

MACHINE LEARNING-EMPOWERED SMART HEALTH WEARABLES FOR  
LONG-TERM HIGH-RESOLUTION HEART RATE AND BLOOD PRESSURE TRACKING

by

Qingxue Zhang



APPROVED BY SUPERVISORY COMMITTEE:

---

Dr. Dian Zhou, Chair

---

Dr. Jin Liu

---

Dr. Lakshman S. Tamil

---

Dr. Mehrdad Nourani

Copyright 2017

Qingxue Zhang

All Rights Reserved

*To my beloved family*

MACHINE LEARNING-EMPOWERED SMART HEALTH WEARABLES FOR  
LONG-TERM HIGH-RESOLUTION HEART RATE AND BLOOD PRESSURE TRACKING

by

QINGXUE ZHANG, BS, MS

DISSERTATION

Presented to the Faculty of  
The University of Texas at Dallas  
in Partial Fulfillment  
of the Requirements  
for the Degree of

DOCTOR OF PHILOSOPHY IN  
ELECTRICAL ENGINEERING

THE UNIVERSITY OF TEXAS AT DALLAS

May 2017

## ACKNOWLEDGMENTS

I would like to sincerely thank my advisor, Dr. Dian Zhou, for his encouragement and guidance during my research. I learned a lot from many discussions with him, and benefited a lot from his broad knowledge and deep insights in various fields. I would also like to extend my gratitude towards my other committee members, Dr. Jin Liu, Dr. Lakshman Tamil and Dr. Mehrdad Nourani. They gave me lots of suggestions and encouragement during my dissertation proposal. Special thanks to my friends and colleagues both in academia and industry for giving me valuable help during this journey. Also many thanks to the Electrical Engineering Department at The University of Texas at Dallas. Finally, my sincere thanks to my family for being always supportive and giving me immeasurable love.

January 2017

MACHINE LEARNING-EMPOWERED SMART HEALTH WEARABLES FOR  
LONG-TERM HIGH-RESOLUTION HEART RATE AND BLOOD PRESSURE TRACKING

Qingxue Zhang, PhD  
The University of Texas at Dallas, 2017

Supervising Professor: Dian Zhou

Wearable computers are paving a promising way for ubiquitous smart health applications. Long-term high-resolution heart rate and blood pressure tracking is highly significant not only for medical purposes such as heart diseases and hypertension management, but also for wellness and fitness applications such as emotion, stress and sport performance monitoring.

It is highly challenging to robustly estimate the high-resolution (second-level) heart rate in long-term application scenarios, due to motion artifacts caused by diverse body movements. We propose two novel electrocardiogram (ECG)-based instantaneous heart rate monitoring frameworks tolerant to severe and continuous motion artifacts. The first framework is phase-domain multi-view dynamic time warping, which transforms the raw ECG signal to the multi-dimensional phase space to reveal more geometrical characteristics of heartbeats, and then performs multi-view dynamic time warping to identify heartbeats similar to a predefined heartbeat template. The heart rate estimation performance greatly outperforms previously reported approaches. The second framework is a two-stage classification approach, which does not sweep all of the signal stream to

find out heartbeats, but instead, just focuses on the signal spikes of interest to classify them as heartbeats or interferential spikes. Since it does not care about the non-spike signal segments, it consumes less computation load. We extract twenty-six features and select out ten features critical to motion artifacts, by a sparse support vector machine (SVM). Then they are used to train a SVM model to perform the heartbeat identification. Afterwards, a refinement engine is introduced to purify the heartbeats. The performance is superior to many well-known approaches.

Moreover, we propose two wearable cuff-less blood pressure monitoring systems, one with single-arm-ECG and photoplethysmogram (PPG) signals and another with ear-ECG and PPG signals. In the first system, the weak single-arm signals are successfully acquired by our bio-potential acquisition platform, and the heartbeats are then identified from these weak signals. A thorough comparative analysis on diverse blood pressure models is also performed to determine an appropriate one for single-arm applications. Experimental results show that the system can robustly estimate the minute-level blood pressure only based on single-arm signals. In the second system, both weak ECG and PPG signals are also successfully acquired by placing all the sensors behind the ear for a super wearability. Moreover, we introduce large amounts of motion artifacts by performing head movements towards practical application scenarios, and then propose an unsupervised learning strategy to automatically evaluate the signal distortion level and perform signal purification. Experimental results also show the effectiveness of the proposed ear-worn blood pressure and heart rate monitoring system. These smart health wearables proposed are expected to contribute to pervasive health, wellness and fitness management.

## TABLE OF CONTENTS

ACKNOWLEDGMENTS .....	v
ABSTRACT.....	vi
LIST OF FIGURES .....	xii
LIST OF TABLES .....	xiv
CHAPTER 1 INTRODUCTION .....	1
CHAPTER 2 HEART RATE MONITORING USING A PHASE-DOMAIN PATTERN RECOGNITION APPROACH.....	6
2.1 INTRODUCTION .....	8
2.2 METHODS .....	10
2.2.1 Platform.....	10
2.2.2 Datasets .....	11
2.2.3 Signal Pre-processing.....	12
2.2.4 Time to Phase Space Transformation.....	13
2.2.5 Time-domain Dynamic Time Warping (TD-DTW).....	16
2.2.6 Phase-domain Multi-view DTW (PD-MDTW) .....	19
2.2.7 Parameter Tuning Phase.....	20
2.2.8 Testing Phase and Performance Evaluation .....	22
2.2.9 Performance Comparison.....	23
2.3 RESULTS .....	24
2.3.1 Domain Transformation .....	24
2.3.2 TD-DTW and PD-MDTW .....	26
2.3.3 Heartbeat Recognition and Heart Rate Estimation .....	30
2.3.4 Performance Comparison.....	35
2.3.5 Limitation Exploration .....	37
2.4 DISCUSSION.....	38
2.5 CONCLUSION.....	41
CHAPTER 3 HEART RATE MONITORING USING A TWO-STAGE CLASSIFICATION FRAMEWORK .....	42

3.1	INTRODUCTION .....	42
3.2	SYSTEM OVERVIEW .....	43
3.3	METHOD .....	44
3.3.1	Platform and Dataset .....	44
3.3.2	Preprocessing .....	44
3.3.3	Heartbeat Candidates Generation.....	45
3.3.4	Feature Extraction .....	46
3.3.5	Feature Selection .....	51
3.3.6	Classification.....	52
3.3.7	Evaluation of Classification performance .....	52
3.3.8	Outlier Removal .....	53
3.3.9	Heartbeat Interpolation and Heart Rate Estimation .....	55
3.4	RESULTS .....	56
3.4.1	Auto-segmentation .....	56
3.4.2	Feature Engineering .....	58
3.4.3	Heartbeat Identification.....	58
3.4.4	Heart Rate Estimation .....	63
3.4.5	Performance Comparison.....	63
3.5	CONCLUSION.....	67
CHAPTER 4 BLOOD PRESSURE MONITORING WITH SINGLE-ARM SIGNALS .....		68
4.1	INTRODUCTION .....	70
4.2	METHODS .....	73
4.2.1	System Overview .....	73
4.2.2	Customized Hardware Platform .....	74
4.2.3	Data Acquisition Protocol.....	75
4.2.4	Signal Pre-processing.....	76
4.2.5	ECG-based Heartbeat Identification .....	77
4.2.6	PPG-based Heartbeat Identification .....	80
4.2.7	Heart Rate Estimation .....	81
4.2.8	Pulse Transit Time .....	81

4.2.9	Blood Pressure Estimation .....	83
4.3	RESULTS .....	85
4.3.1	Signals Acquired .....	85
4.3.2	Heartbeat Identification and Heart Rate Estimation .....	87
4.3.3	Correlation Among PTT, HR and SBP .....	88
4.3.4	Blood Pressure Model Training and Testing.....	90
4.4	DISCUSSION .....	93
4.5	CONCLUSION.....	98
CHAPTER 5 BLOOD PRESSURE MONITORING WITH EAR SIGNALS.....		100
5.1	INTRODUCTION .....	100
5.2	MATERIAL AND METHOD .....	104
5.2.1	System Overview .....	104
5.2.2	Customized Hardware Platform and Sensors Placement .....	104
5.2.3	Dataset Recording .....	106
5.2.4	Supervised Learning of HB Identification .....	107
5.2.5	Unsupervised Learning of Signal Quality Labelling and Purification.....	109
5.2.6	HR Estimation and Supervised Learning of SBP Estimation .....	120
5.3	RESULTS AND DISCUSSION .....	123
5.3.1	Signals Acquired .....	123
5.3.2	Heartbeat Identification.....	125
5.3.3	PPG Segmentation and Template Learning .....	127
5.3.4	PPG Distortion Evaluation and Threshold Learning.....	129
5.3.5	SQI Generation and Heartbeats Purification.....	130
5.3.6	Heart Rate Estimation .....	132
5.3.7	Blood Pressure Estimation .....	133
5.4	CONCLUSION.....	136
CHAPTER 6 CONCLUSION AND FUTURE WORK.....		138
REFERENCES .....		140

BIOGRAPHICAL SKETCH .....147

CURRICULUM VITAE

## LIST OF FIGURES

Figure 2.1. A semi-customized bio-potential acquisition platform. ....	11
Figure 2.2. An example of the TD-DTW algorithm to recognize from the sequence X two segments (x1 to x5, and x7 to x12), with a template Y (y0 to y4).....	17
Figure 2.3. Space transformation for the clean and continuously/heavily corrupted ECG signals. ....	25
Figure 2.4. Similarity sequences measured by the TD-DTW or the PD-MDTW on the ECG signals with different quality levels. ....	27
Figure 2.5. Pattern recognition ability of TD-DTW (a-c) versus PD-MDTW (d-f). ....	29
Figure 2.6. Performance comparison of twelve spaces over different SNRs.....	32
Figure 2.7. Estimated instantaneous heart rate (IHR) @ different SNRs. ....	37
Figure 3.1. The proposed two-stage classification framework for robust instantaneous heart rate monitoring.....	43
Figure 3.2. The ECG stream without motion artifacts to illustrate the characteristic points of the heartbeats. ....	46
Figure 3.3. An example of the auto-segmentation result. ....	57
Figure 3.4. An example of the identified heartbeats. ....	61
Figure 3.5. Heartbeat recognition performance in terms of sensitivity on WECG-based datasets.....	62
Figure 3.6. Heartbeat recognition performance in terms of precision on WECG-based datasets. ....	62
Figure 3.7. Heart rate estimation performance in terms of MAE and RMSE on WECG-based datasets.....	63
Figure 3.8. Heart rate estimation performance comparison in terms of MAE on WECG-based datasets.....	65
Figure 3.9. Heart rate estimation performance comparison in terms of RMSE on WECG-based datasets.....	65
Figure 3.10. Heart rate estimation performance comparison in terms of MAE on FECG-based datasets.....	66

Figure 3.11. Heart rate estimation performance comparison in terms of RMSE on FECG-based datasets.....	66
Figure 4.1. The proposed system for wearable cuff-less SBP and HR monitoring with single-arm-ECG and PPG signals. ....	73
Figure 4.2. Arm-ECG/PPG signals, with the amplitude both scaled to be between 0 and 1 for good readability. ....	77
Figure 4.3. Pulse transit time measured with ECG and PPG signals (This illustration of PTT is based on arm-ECG/PPG signals). ....	82
Figure 4.4. An example of the signal segments acquired (chest-ECG, arm-PPG and arm-ECG). ....	86
Figure 4.5. Heart rate estimation performance.....	87
Figure 4.6. Trend of SBP <sub>cuff</sub> , PTT and HR@arm-ECG, corresponding to the training session of subject 1. ....	89
Figure 4.7. Performance of model 1 and 10 in terms of Bland-Altman plot. ....	91
Figure 5.1. The proposed machine learning-empowered system for wearable cuff-less SBP and HR monitoring from motion artifacts-sensitive ear-ECG/PPG signals.....	105
Figure 5.2. Ear-ECG/PPG Signals. ....	108
Figure 5.3. Pulse transit time (PTT).....	121
Figure 5.4. An example of the ear signal segments acquired.....	124
Figure 5.5. Two examples of heartbeat identification results. ....	126
Figure 5.6. K-medoids clustering-based PPG template learning. ....	128
Figure 5.7. PPG distortion threshold learning.....	129
Figure 5.8. The whole SQI generation process. ....	131
Figure 5.9. Bland-Altman plot for estimated HR (HR <sub>est</sub> ) and reference HR (HR <sub>ref</sub> ). ....	133
Figure 5.10. Bland-Altman plots for SBP model 1 and 10. ....	134

## LIST OF TABLES

Table 2.1. Performance summary of four typical spaces (A0/1/2-40%/3-40).....	33
Table 2.2. Performance comparison of five different approaches.....	35
Table 3.1. Summary of extracted features.....	47
Table 3.2. Learned feature weights on the training data.....	59
Table 3.3. Model comparison in terms of accuracy (%) on the testing data.....	60
Table 4.1. Ten blood pressure models for comparative analysis.....	84
Table 4.2. A summary of the testing performance of ten SBP models.....	93
Table 5.1. Blood pressure estimation methods.....	122
Table 5.2. Blood pressure estimation performance summary.....	135

# **CHAPTER 1**

## **INTRODUCTION**

Traditional health services in hospitals and clinics are usually expensive, time-consuming and inconvenient. Especially for long-term health monitoring, the patients need to go to the facilities frequently, which inevitably introduce lots of inconvenience to their daily activities. There are also very limited health monitoring solutions in the gym or at home, such as the heart rate monitor on the treadmill, the cuff-based blood pressure monitor and the thermometer. However, these monitors are also inconvenient for long-term daily applications.

Nowadays, smart health is an emerging topic attracting more and more attention in different areas [1-7]. Based on the concept of smart health, the health data is first acquired by individual-centric monitors such as wearable computers (watches, wrist bands, etc.), and then transmitted to the body area network to some relays such as cellphones. Afterwards, the data can also be sent to the service providers and medical enterprise for further data mining and statistical analysis purposes. Smart health solutions are expected to provide pervasive monitoring of human health, wellness and fitness in a low cost, time-saving and convenient way.

We take special interest in wearable computers for smart health applications [8-16]. Benefiting from fast development of semiconductor, information and computer technologies, wearable computers are playing more and more roles in medical, wellness and fitness monitoring. For example, in medical monitoring, wearable computers are usually used to provide unobtrusive patient monitoring, diabetes care and vital signals monitoring. In wellness monitoring, wearable computers can be applied to monitor the sleep and emotion status. In fitness monitoring, they can be applied to track the sport performance and monitor daily activities. We take special interest in

long-term high-resolution heart rate and blood pressure monitoring, which are both significant measurements for medical, wellness and fitness monitoring purposes.

Long-term high-resolution heart rate monitoring is highly important. According to WHO, heart disease is the leading cause of death worldwide [17]. The heart rate information tracked in a long-term way can provide valuable insights in heart disease prediction, diagnosis and treatment. Moreover, long-term heart rate monitoring is also significant for emotion, stress and fitness tracking and analysis. The resolution of the heart rate estimates are also important. There are many heart rate variability-related measures need to use second-level heart rate estimates, including both statistical analysis and frequency domain analysis. Although occasionally measured heart rate information can also provide useful information, continuous second-level heart rate tracking is much more valuable to enable heart health, wellness and fitness management.

Long-term high-resolution blood pressure monitoring is necessary to manage the hypertension, which causes 12.8% of total deaths worldwide. Also according to WHO, over 35 % of people have raised blood pressure, which may relates to or result in many cardiovascular, circulatory and cerebrovascular diseases [18].

To provide long-term high-resolution heart rate monitoring and blood pressure monitoring is a challenging goal. When it comes to heart rate monitoring, we can find commercial products such as IWATCH and FITBIT, but they usually ask the user to stay still to get robust heart rate estimates. During body movements, there are lots of motion artifacts, which may make the estimates inaccurate. So these products need to get the estimates based on averaging operations. It is hard to guarantee the second-level temporal-resolution of the heart rate estimates. One of the research topics here is how to provide a high usability and temporal-resolution in heart rate monitoring,

with a motion artifacts-tolerant ability under body movements, which is highly important to enable the long-term usage of these monitors.

Moreover, the cuff-based blood pressure monitor is inconvenient for long-term high-resolution blood pressure monitoring. The commercial products mentioned above do not support blood pressure monitoring because of the difficulty in acquiring enough physiological signals to establish the blood pressure model. They cannot acquire the electrocardiogram (ECG) signal from one single wrist, however, ECG signal is the mostly frequently used signal to build the blood pressure model together with the photoplethysmogram (PPG) signal. Another research topic here is how to provide convenient long-term high-resolution wearable blood pressure monitoring.

To provide long-term high-resolution heart rate monitoring, we propose two frameworks, one of which is names as phase-domain dynamic time warping and another of which is called two-stage classification approach. Besides, we also propose two wearable cuff-less blood pressure monitoring system with single-arm or ear ECG and PPG signals.

Chapter 2 introduces the phase-domain multi-view dynamic time warping for long-term high-resolution heart rate monitoring, based on the ECG signal. The raw ECG signal with large amounts of motion artifacts are firstly transformed to the phase-domain, where more geometrical characteristics of the heartbeats can be revealed and are expected to be more tolerant to motion artifacts. Then the multi-view dynamic time warping approach is used to identify the heartbeats from the raw ECG signal, by sweeping the raw ECG signal in the phase-domain to find out signal segments similar to a predefined ECG heartbeat template. The proposed framework greatly outperforms many previously reported approaches in terms of long-term instantaneous heart rate monitoring.

In chapter 3, a two-stage classification approach is proposed for long-term high-resolution heart rate monitoring. Different from the phase-domain multi-view dynamic time warping which sweeps all the signal stream to identify heartbeats, the two-stage classification approach directly focuses on the signal spikes of interest and classifies them as real heartbeats or interferential spikes due to motion artifacts. Since this method does not care about non-spike signal segments, it consumes less computation load. Specifically, twenty-six features are extracted from multiple domains, and then critical features tolerant to motion artifacts are determined. The selected features are used to train a support vector machine (SVM) classifier for high confident heartbeats identification. Afterwards, a refinement engine further suppresses residual interferential spikes by a rule-based classifier and interpolates missing heartbeats using interpolation techniques. This approach also provides superior performance compared with many other well-known approaches.

Chapter 4 targets cuff-less blood pressure monitoring with a high wearability. The weak single-arm ECG signal and PPG signal are acquired by a bio-potential acquisition platform. Then advanced signal processing and heartbeat identification methods are used to identify the ECG and PPG heartbeats. Diverse blood pressure models are thoroughly compared to distinguish their robustness in wearable applications. Finally, the results show that this single-arm blood pressure monitoring system can effectively estimate the minute-level blood pressure with a high wearability, and may significantly contribute to continuous blood pressure tracking applications.

In chapter 5, we propose another highly wearable blood pressure monitoring system by placing all ECG and PPG sensors behind the ear. We also introduce large amounts of motion artifacts by performing head movements, to consider more practical application scenarios. To deal with the motion artifacts, we present an unsupervised learning strategy to automatically evaluate the

distortion condition of raw heartbeats identified, generate signal quality indices, and then perform signal purification. The proposed system can robustly estimate the blood pressure and heart rate even under head movements

**CHAPTER 2**  
**HEART RATE MONITORING USING A PHASE-DOMAIN**  
**PATTERN RECOGNITION APPROACH**

Authors: Qingxue Zhang, Dian Zhou, Xuan Zeng

The Department of Electrical Engineering, EC33

The University of Texas at Dallas

800 West Campbell Road

Richardson, Texas 75080-3021

© 2016 IEEE. Reprinted, with permission, from Q. Zhang, D. Zhou and X. Zeng, “A Novel Framework for Motion-tolerant Instantaneous Heart Rate Estimation By Phase-domain Multi-view Dynamic Time Warping,” IEEE Transactions on Biomedical Engineering, 2016. DOI: 10.1109/TBME.2016.2640309.

## ABSTRACT

Objective: Long-term wearable instantaneous heart rate (IHR) monitoring is essential to enable pervasive heart health and fitness management. In this paper, a novel framework is proposed to robustly estimate the IHR from electrocardiogram (ECG) signals corrupted by large amounts of daily motion artifacts, which are one of the major impediments against the long-term IHR monitoring. Methods: The corrupted ECG signals are firstly projected to a high-dimensional phase space, where the constructed phase portraits of heartbeats are of many new geometrical properties and are expected to be powerful patterns more immune to the motion artifacts. Afterwards, a multi-view dynamic time warping (MDTW) approach is applied on the constructed phase portraits, to effectively capture motion artifacts-induced inconsistencies and reveal heartbeats-related consistencies from corrupted signals. Finally, the phase portraits of heartbeats in the multi-dimensional phase space can be identified, and then the IHR estimates are achieved. Results: The proposed framework is evaluated on a wrist-ECG dataset acquired by a semi-customized platform and also a public ECG dataset. With a signal-to-noise ratio (SNR) as low as -7 dB, the mean absolute error (MAE) and root mean square error (RMSE) of the estimated IHR are 1.6 beats per minute (BPM) and 5.3 BPM, respectively. Conclusion: These results demonstrate that our framework can effectively identify the heartbeats from ECG signals continuously corrupted by intense and random motion artifacts and estimate the IHR. Significance: The proposed framework greatly outperforms previously reported approaches and is expected to contribute to long-term IHR monitoring.

## 2.1 INTRODUCTION

Heart rate monitoring has been attracting tremendous attentions, not only because of its effectiveness in heart disease prediction and diagnosis, but also due to the boosting requirements on fitness management [19]. The electrocardiogram (ECG) signal is regarded as a gold standard vital sign of the heart and has been widely used to monitor the heart rate. However, to enable robust ECG-based heart rate monitoring in long-term applications, one of the major challenges is how to deal with large amounts of motion artifacts while people perform various body movements. These diverse daily activities usually cause dramatic variations in the electrode-skin contact impedance, and thus induce large amounts of random and heavy motion artifacts and corrupt the ECG signals [20]. Another concern is how to guarantee the instantaneity of the heart rate estimates to reflect the status of the heart in time. As the motion artifacts may continuously corrupt the ECG signals for a long duration, the instantaneous heartbeat identification during this time becomes much more challenging [21]. Therefore, an effective motion artifact-tolerant instantaneous heart rate (IHR) estimation approach is essential for pervasive heart health and fitness management.

Many investigations have been reported to process the low quality ECG signals contaminated by motion artifacts. One simple method is just discarding low quality signal segments but it may cause information loss during continuous motion artifacts [21]. Other advanced techniques like adaptive filtering and independent component analysis, introduce additional reference or redundant channels but they inevitably increase the hardware cost [20, 22]. There are also some model-based methods, like Kalman filtering [23, 24], which continuously predicts and refines the heart rate, through modeling both the process noise and measurement noise with a Gaussian distribution, however, the measurement noise maybe non-Gaussian since motion artifacts are often induced by

intense body movements. The heart rate estimation task can also be formulated as a weighted regularized least square problem [23], nevertheless, when continuous motion artifacts make the raw measurements highly noisy, heavy regularization may cause a non-negligible drift. Besides, the discrete wavelet transformation approach has also been used to de-noise the ECG signals [25]. New signals are reconstructed using specific signal components where most energy of the heartbeat concentrates in, but motion artifacts usually have spectrum completely overlapping that of the ECG so the reconstructed signals are still contaminated with many motion artifacts [20].

A widely applied pattern recognition approach, dynamic time warping (DTW), is also introduced for noisy ECG processing by Kotas [26]. In time series analysis, it usually measures the similarity between a noisy testing sequence and a pre-defined template, aiming to find from the testing sequence one or more segments similar to the template. This time-domain DTW (TD-DTW) can also warp the test sequence in a non-linear approach for template matching, independent of the non-linear variations in the time dimension, which is always the case in ECG signal processing since the beat-to-beat interval is time varying [27]. However, in Kotas's work, ECG signals were not corrupted heavily and continuously [26], which may not be sufficient in long-term wearable IHR monitoring, where severe or extreme motion artifacts usually continuously and heavily corrupt the ECG signals.

In this study, we propose a novel phase-domain multi-view DTW framework (PD-MDTW) for long-term IHR monitoring from continuously and heavily corrupted ECG signals, without discarding signal segments, adding extra hardware, applying an inappropriate model for motion artifacts which are usually hard to be modeled, or extracting signal components of interest where many motion artifacts usually coexist. We firstly transform the low quality ECG stream from the

time domain to the phase domain, where the constructed phase portraits of the quasi-periodic heartbeats now reveal many new geometrical properties (point-wisely, pair-wisely and piecewisely) more immune to motion artifacts. Afterwards, the multi-view DTW (MDTW) is introduced [28], to recognize the multi-dimensional phase portraits of heartbeats in the new domain, and then robust IHR estimates are achieved. The proposed framework are evaluated on an acquired wrist-ECG dataset and also a public dataset, corrupted by heavy and continuous motion artifacts. Furthermore, the comparison with several well-established approaches [23-26], including Kalman filtering, Weighted regularized least square, Discrete wavelet transformation and Time-domain DTW, is also performed. The proposed approach shows a superior ability in IHR estimation.

## 2.2 METHODS

In this section, we provide detailed description of the framework and performance evaluation methods.

### 2.2.1 Platform

To collect a wearable wrist-ECG dataset, we built a semi-customized bio-potential acquisition platform which comprised a *ADS1299EEG-FE evaluation board* and a *Tiva™ C Series LaunchPad* [29-31], as shown in Figure 2.1a. The ADS1299 board (green) is equipped with a low noise and high resolution (24-bit) *analog-to-digital converter* (ADC) for bio-potential measurements, and a *Serial Peripheral Interface* (SPI) for communication. This board is designed by Texas Instruments and particularly useful for acquiring low-level bio-potential signals. The LaunchPad (red) includes an *ARM Cortex M4* microcontroller, as well as a SPI and a *Universal*

*Serial Bus* (USB). We also designed a graphical user interface (GUI) to manage the platform, receive and visualize the data. Firstly, the GUI sends a ‘start’ command via USB to the LaunchPad, which then triggers the ADS1299 board via SPI to acquire the single-channel ECG signal. Afterwards, the ADS1299 board collects the ECG signals (488 Hz, 24-bit) and sends data via SPI to the Launchpad, which then parses the received SPI packets and generates USB packets that are finally sent to and visualized on the PC. In future, this platform can be easily equipped with a wireless module, and also miniaturized by removing unnecessary components which are only for evaluation purpose, to enhance the wearability.

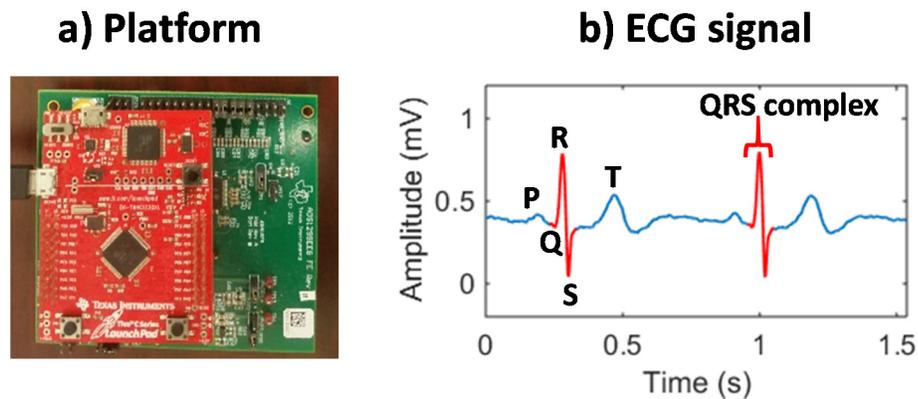


Figure 2.1. A semi-customized bio-potential acquisition platform.

### 2.2.2 Datasets

A dataset including twenty-two 30-minute wrist ECG recordings from healthy people, named as *WECCG*, was acquired using the Lead I configuration [31, 32], with the signal electrode placed on the left wrist, and the reference/bias electrodes on the right wrist. All subjects gave their consent to participate in the data acquisition. To corrupt the ECG signals as shown in Figure 2.1b, the real motion artifacts from *MIT-BIH Noise Stress Test Database* were introduced, considering they have

been widely applied as a test bench for fair algorithm comparison in large amounts of works [33, 34], and can also be used to corrupt other public ECG datasets to further evaluate the generalization ability of our framework. Various signal-to-noise ratios (SNR) (-1, -3, -5, -7 and -9 dB) and phase shifts (0, 0.2, 0.4, 0.6 and 0.8 second) were used to introduce more signal qualities and randomness, generating twenty-five noisy/corrupted datasets (*WECG\_Sm1/3/5/7/9\_SH0/1/2/3/4*). The phase shift is defined as the delay of the motion artifacts when added to the ECG, to further induce the randomness.

To further validate the generalization ability of the proposed framework, a public ECG dataset, *Fantasia*, was also employed, which includes forty recordings from forty people. This dataset was initially collected to study the aging trend of people, and then used in many works [34, 35], since it is a large dataset with diverse age-related changes in both ECG physical morphologies and heart rate variability. Using similar corruption approaches, another twenty-five noisy or corrupted datasets are achieved (each includes forty 30-minute corrupted ECG recordings), named as *FECG\_Sm1/3/5/7/9\_SH0/1/2/3/4*.

To further explore the limitation of the proposed framework, datasets with lower signal conditions are also taken into consideration, which will be introduced in the results section.

### **2.2.3 Signal Pre-processing**

A 6-order Butterworth bandpass filter (2 to 50 Hz) is applied to the corrupted ECG datasets for baseline wander and power line interference removal. The filtered data includes major ECG components and large amounts of motion artifacts since they usually overlap the ECG in frequency domain [20].

## 2.2.4 Time to Phase Space Transformation

We transform the corrupted ECG signals from the time space to the phase space, which not only presents many more meaningful geometrical characteristics of heartbeats in higher dimensions, but also effectively reveals the quasi-periodic nature of heartbeats. This quasi-periodic nature and new geometrical properties make consecutive cardiac cycles overlap and form regular phase portraits, which are much more powerful heartbeat patterns with motion artifact-tolerant ability. Graphical illustrations are given in the next section.

Mathematically, the space transformation  $\phi$  can be expressed as (1), where  $\mathcal{R}$  and  $\mathcal{P}$  denote the time space and the phase space, respectively. If denote an ECG sequence in space  $\mathcal{R}$  as  $\{R(t)|0 \leq t < N - 1\}$  where  $t$  is the time index and  $N$  is the number of samples, the transformation can be expressed as (2), where  $P^0(t)$  to  $P^{L-1}(t)$  correspond to transformed sequences and  $L$  is the number of dimensions. The  $l - th$  sequence  $P^l(t)$  in this space is defined as (3), where  $T^l(\cdot)$  is a basis function and  $S(\cdot)$  is a scaling function.

Therefore, (2) can be expressed as (4) and further as (5), where  $\circ$  is an operator for function combination. If define a transformation matrix  $A$  as (6), then the transformation process can finally be expressed as (7).

$$\phi: \mathcal{R} \rightarrow \mathcal{P} \quad (1)$$

$$\phi[R(t)] = \begin{bmatrix} P^0(t) \\ \vdots \\ P^{L-1}(t) \end{bmatrix}, \quad \forall t \in \{0, 1, \dots, N - 1\} \quad (2)$$

$$P^l(t) = S\{T^l[R(t)]\}, \quad \forall l \in \{0, 1, \dots, L - 1\}, \forall t \quad (3)$$

$$\phi[R(t)] = \begin{bmatrix} S\{T^0[R(t)]\} \\ \vdots \\ S\{T^{L-1}[R(t)]\} \end{bmatrix}, \quad \forall t \quad (4)$$

$$\phi[R(t)] = \begin{bmatrix} S \circ T^0 \\ \vdots \\ S \circ T^{L-1} \end{bmatrix} \circ R(t), \quad \forall t \quad (5)$$

$$A = \begin{bmatrix} S \circ T^0 \\ \vdots \\ S \circ T^{L-1} \end{bmatrix} \quad (6)$$

$$\phi[R(t)] = A \circ R(t), \quad \forall t \quad (7)$$

Now, the space transformation problem has been converted to how to design the transformation matrix  $A$ . The scaling function defined as (8) bounds each transformed sequence between 0 and 1 to balance their contribution in the new space, where  $\alpha^l$  and  $\beta^l$  are the maxima and minima among the sequence  $\{T^l[R(t)] | 0 \leq t \leq N - 1\}$ , respectively. The basis functions need to be designed sophisticatedly to be able to reveal meaningful characteristics of heartbeats and construct phase portraits with a motion artifact-tolerant morphology for pattern recognition. A basis function library  $T^\Theta$  is created as (9) which currently includes three basis functions, i.e., the (point-wise) self-copy basis function as (10), the (pair-wise) first derivative basis function as (11) [36], and the (piece-wise) angle basis function as (12). In (12), the  $\sphericalangle$  operator calculates the angle defined by three points  $R(t - \tau/2)$ ,  $R(t)$  and  $R(t + \tau/2)$ , where the vertex is the point  $R(t)$ , one side is the straight line from  $R(t)$  to  $R(t - \tau/2)$ , and another side is the straight line from  $R(t)$  to  $R(t + \tau/2)$ . Therefore, the process of generating the angle sequence seems like sliding a window of a length of  $\tau$  over the amplitude sequence to extract the piecewise geometrical fluctuation characteristics.

$$S\{T^l[R(t)]\} = \frac{T^l[R(t)] - \beta^l}{\alpha^l - \beta^l}, \quad \forall l, \forall t \quad (8)$$

$$T^\Theta[R(t)] = \begin{cases} T_0^\Theta[R(t)] \\ T_1^\Theta[R(t)] \\ T_2^\Theta[R(t)] \end{cases}, \quad \forall t \quad (9)$$

$$T_0^\Theta[R(t)] = R(t) \quad (10)$$

$$T_1^\Theta[R(t)] = R(t) - R(t - 1) \quad (11)$$

$$T_2^\Theta[R(t)] = \sphericalangle[R(t - \tau/2) \quad R(t) \quad R(t + \tau/2)] \quad (12)$$

$$A0 = [S \circ T_0^\Theta] \quad (13)$$

$$A1 = \begin{bmatrix} S \circ T_0^\Theta \\ S \circ T_1^\Theta \end{bmatrix} \quad (14)$$

$$A2 = \begin{bmatrix} S \circ T_0^\Theta \\ S \circ T_2^\Theta \end{bmatrix} \quad (15)$$

$$A3 = \begin{bmatrix} S \circ T_0^\Theta \\ S \circ T_1^\Theta \\ S \circ T_2^\Theta \end{bmatrix} \quad (16)$$

$$\tau \in \{0.2M, 0.4M, 0.6M, 0.8M, M\} \quad (17)$$

Three space transformation approaches are shown as (14-16), denoted as  $A1$ ,  $A2$  and  $A3$ , compared with  $A0$  in (13) which is actually the original time space.  $A1$  space includes the amplitude and derivative information to reveal both point-wise and pair-wise characteristics;  $A2$  space compromises the amplitude and angle information to cover both point-wise and piece-wise patterns;  $A3$  is expected to present all the above information at the same time. A larger number of dimensions can be easily achieved but also bring too much power burden in wearable applications. In (15-16), the parameter  $\tau$  of the angle basis  $T_2^\Theta[R(t)]$  is chosen as 20%, 40%, 60%, 80% or 100%

of the typical length of the QRS complex as shown in (17) where  $M$  is the typical length of the QRS complexes to be tuned later, to further explore the effectiveness of the angle basis function in constructing phase portraits. Therefore, twelve spaces are presented:  $A0$ ,  $A1$ ,  $A2 - 20\%/40\%/60\%/80\%/100\%$ , and  $A3 - 20\%/40\%/60\%/80\%/100\%$ . We will evaluate all these spaces and determine a proper one for use.

### 2.2.5 Time-domain Dynamic Time Warping (TD-DTW)

The original DTW approach is performed on a time series directly, i.e., in the  $A0$  space without space transformation. This pattern recognition process includes three steps which can be executed in real-time [26]. Firstly, a local distance table is created between a template  $Y$  of length  $M$  as (18) and a testing sequence  $X$  of length  $N$  as (19). The sequence  $X$  is equivalent to the 30-minute ECG stream  $\{R(t)|0 \leq t \leq N - 1\}$  in our case. And the template  $Y$  is chosen to be the most representative QRS complex in the first 5-minute data of  $X$ , which will be introduced in the next parameter tuning phase. The local distance table is a  $N - by - M$  matrix with the element  $d_{i,j}$  (local distance) equaling to the Euclidian distance between sample  $x_i$  in  $X$  and sample  $y_j$  in  $Y$ , defined as (20). An example of the local distance table is given in Figure 2.2 (left part), where  $X$  is a signal sequence including two segments similar to the template  $Y$ . This local distance table is to evaluate all possible point-to-point distance between  $X$  and  $Y$  ( $NM$  possibilities).

$$Y = \{y_j | 0 \leq j \leq M - 1\} \quad (18)$$

$$X = \{x_i | 0 \leq i \leq N - 1\} \quad (19)$$

$$d_{i,j} = |x_i - y_j|, \quad \forall i, \forall j \quad (20)$$

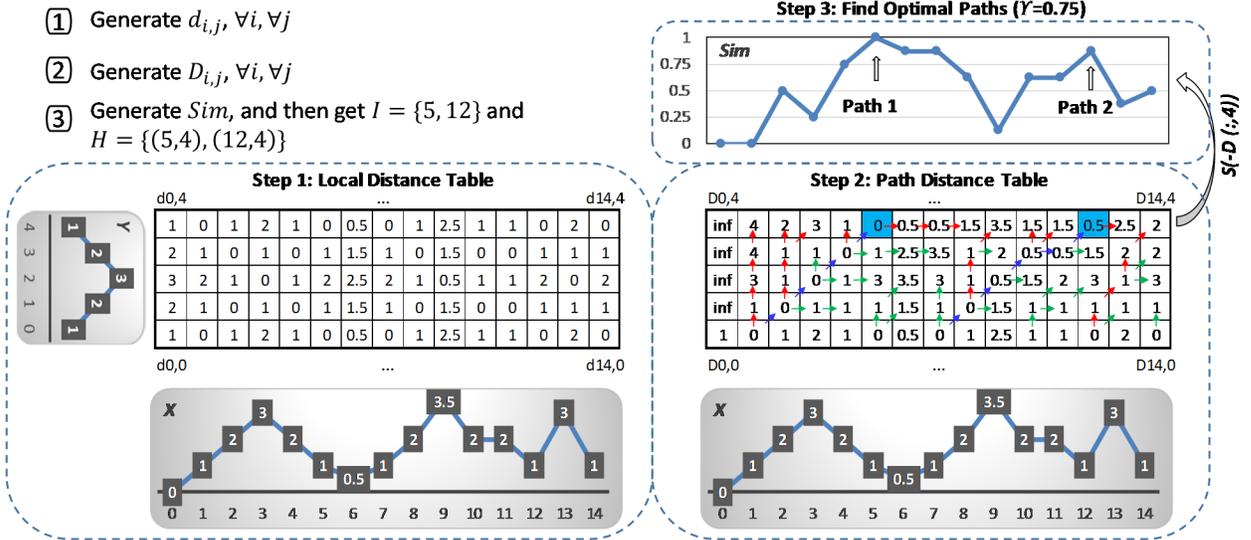


Figure 2.2. An example of the TD-DTW algorithm to recognize from the sequence X two segments (x1 to x5, and x7 to x12), with a template Y (y0 to y4).

$$D_{i,j} = \begin{cases} d_{i,j} + \min \begin{cases} D_{i-1,j} \\ D_{i-1,j-1} \\ D_{i,j-1} \end{cases} & \forall i > 0, \forall j > 0 \\ d_{i,j} & \forall i, j = 0 \\ inf & i = 0, \forall j > 0 \end{cases} \quad (21)$$

Secondly, a path distance table is generated based on the local distance table, with the element  $D_{i,j}$  (except the first row and the first column) equaling to the minima of three preceding path distances plus current local distance  $d_{i,j}$ , as shown in (21). In Figure 2.2 (bottom right part), when establishing a path distance table, two kinds of paths are also generated, the surviving paths (red or blue arrows) which can reach the last element of Y, and non-surviving paths (green arrows). Both red and blue arrows indicated paths can be surviving paths, but only the blue arrows indicated paths are the paths of interest since they give the optimal matching between the testing sequence

and the template, which will be introduced in the next step. Finally, from the surviving paths, one or more optimal paths are identified by a threshold-based approach as (22-23), where  $D_{:,M-1}$  corresponds to the last column of the path distance table,  $Sim$  is a similarity sequence bound by 0 and 1 obtained by scaling the raw similarity sequence  $-D_{:,M-1}$  and  $\alpha^D$  and  $\beta^D$  are its maxima and minima, respectively,  $findpeaks$  is a function to search out from  $Sim$  one or more peaks above a pre-defined similarity threshold  $Y$  ( $0 < Y < 1$ ), and  $I$  is a set of row indices of the found peaks. Thus the ending points of all optimal paths  $H$  in the path distance table, can be expressed as (24), in terms of matrix indices. Now the segments in  $X$  similar to  $Y$  have been recognized.

As shown in Figure 2.2 (top right part), given  $Y = 0.75$ , two peaks are identified ( $D_{5,4}$  and  $D_{12,4}$ ). Consequently, two optimal paths are determined (blue arrows), and thus the peak ( $y_2$ ) in  $Y$  is matched with two peaks ( $x_3$  and  $x_9$ ) in  $X$ , but the interferential peak ( $x_{13}$ ) is successfully eliminated.

This is a simple example to show the mechanism of pattern recognition using the DTW approach. However, in the long-term wearable heart rate monitoring, the ECG signals are of a much poorer signal quality because of large amounts of continuous motion artifacts, therefore an advanced version of DTW called phase-domain multi-view dynamic time warping, is introduced later.

$$Sim = S(-D_{:,M-1}) = \frac{(-D_{:,M-1}) - \beta^D}{\alpha^D - \beta^D} \quad (22)$$

$$I = findpeaks(Sim > Y) \quad (23)$$

$$H = \{(i, j) | i \in I, j = M - 1\} \quad (24)$$

### 2.2.6 Phase-domain Multi-view DTW (PD-MDTW)

With the space transformation, the multi-view DTW (MDTW) approach is performed on the multi-dimensional ECG data in the phase space for QRS complex recognition.  $Y$  in (19) now corresponds to a multi-dimensional QRS complex template represented by a  $L - by - M$  matrix, and  $X$  in (18) is now a  $L - by - N$  matrix, both in the phase domain. Therefore, the local distance  $d_{i,j}$  in (20) is now replaced with (25), where both  $X$  and  $Y$  now have a dimension of  $L$ . The following procedure of MDTW is similar to that of DTW.

$$d_{i,j} = \sqrt{\sum_{l=0}^{L-1} (x_{l,i} - y_{l,j})^2} \quad \forall i, \forall j \quad (25)$$

The contribution of applying MDTW to the phase portraits is two-fold, i.e., excluding the motion artifact-induced interferential spikes and identifying the corrupted QRS complexes. For the former one, each local distance element  $d_{i,j}$  is now calculated considering not only the amplitude information but also other information like derivative and angle. When the interferential spikes highly mimic the QRS complexes in the amplitude dimension, it is less possible for them to also highly mimic the QRS complexes in other dimensions, and vice versa. Therefore, the local distance elements can help capture more inconsistencies between the interferential spikes and the QRS complex template. Subsequently, when generating the path distance table by exhaustively searching over all possible warping paths, these inconsistencies usually lead to non-surviving or non-optimal paths, which is expected to suppress many false positives (incorrectly recognized heartbeats, i.e., interferential spikes).

For the latter one, the local distance element  $d_{i,j}$  can now reveal the characteristics of corrupted QRS complexes from more than one views, therefore, the calculated local distance elements are more robust and can improve the consistencies between the corrupted QRS complexes and the template. When generating the path distance table, these improved consistencies can help provide more opportunities to form surviving paths and also optimal paths for the corrupted QRS complexes, which is expected to suppress many false negatives (missing heartbeats). Graphical results are given in the next section to further demonstrate the improved pattern recognition ability of the PD-MDTW.

### 2.2.7 Parameter Tuning Phase

To guarantee the convenience in practical applications, for each 30-minute recording, only the first 5-minute data is used to tune the parameters related to this recording (offline), and the remaining 25-minute to test the algorithm and tuned parameters (real-time). There is also no dependency among recordings in the datasets for real-time purpose.

The parameters to be tuned include: 1) the parameters  $\alpha^l$  and  $\beta^l$  in (8) to scale each transformed sequence in the phase space, 2) the typical length of the QRS complexes  $M$  in (17), 3) the QRS complex template  $Y$  in (18), 4) the parameters  $\alpha^D$  and  $\beta^D$  in (22) to scale the DTW similarity sequence and 5) the DTW similarity threshold  $Y$  in (23) used to identify target signal segments similar to the template. Considering that the ECG data may be highly diverse, the parameters are not tuned once and then fixed. Instead, the following parameter tuning strategy is designed to be able to adaptively tune these parameters for specific ECG data. In the results section, the proposed strategy is validated on the acquired ECG dataset, and then is well generalized to the public dataset.

The scaling parameters  $\alpha^l$  and  $\beta^l$  are the maxima and minima of the first 5-minute data of the  $l - th$  transformed sequence  $T^l[R(t)]$ , respectively, and  $\alpha^D$  and  $\beta^D$  are also determined in a similar way based on the raw DTW similarity sequence  $-D_{:,M-1}$ . These parameters will be used to scale the remaining 25-minute data and DTW similarity sequence.

The typical length of the QRS complexes  $M$  in one ECG recording, is obtained by averaging the length of  $K$  QRS complexes in the first 5-minute data, i.e., the tuning data of this recording. The template  $Y$  for one recording is determined by a Euclidean distance-based method for simplicity purpose [37], which includes two steps. Firstly, considering the  $K$  QRS complexes in the tuning data may own different lengths, they are all resampled to have a length of  $M$ , by the ‘*interp1*’ function in Matlab with a spline interpolation or decimation operation. The  $K$  resampled QRS complexes are denoted as (26), where  $Q^k$  is the  $k - th$  resampled QRS complex and  $q_{ij}^k$  is the  $j - th$  element in the  $l - th$  dimension of this QRS complex. Secondly, after defining the Euclidean distance between two resampled QRS complexes as (27), the template  $Y$  is determined as (28), which searches an optimal QRS complex from all  $K$  resampled QRS complexes in the tuning data. This optimal QRS complex owns a minimum cumulative Euclidean distance to all the  $K$  QRS complexes in the tuning data. The Euclidean distance is chosen here considering that it is of a low computation cost and can reduce the tuning time.

The optimal DTW similarity threshold  $Y$  is determined based on (29), i.e., the threshold  $th$  that can maximize the average value of the heartbeat recognition sensitivity  $Sens_{th}$  (30) and precision  $Prec_{th}$  (31) on the first 5-minute data of each recording, where  $TP_{th}$ ,  $FN_{th}$  and  $FP_{th}$  are the numbers of true positives (correctly recognized heartbeats), false negatives (missing heartbeats) and false positives (incorrectly recognized heartbeats) corresponding to the threshold  $th$ ,

respectively. A step size of 0.05 is chosen when evaluating the threshold  $th$  between 0 and 1 to find the optimal threshold  $Y$ . This learned threshold is expected to maximize the averaged performance.

$$Q^k = \{q_{lj}^k | 0 \leq l \leq L - 1, 0 \leq j \leq M - 1\},$$

$$\forall k \in \{0, 1, \dots, K - 1\} \quad (26)$$

$$\|Q^i - Q^k\| = \sqrt{\sum_{l=0}^{L-1} \sum_{j=0}^{M-1} (q_{lj}^i - q_{lj}^k)^2} \quad (27)$$

$$Y = Q^i = \underset{0 \leq i \leq K-1}{\operatorname{argmin}} \sum_{k=0}^{K-1} \|Q^i - Q^k\| \quad (28)$$

$$Y = \operatorname{argmax}_{0 < th < 1} \frac{Sens_{th} + Prec_{th}}{2} \quad (29)$$

$$Sen_{th} = \frac{TP_{th}}{TP_{th} + FN_{th}} \quad (30)$$

$$Prec_{th} = \frac{TP_{th}}{TP_{th} + FP_{th}} \quad (31)$$

### 2.2.8 Testing Phase and Performance Evaluation

The testing phase comprises two operations, i.e., heartbeat recognition and heart rate estimation, which are performed in real-time. Firstly, based on tuned parameters, the high confident QRS complexes (R-peaks included) are recognized by TD-DTW in the original space or PD-MDTW in the phase space. The heartbeat recognition performance in the testing phase is also evaluated by sensitivity and precision, denoted as  $Sens_Y$  and  $Prec_Y$  defined in a similar way as (30-31), but using the tuned DTW similarity threshold  $Y$ . Afterwards, the instantaneous heart rate

is estimated based on recognized R-peaks. To further improve the robustness, the heart rate estimates more than 208 or less than 40 beats per minute (BPM) are discarded [38], and then replaced with the previous ones. To ease the comparison with the ground truth, a 2-second window is used to report the instantaneous heart rate, and both mean absolute error (MAE) and the root mean square error (RMSE) in terms of BPM are given.

### 2.2.9 Performance Comparison

To demonstrate the effectiveness of PD-MDTW, not only TD-DTW but *Kalman filtering* (KLMF), *Weighted regularized least square* (WRLS) and *Discrete wavelet transformation* (DWT) approaches are also implemented for comparison [23-25]. For KLMF and WRLS, we employed similar methods as those used by Gribok *et al.* [23], where after applying a well-known Pan & Tompkins algorithm for QRS detection, and an impulse rejection filter for spike-like outlier identification, the IHR is then estimated. In KLMF [23, 24], to get an IHR estimate based on both a predicted value and a measured one, the  $1 \times 1$  state noise covariance matrix  $Q$  is set to 0.1, and the  $1 \times 1$  measurement noise covariance matrix  $M$  is computed as  $M = M_0 \cdot \exp(1/w^2 - 1)$ , where  $w$  is a weighting factor with  $w = 1$  for non-outliers and  $w = 10^{-5}$  for outliers, and  $M_0$  is chosen to minimize the heart rate error in the tuning phase. The tuned parameter is then used on the testing data.

In WRLS [23], the R peak-to-R peak interval (RRI) sequence of length  $N$  is estimated by solving a weighted regularized least square problem, i.e., minimizing  $\|A \cdot W \cdot RRI_{OLS} - A \cdot W \cdot RRI_{WRLS}\|^2 + \lambda^2 \cdot \|L \cdot RRI_{WRLS}\|^2$ , where  $A$  is an  $N \times N$  lower triangular integration matrix,  $RRI_{OLS}$  and  $RRI_{WRLS}$  denote the estimated RRI sequence as a solution to an ordinary least square

problem (OLS) and the one to be sought, respectively,  $W$  is a diagonal  $N \times N$  weighting matrix where each element  $w$  is determined as that in KLMF,  $L$  is a  $(N - 2) \times N$  smoothing matrix representing a second-order derivative, and  $\lambda$  is a regularization parameter to control the smoothness of the fitted curve and is tuned like  $M_0$  in KLMF. In DWT [25], after decomposing the ECG stream into 8 levels using the Daubechies6 wavelet, and then reconstructing it using the 3<sup>rd</sup>/4<sup>th</sup>/5<sup>th</sup> detail coefficients, the QRS complexes are detected by a threshold-based method and the IHR is finally calculated and purified by a median filter with an order of 10. The threshold used in DWT is also selected by the same criterion as the one for  $M_0$  in KLMF. All these frameworks are evaluated on WECG and FECG-based datasets.

## 2.3 RESULTS

### 2.3.1 Domain Transformation

The clean ECG segment (Figure 2.3a-d) and also a corrupted version (Figure 2.3e-h) are illustrated in four spaces ( $A_0$ ,  $A_1$ ,  $A_2 - 40\%$  and  $A_3 - 40\%$ ), respectively. For the clean ECG segment, Figure 2.3a corresponds to the original space, and Figure 2.3b-d show more geometrical characteristics of the QRS complexes in the phase spaces.

In space  $A_1$  (Figure 2.3b1-b4), the QRS complexes form consistent phase portraits leveraging both amplitude and derivative information, which are much more distinguishable patterns than those in  $A_0$ . Similarly, in Figure 2.3c1-c4 and Figure 2.3d1-d4, the phase portraits of the QRS complexes are also shown in  $A_2 - 40\%$  (amplitude and angle dimensions) and  $A_3 - 40\%$  (amplitude, derivative and angle dimensions) spaces, respectively, with well-constructed patterns including many new geometrical properties.

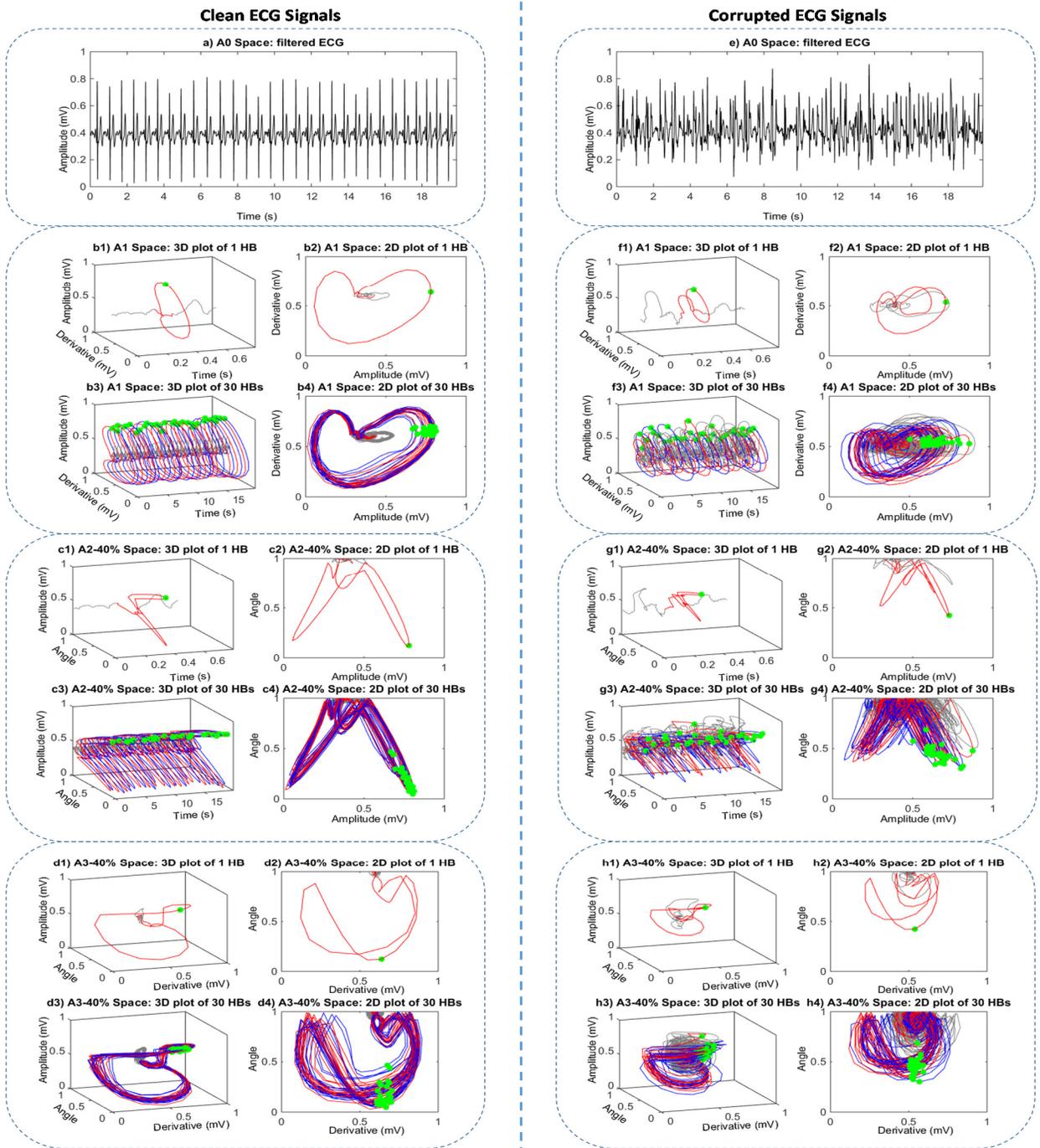


Figure 2.3. Space transformation for the clean and continuously/heavily corrupted ECG signals. Notes. The results are corresponding to a 20-second signal segment, from dataset WECG and WECG\_Sm9\_SH0, respectively. Red and blue lines: QRS complexes; green dots: R peaks (ground-truth heartbeat locations); HB: heartbeat.

The continuously and heavily corrupted ECG signals (Figure 2.3e) make the QRS complexes hard to be identified, however, the phase portraits of them can still be well recognized in the phase spaces (Figure 2.3f-h), where many geometrical properties of heartbeat are revealed not only point-wisely (amplitude) but also pair-wisely (derivative) and piece-wisely (angle). Moreover, many interferential spikes which highly mimic the real QRS complexes are also well suppressed leveraging new geometrical properties of heartbeats. Therefore, these multi-dimensional phase portraits are expected to be effective patterns to enable robust heartbeat recognition, and thus enhance the robustness of the instantaneous heart rate estimation.

### 2.3.2 TD-DTW and PD-MDTW

The similarity sequences measured by TD-DTW or PD-MDTW, are illustrated in Figure 2.4, based on the same signal segment as that used in Figure 2.3, with different signal quality levels. According to the *findpeaks* function in (23), a good similarity sequence is expected to present well-formed QRS complex-related peaks which should be consistently above a similarity threshold and thus can be easily selected out. For the clean ECG signals (Figure 2.4a1), we can see that the similarity curves measured by the PD-MDTW (Figure 2.4a3-a5) are of much better morphologies, i.e., much smoother and with less interferential peaks than that of TD-DTW (Figure 2.4a2). Furthermore, the pattern recognition ability is also proved to be space-specific since the phase portraits own different geometrical properties, and the ranking is:  $A3 - 40\% > A2 - 40\% > A1 > A0$ .

When decreasing the signal quality from -1 dB to -9 dB with a step of -2 dB (Figure 2.4b1-5/c1-5/d1-5/e1-5/f1-5), there are more interferential spikes mimicking or corrupting the QRS

complexes, and the morphologies of similarity curves also get worse. However, the trend is all the same, i.e.,  $A3 - 40\% > A2 - 40\% > A1 > A0$ , which will be further confirmed by the next heart rate estimation results.

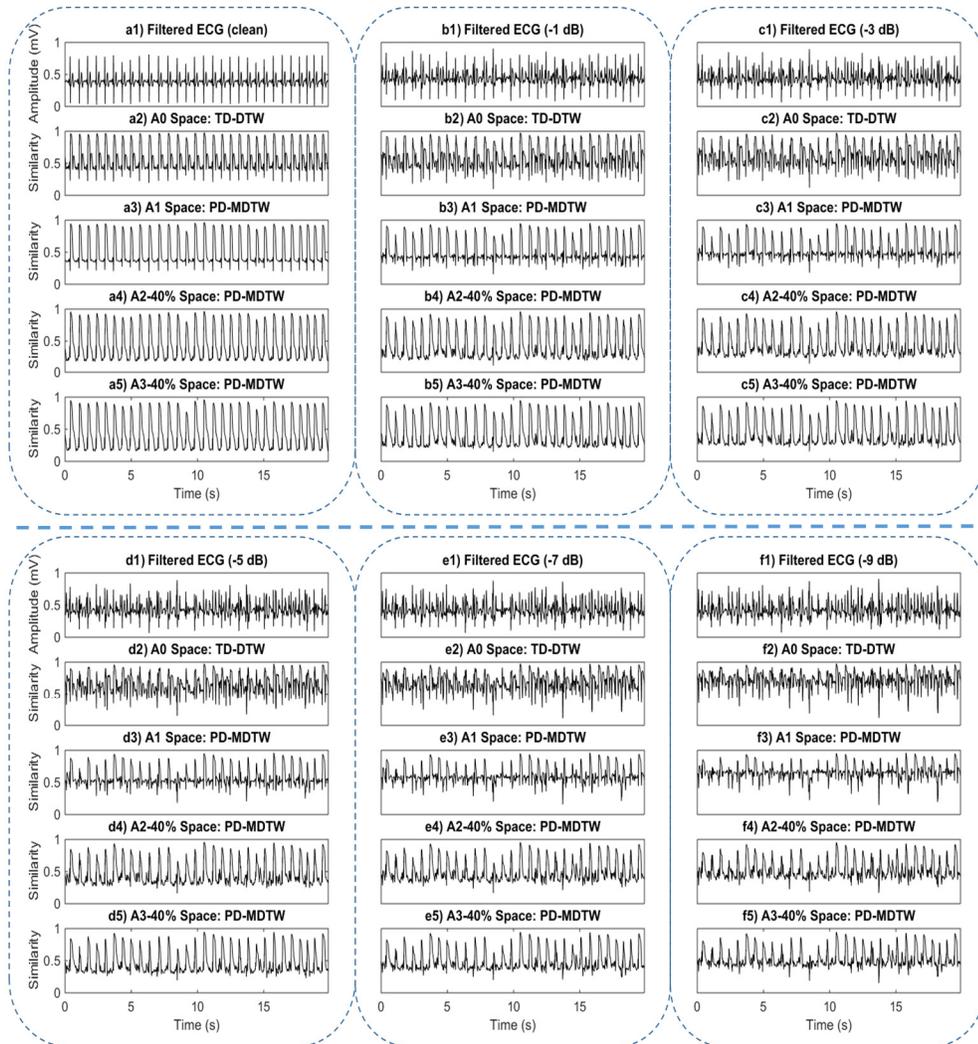


Figure 2.4. Similarity sequences measured by the TD-DTW or the PD-MDTW on the ECG signals with different quality levels.

Notes. The results are corresponding to a 20-second signal segment of one subject, from dataset WECG and WECG\_Sm1/3/5/7/9\_SH0, respectively.

Specially, with an SNR of -9 dB, the similarity curve of  $A1$  transformation is so poor that it is difficult to be used to effectively detect the QRS complex-related similarity peaks through the threshold-based approach.  $A1$  transformation slightly outperforms  $A0$  transformation, but is worse than  $A2 - 40\%$  transformation, indicating that the piece-wise angle information is more effective than the pair-wise derivative information to form the artifact-tolerant phase portraits, while  $A3 - 40\%$  transformation presents the best performance showing that the 3-dimensional phase portraits (of point/pair/piece-wise information) own the strongest motion artifact-tolerant ability. The comparison shows the effectiveness of PD-MDTW in phase-domain QRS complex pattern recognition.

To further compare the pattern recognition ability of TD-DTW and PD-MDTW, Figure 2.5 gives the similarity sequences, path distance tables and local distance tables for both methods, based on the same corrupted signal segment (-9 dB). For TD-DTW, the QRS complex identification results include many false positives (red up-pointing triangles in Figure 2.5a) and also many false negatives (red down-pointing triangles). Taking the red solid boxes (false positives) in Figure 2.5a as an example, their corresponding similarity peaks are above the similarity threshold unexpectedly, because their local distance elements in Figure 2.5c are of relatively small values (the blue color covers most of the table), which means they cannot effectively reflect the difference between the interferential spikes and the QRS complex template only by the amplitude information. Consequently, the low local distance values cause surviving warping paths in Figure 2.5b. These surviving warping paths will be selected as optimal paths (white dashed lines, each optimal path corresponds to a heartbeat detected) if corresponding similarity scores are above the similarity threshold.

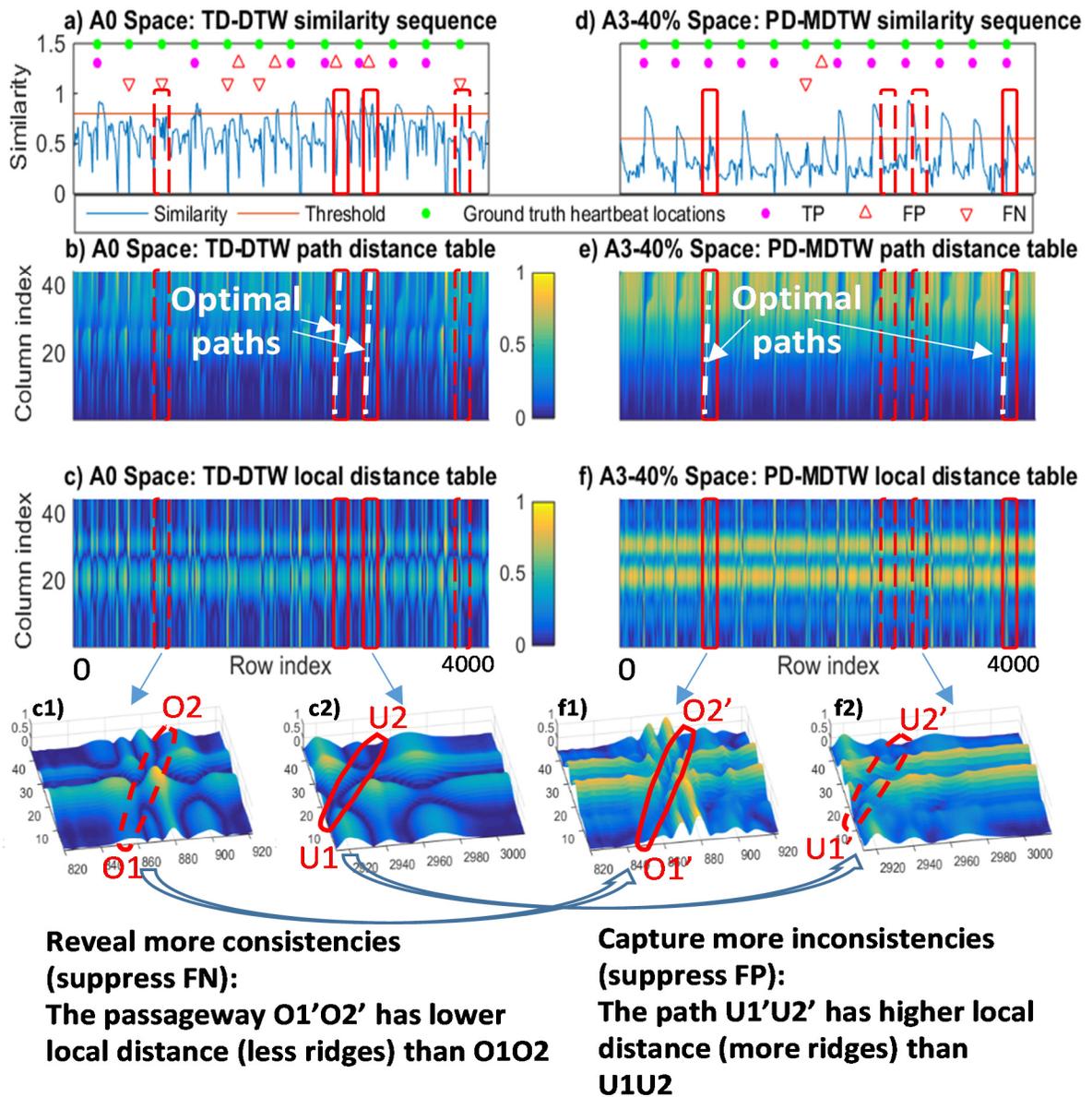


Figure 2.5. Pattern recognition ability of TD-DTW (a-c) versus PD-MDTW (d-f).  
 Notes. The results are based on a corrupted ECG segment from one recording in the dataset WECG\_Sm9\_SH0. Dashed red boxes in (a-c): two false negative examples; solid red boxes in (d-f): two successfully identified heartbeat examples which are missing in (a-c); solid red boxes in (a-c): two false positive examples; dashed red boxes in (d-f): two successfully excluded interferential spike examples which are incorrectly identified as heartbeats in (a-c); yellow color in b/c/e/f): high distance value; blue color in b/c/e/f): low distance value; row index corresponds to the testing data; column index corresponds to the QRS template. All the sequences and tables are scaled to be between 0 and 1.

Moreover, taking the red dashed boxes (false negatives) in Figure 2.5a as an example, their similarity peaks are under the similarity threshold unexpectedly. This is because the low quality local distance elements cannot effectively reveal more characteristics from corrupted QRS complexes, and thus cannot generate optimal paths (Figure 2.5b) for them. This causes missing heartbeats.

However, PD-MDTW (Figure 2.5d-f) has successfully suppressed most of these false positives and false negatives. Leveraging multi-view information and well-formed phase portraits (Figure 2.3), the local distance table can effectively capture more inconsistencies between the interferential spikes and the QRS complex template (e.g., Figure 2.5f2 versus Figure 2.5c2), and thus prevent optimal warping paths (red dashed boxes in Figure 2.5e) for them. On the other hand, the local distance table can also reveal more consistencies between the corrupted QRS complexes and the template (e.g., Figure 2.5f1 versus Figure 2.5c1), and thus can provide corresponding optimal warping paths for them (red solid boxes in Figure 2.5e). This two-fold contribution makes PD-MDTW more powerful than TD-DTW.

### **2.3.3 Heartbeat Recognition and Heart Rate Estimation**

After applying TD-DTW or PD-MDTW to the corrupted ECG signals, the high confident QRS complexes are recognized and consequently R peaks are determined for heart rate estimation. Twelve spaces are all evaluated on five different SNRs (Datasets WECG\_Sm1/3/5/7/9\_SH0), and the heartbeat recognition and the IHR estimation performance are shown in Figure 2.6a-e and Figure 2.6f-j, respectively.

When decreasing the SNR from -1 dB to -9 dB, the performance in the original space ( $A_0$ ) and phase spaces (other eleven spaces) all continuously decrease. Remarkably, nine out of eleven phase

spaces (except  $A2/3 - 100\%$  where the time shift  $\tau$  used in (12) for piece-wise angle information extraction is too large) all outperform the original space ( $A0$ ), indicating the motion artifacts-tolerant capability of these spaces.

Furthermore, in  $A2$  and  $A3$  spaces, the time shift  $\tau$  used for angle information extraction is finally chosen as 40% of the length of the QRS complex template. This is based on the consideration that a larger time shift may not be able to effectively track the local characteristics of the heartbeats, and a smaller one may suffer from the interference from large amounts of motion artifact-induced spikes. Therefore, we take special interest in  $A0$ ,  $A1$ ,  $A2 - 40\%$  and  $A3 - 40\%$  for next exploration.

Then the further evaluation of the above four chosen spaces are implemented on twenty *WECG*-based datasets *WECG\_Sm1/3/5/7/9\_SH0/1/2/3/4* not only with different SNRs and but also with five phase shifts for each SNR, as shown in Table 2.1. The results with the same SNR are further averaged to obtain the SNR-specific performance.

The ranking is basically as follows:  $A3 - 40\% > A2 - 40\% > A1 > A0$ , consistent with Figure 2.4. Specifically, -9 dB-specific sensitivity, precision, MAE and RMSE are improved from 78.9%, 86.6%, 9.7 BPM and 18.5 BPM in  $A0$  space, to 87.9%, 98.3%, 2.5 BPM and 7.0 BPM in  $A3 - 40\%$  space, respectively.

The purpose of Table 2.1 is to thoroughly compare the performance in different spaces and illustrate the effectiveness of space  $A3 - 40\%$ . Other approaches for comparison purpose are given in Table 2.2.

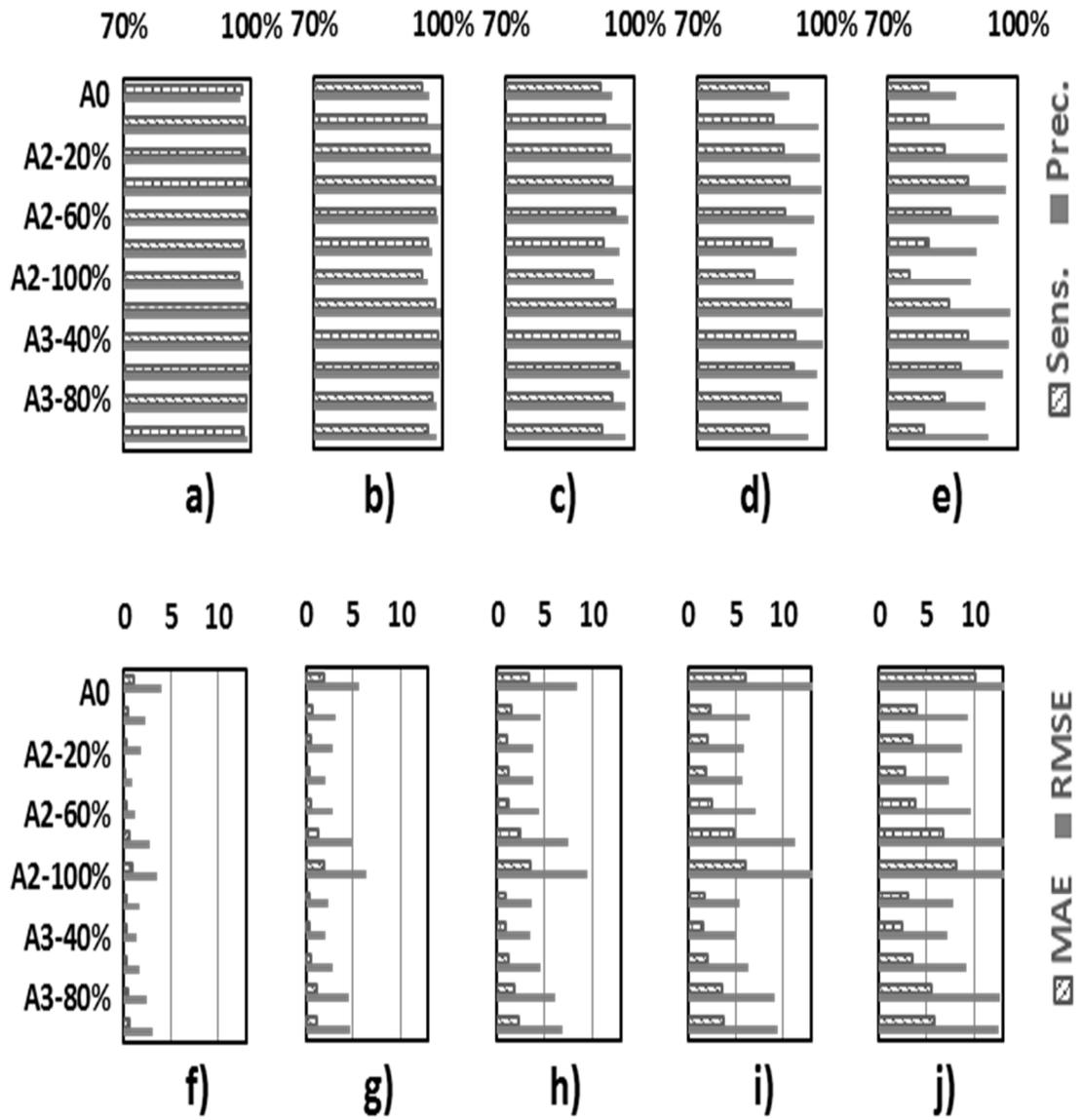


Figure 2.6. Performance comparison of twelve spaces over different SNRs.  
 Notes. Twelve spaces: A0, A1, A2-20%/40%/60%/80%/100%, A3-20%/40%/60%/80%/100%;  
 five SNRs: WECG\_Sm1/3/5/7/9\_SH0; a-e): sensitivity/precision of heartbeat recognition @ five SNRs; f-j): MAE/RMSE of heart rate estimation @ five SNRs.

Table 2.1. Performance summary of four typical spaces (A0/1/2-40%/3-40).

Datasets (WECG_)	A0 Space (TD-DTW)				A1 Space (PD-MDTW)			
	Sens. (%)	Prec. (%)	MAE (BPM)	RMSE (BPM)	Sens. (%)	Prec. (%)	MAE (BPM)	RMSE (BPM)
Sm1_SH0	97.9	97.6	1.1	4.0	98.7	99.7	0.4	2.2
Sm1_SH1	97.7	97.5	1.2	4.5	98.8	99.3	0.5	3.0
Sm1_SH2	98.1	97.3	1.2	4.4	98.9	99.4	0.5	2.8
Sm1_SH3	97.7	98.2	1.0	3.8	98.5	98.8	0.8	3.2
Sm1_SH4	98.1	97.4	1.2	4.3	98.8	96.8	1.7	5.2
<b>Averaged @ -1 dB</b>	<b>97.9</b>	<b>97.6</b>	<b>1.2</b>	<b>4.2</b>	<b>98.7</b>	<b>98.8</b>	<b>0.8</b>	<b>3.3</b>
<b>Averaged @ -3 dB</b>	<b>96.1</b>	<b>96.0</b>	<b>2.2</b>	<b>6.4</b>	<b>96.1</b>	<b>99.4</b>	<b>1.0</b>	<b>4.0</b>
<b>Averaged @ -5 dB</b>	<b>92.6</b>	<b>94.4</b>	<b>3.6</b>	<b>9.2</b>	<b>92.8</b>	<b>99.2</b>	<b>1.5</b>	<b>5.0</b>
<b>Averaged @ -7 dB</b>	<b>87.2</b>	<b>89.9</b>	<b>6.7</b>	<b>14.1</b>	<b>87.6</b>	<b>98.4</b>	<b>2.5</b>	<b>6.8</b>
Sm9_SH0	79.3	85.9	10.1	18.7	79.4	96.9	4.0	9.2
Sm9_SH1	78.0	88.0	8.9	17.5	79.6	96.7	4.2	9.8
Sm9_SH2	79.1	86.5	9.6	18.3	79.5	96.5	4.3	10.2
Sm9_SH3	78.1	86.9	9.8	18.9	79.3	96.5	4.2	9.6
Sm9_SH4	79.8	85.9	10.0	18.9	79.7	96.3	4.4	10.0
<b>Averaged @ -9 dB</b>	<b>78.9</b>	<b>86.6</b>	<b>9.7</b>	<b>18.5</b>	<b>79.5</b>	<b>96.5</b>	<b>4.2</b>	<b>9.8</b>

Notes. For SNR@-3/-5/-7 dB, only averaged results are given for good readability

Table 2.1. (Continued) Performance Summary of Four Typical Spaces.

Datasets (WECG_)	A2-40% Space (PD-MDTW)				A3-40% Space (PD-MDTW)			
	Sens. (%)	Prec. (%)	MAE (BPM)	RMSE (BPM)	Sens. (%)	Prec. (%)	MAE (BPM)	RMSE (BPM)
Sm1_SH0	99.4	99.9	0.2	1.0	99.7	99.8	0.2	1.4
Sm1_SH1	99.4	99.9	0.2	1.2	99.7	99.7	0.2	1.8
Sm1_SH2	99.4	99.9	0.2	1.2	99.7	99.7	0.2	1.3
Sm1_SH3	99.4	99.9	0.2	1.1	99.7	99.7	0.2	1.4
Sm1_SH4	99.2	99.9	0.2	1.4	99.7	99.9	0.1	0.9
<b>Averaged @ -1 dB</b>	<b>99.4</b>	<b>99.9</b>	<b>0.2</b>	<b>1.2</b>	<b>99.7</b>	<b>99.8</b>	<b>0.2</b>	<b>1.4</b>
<b>Averaged @ -3 dB</b>	<b>98.0</b>	<b>99.8</b>	<b>0.5</b>	<b>2.4</b>	<b>98.9</b>	<b>99.8</b>	<b>0.4</b>	<b>2.3</b>
<b>Averaged @ -5 dB</b>	<b>94.8</b>	<b>99.6</b>	<b>1.1</b>	<b>4.0</b>	<b>96.8</b>	<b>99.5</b>	<b>0.9</b>	<b>3.8</b>
<b>Averaged @ -7 dB</b>	<b>91.6</b>	<b>99.0</b>	<b>1.8</b>	<b>5.5</b>	<b>93.3</b>	<b>99.1</b>	<b>1.6</b>	<b>5.3</b>
Sm9_SH0	88.5	97.5	2.6	7.3	88.7	98.1	2.5	7.2
Sm9_SH1	87.4	97.6	2.9	7.8	87.4	98.4	2.5	6.9
Sm9_SH2	88.1	97.0	2.9	7.8	87.5	98.3	2.5	7.0
Sm9_SH3	86.9	98.0	2.7	7.5	88.0	98.2	2.6	7.4
Sm9_SH4	87.5	97.6	2.8	7.6	87.8	98.4	2.4	6.7
<b>Averaged @ -9 dB</b>	<b>87.7</b>	<b>97.5</b>	<b>2.8</b>	<b>7.6</b>	<b>87.9</b>	<b>98.3</b>	<b>2.5</b>	<b>7.0</b>

Notes. For SNR@-3/-5/-7 dB, only averaged results are given for good readability

Table 2.2. Performance comparison of five different approaches.

Datasets	KLMF		WRLS		DWT		TD-DTW		PD-MDTW (Proposed)	
	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE
	(BPM)	(BPM)	(BPM)	(BPM)						
WECG_Sm1_SHx	4.2	6.8	4.6	6.1	3.1	4.8	1.2	4.2	0.2	1.4
WECG_Sm3_SHx	17.5	22.0	11.5	14.1	4.0	7.0	2.2	6.4	0.4	2.3
WECG_Sm5_SHx	45.1	50.9	27.3	31.7	4.9	8.6	3.6	9.2	0.9	3.8
WECG_Sm7_SHx	63.1	68.0	39.0	43.2	7.1	12.8	6.7	14.1	1.6	5.3
WECG_Sm9_SHx	75.3	77.8	47.4	50.8	14.9	23.1	9.7	18.5	2.5	7.0
FECG_Sm1_SHx	2.2	5.0	2.8	4.4	2.0	3.6	1.8	6.6	0.5	2.6
FECG_Sm3_SHx	17.3	25.0	11.6	15.4	2.8	5.9	3.0	9.0	0.6	3.0
FECG_Sm5_SHx	54.4	64.0	33.6	39.5	4.1	8.1	4.4	11.5	0.8	3.5
FECG_Sm7_SHx	69.6	76.7	44.8	50.0	4.5	9.2	7.1	16.0	1.4	4.9
FECG_Sm9_SHx	80.7	84.8	53.4	57.2	7.9	14.9	12.2	22.8	2.2	6.8

Notes. The comparison are performed on noisy datasets (W/FECG\_Sm1/3/5/7/9\_SH0/1/2/3/4) and SNR-specific results are given.

### 2.3.4 Performance Comparison

To further demonstrate the effectiveness of the proposed framework (PD-MDTW in A3 – 40% space is chosen), four well-established approaches including KLMF, WRLS, DWT and TD-DTW are also employed. The performance comparison is performed on noisy datasets, and the SNR-specific results are illustrated in Table 2.2. The framework level performance is given in terms of the estimation errors of final IHR estimates, instead of intermediate results considering that these frameworks have different strategies and intermediate results may not fully reflect their ability. With an SNR of -1 dB, the performance of five approaches are comparable although the PD-MDTW shows a little bit lower estimation errors. When decreasing the SNR by a step of -2 dB to -9 dB, KLMF and WRLS both heavily drift, and DWT, TD-DTW and PD-MDTW also show a rapid deterioration in performance. Nevertheless, the PD-MDTW shows much better performance than the others on both *WECG* and *Fantasia*-based datasets, with an MAE no more than 2.5 BPM

and a RMSE no more than 7.0 BPM, respectively, showing a good generalization ability even to ECG signals with diverse age-related morphologies and beat-to-beat interval variability.

To compare the detailed behaviors of different approaches in instantaneous heart rate (IHR) estimation, an example of the estimated IHR curves is illustrated in Figure 2.7. With an SNR of -1 dB, the PD-MDTW shows the best accuracy and the curve is almost totally overlapped by the ground truth curve, while the other four approaches show worse but still comparable estimates. Nevertheless, with an SNR of -9 dB, the KLMF heavily drifts because of suffering from the assumption that the measurement noise is Gaussian distributed while the actual motion artifacts are usually induced by diverse and intense body movements that cannot be represented by a Gaussian model. The WRLS also drifts a lot because it smooths the curve heavily when the measurements are highly noisy. DWT has an ability to de-noise the data, thus it brings a better accuracy. However, as the spectrum of the motion artifacts usually overlaps that of the ECG signals, it still causes large deterioration in performance [20]. TD-DTW recognizes the heartbeat only in the time domain, therefore, it cannot effectively distinguish the heartbeats from interferential spikes.

The proposed PD-MDTW, however, leveraging the fact that the phase portraits of the heartbeats in the transformed domain more effectively reveal many highly distinguishable geometrical properties (point/pair/piece-wisely), can recognize the heartbeats of multi-dimensional patterns even with an SNR as low as -9 dB and thus can effectively fit the underlying trend of the instantaneous heart rate.

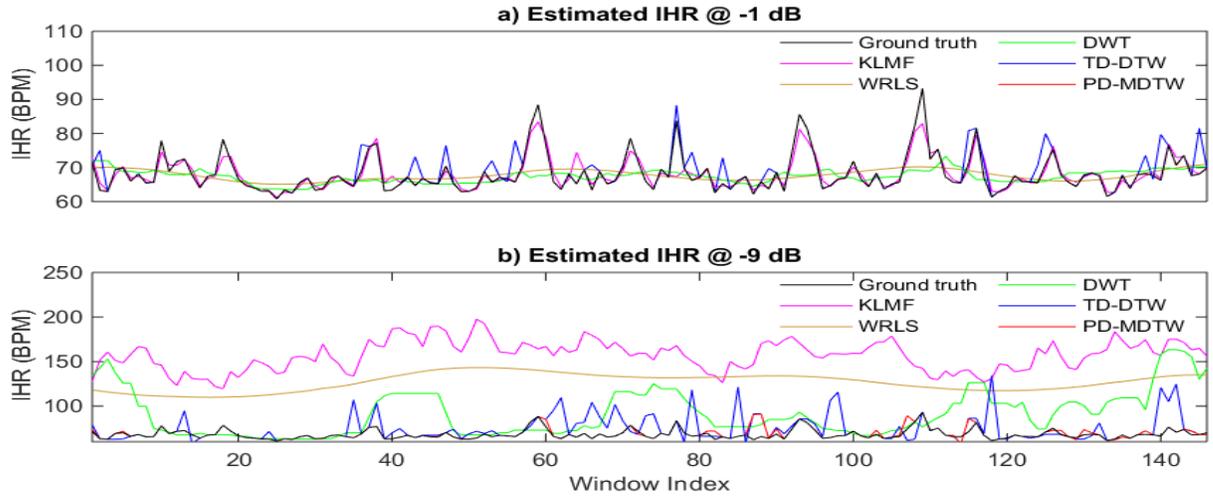


Figure 2.7. Estimated instantaneous heart rate (IHR) @ different SNRs. Notes. The results are corresponding to a 5-minute signal segment (150 2-second windows) from dataset WECG\_Sm1\_SH0 and WECG\_Sm9\_SH0, respectively.

### 2.3.5 Limitation Exploration

A broader range noisy environment is also considered to further evaluate different IHR estimators and explore the limitation of the proposed PD-MDTW. When decreasing the SNR to -11 dB on the WECG dataset, the RMSE results of five approaches KLMF, WRLS, DWT, TD-DTW and PD-MDTW, are 82.9, 54.9, 36.5, 21.8 and 9.0 BPM, respectively. These errors are all above 8.8 BPM, which is 10% of the median of the target heart rate with moderately intense activities [39] and used as a threshold to evaluate if an estimator is robust or not. When further decreasing the SNR to -13 or even -15 dB, the RMSE results for all five approaches are of much lower robustness (e.g., with an SNR of -15 dB, they are 85.0, 56.4, 55.7, 26.0 and 14.5 BPM, respectively).

Therefore, in our opinion, the proposed PD-MDTW may not be able to robustly estimate the IHR with an SNR lower than -9 dB. This limitation is due to the fact that the QRS complex phase

portraits cannot be well preserved when the information in all three dimensions (amplitude, derivative and angle dimensions) is severely corrupted by motion artifacts, and also the fact that the MDTW cannot effectively reveal consistencies from corrupted QRS complexes and capture inconsistencies from interferential spikes. Future work will introduce more information extraction approaches to construct a higher-dimensional phase portrait, and thus enhance its motion artifact-tolerant ability and improve the pattern recognition capacity of the PD-MDTW framework, to deal with more severe application scenarios.

To guarantee the robustness of the heart rate estimates, we set the lowest signal quality as -7 dB. The MAE and RMSE of estimated heart rate are no more than 1.6 and 5.3 BPM, respectively.

## 2.4 DISCUSSION

In wearable instantaneous heart rate monitoring, people may perform diverse activities for a long duration, which could cause dramatic variations in the electrode-skin contact impedance. To thoroughly evaluate the effectiveness of the algorithm in these critical scenarios, the motion artifacts under consideration should be *intense, continuous and random*. *Intense* motion artifacts usually have a spectrum more severely overlapping that of the ECG signals, and also induce many spikes mimicking the heartbeats, therefore, the de-noising techniques like the DWT [25], may not be able to reconstruct good enough signals based on the extracted signal components where many motion artifacts still coexist [20]. The original pattern recognition approach, i.e., TD-DTW [26], is also usually misled by the interferential spikes since it is hard to distinguish the corrupted heartbeats from these spikes only in the time domain. Moreover, when *intense* motion artifacts *continuously* corrupt ECG signals, some approaches may even heavily drift, such as KLMF and WRLS. KLMF assumes a Gaussian model for the measurement noise which is usually not the case

since motion artifacts often come from intense body movements, therefore, the underlying trend of the instantaneous heart rate cannot be well fitted [23, 24]. Besides, WRLS usually needs to perform heavy regulation to fit a smooth curve with large amounts of noisy measurements and thus may cause a large drift [23]. The proposed PD-MDTW, in a different way, can reveal many new geometrical characteristics of the heartbeats immune to the *intense* motion artifacts, by constructing the phase portraits of heartbeats in a high-dimensional space, and then leverage the multi-dimension pattern recognition ability of MDTW for robust heartbeat identification. Furthermore, to prevent drifting when *intense* motion artifacts are *continuous*, the PD-MDTW neither assume an inappropriate model for motion artifacts nor try to fit a heavy smoothed curve, but instead directly identify heartbeats with highest performance leveraging the motion artifact-immune patterns in the phase domain as well as the pattern recognition ability of MDTW.

To introduce *intense, continuous and random* motion artifacts to the ECG signals, we applied the well-known highly-random real motion artifacts from the *MIT-BIH Noise Stress Test Database* which has been used in many previous works and somehow regarded as a gold standard [33, 34]. Many different SNRs are evaluated to induce critical motion artifacts. Meanwhile, to further induce randomness, five phase shifts were also introduced. Many noisy datasets with different signal quality conditions were used to demonstrate the effectiveness of the proposed framework. In future, we will consider more body movements in daily applications.

The space transformation approaches proposed in this study effectively revealed many new geometrical characteristics of the heartbeats in the phase domain. One thing worth noting is that, to balance the contribution of each dimension in the transformed space, the scaling function is introduced to bound the projected data in each dimension between 0 and 1. Although this

dimension-balancing approach [28] has already shown encouraging results, further differentiating each dimension's contribution to the pattern recognition may also be helpful. This is based on the consideration that each dimension like the amplitude, derivative and angle information, may own a more or less different ability to resist the motion artifacts. In future, we will further evaluate the motion artifact-tolerant ability of each dimension, and then applying a weighted multi-view DTW approach to leverage each dimension's ability for the heartbeat recognition task.

Our proposed tuning strategy can adaptively tune all the parameters based on new ECG datasets for a good generalization. One thing worth noting is that the parameter tuning effort can also be lowered when applying the proposed algorithm to new datasets without impacting the performance, For different sampling rates, only 'DTW similarity threshold  $\gamma$  in (23)' needs to be re-tuned, which can actually also be obtained in advance based on resampled data. For different sensors, like wet, dry or non-contact ones, only the DTW similarity threshold  $\gamma$  may need to be re-tuned to cover more signal quality conditions: 1) if  $\gamma$  is tuned on the non-contact sensor, then it can be used on dry or wet sensors without re-tuning, considering that the non-contact sensor often corresponds to more diverse signal quality conditions which can usually covers those based on dry or wet sensors; 2) if  $\gamma$  is tuned on the wet sensor, it may need to be re-tuned for dry or non-contact sensors to cover more signal quality conditions to get optimal performance. For different subjects, the parameters need to be re-tuned, based on the fact that there is usually a large inter-subject ECG variability, which has been leveraged by user authorization and identification applications. In summary, for different sampling rate or sensors, only the DTW similarity threshold  $\gamma$  sometimes need to be re-tuned depending on the specific scenarios. But for different subjects, parameters usually need to be re-tuned.

## 2.5 CONCLUSION

In this study, a novel framework *phase-domain multi-view dynamic time warping* (PD-MDTW), is proposed for long-term robust instantaneous heart rate monitoring, from electrocardiogram (ECG) signals heavily and continuously corrupted by large amounts of daily motion artifacts. The corrupted ECG signals are firstly transformed to the phase domain, and then the multi-view DTW approach is applied to recognize the heartbeats for instantaneous heart rate estimation. There are three key contributions in our framework: 1) New geometrical characteristics of heartbeats are revealed point/pair/piece-wisely in a high-dimensional space where the motion artifact-tolerant phase portraits of heartbeats are constructed; 2) The multi-view DTW enables robust heartbeat identification by revealing more consistencies from corrupted heartbeats and capturing more inconsistencies from the interferential spikes, when screening the corrupted ECG stream with a pre-defined template; 3) The framework is highly scalable, meaning that new dimensions can be easily introduced into the phase portrait establishment and multi-view DTW pattern recognition.

The proposed framework is evaluated on an acquired wrist-ECG dataset and also a large public ECG dataset. With different SNRs, our approach all outperforms four previously reported methods, including the Kalman filtering, weighted regularized least square, discrete wavelet transformation and time-domain DTW. Specifically, it can robustly estimate the underlying heart rate even with an SNR as low as -7 dB (an MAE of 1.6 BPM and a RMSE of 5.3 BPM). Overall, the proposed framework is capable of recognizing high confident heartbeats from the single-lead ECG stream continuously corrupted by random and intense motion artifacts for robust instantaneous heart rate estimation, which is expected to contribute to pervasive heart health, wellness and fitness management.

## CHAPTER 3

### HEART RATE MONITORING USING

#### A TWO-STAGE CLASSIFICATION FRAMEWORK

##### 3.1 INTRODUCTION

To robustly track the instantaneous heart rate, the key factor is to effectively identify the heartbeats corrupted or contaminated by large amounts of diverse motion artifacts. In the previous chapter, we introduced a framework based on a pattern recognition approach, which firstly learns necessary pattern recognition related parameters in the phase domain and then performs the pattern recognition based on the learned parameters to find out high confident heartbeats. This framework sweeps all the signal sequence to find out possible signal segments matching the pre-defined heartbeat template, and thus may consumes too much power in wearable applications. Here we introduce another framework based on a two-stage classification approach, which directly extract characteristic information for signal spikes, and then classify these spikes as real heartbeats or motion artifacts-induced spikes [31]. Because this framework focuses on the signal spikes of interest and does not care about non-spike signal segments, it can provide a higher power efficiency in wearable applications.

Many machine learning classifiers have been successfully used for ECG-based heart disease recognition like myocardial infarction, ventricular fibrillation, arrhythmia, sleep apnea, etc. [40-43]. However, when it comes to long-term wearable heart rate monitoring applications, massive motion artifacts may dramatically corrupt the ECG recordings, which poses new challenges to current algorithms, such as how to perform segmentation on corrupted signals, extract effective features dealing with diverse motion artifacts, enhance the robustness of classification , and so on.

In this chapter, we will firstly give the system overview of the proposed framework, and then give detailed introduction of the algorithm. Finally, we will give the results and conclusion to demonstrate the effectiveness of the proposed framework.

### 3.2 SYSTEM OVERVIEW

The system diagram of the proposed framework is illustrated in Figure 3.1. The raw ECG stream corrupted or contaminated by motion artifacts is sent to the first stage of the framework. In the first stage, for signal spikes, many features extracted in various domains are selected by a spare *Support Vector Machine* (SVM), and then used to train the SVM model. Afterwards, the selected features and the trained model are used to identify the heartbeats on the unseen testing data. The heartbeats identified by the SVM classifier are fed to the second stage of the framework. A rule-based classifier is used to suppress the residual interferential spikes and an interpolation approach is also used to recover the missing heartbeats. Finally, the instantaneous heart rate can be achieved.

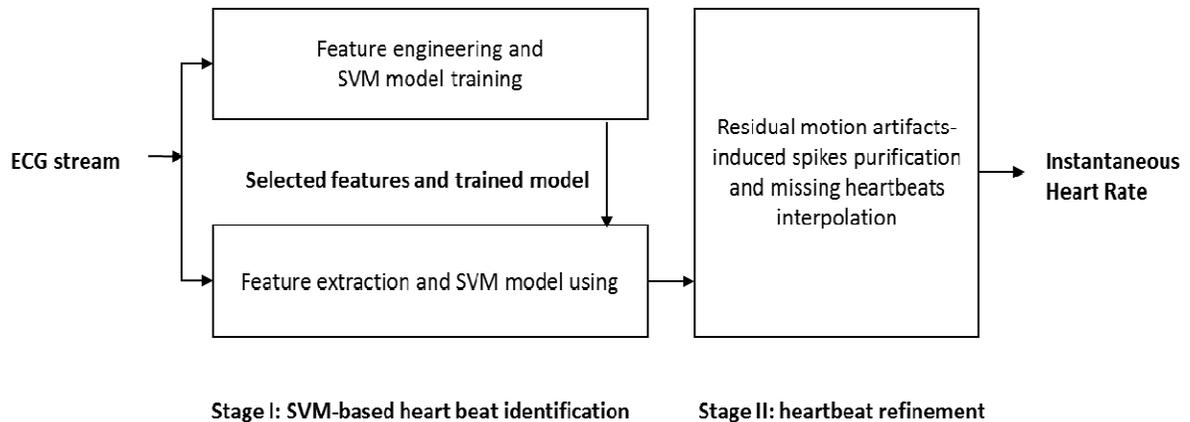


Figure 3.1. The proposed two-stage classification framework for robust instantaneous heart rate monitoring.

### 3.3 METHOD

We will firstly introduce the stage I of the framework and then give detailed introduction of the second stage.

The stage I of the framework includes signal preprocessing, heartbeats candidates generation, feature extraction, feature selection and classification. The stage II includes outlier removal, heartbeat interpolation and heart rate estimation.

#### 3.3.1 Platform and Dataset

Similar approaches are used to corrupt the WECG and Fantasia datasets (FECG), to evaluate the proposed framework. Four different SNRs (-1, -3 -5 and -7 dB) are considered to obtain a thorough analysis for different signal qualities.

Five different phase shifts (0, 0.2, 0.4, 0.6 and 0.8 second) are used when corrupting the ECG signal. The motion artifacts used to corrupt the ECG signal are from the *MIT-BIH Noise Stress Test Database* [33, 34]. The corrupted ECG datasets are named as W/FECG\_Sm1/3/5/7\_SH0/1/2/3/4.

#### 3.3.2 Preprocessing

To remove the baseline wander (low frequency component) and the high frequency noise, a bandpass filter (2 to 50 Hz) is used to process the raw ECG signal.

However, the bandpass operation cannot remove most of the motion artifacts which usually overlap the ECG signal in the frequency domain [44]. Therefore, advanced signal processing and machine learning techniques are used to identify the heartbeats from the filtered ECG signal.

### 3.3.3 Heartbeat Candidates Generation

In the raw ECG stream, there are many signal spikes. They are either heartbeats or motion-artifacts-induced interferential spikes. We use an adaptive threshold-based method to select out the signal spikes as heartbeat candidates, which can wherever possible include the heartbeat spikes. The adaptive threshold is used to track the short-term signal quality to more effectively include the heartbeat spikes. Among the heartbeat candidates generated, there are also large amounts of interferential spikes due to motion artifacts, but they will be suppressed by both the SVM classifier and the rule-based classifier later. The process is defined as the follows:

$$Th_i = Th_0 - \Delta Th_i \quad (1)$$

$$\Delta Th_i = |VPP_{i-1} - VPP_{noMA}| Coef \quad (2)$$

$i$ : the  $i - th$  time window to be considered.

$Th_i$ : the adaptive threshold calculated for the  $i - th$  time window.

$Th_0$ : the threshold for ECG signals not impacted by motion artifacts.

$\Delta Th_i$ : the adjustment quantity to adaptively change the threshold.

$VPP_{i-1}$ : the peak-to-peak voltage of the  $(i - 1) - th$  time window.

$VPP_{noMA}$ : the peak-to-peak voltage of ECG signals not impacted by motion artifacts.

$|VPP_{i-1} - VPP_{noMA}|$ : signal fluctuation term.

$Coef$ : coefficient used to adjust the contribution of the signal fluctuation term.

### 3.3.4 Feature Extraction

We extracted twenty-six features for each heartbeat candidates, which include both real heartbeats and interferential spikes due to motion artifacts. These features are extracted from multiple domains, including the time, frequency, spatial and statistical domains. These features will be further evaluated in the following feature selection step to determine the final feature subset which are critical to motion artifacts. Considering the QRS complex as shown in Figure. 3.2 is the most distinguishable segment of a heartbeat [45], the features are extracted mainly focusing on the characteristics of the QRS complex. We list all the twenty-six features in Table 3.1 and then detailed explanation is given.

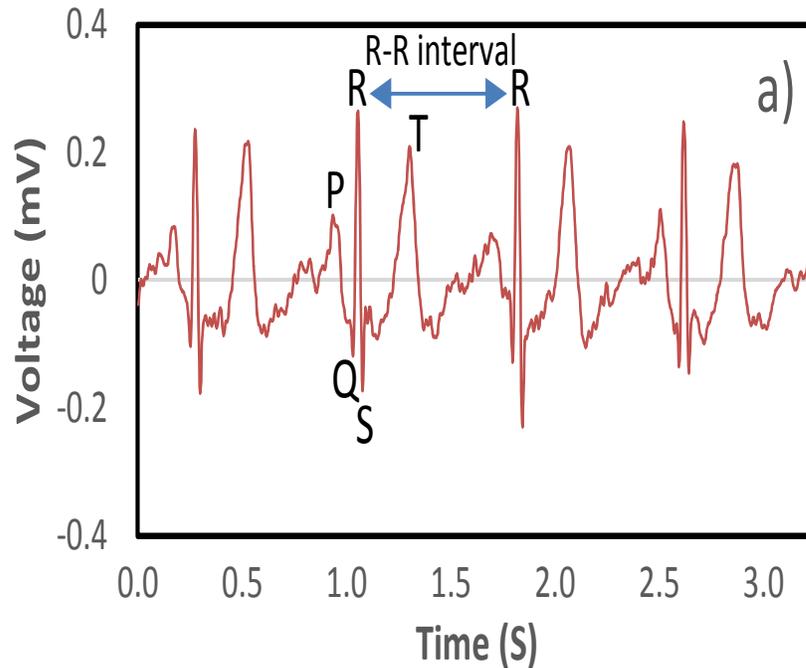


Figure 3.2. The ECG stream without motion artifacts to illustrate the characteristic points of the heartbeats.

Table 3.1. Summary of extracted features.

No.	Name
1	R Angle
2	S Angle
3	RS Slop
4	RS Difference
5	SM Difference
6	SM vs. RS
7	R Symmetry
8	S Symmetry
9	Skewness
10	Kurtosis
11	Auto-regression Coefficient 1
12	Auto-regression Coefficient 2
13	Auto-regression Coefficient 3
14	Auto-regression Coefficient 4
15	Median Frequency
16	Mean Frequency
17	Variance
18	Root Mean Square
19	Modified Median Frequency
20	Modified Mean Frequency
21	Slop Angle -3
22	Slop Angle -2
23	Slop Angle -1
24	Slop Angle 0
25	Slop Angle 1
26	Slop Angle 2

The detailed explanation is given for each feature.

1) R Angle

The R angle feature gives the angle information of the R peak and is defined by three points recorded as Q-R-S, i.e., the Q valley, the R peak and the S valley.

2) S Angle

The S angle feature gives the angle information of the S valley and is also defined by three points recorded as R-S-M, i.e., the R peak, the S valley and the first peak on the right side of the S valley M.

These two spatial features (feature 1 and 2) capture instantaneous morphological information, which may more or less survive even with most part of a heartbeat is corrupted by the motion artifacts.

3) RS Slop

The RS slop feature represents the descent speed of the ECG potential between the R peak and the S valley, and is calculated as follows:

$$RS\_Slop = (Y_S - Y_R)/(S - R) \quad (3)$$

where,  $R$  and  $S$  are the occurrence time of the R peak and S valley, respectively, and their corresponding ECG potentials are  $Y_R$  and  $Y_S$ , respectively. This feature leverages the possibility that even the S wave or the QR segment may be corrupted, the slop information should more or less be reserved since the R peak-to-S valley transition is often sharp and within a small time duration.

4) RS Difference

The voltage difference between S and R.

5) SM Difference

The voltage different between M and S, where M is the first peak on the right side of the S valley.

6) SM vs.RS

The voltage different between M and S divided by the voltage different between S and R.

7) R Symmetry

The R symmetry feature provides the symmetry information of the R wave, equaling to  $\theta_1/\theta_2$ , where,  $\theta_1$  and  $\theta_2$  are angles determined by three-point groups, M1-R-R' and R'-R-M2, respectively. M1 is the point with a maximum slop within the QR region, M2 is the point with the minimum slop within the RS region, and R' is the point on the vertical line going up through the R peak.

8) S Symmetry

The S symmetry feature provides the symmetry information of the S wave, equaling to  $\theta_3/\theta_4$ , where,  $\theta_3$  and  $\theta_4$  are angles determined by three-point groups, M2-S-S' and S'-S-M, respectively. M2 is the point with a minimum slop within the RS region, M is the first peak on the right side of the S valley, and S' is the point on the vertical line going down through the S valley.

These two features (feature 7 and 8) extract the symmetry information of the R wave and S wave, which, to some extent, may not be exactly mimicked by the interferential spikes.

9) Skewness

Skewness is used to measure the lack of symmetry of the distribution of the samples around the R peak region.

10) Kurtosis

Kurtosis is used to measure the distribution of samples around the R peak region.

11) Auto-regression Coefficient 1

Coefficient 1 of the normalized four order auto-regression. (Coefficient 0 is not used since it is a constant)

12) Auto-regression Coefficient 2

Coefficient 1 of the normalized four order auto-regression.

13) Auto-regression Coefficient 3

Coefficient 1 of the normalized four order auto-regression.

14) Auto-regression Coefficient 4

Coefficient 1 of the normalized four order auto-regression.

15) Median Frequency

The median frequency of the power spectrum of the R peak region.

16) Mean Frequency

The mean frequency of the power spectrum of the R peak region.

17) Variance

Signal variance of the sequence from M1 to M2.

18) Root Mean Square

Root mean square value calculated based on the sequence from M1 to M2.

19) Modified Median Frequency

The modified median frequency is the frequency which equally divides the amplitude of the spectrum of the R peak region.

#### 20) Modified Mean Frequency

The modified mean frequency is the frequency which averages amplitude of the spectrum of the R peak region.

#### 21) Slop Angle -3

The angle of the slop at the third sample on the left side of the R peak.

#### 22) Slop Angle -2

The angle of the slop at the second sample on the left side of the R peak.

#### 23) Slop Angle -1

The angle of the slop at the first sample on the left side of the R peak.

#### 24) Slop Angle 0

The angle of the slop at R peak.

#### 25) Slop Angle 1

The angle of the slop at the first sample on the right side of the R peak.

#### 26) Slop Angle 2

The angle of the slop at the second sample on the right side of the R peak.

Above features include both newly extracted ones and some used in other applications such as EMG signal processing [46, 47]. These features extracted from multiple domains will be further evaluated by the following step to select out the critical ones tolerant to motion artifacts.

### **3.3.5 Feature Selection**

There are different feature selection methods, such as the filter method and the wrapper method. The former one is independent of classification algorithms and the latter one usually achieve better performance since it is based on the classification performance [48]. We choose the wrapper

method to perform the feature selection. Because we choose the well-known SVM classifier to identify the heartbeats in the stage I of the proposed framework, we also choose SVM classifier to perform the feature selection.

We introduce a technique, called the absolute shrinkage and selection operator (LASSO) to efficiently evaluate the contribution of the features in SVM model training [49, 50]. The LASSO-based SVM is also named as an  $\ell_1$ -norm SVM which can quickly push insignificant feature weights towards zero, and thus yield a sparse representation of the features.

### **3.3.6 Classification**

A popular classifier, SVM, is used to identify the heartbeats and a linear kernel is chosen for its effectiveness in computation efficiency [51]. An  $\ell_2$ -norm SVM is chosen to train the learning model and perform the heartbeats identification.

The  $\ell_2$ -norm SVM is used considering that although an  $\ell_1$ -norm SVM can effectively push the features weights towards zero, it also usually biases the critical features. The  $\ell_2$ -norm SVM can suppress this bias and thus provides relatively better performance [49]. We will further illustrate this difference in results section.

### **3.3.7 Evaluation of Classification performance**

There are many cross validation methods to evaluate the performance of the classifier. For example [52], (a) intra-subject CV (intra-CV) performs x-fold cross validation for each subject's recording; (b) inter-subject CV (inter-CV) performs x-fold cross validation for the whole dataset; (c) leave-one-subject-out cross validation (LOO-CV) tests each subject's recording based on the model trained on all the remaining recordings in the dataset.

However, in real-world applications, the time series needs to be used sequentially. It means that we should use a small portion of the data at the beginning to train the model and then apply the trained model onto the future continuous data. Therefore, the traditional cross validation methods are not used here. Instead, we used the first 5-minute data of each recording to train the model for that subject, and then test the model on the remaining 25-minute data.

### 3.3.8 Outlier Removal

After the high confident heartbeats identification by the SVM classifier, there may still be some residual interferential spikes due to motion artifacts (false positives). So we introduce a rule-based classifier to remove these outliers.

Some definitions are given here to ease the read. Inputs of outlier removal are from the SVM classifier in the first stage of the framework, including: ‘R peaks’, recoded as  $\{R_i | 1 \leq i \leq N\}$ , and ‘R-R intervals’, recorded as  $\{RRI_i | 1 \leq i \leq N\}$ , where  $R_i$  is the occurrence time of the  $i$ -th ‘R peak’,  $RRI_i$  is the ‘R-R interval’ equaling to  $R_i - R_{i-1}$ , and  $N$  is the number of detected high confident ‘R peaks’ which include both true positives and false positives, identified by SVM with the ET method.

Outputs of outlier removal are the results of the rule-based classifier in the second stage of the framework, including: validity of ‘R peaks’, recoded as  $\{RV_i | 1 \leq i \leq N\}$ , and validity of ‘R-R intervals’, recorded as  $\{RRIV_i | 1 \leq i \leq N\}$ , where  $RV_i$  represents if the  $i$ -th ‘R peak’ is valid or not, and  $RRIV_i$  represents if the  $i$ -th ‘R-R interval’ is valid or not.

To remove the outliers, we defined two attributes to build the classifier. One attribute is called continuity check. Considering the heart rate variability is band-limited [53], we choose a pre-

defined heart rate variability range to check whether the R-R interval change is beyond this range. If yes, the heartbeat is suspected as the interferential spike. The allowed range is  $\pm 60\%$ , which is much more loose than previous works to provide a better generalization ability of the framework [23, 54]. To perform the continuity check for the current R-R interval, it is compared with the reference interval, which is the averaged interval based on previous five intervals to guarantee the robustness. Another attribute is called locality check. This is used to validate the heartbeats when continuous continuity checks fail due to severe motion artifacts and the reference interval cannot reflect the current status of the heart rate. In this case, the consistency of the heartbeats in the upcoming data can still be discovered and be used to validate the heartbeats. To perform the locality check, five local continuity checks for the upcoming five intervals are performed. The local continuity check just checks if one interval is in the range defined by the previous interval. If five upcoming local continuity checks hold, the current interval passes the locality check and is validated even its continuity check fails. The interferential spikes basically cannot make several consecutive local continuity checks hold due to high randomness.

The rule-based classifier is defined by the following four rules, where,  $L_i$  is the locality check result for the current heartbeat,  $C_i$  is the continuity check result for the current heartbeat, and  $RRIV_i$  and  $RV_i$  represent if the current R-R interval and R peak are valid or not.

- 1)  $RL1: (L_i = 1) \rightarrow RRIV_i = yes, RV_i = yes$
- 2)  $RL2: (C_i = 1) \wedge (C_{i+1} = 1) \rightarrow RRIV_i = yes, RV_i = yes$
- 3)  $RL3: (C_i = 0) \wedge (L_i = 0) \rightarrow RRIV_i = no, RV_i = no$
- 4)  $RL4: (L_i = 0) \wedge (C_{i+1} = 0) \rightarrow RRIV_i = no, RV_i = no$

This rule-based classifier is highly scalable. New rules can be easily added depending on the application scenarios.

### 3.3.9 Heartbeat Interpolation and Heart Rate Estimation

As mentioned before, in auto-segmentation, QRS complexes are, wherever possible, reserved by an adaptive threshold-based method, meanwhile, a large amount of interferential spikes are also induced. Afterwards, most of these interferential spikes are removed by the heartbeat identification classifier (SVM), and then the remaining ones are further eliminated by the outlier suppression technique (the rule-based classifier). However, there are also some false negatives (missing heartbeats) resulting in sudden drop of instantaneous heart rate. To deal with them while also keeping power efficiency in mind for wearable solutions, a simple method is proposed here.

We firstly try to use the arithmetic progression method to interpolate the missing heartbeats. This method is expected to track the increasing or decreasing trends of the heart rate curve. When it fails due to continuous severe motion artifacts, we will use a simpler averaging method.

The final set of R peaks  $Rset_{all}$  are given as follows, which includes a set of all valid ‘R peaks’  $Rset_{valid}$  and a set of all interpolated ones  $Rset_{IP}$ . The sort operation rearranges all the peaks according to their occurrence time and  $R_i^{sort}$  is the  $i - th$  beat in the new set  $Rset_{all}$  of a total number of beats  $M$ .

$$\begin{aligned} Rset_{all} &= \{R_i^{sort} \mid R_i^{sort} \in sort(Rset_{valid} \cup Rset_{IP})\}, \\ |Rset_{all}| &= M \end{aligned} \quad (4)$$

Based on the heartbeats set, the heart rate  $HR_i$  is estimated as follows, where  $SR$  is the sampling rate, the unit of  $HR_i$  is beats per minute (BPM) and  $R_0^{sort}$  is 0. Considering the interpolated heartbeats may not form pairs with the gold standard heartbeats, a time window-based heart rate calculation method is used here to ease the performance evaluation. The window length is set to be two seconds to guarantee the existence of at least one heartbeat, since the lowest heart rate is usually about 40 BPM [55]. This window length can represent the real-time heart rate to the utmost possibility. The 2-second window-based heart rate is defined as  $HR_k^{2s}$ , where  $K$  is the number of windows in one ECG recording ( $K = 750$  since the the test data is 25-minute long),  $W_k$  is the end time of the  $k - th$  window ( $W_0 = 0$ ),  $j_k^{start}$  and  $j_k^{end}$  are the indices of the first and last R peaks within this window, respectively.

$$HR_i = 60 SR / (R_i^{sort} - R_{i-1}^{sort}), \quad \forall i \in [1, M] \quad (5)$$

$$HR_k^{2s} = \sum_{j=j_k^{start}}^{j_k^{end}} HR_j / (j_k^{end} - j_k^{start} + 1),$$

$$s.t. \quad j_k^{start} = \arg \min_{R_i^{sort} > W_{k-1}} i, \quad (6)$$

$$j_k^{end} = \arg \max_{R_i^{sort} \leq W_k} i, \quad \forall k \in [1, K]$$

## 3.4 RESULTS

### 3.4.1 Auto-segmentation

The raw ECG signal with motion artifacts is processed by the auto-segmentation step, which selects out heartbeat candidates including both real heartbeats and interferential spikes. The

segmentation parameters are determined empirically. The threshold  $Th_0$  is equal to 70% of the peak-to-peak amplitude of the ECG stream before corrupted by motion artifacts. The factor  $Coef$  is 0.3 and the window size chosen to update the threshold is 20 seconds. In Figure 3.3, the green dots indicate that most of the real heartbeats are successfully selected out by the segmentation approach. Meanwhile, there are many spikes due to motion artifacts are also selected out which will be suppressed by the heartbeat identification and purification steps. The adaptive threshold is also shown to be able to effectively track the signal fluctuations, and wherever possible, help to select out real heartbeats.

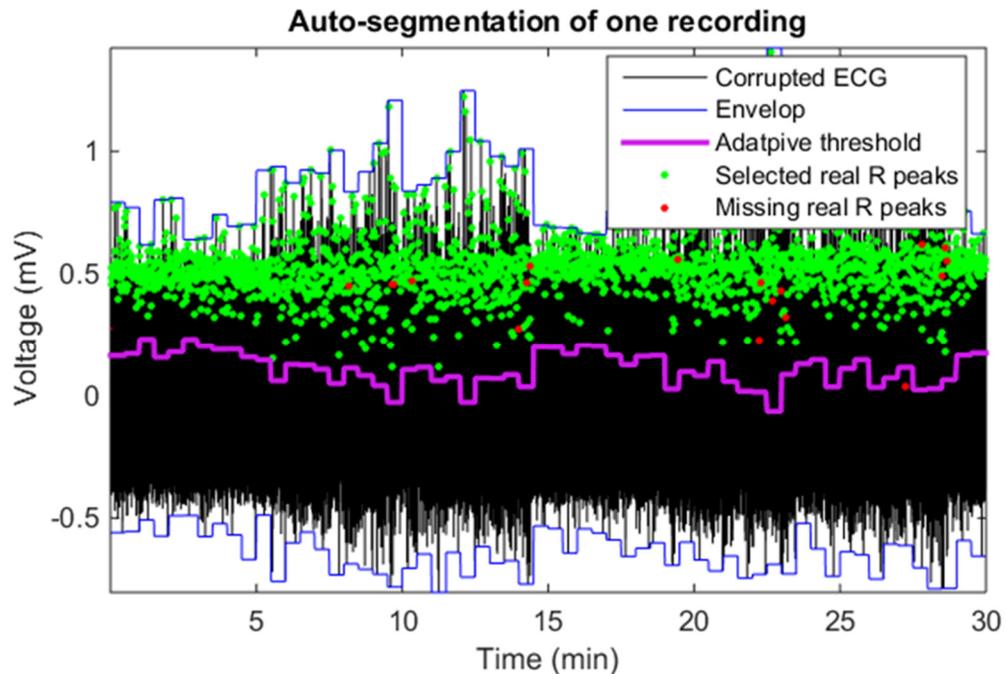


Figure 3.3. An example of the auto-segmentation result.

Notes. The data is from recording 1 in WECG\_Sm7\_SH0. Most of real heartbeats are selected out, and many interferential spikes are also included (not labelled in the figure for a good readability).

### 3.4.2 Feature Engineering

The sparse SVM is used to evaluate the significance of the extracted features. The feature weights learned from the training data of dataset WECG\_Sm7\_SH0 using LIBLINEAR [56] are given in Table 3.2, where the absolute weights are firstly averaged over 22 recordings and then normalized to be between 0 and 1. From the figure we can easily distinguish large feature weights and also small ones close to zero. Large weights mean that corresponding features have bigger contribution and small weights can be removed from the feature set due to their small contribution. We choose the top ten features with the underlines in Table 3.2 as the final feature subset.

In Table 3.3, three SVM models are also compared to show that the  $\ell_2$ -norm SVM using LIBSVM [57] based on ten selected features is slightly better than the  $\ell_1$ -norm SVM based on ten selected features. As mentioned above, this is because the  $\ell_2$ -norm SVM can usually reduce the weights bias compared with the  $\ell_1$ -norm SVM.

### 3.4.3 Heartbeat Identification

Based on the selected critical features, the SVM classifier is used to identify the high confident heartbeats. Afterwards, in stage II of the framework, the refinement engine further suppresses the outliers and interpolates the missing heartbeats. An example of the identified heartbeats is given in Figure 3.4, which includes the results of auto-segmentation, SVM classification, rule-based classifier and heartbeat interpolation. The example shows that the corrupted heartbeats can be robustly identified by the SVM classifier, and remaining interferential spikes can also be effectively suppressed by the rule-based classifier. Missing heartbeats can also be recovered by the interpolation operations.

Table 3.2. Learned feature weights on the training data.

Feature No.	Name	Weight
<u>1</u>	<u>R Angle</u>	<u>1.0</u>
<u>2</u>	<u>S Angle</u>	<u>0.5</u>
3	RS Slop	0.0
<u>4</u>	<u>RS Difference</u>	<u>0.4</u>
5	SM Difference	0.2
6	SM vs. RS	0.0
<u>7</u>	<u>R Symmetry</u>	<u>0.5</u>
<u>8</u>	<u>S Symmetry</u>	<u>0.9</u>
<u>9</u>	<u>Skewness</u>	<u>0.3</u>
10	Kurtosis	0.1
11	Auto-regression Coefficient 1	0.0
12	Auto-regression Coefficient 2	0.0
13	Auto-regression Coefficient 3	0.1
14	Auto-regression Coefficient 4	0.1
15	Median Frequency	0.2
16	Mean Frequency	0.2
<u>17</u>	<u>Variance</u>	<u>0.8</u>
<u>18</u>	<u>Root Mean Square</u>	<u>0.6</u>
19	Modified Median Frequency	0.2
20	Modified Mean Frequency	0.2
<u>21</u>	<u>Slop Angle -3</u>	<u>0.5</u>
22	Slop Angle -2	0.3
23	Slop Angle -1	0.2
24	Slop Angle 0	0.2
25	Slop Angle 1	0.1
<u>26</u>	<u>Slop Angle 2</u>	<u>0.3</u>

Notes. The training data is from the first 5-minute data from recordings in WECG\_Sm7\_SH0.

Table 3.3. Model comparison in terms of accuracy (%) on the testing data.

Recording No.	L1-SVM@26	L1-SVM@10	L2-SVM@10
	features	features	features
1	98.6	98.6	99.8
2	97.3	97.1	99.2
3	98.2	97.4	98.0
4	99.0	98.7	99.2
5	99.7	99.7	99.8
6	99.4	99.1	99.5
7	94.8	96.1	96.5
8	99.1	98.0	99.1
9	98.2	99.3	99.9
10	100.0	100.0	100.0
11	95.3	95.1	96.1
12	97.4	97.0	98.9
13	99.1	99.6	99.9
14	99.7	99.7	99.7
15	99.5	99.5	99.8
16	99.3	99.3	99.3
17	97.1	97.6	98.2
18	95.4	95.8	96.1
19	97.5	97.2	97.0
20	99.5	99.0	99.6
21	98.8	98.7	99.7
22	96.8	95.4	97.5

Notes. The testing data is from the 25-minute data from recordings in WECG\_Sm7\_SH0.

Moreover, the sensitivity and precision of the SVM classifier and the rule-based classifier are also given in Figure 3.5 and Figure 3.6, based on the WECG-based datasets. The results show that the sensitivity has been decreased after the rule-based classification operation, however, the precision has been improved. Here, we have sacrificed some sensitivity to get a better precision when suppressing the residual interferential spikes. This tradeoff is necessary since the residual interferential spikes usually cause sharply increased heart rate estimates. It is also worth noting that the missing heartbeats can be recovered by the interpolation operations.

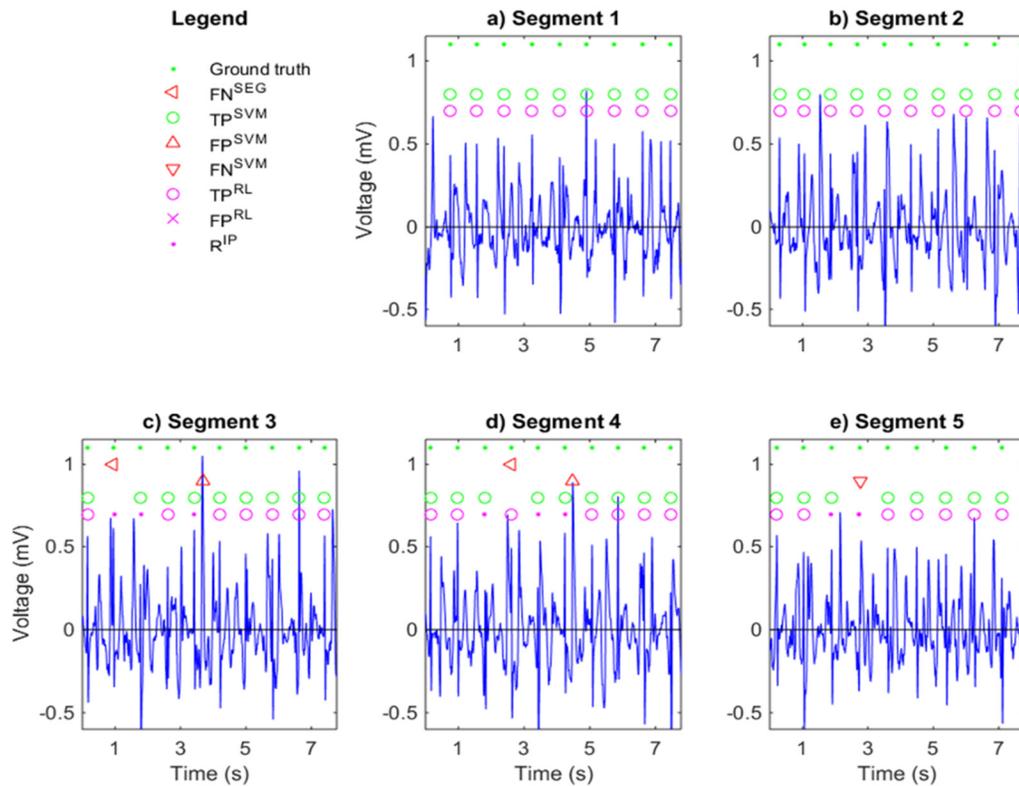


Figure 3.4. An example of the identified heartbeats.

Notes. The data is from recording 1 in WECG\_Sm7\_SH0. FN: false negative; TP: true positive; FP: false positive; SEG: auto-segmentation; SVM: support vector machine; RL: rule-based classifier; IP: interpolation.

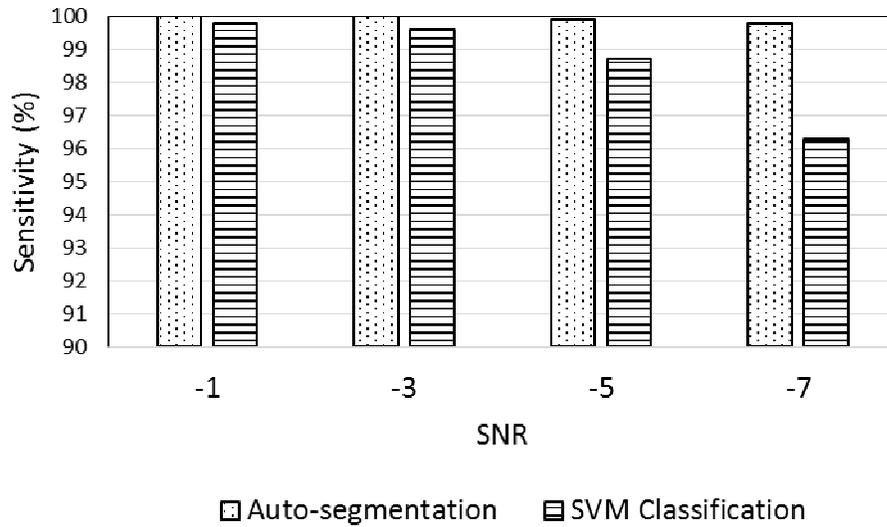


Figure 3.5. Heartbeat recognition performance in terms of sensitivity on WECG-based datasets. Notes. The results have been averaged over five different phase shifts.

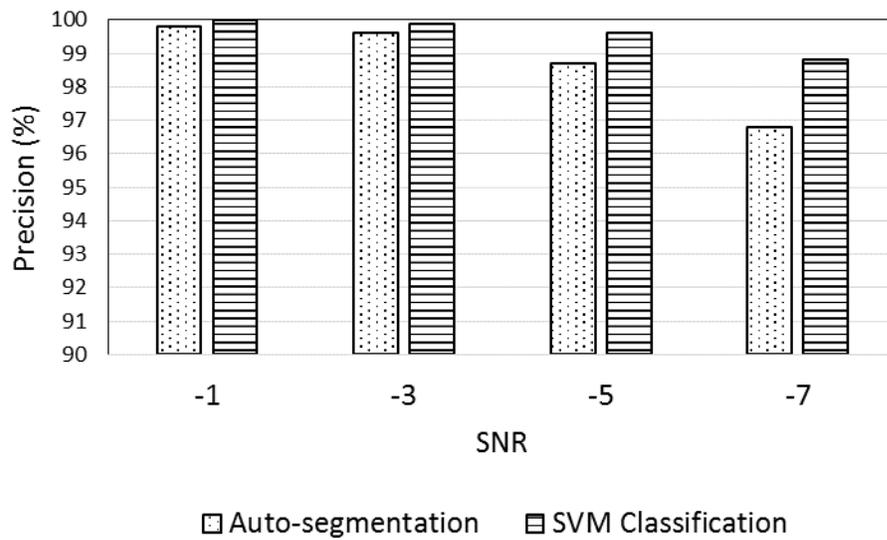


Figure 3.6. Heartbeat recognition performance in terms of precision on WECG-based datasets. Notes. The results have been averaged over five different phase shifts.

### 3.4.4 Heart Rate Estimation

The final heart rate estimates can be achieved based on both validated heartbeats by the rule-based classifier and the interpolated heartbeats. The results on the WECG-based datasets are given in Figure 3.7, where the MAE and RMSE are both given. With an SNR of -7 dB, they are 1.4 and 6.5 BPM, respectively, indicating the effectiveness of the proposed framework even with a low SNR.

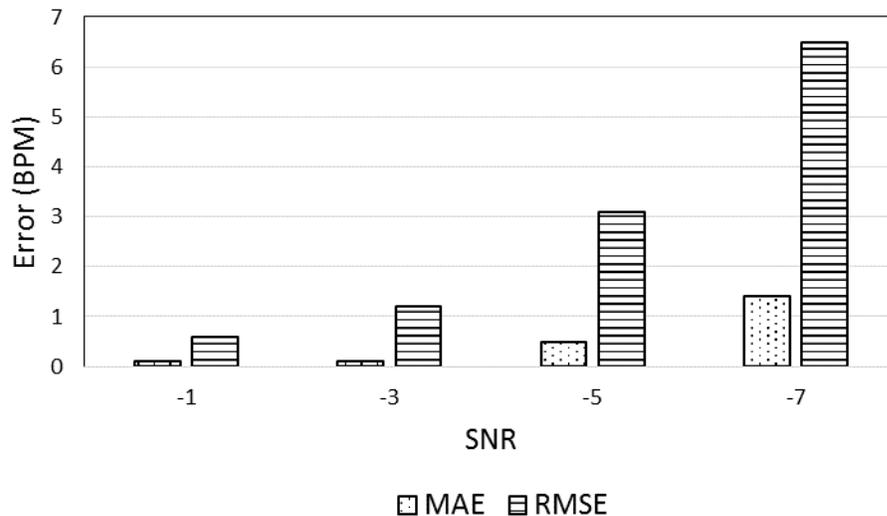


Figure 3.7. Heart rate estimation performance in terms of MAE and RMSE on WECG-based datasets.

Notes. The results have been averaged over five different phase shifts; RMSE: root mean square error; unit of MAE and RMSE: beats per minute (BPM).

### 3.4.5 Performance Comparison

To show the effectiveness of the proposed framework, we also compare the two-stage classification approach with other approaches.

In Figure 3.8, the heart rate estimation performance comparison in terms of MAE based on the WECG dataset is given. The approaches taken into account include Kalman filtering (KLMF), discrete wavelet transformation (DWT), weighted regularized least square (WRLS), time-domain dynamic time warping (TD-DTW), phase-domain multi-view dynamic time warping (PD-MDTW) and the proposed machine learning-enabled framework (MLEF). The comparison indicates that MLEF greatly outperforms KLMF, DWT, WRLS, and TD-DTW especially with a low signal quality, benefitting from the motion artifacts-tolerant critical features and advanced machine learning classifier. Besides, MLEF has similar performance with PD-MDTW.

In Figure 3.9, the heart rate estimation performance comparison in terms of RMSE is also given and owns a similar trend as Figure 3.8.

In Figure 3.10 and Figure 3.11, the heart rate estimation performance comparison on the FECG-based datasets is also given in terms of MAE and RMSE, respectively. They both share the similar trend as Figure 3.8 and Figure 3.9.

The consistent performance trend in these four figures indicates that the proposed MLEF framework owns superior performance compared with KLMF, WRLS, DWT and TD-DTW. Moreover, the MLEF owns similar performance with PD-MDTW, but MLEF only extracts features from signal spikes of interest and thus consumes less computation load. The simulation using MATLAB on Intel CORE i7 shows that MLEF only consumes about 10% of the computation time compared with PD-MDTW. The effectiveness of the MLEF approach mainly benefits from the selected critical features tolerant to motion artifacts, the SVM classifier and the refinement engine to purify the heartbeats results, even with a low signal quality. Specifically, with an SNR of -7 dB, the MAE and RMSE are no more than 1.4 and 6.5 BPM, respectively.

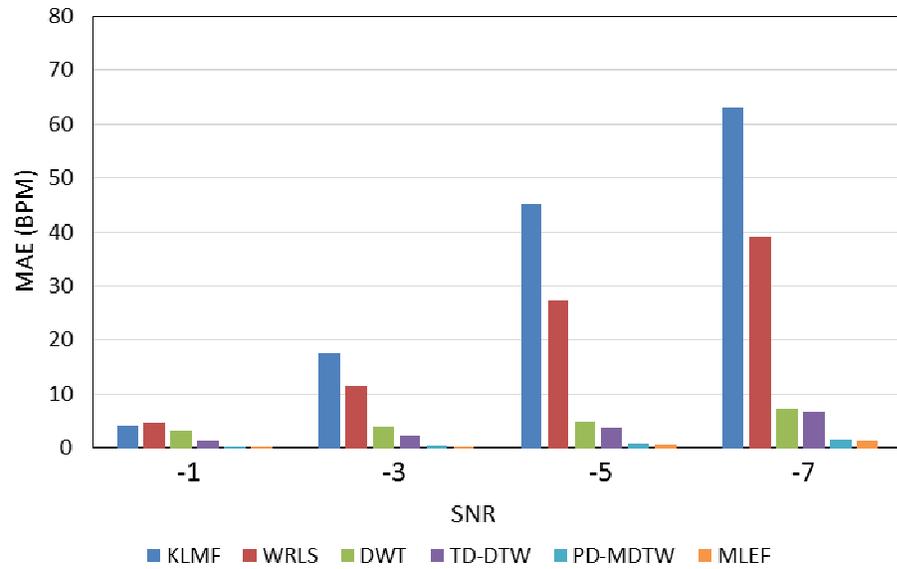


Figure 3.8. Heart rate estimation performance comparison in terms of MAE on WECG-based datasets.

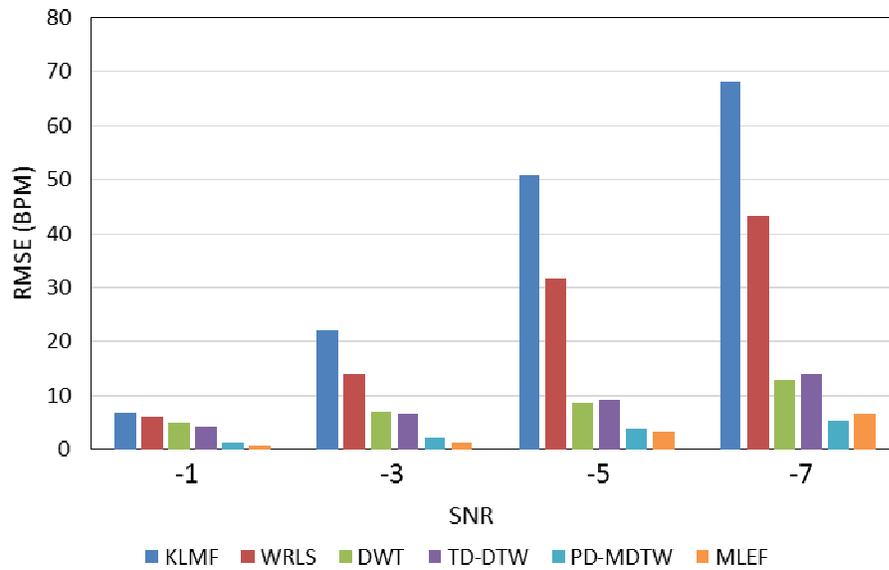


Figure 3.9. Heart rate estimation performance comparison in terms of RMSE on WECG-based datasets.

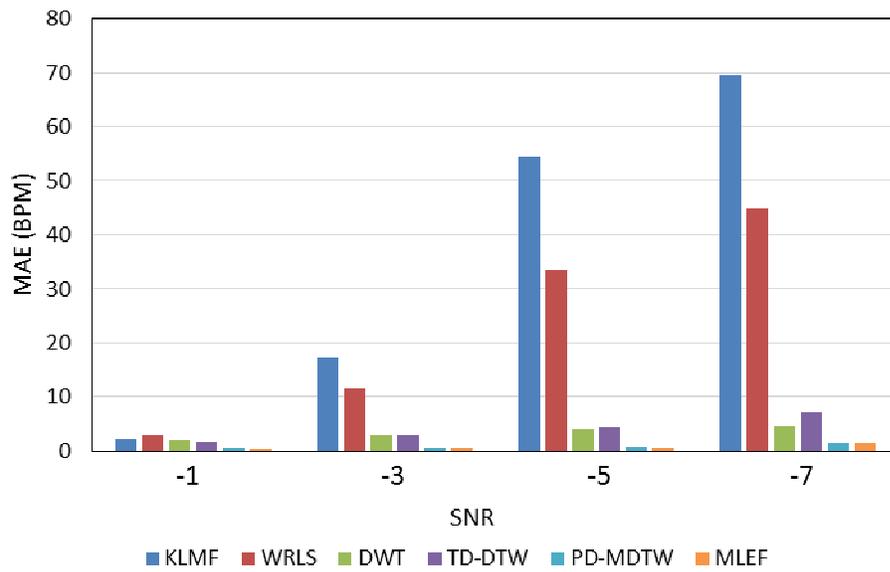


Figure 3.10. Heart rate estimation performance comparison in terms of MAE on FECG-based datasets.

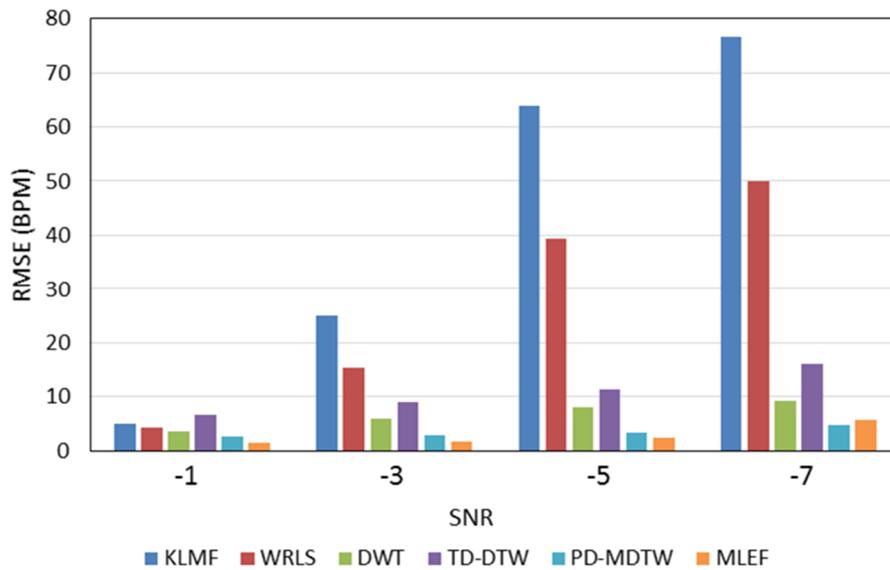


Figure 3.11. Heart rate estimation performance comparison in terms of RMSE on FECG-based datasets.

### 3.5 CONCLUSION

A two-stage classification framework is proposed which can effectively identify high confident heartbeats from signal spikes of interest and then purify the heartbeats results for robust heart rate estimation.

Our contributions are summarized as follows:

- 1) Propose a two-stage framework for real-time one-channel ECG-based instantaneous heart rate monitoring with the motion artifact-tolerant ability.
- 2) Among twenty-six features extracted, finally determine ten critical features tolerant to motion artifacts, which are fed to the SVM classifier for high confident heartbeats recognition.
- 3) Present a refinement engine which includes a rule-based classifier for outlier removal and a heartbeat interpolator.
- 4) Validate the framework under real-world scenarios: each of twenty-four 30-minute recordings are continuously corrupted by motion artifacts with different SNRs, and with multiple phase shifts for critical randomness.

The proposed framework is evaluated on many datasets with different signal quality conditions. Specifically, with an SNR of -7 dB, the heart rate estimation error is no more than 1.4 BPM in terms of MAE and 6.5 BPM in terms of RMSE, indicating the effectiveness of the two-stage-classification framework.

**CHAPTER 4**  
**BLOOD PRESSURE MONITORING WITH SINGLE-ARM SIGNALS**

Authors: Qingxue Zhang, Dian Zhou, Xuan Zeng

The Department of Electrical Engineering, EC33

The University of Texas at Dallas

800 West Campbell Road

Richardson, Texas 75080-3021

© 2017 Biomed Central. Reprinted, with permission, from Q. Zhang, D. Zhou and X. Zeng, “Highly Wearable Cuff-less Blood Pressure and Heart Rate Monitoring with Single-arm-ECG and PPG Signals,” Biomedical Engineering Online, vol. 16, p.23, 2017.

## ABSTRACT

Background: Long-term continuous systolic blood pressure (SBP) and heart rate (HR) monitors are of tremendous value to medical (cardiovascular, circulatory and cerebrovascular management), wellness (emotional and stress tracking) and fitness (performance monitoring) applications, but face several major impediments, such as poor wearability, lack of widely accepted robust SBP models and insufficient proofing of the generalization ability of calibrated models. Methods: Focusing challenges mentioned above, we propose a wearable cuff-less electrocardiography (ECG) and photoplethysmogram (PPG)-based SBP and HR monitoring system. Both ECG/PPG sensors are integrated into a single-arm band to provide a super wearability. A highly convenient but challenging single-lead configuration is proposed for weak single-arm-ECG acquisition, instead of placing the electrodes on the chest, or two wrists. To identify heartbeats and estimate HR from the motion artifacts-sensitive weak arm-ECG, a machine learning-enabled framework is applied. Then ECG-PPG heartbeat pairs are determined for pulse transit time (PTT) measurement. Thirdly, a PTT&HR-SBP model is applied for SBP estimation, which is also compared with many PTT-SBP models to demonstrate the necessity to introduce HR information in model establishment. Fourthly, the fitted SBP models are further evaluated on the unseen data to illustrate the generalization ability. Results: A customized hardware prototype was established and a dataset collected from ten volunteers was acquired to evaluate the proof-of-concept system. Conclusions: The proposed proof-of-concept system is highly wearable, and its robustness is thoroughly evaluated on different modeling strategies and also the unseen data, which are expected to contribute to long-term pervasive hypertension, heart health and fitness management.

## 4.1 INTRODUCTION

Blood pressure (BP) is a key health indicator to diagnose and control hypertension, which impacts over 35 percent of people worldwide, relates to cardiovascular, circulatory and cerebrovascular diseases, and causes 12.8 percent of the total of all deaths [18]. Since BP fluctuates over time, BP should not be measured only at specified times and circumstances. However, the traditional BP measurement approaches are unsuitable for long-term ubiquitous applications, such as the invasive catheterization method and the noninvasive cuff-based oscillometry method [58], which are both time consuming and of a poor wearability.

Nowadays, wearable computers are paving a promising way for ubiquitous BP monitoring by providing convenient and long-term out-of-clinic measurements. Wearable cuff-less BP monitors are usually created leveraging mounting evidence that the pulse transit time (PTT) is reversely related to the BP [58]. To measure the PTT, i.e., the time delay for the pressure wave to propagate between two arterial sites, one popular method is based on two signals, i.e., electrocardiography (ECG) and photoplethysmogram (PPG). The former one is the electrical signal generated by the heart, while the latter one measures fluctuations in the blood volume which are caused by the mechanical pressure pulse and thus changes later than the electrical ECG wave. Currently, the ECG signal is usually measured using a single- or multiple-lead configuration, referring to or modified from the traditional standard 12-lead configuration which can provide strong ECG signals with highly distinguishable morphologies [58, 59]. Nevertheless, these electrodes placement methods may have some limitations in long-term applications, e.g., the chest electrodes placement may be uncomfortable especially when sweating, and the two wrists configuration may be still inconvenient since additional wires or separate devices are inevitable. Likewise, the PPG

sensor is usually placed on the chest which may be uncomfortable, or on the finger where more challenges may be posed to the integration of PPG and ECG sensors [60].

In this study, we propose a single-arm blood pressure monitoring system, which allows for placing the PPG sensor and the ECG electrodes all on the left upper arm, to enable long-term daily applications which have critical requirements on the wearability and comfortableness. Since we put both ECG signal and reference electrodes on the left upper arm which form a non-standard single-lead configuration for super wearability, the potential difference between these two close electrodes due to the heart electric propagation is so small that it is highly challenging to obtain a clear single-arm-ECG signal.

By creating a customized hardware prototype and placing the reference electrode on the top side of the left upper arm and the signal electrode on the bottom side to maximize the distance between these two electrodes, the weak single-arm-ECG signal is successfully acquired, which owns an amplitude much lower than those measured by the standard or modified traditional lead configurations mentioned above, but has a morphology still distinguishable. The single-arm-PPG signal is also acquired by placing the sensor close to the ECG electrodes for a good wearability. Afterwards, to recognize the heartbeats from the weak ECG signal with many interferential spikes induced by motion artifacts and electromyography (EMG) noise, a machine learning-enabled framework is introduced [31]. Based on the identified heartbeats in the ECG signal, the heartbeat pairs in the PPG signal are then determined to obtain the PTT measurements which will be used to build the systolic BP (SBP) model (we take special interest in SBP monitoring in this study). Meanwhile, the heart rate (HR) information is also achieved from the identified ECG heartbeats, which can not only be used in heart health and fitness management, but also contribute to SBP

model enhancement by the correlation between HR and BP [59]. The arm-worn HR monitor is also highly wearable.

Considering both PTT and HR are correlated with SBP, we choose the PTT&HR-based SBP model for SBP estimation. Meanwhile, we have also performed a thorough comparative analysis among PTT&HR-SBP models and PTT-SBP models, to further illustrate the necessity to introduce the HR information in SBP model enhancement. In total, three PTT&HR-SBP models and seven PTT-SBP models are taken into account in this study.

These models are firstly tuned using the training data, and then evaluated on the unseen fresh testing data to show the generalization ability of the tuned models. Therefore, the algorithm had been set before the testing performance evaluation stage and not changed during evaluation.

Moreover, the SBP estimates based on the chest-ECG/arm-PPG signals are also obtained to show the feasibility to replace the strong but inconvenient chest-ECG with the weak arm-ECG, to enable the highly wearable SBP monitoring.

Besides, the participants were asked to perform exercise during some signal periods in data collection to introduce more stress to the signal quality. The exercise stress can not only perturb the SBP to a larger range to increase the diversity, but also introduce more motion artifacts and heart rate variability to the weak arm signals towards practical applications.

Experimental results show that the HR can be robustly estimated from the weak single-arm-ECG signal, and the PTT&HR-SBP with HR enhanced significantly outperform the PTT-SBP models and can be well generalized to the unseen data. Therefore, the single-arm-ECG signal can be a highly convenient and effective alternative to the chest-ECG signal, to enable robust long-term SBP monitoring applications together with the single-arm-PPG signal.

## 4.2 METHODS

### 4.2.1 System Overview

The proposed wearable cuff-less blood pressure and heart rate monitoring system is illustrated as Figure 4.1, where the top part (Figure 4.1a) shows the customized hardware platform for single-arm-ECG and PPG signals acquisition, and the bottom part (Figure 4.1b) gives the flow of the HR and SBP estimation algorithms.

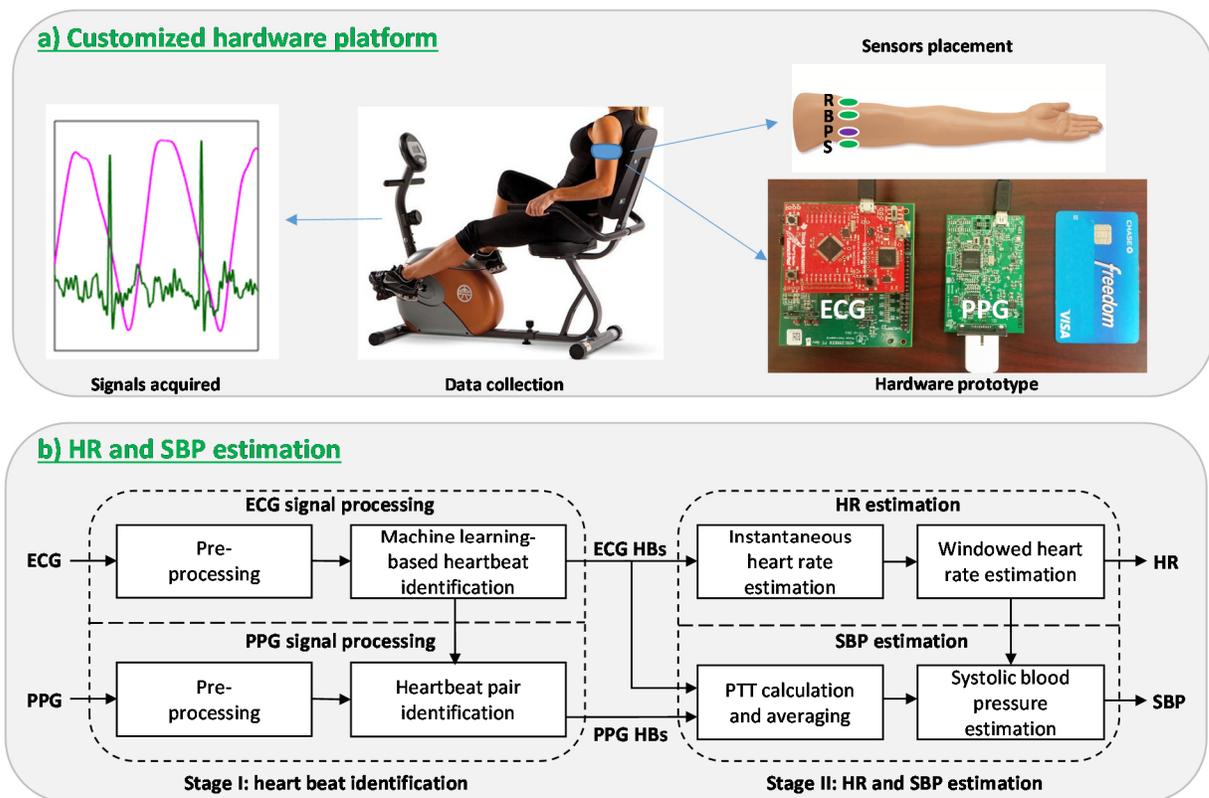


Figure 4.1. The proposed system for wearable cuff-less SBP and HR monitoring with single-arm-ECG and PPG signals.

Notes. R/B/S represent the reference/bias/signal electrodes used for single-lead ECG signal measurement, respectively; P corresponds to the reflective PPG sensor; PTT: pulse transit time; HB: heartbeat; HR: heart rate; SBP: systolic blood pressure.

### 4.2.2 Customized Hardware Platform

The customized hardware platform comprises two parts, i.e., the ECG signal [31] and PPG signal acquisition subsystems, as shown in the right part of Figure 4.1a. The former one includes a TI ADS1299EEG-FE evaluation board (blue one) which is equipped with an ADS1299 24-bit analog-to-digital converter (ADC) for low voltage bio-potential measurement, and a TI Tiva™ C series LaunchPad [30] (red one) which includes an ARM Cortex M4 microcontroller (MCU) to send commands to the ADC, read the measurements from the ADC via the SPI port and give the data to a PC via the USB port. The latter one includes a TI AFE4490SPO2 evaluation board [61] (blue one) which is equipped with an LED transmit section to generate the red or infrared light to illuminate the skin, and a low-noise receiver channel with a 22-bit ADC to measure the time varying light absorption by the tissue to reflect the changes in the blood volume. There is an MSP430F5529IPN MCU embedded on this board to configure the ADCs, fetch the data from the receiver ADC via the SPI port, and send the data to the PC by the USB port. After removing the components only for evaluation purposes and adding a wireless communication module, the proposed prototype can be conveniently used in long-term wearable applications.

The ECG and PPG sensors placement on the left upper arm is illustrated in the right part of Figure 4.1a, where the circles labeled as R/B/S represent the reference/bias/signal electrodes used for single-lead ECG signal measurement, respectively, and the circle labeled as P corresponds to the reflective PPG sensor [62] with the LEDs and photodiode embedded for PPG signal acquisition.

The proposed sensors placement method is highly convenient, since it prevents attaching the ECG electrodes to the chest, or to multiple separate body sites such as two wrists plus one finger

[60]. Moreover, the ECG electrodes and the PPG sensor can be integrated into one single arm band, to further enhance the wearability in long-term daily applications.

### **4.2.3 Data Acquisition Protocol**

The customized platform was used for single-arm-ECG and PPG signals acquisition, with a sampling rate of 500 and 128 Hz, respectively. The chest-ECG signal is collected at the same time for comparison purpose. A higher sampling rate for ECG is based on the consideration that it is used not only for PTT but also for HR estimation. The data collection was performed on ten participants (age:  $35 \pm 14$ ; Weight:  $68 \pm 13$  Kg; height:  $168 \pm 7$  cm; Gender: 7 males and 3 females) to demonstrate our proof-of-concept system. Informed consent was obtained from all individual participants included in the study. For each subject, the data included two sessions collected on the same day using the same data acquisition protocol, i.e., a 26-minute training session used to train the algorithms, such as the heartbeat identification classifier and the SBP models, and another 26-minute testing session, to evaluate the generalization ability to the unseen data of the trained algorithms. Each subject was asked to sit on an IMPEX MARCY ME-709 recumbent exercise bike [63] with armrests. Each session took 26 minutes, including 13 2-minute trials belonging to three parts, i.e., part I (trial 1), part II (trial 2 to 11) and part III (trial 12 to 13). During part I and III, the subject stayed still, and during each trial in part II, the subject rode the bike in the first minute and stayed still in the second minute. The exercise was introduced to perturb the SBP to a larger range referring to protocol used in [64], such that SBP model can be trained and tested both over a larger range of SBP.

Both cuff-based SBP measurement and ECG/PPG-based SBP estimation were performed in the same time duration, i.e., the second minute of each trial when subjects stayed still. We used an

ambulatory blood pressure monitor CONTEC ABPM50 [65] to measure the reference SBP, which consumes about one minute to report one measurement result. Correspondingly, we used the ECG/PPG signals in the second minute of each trial for averaged PTT and HR estimation, which were then used in the SBP model training and testing. In this way, we can guarantee the synchronization between the reference and the estimates. Considering that the reference SBP was measured when the subjects stayed still and put their forearms on the armrests of the exercise bike, the reference SBP can be robustly measured by the ambulatory blood pressure monitor. Therefore, all the data has been used in our analysis.

One thing worth noting is that the exercise stress can also introduce more motion artifacts and heart rate variability to the weak arm signals, to take into account more affecting factors in practical application scenarios.

#### **4.2.4 Signal Pre-processing**

As shown in the left part of Figure 4.1b, the acquired raw ECG signal is pre-processed by a Butterworth bandpass filter with a lower and upper cutoff frequencies of 2 and 30 Hz, respectively, to remove the baseline wander and powerline interference and suppress the motion artifacts. The raw PPG signal is processed by a Butterworth filter with cutoff frequencies of 0.5 and 8 Hz, respectively, and then it is resampled to 500 Hz to own a same time resolution as ECG.

In Figure 4.2, an example of the filtered single-arm-ECG and PPG signals is given, which shows ECG and PPG pulses. Besides, it is found that the exercise usually more or less affects the signal quality and distorts the morphologies, which will be further analyzed in the results section to show why advanced methods are necessary to process the weak arm signals.

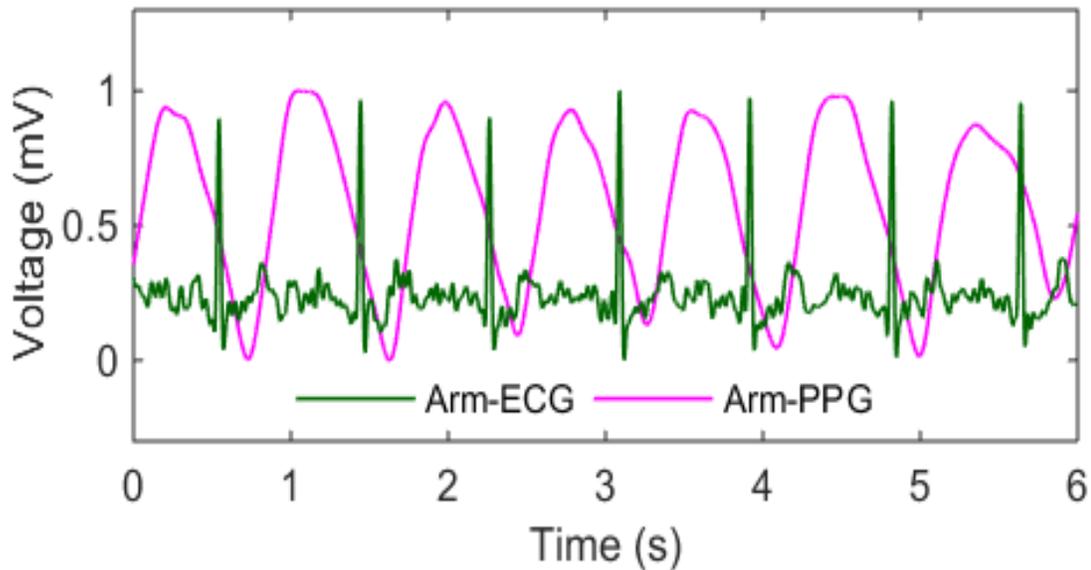


Figure 4.2. Arm-ECG/PPG signals, with the amplitude both scaled to be between 0 and 1 for good readability.

Notes. Arm-ECG signal is actually much weaker than arm-PPG, which will be further analyzed later.

#### 4.2.5 ECG-based Heartbeat Identification

ECG-based heartbeat locations are used in both PTT calculation and HR estimation as shown in Figure 4.1b. To robustly identify heartbeats from the weak arm-ECG signal, our previously reported machine learning-enabled framework (MLEF) is applied [31], which can effectively identify corrupted heartbeats and robustly estimate the heart rate from the wrist-ECG signals based on the support vector machine (SVM) classifier, even with an SNR as low as -7 dB. This heartbeat identification framework includes the following four steps:

Step 1: ECG stream auto-segmentation for heartbeat candidate generation.

Step 2: Feature extraction for each heartbeat candidate.

Step 3: SVM model training on the training data; or SVM model testing on the fresh testing data.

Step 4: Get identified high confident heartbeats.

In this first step, an adaptive threshold is generated based on the time-varying fluctuation of the signal. When there is a larger peak-to-peak voltage in a time window (20 seconds) due to motion artifacts, the vertical fluctuation of the real heartbeats is also increased, and vice versa. Therefore, we introduce an extra item to adaptively adjust a pre-defined fixed threshold to track the signal fluctuation, such that wherever possible, the real heartbeats can be selected as the heartbeat candidates, to guarantee a high sensitivity. Meanwhile, many motion artifacts-induced interferential spikes are also selected, resulting in a low precision, therefore, the next steps will further identify high confident heartbeats from the heartbeat candidates.

In the second step, ten critical motion artifacts-tolerant features are extracted from multiple domains for each candidate, include *R angle* (angle of the R peak), *S angle* (angle of the S valley), *RS Diff* (voltage difference between the R peak and the S valley), *R Symmetry* (the symmetry of the R peak), *S Symmetry* (the symmetry of the S valley), *SKNS* (skewness of the R peak region), *VAR* (variance of the R peak region), *RMS* (root mean square of the R peak region), *alpha -3* (angle of the slop of the third sample on the left side of the R peak), and *alpha 2* (angle of the slop of the second sample on the right side of the R peak). These features are selected from twenty-six raw features by a sparse SVM which can effectively push the non-significant features towards zero [31].

In the third step, an SVM model is trained firstly on the training data, and then tested on the fresh testing data for heartbeat identification. The SVM can constructs a hyperplane to effectively classify the instances into different groups. To train an SVM model, a constrained quadratic

optimization problem is solved. The objective function is composed of two parts, i.e., the regularization part ( $\frac{1}{2} \|w\|^2$ ) and the loss caused by misclassified instances ( $C \sum_{i=1}^M \xi_i$ ), as shown in (1), where  $w$  is a weight vector to be sought,  $C$  is a tradeoff parameter between maximization the separation margin and the minimization of the classification error ( $C$  is chosen as 1 as suggested by [57]), and  $\xi_i$  is the nonnegative slack variables to penalize the misclassified instances (1 to  $M$ ). There are two constraints shown as (2-3), where  $y_i$  is the class label of the instance  $x_i$ ,  $\Phi(x_i)$  is the kernel function to transform the instance  $x_i$  to the kernel space, and  $b$  is the bias to be sought. A linear kernel is chosen to lower the computation load for wearable applications.

$$\min_{w,b} \frac{1}{2} \|w\|^2 + C \sum_{i=1}^M \xi_i \quad (1)$$

$$s. t. \quad y_i(w^T \cdot \Phi(x_i) + b) \geq 1 - \xi_i, \quad \forall x_i \quad (2)$$

$$\xi_i \geq 0 \quad (3)$$

After introducing the Lagrange multipliers  $\alpha_i$ , we now have the dual problem as shown in (4-6), where  $K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j)$  representing the inner production operation between two instances in the transformed space. This dual problem can then be solved by a sequential minimal optimization method [57].

$$\max_{\alpha_i} \sum_{i=1}^M \alpha_i - \frac{1}{2} \sum_{i=1}^M \sum_{j=1}^M \alpha_i \alpha_j y_i y_j K(x_i, x_j) \quad (4)$$

$$s. t. \quad \sum_{i=1}^M \alpha_i y_i = 0 \quad (5)$$

$$C \geq \alpha_i \geq 0, \quad \forall i = 1, \dots, M \quad (6)$$

Based on the learned  $\alpha_i$  and support vectors on the training data, we now can predict a label  $\bar{y}$  for any  $x$  using the following decision function in (7) on the fresh testing data, which means for each heartbeat candidate in the testing session, this function can predict whether it is a heartbeat or motion artifacts-induced interferential spikes. The identified ECG-based heartbeats from the weak arm-ECG stream will then be used in PPG-based heartbeats determination and also heart rate estimation.

$$\bar{y} = \text{sign}(\sum_{i=1}^M \alpha_i y_i K(x_i^T, x) + b) \quad (7)$$

In the fourth step, after we run the SVM classification to identify the high confident heartbeats, we can get all high confident heartbeats from the heartbeat candidates which include both real heartbeats and motion artifacts-induced interferential heartbeat-like spikes. The identified ECG heartbeats are then used for heart rate estimation, and also PPG heartbeat identification.

One thing worth noting is that the SVM-based heartbeat identification algorithm can be run in real-time. After training the SVM classifier based on all training data, the trained SVM model can be applied to each heartbeat candidate in the fresh testing data to predict whether it is a real heartbeat or a motion-artifacts-induced interferential spike.

#### 4.2.6 PPG-based Heartbeat Identification

PPG-based heartbeat arrives later than the ECG-based heartbeat, as shown in Figure 4.3, where a pink dot corresponds to a PPG waveform foot and has been used in many works to represent the PPG-based heartbeat occurrence time [66, 67], and a green dot corresponds to the R peak of an

ECG pulse and represents the ECG-based heartbeat occurrence time. Correspondingly, in our algorithm as shown in Figure 4.1b, the minimum point between two adjacent R peaks are identified as the PPG-based heartbeat locations.

#### **4.2.7 Heart Rate Estimation**

As illustrated in the top right part of Figure 4.1b, after calculating the instantaneous heart rate (denoted as IHR) based on the identified ECG-based heartbeats, the windowed heart rate (denoted as HR or windowed HR) estimates are then achieved by averaging the IHR estimates over each time window, with the window corresponding to the second minute in each two-minute trial during which the reference SBP is measured. The performance of the windowed HR in the testing session will be evaluated in term of mean absolute error (MAE) and root mean square error (RMSE), with a unit of beats per minute (BPM).

#### **4.2.8 Pulse Transit Time**

Pulse transit time reflects the time delay between the pressure pulse flows from the proximal to the distal arterial sites [58]. When using ECG and PPG to estimate the PTT as shown in Figure 4.3, the proximal arterial site usually refers to the thoracic aorta and the corresponding PTT start time is approximately measured as the ECG heartbeat R peaks, while the distal site often means the skin surface where the PPG signal is collected and its waveform foot gives the PTT end time. Therefore, the PTT can be obtained by subtracting the PTT start time from the PTT end time, i.e., the time delay between the R peak in the ECG signal and the waveform foot in the PPG signal, as shown in (8) where  $i$  is the  $i$ -th PTT to be estimated. The PTT values measured in the second minute of each trial are then averaged to obtain the window-based PTT estimates.

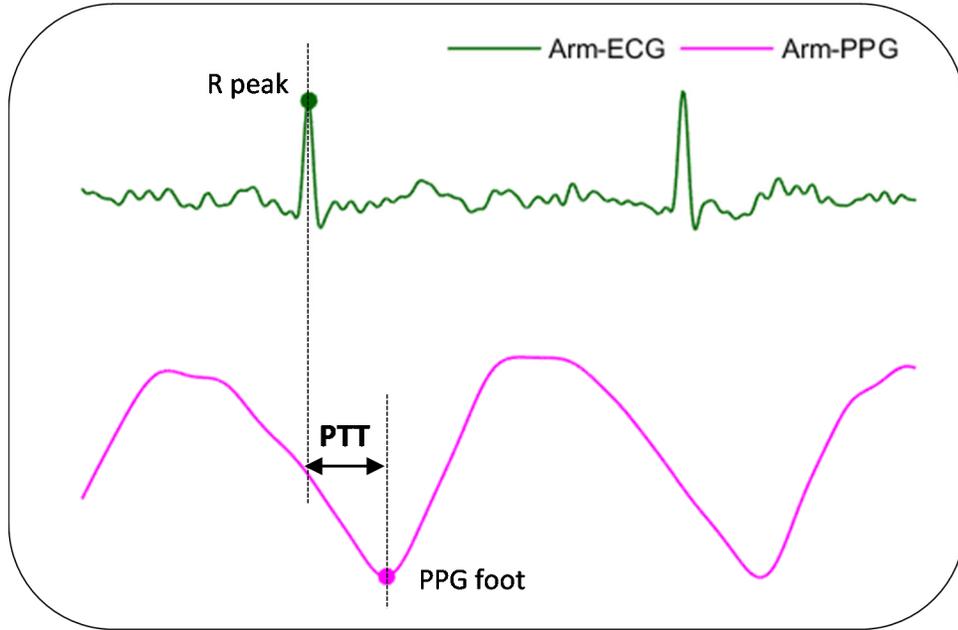


Figure 4.3. Pulse transit time measured with ECG and PPG signals (This illustration of PTT is based on arm-ECG/PPG signals).

$$PTT_i = PPG_i^{foot} - ECG_i^{Rpeak} \quad (8)$$

However, although the above method for PTT estimation has been used in many works for convenient and simplicity purpose, the measurement actually includes another extra item, i.e., the pre-ejection period (PEP) which corresponds to the aortic valve opening time and usually makes the PTT measured significantly larger than its real value [68]. To measure PEP for SBP model improvement, an additional signal usually needs to be acquired, such as the impedance cardiography (ICG) or the phonocardiogram (PCG) [69], which inevitably causes extra hardware burden and impacts the wearability. Instead of measuring PEP to improve the PTT estimate, another strategy for SBP model enhancement is introducing HR to the original PTT-SBP model to form a new PTT&HR-SBP model, leveraging the correlation between HR and SBP [70]. Actually,

introducing HR to the SBP model is natural since the HR information is already carried by the ECG signal and can be robustly measured by appropriate algorithms, such as MLEF in our study, without adding extra signal acquisition hardware.

#### **4.2.9 Blood Pressure Estimation**

The PTT&HR-SBP model is chosen for SBP estimation considering both PTT and HR are correlation with SBP. Meanwhile, we also implemented PTT-SBP models for comparison purpose, which include a bunch of modeling strategies based on different assumptions, such as the linear, quadratic and exponential equations. In total, ten SBP models are evaluated as shown in Table 4.1, and a thorough comparative analysis based on experimental results will be given in the next sections to show that PTT&HR-SBP models are superior to PTT-SBP models.

In Table 4.1, the listed ten blood pressure models not only cover SBP models based on the linear, quadratic and exponential assumptions, but also include SBP models with or without HR information embedded [58, 59, 68-72]. These models are based on different mechanisms and deduction processes.

For example, in model 2, the PTT is reversely related to the SBP since the time delay for the mechanical pressure wave to propagate between the proximal and the distal sites is usually reduced with a higher SBP, and vice versa [58]. In model 7, the relationship between SBP and PTT is demonstrated based on the combined action of the pulse wave and the energy of wave (kinetic and the gravitational potential energy) [69]. In model 10, the embedding of HR is based on the consideration that the cardiac output flow usually increases with HR, and thus SBP would increase with HR if assuming the arteries is purely resistive [70]. More details of these models can be found in [58, 59, 68-72].

To evaluate the generalization ability of the trained SBP models to the unseen data, the SBP models fitted on the data in the first session (training session) is tested on the unseen data in the second session (testing session).

Both training and testing performance is given in terms of many different criterion, including Bland-Altman plot [73], mean error (ME) $\pm$ standard deviation (STD), MAE and RMSE. One thing worth noting is that the testing performance is highly important to show the generalization ability of the learned SBP models.

Table 4.1. Ten blood pressure models for comparative analysis.

No.	Equation	HR information
1	$SBP = a \ln(PTT) + b$	w/o
2	$SBP = a PTT^{-1} + b$	w/o
3	$SBP = a PTT + b$	w/o
4	$SBP = a PTT^2 + b PTT + c$	w/o
5	$SBP = a PTT^2 + b$	w/o
6	$SBP = a e^{b PTT}$	w/o
7	$SBP = a PTT^{-2} + b$	w/o
8	$SBP = a PTT^{-2} + b HR^{-2} + c$	w/
9	$SBP = a \ln(PTT) + b \ln(HR) + c$	w/
10	$SBP = a PTT + b HR + c$	w/

## 4.3 RESULTS

In this section, both the proposed hardware prototype and the HR/SBP estimation algorithms are evaluated in detail, according to the signal processing flow shown in Figure 4.1.

### 4.3.1 Signals Acquired

The signals acquired by the proposed hardware prototype are given in Figure 4.4. In Figure 4.4a-c, three signals, i.e., chest-ECG, arm-PPG and arm-ECG, are compared in terms of several aspects including signal morphology and amplitude. A zoomed in version of the arm-ECG is also given in the bottom right part (Figure 4.4d) to show the details. There are several interesting observations from these illustrations. Firstly, compared with the chest-ECG signal with the electrodes placed close to the heart, the arm-ECG signal has a much lower amplitude (around 10% of that of the chest-ECG signal in this example). This is due to the fact that the arm-ECG electrodes are put not only further from the heart, but also have a small relative distance since they are constrained by the same arm band. Actually, when the electrodes are not close to the heart, the distance between the reference and the signal electrodes is the key factor to be able to observe the ECG signal. In our study, by placing the reference electrode on the top side of the left upper arm and the signal electrode on the bottom side to maximize the distance between these two electrodes, a distinguishable heartbeat wave is finally observed as shown in Figure 4.4d.

Secondly, some portion of the arm-PPG signal is of a clear morphology, indicating that the reflected LED light is quasi-periodically related to the amount of the blood flowing through the arm tissue under the PPG sensor. Thirdly, both the arm-ECG and arm-PPG signals are sensitive to the motion artifacts and EMG noise. This is due to the fact that the arm is almost all surrounded

by muscles and even slight body movements usually relate to random stretching of these muscles, which may cause worse sensor-skin contact and thus induce many artifacts and noise to the signals. Even the signals without riding the bike are chosen to estimate the SBP, there is still considerable amounts of motion artifacts and EMG noise since the body is usually not strictly still, posing more challenges to the following heartbeat identification process.

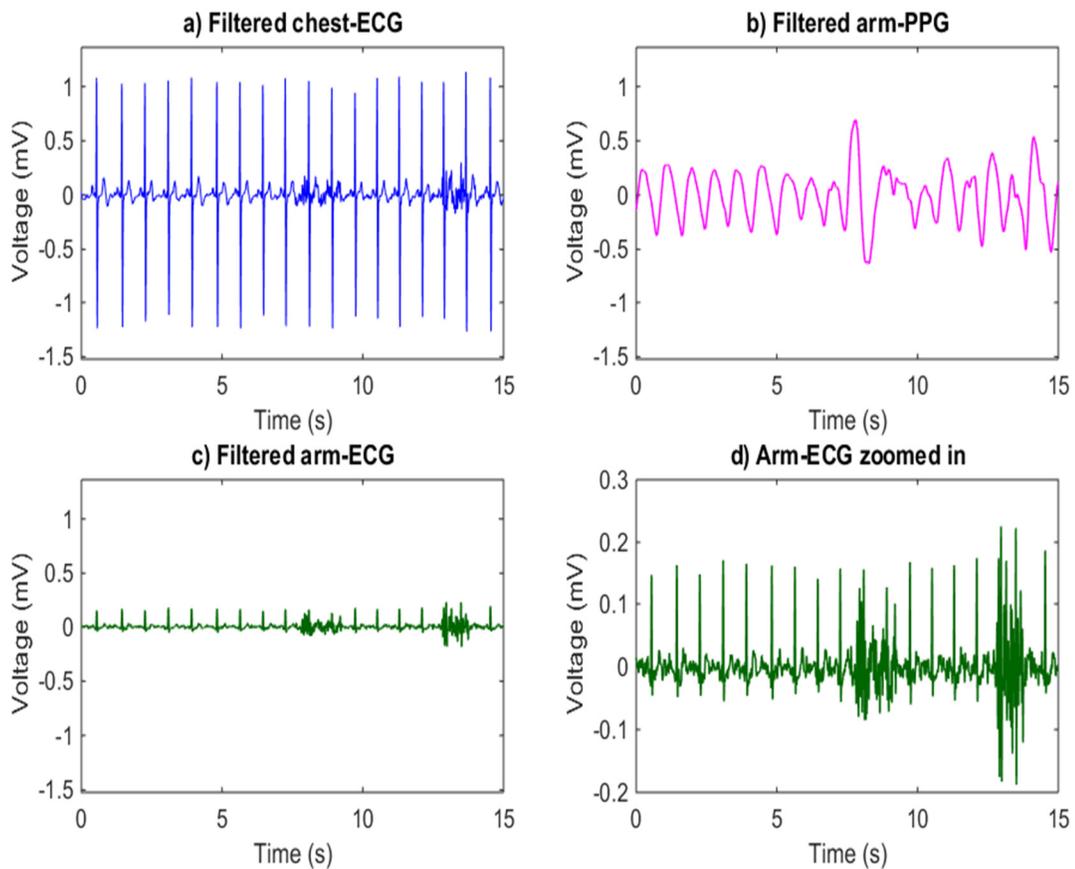


Figure 4.4. An example of the signal segments acquired (chest-ECG, arm-PPG and arm-ECG). Notes. 1) The weak arm-ECG has an amplitude only around 10% of that of chest-ECG; 2) the arm-ECG is so sensitive to motion artifacts and EMG noise that advanced algorithms are necessary for robust heartbeat identification and heart rate estimation.

### 4.3.2 Heartbeat Identification and Heart Rate Estimation

The arm-ECG-based heartbeat locations (R-peaks) are firstly identified by MLEF and then the arm-PPG-based ones (PPG feet) are obtained by searching the minimum point in the arm-PPG signal between two adjacent R-peaks. Afterward, the PTT are estimated with the recognized R-peaks and PPG feet, which will be analyzed later. Meanwhile, the instantaneous heart rate information is obtained from the arm-ECG signal, denoted as IHR@arm-ECG. The IHR@arm-PPG is estimated at the same time for comparison purpose. Besides, to calculate the ground truth IHR, the heartbeats are detected from the strong chest-ECG signal by a simple threshold method and manually checked by us to guarantee the correctness. The obtained ground truth IHR is denoted as IHR@chest-ECG. The correlation matrix of the IHR estimates in the testing session of one subject using these three approaches is illustrated in Figure 4.5a.

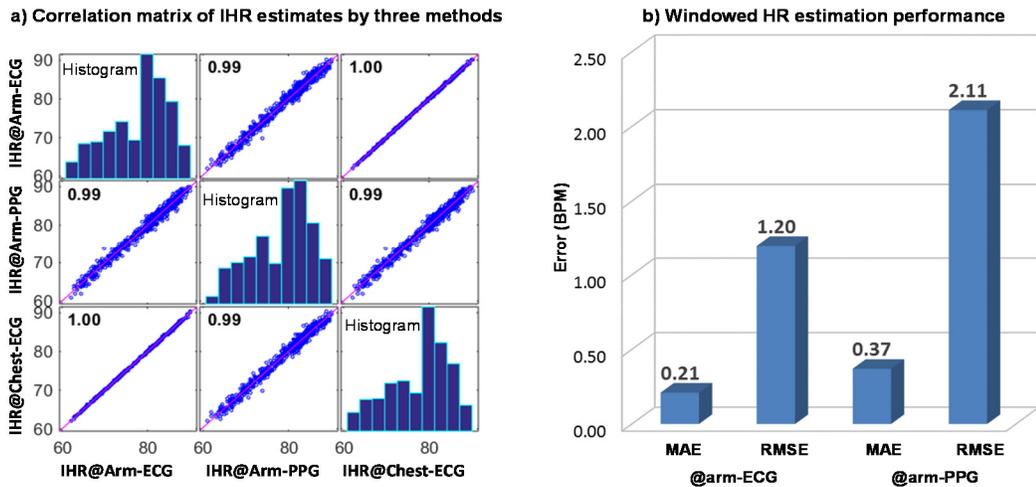


Figure 4.5. Heart rate estimation performance.

Notes. a) Correlation matrix of estimated IHR@arm-ECG, estimated IHR@arm-PPG and ground truth IHR@chest-ECG, in the testing session of subject 1 (unit of IHR: beats per minute, i.e., BPM); b) Performance of windowed HR in the testing session averaged over all subjects. MAE: mean absolute error; RMSE: root mean square error.

In this example, a correlation as high as 1.00 between the IHR@arm-ECG and the IHR@chest-ECG indicates that the ECG heartbeat locations in the weak arm-ECG are successfully identified by the MLEF and thus can be used for robust IHR estimation. On the other hand, the IHR@arm-PPG owns a lower (0.99) correlation with the ground truth IHR@chest-ECG, which indicates that the time varying blood pressure often fluctuates the propagation time the mechanical pulse takes to arrive the tissue under the arm-PPG sensor and thus impacts the IHR estimation. Therefore, we choose the IHR@arm-ECG to get the windowed HR (averaged over the second minute of each trial) for SBP model enhancement, and to report the HR estimation performance.

The performance of windowed HR in the testing session averaged over all subjects is given in Figure 4.5b, where both arm-ECG and arm-PPG-based HR are compared with the ground truth HR@chest-ECG. The HR@arm-ECG corresponds to an MAE of 0.21 BPM and a RMSE of 1.20 BPM, while HR@arm-PPG has higher errors, i.e., an MAE of 0.37 BPM and a RMSE of 2.11 BPM. The results are consistent with the IHR correlation matrix and further indicate that the arm-ECG can be a more robust alternative to chest-ECG in wearable applications.

### **4.3.3 Correlation Among PTT, HR and SBP**

After obtaining the PTT and HR with the identified heartbeats in the arm-ECG and PPG signals, an example of their correlation with the reference SBPcuff is illustrated in Figure 4.6, based on the data in the training session of one subject. Specially, at the beginning of this session (trial 1), a large PTT relates to a low SBP. Afterwards, the SBP gradually increases due to the interventions (exercise) and then stays at a high level (trial 2-11). The PTT also gradually increases, but it does not steadily stay at a high level like the SBP. Finally, at the end of this session (trial 12-13), PTT increases and the SBP decreases over time.

Basically, PTT has a reverse but not completely consistent changing trend compared to the SBPcuff, especially when the SBP is high (due to exercise stress in our study). This limitation gives the opportunity to introduce other variables related to the SBP to enhance the SBP model. As shown in Figure 4.6, the HR@arm-ECG has a similar trend as the SBPcuff, indicating that the introduction of HR information may significantly contribute to the robustness improvement of the SBP model.

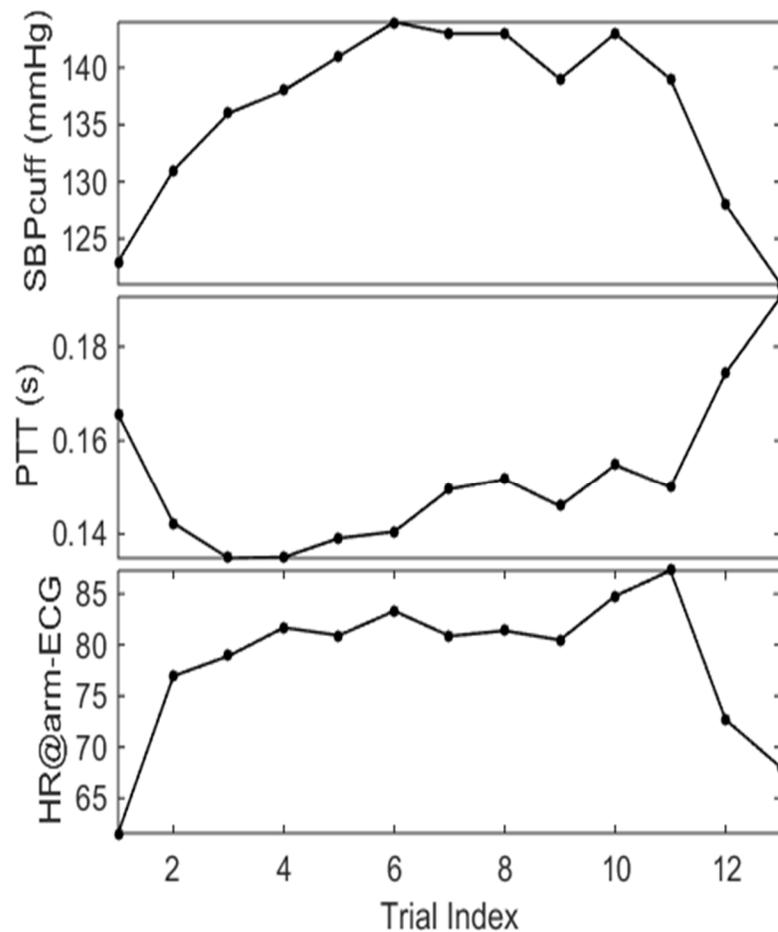


Figure 4.6. Trend of SBPcuff, PTT and HR@arm-ECG, corresponding to the training session of subject 1.

#### 4.3.4 Blood Pressure Model Training and Testing

The SBP models are firstly tuned on the training data, and then evaluated on the unseen fresh testing data. To emphasize the generalization ability of the tuned SBP models, we mainly report the testing performance. As mentioned above, we choose the PTT&HR-SBP model for our SBP estimation task, meanwhile, we have also implemented seven different PTT-SBP models for comparison purpose. One thing worth noting is that, even we have evaluated three PTT&HR-SBP models, we will finally choose the worst case performance among these three models, to fairly illustrate the generalization ability of the PTT&HR-SBP models when introducing the HR information.

To show the different ability of these two kinds of models with or without HR information enhanced, the visualized comparison is given in Figure 4.7, where the left side gives the training performance for model 1 as shown in Figure 4.7a1 (PTT-SBP model) and model 10 as shown in Figure 4.7a2 (PTT&HR-SBP model), and right side illustrates the testing performance on the fresh data for model 1 as shown in Figure 4.7b1 and model 10 as shown in Figure 4.7b2. From the comparison, there are two interesting observations. Firstly, the testing performance on the fresh testing data is worse than the training performance which is based on the same data used for model tuning, indicating that it is highly necessary to apply the tuned models on the unseen fresh data to evaluate their generalization ability.

Secondly, model 10 outperforms model 1 in terms of both training performance and testing performance, indicating that the PTT&HR-SBP is superior to PTT-SBP model and more importantly owns a better generalization ability on the fresh data, leveraging the introduced HR information in model establishment.

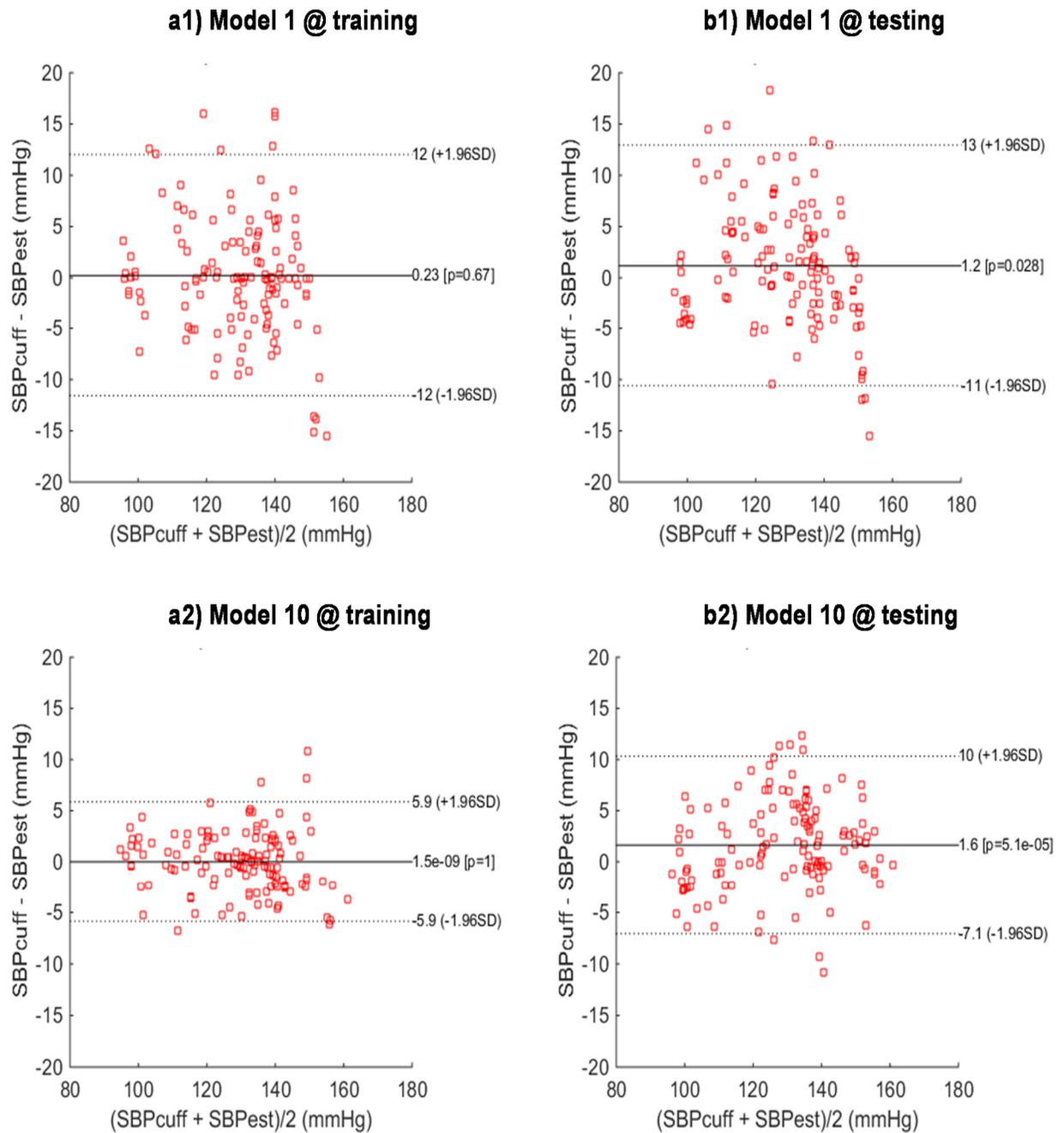


Figure 4.7. Performance of model 1 and 10 in terms of Bland-Altman plot.

Notes. 1) The testing performance on the unseen fresh testing data on the right side is worse than the training performance; 2) PTT&HR-SBP model 10 is superior to PTT-SBP model 1 showing the necessity to include the HR information in SBP modeling.

To effectively illustrate the generalization ability of all ten SBP models, their testing performance is summarized in Table 4.2, which not only includes the performance evaluation for all the SBP models in terms of four criteria (ME, STD, MAE and RMSE), but also illustrates the SBP performance based on the chest-ECG and arm-PPG to show that the arm-ECG can be a robust alternative of the chest-ECG. From the table, there are also several interesting observations. Firstly, for the SBP estimated from arm-ECG and arm-PPG, model 1-7 (PTT-based approaches) own similar testing performance, obtained from the correlation between PTT and SBP as shown in Figure 4.6. However, model 8-10 can achieve much more robust SBP estimates, benefitting from the additionally introduced variable, i.e., HR. Besides, these three models own similar estimation performance, showing that the variables which highly correlates to the SBP, rather than the mechanisms used in the models, dominate the contribution. For the SBP estimated from chest-ECG and arm-PPG, a similar observation also holds. Secondly, there is only little difference between the results based on arm-ECG/PPG and chest-ECG/arm-PPG. For example, using model 10, the testing performance based on arm-ECG and arm-PPG is  $1.63 \pm 4.44$ , 3.68, 4.71 mmHg in terms of  $ME \pm STD$ , MAE and RMSE, respectively, while based on chest-ECG and arm-PPG, they are  $1.61 \pm 4.42$ , 3.62 and 4.69 mmHg, respectively. The results indicate that the weak arm-ECG can effectively replace the chest-ECG in terms of robust SBP estimation, and thus provides a convenient and comfortable sensor placement method to enable a high wearability in long-term applications. According to the Advancement of Medical Instrumentation (AAMI) standard [74], the blood pressure estimation error needs to be below  $5 \pm 8$  mmHg in terms of  $ME \pm STD$ . Although all results in Table 4.2 meet the AAMI standard, the PTT&HR-SBP model 8-10 own much better performance benefiting from the introduced HR information in SBP model

establishment. As mentioned before, the worst case testing performance among three PTT&HR-SBP models (i.e., model 10) is chosen as the final performance for the PTT&HR-SBP models, to fairly illustrate the generalization ability of the SBP models when introducing the HR information.

Table 4.2. A summary of the testing performance of ten SBP models.

No.	<u>With arm-ECG and arm-PPG</u>				<u>With chest-ECG and arm-PPG</u>			
	ME	STD	MAE	RMSE	ME	STD	MAE	RMSE
1	1.17	6.00	4.73	6.09	0.96	5.97	4.66	6.02
2	1.58	6.62	5.08	6.78	1.46	6.60	5.01	6.74
3	0.73	5.88	4.58	5.90	0.59	5.86	4.54	5.87
4	0.29	6.15	4.82	6.14	0.02	6.33	4.96	6.30
5	0.85	5.85	4.59	5.89	0.68	5.86	4.57	5.88
6	0.74	5.88	4.59	5.90	0.61	5.86	4.55	5.86
7	0.79	5.95	4.65	5.98	0.68	5.91	4.60	5.93
<b>8</b>	<b>1.59</b>	<b>4.44</b>	<b>3.69</b>	<b>4.70</b>	<b>1.58</b>	<b>4.41</b>	<b>3.62</b>	<b>4.67</b>
<b>9</b>	<b>1.62</b>	<b>4.43</b>	<b>3.67</b>	<b>4.70</b>	<b>1.61</b>	<b>4.41</b>	<b>3.61</b>	<b>4.67</b>
<b>10</b>	<b>1.63</b>	<b>4.44</b>	<b>3.68</b>	<b>4.71</b>	<b>1.61</b>	<b>4.42</b>	<b>3.62</b>	<b>4.69</b>

#### 4.4 DISCUSSION

As one of the key enabling factors in effective hypertension and heart health management, ubiquitous SBP and HR monitors have been attracting tremendous attentions [18], but at the same

time, still face several major impediments, such as poor wearability, lack of widely accepted robust SBP models, insufficient proofing of the generalization ability of the fitted SBP models, etc. Focusing on these challenges, many efforts have been made in this study.

It is well known that wearability is directly related to the acceptance degree of the long-term SBP and HR monitors. The traditional approaches such as the invasive catheterization method and the noninvasive cuff-based oscillometry method [58], are not suitable for long-term applications. For ECG and PPG-based wearable solutions, although many investigations have already proposed some sensors placement methods, they still suffer from some limitations in terms of the comfortableness and convenience. For example, to collect the ECG signal, the chest placement has been used in many studies [59, 75], since the electrodes are so close to the heart that a strong ECG signal can be easily obtained, which simplifies the processing algorithms and brings a high motion-artifact tolerant ability. However, the chest electrodes placement may induce uncomfortableness, since a chest strap is usually needed to surround the body to fix the sensors and attach them to the skin. Another uncomfortableness may come from the sweat during long-term applications, especially in hot weather. Besides, some works apply a double-wrist-ECG plus finger-PPG setup [60], which prevents a chest strap, but usually needs an additional wire to collect the skin potential difference from two distant sites, which is highly inconvenient in daily applications. This limitation is also critical in overnight applications since the wire may impact the sleep quality. Of course, an extra wireless module may replace the wire, but inevitably introduces more hardware cost. An interesting alternative is using the non-contact ECG electrodes on the T-shirt which may be both comfortable and convenient, however, one critical challenge of this method is that it usually severely suffers from motion artifacts, meaning that the stretch of cloth or a vertical movement

between the skin and the cloth may significantly impact or corrupt the ECG signal [76]. Another thing worth noting is that the PPG sensor may also introduce uncomfortableness with a chest placement method, or pose more challenges to the integration of PPG and ECG sensors with a double-wrist-ECG plus finger-PPG placement approach [60].

To achieve high comfortableness and convenience, we propose a single-arm placement approach, with both ECG and PPG sensors integrated into one arm band. This method avoids the chest strap and the additional wire or wireless module, providing a high level of integration by putting all sensors into one band. On the other hand, the single-arm placement for the ECG signal acquisition is highly challenging, since the reference/signal electrodes are too close to each other and in a similar direction referring to the heart, resulting in a very weak potential difference. By maximizing the distance between these two electrodes, i.e., placing one of them on the top side of the left arm and the other on the bottom side, and reducing the distance between the heart and the electrodes, i.e., putting the electrodes on the upper arm not the forearm, a weak but still distinguishable ECG signal is finally achieved, with an amplitude around only 10% of that of the chest-ECG signal. The acquired weak arm-ECG signal is successfully used in heartbeat locations identification, PTT measurement, and heart rate estimation which has a high correlation with the estimates based on the chest-ECG signal.

Another crucial consideration is the model applied to estimate the SBP. Although there are a large number of diverse SBP models, many efforts are still being made to further enhance the robustness and it is hard to select out one model which can well fit all application scenarios. In our study, we choose PTT&HR-SBP model to perform the SBP estimation task, meanwhile, we have also implemented PTT-SBP models to make a comparative analysis. In total, ten different SBP

models [58, 59, 68-72] are taken into account in terms of the SBP estimation based on the arm-ECG/PPG signals. These models are built upon different mechanisms. For example, the SBP was assumed to be an inverse linear function of the PTT based on the fact that the propagation time of the mechanical pressure wave is reversely related to the SBP [58]. Another work assumed the SBP to be an inverse quadrature function of the PTT based on the combined action of the pulse wave and the energy of wave (kinetic and the gravitational potential energy) [69]. Also, the SBP was regarded as a linear function of both PTT and HR since the observations showed that both PTT and HR correlate to the SBP [70].

To thoroughly compare these SBP models and illustrate the effectiveness of applying the PTT&HR-SBP model to the SBP estimation task, four major considerations are made. Firstly, seven PTT-SBP models without HR embedded are compared with three PTT&HR-SBP models with HR enhanced, to demonstrate that the HR information can significantly contribute to the robustness of the SBP model. A similar changing trend is observed on both HR and SBPcuff curves when perturbing the SBP with exercise stress, which confirms the previous observation that there is a high correlation between the HR and SBP [70]. Leveraging the HR information, the PTT&HR-SBP models greatly outperform the PTT-SBP models, as shown by the Bland-Altman plot in Figure 4.7 and the performance summary in Table 4.2. Besides, the performance is similar among PTT-SBP models or PTT&HR-SBP models, which indicates that the complexity of the models has less contribution to the model improvement compared to HR, i.e., finding more SBP-related variables seems more significant than assuming diverse equations since the underlying accurate mechanism of SBP is highly complicated. Secondly, in model training, some models have already fixed some parameters in the equation based on empirical statistics, however, all the parameters

are trainable in our study considering there is no empirical statistics for those parameters in single-arm scenarios, which is also to provide more general modeling of the SBP. Thirdly, the evaluation of the SBP models is mainly performed on the unseen fresh testing data, which is the key to show the generalization ability of the trained model. The testing performance in Table 4.2 indicates the fitted model can be well generalized to the data that has not been used to train the model. Finally, considering the single-arm-ECG signal is much weaker than the chest-ECG signal and thus sometimes sensitive to the motion artifacts, we also compare the single-arm-ECG signal with the chest-ECG signal in terms of the SBP estimation, to figure out whether the former one can robustly substitute for the latter one. The slightly different MAE and RMSE results between SBP estimation based on these two ECG placement methods shown in Table 4.2 and 4.3, indicate that the single-arm-ECG is both a highly convenient and robust alternative to the chest-ECG signal.

Based on the experimental results on the above key considerations on the SBP model, the PTT&HR-SBP models with HR enhanced can provide a more robust SBP estimation ability compared with the PTT-SBP models, and the trained models own a good generalization ability to the unseen testing data.

Our current study is expected to illustrate the feasibility of the highly wearable single-arm blood pressure monitoring system. We have evaluated the proposed system on the data collected over a short duration and when subjects took a single sitting position, which is usually the strategy used in many wearable blood pressure monitoring studies [59]. Meanwhile, there are a few works focusing on the blood pressure estimation over a longer duration, such as days or months, and proposing some model recalibration techniques, or focusing on the influence of different body postures on the blood pressure, which show that it may be necessary to further include body posture

information in the blood pressure modelling [59]. Our future research will also include recalibrating the blood pressure model over a longer duration and studying the influence of diverse body postures on the blood pressure modelling.

Moreover, in future, we will further evaluate the proposed system on more single-arm-ECG/PPG data acquired. Besides, incorporating the signal quality information to dynamically refine the SBP estimates may also be helpful since the sensor-skin contact status is usually time-varying and has dynamic impacts on the signals acquired [77]. We are also interested in continuously improving the signal quality of the acquired weak single-arm-ECG signal such that it can provide a clearer morphology with well-formed P wave and T wave for medical interpretation purpose [78, 79].

#### **4.5 CONCLUSION**

We propose a wearable cuff-less system for long-term daily blood pressure and heart rate monitoring applications. To enhance the wearability, the ECG and PPG sensors are all integrated into a single arm band. For arm-ECG acquisition, a non-standard single-lead configuration is provided, which is much more convenient and comfortable than normal wearable ECG placements, such as putting the electrodes on the chest (ECG and PPG), or two wrists (ECG) plus one finger (PPG). A weak arm-ECG signal with an amplitude only around 10% of the chest-ECG signal, is successfully obtained by our newly established hardware prototype. This weak arm-ECG signal is then used for heartbeat location identification and heart rate estimation with a machine learning-enabled framework. The HR estimation has a mean absolute error (MAE) and a root mean square error (RMSE) of 0.21 and 1.20 beats per minute (BPM), respectively. Leveraging both arm-ECG and arm-PPG signals, the pulse transit time (PTT) information is then achieved and applied to systolic blood pressure (SBP) estimation.

The PTT&HR-SBP model is applied for SBP estimation considering both PTT and HR are correlated with SBP. Meanwhile, to thoroughly compare various SBP estimation approaches, ten SBP models based on diverse mechanisms are considered, such as applying a linear, quadratic or exponential equation, or leveraging the heart rate (HR) information to enhance the SBP model. Experimental results show that three PTT&HR-SBP models significantly outperform seven PTT-SBP models, indicating the necessary to introduce the HR information in robust SBP estimation. Moreover, the trained SPB models based on the arm-ECG/PPG signals are evaluated on the unseen data to show their high generalization ability. The estimated SBP is comparable to the estimates based on chest-ECG/arm-PPG signals, meaning that the arm-ECG signal can effectively replace the chest-ECG signals in terms of robust SBP estimation. Specially, for the PTT&HR-SBP models, the SBP estimation performance is  $1.63 \pm 4.44$ , 3.68, 4.71 mmHg in terms of  $ME \pm STD$ , MAE and RMSE, respectively. In summary, the proposed wearable cuff-less SBP and HR monitoring system is expected to enable highly convenient SBP and HR monitoring applications and contribute to ubiquitous long-term hypertension, heart health and fitness management.

## CHAPTER 5

### BLOOD PRESSURE MONITORING WITH EAR SIGNALS

In this chapter, we propose another high wearable system for long-term blood pressure and heart rate monitoring, by placing the ECG and PPG sensors all behind the ear. We also introduce large amounts motion artifacts to the signals acquired by performing head movements, and study propose advanced signal processing and machine learning algorithms to deal with the motion artifacts for robustness enhancement purpose.

#### 5.1 INTRODUCTION

High blood pressure (BP), also called hypertension, is a common but dangerous condition, impacting over 35 percent of people, relating to many cardiovascular, circulatory and cerebrovascular diseases, and causing 12.8 percent of total deaths worldwide [18]. Long-term BP monitoring is a key factor in hypertension control for several reasons. Firstly, hypertension is often associated with few or no symptoms which makes timely early diagnosis and treatment highly challenging without daily BP tracking. The more data we have, the more possibility to find some risky patterns as early as possible.

Moreover, a frequent BP measurement is significant to check up how well the treatment of hypertension is going on. Besides, BP always fluctuates over time, making the measurement at specified times and circumstances insufficient for effective BP monitoring and analysis. However, the traditional BP measurement methods cannot well fit the needs of long-term applications, such as the invasive catheterization method sacrificing the comfortableness and the noninvasive cuff-based oscillometry approach lacking of a good wearability [58].

Wearable computers are paving a promising way for pervasive smart health wearables. Specially, for wearable BP monitoring, many investigations have been reported and summarized in several recently published surveys [58, 59]. According to these studies, the most popular BP estimation theories are based on the fact that BP is correlated with the pulse transit time (PTT), i.e., the blood wave propagation time between two arterial sites. To measure the PTT start and end time, the electrocardiography (ECG) and photoplethysmogram (PPG) signals are the most widely used signals, which correspond to the pressure wave occurrence time on the proximal and distal arterial sites, respectively. In these works, the most frequently applied ECG/PPG sensors placement methods are two-wrists/finger, chest/finger, and chest/chest. However, these placement approaches may impact the wearability and comfortableness, considering the former two require extra connection overhead or wearing more than one devices, and the last one may need a chest strap to fix the sensors and suffer from sweating. Some works [80, 81] proposed an in-ear PPG signal monitor to measure HR and other information, but they did not acquire ECG signal and measure BP. Another work [82] proposed placing the PPG sensor behind the left ear and placing two ECG electrodes behind the left ear and neck, respectively. However, the signal quality may be impacted if the collar coat touches the electrode on the neck in long-term daily applications. Moreover, this work did not evaluate the BP estimation performance after obtaining the PTT measurements, and did not consider daily movements-induced motion artifacts.

Another significant concern lacking of enough attention and study is whether BP estimation systems can tolerate to large amounts of motion artifacts [83, 84], since the body movements in long-term daily applications inevitably induce time varying skin-sensor contact variations which usually impact or even corrupt the ECG and PPG signals acquired. Moreover, there are diverse BP

modeling theories and strategies being studied, to deal with the underlying complicated blood pressure wave generation and propagation mechanisms, nevertheless, the comparative analysis of major BP models is rather limited [58, 59].

Besides, the heart rate (HR) information is also highly significant in both direct hypertension management and BP model enhancement, but it is not taken into account in many current wearable blood pressure monitoring studies [70, 83, 84].

In this study, focusing on above-mentioned challenges, we propose a novel machine learning-enabled highly wearable system for pervasive long-term cuff-less BP and HR monitoring, from ear-ECG and ear-PPG signals with a motion artifacts-tolerant ability. Firstly, to meet critical requirements on the wearability and comfortableness in long-term daily applications, a highly convenient sensors placement method is suggested which allows situating the ECG and PPG sensors all behind two ears and the possibility to integrate them into glasses or ear headsets. Using our semi-customized hardware prototype, the weak ECG and PPG signals are successfully acquired with good morphologies.

Secondly, the weak ear signals during BP and HR estimation are inevitably impacted and usually corrupted by many random motion artifacts even the users take a sitting or standing position, including continuous background motion artifacts due to uncontrolled neck muscle and blood vessels movements, and motion artifacts induced by normal head movements such as looking around or up and down. In our study, to provide a high motion artifacts-tolerant ability towards practical scenarios, we introduce much more severe head motions including shaking the head and nodding for one third of each signal period used for BP and HR estimation. After deliberately introducing many head motions, a large portion of signals are highly corrupted which usually

cannot maintain necessary morphological characteristics to represent heartbeat occurrence time. Therefore, we firstly try to identify the raw heartbeats using our previously reported support vector machine (SVM)-based approach [31]. Afterwards, to filter out many residual distorted or faking raw heartbeats, we further propose an unsupervised learning approach to automatically label the heartbeat quality levels and purify the raw heartbeats.

Specifically, we introduce a dynamic time warping (DTW)-based learning approach [85] which can perform non-linear mapping of two time-varying sequences, to measure the dissimilarity between each raw heartbeat and a high quality heartbeat pattern. The measured DTW dissimilarity values are used to quantify the degree of distortion of raw heartbeats and are compared with a distortion threshold, to generate the heartbeat-specific signal quality indices and perform ECG/PPG heartbeats purification. The high quality heartbeat pattern is learned using a K-medoids clustering method [86], and the distortion threshold is learned by a histogram triangle-based method [87].

Thirdly, the HR information is robustly estimated from purified ECG heartbeats, for both heart health management and BP model enhancement purpose. Fourthly, based on HR estimates and the PTT calculated, ten PTT-BP and PTT&HR-BP learning models [58, 59, 68-72] are taken into account to thoroughly compare their difference and determine an effective one for the ear application scenario.

In this study, we take special interest in systolic BP (SBP) estimation. The proposed proof-of-concept system is validated on an acquired ear signal dataset and is expected to contribute to pervasive hypertension, heart health and fitness management.

## **5.2 MATERIAL AND METHOD**

### **5.2.1 System Overview**

The proposed machine learning-enabled system is illustrated in Figure 5.1. The top part (Figure 5.1a) gives the customized hardware prototype and the sensors placement method for ear-ECG and PPG signals acquisition. The bottom part (Figure 5.1b) shows the flow of the signal processing and HB/SBP estimation algorithm including three stages, i.e., stage I – supervised learning of heartbeat (HB) identification, stage II – unsupervised learning of signal quality labelling and signal purification, and stage III – HR estimation and supervised learning of SBP estimation.

### **5.2.2 Customized Hardware Platform and Sensors Placement**

The customized hardware platform shown in Figure 5.1a includes two parts, i.e., the ECG [31] and PPG subsystems. In the former one, the ECG signal is acquired by an ADS1299 24-bit analog-to-digital (ADC) with a sampling rate of 500 Hz on a TI ADS1299EEG-FE evaluation board (green one) [29] and is then sent via the SPI port to a TI Tiva™ C series LaunchPad (red one) [30], which is equipped with an ARM Cortex M4 microcontroller to configure the ADC and relay the signal to a PC via the USB port.

In the latter one, the PPG signal is acquired by a 22-bit ADC with a sampling rate of 128 Hz on a TI AFE4490SPO2 evaluation board [61], which also owns an MSP430F5529IPN microcontroller to configure the ADC and relay the data to the PC. A higher sampling rate for ECG is based on the consideration that it is used for both HR and PTT estimation. This prototype can be conveniently used in long-term applications after removing evaluation-specific components and adding a wireless module.

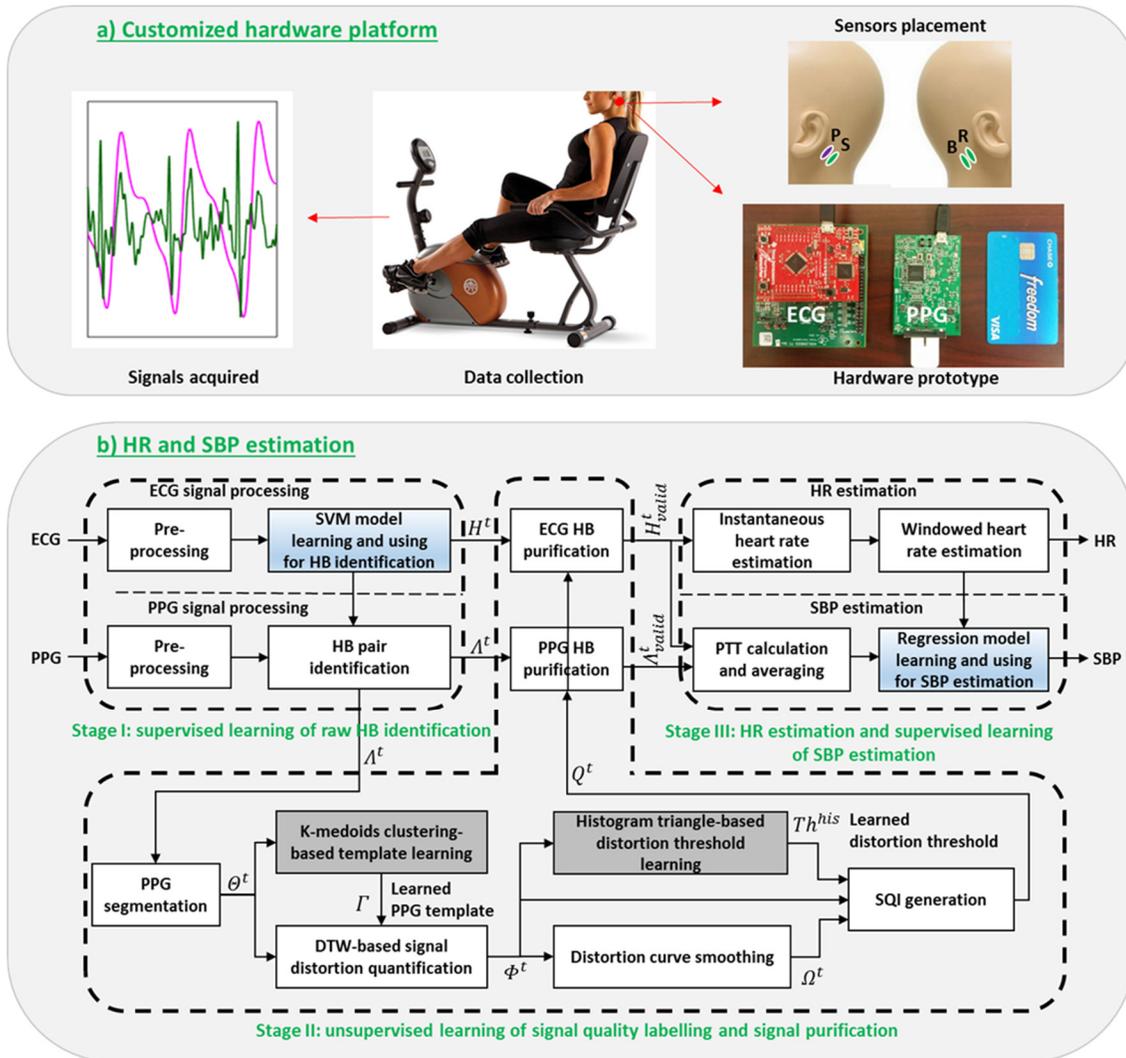


Figure 5.1. The proposed machine learning-empowered system for wearable cuff-less SBP and HR monitoring from motion artifacts-sensitive ear-ECG/PPG signals.

**Algorithm block coloring in b):** white colored blocks are active in both training and testing phases; each blue colored block include both a supervised model learning process and a model using process, which are active in the training phase and the testing phase, respectively; each gray colored block includes an unsupervised model learning process and is active only in the training phase.

**Abbreviations:** R/B/S represent the reference/bias/signal electrodes used for single-lead ECG signal measurement, respectively; P corresponds to the reflective PPG sensor; SVM, support vector machine; HB, heartbeat; DTW, dynamic time warping; SQI, signal quality index; PTT, pulse transit time; HR, heart rate; SBP, systolic blood pressure.

**Definitions of the Greek letters:** refer to section II material and method => unsupervised learning of signal quality labelling and purification => algorithm 1.

The sensors placement method proposed is illustrated in the top right part of Figure 5.1a, where the R/B/S correspond to the reference/bias/signal electrodes for single-lead ECG acquisition, and P represents the sensor for PPG measurement. All the sensors can be integrated into glasses or ear headsets to achieve a much higher wearability and comfortableness, compared with the chest or wrists placement.

### **5.2.3 Dataset Recording**

The customized platform was applied to acquire an ear ECG/PPG dataset from fourteen subjects, to evaluate the effectiveness of the proposed proof-of-concept system. The data collection was approved by the university IRB office and the informed consent was obtained from all participants. The data collection comprises a thirty-minute training session and a thirty-minute testing session for each subject. Each session can be further split to fifteen two-minute trials. During the first seven trials, the subject stayed still to get low SBP conditions, and during each of the other eight trails, the subject rode a recumbent exercise bike in the first minute and stayed still in the second minute, to perturb the SBP to a larger range similar to the methods used in many studies [58, 59]. The reference SBP, denoted as SBP<sub>cuff</sub>, was measured on the left arm in the second minute of each trial, using an ambulatory BP monitor CONTEC ABPM50 [65]. Correspondingly, the ear signals in the second minute of each trial are used for HR and SBP estimation. The chest-ECG signal was also collected to obtain the ground truth of heartbeat occurrence time.

One thing worth mentioning is that, the weak ear signals are already impacted or even corrupted, by normal head movements-induced motion artifacts, and background motion artifacts which are continuously induced by uncontrolled neck muscle and blood vessels movements, especially with exercise stress. To enhance the motion artifacts-tolerant ability towards practical application

scenarios, we further deliberately introduce much more severe head motions for one third of each signal period used for BP and HR estimation. Specifically, in the second minute of each trial, the subject was asked to shake the head during the first ten seconds and nod during the fourth ten seconds. These head movements corrupt a large portion of signals and make heartbeat identification much more challenging.

#### **5.2.4 Supervised Learning of HB Identification**

The stage I of the proposed algorithm in Figure 5.1b performs raw heartbeat identification from both pre-processed ear-ECG and PPG signals. Considering the ECG signal is of relatively richer signal characteristics (especially the QRS complex) than the PPG signal (detailed visualization will be given in the results section), we firstly introduce an advanced supervised machine learning approach for raw ECG heartbeats identification, based on which the raw PPG heartbeat pairs are then determined by a minima searching method [58].

##### **1) Signal Pre-processing**

The raw ear-ECG and PPG signals are both processed by a six-order Butterworth bandpass filter (2-30 Hz and 0.5-8 Hz, respectively). Then PPG is resampled to 500 Hz to obtain a same time resolution as ECG. An example of acquired ear-ECG and PPG signals with only background motion artifacts is given in Figure 5.2, where clear signal morphologies can be observed. More analysis about the signal quality with deliberately introduced severe motion artifacts will be given later.

##### **2) ECG-based and PPG-based Heartbeat Identification**

To identify raw heartbeats from weak ear-ECG signal impacted or corrupted by large amounts of background and deliberately introduced motion artifacts, our previously reported SVM-based

approach is applied [31]. Specifically, after segmenting the ECG stream to heartbeat candidates by an adaptive threshold-based auto-segmentation approach, ten critical multi-domain features are extracted from each candidate and fed into the SVM classifier for raw heartbeat identification. The SVM model is learned based on the ear-ECG training data and tested on the ear-ECG testing data for each subject. The heartbeat locations from the chest-ECG stream collected at the same time are used as reference to label whether the ear-ECG heartbeat candidates are real heartbeats or motion artifacts-induced interferential spikes, to enable a supervised learning process.

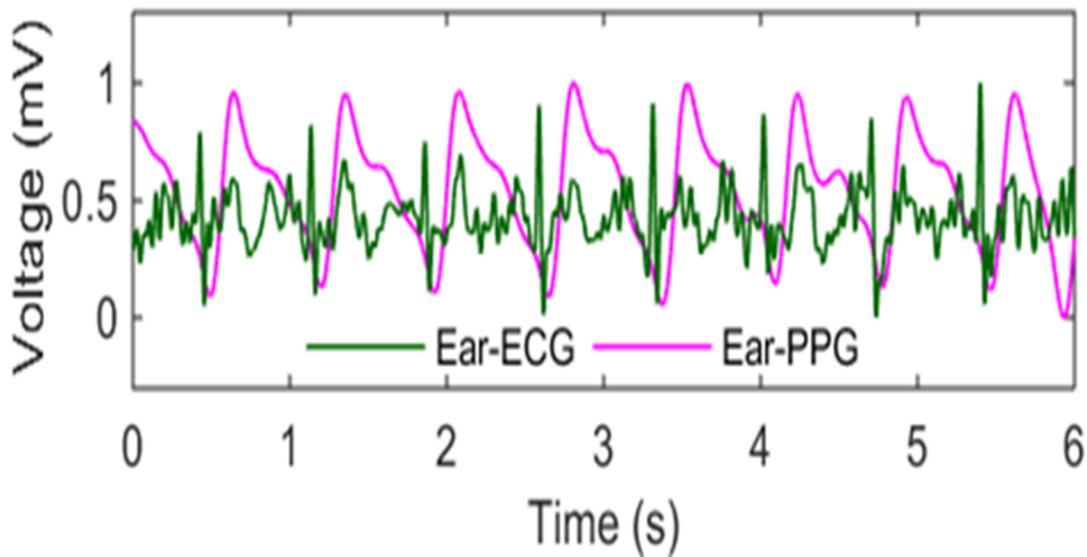


Figure 5.2. Ear-ECG/PPG Signals.

Notes. The amplitude is scaled to be between 0 and 1 for good readability (further analysis on the signal quality such as the signal strength and morphology will be given in the results section).

The SVM decision function is presented by (1), where  $x$  is the ten-dimensional input vector,  $x_i$  is the  $i$ -th support vector in  $m$  support vectors learned and  $y_i$  is its class label,  $\alpha_i$  is the weight learned,  $b$  is the bias learned, and  $k$  is the kernel which is chosen as a linear one to lower

computation load in wearable applications. After identifying the raw ECG heartbeats, the raw PPG heartbeats (PPG feet) can be consequently determined by searching the minimum points between each two adjacent R peaks [58].

$$f(x) = \text{sign}(\sum_{i=1}^m \alpha_i y_i \cdot k(x, x_i) + b) \quad (1)$$

### 5.2.5 Unsupervised Learning of Signal Quality Labelling and Purification

Considering that large amounts of back ground and head movements-induced motion artifacts usually severely corrupt a large portion of weak ear signals, many highly distorted or faking heartbeats in some signal segments need to be suppressed. Therefore, an unsupervised learning approach is proposed to automatically purify the raw heartbeats, as shown in stage II of the proposed algorithm in Figure 5.1b. It can self-learn behaviors of raw PPG heartbeats, automatically generate heartbeat-specific signal quality indices and then purify the raw ECG/PPG heartbeat pairs identified by the stage I of the algorithm. Two major considerations made here include: 1) choosing an unsupervised learning strategy not a supervised one, and 2) further learning motion artifacts-sensitive behaviors of raw PPG heartbeats not ECG heartbeats.

The former one is based on the finding that it is hard to generate ground truth signal quality labels for raw heartbeats (e.g., labelled as a good or poor quality level when using a binary labelling method) which are necessary for supervised signal quality learning. Firstly, the background motion artifacts induced by uncontrolled neck muscle and blood vessels movements, especially with exercise stress, usually occur randomly making it difficult to manually label the signal quality for raw heartbeats. Moreover, the head movements deliberately introduced (shaking and nodding) to generate more critical motion artifacts cannot be strictly controlled, because: 1) it is usually

difficult to precisely control the start and end time of head movements according to the data recording protocol due to the subject-dependent command response delay [88]; 2) the non-linear distortion behaviors of the raw heartbeats cannot be easily quantified, resulting from both high inter-subject and intra-subject variabilities in the swing and speed of head movements. Therefore, we propose an unsupervised learning approach to automatically label the signal quality after self-learning the diverse behaviors of raw heartbeats corrupted by motion artifacts.

Further learning motion artifacts-impacted behaviors of raw PPG heartbeats is based on the observation that the PPG signal is of less signal characteristics and thus more sensitive to motion artifacts than the ECG signal, which will be visualized in the results section. Therefore, the difference between high quality and low quality raw PPG heartbeats is more learnable for an unsupervised learner.

The detailed learning process is given in Algorithm 1. Firstly, the PPG stream is split to raw PPG heartbeat segments (s1-segmentation), which are then fed to a K-medoids clustering-based learner to determine a high quality heartbeat template (s2-template learning). Afterwards, the learned template is used to screen the raw PPG heartbeats to quantify the degree of distortion based on the DTW approach (s3-HB distortion), which can effectively measure the dissimilarity between raw heartbeats owning time varying length/morphology and the learned template using a dynamic programming strategy. The distortion values measured are used to learn by a histogram triangle-based method a distortion threshold (s4-threshold learning), which can be applied to generate binary heartbeat-specific signal quality indices for heartbeat purification purpose (s5-s8).

One thing worth noting is that the step 2) ‘PPG template learning’ and the step 4) ‘distortion threshold learning’ are only active in the training phase, and all the other steps are active in both

training and testing phases. Moreover, in the testing phase, the algorithm can be performed in real time, since two ‘for loops’ in Algorithm 1 can be merged such that the algorithm can be executed trail by trail.

<b>Algorithm 1</b> Unsupervised learning of SQLSP (signal quality labelling and signal purification)	
<b>Input:</b>	
raw PPG HB locations in $T$ trials $\Lambda^t, \forall t \in [1, T]$	
raw ECG HB locations in $T$ trials $H^t, \forall t \in [1, T]$	
number of centroids in K-medoids clustering $K$	
number of points used in distortion smoothing $S$	
shrinkage factor used in distortion threshold learning $\delta$	
lower limit of percentage of HBs to be protected $\rho$	
step size in percentage to adaptively adjust the threshold $\eta$	
<b>Output:</b>	
heartbeat signal quality indices in $T$ trials $Q^t, \forall t \in [1, T]$	
valid PPG HB locations $\Lambda_{valid}^t, \forall t \in [1, T]$	
valid ECG HB locations $H_{valid}^t, \forall t \in [1, T]$	
quality indicator for the $t - th$ BP estimate $BP\_QI^t, \forall t \in [1, T]$	
<b>Procedure:</b>	
<b>for</b> $t = 1$ to $T$ <b>do</b>	▷ sweep trials
$\Theta^t \leftarrow$ PPG segments split by $\Lambda^t$	▷ s1-segmentation
<b>if</b> $t = 1$ <b>then</b>	▷ s2-template learning
$L \leftarrow$ averaged length of segments in $\Theta^t$	
$\widehat{\Theta}^t \leftarrow$ all segments in $\Theta^t$ resampled to a length of $L$	
$\Gamma \leftarrow K\_medoids\_clustering(\widehat{\Theta}^t, K)$	
<b>end if</b>	
$\Phi^t \leftarrow \{\phi_j^t   \phi_j^t = DTW(\Gamma, \theta_j^t), \forall \theta_j^t \in \Theta^t\}$	▷ s3-HB distortion
<b>end for</b>	
$\Psi \leftarrow \Phi^1 \cup \Phi^2 \cup \dots \cup \Phi^T$	
$Th^{his} \leftarrow HIS\_TRI(\Psi, \delta)$	▷ s4-threshold learning
<b>for</b> $t = 1$ to $T$ <b>do</b>	▷ sweep trials
$\Omega^t \leftarrow smooth(\Phi^t, S)$	▷ s5-distortion smoothing
$Q^t, BP\_QI^t \leftarrow SQI(\Phi^t, \Omega^t, Th^{his}, \rho, \eta)$	▷ s6-SQI generation
$\Lambda_{valid}^t \leftarrow$ PPG HBs validated by $Q^t$	▷ s7-HB purification
$H_{valid}^t \leftarrow$ ECG HBs validated by $Q^t$	▷ s7-HB purification
<b>end for</b>	

### 1) PPG Segmentation

The PPG stream is segmented based on the raw PPG heartbeat locations identified in stage I. In the second minute of each 2-minute trial, the foot-to-foot PPG segments  $\theta^t$  are obtained as shown in step 1) ‘segmentation’ in Algorithm 1.

### 2) K-medoids Clustering-based Template Learning

As mentioned above, a PPG heartbeat template with a good morphology is needed by the DTW algorithm, to screen the raw PPG heartbeats to calculate their distortion values used in signal quality labelling. However, there is no pre-labeled signal quality information to directly perform PPG template selection, otherwise, it is unnecessary to perform the unsupervised learning of signal labelling in the proposed system.

It is known that when there are more motion artifacts, there is also more morphological randomness in the raw heartbeats [31], which results in a decreasing consistency among them. If we partition the raw heartbeats into different groups according to beat-to-beat consistency, the high quality heartbeats are more likely to be clustered together benefitting a better inter-beat consistency, and the low quality heartbeats tend to be partitioned into multiple clusters due to much more diverse motion artifacts-induced morphologies. Based on this consideration, a K-medoids clustering approach is introduced to learn a good PPG template from the raw heartbeats. K-medoids clustering is a classical unsupervised machine learning algorithm which breaks the objects (raw heartbeats) up into clusters and attempt to minimize the distance (consistency) between objects belonging to a cluster and the representative object designated as the center (medoid) of that cluster [86]. Therefore, the medoid that represents a highest number of objects is selected as the high quality PPG heartbeat template.

Since the K-medoids clustering method usually applies an iterative strategy to search the final  $K$  representative medoids, to lower the computation load the PPG template is only learned from  $\theta^1$ , i.e., the raw PPG heartbeat segments in the second minute of the first trial in the training session. Moreover, also to lower the computation load, the Euclidean distance is chosen to measure inter-object distance. Since the time-varying raw heartbeats are usually of different lengths, they are all resampled to own a length of  $L$ , which is the averaged length of all raw heartbeats in  $\theta^1$ , to enable the Euclidean distance calculation. The template learning process is given in step 2) ‘template learning’ in Algorithm 1.

A K-means++ algorithm [89] is applied for choosing initial cluster medoid seeds, as shown in (2-3), where  $c_j$  is the  $j$ -th seed to be selected from all resampled heartbeats  $\widehat{\theta}_j^1 \in \widehat{\theta}^1$  with probability  $w_{\widehat{\theta}_j^1}$ ,  $d^2(\widehat{\theta}_j^1, c_p)$  is the Euclidean distance between  $\widehat{\theta}_j^1$  and its closest pre-selected medoid  $c_p, p < j$ ,  $D_p$  is the set of all objects closest to medoid  $c_p$ , and  $\widehat{\theta}_h^1 \in D_p$ . That is, each subsequent medoid seed is selected with a probability proportional to its distance to the closest pre-selected closest seed. The number of medoid seeds  $K$ , is set as 20, considering that a smaller one may not be able to effectively separate diverse signal morphologies of raw heartbeats, and a larger one may over partition the raw heartbeats which makes multiple medoids represent similar numbers of objects and lower the robustness of the selected template.

To solve the K-medoids problem, a partitioning around medoids (PAM) method is chosen [90], which uses a greedy search to iteratively evaluate whether the swapping of each medoid  $c_j$  and each non-medoid  $\widehat{\theta}_j^1$  can decrease the total object-to-medoid dissimilarity  $\xi$  as (4). If yes, then update the medoid until no dissimilarity reduction can be achieved.

$$c_j = \text{Select}(\widehat{\theta}_j^1 | w_{\widehat{\theta}_j^1}, \forall \widehat{\theta}_j^1 \in \widehat{\Theta}^1) \quad (2)$$

$$w_{\widehat{\theta}_j^1} = \frac{d^2(\widehat{\theta}_j^1, c_p)}{\sum_{\{h | \widehat{\theta}_h^1 \in D_p\}} d^2(\widehat{\theta}_h^1, c_p)}, \forall \widehat{\theta}_j^1 \in \widehat{\Theta}^1 \quad (3)$$

$$\xi = \sum_{\{c_p | p=1, \dots, K\}} \sum_{\{h | \widehat{\theta}_h^1 \in D_p\}} d^2(\widehat{\theta}_h^1, c_p) \quad (4)$$

### 3) DTW-based Signal Distortion Evaluation

The DTW approach is introduced to screen all raw PPG heartbeats by the self-learned high quality PPG template to calculate the raw heartbeat-specific distortion value. DTW is a popular pattern learning technique firstly used in speech recognition applications and is seen as a branch of machine learning techniques.

Compared with the Euclidean distance-based method, DTW can more effectively measure the dissimilarity between two time series with time-varying morphologies and lengths [85], which is always the case when processing dynamic signals such as PPG heartbeats. DTW has an ability to warp the testing sequence in a non-linear approach to measure its distance from a pre-defined template sequence. In our study, we use this DTW distance to quantify the degree of distortion for raw PPG heartbeats, as shown in step 3) ‘HB distortion’ in Algorithm 1.

To measure the dissimilarity between a PPG heartbeat template  $\Gamma = \{y_l | l = 0, \dots, L - 1\}$  and a raw PPG heartbeat  $\theta_j^t = \{x_i^{t,j} | i = 0, \dots, N_j^t - 1\}$ , the DTW performs two steps of operation, i.e., the local distance matrix generation, and the path distance matrix generation, respectively. In the former step as (5), a  $N_j^t \times L$  local distance table is created and its each element  $d_{i,l}^{t,j}$  (local distance) equals to the Euclidean distance between sample  $x_i^{t,j}$  in  $\theta_j^t$  and sample  $y_l$  in  $\Gamma$ . This

table gives all possible sample-to-sample dissimilarities between  $\theta_j^t$  and  $\Gamma$ . In the latter step as (6), a path distance table is created by a dynamic programming strategy, with the element  $D_{i,l}^{t,j}$  (accumulated path distance) equaling to the current local distance  $d_{i,l}^{t,j}$  plus the minima of three preceding accumulated path distances.

In such a way, many warping paths are generated which all try to minimize the accumulated path distance, i.e., to find an optimal path to match the testing sequence to the template. Therefore, the last element  $D_{N_j^t,L}^{t,j}$  can effectively represent the minimum accumulated sequence-to-sequence distance, i.e., the sequence level dissimilarity value, which is used as the raw heartbeat-specific distortion level as (7).

Consequently, we can get quantified distortion values for all of the raw PPG heartbeats in each trail, denoted as  $\Phi^t$ , and for all trials, denoted as  $\Psi$ , as shown in step 3) ‘HB distortion’ in Algorithm 1.

$$d_{i,l}^{t,j} = |x_i^{t,j} - y_l|, \quad \forall i, \forall l \quad (5)$$

$$D_{i,l}^{t,j} = \begin{cases} d_{i,l}^{t,j} + \min \begin{cases} D_{i-1,l}^{t,j} \\ D_{i-1,l-1}^{t,j} \\ D_{i,l-1}^{t,j} \end{cases} & \forall i > 0 \ \& \ \forall l > 0 \\ d_{i,l}^{t,j} & i = 0 \ \& \ l = 0 \\ \inf & \text{otherwise} \end{cases} \quad (6)$$

$$\phi_j^t = D_{N_j^t,L}^{t,j} \quad (7)$$

#### 4) Histogram Triangle-based Distortion Threshold Learning

Based on quantified distortion evaluation of the raw PPG heartbeats, the next is to learn an appropriate distortion threshold to differentiate heartbeats with a good or a poor quality. The same consideration used in K-medoids clustering-based template learning is applied here, i.e., low quality raw heartbeats owning much more diverse distorted morphologies due to random motion artifacts. Therefore, statistically, in a distortion histogram, the raw heartbeats with a relatively good quality should concentrate in the low distortion area (the left side of the x-axis), while the ones with gradually worse signal quality conditions usually spread over the higher distortion area (the middle and right side of the x-axis), due to poor consistency induced by random motion artifacts. Leveraging this interesting *left-skewed* histogram, we use an unsupervised learning approach called histogram triangle method [87] to learn the distortion threshold.

Based on a normalized density histogram of the distortion values, the maximum point  $(b_{max}, his(b_{max}))$  is firstly determined, which corresponds to the distortion bin most frequently hit by good quality PPG heartbeats in the low distortion area, where  $b_{max}$  is the histogram bin and  $his(\cdot)$  is a function returning the density value for a given bin. Then a histogram hypotenuse is constructed by connecting the maximum point  $(b_{max}, his(b_{max}))$  and the right boundary of the histogram envelope  $(b_{right}, his(b_{right}))$ , and denote this hypotenuse as  $b_{max} \sim b_{right}$ . A threshold  $Th_{01}^{his}$  is learned as the bin corresponding to a maximum perpendicular distance from this bin to the hypotenuse as (8). After de-normalization based on the maxima and minima of  $\Psi$  (a set of distortion values in all trials in the training session), we get the threshold in the original scale  $Th_{ori}^{his}$  as (9), and after shrinking it by a factor of  $\delta$  (50%) to enhance the robustness, we obtain the final threshold  $Th^{his}$  (10).

$$Th_{01}^{his} = \underset{b_{max} \leq b \leq b_{right}}{\operatorname{argmax}} \operatorname{Dis}\{(b, his(b)), b_{max} \sim b_{right}\} \quad (8)$$

$$Th_{ori}^{his} = Th_{01}^{his} (\max(\Psi) - \min(\Psi)) + \min(\Psi) \quad (9)$$

$$Th^{his} = \delta Th_{ori}^{his} \quad (10)$$

### 5) Distortion Curve Smoothing

To further enhance the robustness before separating the raw PPG heartbeats to binary groups with a good or poor signal quality levels, the raw PPG heartbeats distortion values are smoothed by a 10th order moving average method ( $S = 10$  in Algorithm 1). This is based on the consideration that when some raw PPG heartbeats own high distortion values, they are either real heartbeats highly corrupted by severe motion artifacts, or motion artifacts-induced interferential spikes. Therefore, their neighboring raw PPG heartbeats with lower distortion values may also have a high possibility to be impacted by motion artifacts. The smoothing operation can elevate the low distortion values for these neighboring heartbeats, and help cluster raw heartbeats in suspicious time periods to the low signal quality group more strictly to further enhance the robustness.

### 6) SQI Generation

Based on the learned distortion threshold and the smoothed distortion curve, the raw PPG heartbeats can now be clustered to binary groups with a good or poor quality level. The proposed SQI generation algorithm is shown in Algorithm 2, where two prudential considerations are made to further enhance the robustness. Firstly, although the smoothed distortion curve  $\Omega^t$  in the  $t - th$  trial can help elevate low distortion values when they are close to high distortion values (i.e., suspicious time periods), the smoothing operation usually lowers the high distortion values at the

same time. It means that the unsmoothed distortion curve  $\Phi^t$  can still contribute to highlight the heartbeats with high distortion values. Therefore, we compare not only the smoothed distortion curve but also the unsmoothed one to the learned distortion threshold  $Th^{his}$  for the SQI set generation (step 2 in Algorithm 2).

Secondly, the motion artifacts due to time-varying electrode-skin contact variations are so random that it is impractical to cover all motion artifacts scenarios in the training session. If there happen to be some severe motion artifacts resulting in very high distortion values in the testing session, they may over-elevate many low distortion values in the corresponding suspicious time periods. Consequently, this strict SQI generation procedure may filter out too many raw heartbeats in some trials.

However, based on our observation, even after aggressively introducing twenty-second head movements-induced motion artifacts in each trial (the subjects are usually asked to stay during estimation, but we aggressively asked them to perform movements for one third of each trail time), the heartbeats corrupted are still lower than fifty percent. Leveraging this observation, we introduce a heartbeats protection strategy to protect the best  $\rho$  (20%) heartbeats in each trail, by adaptively increasing the learned threshold  $Th^{his}$  with a step size of  $\eta$  (1%) until at least  $\rho$  heartbeats are labeled with a good quality level (step 3 in Algorithm 2).. To guarantee the consistency of the proposed SQI generation algorithm, this protection operation is also applied to the training session.

The BP estimate quality indicator is also reported to reflect the percentage of raw heartbeats left after step 2 but before heartbeat protection. The indicator can be used to select out high confident BP estimates according to specific application requirements.

After strictly labelling low quality raw PPG heartbeats and performing necessary heartbeats protection operations, the generated SQI set based on raw PPG heartbeats can now be used for heartbeats purification.

<b>Algorithm 2</b> SQI generation: $Q^t \leftarrow SQI(\Phi^t, \Omega^t, Th^{his}, \rho, \eta)$	
<b>Input:</b>	
unsmoothed distortion values in the $t - th$ trial	$\Phi^t = \{\phi_j^t   \forall j\}$
smoothed distortion values in the $t - th$ trial	$\Omega^t = \{\pi_j^t   \forall j\}$
leaned distortion threshold	$Th^{his}$
lower limit of percentage of HBs to be protected	$\rho$
step size in percentage to adaptively adjust the distortion threshold	$\eta$
<b>Output:</b>	
heartbeat signal quality indices in the $t - th$ trial	$Q^t = \{q_j^t   \forall j\}$
quality indicator for the $t - th$ BP estimate	$BP\_QI^t$
<b>Procedure:</b>	
<b>Step 1 – initialize the parameters</b>	
$P \leftarrow \rho \ \Phi^t\ $	▷ # raw heartbeats protected
$\gamma^t \leftarrow Th^{his}$	▷ initialize the adaptive threshold
$\Delta \leftarrow \eta \max\left(\max_{\forall j} \phi_j^t, \max_{\forall j} \pi_j^t\right)$	▷ step size for threshold adjusting
<b>Step 2 – generate the SQI set</b>	
$c \leftarrow 0$	
<b>for</b> $j = 1$ to $\ \Phi^t\ $ <b>do</b>	▷ sweep raw heartbeats
<b>if</b> $\phi_j^t \geq \gamma^t$ <b>and</b> $\pi_j^t \geq \gamma^t$ <b>then</b>	▷ check two conditions
$q_j^t \leftarrow 1$	▷ a good quality heartbeat
$c \leftarrow c + 1$	
<b>else</b>	
$q_j^t \leftarrow 0$	▷ a poor quality heartbeat
<b>end if</b>	
<b>end for</b>	
$BP\_QI^t \leftarrow c / \ \Phi^t\ $	▷ BP quality indicator
<b>Step 3 – adaptively adjust the threshold to protect the best <math>P</math> heartbeats if necessary</b>	
<b>if</b> $c < P$	▷ need to protect the best $P$ heartbeats
$\gamma^t = \gamma^t + \Delta$	▷ increase the threshold
<b>go to Step 2</b>	▷ re-generate the SQI set
<b>else</b>	
<b>stop</b>	
<b>end if</b>	

## 7) PPG and ECG Heartbeats Purification

Considering there are still many residual highly corrupted and faking heartbeats, both raw ECG and PPG heartbeats are purified according to the SQI information, i.e., filtering raw heartbeats with an SQI of 0 and keeping those with an SQI of 1. The purified heartbeats are then sent to the stage III of the proposed HR and SBP estimation algorithm.

### 5.2.6 HR Estimation and Supervised Learning of SBP Estimation

Based on the purified ECG heartbeats, the HR estimates can now be achieved, and together with purified PPG heartbeats, the PTT can also be measured. Afterwards, the SBP model can be firstly calibrated in the training session by a supervised learning process referring to the left arm cuff-based ground truth SBP, and then used for SBP estimation on the unseen data acquired in the testing session.

#### 1) Heart Rate Estimation

As mentioned above, the ECG signal is of a relatively better motion artifacts-tolerant ability than the PPG signal, therefore, the purified ECG heartbeats are used for instantaneous heart rate estimation. Then the windowed heart rate (denoted as HR, with a unit of beats-per-minute, denoted as BPM) estimates can be achieved, where the window corresponds to the second minute in each two-minute trial during which the SBP<sub>cuff</sub> is measured. The performance of the estimated HR will be evaluated in terms of mean error  $\pm$  standard deviation ( $ME \pm STD$ ), mean absolute error (MAE) and root mean absolute error (RMSE).

#### 2) Pulse Transit Time Calculation

Pulse transit time is the time consumed by the pressure pulse to flow from the proximal (PTT start time) to the distal (PTT end time) arterials sites. As shown in Figure 5.3, the ECG R peak

represents when the pulse leaves the proximal site, i.e., the thoracic aorta, and the PPG foot corresponds to when the pulse arrives the distal site, i.e., the skin surface where the PPG sensor is placed on. The measured instantaneous  $PTT_i$ , i.e., the PTT value for the  $i$ -th ECG/PPG heartbeat pair is calculated as (11), where  $PPG_i^{foot}$  is the PPG foot occurrence time and  $ECG_i^{Rpeak}$  corresponds to the ECG R peak occurrence time. Similar to windowed HR, the instantaneous PTT measured in the second minute of each trial is also averaged to obtain the windowed PTT estimates.

$$PTT_i = PPG_i^{foot} - ECG_i^{Rpeak} \quad (11)$$

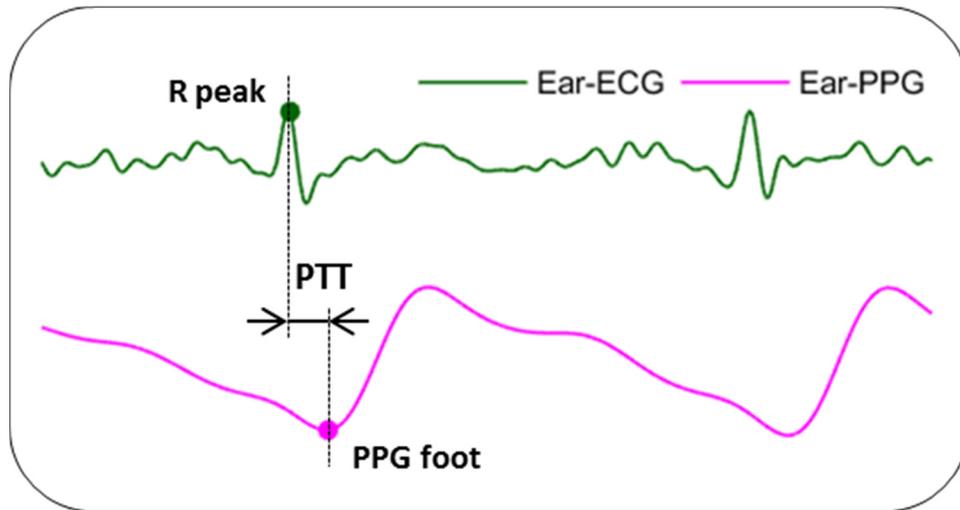


Figure 5.3. Pulse transit time (PTT).  
Note. This illustration of PTT is based on ear-ECG/PPG signals.

### 3) Blood Pressure Model Learning and Testing

Due to the complicated underlying blood pressure wave generation and propagation mechanisms, many SBP learning models have been reported based on diverse assumptions and strategies. To thoroughly compare them and determine an appropriate one for ear application scenarios, ten

popular SBP learning models are taken into account in our study as shown in Table 1, including seven PTT-SBP models, and three PTT&HR-SBP models with HR information enhanced.

Among PTT-SBP models 1 to 7, various styles of equations are applied, such as linear, quadratic, exponential ones and so on, based on different deduction processes. For example, the model 2 reflects the reverse correlation between PTT and SBP shown by large amounts of studies, based on the fact that a high SBP will reduce the time consumed by the pressure pulse to propagate from the proximal to the distal sites, and vice versa [58]. The model 7 is based on the combined action of the pulse wave and the energy of wave.

Table 5.1. Blood pressure estimation methods.

No.	Equation	HR information
1	$SBP = a \ln(PTT) + b$	w/o
2	$SBP = a PTT^{-1} + b$	w/o
3	$SBP = a PTT + b$	w/o
4	$SBP = a PTT^2 + b PTT + c$	w/o
5	$SBP = a PTT^2 + b$	w/o
6	$SBP = a e^{b PTT}$	w/o
7	$SBP = a PTT^{-2} + b$	w/o
8	$SBP = a PTT^{-2} + b HR^{-2} + c$	w/
9	$SBP = a \ln(PTT) + b \ln(PTT) + c$	w/
10	$SBP = a PTT + b HR + c$	w/

Among PTT&HR-SBP models 8-10, the HR information is introduced to model establishment. They are based on the consideration that when HR increases, the cardiac output flow usually increased at the same time which causes a higher SBP, and vice versa. One thing worth noting is that, for simplicity and convenience purpose, the PPT measurement method introduced above actually includes another extra item, i.e., the pre-ejection period (PEP). PEP corresponds to the aortic valve opening time and usually significantly increase the PTT measured. To measure PEP, some studies acquire an additional signal, such as the phonocardiogram (PCG) or the impedance cardiography (ICG), however, it usually results in extra hardware burden and a lower wearability. Therefore, in our study, to further enhance the PTT-SBP model, we only take into account the HR information which has been proved to be correlated with SBP and is already carried by the ECG signal without requiring an additional signal acquisition hardware. The SBP models are learned on the training data and tested on the unseen testing data to show the generalization ability. The left-arm cuff-based SBP is used as reference to enable a supervised learning process. The performance is reported in terms of  $ME \pm STD$ , MAE and RMSE.

## **5.3 RESULTS AND DISCUSSION**

In this section, both the proposed hardware prototype and the HR/SBP estimation algorithms are evaluated in detail, according to the signal processing flow shown in Figure 5.1.

### **5.3.1 Signals Acquired**

After situating the ECG electrodes behind two ears and the PPG sensor behind the left ear, our semi-customized bio-potential acquisition platform successfully collected the weak ear-ECG (Figure 5.4c-d) and ear-PPG (Figure 5.4b) signals, where the chest-ECG signal (Figure 5.4a) is

also given for comparison purpose. The acquired ear-ECG signal is only around 5% of the chest-ECG signal in terms of peak-to-peak voltage, resulting from a much smaller potential difference between the back locations of two ears. Although the ear-ECG is highly weak, it can still show distinguishable heartbeats, especially clear QRS complex morphologies, even with continuous background motion artifacts due to uncontrolled neck muscle and blood vessels movements, indicating the effectiveness of the proposed non-standard highly convenient single lead ECG configuration. Meanwhile, the acquired PPG signal also owns a clear heartbeat morphology leveraging many blood vessels around the back location of the ear.

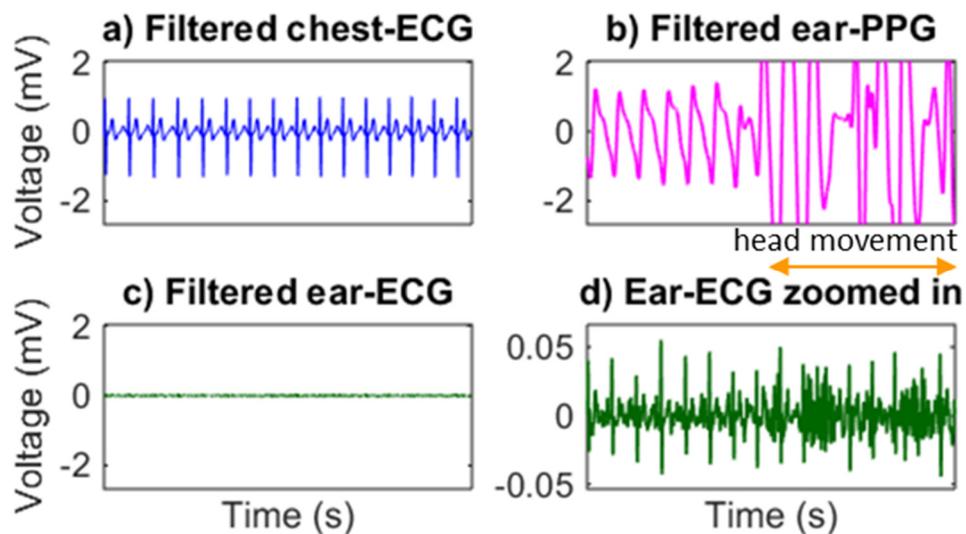


Figure 5.4. An example of the ear signal segments acquired.

Notes. The comparison shows that the weak ear-ECG has a peak-to-peak voltage only around 5% of that of the chest-ECG.

When performing head movements, many motion artifacts are induced to both ear-ECG and ear-PPG signals (Figure 5.4d and 4b) which make heartbeats identification highly challenging. Therefore, advanced signal processing and machine learning algorithms for robust heartbeat recognition are proposed to enable this highly wearable ear signal acquisition solution.

### 5.3.2 Heartbeat Identification

To identify raw heartbeats from motion artifacts-impacted or even corrupted ear signals, our previously proposed SVM-based approach is applied on the weak ear-ECG signal. Two examples of heartbeat identification results in the *testing session* of subject 1 are given in Figure 5.5, where the raw ECG heartbeats are firstly identified and then the PPG heartbeats are determined by a simple minima searching method. There are several interesting observations as follows to support why we firstly identify the raw heartbeats from the ECG signal, and why we need to learn signal quality labelling based on the PPG signal in an unsupervised manner and purify both ECG/PPG heartbeats.

Firstly, the ECG signal acquired with the non-standard signal lead configuration is so weak that it is continuously impacted by the background motion artifacts due to uncontrolled neck muscle and blood vessels movements, as shown in Figure 5.5. Specially, in the signal periods not covered by the wide orange rectangles, the PPG signal owns a better signal morphology compared with the ECG signal, which is even highly corrupted by the background motion artifacts such as the segment in the narrow orange rectangle in Figure 5.5a). However, as shown in the wide rectangle (corresponding to deliberately introduced head movements) in Figure 5.5a, the ECG signal actually owns a relatively better motion artifacts tolerant ability than the PPG signal leveraging sharply changing QRS complexes, although the ECG signal is weaker than the PPG signal. Therefore, we firstly perform raw heartbeat identification from the ECG signal.

Secondly, the PPG signal is highly distorted in the wide orange rectangles due to deliberately introduced motion artifacts by head movements, as shown in wide rectangles in Figure 5.5a-b. Since PPG heartbeat morphological characteristics (PPG feet) during these signal periods cannot

reflect the heartbeat occurrence time for PTT calculation, we need to filter out these signal segments according to the signal quality of these raw PPG heartbeats.

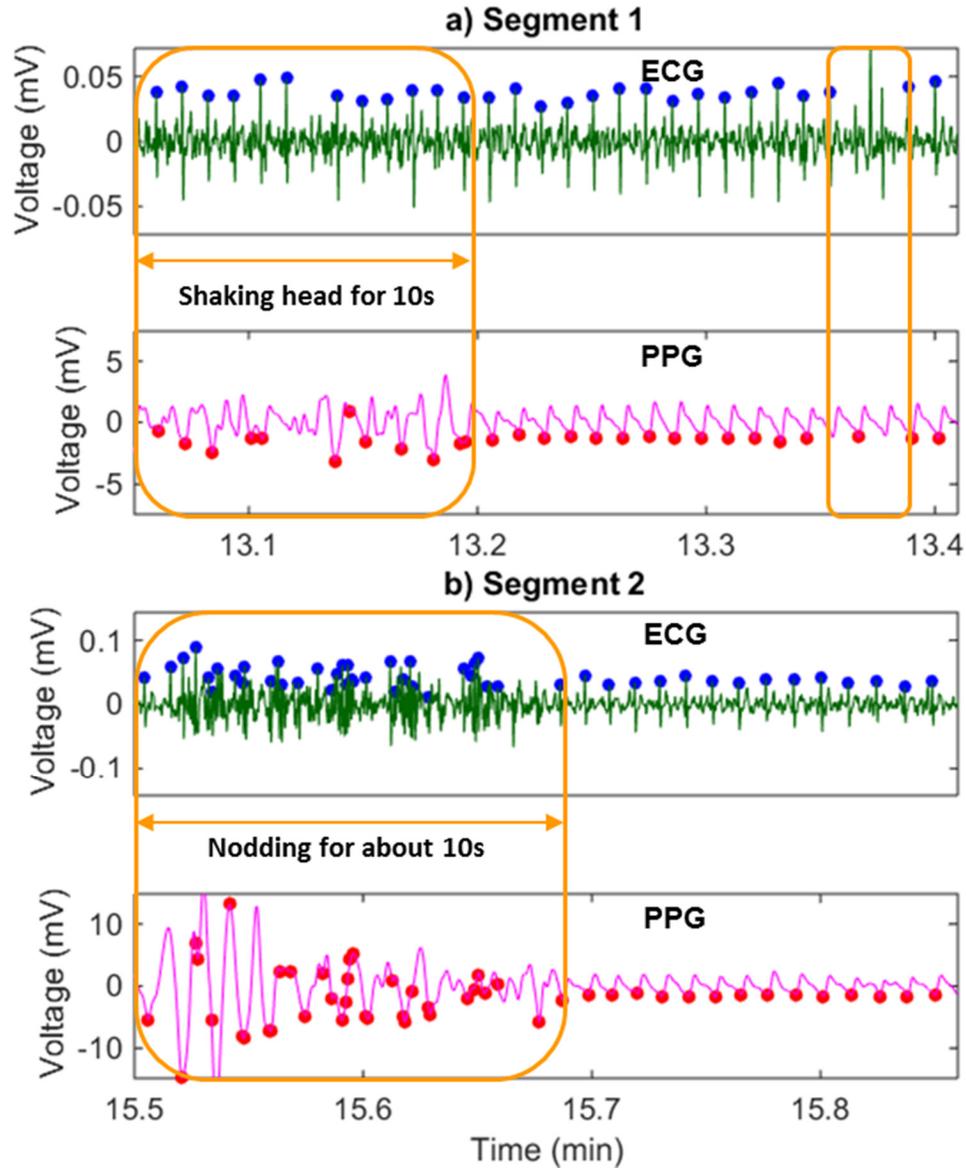


Figure 5.5. Two examples of heartbeat identification results.  
 Notes. Results correspond to the testing session of the subject 1. Blue dots: identified raw ECG heartbeat locations; red dots: identified raw PPG heartbeat locations; wide orange rectangles: signal periods with deliberately introduced motion artifacts due to head movements; narrow orange rectangle: signal period with missing or fake heartbeats due to severe background motion artifacts; all the weak ECG signal is continuously impacted by background motion artifacts.

However, learning how to label the signal quality of the raw PPG heartbeats is not trivial since it is hard to generate ground truth signal quality labels as reference in the *training session* to enable a supervised learning. This is because (using the testing data in Figure 5.5 as an example here): 1) the background motion artifacts are highly random; 2) the start and end time of head movements cannot be precisely controlled due to subject-dependent command response delay as shown in the wide rectangle in segment 2 (Figure 5.5b); 3) the non-linear distortion behaviors of raw PPG heartbeats is so random that it is difficult to manually generate signal quality labels. Based on these considerations, we choose an unsupervised learning approach to learn how to generate the SQI information for each raw PPG heartbeat. One thing worth noting is that the SQI information is also used to purify the ECG heartbeats which may also be corrupted at the same time when the PPG heartbeats are severely corrupted as shown in the wide rectangle in Figure 5.5b.

### 5.3.3 PPG Segmentation and Template Learning

To perform unsupervised learning of signal quality labelling, we apply a DTW method to quantify the degree of distortion for each raw PPG heartbeat. The PPG template for the DTW method is learned by the K-medoids clustering approach on the segmented raw PPG heartbeats. The clustering results are given in Figure 5.6, where raw heartbeats with a relatively good quality concentrate in the M1 cluster, and raw heartbeats with a poor quality are grouped into many other clusters due to the high randomness induced by motion artifacts. Consequently, the medoid in cluster 12 which represents a highest number of instances ( $\# = 65\%$ , i.e., 65% of raw heartbeats in the second minute in the first trial are grouped into cluster 12) is selected as the high quality PPG heartbeat template. One interesting observation is that some slightly distorted heartbeats are also grouped into this cluster since the other clusters correspond to raw heartbeats so randomly

corrupted by severe motion artifacts due to head movements. The high quality PPG template can still be well learned, because the K-medoids clustering algorithm makes each medoid represent the majority of instances in each cluster, i.e., minimizing the object-to-medoid dissimilarity  $\xi$  as (4).

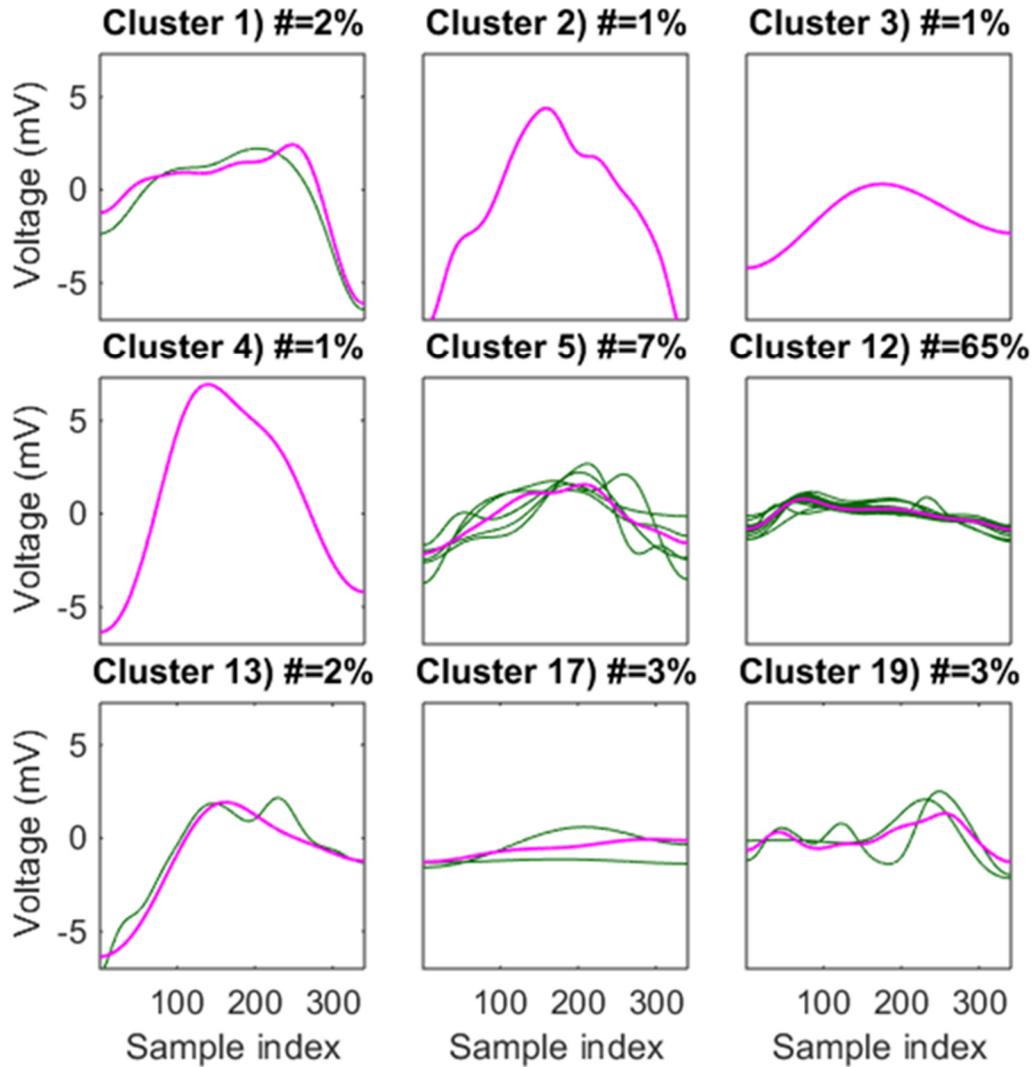


Figure 5.6. K-medoids clustering-based PPG template learning.

Notes. Results correspond to the first trial in the training session of subject 1, with K equals to 20 (only top 9 clusters are visualized here). Pink curve: medoid of each cluster; green curves: instances represented by the medoids; #=: percentage of instances in current cluster.

### 5.3.4 PPG Distortion Evaluation and Threshold Learning

After quantifying the degree of distortion for all raw PPG heartbeats using the DTW method, the histogram triangle-based approach is proposed for PPG distortion threshold learning which will be used to generate the SQI information. An example is given in Figure 5.7, where a skewed intensity histogram of the DTW distances is constructed. The relatively good quality heartbeats concentrate in the low distortion area (the left side of the x-axis) and poor quality heartbeats spread over a larger range resulting from high and diverse distortion values due to random and severe motion artifacts. The global searching process effectively captures the transition point of the histogram and determines the normalized threshold  $Th_{01}^{his}$  in this example as 0.07, which is then de-normalized and multiplied by a shrinkage factor to get the final distortion threshold  $Th^{his}$  which is 11.9 in this example.

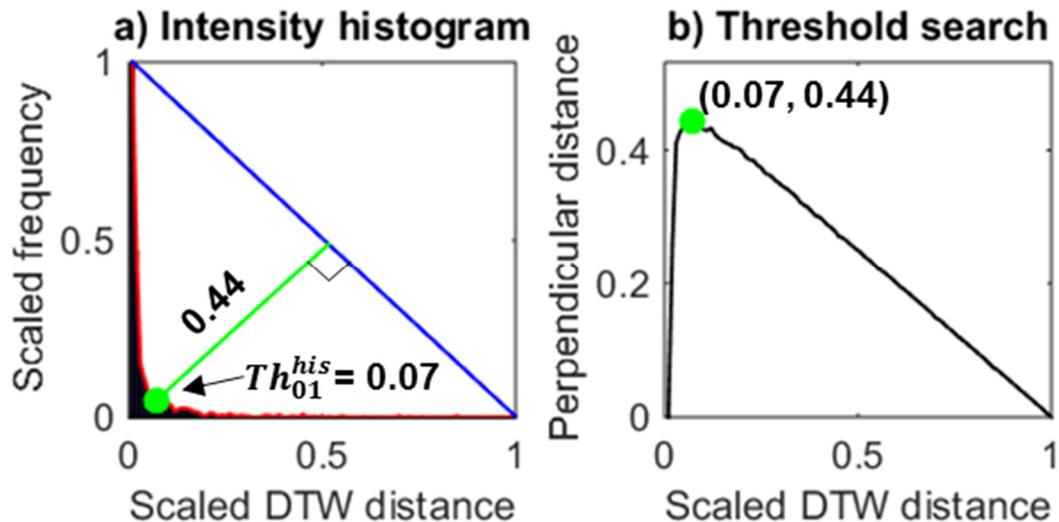


Figure 5.7. PPG distortion threshold learning.

Notes. Results correspond to the first trial in the training session of subject 1. Blue line: the histogram hypotenuse; red curve: the histogram envelope; green line: the maximum perpendicular distance; green dot: the learned normalized (0-1 range) threshold.

### 5.3.5 SQI Generation and Heartbeats Purification

Based on the DTW distance-based distortion values and the learned distortion threshold, we now can generate the SQI information for all PPG raw heartbeats in the training or the testing session. An example of the whole SQI generation process in the testing session of subject 1 is given in Figure 5.8. The acquired thirty-minute ear-ECG and ear-PPG streams are shown in Figure 5.8a and 8b, where fifteen pink segments in each stream corresponding to the second minute in each of fifteen trials.

In each pink segment, there is head shaking movement during the first ten seconds and nodding movement during the fourth ten seconds, resulting in many severe motion artifacts which increase the peak-to-peak voltage. In the last eight black segments, there are exercise-induced signal variations (riding the bike), especially in the ear-PPG stream. In the first seven black segments, there are also some signal variations due to normal head movements. The calculated unsmoothed DTW distance and the smoothed one are given in Figure 5.8c, which shows diverse degree of distortion caused by both background and deliberately introduced motion artifacts. Based on SQI algorithm proposed, the distortion threshold is adaptively elevated if the best  $\rho$  heartbeats need to be protected. Finally, the raw PPG heartbeat-specific SQI information is generated as Figure 5.8d.

Two same segments as those in Figure 5.5 are used here to further illustrate the details of the SQI generation process, as shown in Figure 5.8e and 8f. During deliberately introduced head movements (wide orange rectangles), the DTW distance-based distortion values (black bars) are much higher than those in other time periods. The smoothing operation of the distortion values can make the low distortion values above the distortion threshold and thus pose a more critical distortion evaluation during these highly suspicious periods (wide orange rectangles).

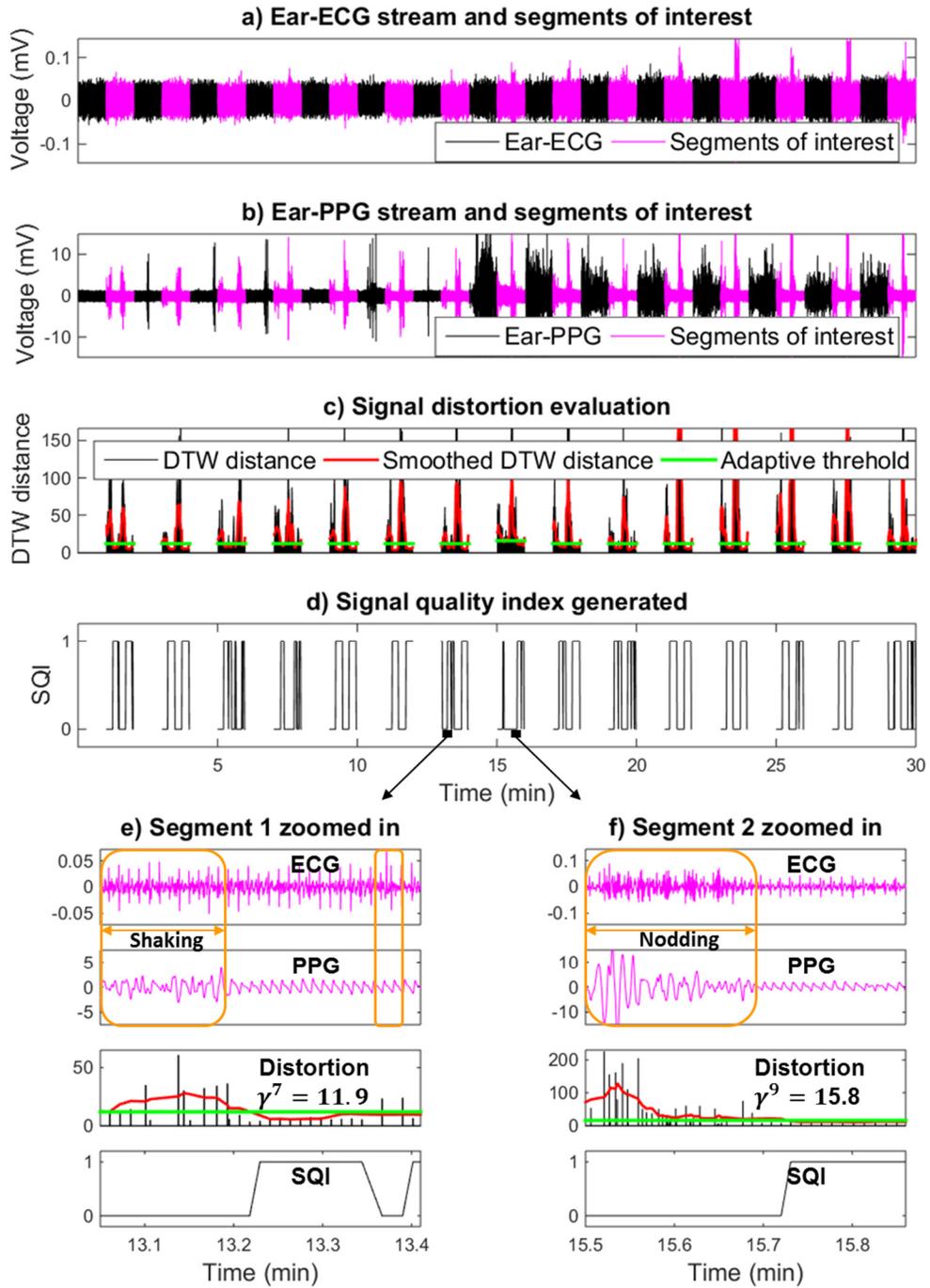


Figure 5.8. The whole SQI generation process.

Notes. Results correspond to the testing session of subject 1, including the signals acquired, quantified degree of distortion for raw PPG heartbeats, the adaptive distortion threshold and the SQI sequence generated.

On the other hand, the unsmoothed distortion values can still highlight the heartbeats with a high distortion condition (around minute 13.37). Besides, the motion artifacts are so random that it is difficult to include all motion artifacts scenarios in the training session.

If there happen to be some highly random motion artifacts shown as in Figure 5.8f, the corrupted PPG morphologies may induce dramatically high distortion values which generate a large range of suspicious period. This shows the necessity to introduce a protection mechanism to adaptively elevate the distortion threshold for some trials to protect the best  $\rho$  (20%) heartbeats (e.g., threshold  $\gamma^9 = 15.8$  in trial 9 is adaptively adjusted to be higher than threshold  $\gamma^7 = Th^{his} = 11.9$  in trial 7).

Based on the proposed SQI generation algorithm, the raw heartbeats are labelled as accepted (SQI = 1) or rejected (SQI = 0), which helps filter out many signal segments highly corrupted by severe motion artifacts such that the remaining purified heartbeats can be used in HR and PTT estimation later.

### 5.3.6 Heart Rate Estimation

The windowed heart rate estimates are achieved based on the ECG heartbeats. A Bland-Altman plot for estimated and reference HR is given in Figure 5.9 to illustrate the HR estimation performance.

It shows that most of the HR estimates concentrate in the low error area, indicating the potential of using ear-ECG for robust long-term HR monitoring applications. Averaged over the acquired ear signal dataset, the  $ME \pm STD$ , MAE and RMSE of HR estimation are  $0.8 \pm 2.7$ , 1.8 and 2.8 BPM, respectively.

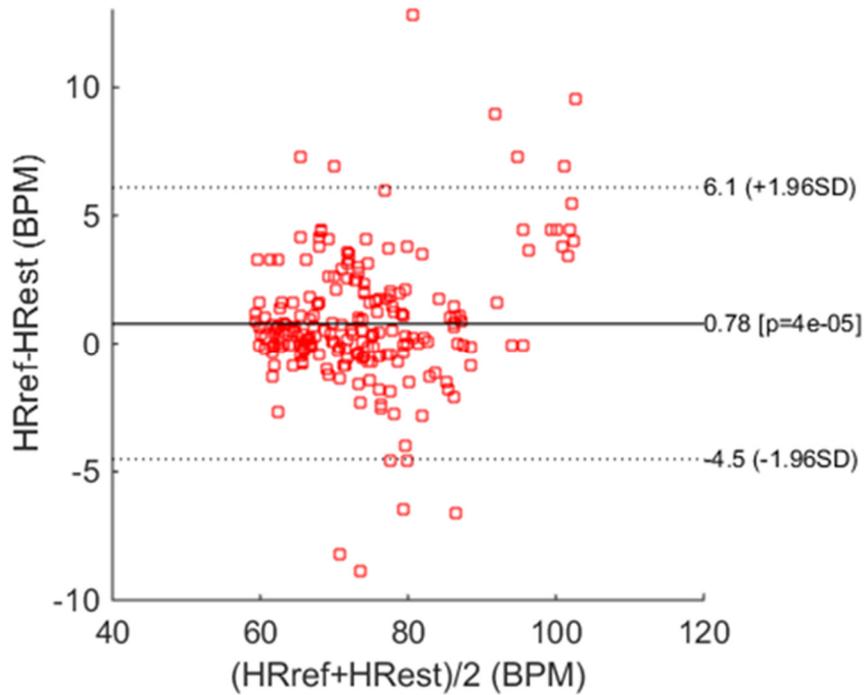


Figure 5.9. Bland-Altman plot for estimated HR (HRest) and reference HR (HRref).

### 5.3.7 Blood Pressure Estimation

Based on HR and PTT estimates, ten diverse SBP models including seven PTT-SBP models (1-7) and three PTT&HR-SBP models (8-10) are thoroughly compared to explore their abilities in SBP estimation. According to the *Advancement of Medical Instrumentation (AAMI) standard* [74], the BP estimation error should be less than  $5.0 \pm 8.0$  mmHg in terms of mean error (ME)  $\pm$  standard deviation (SD). To thoroughly evaluate the SBP estimation performance, we consider four different criteria including ME, STD, MAE and RMSE. Moreover, although many wearable BP estimation studies only reported the performance on the training data [58, 59], we test the proposed algorithm on the unseen testing data to emphasize the generalization ability of the SBP models.

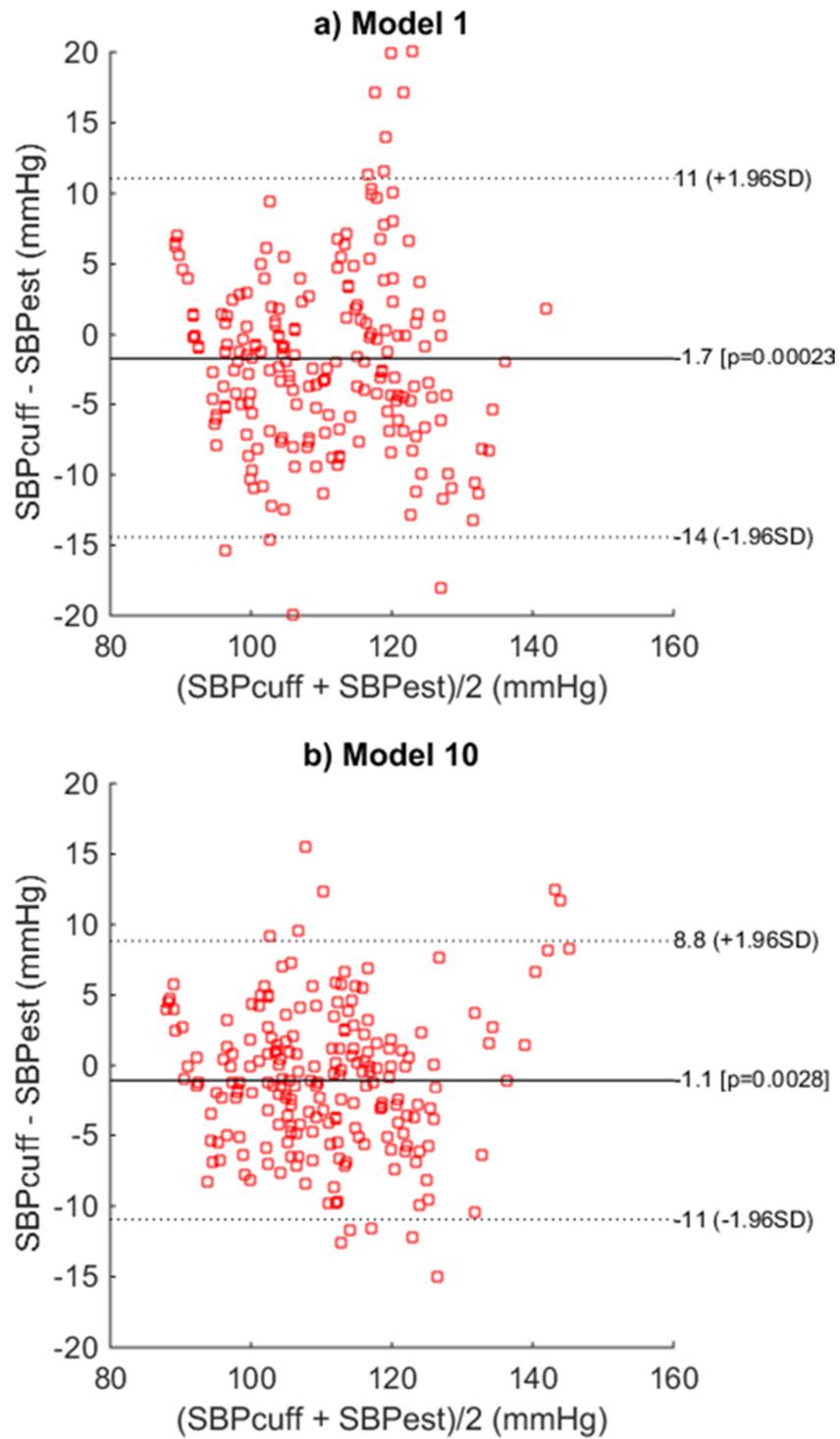


Figure 5.10. Bland-Altman plots for SBP model 1 and 10.

Table 5.2. Blood pressure estimation performance summary.

<b>No.</b>	<b>ME</b>	<b>STD</b>	<b>MAE</b>	<b>RMSE</b>
1	-1.7	6.5	5.2	6.7
2	-1.6	6.6	5.3	6.8
3	-1.8	6.1	5.0	6.4
4	-2.1	6.6	5.4	6.9
5	-1.9	6.4	5.2	6.7
6	-1.8	6.1	5.0	6.3
7	-2.0	6.3	5.2	6.5
<b>8</b>	<b>-1.4</b>	<b>5.2</b>	<b>4.2</b>	<b>5.4</b>
<b>9</b>	<b>-1.2</b>	<b>5.0</b>	<b>4.1</b>	<b>5.2</b>
<b>10</b>	<b>-1.1</b>	<b>5.1</b>	<b>4.0</b>	<b>5.1</b>

The performance summary of ten SBP models on the unseen testing data is summarized in Table 2. Leveraging the additionally introduced HR information, model 8-10 effectively outperform model 1-7, with the  $ME \pm STD$ , MAE and RMSE no more than  $-1.4 \pm 5.2$ , 4.2 and 5.4 mmHg, respectively. For example, the  $ME \pm STD$ , MAE and RMSE for model 1 are  $-1.7 \pm 6.5$ , 5.2 and 6.7 mmHg, respectively, while for model 10, they are  $-1.1 \pm 5.1$ , 4.0 and 5.1 mmHg, respectively.

To further illustrate the performance difference between the PTT-SBP models and the PTT&HR-SBP models, the Bland-Altman plots for model 1 and 10 are given in Figure 5.10. The latter one owns a smaller mean error and a more concentrated distribution (a smaller standard deviation) compared to the former one. This comparison shows that the PTT-SBP model can be

effectively enhanced by the HR information which is also correlated with SBP, yielding a more robust PTT&HR-SBP model.

## 5.4 CONCLUSION

In this study, we propose a machine learning-enabled highly wearable and motion-tolerant system for long-term cuff-less blood pressure and heart rate monitoring from ECG and PPG signals. There are several key contributions in this study. Firstly, we proposed a sensor placement approach to meet the critical requirements on the wearability and comfortableness, which allows for situating all sensors behind two ears and the possibility to integrate them into glasses or ear headsets. Secondly, we successfully acquired the weak ear-ECG and ear-PPG signals with our semi-customized bio-potential acquisition platform, and explored diverse background and head movements-induced motion artifacts towards practical application scenarios. Thirdly, based on the identified raw heartbeats, we proposed an unsupervised learning algorithm to automatically filter out many distorted or faking raw heartbeats. The DTW approach is introduced to quantify the degree of distortion for raw heartbeats, referring to a high quality heartbeat template, which are then compared with a distortion threshold to generate SQI information for heartbeat purification. The high quality heartbeat template and the distortion threshold are learned using the K-medoids clustering approach and the histogram triangle method, respectively. Fourthly, the HR estimates were achieved for both direct heart health management and SBP model enhancement purposes. Fifthly, we demonstrated that the PTT&HR-SBP models are superior to PTT-SBP models.

Based on an acquired ear signal dataset, the  $ME \pm STD$ , MAE and RMSE for HR estimation are  $0.8 \pm 2.7$ , 1.8 and 2.8 BPM, respectively, and for PTT&HR-SBP models, they are no more than -

1.4±5.2, 4.2 and 5.4 mmHg, respectively. The results demonstrate the feasibility of the proof-of-concept system in wearable ear-ECG/PPG acquisition and robust BP/HR estimation, which is expected to contribute to pervasive hypertension, heart health and fitness management. In future, we will acquire data from more subjects, and also further introduce motion artifacts from more scenarios, such as walking, running, sleeping, eating, etc.

## CHAPTER 6

### CONCLUSION AND FUTURE WORK

This dissertation work focused on the long-term high-resolution heart rate and blood pressure monitoring using wearable computers, for heart diseases and hypertension management, and wellness and fitness evaluation.

In long-term high-resolution heart rate monitoring, we proposed two motion artifacts-tolerant heart rate tracking frameworks, which can effectively identify the heartbeats impacted or even corrupted by large amounts of motion artifacts, and then achieve robust heart rate estimates. The phase-domain dynamic time warping framework firstly transforms the signal to the phase domain, and then introduces the multi-view dynamic time warping to find out heartbeats similar to a heartbeat template. The two-stage classification framework includes an SVM-based heartbeat identification stage and a refinement stage. The former one extracts many features in several different domains, selects out the critical ones tolerant to motion artifacts, and then builds an SVM model to perform heartbeat identification. The latter one suppresses the false positives and interpolates missing heartbeats to further enhance the robustness. These frameworks own superior performance compared with previously reported approaches, and are expected to contribute to pervasive heart health and fitness management.

In long-term high-resolution blood pressure monitoring, we proposed a single-arm-worn cuff-less system which can successfully estimate the blood pressure only based on single-arm physiological signals. The weak arm signals are processed by advanced signal processing and machine learning algorithms, to generate the pulse transit time and heart rate information to build

the blood pressure models. Experimental results show that this system can effectively estimate the minute-level blood pressure only based on single-arm signals. We also proposed an ear-worn motion-tolerant blood pressure and heart rate blood pressure monitoring system. To deal with large amounts of motion artifacts induced by head movements, we presented an unsupervised learning strategy to automatically evaluate the distortion condition of the raw heartbeats identified and perform signal quality labelling and signal purification.

Possible future research extensions include:

- 1) The motion artifacts-tolerant ability of the blood pressure monitoring system can be an interesting topic. People may perform different physical activities during daily application scenarios. It will be convenient if robust blood pressure estimates can still be obtained without limiting daily activities.
- 2) Other sensors placement methods for wearable blood pressure monitoring will be helpful to provide more alternatives. It is important to design sophisticated algorithm to process weak signals acquired by convenient by non-standard sensors placement methods.
- 3) Low power consumption of the wearable computers is also important to enable long-term convenient wearable applications. Exploring different low power techniques will contributes to the pervasiveness individual-centric computing systems.

## REFERENCES

- [1] A. Fleury, *et al.*, "SVM-based multimodal classification of activities of daily living in health smart homes: sensors, algorithms, and first experimental results," *IEEE transactions on information technology in biomedicine*, vol. 14, pp. 274-283, 2010.
- [2] B. R. Wang, *et al.*, "Influential factors of smart health users according to usage experience and intention to use," *Wireless Personal Communications*, vol. 79, pp. 2671-2683, 2014.
- [3] T. Suzuki, *et al.*, "Wearable wireless vital monitoring technology for smart health care," in *Medical Information and Communication Technology (ISMICT), 2013 7th International Symposium on*, 2013, pp. 1-4.
- [4] S. Warren, *et al.*, "Designing smart health care technology into the home of the future," in *Workshops on Future Medical Devices: Home Care Technologies for the 21st Century*, 1999, p. 667.
- [5] A. Solanas, *et al.*, "Smart health: a context-aware health paradigm within smart cities," *IEEE Communications Magazine*, vol. 52, pp. 74-81, 2014.
- [6] M. M. Baig, *et al.*, "Smart health monitoring systems: an overview of design and modeling," *Journal of medical systems*, vol. 37, pp. 1-14, 2013.
- [7] A. Lymberis, "Smart wearable systems for personalised health management: current R&D and future challenges," in *Engineering in Medicine and Biology Society, 2003. Proceedings of the 25th Annual International Conference of the IEEE*, 2003, pp. 3716-3719.
- [8] J. Ko, *et al.*, "Wireless sensor networks for healthcare," *Proceedings of the IEEE*, vol. 98, pp. 1947-1960, 2010.
- [9] A. Darwish, *et al.*, "Wearable and implantable wireless sensor network solutions for healthcare monitoring," *Sensors*, vol. 11, pp. 5561-5595, 2011.
- [10] D. O. Olguin, *et al.*, "Wearable sensors for pervasive healthcare management," in *2009 3rd International Conference on Pervasive Computing Technologies for Healthcare*, 2009, pp. 1-4.
- [11] K. Hung, *et al.*, "Wearable medical devices for tele-home healthcare," in *Engineering in Medicine and Biology Society, 2004. IEMBS'04. 26th Annual International Conference of the IEEE*, 2004, pp. 5384-5387.
- [12] D.-O. Kang, *et al.*, "A wearable context aware system for ubiquitous healthcare," in *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*, 2006, pp. 5192-5195.

- [13] E. Foxlin, *et al.*, "Weartrack: A self-referenced head and hand tracker for wearable computers and portable vr," in *Wearable Computers, The Fourth International Symposium on*, 2000, pp. 155-162.
- [14] T. E. Starner, "Wearable computers: No longer science fiction," *IEEE pervasive computing*, vol. 1, pp. 86-88, 2002.
- [15] J. Berzowska, "Electronic textiles: Wearable computers, reactive fashion, and soft computation," *Textile*, vol. 3, pp. 58-75, 2005.
- [16] C. L. Lisetti, *et al.*, "Using noninvasive wearable computers to recognize human emotions from physiological signals," *EURASIP Journal on Advances in Signal Processing*, vol. 2004, pp. 1-16, 2004.
- [17] S. Mendis, *et al.*, *Global atlas on cardiovascular disease prevention and control*: World Health Organization, 2011.
- [18] WHO. *Raised blood pressure* [http://www.who.int/gho/ncd/risk\\_factors/en/](http://www.who.int/gho/ncd/risk_factors/en/).
- [19] R. E. Kleiger, *et al.*, "Heart rate variability: measurement and clinical utility," *Annals of Noninvasive Electrocardiology*, vol. 10, pp. 88-101, 2005.
- [20] N. V. Thakor, *et al.*, "Applications of adaptive filtering to ECG analysis: noise cancellation and arrhythmia detection," *Biomedical Engineering, IEEE Transactions on*, vol. 38, pp. 785-794, 1991.
- [21] G. B. Moody, *et al.*, "Development and evaluation of a 2-lead ECG analysis program," *Computers in cardiology*, vol. 9, pp. 39-44, 1982.
- [22] M. Chawla, "A comparative analysis of principal component and independent component techniques for electrocardiograms," *Neural Computing and Applications*, vol. 18, pp. 539-556, 2009.
- [23] A. V. Gribok, *et al.*, "A robust method to estimate instantaneous heart rate from noisy electrocardiogram waveforms," *Annals of biomedical engineering*, vol. 39, pp. 824-834, 2011.
- [24] Q. Li, *et al.*, "Robust heart rate estimation from multiple asynchronous noisy sources using signal quality indices and a Kalman filter," *Physiological measurement*, vol. 29, p. 15, 2008.
- [25] S. Mukhopadhyay, *et al.*, "Wavelet based qrs complex detection of ECG signal," *International Journal of Engineering Research and Applications*, vol. 2, pp. 2361-2365, 2012.

- [26] M. Kotas, "Application of dynamic time warping to ECG processing," in *Proc. Int. Conf. MIT, Poland, IX*, 2006, pp. 169-175.
- [27] P. E. McSharry, *et al.*, "A dynamical model for generating synthetic electrocardiogram signals," *Biomedical Engineering, IEEE Transactions on*, vol. 50, pp. 289-294, 2003.
- [28] A. Al-Jawad, *et al.*, "Using multi-dimensional dynamic time warping for tug test instrumentation with inertial sensors," in *Multisensor Fusion and Integration for Intelligent Systems (MFI), 2012 IEEE Conference on*, 2012, pp. 212-218.
- [29] Texas Instrument. (2012). *ADS1299EEG-FE evaluation board* Available: <http://www.ti.com/tool/ads1299eegfe-pdk>
- [30] Texas Instrument. (2013). *Tiva™ C Series LaunchPad* Available: <http://www.ti.com/tool/ek-tm4c123gxl>
- [31] Q. Zhang, *et al.*, "A novel machine learning-enabled framework for instantaneous heart rate monitoring from motion-artifact-corrupted electrocardiogram signals," *Physiological Measurement* vol. 37, p. 1945, 2016.
- [32] I. Tomasic, *et al.*, "Electrocardiographic Systems With Reduced Numbers of Leads—Synthesis of the 12-Lead ECG," *Biomedical Engineering, IEEE Reviews in*, vol. 7, pp. 126-142, 2014.
- [33] G. B. Moody, *et al.*, "A noise stress test for arrhythmia detectors," *Computers in cardiology*, vol. 11, pp. 381-384, 1984.
- [34] A. L. Goldberger, *et al.*, "Physiobank, physiokit, and physionet components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, pp. e215-e220, 2000.
- [35] N. Iyengar, *et al.*, "Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, vol. 271, pp. R1078-R1084, 1996.
- [36] E. Plesnik, *et al.*, "Improved removal of electrocardiogram baseline wandering," in *EUROCON, 2013 IEEE*, 2013, pp. 1764-1769.
- [37] E. Turajlic, "A novel algorithm for ECG parametrization and synthesis," in *Biomedical Engineering and Sciences (IECBES), 2012 IEEE EMBS Conference on*, 2012, pp. 927-932.
- [38] H. Tanaka, *et al.*, "Age-predicted maximal heart rate revisited," *Journal of the American College of Cardiology*, vol. 37, pp. 153-156, 2001.
- [39] WHO. (2016). *Target Heart Rates*. Available: <http://www.heart.org/HEARTORG/>

- [40] L. Sun, *et al.*, "ECG analysis using multiple instance learning for myocardial infarction detection," *Biomedical Engineering, IEEE Transactions on*, vol. 59, pp. 3348-3356, 2012.
- [41] Q. Li, *et al.*, "Ventricular fibrillation and tachycardia classification using a machine learning approach," *Biomedical Engineering, IEEE Transactions on*, vol. 61, pp. 1607-1613, 2014.
- [42] F. Alonso-Atienza, *et al.*, "Detection of life-threatening arrhythmias using feature selection and support vector machines," *Biomedical Engineering, IEEE Transactions on*, vol. 61, pp. 832-840, 2014.
- [43] A. H. Khandoker, *et al.*, "Support vector machines for automated recognition of obstructive sleep apnea syndrome from ECG recordings," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 13, pp. 37-48, 2009.
- [44] N. Vyas, "Biomedical Signal Processing," *Pinnacle Technology*, 2011.
- [45] G. M. Friesen, *et al.*, "A comparison of the noise sensitivity of nine QRS detection algorithms," *Biomedical Engineering, IEEE Transactions on*, vol. 37, pp. 85-98, 1990.
- [46] J. Bai, *et al.*, "Tests for skewness, kurtosis, and normality for time series data," *Journal of Business & Economic Statistics*, vol. 23, pp. 49-60, 2005.
- [47] A. Phinyomark, *et al.*, "A novel feature extraction for robust EMG pattern recognition," *arXiv preprint arXiv:0912.3973*, 2009.
- [48] N. Sánchez-Marroño, *et al.*, "Filter methods for feature selection—a comparative study," in *Intelligent Data Engineering and Automated Learning-IDEAL 2007*, ed: Springer, 2007, pp. 178-187.
- [49] T. Hastie, *et al.*, "The Elements of Statistical Learning - Data Mining, Inference, and Prediction," *Springer*, p. 91, 2008.
- [50] J. Bi, *et al.*, "Dimensionality reduction via sparse support vector machines," *The Journal of Machine Learning Research*, vol. 3, pp. 1229-1243, 2003.
- [51] C. J. Burges, "A tutorial on support vector machines for pattern recognition," *Data mining and knowledge discovery*, vol. 2, pp. 121-167, 1998.
- [52] S. Arlot, *et al.*, "A survey of cross-validation procedures for model selection," *Statistics surveys*, vol. 4, pp. 40-79, 2010.
- [53] J. Mateo, *et al.*, "Improved heart rate variability signal analysis from the beat occurrence times according to the IPFM model," *Biomedical Engineering, IEEE Transactions on*, vol. 47, pp. 985-996, 2000.

- [54] M. Vollmer, "Robust detection of heart beats using dynamic thresholds and moving windows," in *Computing in Cardiology 2014*, 2014, pp. 569-572.
- [55] American Heart Association. (2015). *Target Heart Rates*.
- [56] R.-E. Fan, *et al.*, "LIBLINEAR: A library for large linear classification," *The Journal of Machine Learning Research*, vol. 9, pp. 1871-1874, 2008.
- [57] C.-C. Chang, *et al.*, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 2, p. 27, 2011.
- [58] R. Mukkamala, *et al.*, "Toward ubiquitous blood pressure monitoring via pulse transit time: theory and practice," *IEEE Transactions on Biomedical Engineering*, vol. 62, pp. 1879-1901, 2015.
- [59] D. Buxi, *et al.*, "A survey on signals and systems in ambulatory blood pressure monitoring using pulse transit time," *Physiological measurement*, vol. 36, p. R1, 2015.
- [60] Y. Yoon, *et al.*, "Non-constrained blood pressure monitoring using ECG and PPG for personal healthcare," *Journal of medical systems*, vol. 33, pp. 261-266, 2009.
- [61] Texas Instrument. (2014). *AFE4490SPO2 evaluation board* Available: <http://www.ti.com/tool/afe4490spo2evm>
- [62] Covidien. *Nellcor™ Forehead SpO<sub>2</sub> Sensor* <http://www.medtronic.com/covidien/products/pulse-oximetry/nellcor-spo2-forehead-sensor>.
- [63] IMPEX. *MARCY ME-709 recumbent exercise bike* <http://www.marcypro.com/marcy-recumbent-bike>.
- [64] S. Puke, *et al.*, "Blood pressure estimation from pulse wave velocity measured on the chest," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2013, pp. 6107-6110.
- [65] CONTEC. *ABPM50* <http://www.contecm.com/>.
- [66] M. W. Chen, *et al.*, "Continuous estimation of systolic blood pressure using the pulse arrival time and intermittent calibration," *Medical and Biological Engineering and Computing*, vol. 38, pp. 569-574, 2000.
- [67] M. Masè, *et al.*, "Feasibility of cuff-free measurement of systolic and diastolic arterial blood pressure," *Journal of electrocardiology*, vol. 44, pp. 201-207, 2011.
- [68] R. Payne, *et al.*, "Pulse transit time measured from the ECG: an unreliable marker of beat-to-beat blood pressure," *Journal of Applied Physiology*, vol. 100, pp. 136-141, 2006.

- [69] S. Mazaheri, *et al.*, "A comparative review of blood pressure measurement methods using pulse wave velocity," in *Smart Instrumentation, Measurement and Applications (ICSIMA), 2014 IEEE International Conference on*, 2014, pp. 1-5.
- [70] F. S. Cattivelli, *et al.*, "Noninvasive cuffless estimation of blood pressure from pulse arrival time and heart rate with adaptive calibration," in *2009 Sixth international workshop on wearable and implantable body sensor networks*, 2009, pp. 114-119.
- [71] S. Mottaghi, *et al.*, "Cuffless Blood Pressure Estimation during Exercise Stress Test," *International Journal of Bioscience, Biochemistry and Bioinformatics*, vol. 2, p. 394, 2012.
- [72] T. Pereira, *et al.*, "Correlation study between blood pressure and pulse transit time," in *Bioengineering (ENBENG), 2015 IEEE 4th Portuguese Meeting on*, 2015, pp. 1-5.
- [73] D. G. Altman, *et al.*, "Commentary on quantifying agreement between two methods of measurement," *Clinical chemistry*, vol. 48, pp. 801-802, 2002.
- [74] AAMI. (2007). *ANSI/AAMI SP10-1992 standard* <http://www.aami.org/>.
- [75] J. Solà, *et al.*, "Noninvasive and nonocclusive blood pressure estimation via a chest sensor," *IEEE Transactions on Biomedical Engineering*, vol. 60, pp. 3505-3513, 2013.
- [76] Y. M. Chi, *et al.*, "Dry-contact and noncontact biopotential electrodes: methodological review," *IEEE reviews in biomedical engineering*, vol. 3, pp. 106-119, 2010.
- [77] G. Clifford, *et al.*, "Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms," *Physiological measurement*, vol. 33, p. 1419, 2012.
- [78] I. Dotsinsky, *et al.*, "Simple approach for tremor suppression in electrocardiograms," *Bioautomation*, vol. 14, pp. 129-36, 2010.
- [79] I. A. Dotsinsky, *et al.*, "Tremor suppression in ECG," *Biomedical engineering online*, vol. 7, p. 1, 2008.
- [80] S. Vogel, *et al.*, "In-ear vital signs monitoring using a novel microoptic reflective sensor," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, pp. 882-889, 2009.
- [81] B. Venema, *et al.*, "Evaluating innovative in-ear pulse oximetry for unobtrusive cardiovascular and pulmonary monitoring during sleep," *IEEE journal of translational engineering in health and medicine*, vol. 1, pp. 2700208-2700208, 2013.
- [82] D. Da He, *et al.*, "An ear-worn vital signs monitor," *IEEE Transactions on Biomedical Engineering*, vol. 62, pp. 2547-2552, 2015.

- [83] Y.-t. Zhang, *et al.*, "A health-shirt using e-textile materials for the continuous and cuffless monitoring of arterial blood pressure," in *2006 3rd IEEE/EMBS International Summer School on Medical Devices and Biosensors*, 2006, pp. 86-89.
- [84] A. Jadooei, *et al.*, "Adaptive algorithm for continuous monitoring of blood pressure using a pulse transit time," in *Electronics and Nanotechnology (ELNANO), 2013 IEEE XXXIII International Scientific Conference*, 2013, pp. 297-301.
- [85] X. Long, *et al.*, "Sleep and wake classification with actigraphy and respiratory effort using dynamic warping," *IEEE journal of biomedical and health informatics*, vol. 18, pp. 1272-1284, 2014.
- [86] T. Velmurugan, *et al.*, "Computational complexity between K-means and K-medoids clustering algorithms for normal and uniform distributions of data points," *Journal of computer science*, vol. 6, p. 363, 2010.
- [87] G. Zack, *et al.*, "Automatic measurement of sister chromatid exchange frequency," *Journal of Histochemistry & Cytochemistry*, vol. 25, pp. 741-753, 1977.
- [88] A. Pentland, *et al.*, "Modeling and prediction of human behavior," *Neural computation*, vol. 11, pp. 229-242, 1999.
- [89] D. Arthur, *et al.*, "k-means++: The advantages of careful seeding," in *Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms*, 2007, pp. 1027-1035.
- [90] L. Kaufman, *et al.*, "Partitioning around medoids (program pam)," *Finding groups in data: an introduction to cluster analysis*, pp. 68-125, 1990.

## **BIOGRAPHICAL SKETCH**

Qingxue Zhang received his BS and MS degrees in Xi'an Jiaotong University, Xi'an, China, in 2004 and 2007, respectively. From 2007 to 2013, he worked in an information technology company focusing on wireless system R&D. Since 2013, he has been a Ph.D. student in the Department of Electrical Engineering at The University of Texas at Dallas. His current research interests are algorithm design (machine learning, biomedical signal processing) and embedded system design (wearable computers) for smart health applications, towards individual-centric pervasive computing.

## **CURRICULUM VITAE**

**Qingxue Zhang**

Email: qingxue.zhg@gmail.com

### EDUCATION

- PhD in EE                      2013-2017      The University of Texas at Dallas, USA
- MS and BS in EE            2000-2007      Xi'an Jiaotong University, China

### PROFESSIONAL EXPERIENCE

- PhD research                2013-2017      The University of Texas at Dallas, USA  
Study Machine Learning-empowered Smart Health Wearables for 24-hour continuous Blood Pressure & Heart Rate Tracking focusing on unsupported features of on the shelf wearable computers.
- System Architect            2007-2013      Hisilicon Co. Ltd., Shanghai, China  
Research and development of large-scale WCDMA and LTE wireless communication systems.

### RESEARCH INTERESTS

- Machine Learning, Deep Learning, Biomedical/Physiological Signal Processing
- Pervasive Computing, Wearable Computers, Embedded Systems Design
- Smart Health, Brain Computer Interface, Computational Neuroscience

### PUBLICATIONS

- Q. Zhang, D. Zhou, and X. Zeng, "A Novel Framework for Motion-tolerant Instantaneous Heart Rate Estimation By Phase-domain Multi-view Dynamic Time Warping," Transactions on Biomedical Engineering, vol. 99, 2016

- Q. Zhang, D. Zhou, and X. Zeng, "A novel machine learning-enabled framework for instantaneous heart rate monitoring from motion artifact-corrupted electrocardiogram signals," *Physiological Measurement*, vol. 37, 2016
- Q. Zhang, D. Zhou, and X. Zeng, "Highly Wearable Cuff-less Blood Pressure and Heart Rate Monitoring with Single-arm-ECG and PPG Signals," *Biomedical Engineering Online*, vol. 16, p. 23, 2017

#### TECHNICAL SKILLS

- R Studio, Python, C, Matlab, Verilog HDL, Spice, Cadence tools, Synopsys tools

#### PROFESSIONAL MEMBERSHIPS

- IEEE member 2013-present