ViMo: Multi-person Vital Sign Monitoring using Commodity Millimeter Wave Radio

Fengyu Wang, *Student Member, IEEE*, Feng Zhang, *Member, IEEE*, Chenshu Wu, *Member, IEEE*, Beibei Wang, *Senior Member, IEEE*, and K. J. Ray Liu, *Fellow, IEEE*

Abstract—The continuous development of 802.11ad technology provides new opportunities in wireless sensing. In this work, we propose *ViMo*, a calibration-free remote *Vi*tal sign *Mo*nitoring system that can detect stationary/non-stationary users and estimate the respiration rates (RRs) as well as heart rates (HRs) built upon a commercial 60GHz WiFi. The design of ViMo consists of two key components. First, we design an adaptive object detector that can identify static objects, stationary human subjects and human in motion without any calibration. Second, we devise a robust HR estimator, which eliminates the respiration signal from the phase of the channel impulse response (CIR) to remove the interference of the harmonics from breathing and adopts dynamic programming (DP) to resist the random measurement noise. The influence of different settings, including the distance between human and the device, user orientation and incidental angle, blockage material, body movement and conditions of multi-user separation are investigated by extensive experiments. Experimental results show that ViMo monitors user's vital signs accurately, with a median error of 0.19 **BPM** and 0.92 **BPM**, respectively, for RR and HR estimation.

Index Terms—Heart rate estimation, respiration signal elimination, wireless sensing, smoothing spline, 802.11ad technology.

I. INTRODUCTION

Continuous monitoring of respiration as well as heart rate is critical for early detection and prevention of potentially fatal diseases. Current solutions usually require users to wear dedicated devices such as wrist-worn sensors [1] or chest straps [2], which require to contact the human body during the monitoring, making them less convenient and comfortable. With the rapid development of the Internet of Things (IoT), wireless sensing has received increasing attention in recent years because of the ubiquitous deployment of wireless devices [3]–[7]. It has been proved that the presence of human will affect wireless signal propagation, enabling the functionality of wirelessly monitoring human subjects by analyzing the electromagnetic (EM) wave [8]–[11].

Recent works of wireless human vital signs detection leverage the existing communication infrastructure (e.g. WiFi) to provide a pervasive, user-friendly and affordable solution for health status monitoring. Over the past decade, great efforts

F. Wang, F. Zhang, C. Wu, B. Wang and K. J. R. Liu are with the Department of Electrical and Computer Engineering, University of Maryland at College Park, College Park, MD 20742 USA, and also with Origin Wireless, Inc., Greenbelt, MD 20770 USA (e-mail: fywang@umd.edu; zhangfeng0528@gmail.com; cswu@umd.edu; bebewang@umd.edu; kjrliu@umd.edu).

have been put into designing and testing different architectures for robust vital sign monitoring using off-the-shelf WiFi devices [12]–[18]. However, due to the relative low carrier frequency of WiFi systems, the antenna number of 2.4/5GHz WiFi is small, rendering a low spatial resolution. Besides, the narrow bandwidth of WiFi systems results in a coarse range resolution (7.5 meters with bandwidth 20MHz). Therefore, when there is more than one user present, the received radio frequency (RF) signals are reflected by the multiple users and it is hard to extract the vital signs for each of them. Thus, most of the works assume there is a single person [12]–[14], or the breathing rates of different users are distinct [15]–[17]. Moreover, since the perturbation caused by the heartbeat is very small (i.e., $0.2 \sim 0.5$ mm [19]), the embedded heartbeat signal has an extremely low signal-to-noise ratio (SNR). It is extremely difficult, if possible, to use commodity WiFi to estimate the heart rate [18].

In this work, we break down the limitation by leveraging an opportunity in the emerging 60GHz WiFi (e.g., 802.11ad [20]), which is already available in commercial routers [21]. We present **ViMo**, the first system that achieves multi-person stationary/non-stationary detection and vital signs monitoring using an impulse-based commodity 60GHz millimeter wave (mmWave) device. Different from 2.4GHz/5GHz radios, 60GHz WiFi offers high directionality with large phased arrays in small size thanks to millimeter-wavelength and precise time-of-flight measurements brought by the large bandwidth. The advance in 60GHz radios allows higher spatial resolution and range resolution, making it possible to monitor respiration as well as heart rate for multiple persons simultaneously.

However, enabling multi-person vital signs monitoring using 60GHz WiFi is not an easy task. To achieve this goal, we need to deal with multiple challenges. First, it is non-trivial to locate human subjects before vital sign estimation. Due to the high carrier frequency, signals attenuate rapidly over the propagation distance, making it difficult to locate distant targets. Furthermore, the automatic gain control (AGC) module on the chip changes the amplitude of the CIRs over different measurements. To address this challenge, for each measurement, we employ two constant false alarm detectors in range and angle dimension to adaptively estimate the noise level in 3D space, and thus provide an adaptive threshold for target detection.

Second, given the reflections from multiple users and surrounding objects, it is still difficult to differentiate static reflecting objects and stationary/non-stationary users. To overcome this challenge, we devise a novel motion detector by

2327-4662 (c) 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

Authorized licensed use limited to: University of Maryland College Park. Downloaded on November 30,2020 at 19:37:35 UTC from IEEE Xplore. Restrictions apply.

leveraging the sensitivity of CIR phase to the change of travelling distance of EM waves, which can identify the large random body motion (RBM) as well as periodic breathing motion from human subjects and measurement noises from static reflecting objects (e.g., wall, door and furniture).

Third, human chest motions are induced by both respiration and heartbeat, and the distance change caused by heartbeat is magnitude weaker than the respiration signal. Although heart rate is usually higher than the respiration rate, it is hard to distinguish the true heart rate from harmonics of the breathing signal. To address this challenge, we first estimate the waveform of the breathing signal and then eliminate it in the time domain. To further tackle the problem that the heartbeat signal can easily submerge in random measurement noises, we leverage the stationary property of the heart rate and apply dynamic programming (DP) algorithm in spectrogram to obtain an accurate estimation of the heart rate.

We have built a prototype of ViMo by reusing commercial off-the-shelf (COTS) 60GHz WiFi as a radar-like device and conducted extensive experiments to evaluate the performance under different settings, including single-person and multi-person scenarios, LOS and NLOS conditions, etc.. Experimental results show that ViMo achieves accurate estimations, with a median error of 0.19 BPM and 0.92 BPM, respectively, for RR and HR estimation. In addition, ViMo detects multiple users precisely, with a detection rate of 97.86%. We believe ViMo takes an important step towards practical multi-person vital sign monitoring via 802.11ad radios.

The rest of the paper is organized as follows. We review the related works in Section II. We overview the system and theoretical model in Section III, followed by multi-person detection in Section IV and heart rate estimation in Section V. The performance is evaluated in Section VI. We conclude the work in Section VII.

II. RELATED WORK

The RF-based vital signs monitoring utilizes wireless signal to track chest movement, which can mainly been categorized as 2.4GHz/5GHz WiFi based and radar based method.

2.4GHz/5GHz WiFi based: Compared with the conventional methods which require a user to wear dedicated sensors [22], WiFi-based monitoring solutions can use ubiquitous off-the-shelf devices to estimate vital signs contactlessly. Due to the availability of the received signal strength (RSS) measurement on most WiFi devices, UbiBreathe [23] is proposed to estimate respiration rate using RSS. However, since RSS is not sensitive to the minute chest movements, the setting should be well designed to get a good accuracy. Compared to RSS, channel state information (CSI) is a fine-grained information that can portrait the EM wave propagation and is more sensitive to the minute chest movement [12]–[17]. However, due to the omni-directional propagation and the narrow bandwidth, it is impossible to isolate each individual's breathing signal. So most of the previous works either study single person scenario [12]–[14] or assume the breathing rates of each individuals are distinct [15]–[17]. Besides, since the perturbation caused by the heartbeat ($0.2 \sim 0.5$ mm) [19] is

much smaller than the wavelength of 2.4/5GHz WiFi system ($60 \sim 120$ mm), the phase change caused by the heartbeat is very small, resulting in a low SNR. Thus, most of the conventional WiFi-based systems cannot estimate heart rates.

Radar based: To overcome the coarse range as well as spatial resolution limitation in WiFi-based systems, researchers try to build dedicated radar systems to remotely monitor vital signs. Ultra-wideband (UWB) [24]–[26] and Doppler radar [27]–[36] either directly measure the distance between chest and device or the relative speed change of chest movement to get the estimation of chest displacement caused by vital signs. However, the assumption of a single user in these works limits the further deployment of the system. Frequency Modulated Carrier Waves (FMCW) is built in [37] [38] to measure both the respiration rate (RR) and heart rate (HR). Leveraging the fact that different users may locate in distinct bins (a.k.a, range buckets), the vital signs of multiple people can be monitored simultaneously. Finer spatial resolution is achieved in mmVital [39] by using a pair of horn antennas, which investigates the use of 60GHz mmWave signal to simultaneously monitor vital signs in a multi-user case. However, most of these works [25]–[28], [37]–[39] try to directly utilize frequency analysis and bandpass filter (BPF) to estimate the heart rate. These methods are easy to fail when the subject's heart rate is close to the respiration harmonics. The polynomial fitting has been used [30] [31] [35] to remove respiration motion. However, the order needs to be carefully selected by empirical experience, and under-fitting or over-fitting can be easily triggered when the experimental setting is changed (e.g., change of sampling rate or window length).

In this work, we present ViMo, a multi-user **Vi**tal sign **Mo**nitoring system using a commercial 60GHz WiFi. The system can locate stationary/non-stationary users without cumbersome calibration. Furthermore, we eliminate the waveform of the breathing signal before estimating the heart rates in the time domain by using the smoothing spline, which can get rid of the harmonics of the respiration without dedicated choose of hyperparameter as in polynomial fitting. To further reduce the impact of the measurement noise and make robust estimations of heart rate, ViMo leverages the stationary property of heart rate and utilizes both the time and frequency diversity for heart rate estimation.

III. SYSTEM DESIGN AND THEORETICAL MODEL

A. System Overview

ViMo is a wireless system that can accurately detect human subjects and estimate their vital signs by using purely the reflections of RF signals off the users' bodies. The processing flow of ViMo is shown in Fig. 1.

Enabling multi-person contactless vital sign monitoring using 60GHz WiFi faces several challenges. First, due to the fast attenuation of 60GHz RF signal [40], the strength of signal reflected at a large distance is much smaller than that at a short distance. Therefore, it is hard to detect human subjects without prior calibration, let alone detecting the stationary/non-stationary status of human subjects. Second, the minute heartbeat signals are easily corrupted by measurement
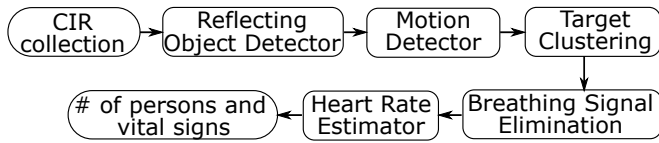
Fig. 1: An overview of ViMo

(a) Coordinate (b) Phase signal

Fig. 2: Coordinate system and typical signal of vital sign

noises and concealed by the large scale respiration signals. Thus, dedicated systems should be designed to resist the interference from respiration and measurement noises when estimating the heart rate.

In order to detect human subjects at various distances, we apply a *reflecting object detector* that adaptively estimates the noise level at various distances and thus detects the presence of reflecting objects. To further differentiate the human subjects from static objects, we design a *motion detector* that identifies static objects, stationary human subjects and human with large body motion. A *target clustering* module is implemented to further identify the number of human subjects and their respective locations. Moreover, to make a robust estimate of the heart rate, we first devise a *breathing signal eliminator* to reduce the interference from the respiration signal after the breathing rate is estimated. The eliminator can remove the harmonics of the breathing signal, as well as deal with the spread of the breathing frequency component when the breathing period slightly changes. To tackle with the random measurement noise, we leverage the stationary property of the heart rate and apply dynamic programming to estimate the heart rate utilizing both the frequency and time diversity.

B. CIR Modeling with Vital Sign Impact

Assume the travelling distance of the EM wave reflected by human chest is $d(t)$, then the CIR between Tx antenna m and Rx antenna n can be expressed as

$$h_{m,n}(t) = a_{m,n}(t) \exp(-j2\pi \frac{d_{m,n}(t)}{\lambda_c}), \quad (1)$$

where $a_{m,n}(t)$ is the complex channel gain, λ_c denotes the wavelength of the carrier. Due to the modulation of the vital signs, i.e., respiration and heartbeat, $d_{m,n}(t)$ appears to be a combination of two periodic signals, which can be further expressed as

$$d_{m,n}(t) = d_0(m,n) + s_r(t) + s_h(t), \quad (2)$$

where $s_r(t)$ and $s_h(t)$ denote the distance change due to respiration and heartbeat, and $d_0(m,n)$ denotes the constant travelling distance. Considering the fast-time resolution of the device, the reflected signal will fall into the l-th tap if we have $d_0(m,n) = lT_s c + \Delta d(m,n)$, where $T_s = 1/B$ denotes the fast time resolution and B stands for the system bandwidth. $\Delta d(m,n)$ denotes the distance residual and we have $\Delta d(m,n) = d_0(m,n) - lT_s c < T_s c$.
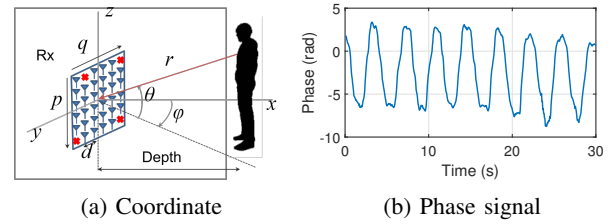
If the reflected signal falls into the l-th tap of the measured CIR with residual $\Delta d(m,n)$, the CIR at tap l, denoting as $\mathbf{h}_l(t) = [h_{1,1}(t), h_{1,2}(t), \ldots, h_{M,N}(t)]^T$, can be expressed as

$$\mathbf{h}_l(t) = \mathbf{a}(t) \odot \exp(-j2\pi \frac{\mathbf{\Delta d} + s_r(t) + s_h(t)}{\lambda_c})$$
$$= \tilde{\mathbf{a}} \exp(-j2\pi \frac{s_r(t) + s_h(t)}{\lambda_c}), \quad (3)$$

where $\mathbf{\Delta d} = [\Delta d(1,1), \Delta d(1,2), \ldots, \Delta d(M,N)]^T$, $\mathbf{a}(t) = [a_{1,1}(t), a_{1,2}(t), \ldots, a_{M,N}(t)]^T$, and $\odot$ denotes elementwise product. We assume $\mathbf{a}(t)$ is time-invariant due to the tiny movement of the subject, and the common phase shift is absorbed in the term $\tilde{\mathbf{a}}$.

To further get the CIR at a specific angle, ViMo performs conventional beamforming at both Tx and Rx. In specific, at Tx side, to create a beam toward to angle (θ, ϕ), we perform conventional beamforming, where the coefficient of the m-th antenna of steering vector $\mathbf{s}_{\text{Tx}}$ is

$$s_{\text{Tx},m}(\theta, \phi) = \exp\left(-j2\pi \frac{d_{m,x} \cos\theta \sin\phi + d_{m,y} \cos\theta \cos\phi}{\lambda_c}\right). \quad (4)$$

$d_{m,y}$ and $d_{m,z}$ are the horizontal and vertical distance between antenna m and the origin, as shown in Fig. 2 (a). Similarly, at Rx side, to detect the angle of arrival (AoA) of the signal, we use the conventional beamforming, where the coefficient of the n-th antenna of steering vector $\mathbf{s}_{\text{Rx}}$ is

$$s_{\text{Rx},n}(\theta, \phi) = \exp\left(-j2\pi \frac{d_{n,x} \cos\theta \sin\phi + d_{n,y} \cos\theta \cos\phi}{\lambda_c}\right). \quad (5)$$

$d_{n,y}$ and $d_{n,z}$ are the horizontal and vertical distance between antenna n and the origin. The CIR after performing beamforming can be expressed as

$$h_{\theta,\phi,l}(t) = \mathbf{s}^H(\theta, \phi)\mathbf{h}_l(t) + \epsilon(t), \quad (6)$$

where $\epsilon(t)$ stands for additive white Gaussian noise which is independent and identically distributed (I.I.D) for different links. $\mathbf{s}^H(\theta, \phi)$ is the steering vector pointing to the direction (θ, ϕ), which can be expressed as the vector form of the Kronecker product between $\mathbf{s}_{\text{Rx}}$ and $\mathbf{s}_{\text{Tx}}$, i.e.,

$$\mathbf{s}(\theta, \phi) = \mathbf{Vec}\left[\mathbf{s}_{\text{Tx}}(\theta, \phi) \otimes \mathbf{s}_{\text{Rx}}(\theta, \phi)\right]. \quad (7)$$

It is apparent that the phase of the CIR measurement changes periodically in slow time due to the periodic motions of respiration and heartbeat, as shown in (3). Fig. 2 (b) shows a typical phase signal containing vital signs collected by our system.
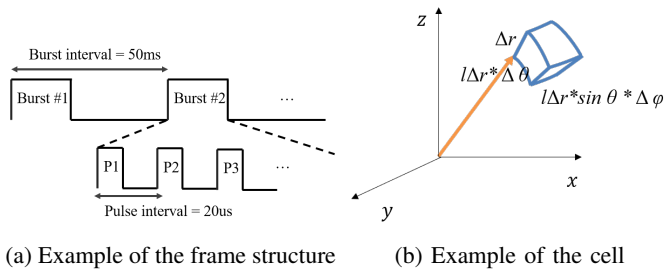
(a) Example of the frame structure   (b) Example of the cell

Fig. 3: Example of the frame structure and the cell

Fig. 4: CFAR window.

In ViMo, we perform on-chip beamforming at both Tx and Rx. The frame structure is shown in Fig. 3 (a). In each burst, Tx will send a series of known pulses toward to different sectors (θ, ϕ), and Rx will also perform beamforming with corresponding steering vector towards to angle (θ, ϕ). Therefore, the space can be separated as cells, as shown in Fig. 3 (b). The volume of each cell can be approximated as $\Delta V \approx \Delta r * (l\Delta r\Delta\theta) * (l\Delta r \sin\theta\Delta\phi)$, where $\Delta r = \frac{c}{2B}$ is the range resolution of the device, and l is the tap index. $\Delta\theta$ and $\Delta\phi$ denote the elevation and azimuth separation respectively. Thus, in each burst, we will get information of all cells. Considering the volume of each cell and the size of the human subject, each cell is assumed to only contain one person[1].

IV. Target Detection

Since various indoor objects (e.g., wall, desk, etc.) reflect the EM wave, before starting monitoring vital signs, we first need to detect human subjects in the vicinity of the Tx and the Rx. Note that the human subjects may have body motion and thus will change his/her location in the long run, ViMo divides the duration of measurements into multiple blocks, where each block consists CIR measurements of W seconds. Two adjacent blocks overlap by $W - W_s$ seconds, where W_s is the sliding window length.

A. Detecting Reflecting Objects

Since the RF signal at 60GHz attenuates severely with distance [40], the reflected energy from the same object varies with distance. To locate the human subject, we first need to identify which spatial cell has reflecting objects.

The CIR measurement for the case when there is no reflecting object and the case when there is a static reflecting object at cell (θ, ϕ, l) can be expressed as

$$h_{\theta,\phi,l}^{\text{empty}}(t) = \epsilon(t), \tag{8}$$

and

$$h_{\theta,\phi,l}^{\text{static}}(t) = \mathbf{s}^H(\theta,\phi)[\mathbf{a} \odot \exp(-j2\pi\frac{\mathbf{d}_0}{\lambda_c})] + \epsilon(t), \tag{9}$$

[1]Note that, in 60GHz WiFi, the multipath effect is not as severe as in commercial 5GHz WiFi. Due to the large bandwidth, each range tap corresponds to probably 1-2 multipaths [41]. By further utilizing the beamforming on both Tx and Rx, the angular separation of each cell is small enough such that the multipaths corresponding to each cell, if any, are similar. Therefore, multipath signals should exhibit similar periodicity in the same cell. For vital sign monitoring, what we value most is the periodicity of the vital signs. Thus, even there are multipaths in the same cell, they will not affect the estimation performance.
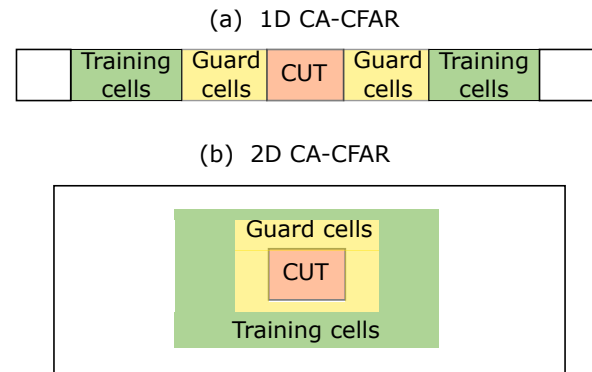
respectively. It is obvious that the power response when there is a reflecting object is much larger than the empty tap. However, it is impossible to find a universal predefined threshold for target detection. According to the propagation laws of EM wave, for the same reflecting object, a shorter distance corresponds to a larger reflecting energy. Furthermore, due to the automatic gain control (AGC) module, the amplitude of the CIRs will change for different measurements.

In order to find the adaptive power threshold for each block, ViMo utilizes constant false alarm rate (CFAR) algorithm [42] for target detection. The power of the noise level for the cell under test (CUT) is estimated by averaging the power of neighboring cells. Furthermore, the guard cell is used to avoid corrupting estimates with power from the CUT itself.

In specific, for each block, the input of CFAR detector is the time-averaging amplitude of all the CIR measurements, i.e., $h(\theta, \phi, l) = \frac{1}{WF_s}\sum_t |h_{\theta,\phi,l}(t)|$, where F_s is the sampling rate. Considering the attenuation property of EM wave, where the reflected signal strength at a different distance of the same object will be different, to determine the range of the reflecting objects, 1D-CFAR is adopted, as shown in Fig. 4 (a). For each sector (θ, ϕ), ViMo convolves CIR measurements $h_{\theta,\phi}(l)$ with the CFAR window to get the estimation of noise level $\hat{n}_{\theta,\phi}(l)$. A scaling factor α is applied to scale the estimated noise level. The detection threshold is thus set to be $\alpha\hat{n}_{\theta,\phi}(l)$, and the taps with reflecting objects should be those whose amplitude are above the detection threshold, as shown in Fig. 5. To determine the noise level at the direction (θ, ϕ), we further employ 2D-CFAR for tap l, where the noise level $\hat{n}_l(\theta, \phi)$ is estimated by convolving CIR measurements $h_l(\theta, \phi)$ with the 2D-CFAR window as shown in Fig. 4 (b). Scaling factor β is applied to scale the estimated noise level. The reflecting object should be in the cell (l, θ, ϕ) whose CIR measurement $h(\theta, \phi, l)$ is above detection threshold $\alpha\hat{n}_{\theta,\phi}(l)$ and $\beta\hat{n}_l(\theta, \phi)$ simultaneously. Here, we define the indicator of reflecting object $\mathbf{1}_R(l, \theta, \phi)$ as

$$\mathbf{1}_R(l, \theta, \phi) = \mathbb{1}\{|h(\theta, \phi, l)| > \max(\alpha\hat{n}_{\theta,\phi}(l), \beta\hat{n}_l(\theta, \phi))\}, \tag{10}$$

where $\mathbb{1}\{\cdot\}$ is the indicator function. Fig. 7 shows the cells with reflecting objects in blue circles.
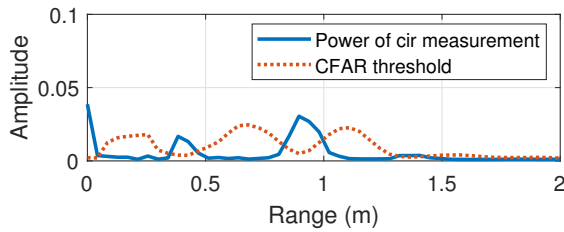
Fig. 5: Example of 1D-CFAR.

B. Motion Detector

Although the CFAR detector can identify which cell is occupied by reflecting objects, it cannot differentiate whether the reflection comes from human or not. Note that human subjects always accompany motion (either from breathing or RBM), which is a specific characteristic different from static objects, we can design a motion detector to identify human subjects. Furthermore, most of the wireless vital sign monitoring systems assume there is only one human subject and no RBM during the measurement, and thus the procedure of finding the human subjects is omitted [30]–[36], which is neither natural nor realistic for practical deployment. So in this part, we design a motion detector, which enables ViMo to identify static reflecting objects, stationary human subjects and human with RBM.

1) Static reflecting objects detection: Note that even stationary human subjects can introduce motion due to respiration and heartbeat, and the distance change caused by respiration can be discerned by phase change according to (3), we calculate the variation of the phase of the CIR measurement $V_t(\theta, \phi, l)$ for each candidate cell (θ, ϕ, l) selected by *reflecting objects detector* discussed in Section IV-A, which is defined as

$$V_t(l, \theta, \phi) = \text{Var}_t[\angle h_{\theta,\phi,l}(t)], \tag{11}$$

where $\text{Var}_t[\cdot]$ denotes the variance over parameter t and $\angle$ denotes the phase of a complex value. As shown in (9), for a static reflecting objects, $V_t(\theta, \phi, l)$ would be small, but for the cell with human subjects, either respiration or RBM will contribute a large $V_t(\theta, \phi, l)$. ViMo utilizes a predefined threshold $\eta_{\min}$ to identify a static reflecting objects if $V_t(\theta, \phi, l) < \eta_{\min}$ [2]. The phase signal and its variance are shown in Fig. 6 (a) and Fig. 6 (d) respectively.

2) Stationary human subjects detection: For a stationary human subject, periodic breathing signal can be observed in the phase measurement according to (1), and Fig. 6 (b) gives an example of a phase measurement with a stationary subject. A bin with a stationary subject would have $V_t(\theta, \phi, l) > \eta_{\min}$ and a periodic phase signal whose frequency within $[b_{\min}, b_{\max}]$.

Note that spectrum analysis is widely used to evaluate the period of respiration in previous works [44]. However, the frequency resolution is $\Delta f = \frac{60}{W}$ breath per minute (BPM), where W is the window length in seconds. Therefore,

---

[2]The chest movement caused by tidal breathing ranges from 4-12mm [43]. Considering the wavelength of the 60GHz WiFi, the phase variance is 6.3-56.8. In ViMo, we set $\eta_{\min} = 3$.

to get an acceptable estimation accuracy of the respiration rate, the window length should be long enough, which will cause a large delay. In our system, we adopt a statistical approach by examining the auto-correlation function (ACF) of the candidate CIR phase to evaluate the periodicity.

Here we denote the time-variant part of CIR phase measurement as

$$y(t) = s_r(t) + s_h(t) + n(t), \tag{12}$$

where $n(t)$ is the random phase offset introduced by noise, and is also a random variable independent in time instances. Thus the ACF of $y(t)$ can be calculated as

$$\rho(\tau) = \frac{\text{Cov}[y(t), y(t+\tau)]}{\text{Cov}[y(t), y(t)]}, \tag{13}$$

where τ denotes the time lag, and $\text{Cov}[\cdot]$ denotes the convariance operator. Assume that the distance change caused by heartbeat $s_h(t)$ is uncorrelated with the distance change caused by respiration $s_r(t)$, then $\rho(\tau)$ can be expressed as

$$\rho(\tau) = \frac{\text{Var}[s_r(t)]}{\text{Var}[y(t)]}\rho_r(\tau) + \frac{\text{Var}[s_h(t)]}{\text{Var}[y(t)]}\rho_h(\tau) + \frac{\text{Var}[n(t)]}{\text{Var}[y(t)]}\rho_n(\tau), \tag{14}$$

where $\text{Var}[y(t)] = \text{Var}[s_r(t)] + \text{Var}[s_h(t)] + \text{Var}[n(t)]$. $\rho_r(\tau), \rho_h(\tau)$ and $\rho_n(\tau)$ denote the ACF of respiration, heartbeat and noise respectively. Since we have $\text{Var}[s_r(t)] \gg \text{Var}[s_h(t)]$ and $\text{Var}[s_r(t)] \gg \text{Var}[n(t)]$, we have the approximation that $\rho(\tau) \approx \rho_r(\tau)$. The ACF will have a definite peak at a certain delay which corresponds to the breathing cycle as shown in Fig. 6 (e).

3) Motion detection: Random body motion (RBM) has been one of the most difficult technical challenges to wireless vital sign monitoring. Compared with the millimeter-scale chest movement caused by heartbeats, the scale of RBM can be tens of centimeters. The time-variant part of CIR phase measurement with RBM can be modelled as

$$y(t) = s_m(t) + s_r(t) + s_h(t) + n(t), \tag{15}$$

where $s_m(t)$ is the distance change caused by motion. Fig. 6 (c) shows an example of the phase measurement with motion. Note that when the scale of RBM is much larger than the respiration signal, the variation $\text{Var}[s_m(t)] \gg \text{Var}[s_r(t)] \gg \text{Var}[s_h(t)]$, and thus $V_t(l, \theta, \phi) > \eta_{\max}$, where $\eta_{\max}$ is a predefined threshold. When the subjects have moderate RBM, the variance of phase may be within the threshold, however, since RBM lacks periodicity in most case, we **cannot** observe a peak in $\rho(\tau)$ as the stationary case as shown in Fig 6 (f). Therefore, we have the motion indicator $\mathbb{1}_M(\cdot)$ defined as

$$\mathbb{1}_M(\theta, \phi, l) = \mathbb{1}(V_t(\theta, \phi, l) > \eta_{\max} \cup \rho(\tau_b) < \Gamma), \tag{16}$$

where $\mathbb{1}(\cdot)$ is the indicator function, τ_b is the first peak of ACF $\rho(\tau)$, and Γ is a predefined threshold.

C. Cell Merging/Clustering

Due to the fact that more than one cells have the RF signals reflected off a single human subject, a target clustering method should be employed before determining the target number

(a) Phase with static object

(b) Phase with stationary subject

(c) Phase with RBM

State	Phase Variance
Static object	0.099
Stationary subject	12.306
Subject with RBM	199.975

(d) Phase Variance of signals in (a)-(c)
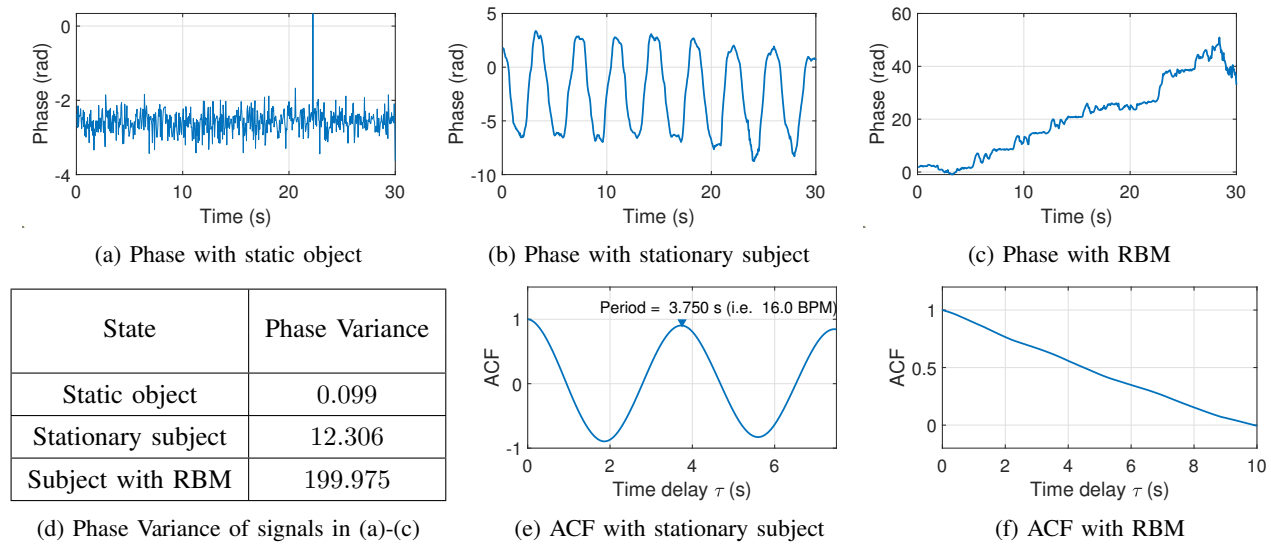
(e) ACF with stationary subject

(f) ACF with RBM

Fig. 6: Example phase and ACF for target detection. (a)-(c) are the phase signals with static object, stationary subject, and RBM subject, (d) shows the phase variance of signals in (a)-(c), (e) and (f) show the ACF of (b) and (c) respectively. The reference RR of the stationary human subject is 16BPM.
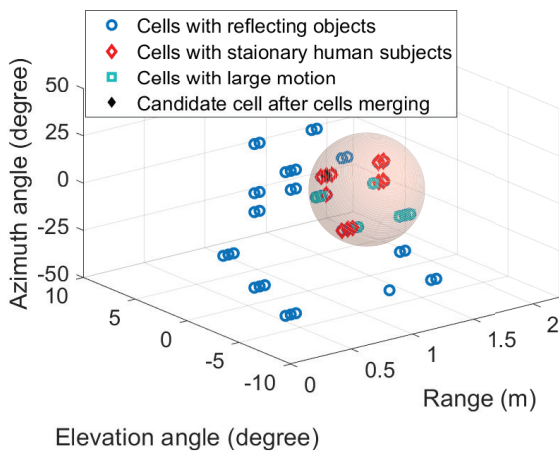

Fig. 7: Example of cell merging.

and vital sign monitoring. Considering the size of the human body, we can merge them into a cluster if the spacial distance between these cells is within the threshold $d_{\min}$. In ViMo system, we set the $d_{\min}$ as 0.8 m considering the typical body volume of human subjects. The cluster center of stationary cells is the cell with the largest ACF peak, corresponding to the cell with the human chest. The center of the RBM cells for each cluster is the cell with the largest $V_t(\theta, \phi, l)$, corresponding to the cell with the largest body motion. Note that even for a stationary person, he/she can have body motion from the body part away from the chest. So when the distance between stationary cells and RBM cells is smaller than the threshold $d_{\min}$, then these cells belong to the same person, and the center of the cluster should be the representative of stationary cells. The number of people is estimated by the number of clusters, where the location of each person is the center of its corresponding cluster.

Fig. 7 shows an exemplar for cells merging, where the ground truth is that a human subject sits at 1 m away from the device in a typical indoor office. The reflecting objects detected by the CFAR detector is shown in blue circles. The motion detector further differentiates cells with stationary subjects and RBM, shown as red diamonds and green squares respectively. The representative of the target is shown in solid black diamond.

V. HEART RATE ESTIMATION

In ViMo, we enable the heart rate estimation module once a stationary subject has been detected. Since we check periodicity using ACF to determine whether the cell contains a stationary respiration signal as shown in Section IV-B2, we can easily determine the breathing cycle by finding the peak location τ_r of $\rho(\tau)$, and the breathing rate should be

$$f_r = \frac{60}{\tau_r} \qquad (17)$$

breath per minute (BPM).

Note that heartbeats can introduce minute movements of the chest [19], which can be detected as small peaks in the unwrapped phase as shown in Fig. 8 (a). Past works [37] [39] [45] try to directly utilize frequency analysis and bandpass filter (BPF) to estimate the heart rate. However, due to the harmonics introduced by respiration, it is easy to pick up the wrong peak for estimation as shown in the blue line in Fig. 8 (c). Thus, in order to get a higher estimation accuracy, we first eliminate breathing signal before heart rate estimation.

A. Breathing Interference Elimination

Eliminating the breathing signal can improve the signal-to-interference-plus-noise ratio (SINR) of the heartbeat signal, and thus improve the estimation accuracy. The polynomial
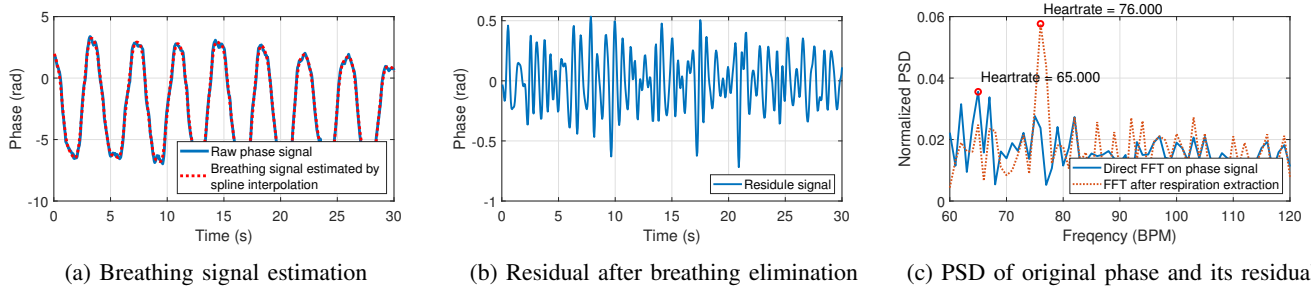
Fig. 8: Example of breathing interference elimination. The ground truth of heart rate is 76 BPM. (a) shows the smoothing spline estimate of breathing signal, (b) shows the residual after breathing signal elimination, (c) shows the power spectrum density (PSD) of original phase and residual signal.

fitting has been used [30] [31] [35] to remove respiration motion. However, one of the main drawbacks of the polynomial fitting is the order selection. In previous works, the order is carefully selected by empirical experience, but under-fitting or over-fitting can be easily triggered when the experimental setting is changed (e.g., change of sampling rate or window length). Besides, the elimination effect is also related to the breathing rate. In other words, in order to achieve a similar elimination effect, the polynomial order should adapt to the user's breathing rate, which is not practical for robust daily deployment. To avoid this effect, ViMo adopts smoothing spline to estimate the breathing signal.

Let $\{t_i, y(t_i) : i = t_0, t_0 + T_s, \cdots, t_0 + W\}$ to be a set of observation in the current window, where $T_s = \frac{1}{F_s}$ is the time interval between two adjacent samples, t_0 is the initial time of the observation window, and W is the window length. Compared to the heartbeat signal, the respiration signals have larger distance change and lower frequency, thus, the estimate of the breathing signal $s_r(t)$ should be the solution of

$$\min_{\hat{f}} \sum_{i=t_0}^{t_0+W} \{y(t_i) - \hat{f}(t_i)\}^2 + \lambda \int \hat{f}''(t)^2 \, dt, \quad (18)$$

where $\lambda \geq 0$ is a smoothing parameter. The second term evaluates the smoothness of a function. The smoothing parameter controls the trade-off between fidelity to the data and smoothness of the function estimate. $\hat{f}$ is the estimate of $s_r(t)$, defined as

$$\hat{f}(t) = \sum_{t=t_0}^{t_0+W} \hat{f}(t_i) f_i(t), \quad (19)$$

where $f_i(t)$ are a set of spline basis function. In this work, we use B-spline as the spline basis, and the detail of the definition can be referred to [46].

To get the optimum solution of (18), we first define the vector $\hat{\mathbf{m}} = [\hat{f}(t_0), \cdots, \hat{f}(t_0 + W)]^\top$, and the roughness penalty has the form

$$\int \hat{f}''(t)^2 \, dt = \hat{\mathbf{m}}^\top \mathbf{A} \hat{\mathbf{m}}, \quad (20)$$

where the elements of $\mathbf{A}$ are $\int f_i''(t) f_j''(t) \, dt$. Therefore, the penalized sum-of-squares can be written as

$$\min_{\hat{\mathbf{m}}} \{\mathbf{y} - \hat{\mathbf{m}}\}^\top \{\mathbf{y} - \hat{\mathbf{m}}\} + \lambda \hat{\mathbf{m}}^\top \mathbf{A} \hat{\mathbf{m}}, \quad (21)$$
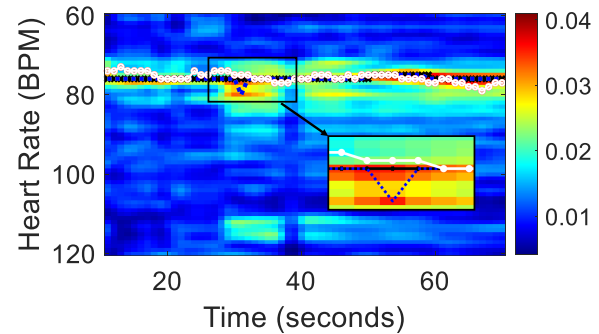

Fig. 9: Spectrogram of residual signal.

where $\mathbf{y} = (y(t_0), \cdots y(t_0+W))^\top$. The minimizer of problem (21) is thus

$$\hat{\mathbf{m}}^* = (\mathbf{I} + \lambda \mathbf{A})^{-1} \mathbf{y}. \quad (22)$$

The heartbeat after elimination of the breathing signal is thus

$$\hat{s}_h(t) = y(t) - \hat{\mathbf{m}}^{*\top} \mathbf{f}(\mathbf{t}), \quad (23)$$

where $\mathbf{f}(\mathbf{t})$ is the vector form of spline basis functions.

The dashed line in Fig. 8 (a) shows the estimation of the breathing signal. After breathing signal elimination, the spectrum of the residual signal after applying a band-pass filter (BPF) with passing band $[h_{\min} \ h_{\max}]$ is shown in the orange dashed line in Fig. 8 (c). The spectrum of the phase without eliminating respiration signal using the same BPF is shown in the blue solid line. It is obvious that the signal-to-interference-plus-noise ratio (SINR) of the heartbeat signal after breathing elimination is boosted. Specifically, the SINR is boosted from 1.65 dB to 5.65 dB by eliminating the respiration signal [3].

B. Heart Rate Estimation using Spectrogram

Breathing signal elimination can enhance the SINR of the heartbeat signal, and thus, increase the accuracy of heart rate estimation. However, the random measurement noises can still corrupt the estimation at some time instances. To further

---

[3]The SINR is calculated from the frequency domain. In specific, the power of the heartbeat signal can be obtained from the power spectrum, i.e., the power at the frequency point w.r.t. the true heart rate. The power of interference and noise is the summation of the power at the rest of the frequency points.

increase the estimation accuracy, in ViMo, we leverage the stationary property of heart rate and utilize the diversity in both frequency and time domains for reliable estimation.

Note that the heart rate can smoothly change over time, we model the heart rate as a Markov process, where the variation of heart rate between two adjacent time bins follows a normal distribution $\mathcal{N}(0, \sigma^2)$, and the probability density function (PDF) is denoted as $p(f)$. After breathing signal elimination, we perform *Fast Fourier Transform* (FFT) on the residual and concatenate the PSD of each window to get a spectrogram as shown in Fig. 9.

Since the operation of FFT automatically discretizes the continuous frequency in the range of $[h_{\min}, h_{\max}]$ [4] into $|\mathcal{Q}|$ frequency components, where $|\mathcal{Q}|$ means the cardinality of set $\mathcal{Q}$, the heart rate can be modelled as a Markov chain, and the transition probability matrix is denoted as $\mathbf{P} \in \mathbb{R}^{|\mathcal{Q}|} \times \mathbb{R}^{|\mathcal{Q}|}$, which is defined as

$$\begin{aligned} \mathbf{P}(q, q^{'}) &= \mathbf{P}(g(n) = q^{'} | g(n-1) = q) \\ &= \int_{(q^{'} - q - \frac{1}{2}) * \Delta f}^{(q^{'} - q + \frac{1}{2}) * \Delta f} p(f) \mathbf{d}f, \end{aligned} \quad (24)$$

where $\forall q, q^{'} \in \mathcal{Q}$. Here, $g : [1, N] \rightarrow \mathcal{Q}$ is a mapping indicating the frequency component at the given time, and N is the total time instances of a given spectrogram.

In principle, the heartbeat signal is more periodic than noise and other motion interference. Thus, it is more likely to be observed as peaks in most of the time. Moreover, considering that one's heart rate will not fluctuate a lot within a short period, estimations of heart rates should form a trace that achieves a good balance between frequency power and temporal smoothness.

The most probable heart rate trace can be found by solving

$$\mathbf{g}^{*} = \arg\max_{\mathbf{g}} \; E(\mathbf{g}) - \kappa C(\mathbf{g}), \quad (25)$$

where κ is a regularization factor. $\mathbf{g}$ is denoted as a trace, where

$$\mathbf{g} = (n, g(n))_{n=1}^{N}. \quad (26)$$

$E(\mathbf{g})$ is the power of trace $\mathbf{g}$, defined as

$$E(\mathbf{g}) = \sum_{n=1}^{N} \mathcal{E}(n, g(n)), \quad (27)$$

where $\mathcal{E}(n, q)$ denotes the energy at time bin n and frequency component q. The smoothness of the trace can be evaluated by a cost function $\mathbf{C}(\mathbf{g})$, defined as

$$\mathbf{C}(\mathbf{g}) \triangleq -\log \mathbf{P}(g(1)) - \sum_{n=2}^{N} \log \mathbf{P}(g(n-1), g(n)), \quad (28)$$

where the frequency transition probability $\mathbf{P}(g(n-1), g(n))$ can be calculated by (24). Without loss of generality, we assume a uniform prior distribution, i.e., $\mathbf{P}(g(1)) = \frac{1}{|\mathcal{Q}|}$.

---

[4]In ViMo, we set $h_{\min}$ and $h_{\max}$ as 60BPM and 120BPM respectively.

This problem can be solved by dynamic programming [47]. For clarity, we first define the score at bin (n, q) as the maximum achievable regularized energy, i.e.,

$$\begin{aligned} S(n, q) &= \mathcal{E}(n, q) + \max_{\forall q' \in Q} \{ S(n-1, q^{'}) + \lambda \log \mathbf{P}(q^{'}, q) \}, \\ & n = 2, 3.., N, \; \forall \, q, q^{'} \in \mathcal{Q}, \end{aligned} \quad (29)$$

where $S(1, q) = \mathcal{E}(1, q) + \lambda \log \mathbf{P}(g(1) = q)$. Furthermore, since $S(n, q)$ consider both the smoothness and regularized energy of the previous trace, the process of calculating the score also determines the optimal trace passing through bin (n, q). The entire optimal heart rate trace can be found by backtracking the bins $(N, g^{*}(N))$ that contribute to the maximum score of the last timestamp. For the rest of the heart rate trace in the observation window, i.e., $\forall \, n = N-1, N-2, ..., 1$, we have

$$g^{*}(n) = \arg\max_{\forall q \in Q} \; S(i, q) + \lambda \; \log \mathbf{P}(q, g^{*}(n+1)). \quad (30)$$

The backtracking procedure in (30) gives the optimal trace $\mathbf{g}^{*}$ for a given spectrum, which is the optimal solution for (25). The result of heart rate trace estimation is shown as the black line in Fig. 9, where the reference measured by a gold standard electrocardiogram (ECG) sensor [48] is marked as the white line. The estimation result of directly using the location of the highest peak of the spectrum as the heart rate estimation is shown as the blue dashed line. We can see that when the noise is too large (e.g., at time instance $t = 32$ s), the estimation without dynamic programming will locate at the wrong peak, resulting in a large estimation error. By using dynamic programming, the maximum estimation error decreases from 5BPM to 3BPM, as shown in Fig. 9.

VI. EXPERIMENT EVALUATION

In this section, we evaluate ViMo in practical settings using a commodity 802.11ad chipset in a typical office of size 3.5 m $\times$ 3.2 m as shown in Fig. 10. We embed ViMo in a commodity off-the-shelf 60GHz WiFi [21] as shown in Fig. 10 (a). Specifically, the chipset we used for ViMo has 32 antennas assembled in a 6×6 layout with a form factor of 1.8 cm$\times$ 1.8 cm for both the transmitter (Tx) and the receiver (Rx). The chipset operates at 60GHz center frequency with 3.52GHz bandwidth, providing a range resolution of 4.26 cm. To extract channel impulse response (CIR), the Tx transmits a known pulse composed of a complementary Golay sequence. A Golay correlator is implemented in the Rx hardware and the correlation result corresponds to the CIR. We enroll 8 participants (4 male and 4 female) aging from 22 to 35 for testing. The ground truth is provided by a commercial ECG sensor with a chest strap [48]. The sampling rate of ViMo is 20Hz.

To further evaluate our system, we compare it with the mmVital [39], which is the state-of-art wireless vital sign monitoring system using impulse-based mmWave radio. mmVital leverages the RSS from a pair of horn antenna and finds the highest magnitude peak as well as its adjacent bins in the frequency domain to form a custom narrow BPF, and then counts peaks of the time-domain filtered signal to determine

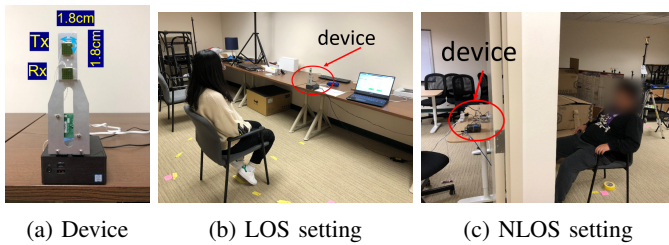(a) Device   (b) LOS setting   (c) NLOS setting

Fig. 10: Experiment setup

the breathing rates as well as heart rate. In order to make a fair comparison, same as ViMo, phases of CIRs from the detected cells are used as the input of mmVital algorithm, rather than the coarse information of RSS. To estimate both respiration and heart rate, the adaptive narrow BPF and IFFT are implemented as illustrated in mmVital [39]. The window length for both mmVital and ViMo are set to be 60 seconds, and systems give output every second. Note that mmVital [39] estimates vital signs according to the number of peaks in a time window (i.e., the estimation is an integer), the resolution of its breathing rate as well as heart rate estimation is 1 beat/breath per minute (BPM).

A. Overall performance

We first report the overall performance of ViMo in the measurement of respiration rate (RR) and heart rate (HR). The accuracy is calculated over 17 experimental runs of 3 minutes for all 8 participants. During the experiment, participants wear casual clothes, such as sweaters and shirts, sit in front of the device and breathe normally. Various factors including user and device placement diversity (e.g., distance, orientation, incident angle and blockage), motion interference and multiple users' position are considered.

The detection rate of the system is 97.86% and the overall median error of RR and HR evaluated by ViMo is 0.19 BPM and 0.92 BPM respectively. mmVital achieves similar performance w.r.t. RR, but its median error of HR is 1.6 BPM, 73.91% worse than ViMo. We further compare the performance of ViMo with Vital-Radio [37], which is the representative multi-user vital sign monitoring system build upon FMCW radar. Vital-Radio achieves 98.5% median accuracy of HR estimation (0.95 BPM) when the person is 1 m away from the device, whereas, the median error of HR estimation for ViMo is 0.58 BPM. Note that although the SNR of our device is lower than FMCW radar, ViMo still achieves better performance. The advantage of our system benefits from the breathing signal elimination module and dynamic programming (DP) algorithm when estimating the HR, which can increase the SINR of the heartbeat signal.

Moreover, experimental results show that ViMo can effectively detect stationary/non-stationary state of human subjects, and can make accurate estimates of both RR and HR when slight user motion incurs (e.g., shaking head). Comparing with mmVital, which does not take users' motion into consideration, ViMo makes an important improvement towards practical deployment. The details will be discussed in the following sections.

B. Impact of distance

In this section, we investigate the effect of the distance between the device and human subject on the estimation accuracy. Participants sit at different distances facing the device as shown in Fig. 11 (a). The empirical cumulative distribution function (CDF) of the absolute error of RR and HR estimation are shown in Fig. 11 (b) and Fig. 11 (c) respectively, where the performance of ViMo and mmVital are shown in solid lines and dash lines respectively. To account for the misdetection, we set the estimation to be 0 BPM when the target is missed.

As expected, the performance degrades with distance due to the signal-to-noise ratio (SNR) degradation. The median error for RR of ViMo is within 0.15 BPM when the distance is within 1.5 m and it increases to 0.22 BPM when the distance increases to 2 m. For HR estimation, the median error of ViMo increases from 0.42 BPM to 0.9 BPM when the distance increases from 0.5 m to 2 m. Furthermore, we can see that the degradation of RR estimation is less than the HR estimation due to the higher SNR of the breathing signal.

The CDF of RR estimation using mmVital algorithm is stepwise since the resolution of both ground truth and estimation is 1 BPM. It is obvious that both algorithms achieve similar performance as for RR estimation, but ViMo achieves a higher resolution. Moreover, for HR estimation, ViMo outperforms mmVital for all the 4 settings, and the performance gap becomes larger with the increment of distance. The main reason is that the breathing signal elimination helps to improve the SINR of the heartbeat signal as discussed in Section V-A. Besides, DP algorithm in ViMo also leverages the time diversity besides the frequency diversity to make estimations, which can further alleviate the impact of the measurement noises.

C. Impact of orientation

In this study, we investigate the impact of human orientation on estimation accuracy. The orientation corresponds to the closest part of the user w.r.t. the device as shown in Fig. 12 (a). The distance from the user to the device is set to be 1 m. Fig. 12 (b) and Fig. 12 (c) show the estimation performance of RR and HR respectively.

It is shown that the "front" setting achieves the best performance, whereas, the "back" setting has the worst performance, for both RR and HR estimation. This result is due to the distinct displacement of reflecting part caused by respiration in different orientations. Since smaller displacement means lower SNR of breathing signal, when the displacement is too small, mis-detection occurs. The detection rate when the subject sits facing the device is 100%, and it degrades to 99.06% and 99.37% when the left and right side of the chest facing the device. The detection rate drops to 83.83% when human subjects sit back to the device. It is worth noting that even similar detection rates are achieved when participants sitting at the left and right orientation, the HR estimation performance is distinct, where the "left" setting outperforms the "right" setting. This is due to the physiological structure of human beings, where the vibration caused by the heartbeat
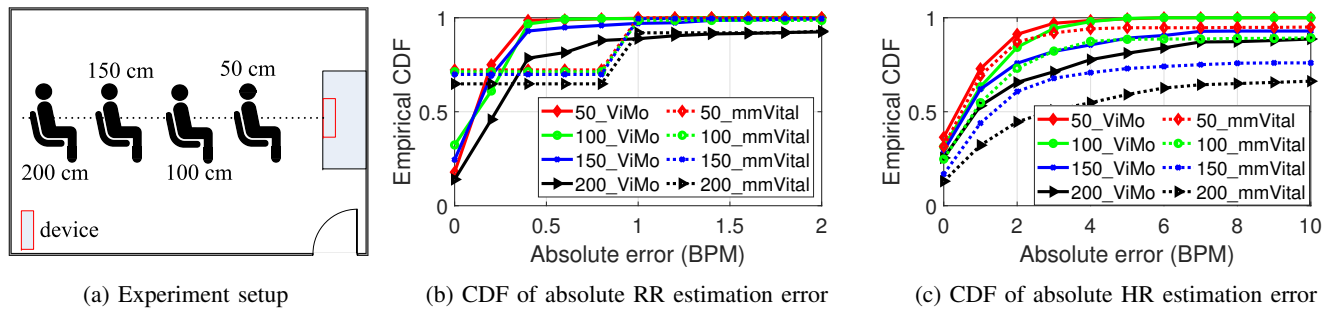
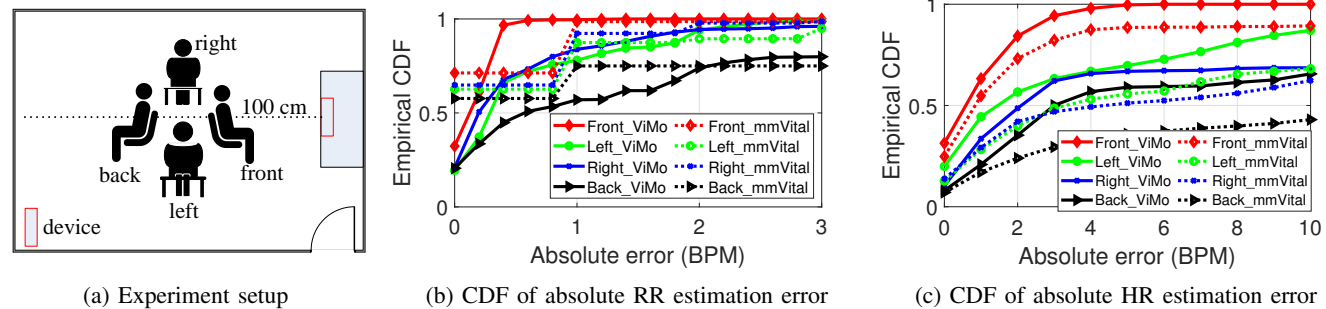Fig. 11: Experiment setup and result for the impact of distance.

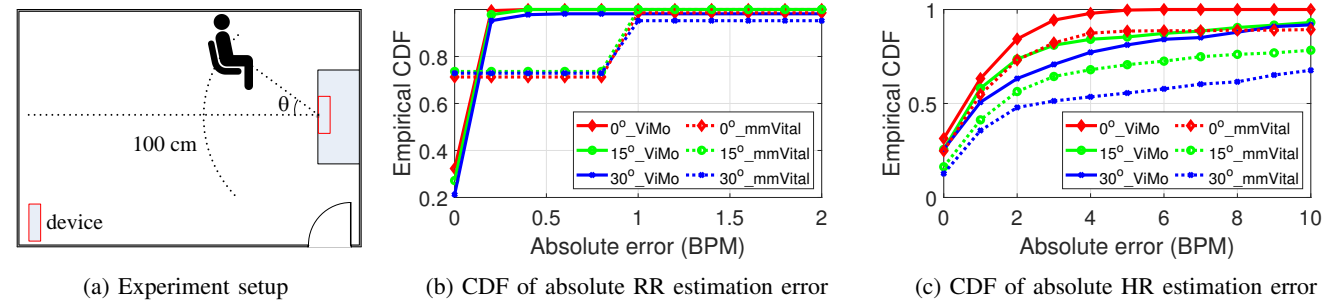Fig. 12: Experiment setup and result for the impact of orientation.

Fig. 13: Experiment setup and result for the impact of incident angle.

is larger on the left side of the chest. Similarly, ViMo has equivalent performance in terms of RR estimation compared with mmVital, however, it has much better performance of HR estimation for all the 4 settings, as shown in Fig. 12 (c).

D. Impact of incident angle

In this part, we investigate the impact of the incident angle on the estimation performance, where human subjects are asked to sit at angles $[0°, 15°, 30°]$, and the distance between human and device is 1 m, as shown in Fig. 13 (a). The CDF of the absolute estimation error of RR and HR with different incident angles are shown in Fig. 13 (b) and Fig. 13 (c) respectively. We can see that for both RR and HR estimation, the accuracy decreases with the increment of the incident angle. The reason is that the reflection loss is dependent on the incident angle, and increment in incident angle increases the reflection loss, rendering lower SNR of the reflected signal. However, since the SNR of the breathing signal is much higher than the heartbeat signal, the performance degradation of RR

estimation is not as severe as HR estimation. Furthermore, we can see that the performance of ViMo is much better compared with mmVital in terms of HR estimation, especially in the case of a large incident angle.

E. NLOS Case

The RR and HR estimation accuracy are evaluated for the through-the-wall case, and the experiment setup is shown in Fig. 14 (a). Participants are asked to sit on the other side of a drywall, and the distance between the device and the human subject is 1 m. The median error of RR estimation increases from 0.15 BPM to 0.25 BPM due to the penetration loss, and the median error of HR estimation increases from 0.6 BPM to 1.4 BPM, as shown in Fig. 14 (b) and Fig. 14 (c) respectively.

In order to further investigate the influence of blocking material (corresponding to different penetration loss), we conduct a set of experiments, where different commonly used materials are used to block the LOS path, as shown in Fig. 15. Since the penetration loss is distinct for different material, the
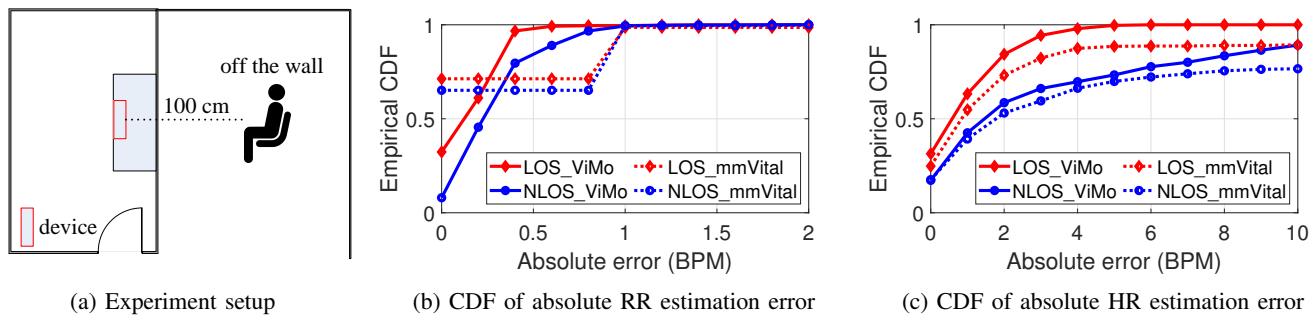
(a) Experiment setup   (b) CDF of absolute RR estimation error   (c) CDF of absolute HR estimation error

Fig. 14: Experiment setup and result for the impact of blockage.

Glass   Cotton   Wood   White board
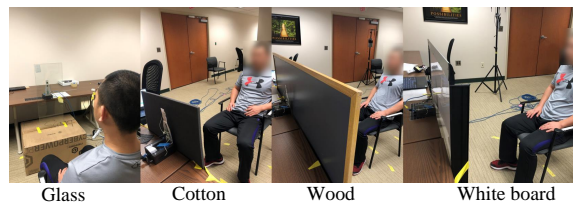
Fig. 15: Experiment setup with different blockage materials

(a) RR estimation   (b) HR estimation

Fig. 16: Impact of user heterogeneity on estimation accuracy

Blocking material	None (LOS)	Glass	Cotton pad	Wood panel	White board	Drywall
Mean RR error (BPM)	0.14	0.23	0.24	0.26	0.28	0.29
Mean HR error (BPM)	1.29	2.66	3.45	4.82	4.85	5.95

TABLE I: Performance with different blockage materials

Body state	Stationary	Shaking head (left-right)	Moving body (left-right)	Speaking
Mean RR error (BPM)	0.14	0.28	0.51	1.22
Mean HR error (BPM)	1.29	4.16	3.06	6.31

TABLE III: Performance for different motion states

performance drop is different. The mean absolute error (MAE) of RR and HR estimation is shown in Tab. I.

Besides, note that clothes can be considered as a cotton layer between the human subject and the device, and the penetration loss is related to the thickness of clothes. To investigate the influence of the thickness of clothes, we conduct a similar experiment as that in Fig. 15, where participants are asked to wear T-shirts, sweaters and winter jackets during the experiment. The median absolute error of RR and HR are shown in Tab. II. We can see that the estimation accuracy decreases with the increment of the thickness of clothes. However, for the general indoor wearing (e.g., T-shirt and sweater), the degradation can be negligible.

F. Impact of user heterogeneity

In this part, we investigate the impact of the user heterogeneity on the performance. The data of all the settings above are used to get the performance of each user. The difference in error distribution can be caused by various factors, such as reflection loss and heartbeat strength, etc. Fig. 16 (a) gives the

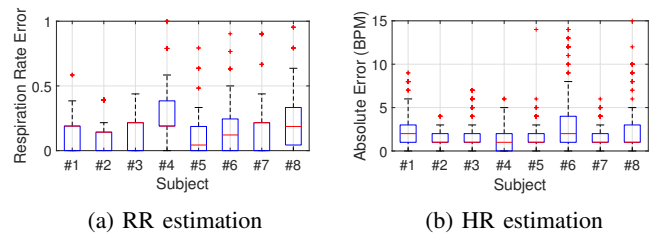| Clothes Type	T-shirt	Sweater	Winter jacket
Median absolute error of RR (BPM)	0.125	0.142	0.164
Median absolute error of HR (BPM)	0.5	0.58	1.13

TABLE II: Impact of thickness of clothes

RR estimation performance, with the maximum median error within 0.2 BPM for all participants. Fig. 16 (b) shows the error distribution of absolute HR error of all 8 subjects, where all of them have a median error within 2 BPM.

G. Impact of body movement

We evaluate the performance of ViMo when users have different motion states. Participants are asked to shake head ($1 \sim 3$ cm) and move the body ($4 \sim 5$ cm) every 20 seconds. The distance from the device to the user is 1 m with incident angle $0°$. We also evaluate the performance when users answer phone with headset (continuously talking). The MAE of RR and HR are shown in Tab. III. All the cases achieve more than 99.7% detection rate, where for the case of moving body, in 27% of the duration we detect large body motion, and thus the vital signs estimation module will not be triggered. As for the time that body motion is within the detection threshold (a.k.a, stationary period), the vital sign estimation module is triggered, and the mean HR error is 3.06 BPM for the case of moving body (the relative error is 4%). However, for the case when people are answering the phone, the chest will involve RBM caused by speaking more frequently, resulting in the worst performance for all the test cases.

H. Multi-user case

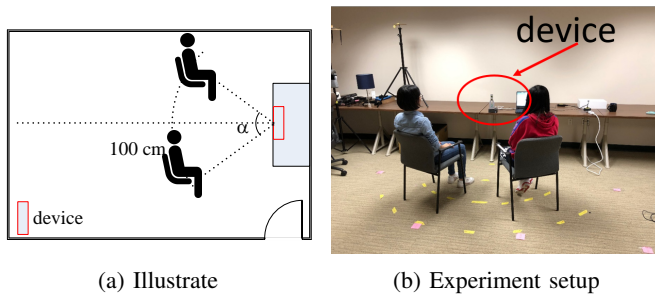In this part, we first study the impact of the angular separation between users, where two users sit at a distance

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.2020.3004046, IEEE Internet of Things Journal

12

(a) Illustrate   (b) Experiment setup

Fig. 17: Experiment setup for the impact of separation angle

Separation angle	DI	FI	Med. error of breathing	Med. error of heart rate
30°	0.84	0	(1.14;0.15)	(2;1)
45°	0.98	0	(0.22;0.14)	(1;1)
60°	1	0	(0.21;0.14)	(1;1)
75°	1	0	(0.21;0.12)	(1;1)

TABLE IV: Performance for different separation angles

of 1 m away from the device with different separation angles as shown in Fig. 17. We define the detection index (DI) of a separation angle as the ratio between the number of samples when the number of detected targets matches the ground truth and the total number of samples. We also define the false-alarm index (FI) of a separation angle as the ratio between the number of samples when the number of detected targets is larger than the ground truth and the total number of samples. Tab. IV shows the median error of RR and HR estimation for both users.

Compared to the single-user scenario, the performance degrades at small separation angles (i.e., 30°), but the performance is similar to the single-user scenario if the separation angels are large enough (i.e., larger than 45°). This is because when the distance of two targets is small enough, the distance of the candidate cells with each user can be smaller than the predefined threshold $d_{\min}$. Thus, the two clusters will be merged together and there will be only one representative cell left, resulting in a mis-detection. Besides, the cells with high SNR signals of one user can be merged with the other user's, therefore, the SNR of the representative cell for vital signs estimation can drop, resulting in degradation of the performance.

To further evaluate ViMo's accuracy for multi-user vital sign estimation, we perform controlled experiments, where we ask 3 users to sit in parallel as shown in Fig. 18. ViMo detects the location of each target and simultaneously estimate their vital signs. When mis-detection happens, we define the relative error as 1 as before. Fig. 19 shows the mean relative accuracy of RR and HR as well as the detection rate at each location. We can see that for all the 3 locations, ViMo achieves the mean accuracy of both RR and HR over 92.8%. As for the detection rate, since the separations between the middle location and the other two locations are not large enough, and the middle location is more distant, the detection rate drops at the middle location. However, the overall detection rate over time is still above 92.7% during the testing.

Note that the system capacity is related to the range as well

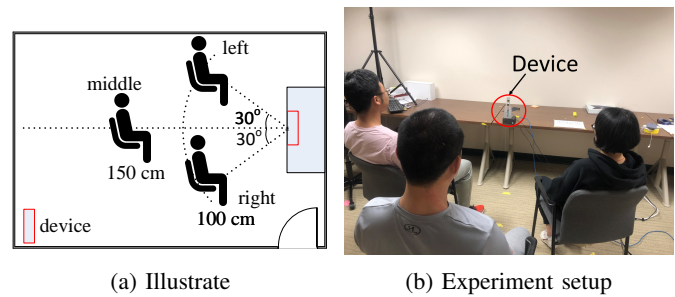(a) Illustrate   (b) Experiment setup

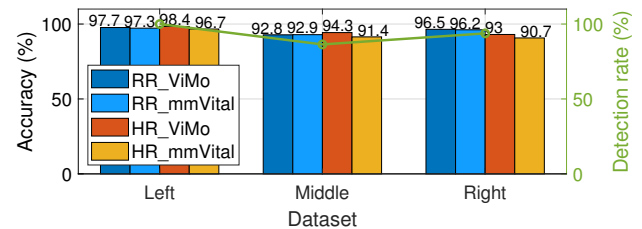Fig. 18: Multi-user experiment setup

Fig. 19: Multi-user accuracy and detection performance.

as the field-of-view (FoV) of the device, and the volume of the human body. Besides, due to the large penetration loss of human subjects, human subjects should not block each other. The FoV of the device is 100°, and the range of the human subjects should be within 2.5 m to ensure acceptable SNR. The human subject is simulated as a circle with $r = 0.4$ m. The location of human subjects when achieving the maximum capacity is shown in black circles in Fig. 20. Note that the human subject with a shorter distance (e.g. blue circle) occupies a larger effective space (e.g., the space between the dashed blue lines), thus, to achieve maximum support, human subjects should locate at the boundary of the coverage area as shown in black circles in Fig. 20. We can see that the system can support at most 5 users at the same time. However, considering the reflection loss increases with the incident angle as shown in Section VI-D, we recommend 3 users setup as shown in Fig. 18.

VII. Conclusion

This paper presents ViMo, a multi-person **Vi**tal sign **Mo**nitoring system using a single commercial 802.11ad device. We devise a multi-object stationary/non-stationary detec-
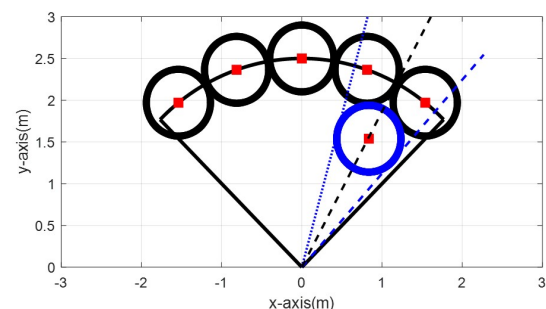

Fig. 20: Human location for maximum support

tion algorithm to locate and count human targets without any prior calibration. In addition to the instantaneous estimating breathing rates using ACF with high accuracy, we further design a robust heart rate estimator, which eliminates the interference of the breathing signal and then estimates the heart rate leveraging both the time and frequency diversity. We evaluate the performance of ViMo by various settings, including NLOS and motion artifacts, the most challenging scenarios for wireless vital signs monitoring. Experiment results show that ViMo can accurately monitor vital signs, with a median error of 0.19 BPM and 0.92 BPM, respectively, for RR and HR estimation.

REFERENCES

[1] D. Biswas, N. Simões-Capela, C. Van Hoof, and N. Van Helleputte, "Heart rate estimation from wrist-worn photoplethysmography: A review," *IEEE Sensors Journal*, 2019.

[2] G. Konstantinos, L. Andreas V, K. Nehal N, A. Ghadah I, A. Yasser A, and G. Elias J, "Can wearable devices accurately measure heart rate variability? a systematic review," *Folia medica*, 2018.

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

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

[5] Y. Cheng, X. Ji, T. Lu, and W. Xu, "Dewicam: Detecting hidden wireless cameras via smartphones," in *Proceedings of the 2018 on Asia Conference on Computer and Communications Security*, 2018.

[6] W. Xu, C. Yan, W. Jia, X. Ji, and J. Liu, "Analyzing and enhancing the security of ultrasonic sensors for autonomous vehicles," *IEEE Internet of Things Journal*, 2018.

[7] X. Ji, and M. Liu J. Wang, Y. Yan, P. Yang, and Y. Liu, "Hitchhike: A preamble-based control plane for snr-sensitive wireless networks," *IEEE Transactions on Wireless Communications*, 2015.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] Y. Zeng, D. Wu, R. Gao, T. Gu, and D. Zhang, "FullBreathe: Full human respiration detection exploiting complementary of CSI phase and amplitude of WiFi signals," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2018.

[14] C. Wu, Z. Yang, Z. Zhou, X. Liu, Y. Liu, and J. Cao, "Non-invasive detection of moving and stationary human with WiFi," *IEEE Journal on Selected Areas in Communications*, 2015.

[15] C. Chen, Y. Han, Y. Chen, and K. J. R. Liu, "Multi-person breathing rate estimation using time-reversal on WiFi platforms," in *2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2016.

[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] D. Zhang, Y. Hu, Y. Chen, and B. Zeng, "BreathTrack: Tracking indoor human breath status via commodity WiFi," *IEEE Internet of Things Journal*, 2019.

[18] J. Liu, Y. Chen, Y. Wang, X. Chen, J. Cheng, and J. Yang, "Monitoring vital signs and postures during sleep using WiFi signals," *IEEE Internet of Things Journal*, 2018.

[19] 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.

[20] IEEE P802. 11ad, "Part 11: Wireless LAN medium access control (MAC) and physical layer (PHY) specifications amendment 3: Enhancement for very high throughput in the 60 GHz band," 2012.

[21] Qualcmm 802.11ad 60GHz WiFi, https://www.qualcomm.com/products/features/80211ad.

[22] Z. Zhang, Z. Pi, and B. Liu, "TROIKA: A general framework for heart rate monitoring using wrist-type photoplethysmographic signals during intensive physical exercise," *IEEE Transactions on Biomedical Engineering*, 2015.

[23] Heba Abdelnasser, Khaled A. Harras, and Moustafa Youssef, "UbiBreathe: A ubiquitous non-invasive WiFi-based breathing estimator," in *ACM International Symposium on Mobile Ad Hoc Networking and Computing*, 2015.

[24] 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.

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

[26] Lingyun Ren, Yun Seo Koo, Yazhou Wang, and Aly E Fathy, "Noncontact heartbeat detection using UWB impulse Doppler radar," in *2015 IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing Systems (BioWireleSS)*, 2015.

[27] 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.

[28] 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.

[29] 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.

[30] 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.

[31] 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.

[32] 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.

[33] Lingyun Ren, Haofei Wang, K. Naishadham, Q. Liu, and A. E. Fathy, "Non-invasive 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.

[34] 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.

[35] T. Ohtsuki and E. Mogi, "Heartbeat detection with Doppler radar based on estimation of average R-R interval using Viterbi algorithm," in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2016.

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

[37] Fadel Adib, Hongzi Mao, Zachary Kabelac, Dina Katabi, and Robert 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.

[38] Marco Mercuri, Ilde Rosa Lorato, Yao-Hong Liu, Fokko Wieringa, Chris Van Hoof, and Tom Torfs, "Vital-sign monitoring and spatial tracking of multiple people using a contactless radar-based sensor," *Nature Electronics*, 2019.

[39] Zhicheng Yang, Parth H. Pathak, Yunze Zeng, Xixi Liran, and Prasant Mohapatra, "Monitoring vital signs using millimeter wave," in *Proceedings of the 17th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, 2016.

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

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.2020.3004046, IEEE Internet of Things Journal

14

[41] M. Comiter, M. Crouse, H. T. Kung, J. Tarng, Z. Tsai, W. Wu, T. Lee, M. C. F. Chang, and Y. Kuan, "Millimeter-wave field experiments with many antenna configurations for indoor multipath environments," in *2017 IEEE Globecom Workshops (GC Wkshps)*, 2017.

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

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

[44] L. Ren, H. Wang, K. Naishadham, O. Kilic, and A. E. Fathy, "Phase-based methods for heart rate detection using UWB impulse Doppler radar," *IEEE Transactions on Microwave Theory and Techniques*, 2016.

[45] S. Bakhtiari, T. W. Elmer, N. M. Cox, N. Gopalsami, A. C. Raptis, S. Liao, I. Mikhelson, and A. V. Sahakian, "Compact millimeter-wave sensor for remote monitoring of vital signs," *IEEE Transactions on Instrumentation and Measurement*, 2012.

[46] *Cardinal Spline Interpolation, Chapter 2, The Basis Property of B-Splines*.

[47] R. L. Rivest T. H. Cormen, C. E. Leiserson, *Introduction to algorithms*, MIT press, 2009.

[48] Polar H10 heart rate monitors, https://www.polar.com/us-en/products/accessories/h10_heart_rate_sensor.

Beibei Wang (SM'15) received the B.S. degree in electrical engineering (with the highest honor) from the University of Science and Technology of China, Hefei, in 2004, and the Ph.D. degree in electrical engineering from the University of Maryland, College Park in 2009. She was with the University of Maryland as a research associate in 2009-2010, and with Qualcomm Research and Development in 2010-2014. Since 2015, she has been with Origin Wireless Inc., where she is currently the Vice President of Research. She is also affiliated with the University of Maryland, College Park. Her research interests include Internet of Things, mobile computing, wireless sensing and positioning, and communications and networking. Dr. Wang received the Graduate School Fellowship, the Future Faculty Fellowship, and the Dean's Doctoral Research Award from the University of Maryland, and the Overview Paper Award from the IEEE Signal Processing Society in 2015. She is a co-author of Cognitive Radio Networking and Security: A Game-Theoretic View (Cambridge University Press, 2010) and Wireless AI: Wireless Sensing, Positioning, IoT, and Communications (Cambridge University Press, 2019).

Fengyu Wang (S' 19) received her B.S. and M.S. degrees from the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China, in 2014 and 2017, respectively. She is now pursuing the Ph.D. degree from the Department of Electrical and Computer Engineering, University of Maryland, College Park. Her current research interests include Internet of Things, wireless sensing and mobile health.

Feng Zhang (S'12-M'19) received his B.S. and M.S. degrees from the Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei, in 2011 and 2014, respectively. He received his Ph.D. degree from the Department of Electrical and Computer Engineering, University of Maryland, College Park in Dec. 2018. He is currently with Origin Wireless, Inc. His research interests include wireless sensing, statistical signal processing, and wireless indoor localization. He was the recipient of Distinguished TA award from the University of Maryland and the State Scholarship from the University of Science and Technology of China.

K. J. Ray Liu (F'03) is a Distinguished University Professor and a Distinguished Scholar-Teach of University of Maryland, College Park, where he is also Christine Kim Eminent Professor of Information Technology. He leads the Maryland Signals and Information Group conducting research encompassing broad areas of information and communications technology with recent focus on wireless AI for indoor tracking and wireless sensing.

Dr. Liu was the recipient of two IEEE Technical Field Awards: the 2021 IEEE Fourier Award for Signal Processing and the 2016 IEEE Leon K. Kirchmayer Graduate Teaching Award, IEEE Signal Processing Society 2009 Technical Achievement Award, IEEE Signal Processing Society 2014 Society Award, and over a dozen of best paper/invention awards. Recognized by Web of Science as a Highly Cited Researcher, Dr. Liu is a fellow of IEEE, AAAS, and U.S. National Academy of Inventors. As the founder of Origin Wireless, his invention won the 2017 CEATEC Grand Prix and CES 2020 Innovation Award.

Dr. Liu was IEEE Vice President, Technical Activities, and a member of IEEE Board of Director as Division IX Director. He has also served as President of IEEE Signal Processing Society, where he was Vice President – Publications and Editor-in-Chief of IEEE Signal Processing Magazine.

He also received teaching and research recognitions from University of Maryland including university-level Invention of the Year Award; and college-level Poole and Kent Senior Faculty Teaching Award, Outstanding Faculty Research Award, and Outstanding Faculty Service Award, all from A. James Clark School of Engineering.

Chenshu Wu received his B.E. degree in School of Software in 2010 and Ph.D. degree in Department of Computer Science in 2015, both from Tsinghua University, Beijing, China. He is now an Assistant Research Scientist in the Department of Electrical and Computer Engineering at the University of Maryland, College Park. He is also a Principal Scientist at Origin Wireless Inc. His research interests include Internet of Things, wireless sensing, ubiquitous computing, and mobile health. He is a member of the IEEE and the ACM.