyzed. The analysis was performed on the subject-by-subject basis.

First, the outliers in the characteristic features were extracted from the BCG and PPG signals as follows. In each of the nine rest and intervention periods associated with each subject, we examined the time series sequences of the characteristic features. Each 3 consecutive samples in the time series were inspected for possible outliers in a 9-sample window (including 3 samples before and 3 samples after the inspected samples). An outlier was identified if a sample was outside of 3 scaled median absolute deviations around the median of the 9 characteristic feature samples. If >75% of the beats in a period were removed, the period itself was excluded from subsequent analysis. Subjects in which <6 rest and intervention periods are available for analysis was also excluded from subsequent analysis.

Second, we increased the sample size of the characteristic features using the bootstrap technique similar to prior work [40,41] so as to conduct robust regression analysis (i.e., to reliably determine the coefficients in the regression models). More specifically, in each of the nine rest and intervention periods associated with each subject, the time intervals at which the CV parameters and the characteristic features attained stable extrema were determined (see Table 2 for the definition of the extrema). Then, 11 samples in the vicinity of the extrema were taken, the average of which were used as the representative CV parameter and characteristic feature values associated with the period. In addition, each of the CV parameters and characteristic features were approximated as the corresponding parametric bootstrap based on the mean and standard deviation of the 11 samples. Then, 100 bootstrap samples were created using the Monte Carlo method. Each bootstrap sample was created by (i) creating 11 random Monte Carlo samples and (ii) taking their average. Hence, up to 900 bootstrap samples (corresponding to the nine rest and intervention periods) were created in each subject. In each subject, the bootstrap samples of the CV parameters and characteristics features associated with all the rest and intervention periods were merged for multivariate regression analysis.

Table 2. Extremum regions of cardiovascular (CV) parameters in individual rest and intervention periods.

	R1	CP	R2	MA	R3	SB	R4	BH	R5
DP	Min	Max	Min	Max	Min	Min	Min	Max	Min
PP	Min	Max	Min	Max	Min	Min	Min	Max	Min
SP	Min	Max	Min	Max	Min	Min	Min	Max	Min
SV	Max	Min	Max	Min	Max	Min	Max	Min	Max
CO	Min	Max	Min	Max	Min	Min	Min	Min	Min
TPR	Min	Max	Min	Max	Min	Min	Min	Max	Min

Third, multivariate linear regression analysis was conducted at the individual subject level to investigate the potential of the weighing scale BCG for unobtrusive estimation of CV parameters. First, multivariate linear regression models associated with each of the CV parameters were developed using the bootstrap samples. Then, the validity of these models was tested using the representative CV parameters and characteristic features at the extrema associated with all the available rest and intervention periods of the subject (≤ 9 ; Figure 1c). The goal of the multivariate regression analysis was to determine (i) the most predictive characteristic features for the CV parameters as well as (ii) the number of characteristic features required to achieve high degree of correlation ($r \geq 0.7$) with the CV parameters for accurate estimation. Hence, we considered all possible combinations of the characteristics features exhaustively, and selected the models exhibiting high degree of correlation and equipped with physiologically relevant characteristic features (e.g., as suggested by our prior work [16,42]). The Pearson's correlation coefficient was used for determining the univariate characteristics features closely correlated with the CV parameters as well as for assessing the performance of the multivariate linear regression models.

2.4. Analysis of Armband BCG for CV Parameter Estimation

The armband BCG was analyzed to investigate its association with the CV parameters in the following steps: (i) transformation of the armband BCG to the weighing scale BCG, (ii) feature extraction, and (iii) multivariate regression analysis.

2.4.1. Transformation of Armband BCG to Weighing Scale BCG

The armband BCG and the weighing scale BCG are distinct in waveform morphology due to the difference in the measurement modality involved: the former is an acceleration measurement whereas the latter is a displacement measurement. Our prior work on the physical mechanisms and implications of the BCG [16,42] suggests that the relationship between the upper-limb acceleration BCG and CV parameters is obscure due to the mechanical body filtering effect compared with the lower-limb

displacement BCG. Hence, the armband BCG was transformed into an equivalent weighing scale BCG. Given that the primary source of the discrepancy between the armband BCG and the weighing scale BCG is the measurement modality (i.e., accelerometer versus strain gauge) if the body is assumed to be rigid, this was accomplished by applying two integrations to the armband BCG (Figure 2b). More specifically, the armband BCG was integrated in time twice using the trapezoidal method to yield the synthetic weighing scale BCG. Then, the synthetic weighing scale BCG was zero-phase filtered using a 4th-order Butterworth high-pass filter to remove the low-frequency drift therein. The cut-off frequency of the filter was determined so that the power spectra (especially in terms of the primary spectral peaks) associated with the weighing scale BCG and the synthetic weighing scale BCG were made consistent. The comparison of the power spectra associated with the weighing scale BCG and the synthetic weighing scale BCG showed that the latter exhibited largely higher spectra up to the 2nd spectral peak compared to the former (Figure 2b). Hence, the cut-off frequency was determined empirically as the average of the 2nd and 3rd peaks in the BCG power spectrum. Practically, the cut-off frequency can be computed easily from the heart rate as 2.5 times the heart rate, since the spectral peaks in the BCG represent the heart rate and its harmonics. The above-described procedure was performed in each subject on a period-by-period basis.

We quantitatively assessed the beat-by-beat quality of the weighing scale BCG and the synthetic weighing BCG calculated from the armband BCG via the following criteria: (1) $\frac{\|s[i] - \bar{s}\|}{\|\bar{s} - m\|} > 1$, where $s[i]$ is individual BCG beat in a (rest or intervention) period, \bar{s} is the ensemble average of all beats in the period, m is the mean of \bar{s} ; (2) correlation coefficient between $s[i]$ and \bar{s} less than 0.5 in each period; (3) a peak with a prominence [43,44] of >0.25 is detected from the 2nd derivative of the BCG waveform from I wave to K wave as a measure of distortion in the BCG waveform. All beats not fulfilling these criteria were removed from further analysis.

2.4.2. Feature Extraction and Data Analysis

Feature extraction and data analysis were conducted in the same way as the weighing scale BCG, as described in detail in Section 2.3.

3. Results

3.1. Experimental Data

Figure 3 shows the trends of the changes in the CV parameters in response to the hemodynamic interventions employed in this work. DP, PP, SP, and TPR increased in response to CP, MA, and BH, while it decreased in response to SB. Likewise, CO increased in response to CP and MA. But, it increased modestly in response to SB and decreased modestly in response to BH. SV decreased in response to all the hemodynamic interventions. Noting that CO increased in CP and MA, the decrease in SV may be attributed to a large increase in heart rate which shortens the left ventricular ejection time yet still increases CO [45]. On the other hand, the decrease in SV in SB and BH may be associated with the marginal change in CO and decrease in heart rate, which is consistent with the findings of prior studies [46–48]. These trends were used in defining the extrema associated with the CV parameters in Table 2.

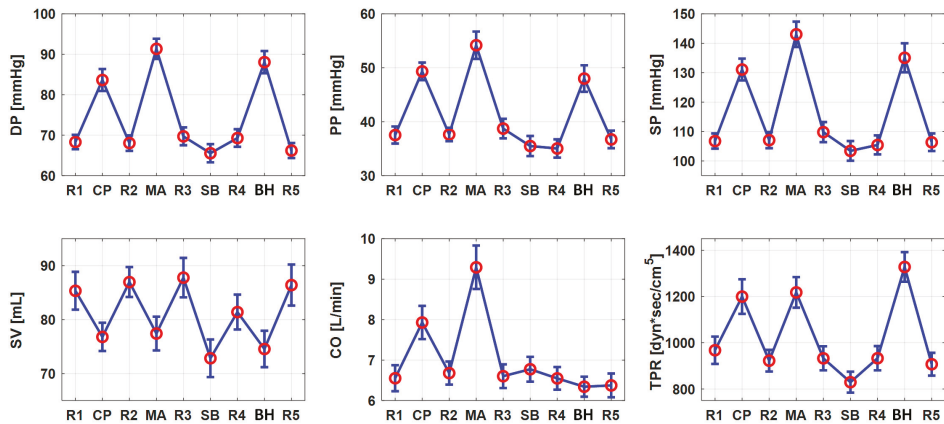


Figure 3. Group-average changes in the cardiovascular parameters in response to hemodynamic interventions (mean \pm standard error (SE)).

3.2. CV Parameter Estimation with Weighing Scale BCG

The number of subjects available for multivariate linear regression analysis after the outlier removal (i.e., subjects with ≥ 6 rest and intervention periods available for analysis; see Section 2.3 for details) was ≥ 14 for all the CV parameters associated with the weighing scale BCG. Multivariate linear regression analysis suggested that each of the CV parameters of interest may be accurately estimated by a combination of as few as two characteristic features. In contrast, the best correlation coefficients achieved by univariate characteristic features were on the average high for DP (0.81) and SP (0.82) but not sufficiently high for the remaining CV parameters (< 0.65). For the weighing scale BCG at the univariate level, DP was correlated well with PTT_I ($r = -0.81 \pm 0.02$) and PTT_J ($r = -0.69 \pm 0.04$), PP was correlated reasonably with PTT_I ($r = -0.65 \pm 0.05$) and PTT_J ($r = -0.57 \pm 0.07$) as well as A_{JK} (0.54 ± 0.07) and A_{IJ} (0.53 ± 0.07), and SP was correlated well with PTT_I ($r = -0.82 \pm 0.02$) and PTT_J ($r = -0.72 \pm 0.04$). SV was correlated most strongly with A_J ($r = 0.50 \pm 0.09$). CO was correlated with T_{JJ} ($r = -0.57 \pm 0.11$), and to a lesser extent, with PTT_I and PTT_J . TPR was likewise correlated with PTT_I ($r = -0.58 \pm 0.07$) and PTT_J ($r = -0.52 \pm 0.09$) but also with T_{JJ} ($r = 0.54 \pm 0.07$). Table 3 shows the best-performing univariate and bivariate regression models associated with the weighing scale BCG. Figure 4a shows the correlation plot and Figure 4b the Bland–Altman plot between measured versus regressed CV parameters associated with the weighing scale BCG.

3.3. CV Parameter Estimation with Armband BCG

The signal processing procedure (Figure 2) drastically improved the correlation between the measured versus synthetic weighing scale BCG compared to the correlation between the measured weighing scale versus armband BCG, both at all the individual rest and intervention states as well as across all the rest and intervention states ($r = 0.70$ versus $r = 0.52$ on the average).

The number of subjects available for multivariate linear regression analysis after the outlier removal was ≥ 14 for all the CV parameters associated with the armband BCG except SV (12 subjects). Multivariate linear regression analysis suggested that each of the CV parameters of interest may be accurately estimated by a combination of as few as two characteristic features. In contrast, the best correlation coefficients achieved by univariate characteristic features were in general low (< 0.57) for all CV parameters. For the armband BCG at the univariate level, DP was correlated with PTT_J ($r = -0.36 \pm 0.12$) and PTT_I ($r = -0.34 \pm 0.15$), PP was correlated with PTT_J ($r = -0.53 \pm 0.06$) and PTT_I ($r = -0.48 \pm 0.08$), and SP was correlated with PTT_J ($r = -0.42 \pm 0.11$). SV, CO, and TPR were most strongly correlated with T_{JJ} ($r = 0.34 \pm 0.10$, -0.57 ± 0.10 , and 0.50 ± 0.10). Table 3 shows the best-performing univariate and bivariate regression models associated with the synthetic weighing

scale BCG. Figure 5a shows the correlation plot and Figure 5b the Bland–Altman plot between measured versus regressed CV parameters associated with the synthetic weighing scale BCG.

Table 3. Representative univariate and bivariate regression models associated with weighing scale ballistocardiogram (BCG) and synthetic weighing scale BCG transformed from armband BCG.

	DP	PP	SP	SV	CO	TPR
Weighing Scale BCG: Univariate (r: mean ± SE)						
Features	PTT _I	PTT _I	PTT _I	A _J	T _{JJ}	PTT _I
r	0.81 ± 0.02	0.65 ± 0.05	0.82 ± 0.02	0.50 ± 0.09	0.57 ± 0.11	0.58 ± 0.07
Synthetic Weighing Scale BCG: Univariate (r: mean ± SE)						
Features	PTT _J	PTT _J	PTT _J	T _{JJ}	T _{JJ}	T _{JJ}
r	0.36 ± 0.12	0.53 ± 0.06	0.42 ± 0.11	0.34 ± 0.10	0.57 ± 0.10	0.50 ± 0.10
Weighing Scale BCG: Bivariate (r: mean ± SE)						
Features	PTT _I , A _I	PTT _I , A _{IJ}	PTT _I , A _{JK}	A _J , A _{JK}	T _{JJ} , PTT _J	T _{JJ} , A _{JK} ·PTT _I ²
r	0.85 ± 0.02	0.85 ± 0.02	0.86 ± 0.02	0.73 ± 0.04	0.76 ± 0.05	0.77 ± 0.03
Synthetic Weighing Scale BCG: Bivariate (r: mean ± SE)						
Features	PTT _I , A _I	T _{JJ} , PTT _I	A _{JK} , A _{IJ} ·PTT _I ²	T _{JJ} , A _{IJ} ·PTT _I ²	T _{JJ} , PTT _J	T _{JJ} , A _{JK} ·PTT _I ²
r	0.73 ± 0.04	0.74 ± 0.04	0.73 ± 0.04	0.64 ± 0.06	0.76 ± 0.04	0.75 ± 0.05

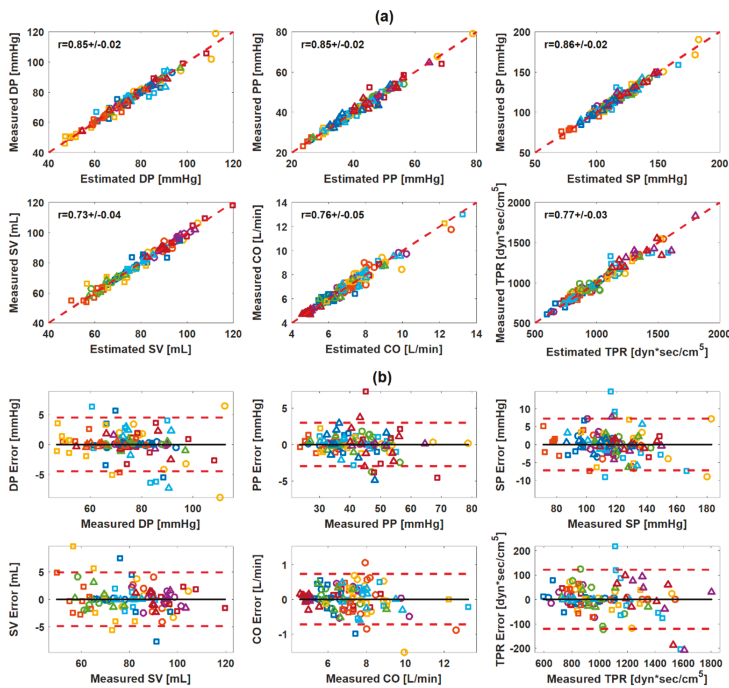


Figure 4. Correlation and Bland–Altman plots between measured versus regressed cardiovascular parameters: weighing scale ballistocardiogram (BCG). (a) Correlation plots. (b) Bland–Altman plots. Black solid line: bias. Red dashed lines: confidence interval.

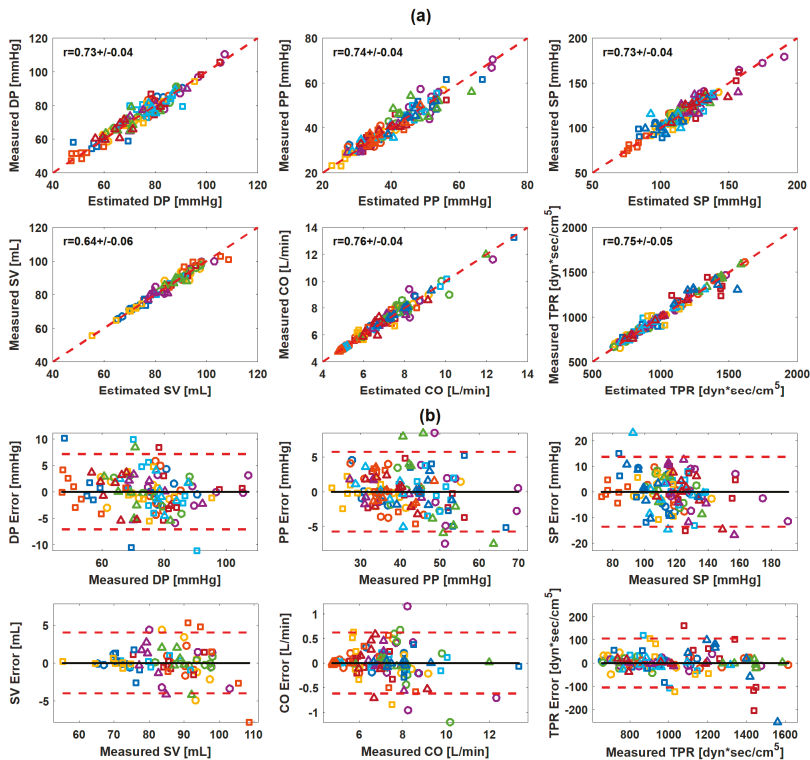


Figure 5. Correlation and Bland–Altman plots between measured versus regressed cardiovascular parameters: synthetic weighing scale ballistocardiogram (BCG) transformed from armband BCG. (a) Correlation plots. (b) Bland–Altman plots.

4. Discussion

Direct measurement of the CV parameters necessitates inconvenient and costly equipment and procedures as well as trained operators. The BCG is closely associated with the aortic BP. Considering the prior success with the pulse contour techniques in deriving the CV parameters from arterial BP waveforms, the BCG may have potential value in estimating the CV parameters. Yet, prior work to investigate the feasibility of estimating the CV parameters from the BCG is quite rare. This work rigorously examined, perhaps for the first time, the relationship between the characteristic features in the limb BCG and the CV parameters.

4.1. Potential of Scale and Armband BCG in CV Parameter Estimation

The results from the regression analysis suggest that the limb BCG may have the potential to enable unobtrusive CV parameter estimation. For the weighing scale BCG, the pair of two features could achieve close correlations with CV parameters ($r \geq 0.85$ for all BP and $r \geq 0.73$ for SV, CO, and TPR on the average; Table 3). For the armband BCG, the pair of two features extracted from the synthetic weighing scale BCG transformed from the armband BCG could likewise achieve close correlations with CV parameters ($r \geq 0.73$ for all BP, $r \geq 0.75$ for CO and TPR, and $r = 0.64$ for SV on the average; Table 3). In general, the weighing scale BCG outperformed the armband BCG. This may be attributed to (i) the more stable measurement setting for the weighing scale BCG relative to the armband BCG and (ii) the errors induced by the transformation of the armband BCG to the synthetic weighing scale BCG (see Section 4.4 for details). Indeed, the upper limb may be more susceptible to involuntary movement

than the lower limb in contact with the weighing scale. Furthermore, the synthetic weighing scale BCG transformed from the armband BCG is not exactly identical to the weighing scale BCG (which may also explain why the features selected for weighing scale BCG and synthetic weighing scale BCG were not identical in Table 3). Combined, these artifacts may result in the deterioration in efficacy of the armband BCG relative to the weighing scale BCG in estimating the CV parameters. Regardless, the degree of correlation between the armband BCG and the CV parameters was still adequate.

The adequate correlation between the armband BCG and the CV parameters appears to have benefited from the signal processing procedure developed in this work to transform the armband BCG to weighing scale BCG. Considering the distinct waveform morphology associated with the weighing scale BCG versus the armband BCG, the efficacy of the signal processing procedure may have a significant implication on the feasibility of standardized analysis of both the BCG. Arguably, the improvement in the correlation between the measured versus synthetic weighing scale BCG compared to the correlation between the measured weighing scale versus armband BCG may suggest that the armband BCG may now be analyzed in the same way as the weighing scale BCG, the analysis method for which is much more established in the sense that the weighing scale BCG may approximately represent the whole-body BCG (i.e., the BCG associated with the movement of the main trunk) [16,42].

4.2. Physiological Relevance of Weighing Scale BCG Features

The characteristic features in the weighing scale BCG exhibiting close correlation with the CV parameters were physiologically relevant as described below (Table 3 and Figure 6).

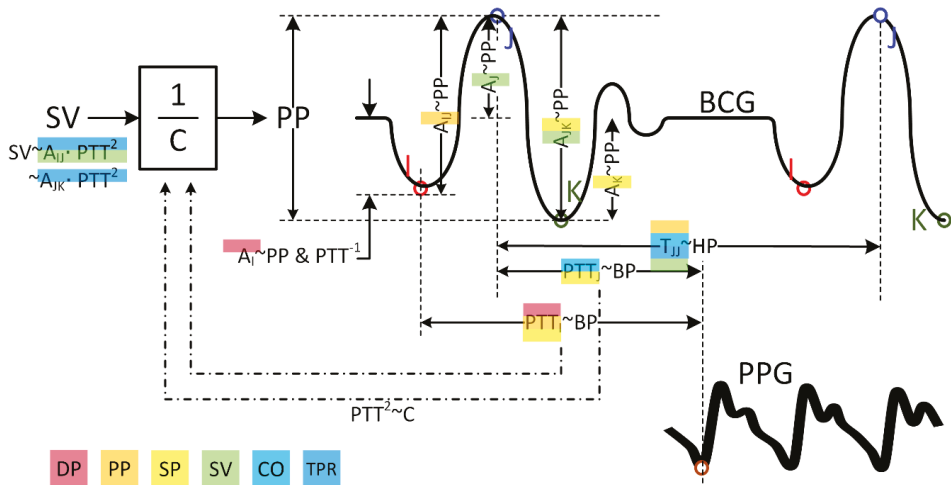


Figure 6. Relationship between the characteristic features in the weighing scale and armband BCG and the CV parameters. BP: blood pressure. DP: diastolic BP. PP: pulse BP. SP: systolic BP. SV: stroke volume. CO: cardiac output. TPR: total peripheral resistance. HP: heart period. C: arterial compliance. “X~Y” means that X and Y are proportional.

First, physiologically relevant weighing scale BCG features exhibited close correlation to the CV parameters in the univariate regression analysis. The correlation of DP with PTT_I and PTT_J is consistent with the established fact that DP is correlated closely to PTT [49]. The correlation of PP with PTT_I and PTT_J may be understood by the fact that PP may be (at least in a local sense) inversely proportional to PTT [50–53]. The correlation of PP with A_{JK} and A_{JJ} may be understood by the fact that the amplitude features A_J and A_{JK} may be the surrogates of ascending aortic and descending aortic PP [16] as well as the fact that an increase in PP may lead to an increase in the overall BCG amplitude. The correlation of SP with PTT_I and PTT_J may be understood from the correlation between DP and PTTs in conjunction

with the fact that the hemodynamic interventions considered in this work elicited concurrent increases in both DP and SP. Likewise, physiologically relevant weighing scale BCG features were properly correlated with SV, CO, and TPR in the univariate regression analysis, though not as strong as BP. The correlation of SV with A_j is reasonable in that A_j may be the surrogate of ascending aortic PP and that SV and PP are proportional to each other if the arterial compliance (AC) does not change largely ($PP = SV/AC$) [29]. The correlation of CO with T_{jj} , and to a lesser extent, with PTT_1 and PTT_j may also be reasonable by noting that CO is the product of SV and heart rate, T_{jj} is a surrogate of heart rate, and PTT_1 and PTT_j are correlated with PP (which is proportional to SV). The negative correlation of TPR and PTT_1 and PTT_j appears reasonable given that the changes in BP and TPR are in phase (Figure 3). In contrast, the positive correlation of TPR with T_{jj} is counter-intuitive in that BP, heart rate, and TPR mostly change in the same direction, except in BH (the change in heart rate may be deduced from SV and CO in Figure 3 as CO/SV). It is speculated that the large inverse change in TPR and heart rate in BH appears to dominate the relatively small in-phase changes in the remaining hemodynamic interventions and, thereby, yielded the positive correlation between TPR and T_{jj} . Hence, the positive correlation between TPR and T_{jj} as observed in this work may not generalize.

Second, the weighing scale BCG features selected in the bivariate regression analysis were also quite physiologically relevant (Figure 6). DP was regressed with PTT_1 and A_1 ($r = 0.85 \pm 0.02$), which is relevant in that A_1 may be inversely proportional to BP since a decrease in PTT (corresponding to an increase in BP) may be associated with a decrease in A_1 [16,42]. PP was regressed with PTT_1 and A_{jj} ($r = 0.85 \pm 0.02$) consistently to the univariate regression analysis. SP was regressed with PTT_1 and A_{jk} ($r = 0.86 \pm 0.02$), which is relevant in that A_{jk} may represent PP as mentioned above [16]. SV was regressed with A_j and A_{jk} ($r = 0.73 \pm 0.04$), which is supported by the close relationship between these amplitude features and PP [16] and the proportionality between PP and SV under small AC change [29]. SV was also regressed well with T_{jj} and RMS ($r = 0.73 \pm 0.04$), which may be due to the inversely proportional change between SV and HR (Figure 3; which may be specific to the data analyzed in our work due to large changes in HR and thus may not generalize) and the proportional association between the amplitude features and RMS. CO was regressed with T_{jj} (which is consistent with the univariate regression case) and PTT_j ($r = 0.76 \pm 0.05$). The correlation between CO and PTT_j appears relevant because CO and SV are proportional, SV and PP may be proportional, and PP is locally inversely proportional to PTT (as stated above) [50–53]. TPR was regressed with T_{jj} (consistent with the univariate regression analysis) and $A_{jj} \cdot PTT_1^2$ ($r = 0.77 \pm 0.03$). Considering that A_{jj} may serve as a surrogate of PP (as stated above) and that PTT^2 is proportional to AC according to the wave speed equation [49], $A_{jj} \cdot PTT_1^2$ may be regarded as a surrogate of SV. Hence, it may qualify for a feature to track the trend of TPR given its inversely proportional relationship to TPR ($r = -0.32 \pm 0.08$; Figure 3).

4.3. Physiological Relevance of Armband BCG Features

The characteristic features in the synthetic weighing scale BCG transformed from the armband BCG exhibiting close correlation with the CV parameters were physiologically relevant to a large extent as described below (Table 3 and Figure 6).

First, many physiologically relevant synthetic weighing scale BCG features exhibited correlation to the CV parameters in the univariate regression analysis consistently to the weighing scale BCG. However, the degree of correlation was not as strong as the weighing scale BCG.

Second, the synthetic weighing scale BCG features selected in the bivariate regression analysis were likewise quite physiologically relevant and largely consistent with the weighing scale BCG (Figure 6). DP was regressed with PTT_1 and A_1 ($r = 0.73 \pm 0.04$). PP was regressed with PTT_1 and T_{jj} ($r = 0.74 \pm 0.04$). PTT_1 may have been selected since it changed in the opposite direction to DP, PP, and SP in this work. T_{jj} may have been selected since it exhibited a positive correlation with SV in this work (which may be deduced from SV and CO in Figure 3). SP was best regressed with A_{jk} and $A_{jj} \cdot PTT_1^2$ ($r = 0.73 \pm 0.04$). This correlation may be understood in that SP and PP mostly changed in the same direction in response to the interventions considered in this work (Figure 3).

However, SP was also well regressed with the pair of PTT_J and an amplitude feature (e.g., $PTT_J \cdot A_K$: $r = 0.72 \pm 0.05$). These correlations may be readily interpreted in that PTT and amplitude features may represent DP and PP, respectively [1,16]. SV was regressed with T_{JJ} and $A_{JJ} \cdot PTT_1^2$ ($r = 0.64 \pm 0.06$). T_{JJ} may have been selected since it exhibited a positive correlation with SV in this work as stated above, while $A_{JJ} \cdot PTT_1^2$ may be a meaningful surrogate of SV as stated earlier. CO was regressed with T_{JJ} and PTT_J ($r = 0.76 \pm 0.04$), which may be relevant in that T_{JJ} and PTT_J may represent heart rate and PP (which in general correlates with SV; also DP and PP varied in the same direction in response to the hemodynamic interventions considered in this work as shown in Figure 3), respectively. TPR was regressed with T_{JJ} and $A_{JK} \cdot PTT_1^2$ ($r = 0.75 \pm 0.05$) similarly to the weighing scale BCG.

4.4. Summarizing Remarks and Study Limitations

In summary, the results obtained from this work provide several important implications. First, the characteristic features in the limb BCG have the potential for unobtrusive estimation of CV parameters. Indeed, for both the weighing scale and armband BCG, the pair of as few as two features could achieve close correlation with the CV parameters. Second, the characteristic features selected by the multivariate regression analysis appeared to be largely interpretable (meaning that the selected characteristic features were to a large extent congruent with the physiological insights [16,42]). Indeed, despite the fact that the multivariate regression analysis conducted in this work was predominantly a data-mining exercise, the majority of the characteristic features selected by the analysis were physiologically relevant and consistent with the findings derived from our prior mathematical model-based analysis of the BCG [16,42] (see Sections 4.2 and 4.3 for details). Hence, the BCG features identified to exhibit close association with CV parameters in this work may be generalizable to other independent datasets. Third, PTT may make significant contributions in CV parameter estimation. Indeed, PTT was selected in all the bivariate regression analyses derived for BP (DP, PP, and SP) in this work. In comparison with our prior work that investigated the association between the characteristic features in the wrist BCG and BP ($r = 0.75$ for both DP and SP on the average when three predictors were employed) [14], this work achieved much higher correlation with less number of predictors (i.e., two) by including PTT. From this standpoint, it may be of interest to see the potential value of pulse arrival time (PAT) in further improving the association between the limb BCG and CV parameters. In fact, existing work suggests that PAT may serve as a good characteristic feature for SP [49,50,54] as well as CV parameters via the pre-ejection period (which has implications on the heart contractility). One practical consideration may be that the use of PAT necessitates the measurement of the ECG, which generally requires conventional electrodes or two-handed user maneuvers [55]. In this regard, an accuracy–convenience trade-off may need to be made.

This study has a few limitations. First, the signal processing procedure for transforming the armband BCG to the synthetic weighing scale BCG was empiric. The application of two integrations to the armband BCG to yield the armband displacement can be well justified in that the armband displacement may be identical to the weighing scale BCG (which is essentially a displacement measurement) if the body is assumed to be perfectly rigid. However, our choice of the cut-off frequency for the post-integration high-pass filtering of the synthetic weighing scale BCG may not be optimal: although the cut-off frequency was chosen consistently using a set procedure (Figure 2), the procedure was primarily based on empiric attempts to make the power spectra associated with the weighing scale BCG and the synthetic weighing scale BCG look comparable in a qualitative sense (i.e., the amplitudes and locations of the spectral peaks). In future work, this weakness needs to be investigated, so that more effective and robust signal processing procedure for transforming the armband BCG to the weighing scale BCG can be conceived and developed. One possibility may be to explicitly account for the body compliance in the signal processing procedure, by incorporating into it a mathematical model of the body biomechanics that can predict the alteration of the limb BCG waveform due to the compliance and elasticity of the tissues and joints. Second, the study participants were quite homogenous in terms of age and CV health. It is important to investigate if the findings from this work remain valid in

a wider group of subjects (e.g., subjects with pacemakers [56] and cardiovascular disease [57]). Third, the CV parameters were measured using a non-invasive device (ccNexfin). Despite the widespread use of the device used in this study in research and its demonstrated accuracy [26], it is possible that the reference CV parameters were associated with inaccuracies due to, e.g., group-average formula used in its pulse contour algorithm [26]. Fourth, this work only investigated the feasibility of estimating the trend of CV parameters. In addition, the regression analysis was conducted in a subject-specific setting. To leverage the results of this work practically in real-world conditions, the characteristic features in the BCG must be calibrated to CV parameters. Hence, future work must investigate subject-specific calibration of the characteristic features in the BCG to CV parameters.

5. Conclusions

In this work, we demonstrated that (i) the characteristic features in the limb BCG exhibit close correlation with the CV parameters, and (ii) the characteristic features representative of the CV parameters are largely relevant from a physiological standpoint. Future work must be conducted to translate the findings of this work to more realistic CV parameter estimation techniques.

Author Contributions: Conceptualization, J.H., R.M., and L.X.; methodology, J.H. and Y.Y.; software, Y.Y. and S.S.; validation, Y.Y.; formal analysis, Y.Y.; investigation, Y.Y., J.H., R.M., and L.X.; resources, J.H., C.K., and L.X.; data curation, J.H., A.M., S.S., and C.K.; writing—original draft preparation, J.H. and Y.Y.; writing—review and editing, Y.Y., S.S., A.M., C.K., L.X., R.M., and J.H.; visualization, Y.Y. and J.H.; supervision, J.H., R.M., and L.X.; project administration, J.H.; funding acquisition, J.H. and L.X.

Funding: This work was supported in part by the University of Maryland under the UM Ventures Seed Grant, the China Scholarship Council, and the National Natural Science Foundation of China under Grant 61773110. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the University of Maryland.

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

References

- Kim, C.-S.; Carek, A.M.; Inan, O.T.; Mukkamala, R.; Hahn, J.-O. Ballistocardiogram-Based Approach to Cuffless Blood Pressure Monitoring: Proof of Concept and Potential Challenges. *IEEE Trans. Biomed. Eng.* **2018**, *65*, 2384–2391. [[CrossRef](#)] [[PubMed](#)]
- Martin, S.L.; Carek, A.M.; Kim, C.; Ashouri, H.; Inan, O.T.; Hahn, J.-O.; Mukkamala, R. Weighing Scale-Based Pulse Transit Time is a Superior Marker of Blood Pressure than Conventional Pulse Arrival Time. *Sci. Rep.* **2016**, *6*, 39273. [[CrossRef](#)] [[PubMed](#)]
- Ashouri, H.; Orlandic, L.; Inan, O.T. Unobtrusive Estimation of Cardiac Contractility and Stroke Volume Changes using Ballistocardiogram Measurements on a High Bandwidth Force Plate. *Sensors* **2016**, *16*, 787. [[CrossRef](#)] [[PubMed](#)]
- Inan, O.T.; Etemadi, M.; Wiard, R.M.; Giovangrandi, L.; Kovacs, G.T.A. Robust Ballistocardiogram Acquisition for Home Monitoring. *Physiol. Meas.* **2009**, *30*, 169–185. [[CrossRef](#)] [[PubMed](#)]
- Inan, O.T.; Park, D.; Member, S.; Giovangrandi, L.; Kovacs, G.T.A. Noninvasive Measurement of Physiological Signals on a Modified Home Bathroom Scale. *IEEE Trans. Biomed. Eng.* **2012**, *59*, 2137–2143. [[CrossRef](#)]
- Shin, J.H.; Lee, K.M.; Park, K.S. Non-Constrained Monitoring of Systolic Blood Pressure on a Weighing Scale. *Physiol. Meas.* **2009**, *30*, 679–693. [[CrossRef](#)] [[PubMed](#)]
- Campo, D.; Khettab, H.; Yu, R.; Genain, N.; Edouard, P.; Buard, N.; Boutouyrie, P. Measurement of Aortic Pulse Wave Velocity With a Connected Bathroom Scale. *Am. J. Hypertens.* **2017**, *30*, 876–883. [[CrossRef](#)]
- Jung, D.W.; Hwang, S.H.; Yoon, H.N.; Lee, Y.G.; Jeong, D.; Park, K.S. Nocturnal Awakening and Sleep Efficiency Estimation Using Unobtrusively Measured Ballistocardiogram. *IEEE Trans. Biomed. Eng.* **2014**, *61*, 131–138. [[CrossRef](#)]
- Shin, J.H.; Chee, Y.J.; Jeong, D.; Park, K.S. Nonconstrained Sleep Monitoring System and Algorithms Using Air-Mattress With Balancing Tube Method. *IEEE J. Biomed. Heal. Inform.* **2010**, *14*, 147–156.
- Pinheiro, E.; Postolache, O.; Girão, P. Non-Intrusive Device for Real-Time Circulatory System Assessment with Advanced Signal Processing Capabilities. *Meas. Sci. Rev.* **2010**, *10*, 166–175. [[CrossRef](#)]

11. Akhbardeh, A.; Junnila, S.; Koivistoinen, T.; Varri, A. An intelligent Ballistocardiographic Chair using a Novel SF-ART Neural Network and Biorthogonal Wavelets. *J. Med. Syst.* **2007**, *31*, 69–77. [[CrossRef](#)] [[PubMed](#)]
12. Alametsä, J.; Palomäki, A.; Viik, J. Short and Longer Term Repeatability of Ballistocardiography in a Sitting Position with EMFi Sensor. *Med. Biol. Eng. Comput.* **2011**, *49*, 881–889. [[CrossRef](#)] [[PubMed](#)]
13. Da He, D.; Winokur, E.S.; Sodini, C.G. An Ear-Worn Vital Signs Monitor. *IEEE Trans. Biomed. Eng.* **2015**, *62*, 2547–2552. [[CrossRef](#)] [[PubMed](#)]
14. Yousefian, P.; Shin, S.; Mousavi, A.; Kim, C.S.; Mukkamala, R.; Jang, D.G.; Ko, B.H.; Lee, J.; Kwon, U.K.; Kim, Y.H.; et al. Data Mining Investigation of the Association between a Limb Ballistocardiogram and Blood Pressure. *Physiol. Meas.* **2018**, *39*, 075009. [[CrossRef](#)]
15. Wiens, A.D.; Etemadi, M.; Roy, S.; Klein, L.; Inan, O.T. Toward Continuous, Noninvasive Assessment of Ventricular Function and Hemodynamics: Wearable Ballistocardiography. *IEEE J. Biomed. Heal. Inform.* **2015**, *19*, 1435–1442. [[CrossRef](#)]
16. Kim, C.-S.; Ober, S.L.; McMurtry, M.S.; Finegan, B.A.; Inan, O.T.; Mukkamala, R.; Hahn, J.-O. Ballistocardiogram: Mechanism and Potential for Unobtrusive Cardiovascular Health Monitoring. *Sci. Rep.* **2016**, *6*, 31297. [[CrossRef](#)]
17. Mueller, R.L.; Sanborn, T.A. The History of Interventional Cardiology: Cardiac Catheterization, Angioplasty, and Related Interventions. *Am. Heart J.* **1995**, *129*, 146–172. [[CrossRef](#)]
18. Wesseling, K.H.; De Wit, B.; van der Hoeven, G.M.A.; van Goudoever, J.; Settels, J.J. Physiological, Calibrating Finger Vascular Physiology for Finapres. *Homeostasis* **1995**, *36*, 67–82.
19. Imholz, B.P.; Wieling, W.; van Montfrans, G.A.; Wesseling, K.H. Fifteen Years Experience with Finger Arterial Pressure Monitoring: Assessment of the Technology. *Cardiovasc. Res.* **1998**, *38*, 605–616. [[CrossRef](#)]
20. Eckerle, J.S. Tonometry, Arterial. In *Encyclopedia of Medical Devices and Instrumentation*; Webster, J.G., Ed.; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2006; pp. 402–410.
21. Hamilton, W.F.; Riley, R.L.; Attyah, A.M.; Cournand, A.; Fowell, D.M.; Himmelstein, A.; Noble, R.P.; Remington, J.W.; Richards, D.W., Jr.; Wheeler, N.C.; et al. Comparison of the Fick and Dye-Injection Methods of Measuring the Cardiac Output in Man. *Am. J. Physiol.* **2016**, *153*, 309–321. [[CrossRef](#)]
22. Schiller, N.B.; Acquatella, H.; Ports, T.A.; Drew, D.; Goerke, J.O.N.; Ringertz, H.; Silverman, N.H.; Brundage, B.; Botvinick, E.H.; Boswell, R.; et al. Left Ventricular Volume from Paired Biplane Two-Dimensional Echocardiography. *Circulation* **1979**, *60*, 547–555. [[CrossRef](#)] [[PubMed](#)]
23. Bernstein, D.P.; Lemmens, H.J.M. Stroke Volume Equation for Impedance Cardiography. *Med. Biol. Eng. Comput.* **2005**, *43*, 443–450. [[CrossRef](#)] [[PubMed](#)]
24. Braun, F.; Proenc, M.; Adler, A.; Riedel, T.; Thiran, P.; Solà, J. Accuracy and Reliability of Noninvasive Stroke Volume Monitoring via ECG-Gated 3D Electrical Impedance Tomography in Healthy Volunteers. *PLoS ONE* **2018**, *13*, e0191870. [[CrossRef](#)] [[PubMed](#)]
25. Reinsner, A.T.; Xu, D.; Ryan, K.L.; Convertino, V.A.; Rickards, C.A.; Mukkamala, R. Monitoring Non-Invasive Cardiac Output and Stroke Volume during Experimental Human Hypovolaemia and Resuscitation. *Br. J. Anaesth.* **2011**, *106*, 23–30. [[CrossRef](#)] [[PubMed](#)]
26. Truijen, J.; Van Lieshout, J.J.; Wesselink, W.A.; Westerhof, B.E. Noninvasive Continuous Hemodynamic Monitoring. *J. Clin. Monit. Comput.* **2012**, *26*, 267–278. [[CrossRef](#)] [[PubMed](#)]
27. Bogert, L.W.J.; Lieshout, J.J. Van Non-Invasive Pulsatile Arterial Pressure and Stroke Volume Changes from the Human Finger. *Exp. Physiol.* **2005**, *90*, 437–446. [[CrossRef](#)] [[PubMed](#)]
28. Su, C.-H.; Liu, S.-H.; Tan, T.-H.; Lo, C.-H. Using the Pulse Contour Method to Measure the Changes in Stroke Volume during a Passive Leg Raising Test. *Sensors* **2018**, *18*, 3420. [[CrossRef](#)]
29. Fazeli, N.; Hahn, J.O. Estimation of Cardiac Output and Peripheral Resistance Using Square-Wave-Approximated Aortic Flow Signal. *Front. Physiol.* **2012**, *3*, 298. [[CrossRef](#)] [[PubMed](#)]
30. Conn, N.J.; Schwarz, K.Q.; Borkholder, D.A. In-Home Cardiovascular Monitoring System for Heart Failure: Comparative Study. *JMIR MHealth UHealth* **2019**, *7*, e12419. [[CrossRef](#)]
31. Kim, C.-S.; Carek, A.M.; Mukkamala, R.; Inan, O.T.; Hahn, J.-O. Ballistocardiogram as Proximal Timing Reference for Pulse Transit Time Measurement: Potential for Cuffless Blood Pressure Monitoring. *IEEE Trans. Biomed. Eng.* **2015**, *62*, 2657–2664. [[CrossRef](#)]
32. Lee, K.J.; Roh, J.; Cho, D.; Hyeon, J.; Kim, S. A Chair-Based Unconstrained/Nonintrusive Cuffless Blood Pressure Monitoring System Using a Two-Channel Ballistocardiogram. *Sensors* **2019**, *19*, 595. [[CrossRef](#)] [[PubMed](#)]

33. Javaid, A.Q.; Ashouri, H.; Tridandapani, S.; Inan, O.T. Elucidating the Hemodynamic Origin of Ballistocardiographic Forces: Toward Improved Monitoring of Cardiovascular Health at Home. *IEEE J. Transl. Eng. Heal. Med.* **2016**, *4*, 1900208. [[CrossRef](#)] [[PubMed](#)]
34. Martín-Yebra, A.; Landreani, F.; Casellato, C.; Pavan, E.; Migeotte, P.-F.; Frigo, C.; Martínez, J.P.; Caiani, E.G. Evaluation of Respiratory- and Postural-Induced Changes on the Ballistocardiogram Signal by Time Warping Averaging. *Physiol. Meas.* **2017**, *38*, 1426–1440. [[CrossRef](#)] [[PubMed](#)]
35. Scarborough, W.R.; Folk III, E.F.; Smith, P.M.; Condon, J.H. The Nature of Records from Ultra-Low Frequency Ballistocardiographic Systems and Their Relation to Circulatory Events. *Am. J. Cardiol.* **1958**, *2*, 613–641. [[CrossRef](#)]
36. Rappaport, M.B. Displacement, Velocity, and Acceleration Ballistocardiograms as Registered with an Undamped Bed of Ultralow Natural Frequency: II. Instrumental Considerations. *Am. Heart J.* **1956**, *52*, 643–652. [[CrossRef](#)]
37. Starr, I.; Horwitz, O.; Mayock, R.L.; Krumbhaar, E.B. Standardization of the Ballistocardiogram by Simulation of the Heart's Function at Necropsy; With a Clinical Method for the Estimation of Cardiac Strength and Normal Standards for It. *Circulation* **1950**, *1*, 1073–1096. [[CrossRef](#)]
38. Leys, C.; Ley, C.; Klein, O.; Bernard, P.; Licata, L. Detecting Outliers: Do Not Use Standard Deviation around the Mean, Use Absolute Deviation around the Median. *J. Exp. Soc. Psychol.* **2013**, *49*, 764–766. [[CrossRef](#)]
39. Gaddum, N.R.; Alastruey, J.; Beerbaum, P.; Chowienzyk, P.; Schaeffter, T. A Technical Assessment of Pulse Wave Velocity Algorithms Applied to Non-Invasive Arterial Waveforms. *Ann. Biomed. Eng.* **2013**, *41*, 2617–2629. [[CrossRef](#)]
40. Lee, S.; Chang, J.H. Oscillometric Blood Pressure Estimation Based on Deep Learning. *IEEE Trans. Ind. Inform.* **2017**, *13*, 461–472. [[CrossRef](#)]
41. Lee, S.; Chang, J. Deep Boltzmann Regression With Mimic Features for Oscillometric Blood Pressure Estimation. *IEEE Sens. J.* **2017**, *17*, 5982–5993. [[CrossRef](#)]
42. Yousefian, P.; Shin, S.; Mousavi, A.S.; Kim, C.-S.; Finegan, B.; McMurtry, M.S.; Mukkamala, R.; Jang, D.-G.; Kwon, U.; Kim, Y.H.; et al. Physiological Association between Limb Ballistocardiogram and Arterial Blood Pressure Waveforms: A Mathematical Model-Based Analysis. *Sci. Rep.* **2019**, *9*, 5146. [[CrossRef](#)] [[PubMed](#)]
43. Helman, A. *The Finest Peaks-Prominence and Other Mountain Measures*; Trafford Publishing: Victoria, BC, Canada, 2005; ISBN 1-4120-5995-x.
44. Griffié, J.; Boelen, L.; Burn, G.; Cope, A.P.; Owen, D.M. Topographic Prominence as a Method for Cluster Identification in Single-Molecule Localisation Data. *J. Biophotonics* **2015**, *8*, 925–934. [[CrossRef](#)] [[PubMed](#)]
45. Allen, M.T.; Obrist, P.A.; Sherwood, A.; Crowell, M.D. Evaluation of Myocardial and Peripheral Vascular Responses during Reaction Time, Mental Arithmetic, and Cold Pressor Tasks. *Psychophysiology* **1987**, *24*, 648–656. [[CrossRef](#)] [[PubMed](#)]
46. Turankar, A.V.; Jain, S.; Patel, S.B.; Sinha, S.R.; Joshi, A.D.; Vallish, B.N.; Mane, P.R.; Turankar, S.A. Effects of Slow Breathing Exercise on Cardiovascular Functions, Pulmonary Functions & Galvanic Skin Resistance in Healthy Human Volunteers—A Pilot Study. *Indian J. Med. Res.* **2013**, *137*, 916–921. [[PubMed](#)]
47. Adhana, R.; Agarwal, M.; Gupta, R.; Divedi, J. Effect of Slow Breathing Training on Heart Rate, Spontaneous Respiratory Rate and Pattern of Breathing. *Int. J. Res. Med. Sci.* **2016**, *4*, 1027–1030. [[CrossRef](#)]
48. Grunovas, A.; Trinkunas, E.; Buliuolis, A.; Venskaityte, E.; Poderys, J.; Poderiene, K. Cardiovascular Response to Breath-Holding Explained by Changes of the Indices and their Dynamic Interactions. *Biol. Syst. Open Access* **2016**, *5*, 152. [[CrossRef](#)]
49. Mukkamala, R.; Hahn, J.; Inan, O.T.; Mestha, L.K.; Kim, C.; Hakan, T. Toward Ubiquitous Blood Pressure Monitoring via Pulse Transit Time: Theory and Practice. *IEEE Trans. Biomed. Eng.* **2015**, *62*, 1879–1901. [[CrossRef](#)]
50. Ding, X.; Zhang, Y.; Liu, J.; Dai, W.; Tsang, H.K. Continuous Cuffless Blood Pressure Estimation Using Pulse Transit Time and Photoplethysmogram Intensity Ratio. *IEEE Trans. Biomed. Eng.* **2016**, *63*, 964–972. [[CrossRef](#)]
51. Seo, J.; Pietrangolo, S.J.; Lee, H.-S.; Sodini, C.G. Noninvasive Arterial Blood Pressure Waveform Monitoring Using Two-Element Ultrasound System. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2015**, *62*, 776–784. [[CrossRef](#)]

52. Beulen, B.W.; Bijmens, N.; Koutsouridis, G.G.; Brands, P.J.; Rutten, M.C.; van de Vosse, F.N. Toward Noninvasive Blood Pressure Assessment in Arteries by Using Ultrasound. *Ultrasound Med. Biol.* **2011**, *37*, 788–797. [[CrossRef](#)]
53. Vappou, J.; Luo, J.; Okajima, K.; Di Tullio, M.; Konofagou, E.E. Non-Invasive Measurement of Local Pulse Pressure by Pulse Wave-Based Ultrasound Manometry (PWUM). *Physiol. Meas.* **2011**, *32*, 1653–1662. [[CrossRef](#)] [[PubMed](#)]
54. Miao, F.; Fu, N.; Zhang, Y.T.; Ding, X.R.; Hong, X.; He, Q.; Li, Y. A Novel Continuous Blood Pressure Estimation Approach Based on Data Mining Techniques. *IEEE J. Biomed. Heal. Inform.* **2017**, *21*, 1730–1740. [[CrossRef](#)] [[PubMed](#)]
55. Kang, J.M.; Yoo, T.; Kim, H.C. A wrist-worn integrated health monitoring instrument with a tele-reporting device for telemedicine and telecare. *IEEE Trans. Instrum. Meas.* **2006**, *55*, 1655–1661. [[CrossRef](#)]
56. Theorell, T.; Edhag, O.; Fagrell, B. Non-Invasive Methods for Evaluating the Importance of Heart Rate and Atrial Activity in Cardiac Pacing. *Acta Med. Scand.* **1978**, *203*, 497–502. [[CrossRef](#)] [[PubMed](#)]
57. Theorell, T.; Rahe, R.H. Life Change Events, Ballistocardiography and Coronary Death. *J. Human Stress* **1975**, *1*, 18–24. [[CrossRef](#)] [[PubMed](#)]



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).

Article

In-Ear Pulse Rate Measurement: A Valid Alternative to Heart Rate Derived from Electrocardiography?

Stefanie Passler *, Niklas Müller and Veit Senner

Technical University of Munich, Department of Mechanical Engineering, Professorship of Sport Equipment and Materials, Boltzmannstraße 15, D-85747 Garching, Germany

* Correspondence: stefanie.passler@tum.de; Tel.: +49-89-289-15380; Fax: +49-89-289-15389

Received: 10 July 2019; Accepted: 19 August 2019; Published: 21 August 2019

Abstract: Heart rate measurement has become one of the most widely used methods of monitoring the intensity of physical activity. The purpose of this study was to assess whether in-ear photoplethysmographic (PPG) pulse rate (PR) measurement devices represent a valid alternative to heart rate derived from electrocardiography (ECG), which is considered a gold standard. Twenty subjects (6 women, 14 men) completed one trial of graded cycling under laboratory conditions. In the trial, PR was recorded by two commercially available in-ear devices, the Dash Pro and the Cosinuss^oOne. They were compared to HR measured by a Bodyguard2 ECG. Validity of the in-ear PR measurement devices was tested by ANOVA, mean absolute percentage errors (MAPE), intra-class correlation coefficient (ICC), and Bland–Altman plots. Both devices achieved a MAPE $\leq 5\%$. Despite excellent to good levels of agreement, Bland–Altman plots showed that both in-ear devices tend to slightly underestimate the ECG's HR values. It may be concluded that in-ear PPG PR measurement is a promising technique that shows accurate but imprecise results under controlled conditions. However, PPG PR measurement in the ear is sensitive to motion artefacts. Thus, accuracy and precision of the measured PR depend highly on measurement site, stress situation, and exercise.

Keywords: photoplethysmography; heart rate; consumer-wearable devices; in-ear; validation; optical pulse rate monitoring; pulse rate

1. Introduction

As a result of the development of mobile heart rate monitors, heart rate has become one of the most widely used methods for controlling the general state of health and, particularly, the intensity of physical activity [1]. In 1938, Hertzman [2] first introduced photoplethysmography (PPG) as an alternative to electrocardiographic (ECG) heart rate monitoring. Since then, PPG has gained increasing popularity [3] and, with increasing technological improvements, it is used as an alternative pulse rate measurement method in wearable devices. In his review article, Toshiyo Tamura [4] provides an overview of the parameters that influence PPG signals. Besides the wavelength of light, contact force, motion artefacts, ambient temperature, and light intensity, the anatomical measurement location also influences PPG signals. Currently, PPG signals can be measured at the wrist [5–7], upper and lower arm [8–10], finger [11], esophageal region [12], forehead [13,14], and the ear, respectively. Ear-worn devices are defined as devices worn in or on the ear. Specifically, pulse rate measurement at the earlobe [15], the external ear cartilage [16,17], the superior and inferior auricular region [18–21], and the external auditory canal [22–24] has already been discussed in several studies as an alternative to ECG heart rate monitoring. Besides wrist-worn devices, ear-worn devices are probably the most common application of PPG.

The measuring principle of a PPG sensor is based on optical variations in pulsatile blood-flow volume. A PPG sensor includes a light-emitting diode (LED) and a photodetector (PD). The arrangement

of these two components determines the mode of PPG. There are two modes, transmission and reflectance. In transmission mode, the PD, located opposite to the LED, captures the light transmitted through the tissue. Especially in medicine, this mode is frequently used. In it, the finger is the most common measurement location. A clip enables the opposite positioning of LED and PD. However, blood flow in the extremities, e.g., in the fingers, can be severely impaired due to circulatory disorders, for example. This may lead to unreliable and invalid measurement of the pulse rate. In reflectance mode, the PD captures the light reflected from bony structures, tissue, and blood vessels. Moreover, LED and PD are located next to each other. Thus, the reflectance mode is hardly limited to certain anatomical measurement locations [25,26]. Obviously, this is a reason for the suitability of PPG's reflectance mode in wearable devices. However, the intensity of the reflected or backscattered light is strongly dependent on the anatomical conditions of the measuring location [13]. On the forehead, where skin is very thin, but many blood vessels are present, a reliable signal can usually be recorded. On anatomical locations with a lower density of blood vessels and bony structures, the detected light intensity is usually lower.

Common wavelengths of PPG are between 500 nm and 1100 nm. This corresponds to the range from green-yellow to infrared light. The absorption rate of the light is mainly influenced by the pigment melanin and the water content in the tissue. The correlation between melanin and the absorption rate can be described as inversely proportional; the longer the wavelength, the more light is absorbed by melanin [27]. Water absorbs light in the ultraviolet and upper infrared range, whereas red and near infrared (NI) light is slightly absorbed by water [28]. This shows that the spectrum from NI to infrared light should be used for measuring deeper tissue structures. Hence, the transmission mode of PPG often uses wavelengths of 500–600 nm. Because of the stronger plethysmography signal, the green spectrum of light is more suited for optical pulse rate measurement [9]. The use of green LEDs allows a more accurate detection of the pulse rate [9,29–31].

The main difference between heart rate and pulse rate is the duration the pulse wave takes to complete the distance from the heart to the measurement site. This time is named the pulse transit time (PTT). It is determined as the time lag between the peak of the R-wave on the ECG and the peak value of the corresponding pulse at the measurement site, measured by PPG.

Altogether, several studies show that PPG is a promising method with regard to its use in wearable devices, especially when applied on the ear. However, most of the above-mentioned studies evaluated non-commercial devices, applied on the ear. So far, to our knowledge, no comprehensive studies have been published on the validity of commercial in-ear pulse rate monitoring devices using PPG. Thus, detailed validation of in-ear pulse rate monitoring devices using PPG is still pending.

The purpose of this study is to evaluate the validity of two consumer-wearable in-ear devices with respect to an ECG. Within the scope of this study, heart rate and PPG signals were first recorded simultaneously and then the rate variability of the two methods was compared.

2. Materials and Methods

2.1. Participants

Twenty healthy subjects (men = 14, women = 6) participated in the study and provided their written informed consent. Study protocol was proceeded under medical supervision of the Department of Prevention, Rehabilitation, and Sports Medicine of the Technical University of Munich. Additionally, the study was conducted in accordance with the Declaration of Helsinki. Participants who were eligible received detailed information on the purpose and methods of the study, as well as on data treatment and confidentiality according to the General Data Protection Regulation (2016/679) of the European Parliament and the Council of 27 April 2016 [32] and its Corrigendum of 23 May 2018 [33]. The characteristics of the sample population are shown in Table 1.

Table 1. Participant characteristics. Values are means \pm standard deviation.

	Male (N = 14)	Female (N = 6)	All Participants (N = 20)
Age (years)	22.1 \pm 1.8	22.5 \pm 2.1	22.3 \pm 2.0
Body mass (kg)	74.6 \pm 8.4	57.8 \pm 4.2	69.6 \pm 11.0

2.2. Instruments

Within the scope of this study, the validity of the Cosinuss^oOne and the Dash Pro in-ear pulse rate measurement devices was investigated.

2.2.1. The Dash Pro

The Dash Pro (Bragi, Munich, Germany) is a wireless headset, equipped with sensor technology that is able to provide real-time feedback of recorded movements and pulse rate. The device consists of left and right headphones that communicate wirelessly with each other. The Dash Pro measures pulse rate in the external auditory canal using infrared light by reflection measurement. The device itself is available in one size, but can be fitted to the user's ear with interchangeable silicone caps in sizes XS to L. Figure 1a shows the Dash Pro with a silicone cap in size M. The blue arrows mark the diodes that allow heart rate detection by reflection measurement. Figure 1b shows the right device, worn in the right ear of a participant.

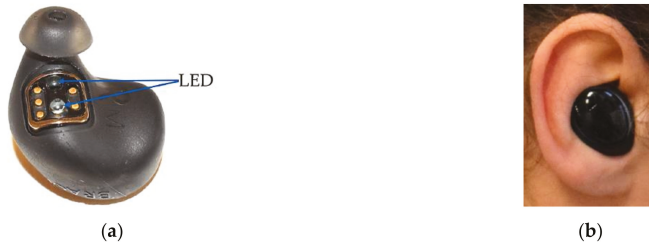


Figure 1. (a) The Dash Pro with silicone cap in size M; blue arrows show the LEDs for pulse rate detection. (b) Right Dash Pro, worn in the right ear.

For the comparison of the two in-ear pulse rate measurement devices, they had to be worn simultaneously. Therefore, the right Dash Pro has a self-sufficient and independent single-use mode.

2.2.2. Cosinuss^oOne

The second in-ear pulse rate measurement device evaluated in this study was the Cosinuss^oOne (Cosinuss, Munich, Germany). For comparison with the Dash Pro, the Cosinuss^oOne was used in the participant's left ear. Cosinuss^oOne measures the pulse rate in the external auditory canal by means of reflection measurement. Green light is used for this purpose. Measurement accuracy is ± 1 bpm, as specified by the manufacturer. To optimize the fit and minimize movement artefacts, the device is available in sizes S to L. Individual size could be determined by means of the Cosinuss^o app, using a value representing the received signal strength of the PO sensor. At a signal quality above 60%, size can be considered suitable. Figure 2a shows the Cosinuss^oOne. The two LEDs used for the reflection measurement in the external auditory canal are marked with blue arrows. Figure 2b shows the Cosinuss^oOne worn in the left ear of a participant.

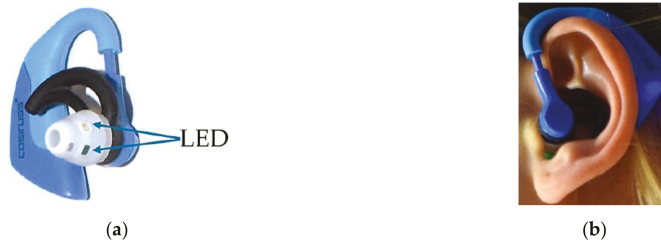


Figure 2. (a) Cosinuss°One; blue arrows show the LEDs for pulse rate detection. (b) Left Cosinuss°One, worn in the left ear.

2.2.3. Criterion Measure ECG—Bodyguard 2

Bodyguard 2 (Firstbeat Technologies Oy, Jyväskylä, Finland) was used as criterion measure. This ECG is not a medical device, but a sports-oriented electrocardiogram suitable for long-term and exercise ECG. Heart rate is recorded with two electrodes and processed with an integrated algorithm to correct artefacts. Compared to a clinical standard ECG, the Bodyguard 2 indicates 99.98% agreement [34].

2.3. Experimental Protocol

In order to evaluate the two in-ear pulse rate measurement devices as comprehensively as possible, their validity was evaluated during rest and under stress on a bicycle ergometer (Excalibur sport, Lode, Groningen, The Netherlands). Protocol started with a 10-min rest measurement in lying position, followed by a standardized, self-designed exercise protocol in a controlled laboratory setting. Prior to measurement under stress, each participant was instructed to cycle at a stress of 50 W for a 3-min warm-up period. Measurement under stress then started. The stress pattern of the protocol depended on the subject and was calculated from his/her weight in order to achieve an appropriate, comparable increase in intensity and thus heart rate for each subject. The number of watts per kilogram of body weight per minute increased uniformly by 0.4 W and 0.3 W in male and female subjects, respectively, starting from 50 W. The aim of this test protocol was to record a stress phase of at least 10 min and to measure a heart rate range of approximately 100–170 bpm. In order to ensure that participants did not stop prematurely because of maximum exhaustion, the duration of the protocol was set to 20 min. Formulas (1) and (2) show the calculation of the individual target stress of male and female participants.

$$\text{Target stress for males} = 50 \text{ W} + (\text{weight} \times 0.4 \times 20 \text{ min}), \quad (1)$$

$$\text{Target stress for females} = 50 \text{ W} + (\text{weight} \times 0.3 \times 20 \text{ min}). \quad (2)$$

The participants were instructed to cycle at a self-chosen number of rounds per minute (rpm) up to the individual maximum stress and to stop by hand signal in case of exhaustion. Upon completion, participants cooled down by cycling at a stress of 50 W for 3 min. During the entire data recording, participants were instructed to speak as little as possible, since jaw movements can lead to movement artefacts [21].

Regarding the use of all devices, care was taken to follow user guidelines as suggested by the manufacturers.

2.4. Data Analysis

ECG data were sampled at 1000 Hz and the heart rate was calculated from the time between R-R intervals, then exported as a text file at 1 s intervals. Data from optical pulse rate measurement of the in-ear devices was sampled at 100 Hz. The Cosinuss°One and the Dash Pro report the currently measured pulse rate to the respective mobile device app. For further analyses, data files of both in-ear

devices were downloaded at 5 s intervals. Afterwards, data files of the in-ear and ECG devices were synched using the respective timestamps of each data acquisition.

To ensure the synchronization of the in-ear and the ECG devices, their timestamps have to be reliable and identical. Therefore, all data files were recorded in the Unix timestamp format (UTC). This format counts time in milliseconds since 1 January 1970. In contrast, the Dash Pro counts time in milliseconds since 1 January 2015. This represents an overall time discrepancy of 45 years or a shift by up to 40 s within 24 h. This correction was carried out immediately before each examination.

Time-synched data from each device were concurrently and continuously acquired for each participant throughout the entire test protocol. In accordance to the validation study of Spierer et al. [31], a 5 s time interval was defined as sufficiently accurate for detecting significant variations in heart rate measurement. Hence, every fifth value of the heart rate was used for further analysis.

Figure 3 presents heart and pulse rate during the entire test protocol, including the change from lying position to cycling.

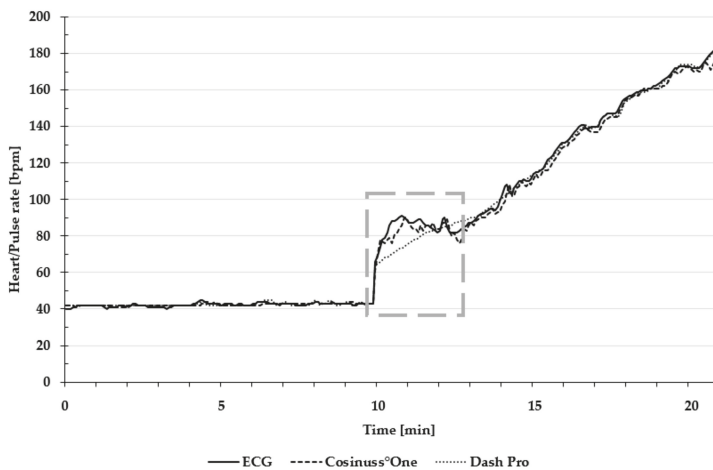


Figure 3. Exemplary presentation of pulse and heart rate during the entire test protocol. The dashed frame indicates the change from lying position to cycling. Heart rate of the ECG is depicted as a solid line; the pulse rate of Cosinuss°One is depicted as the dashed line; and the pulse rate of Dash Pro is depicted as the dotted line.

Motion artefacts, due to the change of body position and the re-adjustment of the sensors, resulted in strong signal noise, which can be seen in the dashed frame of Figure 3. Consequently, this data was not taken into account in the statistical evaluation.

Statistical analyses were conducted using IBM's SPSS Statistics software version 24 (IBM, Armonk, NY, USA). Descriptive statistics were used to characterize the sample population. The validity of the in-ear pulse rate measurement devices was determined by means of several statistical tests. Looking more closely at the term of validity, a distinction should be made between accuracy and precision. In the present study, accuracy was tested by MAPE and ICC values, whereas precision was identified by the limits of agreement of the Bland–Altman analysis.

Mean absolute percentage errors (MAPE) compared to the criterion measure were calculated as indicators of measurement error. MAPE, representing the error as a percentage of the overall mean relative to the ECG, does not have a standardized threshold for determining the accuracy of measurements. In the present study, a MAPE of $\leq 5\%$ [35] was used as the criterion value for accuracy. To further investigate the level of agreement, Bland–Altman plots [36] were prepared. These plots serve as a visual illustration of variance and over- or underestimated measurement ranges of the

investigated in-ear devices. For this, limits of agreement were set to 95%. Maximum and minimum pulse rates measured by means of the in-ear devices were compared with the ECG results using one-way repeated-measures ANOVA. An alpha of 0.05 was used to determine statistical significance. In addition, the agreement of maximum rate between the ECG and the tested devices was defined by means of the Intra-Class Correlation Coefficient (ICC) according to Liu et al. [37]. Excellent, good, moderate, and low agreement thresholds were defined as ICC values of ≥ 0.90 ; 0.75–0.90; 0.60–0.75; and ≤ 0.60 , suggested by Fokkema et al. [35].

3. Results

3.1. Preliminary Analysis

The Kolmogorov–Smirnov test as well as visual data plotting of the criterion measure (Bodyguard 2), Cosinuss°One, and the Dash Pro revealed that the overall rates among all participants were significantly different from a normal distribution (resting conditions: $D(2376) = 0.066$, $p < 0.001$ for ECG and $D(2376) = 0.080$, $p < 0.001$ for Cosinuss°One and $D(2376) = 0.081$, $p < 0.001$; stress conditions: $D(2547) = 0.079$, $p < 0.001$ for ECG and $D(2547) = 0.075$, $p < 0.001$ for Cosinuss°One and $D(2547) = 0.073$, $p < 0.001$ for Dash Pro). To account for the differences between criterion measure and alternative method, Bland and Altman [38] suggest to investigate the variances for normal distribution, too. The preliminary applied Kolmogorov–Smirnov test on differences between criterion measure and alternative method did significantly deviate from a normal distribution. However, visual inspection indicated mostly normally distributed data (resting conditions: $D(2376) = 0.172$, $p < 0.001$ for Cosinuss°One and $D(2376) = 0.192$, $p < 0.001$ for Dash Pro; stress conditions: $D(2547) = 0.280$, $p < 0.001$ for Cosinuss°One and $D(2547) = 0.216$, $p < 0.001$ for Dash Pro). Hence, the procedure suggested by Bland and Altman [38,39] was implemented.

3.2. Resting Heart/Pulse Rate and Heart/Pulse Rate ≤ 90 bpm

The differences between the resting rates and the rates ≤ 90 bpm of the investigated in-ear devices and the ECG are provided in Table 2.

Table 2. Comparison among different devices for resting heart/pulse rate and heart/pulse rate ≤ 90 bpm. Values are mean \pm standard deviation (SD), intra-class correlation coefficient (ICC), mean absolute error (MAE) \pm standard deviation (SD), and mean absolute percentage error (MAPE).

	Resting Heart/Pulse Rate		Heart/Pulse Rate ≤ 90 bpm	
	Mean \pm SD (bpm)	ICC	MAE \pm SD (bpm)	MAPE (%)
Cosinuss°One	53.6 \pm 8.3	0.94	1.5 \pm 1.8	2.5
Dash Pro	55.0 \pm 9.7	0.98	2.0 \pm 2.5	3.2
ECG	54.9 \pm 10.1			

On average, the participants achieved a resting heart rate of 54.9 ± 10.1 bpm in the ECG examinations. The average in-ear measured resting pulse rate of the Cosinuss°One and the Dash Pro is 53.6 ± 8.3 bpm and 55.0 ± 9.7 bpm, respectively.

Figure 4 is an exemplary presentation of the in-ear devices' pulse rate and the ECG's heart rate, recorded during the 10 min rest measurement in a lying position.

For the main analysis, one-way repeated-measures ANOVA was conducted. Mauchly's tests indicated that the assumption of sphericity had not been violated, $X^2(2) = 4.27$, $p > 0.05$, therefore non-corrected tests are reported. The results revealed that the minimum rates were not significantly affected by the measurement device, $F(2, 38) = 3.17$, $p > 0.05$.

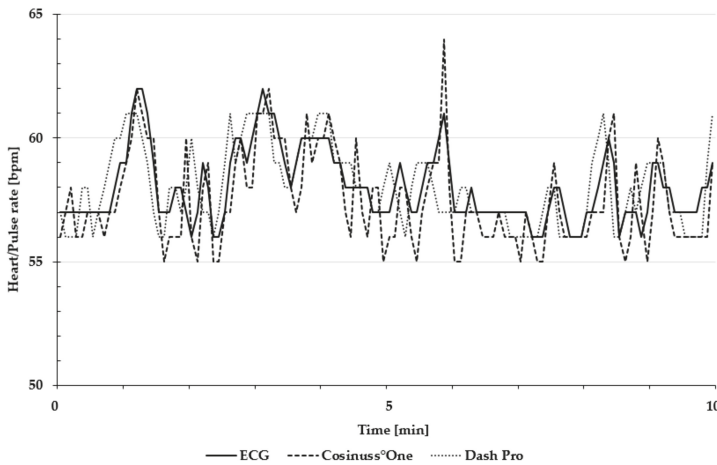


Figure 4. Exemplary presentation of pulse and heart rate during the 10 min rest measurement in a lying position. The heart rate of ECG is depicted as the solid line; the pulse rate of Cosinuss°One is depicted as the dashed line; and the pulse rate of Dash Pro is depicted as the dotted line.

Both in-ear devices are quite similar in terms of mean absolute error and MAPE, whereas the Dash Pro indicates higher values than the Cosinuss°One. In addition, the Cosinuss°One and the Dash Pro show excellent agreement in comparison to ECG, $R = 0.94$ for Cosinuss°One and $R = 0.98$ for Dash Pro.

Figure 5 shows Bland–Altman plots of the Cosinuss°One and the Dash Pro in-ear devices in comparison to the ECG with Bodyguard 2.

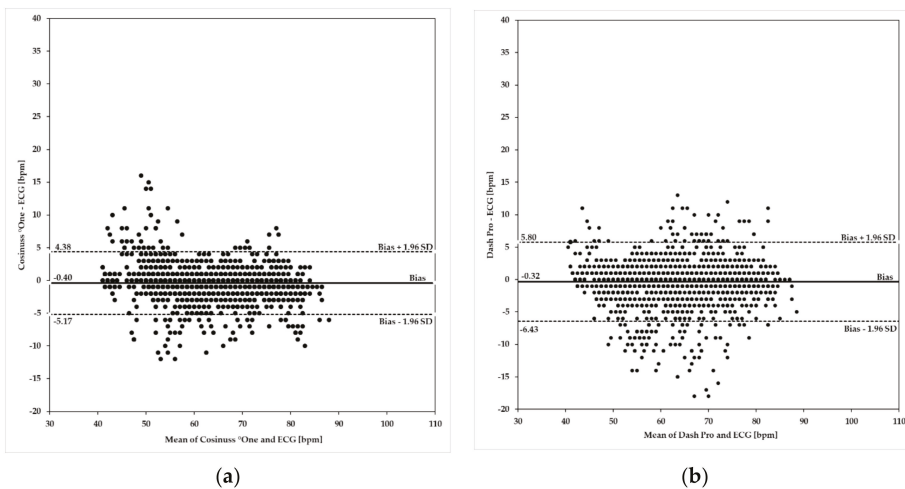


Figure 5. Bland–Altman plots using heart/pulse rates ≤ 90 bpm. Investigated in-ear devices: (a) Cosinuss°One; (b) Dash Pro. Plots indicate differences of the rate values on the y-axis relative to the mean of the two methods (ECG and in-ear measurement) on the x-axis. Limits of agreement (LoA) were calculated as mean $\pm 1.96 \times$ SD. Biases are depicted as a solid line; LoA are depicted as dashed lines.

Upper and lower limits of agreement (ULoA, LLoA) as well as the mean differences of the in-ear devices' PR ≤ 90 bpm compared to the ECG's HR are labeled in Figure 5a,b. Both in-ear devices show

considerable deviations in scattering when compared to the ECG and tend to underestimate the heart rate by 0.40 bpm for Cosinuss°One and by 0.32 bpm for the Dash Pro. Bland–Altman analysis of the Cosinuss°One (Figure 5a) and the Dash Pro (Figure 5b) show that variability occurred across the spectrum of rates ≤ 90 bpm. The Cosinuss°One indicates more variability between 40 bpm and 60 bpm, while the Dash Pro shows higher variability across the midrange rates. In addition, the differences in variance are visualized. The Cosinuss°One (ULoA–LLoA: 9.55 bpm) shows lower scattering among its measurements when compared to the Dash Pro (ULoA–LLoA: 12.23 bpm). The Cosinuss°One had 95% of differences within +4.38 bpm and -5.17 bpm of the ECG, while the Dash Pro had 95% of differences within +5.80 bpm and -6.43 bpm.

3.3. Heart/Pulse Rate ≥ 100 bpm

The mean absolute error and the MAPE of the investigated in-ear devices' pulse rate ≥ 100 bpm are provided in Table 3.

Table 3. Comparison among in-ear devices for pulse rates ≥ 100 bpm. Values are mean absolute error (MAE) \pm standard deviation (SD), mean absolute percentage error (MAPE).

	Pulse Rate ≥ 100 bpm	
	MAE \pm SD (bpm)	MAPE (%)
Cosinuss°One	1.8 \pm 2.8	1.3
Dash Pro	1.8 \pm 2.8	1.4

Both in-ear devices are quite similar in terms of mean absolute error and MAPE. Figure 6a is an exemplary presentation of the in-ear devices' pulse rate and the ECG's heart rate, recorded during the stress protocol to the individual exhaustion. Figure 6b illustrates high differences between ECG and the in-ear devices around 100 bpm.

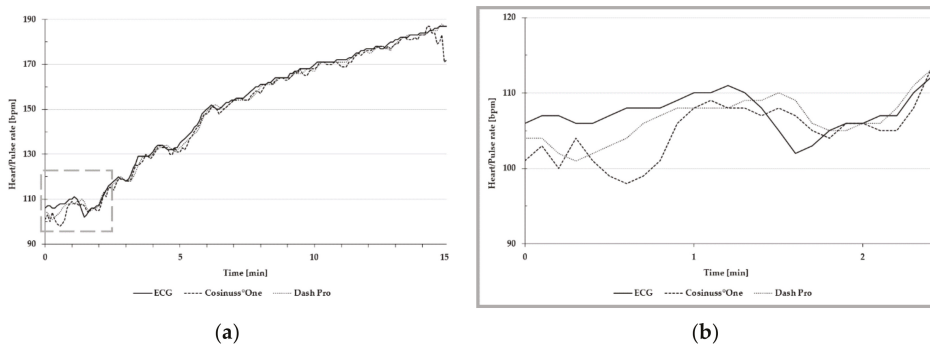


Figure 6. Exemplary presentation of pulse and heart rate (a) during the entire stress protocol; (b) around 100 bpm. The heart rate of ECG is depicted as the solid line; the pulse rate of Cosinuss°One is depicted as the dashed line; and the pulse rate of Dash Pro is depicted as the dotted line.

Figure 7 shows Bland–Altman plots of the Cosinuss°One and the Dash Pro in-ear devices in comparison to the ECG with Bodyguard 2.

Upper and lower limits of agreement (ULoA, LLoA) as well as the mean differences of the in-ear devices' PR ≥ 100 bpm compared to the ECG's HR are labeled in Figure 7a,b. Both in-ear devices show considerable deviations in scattering when compared to the ECG and tend to underestimate heart rate by 1.60 bpm for Cosinuss°One and by 0.51 bpm for the Dash Pro. Bland–Altman analysis of the Cosinuss°One (Figure 7a) and the Dash Pro (Figure 7b) shows that variability occurred across the spectrum of rates ≥ 100 bpm. The Cosinuss°One indicates more variability between 100 bpm and

150 bpm, while the Dash Pro shows higher variability between 100 bpm and 115 bpm. In addition, the differences in variance are visualized. The Cosinuss°One (ULoA-LLoA: 11.48 bpm) shows lower scattering among its measurements when compared to the Dash Pro (ULoA-LLoA: 12.67 bpm). The Cosinuss°One had 95% of differences within +4.14 bpm and -7.34 bpm of the ECG, while the Dash Pro had 95% of differences within +5.82 bpm and -6.85 bpm.

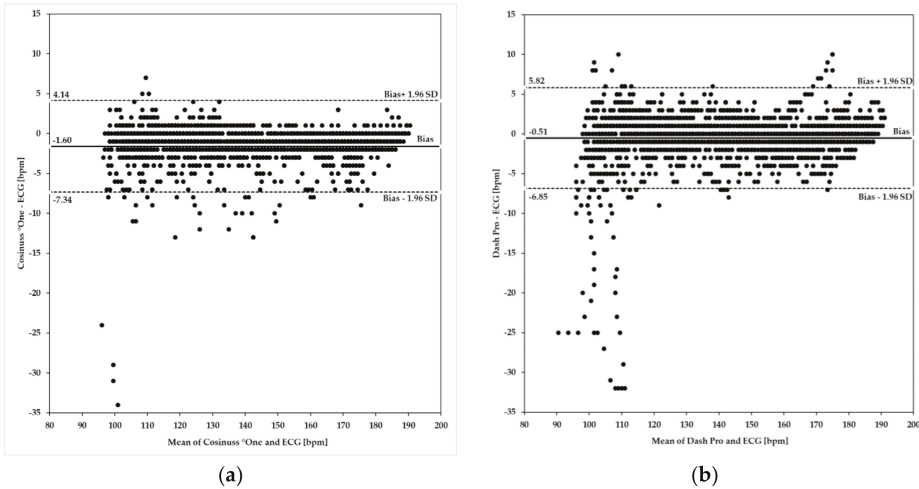


Figure 7. Bland–Altman plots using heart/pulse rate ≥ 100 bpm. Investigated in-ear devices: (a) Cosinuss°One; (b) Dash Pro. Plots indicate differences of the rate values on the y-axis relative to the mean of the two methods (ECG and in-ear measurement) on the x-axis. Limits of agreement (LoA) were calculated as mean $\pm 1.96 \times$ SD. Biases are depicted as a solid line; LoA are depicted as dashed lines.

3.4. Maximum Heart/Pulse Rate

The differences between the maximum heart/pulse rate of the investigated in-ear devices and the ECG are provided in Table 4.

Table 4. Comparison among different devices for maximum heart/pulse rate. Values are mean \pm standard deviation (SD), intra-class correlation coefficient (ICC).

	Maximum Heart/Pulse Rate	
	Mean \pm SD (bpm)	ICC
Cosinuss°One	181.6 \pm 6.4	0.84
Dash Pro	183.7 \pm 4.8	0.83
ECG	183.0 \pm 5.1	

On average, participants achieved a HR_{max} of 183.0 ± 5.1 bpm in the ECG examinations. The average in-ear measured PR_{max} of the Cosinuss°One and the Dash Pro is 181.6 ± 6.4 bpm and 183.7 ± 4.8 bpm, respectively.

For the main analysis, one-way repeated-measures ANOVA was conducted. Mauchly’s test indicated that the assumption of sphericity had not been violated, $X^2(2) = 3.66, p > 0.05$, therefore non-corrected tests are reported. In addition, the results revealed that the maximum rates were significantly affected by the measurement device, $F(2, 38) = 3.80, p \leq 0.05$. In addition, the Cosinuss°One and the Dash Pro show good agreement in comparison to ECG, $R = 0.84$ for Cosinuss°One and $R = 0.83$ for Dash Pro.

3.5. Motion Artefacts

To illustrate the influence of jaw movements on the PPG signal, one participant was asked to chew chewing gum as well as to talk throughout the data recording. The effects on the PPG signal and the pulse rate can be clearly seen in Figure 8a,b. Figure 8a illustrates the pulse rate of the Cosinuss°One, Figure 8b shows the spectrogram of the PPG signal of the Dash Pro.

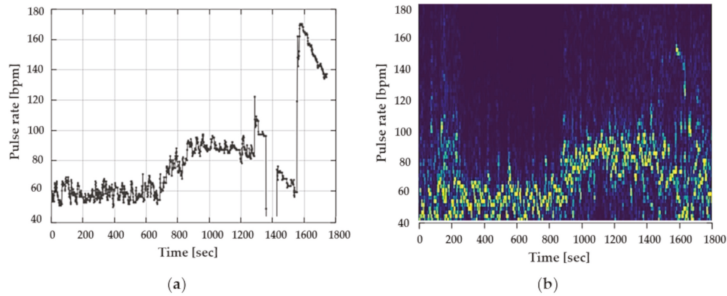


Figure 8. Influence of jaw movements and talking on the PPG signal of both in-ear devices: (a) Cosinuss°One; (b) spectrogram of PPG signal of the Dash Pro.

Both figures depict the significant influence of motion on the PPG signal and the effect on the pulse rate. In both devices, signal interference was too intense to determine a precise pattern of the pulse rate.

4. Discussion

The present study examined the accuracy and precision of two ear-worn pulse rate measurement devices using PPG technology.

Systematic differences should be assessed using the MAPE. According to Nelson et al. [40], wearable devices should not exceed a MAPE threshold of $\leq 10\%$ in order to be considered accurate. Fokkema et al. [35] suggest a threshold of $\leq 5\%$. In the present study, both devices indicated smaller MAPE scores at rates ≥ 100 bpm than at rates ≤ 90 bpm, but not exceeding the threshold of 5%. Thus, both in-ear devices can be classified as accurate almost completely within heart rates of 60–190 bpm, even according to Fokkema et al. [35].

To investigate the level of agreement between the in-ear devices and the criterion measure ECG, Bland–Altman plots were prepared according to Bland and Altman [36]. In resting condition as well as within the range of pulse rates ≥ 100 bpm, the Cosinuss°One revealed a narrower 95% limit of agreement than the Dash Pro. Both devices tended to slightly underestimate heart rate values within 60–190 bpm and showed high deviations from the ECG around 100 bpm. Although the limits of agreement might seem pretty narrow, these results have to be considered carefully. Heart and pulse rate data are a very sensitive source, e.g., for app-based training programs, and therefore imprecise and inaccurate raw data can lead to fatal results.

To determine the level of agreement, the ICC values of resting heart and pulse rates as well as maximum heart and pulse rates were examined. Under resting conditions, the Dash Pro and the Cosinuss°One show excellent agreement in comparison to ECG, whereas at maximum heart rates, a decrease to good agreement of both devices could be observed. Despite not exceeding the MAPE threshold of $\leq 5\%$ and good to excellent levels of agreement, the results revealed that the maximum rates were significantly affected by the measurement device.

From a scientific point of view, there is only one commercial in-ear device, Bose Sound Sport (BSS; Bose Corporation, Framingham, MA, USA), which has been validated with respect to PPG pulse rate measurement to date. This underlines the lack of scientific validation studies. Pulse rate is a

physiological parameter used for training control in sports or for monitoring the general state of health. Consumers must be protected, especially if activity trackers are to be used increasingly in the health and fitness sector. However, this access can only be defended or considered responsible if the current lack of transparency of the activity trackers industry is remedied through high-quality research, which can also help define general standards for these devices. The present study contributes to fill this gap. It shows the potential and the weaknesses of two commercially available in-ear devices in terms of measuring pulse rate.

Boudreaux et al. [41] conducted a study very similar to the present research. It is the first and, so far, the only study that investigated the accuracy and precision for pulse rate sensors in commercial headphones. The difference of this to the present study is the measurement site of PPG pulse rate detection. The measurement site of the Bose Sound Sport is the auricle. Both in-ear PPG devices validated in the present study use the external auditory canal for pulse rate measurement. Considering the measurement site and the consumer market, it should be noted that the present study is the first one to validate commercially available pulse rate monitors using PPG in the external auditory canal.

In the study of Boudreaux et al. [41], eight wearable devices, including the ear-worn Bose Sound Sport PPG device, were compared to a six-channel ECG regarding heart rate and caloric expenditure measurements. In addition to other stress situations, the BSS was also validated during cycling. For this purpose, a graded stress protocol was used, starting at rest and ending at a maximum stress of 200 W. The BSS slightly overestimates heart rate during low cycling intensities and underestimates heart rate during higher intensities. The latter agrees with the results of the present study. On average, the MAPE of the BSS was 7.4%. Under resting conditions, the MAPE of the BSS was 3.2%, very close to the Dash Pro and the Cosinuss°One, with MAPE values of 3.2% and 2.5%, respectively.

Considering the level of agreement, Boudreaux et al. [41] observed decreasing ICC values with higher exercise intensity and higher heart rates, respectively. This is in line with the results of the present study. The Cosinuss°One and the Dash Pro indicate a decrease from excellent to good agreement.

On the basis of MAPE values $\leq 5\%$ and excellent to good levels of agreement, it can be stated that both the Dash Pro and the Cosinuss°One deliver accurate pulse rate values. In addition to accuracy, precision must also be considered as key factor when testing the validity of wearable devices. Due to high variance measured over the entire spectrum of heart/pulse rates, both in-ear devices of the present study have to be considered too imprecise as to be used as an alternative to the ECG.

Other than the mentioned study of Boudreaux et al. [41] and the present research, where commercially available ear-worn PPG devices were tested, the studies of Tigges et al. [23], Budidha and Kyriacou [16], and Leboeuf et al. [42] investigated self-developed pulse rate measurement systems based on PPG-technology. These devices are not commercially available. Tigges et al. [23] built such a device for scientific purposes and tested the device in the spectrum of 50–125 bpm under resting conditions. The Bland–Altman analysis showed a bias of -0.03 bpm, with 95% of the data lying within the boundaries of $+2.88$ bpm and -2.94 bpm. The study of Leboeuf et al. [42] showed a bias of -0.2 bpm, with the sensor slightly underestimating heart rate data. Similar to these results, where the data showed excellent agreement up to 200 bpm within varying activities, Budidha and Kyriacou [16] also concluded that the ear canal might be a suitable site for pulse rate measurement conducted by PPG sensors.

In comparison to the in-ear pulse rate PPG devices, considerably more wrist-worn pulse rate devices were validated [7,43–53]. Both fitness trackers and sports watches were tested under different exercise conditions. In order to ensure comparability to the presented results of the in-ear devices, only the literature in which validation was conducted on both a bicycle ergometer and an ECG was used as criterion measure will be discussed further. Within the framework of these studies [43–46], the validity of the wrist-worn devices, e.g., Apple Watch, Fitbit Charge HR, Basis Peak, Samsung Gear S, Polar M600, was mostly examined at rest, low, and high intensities. All devices tended to underestimate the heart rate regardless of intensity under both resting and cycling conditions. For instance, Wallen et al. [43] and Horton et al. [46] indicated that increasing physical effort leads to

decreasing accuracy of the heart rate measurement. The results of Wallen et al. [43] revealed that heart rate underestimation ranged from -0.52 bpm (Basis Peak) to -12.67 bpm (Fitbit Charge HR) at low intensities, and -7.42 bpm (Basis Peak) to -14.20 bpm (Fitbit Charge HR) at high intensities, depending on the wrist-worn device used. Horton et al. [46] compared the sports watch Polar M600 pulse rate with the ECG heart rate during rest and cycling. They demonstrated a heart rate underestimation of -0.1 bpm during rest and -1.9 bpm during cycling. These findings are in agreement with the present results that show heart rate underestimation ranged from -0.32 bpm (Dash Pro) to -0.40 bpm (Cosinuss°One) under resting conditions and from -0.51 bpm (Dash Pro) to -1.60 bpm (Cosinuss°One) under cycling conditions.

The results presented in this study suggest the weakness of the PPG pulse rate monitoring devices, both in-ear and the wrist-worn.

PPG is a low-cost medical technique applied to the skin that uses the transmission and reflection of light into the skin to measure changes in blood volume within a specific tissue [9]. Previous research [1,4,54] suggests that PPG devices may have limitations in measuring PR that arise from the continuous increase and decrease of compression of a wearable device's PR sensor on the skin. In addition to conditions of low blood circulation and movements of the sensor on the skin, ambient light, which falsifies the values of the receiving diode, can also be seen as a source of error in PPG heart rate measurement [55]. In-ear PPG pulse rate monitors were developed to solve these main problems. As Vogel et al. [21] were able to show, a PPG sensor in the auricle near the ear canal, which irradiates light into the skin by reflection measurement, is resistant to most of the above-mentioned interfering factors. Circulatory disturbances and ambient light were minimized by the placement inside the ear. The inner ear is supplied by the same artery as the brain, which is why the blood flow and thus the strength of the PPG signal received in the inner ear remains constant, in contrast to PPG sensors on the extremities [16].

Due to pulse transit time, differences between heart rate and pulse rate may occur in general. Time lag between the peak of R-wave on the ECG and the peak value of the corresponding pulse in the ear, measured by PPG, may be a reason for underestimating heart rate.

A main remaining weakness of PPG is an unreliable signal due to motion artefacts. For in-ear PPG pulse rate measurement, jaw movements are the most dominant artefact [21]. In the present study, the external auditory canal as an application site for commercially available PPG pulse rate measurements was therefore tested as an alternative to electrocardiographic heart rate measurement.

While it was shown that the external auditory canal is a limited alternative measurement site for pulse rate measurement, the specific conditions in this study have to be taken into consideration. The study was conducted under controlled testing conditions, as the subjects were instructed to reduce jaw movement, e.g., by talking as little as possible to deliberately minimize motion artefacts. In addition, motion artefacts, such as head movement and vibrations due to uneven terrain, were not taken into account in this study. Despite the instruction to cycle at a self-chosen number of rounds per minute (rpm), most of the participants cycled at 80–100 rpm. This frequency may have increased the deviation of both of the in-ear devices from the ECG at a heart rate around 100 bpm.

In his review article, Toshiyo Tamura [4] suggests various methods to eliminate motion artefacts.

A conclusion regarding the higher validity of PPG due to different wavelengths when comparing both devices in the present study could not be made. Even though the Cosinuss°One measures with green light while the Dash Pro uses infrared light, significant differences in the measurement accuracy and precision could not be shown, opposing the suggestion that the use of green LEDs allows a more accurate and precise detection of the pulse rate [9,29–31].

The subject group represents just a certain and small population. Due to the young age and the healthy condition of the subjects, conclusions that consider older or diseased people cannot be drawn. In addition, a gender-specific differentiation of the results cannot be made due to the unbalanced distribution of fourteen men and six women.

5. Conclusions

It can be concluded that PPG measurement in the ear is a promising technique in resting positions. For future improvement, however, motion artefacts should be significantly reduced in order to ensure accuracy and precision in in-ear PPG pulse rate measurement during activity. Therefore, extensive studies under real life conditions should be carried out for a better understanding of the interaction between measurement accuracy as well as precision and the effects of artefacts from motion. This requires a considerable contribution from the activity trackers industry. The existing lack of transparency, e.g., regarding the algorithms used, complicates the analysis of the problems and thus the development of solutions.

The results of this study give an insight into the validity of the PPG technology used in the ear. Particularly in consumer wearable devices, a device worn in the ear can have many advantages over chest- or wrist-worn devices, as sports-specific parameters could be broadcasted to the athlete without interfering with the current activity. The combination of hearing aids and their ability to measure pulse rate simultaneously, could be a future application of in-ear PPG technology to promote the safety of elders. For use in health care, general standards and guidelines, with respect to measurement accuracy and precision, have to be defined and introduced to the activity trackers industry.

Author Contributions: Conceptualization, S.P. and N.M.; methodology, S.P. and N.M.; formal analysis, S.P. and N.M.; investigation, S.P. and N.M.; writing—original draft preparation, S.P. and N.M.; writing—review and editing, S.P., N.M. and V.S.; visualization, S.P.; supervision, S.P. and V.S.

Funding: This research received no external funding.

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

References

1. Achten, J.; Jeukendrup, A.E. Heart rate monitoring: Applications and limitations. *Sports Med.* **2003**, *33*, 517–538. [[CrossRef](#)]
2. Hertzman, A.B. The blood supply of various skin areas as estimated by the photoelectric plethysmograph. *Am. J. Physiol. Content* **1938**, *124*, 328–340. [[CrossRef](#)]
3. Sanders, J.; Wark, P.; Winfree, K.; Fallahzadeh, R.; Fernández, C.; Henriksen, A.; Mikalsen, M.H.; Woldaregay, A.Z.; Muzny, M.; Hartvigsen, G.; et al. Using Fitness Trackers and Smartwatches to Measure Physical Activity in Research: Analysis of Consumer Wrist-Worn Wearables. *J. Med. Internet Res.* **2018**, *20*, e110.
4. Tamura, T. Current progress of photoplethysmography and SPO₂ for health monitoring. *Biomed. Eng. Lett.* **2019**, *9*, 21–36. [[CrossRef](#)] [[PubMed](#)]
5. Jung, J.Y.; Lee, L.W. Zigbee device access control and reliable data transmission in Zigbee based health monitoring. In Proceedings of the 2008 10th International Conference on Advanced Communication Technology, Gangwon-Do, Korea, 17–20 February 2008; pp. 795–797.
6. Lee, Y.K.; Jo, J.; Shin, H.S. Development and Evaluation of a Wristwatch-Type Photoplethysmography Array Sensor Module. *IEEE Sens. J.* **2013**, *13*, 1459–1463. [[CrossRef](#)]
7. Cadmus-Bertram, L.; Gangnon, R.; Wirkus, E.J.; Thraen-Borowski, K.M.; Gorzelitz-Liebhauser, J. The Accuracy of Heart Rate Monitoring by Some Wrist-Worn Activity Trackers. *Ann. Intern. Med.* **2017**, *166*, 610–612. [[CrossRef](#)] [[PubMed](#)]
8. Maeda, Y.; Sekine, M.; Tamura, T. Relationship between measurement site and motion artifacts in wearable reflected photoplethysmography. *J. Med. Syst.* **2011**, *35*, 969–976. [[CrossRef](#)]
9. Maeda, Y.; Sekine, M.; Tamura, T. The Advantages of Wearable Green Reflected Photoplethysmography. *J. Med. Syst.* **2011**, *35*, 829–834. [[CrossRef](#)]
10. Maguire, M.; Ward, T. *The Design and Clinical Use of a Reflective Brachial Photoplethysmograph*; Maynooth University: Kildare, Ireland, 2002.
11. Rhee, S.; Yang, B.-H.; Asada, H. Artifact-resistant power-efficient design of finger-ring plethysmographic sensors. *IEEE Trans. Biomed. Eng.* **2001**, *48*, 795–805. [[CrossRef](#)] [[PubMed](#)]

12. Kyriacou, P.A.; Moye, A.R.; Gregg, A.; Choi, D.M.A.; Langford, R.M.; Jones, D.P.; Kyriacou, P. A system for investigating oesophageal photoplethysmographic signals in anaesthetised patients. *Med. Biol. Eng.* **1999**, *37*, 639–643. [[CrossRef](#)]
13. Mendelson, Y.; Pujary, C. Measurement site and photodetector size considerations in optimizing power consumption of a wearable reflectance pulse oximeter. In Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Cancún, Mexico, 17–21 September 2003; pp. 3016–3019.
14. Dassel, A.C.M.; Graaff, R.; Sikkema, M.; Meijer, A.; Zijlstra, W.G.; Aarnoudse, J.G. Reflectance pulse oximetry at the forehead improves by pressure on the probe. *J. Clin. Monit.* **1995**, *11*, 237–244. [[CrossRef](#)] [[PubMed](#)]
15. Poh, M.-Z.; Swenson, N.C.; Picard, R.W. Motion-tolerant magnetic earring sensor and wireless earpiece for wearable photoplethysmography. *IEEE Trans. Inf. Technol. Biomed.* **2010**, *14*, 786–794. [[CrossRef](#)] [[PubMed](#)]
16. Budidha, K.; Kyriacou, P.A. The human ear canal: Investigation of its suitability for monitoring photoplethysmographs and arterial oxygen saturation. *Physiol. Meas.* **2014**, *35*, 111–128. [[CrossRef](#)] [[PubMed](#)]
17. Poh, M.-Z.; Kim, K.; Goessling, A.; Swenson, N.; Picard, R. Cardiovascular Monitoring Using Earphones and a Mobile Device. *IEEE Pervasive Comput.* **2012**, *11*, 18–26. [[CrossRef](#)]
18. Patterson, J.A.C.; McIlwraith, D.C.; Yang, G.-Z. A Flexible, Low Noise Reflective PPG Sensor Platform for Ear-Worn Heart Rate Monitoring. In Proceedings of the 2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks, Berkeley, CA, USA, 3–5 June 2009; pp. 286–291.
19. Wang, C.-Z.; Zheng, Y.-P. Home-Telecare of the Elderly Living Alone Using an New Designed Ear-wearable Sensor. In Proceedings of the 5th International Workshop on Wearable and Implantable Body Sensor Networks, Hong Kong, China, 1–3 June 2008; pp. 71–74.
20. Wang, L.; Lo, B.P.L.; Yang, G.-Z. Multichannel Reflective PPG Earpiece Sensor with Passive Motion Cancellation. *IEEE Trans. Biomed. Circuits Syst.* **2007**, *1*, 235–241. [[CrossRef](#)] [[PubMed](#)]
21. Vogel, S.; Hülsbusch, M.; Hennig, T.; Blazek, V.; Leonhardt, S. In-ear vital signs monitoring using a novel microoptic reflective sensor. *IEEE Trans. Inf. Technol. Biomed.* **2009**, *13*, 882–889. [[CrossRef](#)] [[PubMed](#)]
22. Goverdovsky, V.; Von Rosenberg, W.C.; Nakamura, T.; Looney, D.; Sharp, D.J.; Papavassiliou, C.; Morrell, M.J.; Mandic, D.P. Hearables: Multimodal physiological in-ear sensing. *Sci. Rep.* **2017**, *7*, 6948. [[CrossRef](#)] [[PubMed](#)]
23. Tigges, T.; Büchler, T.; Pielmuş, A.; Klum, M.; Feldheiser, A.; Hunsicker, O.; Orglmeister, R. Assessment of In-ear Photoplethysmography as a Surrogate for Electrocardiography in Heart Rate Variability Analysis. In *World Congress on Medical Physics and Biomedical Engineering 2018*; Lhotska, L., Sukupova, L., Lacković, I., Ibbott, G.S., Eds.; Springer: Singapore, 2019; pp. 293–297.
24. Shin, K.; Kim, Y.; Bae, S.; Park, K.; Kim, S. A Novel Headset with a Transmissive PPG Sensor for Heart Rate Measurement. In Proceedings of the 13th International Conference on Biomedical Engineering, Singapore, 3–6 December 2008; Lim, C.T., Goh, J.C.H., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; pp. 519–522.
25. Agashe, G.S.; Coakley, J.; Mannheimer, P.D. Forehead pulse oximetry: Headband use helps alleviate false low readings likely related to venous pulsation artifact. *Anesthesiology* **2006**, *105*, 1111–1116. [[CrossRef](#)]
26. Mendelson, Y.; Ochs, B.D. Noninvasive pulse oximetry utilizing skin reflectance photoplethysmography. *IEEE Trans. Biomed. Eng.* **1988**, *35*, 798–805. [[CrossRef](#)]
27. Anderson, R.R.; Parrish, J.A. The Optics of Human Skin. *J. Investig. Dermatol.* **1981**, *77*, 13–19. [[CrossRef](#)] [[PubMed](#)]
28. Allen, J. Photoplethysmography and its application in clinical physiological measurement. *Physiol. Meas.* **2007**, *28*, R1–R39. [[CrossRef](#)]
29. Delgado-Gonzalo, R.; Parák, J.; Tarniceriu, A.; Renevey, P.; Bertschi, M.; Korhonen, I. Evaluation of accuracy and reliability of PulseOn optical heart rate monitoring device. In Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015; pp. 430–433.
30. Lee, S.; Shin, H.; Hahm, C. Effective PPG sensor placement for reflected red and green light, and infrared wristband-type photoplethysmography. In Proceedings of the 2016 18th International Conference on Advanced Communication Technology (ICACT), Pyeongchang, Korea, 31 January–3 February 2016.
31. Spierer, D.K.; Rosen, Z.; Litman, L.L.; Fujii, K. Validation of photoplethysmography as a method to detect heart rate during rest and exercise. *J. Med. Eng. Technol.* **2015**, *39*, 264–271. [[CrossRef](#)] [[PubMed](#)]

32. European Union. Regulation (EU) 2016/679 of the European Parliament and the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Off. J. Eur. Union* **2016**, *59*, 1–88.
33. European Union. Corrigendum to Regulation (EU) 2016/679 of the European Parliament and the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Off. J. Eur. Union* **2018**, *61*, 2–5.
34. Parak, J. Accuracy of Firstbeat Bodyguard 2 Beat-To-Beat Heart Rate Monitor; (whitepaper). 2013. Available online: https://assets.firstbeat.com/firstbeat/uploads/2015/11/white_paper_bodyguard2_final.pdf (accessed on 20 August 2019).
35. Fokkema, T.; Kooiman, T.J.M.; Krijnen, W.P.; Van der Schans, C.P.; Groot, M. Reliability and Validity of Ten Consumer Activity Trackers Depend on Walking Speed. *Med. Sci. Sports Exerc.* **2017**, *49*, 793–800. [[CrossRef](#)] [[PubMed](#)]
36. Bland, J.M.; Altman, D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1986**, *327*, 307–310. [[CrossRef](#)]
37. Liu, J.; Tang, W.; Chen, G.; Lu, Y.; Feng, C.; Tu, X.M. Correlation and agreement: Overview and clarification of competing concepts and measures. *Shanghai Arch Psychiatry* **2016**, *28*, 115–120.
38. Bland, J.M.; Altman, D.G. Applying the right statistics: Analyses of measurement studies. *Ultrasound Obstet. Gynecol.* **2003**, *22*, 85–93. [[CrossRef](#)]
39. Bland, J.M.; Altman, D.G. Measuring agreement in method comparison studies. *Stat. Methods Med. Res.* **1999**, *8*, 135–160. [[CrossRef](#)]
40. Nelson, M.B.; Kaminsky, L.A.; Dickin, D.C.; Montoye, A.H. Validity of Consumer-Based Physical Activity Monitors for Specific Activity Types. *Med. Sci. Sports Exerc.* **2016**, *48*, 1619–1628. [[CrossRef](#)]
41. Boudreaux, B.D.; Hebert, E.P.; Hollander, D.B.; Williams, B.M.; Cormier, C.L.; Naquin, M.R.; Gillan, W.W.; Gusew, E.E.; Kraemer, R.R. Validity of Wearable Activity Monitors during Cycling and Resistance Exercise. *Med. Sci. Sports Exerc.* **2018**, *50*, 624–633. [[CrossRef](#)] [[PubMed](#)]
42. Leboeuf, S.F.; Aumer, M.E.; Kraus, W.E.; Johnson, J.L.; Duscha, B. Earbud-based sensor for the assessment of energy expenditure, HR, and VO₂max. *Med. Sci. Sports Exerc.* **2014**, *46*, 1046–1052. [[CrossRef](#)] [[PubMed](#)]
43. Wallen, M.P.; Gomersall, S.R.; Keating, S.E.; Wisloff, U.; Coombes, J.S. Accuracy of Heart Rate Watches: Implications for Weight Management. *PLoS ONE* **2016**, *11*, e0154420. [[CrossRef](#)] [[PubMed](#)]
44. Jo, E.; Lewis, K.; Directo, D.; Kim, M.J.; Dolezal, B.A. Validation of Biofeedback Wearables for Photoplethysmographic Heart Rate Tracking. *J. Sports Sci. Med.* **2016**, *15*, 540–547. [[PubMed](#)]
45. Benedetto, S.; Caldato, C.; Bazzan, E.; Greenwood, D.C.; Pensabene, V.; Actis, P. Assessment of the Fitbit Charge 2 for monitoring heart rate. *PLoS ONE* **2018**, *13*, e0192691. [[CrossRef](#)] [[PubMed](#)]
46. Horton, J.F.; Stergiou, P.; Katz, L.; Fung, T.S. Comparison of Polar M600 Optical Heart Rate and ECG Heart Rate during Exercise. *Med. Sci. Sports Exerc.* **2017**, *49*, 2600–2607. [[CrossRef](#)] [[PubMed](#)]
47. Abt, G.; Bray, J.; Benson, A.C. The validity and inter-device variability of the Apple Watch™ for measuring maximal heart rate. *J. Sports Sci.* **2018**, *36*, 1447–1452. [[CrossRef](#)]
48. Støve, M.P.; Haucke, E.; Nymann, M.L.; Sigurdsson, T.; Larsen, B.T. Accuracy of the wearable activity tracker Garmin Forerunner 235 for the assessment of heart rate during rest and activity. *J. Sports Sci.* **2019**, *37*, 895–901. [[CrossRef](#)]
49. Wang, R.; Blackburn, G.; Desai, M.; Phelan, D.; Gillinov, L.; Houghtaling, P. Accuracy of Wrist-Worn Heart Rate Monitors. *JAMA Cardiol.* **2017**, *2*, 104–106. [[CrossRef](#)]
50. Stahl, S.E.; An, H.-S.; Dinkel, D.M.; Noble, J.M.; Lee, J.-M. How accurate are the wrist-based heart rate monitors during walking and running activities? Are they accurate enough? *BMJ Open Sport Exerc. Med.* **2016**, *2*, e000106. [[CrossRef](#)]
51. Reddy, R.K.; Pooni, R.; Zaharieva, D.P.; Senf, B.; El Youssef, J.; Dassau, E.; Doyle Iii, F.J.; Clements, M.A.; Rickels, M.R.; Patton, S.R.; et al. Accuracy of Wrist-Worn Activity Monitors During Common Daily Physical Activities and Types of Structured Exercise: Evaluation Study. *JMIR mHealth uHealth* **2018**, *6*, e10338. [[CrossRef](#)] [[PubMed](#)]

52. Gillinov, S.; Etiwy, M.; Wang, R.; Blackburn, G.; Phelan, D.; Gillinov, A.M.; Houghtaling, P.; Javadikasgari, H.; Desai, M.Y. Variable Accuracy of Wearable Heart Rate Monitors during Aerobic Exercise. *Med. Sci. Sports Exerc.* **2017**, *49*, 1697–1703. [[CrossRef](#)] [[PubMed](#)]
53. Shcherbina, A.; Mattsson, C.M.; Waggott, D.; Salisbury, H.; Christle, J.W.; Hastie, T.; Wheeler, M.T.; Ashley, E.A. Accuracy in Wrist-Worn, Sensor-Based Measurements of Heart Rate and Energy Expenditure in a Diverse Cohort. *J. Pers. Med.* **2017**, *7*, 3. [[CrossRef](#)] [[PubMed](#)]
54. Tamura, T.; Maeda, Y.; Sekine, M.; Yoshida, M. Wearable Photoplethysmographic Sensors—Past and Present. *Electronics* **2014**, *3*, 282–302. [[CrossRef](#)]
55. Trivedi, N.S.; Ghouri, A.F.; Shah, N.K.; Lai, E.; Barker, S.J. Effects of motion, ambient light, and hypoperfusion on pulse oximeter function. *J. Clin. Anesth.* **1997**, *9*, 179–183. [[CrossRef](#)]



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).

Article

Assessment of Ultra-Short Heart Variability Indices Derived by Smartphone Accelerometers for Stress Detection

Federica Landreani ¹, Andrea Faini ², Alba Martin-Yebra ^{1,3}, Mattia Morri ¹, Gianfranco Parati ^{2,4} and Enrico Gianluca Caiani ^{1,5,*}

¹ Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, 20133 Milan, Italy

² Istituto Auxologico Italiano, IRCCS, Department of Cardiovascular Neural and Metabolic Sciences, S. Luca Hospital, 20149 Milan, Italy

³ Department of Biomedical Engineering, Lund University, 22100 Lund, Sweden

⁴ Department of Medicine and Surgery, Università di Milano-Bicocca, 20126 Milan, Italy

⁵ Consiglio Nazionale delle Ricerche, Istituto di Elettronica e di Ingegneria dell'Informazione e delle Telecomunicazioni, 20133 Milan, Italy

* Correspondence: enrico.caiani@polimi.it; Tel.: +39-02-2399-3390

Received: 28 June 2019; Accepted: 23 August 2019; Published: 28 August 2019

Abstract: Body acceleration due to heartbeat-induced reaction forces can be measured as mobile phone accelerometer (m-ACC) signals. Our aim was to test the feasibility of using m-ACC to detect changes induced by stress by ultra-short heart rate variability (USV) indices (standard deviation of normal-to-normal interval—SDNN and root mean square of successive differences—RMSSD). Sixteen healthy volunteers were recruited; m-ACC was recorded while in supine position, during spontaneous breathing at rest conditions (REST) and during one minute of mental stress (MS) induced by arithmetic serial subtraction task, simultaneous with conventional electrocardiogram (ECG). Beat occurrences were extracted from both ECG and m-ACC and used to compute USV indices using 60, 30 and 10 s durations, both for REST and MS. A feasibility of 93.8% in the beat-to-beat m-ACC heart rate series extraction was reached. In both ECG and m-ACC series, compared to REST, in MS the mean beat duration was reduced by 15% and RMSSD decreased by 38%. These results show that short term recordings (up to 10 s) of cardiac activity using smartphone's accelerometers are able to capture the decrease in parasympathetic tone, in agreement with the induced stimulus.

Keywords: ballistocardiography; seismocardiography; ultra-short heart rate variability; stress evaluation; smartphone; accelerometers

1. Introduction

Technology developments and device miniaturization have opened the possibility for hand-held devices such as smartphones to be used for physiological data collection. Through their embedded tri-axial accelerometers, the mobile phone is sensitive enough to record the vibrations generated by the beating heart, as accelerometer signal (m-ACC) of milligravity (mg) level. In this way, the movements along the lateral, the normal, and the longitudinal direction can be detected.

Depending on the accelerometers position, these signals usually resemble:

- (1) The ballistocardiogram (BCG), measuring the displacement of the mass of ejected blood from the ventricles through the aorta and then towards the peripheral circulation, represented by a series of systolic (I, J, K) waves describing the forces associated to the shifting of the center of body mass [1–3];

- (2) The seismocardiogram (SCG), capturing the sequence of mechanical cardiac events known as isovolumetric contraction (IVC), aortic valve opening (AO), and aortic valve closure (AC) relevant to the systolic period [3–6].

The heartbeat fiducial points on the m-ACC signal associated to the sharp cardiac vibration waves in concomitance to the systolic activity can be detected using an electrocardiogram (ECG)-independent processing algorithm, and used to compute the cardiac interbeat interval, thus obtaining corresponding beat-to-beat time series. The feasibility and accuracy of measuring the beat-to-beat heart rate using smartphone accelerometers has been recently demonstrated [7–13].

Heart rate variability (HRV) analysis is usually applied to the series of time intervals between consecutive R-wave peaks (RR) extracted from the ECG, thus providing quantitative markers to evaluate the influence of the autonomic nervous system (ANS) on the heart rate [14–17]. The HRV approach considers monitoring periods that may range from 5 min up to 24 h [15], providing information that may be related to physiological status such as diabetic neuropathy [17,18], myocardial dysfunction [19], or stress conditions [20]. The feasibility of applying HRV analysis to beat-to-beat series obtained by accelerometric recordings of 5 minutes length has been already proven [12].

In the context of self-monitoring individual's health and well-being status, the interest in using shorter recordings (<5 min) in stationary conditions of real-life scenarios suitable for HRV analysis is emerging, thus increasing user compliance and reliability of measurement. To this purpose, the ultra-short heart rate variability (USV) time domain indices—the standard deviation of normal-to-normal intervals (SDNNs) and the root mean square of successive differences (RMSSD)—have been proposed as a surrogate to assess the ANS influence on the heart rate [21–24].

We hypothesized that the USV analysis could be applied to assess the level of stress from the accelerometric signals acquired for short periods using a mobile phone, thus facilitating this self-assessment procedure without the need of other wearables or sensors, and overcoming the main limitation of keeping in position the device for longer periods.

Accordingly, our aim was to test the feasibility of detecting changes in the ANS state provoked by a mental task by using the beat-to-beat series from short recordings (<1 min) extracted by a mobile phone m-ACC signal. To attain that objective the USV indices were computed and compared with the indices obtained by the conventional ECG-RR series, considered the gold standard, simultaneously extracted. In addition, the ability to detect these changes using sub-segments of shorter durations (up to 10 s) was explored.

2. Materials and Methods

2.1. Study Population

A total of 16 subjects (age range 19–28, six females) were recruited, whose anthropometrics data are reported in Table 1. The experimental procedures described in this paper agreed with the ethics defined in the Helsinki Declaration of 1975, as revised in 2000. Each subject also provided voluntary written, informed consent to participate in the experimental protocol approved by the Ethical Committee of the Ospedale San Luca in Milan.

Table 1. Anthropometric characteristics of the population in terms of age, weight, height, and body mass index (BMI) expressed as the median (25th–75th percentiles).

Age (y)	Weight (Kg)	Height (m)	BMI (Kg/m ²)
22	67.5	180	21.5
(21–23)	(61–75.5)	(169–184)	(20.3–22.9)

2.2. Accelerometric Signal Acquisition

Each volunteer was studied in supine position using a smartphone (iPhone 6s, Apple Inc., Cupertino, CA, United States), positioned directly on the navel, with the phone top towards the head (Figure 1). The 3-orthogonal axis accelerometric signals (m-ACC, $fs = 100$ Hz, accelerometer sensitivity of 0.001 g) were acquired using the app 'SensorLog' v.2.4, resulting in three oriented channels corresponding to lateral (X), longitudinal (Y), and normal (Z) directions, simultaneously with a 6-leads electrocardiogram (ECG, Nexfin HD monitor, BMEYE, Amsterdam; $fs = 1000$ Hz). Despite the morphology of the m-ACC signal depends on the device position on the subject's body [9], Y and Z components showed a major informative content relevant to heartbeat occurrence. For this reason, they were chosen to be processed with the ECG-free heartbeat detection.

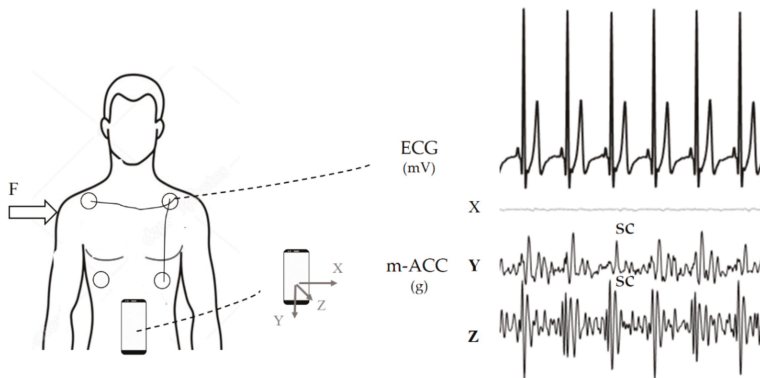


Figure 1. The electrodes and a smartphone were positioned on the subject: the electrocardiogram (ECG) signals were acquired simultaneously and synchronized by a motion artifact caused by an impulsive force (F) impressed on the subject's shoulder. The ECG and the simultaneously acquired mobile phone tri-axial accelerometric signals (m-ACC) are shown: While the lateral (X) component does not project any heartbeat vibration, the longitudinal (Y) and normal (Z) components show a clear periodic complex (SC) related to cardiac heartbeat activity.

The signals were synchronized by applying a lateral impulsive force stimulus applied on to the subject's shoulder, which was detected both by the smartphone's accelerometers and by the ECG electrodes (as a movement artifact). After a 10 min acclimation period in supine posture, the experimental protocol included two acquisitions performed sequentially (see Figure 2): The former, lasting 3 minutes with the subject breathing normally in resting conditions (REST) and, after 1 minute of readjustment, the latter lasting 1 minute with a mental stress (MS) condition. As mental arithmetic is one of the most commonly utilized laboratory psychological stressors able to increase heart rate (HR) [25–27], stress was provoked as follows: The subject was instructed before the beginning of the experiment on how this phase would have been performed: Given a starting 4-digit number told to the subject at the beginning of the mental stress acquisition, the subject had to perform silently arithmetic serial subtractions of seven from that number at a pace of one every ten seconds, and communicate the final result only at the end of the 1-min acquisition. Preventing subjects speaking during the acquisition was a design choice, as communication between the subject and the researchers would have hindered the signal quality, as the acquired signal is sensitive to vibrations that are generated by speaking.

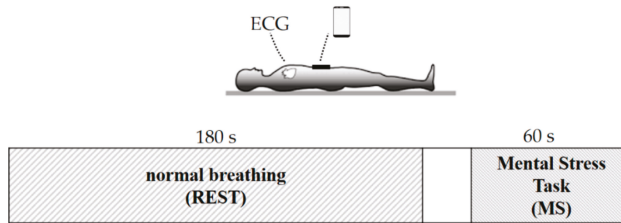


Figure 2. Schematic representation of the experimental protocol, where the subject was acquired in supine posture with a smartphone positioned on the belly above the navel. The protocol consisted in two steps: Normal breathing (REST) and the mental stress task (MS) where the subject was asked to perform arithmetic serial subtractions starting from a 4-digit number.

2.3. Signal Processing

2.3.1. Pre-Processing

To synchronize the smartphone signals with the ECG, the spike motion artifact introduced during the measurement was identified on the X component of the m-ACC signal and on the ECG lead I. After this step, the Y and Z components of the m-ACC signal were both band-pass filtered (4th order Butterworth filter): Cut-off frequencies of 5 and 25 Hz were used for Z, while 1 and 30 Hz for Y. This approach removed the out-of-band noise and breathing activity-related motion artifacts. The chest wall vibrations and the body acceleration due to the heartbeat-induced recoil forces were thus retained considering these different band-pass ranges [8].

2.3.2. Heartbeat Detection and Algorithm Performance

On the ECG signal, R peaks were detected using Pan-Tompkins algorithm [28,29] and used to derive the RR series for comparison purposes with smartphone-derived series.

On both Y and Z components of the m-ACC signal, to detect the systolic complex (SC), an ECG-free algorithm based on template matching technique was applied [30], performing the following steps:

- Signals were divided into 30 s segments (Figure 3a);
- For each segment, centered at the absolute maximum within the first 10 s, a template of 400 ms duration was extracted (Figure 3b): Such duration was chosen keeping into consideration the possible minimum duration of the heart cycle interval in case of a fast heart rate (~150 bpm);
- Computation of the cross-correlation between the identified template and the 30 s signal segment (Figure 3c);
- Searching windows centered at each maxima of the cross-correlation function (Figure 3c), dashed red) were defined and used to precisely locate each SC (red dot) as the wave with the maximum absolute amplitude on the m-ACC signal (black; Figure 3d). The length of the search window was calculated taking into consideration the mean of the previous three heartbeats durations.

Then, beat-to-beat duration SC series were calculated as the distance between two consecutive SCs, for both Y and Z components. Then, the algorithm automatically selected the optimal (OPT) component as follows (see Figure 4):

- Least-squares fitting with a 5th order polynomial $P(X)$ of the calculated SCs series for the Y and Z components (see Figure 4a,b, respectively);
- For each component, computation of the mean distance value between the SCs and the $P(X)$ corresponding value;
- The component with the lowest distance was selected for further processing.

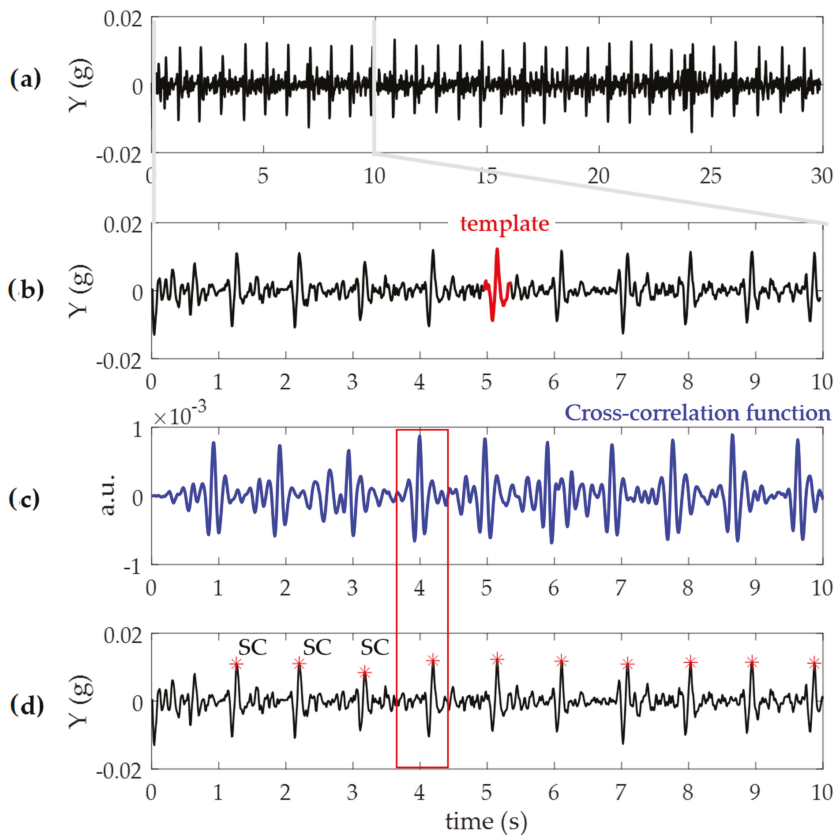


Figure 3. Schematic representation of the ECG-free heartbeat detection algorithm. (a): A 30 s segment of the m-ACC signal recording (black); (b) within the first 10 s segment, a template (red) is automatically selected; (c) the cross-correlation function (in blue) was obtained using the template with the 30 s signal (black) and the position of maximum values of cross-correlation were used to identify a search window (in red, minus sign) for each heartbeat; and (d) the windows were thus used to detect the systolic complex (SC, red dots) on the m-ACC signal (black).

This step allowed us to automatically select the series with fewer outliers in presence of possible artifacts and missed or wrong detections, thus obtaining at least a beat-to-beat series for each subject and condition, from which to proceed with the USV indices computation.

2.3.3. Ultra-Short Heart Rate Variability Indices

The SDNN and RMSSD were computed as USV time domain indices. The SDNN estimates overall HRV, while RMSSD is actually an estimate of high frequency variation in HR led by the vagal tone activity of the ANS [15,21,31,32]. They were calculated using the most central 60 s, 30 s, and 10 s signal segments of the REST and MS recordings, from both RR and OPT series. In addition, to evaluate the possible implications of different selection criteria, the USV indices were computed also for the initial and final (30 s and 10 s length) segments on the captured series.

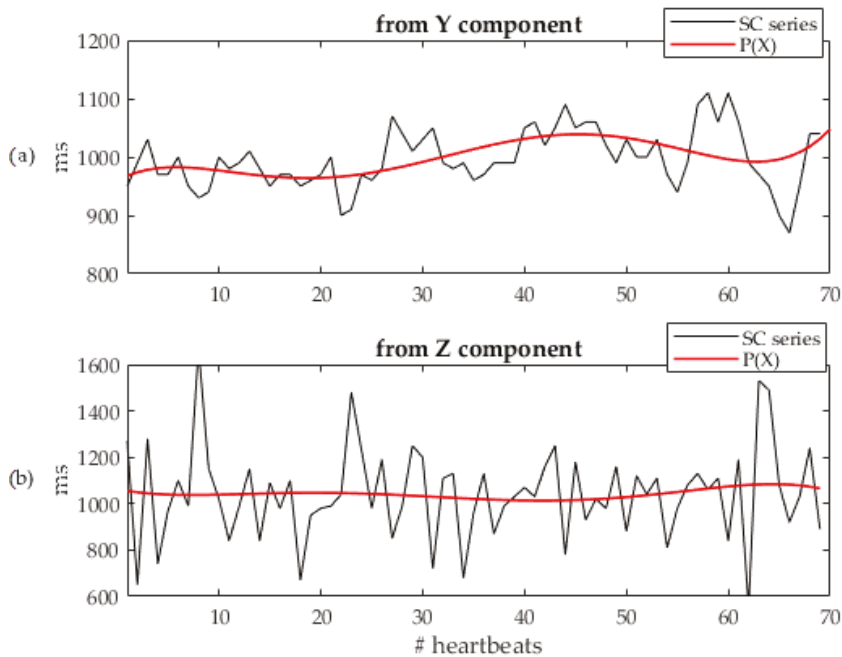


Figure 4. Polynomial interpolation series $P(X)$ (in red) over imposed to the corresponding systolic complex beat-by-beat duration series (SCs; in black), separately for Y (a) and Z (b) acceleration components. See text for more details.

2.4. Statistical Analysis

Results are presented as medians (25th–75th percentile).

The algorithm's feasibility was computed as the number of acquisitions in which at least one optimal series for the ultra-short heart rate variability analysis was obtained, with respect to the total number of acquisitions. The sensitivity was calculated as the percentage ratio between the SC detected with respect to the corresponding R-wave ECG peaks. To evaluate the accuracy of the applied ECG-free detection algorithm, all the peaks were visually inspected together with ECG annotations: The misdetections, as double detection or incorrect peaks, were identified and categorized as false positive (FP), while the missing detections were considered as false negative (FN), where the true positive (TP) corresponded to a detection in the correct position (see Equation (1)).

$$\text{accuracy} = \frac{\text{TP}}{\text{FP} + \text{FN} + \text{TP}}. \quad (1)$$

In order to test if the OPT series could represent a valid surrogate for electrodes-free heart beat duration extraction, linear correlation and Bland-Altman analyses were computed by comparison with the corresponding RR series.

A Mann-Whitney unpaired test (*: $p < 0.05$) was applied to compare the OPT versus RR series parameters (cardiac cycle median duration and USV indices) to support the hypothesis that the obtained parameters represent the same information, both in REST and MS.

Friedman test and multi-comparison Bonferroni correction were applied to test whether heartbeat mean duration and USV indices obtained from different length recordings (60 s, 30 s, and 10 s) were representative of the same distribution (#: $p < 0.05$), to support the hypothesis of using the shortest

possible acquisition length, and to evaluate the effect of the segment position (initial, middle, and final) on the obtained results ($t: p < 0.05$).

A non-parametric Wilcoxon paired test ($^*: p < 0.05$) was used to test significant differences in ANS sympatho-vagal status evidenced by the USV parameters between REST and MS, separately for the ECG and OPT series.

3. Results

3.1. Algorithm Performance

The m-ACC signals from one subject in REST and one in MS were discarded due to poor signal-to-noise ratio, thus resulting in a feasibility of the beat-to-beat heart rate series extraction of 93.8%. The algorithm automatically selected the Z component as the best component for heartbeat detection in 10/15 subjects for REST and in 11/15 subjects for MS.

Compared to the 1784 R-ECG peaks in REST and 1271 in MS, the ECG-free detection algorithm detected correctly 1754 beats in REST and 1246 beats in MS, thus resulting in an algorithm sensitivity of 98.3% and 98%, respectively, with high accuracy (98%) reached.

In addition, due the presence of ectopics beats in both REST and MS OPT series that could result in erroneous USV indices, another subject was discarded. Accordingly, to allow a paired comparison, 13 subjects were considered for further analysis.

3.2. Cardiac Cycle Duration

In Figure 5, an example of the RR and OPT beat-to-beat variability series derived from a representative subject at REST and during MS are shown superimposed: It is possible to appreciate in both conditions the correspondence of the measured heart cycle durations.

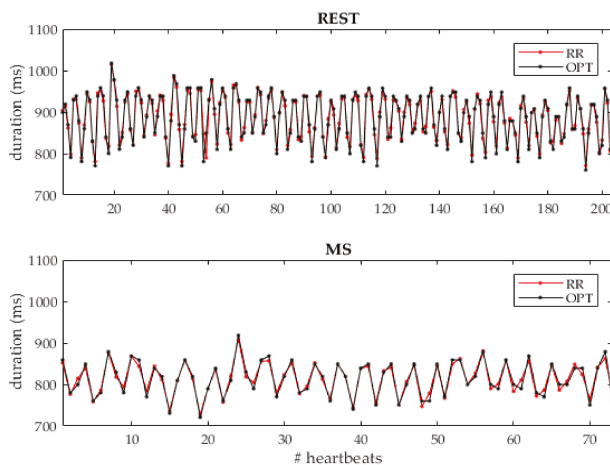


Figure 5. Example of the beat-to-beat RR series (red) and optimal (OPT; black) heart cycle duration series obtained at REST and MS for a representative subject. At REST, the measurement lasts 180 s, while the MS lasts 60 s.

In Figure 6, the result of linear correlation and Bland-Altman analyses by comparing the gold standard RR versus the corresponding beat-by-beat OPT series are presented. It is possible to appreciate the high r^2 value (0.99) and the narrow confidence interval (CI = ± 33 ms, ± 2 SD) achieved by the proposed method, tested globally in a range of heartbeat duration from 447 ms up to 1337 ms.

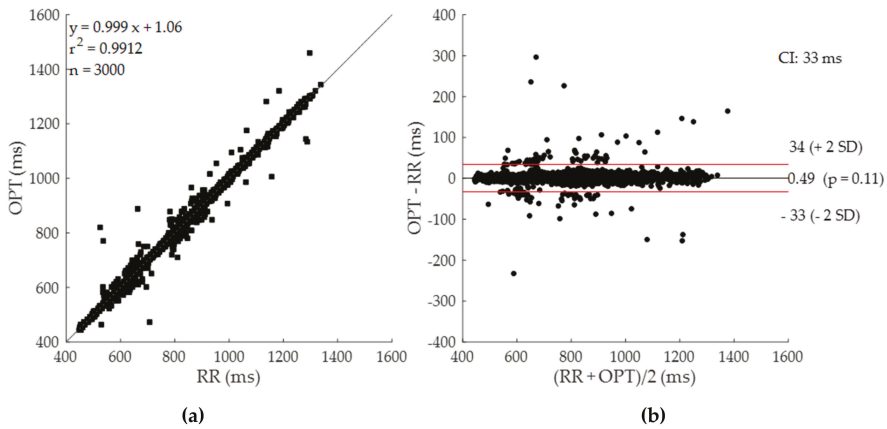


Figure 6. Linear correlation (a) and Bland-Altman (b) analyses obtained considering all the series (RR and OPT) obtained at REST and MS for a total of 3000 heartbeats. R^2 coefficient and confidence’s intervals (CI = ± 2 SD) are shown.

In Table 2 and Figure 7, the results relevant to the heart cycle duration obtained considering different durations and segment positions on the ECG (RR) and the optimal (OPT) series by the smartphone’s accelerometer are presented. No statistical differences (Mann-Whitney and Friedman test) were found between RR and OPT for each duration and position on captured series, in both conditions. Changes induced by the stress condition were visible independently from the duration and segment position in both RR and OPT. In particular, for a 60 s duration, both the RR and OPT were found reduced by 16% (10%–26%); by considering the most central segment the RR and OPT were reduced by 17% (10%–27%) for 30 s duration, while using 10 s duration the RR and OPT were reduced by 20% (9%–27%). Interestingly, a trend of decrease in heartbeat duration during MS was visible within the 60 s period.

Table 2. Heart beat duration (ms) obtained from the ECG (RR) and the optimal (OPT) series by the smartphone’s accelerometer expressed as a median (25th–75th percentiles), for 60 s, 30 s, and 10 s duration and different segments (initial, central, final), in control (REST) and mental stress (MS) conditions. *: $p < 0.05$ REST vs. MS (Wilcoxon test).

	Heart Cycle Duration (ms)					
	REST	MS	REST	MS	REST	MS
	60 "					
RR			887 (780–985)	740 * (580–827)		
OPT			887 (780–987)	741 * (584–830)		
	30 "					
	initial		central		final	
RR	909 (813–967)	814 * (630–852)	888 (779–992)	748 * (570–832)	869 (769–1002)	667 * (573–850)
OPT	912 (813–968)	814 * (631–852)	888 (779–993)	747 * (571–832)	869 (768–1002)	668 * (576–850)
	10 "					
RR	893 (827–993)	792 * (640–824)	886 (774–982)	737 * (575–822)	867 (780–1029)	662 * (573–898)
OPT	892 (828–924)	792 * (640–823)	887 (774–982)	740 * (576–822)	866 (780–1029)	659 * (575–898)

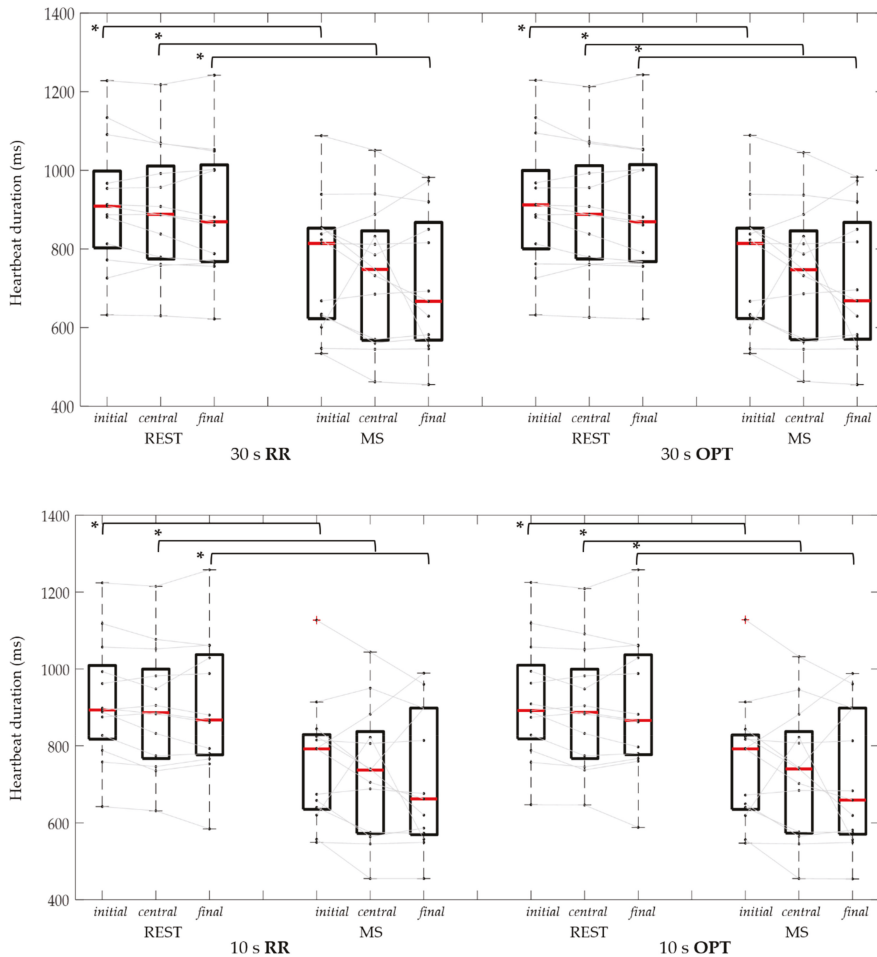


Figure 7. Heart beat duration distributions and individual data of RR and OPT series, in control (REST) and mental stress (MS) conditions, using 30 s (upper panel) and 10 s (bottom panel) signal segments considered at initial, central, and final position on the considered series. *: $p < 0.05$ REST vs. MS (Wilcoxon test).

3.3. Ultra-Short Heart Rate Variability Indices

In Table 3 and Figure 8, and Table 4 and Figure 9, the results of USV indices (SDNN and RMSSD, respectively) obtained considering different durations and positions of RR and OPT series are reported. It is possible to notice that results obtained from RR and OPT series were very similar, in each tested condition.

Regarding SDNN, no statistical difference was found between REST and MS, both using ECG and OPT, when considering the whole 60 s duration, or the initial and central portions of the shorter durations. On the contrary, the SDNN was significantly reduced from CTRL to MS, for both RR and OPT series, in the final 30 s and 10 s segments, previously associated to shortened mean heart duration.

The Friedman test showed that SDNN had values proportional with the duration of the RR and OPT series during MS, with statistical significance when comparing the 60 s with the central 30 s and 10 s indices.

Table 3. SDNN values (ms) obtained from the ultra-short heart rate variability analysis using RR and optimal (OPT) series. SDNN are expressed as medians (25th–75th percentiles), for 60 s, 30 s, and 10 s signal segments duration and different segments (initial, central, and final), both in control (REST) and mental stress (MS) conditions. *: $p < 0.05$ REST vs. MS (Wilcoxon test); #: $p < 0.016$ vs. 60 s and †: $p < 0.016$ between segment position (Friedman test and Bonferroni correction).

	REST		MS		SDNN (ms)	
	REST	MS	REST	MS	REST	MS
	60 "					
RR			48 (41–64)	42 (27–72)		
OPT			54 (43–66)	45 (32–72)		
	initial		30 " central		final	
RR	44 (36–62)	44 † (24–54)	42 (35–68)	39 #, † (28–61)	47 (34–58)	26 *, † (19–34)
OPT	46 (39–73)	47 † (24–55)	47 (40–65)	49 (27–51)	49 (37–69)	29 *, † (21–38)
	10 "					
RR	45 (31–65)	24 (19–49)	39 (19–47)	26 # (16–42)	44 (29–78)	21 * (18–28)
OPT	47 (36–62)	40 (22–50)	38 (35–66)	31 # (21–44)	46 (33–79)	25 * (21–30)

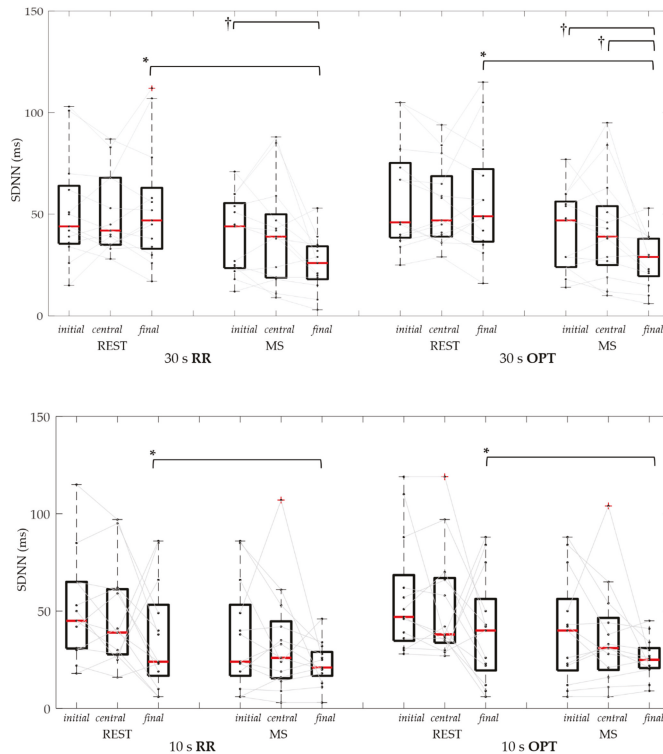


Figure 8. Standard deviation of normal-to-normal interval (SDNN) distributions and individual data of RR and OPT series, in control (REST) and mental stress (MS) conditions, using 30 s (upper panel) and 10 s (bottom panel) signal segments considered at the initial, central, and final series portion. *: $p < 0.05$ REST vs. MS (Wilcoxon test); †: $p < 0.016$ among segment positions (Friedman test and Bonferroni correction).

Table 4. RMSSD values (ms) obtained from the ultra-short heart rate variability analysis using RR and optimal (OPT) series. RMSSD are expressed as medians (25th–75th percentiles), for 60 s, 30 s, and 10 s (initial, central, and final) signal segments duration, both in control (REST) and mental stress (MS) conditions. *: $p < 0.05$ REST vs. MS (Wilcoxon test); #: $p < 0.016$ vs. 60 s (Friedman test and Bonferroni correction); †: $p < 0.016$ between segment position (Friedman test and Bonferroni correction).

		RMSSD (ms)					
		REST	MS	REST	MS	REST	MS
		60 "					
RR				39 (30–65)	23* (14–23)		
OPT				48 (38–76)	36* (25–48)		
		30 "					
		initial		Central		final	
RR		57 (30–63)	31*† (13–40)	40 (24–67)	22* (10–30)	49 (29–66)	19*† (11–27)
OPT		60† (38–91)	31* (21–49)	51† (38–73)	36* (24–45)	48# (38–77)	27* (18–39)
		10 "					
RR		54 (36–69)	32* (15–48)	41 (27–61)	20* (11–30)	57 (35–65)	21* (14–29)
OPT		61 (38–77)	36* (24–52)	47 (36–78)	27* (20–35)	56 (38–70)	26* (21–24)

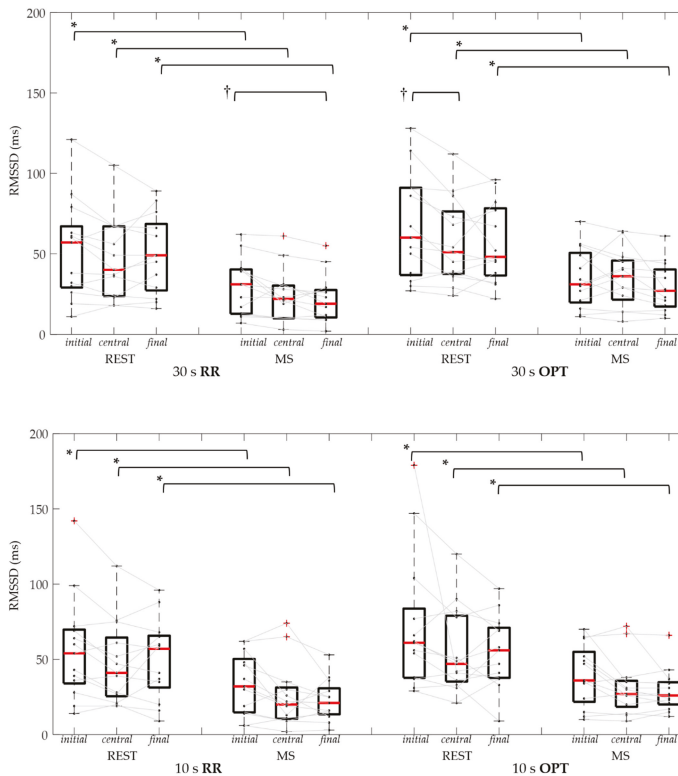


Figure 9. Root mean square of successive differences (RMSSD) distributions and individual data of RR and OPT series, in control (REST) and mental stress (MS) conditions, using 30 s (upper panel) and 10 s (bottom panel) signal segments considered at initial, central, and final series portion. *: $p < 0.05$ REST vs. MS (Wilcoxon test); †: $p < 0.016$ vs. initial portion (Friedman test and Bonferroni correction).

Regarding RMSSD, independently from the duration or the position of the segment, the mental stress condition resulted in reduced values of the computed parameter compared to REST for both RR and OPT series. In particular, for the 60 s segment the RMSSD in RR was reduced by 38% (26%–71%), and by 40% (8%–68%) in OPT. Considering the central 30 s segment, it was reduced by 45% (27%–72%) in RR, and by 46% (22%–64%) in OPT, while for the central 10 s segments it was reduced by 53% (13%–73%) in RR, and by 49% (8%–66%) in OPT.

4. Discussion

Current smartphone technology and embedded sensors have the potential to acquire signals related to cardiac activity. In addition to the use of the on-board camera to derive the pulsation of the skin capillary blood flow in the fingertips from which to obtain the pulse rate [33], micro-electro-mechanical systems technology embedded in smartphones potentially allows measuring heart mechanical activity by acquiring vibrational signals when positioned on the body [9]. The potential of these approaches in using the smartphone as a source of vital parameters stands in the ability to acquire these signals anytime and anywhere, without the need of additional peripherals, thus improving patient empowerment through self-measurement.

However, in order to be accepted by the medical community, the potential value relevant to the use of these technologies needs to be proved. While the use of smartphone cameras using photoplethysmography for early detection of atrial fibrillation, based on the analysis of beat-by-beat duration variability series, has initially proved its value in a prospective two-center, international clinical validation study [33], the validation of using smartphone accelerometers is still limited and again focusing on beat-by-beat duration variability only [8,10,34].

Our hypothesis was that the beat-to-beat heart rate series derived by a smartphone without any peripheral were suitable for the USV analysis, potentially useful for stress evaluation from a short time series. To this purpose, the feasibility of detecting changes in the ANS sympatho-vagal state provoked by a mental task in normal volunteers was tested.

The choice of acquiring the accelerometric signal from the navel with the subject in the supine position was guided from previous results we obtained in a preliminary analysis [8], in which the feasibility of acquisitions and accuracy of the beat-by-beat measurements were tested with the smartphone positioned on the thorax around the cardiac apex position and on the navel, with subjects in the supine and standing posture. With the aim of including both male and females volunteers in this study, thus overcoming possible gender-related limitations with the thoracic position, and in the perspective of a real applicability scenario including daily self-assessment in patients performed in supine resting condition in the morning, when clinical guidance suggest to obtain these kind of measurement [35,36], the navel position was chosen as it guarantees easy reproducibility and is not influenced by gender-related body morphology. Moreover, in this position the smartphone does not require any external accessory to be worn.

In our work, the smartphone was used as a sensor for accelerometric data acquisition, showing a good feasibility (93.8%). Signal quality allowed further analysis, including automatic parameters extraction at least in one component of the smartphone tri-axial accelerometers. Acceptable limits of agreement corresponding to ± 10 bpm for the fastest heart rate analyzed (134 bpm) and to ± 1 bpm for the lowest one (45 bpm) were obtained, in agreement with previous studies [8,30,37].

The different durations of considered signal segments were selected as a compromise, taking into consideration an acquisition as short as possible in a hypothetical user-driven scenario, but at the same time being able to record a reasonable amount of data to provide reliable measurements. As already stated, for mobile applications short-term measurements are desirable for USV analysis, since the conventional five minute long recordings might be inadequately long and prone to artifacts.

The results of this study are in agreement with our preliminary findings [30] obtained in only six subjects, thus confirming the feasibility of applying USV to SC beat-to-beat measurements derived by the smartphone accelerometers. From the obtained results, we showed that the median heart

rate could be accurately estimated from very short segments (even from 10 s acquisition) of m-ACC signals, without differences when compared to the ECG results. Through the mental stress task, the median heartbeat duration was found significantly shortened when compared to the rest condition, as physiological expected, and in line with what observed using the ECG derived series.

For both RR and OPT series, the obtained results for USV parameters showed a decrease in RMSSD induced by the stress condition visible for each signal duration and independently from the criterion used to choose the segment position. In SDNN, this was visible only when considering the last portion of the signal, for both 30 s and 10 s durations. As mental stress is known to increase sympathetic activity, as revealed by the increased heart rate and reduced SDNN, the mental exercise induced a significant decrease in parasympathetic activation, in agreement with the induced stimulus, reflected by the significant decrease in RMSSD in both ECG and OPT series. These results highlight the potential to use the smartphone's accelerometers to derive cardiac beat-to-beat measurements, able to monitor a stress-induced situation as a decrease from a baseline value, using very short acquisitions (even from 10 s).

Once confirmed in a larger number of subjects, a 10 s or 30 s acquisition could be considered as an easy, non-invasive way for the self-evaluation of stress using accelerometers already embedded in the mobile phone. This technology could have potential benefits in both cardiac disease prevention and self-assessment of patients with chronic disease (such as diabetes [17,38], or in patients where an imbalance of cardiac autonomic activity plays an important role, such as in coronary heart disease [39–42]), where simple but effective monitoring tools are needed in order to have reliable at-home measurements, managed directly by the patient.

A possible limitation of this study is constituted by the low sampling rate (100 Hz) of the accelerometric signal obtained from the mobile phone, which could reflect a lower precision in the computed USV parameters. However, as visible from Tables 3 and 4, the reported changes in values of SDNN and RMSSD between REST and MS in OPT series were very similar to those obtained by the ECG sampled at 1000 Hz, but with a distribution of values in REST and MS, both for SDNN and RMSSD, that was always higher than those of the ECG, related to the lower sampling rate. In addition, previous studies [43] showed that a 100 Hz sampling frequency could be considered still appropriate to produce acceptable results for time-domain heart-variability analysis (but not for frequency domain parameters).

5. Conclusions

The beat-to-beat heart rate variability series derived by a smartphone's accelerometers were able to detect the changes in ultra-short term HRV indices relevant to a change in the sympatho-vagal balance activation, induced by a stressor stimulus. In particular, the USV feature of RMSSD obtained by the m-ACC signal of up to a 10 sec duration could be used as a potential marker to estimate the stress level compared to a control value. This simple approach and its potential application in stress evaluation generate a new value of using the embedded smartphone accelerometers as a new tool for self-tracking cardiac activity.

Author Contributions: Conceptualization, E.G.C., F.L. formal analysis, F.L., M.M.; data curation, F.L.; writing—original draft preparation, F.L., E.G.C.; writing—review and editing, F.L., E.G.C., A.M.-Y.; G.P; A.F.; visualization, F.L., E.G.C., A.M.-Y., G.P., A.F.; supervision, E.G.C.; project administration, E.G.C.

Funding: This research was partially supported by the Italian Space Agency, contract 2018-7-U.0, PI Enrico Caiani.

Acknowledgments: We would like to thank the volunteers that participated to this study.

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

References

1. Starr, I.; Rawson, A.J.; Schroeder, H.A.; Joseph, N.R. Studies on the estimation of cardiac output in man, and of abnormalities in cardiac function, from the heart's recoil and the blood's impacts; The ballistocardiogram. *Am. J. Physiol-Legacy Content* **1939**, *127*, 1–28. [[CrossRef](#)]
2. Gubner, R.S.; Rodstein, M.; Ungerleider, H.E. Ballistocardiography; An appraisal of technic, physiologic principles, and clinical value. *Circulation* **1953**, *7*, 268–286. [[CrossRef](#)] [[PubMed](#)]
3. Inan, O.T.; Migeotte, P.F.; Park, K.S.; Etemadi, M.; Tavakolian, K.; Casanella, R.; Di Rienzo, M. Ballistocardiography and seismocardiography: A review of recent advances. *IEEE J-BHI* **2015**, *19*, 1414–1427. [[CrossRef](#)] [[PubMed](#)]
4. Salerno, D.M.; Zanetti, J. Seismocardiography for monitoring changes in left ventricular function during ischemia. *Chest* **1991**, *100*, 991–993. [[CrossRef](#)] [[PubMed](#)]
5. Zanetti, J.M.; Salerno, D.M. Seismocardiography: A technique for recording precordial acceleration. In Proceedings of the Fourth Annual IEEE Symposium on Computer-Based Medical Systems, Baltimore, MD, USA, 12–13 May 1991; pp. 4–9. [[CrossRef](#)]
6. Zanetti, J.M.; Tavakolian, K. Seismocardiography: Past, present and future. In Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka, Japan, 3–7 July 2013; pp. 7004–7007. [[CrossRef](#)]
7. Landreani, F.; Martin-Yebra, A.; Casellato, C.; Pavan, E.; Frigo, C.; Migeotte, P.F.; Caiani, E.G. Feasibility study for beat-to-beat heart rate detection by smartphone's accelerometers. In Proceedings of the E-Health and Bioengineering Conference (EHB), Iasi, Romania, 19–21 November 2015; pp. 3–6. [[CrossRef](#)]
8. Landreani, F.; Martin-Yebra, A.; Casellato, C.; Frigo, C.; Pavan, E.; Migeotte, P.F.; Caiani, E.G. Beat-to-beat heart rate detection by smartphone's accelerometers: Validation with ECG. In Proceedings of the 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 16–20 August 2016; pp. 525–528. [[CrossRef](#)]
9. Landreani, F.; Caiani, E.G. Smartphone accelerometers for the detection of heart rate. *Expert Rev. Med. Devic.* **2017**, *14*, 935–948. [[CrossRef](#)] [[PubMed](#)]
10. Sיעiński, S.; Kostka, P. Determining Heart Rate Beat-to-Beat from Smartphone Seismocardiograms: Preliminary Studies. In *Proceedings of the Innovations in Biomedical Engineering (IBE), Zabrze, Poland, 19–20 October 2017*; Gzik, M., Tkacz, E., Paszenda, Z., Piętka, E., Eds.; Springer: Cham, Switzerland, 2018; Volume 623, pp. 133–140. [[CrossRef](#)]
11. Gavriel, C.; Parker, K.H.; Faisal, A.A. Smartphone as an ultra-low cost medical tricorder for real-time cardiological measurements via ballistocardiography. In Proceedings of the IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN), Cambridge, MA, USA, 9–12 June 2015. [[CrossRef](#)]
12. Ramos-Castro, J.; Moreno, J.; Miranda-Vidal, H.; Garcia-Gonzalez, M.A.; Fernandez-Chimeno, M.; Rodas, G.; Capdevila, L. Heart rate variability analysis using a seismocardiogram signal. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, USA, 28 August–1 September 2012; pp. 5642–5645. [[CrossRef](#)]
13. Koivisto, T.; Lahdenoja, O.; Hurnanen, T.; Knuutila, T.; Vasankari, T.; Kiviniemi, T.; Pankaala, M. Detecting atrial fibrillation via existing smartphones without any add-ons. In Proceedings of the ESC Congress, Rome, Italy, 27–31 August 2016.
14. Rodriguez, J.; Blaber, A.P.; Kneihsl, M.; Ruedl, R.; Green, D.A.; Broadbent, J.; Goswami, N. Poststroke alterations in heart rate variability during orthostatic challenge. *Medicine* **2017**, *96*, e5989. [[CrossRef](#)] [[PubMed](#)]
15. Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *Eur. Heart J.* **1996**, *17*, 354–381. [[CrossRef](#)]
16. Thayer, J.F.; Yamamoto, S.S.; Brosschot, J.F. The relationship of autonomic imbalance, heart rate variability and cardiovascular disease risk factors. *Int. J. Cardiol.* **2010**, *141*, 122–131. [[CrossRef](#)] [[PubMed](#)]
17. Chessa, M.; Butera, G.; Lanza, G.A.; Bossone, E.; Delogu, A.; De Rosa, G.; Carminati, M. Role of heart rate variability in the early diagnosis of diabetic autonomic neuropathy in children. *Herz* **2002**, *27*, 785–790. [[CrossRef](#)]

18. Pagani, M.; Malfatto, G.; Pierini, S.; Casati, R.; Masu, A.M.; Poli, M.; Malliani, A. Spectral analysis of heart rate variability in the assessment of autonomic diabetic neuropathy. *J. Auton. Nervous Syst.* **1988**, *23*, 143–153. [CrossRef]
19. Nolan, J.; Flapan, A.D.; Capewell, S.; MacDonald, T.M.; Neilson, J.M.; Ewing, D.J. Decreased cardiac parasympathetic activity in chronic heart failure and its relation to left ventricular function. *Br Heart J.* **1992**, *67*, 482–485. [CrossRef] [PubMed]
20. Taelman, J.; Vandepuit, S.; Spaepen, A.; Van Huffel, S. Influence of Mental Stress on Heart Rate and Heart Rate Variability. In *Proceedings of the 4th European Conference of the International Federation for Medical and Biological Engineering (IFMBE)*; Vander Sloten, J., Verdonck, P., Nyssen, M., Hauelsen, J., Eds.; Springer: Berlin/Heidelberg, Germany, 2009; Volume 22, pp. 395–400. [CrossRef]
21. Munoz, M.L.; Van Roon, A.; Riese, H.; Thio, C.; Oostenbroek, E.; Westrik, I.; Snieder, H. Validity of (Ultra-)Short recordings for heart rate variability measurements. *PLoS ONE* **2015**, *10*, 1–15. [CrossRef] [PubMed]
22. Baek, H.J.; Cho, C.-H.; Cho, J.; Woo, J.-M. Reliability of Ultra-Short-Term Analysis as a Surrogate of Standard 5-Min Analysis of Heart Rate Variability. *Telemed. E-Health* **2015**, *21*. [CrossRef] [PubMed]
23. Pecchia, L.; Castaldo, R.; Montesinos, L.; Melillo, P. Are ultra-short heart rate variability features good surrogates of short-term ones? State-of-the-art review and recommendations. *Heal. Tech. Lett.* **2018**, *5*, 94. [CrossRef] [PubMed]
24. Castaldo, R.; Montesinos, L.; Melillo, P.; James, C.; Pecchia, L. Ultra-short term HRV features as surrogates of short term HRV: A case study on mental stress detection in real life. *BMC Med. Inform. Decis. Mak.* **2019**, *19*, 12. [CrossRef] [PubMed]
25. Sloan, R.P.; Korten, J.B.; Myers, M.M. Components of heart rate reactivity during mental arithmetic with and without speaking. *Physiol. Behav.* **1991**, *50*, 1039–1045. [CrossRef]
26. Brown, T.G.; Szabo, A.; Seragianian, P. Physical versus psychological determinants of heart rate reactivity to mental arithmetic. *Psychophysiology* **1988**, *25*, 532–537. [CrossRef]
27. Sharpley, C.F.; Kamen, P.; Galatsis, M.; Heppel, R.; Veivers, C.; Claus, K. An examination of the relationship between resting heart rate variability and heart rate reactivity to a mental arithmetic stressor. *Appl. Psychophysiol. Biofeedback* **2000**, *25*, 143–153. [CrossRef]
28. Pan, J.; Tompkins, W.J. A Real-Time QRS Detection Algorithm. *IEEE Trans. Biomed. Eng* **1985**, *BME-32*, 230–236. [CrossRef]
29. Sedghamiz, H. Matlab Implementation of Pan Tompkins ECG QRS Detector. 2014. Available online: https://www.researchgate.net/publication/313673153_Matlab_Implementation_of_Pan_Tompkins_ECG_QRS_detector. (accessed on 25 August 2019).
30. Landreani, F.; Morri, M.; Martin-Yebra, A.; Casellato, C.; Pavan, E.; Frigo, C.; Caiani, E.G. Ultra-short-term heart rate variability analysis on accelerometric signals from mobile phone. In *Proceedings of the E-Health and Bioengineering Conference (EHB)*, Sinaia, Romania, 22–24 June 2017; pp. 241–244. [CrossRef]
31. Salahuddin, L.; Cho, J.; Jeong, M.G.; Kim, D. Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings. In *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, Lyon, France, 22–26 August 2007; pp. 4656–4659. [CrossRef]
32. Elghozi, J.-L.; Julien, C. Sympathetic control of short-term heart rate variability and its pharmacological modulation. *Fundam. Clin. Pharmacol.* **2007**, *21*, 337–347. [CrossRef]
33. Brasier, N.; Raichle, C.J.; Dörr, M.; Becke, A.; Nohturfft, V.; Weber, S.; Eckstein, J. Detection of atrial fibrillation with a smartphone camera: First prospective, international, two-centre, clinical validation study (DETECT AF PRO). *Europace* **2019**, *21*, 41–47. [CrossRef] [PubMed]
34. Mohamed, R.; Youssef, M. HeartSense: Ubiquitous Accurate Multi-Modal Fusion-based Heart Rate Estimation Using Smartphones. In *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*; ACM: New York, NY, USA, 2017; Volume 1, pp. 1–18. [CrossRef]
35. Böhm, M.; Reil, J.-C.; Deedwania, P.; Kim, J.B.; Borer, J.S. Resting heart rate: risk indicator and emerging risk factor in cardiovascular disease. *Am. J. Medic.* **2015**, *128*, 219–228. [CrossRef]
36. Hamill, V.; Ford, I.; Fox, K.; Böhm, M.; Borer, J.S.; Ferrari, R.; Swedberg, K. Repeated heart rate measurement and cardiovascular outcomes in left ventricular systolic dysfunction. *Am. J. Medic.* **2015**, *128*, 1102–1108.e6. [CrossRef] [PubMed]

37. Landreani, F.; Hossein, A.; Golier, D.; Caiani, E.; Van de Borne, P.; Rabineau, J.; Migeotte, P.-F. Heartbeat detection using three-axial seismocardiogram acquired by mobile phone. In Proceedings of the Computing in Cardiology Conference (CinC), Maastricht, The Netherlands, 23–26 September 2018. [[CrossRef](#)]
38. Nussinovitch, U.; Cohen, O.; Kaminer, K.; Ilani, J.; Nussinovitch, N. Evaluating reliability of ultra-short ECG indices of heart rate variability in diabetes mellitus patients. *J. Diabetes Complicat.* **2012**, *26*, 450–453. [[CrossRef](#)] [[PubMed](#)]
39. Stein, P.K.; Carney, R.M.; Freedland, K.E.; Skala, J.A.; Jaffe, A.S.; Kleiger, R.E.; Rottman, J.N. Severe depression is associated with markedly reduced heart rate variability in patients with stable coronary heart disease. *J. Psychosom. Res.* **2000**, *48*, 493–500. [[CrossRef](#)]
40. La Rovere, M.T.; Pinna, G.D.; Maestri, R.; Mortara, A.; Capomolla, S.; Febo, O.; Cobelli, F. Short-term heart rate variability strongly predicts sudden cardiac death in chronic heart failure patients. *Circulation* **2003**, *107*, 565–570. [[CrossRef](#)] [[PubMed](#)]
41. Carney, R.M.; Saunders, R.D.; Freedland, K.E.; Stein, P.; Rich, M.W.; Jaffe, A.S. Association of depression with reduced heart rate variability in coronary artery disease. *Am. J. Cardiol.* **1995**, *76*, 562–564. [[CrossRef](#)]
42. Kupari, M.; Virolainen, J.; Koskinen, P.; Tikkanen, M.J. Short-term heart rate variability and factors modifying the risk of coronary artery disease in a population sample. *Am. J. Cardiol.* **1993**, *72*, 897–903. [[CrossRef](#)]
43. Kwon, O.; Jeong, J.; Kim, H.B.; Kwon, I.H.; Park, S.Y.; Kim, J.E.; Choi, Y. Electrocardiogram sampling frequency range acceptable for heart rate variability analysis. *Healthc. Inform. Res.* **2018**, *24*, 198. [[CrossRef](#)]



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).

Article

Sensors for Expert Grip Force Profiling: Towards Benchmarking Manual Control of a Robotic Device for Surgical Tool Movements [†]

Michel de Mathelin, Florent Nageotte, Philippe Zanne and Birgitta Dresp-Langley *

ICube Lab, UMR 7357 CNRS, Robotics Department, University of Strasbourg, 6700 Strasbourg, France; demathelin@unistra.fr (M.M.); nageotte@unistra.fr (F.N.); zanne.philippe@unistra.fr (P.Z.)

* Correspondence: birgitta.dresp@unistra.fr

[†] This paper is an extended version of our previous work published as “Batmaz, A.U.; Falek, A.M.; Zorn, L.; Nageotte, F.; Zanne, P.; de Mathelin, M.; Dresp-Langley, B. Novice and expert behavior while using a robot controlled surgery system, IEEE Proceedings of BioMed2017, Innsbruck, Austria, 20–21 February 2017”.

Received: 26 July 2019; Accepted: 17 October 2019; Published: 21 October 2019

Abstract: *STRAS* (Single access Transluminal Robotic Assistant for Surgeons) is a new robotic system based on the Anubis[®] platform of Karl Storz for application to intra-luminal surgical procedures. Pre-clinical testing of *STRAS* has recently permitted to demonstrate major advantages of the system in comparison with classic procedures. Benchmark methods permitting to establish objective criteria for ‘expertise’ need to be worked out now to effectively train surgeons on this new system in the near future. *STRAS* consists of three cable-driven sub-systems, one endoscope serving as guide, and two flexible instruments. The flexible instruments have three degrees of freedom and can be teleoperated by a single user via two specially designed master interfaces. In this study, small force sensors sewn into a wearable glove to ergonomically fit the master handles of the robotic system were employed for monitoring the forces applied by an expert and a trainee (complete novice) during all the steps of surgical task execution in a simulator task (*4-step-pick-and-drop*). Analysis of grip-force profiles is performed sensor by sensor to bring to the fore specific differences in handgrip force profiles in specific sensor locations on anatomically relevant parts of the fingers and hand controlling the master/slave system.

Keywords: robotic assistant systems for surgery; expertise; pick-and-drop simulator task; grip force profiles; grip force control

1. Introduction

Flexible systems such as endoscopes are widely used for performing minimally invasive surgical interventions, as in intraluminal procedures or single port laparoscopy. Surgical platforms have been developed by companies and by laboratories to improve the capabilities of these flexible systems, for instance by providing additional degrees of freedom (DoF) to the instruments or triangulation configurations [1,2]. In classic intraluminal procedures, the high number of DoF to be controlled represents a constraint, where several expert technicians, including the surgeon, have to work together in a complex environment. Robot assistance has been identified as a solution to this problem relative to the use of flexible systems in minimally invasive surgery [3], which explains the motivation for developing the new, teleoperated robotic system put to work in this study here. The goal of *STRAS* is to optimally assist the expert surgeon in minimally invasive procedures [4], and the design is based on the Anubis[®] platform developed by Karl Storz and the IRCAD [5]. Previous studies on *STRAS* were focused on the system architecture and the control theory of the application [4–6]. In minimally invasive surgical systems for endoscopic surgery, surgeons need to operate master interfaces to control

the endoscope and surgical instruments. They need to be able to have optimal skills in controlling the system and the user interface for targeted manipulation of the remote-controlled slave system as well as to cope with the overall complexity of the design. Such expertise can only be achieved by learning to optimally master the control mechanisms through practice in a simulator task and in vivo. Human control of endoscopic surgical systems may benefit from robotic surgical assistance [7]. Previous studies were focused on tool-tip pressures and tactile feedback effects, rather than on the grip forces applied during manipulation of the handles [8]. The system described here was designed without force feedback, and maneuver control is therefore based solely on visual feedback from the 2D images provided by an endoscopic fisheye camera and displayed on a screen. Anthropometric data from the literature suggest that, with or without force feedback, dynamic changes in perceptual hand and body schema representations and cognitive motor programming occur inevitably after repeated tool use [9,10]. These cognitive changes reflect the processes which highly trained surgeons go through in order to adapt to the visual and tactile constraints of laparoscopic surgical interventions. Experts perform tool-mediated image-guided tasks significantly quicker than trainees, with significantly fewer tool movements, shorter tool trajectories, and fewer grasp attempts [11]. Additionally, an expert tends to focus attention mainly on target locations, while novices split their attention between trying to focus on the targets and, at the same time, trying to track the surgical tools. This reflects a common strategy for controlling goal-directed hand movements in non-trained operators in various goal-directed manual tasks [12], often considerably affecting task execution times. Such strategy variables are also likely to influence grip forces while manipulating the control sticks of a robotic device [13]. This work here is focused on the analysis of expertise and sensor specific force profiles during execution of a four-step pick-and-drop task with the telemanipulation system of STRAS. Pre-clinical testing of the STRAS robotic system has permitted to demonstrate that an expert surgeon on his own can successfully perform all the steps of a complex endoscopic surgery task (colorectal endoscopic submucosal dissection) with the telemanipulation system [14,15]. Previously [16], we had shown that proficiency (expertise) in the control of the STRAS master/slave system is reflected by a lesser grip force during task execution as well as by a shorter task execution time. In the meantime, pre-clinical testing of the STRAS robotic system has permitted to demonstrate major advantages of the system for expert endoscopic surgeons in comparison with classic procedures [14,15], and benchmark measures permitting to establish objective criteria for expertise in using the system need to be found to ensure effective training of future surgeons on the system. Experimental studies of grip force strength and control for lifting and manipulating objects strategically have provided an overview of the contributions of each finger to overall grip strength and fine grip force control [17]. While the middle finger is the most important contributor to the gross total grip force and, therefore, most important for getting a good grip of heavy objects to lift or carry, the ring finger and the small (pinky) finger are most important for the fine control of subtle grip force modulations [17], as those required for effectively manipulating the control handles of STRAS. Moreover, it is well-documented in the literature that grip force is systematically stronger in the dominant hand compared with the non-dominant hand [16,18]. In this study here, the grip force profiles correspond to measurements collected from specific sensor positions on these anatomically relevant parts of the finger and hand regions of the dominant and non-dominant hands. The grip force profiles of an expert in controlling the master/slave system are compared to those of an absolute beginner, who manipulated the robotic device for the first time. The wireless sensor glove hardware-software system described in [16], was improved and employed in this study here to collect force data from a novice trainee and an expert in various anatomical locations in the palm and on the phalanges of fingers of the right and left hands for detailed analyses in terms of sensor-specific grip force profiles.

2. Materials and Methods

2.1. Slave Robotic System

The slave robotic system is built on the Anubis® platform of Karl Storz. This system consists of three flexible, cable-driven sub-systems (for more information, [4]): One main endoscope and two lateral flexible instruments. The endoscope carries the camera providing the visual feedback at its tip, and has two lateral channels which are deviated from the main direction by two flaps at the distal extremity. The instruments have bending extremities (one direction) and can be inserted inside the channels of the endoscope. This system has a tree-like architecture and the motions of the endoscope act also upon the position and orientation of the instruments. Two kinds of instruments are available: Electrical instruments and mechanical instruments. Overall, the slave system has 10 motorized DoF. The main endoscope can be bent in two orthogonal directions. This allows moving the endoscopic view, respectively from left to right and from up to down, as well as forward/backward. Each instrument has three DoF: Translation (tz) and rotation (θz) in the endoscope channel, and deflection of the active extremity (angle β). The deflection is actuated by cables running through the instrument body from the proximal part up to the distal end. The mechanical instruments can be opened and closed.

2.2. Master/Slave Control

The slave robot is controlled at the joint level by a position loop running at 1000 Hz on a central controller. The master side consists of two specially designed interfaces, which are passive mobile mechanical systems. The user grasps two handles, each having three DoF: They can translate for controlling instrument insertion, rotate around a horizontal axis for controlling instrument rotation, and rotate around a final axis (moving with the previous DoF) for controlling instrument bending. These DoFs are similar to the possible motions of the instruments as demonstrated in preclinical trials [15]. Each handle is also equipped with a trigger and with a small four-way joystick for controlling additional DoF. In the experiments here, the trigger is operated with the index finger of a given hand for controlling grasper opening and closing, the small joysticks for moving the endoscope are not used. Since there is no force measurement on the slave side, no force effects are reproduced on the master side.

A high-level controller running on a computer under a real-time Linux OS communicates with the master interfaces and provides reference joint positions to the slave central controller. The user sits in front of the master console and looks at the endoscopic camera view displayed on the screen in front of him/her at a distance of about 80 cm while holding the two master handles, which are about 50 cm away from each other. Seat and screen heights are adjustable to optimal individual comfort. The two master interfaces are identical and the two slave instruments they control are also identical. Therefore, for a given task the same movements need to be produced by the user whatever the hand he/she uses (left or right). The master interfaces are statically balanced and all joints exhibit low friction, and therefore only minimal forces are required to produce movements in any direction. A snapshot view of a user wearing the sensor gloves while manipulating the handles of the system is shown in Figure 1a above. The master-slave control chart of the master/slave system is displayed in Figure 1b. Figure 1c shows the different directions and types of tool-tip and control movements.

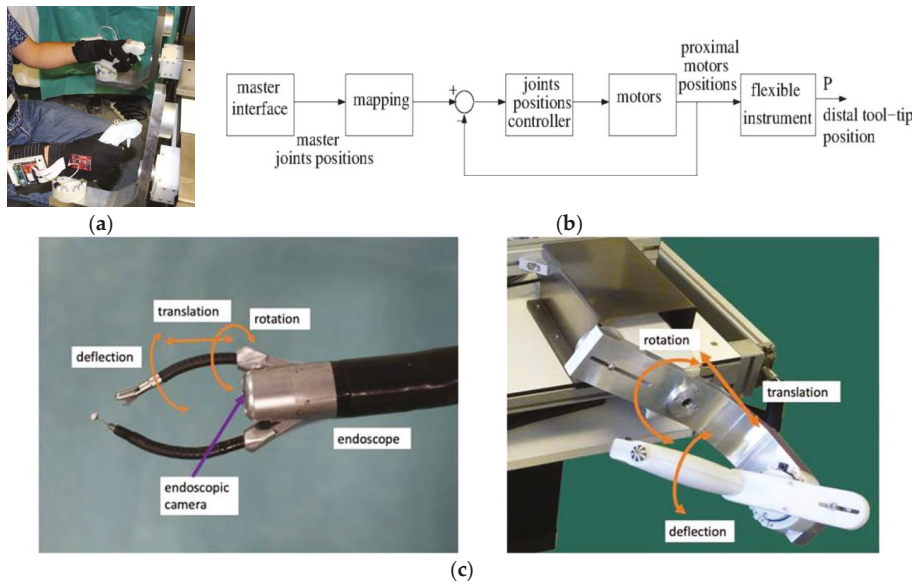


Figure 1. (a) Expert wearing the sensor gloves while manipulating the handles of the robotic master/slave system. (b) Master-slave control chart of the system. (c) Direction and type of tool-tip and control movements.

2.3. Glove Design for Grip Force Profiling

STRAS has its own grip style design and, therefore, specific gloves, one for each hand, with inbuilt force sensitive resistors (FSRs) were developed to measure the two male individuals' left and right hand grip forces applied to the two handles of STRAS for controlling and operating the master/slave system. The hardware and software configurations are described below.

2.3.1. Hardware

The gloves designed for the study contain 12 FSR, in contact with specific locations on the inner surface of the hand as given in Figure 2. Two layers of cloth were used and the FSRs were inserted between the layers. The FSRs did not interact, neither directly with the skin of the subject, nor with the master handles, which provided a comfortable feel when manipulating the system. FSRs were sewn into the glove with a needle and thread. Each FSR was sewn to the cloth around the conducting surfaces (active areas). The electrical connections of the sensors were individually routed to the dorsal side of the hand and brought to a soft ribbon cable, connected to a small and very light electrical casing, strapped onto the upper part of the forearm and equipped with an Arduino microcontroller. Eight of the FSR, positioned in the palm of the hand and on the finger tips, had a 10 mm diameter, while the remaining four located on middle phalanxes had a 5 mm diameter. Each FSR was soldered to 10 K Ω pull-down resistors to create a voltage divider, and the voltage read by the analog input of the Arduino is given by Equation (1)

$$V_{out} = R_{PD}V_{3.3}/(R_{PD} + R_{FSR}) \quad (1)$$

where R_{PD} is the resistance of the pull-down resistor, R_{FSR} is the FSR resistance, and 3.3 V is the 3.3 V supply voltage. FSR resistances can vary from 250 Ω when subject to 20 Newton (N) to more than 10 M Ω when no force is applied at all. The generated voltage varies monotonically between 0 and 3.22 V, as a function of the force applied, which is assumed uniform on the sensor surface. In the experiments here, forces applied did not exceed 10 N, and voltages varied within the range

of [0; 1500] mV. The relation between force and voltage is almost linear within this range. It was ensured that all sensors provided similar calibration curves. Thus, all following comparisons are directly between voltage levels at the millivolt scale. Regulated 3.3 V was provided to the sensors from the Arduino. Power was provided by a 4.2 V Li-Po battery enabling use of the glove system without any cable connections. The battery voltage level was controlled during the whole duration of the experiments by the Arduino and displayed continuously via the user interface. The glove system was connected to a computer for data storage via Bluetooth enabled wireless communication running 115,200 bits-per-second (bps).

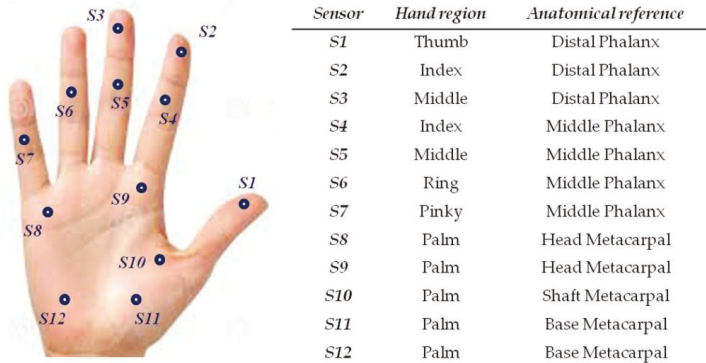


Figure 2. Sensor locations on the inner surface of the hand.

2.3.2. Software

The software of the glove system was divided into two parts: One running on the gloves, and one running on the computer algorithm for data collection. The general design of the glove system is described as follows. Each of the two gloves was sending data to the computer separately, and the software read the input values and stored them on the computer according to their header values indicating their origin. The software running on the Arduino was designed to acquire analog voltages provided by the FSR every 20 ms (50 Hz). In every loop, input voltages were merged with their time stamps and sensor identification. This data package was sent to the computer via Bluetooth, which was decoded by the computer software. The voltages were saved in a text file for each sensor, with their time stamps and identifications. Furthermore, the computer software monitored the voltage values received from the gloves via a user interface showing the battery level. In case the battery level drops below 3.7 V, the system warns the user to change or charge the battery. Such an event did not occur during the experiments reported here. Figure 3 below shows a snapshot view of the right-hand glove in action (a) and the general design chart of hard-to-software operating system (b).

2.4. Experimental Task Design

For this user grip force profiling study, a 4-step pick-and-drop task was designed. Four snapshot views of the four task steps are shown in Figure 4. During the experiments, only one of the two instruments controlling the tool-tips (left or right, depending on the task session) was moved, while the main endoscope and its image remained still. The experiments started with the right or left hand gripper being pulled back. Then, the user had to approach the object (step 1) with the distal tool extremity by manipulating the handles of the master system effectively.

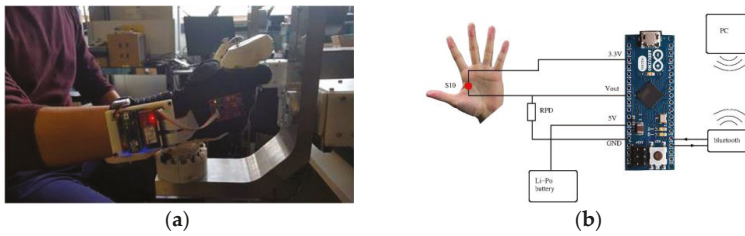


Figure 3. (a) Snapshot view of right-hand glove in action. (b) Design chart of the single-sensor-to-software operating system.

Then, the object had to be grasped with the tool (*step 2*). Once firmly held by the gripper, the object had to be moved to a position on top of the target box (*step 3*) with the distal extremity of the tool in the correct position for dropping the object into the target box without missing (*step 4*). To drop the object, the user had to open the gripper of the tool. The user started and ended a given task session by pushing a button, wirelessly connected to the computer. One expert, who had been practicing with the system since its manufacturing and who is currently the most proficient user, and one complete novice who had never used the system and had no prior experience with any similar surgical system participated in the grip force profiling experiments.

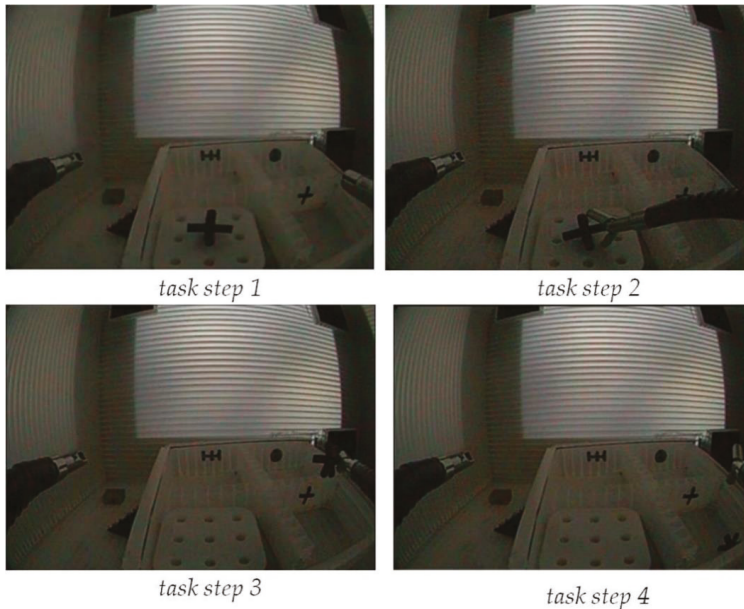


Figure 4. Snapshot views of the four successive steps of the pick-and-drop task when executed with the right hand by manipulating the corresponding instrument of the robotic system.

The expert and the novice's hand sizes were about the same, and the sensor gloves were developed specifically to fit the hands of these two individuals. The expert user was left handed and the novice user was right handed. As explained earlier, the left and right interfaces are identical, and the same task is realized with both hands. Between nine and eleven consecutive sessions were recorded for the two subjects, for task execution with their right and left hands. Before the experiment started,

the novice user was made familiar with the buttons and the running of the system. Force data were collected from all twelve sensor locations for both individuals and both hands, left and right.

3. Results

In a first preliminary analysis, the data from all sensors from left-hand and right-hand task execution were plotted as a function of the individual total number of grip force measures in millivolt sampled in time across all individual sessions. The function that links the sensor output voltage to force is almost linear within the measured range, as shown below in Figure 5.

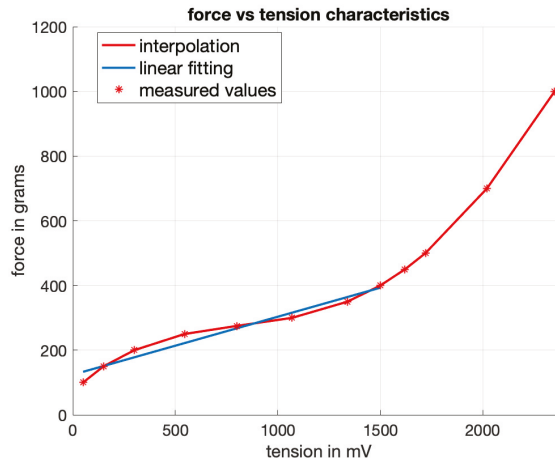


Figure 5. Static force in grams as a function of the tension output in mV of the sensors. The relation is almost linear between 50 and 1500 mV.

The total number of sessions differs slightly between the two subjects and between hand conditions. The results from these preliminary, descriptive analyses in terms of means and standard deviations of individual grip force data for task execution with the dominant or non-dominant hand, collected across sessions, are shown in Table 1. It is shown that the sensors S5, S6, S7, and S10 produced reliable and consistent output values across sessions and hand conditions for the expert and the novice allowing for a statistical comparison between their individual grip force profiles. The means and standard deviations shown here for the expert were computed on the basis of a total of 5117 grip force data for the non-dominant right hand recorded in 12 successive sessions and a total of 4442 grip force data for the dominant left hand recorded in 10 successive sessions. The statistics for the novice were computed on the basis of a total of 6497 grip force data for the non-dominant left hand recorded from 10 successive sessions, and a total of 8483 data for the dominant right hand, recorded from 11 successive task sessions.

Sensors S5, S6, S7, and S10 were positioned on regions of the fingers and the hand that are particularly critical for general grip force and/or subtle grip force control during task execution. Sensor 5 was positioned on the middle phalanx of the middle finger, critical for strong grip force control, Sensor 6 on the middle phalanx of the ring finger, contributes to force modulation less critical for strong grip control, Sensor 7 on the middle phalanx of the small finger (pinky) is highly critical for subtle, finely tuned grip force control. Experts like surgeons use their pinkies strategically for the fine-tuning of hand-tool interactions [17,19,20]. Sensor 10 was positioned on the metacarpal that joins the thumb to the wrist, important in the general grip control (getting hold of the device handles), but not critical for subtle strategic grip force control [17].

Table 1. Means and standard deviations of the expert's and the novice's individual grip force data (in millivolts (mV)) collected from each sensor during task execution in successive sessions with the dominant and non-dominant hand.

Dominant Hand													
	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	Sensor
Expert	0	1.4	4.5	2	99	452	587	0	0.5	474	0	1.2	Mean
	0	0.7	1.6	1.2	89	102	53	0	7.7	70	0	1.7	Std.dev.
Novice	0	23	674	0.7	754	498	85	651	1132	617	847	858	Mean
	0	150	207	5.5	188	74	49	192	483	312	418	280	Std.dev.
Non-dominant hand													
	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	Sensor
Expert	1	0	0	9	364	371	71	109	90	160	825	418	Mean
	1.5	0	0	27	107	68	37	118	170	138	450	250	Std.dev.
Novice	0	69	0	0	296	1063	526	233	0	500	0	0.4	Mean
	0	27	0	0	148	120	64	257	2	365	0	0.5	Std.dev.

The sensors, S1, S2, S3, and S4 were all placed on distal phalanxes, which were not needed for producing task-critical tool-movements here in this task. The index finger is minimally needed for triggering movements relative to opening and closing the grippers of the system. S8, S9, S11, and S12 produced markedly different grip force data across the two individuals in all the sessions and across hand conditions. These differences in grip-force profiles are explained by the fact that the corresponding finger or hand regions were not used in the same way, for either general grip force control or strategic grip force deployment in the manipulation of the robotic system, by the expert and the novice, who was an absolute beginner and had no experience at all with the system. These findings show promising differences that could be exploited in future studies on larger study populations, where the expert grip force profiles for strategically selected sensor locations may serve as benchmarks for assessing the skill status or evolution of novices and/or absolute beginners from different sample populations. For now, the data here will be exploited in further statistical analyses that are to highlight some of the key aspects of the differences in the grip-force profiles between an expert user and an absolute beginner. For now, data from sensor locations that produced zero-signal profiles in any (one or more) of the conditions tested were not taken into account for further statistical comparisons. As a consequence, the analyses shown and discussed here below are selective to sensor locations S5, S6, S7, and S10. They are to serve as examples, and to provide guidance for further benchmarking in future studies on larger sample populations.

3.1. Dynamic Grip Force Range in the Expert and Novice Data at Selected Sensor Locations

To gain a descriptive overview of the dynamic range (upper and lower limits) of the grip force distributions from each of the four relevant sensor positions S5, S6, S7, and S10 during task execution with the dominant and non-dominant hands across sessions, these sensor data were represented in terms of box-plots for each subject (expert and novice) and for each of their hands (dominant and non-dominant). These box-plots are shown below in Figure 6.

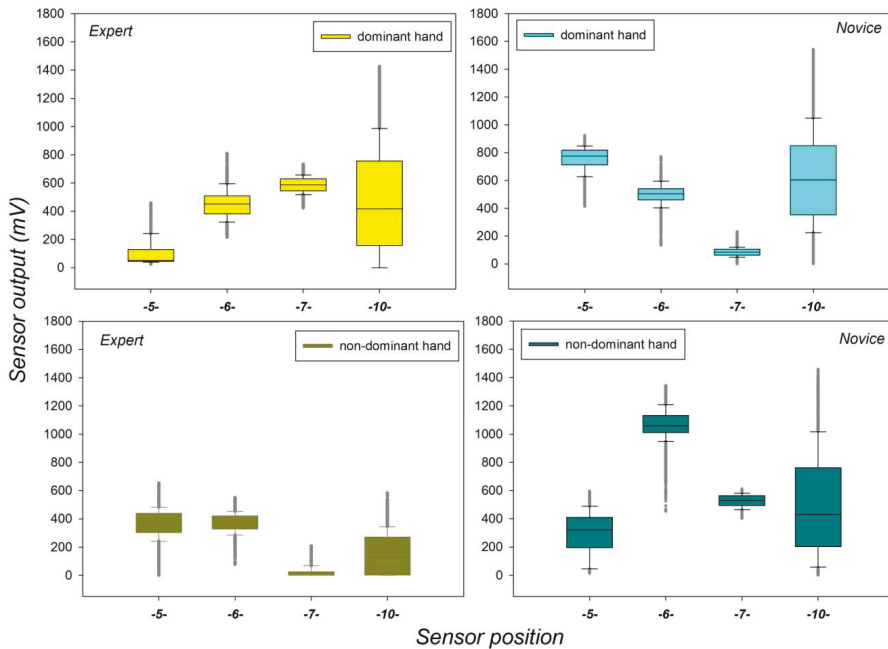


Figure 6. Dynamic range of grip force data recorded for the expert and the novice from each task-relevant sensor position in successive sessions with the dominant and non-dominant hands.

The box-plots reveal marked differences in upper and lower limits of the grip force distributions from the different sensors. The output range is found to vary as a function of expertise, sensor position, and handedness. The largest amount of grip force variability is found in Sensor 10, positioned on the metacarpal that joins the thumb to the wrist, important in the general grip control for getting a good hold of the device handles, but not critical for subtle strategic grip force control for finely tuned task maneuvers [17]. The smallest amount of grip force variability is found in sensor S7, positioned on the middle phalanx of the small finger (pinky) and highly critical for subtle grip force control during task maneuvers. The most noticeable differences between the grip force distributions of the expert and the novice are observed in sensor positions S7 and S5 on the middle finger, critical for strong grip force modulation and control. The total amount of grip force deployed on a given sensor is shown to depend on expertise, and on whether the dominant or the non-dominant was used to manipulate the robotic system. These observations suggest complex interactions between factors warranting a series of detailed analysis of variance (ANOVA). The design and outcome of these statistical analyses are described in the following subsections.

3.2. Individual Sensor-Specific Effects of Task Session and Handedness with Interactions

In the next step, we analyzed the individual grip force profiles of the expert and the novice for task execution with each hand, the dominant and the non-dominant one, across the successive individual sessions for each of the four relevant sensors S5, S6, S7, and S10. The data from this analysis are shown in Figure 7.

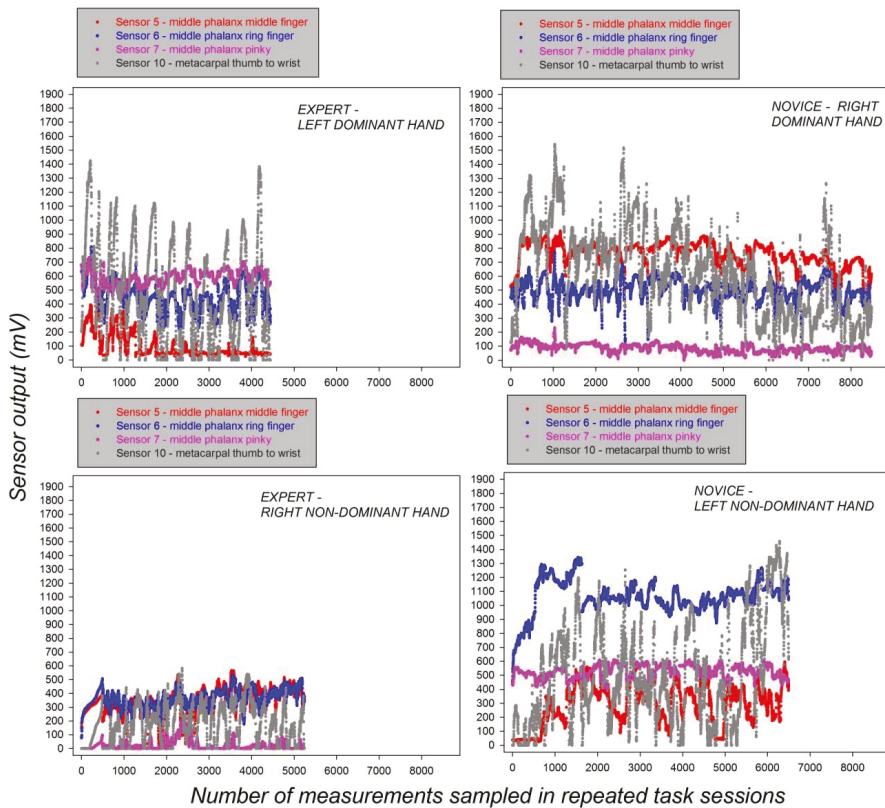


Figure 7. Grip force profiles of the expert (left) and the novice (right) from the relevant sensors for task execution with the dominant and the non-dominant hand across successive individual sessions.

The data plotted in Figure 7 reveal distinct grip force profiles of the novice and the expert, depending on whether they use their dominant or non-dominant hands. The grip force profiles of the novice systematically display stronger grip forces in the non-dominant hand compared with those of the expert's non-dominant hand irrespective of sensor position. A marked dependency of individual grip forces on sensor position is seen when comparing the grip force profiles of the two subjects using their dominant hands. While the expert's dominant hand systematically displays noticeably stronger grip forces compared with the novice in the profiles for Sensor 7 positioned on the pinky finger, exactly the reverse is observed for sensor position S5 on the middle finger, where the expert's dominant hand displays minimal grip force strength, while the profile of the novice displays forces up to eight times the strength of those of the expert, especially towards the end of the successive task sessions.

To assess the statistical significance of these effects, analysis of variance on the raw data were performed. In each of the four relevant sensors S5, S6, S7, and S10, and in each subject, the expert and the novice, we tested for statistically significant effects of task repetition (training), reflected by the factor task session, and of handedness on the individual grip force data. To this effect, the individual raw data of each subject were submitted to several two-way analysis of variance using the general linear model for $Subject_1 \times Sensor_1 \times Hand_2 \times Session_{10}$, with one level of the 'subject' factor, two levels of the *Hand* factor ('dominant' versus 'non-dominant'), and ten levels of the *Session* factor. Given that we have a total of 10 sessions for the dominant and a total of 11 sessions for the non-dominant hand in the case of the expert, and a total of 11 sessions for the dominant hand, and a total of 10 sessions

for the non-dominant hand in the case of the novice, the analysis of variance were performed on the first ten successive sessions for each hand in each subject. This Cartesian analysis plan enabled the computation of interactions between the *Session* and *Hand* factors for each subject and sensor. Grip force data in terms of means (M_{1-10}) and their standard errors (SEM) from sensor-specific individual two-way analysis of variance for ten successive individual task sessions are summarized in Table S1 of the Supplementary Materials.

The F-statistics relative to the effects of the *Hand* and *Session* factors and their interactions, with their corresponding probability limits (p), are shown in Table 2 above. Since the effect sizes in terms of differences between means are difficult to grasp from looking at the tables, we represented the average results graphically in Figure 8 below, which also permits highlighting interactions visually.

Table 2. F-statistics and probability limits relative to the effects of the *Hand* and *Session* factors and their interactions from the two-way analysis of variance for each subject and sensor.

Expert				
	S5	S6	S7	S10
<i>Hand</i>	F (1,9685) = 3017; $p < 0.001$	F (1,9685) = 3734; $p < 0.001$	F (1,9685) = 5745; $p < 0.001$	F (1,9685) = 3339; $p < 0.001$
<i>Session</i>	F (9,9685) = 213; $p < 0.001$	F (9,9685) = 209; $p < 0.001$	F (9,9685) = 307; $p < 0.001$	F (9,9685) = 47; $p < 0.001$
<i>Hand</i> × <i>Session</i>	F (9,9685) = 747; $p < 0.001$	F (9,9685) = 277; $p < 0.001$	F (9,9685) = 295; $p < 0.001$	F (9,9685) = 138; $p < 0.001$
Novice				
	S5	S6	S7	S10
<i>Hand</i>	F (1,13613) = 8655; $p < 0.001$	F (1,13613) = 1499; $p < 0.001$	F (1,13613) = 6755; $p < 0.001$	F (1,13613) = 200; $p < 0.001$
<i>Session</i>	F (9,13613) = 413; $p < 0.001$	F (9,13613) = 102; $p < 0.001$	F (9,13613) = 188; $p < 0.001$	F (9,13613) = 201; $p < 0.001$
<i>Hand</i> × <i>Session</i>	F (9,13613) = 677; $p < 0.001$	F (9,13613) = 152; $p < 0.001$	F (9,13613) = 359; $p < 0.001$	F (9,13613) = 1073; $p < 0.001$

The effect sizes from the two-way analysis of variance, shown graphically in Figure 8 and reflected by differences in average grip forces, reveal significant trends towards a decrease in the dominant hand as the sessions progress, and towards an increase in the non-dominant hand as the sessions progress, especially for Sensor 10 on the metacarpal between the thumb and wrist of the novice, presumably as a result of task fatigue: The novice presses harder with the wrist to compensate for lack of fine grip force control. The statistically significant effect of handedness expresses itself differently in the two subjects; while the expert mostly deploys stronger average grip force with his dominant hand, at most sensor positions, the novice deploys stronger average grip force with his non-dominant hand at most sensor positions. While the force profiles from Sensor 7 of the expert's dominant hand indicate proficient use of the pinky finger for subtle grip force control with the dominant hand, the weak average grip forces from Sensor 7 in his non-dominant hand, resembling those of the novice's dominant hand, reflect the expert's lack of fine grip force control in the non-dominant hand: The expert is, indeed, not trained in performing the task with his non-dominant hand. Expertise, or proficiency, is reflected in the individual grip force profiles by a strategic and consistently parsimonious deployment of grip forces in the hand that has been trained for controlling the robotic device. The novice, on the other hand, deploys forces non-strategically, in the dominant and the non-dominant hand. Significant interactions between the *Hand* factor and the *Session* factor' are found in each of the two subjects, and in each of the four sensors. This leads to conclude on a complex joint influence of handedness and repeated training on an individual's grip force profiles. Interactions between handedness and repeated training are, as would be expected, partly independent of an individual's proficiency level, i.e., whether he is an

expert at performing the given task or not, as even expert performance is subject to variations and likely to evolve.

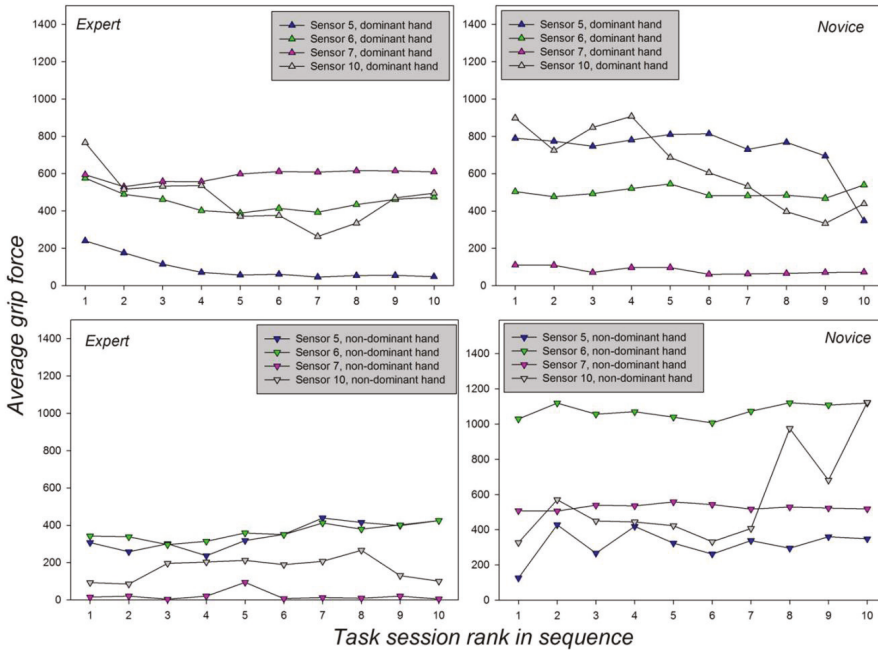


Figure 8. Average grip forces, reflected by sensor output in mV, are plotted as a function of the session and hand factor for the expert (left) and the novice (right) performing the task with their dominant (top) and non-dominant (bottom) hands.

3.3. Effects of Expertise as a Function of Sensor Position

The grip force profiles from the hand of the expert trained in the specific task given, which is generally the dominant hand, allow to bring to the fore specific characteristics of expert performance compared with that of an absolute beginner, i.e., a novice performing the same task under the same conditions for the first time. The next step of the analysis is aimed at further highlighting critical strategy relevant differences in the grip force profiles, comparing the expert performing the task with his dominant hand to the novice performing the task with his dominant hand. To test for significant differences between subjects (proficiency level) and sensors, and for significant interactions between proficiency level and sensor location, the raw force data of the expert and the absolute novice performing the task with their dominant hands only, in all sessions, were considered. The data from both subjects, all four sensors, and all sessions, were submitted to a single two-way analysis of variance using the general linear model for the analysis plan $Subject_2 \times Sensor_4$, with two levels of the *Subject* factor (*expertise* factor) and four levels of the *Sensor* factor. The F-statistics relative to the effects of the *Subject* and *Sensor* factors and their interactions, with their corresponding probability limits (*p*), are shown in Table 3. Statistics from the post-hoc comparisons for effects of expertise within each sensor are shown in Table 4.

Table 3. F-statistics from the two-way analysis of variance for effects of expertise and sensor.

Subject (expertise)	F (1,51692) = 2680; $p < 0.001$
Sensor	F (3,51692) = 2840; $p < 0.001$
Subject× Sensor	F (3,51692) = 2569; $p < 0.001$

Table 4. Statistics and probability limits from the Holm-Sidak post-hoc comparisons for effects of expertise within each sensor.

	Diff of Means	t	Unadjusted P	Critical Level
Expertversus Novice	655.653	198.175	<0.001	0.05
Sensor 6	Diff of Means	t	Unadjusted P	Critical Level
Expertversus Novice	46.167	13.954	<0.001	0.05
Sensor 7	Diff of Means	t	Unadjusted P	Critical Level
Expertversus Novice	502.306	151.825	<0.001	0.05
Sensor 10	Diff of Means	t	Unadjusted P	Critical Level
Expertversus Novice	14.069	43.24	< 0.001	0.05

The results from this analysis show a statistically significant effect of the *Subject (expertise)* factor, a statistically significant effect of the *Sensor* factor, and a statistically significant interaction between these two factors. This leads to conclude that the expert and the absolute novice do not use the different anatomical locations on which the sensors were placed in the same way, but employ significantly different grip force control strategies, reflected by the statistically significant differences in the sensor output data. These effects and their interaction are shown graphically in Figure 9, where individual force data of the expert and the absolute novice were plotted for each sensor and individual task session in time.

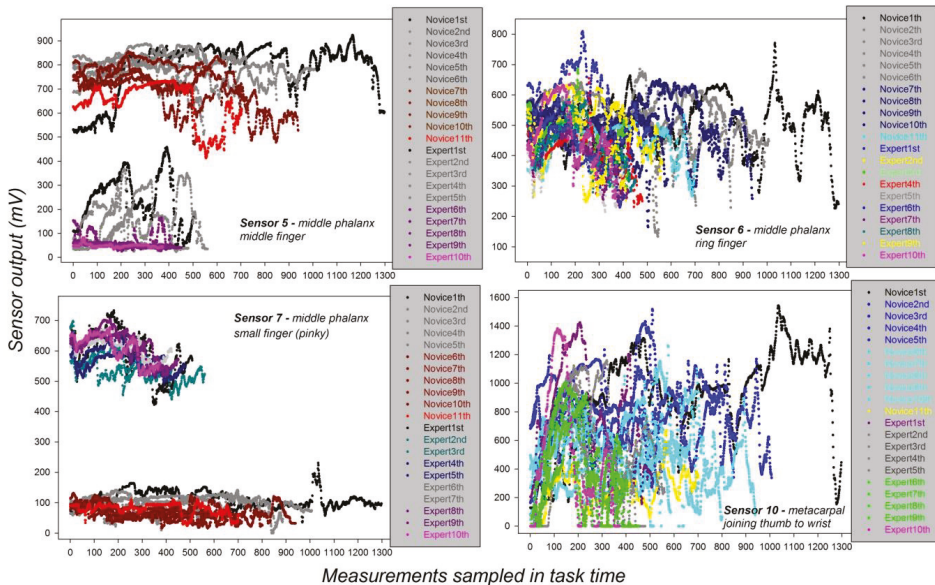


Figure 9. Distinct grip force profiles, plotted separately for each of the four specific sensor locations, from the successive sessions of the expert and the novice performing the *pick-and-drop* task on the robotic system with their dominant hands.

The graphs in Figure 9 below display distinct grip force profile telling apart the control strategies of the expert and the novice. The middle finger, preferentially used for controlling subtle grip force modulations in surgery and other manual tasks, is considerably solicited by the novice in all sessions, significantly less by the expert, as clearly shown here on the basis of the individual grip force profiles for sensor location 5, on the middle phalanx of the middle finger (top left graph). The small finger (pinky), strategically used by surgeons and other experts in various manual tasks for strong grip force control, is consistently solicited by the expert and significantly less by the novice, who almost applied no force at all to that finger region, as shown here on the basis of the individual grip force profiles for sensor location 7, on the middle phalanx of the pinky finger (bottom left graph). The ring finger contributes largely to grip force control, much less to total force magnitude than the middle and small fingers. The individual grip force profiles for the corresponding sensor location 8 show that the expert applies force to that finger region in a different manner compared with the novice (top right graph). Although the differences between the two subjects here may appear small in the graph, they are systematic and statistically significant in the light of 2×2 post-hoc comparisons shown in Table 4 above. The post-hoc tests were performed using the Holm-Sidak method. Similar conclusions are drawn from the individual grip force profiles for the sensor location 10, on the metacarpal joining the thumb to the wrist. This hand region is exploited for maintaining constant grip force on the handles of the robotic device, and the fully proficient expert needs to apply less force to achieve that goal than a novice, as consistently shown here.

The next analysis was aimed at further highlighting subtle ways in which the interaction between task proficiency (expertise) and sensor location expresses itself at the beginning and at the end of the repeated task sessions. To that effect, the individual force data of the expert and the novice were plotted, for each sensor, showing recordings from the first half of the first sessions, and the last half of the last sessions only, for a comparison. The results of this analysis are shown in Figure 10 below.

The results displayed in Figure 10, show that the novice deploys excessive grip force on sensor 5 (Figure 10, top) of the middle finger, which plays an important role in generating total grip force magnitude for lifting a heavy object with the hand, but is not useful to subtle grip force control of the robotic system. At the end of the last session, the novice's grip force is still about five times that of the expert, who deploys only a minimal amount of grip force on Sensor 5. The forces deployed by the expert on Sensor 6 positioned on the ring finger also evolve differently from the beginning to the end of the sessions compared with those deployed by the novice on that sensor (Figure 10, upper middle). The ring finger contributes largely to grip force control, much less to total force magnitude. The expert deploys stronger grip force on this specific sensor location at the beginning of the sessions, when adjusting the tool and getting in to the swing of the control process, considerably less at the end of the sessions. The novice, on the other hand, does not use the ring finger in a noticeably differentiated manner: The grip forces deployed on Sensor 6 at the beginning of the sessions are about the same as towards the end of the last session. The largest difference in grip force strategy between the expert and the novice is reflected by the grip force profiles of Sensor 7, positioned on the small (pinky) finger of the dominant hand (Figure 10, lower middle). While the expert deploys grip force consistently on that sensor across all sessions from the beginning to the end, within a moderately narrow range of variability for subtle grip force control while steering the robotic handles, the novice hardly deploys any grip force at all on that sensor throughout all sessions. Towards the end, he still shows no signs of expertise in subtle grip force control through strategically deployed force variations in the small finger.

Finally, the grip force profiles that produced the least distinctive characteristics between the expert and the novice are those from Sensor 10, positioned on the base metacarpal between thumb and wrist (Figure 10, bottom), a hand region that plays no major role in subtle grip force control, but is important for pushing heavy objects.

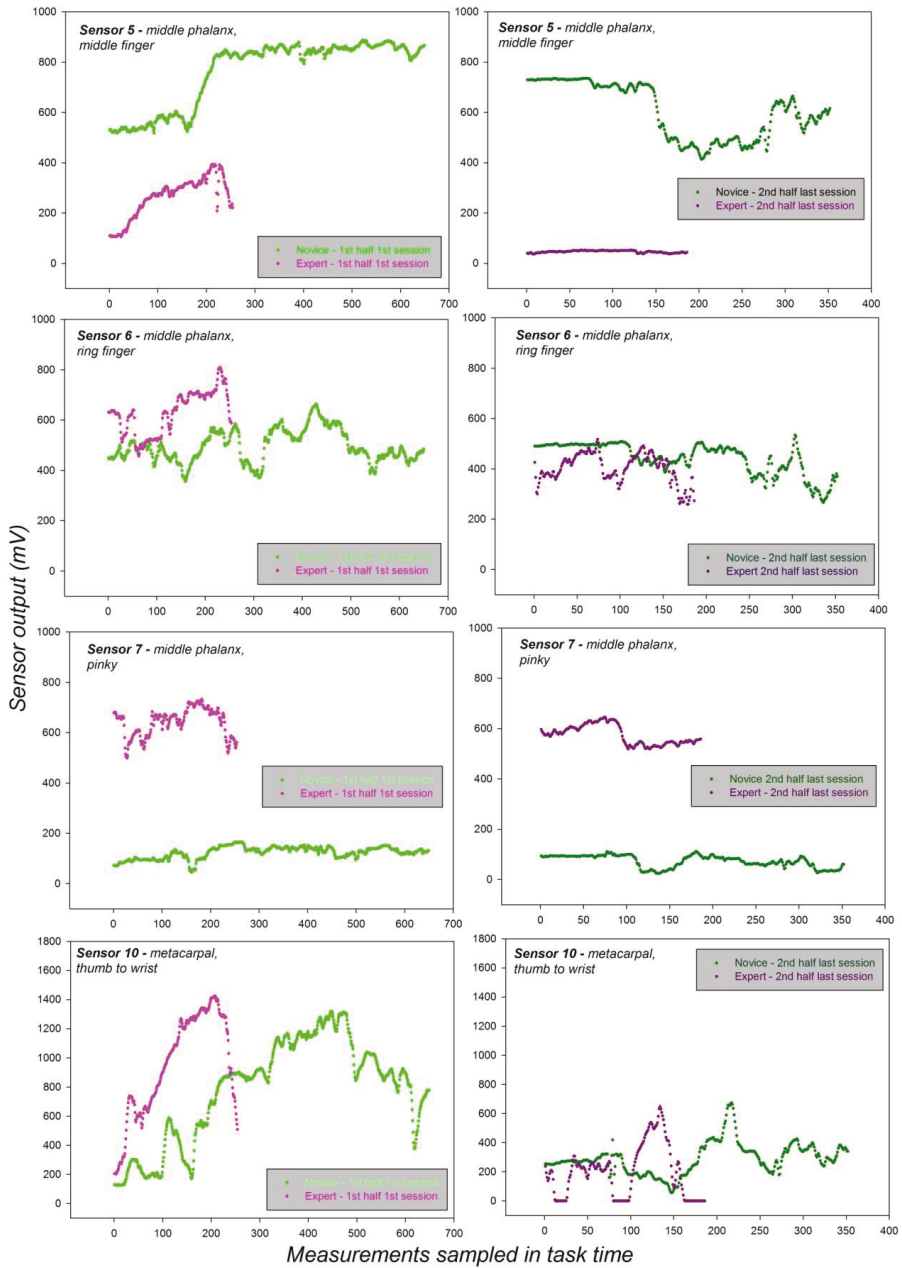


Figure 10. Grip force profiles for each of the four specific sensor locations from the first half of the first task sessions, and from the last half of the last task sessions of the expert and the novice performing the task on the robotic system with their dominant hands.

3.4. Task Times and Quantitative/Qualitative Analysis of the Task Videos

In an additional analysis, the task times for each individual session of the expert and the absolute novice were compared and submitted to analysis of variance. The video sequences captured by the endoscopic camera attached at the distal side of the robotic endoscope and filming the expert and the novice performing the *four step pick-and-drop task* with their dominant hands across the full sequence of individual sessions were analyzed subsequently. The times taken by the expert and the novice in each session to accomplish the robotic system task with the dominant hand are shown in Figure 11.

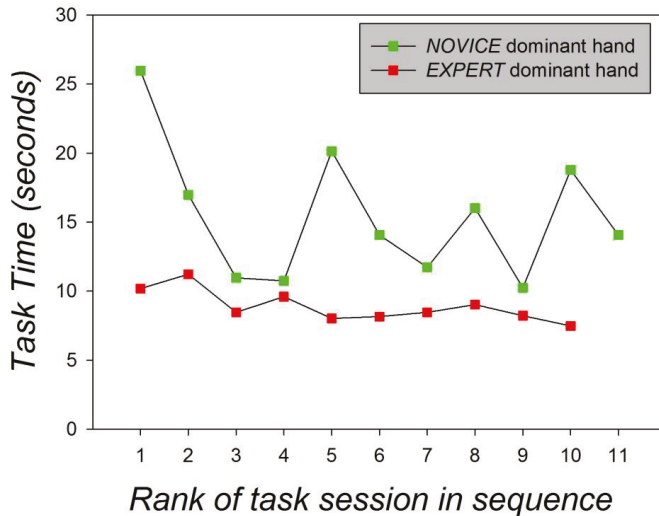


Figure 11. Task times from the sessions of the expert and the novice performing the task on the robotic system with their dominant hands.

As shown in Figure 11, the task times of the expert are systematically, often considerably, shorter than the task time of the novice and display very little variance, which is characteristic of highly proficient operators in general. The expert's task times vary between 7 and 11 s, while those of the novice vary markedly between 7 and 27 s. The expert's task time in the last session is at the minimum of 7 s, while the novice takes twice as long (14 s) in the last task session with his dominant hand. The difference in task times between the expert was found to be statistically significant on the grounds of statistical paired comparison (Student's *t*-test). The results are displayed below in Table 5.

Table 5. Paired comparison statistics (Student's *t*) for the task times of the expert and the novice from ten and eleven successive task sessions, respectively.

Group Name	N	Missing	Mean	Stand. Dev.	SEM
Time Novice	11	0	15.424	4.832	1.457
Time Expert	10	0	8882	1.141	0.361

Difference between means: 6.542
t = 4.168 with 19 degrees of freedom; $p \leq 0.001$.
95 percent confidence interval for difference of means: 3.257 to 9.827

In a final analysis, quantitative and qualitative criteria for task precision distinguishing the dominant-hand task performance of the expert from that of the novice in the four-step pick-and-drop task were determined. This was possible on the basis of analyses of the video data relative to the individual task sequences filmed by the endoscopic camera of the robotic system. Copies of the original

videos from which these analyses were drawn are made available in the Supplementary Materials Videos S1 and S2. The results of these analyses are shown below in Table 6.

Table 6. Quantitative performance analysis relative to task precision of the expert and the novice.

	Expert	Novice
Tool trajectories towards target requiring adjustment	3	20
Number of times object was accidentally released	0	1
Unsuccessful attempts to grasp object	1	8
Tool collisions with task-space boundary	0	2
Number of times object was dropped outside target area	0	1

The results from Table 6 show clear quantitative differences in task precision between the expert and the novice, who has to adjust the tool trajectories many more times in all sessions, and scores a much higher number of unsuccessful grasp attempts compared with the expert, who delivers a close to optimal precision performance. In addition to these quantitative differences, it was noted that the expert adjusted tool movements only very slightly at the beginning of a trajectory. This happened no more than three times in the entire sequence of all eleven task sessions. The novice adjusted tool movements mostly at the end of trajectories, often by multiple adjustments, which happened 20 times in the sequence of his ten task sessions with the dominant hand.

4. Discussion

Sensor specific grip force profiles of an expert and an absolute beginner allow for a precise, quantitative, and qualitative characterization of expert control strategies in dominant-hand manipulation of the robotic device handles. Expert force control is characterized here by a marked differential middle and small finger grip force strategy, subtle force modulation using the ring finger, and parsimonious grip force deployment on the base metacarpal that joins the thumb to the wrist. The novice uses inappropriate grip forces in these strategic finger-hand regions, producing either excessive or insufficient force at the critical sensor locations. The grip force profiles are consistent with anatomical data relative to the evolution of strategic use of functionally specific finger and hand regions for grip and grip force control in manually executed precision tasks [17–20]. Preconceptions that grip forces may be universally stronger in the dominant hand [18] require a reconsideration in the light of the grip force profiles from this study, which show that the level of expertise and task-device specific factors determine the total amount of grip force deployed by either the dominant or the non-dominant hand. The robotic system exploited here is designed without force feedback, and the maneuver control is therefore based solely on visual feedback from the 2D screen images generated by the endoscopic fisheye camera. Different camera systems produce image solutions of varying quality, and image quality affects image-guided performance [21–23]. It will be useful in the future to exploit individual grip force profiling in the context of comparisons between different camera systems for robot-assisted surgery systems. Finally, the sensor glove system developed for this study will need to be improved for further investigation on larger sample populations. This will allow us to refine expertise benchmarks, and to produce increasingly objective performance criteria for monitoring surgical skill evolution during the training of surgeons on STRAS. The new glove device will ensure that sensor positions can be adapted with high precision to the left and right hands of any group of surgeons or surgical trainees, including women surgeons, whose hands are statistically smaller, as discussed previously in systematic anthropometric studies [20].

Supplementary Materials: The following are available online at <http://www.mdpi.com/1424-8220/19/20/4575/s1>, **Table S1:** Grip force data in terms of means (M_{1-10}) and their standard errors (SEM) from sensor-specific individual two-way analysis of variance for ten successive individual task sessions. **Videos S1 and S2:** Video sequences captured by the endoscopic camera filming the expert and the novice performing the *four step pick-and-drop task* with their dominant hands.

Author Contributions: Conceptualization, B.D.-L., F.N., and M.d.M.; Methodology, B.D.-L. and F.N.; Software, P.Z.; Validation, F.N., P.Z., and B.D.-L.; Formal analysis, B.D.-L. and M.d.M.; Investigation, F.N. and B.D.-L.; Resources, F.N. and P.Z.; Data curation, F.N.; Writing—original draft preparation, B.D.-L., F.N., P.Z., and M.d.M.; Writing—review and editing, B.D.-L., F.N., and M.d.M.; Visualization, F.N. and P.Z.; Internal funding acquisition, M.d.M., B.D.-L., and F.N.

Funding: This research received no external funding.

Acknowledgments: L. Zorn participated in the design of STRAS; A.U. Batmaz, and M.A. Falek provided technical assistance with the sensor glove design, and helped in collecting data at an earlier stage of this project. Their respective contributions are gratefully acknowledged.

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

References

1. Thompson, C.C.; Ryou, M.; Soper, N.J.; Hungess, E.S.; Rothstein, R.I.; Swanstrom, L.L. Evaluation of a manually driven, multitasking platform for complex endoluminal and natural orifice transluminal endoscopic surgery applications (with video). *Gastrointest. Endosc.* **2009**, *70*, 121–125. [[CrossRef](#)] [[PubMed](#)]
2. Swanstrom, L.L.; Kozarek, R.; Pasricha, P.J.; Gross, S.; Birkett, D.; Park, P.O.; Saadat, V.; Ewers, R.; Swain, P. Development of a new access device for transgastric surgery. *J Gastrointest. Surg.* **2005**, *9*, 1129–1137. [[CrossRef](#)] [[PubMed](#)]
3. Phee, S.J.; Low, S.C.; Huynh, V.A.; Kencana, A.P.; Sun, Z.L.; Yang, K. Master and slave transluminal endoscopic robot (MASTER) for natural orifice transluminal endoscopic surgery (NOTES). In Proceedings of the 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society: Engineering the Future of Biomedicine, Minneapolis, MN, USA, 3–6 September 2009; pp. 1192–1195.
4. Zorn, L.; Nageotte, F.; Zanne, P.; Legner, A.; Dallemagne, B.; Marescaux, J.; de Mathelin, M. A novel telemanipulated robotic assistant for surgical endoscopy: Preclinical application to ESD. *IEEE Trans. Biomed. Eng.* **2018**, *65*, 797–808. [[CrossRef](#)] [[PubMed](#)]
5. Dallemagne, B.; Marescaux, J. The ANUBISTM Project. *Minim. Invasive Ther. Allied Technol.* **2010**, *19*, 257–261. [[CrossRef](#)] [[PubMed](#)]
6. De Donno, A.; Nageotte, F.; Zanne, P.; Zorn, L.; de Mathelin, M. Master/slave control of flexible instruments for minimally invasive surgery. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Tokyo, Japan, 3–7 November 2013; pp. 483–489.
7. Chitwood, W.R.; Nifong, L.W.; Chapman, W.H.; Felger, J.E.; Bailey, B.M.; Ballint, T.; Mendelson, K.G.; Kim, V.B.; Young, J.; Albrecht, R. Robotic surgical training in an academic institution. *Ann. Surg.* **2001**, *234*, 475–484. [[CrossRef](#)] [[PubMed](#)]
8. King, C.H.; Culjat, M.O.; Franco, M.L.; Lewis, C.E.; Dutson, E.P.; Grundfest, W.S.; Bissley, J.W. Tactile feedback induces reduced grasping force in robot-assisted surgery. *IEEE Trans. Haptics* **2009**, *2*, 103–110. [[CrossRef](#)] [[PubMed](#)]
9. Famè, A.; Lådavas, E. Dynamic size-change of hand peripersonal space following tool use. *Neuroreport* **2000**, *11*, 1645–1649. [[CrossRef](#)] [[PubMed](#)]
10. Maravita, A.; Husain, M.; Clarke, K.; Driver, J. Reaching with a tool extends visual–tactile interactions into far space: Evidence from cross-modal extinction. *Neuropsychologia* **2001**, *39*, 580–585. [[CrossRef](#)]
11. Wilson, M.R.; McGrath, J.S.; Vine, S.J.; Brewer, J.; Defriend, D.; Masters, R.S.W. Perceptual impairment and psychomotor control in virtual laparoscopic surgery. *Surg. Endosc.* **2011**, *25*, 2268–2274. [[CrossRef](#)] [[PubMed](#)]
12. Sarlegna, F.; Blouin, J.; Bresciani, J.P.; Bourdin, C.; Vercher, J.L.; Gauthier, G.M. Target and hand position information in the online control of goal-directed arm movements. *Exp. Brain Res.* **2003**, *151*, 524–535. [[CrossRef](#)] [[PubMed](#)]
13. Pereira, R.; Moreira, A.H.J.; Leite, M.; Rodrigues, P.L.; Queirós, S.; Rodrigues, N.F.; Leão, P.; Vilaça, J.L. Hand-held robotic device for laparoscopic surgery and training. In Proceedings of the IEEE International Conference on Serious Games and Applications for Health, Rio de Janeiro, Brazil, 14–16 May 2014; pp. 1–8.
14. Legner, A.; Diana, M.; Halvax, P.; Liu, Y.; Zorn, L.; Zanne, P.; Nageotte, F.; de Mathelin, M.; Dallemagne, B.; Marescaux, J. Endoluminal surgical triangulation 2.0: A new flexible surgical robot. Preliminary pre-clinical results with colonic submucosal dissection. *Int. J. Med. Robot. Comput. Assist. Surg.* **2017**, *13*, e1819. [[CrossRef](#)] [[PubMed](#)]

15. Mascagni, P.; Lim, S.G.; Fiorillo, C.; Zanne, P.; Nageotte, F.; Zorn, L.; Peretta, S.; de Mathelin, M.; Marescaux, J.; Dallemagne, B. Democratizing endoscopic submucosal dissection: Single-operator fully robotic colorectal ESD in a pig model. *Gastroenterology* **2019**, *156*, 1559–1571. [[CrossRef](#)] [[PubMed](#)]
16. Batmaz, A.U.; Falek, A.M.; Zorn, L.; Nageotte, F.; Zanne, P.; de Mathelin, M.; Dresch-Langley, B. Novice and expert behavior while using a robot controlled surgery system. In Proceedings of the 2017 13th IASTED International Conference on Biomedical Engineering (BioMed), Innsbruck, Austria, 20–21 February 2017; pp. 94–99.
17. Cha, S.M.; Shin, H.D.; Kim, K.C.; Park, J.W. Comparison of grip strength among six grip methods. *J. Hand Surg.* **2014**, *39*, 2277–2284. [[CrossRef](#)] [[PubMed](#)]
18. Bohannon, R.W. Grip strength: A summary of studies comparing dominant and non-dominant limb measurements. *Percept. Mot. Skills* **2003**, *96*, 728–730. [[CrossRef](#)] [[PubMed](#)]
19. Young, R.W. Evolution of the human hand: The role of throwing and clubbing. *J. Anat.* **2003**, *202*, 165–174. [[CrossRef](#)]
20. González, A.G.; Rodríguez, D.R.; Sanz-Calcedo, J.G. Ergonomic analysis of the dimension of a precision tool handle: A case study. *Procedia Manuf.* **2017**, *13*, 1336–1343. [[CrossRef](#)]
21. Batmaz, A.U.; de Mathelin, M.; Dresch-Langley, B. Seeing virtual while acting real: Visual display and strategy effects on the time and precision of eye-hand coordination. *PLoS ONE* **2017**, *12*, e0183789. [[CrossRef](#)] [[PubMed](#)]
22. Dresch-Langley, B. Principles of perceptual grouping: Implications for image-guided surgery. *Front. Psychol.* **2015**, *6*, 1565. [[CrossRef](#)] [[PubMed](#)]
23. Wandeto, J.M.; Dresch-Langley, B. The quantization error in a Self-Organizing Map as a contrast and color specific indicator of single-pixel change in large random patterns. *Neural Netw.* **2019**, *119*, 273–285. [[CrossRef](#)] [[PubMed](#)]



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).

Article

SeisMote: A Multi-Sensor Wireless Platform for Cardiovascular Monitoring in Laboratory, Daily Life, and Telemedicine

Marco Di Rienzo *, Giovannibattista Rizzo, Zeynep Melike Işilay and Prospero Lombardi

IRCCS Fondazione Don Carlo Gnocchi, 20148 Milano, Italy; frizzo@dongnocchi.it (G.R.); zisilay@dongnocchi.it (Z.M.I.); plombardi@dongnocchi.it (P.L.)

* Correspondence: mdirienzo@dongnocchi.it; Tel.: +39-02-40308-541

Received: 20 November 2019; Accepted: 24 January 2020; Published: 26 January 2020

Abstract: This article presents a new wearable platform, SeisMote, for the monitoring of cardiovascular function in controlled conditions and daily life. It consists of a wireless network of sensorized nodes providing simultaneous multiple measures of electrocardiogram (ECG), acceleration, rotational velocity, and photoplethysmogram (PPG) from different body areas. A custom low-power transmission protocol was developed to allow the concomitant real-time monitoring of 32 signals (16 bit @200 Hz) from up to 12 nodes with a jitter in the among-node time synchronization lower than 0.2 ms. The BluetoothLE protocol may be used when only a single node is needed. Data can also be collected in the off-line mode. Seismocardiogram and pulse transit times can be derived from the collected data to obtain additional information on cardiac mechanics and vascular characteristics. The employment of the system in the field showed recordings without data gaps caused by transmission errors, and the duration of each battery charge exceeded 16 h. The system is currently used to investigate strategies of hemodynamic regulation in different vascular districts (through a multisite assessment of ECG and PPG) and to study the propagation of precordial vibrations along the thorax. The single-node version is presently exploited to monitor cardiac patients during telerehabilitation.

Keywords: body sensor network; wearable sensor; telemedicine; telerehabilitation; seismocardiogram; acceleration; electrocardiogram; cardiac mechanics; photoplethysmogram; pulse transit time

1. Introduction

Over the years, there has been a growing demand for wearable systems able to monitor the cardiovascular function out of laboratory settings in ambulant subjects. The electrocardiogram (ECG) was the first signal to be monitored by this class of devices since the early 1960s (ECG Holter monitors) [1]. More recently, additional signals have also been considered for the evaluation of cardiac function in daily life. One of them is the seismocardiogram (SCG); this is the measure of minute thorax accelerations produced by the beating heart and can be simply detected by placing an accelerometer on the chest surface [2]. Usually, only the dorso-ventral component of the acceleration (corresponding to the z-axis of our sensor) is considered for SCG measurement. From the analysis of this signal, it is possible to obtain information on different mechanical events of the cardiac cycle including opening and closing of the aortic and mitral valves, atrial systole, and isovolumic contraction and relaxation [3–6].

Traditionally, cardiac mechanics are evaluated by ultrasound (US) techniques. This methodology offers a detailed investigation of heart performance, but it cannot be exploited for obtaining measurements outside of laboratory settings, and it cannot be used for studying the dynamic features of cardiac mechanics over time. This is because it provides snapshot measurements and also because of the considerable size of the device and the complexity of the assessments requiring expert operators. Conversely, monitoring by SCG allows repeated estimations of mechanical cardiac indexes either in

controlled conditions and during outdoor living or at home through wearable devices that can be easily self-managed [7]. Because of all these features, the use of an SCG signal opens new opportunities for the investigation of cardiac mechanics in research and clinics [8–11].

The typical SCG waveform is illustrated in the middle panel of Figure 1. Further details on this signal and its derived parameters may be found in Reference [7].

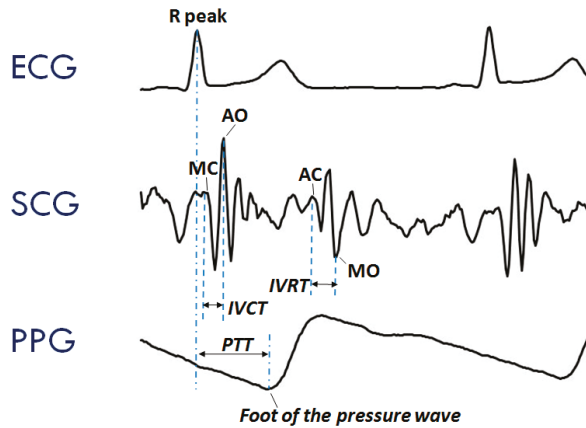


Figure 1. Typical waveforms of the seismocardiogram (SCG) and photoplethysmogram (PPG) as compared with the electrocardiogram (ECG). (**Upper Panel**) ECG signal with indication of the *R peak* fiducial point used for the estimation of the pulse transit time (PTT). (**Mid Panel**) SCG signal with indication of the fiducial points associated with the Opening and Closing of the Aortic and Mitral valves (AO, AC, MO, and MC) to be considered for the estimation of the isovolumic contraction time (IVCT) and isovolumic relaxation time (IVRT), two clinical indexes of cardiac contractility and relaxation. (**Lower Panel**) PPG signal with indication of the timing for the PTT estimate.

Another signal frequently considered in the assessment of cardiovascular performance is the photoplethysmogram (PPG), namely, the measure of the light absorbed by the blood flowing into the arteries. An example of a PPG waveform is shown in the lower panel of Figure 1. This signal is commonly employed for estimating the blood oxygen saturation using red and infrared LED lights [12]. The PPG may also be exploited to track changes in the vessel diameter caused by the travelling of the blood pressure pulse. This information can be used to investigate features of the pressure pulse waveform and detect the pulse arrival at a given location of the vascular tree [13].

Even though ECG, SCG, and PPG individually provide rich information on cardiac and vascular performance, additional information may be derived when two or all three of these signals are simultaneously recorded and their reciprocal relationship investigated. For example, when PPG and ECG are concurrently recorded, the pulse transit time (PTT) can be measured. As schematized in the lower panel of Figure 1, this parameter is commonly estimated as the time delay from the R wave of the ECG and the arrival of the pressure pulse at a distal arterial site, usually the fingertip, earlobe or forehead (as detected by the PPG). The PTT may also be estimated by placing two PPG sensors on two different arterial sites and measuring the transit time of the pulse wave between sensors. The PTT inversely depends on the vascular stiffness, peripheral arterial resistance, and blood pressure, and its assessment provides us with integrated information on the vascular characteristics [14,15]. Further examples of simultaneous measures are illustrated in Section 3.2. At present, when concomitant recordings of multiple signals are needed, they are obtained by combining data from independent devices. The handling of multiple systems may lead to difficulties in the subject's instrumentation, data collection, and time synchronization among signals.

As part of our research activity in the cardiovascular area, we recently activated a project requiring the simultaneous measure of the above three signals for the monitoring of healthy subjects and heart failure patients. For this purpose, a specific acquisition platform named SeisMote was developed. In the following, we describe the new system and its performances and illustrate examples of its current applications.

2. Methods

The SeisMote system consists of a wireless network of 12 sensorized nodes. The overall architecture was designed to (1) be wearable and unobtrusive during daily activities and sleep; (2) allow a possible simultaneous assessment of each signal from different body sites (by placing more nodes containing the same type of sensor on different body spots), (3) provide a time synchronization among different nodes with a maximal error of 1 ms; (4) guarantee at least 10 h of continuous recording; and (5) facilitate the possible future inclusion of additional types of sensors into the nodes.

In the development of the system, particular attention was paid to the efficiency of the data transmission so as to maintain low power consumption and guarantee the node connectivity with the proper time synchronization. As detailed in the following, none of the commercially available transmission protocols met all our needs and, thus, a custom protocol was developed.

The system is composed of the wireless nodes, a USB dongle, which acts as network receiver, and the wireless battery recharger (see Figure 2), plus a software suit which includes a configuration/visualization program, a network file manager, and an Android app. Before each monitoring session the nodes are configured via software to select the signals to be acquired and one of the following monitoring modes:

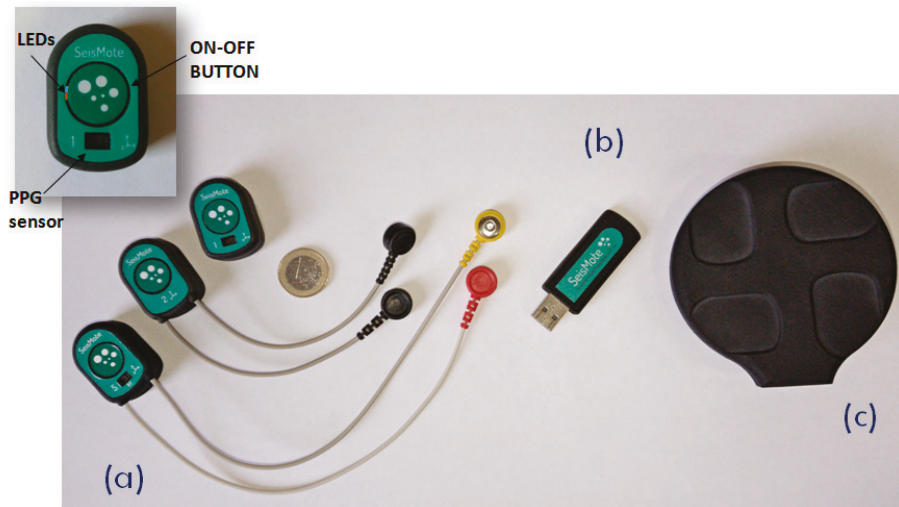


Figure 2. The hardware components of the SeisMote platform. (a) The sensorized nodes allowing the measure of different combinations of signals; from left to right: ECG–PPG–SCG, ECG–SCG, SCG–PPG. (b) The USB dongle (the hub). (c) The wireless recharger. Inset: A detail of a node and the position of the LEDs, switching button, and PPG sensor.

- Real-Time mode (RT). In this mode, data are collected by multiple nodes and sent to the receiver (in the following “the hub”) which re-transmits them to a computer for a real-time visualization, analysis, and storage. In RT mode, the hub provides time synchronization to all nodes as detailed in the subsequent sections;

- The Off-Line mode (OL). In this mode, multiple nodes are used but data are locally stored on the memory card of each node;
- The Bluetooth Real-Time mode (BRT). A single node in the system may also be configured to allow connection to smartphones and tablets with the BluetoothLE (BLE) protocol. However, as hereafter detailed, BLE would not guarantee the proper time synchronization among different nodes and, thus, only one node at a time can be used in this mode.

2.1. The Hardware Architecture

2.1.1. The Sensorized Node

Each node in the system has a size of $38 \times 25 \times 15$ mm and weighs 10 grams.

As schematized in Figure 3, the node's internal structure is composed of a motherboard and a daughterboard. The microcontroller (CC2650, Texas Instrument) is included in the motherboard. This component is based on the ARM-Cortex technology, has 8 kB of SRAM, 128 k of programmable flash memory, and an embedded 2.4 GHz RF transceiver. The motherboard also contains a secure digital memory card and the electronics for the wireless battery recharge and power supply.

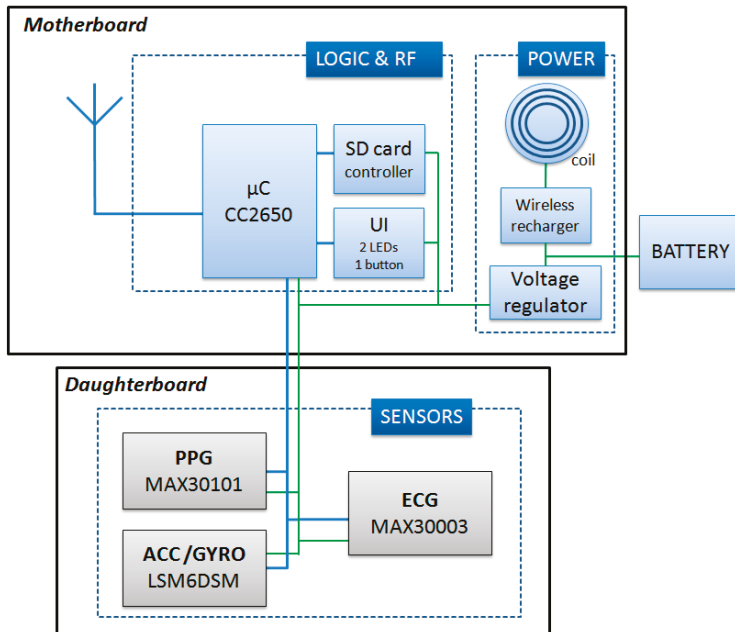


Figure 3. Scheme of the node circuit. UI = User interface; SD = Secure Digital.

The daughter board is 20×10 mm and is stacked on top of the motherboard. It contains a one-lead ECG front-end, a triaxial accelerometer (for the SCG measure), a triaxial gyroscope, and a green/red/infrared photoplethysmograph. Accelerations and rotational velocities are detected by the inertial unit LSM6DSM (ST microelectronics). For the SCG assessment, we need acceleration data with a resolution of 0.5 mg; the LSM6DSM component, when set with a full scale of ± 2 g provides acceleration measures with a sensitivity of 0.061 mg/LSB, namely, with a resolution approximately ten times higher than needed by our applications. The PPG is detected by the MAX30101 (Maxim integrated) component, and the ECG analog front end is managed by the MAX30003 (Maxim integrated) chip which provides a clinical-grade signal. The node is powered by a polymer lithium-ion rechargeable battery with a capacity of 150 mAh. As shown in Figure 2a, nodes are available with various combinations of the

above sensors. All signals are sampled at 200 Hz on 16 bits. Before transmission, sampled data may be encoded by the Adaptive Differential Pulse Code Modulation (ADPCM) algorithm [16] to improve the throughput of the network. The applicability and validity of this compression algorithm for the monitoring of biological signals has been previously tested [17,18].

2.1.2. The Hub

When the system is functioning in the RT mode, a hub is needed to coordinate nodes and receive their data. In our platform, this role is played by a custom USB dongle (Figure 2b) containing the same CC2650 microcontroller with RF transceiver used in the node motherboard; the power supply is taken from the USB port. The dongle has the master role in the network; it regularly broadcasts a timestamp for the node synchronization, handles possible transmission errors, receives data from all nodes, creates a unique synchronized data stream, and send it to the PC via the USB port.

2.2. The Network Real-Time Mode

One challenging task of this project has been the development of the transmission protocol to be used when the system operates in the RT mode. Indeed, we needed a low-energy protocol able to connect up to 12 nodes with sufficient bandwidth and, importantly, capable to keep the time synchronization among nodes with a maximal error <1 ms. From a preliminary market survey, we soon discovered that none of commercially available low-power wireless technologies (e.g., ANT, ZigBee, Z-Wave, Bluetooth, BluetoothLE) fitted our requirements and, thus, we decided to design an ad-hoc protocol. Its details are provided in the following: Sections 2.2.1–2.2.3.

2.2.1. Data Transmission

The unlicensed 2.4 GHz ISM (Industrial Scientific and Medical) band was used for the data transmission, and the star topology was adopted to transfer data from multiple slave nodes to the coordinator of the network, the hub, and vice versa.

To minimize the dimension of the nodes, it was decided that the transmission protocol had to be managed by the same microcontroller governing all the remaining node functions and not by a dedicated component. This policy required an optimization of CPU scheduling and memory resources to allow the concurrent running of the processes controlling data acquisition and data transmission.

Coordination among nodes was achieved by the TDMA (Time Domain Multiple Access) methodology [19–21]. In TDMA, a certain RF channel is allocated for the access of one master and N slaves. In the channel, only one device at a time is allowed to transmit data; thus, different devices have dedicated time slots during which they can exclusively transmit. In our TDMA implementation, the transmission frame (in the following called “connection event” (CE)) is subdivided into 15 time slots, each lasting 1 ms (see Figure 4). The first slot, S0, is allocated to the master (the hub) to transmit a beacon packet to all nodes. This specific data packet contains a timestamp used for the time synchronization among nodes and possible additional network commands. Slots S2–S13 are allocated to the nodes 1–12 for data transmission. Slots S1 and S14 are not used. In particular, S1 is reserved to leave sufficient time to node 1 for the execution of possible commands received from the master before the data transmission in S2. Slot S14 is reserved to allow the actuation of the frequency hopping (see hereafter). When all 12 nodes are active in the network, each node has one slot assigned. In the case of fewer nodes, a single node may have up to 3 slots assigned for every CE to be used for the transmission of more signals or for allowing more time for data retransmission in case of error recovery.

Parenthetically, the adoption of the TDMA technique also leads to a reduction in power consumption. Indeed, as shown in Figure 4, during the beacon transmission, all slaves listen to the master, but in the subsequent phase, each node transmits data in the assigned time slot and remains in the idle state for the remaining time. This means that from a CE to the subsequent one, the RF subsystem of each node remains in the idle state for most of the time.

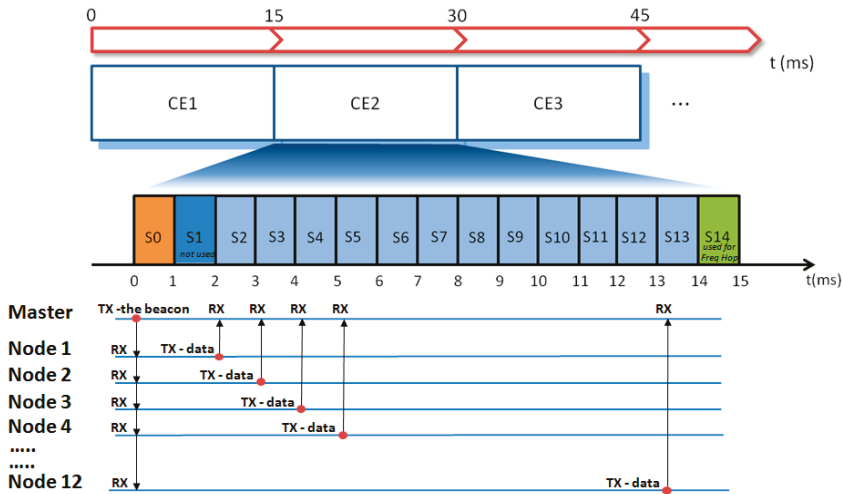


Figure 4. Diagram illustrating the TDMA (Time Domain Multiple Access) implementation and the data flow between the master (hub) and nodes at every connection event (CE). Freq Hop = Frequency Hopping.

The 2.4 GHz ISM band is also used by WiFi, Bluetooth, and proprietary protocols. Thus, it is possible that other devices are working exactly in the same RF channel used by our platform; in this case, the transmitted data may be corrupted. To limit the negative effects of this eventuality, the transmission frequencies are changed over time through the frequency hopping technique [19,22,23] at every CE. Frequency hopping is also useful to counteract the effect of the multi-path fading, namely, the attenuation of the received RF power due to the multiple reflections of the electromagnetic waves [19]. Indeed, when transmitter and receiver do not operate in line-of-sight, the transmitted signal is reflected by walls or obstacles in the ambient. In this case, the signal can reach the receiver multiple times with different time delays due to the different propagation paths of the RF waves. In specific configurations of the obstacles, this phenomenon may result in a negative interference which reduces the power of the received signal. For any given ambient configuration, the extent of the power reduction depends on the frequency; thus, the possible fading effect can be mitigated over time by a regular change of the transmission channel.

2.2.2. Data Fragmentation and Error Recovery

It is desired that the protocol be compatible with Application Packets of variable length also in view of possible future developments. For this purpose, a data fragmentation policy is adopted in our protocol (see Figure 5). It consists of the partitioning of the Application Packets into smaller fixed-length data chunks, the Link Layer Packets, which can be transmitted over multiple CEs. These chunks are then reassembled on the receiver side by a defragmentation procedure to reconstruct the original message. In the present implementation of the protocol, the Link Layer Packet has a size of 36 bytes.

Some transmitted packets may be lost in reception, however. This event may occur for different reasons: (1) because of the abovementioned fading effect; (2) because of the absorption phenomena of electromagnetic waves, e.g., due to the water content of the human body [24]; and (3) because of interference with other wireless devices using the same RF channel, notwithstanding the frequency hopping. A retransmission of the lost packets is implemented in the stack.

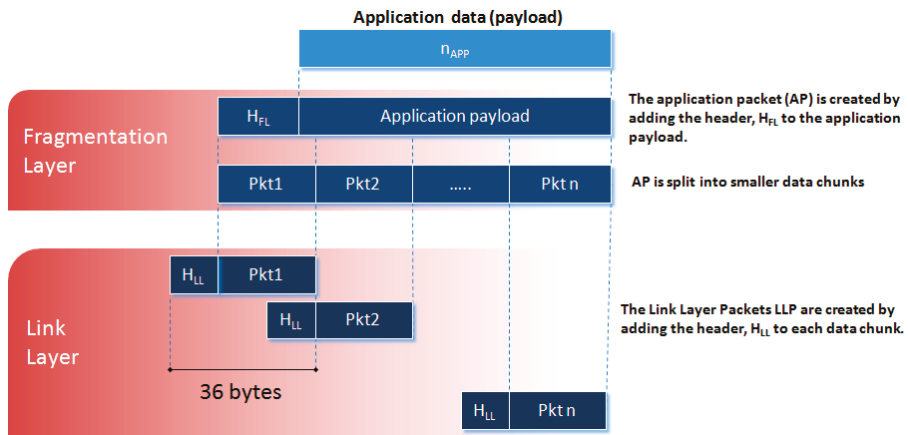


Figure 5. Diagram of the data fragmentation.

2.2.3. Time Synchronization

In our applications, some cardiac parameters are computed as time delays among displacements occurring in signals collected by different nodes. Thus, we need that every node has the same vision of time with an error <1 ms. The quartzes of our nodes have a nominal frequency of 4 MHz with a tolerance of 10 ppm. This means that in the worst case and in absence of a time resynchronization, the accumulated time error between two nodes may reach 1 ms after only 50 s of transmission. This value may be reached considering only the possible intrinsic deviation in the nominal frequency of the quartzes. Additional changes in the quartz frequencies may also be expected from other factors including temperature. Therefore, the clock of every node in the network must be periodically adjusted to retain the same timekeeping. In the literature, there are three main types of synchronization within a network [25]: unidirectional, bidirectional, and reference broadcasting [26]. We preferred to keep the exchange of synch messages among the nodes at a minimal level; thus, we adopted the unidirectional synchronization procedure. By this method, N slaves (the nodes) are calibrated simultaneously through a single broadcast message sent by the master (the hub) which contains the timestamp. This message is embedded into the beacon. At every CE, each node receives the updated timestamp and adjusts the local clock.

2.3. The Network Off-Line Mode

When the system is in the OL mode, the hub is not used, and sensors' data are locally stored in each node. At the end of the recording session data are downloaded from each node, pooled together and aligned over time by a custom program. This mode allows the data collection far from the receiver. If more than one node is used, during the initial configuration procedure, one of the nodes of the network is assigned the role of master to control the timekeeping. During the monitoring the master node transmits the beacon with the reference time to all the other nodes in the network every CE. The slave nodes receive the beacon and synchronize their own clock but do not transmit the sensor data to the master and rather store them on the local memory card. If only a single node is used, obviously no synch is required, and the wireless communication is disabled.

2.4. The BLE Real-Time Mode

This further functioning mode has been added to allow a real-time monitoring using a smartphone or tablet with a BLE connection. As already mentioned, the time synchronization among nodes cannot be guaranteed by the BLE protocol, and the throughput is limited; thus, only one node at a time can

be used in this modality. In the designated nodes, the BLE stack has been included in the firmware replacing the code of the proprietary protocol.

A schematization of the three monitoring modes is illustrated in Figure 6.

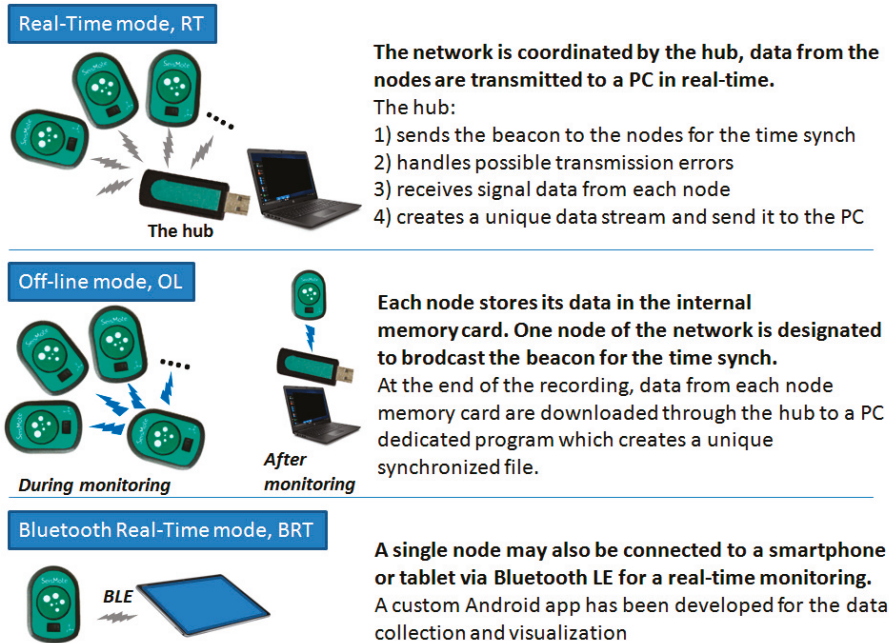


Figure 6. Scheme of the three monitoring modes.

3. Results

An example of collected data is shown in Figure 7. Data were recorded in the setting illustrated in Figure 9.

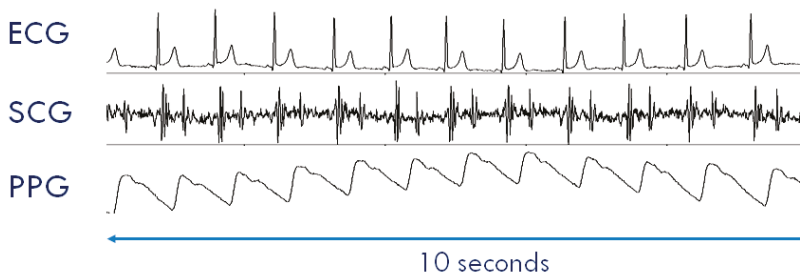


Figure 7. Example of ECG, SCG, and PPG collected by a SeisMote node. Data were collected in the setting illustrated in Figure 9.

3.1. System Performance

3.1.1. Node Current Consumption

The total current consumption measured while the node was acquiring and transmitting data from all sensors was 9.4 mA which corresponded to approximately 16 h of continuous data monitoring for each battery recharge. The latter was obtained through a wireless charger in 2.5 h.

3.1.2. Network Throughput

The bitrate of the radio chip was 500 kbps. As indicated in Section 2.2.1, at each CE, namely, every 15 ms, the node transmitted a Link Layer Packet (5 bytes for the header + 31 bytes of payload) which carried a chunk of application data. In a second, corresponding to 66.666 CEs, the throughput for each node was $31 \times 66.666 = 2.07$ kbps. For the whole network of 12 nodes, the total TP was $12 \times 2.07 = 24.8$ kbps = 198.7 kbps with an efficiency of 39.7%.

To estimate the TP, also in terms of the application data, it should be considered that in our protocol, the Application Packets have a header, H_{app} , of 7 bytes independent from the length of the payload. Because of the splitting of the application packets into Link Layer Packets, in a single node, the link between application TP, TP_{app} , and length of the application payload, n_{app} , is described by the following formula:

$$TP_{app}(n_{app}) = n_{app}/\text{round}((n_{app} + H_{app})/LLP) \times 66.666 \quad ()$$

where LLP is the length of the Link Layer payload (31 bytes), and the round function rounds the argument up to the nearest integer.

The relationship between TP_{app} and n_{app} has a sawtooth behavior as reported in Figure 8. We set the length of our application packets to 179 bytes; thus, from the above formula, the single node TP_{app} is 1.99 kbps corresponding to a network TP_{app} of $12 \times 1.99 = 23.88$ kbps = 190.9 kbps with an efficiency of 38.18%. For a full 12 node configuration, this setting allows each node to transmit up to three signals (16 bit @200 Hz). In the case of configurations with fewer nodes, the number of signals that can be transmitted progressively increases up to nine signals per node for a configuration of four nodes. In terms of error recovery, this setting allows a safe re-transmission of 6% of packets when all 36 signals are collected. Obviously, the allowed re-transmission rate increases if the number of collected signals is reduced.

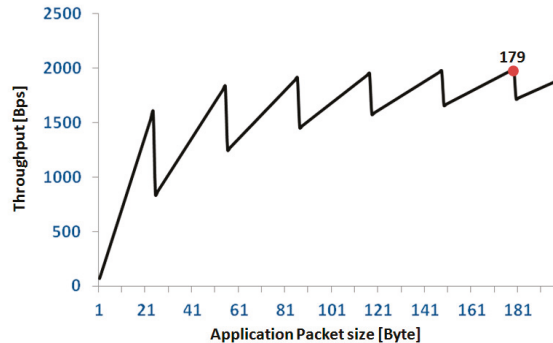


Figure 8. Relationship between application throughput and size of the application packet in a single node. The red spot indicates the application packet size selected for our protocol.

We checked the quality of transmission in our laboratory. In all trials, we considered the full network configuration (i.e., 12 nodes), each transmitting three signals (we arbitrarily selected the x -, y - and z -axis of the acceleration); the nodes and receiver were in line-of-sight. Two types of tests were performed. In the first type (static), the 12 nodes were positioned close to each other on a tray, and the tray was located at 2, 5, 8, and 10 meters from the hub. Measures were taken for 5 min at each distance. The test was repeated three times. In the second group of tests (dynamic), measures were taken in a subject wearing the nodes and walking at 2, 5, 8, and 10 meters from the receiver for 3 min. Six nodes were placed on the front wall of the chest and the remaining six nodes on the back wall at the level of the 6th rib (just under the pectoral muscles). All nodes were in direct contact with the skin and were fastened by an elastic strap. Also, this test was repeated three times. During all recordings, WiFi was active in the area. For each test and each node we measured: (1) the received signal strength indicator

(RSSI); (2) the percentage of Link Layer Packets retransmitted for the recovery of transmission errors; and (3) the number of Application Packets lost because of failure in the error recovery; lost application packets produce gaps inside the recording. Table 1 illustrates the results obtained from each test type. Values of RSSI and the percentage of retransmitted packets were averaged over the 12 nodes and the three test repetitions; the number of lost packets is the cumulative number of Application Packets lost by all nodes. It was apparent that 8 m is the maximal distance between subject and receiver for a good quality transmission.

Table 1. The results of the trial aimed at checking the quality of the transmission. RSSI = received signal strength indicator; %RP = percentage of retransmitted link layer packets.

Static Tests			
	RSSI	%RP	Lost packets
2 m	−55 dBm	0	0
5 m	−60 dBm	0	0
8 m	−63 dBm	0	0
10 m	−68 dBm	1.5%	0
Dynamic Tests			
2 m	−65 dBm	0	0
5 m	−73 dBm	2%	0
8 m	−78 dBm	5%	0
10 m	−83 dBm	9%	2

3.1.3. Time Synchronization

The jitter in the time synchronization among nodes was evaluated by measuring at every CE the discrepancy between the node local clock and the timestamp just transmitted by the master. In each node, the time was kept by a quartz-controlled counter advancing of a tick every 50 μ s, thus our time measures were expressed in number of ticks. For the test, the node was programmed with a modified version of the firmware: the sensor signals were sampled (3 signals 16 bit @200 Hz) and the application packets prepared as usual, but at variance from the standard version, now these packets were not transmitted but rather trashed. Instead, a new packet with the difference in the timings was sent to the master. In this way, the processing load for the CPU was kept close to the real monitoring condition as much as possible. The measurement was taken with five different nodes, and each test lasted 20 min.

Results indicated that in every node the time discrepancy between consecutive resynchronization events ranged from 0 to 1 tick. These values correspond to a maximal jitter between nodes of 200 μ s, i.e., well below the 1 ms threshold set in the project specification.

3.1.4. Wearability

Each node in the system can be directly positioned on the body's surface by adhesive tape, elastic straps, clips, or via integration into clothing. In our current application, where ECG and SCG were measured, the node/s were placed on the thorax. In this case, we usually applied a small piece of medical plaster to the chest and then attached the node to this substrate using a bi-adhesive tape. With this arrangement, the tape was not in direct contact with the skin and a strong adhesion tape may be used without risking possible skin irritations. This strategy was found to provide a comfortable and efficient bonding of the node to the body during movement and sweating. In addition, the small mass of the node makes it imperceptible while wearing, even during sleep.

When the PPG is used, two different scenarios are possible. First, the subject stays still; a single node is used, and the node is placed on the sternum with the bi-adhesive tape for the ECG and SCG measurements (this arrangement is currently used for the monitoring of cardiac patients in the frame of the telerehabilitation project illustrated in the next section). In this case, PPG can be measured for short

time periods by just putting the finger on the PPG sensor as shown in Figure 9. For a more general PPG measurement, we rotate the node so as to have the PPG sensor in contact with the skin, and specific adapters are used to keep the node adherent to the body. As shown in Figure 10, three types of adapters have been developed. The first is a clip for the PPG measure at the earlobe, and the second and the third are straps of different lengths that fasten the node to a finger and forehead, respectively.

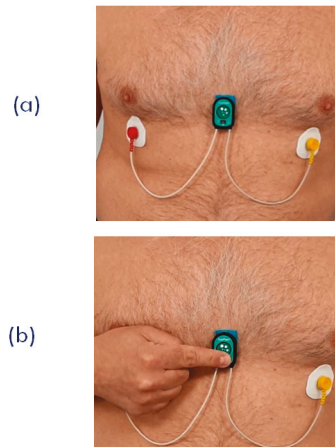


Figure 9. (a) Setting of the single node for the joint measurement of ECG, SCG, and PPG. This arrangement was used for the remote signal monitoring during the SideraB telerehabilitation program (see text). (b) Finger positioning for the PPG measurement.

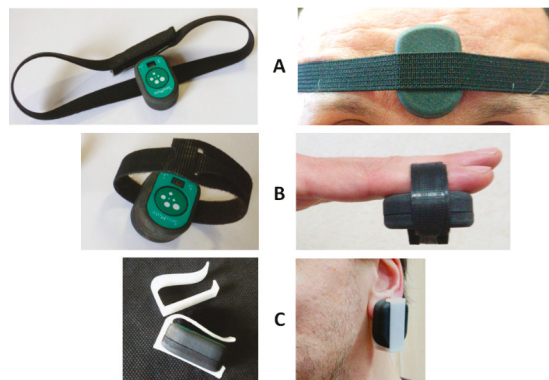


Figure 10. Adapters for the PPG measurement at the forehead (A), fingertip (B), and earlobe (C).

3.2. Applications

The system is currently used in both the single- and multi-node configurations for the monitoring of healthy subjects and cardiac patients in a laboratory environment, telemedicine, and during sleep. The RT feature of the platform is now exploited to investigate the differences in the PTT dynamics when measured in various vascular districts simultaneously. Four nodes are used for this protocol. The first node is placed on the chest to measure ECG and SCG; the second, third, and fourth node detect PPGs, respectively, at the fingertip, earlobe, and forehead. An example of a PPG multisite measurement is shown in Figure 11. The figure also illustrates how the PTTs are estimated from the ECG and PPG signals. Through the analysis of these data, we are now studying the strategies of local vascular blood pressure regulation.

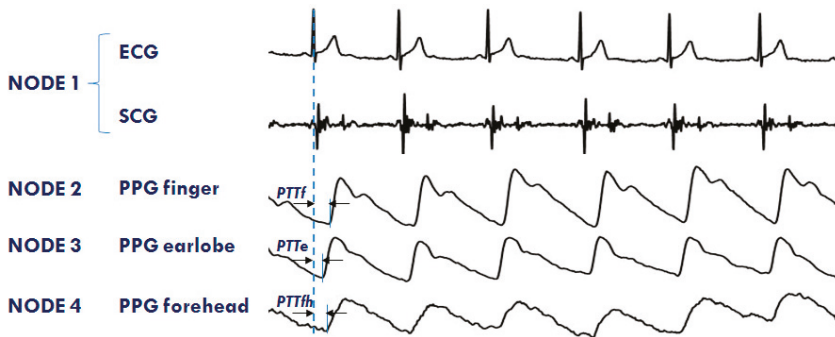


Figure 11. Example of signals from the 4 node configuration of the system aimed at investigating the PTT in different arterial sites: finger (PTTf), earlobe (PTTe), and forehead (PTTfh) by a multisite measurement of PPG.

The second application of the system in the RT mode refers to the assessment of SCG from different precordial locations. The study was triggered by the observation that doctors auscultate the heart sounds from various chest sites to evaluate the different features of the heart’s performance. Similarly, it can be hypothesized that multisite measurement of SCG might provide more details on heart mechanics than a mono-site measurement. For this investigation, three nodes were used (additional nodes will be used in the future). They were placed on the lower part of the sternum (the traditional assessment site for SCG) in correspondence with the 2nd right intercostal space (the position for the aortic valve auscultation by the doctors) and on the heart apex (the position for the mitral valve auscultation). A diagram of the node position and an example of collected data are shown in Figure 12. It is apparent that, although common patterns are present in all signals, each individual SCG waveform is also characterized by peculiar features. We are now investigating the correlations between the SCG morphologies obtained from the multisite assessment and the real heart mechanical events visualized by ultrasound images.

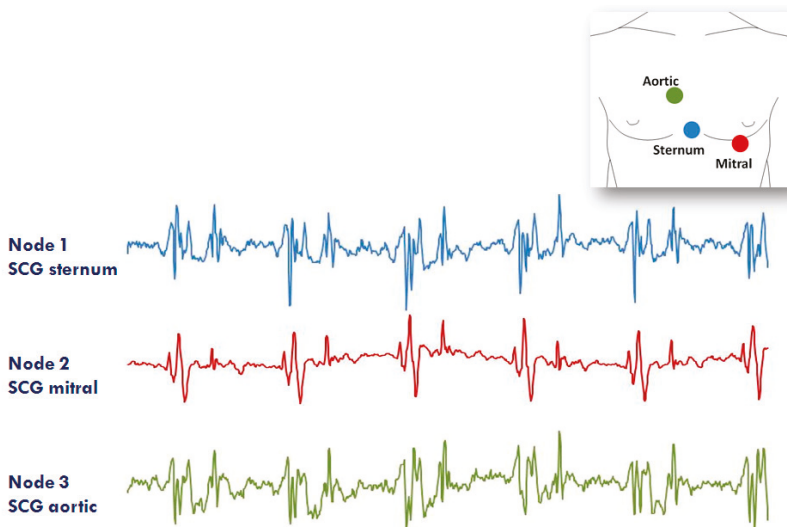


Figure 12. Example of signals from the 3 node configuration of the system aimed at investigating the SCG from different thorax sites (i.e., lower sternum, mitral, and aortic auscultation sites). Inset: positioning of the three nodes for the measure.

Further applications of the system make use of single nodes. In the first ongoing study, the nodes are used to monitor ECG and SCG during sleep. The aim of the study is to investigate the correlation between heart rate variability and dynamic characteristics of the cardiac mechanics in this condition. For this protocol, the node is working in OL mode; thus, during monitoring, data are locally stored on the memory card. So far, 10 sleep recordings have been performed.

In the second application, single nodes were used in BRT mode to remotely monitor patients with heart failure during their telerehabilitation program at home. The study is part of a wide research project, SIDERA-B, financed by the Italian regional government of Lombardy, Regione Lombardia (POR FESR, id 232549), and aimed at testing new methodologies for the telerehabilitation of patients after hospital discharge. Each patient is guided by a tablet to do a series of physical activities and take a number of biomedical self-measurements every day. Also, once a week, a 3 min recording of ECG, SCG, and PPG from our device is taken. Signals are recorded by a single node placed on the sternum by adhesive tape as indicated in Section 3.1.4. while the patient is sitting. For the first two minutes, only the ECG and SCG are recorded, while, in the last minute, the patient places his/her finger on the PPG sensor (Figure 10) and also this signal is recorded. All devices, including our node, transmit data to the tablet via a BluetoothLE connection. Data are then re-transmitted to a central server that automatically prepares the reports and sends them to the cardiologists. The integration of the SeisMote node into the telerehabilitation platform, through the joint provision of ECG, indexes of cardiac mechanics and PTT, is intended to augment information on the patient's health status. This should facilitate the evaluation of the effects of rehabilitation on the cardiovascular performance and a fast tuning of the exercise load on the basis of possible changes in the patient condition. This study is still in progress, and 48 recordings have so far been received and analyzed.

The above employment of the system is characterized by recordings without data gaps caused by transmission errors, and battery durations exceeding 16 h.

4. Discussion

In this article, a new wireless platform for the monitoring of cardiovascular performance was presented. The system was designed to guarantee flexibility of use in terms of the type of signals to be monitored, number of nodes, and functioning modalities.

SeisMote implements a wireless body area network (WBAN). Several platforms based on this paradigm have been proposed in the literature for the monitoring of vital signs (recent surveys may be found in References [27–29]). However, to the best of our knowledge, none of those platforms are characterized by the features we needed in terms of low-power consumption, number of nodes, and synch jitter. Two interesting solutions are commercially available, but one of them is based on the Bluetooth piconet thus limiting the number of connectable nodes to seven [30], and the other allows the connection to only three nodes [31]. None of the WBAN systems provides a SCG measure. Some of the systems currently available for SCG assessment may transmit sensor data to a receiver via a wireless connection [10,32,33], but they are essentially single nodes and are not part of a WBAN.

Thus, the SeisMote's ability to measure up to 36 signals by dislocating 12 sensorized nodes in different parts of the body with an accuracy in the time synchronization better than 200 μ s represents a unique feature. In particular, the latter characteristic allows a solid estimation of important biological parameters, such as PTT, based on the measure of time delays among signals collected by different nodes.

Another feature of the system is the possibility to have a multisite measure of the same signal. This aspect paves the way for interesting experimental applications. Two of them, namely, the multisite measure of accelerations and PPG (from which SCG and PTT were derived) have been described in Section 3.2. However, the possibility of obtaining simultaneous measures of single-lead ECGs by multiple nodes placed in different locations may also have practical relevance. Indeed, evidence is emerging that standard ECG leads, such as the Einthoven leads I, II, and III, might be synthesized by multiple single-lead ECG measurements [34].

Finally, a word on the layout of the node electronics. At present, the nodes include sensors for ECG, acceleration, gyroscope, and PPG measurement. However, their hardware architecture was designed to ease the integration of additional sensors by only the change of the daughterboard and keeping untouched the motherboard containing the microcontroller with the RF section, storage, and power supply electronics.

Future developments: In the current form, SeisMote allows the monitoring of signals, while the data analysis is performed offline. The next enhancement of the platform will include a DSP (Digital Signal Processor) chip and more memory in the circuit of the node so as to provide also real-time computation of the derived parameters (such as the PTT and the indexes of cardiac mechanics). The second planned improvement will be an increase in the battery's duration in order to allow monitoring over 24 h.

Author Contributions: Conceptualization, M.D.R. Data curation, G.R and Z.M.I.; Formal analysis, Z.M.I.; Investigation, Z.M.I. and P.L.; Methodology, M.D.R. and P.L.; Resources, P.L.; Supervision, G.R.; Validation, M.D.R., G.R. and P.L.; Writing—original draft, M.D.R.; Writing—review and editing, P.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

- Kennedy, H.L. The History, Science, and Innovation of Holter Technology. *Ann. Noninvasive Electrocardiol.* **2005**, *10*, 1–10. [[CrossRef](#)] [[PubMed](#)]
- Crow, R.S.; Hannan, P.; Jacobs, D.; Hedquist, L.; Salerno, D.M. Relationship between seismocardiogram and echocardiogram for the events in the cardiac cycle. *Am. J. Noninvasive Cardiol.* **1994**, *8*, 39–46. [[CrossRef](#)]
- Inan, O.; Migeotte, P.F.; Park, K.S.; Etemadi, M.; Tavakolian, K.; Casanella, R.; Zanetti, J.; Tank, J.; Funtova, I.; Prisk, G.K.; et al. Ballistocardiography and Seismocardiography: A Review of Recent Advances. *J. Biomed. Health Inform.* **2015**, *19*, 1414–1427. [[CrossRef](#)] [[PubMed](#)]
- Taebi, A.; Solar, B.E.; Bomar, A.J.; Sandler, R.H.; Mansy, H.A. Recent Advances in Seismocardiography. *Vibration* **2019**, *2*, 64–86. [[CrossRef](#)]
- Sorensen, K.; Schmidt, S.E.; Jensen, A.S.; Søgaard, P.; Struijk, J.J. Definition of Fiducial Points in the Normal Seismocardiogram. *Sci. Rep.* **2018**, *8*, 15455.
- Paukkunen, M.; Parkkila, P.; Hurnanen, T.; Pankaala, M.; Koivisto, T.; Nieminen, T.; Kettunen, R.; Sepponen, R. Beat-by-Beat Quantification of Cardiac Cycle Events Detected From Three-Dimensional Precordial Acceleration Signals. *IEEE J. Biomed. Health Inform.* **2016**, *20*, 435–439. [[CrossRef](#)]
- Di Rienzo, M.; Vaini, E.; Castiglioni, P.; Merati, G.; Meriggi, P.; Parati, G.; Faini, A.; Rizzo, F. Wearable seismocardiography: Towards a beat-to-beat assessment of cardiac mechanics in ambulant subjects. *Auton. Neurosci.* **2013**, *178*, 50–59. [[CrossRef](#)]
- Becker, M.; Roehl, A.B.; Siekmann, U.; Koch, A.; de la Fuente, M.; Roissant, R.; Radermacher, K.; Marx, N.; Hein, M. Simplified detection of myocardial ischemia by seismocardiography. *Herz* **2014**, *39*, 586–592. [[CrossRef](#)]
- Tavakolian, K.; Dumont, G.A.; Houlton, G.; Blaber, A.P. Precordial vibrations provide noninvasive detection of early-stage hemorrhage. *Shock* **2014**, *41*, 91–96. [[CrossRef](#)]
- Hurnanen, T.; Lehtonen, E.; Tadi, M.J.; Kuusela, T.; Kiviniemi, T.; Saraste, A.; Vasankari, T.; Airaksinen, J.; Koivisto, T.; Pankaala, M. Automated Detection of Atrial Fibrillation Based on Time-Frequency Analysis of Seismocardiograms. *IEEE J. Biomed. Health Inform.* **2017**, *21*, 1233–1241. [[CrossRef](#)]
- Inan, O.T.; Baran Pouyan, M.; Javaid, A.Q.; Dowling, S.; Etemadi, M.; Dorier, A.; Heller, J.A.; Bicen, A.O.; Roy, S.; De Marco, T.; et al. Novel Wearable Seismocardiography and Machine Learning Algorithms Can Assess Clinical Status of Heart Failure Patients. *Circ. Heart Fail.* **2018**, *11*, e004313. [[CrossRef](#)] [[PubMed](#)]
- Sinex, J.E. Pulse oximetry: Principles and limitations. *Am. J. Emerg. Med.* **1999**, *17*, 59–66. [[CrossRef](#)]
- Allen, J. Photoplethysmography and its application in clinical physiological measurement. *Physiol. Meas.* **2007**, *28*, R1–R39. [[CrossRef](#)] [[PubMed](#)]

14. Chiu, Y.C.; Arand, P.W.; Shroff, S.G.; Feldman, T.; Carroll, J.D. Determination of pulse wave velocities with computerized algorithms. *Am. Heart J.* **1991**, *121*, 1460–1470. [[CrossRef](#)]
15. Kei Fong, M.W.; Ng, E.Y.K.; Er Zi Jian, K.; Hong, T.J. SVR ensemble-based continuous blood pressure prediction using multi-channel photoplethysmogram. *Comput. Biol. Med.* **2019**, *113*, 103392. [[CrossRef](#)]
16. Cummiskey, P.; Jayant, N.S.; Flanagan, J.L. Adaptive quantization in differential PCM coding of speech. *Bell Syst. Tech. J.* **1973**, *52*, 1105–1118. [[CrossRef](#)]
17. Norris, J.F.; Lovely, D.F. Real-time compression of myoelectric data utilising adaptive differential pulse code modulation. *Med. Biol. Eng. Comput.* **1995**, *33*, 629–635. [[CrossRef](#)]
18. Farshchi, S.; Pesterev, A.; Nuyujukian, P.H.; Mody, I.; Judy, J.W. Bi-Fi: An Embedded Sensor/System Architecture for Remote Biological Monitoring. *IEEE Trans. Inform. Tech. Biomed.* **2007**, *11*, 611–618. [[CrossRef](#)]
19. Rappaport, T.S. Wireless communications: Principles and Practice. In *Prentice Hall Communications Engineering and Emerging Technologies Series*, 2nd ed.; Prentice Hall: Upper Saddle River, NJ, USA, 1998.
20. Rubin, I. Access-Control Disciplines for Multi-Access Communication Channels: Reservation and TDMA Schemes. *IEEE Trans. Inform. Theory* **1979**, *25*, 516–536. [[CrossRef](#)]
21. Simonot, F.; Song, Y.Q. Real-time communications using TDMA-based multi-access protocol. *Comput. Commun.* **1997**, *20*, 435. [[CrossRef](#)]
22. Ephremides, A.; Wieselthier, J.E.; Baker, D.J. A design concept for reliable mobile radio networks with frequency hopping signaling. *Proc. IEEE* **1987**, *75*, 56–73. [[CrossRef](#)]
23. Popovski, P.; Yomo, H.; Prasad, R. Strategies for Adaptive Frequency Hopping in the Unlicensed Bands. *IEEE Wirel. Commun.* **2006**, *13*, 60–67. [[CrossRef](#)]
24. Kuster, N.; Balzano, Q. Energy absorption mechanism by biological bodies in the near field of dipole antennas above 300 MHz. *IEEE Trans. Veh. Technol.* **1992**, *41*, 17–23. [[CrossRef](#)]
25. Romer, K.; Blum, P.; Meier, L. *Handbook of Sensor Network: Algorithms and Architectures*; Stojmenovic, I., Ed.; Wiley: Hoboken, NJ, USA, 2005.
26. Elson, J.; Girod, L.; Estrin, D. Fine-grained network time synchronization using reference broadcast. In Proceedings of the 5th Symposium on Operating Systems Design and Implementation, Boston, MA, USA, 9–11 December 2002; pp. 147–163.
27. Majumder, S.; Mondial, T.; Deen, M.J. Wearable sensors for remote health monitoring. *Sensors* **2017**, *17*, 130. [[CrossRef](#)] [[PubMed](#)]
28. Hasan, K.; Biswas, K.; Ahmed, K.; Nafi, N.S. A comprehensive review of wireless body area network. *J. Netw. Appl. Technol.* **2019**, *143*, 178–198. [[CrossRef](#)]
29. Kan, R.A.; Pathan, A.K. The state-of-the-art wireless body area sensor networks: A survey. *Int. J. Distrib. Sens. Netw.* **2018**, *14*, 1–23. [[CrossRef](#)]
30. Shimmer. Available online: <https://shimmersensing.com/> (accessed on 25 January 20).
31. Biopac Bionomadix. Available online: <https://www.biopac.com/product/bionomadix-smart-sys/> (accessed on 25 January 20).
32. Heartkinetics. Available online: <http://heartkinetics.com/> (accessed on 25 January 20).
33. Heartforce. Available online: <http://heartforce.com/> (accessed on 25 January 20).
34. Samol, A.; Bischof, K.; Luani, B.; Pascut, D.; Wiemer, M.; Kaese, S. Recording of Bipolar Multichannel ECGs by a Smartwatch: Modern ECG Diagnostic 100 Years after Einthoven. *Sensors* **2019**, *19*, 2894. [[CrossRef](#)]



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).



Article

A Novel Adaptive Recursive Least Squares Filter to Remove the Motion Artifact in Seismocardiography

Shuai Yu ¹ and Sheng Liu ^{1,2,3,*}

¹ School of Mechanical Science and Engineering, Huazhong University of Science and Technology, 1037 Luoyu Road, Wuhan 430074, China; yushuai91@hust.edu.cn

² Key Lab for Hydropower Transients of Ministry of Education, School of Power and Mechanical Engineering, Wuhan University, 8 East Lake South Road, Wuhan 430072, China

³ Institute of Technological Sciences, Wuhan University, 8 East Lake South Road, Wuhan 430072, China

* Correspondence: shengliu@whu.edu.cn; Tel.: +86-27-8754-2604; Fax: +86-27-8755-7074

Received: 9 February 2020; Accepted: 9 March 2020; Published: 13 March 2020

Abstract: This paper presents a novel adaptive recursive least squares filter (ARLSF) for motion artifact removal in the field of seismocardiography (SCG). This algorithm was tested with a consumer-grade accelerometer. This accelerometer was placed on the chest wall of 16 subjects whose ages ranged from 24 to 35 years. We recorded the SCG signal and the standard electrocardiogram (ECG) lead I signal by placing one electrode on the right arm (RA) and another on the left arm (LA) of the subjects. These subjects were asked to perform standing and walking movements on a treadmill. ARLSF was developed in MATLAB to process the collected SCG and ECG signals simultaneously. The SCG peaks and heart rate signals were extracted from the output of ARLSF. The results indicate a heartbeat detection accuracy of up to 98%. The heart rates estimated from SCG and ECG are similar under both standing and walking conditions. This observation shows that the proposed ARLSF could be an effective method to remove motion artifact from recorded SCG signals.

Keywords: adaptive recursive least squares filter (ARLSF); Seismocardiography (SCG); motion artifact; Electrocardiogram (ECG); heart rate

1. Introduction

Seismocardiography (SCG) is a non-invasive measurement that records the local vibrations of the chest wall in response to the heartbeat [1]. SCG was first discovered in 1961 [1], and its first clinical application was used 30 years later in 1991 [2]. With the development of an accelerometer with high-sensitivity, low-noise, small-size, high-efficiency, and high-robustness signal-processing technology, SCG has shown its great potential to be used by wearables. Consequently, it is now feasible to use the information in clinical applications [3,4].

However, there are still some limitations in SCG measurements and assessments. Motion artifact is one of the major limitations. As a result, SCG research on motion artifact has been very active in recent years. Motion artifact is usually irregular and it is mixed with the heartbeat signals in the time and frequency domains. The mixture makes it difficult to separate the heartbeat signal from the mixed signal [3–5]. Some researchers tried to use multiple sensors to remove the motion artifact from the recorded signals [6–9]. The tri-axis acceleration data collected from a chest-worn accelerometer were utilized to remove motion artifact in an electrocardiogram (ECG) signal in 2003 [6], 2008 [7] and 2010 [8]. An independent component analysis approach and a normalized least means square (NLMS) adaptive filter to motion artifact cancellation of the SCG signal using two accelerometers were developed in [9] and [10], respectively. In these studies, one accelerometer was placed at the center of the sternum and the other was attached to the right side of the back of the subjects. The results

were promising, but multiple sensors used in the experiments increased the complexity of the SCG measurement and assessment.

Meanwhile, other researchers developed several algorithms to remove motion artifact from the SCG signal by using only one accelerometer [11–17]. An ensemble empirical mode decomposition method was developed to remove white noise from a synthetic vibrocardiographic signal in [11] and the same method was successfully utilized to reduce the motion artifacts generated due to walking at normal and moderately fast speeds at a treadmill [12]. However, SCG signal could not be well recovered from corrupted signal. Rienzo et al. utilized an accelerometer to record a 24-hour SCG signal from freely moving subjects [13]. A movement-free SCG was extracted from the recorded accelerometer data using a continuous 5-second segment-based method. The results were promising, but the physiological parameters were not extracted from the signal. Pandia et al. designed a Savitzky Golay-based polynomial smoothing algorithm to extract the primary heart sound from the accelerometer data during walking [14]. The primary heart sound detection rate was up to 99.36%, but the graph of the extracted SCG signal could not be recovered. In another approach, the motion-free SCG signal was successfully extracted using a time delay-based normalized least mean square (NLMS) adaptive filter [17]. However, an extra moving average method was utilized to obtain the heart rate as the primary heart sound graph was not clear in the extracted motion-free SCG signal.

To solve this problem, we present a novel adaptive recursive least squares filter (ARLSF) for motion artifact removal in the SCG signal that was obtained by one accelerometer. The primary heartbeat signal graph is very clear in the motion artifact removed SCG signal without any other signal-processing procedures.

In Section 2 of this paper, the main idea of ARLSF is introduced and two major parameters of ARLSF are discussed. The measurement system including the hardware system, experimental setup and software system are discussed in Section 3. Section 4 shows the results and Section 5 concludes this paper.

2. Theory of Adaptive Recursive Least Squares Filter (ARLSF)

2.1. The Principle of Adaptive Recursive Least Squares Filter

Figure 1 illustrates the block diagram of ARLSF. An RLS filter is a finite impulse response (FIR) filter of length M with coefficients $\mathbf{w}(n)$ [18,19]. The input vector $\mathbf{u}(n)$ is passed through the FIR filter to produce the output vector $y(n)$. At each time-step, the coefficients are updated through the adaptive control unit using the input vector $\mathbf{u}(n)$. The prior estimation error $\xi(n)$ is described in Figure 2. All the parameters are defined in Equations (1)–(4) where $d(n)$ is the desired signal.

$$\mathbf{w}(n) = [w_0(n), w_1(n), \dots, w_{M-1}(n)]^T \quad (1)$$

$$\mathbf{u}(n) = [u(n), u(n-1), \dots, u(n-M+1)]^T \quad (2)$$

$$y(n) = \mathbf{w}^H(n-1)\mathbf{u}(n) \quad (3)$$

$$\xi(n) = d(n) - y(n) \quad (4)$$

The adaptive control unit updates the coefficients using the input vector and the prior estimation error. The detailed information can be described in Equation (5):

$$\hat{\mathbf{w}}(n) = \hat{\mathbf{w}}(n-1) + \mathbf{k}(n)\xi(n) \quad (5)$$

where $\mathbf{k}(n)$ is the gain vector that is described in Equation (6):

$$\mathbf{k}(n) = \frac{\lambda^{-1}\mathbf{P}(n-1)\mathbf{u}(n)}{1 + \lambda^{-1}\mathbf{u}^H\mathbf{P}(n-1)\mathbf{u}(n)} \quad (6)$$

where λ is the forgetting factor and $\mathbf{P}(n)$ is a covariance matrix of the noise which can be updated by Equation (7):

$$\mathbf{P}(n) = \lambda^{-1}\mathbf{P}(n-1) - \lambda^{-1}\mathbf{k}(n)\mathbf{u}(n)\mathbf{P}(n-1) \tag{7}$$

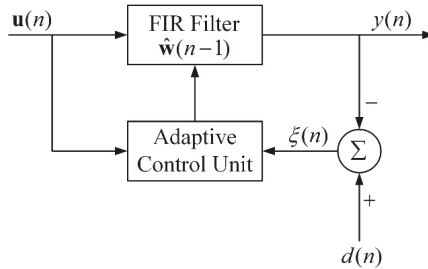


Figure 1. Block diagram of adaptive recursive least squares filter (ARLSF).

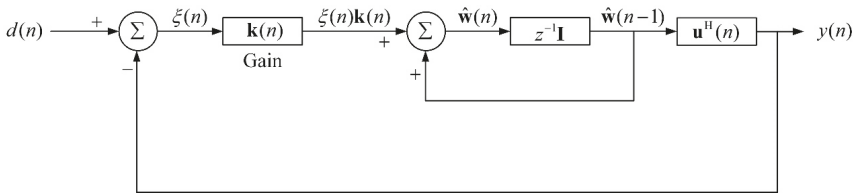


Figure 2. Signal-flow graph of ARLSF.

2.2. Discussion of the Desired Signal

The design of the desired signal is very important in the RLS filter when the desired signal cannot be directly obtained from an aiding sensor. The RLS filter works under the premise that the desired signal is linearly correlated to the input signal and orthogonal to the estimation error. The RLS filter will perform better with the higher linear correlation and the stronger orthogonality mentioned above [18,19].

The desired signal and the estimation error are the motion artifact and the SCG signal, respectively. They are collected from a single accelerometer simultaneously and they are aliased in the frequency domain with different frequency characteristics. The collected acceleration data contains the heartbeat signal, the motion artifact, the respiratory component, the noise of the hardware system and sounds from the other organs. The frequency of the respiratory component is less than 1 Hz [20,21], and the frequency of the heartbeat signal can be up to 25 Hz [22,23]. As a result, the collected acceleration data is band-pass filtered from 1 to 25 Hz with a 32th order FIR filter to remove the gravity component, the respiratory component and the high frequency noise, and the filtered data composed of the expected SCG signal and motion artifact within a frequency range from 1 to 25 Hz can be set as the input of ARLSF.

In addition, the maximum heart rate of a healthy adult is less than 210 beats per minute (bpm) [24–26] which illustrate the maximum frequency of an adult’s heart rate is 3.5 Hz. In order to further analyze the frequency of the motion signal and the heartbeat signal, a similar experiment has been conducted to record the motion signal and the heartbeat signal [9]. Firstly, the SCG recorder system is attached to the chest wall of the subject to record the heartbeat signal under the condition of stand-up without moving as shown in Figure 3a; z-axis acceleration data is collected at a sampling rate of 800 Hz for about 120 s. Afterward, the same SCG recorder system is placed at the right side of the back of the same subject to record the motion artifact at a sampling rate of 800 Hz as shown in Figure 3b. The subject is asked to walk on a treadmill that works at a low speed (3–5 km/h) and the tri-axis acceleration data are collected for about 120 s. Four continuous 60-second signals are

selected from the middle of each recorded acceleration data respectively and they are analyzed in time and frequency domain. Figure 4a illustrates that the heart rate frequency (1.2 Hz) lies in the low-frequency range (<3.5 Hz) while high frequency heart sound signal lies in the range from 4 Hz to 25 Hz. An obvious low-frequency point (1.82 Hz) representing the footsteps frequency is marked in Figure 5b, and the high-frequency component of the motion concentrates on the range from 4 Hz to 10 Hz [15,27]. It can be observed that the heartbeat signal and the motion signal are overlapped in the frequency domain and cannot be separated by bandpass filters.

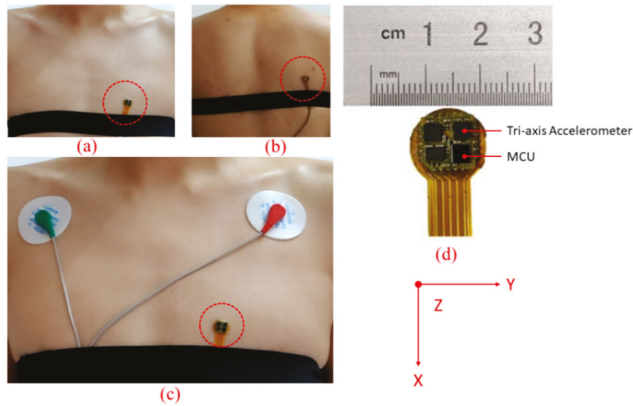


Figure 3. (a,b) Seismocardiography (SCG) recorder system placement for measuring heartbeat signal and motion signals, respectively. (c) A pair of electrocardiogram (ECG) lead I electrodes placement and SCG recorder system placement. (d) An image of the SCG recorder system which shows the dimensions. X-axis, y-axis and z-axis describe the head to foot, shoulder to shoulder and dorsoventral direction, respectively.

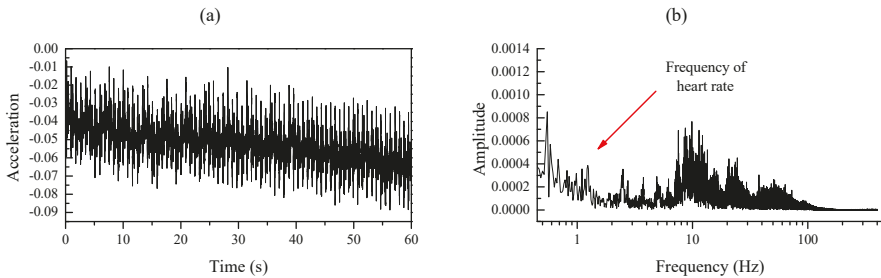


Figure 4. (a,b) Time plot and frequency plot of heartbeat signal respectively.

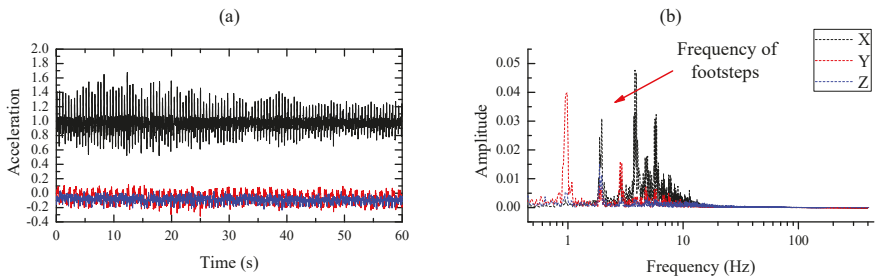


Figure 5. (a,b) Time plot and frequency plot of the tri-axis motion signal respectively.

Further process procedures including bandpass filtering and correlation analysis are performed on the z-axis of the motion signal and the heartbeat data. The z-axis of the motion signal and heartbeat signal are bandpass filtered from 1 Hz to 25 Hz, and the filtered signals are plotted in Figure 6a,c respectively. Figure 6b represents the filtered signal of the z-axis of the motion signal through the FIR bandpass filter of the same type as Figure 6a,c, but at different cutoff frequencies from 3.5 Hz to 25 Hz. Meanwhile, the correlation coefficient between signals in Figure 6a,b is calculated to be 0.99978 and the counterpart between signals in Figure 6a,c is 0.15374, which proved well the high linear correlation of the signals in Figure 6a,b, and the strong orthogonality of the signals in Figure 6a,c [28]. Therefore, the desired signal can be obtained by bandpass filtering the recorded acceleration signal from 3.5 to 25 Hz.

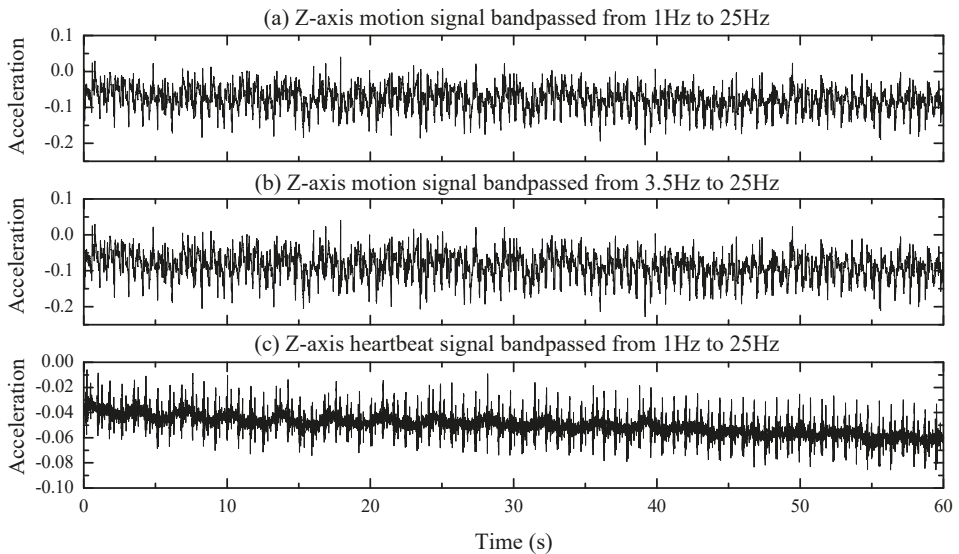


Figure 6. (a) Z-axis motion signal band passed from 1 to 25 Hz. (b) The Z-axis motion signal band passed from 3.5 to 25 Hz. (c) The Z-axis heartbeat signal band passed from 1 to 25 Hz.

2.3. Discussion of the Forgetting Factor

The forgetting factor plays an important role in the behavior of the RLS algorithm under non-stationary conditions. In a classical RLS algorithm, the value of the forgetting factor is fixed between 0 and 1. For the value of the forgetting factor closer to 0, the RLS algorithm has not only a smaller memory length and fast-tracking ability but also a reduced convergence speed and stability. On the other hand, when the forgetting factor is closer to 1, the RLS algorithm has fast convergence and good stability, but the tracking ability suffers and the memory length becomes longer [18,19]. In order to meet the conflicting requirements in non-stationary conditions, the forgetting factor is set between 0.98 and 1.0 [29,30]. In theory, the optimal value of the forgetting factor in non-stationary conditions can be defined by Equation (8) [19]:

$$\lambda \approx 1 - \frac{1}{\sigma_v} \left(\frac{\text{tr}[R_\omega]}{\text{tr}[R_u^{-1}]} \right)^{1/2} \quad (8)$$

where σ_v^2 , R_ω , R_u are the measurement noise variance, process noise correlation matrix, and input vector correlation matrix respectively, $\text{tr}[\cdot]$ denotes the trace of the matrix.

3. Measurement Technique

3.1. Hardware System

Figure 3d shows the prototype of the SCG Recorder System (SRS), which integrates a commercial tri-axis accelerometer (ICM-20602 manufactured by InvenSense) and a microprocessor (STM32F411CEY6 manufactured by STMicroelectronics). The size of the device is less than 1 cm². The tri-axis accelerometer is used to capture acceleration data including SCG signal and motion information within a range of $\pm 2g$. The micro controller unit (MCU) collects the acceleration data from the accelerometer via a serial peripheral interface (SPI) at a rate of 800 Hz. The SRS is attached to the chest wall and placed at the left of the sternum as shown in Figure 3c. The z-axis of SRS describes the dorsoventral direction of the subject, and the x-axis and y-axis describe the head to foot direction and the shoulder to shoulder direction respectively.

In addition to the acceleration data captured by the tri-axis accelerometer, a standard ECG system simultaneously collects a standard ECG lead I signal at a rate of 512 Hz. The two electrodes are placed at the right arm and the left arm, respectively, as shown in Figure 3c. Both the SCG Recorder System and ECG system are connected to a host PC via serial cables for data transmission and synchronization.

3.2. Experiment Setup

The hardware system described above was used on sixteen subjects whose ages ranged from 24 to 35. The experiment was conducted on a treadmill and the subjects were asked to keep standing for at least 120 s before walking at least 180 s. After walking, the subjects stood for at least 60 s. The subjects could breathe freely during the whole experiment. The walking speed of the subjects was limited to less than 1.5 m/s by setting the treadmill at a low speed (3–5 km/h). The SCG and ECG signals including sampling time and sensor data were collected and transmitted to the PC for further analysis in MATLAB (R2016a).

3.3. Software System

MATLAB (R2016a) was used to analyze all the data. The processing procedure consists of three parts: signal preprocessing to obtain the primary and reference channel of the ARLSF, the ARLSF and feature extraction.

3.3.1. Signal Preprocessing

The collected acceleration data is a mixed collection of signals in both the time and frequency domains, which contain the heartbeat signal, the motion artifact, the respiratory component, the noise of the hardware system and sounds from the other organs.

Firstly, the collected acceleration data is bandpass filtered from 1 to 25 Hz with a 32th order FIR filter to remove the gravity component, the respiratory component and the high-frequency noise. This procedure leaves the expected SCG signal and motion artifact within a frequency range from 1 to 25 Hz in the filtered signal, which is fed to the primary channel of the ARLSF.

The second filtered signal is obtained by bandpass filtering the acceleration data from 3.5 to 25 Hz with the same order and type of FIR filter. The correlation coefficient between the filtered signal of the primary channel and the second filter signal is calculated to be 0.987 which proved well the high linear correlation of these two filtered signals as discussed in Section 2.2. Therefore, the second filtered signal can be fed into the reference channel of the ARLSF. The raw data of the collected acceleration data, the primary channel obtained, and the reference channel are plotted in Figure 7a–c, respectively.

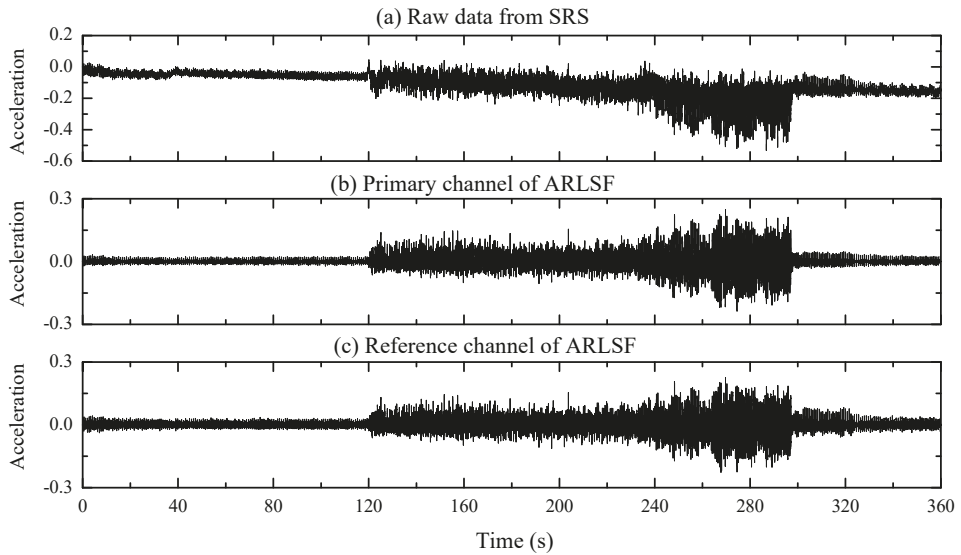


Figure 7. (a) Raw data from SCG Recorder System (SRS). (b,c) The primary and reference channel of ARLSF respectively.

3.3.2. ARLSF

The same order and same type of FIR filter used in signal preprocessing section creates the same delay to both the primary channel and the reference channel. As a result, the primary channel and the reference channel are synchronized in the time domain. In addition, the forgetting factor is calculated to be 0.9908 based on Equation (8). The variable $\xi(n)$ described in Equation (4) is the estimated heartbeat signal.

3.3.3. Feature Extraction

Features extracted from the ECG signal and filtered SCG signal are R peaks and aortic valve opening (AO) peaks, respectively. R peaks can be extracted from the ECG signal using the classical Pan Tompkin algorithm [31], while the extraction algorithm for AO peaks is different. The formula is described in Equations (9) and (10):

$$[loc_min, val_min] = \min(\xi(t - 0.3 : t)) \quad (9)$$

$$[loc_max, val_max] = \max(\xi(t - 0.3 : t)) \quad (10)$$

The estimated heartbeat signal $\xi(n)$ is divided into several segments with a length of 0.3s [32]. $\xi(t - 0.3 : t)$ is the segment at time t and loc_min , val_min , loc_max and val_max are the timestamp of the minimum value, the minimum value, the timestamp of maximum value and the maximum value of the segment at time t respectively. In addition, some constraints are used to avoid incorrect extracted AO peaks:

$$\begin{cases} val_min < -0.007 \\ val_max > 0.01 \\ |loc_max - loc_min| < 0.02 \end{cases} \quad (11)$$

A continuous 0.3 s segment including a correct AO peak is picked from the estimated heartbeat signal and plotted in Figure 8. val_max represents the magnitude of the AO peak and val_min represents the magnitude of the successive isovolumic moment (IM) peak or maximum acceleration (MA) peak.

Constant values -0.007 and 0.01 in Equation (11) are the maximum amplitude of the IM envelope peaks and minimum amplitude of the AO envelope peaks respectively [32]. In addition, the time interval between val_max and val_min should be restricted shorter than $0.02s$ [33]. When loc_min , val_min , loc_max and val_max meet the restrictions in Equation (11), loc_max and val_max are the timestamp and the value of the extracted AO peak respectively. For further analysis, the heart rate can be calculated from the R–R and AO–AO interval.

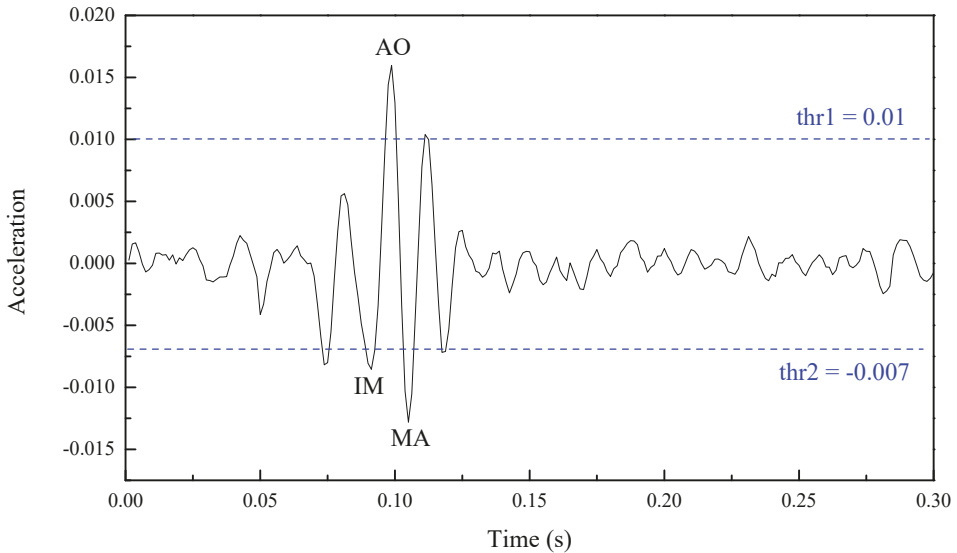


Figure 8. A correct aortic valve opening (AO) peak from a continuous 0.3s segment.

4. Results

Figure 9 displays the raw data and processing results. Figure 9a shows the raw acceleration data collected from the SRS. The data show that the heartbeat signals are contaminated by the motion artifact, and the features and graphs of the heartbeat signals cannot be identified during the walking period. Figure 9b,c show the extracted heartbeat signals using Savitzky Golay-based polynomial smoothing [14] and ARLSF. Features and graphs of heartbeat signals are not clear from 120 s to 300 s in Figure 9b. It can be observed that the proposed ARLSF outperforms the Savitzky Golay-based polynomial smoothing. For better visualization, six segmented signals selected from the standing period before walking, transitions from standing to walking, the first half of walking, the second half of walking, transitions from walking to standing, and standing after walking are plotted in Figure 10a–f respectively. It is obvious that the features and graphs of the heartbeat signals are visually noticeable after ARLSF. Thus, any other signal-processing procedures are unnecessary on the heartbeat signals.

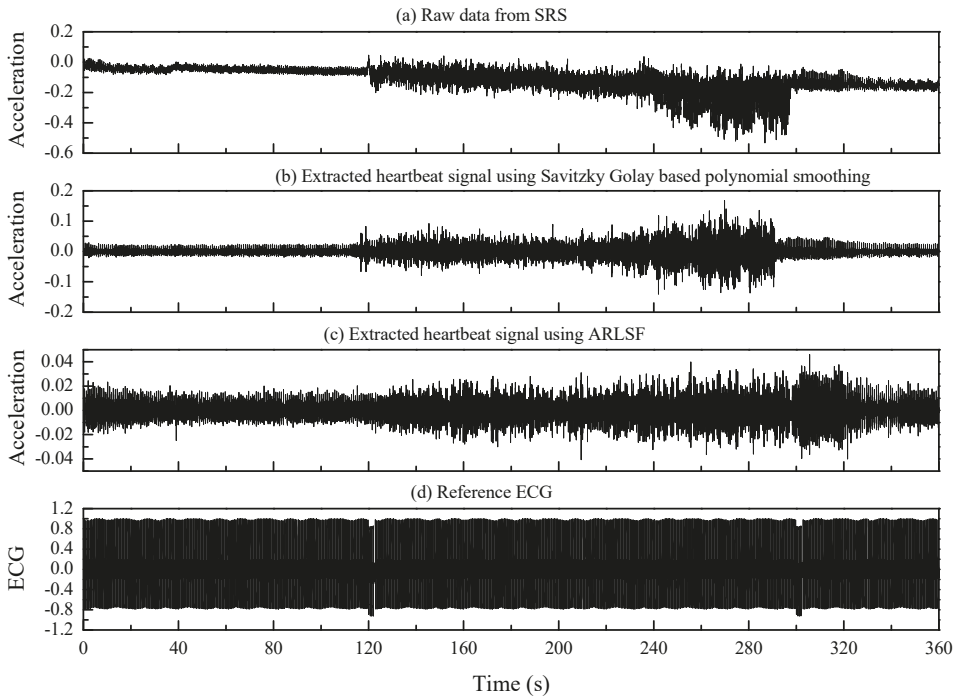


Figure 9. (a) Raw data from SRS. (b) Extracted heartbeat signal using Savitzky Golay filter. (c) Extracted heartbeat signal using ARLSF. (d) The reference ECG lead I signal. (The units of acceleration and ECG are g and mV, respectively.)

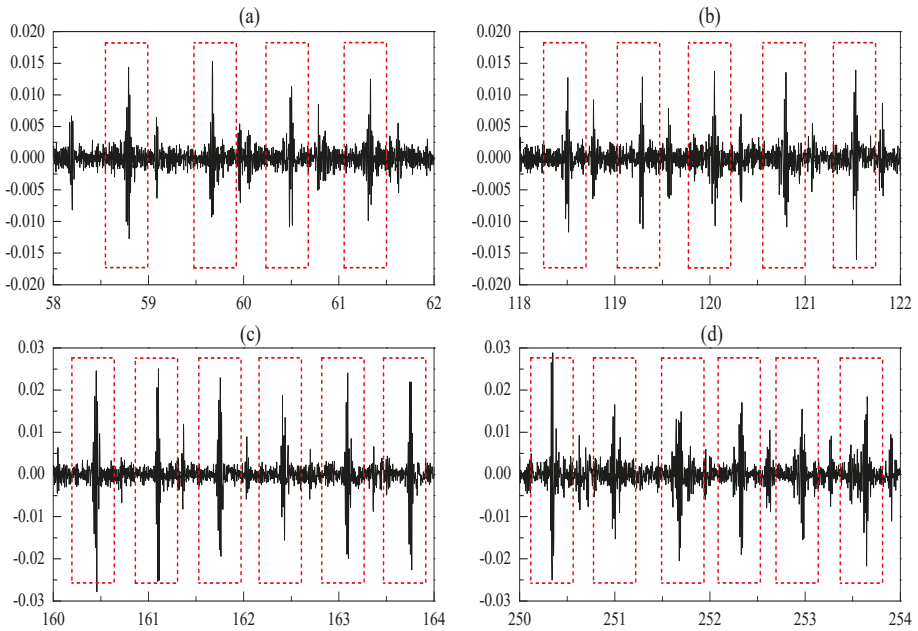


Figure 10. Cont.

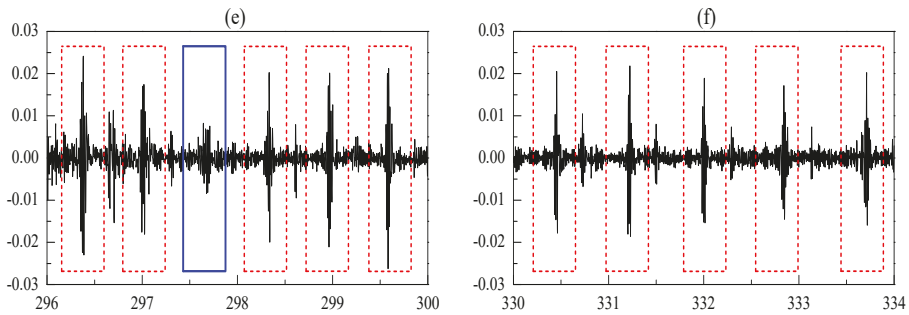


Figure 10. (a–f) Segmented signal during standing before walking, transitions from standing to walking, the first half of walking, the second half of walking, transitions from walking to standing, standing after walking. (The x -axis and y -axis are time in seconds and acceleration in gravity respectively.)

4.1. Heartbeat Detection Accuracy

The standard ECG lead I recordings are the baseline references for heartbeat signals. Heartbeat detection accuracy is defined as the ratio of the detected SCG peaks divided by the detected ECG peaks. The detected SCG peaks are visually noticeable and marked with the red dotted line in Figure 11. In addition, the missing SCG peaks are also considered to give a more precise estimate of detection accuracy. These peaks represent the undetected and false-positive detected SCG peaks. The blue rectangular, which is marked in Figure 10, can be considered as an undetected SCG peak using the rule of feature extraction described previously. This method is used even though the features and graphs of the signal are visually noticeable. Numbers in the brackets represent the missing peaks that occur at the beginning and end of walking in Table 1. It can be observed that above 60% of the missing SCG peaks occur at the beginning and end of walking when MA changes rapidly. The detection rates, which are shown in Table 1, are higher than 98%.

Table 1. Heartbeat detection accuracy.

Subject No.	ECG Peaks Detected	SCG Peaks Detected	SCG Peaks Missing	Accuracy
1	483	478	5(2)	98.9%
2	475	468	7(4)	98.5%
3	472	468	4(3)	99.1%
4	480	475	5(3)	98.9%
5	488	483	5(3)	98.9%
6	478	473	5(2)	98.9%
7	492	483	9(6)	98.1%
8	501	496	5(2)	99.0%
9	495	490	5(2)	98.9%
10	490	486	4(1)	99.1%
11	477	471	6(5)	98.7%
12	486	483	3(2)	99.3%
13	496	491	5(3)	98.9%
14	491	489	2(2)	99.5%
15	485	477	8(4)	98.3%
16	482	474	8(6)	98.3%

4.2. Heart Rate Estimations

The detected SCG peaks and ECG peaks evaluate the heartbeat signals from the perspective of the signal graph without considering the correctness. Heart rate can be an effective factor to verify the correctness of the detected peaks. As a gold standard in the clinical field, the heart rate estimated from the detected ECG peaks can be considered as the reference for the ground truth. Figure 11 illustrates the heart rates estimated from the detected SCG peaks and ECG peaks. These peaks are marked as black stars and red points, respectively. The x-axis represents the time that can be divided into three regions: standing, walking and return to standing. It can be observed that the heart rate is stable and low while standing, increases and then stabilizes at a relatively high level during walking, and then decreases to a stable and similar low magnitude when the subject stands again. The detailed difference between the heart rates from ECG and SCG with a mean value of 0.08 bpm and a standard deviation value of 2.08 bpm is plotted in Figure 12. It can be observed that the heart rates estimated from the detected SCG peaks and ECG peaks match very well.

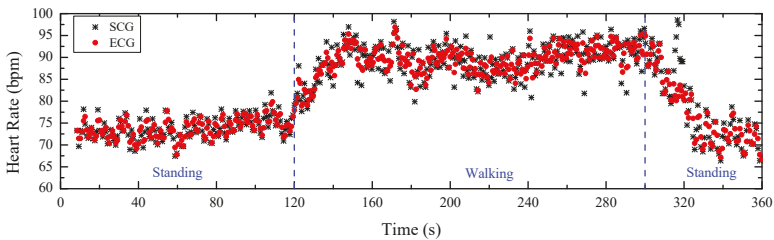


Figure 11. Heart rates estimated from ECG and SCG.

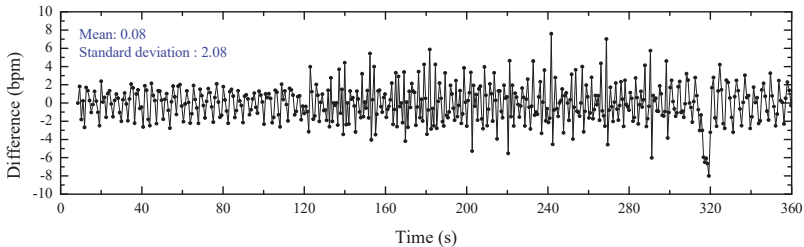


Figure 12. Heart rate difference between ECG and SCG.

4.3. Bland–Altman Analyzation

To further analyze the agreement between the heart rates estimated from the detected SCG peaks and ECG peaks, a Bland–Altman plot [34] is used and shown in Figure 13. Based on the definition of a Bland–Altman plot, the x-axis represents the average heart rates from SCG and ECG, and the y-axis represents their differences. A 95% confidence region is marked with blue dashed lines that have an upper threshold of 4.9 and a lower threshold of -4.6 . It is observed that there are a few outliers, but overall most measurements lie in the 95% confidence region.

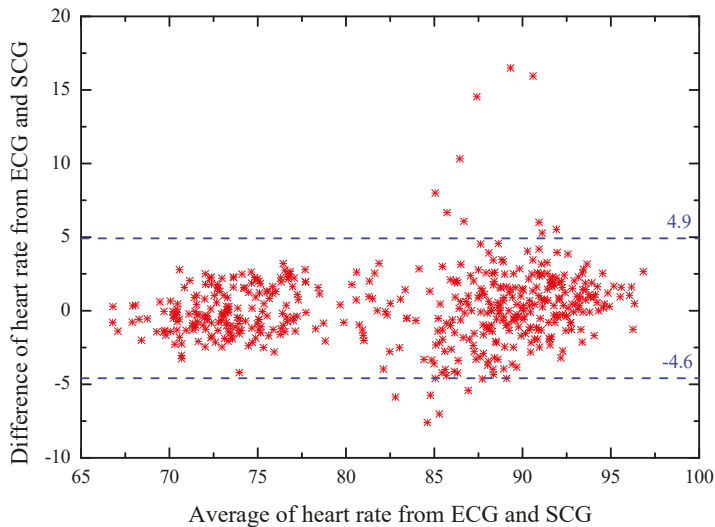


Figure 13. Bland–Altman plot of the heart rate measurements from ECG and SCG.

5. Discussion and Conclusions

In this paper, we proposed a novel method based on an adaptive recursive least squares filter to remove the motion artifact of the acceleration data recorded by only one accelerometer. The primary channel of ARLSF containing the heartbeat signal and motion artifact signal within a frequency range of 1 to 25 Hz was obtained by bandpass filtering the recorded acceleration data from 1 Hz to 25 Hz. Then the same acceleration data was bandpass filtered from 3.5 to 25 Hz to remove the heart rate frequency and footsteps frequency component from the heartbeat and motion artifact signals, respectively. The filtered signal and the signal of the primary channel were proved to have high correlation, and the heartbeat and motion artifact signal were proved to have strong orthogonality, which proved well that the filtered data could be fed into the reference channel of the ARLSF. The heartbeat signal graph of the extracted SCG signal was very clear without the need for any other signal recovery procedures. The heartbeat detection accuracy was up to 98% and heart rates estimated from the SCG and ECG matched well under both the standing and walking conditions.

At present these results are limited by a particularly rapidly changing motion artifact which will lead to missing SCG peaks' detection. The forgetting factor is set at a fixed value at a balance of the convergence, stability and tracking ability which leads to good convergence and stability but poor tracking ability of the proposed ARLSF. Moreover, the forgetting factor calculated by Equation (8) was a globally optimal solution but not a locally optimal solution especially when motion artifact changes rapidly. To obtain better performance, future work on the ARLSF proposed in this paper will focus on the optimization of the forgetting factor by using an adaptive method. That method will change the forgetting factor in real time according to the changes of the motion and promote the performance of ARLSF.

In addition, the experimental setup will be improved in future work. More subjects with a wider range of age and different health conditions, and more dynamic conditions including jumping and running, will be considered to evaluate the filter performance. In order to simplify the sensor placement before experiment and sensor data collection during the experiment, a miniature system-integrated SCG sensor, ECG sensor, MCU-embedded real-time filter algorithm and Bluetooth module will be developed in the future.

Author Contributions: S.Y. proposed the signal processing algorithm, developed the homemade prototype of SCG recorder system, designed the experiments and wrote this paper; S.L. gave some valuable suggestions and revised the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Hubei Provincial Major Program of Technological Innovation (Grant No.2017AAA121).

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

References

1. Bozhenko, B.S. Seismocardiography—a new method in the study of functional conditions of the heart. *Ter Arkh* **1961**, *33*, 55–64.
2. Salerno, D.M.; Zanetti, J. Seismocardiography for monitoring changes in left ventricular function during ischemia. *Chest J.* **1991**, *100*, 991–993. [[CrossRef](#)]
3. Inan, O.T.; Migeotte, P.F.; Park, K.S.; Etemadi, M.; Tavakolian, K.; Casanella, R.; Zanetti, J.; Tank, J.; Funtova, L.; Prisk, G.K.; et al. Ballistocardiography and seismocardiography: A review of recent advances. *IEEE J. Biomed. Health Inform.* **2014**, *19*, 1414–1427. [[CrossRef](#)]
4. Taebi, A.; Solar, B.E.; Bomar, A.J.; Sandler, R.H.; Mansy, H.A. Recent Advances in Seismocardiography. *Vibration* **2019**, *2*, 5. [[CrossRef](#)]
5. Lee, H.; Lee, H.; Whang, M. An enhanced method to estimate heart rate from seismocardiography via ensemble averaging of body movements at six degrees of freedom. *Sensors* **2018**, *18*, 238. [[CrossRef](#)]
6. Ng, J.; Sahakian, A.V.; Swiryn, S. Accelerometer-based body- position sensing for ambulatory electrocardiographic monitoring. *Biomed Instrum. Technol.* **2003**, *37*, 338–346.
7. Yoon, S.; Min, S.; Yun, Y.; Lee, S.; Lee, M. Adaptive motion artifacts reduction using 3-axis accelerometer in e-textile ecg measurement system. *J. Med. Syst.* **2008**, *2*, 101–106. [[CrossRef](#)]
8. Liu, S.H. Motion artifact reduction in electrocardiogram using adaptive filter. *J. Med. Bio Log. Eng.* **2011**, *31*, 67–72. [[CrossRef](#)]
9. Yang, C.; Tavassolian, N. An Independent Component Analysis Approach to Motion Noise Cancellation of Cardio-Mechanical Signals. *IEEE Trans. Biomed. Eng.* **2019**, *3*, 66. [[CrossRef](#)]
10. Yang, C.; Tavassolian, N. Motion noise cancellation in seismocardiogram of ambulant subjects with dual sensors. In Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society, Orlando, FL, USA, 16–20 August 2016; pp. 5881–5884.
11. Taebi, A.; Mansy, H.A. Noise cancellation from vibrocardiographic signals based on the ensemble empirical mode decomposition. *J. Appl. Biotechnol. Bioeng.* **2017**, *2*, 49–54. [[CrossRef](#)]
12. Javaid, A.Q.; Ashouri, H.; Dorier, A.; Etemadi, M.; Heller, J.A.; Roy, S.; Inan, O.T. Quantifying and Reducing Motion Artifacts in Wearable Seismocardiogram Measurements During Walking to Assess Left Ventricular Health. *IEEE Trans. Biomed. Eng.* **2017**, *64*, 1277–1286. [[CrossRef](#)]
13. Di Rienzo, M.; Meriggi, P.; Vaini, E.; Castiglioni, P.; Rizzo, F. 24h seismocardiogram monitoring in ambulant subjects. In Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, USA, 28 August–1 September 2012; pp. 5050–5053.
14. Pandia, K.; Ravindran, S.; Cole, R.; Kovacs, G.; Giovangrandi, L. Motion artifact cancellation to obtain heart sounds from a single chest-worn accelerometer. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, Dallas, TX, USA, 14–19 March 2010; pp. 590–593.
15. Kumar Jain, P.; Kumar Tiwari, A. A novel method for suppression of motion artifacts from the seismocardiogram signal. In Proceedings of the IEEE International Conference on Digital Signal Processing, Beijing, China, 16–18 October 2016; pp. 6–10.
16. Choudhary, T.; Sharma, L.N.; Bhuyan, M.K. Heart Sound Extraction from Sternal Seismocardiographic Signal. *IEEE Signal Process. Lett.* **2018**, *25*, 482–486. [[CrossRef](#)]
17. Yang, C.; Tavassolian, N. Motion artifact cancellation of seismocardiographic recording from moving subjects. *IEEE Sens. J.* **2016**, *16*, 5702–5708. [[CrossRef](#)]
18. Diniz, P.S.R. *Adaptive Filtering: Algorithms and Practical Implementation*, 4th ed.; Springer: New York, NY, USA, 2013; pp. 209–248.
19. Haykin, S.O. *Adaptive Filter Theory*, 5th ed.; Pearson: Upper Saddle River, TX, USA, 2013; pp. 449–472.

20. Reinvuol, T.; Hannula, M.; Sorvoja, H.; Alasaarela, E.; Myllyla, R. Measurement of respiratory rate with high-resolution accelerometer and emfit pressure sensor. In Proceedings of the IEEE Sensors Applications Symposium, Houston, TX, USA, 22–25 October 2006; pp. 192–195.
21. Morillo, D.S.; Ojeda, J.L.R.; Foix, L.F.P.; Jimenez, A.L. An accelerometer-based device for sleep apnea screening. *IEEE Trans. Inf. Technol. Biomed.* **2010**, *14*, 491–499. [CrossRef]
22. Marcus, F.; Sorrell, V.; Zanetti, J.; Bosnos, M.; Baweja, G.; Perlick, D.; Ott, P.; Indik, J.; He, D.S.; Gear, K. Accelerometer-derived time intervals during various pacing modes in patients with biventricular pacemakers: Comparison with normals. *PACE* **2007**, *30*, 1476–1481. [CrossRef]
23. Zanetti, J.M.; Tavakolian, K. Seismocardiography: Past, present and future. In Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society, Osaka, Japan, 3–7 July 2013; pp. 7004–7007.
24. Roskamm, H. Optimum patterns of exercise for healthy adults. *Can. Med. Assoc. J.* **1967**, *96*, 895–900.
25. Robergs, R.A.; Landwehr, R. The surprising history of the "HRmax=220-age" equation. *J. Exerc. Physiol. Online* **2002**, *5*, 1–10.
26. Tanaka, H.; Monahan, K.D.; Seals, D.R. Age-predicted maximal heart rate revisited. *J. Am. Coll. Cardiol.* **2001**, *37*, 153–156.
27. Pachi, A.; Ji, T. Frequency and velocity of people walking. *Struct. Eng.* **2005**, *83*, 36–40.
28. Ratner, B. The correlation coefficients: Its values range between or do they? *J. Target. Meas. Anal. Mark.* **2009**, *17*, 139–142. [CrossRef]
29. Suresh, H.N. New Technique for Recursive Least Square Adaptive Algorithm for Acoustic Echo Cancellation of Speech signal in an auditorium. *Int. Research J. Eng. Technol.* **2015**, *36*, 110–117.
30. LabVIEW 2013 System Identification Toolkit Help. Available online: http://zone.ni.com/reference/en-XX/help/372458D-01/lvsysidconcepts/recursive_ls/ (accessed on 28 February 2020).
31. Pan, J.; Tompkins, W.J. A Real-Time QRS Detection Algorithm. *IEEE Trans. Biomed. Eng.* **1985**, *32*, 230–236. [CrossRef] [PubMed]
32. Khosrow, K.F.; Tavakolian, K.; Menon, C. Moving toward automatic and standalone delineation of seismocardiogram signal. In Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015; pp. 7163–7166.
33. Sørensen, K.; Schmidt, S.E.; Jensen, A.S. Definition of Fiducial Points in the Normal Seismocardiogram. *Sci. Rep.* **2018**, *8*, 15455. [CrossRef] [PubMed]
34. Bland, J.M.; Altman, D.G. Measuring agreement in method comparison studies. *Stat. Methods Med. Res.* **1999**, *8*, 135–160. [CrossRef] [PubMed]



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).

Article

Machine-Learning Analysis of Voice Samples Recorded through Smartphones: The Combined Effect of Ageing and Gender

Francesco Asci ^{1,†}, Giovanni Costantini ^{2,†}, Pietro Di Leo ², Alessandro Zampogna ¹, Giovanni Ruoppolo ³, Alfredo Berardelli ^{1,4}, Giovanni Saggio ² and Antonio Suppa ^{1,4,*}

¹ Department of Human Neurosciences, Sapienza University of Rome, 00185 Rome, Italy; francesco.asci@uniroma1.it (F.A.); alessandro.zampogna@uniroma1.it (A.Z.); alfredo.berardelli@uniroma1.it (A.B.)

² Department of Electronic Engineering, University of Rome Tor Vergata, 00133 Rome, Italy; costantini@uniroma2.it (G.C.); pietero.dileo@alumni.uniroma2.eu (P.D.L.); saggio@uniroma2.it (G.S.)

³ Department of Sense Organs, Otorhinolaryngology Section, Sapienza University of Rome, 00185 Rome, Italy; giovanni.ruoppolo@uniroma1.it

⁴ IRCCS Neuromed, 86077 Pozzilli (IS), Italy

* Correspondence: antonio.suppa@uniroma1.it; Tel.: +39-06-49914544

† These authors have equally contributed to the manuscript.

Received: 20 July 2020; Accepted: 2 September 2020; Published: 4 September 2020

Abstract: Background: Experimental studies using qualitative or quantitative analysis have demonstrated that the human voice progressively worsens with ageing. These studies, however, have mostly focused on specific voice features without examining their dynamic interaction. To examine the complexity of age-related changes in voice, more advanced techniques based on machine learning have been recently applied to voice recordings but only in a laboratory setting. We here recorded voice samples in a large sample of healthy subjects. To improve the ecological value of our analysis, we collected voice samples directly at home using smartphones. Methods: 138 younger adults (65 males and 73 females, age range: 15–30) and 123 older adults (47 males and 76 females, age range: 40–85) produced a sustained emission of a vowel and a sentence. The recorded voice samples underwent a machine learning analysis through a support vector machine algorithm. Results: The machine learning analysis of voice samples from both speech tasks discriminated between younger and older adults, and between males and females, with high statistical accuracy. Conclusions: By recording voice samples through smartphones in an ecological setting, we demonstrated the combined effect of age and gender on voice. Our machine learning analysis demonstrates the effect of ageing on voice.

Keywords: ageing; gender; machine learning; support vector machine; voice analysis

1. Introduction

Human voice represents a complex biological signal resulting from the dynamic interaction of vocal folds adduction/vibration with pulmonary air emission and flow through resonant structures [1]. Physiologic ageing leads to specific changes in the anatomy and physiology of all structures involved in the production and modulation of the human voice [2–14]. Hence, a possible approach to evaluate the effect of physiological ageing in humans would include the analysis of voice.

Early seminal studies aimed to characterize age-related changes in voice have used qualitative tools consisting of a perceptual examination of voice recordings [3]. These studies have demonstrated that physiologic ageing induces a variable combination of effects on voice including reduced intensity and phonation time, and a general worsening of voice quality due to hoarseness and vocal fatigue [1,15–17].

Some authors have also used more advanced quantitative tools for recording and analyzing voice and thus for achieving an objective examination of age-related changes of voice [1]. Objective voice analysis commonly includes several acoustic parameters calculated in the time-domain such as the jitter, the shimmer, the signal to noise ratio (SNR) and the harmonic to noise ratio (HNR) [18] or spectral analysis measures calculated in the frequency-domain such as the fundamental frequency (fo) [19,20]. More recently, cepstral analysis has been recognized as a methodologic evolution of the spectral analysis resulting from a mathematical transformation from the domain of frequency to quefrency. The cepstral analysis allows for calculating innovative variables such as the cepstral prominence peak smoothed (CPPs) [21,22]. Spectral and cepstral analyses have demonstrated that physiological ageing induces changes in several voice parameters including the fo, the SNR, the HNR, and finally the CPPs [1,20,23]. However, although spectral/cepstral analysis allows measuring age-related changes in specific voice features, it failed to provide a detailed examination of the complex and dynamic interaction of voice features which characterize the physiologic ageing of voice [1,23].

The most recent approach used to assess physiologic ageing in healthy subjects consists of the objective voice analysis based on machine learning algorithms [24–28]. Machine-learning is a novel and robust method commonly applied to classify complex variables obtained from large datasets [29–31]. More in detail, machine learning can be applied to predict outcomes from recurring patterns of features within various types of multidimensional data sets [32]. Several authors have applied automatic classifiers based on machine learning analysis on voice recordings to classify healthy subjects according to their age and gender [24–28,33–38]. More recently, to further improve the overall accuracy of the machine learning analysis, several studies have included an increasing number of voice features in the datasets [24–28] and compared the performance of different machine learning algorithms [37,38].

In this study, we examined the combined effect of the age- and gender-related factors on voice features through machine learning. Also, previous studies have not compared the performances of the machine learning analysis of voice samples obtained during the sustained emission of a vowel or a sentence, by using the receiver operating characteristic (ROC) curve. So far voice samples have been only collected in a laboratory setting by using dedicated technological instruments consisting of hi-tech audio recorders which require expert supervision [1]. Currently available smartphones and information technology (IT) services have allowed to record and analyze a large number of health parameters in free-living scenarios [39]. The use of a smartphone to record high-quality voice samples would simplify the procedures of recordings, allowing to acquire and analyze a large amount of data. Further advantages of doing recordings using smartphone consist of the building up of a more ecologic scenario compared to the laboratory setting, thus helping to overcome possible voice changes due to supervised conditions.

In this cross-sectional study, we collected voice samples recorded through smartphones in two independent groups of healthy participants with different ages. We used machine learning algorithms to investigate the effect of physiologic ageing on voice. To evaluate the combined effect of age and gender on voice, we also examined the voice samples recorded by females and males from different ages, using machine learning. To verify whether age-related changes of the voice depends on specific speech tasks, we examined and compared the voice recordings during the sustained emission of a vowel and a sentence. All analyses included ROC curves and a detailed description of the statistical output including accuracy, sensibility, specificity, and area under the curve (AUC).

2. Materials and Methods

2.1. Subjects

We recruited an overall group of 261 healthy subjects (HS) (112 males and 149 females; mean age \pm SD 41.0 \pm 18.7 years, range 15–85). Subjects were then divided into two independent sex-matched groups according to age: younger adults (YA) (number 138; 65 males and 73 females; mean age \pm SD 25.1 \pm 3.1 years, range 15–30), and older adults (OA) (number 123; 47 males and

76 females; mean age \pm SD 58.9 ± 11.0 years, range 40–85). All the participants were recruited at the Department of Human Neurosciences, Sapienza University of Rome, Italy. All subjects were non-smokers, native Italian-speakers. Participants did not manifest cognitive or mood impairment nor bilateral/unilateral hearing loss, respiratory disorders, and other disorders affecting the vocal cords. Also, participants did not manifest gastro-esophageal reflux disease, acute or chronic gastritis, or other gastrointestinal disorders possibly affecting the emission of the voice. At the time of the study, all the YA completed the pubertal development. Participants took no drugs acting over the central nervous system at the time of the study. Participant demographic features are summarized in Table 1 and reported in detail in Supplementary Materials Tables S1 and S2. Participants gave consent to the study, which was approved by the institutional review board following the Declaration of Helsinki.

Table 1. Demographic and clinical characteristics of the participants.

Group	Age (years)	Weight (Kg)	Height (cm)	BMI
YA	25.1 \pm 3.1	64.5 \pm 12.4	171.4 \pm 8.5	21.8 \pm 3.1
OA	58.9 \pm 11.0	66.9 \pm 11.9	166.5 \pm 9.8	25.2 \pm 4.1
YA ₂₅	22.9 \pm 2.2	61.4 \pm 10.3	171.0 \pm 8.1	20.9 \pm 2.5
OA ₅₅	66.4 \pm 8.1	68.6 \pm 11.9	163.0 \pm 9.1	25.8 \pm 4.3
YA _f	24.7 \pm 3.0	56.5 \pm 7.6	166.2 \pm 5.7	20.5 \pm 2.7
YA _m	25.5 \pm 3.2	73.4 \pm 10.7	177.2 \pm 7.1	23.3 \pm 2.8
OA _f	59.8 \pm 10.5	65.7 \pm 11.3	161.2 \pm 7.3	25.4 \pm 4.7
OA _m	58.1 \pm 11.3	76.4 \pm 9.6	175.0 \pm 6.9	25.0 \pm 3.1

OA: older adult; OA_f: female OA; OA_m: male OA; YA_f: female YA; YA_m: male YA; YO: younger adult; YO₂₅: younger adult \leq 25 years; OA₅₅: older adult \geq 55 years. Values are expressed as average \pm standard deviation.

2.2. Voice Recordings

The recording session started by asking participants to sit on a chair in the middle of a silent room. Subjects were instructed to handle and face a smartphone at about 30 cm from the mouth and then to speak with their usual voice intensity, pitch, and quality. Smartphones currently available in the market (various brands including Apple®, Samsung®, Huawei®, Xiaomi® and Asus®) were used for voice recordings. The recording session consisted of two separate speech-tasks, the former including the sustained emission of a vowel and the latter consisting of a sample of connected-speech. More in detail, patients were first asked to produce the sustained emission of the vowel/e/for 5 s and then to read the following Italian phonetically balanced sentence: “Nella casa in riva al mare maria vide tre cani bianchi e neri.” To simplify the procedures of home-made audio recording, all participants were asked to save the audio tracks in mp4 format at the end of the recording session. Participants were then asked to send voice samples by e-mail to our institutional mail server, which was protected and accessible only by the authors. Lastly, voice recordings were separated in audio tracks containing each of the two speech-tasks, through a segmentation procedure included in dedicated software for audio-editing (Audacity®) [40].

2.3. Machine-Learning Analysis

The machine-learning analysis consisted of specific and standardized algorithms of artificial intelligence [41–44]. We converted all the audio tracks from mp4 into Wav format (sampling frequency: 44.1 kHz; bit depth: 16 bit), before submitting data to OpenSMILE, a dedicated software for the pre-process of feature extraction (OpenSMILE; audEERING GmbH, Munich, Germany) [45]. For each voice sample, 6139 voice features were extracted by using a modified INTERSPEECH2016 Computational Paralinguistics Challenge (IS ComParE 2016) feature dataset [44]. IS ComParE 2016 contains voice features calculated using computational functionals (e.g., mean, quartiles, percentiles, position of max/min, linear regression) over acoustic low-level descriptors (LLDs), including those related to the energy, spectrum, cepstrum of the signal [44,46], and also including the Mel-Frequency Cepstral Coefficients [47,48], RASTA-PLP Coefficients [49], jitter, shimmer, sound quality descriptors,

and prosodic features. Given that the IS ComParE 2016 features dataset does not contain the CPPs, the HNR, and SNR, we additionally extracted these features through specific home-made algorithms (MATLAB, The Math Works, Inc., Version R2020a, Natick, MA, USA, 2020) [21,50,51]. Then, the CPPs, HNR, and SNR were added to the IS ComParE 2016 feature dataset using Wolfram Mathematica (Wolfram Research, Inc., Mathematica, Version 12.1, Champaign, IL, USA, 2020).

To identify a small subset of relevant features for the objective analysis of voice ageing [52], the extracted voice features underwent feature selection using the correlation features selection (CFS) algorithm [53]. Through CFS, we selected voice features highly correlated with the class, thus removing the irrelevant and redundant features from the original dataset. Selected features were ranked by using the correlation attribute evaluation (CAE) algorithm, which evaluates and ranks all the attributes in order of relevance, according to Pearson's correlation method. To further increase the accuracy of results, we applied the Fayyad & Irani's discretization method to the features' values [54]. Discretization is an optimization procedure consisting in modifying the values and the distribution of the features, by calculating the best splitting point from the two classes and assigning a binary value to the features, in two groups.

After pre-processing procedures, we started the machine learning analysis by using the support vector machine (SVM) classifier. To train the SVM, we considered only the first twenty most relevant features ranked by the CAE. This approach was applied to reduce the number of selected features needed to perform the machine learning analysis. Specifically, the SVM was trained using the sequential minimal optimization (SMO) method, which is considered a fast and efficient machine learning algorithm to implement an SVM classifier [55]. All the classifications were made using a 5-or 10-folds cross-validation, depending on the number of the instances (voice samples) contained in the examined dataset. Both the feature selection and the classification were performed by dedicated software that contains a collection of algorithms for data analysis and predictive modelling (Weka, Waikato Environment for Knowledge Analysis, University of Waikato, New Zealand) [53,56]. The experimental procedures are summarized in Figure 1.

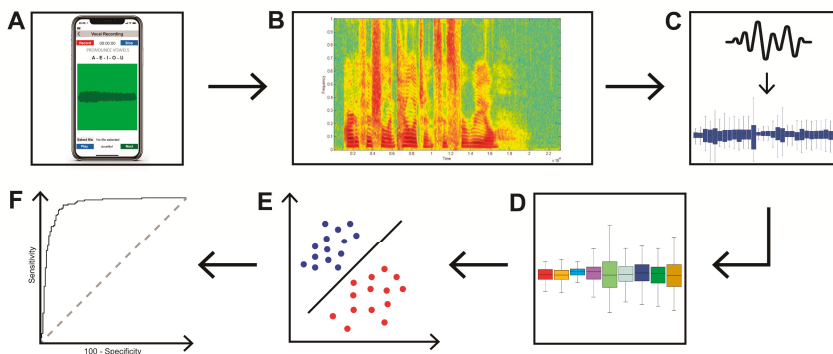


Figure 1. Experimental procedures. (A) Smartphone recording of voice samples of sustained emission of vowel and a sentence. (B) Acoustic voice spectrogram. (C) Procedures of features extraction, (D) features selection, and (E) classification obtained through the SVM. (F) Receiver operating characteristic (ROC) analysis used to perform the statistics.

2.4. Statistical Analysis

The normality of the demographic and anthropometric variables in YA and OA was assessed using the Kolmogorov-Smirnov test. Mann-Whitney U test was used to compare demographic scores in YA and OA. ROC analyses were performed to identify the optimal diagnostic cut-off values of SMO (selected features), calculated during the sustained emission of the vowel as well as during the emission of the sentence, for discriminating between (1) YA and OA; (2) female YA and OA; (3) male YA and

OA; (4) male and female YA and finally; (5) male and female OA. Cut-off values were calculated as the point of the curves with the highest Youden index (sensitivity + specificity – 1) to maximize the sensitivity and specificity of the diagnostic tests. The positive and negative predictive values were also calculated. According to standardized procedures [57], we compared the area under the curves (AUCs) in the ROC curves calculated from SMO (selected features) to verify the optimal test for discriminating within the subgroups. All ROC analyses were performed using WEKA and Wolfram Mathematica. $p < 0.05$ was considered statistically significant. Unless otherwise stated, all values are presented as mean \pm standard deviation (SD). Statistical analyses were performed using Statistica version 10 (StatSoft, Inc) and Wolfram Mathematica.

2.5. Data Availability

The anonymized database used in the current study is available from the corresponding author on reasonable request for a limited time-window of 3 months after publication.

3. Results

The Kolmogorov-Smirnov test showed that demographic and anthropometric parameters were normally distributed in the YA and OA as well as in female and male YA and OA subjects ($p > 0.05$ for all analyses). Mann-Whitney U test showed increased weight and BMI and decreased height values in OA subjects compared with YA ($p < 0.05$ for all comparisons)(Table 1, Supplementary Materials Tables S1 and S2).

3.1. YA and OA

When discriminating YA and OA, the artificial classifier based on SMO using selected features allowed us to achieve a significant diagnostic performance of our test. When comparing the 20 most relevant selected features extracted from the sustained emission of the vowel, ROC curve analyses identified an optimal diagnostic threshold value of 0.50 (associated criterion), when applying discretization and 10-folds cross-validation ($Y.I = 0.72$). Using this cut-off value, the performance of our diagnostic test was: sensitivity = 86.9%, specificity = 85.2%, PPV = 86.9%, NPV = 85.2%, accuracy = 86.1%, and AUC = 0.931 (Figure 2A, Table 2). Furthermore, when comparing 20 selected features extracted from the sustained emission of the sentence, ROC curve analyses identified an optimal diagnostic threshold value of 0.50, when applying discretization and 10-folds cross-validation ($Y.I = 0.77$). Using this cut-off value, the performance of our diagnostic test was: sensitivity = 89.1%, specificity = 87.7%, PPV = 89.1%, NPV = 87.7%, accuracy = 88.5%, and AUC = 0.938 (Figure 2B, Table 2). The two ROC curves obtained during the emission of the vowel and the sentence were comparable (the difference between AUCs = -0.007 , $z = -0.314$, $SE = 0.022$, $p = 0.75$) (Figure 2C).

To reduce excessive age dispersion, and thus perform a more consistent analysis of voice ageing, in a further analysis we compared the voice recordings collected from two subgroups of YA and OA. Moreover, in detail, among YA, we considered a subgroup of 79 YA with age ≤ 25 years (YA_{25}) (31 males and 41 females; mean age \pm SD 22.9 ± 2.2 years, range 15–25), whereas, among OA, we selected a subgroup of 71 OA with age ≥ 55 years (OA_{55}) (21 males and 50 females; mean age \pm SD 66.4 ± 8.1 years, range 55–85). When comparing the sustained emission of the vowel and the sentence in YA_{25} and OA_{55} we achieved further improvement in the results as shown by the ROC curve analysis. More in detail, when comparing 20 selected features extracted from the sustained emission of the vowel, ROC curve analyses identified optimal diagnostic threshold value of 0.59, when applying discretization and five-folds cross-validation ($Y.I = 0.86$). Using this cut-off value, the performance of our diagnostic was: sensitivity = 93.6%, specificity = 92.9%, PPV = 93.6%, NPV = 92.9%, accuracy = 93.2%, and AUC = 0.966 (Figure 2D, Table 2). Also, when comparing 20 selected features extracted from the sustained emission of the sentence, ROC curve analyses identified an optimal diagnostic threshold value of 0.52, when applying discretization and five-folds cross-validation ($Y.I = 0.91$). Using this cut-off value, the performance of our diagnostic test was: sensitivity = 92.8%, specificity = 98.5%, PPV = 98.7%,

NPV = 91.4%, accuracy = 95.3%, and AUC = 0.984 (Figure 2E, Table 2). Again, the two ROC curves obtained during the emission of the vowel and the sentence were comparable (the difference between AUCs = 0.018, $z = 0.753$, SE = 0.024, $p = 0.45$) (Figure 2F).

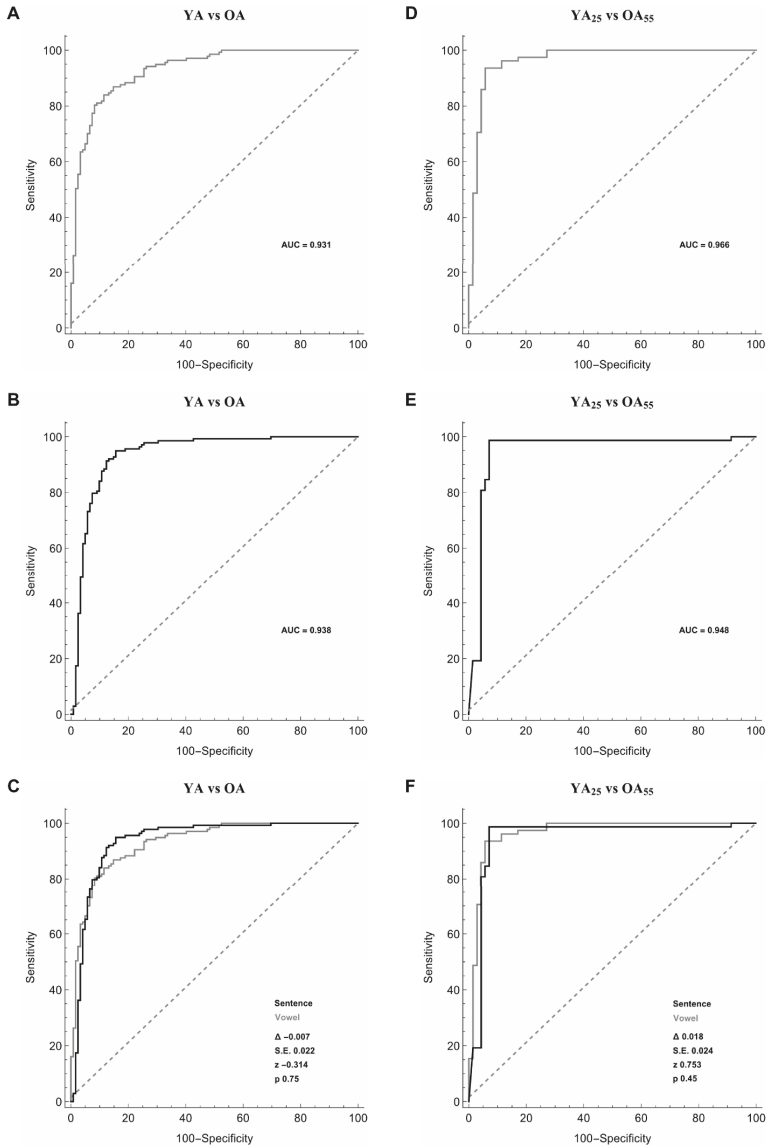


Figure 2. Receiver operating characteristic (ROC) curves used to differentiate younger adults (YA) vs. older adults (OA) (left column, panels (A–C)) and younger adults ≤ 25 years (YA₂₅) vs. older adults ≥ 55 years (OA₅₅) (right column, panels (D–F)) during the sustained emission of a vowel (grey line) (panels (A,D)), the sentence (black line) (panels (B,E)) and the comparison between the vowel and the sentence (panels (C,F)).

Table 2. Performance of the machine-learning algorithm in all the comparisons.

Comparisons	Speech-Task	Number of Instances	Cross-Validation	Assoc. Criterion	Youden Index	Se (%)	Sp (%)	PPV (%)	NPV (%)	Acc (%)	AUC
YA vs. OA	Vowel	259	10 folds	0.50	0.72	86.9	85.2	86.9	85.2	86.1	0.961
	Sentence	260	10 folds	0.50	0.77	89.1	87.7	89.1	87.7	88.5	0.938
YA ₂₅ vs. OA ₅₅	Vowel	148	5 folds	0.59	0.86	93.6	92.9	93.6	92.9	93.2	0.966
	Sentence	148	5 folds	0.52	0.91	92.8	98.5	98.7	91.4	95.3	0.984
YA _y vs. OA _y	Vowel	147	5 folds	0.57	0.81	90.3	90.7	90.3	90.7	90.5	0.958
	Sentence	148	5 folds	0.66	0.85	91.9	93.2	93.2	92.0	92.6	0.962
YA _m vs. OA _m	Vowel	111	5 folds	0.53	0.82	91.0	90.9	93.8	87.0	91.0	0.962
	Sentence	111	5 folds	0.52	0.87	91.3	95.2	96.9	87.0	92.8	0.958
YA _m vs. YA _f	Vowel	134	5 folds	0.69	0.91	95.4	95.7	95.4	95.7	95.5	0.965
	Sentence	135	5 folds	0.61	0.89	90.3	98.4	98.5	89.9	94.1	0.966
OA _m vs. OA _f	Vowel	120	5 folds	0.74	0.87	89.4	97.1	95.5	93.2	94.2	0.969
	Sentence	120	5 folds	0.63	0.86	89.8	95.8	93.6	93.2	93.3	0.975

Acc: accuracy; AUC: area under the curve; NPV: negative predictive value; OA: older adult; OA₅₅: older adult ≥ 55 years; OA_f: female OA; OA_m: male OA; PPV: positive predictive value; Se: sensitivity; Sp: specificity; YA_f: female YA; YA_m: male YA; YO: younger adult; YO₂₅: younger adult ≤ 25 years. Instances refer to the number of subjects considered in each comparison. Cross-validation refers to standardized procedures of a machine learning algorithm (see the text for details).

3.2. Female YA and Female OA

In the comparison of female YA and OA, the artificial classifier based on SMO achieved a significant diagnostic performance. More in detail, when comparing 20 selected features extracted from the sustained emission of the vowel, ROC curve analyses identified an optimal diagnostic threshold value of 0.57, when applying discretization and five-folds cross-validation (Y.I = 0.81). Using this cut-off value, the performance of our diagnostic test was: sensitivity = 90.3%, specificity = 90.7%, PPV = 90.3%, NPV = 90.7%, accuracy = 90.5% and AUC = 0.958 (Figure 3A, Table 2). Also, when examining the sustained emission of the sentence, ROC curve analyses identified optimal diagnostic threshold value of 0.66, when applying discretization and five-folds cross-validation (Y.I = 0.85). Using this cut-off value, the performance of our diagnostic test was: sensitivity = 91.9%, specificity = 93.2%, PPV = 93.2%, NPV = 92.0%, accuracy = 92.6%, and AUC = 0.962 (Figure 3B, Table 2). The two ROC curves obtained during the emission of the vowel and the sentence were similar (the difference between AUCs = -0.004 , $z = -0.164$, $SE = 0.024$, $p = 0.87$) (Figure 3C).

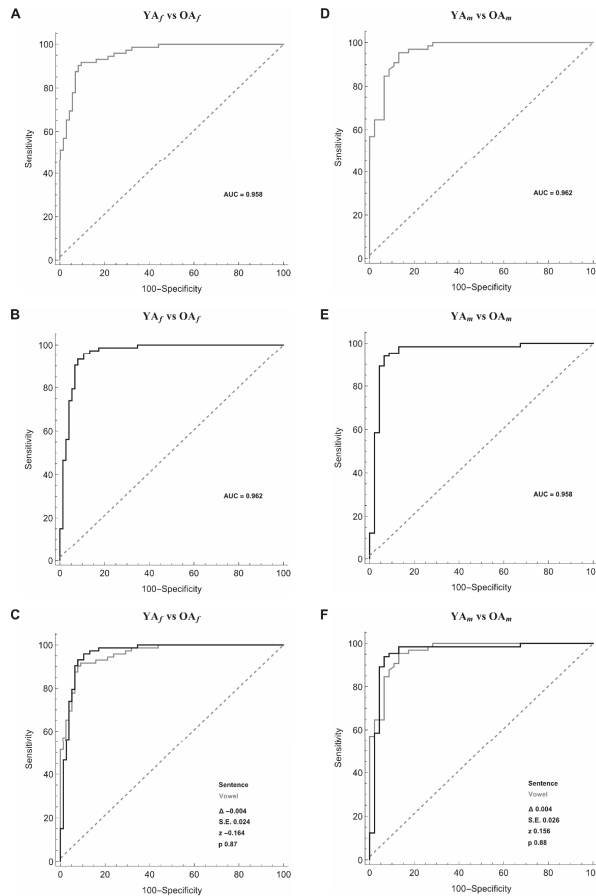


Figure 3. Receiver operating characteristic (ROC) curves used to differentiate female younger adults (YA_f) and female older adults (OA_f) (left column, panels (A–C)) and male younger adults (YA_m) and male older adults (OA_m) (right column, panels (D–F)) during the sustained emission of a vowel (grey line) (panels (A,D)), the sentence (black line) (panels (B,E)), and the comparison between the vowel and the sentence (panels (C,F)).

3.3. Male YA and Male OA

In the comparison of male YA and OA, the artificial classifier based on SMO using 20 selected features achieved a significant diagnostic performance. When comparing selected features extracted from the sustained emission of the vowel, ROC curve analyses identified optimal diagnostic threshold value of 0.53, when applying discretization and five-folds cross-validation (Y.I = 0.82). Using this cut-off value, the performance of our diagnostic test was: sensitivity = 91.0%, specificity = 90.9%, PPV = 93.8%, NPV = 87.0%, accuracy = 91.0% and AUC = 0.962 (Figure 3D, Table 2). Also, when examining the sustained emission of the sentence, ROC curve analyses identified an optimal diagnostic threshold value of 0.52, when applying discretization and five-folds cross-validation (Y.I = 0.87). Using this cut-off value, the performance of our diagnostic test was: sensitivity = 91.3%, specificity = 95.2%, PPV = 96.9%, NPV = 87.0%, accuracy = 92.8%, and AUC = 0.958 (Figure 3E, Table 2). The difference between the two ROC curves obtained during the emission of the vowel and the sentence was not significant (the difference between AUCs = 0.004, $z = 0.156$, SE = 0.026, $p = 0.88$) (Figure 3F).

3.4. Male and Female YA

In the analysis of male vs. female YA, the artificial classifier based on SMO achieved a significant diagnostic performance. More in detail, when comparing 20 selected features extracted from the sustained emission of the vowel, ROC curve analyses identified an optimal diagnostic threshold value of 0.69, when applying discretization and 5-folds cross-validation (Y.I = 0.91). Using this cut-off value, the performance of our diagnostic test was: Sensitivity = 95.4%, Specificity = 95.7%, PPV = 95.4%, NPV = 95.7%, Accuracy = 95.5% and AUC = 0.965 (Figure 4A, Table 2). Also, when analyzing the sustained emission of the sentence, ROC curve analyses identified an optimal diagnostic threshold value of 0.61, when applying discretization and 5-folds cross-validation (Y.I = 0.89). Using this cut-off value, the performance of our diagnostic test was: sensitivity = 90.3%, specificity = 98.4%, PPV = 98.5%, NPV = 89.9%, accuracy = 94.1%, and AUC = 0.966 (Figure 4B, Table 2). The two ROC curves obtained during the emission of the vowel and the sentence were comparable (the difference between AUCs = -0.001 , $z = -0.043$, SE = 0.023, $p = 0.97$) (Figure 4C).

3.5. Male and Female OA

When differentiating male and female OA, the artificial classifier based on SMO achieved a significant diagnostic performance. More in detail, when comparing 20 selected features extracted from the sustained emission of the vowel, ROC curve analyses identified an optimal diagnostic threshold value of 0.74, when applying discretization and five-folds cross-validation (Y.I = 0.87). Using this cut-off value, the performance of our diagnostic test was: sensitivity = 89.4%, specificity = 97.1%, PPV = 95.5%, NPV = 93.2%, accuracy = 94.2%, and AUC = 0.969 (Figure 4D, Table 2). Also, when examining the sustained emission of the sentence, ROC curve analyses identified an optimal diagnostic threshold value of 0.63, when applying discretization and five-folds cross-validation (Y.I = 0.86). Using this cut-off value, the performance of our diagnostic test was: sensitivity = 89.8%, specificity = 95.8%, PPV = 93.6%, NPV = 93.2%, accuracy = 93.3%, and AUC = 0.975 (Figure 4E, Table 2). The two ROC curves obtained during the emission of the vowel and the sentence were comparable (the difference between AUCs = -0.006 , $z = -0.245$, SE = 0.025, $p = 0.81$) (Figure 4F).

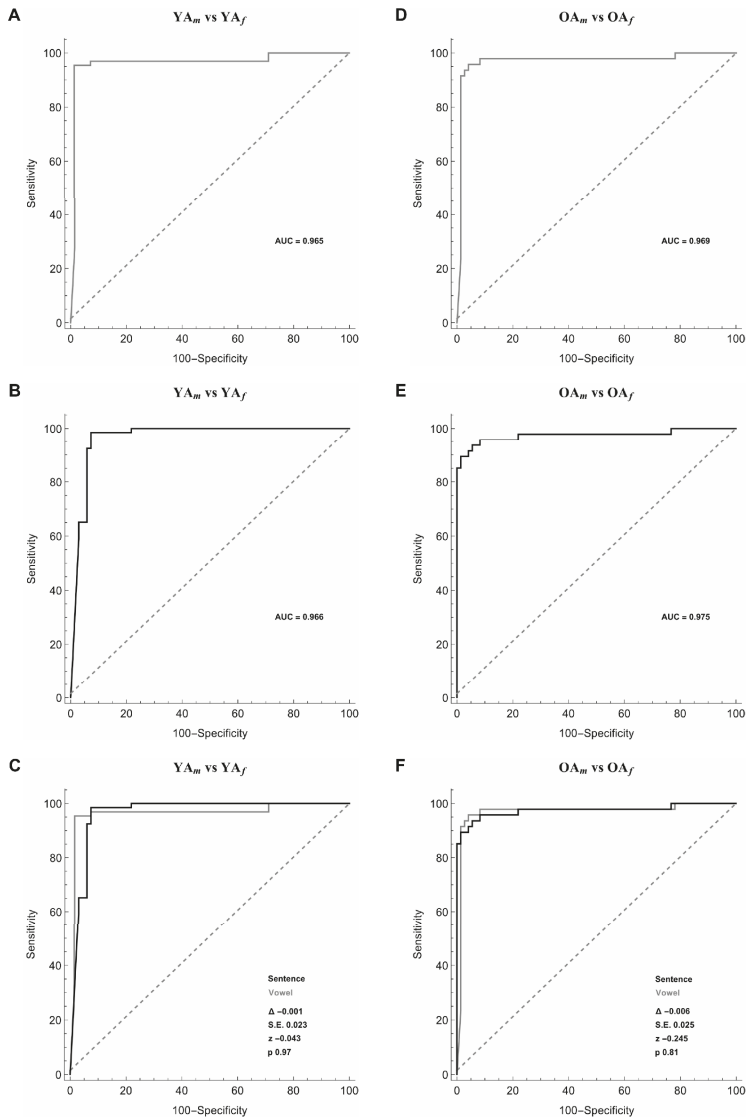


Figure 4. Receiver operating characteristic (ROC) curves used to differentiate female Younger Adults (YA_f) and male Younger Adults (YA_m) (left column, panels (A–C)) and female older adults (OA_f) and male older adults (OA_m) (right column, panels (D–F)) during the sustained emission of a vowel (grey line) (panels (A,D)), the sentence (black line) (panels (B,E)) and the comparison between the vowel and the sentence (panels (C,F)).

4. Discussion

In this study, we found that machine learning analysis of voice samples recorded through smartphones correctly discriminates between YA and OA. We have also demonstrated that our voice analysis accurately discriminates females and males in both groups. By comparing male and female YA, as well as male and female OA, we have also examined in detail the combined effect of age and

gender on voice. Accordingly, by using machine learning analysis, in this study we have demonstrated the effect of ageing and gender on voice.

To collect homogeneous and high-quality recordings, we have carefully controlled for several methodological factors. All participants were native Italian speakers. To exclude confounding related to the acute and chronic effects of smoking on the physiology of the vocal folds, lungs, and resonant structures, we have included in the study only non-smokers. By contrast, we excluded subjects with cognitive or mood impairment or those taking drugs acting on the central nervous system at the time of the study. We also excluded from the study cohort subjects with bilateral/unilateral hearing loss, respiratory disorders, and other pathological conditions directly or indirectly affecting the vocal cords. The age range considered for the YA group was based on the definition of young subjects provided by the World Health Organization [58]. Accordingly, all the YA participants completed the pubertal development. By contrast, the age range considered for the OA group was set to include subjects in the middle and late adulthood [59]. In this study, we excluded voice recordings from subjects in the early adulthood (30–40 years) in order to better separate the study cohort into two independent subgroups of different ages. Lastly, all voice samples were collected through smartphones able to save audio tracks in mp4 format.

The main novelty of the study consists of the acquisition and analysis of voice samples collected through smartphones. Indeed, although a few studies have previously used smartphones to collect voice samples in patients with voice disorders [60–62], so far no authors have used this methodological approach to examine age-related changes of voice. The use of smartphones allows a simplified procedure of voice recordings and open to the acquisition of a large amount of data collected in an ecologic scenario.

4.1. The Effect of Ageing on Voice

The first finding of our study is that the objective voice analysis based on machine learning can distinguish YA and OA subjects, with a high level of accuracy as demonstrated by our ROC curve analyses. The accuracy of the algorithm tended to improve further when comparing the YA and OA subjects with a narrower age-band (YA₂₅ and OA₅₅). Furthermore, to investigate age-related changes in the human voice in more detail, we have also compared gender-matched groups of YA and OA subjects. Indeed, by comparing females included in the YA and OA groups as well as males included in the YA and OA groups, in separate analyses, we have examined the pure effect of ageing on voice. Our findings fully agree with previous reports demonstrating the effect of ageing on the human voice [24–28,33–38]. Early studies based on the qualitative/perceptual evaluation of voice recordings have demonstrated that physiologic ageing leads to several changes in specific characteristics of the human voice [1]. Indeed, as a result of physiologic ageing, voices progressively manifest increased breathiness and hoarseness, reduced speech intensity as well as maximum phonation time [2–4,15]. Experimental studies using spectral analysis have confirmed age-related changes in voice by providing new objective measures in the time-domain as well as in the frequency-domain. For instance, both the jitter and the shimmer were higher in OA than in YA subjects [1], the former reflecting the degree of voice hoarseness [63], whereas the latter relates to the degree of the breathiness of the voice [1]. Also, the N/H ratio, which reflects the level of noise of an acoustic signal, also increases in the elderly [18]. Finally, concerning measures in the frequency domain, previous studies using spectral analysis have also shown age-related changes in voice even though with some inconsistency. For instance, in the elderly, the fundamental frequency (f₀) decreased [64–67], increased [68–70], or even remain unchanged [71–73].

In our study, by applying the ROC curve analysis, we demonstrated in detail the high accuracy of our machine learning analysis in demonstrating age-related changes in the human voice. Our results fit in well with previous studies applying automatic classifiers based on machine learning analysis [24–28,33–38]. More in detail, our machine learning algorithm has achieved higher results than those obtained on the INTERSPEECH 2010 age and gender sub-challenge feature set [33,34]. Among machine learning algorithms, the standard and hybrid versions of the SVM (e.g., SVM-GMM)

are thought to be both consistent and accurate [33–35,38,73]. In our study, SVM achieved relatively high performance with an accuracy of 95.3% in age recognition and of 95.5% in gender recognition, showing comparable or even better results than those obtained in previous reports [33–35,38,73]. When comparing our methodological approach to those previously used, it is important to consider that we started with a large dataset of features (more than 6000), adopting dedicated ranking and feature selection algorithms [33–38,73]. The advantages of applying those algorithms consist of obtaining smaller dataset of features (only 20 features in our study), easier math handled and with shorter computation time. Moreover, all the previous studies considered only MFCC-, f₀-, pitch-, energy-, jitter-, and shimmer-related features [24–28,33–37], with only a study considering non-traditional features including RASTA-PLP coefficients [38]. In addition to the traditional frequency-, jitter-, shimmer-, energy-, spectral, and cepstral-related features, we have also included MFCC and RASTA-PLP coefficients and three additional representative features (HNR, SNR, and CPPs). The inclusion of HNR, SNR, CPPs, and RASTA-PLP coefficients to the general dataset of LLDs allowed us to achieve a more robust analysis. Indeed, these features were frequently included in the 20 most relevant selected features in all the comparisons made by our machine learning algorithm. Also, SNR, CPPs, MFCC-, RASTA coefficients-, f₀-, spectral-, and energy-related features specifically changed in the human voice according to physiologic ageing (see Table S3 in supplementary material for a detailed list of the first 20 selected features during the comparison between YA and OA). In our case, particularly the RASTA filtering technique has allowed reducing the irrelevant information introduced into the signal by the microphones or by the background noise [49]. Since in our study each vocal sample was recorded with a different smartphone the use of RASTA filtering made possible to eliminate the effect due to the use of different microphones.

Several age-related changes in physiological functions may explain our findings. The physiological basis underlying our results and those previously obtained with the perceptual and standard objective analysis are prominently related to age-related changes of the phonatory apparatus. These changes are secondary to: Loss of elasticity and tone of the vocal folds and the pharyngeal walls; increase of fat distribution in the neck and the parapharyngeal space; progressive reduction of the secretion of the salivary and mucous glands; thinning of the tongue and loss of teeth with relevant changes in shape and diameter of the oral cavity [5]. Moreover, at a cellular and molecular level, physiological ageing leads to thinning of the laryngeal epithelium, loss of the elastic chord component, and increase in the collagen fibers/elastic fibers ratio which in turn decrease vocal folds viscoelasticity [6–14]. Also, the myelin fiber density of the superior and recurrent laryngeal nerve progressively reduces with age leading to an alteration of the intrinsic reflex tone and muscle flaccidity [74,75]. Besides age-related changes in specific components of the phonatory apparatus, voice can be influenced also by additional anthropometric factors including weight and height of the subjects. In this study, we found that OA subjects had increased weight and BMI and decreased height values compared with YA. Although our methodological approach does not allow to clarify the link between any of the voice features selected by the SMO and age-related changes in specific components of the phonatory apparatus or anthropometric factors, we believe that our machine learning analysis of the human voice provides objective evaluation of the human ageing.

4.2. The Effect of Gender on Voice

Our machine learning analysis allowed us also to examine in detail the effect of gender on voice. Our machine learning analysis differentiated female and male YA as well as female and male OA with high accuracy. It is known that gender leads to additional sources of variability in voice features. Previous perceptual and objective studies of the human voice have shown that before the pubertal age, males and females have a rather similar vocal pitch. During puberty, the male voice typically deepens an octave, while the female voice usually deepens only by a few tones. Thus, before puberty, the voice examination does not show any difference between males and females, whereas, in the adulthood, the examiner can usually recognize the gender of the speaker [18,63–65,67,68,71–73]. The physiologic

basis of differences in voice parameters between males and females relies on several physiologic and anatomic issues. The hormones grow the larynx and the vocal folds in both males and females, but in males, the growth is rather prominent. Then, in women during the menopausal phase, the level of estrogen hormone decreases along with an increase in androgens. As a result, the thickness of the vocal cords increases and leads to a deeper tone of voice. A complementary phenomenon occurs in males during andropause, characterized by a drop in the level of androgens and a relative increase of the estrogen/androgen ratio [5,76]. Our findings agree with previous findings from perceptible and quantitative voice studies further demonstrating that voice objectively differs in females and males [1]. However, our machine learning analysis does not provide evidence for a strict relation between any of the voice features here considered and specific gender-related changes in the phonatory apparatus.

Another important finding of our study concerns the comparable results achieved when examining voice samples collected during the emission of the vowel and the sentence [24,77]. This finding suggests the comparable ability of machine learning to recognize voice changes due to the combined effect of ageing and gender, during the sustained emission of a vowel as well as a sentence. We suggest, however, that compared to the recording of a sentence, voice samples including the sustained emission of a vowel would be more practical and more reliable thus improving voice analyses among the different languages.

A final comment concerns how relevant is the objective evaluation of ageing processes in humans [78]. Age can be classified into “chronological” and “biological” components [79], the former referring to the actual amount of years of a subject, whereas the latter reflects the amount of age-related changes in various physiological functions in the same subject. The physiologic ageing represents a gradual and continuous process reflecting the interaction between genetic and environmental factors, and leading to the progressive decline of physical, psychological, and social functions [80]. To date, no standardized biomarkers of physiologic ageing are currently available. We, therefore, believe that our voice analysis with machine learning would provide a novel and advanced tool possibly helpful for quantifying the individual “biological” age of a single subject [81,82]. The objective voice analysis would also allow to better discriminate and monitor processes of physiological as well as pathological ageing.

A possible limitation of this study is the reduced sample of voice recordings undergoing machine learning analysis. However, the level of significance of our results in all the comparison is relatively high. We did not record voices in young females under different phases of the menstrual cycle thus not excluding the possible effect of hormones on voices. The intrinsic variability in the brand and model of the smartphones used to record voice samples (e.g., variability related to microphones and recording algorithms) would have affected our results. For instance, depending on the specific smartphone used, mp4 audio files can be compressed through different audio coding standards for lossy or lossless digital audio compression (e.g., AAC—advanced audio coding; Apple Lossless Audio Codec—ALEC, or Free Lossless Audio Codec—FLAC). Hence, we cannot exclude that the heterogeneity in the brand and model of the smartphones also increased the variability of our data. Also, since in the present study we did not record voice samples serially, we cannot exclude variability in voice recordings due to daily fluctuations in voice parameters. Furthermore, our study did not include the longitudinal evaluation of voice recordings in the same subjects. This study design although theoretically feasible is technically difficult. Hence, in the present study, the lack of a follow-up evaluation of voice recordings did not allow us to clarify intra-subject age-related changes in the human voice. Lastly, we cannot fully exclude that the increased weight and BMI, and the decreased height observed in OA subjects would have contributed at least in part to our findings [83].

5. Conclusions

Advanced voice analysis based on machine-learning performed on voice samples collected using smartphones can distinguish between younger and older healthy subjects, thus objectively evaluating the effect of physiologic ageing on the voice in humans. Our voice analysis is also able to discriminate

between females and males from YA and OA groups, thus demonstrating the interaction between ageing- and gender-related factors in determining the human voice. Future cohort studies comparing voice recordings in a larger number of samples of different ages (e.g., large samples of subjects in early, middle and late adulthood) will better examine whether age-related changes in voice can be considered biomarkers of human ageing. Furthermore, we believe that our study would provide new helpful information to clinicians to better distinguish physiologic ageing from pathological changes of the human voice in subjects affected by various speech disorders [77,84].

Supplementary Materials: The following are available online at <http://www.mdpi.com/1424-8220/20/18/5022/s1>, Table S1: Demographic and anthropometric characteristics of younger adults. Table S2: Demographic and anthropometric characteristics of older adults. Table S3. Ranking of the first 20 features (functionals applied to low-level descriptors) extracted using OpenSMILE and selected using CAE for the comparison between YA and OA, during the emission of the vowel and the sentence. Each feature is identified by four items: (1) family of low-level descriptor (LLD), (2) LLD, (3) functional used to calculate that specific feature and, (4) the value of relevance calculated through CAE algorithm.

Author Contributions: Conceptualization, F.A., G.C., G.S., and A.S.; data curation, F.A., G.C., and A.Z.; formal analysis, F.A. and G.C.; investigation, F.A., G.C., A.Z., and A.S.; methodology, F.A., P.D.L., G.R., and A.S.; software, P.D.L. and A.Z.; supervision, G.S. and A.S.; validation, A.S.; writing—original draft, F.A. and P.D.L.; writing—review and editing, G.C., G.R., A.B., G.S., and A.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

- Goy, H.; Fernandes, D.N.; Pichora-Fuller, M.K.; van Lieshout, P. Normative Voice Data for Younger and Older Adults. *J. Voice* **2013**, *27*, 545–555. [[CrossRef](#)]
- Kendall, K. Presbyphonia: A review. *Curr. Opin. Otolaryngol. Head Neck Surg.* **2007**, *15*, 137–140. [[CrossRef](#)]
- De Araújo Pernambuco, L.; Espelt, A.; Balata, P.M.M.; de Lima, K.C. Prevalence of voice disorders in the elderly: A systematic review of population-based studies. *Eur. Arch. Otorhinolaryngol.* **2015**, *272*, 2601–2609. [[CrossRef](#)]
- Mezzedimi, C.; Di Francesco, M.; Livi, W.; Spinosi, M.C.; De Felice, C. Objective Evaluation of Presbyphonia: Spectroacoustic Study on 142 Patients with Praat. *J. Voice* **2017**, *31*, 257.e25–257.e32. [[CrossRef](#)]
- Bruzzi, C.; Salsi, D.; Minghetti, D.; Negri, M.; Casolino, D.; Sessa, M. Presbiphonya. *Acta Biomed.* **2017**, *88*, 6–10. [[CrossRef](#)]
- Gonçalves, T.M.; Dos Santos, D.C.; Pessin, A.B.B.; Martins, R.H.G. Scanning Electron Microscopy of the Presbylarynx. *Otolaryngol. Head Neck Surg.* **2016**, *154*, 1073–1078. [[CrossRef](#)]
- Hirano, S.; Minamiguchi, S.; Yamashita, M.; Ohno, T.; Kanemaru, S.-I.; Kitamura, M. Histologic characterization of human scarred vocal folds. *J. Voice* **2009**, *23*, 399–407. [[CrossRef](#)]
- Sato, K.; Hirano, M. Histologic investigation of the macula flava of the human newborn vocal fold. *Ann. Otol. Rhinol. Laryngol.* **1995**, *104*, 556–562. [[CrossRef](#)]
- Chan, R.W.; Gray, S.D.; Titze, I.R. The importance of hyaluronic acid in vocal fold biomechanics. *Otolaryngol. Head Neck Surg.* **2001**, *124*, 607–614. [[CrossRef](#)]
- Chen, X.; Thibeault, S.L. Characteristics of age-related changes in cultured human vocal fold fibroblasts. *Laryngoscope* **2008**, *118*, 1700–1704. [[CrossRef](#)]
- Allah, R.; Dkhil, M.; Farhoud, E. Fibroblasts in the human vocal fold mucosa: An ultrastructural study of different age groups. *Singap. Med. J.* **2009**, *50*, 201–207.
- Hammond, T.H.; Gray, S.D.; Butler, J.; Zhou, R.; Hammond, E. Age- and gender-related elastin distribution changes in human vocal folds. *Otolaryngol. Head Neck Surg.* **1998**, *119*, 314–322. [[CrossRef](#)]
- McMullen, C.A.; Andrade, F.H. Contractile dysfunction and altered metabolic profile of the aging rat thyroarytenoid muscle. *J. Appl. Physiol.* **2006**, *100*, 602–608. [[CrossRef](#)]
- Clafin, D.R.; Faulkner, J.A. Shortening velocity extrapolated to zero load and unloaded shortening velocity of whole rat skeletal muscle. *J. Physiol.* **1985**, *359*, 357–363. [[CrossRef](#)]

15. Vaca, M.; Mora, E.; Cobeta, I. The Aging Voice: Influence of Respiratory and Laryngeal Changes. *Otolaryngol. Head Neck Surg.* **2015**, *153*, 409–413. [[CrossRef](#)]
16. Hodge, F.S.; Colton, R.H.; Kelley, R.T. Vocal Intensity Characteristics in Normal and Elderly Speakers. *J. Voice* **2001**, *15*, 503–511. [[CrossRef](#)]
17. Prakup, B. Acoustic Measures of the Voices of Older Singers and Nonsingers. *J. Voice* **2012**, *26*, 341–350. [[CrossRef](#)]
18. Ferrand, C.T. Harmonics-to-Noise Ratio. *J. Voice* **2002**, *16*, 480–487. [[CrossRef](#)]
19. Baughman, R.P.; Loudon, R.G. Sound spectral analysis of voice-transmitted sound. *Am. Rev. Respir. Dis.* **1986**, *134*, 167–169. [[CrossRef](#)] [[PubMed](#)]
20. Titze, I.R.; Baken, R.J.; Bozeman, K.W.; Granqvist, S.; Henrich, N.; Herbst, C.T.; Howard, D.M.; Hunter, E.J.; Kaelin, D.; Kent, R.D.; et al. Toward a consensus on symbolic notation of harmonics, resonances, and formants in vocalization. *J. Acoust. Soc. Am.* **2015**, *137*, 3005–3007. [[CrossRef](#)]
21. Hillenbrand, J.; Houde, R.A. Acoustic correlates of breathy vocal quality: Dysphonic voices and continuous speech. *J. Speech Hear. Res.* **1996**, *39*, 311–321. [[CrossRef](#)]
22. Hillenbrand, J.; Cleveland, R.A.; Erickson, R.L. Acoustic correlates of breathy vocal quality. *J. Speech Hear. Res.* **1994**, *37*, 769–778. [[CrossRef](#)]
23. Delgado-Hernández, J.; León-Gómez, N.M.; Izquierdo-Arteaga, L.M.; Llanos-Fumero, Y. Cepstral analysis of normal and pathological voice in Spanish adults. Smoothed cepstral peak prominence in sustained vowels versus connected speech. *Acta Otorrinolaringol. Esp.* **2018**, *69*, 134–140. [[CrossRef](#)]
24. Li, M.; Han, K.; Narayanan, S. Automatic Speaker Age and Gender Recognition Using Acoustic and Prosodic Level Information Fusion. *Comput. Speech Lang.* **2012**, *27*. [[CrossRef](#)]
25. Spiegl, W.; Stemmer, G.; Lasarczyk, E.; Kolhatkar, V.; Cassidy, A.; Potard, B.; Shum, S.; Song, Y.; Xu, P.; Beyerlein, P.; et al. Analyzing Features for Automatic Age Estimation on Cross-Sectional Data. In Proceedings of the INTERSPEECH 2009, 10th Annual Conference of the International Speech Communication Association, Brighton, UK, 6–10 September 2009; pp. 2923–2926.
26. Stolcke, A.; Kajarekar, S.S.; Ferrer, L.; Shrinberg, E. Speaker Recognition with Session Variability Normalization Based on MLLR Adaptation Transforms. *IEEE Trans. Audio Speech Lang. Process.* **2007**, *15*, 1987–1998. [[CrossRef](#)]
27. Berardi, M.L.; Hunter, E.J.; Ferguson, S.H. Talker age estimation using machine learning. *Proc Meet Acoust* **2017**, *30*. [[CrossRef](#)]
28. Zhavoronkov, A.; Li, R.; Ma, C.; Mamoshina, P. Deep biomarkers of aging and longevity: From research to applications. *Aging* **2019**, *11*, 10771–10780. [[CrossRef](#)]
29. Deo, R.C. Machine Learning in Medicine. *Circulation* **2015**, *132*, 1920–1930. [[CrossRef](#)]
30. Costantini, G.; Todisco, M.; Perfetti, R.; Basili, R.; Casali, D. Svm Based Transcription System with Short-Term Memory Oriented to Polyphonic Piano Music. In Proceedings of the MELECON 2010—2010 15th IEEE Mediterranean Electrotechnical Conference, La Valletta, Malta, 26–28 April 2010; pp. 196–201.
31. Costantini, G.; Casali, D.; Todisco, M. An SVM Based Classification Method for EEG Signals. In Proceedings of the 14th WSEAS international conference on Circuits, Corfu Island, Greece, 22–24 July 2010; Volume 2224.
32. Van Calster, B.; Wynants, L. Machine Learning in Medicine. *N. Engl. J. Med.* **2019**, *380*, 2588. [[CrossRef](#)]
33. Kockmann, M.; Burget, L.; Černocký, J. Brno University of Technology System for Interspeech 2010 Paralinguistic Challenge. In Proceedings of the INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association, Makuhari, Chiba, Japan, 26–30 September 2010.
34. Meinedo, H.; Trancoso, I. Age and Gender Classification Using Fusion of Acoustic and Prosodic Features. In Proceedings of the INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication Association, Makuhari, Chiba, Japan, 26–30 September 2010.
35. Pribil, J.; Přebilová, A.; Matoušek, J. GMM-based speaker age and gender classification in Czech and Slovak. *J. Electr. Eng.* **2017**, *68*, 3–12. [[CrossRef](#)]
36. Grzybowska, J.; Kacprzak, S. Speaker Age Classification and Regression Using i-Vectors. In Proceedings of the INTERSPEECH 2016, 16th Annual Conference of the International Speech Communication Association, San Francisco, CA, USA, 8–12 September 2016; pp. 1402–1406.
37. Sedaghi, M. A Comparative Study of Gender and Age Classification in Speech Signals. *Iran. J. Electr. Electron. Eng.* **2009**, *5*, 1–12.

38. Barkana, B.D.; Zhou, J. A new pitch-range based feature set for a speaker's age and gender classification. *Appl. Acoust.* **2015**, *98*, 52–61. [\[CrossRef\]](#)
39. Higgins, J.P. Smartphone Applications for Patients' Health and Fitness. *Am. J. Med.* **2016**, *129*, 11–19. [\[CrossRef\]](#)
40. Alameen, G. Review of Audacity computer software. *TESL-EJ* **2007**, *11*, 1–11.
41. Russell, S.J.; Norvig, P.; Davis, E. *Artificial Intelligence: A Modern Approach*, 3rd ed.; Prentice Hall Series in Artificial Intelligence; Prentice Hall: Upper Saddle River, NJ, USA, 2010; ISBN 978-0-13-604259-4.
42. Specht, D. A General Regression Neural Network. *IEEE Trans. Neural Netw.* **1991**, *2*, 568–578. [\[CrossRef\]](#)
43. Alpaydin, E. *Introduction to Machine Learning*, 2nd ed.; Adaptive Computation and Machine Learning; MIT Press: Cambridge, MA, USA, 2010; ISBN 978-0-262-01243-0.
44. Schuller, B.; Steidl, S.; Batliner, A.; Hirschberg, J.; Burgoon, J.K.; Baird, A.; Elkins, A.; Zhang, Y.; Coutinho, E.; Evanini, K. The INTERSPEECH 2016 Computational Paralinguistics Challenge: Deception, Sincerity and Native Language. In Proceedings of the INTERSPEECH 2016, 16th Annual Conference of the International Speech Communication Association, San Francisco, CA, USA, 8–12 September 2016; pp. 2001–2005.
45. Eyben, F.; Wenginger, F.; Gross, F.; Schuller, B. Recent Developments in openSMILE, the Munich Open-Source Multimedia Feature Extractor. In *Proceedings of the 21st ACM International Conference on Multimedia—MM'13*; ACM Press: Barcelona, Spain, 2013; pp. 835–838.
46. Schuller, B.; Steidl, S.; Batliner, A.; Vinciarelli, A.; Scherer, K.; Ringeval, F.; Chetouani, M.; Wenginger, F.; Eyben, F.; Marchi, E.; et al. The INTERSPEECH 2013 Computational Paralinguistics Challenge: Social Signals, Conflict, Emotion, Autism. In Proceedings of the INTERSPEECH 2013, 13th Annual Conference of the International Speech Communication Association, Lyon, France, 25–29 August 2013; p. 152.
47. Davis, S.; Mermelstein, P. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Trans. Acoust. Speech Signal Process.* **1980**, *28*, 357–366. [\[CrossRef\]](#)
48. Young, S.; Kershaw, D.; Odell, J.; Ollason, D.; Valtchev, V.; Woodland, P. The HTK Book. *Camb. Univ. Eng. Dep.* **2002**, *3*, 12.
49. Hermansky, H.; Morgan, N. RASTA processing of speech. *IEEE Trans. Speech Audio Process.* **1994**, *2*, 578–589. [\[CrossRef\]](#)
50. Heman-Ackah, Y.D.; Michael, D.D.; Goding, G.S. The relationship between cepstral peak prominence and selected parameters of dysphonia. *J. Voice* **2002**, *16*, 20–27. [\[CrossRef\]](#)
51. Heman-Ackah, Y.D.; Sataloff, R.T.; Laureyns, G.; Lurie, D.; Michael, D.D.; Heuer, R.; Rubin, A.; Eller, R.; Chandran, S.; Abaza, M.; et al. Quantifying the cepstral peak prominence, a measure of dysphonia. *J. Voice* **2014**, *28*, 783–788. [\[CrossRef\]](#)
52. Hall, M.A.; Smith, L.A. Practical Feature Subset Selection for Machine Learning. In Proceedings of the 21st Australasian Computer Science Conference ACSC'98, Perth, Australia, 4–6 February 1998; pp. 181–191.
53. Hall, M. Correlation-Based Feature Selection for Machine Learning. *Dep. Comput. Sci.* **2000**, *19*, 1–198.
54. Fayyad, U.M.; Irani, K.B. On the handling of continuous-valued attributes in decision tree generation. *Mach. Learn.* **1992**, *8*, 87–102. [\[CrossRef\]](#)
55. Platt, J. Fast Training of Support Vector Machines Using Sequential Minimal Optimization. In *Advances in Kernel Methods: Support Vector Learning*; MIT Press: Cambridge, MA, USA, 1999; pp. 185–208. ISBN 978-0-262-19416-7.
56. Frank, E.; Hall, M.; Holmes, G.; Kirkby, R.; Pfahringer, B.; Witten, I.H.; Trigg, L. Weka-A Machine Learning Workbench for Data Mining. In *Data Mining and Knowledge Discovery Handbook*; Maimon, O., Rokach, L., Eds.; Springer: Boston, MA, USA, 2009; pp. 1269–1277, ISBN 978-0-387-09822-7.
57. DeLong, E.R.; DeLong, D.M.; Clarke-Pearson, D.L. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics* **1988**, *44*, 837–845. [\[CrossRef\]](#)
58. Assembly, U.G. Convention on the Rights of the Child. *U. N. Treaty Ser.* **1989**, *1577*, 1–15.
59. Livingston, G.; Huntley, J.; Sommerlad, A.; Ames, D.; Ballard, C.; Banerjee, S.; Brayne, C.; Burns, A.; Cohen-Mansfield, J.; Cooper, C.; et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet* **2020**, *396*, 413–446. [\[CrossRef\]](#)
60. Hegde, S.; Shetty, S.; Rai, S.; Dodderi, T. A Survey on Machine Learning Approaches for Automatic Detection of Voice Disorders. *J. Voice* **2019**, *33*, 947.e11–947.e33. [\[CrossRef\]](#)

61. Zhan, A.; Mohan, S.; Tarolli, C.; Schneider, R.B.; Adams, J.L.; Sharma, S.; Elson, M.J.; Spear, K.L.; Glidden, A.M.; Little, M.A.; et al. Using Smartphones and Machine Learning to Quantify Parkinson Disease Severity: The Mobile Parkinson Disease Score. *JAMA Neurol.* **2018**, *75*, 876–880. [[CrossRef](#)]
62. Arora, S.; Venkataraman, V.; Zhan, A.; Donohue, S.; Biglan, K.M.; Dorsey, E.R.; Little, M.A. Detecting and monitoring the symptoms of Parkinson’s disease using smartphones: A pilot study. *Parkinsonism Relat. Disord.* **2015**, *21*, 650–653. [[CrossRef](#)]
63. Hakkesteegt, M.M.; Brocaar, M.P.; Wieringa, M.H.; Feenstra, L. Influence of Age and Gender on the Dysphonia Severity Index. *Folia Phoniatr. Logop.* **2006**, *58*, 264–273. [[CrossRef](#)] [[PubMed](#)]
64. Awan, S.N. The aging female voice: Acoustic and respiratory data. *Clin. Linguist. Phon.* **2006**, *20*, 171–180. [[CrossRef](#)]
65. Ma, E.P.-M.; Love, A.L. Electroglottographic Evaluation of Age and Gender Effects During Sustained Phonation and Connected Speech. *J. Voice* **2010**, *24*, 146–152. [[CrossRef](#)]
66. de Aquino, F.S.; Ferreira, L.P.; de Aquino, F.S.; Ferreira, L.P. Vocal Parameters of Elderly Female Choir Singers. *Int. Arch. Otorhinolaryngol.* **2016**, *20*, 25–29. [[CrossRef](#)]
67. Deliyski, D. Effects of aging on selected acoustic voice parameters: Preliminary normative data and educational implications. *Educ. Gerontol.* **2001**, *27*, 159–168. [[CrossRef](#)]
68. Decoster, W.; Debruyne, F. The ageing voice: Changes in fundamental frequency, waveform stability and spectrum. *Acta Otorhinolaryngol. Belg.* **1997**, *51*, 105–112.
69. Harnsberger, J.D.; Shrivastav, R.; Brown, W.S.; Rothman, H.; Hollien, H. Speaking rate and fundamental frequency as speech cues to perceived age. *J. Voice* **2008**, *22*, 58–69. [[CrossRef](#)]
70. Benjamin, B.J. Frequency variability in the aged voice. *J. Gerontol.* **1981**, *36*, 722–726. [[CrossRef](#)]
71. Orlikoff, R.F. The Relationship of Age and Cardiovascular Health to Certain Acoustic Characteristics of Male Voices. *J. Speech Lang. Hear. Res.* **1990**, *33*, 450–457. [[CrossRef](#)]
72. Ramig, L.A.; Ringel, R.L. Effects of Physiological Aging on Selected Acoustic Characteristics of Voice. *J. Speech Lang. Hear. Res.* **1983**, *26*, 22–30. [[CrossRef](#)]
73. Zraick, R.I.; Smith-Olinde, L.; Shotts, L.L. Adult Normative Data for the KayPENTAX Phonatory Aerodynamic System Model 6600. *J. Voice* **2012**, *26*, 164–176. [[CrossRef](#)]
74. Connor, N.P.; Suzuki, T.; Sewall, G.K.; Lee, K.; Heisey, D.M. Neuromuscular Junction Changes in Aged Rat Thyroarytenoid Muscle. *Ann. Otol. Rhinol. Laryngol.* **2002**, *111*, 579–586. [[CrossRef](#)]
75. Tiago, R.S.L.; de Pontes, P.A.L.; de Brasil, O.O.C. do Quantitative analysis of myelinic fibers in human laryngeal nerves according to age. *Braz. J. Otorhinolaryngol.* **2008**, *74*, 45–52. [[CrossRef](#)]
76. D’haeseleer, E.; Depypere, H.; Claeys, S.; Baudonck, N.; Van Lierde, K. The Impact of Hormone Therapy on Vocal Quality in Postmenopausal Women. *J. Voice* **2012**, *26*, 671.e1–671.e7. [[CrossRef](#)]
77. Suppa, A.; Asci, F.; Saggio, G.; Marsili, L.; Casali, D.; Zarezadeh, Z.; Ruoppolo, G.; Berardelli, A.; Costantini, G. Voice analysis in adductor spasmodic dysphonia: Objective diagnosis and response to botulinum toxin. *Parkinsonism Relat. Disord.* **2020**, *73*, 23–30. [[CrossRef](#)]
78. Bai, X. Biomarkers of Aging. *Adv. Exp. Med. Biol.* **2018**, *1086*, 217–234. [[CrossRef](#)]
79. Jylhävä, J.; Pedersen, N.L.; Hägg, S. Biological Age Predictors. *EBioMedicine* **2017**, *21*, 29–36. [[CrossRef](#)]
80. Khan, S.S.; Singer, B.D.; Vaughan, D.E. Molecular and physiological manifestations and measurement of aging in humans. *Aging Cell* **2017**, *16*, 624–633. [[CrossRef](#)]
81. Hamczyk, M.R.; Nevado, R.M.; Baretino, A.; Fuster, V.; Andrés, V. Biological Versus Chronological Aging. *J. Am. Coll. Cardiol.* **2020**, *75*, 919–930. [[CrossRef](#)]
82. Mueller, P. The Aging Voice. *Semin Speech Lang.* **1997**, *18*, 159–169. [[CrossRef](#)]
83. de Souza, L.B.R.; Santos, M.M.D. Body mass index and acoustic voice parameters: Is there a relationship? *Braz. J. Otorhinolaryngol.* **2018**, *84*, 410–415. [[CrossRef](#)]
84. Suppa, A.; Marsili, L.; Giovannelli, F.; Stasio, F.D.; Rocchi, L.; Upadhyay, N.; Ruoppolo, G.; Cincotta, M.; Berardelli, A. Abnormal motor cortex excitability during linguistic tasks in adductor-type spasmodic dysphonia. *Eur. J. Neurosci.* **2015**, *42*, 2051–2060. [[CrossRef](#)]



Article

Thin-Film Flexible Wireless Pressure Sensor for Continuous Pressure Monitoring in Medical Applications

Muhammad Farooq ^{1,*}, Talha Iqbal ¹, Patricia Vazquez ¹, Nazar Farid ²,
Sudhin Thampi ³, William Wijns ¹ and Atif Shahzad ¹

¹ Smart Sensors Lab, School of Medicine, National University of Ireland Galway, Galway H91 TK33, Ireland; t.iqbal1@nuigalway.ie (T.I.); patricia.vazquez@nuigalway.ie (P.V.); william.wyns@nuigalway.ie (W.W.); atif.shahzad@nuigalway.ie (A.S.)

² School of Physics, National University of Ireland Galway, Galway H91 TK33, Ireland; nazar.farid@nuigalway.ie

³ Lambe Institute of Translational Research, National University of Ireland Galway, Galway H91 TK33, Ireland; sudhin.thampi@nuigalway.ie

* Correspondence: m.farooq1@nuigalway.ie

Received: 26 October 2020; Accepted: 17 November 2020; Published: 20 November 2020

Abstract: Physiological pressure measurement is one of the most common applications of sensors in healthcare. Particularly, continuous pressure monitoring provides key information for early diagnosis, patient-specific treatment, and preventive healthcare. This paper presents a thin-film flexible wireless pressure sensor for continuous pressure measurement in a wide range of medical applications but mainly focused on interface pressure monitoring during compression therapy to treat venous insufficiency. The sensor is based on a pressure-dependent capacitor (C) and printed inductive coil (L) that form an inductor-capacitor (LC) resonant circuit. A matched reader coil provides an excellent coupling at the fundamental resonance frequency of the sensor. Considering varying requirements of venous ulceration, two versions of the sensor, with different sizes, were finalized after design parameter optimization and fabricated using a cost-effective and simple etching method. A test setup consisting of a glass pressure chamber and a vacuum pump was developed to test and characterize the response of the sensors. Both sensors were tested for a narrow range (0–100 mmHg) and a wide range (0–300 mmHg) to cover most of the physiological pressure measurement applications. Both sensors showed good linearity with high sensitivity in the lower pressure range <100 mmHg, providing a wireless monitoring platform for compression therapy in venous ulceration.

Keywords: pressure sensors; compression therapy; thin-film sensors; wireless sensors; medical pressure monitoring; capacitive sensors; flexible sensors; LC sensor; wound monitoring

1. Introduction

Physiological pressure, including intraocular, intracranial, and cardiovascular pressure, is a key parameter for the assessment of human health and provides opportunities for early diagnosis, personalized therapy, and preventive healthcare [1]. Pressure monitoring has been used in diagnosing lower limb problems, muscle rehabilitation, and wound monitoring [1–4]. A common medical application of non-invasive pressure sensing is the monitoring of compression therapy to treat venous leg ulcers. Venous insufficiency occurs when blood is unable to return to the heart and accumulates in the lower limbs. Chronic venous insufficiency (CVI) may cause swelling, pain, edema, and ulcerations [5,6]. The most effective treatment for CVI is compression therapy, in which a compression bandage is used to apply gradual pressure between the ankle and knee to improve the circulation of blood in the lower

limb [7,8]. The typical pressure range for compression therapy is between 10 and 50 mmHg, where the bandage pressure has direct impact on the healing of ulcer [4,9]. To improve the healing process of venous ulcers, continuous monitoring of applied pressure is essential and has become the focus of current research and commercial solutions. Clinical evidence suggests that compression therapy becomes more effective with a feedback sensing system. This feedback system is achieved by using a pressure sensor. Existing solutions that are commonly used in clinical practice are accurate and robust, but they are mostly tethered, rigid, bulky, and require an additional power supply [2,10,11].

The need for wireless, small scale, lightweight, and mobile sensing solutions has led current research to focus on miniaturized thin-film and microelectromechanical (MEMS) pressure sensing devices [12,13]. Current pressure monitoring technologies are generally based on either pneumatic, fluid-filled, piezoelectric, resistive, or capacitive working principles [14,15]. In a pneumatic pressure sensing system, the force of compression bandage is transferred to air pressure and later this air pressure is converted into an electrical signal for further processing [16]. Pneumatic sensors are cheap, flexible, and thin but they are not suitable for dynamic pressure applications and are prone to temperature drift and hysteresis [7,15]. Fluid-filled pressure sensors are similar to pneumatic pressure sensors, where water or oil is used instead of air [17]. The main drawbacks of a fluid-filled sensing system are the air bubbles in the fluid, leakage risk, and bulkiness [16]. In the piezoelectric sensing technology, when pressure is applied on a piezoelectric material, it gets polarized and generates a voltage differential across the device. The piezoelectric effect is proportional to the applied pressure on the device. Thin-film piezoelectric pressure sensors are used for arterial pulse monitoring, respiratory rate, and integrated with a catheter for intravascular pressure measurements, and biomedical implants [18–20]. Piezoelectric pressure sensors are self-powered, low cost, and good for dynamic pressure applications, but they are not suitable for static pressure measurements due to current leakages [21,22]. In resistive pressure sensing technology, the contact area between the active thin-film resistive layer and the electrodes changes with the applied pressure so the effective resistance of the sensor changes [23,24]. Resistive pressure sensors are easier to fabricate, faster in response, and less expensive than piezoelectric pressure sensors; however, an active power source with additional adapting circuitry is required to enable pressure sensing and they are very sensitive to temperature [25]. In capacitive pressure sensing technology, the distance between capacitor electrodes is a function of the applied pressure. A capacitive sensor can be either an active sensing device where the applied pressure can be measured by the changes in capacitance or more often a passive wireless sensing device by combining it with an inductor coil [8]. The combination of the sensing capacitor and inductor coil makes an inductor-capacitor (LC) resonant tank circuit, which makes it suitable for wireless sensing via inductive coupling with an external antenna. The pressure is measured from relative changes in the resonance frequency of the LC resonant tank [26–28]. Due to this wireless communication between sensor and reader coil, capacitive pressure sensors are more practical for wearable and implantable applications as compared to resistive and piezoelectric sensing technologies that demand wired connection to communicate. Capacitive pressure sensing technology is generally used in MEMS and thin-film pressure sensors. MEMS-based sensors are accurate, miniaturized, wirelessly powered, and are widely used in wearable and implantable applications [29]. However, MEMS sensors are generally rigid and have a complex fabrication process that requires specialized equipment. On the other hand, thin-film based capacitive pressure sensors are flexible, less expensive, and simple to fabricate [30].

In past decades, many commercial solutions have been developed for pressure monitoring during compression therapy with growing research focused on lightweight, flexible, and wireless sensing systems. PicoPress (Microlab Electronica, Ponte S. Nicolo, Italy), air-pack type analyzer (AMI Techno, Tokyo, Japan), Kikuhime pneumatic transducer (Advancis Medical, Nottinghamshire, UK), Medical stocking tester (MST, Salzmann AG (SAG), St. Gallen, Switzerland), SIGaT tester (Ganzoni-Sigvaris, St. Gallen, Switzerland), and Oxford pressure monitor MK II (Talley Ltd., Romsey, UK) are available pneumatic sensor-based solutions to monitor the pressure during compression therapy [31–34].

A comparative study has confirmed that PicoPress and Kikuhime are more accurate compared to SIGaT [31]. PicoPress, Kikuhime, MST, and SIGaT are the most common medical devices focused on clinical applications, with relatively higher costs compared to stand-alone sensors [4,8]. Because of the pneumatic sensing principle, these systems are not appropriate for continuous dynamic pressure measurements [9,34].

On the other hand, Quantum tunneling composite (QTC, Peratech, Richmond, UK), ThruMode Force Sensing Resistor (FSR, Sensitronics, Bow, WA, USA), Interlink FSR (Interlink Electronics Inc., Camarillo, CA, USA), F-Scan (Tekscan, Inc., Boston, MA, USA) and Tactilus (Sensor Products Inc., Madison, NJ, USA) are commercially available piezoresistive pressure sensors being widely used to measure interface pressure during compression therapy [35]. Although these sensors are low-cost, thin, and flexible, they require a wired connection and additional electronics to work which makes the system bulky and impractical for real-time pressure measurements [35,36].

In addition to the commercially available compression therapy monitoring solutions, several research studies on pressure sensors and systems have been reported in the literature. Raj et al. [37] used water-filled polymerizing vinyl chloride (PVC) envelopes connected to an electrical pressure transducer to measure the interface pressure at four positions and reported that only within 6–8 h of daily routine applied pressure falls significantly. Hafner et al. [38] reported a silicone oil-filled pressure sensing system to train healthcare staff for an optimal compression in venous ulcer patient management; however, no details about the effect of temperature, hysteresis, and dynamic pressure are reported. Barbenel et al. [16] demonstrated a pressure sensing system for interface pressure using PVC probes filled with vegetable oil and was only limited to a pressure range of 0–37.5 mmHg. Burke et al. [4] developed an interface pressure monitoring system using four commercially available force sensors after integrated with a microcontroller and was capable to work in a range of 0–96 mmHg. However, there was observed a large hysteresis and lack of repeatability. Mehmood et al. [39] reported a telemetric mobile-based sub-bandage for monitoring the pressure and moisture of wounds but because of improper integration of commercial sensors, the system size was big. Casey et al. [8] reported a wearable capacitive flexible pressure sensing technology for sub-acute compression therapy monitoring. This flexible sensor array is built on active capacitor-based pressure sensing. Therefore, it requires a connected power supply and control unit. Farooqui et al. [40] reported a low-cost inkjet-printed wireless sensing system for chronic wound monitoring by measuring the pH level and physical pressure at the wound site. Rahimi et al. [41] has proposed an LC wireless strain sensor for wound monitoring by directly printing the conductive traces on the wound dressing but linearity was limited to 35% strain and no details reported about the repeatability and reliability of the system. Deng et al. [42] fabricated an LC wireless sensor for wound monitoring with a sensitivity of 270 kHz/mmHg in the range between 0 and 200 mmHg.

The majority of implantable or wearable sensors are based in LC systems due to the wireless communication between sensor and reader coil. Fonseca et al. [43] presented a very flexible wireless LC pressure sensor that was rollable and foldable to a compact shape for catheter-based delivery. This sensor was tested acutely *in vivo* for greater than 30 days in canine models simulating abdominal aortic aneurysms (AAA). Li et al. [44] reported a low power flexible sensor for intracranial pressure (ICP) monitoring, with a dual-mode operation in piezoelectric and capacitive modes, accuracy and reliability can be improved using dual-mode capability. Chen et al. [45] presented a wireless pressure sensor for continuous intraocular pressure monitoring of glaucoma patients with a long sensing distance and small physical form factor. Lei et al. [3] reported a flexible capacitive pressure sensor for plantar pressure measurements, different ratios of polydimethylsiloxane (PDMS) prepolymer and curing agent were mixed to improve the linearity by tweaking the stiffness factor.

The work presented here shows a flexible thin-film capacitive pressure sensor that can be fabricated using a simple and cost-effective etching process. The proposed sensor can be used in a wide range of medical applications, including intra and extracranial pressure, wound healing, and muscle

rehabilitation monitoring; although, in this instance it has been designed mainly for interface pressure monitoring during compression therapy.

Considering varying ulcer sizes and lower limb curvatures, as well as different positions, two versions of the sensor with different sizes were fabricated, after optimization of their design parameters for best quality factor and resonance frequencies. Nevertheless, both sensors are LC resonant tank circuits and work on a capacitive sensing mechanism. The optimization of such parameters is reported as analytical results. In the experimental work, the performance of these sensors was evaluated over a pressure range of 0–100 mmHg. In addition, both sensors were also tested for a wider pressure range of 0–300 mmHg, as to suit a varying range of medical applications.

The rest of the paper is organized as follows: Section 2 describes the methodology, including the design, fabrication, and validation of the sensor; Section 3 presents the results obtained (analytical and experimental); Sections 4 and 5 provide the final discussion and conclusions, respectively.

2. Materials and Methods

The proposed sensor is based on an LC resonance circuit, where the resonance frequency of the LC circuit is proportional to the applied pressure. The schematic diagram of the wireless sensing system is presented in Figure 1a. By placing multiple sensors under compression bandage as shown in Figure 1b, an array of wireless sensors can be formed to help in delivering a more controlled personalized compression therapy for the fast recovery of venous ulcers. A wearable readout band can keep records of pressure profiles during the daily routine.

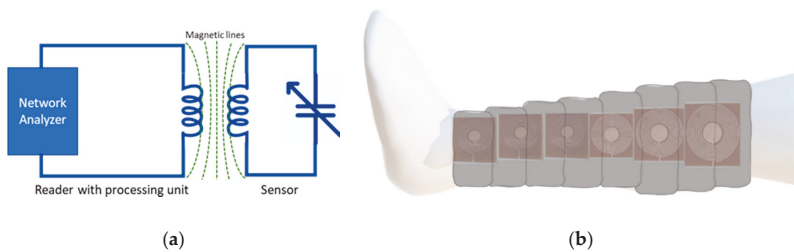


Figure 1. (a) Schematic diagram of wireless LC sensing system showing sensor and reader coil connected with a vector network analyzer (b) An application demonstration using flexible pressure sensors under the compression bandage.

2.1. Sensor Design

The LC sensor is designed as a disc capacitor made of two parallel circular plates, and the inductor is a planar circular spiral coil located around one of the capacitor electrodes suited for a flexible design for a bandage–skin interface. A geometrical representation of the sensor and reader are shown in Figure 2a,b respectively. The resonance frequency (f_o) of the proposed LC sensor depends on the inductance (L_s) and capacitance (C_s) of the sensor, as given in Equation (1):

$$f_o = \frac{1}{2\pi \sqrt{L_s C_s}} \quad (1)$$

The capacitance of the sensor can be calculated as in Equation (2):

$$C_s = \frac{\epsilon_0 \epsilon_r \pi r^2}{d} \quad (2)$$

where ϵ_0 is the permittivity of free space, ϵ_r is the relative permittivity of dielectric material in the capacitor, and r is the radius of the disk capacitor. The inductance of the planar spiral inductor

is calculated using its current sheet expression [46], which depends on the inner diameter d_{in} , outer diameter d_{out} and number of turns N , as given in Equation (3):

$$L_s = \frac{\mu_0 N^2 d_{avg} C_1}{2} \left(\ln(C_2/\tau) + C_3\tau + C_4\tau^2 \right) \quad (3)$$

where μ_0 is the permeability of free space, N is the number of turns, $d_{avg} = \frac{(d_{in}+d_{out})}{2}$, $\tau = \frac{(d_{out}-d_{in})}{(d_{out}+d_{in})}$, and C_1 , C_2 , C_3 and C_4 are the coefficients for the current sheet expression, which are 1, 2.46, 0, and 0.2 for a circular design [46].

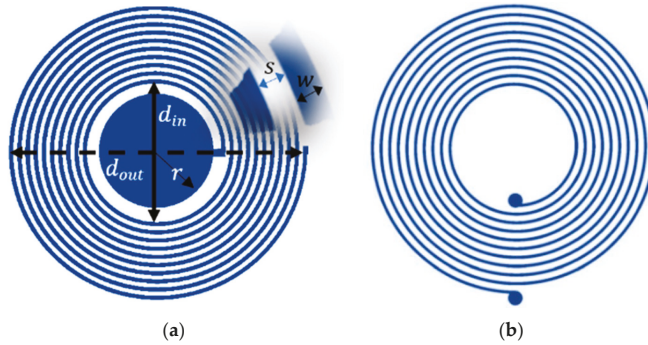


Figure 2. Geometrical representation of the proposed LC system: (a) LC sensor with a capacitor of the radius (r) and planar inductor with an inner diameter (d_{in}) shown with a solid line, outer diameter (d_{out}) shown with a dotted line, trace separation (s) and trace width (w); (b) Reader antenna with the same design parameters (d_{in} , d_{out} , s , w).

Parasitic Components

The inductive part of the sensor, consisting of circular spirals, can be modeled accurately using lumped elements. Its elements are an inductor (L_s), a parasitic resistance (R_{tot}), and parasitic capacitance (C_p), where L_s and R_{tot} are in series in parallel to C_p as shown in Figure 3a.

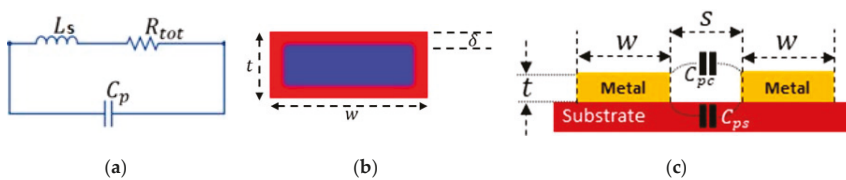


Figure 3. Parasitic effects: (a) Lumped model of the spiral inductor showing the inductor (L_s), parasitic resistance and capacitance; (b) Skin effect on a rectangular conductor, with current flowing only in the red area; (c) The parasitic capacitance is due to the air gap between coil turns (C_{pc}) and the substrate material (C_{ps}).

One of the major parasitic effects that play a major role in the quality factor of the inductor is the series resistance, which is modeled as R_{tot} in this paper. A large R_{tot} will result in a poor quality factor of the inductor in the sensor, as well as in the reader coil. This R_{tot} can be represented by Equation (4), which includes direct current resistance (R_{dc}) and alternating current resistance (R_{ac}).

$$R_{tot} = R_{dc} + R_{ac} \quad (4)$$

R_{dc} can be calculated according to Equation (5), where ρ is the resistivity of the conductor, l is the length of the spiral conductor, w is the trace width and t is the trace thickness.

$$R_{dc} = \frac{\rho l}{wt} \quad (5)$$

For a spiral inductor with N number of turns, outer and inner diameters d_{out} and d_{in} , the length of the conductive traces can be calculated using Equation (6).

$$l = \frac{\pi N(d_{in} + d_{out})}{2} \quad (6)$$

The component R_{ac} in Equation (4) is affected by the values of R_{skin} and R_{prox} , which correspond to the skin effect and proximity effect, respectively:

$$R_{ac} = R_{skin} + R_{prox} \quad (7)$$

The skin effect occurs at higher frequencies when current does not flow through the complete cross-sectional area of the conductor, and it starts flowing only through its surface as shown in Figure 3b, which increases the effective resistance. In Figure 3b, the red color represents the skin depth (δ) for current flow and the blue color shows the area without electric current. This effect is represented by the skin depth δ . The mathematical expression to compute R_{skin} is given in Equation (8) [47]. Here μ_o is the permeability constant and μ_r is the relative permeability of the conductor and f is the operational frequency.

$$R_{skin} = \frac{\rho l}{w\delta\left(1 - e^{-\frac{t}{\delta}}\right)\left(1 + \frac{t}{w}\right)}, \text{ where } \delta = \sqrt{\frac{\rho}{\pi\mu_o\mu_r f}} \quad (8)$$

The proximity effect is another major contributor to R_{ac} that becomes significant above a frequency specific to the design, known as crowding frequency, f_{crit} . In the signal frequencies above f_{crit} , magnetic forces surrounding the conductor become significant and result in a nonuniform current flow through the conductor. This redistribution of the current causes an increase in effective resistance and can be calculated through Equation (9) [48].

$$R_{prox} = \frac{R_{dc}}{10} \left(\frac{f}{f_{crit}}\right)^2, \text{ where } f_{crit} = \frac{3.1(w + s)\rho}{2\pi\mu_o w^2 t} \quad (9)$$

The parasitic capacitance between the nearby turns can be computed from Equation (10) [49,50], where α and β are 0.9 and 0.1, respectively, and represent the parasitic contribution due to the air gap between the coil turns, and the gap between the metallic tracks and the substrate, as shown in Figure 3c. ϵ_{rc} and ϵ_{rs} are the relative permittivity of air and substrate material respectively.

$$C_p = C_{pc} + C_{ps} = \frac{lt\epsilon_o}{s} (\alpha\epsilon_{rc} + \beta\epsilon_{rs}) \quad (10)$$

The value of the self-resonance frequency f_{SRF} of an inductor is critical, as above this frequency the parasitic capacitance of the inductor becomes dominant. The f_{SRF} can be calculated using Equation (11) [50].

$$f_{SRF} = \frac{1}{2\pi \sqrt{L_s C_p}} \quad (11)$$

Finally, the quality factor of the LC sensor is given by Equation (12) [51].

$$QF = \frac{1}{R_{tot}} \sqrt{\frac{L_s}{C_s}} \quad (12)$$

2.2. Device Fabrication

After the optimization of design parameters that is discussed in Section 3.1, a wet etching process was used to fabricate the two different sensors and their reader antennas. Figure 4 shows the stages in the fabrication process. In step I, as shown in Figure 4a, the mask of the sensor was directly printed on a 50 μm thick copper-coated polyimide film (Flexible isolating circuit 50 μm -coppered 35 μm -1 side, CIF, Buc, France) with a LaserJet printer (HP M553, HP Technology, Dublin, Ireland). In step II, the printed copper sheets were immersed in an etchant solution (CIF, Boosted ferric chloride solution). After manual stirring for 15 min at room temperature, all the unwanted copper was removed as shown in Figure 4b, and the patterned sheet was washed with hot water. Acetone was used to remove the ink particles from the copper surface after the etching process. In the next step, a polydimethylsiloxane (PDMS) layer (Ultra-thin film, 30° shore A hardness, Silex Ltd., Bordon, UK) of 200 μm thickness was cut into a circular shape equal to the diameter of the capacitor electrodes and was placed on the bottom electrode as shown in Figure 4c. PDMS is widely used as a dielectric layer in capacitive pressure sensors due to its low Young's modulus and compressibility. An adhesive layer composed of polypropylene and synthetic rubber of 90 μm thickness (Tesa64621, Tesa, Norderstedt, Germany) was placed around the PDMS layer as shown in Figure 4d. In the final step, the top layer of the sensor was folded onto the PDMS layer for the final assembly of the sensor. Figure 4e,f shows the top and bottom views of the fabricated sensor. The reader antenna was also fabricated by the same etching procedure, and flexible multithread wires were soldered to connect with a Sub-Miniature version A (SMA) connector.

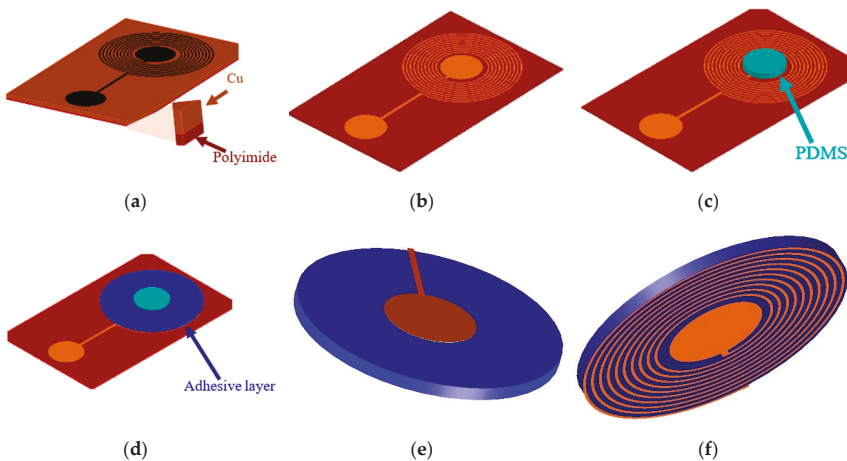


Figure 4. Fabrication process: (a) Copper-coated polyimide film with ink printed mask. (b) Etched pattern showing capacitor electrodes and planar inductor spirals. (c) Dielectric layer of PDMS elastomer. (d) Adhesive layer placement around the PDMS dielectric. (e) Top view and (f) bottom view of the LC sensor.

2.3. Device Validation

To test the fabricated system (sensor with reader coils), a bench-test model was developed using a vector network analyzer (VNA E5063, Keysight Technologies Inc., Santa Rosa, CA, USA),

a high-pressure glass bottle (Pressure+ 1000, Duran, Mainz, Germany), and a digital pressure gauge (Traceable 3462, Fisher Scientific Ltd., Loughborough, UK), as shown in Figure 5. The sensor was placed inside the pressure chamber and its response recorded using the reader antenna, which was placed outside the wall of the chamber. The pressure was varied using a vacuum pump (FB70155 Pump, Fisher Scientific Ltd., Loughborough, UK) to produce positive pressure inside the chamber, which was measured as well by the digital pressure gauge. The input impedance of the VNA was 50Ω . A frequency sweep was generated from the VNA to observe the variation in resonance frequency against the varying pressure, and the S parameters of the sensor were recorded simultaneously.

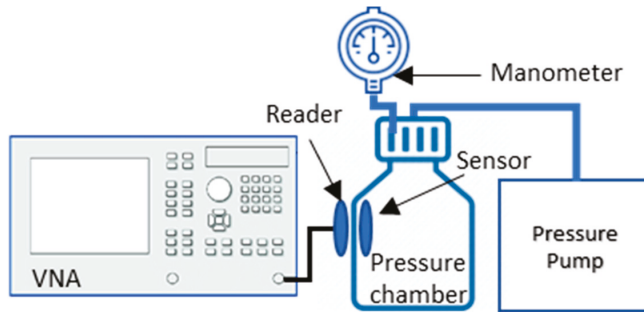


Figure 5. Bench test setup for sensor validation where reader coil is connected to vector network analyzer and sensor is kept inside the pressure chamber and pressure is varied using pressure pump.

3. Results

The results presented in this paper comprise of the outcomes of two types of investigation: analytical investigations (Section 3.1) and experimental investigations (Section 3.2). The analytical investigations are performed for optimization of design parameters (d_{out} , N , w , s) to achieve the best quality factor (QF), and lower resonance frequencies (f_o). The experimental investigations are performed to test and characterize the performance of the two fabricated prototype sensors on suitable testbeds.

3.1. Analytical Results: Numerical Estimation of Sensor Parameters

Sensor optimization was done in two steps. In the first step, the outer diameter (d_{out}) and the number of turns (N) of the inductor were optimized while keeping the trace width (w) and trace separation (s) constant. In the second step, after selecting the optimal values of d_{out} and N , both s and w were adjusted to achieve the best quality factor (QF) with a low resonance frequency (f_o).

3.1.1. Optimization of Outer Diameter (d_{out}) and Number of Turns (N)

Before the fabrication stage of the sensor, MATLAB numerical modeling was performed to achieve the best quality factor (QF) within low resonance frequency (f_o) range to achieve a better signal to noise ratio (SNR). The two different designs of the sensor, sensor 1 (S_1) and sensor 2 (S_2), were characterized according to their individual parameters. S_1 was modeled for different d_{out} values, between 36 and 45 mm, and a varying N from 1 to 10, while keeping $s = w = 500 \mu\text{m}$. As can be seen from the data point shown in Figure 6, the best QF was 106.4, with a correspondent resonance frequency of 17.147 MHz, when d_{out} and N were 45 mm and 10 respectively. However, to keep the sensor size small, we selected $d_{out} = 40$ mm and $N = 10$ for the fabrication as there was no significant loss in QF (97.46), and f_o was also low (19.188 MHz).

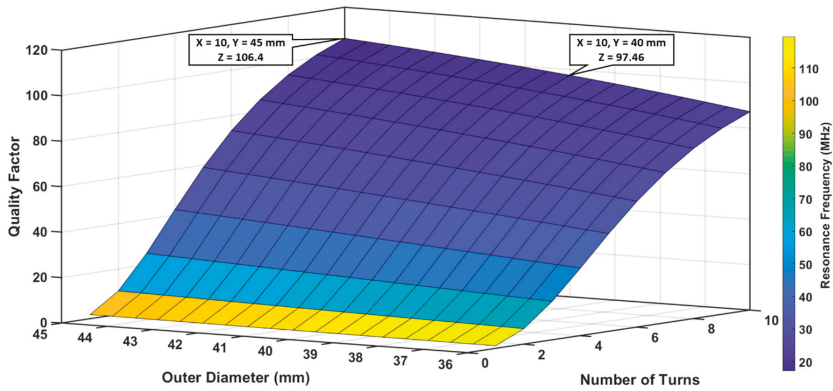


Figure 6. Analysis of S_1 quality factor and resonance frequency for different number of turns and outer diameters, when trace separation and width were kept constant at 500 μm .

A similar model was computed for S_2 as shown in Figure 7. In this case, the objective was to design a relatively small sensor; therefore, d_{out} was varied between 10 and 14 mm and N between 1 and 5 turns, while s and w were kept constant at 500 and 200 μm respectively. For S_2 , the highest QF was ~ 32 , with a f_0 of 222.4 MHz for $d_{out} = 14$ mm and $N = 5$; however, we selected $d_{out} = 12$ mm and $N = 5$ to achieve an optimal set of QF (23.93) and f_0 (259.44 MHz) against the size of the sensor.

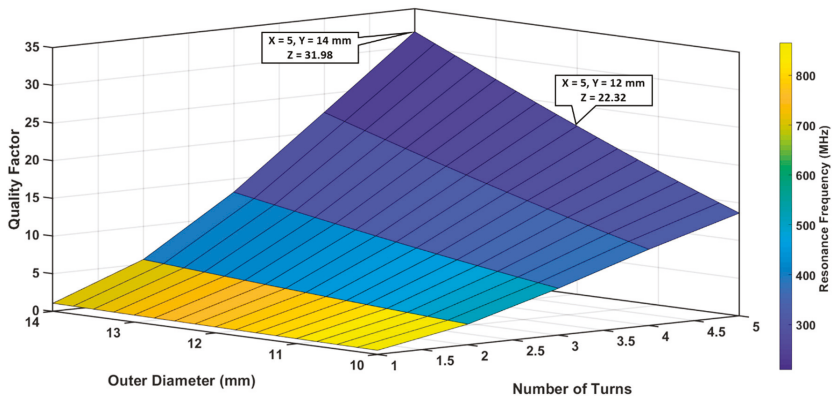


Figure 7. Analysis of S_2 quality factor and resonance frequency for different number of turns and outer diameters, when trace separation was 500 μm and trace width was 200 μm .

3.1.2. Optimization of Trace Width (w) and Trace Separation (s)

Trace width and trace separation also affect the QF and resonance frequency; therefore, complete numerical modeling was performed for the selection of the trace geometry. QF and f_0 were analyzed for different values of s and w , while the number of turns and d_{out} were fixed this time. Both trace width (w) and trace gap (s) were varied within the maximum allowable range to fit within the limits of given sensor size and number of turns. For S_1 , values of s and w were modeled between 200 and 600 μm and d_{out} and N were 40 mm and 10, respectively. As shown in Figure 8, the highest QF (103.5) was observed for $s = 325$ μm and $w = 400$ μm , with a resonance frequency of 16.82 MHz. For an equally distributed pattern with a trace width (w) and trace gap (s) of 500 μm , a very small loss in QF ($\sim 5\%$) was observed, therefore, $w = s = 500$ μm were chosen for the design of S_1 .

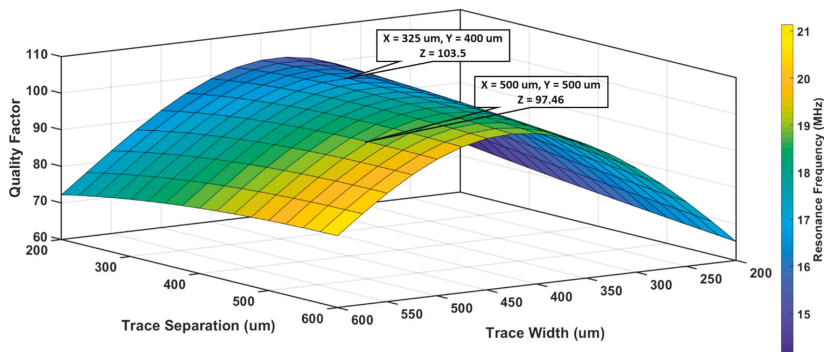


Figure 8. Analysis of S_1 quality factor and resonance frequency for different trace separation and trace width, when the number of turns and outer diameter were 10 and 40 mm, respectively.

S_2 was modeled by varying s between 200 and 300 μm and w from 200 and 500 μm , while keeping $d_{out} = 12$ mm and $N = 5$ fixed, as shown in Figure 9. Maximum QF was 23.93 with a resonance frequency of 259.44 MHz for a combination of $s = 500$ μm and $w = 200$ μm . As both the QF and the resonance frequency of S_2 were very sensitive to trace width and separation, the combination of s and w that produced the best QF were chosen for S_2 .

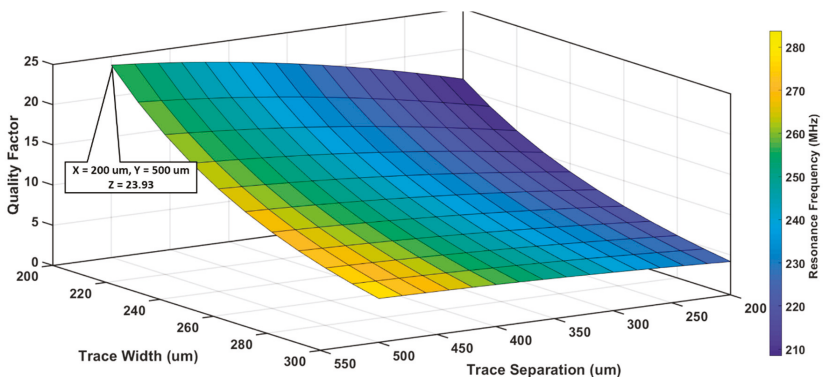
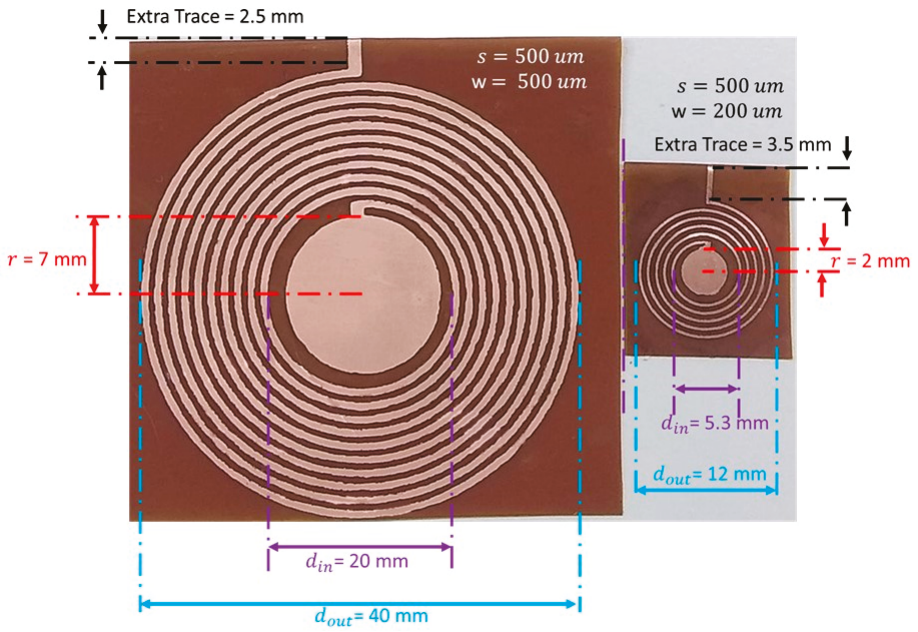


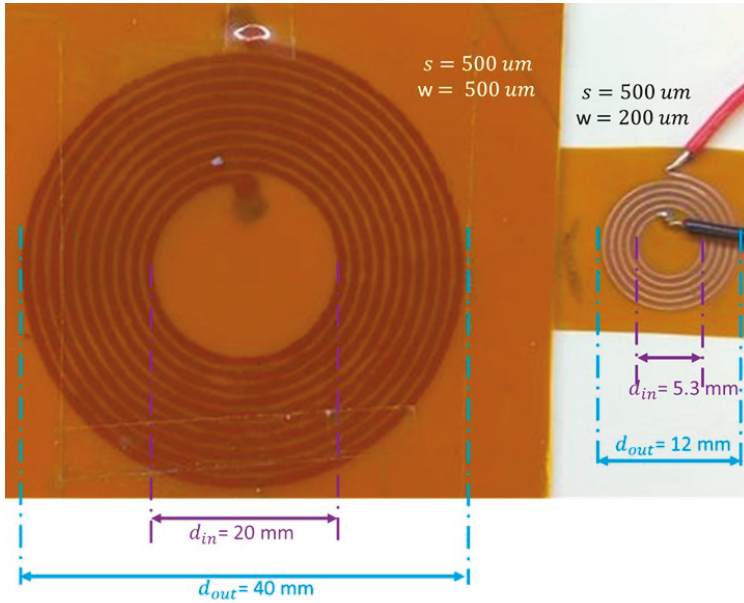
Figure 9. S_2 quality factor and resonance frequency analysis for different trace separation and trace width when the number of turns and outer diameter were 5 and 12 mm, respectively.

3.2. Experimental Prototype and Results

After selecting the optimized design parameters (d_{out} , N , s , and w), two sensors, of outer diameters 40 and 12 mm (shown in Figure 10), were fabricated and tested using the test-bench described in Section 2.3. The key design parameters, results, and operating frequencies for both sensors and respective reader coils are listed in Table 1.



(a)



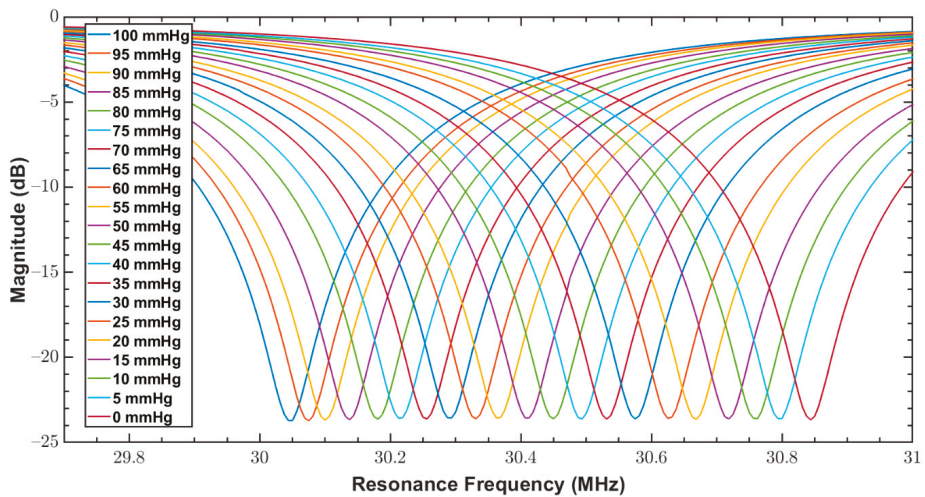
(b)

Figure 10. Fully labeled images of final prototypes, (a) Fabricated wireless LC resonance sensors: left, S_1 (40 mm in diameter), and right, S_2 (12 mm diameter); (b) Reader coils for S_1 (left) and S_2 (right).

Table 1. Key design parameters and results for both sensors (S_1, S_2) and their readers (R_1, R_2).

Parameter/Result	S_1/R_1 (40 mm)	S_2/R_2 (12 mm)
Outer diameter, d_{out} (mm)	40	12
Inner diameter, d_{in} (mm)	20	5.3
Trace width, w (mm)	0.5	0.2
Trace separation, s (mm)	0.5	0.5
Number of turns, N	10	5
Capacitor electrode radius, r (mm)	7	2
Dielectric layer thickness, d (μm)	200	200
Spiral length, l (mm)	942.5	135.9
Skin depth, δ (μm)	15.15	4.12
DC resistance, R_{dc} (Ω)	0.9371	0.3377
AC resistance, R_{ac} (Ω)	3.7755	17.0536
Relative permittivity of PDMS [3], ϵ_{r_PDMS}	2.65	2.65
Calculated resonance frequency, f_{o_cal} (MHz)	19.188	259.44
Measured resonance frequency, f_{o_meas} (MHz)	30.843750	274.10
Sensitivity, m (kHz/mmHg)	-8.110	-65.48
Capacitance, C_s (pF)	18.06	1.4743
Parasitic capacitance, C_p (pF)	0.72434	0.10442
Inductance, L_s (μH)	3.8095	0.25572
Self-resonance frequency, f_{SRF} (MHz)	95.81	974.80
Quality factor, QF at f_o	97.4584	23.92

As discussed in Section 1, since bandage pressure varies between 10 and 60 mmHg during compression therapy, both fabricated sensors were tested for a pressure range of 0 to 100 mmHg. The reader coil connected with the network analyzer was magnetically coupled with the sensor, and the response of the sensor over varying pressure was measured. The measurements from VNA were triggered at an interval of 5 mmHg for a narrow range of 0–100 mmHg. These measurements are the reflection coefficients (S_{11} parameter) and are shown in Figures 11 and 12 for the sensors S_1 and S_2 , respectively.

**Figure 11.** Reflection coefficients (S_{11} parameter) of S_1 for a pressure range of 0 to 100 mmHg.

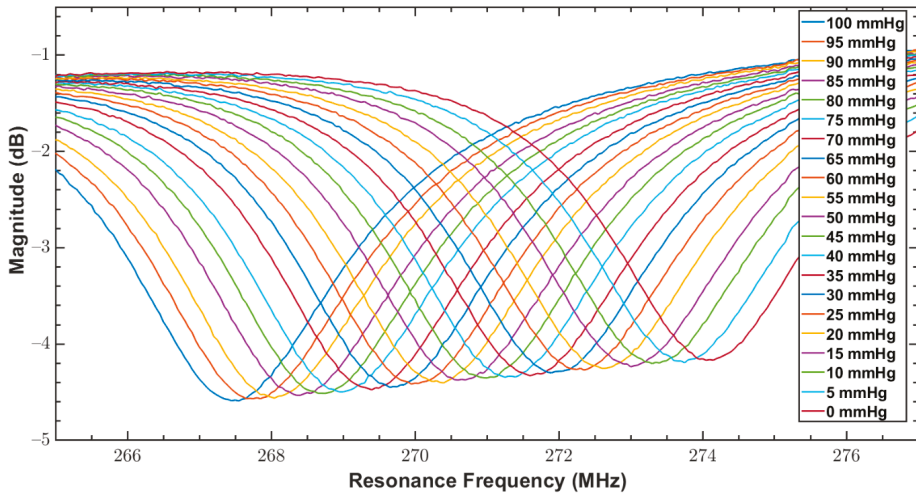


Figure 12. Reflection coefficients (S11 parameter) of S2 for a pressure range of 0 to 100 mmHg.

In addition to the compression therapy monitoring, the proposed sensors could be used for other medical applications, including physiological pressure measurement. Therefore, both sensors were also tested over a wider range of 0 to 300 mmHg that covers almost the entire physiological pressure range. The measurements from VNA were triggered at an interval of 25 mmHg for a wide range of 0–300 mmHg. Figures 13 and 14 show the measured reflection coefficients (S11 parameter) of S₁ and S₂ over this broad pressure range.

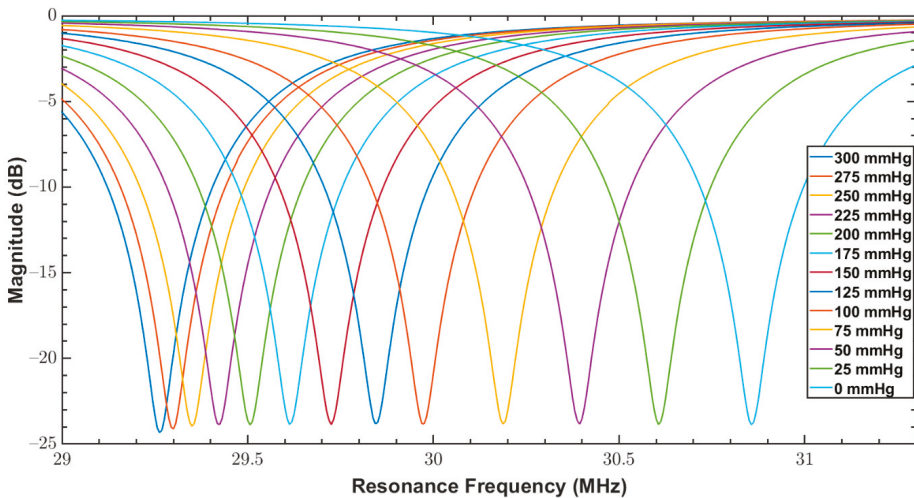


Figure 13. Reflection coefficients (S11 parameter) of S₁ for a pressure range between 0 to 300 mmHg.

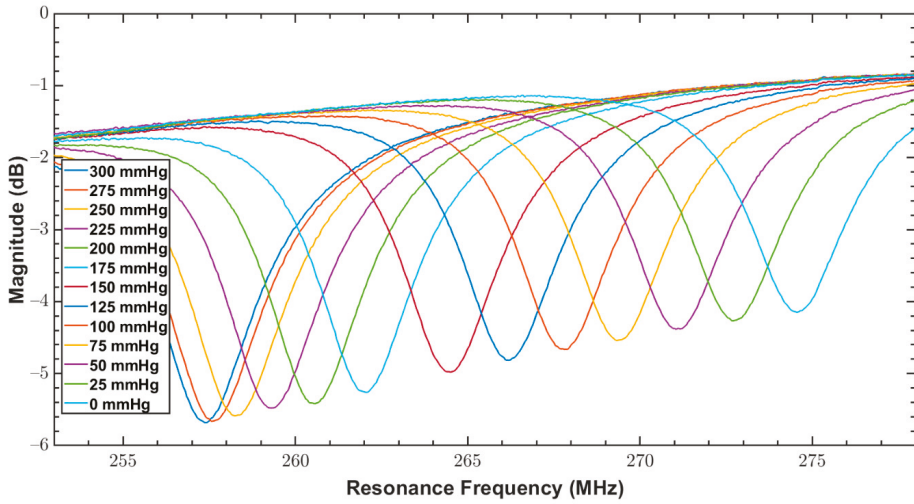


Figure 14. Reflection coefficients (S_{11} parameter) of S_2 for a pressure range between 0 to 300 mmHg.

As the response of both the sensors was linear within the targeted pressure range of 0 to 100 mmHg, a first-order polynomial was fitted over the measured response of the sensors. The coefficients of the linear fitted model are given in Table 2. The measure sensor response (dotted) and fitted curve (solid) for both sensors are shown in Figure 15.

Table 2. Coefficients of the polynomial equation ($f(P) = m \times P + \beta$; where P is pressure) curve fitting between measured resonance frequencies and applied pressure.

Parameters	S_1 (40 mm)	S_2 (12 mm)
m (sensitivity)	-8.11×10^3	-65.48×10^3
β (f_o)	3.083×10^7	2.74×10^8
R^2 (goodness of fit)	0.9977	0.9989

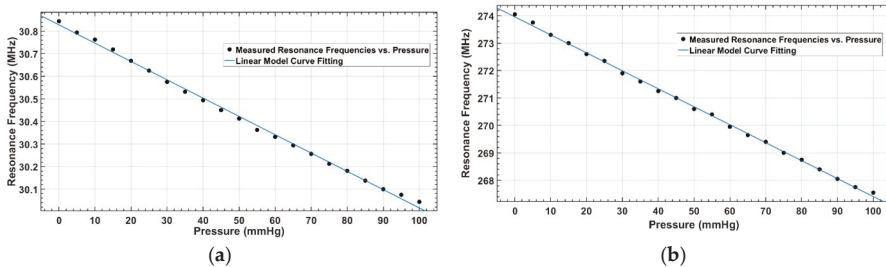


Figure 15. Linear model fitting for a pressure range of 0 to 100 mmHg: (a) Measured response (dotted) and linear fit (solid) of S_1 ; (b) Measured response (dotted) and linear fit (solid) of S_2 .

In the sensor response over a wide range of pressure up to 300 mmHg, a nonlinearity, associated with compression saturation of the dielectric layer, was observed at higher pressures as shown in Figure 16. Therefore, a second-order polynomial function was fitted to the measured response to obtain a model relating the resonance frequency to the pressure. The values of R-square (goodness of fit) and the model coefficients are listed in Table 3.

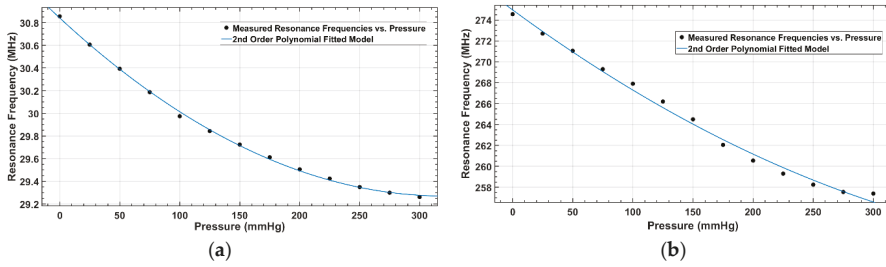


Figure 16. Nonlinear model fitting for a pressure range of 0 to 300 mmHg: (a) Measured response (dotted) and linear fit (solid) of S_1 ; (b) Measured response (dotted) and linear fit (solid) of S_2 .

Table 3. Coefficients of 2nd order polynomial ($f(P) = a \times P^2 + b \times P + \beta$; where P is pressure) curve fitting between measured resonance frequencies and applied pressure.

Parameters	S_1 (40 mm)	S_2 (12 mm)
a	15.27	77.12
b	-9.797×10^3	-8.439×10^4
$\beta (f_0)$	3.084×10^7	2.75×10^8
R^2 (goodness of fit)	0.9991	0.993

To assess the repeatability of pressure measurement with both sensors, the response of the sensors for six different pressure points between 0 and 100 mmHg was measured repeatedly for 10 cycles. Figures 17 and 18 show the repeatability of S_1 and S_2 , respectively. The mean values of the frequency response against applied pressure (f_u) and standard deviation (σ) of 10 repeated measurements at 6 pressure points (100, 80, 60, 40, 20, 0) mmHg are given in Table 4.

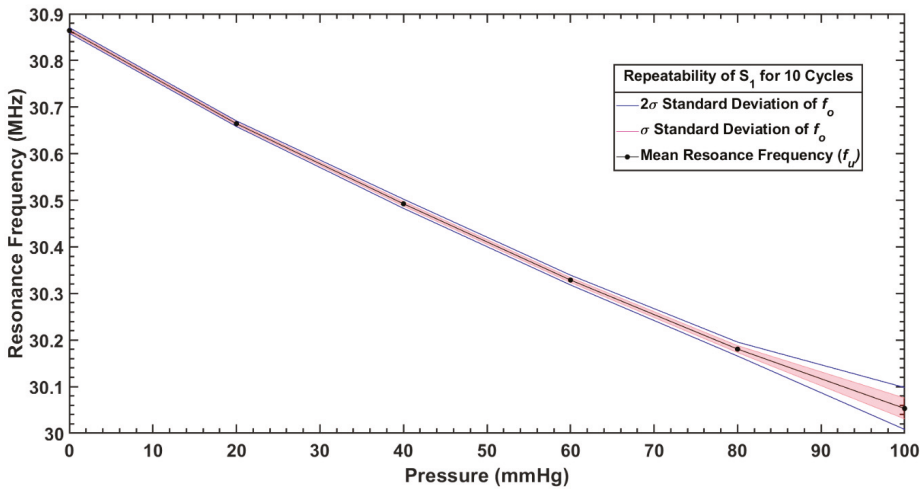


Figure 17. Repeatability of measurements with S_1 over 10 cycles.

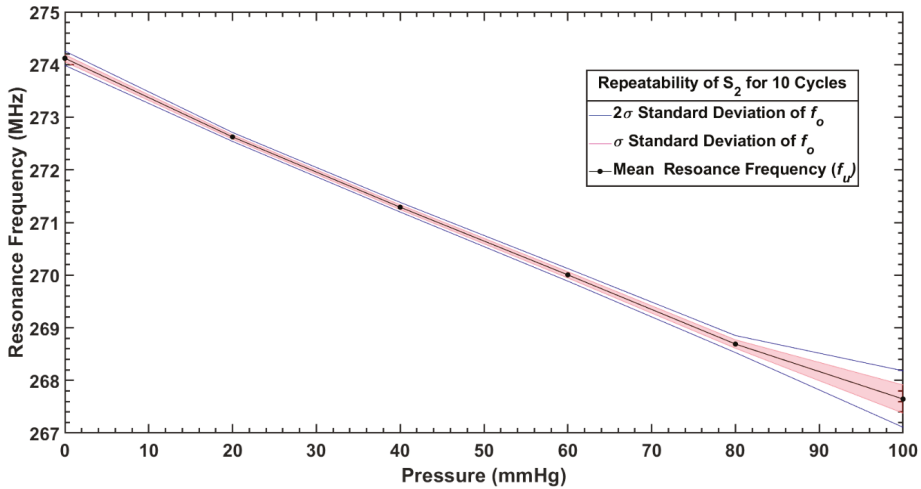


Figure 18. Repeatability of measurements with S_2 over 10 cycles.

Table 4. Mean value and standard deviation of measured resonance frequencies when both sensors were tested under different pressures for 10 cycles.

Sensor	Parameter	Pressure (mmHg)					
		100	80	60	40	20	0
S_1	Mean, f_u (MHz)	30.05	30.18	30.32	30.49	30.66	30.86
	Standard deviation, σ (kHz)	22.48	7.48	5.27	4.93	3.01	3.01
S_2	Mean, f_u (MHz)	267.64	268.69	270.00	271.29	272.62	274.12
	Standard deviation, σ (kHz)	26.81	80.96	59.86	45.94	42.49	67.49

4. Discussion

An LC pressure sensing system is developed to measure the pressure in compression therapy due to wireless communication between sensor and reader coil. Optimization of the sensors is essential to achieve the best quality factor and resonance frequency while keeping the sensor size limited. Optimized values of outer diameter (d_{out}) and the number of turns (N), trace width (w), and trace separation (s) are listed in Table 1. The parasitic components of the sensor which are parasitic capacitance and parasitic resistance at resonance frequency were analyzed through numerical modeling and their values are reported in Table 1. The reported sensors were fabricated using a wet etching process, which is cost-effective and very simple but comes at the cost of less control on trace widths. In these circumstances, the thinnest trace width achieved was 200 μm . Both sensors were characterized using a bench test setup that was developed during this research work. Both sensors showed good linearity and repeatability for a pressure <100 mmHg.

As shown in Figure 15, the response of both designed sensors was linear over a pressure range of 0–100 mmHg, with a sensitivity of 8 kHz/mmHg for S_1 and 65 kHz/mmHg for S_2 . The sensor response was observed as nonlinear at the higher pressure range of 0 to 300 mmHg, as shown in Figure 16. This is due to the nonlinear effect of the compression saturation of the dielectric layer of the capacitor in the sensor. Up to 100 mmHg, the sensitivity of S_1 was 8.11 kHz/mmHg, which was reduced at higher pressure due to the dielectric layer saturation. Similar behavior was noticed for S_2 , where sensitivity was 65.48 kHz/mmHg up to 100 mmHg, and was reduced when the sensor was loaded with higher values of pressure.

Both sensors offered good repeatability as shown in Figures 17 and 18, for a pressure range <80 mmHg; however, variability in measurements started growing in the sensor response for higher applied pressures (>80 mmHg), due to the already mentioned hysteresis of the dielectric layer. As it can be noticed from Table 4 that the average repeatability for both the sensors over the pressure range of 0–100 mmHg is slightly larger than the sensitivity per mmHg, the measurement uncertainty is estimated as less than ± 1 mmHg.

From Table 1, it can be noticed that QF of S_1 was better than S_2 , which is due to the exponential increase of the ac resistance at higher frequencies for S_2 caused by the skin effect. In addition, by comparing the amplitude of S parameters of both sensors in Figures 13 and 14, it is quite clear that S_1 has a better signal to noise ratio (SNR) compared to S_2 .

There was noticed a difference between the calculated and measured resonance frequencies of both sensors (S_1 and S_2) was due to numerous possible reasons. The first possible reason might be the value of the PDMS dielectric constant (ϵ_{r_PDMS}), which reported between 2.3 and 2.8 in literature [52]; however, for this research ϵ_{r_PDMS} was selected 2.65 as stated in Table 1. The second reason for this difference might be due to the roughness of conductive traces caused by an over-etching effect during the fabrication process. This difference was greater for S_1 due to the uneven distribution of the dielectric layer and air gaps between the capacitor plates, which were relatively bigger as compared to S_2 . In the future, a more controlled fabrication process can be used to improve the etching process and dielectric layer deposition to overcome the mismatch between analytical and real values of sensor parameters.

A comparison of the developed sensors with previously reported systems is given in Table 5. It includes sensors developed explicitly for wound compression therapy, and as an extension, implantable sensors that measure bodily pressures in different locations. Although not designed specifically for the application targeted in this work, these implantable sensors are based on the same sensing concept of LC systems and operate in similar pressure ranges (as shown in Table 5). From the observation of the values listed in the table, it is noticeable that the sensitivity of the S_2 sensor, 65.48 kHz/mmHg, is comparable with the prototypes reported in the literature. This is, in the author's view, a noteworthy achievement, considering the fact that the sensor proposed here is based on a very simple and non-expensive fabrication method. By contrast, most states of the art sensors are based on microfabrication techniques, which are very expensive and laborious.

Table 5. Comparative analysis of this study with the literature.

Study	Sensing Mechanism	Methodology	Linear Operational Range	Sensitivity	Fabrication Cost	Application
This study	Capacitive	Laserjet printing, Copper etching, Sandwicing of PDMS layer	0–100 mmHg	8.11 kHz/mmHg 65.48 kHz/mmHg	Low	Interface pressure monitoring during compression therapy
Deng et al. [42]	Capacitive	Si wafer moulding, PDMS casting, Conductive printing, Packaging	0–200 mmHg	270 kHz/mmHg	High	Wound monitoring
Casey et al. [8]	Capacitive	Micro-machining, Si wafer moulding, Electrodes patterning, Component mounting	10–80 mmHg	N/A	High	Sub-bandage pressure measurements
Farooqui et al. [40]	Capacitive	Screen printing, Conductive printing, Component mounting, Packaging	5–100 mmHg	0.0523 pF/mmHg	Medium	Smart bandage for chronic wounds
Fonseca et al. [43]	Capacitive	Standard lithography, Wet-chemical etching, Laser-cutting	70–120 mmHg	5.76 kHz/mmHg	High	Implantable pressure sensing
Chen et al. [45]	Capacitive	Oxide patterning, Coating, and patterning, Metal deposition, Deep Si etching	0–100 mmHg	160 kHz/mmHg	High	Intraocular pressure monitoring
Li et al. [44]	Capacitive	Microfabrication	0–50 mmHg	0.419 kHz/mmHg	High	Intracranial blood pressure monitoring
Rahimi et al. [41]	Inductive	Corona treatments, Laser patterning of mask, Screen printing of electrodes, Temperature curing	0–35%	150 kHz/% strain	High	Wound monitoring
Mehmood et al. [39]	Resistive (FSR)	Off the shelf sensors integration with electronics, Biocompatible coating	0–60 mmHg	N/A	Medium	Sub-bandage pressure and wound moisture
Burke et al. [4]	Resistive (FSR)	Off the shelf sensors integration with electronics	0–96 mmHg	31.27 mV/mmHg	Medium	Sub-bandage pressure during venous compression therapy

5. Conclusions

This work presented the design of a wireless capacitive pressure sensor of low-cost fabrication for medical applications. In particular, the sensor is designed to be used for monitoring of compression therapy in venous leg ulcers. The sensor design was optimized to achieve an optimal quality factor and resonance frequency by numerical modeling of the design parameters. The proposed thin-film flexible wireless pressure sensor was fabricated using a simple and cost-effective fabrication method. Two versions of the sensors, with 40 and 12 mm outer diameters respectively, were developed and characterized between 0–100 and 0–300 mmHg to cover the pressure range of compression therapy and the nominal range of all other physiological applications. A bench test setup was also developed for sensor validation using a glass pressure bottle, pressure pump, and a network analyzer. Both sensors showed good sensitivity, linearity, and repeatability for the lower pressure regime (0–100 mmHg). A MATLAB curve-fitting tool was used to model the relationship between the shift in resonance frequency and the change in pressure.

The focus of this research work was on the early prototype development of the sensor, which is characterized by the benchtop model. However, in the future, improved and miniaturized prototypes will be fabricated by a more controlled fabrication process, and an extensive study will be performed on human subjects to validate the effectiveness. The miniaturization and replacement of the dielectric material used in the proposed sensors with other elastomeric polymers, can improve the linearity, sensitivity, and repeatability of the sensor and will make it more suitable for numerous medical applications.

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

Funding: The research leading to this publication was funded by the Science Foundation Ireland Research Professorship Award (grant no. 15/RP/2765), the Government of Ireland Disruptive Technology Innovation Fund (grant no. DT20180031A), and Enterprise Ireland (grant no. CF-2019-1125).

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

References

1. Trung, T.Q.; Lee, N.E. Flexible and Stretchable Physical Sensor Integrated Platforms for Wearable Human-Activity Monitoring and Personal Healthcare. *Adv. Mater.* **2016**, *28*, 4338–4372. [[CrossRef](#)] [[PubMed](#)]
2. Liu, Y.; Pharr, M.; Salvatore, G.A. Lab-on-Skin: A Review of Flexible and Stretchable Electronics for Wearable Health Monitoring. *ACS Nano* **2017**, *11*, 9614–9635. [[CrossRef](#)] [[PubMed](#)]
3. Lei, K.F.; Lee, K.F.; Lee, M.Y. Development of a flexible PDMS capacitive tactile sensor with adjustable measurement range for plantar pressure measurement. *Microsyst. Technol.* **2014**, *20*, 1351–1358. [[CrossRef](#)]
4. Burke, M.; Murphy, B.; Geraghty, D. Measurement of sub-bandage pressure during venous compression therapy using flexible force sensors. In Proceedings of the IEEE SENSORS Conference, Valencia, Spain, 2–5 November 2014; pp. 1623–1626. [[CrossRef](#)]
5. Ghosh, S.; Mukhopadhyay, A.; Sikka, M.; Nagla, K.S. Pressure mapping and performance of the compression bandage/garment for venous leg ulcer treatment. *J. Tissue Viability* **2008**, *17*, 82–94. [[CrossRef](#)]
6. Valencia, I.C.; Falabella, A.; Kirsner, R.S.; Eaglstein, W.H. Chronic venous insufficiency and venous leg ulceration. *J. Am. Acad. Dermatol.* **2001**, *44*, 401–424. [[CrossRef](#)]
7. Partsch, H.; Clark, M.; Bassez, S.; Benigni, J.; Becker, F.; Blazek, V.; Caprini, J.; Cornu-Thénard, A.; Hafner, J.; Flour, M.; et al. Measurement of lower leg compression in vivo: Recommendations for the performance of measurements of interface pressure and stiffness. *Dermatol. Surg.* **2006**, *32*, 224–233. [[CrossRef](#)]

8. Casey, V.; McAree, B.; Moloney, M.C.; Grace, P. Wearable sub-bandage pressure measurement system. In Proceedings of the 2010 IEEE Sensors Applications Symposium (SAS), Cancun, Mexico, 9–11 November 2016; pp. 41–45. [\[CrossRef\]](#)
9. Flaud, P.; Bassez, S.; Counord, J.L. Comparative in vitro study of three interface pressure sensors used to evaluate medical compression hosiery. *Dermatol. Surg.* **2010**, *36*, 1930–1940. [\[CrossRef\]](#)
10. Hegarty, M.S.; Grant, E.; Reid, L. An overview of technologies related to care for venous leg ulcers. *IEEE Trans. Inf. Technol. Biomed.* **2010**, *14*, 387–393. [\[CrossRef\]](#)
11. Rahimi, M.; Blaber, A.P.; Menon, C. Towards the evaluation of force-sensing resistors for in situ measurement of interface pressure during leg compression therapy. In Proceedings of the 2016 IEEE Healthcare Innovation Point-Of-Care Technologies Conference (HI-POCT), Cancun, Mexico, 9–11 November 2016; pp. 25–28. [\[CrossRef\]](#)
12. Kaisti, M.; Panula, T.; Leppänen, J.; Punkkinen, R.; Tadi, M.J.; Vasankari, T.; Jaakkola, S.; Kiviniemi, T.; Airaksinen, J.; Kostiaainen, P.; et al. Clinical assessment of a non-invasive wearable MEMS pressure sensor array for monitoring of arterial pulse waveform, heart rate and detection of atrial fibrillation. *NPJ Digit. Med.* **2019**, *2*, 1–10. [\[CrossRef\]](#)
13. Brown, M.S.; Ashley, B.; Koh, A. Wearable technology for chronic wound monitoring: Current dressings, advancements, and future prospects. *Front. Bioeng. Biotechnol.* **2018**, *6*, 1–21. [\[CrossRef\]](#)
14. Zang, Y.; Zhang, F.; Di, C.A.; Zhu, D. Advances of flexible pressure sensors toward artificial intelligence and health care applications. *Mater. Horiz.* **2015**, *2*, 140–156. [\[CrossRef\]](#)
15. Fraulini, K.E. Interface pressure. *Decubitus* **2014**, *3*, 1–12. [\[CrossRef\]](#)
16. Barbenel, J.C.; Sockalingham, S. Device for measuring soft tissue interface pressures. *J. Biomed. Eng.* **1990**, *12*, 519–522. [\[CrossRef\]](#)
17. Van Pijkeren, T.; Naeff, M.; Kwee, H.H. A new method for the measurement of normal pressure between amputation residual limb and socket. *Bull. Prosthet. Res.* **1980**, *10*, 31–34.
18. Bu, N.; Ueno, N.; Fukuda, O. Monitoring of respiration and heartbeat during sleep using a flexible piezoelectric film sensor and empirical mode decomposition. In Proceedings of the 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Lyon, France, 22–26 August 2007; pp. 1362–1366. [\[CrossRef\]](#)
19. Han, M.; Wang, H.; Yang, Y.; Bai, W.; Yan, Z.; Li, H.; Xue, Y.; Wang, X.; Akar, B.; Zhao, H.; et al. Three-dimensional piezoelectric polymer microsystems for vibrational energy harvesting, robotic interfaces and biomedical implants. *Nat. Electron.* **2019**, *2*, 26–35. [\[CrossRef\]](#)
20. Sharma, T.; Je, S.S.; Gill, B.; Zhang, J.X.J. Patterning piezoelectric thin film PVDF-TrFE based pressure sensor for catheter application. *Sens. Actuators A Phys.* **2012**, *177*, 87–92. [\[CrossRef\]](#)
21. Gautschi, G. *Piezoelectric Sensors*; Springer: Berlin/Heidelberg, Germany, 2002; Volume 3.
22. Chen, X.; Shao, J.; An, N.; Li, X.; Tian, H.; Xua, C.; Ding, Y. Self-powered flexible pressure sensors with vertically well-aligned piezoelectric nanowire arrays for monitoring vital signs. *J. Mater. Chem. C* **2015**, *3*, 11806–11814. [\[CrossRef\]](#)
23. Gong, S.; Schwalb, W.H.; Wang, Y.; Chen, Y.; Tang, Y.; Si, K.J.; Shirinzadeh, B.; Cheng, W. A wearable and highly sensitive pressure sensor with ultrathin gold nanowires. *Nat. Commun.* **2014**, *5*, 1–8. [\[CrossRef\]](#)
24. Ng, K.C.-K.; Leung, Y.-F.; Lee, Y.-L. An ultra-sensitive resistive pressure sensor based on hollow-sphere microstructure induced elasticity in conducting polymer fil. *Hand Surg.* **2014**, *19*, 423–425. [\[CrossRef\]](#)
25. Kenry, J.; Yeo, C.; Lim, C.T. Emerging flexible and wearable physical sensing platforms for healthcare and biomedical applications. *Microsyst. Nanoeng.* **2016**, *2*. [\[CrossRef\]](#)
26. Chitnis, G.; Maleki, T.; Samuels, B.; Cantor, L.B.; Ziaie, B. A minimally invasive implantable wireless pressure sensor for continuous IOP monitoring. *IEEE Trans. Biomed. Eng.* **2013**, *60*, 250–256. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Boutry, C.M.; Beker, L.; Kaizawa, Y.; Vassos, C.; Tran, H.; Hinckley, A.C.; Pfattner, R.; Niu, S.; Li, J.; Claverie, J.; et al. Biodegradable and flexible arterial-pulse sensor for the wireless monitoring of blood flow. *Nat. Biomed. Eng.* **2019**, *3*, 47–57. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Takahata, K.; Dehennis, A.; Wise, K.D.; Gianchandani, Y.B. A Wireless Microsensor For Monitoring Flow and Pressure in a Blood Vessel Utilizing a Dual-Inductor Antenna Stent and Two Pressure Sensors. In Proceedings of the 17th IEEE International Conference on Micro Electro Mechanical Systems, Maastricht, The Netherlands, 25–29 January 2004; pp. 5–8.

29. Park, J.; Kim, J.K.; Patil, S.J.; Park, J.K.; Park, S.A.; Lee, D.W. A wireless pressure sensor integrated with a biodegradable polymer stent for biomedical applications. *Sensors* **2016**, *16*, 809. [[CrossRef](#)] [[PubMed](#)]
30. Chhetry, A.; Yoon, H.; Park, J.Y. A flexible and highly sensitive capacitive pressure sensor based on conductive fibers with a microporous dielectric for wearable electronics. *J. Mater. Chem. C* **2017**, *5*, 10068–10076. [[CrossRef](#)]
31. Al Khaburi, J.A.J. Pressure Mapping of Medical Compression Bandages Used for Venous Leg Ulcer Treatment. 2010. Available online: http://theses.whiterose.ac.uk/3386/1/Thesis_Corrected.pdf (accessed on 10 September 2020).
32. Partsch, H. The Static Stiffness Index: A Simple Method to Assess the Elastic Property of Compression Material In Vivo. *Dermatol. Surg.* **2006**, *31*, 625–630. [[CrossRef](#)]
33. Van den Kerckhove, E.; Fieuws, S.; Massagé, P.; Hierner, R.; Boeckx, W.; Deleuze, J.-P.; Laperre, J.; Anthonissen, M. Reproducibility of repeated measurements with the Kikuhime pressure sensor under pressure garments in burn scar treatment. *Burns* **2007**, *33*, 572–578. [[CrossRef](#)]
34. Finnie, A. Interface pressure measurements in leg ulcer management. *Br. J. Nurs.* **2000**, *9* (Suppl. 1), S8–S18. [[CrossRef](#)]
35. Parmar, S.; Khodasevych, I.; Troynikov, O. Evaluation of flexible force sensors for pressure monitoring in treatment of chronic venous disorders. *Sensors* **2017**, *17*, 1923. [[CrossRef](#)]
36. Khodasevych, I.; Parmar, S.; Troynikov, O. Flexible sensors for pressure therapy: Effect of substrate curvature and stiffness on sensor performance. *Sensors* **2017**, *17*, 2399. [[CrossRef](#)]
37. Rai, T.B.; Goddard, M.; Makin, G.S. How long do compression bandages maintain their pressure during ambulatory treatment of varicose veins. *Br. J. Surg.* **1980**, *67*, 122–124.
38. Hafner, J.; Lüthi, W.; Hänssle, H.; Kammerlander, G.; Burg, G. Instruction of compression therapy by means of interface pressure measurement. *Dermatol. Surg.* **2000**, *26*, 481–488. [[CrossRef](#)] [[PubMed](#)]
39. Mehmood, N.; Hariz, A.; Templeton, S.; Voelcker, N.H. An improved flexible telemetry system to autonomously monitor sub-bandage pressure and wound moisture. *Sensors* **2014**, *14*, 21770–21790. [[CrossRef](#)] [[PubMed](#)]
40. Farooqui, M.F.; Shamim, A. Low cost inkjet printed smart bandage for wireless monitoring of chronic wounds. *Sci. Rep.* **2016**, *6*, 28949. [[CrossRef](#)] [[PubMed](#)]
41. Rahimi, R.; Ochoa, M.; Zieger, M.; Sood, R.; Ziaie, B. A wireless strain sensor for wound monitoring with direct laser-defined patterning on a commercial dressing. In Proceedings of the 2016 IEEE 29th International Conference on Micro Electro Mechanical Systems (MEMS), Shanghai, China, 24–28 January 2016; pp. 481–484. [[CrossRef](#)]
42. Deng, W.J.; Wang, L.F.; Dong, L.; Huang, Q.A. LC Wireless Sensitive Pressure Sensors with Microstructured PDMS Dielectric Layers for Wound Monitoring. *IEEE Sens. J.* **2018**, *18*, 4886–4892. [[CrossRef](#)]
43. Fonseca, M.A.; Allen, M.G.; Kroh, J.; White, J. Flexible Wireless Passive Pressure Sensors for Biomedical Applicators. In Proceedings of the Solid-State Sensors, Actuators, and Microsystems Workshop, Hilton Head Island, SC, USA, 4–8 June 2006; pp. 37–42.
44. Li, C.; Wu, P.M.; Shutter, L.A.; Narayan, R.K. Dual-mode operation of flexible piezoelectric polymer diaphragm for intracranial pressure measurement. *Appl. Phys. Lett.* **2010**, *96*, 2008–2011. [[CrossRef](#)]
45. Chen, P.-J.; Saati, S.; Varma, R.; Humayun, M.S.; Tai, Y.-C. Wireless Intraocular Pressure Sensing Using Microfabricated Minimally Invasive Flexible-Coiled LC Sensor Implant. *J. Microelectromech. Syst.* **2010**, *19*, 721–734. [[CrossRef](#)]
46. Mohan, S.S.; Hershenson, M.D.M.; Boyd, S.P.; Lee, T.H. Simple accurate expressions for planar spiral inductances. *IEEE J. Solid-State Circuits* **1999**, *34*, 1419–1420. [[CrossRef](#)]
47. Schormans, M.; Valente, V.; Demosthenous, A. Practical Inductive Link Design for Biomedical Wireless Power Transfer: A Tutorial. *IEEE Trans. Biomed. Circuits Syst.* **2018**, 1–19. [[CrossRef](#)]
48. Kuhn, W.B.; Ibrahim, N.M. Analysis of current crowding effects in multiturn spiral inductors. *IEEE Trans. Microw. Theory Tech.* **2001**, *49*, 31–38. [[CrossRef](#)]
49. Ammouri, A.; Salah, T.B.; Morel, H. A spiral planar inductor: An experimentally verified physically based model for frequency and time domains. *Int. J. Numer. Model. Electron. Netw. Devices Fields* **2018**, *31*, 1–13. [[CrossRef](#)]
50. Islam, A.B.; Islam, S.K.; Tulip, F.S. Design and Optimization of Printed Circuit Board Inductors for Wireless Power Transfer System. *Circuits Syst.* **2013**, *4*, 237–244. [[CrossRef](#)]

51. Huang, Q.A.; Dong, L.; Wang, L.F. LC Passive Wireless Sensors Toward a Wireless Sensing Platform: Status, Prospects, and Challenges. *J. Microelectromech. Syst.* **2016**, *25*, 822–841. [[CrossRef](#)]
52. MIT.edu. Material: PDMS (Polydimethylsiloxane). Available online: <http://www.mit.edu/~6.777/matprops/pdms.htm> (accessed on 24 April 2019).

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).

MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland
Tel. +41 61 683 77 34
Fax +41 61 302 89 18
www.mdpi.com

Sensors Editorial Office
E-mail: sensors@mdpi.com
www.mdpi.com/journal/sensors



MDPI
St. Alban-Anlage 66
4052 Basel
Switzerland

Tel: +41 61 683 77 34
Fax: +41 61 302 89 18

www.mdpi.com



ISBN 978-3-0365-0975-4