mmHRV: Contactless Heart Rate Variability Monitoring using Millimeter-Wave Radio

Fengyu Wang, *Student Member, IEEE*, Xiaolu Zeng, *Student Member, IEEE*, Chenshu Wu, *Member, IEEE*, Beibei Wang, *Senior Member, IEEE*, and K. J. Ray Liu, *Fellow, IEEE*
Department of Electrical and Computer Engineering, University of Maryland, College Park, MD 20742
Origin Wireless Inc., Greenbelt, MD 20770 USA,
Email: {fywang, xlzeng09, cswu, bebewang, kjrliu}@umd.edu.

***Abstract*—Heart Rate Variability (HRV), which measures the fluctuation of heartbeat intervals, has been considered as an important indicator for general health evaluation. To alleviate the user burden and explore the usability for long-term health monitoring, non-contact methods for HRV monitoring have drawn tremendous attention. In this paper, we present mmHRV, the first contact-free multi-user HRV monitoring system using commercial millimeter-Wave (mmWave) radio. The design of mmHRV consists of two key components. First, we develop a calibration-free target detector to identify each user's location. Second, a heartbeat signal extractor is devised, which can optimize the decomposition of the phase of the channel information modulated by the chest movement, and thus estimate the heartbeat signal. The exact time of heartbeats is estimated by finding the peak location of the heartbeat signal while the Inter-Beat Intervals (IBIs) can be further derived for evaluating the HRV metrics of each target. We evaluate the system performance and the impact of different settings including the distance between human and the device, user orientation, incidental angle and blockage. Experimental results show that mmHRV can measure the HRV accurately with a median IBI estimation error of 28ms (w.r.t. 96.16% accuracy). In addition, the Root-Mean-Square-Error (RMSE) measured in the Non-Line-of-Sight (NLOS) scenarios is 31.71ms based on the experiments with 11 participants. The performance of the multi-user scenario is slightly degraded compared with the single-user case, however, the median error of the 3-user case is within 52ms for all 3 tested locations.**

***Index Terms*—Heart Rate Variability (HRV), heart beat estimation, wireless sensing, millimeter-wave radio.**

I. Introduction

Heart Rate Variability (HRV), defined as the variation of the periods between consecutive heartbeats, i.e., Inter-Beat Intervals (IBI), is an important indicator of the overall health status of an individual [1]. Analysis of the HRV has been proved to be a powerful tool to assess cardiac health and evaluate the state of the Autonomic Nervous System (ANS) [2]. High-accuracy HRV monitoring is required in numerous applications such as early diagnose of cardiovascular disease, stress evaluation, emotions recognition and anxiety treatment, etc. [3]–[7].

Traditional measurements of the HRV are obtained by continuously measuring the IBIs using the electrocardiogram (ECG) or photoplethysmogram (PPG) sensors, both of which are dedicated medical devices and have to be physically contacted with the human skin. However, using ECG or PPG is uncomfortable for users and sometimes may cause skin allergies. To avoid the direct contact with users' skin, other wearable devices such as Inertial Measurement Units (IMUs) have been explored to measure the movements of the chest surfaces to determine the IBIs and then measure the HRV [8]. Although some of the aforementioned methods are less invasive than ECG and PPG-based approaches, all of them require users to wear dedicated devices, which is cumbersome and usually expensive for daily usage. Therefore, how to monitor the HRV in a non-contact way has become an important topic for both academia and industry.

To this end, Radio Frequency (RF) based sensing has become one of the most promising candidates, because the presence of a human subject will affect the RF signal propagation [9]–[11], e.g., RF signals reflected from the human body will be modulated by the body movement such as the chest movement caused by respiration and heartbeat, vital information of the human subject can be unveiled by analyzing the channel propagation characteristics [11]–[14]. While many existing works have validated the feasibility of vital sign monitoring using RF signal, most of them can only estimate the Respiration Rate (RR) [15]–[17] and the Heart Rate (HR) [18]–[22], from which one cannot obtain the HRV without the precise timing of each heartbeat. As a result, accurate RF-based HRV monitoring needs to be further investigated.

Technically, accurate HRV estimation is much more difficult than HR estimation. The existing HR estimating systems usually take multiple samples in the time domain to achieve higher HR estimation accuracy [17]–[21], which equals to averaging the heartbeats over a certain time window. However, they are not applicable for HRV estimation which needs the exact time of each heartbeat and entails the following challenges. First, RF signals reflected by human chests are modulated by both respiration and heartbeats in which the distance change caused by respiration is a magnitude greater than that caused by heartbeats. In signal process terminology, the Signal-to-Interference-plus-Noise Ratio (SINR) is very low to recover and separate the heartbeat wave from the compound signal. Second, the heart pumping motion has to reach the chest wall through bones and tissues first and then be detected by the RF signal. As a result, the bones and tissues of a human body act as a filter and thus dampen the signal. Therefore, the heartbeat wave captured by RF signals lacks sharp peaks as those in ECG signals, making it harder to identify IBIs. Furthermore, to provide a robust system for HRV estimation, it is necessary to determine the number of targets and their locations before

estimating HRV for each human subject, which is non-trivial as well.

To address the above challenges, we present mmHRV, the first multi-person HRV estimation system using Commodity Off-The-Shelf (COTS) millimeter-Wave (mmWave) radio. First, a *target detector* is devised to identify the number of users and their locations without any prior calibration. Note that due to the fast attenuation of the mmWave RF signal [23], the strength of the signal decreases as it traverses a longer distance. To detect human subjects at various distances, we employ a two-dimension constant false alarm detector in the range-azimuth plane to estimate the noise level, and thus provide an adaptive threshold for target detection. The phase information is further used to filter out the static objects (e.g., walls, furniture). Note that there are usually more than one reflecting point for a single human subject. As a result, to determine the number of targets, we further employ a non-parametric clustering to identify the range-azimuth bins corresponding to each human subject.

After target detection, to estimate the HRV, the heartbeat wave needs to be extracted from the composite received signal whose phase consists of the whole chest motion including both the respiration and heartbeat movements. Note that the respiration movement ranges from $4-12$mm with a frequency of $6-30$ Breaths Per Minute (BPM) [24] while the heartbeat movement ranges from $0.2-0.5$mm with a frequency of $50-120$ BPM, both of which are quasi-periodic signals. Leveraging this property, we develop a *heartbeat wave extractor*, which optimizes the decomposition of the composite signal to several band-limited signal components. Among the decomposed signal components, the heartbeat wave will be the one whose amplitude and frequency satisfies the requirement of a typical heartbeat signal. Compared with most of the works [25]–[27], where the composite signal is decomposed successively, mmHRV can avoid the error propagation problem by concurrently decomposing the signal components. The peaks of the estimated heartbeat wave are then recognized to identify the exact time of each heartbeat. Consequently, the IBIs can be further derived and used for calculating the commonly used HRV metrics such as the Root Mean Square of Successive Differences (RMSSD), the standard deviation of all the IBIs (SDRR) and the percentage of successive IBIs that differ by more than 50ms (pNN50).

To evaluate the performance of mmHRV, 11 participants aging from 20 to 60 are asked to perform extensive experiments under different settings, including different distances, orientation and incidental angles. The Non-Light-of-Sight (NLOS) scenario and multi-person scenario are also investigated. Experimental results show that mmHRV achieves accurate estimations with a medium error of about 28ms for IBI estimations (w.r.t. 96.16% accuracy). The performance slightly degrades for the NLOS and the multi-user scenarios, however, the Root-Mean-Square-Error (RMSE) of the NLOS and the multi-user case are still within 32ms and 69ms respectively. The HRV metrics are also evaluated, which show a better performance compared with the state-of-art works. It is shown that mmHRV can achieve 3.89ms average error of mean IBI, 6.43ms average error of RMSSD, 6.44ms average error of SDRR and 2.52% average error of the pNN50 when users sit 1m away from the device.

The rest of the paper is organized as follows. We review the related works in Section II. The system overview and theoretical model are presented in Section III, followed by the target detection in Section IV and the heartbeat extraction and HRV estimation in Section V. The performance is evaluated in Section VI. We discuss the future work in Section VII and conclude the paper in Section VIII.

II. Related Work

Continuous monitoring of HRV is critical for early detection and prevention of potentially fatal disease. Compared with the conventional methods relying on wearable devices, the contactless method can alleviate users' burden and reduce the device cost. Existing approaches for contactless HRV estimation can be classified into two categories: video-based method and RF-based method.

Video-based: In the past decade, researches have been conducted using the video signal to estimate HRV [28]–[30]. These systems focus on the measurement of the small changes in skin color caused by blood perfusion, so that the exact time of each heartbeat can be derived for HRV analysis. However, one of the main drawbacks of these systems is the sensitivity to the lighting conditions. The system cannot work in the NLOS scenario, and the privacy invasion also hinders the wide deployment of these systems.

RF-based: Compared with the video-based method, RF signals can penetrate nonmetallic obstacles, and is more tolerable to environmental conditions such as light and temperature, making it a great candidate for non-contact HRV estimation. Dedicated radars [31]–[39] and Ultra-Wideband (UWB) [40]–[42] have been used to measure the distance change between the chest and the device to estimate heart rate (HR). However, those systems mainly rely on the frequency-domain spectral analysis to estimate the HR which may take a couple of seconds, making it impossible to estimate the precise timing of each heartbeat for calculating the HRV.

To achieve robust HRV estimation, a preliminary work [43] tries to eliminate the respiration effect by asking users to hold their breaths. However, holding breath will impact the estimation of the HRV and thus the HRV metrics estimated in this condition cannot indicate the users' health condition accurately. To extract the heartbeat wave, the 2nd-derivative of the distance change (i.e., acceleration) has been considered in [7], which is equivalent to a High-Pass Filter (HPF). However, the residual signal is too noisy for heartbeat extraction. To identify the exact time of each heartbeat, it is assumed that the heartbeat signal is the successive multiple copies of a heartbeat template with different time scales. However, as shown in [25], using a single template is insufficient. To get a high detection accuracy, multiple templates are needed, and training is required before estimation. A simple approach based on two Band-Pass Filters (BPFs) is employed for separating the respiration and the heartbeat signal in [26] [27], however, since undesired peaks caused by the interference (e.g., harmonics of respiration) may overlap with the dominant frequency band of the heartbeat, dedicated systems need to be employed to

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2021.3075167, IEEE Internet of Things Journal

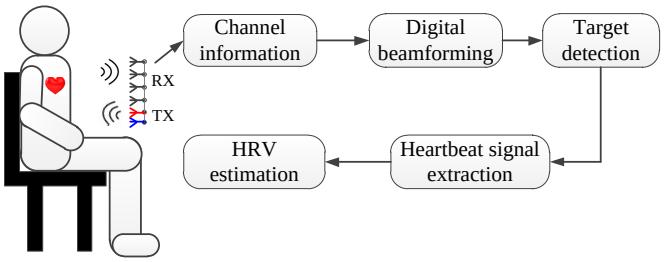


Fig. 1: Processing flow of mmHRV.

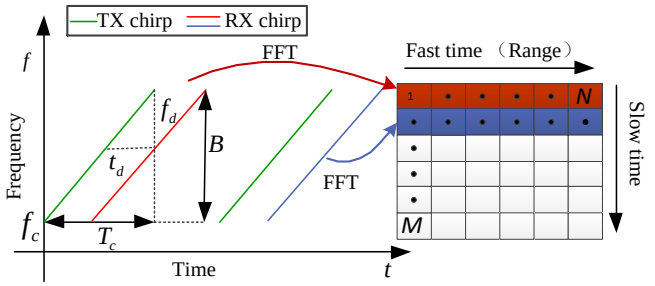


Fig. 2: FMCW radar system.

remove the false peaks. Ensemble Empirical Mode Decomposition (EEMD) is applied in [26] to remove the noise and interference, however, the mode mixing problem as well as selecting the optimal decomposed Intrinsic Mode Functions (IMFs) is not easy to solve in real applications. IBIs are assumed to not change much in [27], and the auto-correlation is used to remove the false peaks. However, the system can only perform well when the passing band does not overlap with the strong interference. To reduce the interference, the Band-Pass Filter Bank (BPFB) is applied in [44], where the HR is first estimated and the heartbeat signal is then filtered by using the BPF with a center frequency at HR. However, a large error will occur once the HR is not estimated correctly.

The above methods try to decompose the composite signal successively [25]–[27], which causes error propagation once the interference is not removed correctly. In our system, to avoid the problem, we try to decompose the composite signal concurrently leveraging the fact that the chest movement is composed of several band-limited signals. The experiment results show the superiority of our method compared with the state-of-art works. Besides, all the related works only work for the single-user case, and the target detection procedure is omitted. In mmHRV, by properly using the channel information, the estimation of HRV for multi-users can be obtained.

III. System Design and Theoretical Model

A. System Overview

mmHRV is a wireless system that can accurately detect the heartbeat signal of human subjects and estimate their HRV by purely using the RF signals reflected off the users' bodies. The processing flow of mmHRV is shown in Fig. 1. First, a Frequency-Modulated Continuous Wave (FMCW) radar transmits the RF signal and captures the reflections of human subjects and static objects. In order to detect human subjects at different locations, the ***Bartlett beamformer*** is applied to get the channel information at different azimuth-range bins. Then, we devise a ***target detector*** that adaptively estimates the noise level at various distances and azimuth angles and thus detects the presence of reflecting objects. The variance of phase is further utilized to distinguish human subjects and static objects. To identify the number of target and their locations, a non-parametric clustering algorithm is employed. To extract the heartbeat signal from the phase information that is modulated by both respiration and heartbeat, we devise a *heartbeat signal extractor*, which can decompose the phase signal into several narrow-band signals concurrently and give an estimate of heartbeat wave. HRV can be further analyzed based on the Inter-Beat Intervals (IBIs) derived from the estimated heartbeat signals.

B. Signal Model

A chirp signal is transmitted by the FMCW radar, where the instantaneous transmitting frequency is a periodic linearly-increasing signal as shown in Fig. 2, and it can be expressed as [20]

$$f_t = f_c + \frac{B}{T_c}t, \tag{1}$$

where f_c is the chirp starting frequency, T_c is the chirp duration and B is the bandwidth. According to Frequency Modulation (FM), the transmitted signal $x_T(t)$ can be expressed as

$$\begin{aligned} x_T(t) &= A_T \exp\{-j[2\pi \int_0^t f_t(\tau)d\tau]\} \\ &= A_T \exp\{-j[2\pi f_c t + \pi \frac{B}{T_c}t^2]\}, \end{aligned} \tag{2}$$

where A_T is the transmitting power. When the electromagnetic (EM) wave is reflected by human chest at distance $d(t)$, the reflected signal $x_R(t)$ can be expressed as

$$x_R(t) = A_R \exp\{-j[2\pi f_c(t-t_d) + \pi \frac{B}{T_c}(t-t_d)^2]\}, \tag{3}$$

where A_R is the amplitude of the receiving signal. t_d stands for the round-trip delay and can be denoted as $t_d = \frac{2d(t)}{c}$, where c is the speed of light.

Mixing the received signal with a replica of the transmitted signal and following a low-pass filter, the channel information $h(t)$ can be expressed as

$$h(t) = A \exp\{-j(2\pi \frac{Bt_d}{T_c}t + 2\pi f_c t_d - \pi \frac{B}{T_c}t_d^2)\}. \tag{4}$$

Note that the term $\pi \frac{B}{T_c}t_d^2$ is negligible, especially in short-range scenarios. Therefore, the $h(t)$ can be written as

$$h(t) = A \exp\{-j(2\pi \frac{Bt_d}{T_c}t + 2\pi f_c t_d)\}, \tag{5}$$

which is a sinusoidal signal whose frequency $f_b \triangleq \frac{Bt_d}{T_c} = \frac{2Bd(t)}{cT_c}$ depends on the target's distance. For each chirp, the baseband signal $h(t)$ is digitized by Analog-to-Digital Converter (ADC)[1], producing N samples per chip, referred to as

[1]IWR1843 chipset employs a 10-bit high-speed successive approximation (SAR) ADC, where the throughput rate is 625 Kilosamples per second (Ksps) [45].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2021.3075167, IEEE Internet of Things Journal

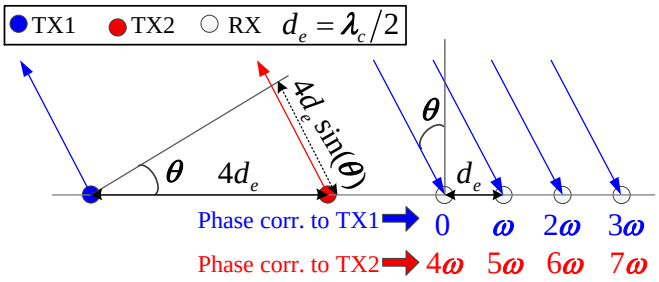


Fig. 3: Antenna Deployment.

fast time. The time corresponding to the transmission of chirps is referred to as slow time, as shown in Fig. 2. Therefore, the digitized channel information for the n^{th} ADC sample and m^{th} chirp can be expressed as

$$h(n,m) = A\exp\{-j(2\pi f_b n T_f + \frac{4\pi d(nT_f + mT_s)}{\lambda_c})\}, \quad (6)$$

where T_f and T_s are the time interval in fast time and slow time respectively[2]. λ_c denotes the wavelength of the chirp. In our system, we take advantage of the multiple antennas of the chipset, and use 2 Tx antennas and 4 Rx antennas, as shown in Fig. 3. To increase the azimuth resolution, the chirps are transmitted in the time-division multiplexing (TDM) mode [47] by transmitting sequentially through two Tx antennas. This is equivalent to the 8-element virtual array as shown in Fig. 3. Therefore, for channel l, the channel information can be rewritten as

$$h(l,n,m) = A\exp\{-j(2\pi f_b n T_f + \frac{4\pi d(nT_f + mT_s)}{\lambda_c} + 2\pi\frac{d_l \sin\theta}{\lambda_c})\}, \quad (7)$$

where d_l is the relative distance introduced by virtual antenna l. θ is the azimuth angle of the target as shown in Fig. 3.

It is apparent that the phase of the channel information changes periodically in slow time due to the periodic motions of respiration and heartbeat. Fig. 5 (a) shows a typical phase signal containing vital signs collected by our system.

IV. Target Detection

For practical application, target detection needs to be performed before vital sign detection, which has been omitted in many works. The target detection is hard to achieve, especially in the indoor scenario, where there are various objects (e.g., wall, desk, metal objects, etc.) with strong reflections of EM waves.

A. Range-FFT and Digital Beamforming

The channel information for the case when there is a static object is

$$h(l,n,m) = A\exp\{-j(2\pi f_b n T_f + \frac{4\pi d_0}{\lambda_c} + 2\pi\frac{d_l \sin\theta}{\lambda_c})\}, \quad (8)$$

[2]Note that the chirp duration is very small, and the distance change caused by vital signs is small ($<$ 12 mm) [24] [46] with low frequency ($<$ 3 Hz), therefore, there is no appreciable change in center frequency and phase during chirp time.

where d_0 is the distance between the object and the device, which stays constant in slow time.

Note that the channel information corresponding to the reflecting object is a periodic signal in fast time, and the periodicity is related to the distance as shown in Equ. (6) and Equ. (8). To determine the range information of reflecting objects, the *Fast Fourier Transform* (FFT) is performed over the fast time for each chirp, i.e., range-FFT, and the channel information can be written as $h_r(l,m)$, where r is the range tap index. The range taps corresponding to the reflecting objects would observe larger energy compared with that without reflecting objects.

To further determine the azimuth angles of the reflecting objects, digital beamforming is performed over all antenna elements for each range tap, and the channel information corresponding to range r and azimuth angle θ can be expressed as

$$h_{r,\theta}(m) = \mathbf{s}^H(\theta)\mathbf{h_{r,l}}(m) + \epsilon(m), \quad (9)$$

where $\mathbf{s}^H(\theta)$ is the steering vector towards angle θ. In mmHRV, Bartlett beamformer [48] is adopted, where the coefficient of the $l-$th antenna is

$$s_l(\theta) = \exp(-j2\pi\frac{d_l \sin\theta}{\lambda_c}). \quad (10)$$

$\epsilon(m)$ is the additive white Gaussian noise assumed to be independent and identically distributed (I.I.D) for different range-azimuth bins. $\mathbf{h_{r,l}}(m) = [h_{r,1}(m), h_{r,2}(m), \ldots, h_{r,L}(m)]$ is the channel information vector at range tap r overall all antenna elements. Therefore, for each sample m in slow time, we will have a channel information matrix $h(r,\theta)$, which contains channel information at different location bins with range r and azimuth angle θ. Fig. 4 (b) shows the amplitude of the channel information at the range-azimuth plane.

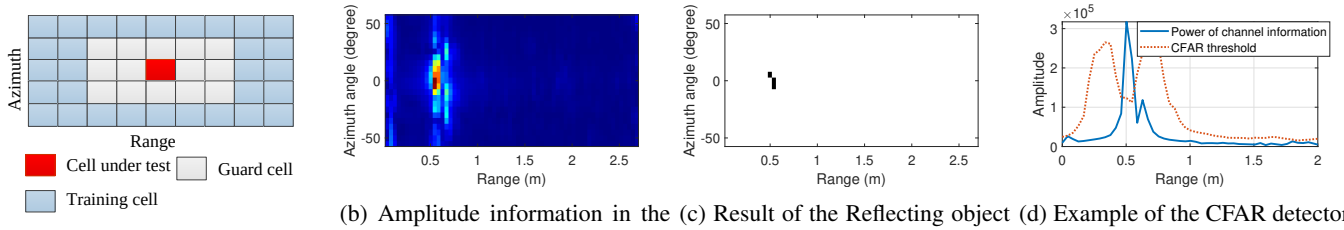
B. Reflecting Object Detector

To locate human subjects, we first need to identify the range-angle bins with reflecting objects. Note that the channel information for the bins without any reflecting object only contains noise, and thus, the energy of channel information for the bins with reflecting objects is larger than those without any reflecting objects, as shown in Equ. (6) and Equ. (8) respectively. However, it is impossible to find a universal predefined threshold for target detection. According to the propagation laws of EM wave, for the same reflecting objects, a shorter distance corresponding to a larger reflecting energy. In mmHRV, we utilize the Constant False Alarm Rate (CFAR) [49] detector, which can estimate the noise level by convolving the CFAR window (shown in Fig. 4 (a)) with the channel information at the range-azimuth plane (shown in Fig. 4 (b)), and the location bins with reflecting objects are those whose energy is above the noise level, as shown in Fig. 4 (c). Fig. 4 (d) shows the example of CFAR detection in the range domain, where the threshold is shown in the dashed line.

C. Human Subjects Detector

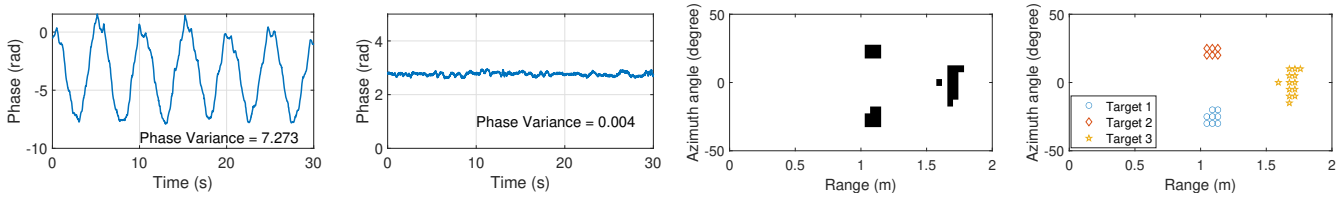
Although *Reflecting object detector* can filter out the empty taps, it cannot distinguish human subjects from static reflecting

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2021.3075167, IEEE Internet of Things Journal



(a) 2D CFAR window (b) Amplitude information in the range-azimuth plane (c) Result of the Reflecting object detector (d) Example of the CFAR detector in range domain

Fig. 4: Example of the reflecting object detector. The ground truth is that there is a human subject sit at 0.5m away from the device, with azimuth angle $0°$. (a) is the 2D-CFAR window, (b) is the amplitude of channel information in the range-azimuth plane, (c) is the reflecting result of reflecting object detector, where the black spot corresponding to the reflecting object, and (d) shows the CFAR threshold and the amplitude of channel information in range domain.



(a) Phase w.r.t. a human subject (b) Phase w.r.t. a static object (c) Bins w.r.t. human subjects (d) Target detection result

Fig. 5: Example of the human subject detector. The ground truth is that there are 3 human subjects, one of which sits at 1.5m away from the device, with azimuth angle $0°$, and the other two sit at 1m away from the device with azimuth angle $30°$ and $-30°$ respectively. (a) is phase information corresponding to a human subject, (b) is phase information corresponding to a static reflecting object, (c) is the result of the human subject detector, where the black spots corresponding to human subjects, and (d) shows the clustering result for each target.

objects. Note that different from static objects, the distance between human subjects and the device will change over slow time due to motions (e.g., respiration and heartbeat), and thus result in a phase change as shown in Fig. 5 (a). Therefore, to further filter out the static reflecting objects, we leverage the phase information of the candidate bins selected by the *Reflecting object detector*.

In specific, when the EM wave is reflected by a human subject, the phase will change over slow time due to the modulation of human motions. Therefore, there is a large phase variance for the bins corresponding to human subjects. However, for bins corresponding to the static objects (e.g. desk, wall, etc.), the phase variance will be much smaller, as shown in Fig. 5 (a) and Fig. 5 (b). So in mmHRV, to filter out the static objects, we check the variance of the phase information over slow time, and the bins corresponding to a human subject are those whose phase variance above a certain threshold.

Note that there will be more than one bin corresponding to a human subject considering the volume of a human subject, as shown in Fig. 5 (c). To identity the target number, mmHRV utilizes a non-parametric clustering method, Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm, to cluster the candidate bins without prior knowledge of cluster number. The clustering result is shown in Fig. 5 (d). The representative of each cluster is the bin with the best periodicity. In specific, the bin with the highest peak for the first peak of the auto-correlation is selected, which corresponds to the bin with the highest SNR of the vital signs [18].

V. Heartbeat Extraction and HRV Estimation

Estimating HRV requires accurate estimation of Inter-Beat Intervals (IBIs), therefore, we need to extract the displacement change caused by heartbeats (a.k.a., heartbeat wave) from the compound displacement change of chest wall and detect moments in which heartbeats occur.

A. Heartbeat Extraction Algorithm

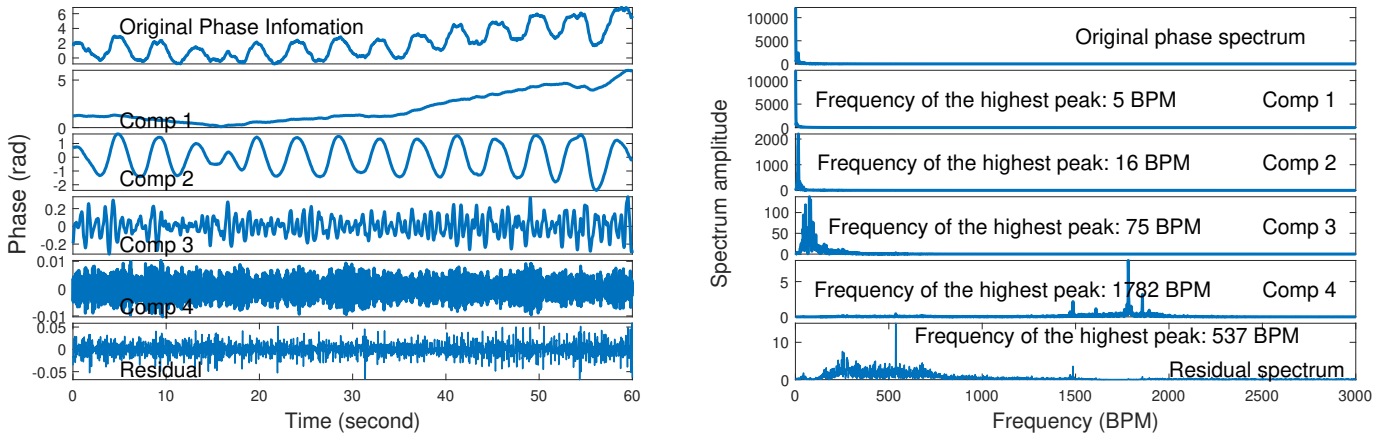
1) Problem formulation: Recall that the phase information reflects the distance change caused by vital signs. For simplicity, we directly use the analog form of signals, and the distance change of the human chest can be written as

$$y(t) = s_m(t) + s_r(t) + s_h(t) + n(t), \tag{11}$$

where $s_m(t)$ denotes the distance change caused by body motion. $s_r(t)$ and $s_h(t)$ denote the distance change caused by respiration and heartbeat, respectively. $n(t)$ is the random phase offset introduced by noise, which is independent with the phase change caused by vital signs.

Note that both $s_r(t)$ and $s_h(t)$ are quasi-periodic signals, where the period can slightly change over time. Besides, we assume the body motion introduces few oscillations, i.e., a base-band signal. Thus, the signals related with the human subject are sparse in the spectral domain and we can reconstruct these signals with a few band-limited signals. In specific, each component $u_k(t)$ is assumed to be compact around a center pulsation ω_k, which is to be determined along with the decomposition. Moreover, the decomposition should achieve

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2021.3075167, IEEE Internet of Things Journal



(a) Decomposition of a typical phase signal in time domain

(b) Corresponding spectrum of the decomposed component

Fig. 6: Example of heartbeat extractor. (a) is the decomposition result in the time domain, (b) is the corresponding spectrum of each component. In this example, the 1st component corresponds to the body motion, the 2nd component corresponds to the respiration and the 3rd component corresponds to the heartbeat. The noise falls into the 4th component and the residual.

the spectrum sparsity and data fidelity at the same time, which is modeled as

$$\min_{u_k\in\mathcal{U},\omega_k\in\Omega}\alpha\sum_{k=1}^{K}\left\|\partial t\ [(\delta(t)+\frac{j}{\pi t})*u_k(t)]\exp(-j\omega_k t)\right\|_2^2 + \left\|y(t)-\sum_{k=1}^{K}u_k(t)\right\|_2^2, \tag{12}$$

where the first term evaluates the bandwidth of the analytic signal associated with each component, and the second term evaluates the data fidelity. K is the total number of decomposition components, where $\mathcal{U}=\{u_1(t),\ldots,u_K(t)\}$ and $\Omega=\{\omega_1,\ldots,\omega_K\}$ are the set for all components and their center frequencies, respectively. α is a parameter for balancing the bandwidth constraint and data fidelity.

Once the hyper-parameters are known, the optimization problem in Equ.(12) can be solved by alternatively updating $u_k(t)$ and ω_k until convergence [50]. To update u_k, the subproblem can be written as

$$u_k(t)=\arg\min_{u_k(t)}\left\|\partial t\ [(\delta(t)+\frac{j}{\pi t})*u_k(t)]\exp(-j\omega_k t)\right\|_2^2 + \left\|y(t)-\sum_{i=1}^{K}u_i(t)\right\|_2^2. \tag{13}$$

By using the Parseval theorem, the problem can be rewritten as

$$\sqcap_k(\omega)=\arg\min_{\sqcap_k(\omega)}\alpha\left\|j\omega[(1+\mathrm{sgn}(\omega+\omega_k))\sqcap_k(\omega)]\right\|_2^2 + \left\|\dagger(\omega)-\sum_{i=1}^{K}\sqcap_i(\omega)\right\|_2^2, \tag{14}$$

where $\sqcap_k(\omega)$ and $\dagger(\omega)$ are the Fourier transfer of $u_k(t)$ and $y(t)$ respectively. After taking integrals over frequency and performing a change of variable, we can get the updating formula, where

$$\sqcap_k(\omega)=\frac{\dagger(\omega)-\sum_{i,i\neq k}\sqcap_i(\omega)}{1+2\alpha(\omega-\omega_k)^2}. \tag{15}$$

Note that the center frequencies ω_k only appears in the bandwidth constraint and thus the subproblem can be written as

$$\omega_k=\arg\min_{\omega_k}\left\|\partial t\ [(\delta(t)+\frac{j}{\pi t})*u_k(t)]\exp(-j\omega_k t)\right\|_2^2. \tag{16}$$

As before, we find the optimum in Fourier domain, and we have

$$\omega_k=\arg\min_{\omega_k}\int_0^\infty(\omega-\omega_k)^2|\sqcap_k(\omega)|^2d\omega. \tag{17}$$

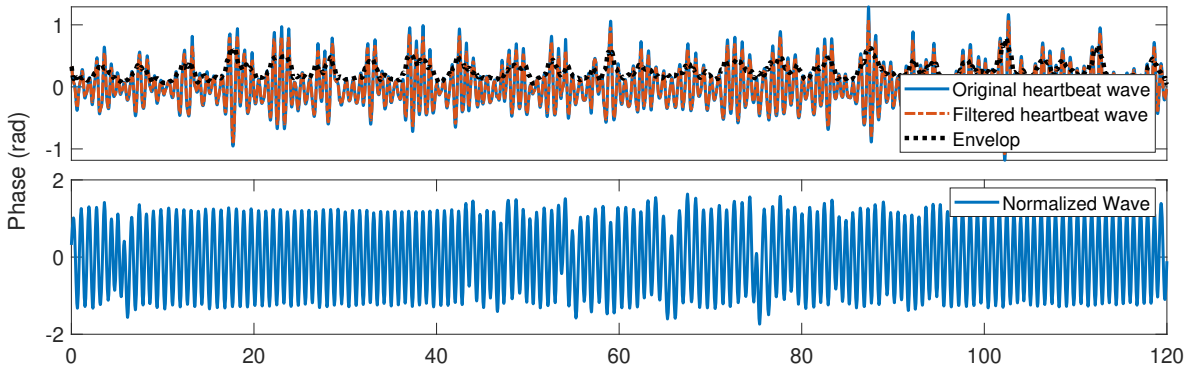
The minimizer of the above quadratic problem is

$$\omega_k=\frac{\int_0^\infty\omega|\sqcap_k(\omega)|^2d\omega}{\int_0^\infty|\sqcap_k(\omega)|^2d\omega}. \tag{18}$$

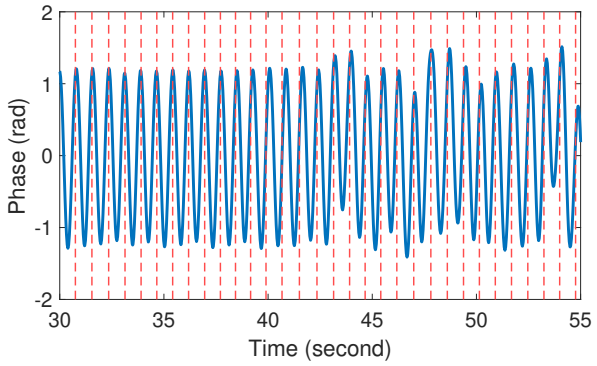
Fig. 6 illustrates the decomposition of a typical one-minute phase signal from the experiment, where the original phase information has been decomposed into 4 components. The first component reflects the body motion of the human subject, the second component is the respiration motion, and the third component is the heartbeat wave. Since the noise has different vibration characteristics as vital signals, it falls into a different mode as well as in the residual of the decomposition of the signal, as shown in Fig. 6.

2) Algorithm design: It has been proven that the decomposition problem can be solved once the hyper-parameters are properly defined. However, it is hard to predefine these hyper-parameters in real applications for heartbeat wave extraction. First, the human motion does not always exist and the human respiration sometimes will have a strong second harmonic component, making it even harder to determine the component number. Furthermore, the hyper-parameter α also influences the decomposition performance. Before discussing how to

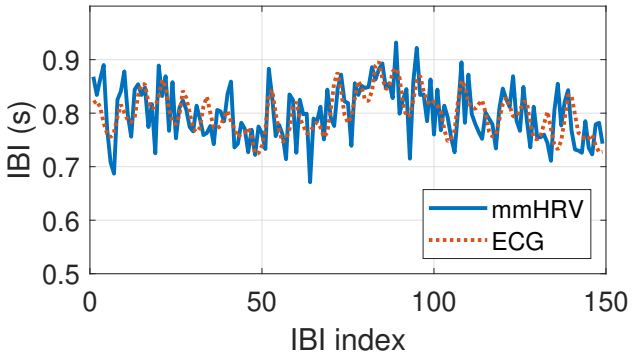
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2021.3075167, IEEE Internet of Things Journal



(a) Heartbeat wave normalization



(b) Normalized heartbeat wave v.s. ECG ground truth



(c) IBI estimation of mmHRV v.s. ECG ground truth

Fig. 7: Example of IBI estimation. (a) shows the normalization of the estimated heartbeat wave to increase accuracy in locating peaks caused by heartbeat, (b) shows the normalized heartbeat signal compared with the ECG sensor, where the ground-truth from ECG are marked as dashed lines, (c) shows the estimated IBI compared with the ground-truth from ECG sensor.

choose the hyper-parameter, we first discuss their influence on the decomposition result.

In specific, for the case that α is too small, i.e., the bandwidth constraint is too loose, when K is too small, the mixing problem will happen so that two signals may merge to a single decomposed component. However, when K is too large, some of the decomposed components may consist of noise. For the case that α is too large, i.e., the bandwidth constraint is too tight, when K is too small, some target signals may be discarded in noise. However, when K is too large, some important parts of the signal may be separated into two or more decomposed components.

In mmHRV, to accurately decompose the signal and get the component we are interested, i.e., the heartbeat wave, we are trying to adaptively change the component number K and α for different datasets. Here, we introduce a heuristic method to change K and α as the iteration proceeds to get proper decomposition result. Since the distance change caused by heartbeat is much smaller than the distance change caused by respiration and human motion[3], once the component corresponding to the heartbeat is decomposed, the component corresponding to respiration and motion should be decomposed as well, considering the data fidelity constraint in the objective function. Therefore, the algorithm will terminate once we get the component corresponding to the heartbeat. The details about the algorithm are shown in Algorithm 1.

[3]The chest movement caused by heartbeat is 0.2-0.5 mm [51] with frequency of 50-120 BPM [21]. Whereas, the chest movement caused by respiration is 4–12 mm [24] with frequency 6-30 BPM [21].

B. HRV Estimation

Once the heartbeat wave is extracted, the exact time corresponding to each heartbeat can be identified by the peaks of the heartbeat wave. To further increase the accuracy, normalization is performed before peak extraction.

In specific, the envelope of the heartbeat wave is estimated by taking moving average to the absolute value of the heartbeat component, shown as the dashed line in Fig. 7 (a). We further perform a moving average filter to the original heartbeat wave to reduce the noise. The normalized wave is the ratio between the filtered heartbeat wave and the estimated envelope. IBIs can thus be derived by calculating the time duration between two adjacent heartbeats. Fig. 7 (b) shows a segment of heartbeat wave and its ECG ground-truth, where the dashed lines show the exact time of each heartbeat from a commercial ECG sensor [52]. The peaks of normalized heartbeat wave match with the ground-truth, and Fig. 7 (c) shows the estimated IBIs and the ECG ground-truth.

The HRV features can be further obtained from the IBI sequence. In mmHRV, we use the three most widely used metrics to evaluate the HRV [2]. One is the Root Mean

Algorithm 1 Heartbeat wave extraction algorithm

Input $y(t)$
Set $\alpha \leftarrow \alpha_{\min}$, $K \leftarrow K_{\min}$
repeat
  repeat
    Initialize $\mathcal{U}$ and Ω, flag $\leftarrow 0$, $n \leftarrow 0$
    repeat
      $n \leftarrow n+1$
      for $k = 1 : K$ **do**
        update $\sqcap_k(\omega)$ using equ.(15)
        update ω_k using equ.(18)
      end for
    until convergence:$\sum_{k=1}^{K} \left\|\sqcap_k^{n+1} - \sqcap_k^n\right\|_2^2 / \left\|\sqcap_k^n\right\| < \epsilon$ or $n > n_{\max}$
    if exist $\omega_k \in [h_{\min}, h_{\max}]$ and $Range(IFFT(\sqcap_k(\omega))) < r_{\max}$ **then**
      flag $\leftarrow 1$, break;
    else
      $K \leftarrow K+1$
    end if
  until $K > K_{\max}$
  if flag == 1 **then**
    break;
  else
    $\alpha = 2\alpha$
  end if
until $\alpha > \alpha_{\max}$

Square of Successive Differences (RMSSD), which measures the successive IBI changes, and can be calculated by

$$\text{RMSSD} = \sqrt{\frac{1}{N_{\text{IBI}} - 1} \sum_{i=2}^{N_{\text{IBI}}} (\text{IBI}(i) - \text{IBI}(i-1))^2}, \quad (19)$$

where N_{IBI} is the total number of IBIs of the measurement. The standard deviation of all the IBIs (SDRR) measures the variation of the IBIs, which can be calculated as

$$\text{SDRR} = \sqrt{\frac{1}{N_{\text{IBI}}} \sum_{i=1}^{N_{\text{IBI}}\text{IBI}} (\text{IBI}(i) - \overline{\text{IBI}})^2}, \quad (20)$$

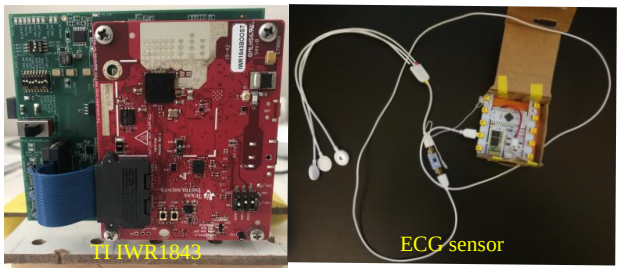
where $\overline{\text{IBI}}$ is the empirical mean of the IBIs of each measurement. The metric pNN50 measures the percentage of successive IBI that differ by more than 50 milliseconds (ms), which can be calculated by

$$\text{pNN50} = \frac{\sum_{i=2}^{N_{\text{IBI}}} \mathbb{1}\{(\text{IBI}(i) - \text{IBI}(i-1)) > 50\text{ms}\}}{N_{\text{IBI}}}, \quad (21)$$

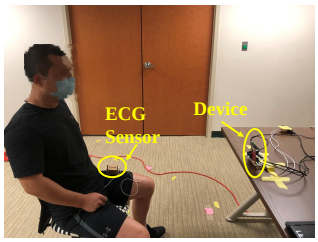
where $\mathbb{1}\{\cdot\}$ is the indicator function.

VI. Experiment Evaluation

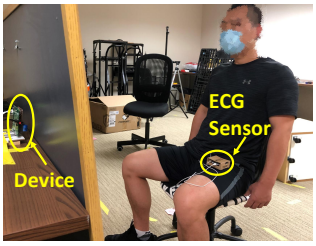
This section introduces the evaluation details of the proposed mmHRV, including practical system implementation, experiment setup, performance analysis and also the comparison with the state-of-art work.



(a) Hardware



(b) LOS setting



(c) NLOS setting

Fig. 8: Experiment setup.

A. Methodology

We prototype the mmHRV system by leveraging a commodity mmWave FMCW radar [45] in a typical office of size 3.5 m $\times$ 3.2 m as shown in Fig. 8. By configuring the 2 Tx antennas and 4 Rx antennas into TDM-MIMO mode as introduced in Section III-B, the system can achieve a theoretical azimuth resolution of 15°. The Field of View (FoV) is 100° in the horizontal plane with a radius of about 4m [21], which is sufficient to cover typical rooms. To get the true heartbeat signal, an ECG sensor [52] (shown in Fig. 8 (a)) is used to collect the ground-truth simultaneously with the mmHRV during the experiment.

In total, 11 participants (6 males and 5 females) aging from 20 to 60 are invited to conduct experiments in both LOS and NLOS scenario as shown in Fig. 8 (b) and Fig. 8 (c). We conduct the experiments with a variety of settings including different distances, incidental angles, orientations and blockages between the human subject and the radar.

To further evaluate the performance of the proposed system, we compare mmHRV with the state-of-the-art HRV estimation technique using Band-Pass-Filter-Bank (BPFB) [44], where the BPF is used to eliminate respiration interference before heartbeat wave estimation. The HR is then estimated and the heartbeat signal is estimated by using the narrow BPF whose passing band contains HR. Finally, the zero-crossing technique is applied to extract the IBI estimations from the heart rate signal.

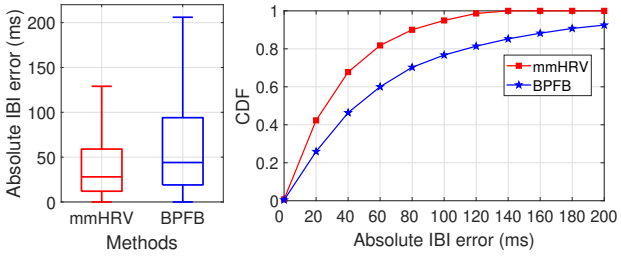
B. Overall Performance

Fig. 9 shows the overall IBI estimation accuracy of the proposed mmHRV and BPFB methods. The experiment consists of 11 participants while 15 different experiment settings (e.g., different distances, incidental angle, orientation and blockages) are conducted for each participant. As shown in Fig. 9, BPFB yields about 44ms medium error while the 90-percentile error is about 200ms. The proposed mmHRV achieves a medium

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2021.3075167, IEEE Internet of Things Journal

TABLE I: HRV estimation results in terms of mean IBI, RMSSD, SDRR and pNN50 for 11 subjects.

Metrics	Methods		User ID										
			1	2	3	4	5	6	7	8	9	10	11
Mean IBI	Value (ms)	ECG	899.4	789.9	723.2	854.6	654.5	822.9	645.2	890.1	564.9	728.1	763.8
		mmHRV	906.3	790.4	725.6	848.6	652.4	828.3	644.2	888.1	574.2	722.7	762.6
		BPFB	881.5	784.2	781.5	842.1	676.6	821.7	651.5	878.4	579.1	719	773.5
	Error (ms)	**mmHRV**	6.95	0.45	2.47	5.92	2.17	5.4	0.99	1.97	9.33	5.38	1.2
		BPFB	17.87	5.7	58.36	12.44	22.01	1.25	6.31	11.66	14.21	9.16	9.66
RMSSD	Value (ms)	ECG	38.59	10.85	37.56	31.49	34.05	35.1	16.88	27.52	5.26	23.28	31.16
		mmHRV	33.52	16.53	39.08	35.26	20.29	39.72	18.14	26.06	27.8	30.52	34.92
		BPFB	59.34	54.26	53.83	52.94	78.57	65.63	95.09	45.56	140.36	59.61	47.92
	Error (ms)	**mmHRV**	5.08	5.68	1.52	3.77	13.76	4.62	1.26	1.46	22.53	7.25	3.76
		BPFB	20.75	43.41	16.27	21.45	44.53	30.52	78.21	18.04	135.1	36.34	16.76
SDRR	Value (ms)	ECG	56.28	22.91	50.54	35.35	33.61	48.55	23.24	32.66	12.25	35.83	50.87
		mmHRV	43.22	27.25	53.3	45.88	33.54	48.53	25.49	37.43	38.66	37.15	45.51
		BPFB	71.01	47.28	110.29	58.92	69.68	55.11	67.61	50.44	118.41	47.92	63.94
	Error (ms)	**mmHRV**	13.07	4.34	2.76	10.53	0.07	0.02	2.24	4.78	26.42	1.31	5.36
		BPFB	14.72	24.37	59.74	23.57	36.07	6.55	44.37	17.78	106.16	12.09	13.07
pnn50	Value (%)	ECG	11.54	0	9.15	4.32	1.14	6.29	0.55	3.76	0	0.61	4.49
		mmHRV	8.46	1.33	7.93	5.76	2.2	6.99	2.17	2.26	4.83	6.71	6.41
		BPFB	19.4	18.54	14.57	20	14.2	22.92	18.13	12.59	10.24	12.8	12.42
	Error (ms)	**mmHRV**	3.08	1.33	1.22	1.44	1.05	0.7	1.62	1.5	4.83	6.09	1.92
		BPFB	7.86	18.54	5.42	15.68	13.06	16.62	17.58	8.83	10.24	12.19	7.93



(a) (b)

Fig. 9: Over all performance of the IBI estimation error.

error of about 28ms, with the 80ms of the 90-percentile error, which outperforms the BPFB about 60%. To thoroughly evaluate the HRV estimation accuracy, Table. I shows the estimated HRV features in terms of mean IBI, RMSSD, SDRR and pNN50 of 11 participants, where the distance between user and device is about 1m. It is shown that mmHRV can achieve 3.89ms average error of mean IBI, 6.43ms average error of RMSSD, 6.44ms average error of SDRR and 2.52% average error of the pNN50. Correspondingly, the average estimation error of BPFB is 15.33ms of mean IBI, 41.94ms of RMSSD, 32.59ms of SDRR and 12.17% of the pNN50 estimations.

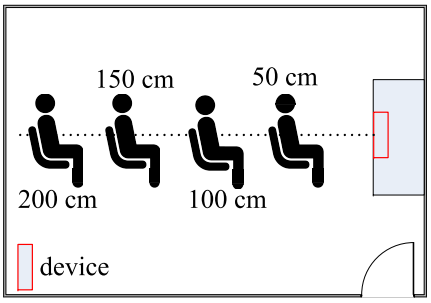
C. Impact of Distance

In this section, we explore the impact of the distance between the human subject and the device. As shown in Fig. 10 (a), the participants are asked to face towards the device and sit in four different locations ranging from 50cm to 200cm. The empirical Cumulative Distribution Function (CDF) of the absolute IBI estimation error is shown in Fig. 10 (b), while Fig. 10 (c) lists the Root Mean Square Error (RMSE) of the absolute IBI estimation error versus distance.

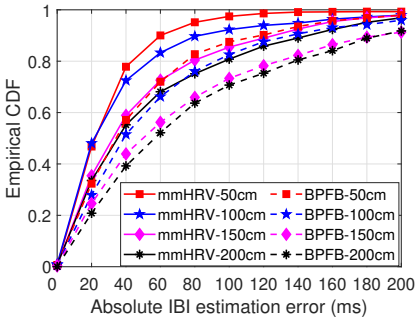
Fig. 10 (b) shows that the medium errors of mmHRV are 22ms, 22ms, 30ms and 33ms corresponding to the distance of 50cm, 100cm, 150cm and 200cm. Similarly, when the human subject moves away from the device gradually, the RMSE of the IBI estimation error increases from 26.06ms to 68.974ms as shown in Fig. 10 (c). It is clear that the IBI estimation accuracy degrades with the increment of distance, which is due to the attenuation property of the mmWave signals, as a longer propagation distance results in a lower Signal-to-Noise-Ratio (SNR).

The medium error of BPFB increase from 40ms to 60ms with the distance increasing from 50cm to 200cm. In all the four settings, mmHRV shows better performance than the benchmark BPFB in both CDF and RMSE. This is because that mmHRV directly extracts the heartbeat signal from the composite signal by optimizing the decomposition, so that the error propagation from breathing as well as random body motion elimination can be avoided. Besides, the accurate heart rate estimation is necessary for BPFB method, which however is vulnerable to noise and interference from other signal components.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2021.3075167, IEEE Internet of Things Journal



(a) Experiment setup

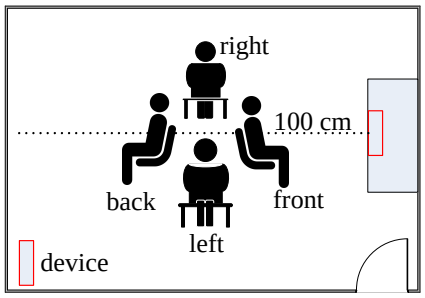


(b) CDF of absolute IBI estimation error

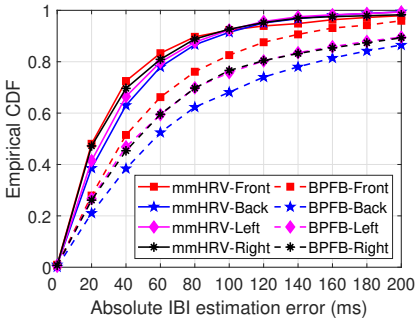
Distance (cm)	BPFB (ms)	**mmHRV** (ms)
50	43.528	**26.060**
100	48.878	**29.850**
150	57.054	**36.513**
200	61.376	**68.974**

(c) RMSE of absolute IBI estimation error

Fig. 10: Experiment setup and the absolute IBI estimation results versus distance.



(a) Experiment setup



(b) CDF of absolute IBI estimation error

Orientation	BPFB (ms)	**mmHRV** (ms)
Front	48.878	**29.850**
Back	81.906	**55.629**
Left	53.013	**36.186**
Right	56.542	**43.144**

(c) RMSE of absolute IBI estimation error

Fig. 11: Experiment setup and the absolute IBI estimation results versus orientation.

D. Impact of Orientation

Considering the real case that the user may not strictly face towards the device, this section studies the impact of users' orientation on IBI estimation accuracy. As shown in Fig. 11 (a), four normal orientations including front, back, left and right are investigated. For all the orientations, the human subject sits 1m away from the device and Fig. 11 (b) and Fig. 11 (c) show the IBI estimation accuracy in terms of CDF and RMSE, respectively.

As shown in Fig. 11 (b) and Fig. 11 (c), the median absolute IBI estimating error of mmHRV are 21ms, 22ms, 25ms and 28ms corresponding to the front, right, left and back settings. BPFB shows larger IBI errors with 40ms, 42ms, 43ms, 55ms correspondingly. Averagely, mmHRV outperforms BPFB of about 18.883ms in terms of RMSE as shown in Fig. 11 (c). However, for both methods, the "front" setting shows the best performance while the "back" one yields the largest IBI estimation error. This is due to the physiological structure of a human body, where the vibration caused by the heartbeat is larger in the front chest than in the back.

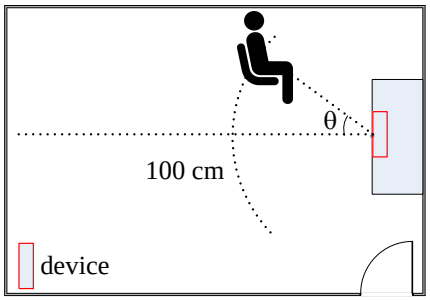
E. Impact of Incident Angle

In this section, we investigate the impact of incident angle denoted by θ in Fig. 12 (a). Specifically, the incident angle θ is set as $\theta \in \{0°, 15°, 30°\}$ while the distance between the user and device is fixed at 1m. The IBI estimation errors are shown by the CDF and RMSE in Fig. 12 (b) and Fig. 12 (c), respectively. As expected, for both methods, the performance degrades with the increment of θ in both CDF and RMSE. This is because the effective reflection area decreases when the human subject deviates from the device from $0°$ to $30°$. Moreover, according to the array signal processing theorem, the beam width will also increase with the increment of the incident angle, which reduces the directionality of the receiving signal. As a result, the SNR of the received signal decreases when the incident angle rises from $0°$ to $30°$, thus resulting in larger IBI estimation errors. However, mmHRV still outperforms BPFB of an average about 14.544ms in RMSE, which benefits from its optimization in signal decomposition for heartbeat signal extraction as introduced in Section. VI-C.

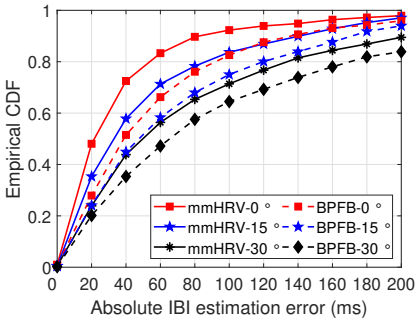
F. LOS vs NLOS

This section evaluates the estimating performance when the user and the device are blocked by a wood panel as shown in Fig. 13 (a). The distance between the participant and the device is set as 1m while the user is asked to face towards the device. As shown in Fig. 13 (b), the medium estimating error of IBI of mmHRV increases from 22ms to 24ms if the blockage happens. Correspondingly, the medium error of IBI of BPFB increases from 40ms to 48ms when the blockage occurs (see Fig. 13 (b)). The performance degradation in the blockage setting is because that the EM signal further attenuates when it penetrates the wood panel, thus rendering the decrements of SNR in the received signal. However, the HRV estimation still can work in NLOS scenario, and Fig. 13 (c) shows that the RMSE of the IBI estimation error degrades only about 2.1ms and 5ms for mmHRV and BPFB respectively.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2021.3075167, IEEE Internet of Things Journal



(a) Experiment setup

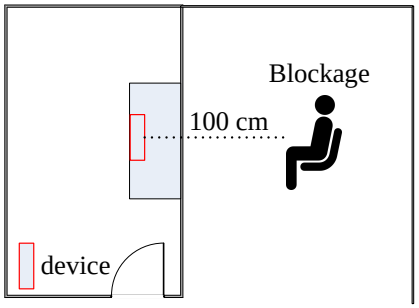


(b) CDF of absolute IBI estimation error

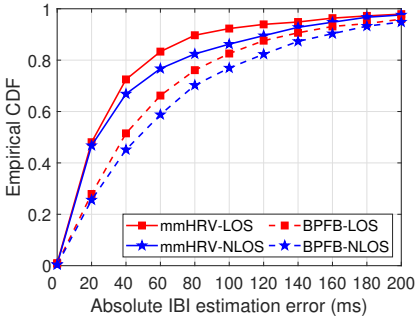
Incident angle (○)	BPFB (ms)	**mmHRV** (ms)
0	48.878	**29.850**
15	56.477	**44.981**
30	62.165	**49.027**

(c) RMSE of absolute IBI estimation error

Fig. 12: Experiment setup and the absolute IBI estimation results versus incident angle.



(a) Experiment setup



(b) CDF of absolute IBI estimation error

Bloc condition	BPFB (ms)	**mmHRV** (ms)
LOS	48.878	**29.850**
NLOS	53.736	**31.716**

(c) RMSE of absolute IBI estimation error

Fig. 13: Experiment setup and the absolute IBI estimation results versus blockage.

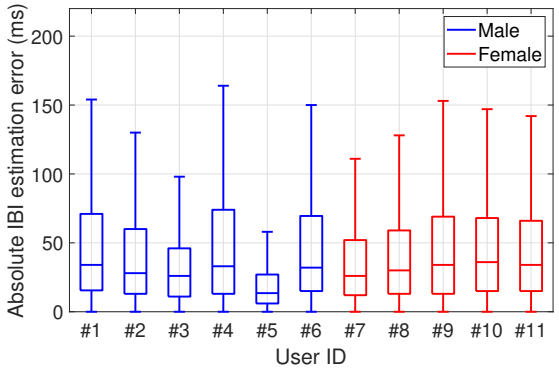


Fig. 14: Impact of user heterogeneity.

G. Impact of User Heterogeneity

To validate the robustness of mmHRV over different users, Fig. 14 summaries the absolute IBI estimation error distribution for all the 11 users of different settings (including different distance, incident angles, orientations and blockage scenario). Fig. 14 shows the error distribution of each user, where the first 6 users are males denoted by the blue box and the last 5 users are females denoted by the red box. Evidently, mmHRV demonstrates different IBI estimation errors for different users in which the medium error varies from 13.5ms to 37ms. This can be caused by several reasons such as different body shapes and heartbeat strengths over different users. It is shown that the 75-percentile error of all the uses are smaller than 75ms, which indicates great robustness of mmHRV over different subjects.

H. Multiple-User Case

In this section, we investigate the accuracy of mmHRV in a multiple-user scenario. As shown in Fig. 15 (a) and Fig. 15 (b), the participant in the middle is 1.5 meters away from the device with incident angle 0°, while the other 2 users in the left and right are 1m away from the device at incidental angle ±30°. Fig. 15 (c) and Fig. 15 (d) depicts the mean and RMSE of the absolute IBI estimation error for mmHRV. The target detection result is shown in Fig. 5 (d). Overall, mmHRV can work robustly for the 3-user setup, where the mean of the IBI error is less than 51.83ms for all the 3 locations, as shown in Fig. 15 (c). The RMSE of the absolute IBI estimation error is within 70ms for all the locations. We can see that mmHRV achieve higher accuracy in the middle location than that of either the left or the right location. This is because that the participant at the middle location enjoys the larger reflection area and thus achieves higher SNR in the received signal correspondingly. This result coincides with the experiment result as shown in Section. VI-C and Section. VI-E. However, the accuracy decreases compared with the single-user scenario for all the 3 locations. The main reason is that mmHRV utilizes the digital beamforming, and thus the reflections from other people, although suppressed by digital beamforming, act as extra interference compared with the single-user case.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JIOT.2021.3075167, IEEE Internet of Things Journal

(a) Experiment setup (b) Experiment snapshot (c) Mean IBI error. (d) RMSE of IBI error.

Fig. 15: Mean and RMSE of the absolute IBI estimation error of multiple users.

VII. Discussion and Future Work

mmHRV takes an important step towards contactless vital sign monitoring. With the miniaturization of antennas and chips, mmWave is expected to be widely available on home routers, smartphones, vehicles, e.t.c., and mmHRV enables ubiquitous and pervasive vital sign monitoring by reusing a mmWave device wherever it is already available. Note that mmHRV is designed for the common clinical setting with people stay stationary. For more general use case, handling the large body motion and enabling the vital sign monitoring when the human body is in motion is an immediate next step. Besides, given the HRV results already achieved, it is of interest to study "wireless health evaluation" problems such as stress evaluation, emotion recognition, and sleep quality evaluation.

VIII. Conclusion

In this paper, we devise mmHRV, a contact-free multi-user HRV estimating system built upon a commercial mmWave radio. To identify the number of users and their locations, a target detector is first designed to locate each user without any prior calibration. The heartbeat wave of each user is then estimated by optimizing the decomposition of the composite phase information consisting respiration, heartbeats and random body motion. The exact time of heartbeats is extracted from the estimated heartbeat wave to further evaluate the IBIs and HRV metrics. Extensive experiments are conducted, where 11 participants aging from 20 to 60 are asked to sit at different locations (distance, incidental angle, orientation, and NLOS scenario) for HRV evaluation. Experimental results show that mmHRV achieves a median error of 28 ms for the IBI estimation, outperforming the state-of-art work.

References

[1] U. Rajendra Acharya, K. Paul Joseph, N. Kannathal, C. M. Lim, and J. S. Suri, "Heart rate variability: a review," *Medical and biological engineering and computing*, 2006.

[2] F. Shaffer and J. P. Ginsberg, "An overview of heart rate variability metrics and norms," *Frontiers in public health*, 2017.

[3] J. A. Chalmers, D. S. Quintana, M. J.-A. Abbott, and A. H. Kemp, "Anxiety disorders are associated with reduced heart rate variability: a meta-analysis," *Frontiers in psychiatry*, 2014.

[4] J. Morales et al., "Use of heart rate variability in monitoring stress and recovery in judo athletes," *The Journal of Strength & Conditioning Research*, 2014.

[5] Y. Han, T. Lauteslager, T. S. Lande, and T. G. Constandinou, "UWB radar for non-contact heart rate variability monitoring and mental state classification," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019.

[6] S. Suzuki, T. Matsui, H. Imuta, M. Uenoyama, H. Yura, M. Ishihara, and M. Kawakami, "A novel autonomic activation measurement method for stress monitoring: non-contact measurement of heart rate variability using a compact microwave radar," *Medical & biological engineering & computing*, 2008.

[7] M. Zhao, F. Adib, and D. Katabi, "Emotion recognition using wireless signals," in *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*, 2016.

[8] G. Shafiq and K. C. Veluvolu, "Surface chest motion decomposition for cardiovascular monitoring," *Scientific reports*, 2014.

[9] F. Zhang, C. Chen, B. Wang, and K. J. R. Liu, "WiSpeed: A statistical electromagnetic approach for device-free indoor speed estimation," *IEEE Internet of Things Journal*, 2018.

[10] F. Zhang, C. Wu, B. Wang, H. Lai, Y. Han, and K. J. Ray Liu, "WiDetect: Robust motion detection with a statistical electromagnetic model," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2019.

[11] K. J. R. Liu and B. Wang, *Wireless AI: Wireless Sensing, Positioning, IoT, and Communications*, Cambridge University Press, 2019.

[12] E. Cianca, M. De Sanctis, and S. Di Domenico, "Radios as sensors," *IEEE Internet of Things Journal*, 2017.

[13] B. Wang, Q. Xu, C. Chen, F. Zhang, and K. J. R. Liu, "The promise of radio analytics: A future paradigm of wireless positioning, tracking, and sensing," *IEEE Signal Processing Magazine*, 2018.

[14] F. Wang, F. Zhang, C. Wu, B. Wang, and K. J. Ray Liu, "Respiration tracking for people counting and recognition," *IEEE Internet of Things Journal*, 2020.

[15] F. Zhang, C. Wu, B. Wang, M. Wu, D. Bugos, H. Zhang, and K. J. R. Liu, "SMARS: Sleep monitoring via ambient radio signals," *IEEE Transactions on Mobile Computing*, 2019.

[16] C. Chen, Y. Han, Y. Chen, H. Lai, F. Zhang, B. Wang, and K. J. R. Liu, "TR-BREATH: Time-reversal breathing rate estimation and detection," *IEEE Transactions on Biomedical Engineering*, 2018.

[17] F. Adib, H. Mao, Z. Kabelac, D. Katabi, and R . C. Miller, "Smart homes that monitor breathing and heart rate," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015.

[18] F. Wang, F. Zhang, C. Wu, B. Wang, and K. J. R. Liu, "Vimo: Multi-person vital sign monitoring using commodity millimeter wave radio," *IEEE Internet of Things Journal*, 2020.

[19] Z. Yang, P. H. Pathak, Y. Zeng, X. Liran, and P. Mohapatra, "Monitoring vital signs using millimeter wave," in *Proceedings of the 17th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, 2016.

[20] M. Mercuri, I. R. Lorato, Y. Liu, F. Wieringa, C. Van Hoof, and T. Torfs, "Vital-sign monitoring and spatial tracking of multiple people using a contactless radar-based sensor," *Nature Electronics*, 2019.

[21] A. Ahmad, J. C. Roh, D. Wang, and A. Dubey, "Vital signs monitoring of multiple people using a fmcw millimeter-wave sensor," in *2018 IEEE Radar Conference (RadarConf18)*, 2018.

[22] I. Mostafanezhad, E. Yavari, O/ Boric-Lubecke, V. M. Lubecke, and D. P. Mandic, "Cancellation of unwanted doppler radar sensor motion using empirical mode decomposition," *IEEE Sensors Journal*, 2013.

[23] M. Marcus and B. Pattan, "Millimeter wave propagation: spectrum management implications," *IEEE Microwave Magazine*, 2005.

[24] A. De Groote, M. Wantier, G. Cheron, M. Estenne, and M. Paiva, "Chest wall motion during tidal breathing," *Journal of Applied Physiology*, 1997.

[25] C. Will, K. Shi, R. Weigel, and A. Koelpin, "Advanced template matching algorithm for instantaneous heartbeat detection using continuous wave radar systems," in *2017 First IEEE MTT-S International Microwave Bio Conference (IMBIOC)*, 2017.

[26] W. Hu, Z. Zhao, Y. Wang, H. Zhang, and F. Lin, "Noncontact accurate measurement of cardiopulmonary activity using a compact quadrature doppler radar sensor," *IEEE Transactions on Biomedical Engineering*, 2013.

[27] N. T. Phuong Nguyen, P.-Y. Lyu, M. H. Lin, C.-C. Chang, and S.-F. Chang, "A short-time autocorrelation method for noncontact detection of heart rate variability using cw doppler radar," in *2019 IEEE MTT-S International Microwave Biomedical Conference (IMBioC)*, 2019.

[28] D. McDuff, S. Gontarek, and R. Picard, "Remote measurement of cognitive stress via heart rate variability," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014.

[29] A. Melchor Rodríguez and J. Ramos Castro, "Pulse rate variability analysis by video using face detection and tracking algorithms," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015.

[30] A. Melchor Rodríguez and J. Ramos-Castro, "Video pulse rate variability analysis in stationary and motion conditions," *Biomedical engineering online*, 2018.

[31] C. Gu, G. Wang, Y. Li, T. Inoue, and C. Li, "A hybrid radar-camera sensing system with phase compensation for random body movement cancellation in doppler vital sign detection," *IEEE Transactions on Microwave Theory and Techniques*, 2013.

[32] C. Gu, G. Wang, T. Inoue, and C. Li, "Doppler radar vital sign detection with random body movement cancellation based on adaptive phase compensation," in *IEEE MTT-S International Microwave Symposium Digest (MTT)*, 2013.

[33] Q. Lv, L. Chen, K. An, J. Wang, H. Li, D. Ye, J. Huangfu, C. Li, and L. Ran, "Doppler vital signs detection in the presence of large-scale random body movements," *IEEE Transactions on Microwave Theory and Techniques*, 2018.

[34] J. Wang, X. Wang, Z. Zhu, J. Huangfu, C. Li, and L. Ran, "1-D microwave imaging of human cardiac motion: An Ab-Initio investigation," *IEEE Transactions on Microwave Theory and Techniques*, 2013.

[35] M. Mercuri, Y. Liu, I. Lorato, T. Torfs, F. Wieringa, A. Bourdoux, and C. Van Hoof, "A direct phase-tracking Doppler radar using wavelet independent component analysis for non-contact respiratory and heart rate monitoring," *IEEE Transactions on Biomedical Circuits and Systems*, 2018.

[36] K. Qian, C. Wu, F. Xiao, Y. Zheng, Y. Zhang, Z. Yang, and Y. Liu, "Acousticcardiogram: Monitoring heartbeats using acoustic signals on smart devices," in *IEEE Conference on Computer Communications (INFOCOM)*, 2018.

[37] L. Ren, H. Wang, K. Naishadham, Q. Liu, and A. E. Fathy, "Noninvasive detection of cardiac and respiratory rates from stepped frequency continuous wave radar measurements using the state space method," in *IEEE MTT-S International Microwave Symposium*, 2015.

[38] V. R. Radzicki, D. Boutte, P. V. Taylor, and H. Lee, "Standoff CW radar for through-the-wall detection of human heartbeat signatures," in *IEEE Radar Conference (RadarConf)*, 2016.

[39] E. Mogi and T. Ohtsuki, "Heartbeat detection with Doppler radar based on spectrogram," in *IEEE International Conference on Communications (ICC)*, 2017.

[40] H. Shen, C. Xu, Y. Yang, L. Sun, Z. Cai, L. Bai, E. Clancy, and X. Huang, "Respiration and heartbeat rates measurement based on autocorrelation using IR-UWB radar," *IEEE Transactions on Circuits and Systems II: Express Briefs*, 2018.

[41] F. Khan and S. Cho, "A detailed algorithm for vital sign monitoring of a stationary/non-stationary human through IR-UWB radar," *Sensors*, 2017.

[42] L. Ren, Y. S. Koo, Y. Wang, and A. Fathy, "Noncontact heartbeat detection using UWB impulse Doppler radar," in *2015 IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing Systems (BioWireleSS)*, 2015.

[43] T. Sakamoto, R. Imasaka, H. Taki, T. Sato, M. Yoshioka, K. Inoue, T. Fukuda, and H. Sakai, "Feature-based correlation and topological similarity for interbeat interval estimation using ultrawideband radar," *IEEE Transactions on Biomedical Engineering*, 2015.

[44] V. L. Petrovic, M. M. Jankovic, A. V. Lupsic, V. R. Mihajlovic, and J. S. Popovic-Bozovic, "High-accuracy real-time monitoring of heart rate variability using 24 ghz continuous-wave doppler radar," *IEEE Access*, 2019.

[45] TI Inc, "TI product," https://www.ti.com/product/IWR1843, 2020, [Online; accessed 23-August-2020].

[46] G. Shafiq and K. C. Veluvolu, "Surface chest motion decomposition for cardiovascular monitoring," *Scientific reports*, 2014.

[47] MIMO Radar, https://www.ti.com.cn/cn/lit/an/swra554a/swra554a.pdf?ts=1598454063259.

[48] C. S. Marino and P. M. Chau, "High-resolution DOA estimation from synthetic aperture beamforming," in *2005 IEEE Antennas and Propagation Society International Symposium*, 2005, vol. 3A, pp. 279–282 vol. 3A.

[49] M.A. Richards, *Fundamentals of Radar Signal Processing, Second Edition*, McGraw-Hill Education, 2014.

[50] K. Dragomiretskiy and D. Zosso, "Variational mode decomposition," *IEEE Transactions on Signal Processing*, 2014.

[51] G. Ramachandran and M. Singh, "Three-dimensional reconstruction of cardiac displacement patterns on the chest wall during the P, QRS and T-segments of the ECG by laser speckle inteferometry," *Medical and Biological Engineering and Computing*, 1989.

[52] BITALINO Inc, "Bitalino Revolution Plugged Kit BT," https://bitalino.com/en/plugged-kit-bt, 2020, [Online; accessed 23-August-2020].