

Driver Vital Signs Monitoring Using Millimeter Wave Radio

Fengyu Wang^{ID}, Member, IEEE, Xiaolu Zeng^{ID}, Chenshu Wu^{ID}, Senior Member, IEEE,
Beibei Wang^{ID}, Senior Member, IEEE, and K. J. Ray Liu^{ID}, Fellow, IEEE

Abstract—As automobiles have become an essential part to facilitate our daily life, advanced driver assistance systems (ADASs) have been gaining more and more interest in assisting drivers to enhance both safety and convenience. To respond timely in case of an emergency, ADAS needs to keep track of the driver's health/consciousness, which is generally achieved by monitoring the driver's vital signs, including respiration rate (RR), heart rate (HR), and heart rate variability (HRV). However, most of the state-of-art solutions need to assume that the human is stationary, which does not hold in practical driving scenarios. To tackle the problem, we propose a novel system, which can estimate driver's RR, HR, and interbeat intervals (IBIs) in the presence of driver's motion artifacts using commercial millimeter-wave (mmWave) radio. The system consists of two key components. First, to extract the reflection signals containing vital signals, the motion artifacts are first removed by a novel motion compensation module, followed by the periodicity check to identify the components with vital signals. Second, the respiration and heartbeat signals are reconstructed by jointly optimizing the decomposition of all the extracted compound vital signals over different range-azimuth bins. We evaluate the system performance in a real driving environment and investigate the impact of different parameters, including the device locations, pavement conditions, and motion types. The experimental results show that the proposed system can achieve a median error of 0.16 respiration per minute (RPM), 0.82 beat per minute (BPM), and 46 ms for RR, HR, and IBI estimations, corresponding to the relative accuracy of 99.17%, 98.94%, and 94.11%, respectively.

Index Terms—Advanced driver assistance systems (ADASs), driver vital signs monitoring, millimeter-wave radio, motion artifacts elimination, wireless sensing.

Manuscript received May 19, 2021; revised August 7, 2021, September 6, 2021, and October 21, 2021; accepted November 4, 2021. Date of publication November 16, 2021; date of current version June 23, 2022. (Corresponding author: Fengyu Wang.)

Fengyu Wang was with Origin Wireless Inc., Greenbelt, MD 20770 USA. She is now with the School of Artificial Intelligence, Beijing University of Posts and Telecommunications, Beijing 100876, China (e-mail: fywang@umd.edu).

Xiaolu Zeng is with the Department of Electrical and Computer Engineering, University of Maryland at College Park, College Park, MD 20742 USA.

Chenshu Wu is with the Department of Computer Science, The University of Hong Kong, Hong Kong, and also with Origin Wireless Inc., Greenbelt, MD 20770 USA.

Beibei Wang is with Origin Wireless Inc., Greenbelt, MD 20770 USA.

K. J. Ray Liu is 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.

Digital Object Identifier 10.1109/JIOT.2021.3128548

I. INTRODUCTION

Automobiles have become a daily necessity in the current fast-paced world due to its mobility, convenience, and comfortableness. Statistics show that the number of worldwide automobiles on-the-road has reached 1.2 billion by 2015 [1]. However, in the meanwhile, road traffic crashes result in about 1.35 million deaths around the world each year and leave between 20 and 50 million people with nonfatal injuries [2], according to the World Health Organization.

To reduce the number of road accidents and enhance the driving safety, automobile manufacturers, and researchers have been working on more and more advanced driver assistance systems (ADASs). Among many popular topics in autonomous driving, driver's vital sign monitoring is one of the essential components. Continuously monitoring driver's status makes it possible to allow the ADAS to take control of the automobiles in case of emergency, such as when the driver encounters a sudden heart attack, stroke, or fatigue, which can be predicted/indicated by using the driver's heart rate variability (HRV), i.e., the variation of the interbeat intervals (IBIs). HRV, in combination with heart rate (HR) [3] and respiration rate (RR) [4], has been well established as a good indicator of cardiac arrhythmia, alcohol usage [5], mental stress [6] and drowsiness [7], and thus predicts the human alertness well.

Traditional driver vital signs monitoring solutions mainly include two categories: 1) sensor-based methods and 2) vision-based methods. The sensor-based methods require a driver to wear physiological sensors, such as photoplethysmography (PPG) [8], electrocardiogram (ECG) [9], [10], and electroencephalography (EEG) [11], [12] to monitor vital signs. However, it is cumbersome and uncomfortable to wear these dedicated sensors in the daily commute. Moreover, wearing sensors may distract driver's attention, degrading the safety, and user experience. As a less intrusive solution, vision-based methods utilize image sequences to detect the vital signs, including RR [13], HR [14], and HRV [15], [16]. However, the main drawbacks such as its poor performance in low-light scenarios and the privacy concerns hinder the wide deployment of the vision-based systems.

With the development of wireless sensing [17]–[19], radio frequency (RF)-based methods have become one of the most promising candidates. Intuitively, the presence of a human subject will affect the RF propagation [20]–[23], i.e., RF signals reflected off human subjects will be modulated by the body movement, including chest movement due to respiration and

heartbeat. As a result, RF-based systems can estimate vital signs without any physical contact, while preserving the user privacy and operating robustly regardless of the light conditions. Many existing work have validated the feasibility of RR [24]–[26], HR [27]–[30], and HRV [31]–[35] monitoring using RF signal. However, most of these solutions focus on indoor scenarios with stationary human subjects, which cannot deal with the noisy in-car environment with engine vibrations, road vibrations, and human body motion. Therefore, accurate RF-based driver vital sign monitoring needs to be further investigated.

Technically, it is nontrivial to enable RF-based driver vital signs monitoring. First, during driving, the driver exhibits frequent and unpredictable motion (e.g., control the steering wheel, head movement to keep track of the car, body roaming due to acceleration or brake, etc.), which frequently cause dominant motion larger than respiration and heartbeat, and can easily corrupt the periodic variations induced by vital signals. Therefore, it is hard to distill the minute motion caused by vital signals through the raw RF signal.

To overcome the problem, we design a two-step motion compensation algorithm. Note that the reflection profile of the driver stays similar considering the resolution of the system and the size of the target. Given such an observation, in the first step, the location change of the driver is compensated based on the cross-correlation between consecutive channel impulse response (CIR). After that, the reflections corresponding to the same part of the human body will be aligned in the same range-azimuth bin over time. To further remove the fine motion artifacts and recover the periodicity of vital signals revealed in the phase measurement, in the second step, the motion trend is further estimated by smoothing spline and then eliminated.

Second, even if we have eliminated most of the effect of body motion, it is still challenging to extract individual heartbeats from the compound vital signals. This is because that the distance change caused by heartbeat is an order of magnitude smaller than that caused by respiration, and the heartbeat signal is easily to be submerged. Moreover, these subtle cardiogenic body movements lack sharp peak feature as in ECG signals, making it harder to accurately pinpoint the exact timing of heartbeats for HRV estimation.

To tackle these challenges and reconstruct respiration as well as heartbeat signals from the RF reflections, in this article, we propose a joint decomposition method by exploring the following properties of the vital signals.

- 1) Both respiration and heartbeat signals are quasi-periodic signals, where the normal frequency of respiration and heartbeat are 6–30 respiration per minute (RPM) and 50–120 beat per minute (BPM), respectively.
- 2) The reflections from the human chest would occupy different range taps and azimuth angles (known as range-azimuth bins as shown in Fig. 1) considering the range-azimuth resolution of the device and the size of the human body. Hence, the vital information contained in multiple range-azimuth bins can be jointly optimized to improve the estimation accuracy.
- 3) The frequency of vital signals reflected by different parts of the human chest (corresponding to different

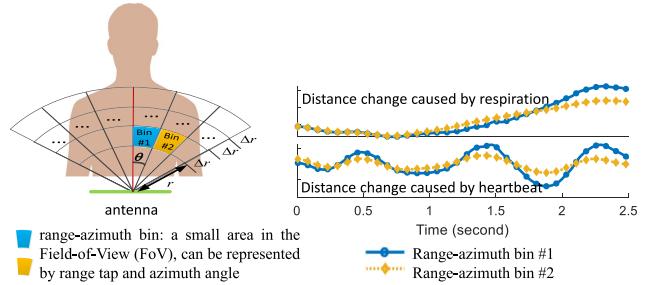


Fig. 1. Vital signals in different range-azimuth bins.

range-azimuth bins) stay the same because the reflections come from the same human subject. However, the distance change caused by respiration and heartbeat can be distinct in different parts of human body due to the physiological structure as shown in Fig. 1. Therefore, for all the range-azimuth bins containing vital signals, we would observe periodic signals with the same frequency but different amplitude in the phase measurement.

Leveraging the aforementioned properties, we jointly optimize the decomposition of the vital signals in different range-azimuth bins as an ensemble of band-limited signals. The respiration and heartbeat signals can be further reconstructed by using the empirical mean of the corresponding component overall range-azimuth bins for RR, HR, and IBI estimations.

We prototype our system using a single commodity off-the-shelf (COTS) millimeter-wave (mmWave) radio and conduct extensive on-road tests to evaluate the performance. We recruit four volunteers (two males and two females) to help on the data collection, and the testing route is a cycle of 50.7 miles, including local routes and highway with different road conditions. The impact of different factors, including the pavement condition, the device location, and user heterogeneity are investigated. The experimental results show that the proposed system can achieve accurate estimations with the median errors of RR, HR, and IBI estimation being 0.16 RPM, 0.82 BPM, and 46 ms, respectively. We also compare the proposed system with the state-of-the-art works, which validates the superiority of our system in accuracy and robustness. To the best of our knowledge, it is the first RF-based driver vital sign monitoring system that can achieve accurate HRV estimation with motion artifacts.

The remainder of this article is organized as follows. We review the related works in Section II. The system and theoretical model are introduced in Section III, followed by vital motion extraction in Section IV and vital signs estimation in Section V. Section VI evaluates the performance of the system. The future work is discussed in Section VII and the whole work is concluded in Section VIII.

II. RELATED WORK

With the proliferation of automobiles, there has been a surging demand in monitoring driver's vital signs to provide driving assistance and safety enhancement. Compared with

the traditional method that requires users to wear contact sensors, the contactless method is less intrusive and can reduce the driver distracting issue. Existing approaches for contactless driver's vital signs monitoring can be classified into two categories: 1) the vision-based method and 2) the RF-based method.

A. Vision Based

Vision-based methods utilize the image sequences to monitor vital signs. In principle, the breathing process causes involuntary quasi-periodic thoracic and abdominal movements, which can be captured by video stream and thus utilized for the RR estimation [13]. It has also been studied that the skin color changes caused by blood perfusion can be used as a good feature to estimate HR [14] and HRV [15], [16]. However, the vulnerability to the lighting condition and the privacy invasions are the main drawbacks hindering the widely deployment of the vision-based system.

B. RF-Based

Compared with the vision-based method, the RF-based method is more robust in handling environment change (e.g., the light and temperature condition). However, note that the displacement caused by body roaming is usually larger than that caused by respiration and heartbeat. As a result, handling the motion artifacts is one of the biggest challenges to achieve RF-based driver's vital sign monitoring. Preliminary works [36], [37] try to get rid of the driver's motion by carefully placing the device (e.g., embedded in the car seat or the seat belt) and simply utilize the basic time-frequency analysis [e.g., the fast Fourier transform (FFT)] to estimate the vital signs. However, the assumption that the dominant frequency component is caused by vital signals may not hold in the real driving scenario when the user encounters large body motion (e.g., body roaming due to acceleration or brake).

To eliminate body movement, multiple transceivers have been deployed at opposite sides of a human body [38], [39]. However, it significantly increases the system complexity and deployment cost, and thus making it hard to implement in practice. The correlation of range taps between different time blocks is used in [40]–[42] to remove the body movement in the system with a single transceiver. However, this method can only remove specific body movement that is larger than the range resolution, and the motion artifacts within the range resolution still remains, thus reducing the estimation accuracy. To remove the motion artifacts located in the same range tap, polynomial fitting is used in [43] to estimate the displacement caused by body motion. However, the order of the polynomial fitting needs to be carefully selected for different motion types, which is not robust in practical use. Note that the above systems can only estimate HR, and the residual signal after motion elimination is too noisy to extract the exact time of heartbeats for further HRV estimation. In contrast, the proposed system can estimate the driver's vital signs including RR, HR, and more importantly HRV regardless of motion artifacts, which is not achievable in existing works.

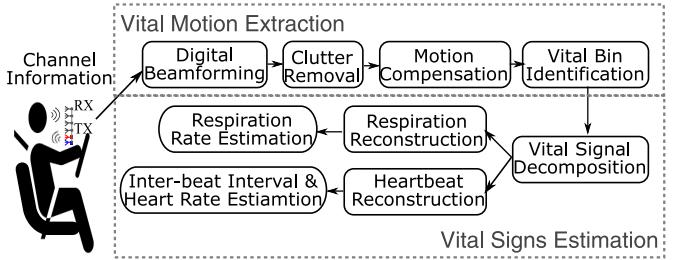


Fig. 2. Processing flow of system.

III. SYSTEM OVERVIEW AND THEORETICAL MODEL

A. System Overview

The proposed system aims at noncontact driver's vital sign monitoring in practical driving scenarios with inevitable random motions by using a single commodity frequency-modulated continuous-wave (FMCW) radar. The pipeline of the system is shown in Fig. 2, which consists of two main modules: 1) vital motion extraction and 2) vital signs estimation.

In the first stage, the *vital motion extraction* module extracts the bins containing vital signals from the channel information. To begin with, conventional beamforming is performed on the channel information to get the CIR at different range-azimuth bins. Then, the clutter removal is performed to subtract the background reflections. However, vital signals cannot be directly extracted even after background subtraction because the driver's location w.r.t. radar can change over time (e.g., body roaming due to acceleration or brake) during driving. As a result, the vital signals will spread over multiple range bins. Therefore, a motion compensation algorithm is devised to eliminate the effect of large body movement. The location change of the driver is first roughly compensated between consecutive CIRs based on correlation of the CIR amplitude. Then, the subtle motion within the range bin are estimated and eliminated from the CIR phase utilizing a smoothing spline. After motion compensation, the range-azimuth bins containing vital signals (*also known as vital bins*) will show periodic pattern, and the CIR of these bins will be exported for further vital signs estimation.

In the second stage, the *vital signs estimation* module estimates drivers' RR, HR, and HRV using the vital signals exported by the previous module. To enable HRV analysis, a heartbeat wave needs to be reconstructed to get the exact time of each heartbeat. However, it is nontrivial to extract the heartbeat signal from the compound vital signals including both respiration and heartbeat movements. To accurately recover the respiration as well as heartbeat signal, we optimize the decomposition of vital signals in all vital bins with multiple band-limited signals concurrently. And the extracted respiration and heartbeat signals in all the vital bins are further combined to give an estimate of the respiration and heartbeat wave for RR, HR, and IBI estimation.

B. Signal Model

Our system is built upon an FMCW radar, which transmits a signal with periodic linearly increasing frequency ramps,

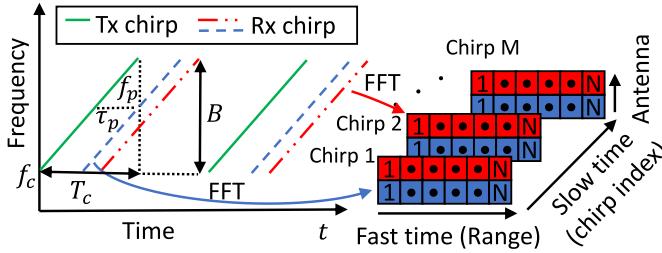


Fig. 3. FMCW radar system.

as shown in Fig. 3. A chirp is a single transmission and the transmitted signal of the m th chirp can be expressed as

$$x_T^m(\tau) = A_T \exp\left\{-j\left[2\pi f_c \tau + \pi \frac{B}{T_c} \tau^2\right]\right\} \quad (1)$$

where f_c is the chirp starting frequency, T_c is the chirp duration, B is the bandwidth, and A_T is the transmitting power. The reflected signal $x_R^m(\tau)$ can be expressed as

$$x_R^m(\tau) = \sum_{p=1}^P A_R \exp\left\{-j\left[2\pi f_c(\tau - \tau_p) + \pi \frac{B}{T_c}(\tau - \tau_p)^2\right]\right\} \quad (2)$$

where A_R is the amplitude of the receiving signal. τ_p stands for the round-trip delay of p th reflecting path and can be denoted as $\tau_p = (2d_p/c)$, where d_p is the distance between the reflecting object and the device, c is the speed of light. P denotes the total number of reflecting points in the environment.

Mixing the received signal with a replica of the transmitted signal and following a low-pass filter, the channel information at time instance m can be expressed as:

$$h^m(\tau) = \sum_{p=1}^P A \exp\left\{-j\left(2\pi \frac{B\tau_p}{T_c}\tau + 2\pi f_c\tau_p - \pi \frac{B}{T_c}\tau_p^2\right)\right\} \quad (3)$$

where A denotes the channel gain. Note that τ_p is in nanosecond for the short-range applications, and the term $\pi(B/T_c)\tau_p^2$ is negligible, therefore, the $h^m(\tau)$ can be written as

$$h^m(\tau) = \sum_{p=1}^P A \exp\left\{-j\left(2\pi \frac{B\tau_p}{T_c}\tau + 2\pi f_c\tau_p\right)\right\} \quad (4)$$

which is a summation of P sinusoidal signals, whose frequency $f_p \triangleq (B\tau_p/T_c) = (2Bd_p/cT_c)$ depends on the target's distance. Besides, by leveraging multiple antennas of the chipset to increase angle resolution, the channel information can be further denoted as

$$h^m(\tau, l) = \sum_{p=1}^P A \exp\left\{-2\pi j\left(f_p\tau + f_c\tau_p + \frac{d_l \sin \theta}{\lambda_c}\right)\right\} \quad (5)$$

where λ_c denotes the wavelength of the chirp, d_l is the relative distance introduced by the l th antenna, and θ is the azimuth angle of the target. This channel information can be converted to CIR by FFT of $h^m(\tau, l)$, a.k.a Range-FFT, which can be denoted as

$$h_{r,l}(m) = \sum_{n=1}^N h^m(n, l) \exp\left\{-j2\pi \frac{rn}{N}\right\} \quad (6)$$

where $h_{r,l}(m)$ denotes the CIR of l th antenna element and r th range tap r at time instance m . n denotes the sample index after digitizing the $h^m(\tau, l)$ over fast time τ , and N is the total number of samples per chirp.

IV. VITAL MOTION EXTRACTION

In a real-world setting, extracting vital motions from the RF signal is not trivial. Due to the presence of various clutters in a car (e.g., chairs, metal objects, ceilings, etc.), it is hard to filter the RF reflections off human body. Moreover, since body motion will be involved during driving, the periodicity of the reflected signal caused by vital motions can be corrupted, complicating the detection of vital signals.

A. Digital Beamforming

To determine the range and the direction of the reflecting objects, the system employs digital beamforming overall antennas for each range tap. In this work, the Bartlett beamformer [44] is used, where the coefficient of the l th antenna toward azimuth angle θ is

$$s_l(\theta) = \exp\left(-2\pi j \frac{d_l \sin \theta}{\lambda_c}\right). \quad (7)$$

The beamformed CIR corresponding to range r and azimuth angle θ can be expressed as

$$h(r, \theta, m) = \mathbf{s}^H(\theta) \mathbf{h}_{r,l}(m) + \epsilon(m) \quad (8)$$

where $\mathbf{s}(\theta) = [s_1(\theta), s_2(\theta), \dots, s_L(\theta)]^\top$ is the steering vector toward angle θ . $\mathbf{h}_{r,l}(m) = [h_{r,1}(m), h_{r,2}(m), \dots, h_{r,L}(m)]^\top$ is the channel information vector at range tap r . $\epsilon(m)$ is the additive white Gaussian noise assumed to be independent and identically distributed (I.I.D) for different range-azimuth bins.

B. Clutter Removal

To locate the range-azimuth bins corresponding to the driver and reduce the impact of reflections from static objects in the vehicle, the system deploys a clutter removal algorithm to subtract the CIR from the background. Note that the reflections from the static object is reasonably assumed to be invariant within a certain period of time, while the reflections from the driver change over time due to human motion (including body motion and motion caused by vital signals). The background profile can be estimated by taking an average of the CIR over slow-time, and the calibrated CIR can be denoted as

$$\hat{h}(r, \theta, m) = h(r, \theta, m) - \frac{1}{M} \sum_{i=1}^M h(r, \theta, m-i) \quad (9)$$

where M is the number of samples used for clutter removal. Fig. 4 shows the effect of the background cancellation, where the raw CIR before clutter removal is shown in Fig. 4(a), and the corresponding calibrated CIR after clutter removal is shown in Fig. 4(b). As can be seen, clutter removal reduces the background noise significantly.

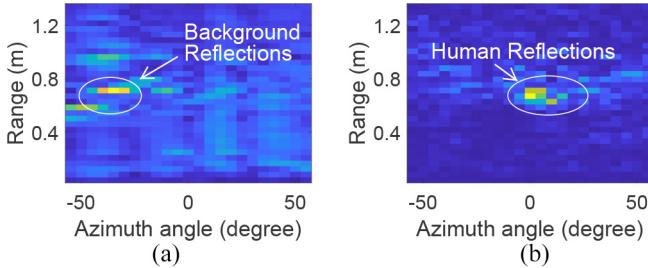


Fig. 4. Example of clutter removal (a) is the CIR amplitude before clutter removal, where the reflections from the driver are submerged in the background reflections and (b) shows the CIR amplitude after clutter removal, where the reflections corresponding to the driver can be easily identified.

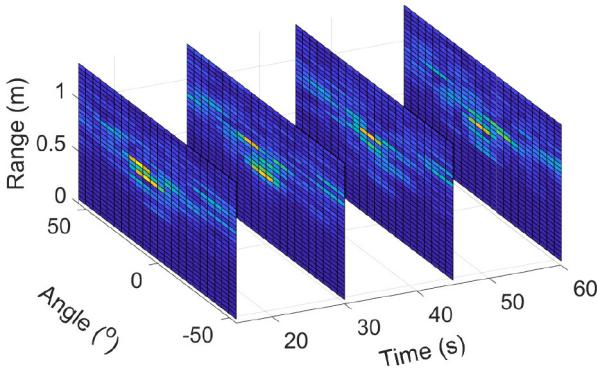


Fig. 5. Example of the consecutive frame after clutter removal. The ground truth is that there is a human subject sit at around 0.5-m away from device at azimuth angle 0°. The human subject sways the body back-and-forth during the experiment. The amplitude of CIR measurement is shown every 15 s in this example.

C. Motion Compensation

After extracting the dynamic CIR corresponding to the driver, we would like to get the range-azimuth bins contributed by the vital signals (*also known as* vital bins). The vital bins can be easily identified by checking the periodicity of the phase signal if the human subject stays stationary as studied in previous works [27]. However, the assumption of the stationary human subject barely holds in the driving scenario. To recover the periodic vital signals from the CIR involving human motion, we design a two-step motion compensation algorithm.

1) *Large Body Movement Compensation:* Note that when there is a large body motion, the location of range-azimuth bins corresponding to the human subject will change, as shown in Fig. 5, where the human subject sit at around 0.5-m away from device at azimuth angle 0°. The human subject sways the body back-and-forth, resulting in the change of reflecting locations. However, the profile of human reflections stays similar, as shown in Fig. 5. Therefore, to remove body movement, the 2-D cross correlation [45] between consecutive CIRs is calculated. Then the CIR at each time instance is circularly shifted to the point corresponding to the maximum cross correlation.

Fig. 6 shows the amplitude of 1-min CIR before and after body movement compensation. For visualization, we plot the CIR at azimuth angle 0° over range [0, 0.9] m. It is shown that after the large body movement compensation, the bins

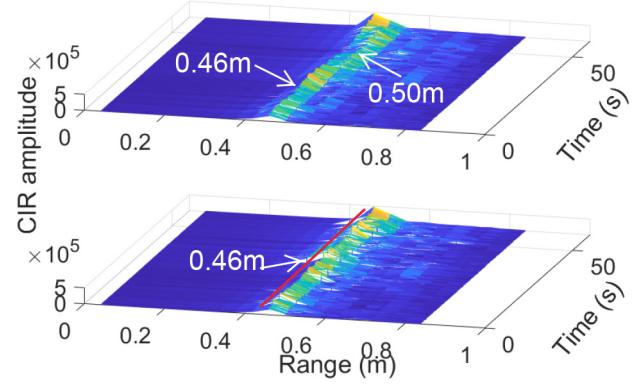


Fig. 6. Example of large body movement compensation. The upper figure shows 1-min CIR amplitude at azimuth angle 0° over range [0, 0.9] m, where the distance between the human subject and device changes over time. The lower figure shows the corresponding CIR amplitude after large body movement compensation, where the range tap of the human subject stays the same.

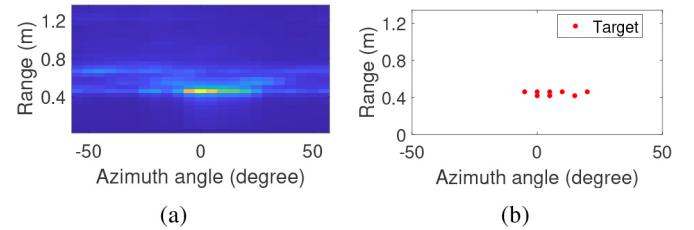


Fig. 7. Example of target detection (a) is the average of the CIR amplitude over 1-min window after large body motion compensation and (b) shows bins corresponding to the driver by using CFAR detector.

correspond to the human subject have been aligned. The 2-D constant false alarm rate (CFAR) detector will be further applied over the CIR after aligning the human subject, and the candidate bins with the human subject can be selected as shown in Fig. 7. Due to the limit of space, the details of the CFAR detector are omitted here, and readers can refer to [46] and the references therein for details.

2) *Fine Movement Elimination:* Although the candidate range-azimuth bins corresponding to the human subject have been aligned and selected in the first step, it is still hard to locate those bins reflected by the chest with periodic vital signals. The reason is that the first step can only remove the motion artifacts that is larger than the range-azimuth resolution, however, it cannot deal with the fine movements within the range-azimuth resolution. Fig. 8(a) shows an example of the unwrapped phase measurement of the candidate range-azimuth bins after large body movement compensation in solid lines, where the slow trend is caused by the fine movements. To recover the periodicity of vital signals, we need to further eliminate the impact of these fine movements.

Let $\mathbf{y}_{r,\theta} = [y_{r,\theta}(1), y_{r,\theta}(2), \dots, y_{r,\theta}(M)]$ to be the unwrapped phase sequence corresponding to the range r and the azimuth angle θ at the observation window, where M is the total number of samples. $[t_1, t_2, \dots, t_M]$ denotes the time corresponding to each observation. Note that the operation of the fine movement elimination is performed within the same

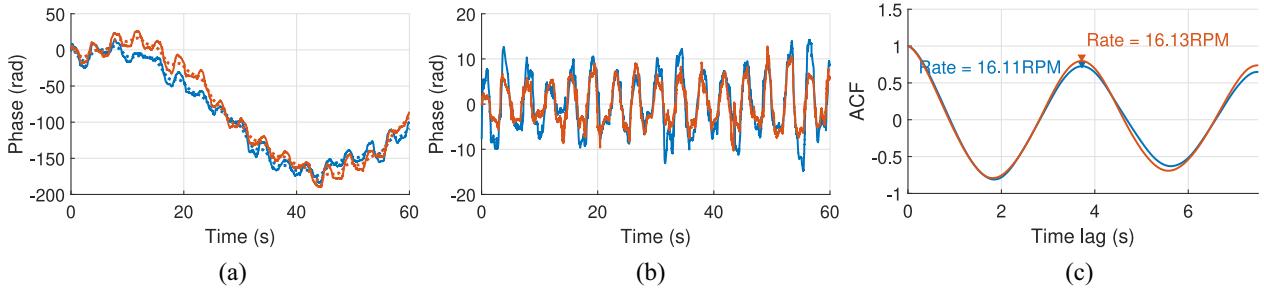


Fig. 8. Example of fine movement cancelation (a) shows the original unwrapped phase measurement from two different vital bins (in solid lines) and the corresponding estimated motion (in dashed lines) (b) is the phase measurement after fine movement cancelation and (c) shows the ACF of the calibrated phase measurement.

range-azimuth bin over slow time, for simplicity, subscripts (r, θ) are omitted in the following analysis.

To remove the motion artifacts that have larger distance change and lower frequency compared to the vital motions, the estimation of the phase change caused by motion artifacts can be obtained by

$$\min_{\hat{f}} \sum_{m=1}^M \left\{ y(m) - \hat{f}(t_m) \right\}^2 + \lambda \int \hat{f}''(t)^2 dt \quad (10)$$

where $\lambda \geq 0$ is a smoothing parameter. The second term evaluates the smoothness of a function. \hat{f} is the estimate of the phase change caused by motion, defined as

$$\hat{f}(t) = \sum_{m=1}^M \hat{f}(t_m) f_m(t) \quad (11)$$

where $f_m(t)$ 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 [47]. Let $\hat{\mathbf{P}} = [\hat{f}(t_1), \dots, \hat{f}(t_M)]^\top$, and the roughness penalty has the form

$$\int \hat{f}''(t)^2 dt = \hat{\mathbf{P}}^\top \mathbf{A} \hat{\mathbf{P}} \quad (12)$$

where the elements of \mathbf{A} are $\int f_i''(t) f_j''(t) dt$. Therefore, we can rewrite (10) as

$$\min_{\hat{\mathbf{P}}} \left\{ \mathbf{y} - \hat{\mathbf{P}} \right\}^\top \left\{ \mathbf{y} - \hat{\mathbf{P}} \right\} + \lambda \hat{\mathbf{P}}^\top \mathbf{A} \hat{\mathbf{P}} \quad (13)$$

where the minimizer of (13) is obtained as

$$\hat{\mathbf{P}}^* = (\mathbf{I} + \lambda \mathbf{A})^{-1} \mathbf{y}. \quad (14)$$

The estimation of motion can be obtained by

$$\hat{f}(t) = \hat{\mathbf{P}}^{*\top} \mathbf{f}(t) \quad (15)$$

where $\mathbf{f}(t)$ is the vector form of the spline basis function. The estimated motion artifacts can then be removed to get the clean phase revealing the vital information. Fig. 8 illustrates the effect of fine movement cancelation, where the dashed lines in Fig. 8(a) show the estimated phase measurement caused by body movement. Fig. 8(b) shows the phase measurement after we remove the motion artifacts, where the periodicity caused by vital signals appears. The above fine movement elimination is performed overall candidate bins selected by the CFAR detector and the cleaned phase of each candidate bin is saved for further analysis.

D. Vital Bin Identification

Note that after motion compensation in Section IV-C, the phase information corresponding to the human chest show periodicity due to the modulation of both respiration and heartbeat, as shown in Fig. 8(b). To filter out the bins reflected by other parts of the human body (i.e., bins dominated by motion), we check the periodicity of the phase signals over slow time by examining their auto-correlation function (ACF) [27]. The reason is that when the phase measurement contains vital signals, a peak can be observed at τ^* in its corresponding ACF, which reveals the time duration of a breathing cycle [27]. Fig. 8(c) shows an example of the ACF of the phase measurement corresponding to the human chest, where the time duration of a breathing cycle is about 3.7 s, correspond to 16.1 RPM. We check the periodicity overall candidate bins corresponding to the human subject, and those bins whose peak located within the range of normal human RR are identified as vital bins for further analysis.

V. VITAL SIGNS ESTIMATION

The vital bins identified by the previous module can only reflect the compound distance change caused by respiration and heartbeat. To further estimate the vital signs, including RR, HR, and HRV, we need to reconstruct the distance change caused by respiration and heartbeat, respectively. For simplicity, in the following analysis, we directly use the analog form of the signal model.

Let $\mathbf{y}(t) = [y_1(t), y_2(t), \dots, y_B(t)]^\top$ denote the vector of the phase signals of all the B vital bins. Recall that the phase signal after movement elimination is a mixture of vital signals, we have

$$\mathbf{y}(t) = \mathbf{s}_r(t) + \mathbf{s}_h(t) + \mathbf{n}(t) \quad (16)$$

where $\mathbf{s}_r(t)$ and $\mathbf{s}_h(t)$ denote the vector of respiration and heartbeat signal, respectively, $\mathbf{n}(t)$ is the random phase offset introduced by noise, which is independent with the phase change caused by vital signs. To decompose the phase and get the estimate of vital signs, we leverage the following properties.

- 1) Both respiration and heartbeat are quasi-periodic signals, whose periodicity changes slightly over time.

- 2) The periodicity of signals corresponding to respiration and heartbeat should stay the same in different vital bins since these signals are modulated by the same person.
- 3) The distance change caused by respiration and heartbeat can be different in different parts of the human body due to the physiological structure (i.e., the distance change in different vital bins can be distinct) [48].

The phase signal, therefore, can be decomposed as an ensemble of band-limited signals, denoted as $\{\mathbf{u}_k(t)\}_{k=1}^K$, where for each component $\mathbf{u}_k(t) = [u_{\{k,1\}}(t), u_{\{k,2\}}(t), \dots, u_{\{k,B\}}(t)]^T$, the decomposed signals w.r.t. all vital bins should be compact around the same center frequency ω_k (corresponding to the properties 1 and 2). Moreover, the distance change in different vital bins should be optimized separately (corresponding to property 3). The decomposition is modeled as [49]

$$\begin{aligned} \min_{u_{k,b} \in \mathcal{U}, \omega_k \in \Omega} \alpha & \sum_{k=1}^K \sum_{b=1}^B \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_{k,b}(t) \right] \right. \\ & \times \exp(-j\omega_k t) \left. \right\|_2^2 \\ & + \sum_{b=1}^B \left\| y_b(t) - \sum_{k=1}^K u_{k,b}(t) \right\|_2^2 \end{aligned} \quad (17)$$

where $\mathcal{U} = \{u_{1,1}, u_{1,2}, \dots, u_{1,B}, \dots, u_{K,B}\}$ and $\Omega = \{\omega_1, \dots, \omega_K\}$ denote the set for all components and their center frequencies, respectively. The first term in (17) represents the bandwidth constraint, which is measured by the sum of the L_2 norm of the gradient of the analytic signal corresponding to each component. The second term is the fidelity constraint, which is evaluated by the quadratic penalty w.r.t. reconstruction. α is a parameter for balancing the bandwidth constraint and data fidelity. The optimization problem in (17) can be solved by alternatively updating \mathcal{U} and Ω until convergence.

A. Minimization w.r.t. $u_{k,b}$

To update the k th component for vital bin b , the subproblem can be written as

$$\begin{aligned} u_{k,b}(t) = \arg \min_{u_{k,b}(t)} \alpha & \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_{k,b}(t) \right] \exp(-j\omega_k t) \right\|_2^2 \\ & + \left\| y_b(t) - \sum_{i=1}^K u_{i,b}(t) \right\|_2^2. \end{aligned} \quad (18)$$

By using the Parseval theorem, the problem is equivalent to

$$\begin{aligned} \square_{k,b}(\omega) = \arg \min_{\square_{k,b}(\omega)} \alpha & \| j\omega [(1 + \text{sgn}(\omega + \omega_k)) \square_{k,b}(\omega + \omega_k)] \|_2^2 \\ & + \left\| \dagger_b(\omega) - \sum_{i=1}^K \square_{i,b}(\omega) \right\|_2^2 \end{aligned} \quad (19)$$

where $\square_{k,b}(\omega)$ and $\dagger_b(\omega)$ are the Fourier transfer of $u_{k,b}(t)$ and $y_b(t)$, respectively. After performing a change of variables $\omega \leftarrow \omega - \omega_k$ in the first term, and using the Hermitian symmetry of the real signals in the spectrum for the second term,

the above problem can be rewritten as

$$\begin{aligned} \square_{k,b}(\omega) = \arg \min_{\square_{k,b}(\omega)} & \int_0^\infty 4\alpha(\omega - \omega_k)^2 |\square_{k,b}(\omega)|^2 \\ & + 2 \left| \dagger_b(\omega) - \sum_{i=1}^K \square_{i,b}(\omega) \right|^2 d\omega. \end{aligned} \quad (20)$$

The updated solution can be expressed as

$$\square_{k,b}(\omega) = \frac{\dagger_b(\omega) - \sum_{i,i \neq k} \square_{i,b}(\omega)}{1 + 2\alpha(\omega - \omega_k)^2}. \quad (21)$$

B. Minimization w.r.t. ω_k

The center frequencies ω_k only appear in the bandwidth constraint and thus the updating function can be written as

$$\omega_k = \arg \min_{\omega_k} \sum_{b=1}^B \left\| \partial_t \left[\left(\delta(t) + \frac{j}{\pi t} \right) * u_{k,b}(t) \right] \exp(-j\omega_k t) \right\|_2^2. \quad (22)$$

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

$$\omega_k = \arg \min_{\omega_k} \sum_{b=1}^B \int_0^\infty (\omega - \omega_k)^2 |\square_{k,b}(\omega)|^2 d\omega. \quad (23)$$

The minimizer of the above problem is

$$\omega_k = \frac{\sum_b \int_0^\infty \omega |\square_{k,b}(\omega)|^2 d\omega}{\sum_b \int_0^\infty |\square_{k,b}(\omega)|^2 d\omega}. \quad (24)$$

Fig. 9 shows an example of vital signals decomposition, where the time and frequency domain of the original phase as well as the decomposition components are shown in Fig. 9(a) and (b), respectively. The information of three different vital bins are distinguished by the color of lines, and it is clear to see that although the distance change of different vital bins is distinct, as shown in Fig. 9(a), the periodicity of the signal of each component stays the same, as shown in Fig. 9(b). In other words, components corresponding to vital signals are perfectly aligned overall vital bins, e.g., the first component represents the distinct displacement caused by respiration over different vital bins, and the second component represents the distinct displacement caused by heartbeat over different vital bins. The residue of the decomposition contains noise including car vibrations, as shown in Fig. 9.

C. Vital Signals Reconstruction

To further reduce the noise impact, we reconstruct the vital signals by combining the signals of all vital bins using empirical mean, i.e., $s_r(t) = (1/B) \sum_b u_{i,b}(t)$ and $s_h(t) = (1/B) \sum_b u_{j,b}(t)$, where the i th and j th components correspond to the respiration signal and heartbeat signal, respectively. The RR is estimated by finding the first peak of the ACF of the estimated respiration signal, as shown in Fig. 8(c). Besides, the FFT is further performed on the estimated heartbeat signal to get the estimation of HR. Moreover, the exact time of each heartbeat can be further extracted from the reconstructed heartbeat wave to estimate the IBI.

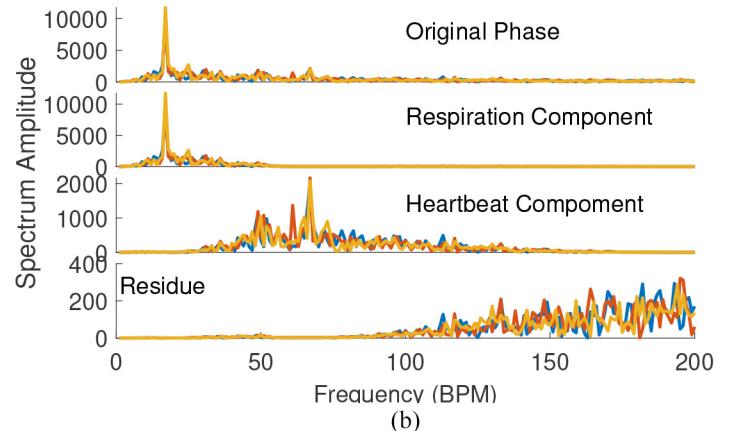
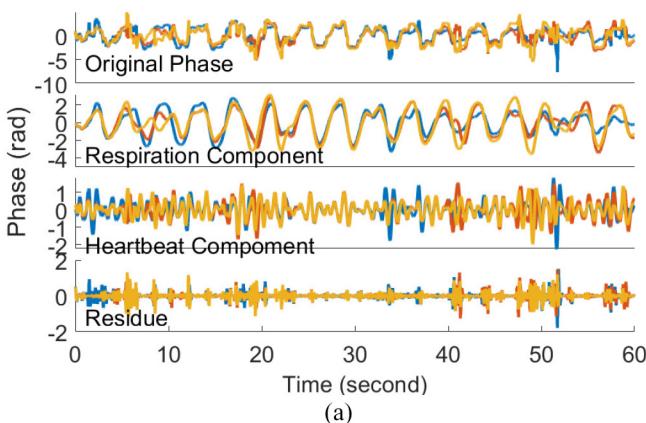


Fig. 9. Example of phase decomposition of three vital bins, where the information in different vital bins are distinguished by the color of lines (a) is the decomposition result in the time domain, where the first subfigure shows the phase measurement after motion cancelation. The respiration and heartbeat component are shown in the second and the third subfigures, respectively. The fourth subfigure shows the decomposition residue. (b) Shows the corresponding spectrum of each component.

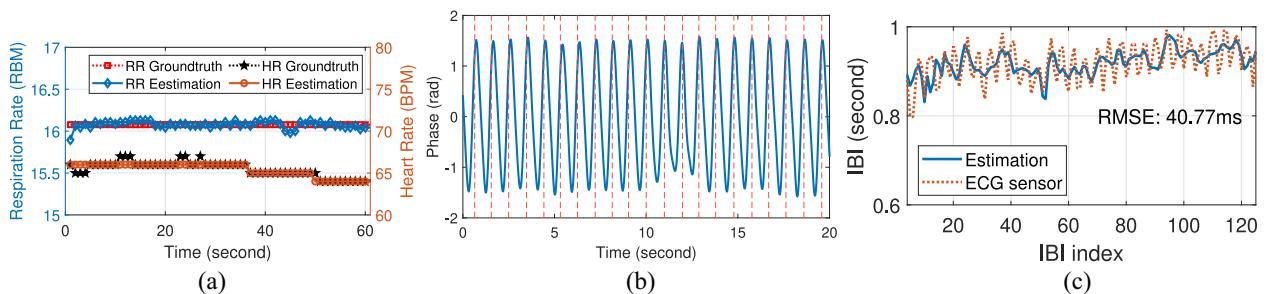


Fig. 10. Example of estimated result versus ground truth. (a) Shows the RR and HR estimation result. (b) Shows the estimated heartbeat signal compared with the ECG sensor result, where the ground truth from ECG sensor is marked as vertical dashed lines. (c) Show the estimated IBI compared with ground truth from the ECG sensor.

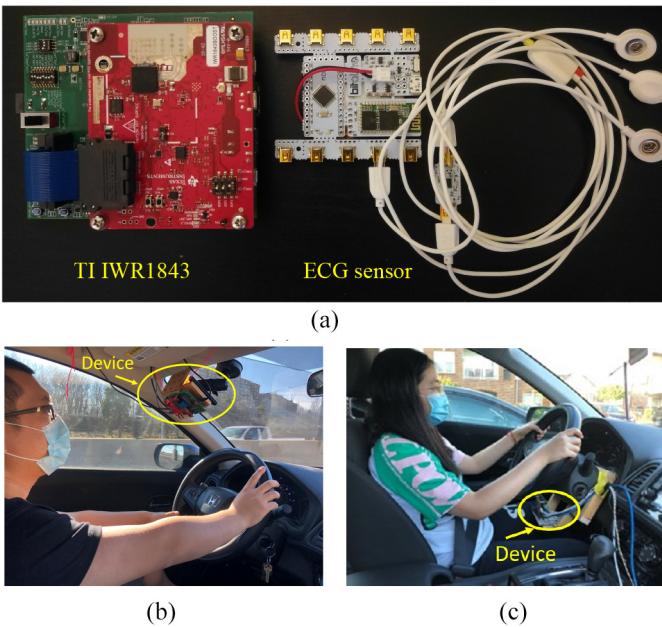


Fig. 11. Experiment setup. (a) Hardware. (b) Device mount on windshield. (c) Device under steering wheel.

Fig. 10 shows the estimated vital signs versus their ground truths of a 2-min data set, where a 1-min window is employed

TABLE I
PARAMETERS USED

System Parameters	Value	System Parameters	Value
Starting Frequency	77 GHz	Max. Range	1.2 m
Slow Time Sampling	1000 chirps/sec	Range Resolution	3.75 cm
Chirp Duration	57.14 us	Field-of-View	$[-50^\circ, 50^\circ]$

for the time-frequency domain transform (i.e., ACF and FFT). The estimated RR and HR are shown in solid lines in Fig. 10(a), which match with the ground truth, shown as dashed lines in Fig. 10(a). Fig. 10(b) shows a segment of the estimated heartbeat wave, and the ground truth of the exact time of each heartbeat is marked as vertical dashed lines. The estimated IBIs of the whole data and their corresponding ground truth are shown in Fig. 10(c). Clearly, the proposed system achieves high accuracy in vital signs estimation, and the root-mean-squared-error (RMSE) of IBI estimation in Fig. 10(c) is 40.77 ms, corresponding to the 96% relative accuracy.

VI. EXPERIMENT EVALUATION

In this section, extensive experiments are performed to evaluate the performance of the proposed system. We compare the performance with the state-of-art work under different experimental settings.

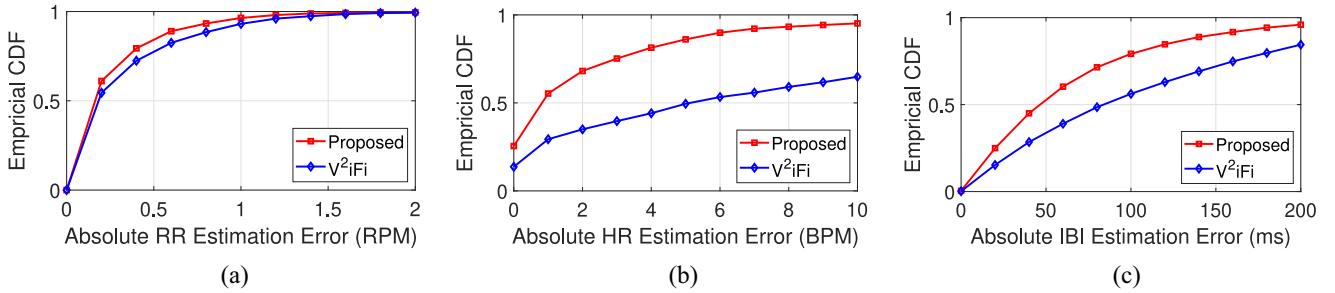


Fig. 12. Comparison of vital sign estimation performance between proposed method and V^2iFi . (a) CDF of RR estimation error. (b) CDF of HR estimation error. (c) CDF of IBI estimation error.

TABLE II
INFORMATION OF THE PARTICIPANTS

Subject ID	1	2	3	4
Gender	M	M	F	F
Height (cm)	174	172	160	166
Weight (Kg)	79.8	70	61.4	50
BMI	26.36	23.66	23.98	18.14

A. Methodology

We conduct experiments using a COTS mmWave radar, IWR1843BOOST [50], as shown in Fig. 11(a), where the two Tx antennas and four Rx antennas are configured in TDM-MIMO mode [51]. The device can achieve a theoretical azimuth resolution of 15° , and the Field-of-View (FoV) is 100° in the horizontal plane, which is large enough to cover the driver. The parameters corresponding to the FMCW radar setting are listed in Table I. The ground truth of heartbeat is captured by a commercial ECG sensor [52], as shown in Fig. 11(a), and the ground truth of breathing is measured by a respiration belt [53].

We recruit four volunteers (two males and two females) to help on the data collection including two different device locations as shown in Fig. 11(b) and (c). All of the participants do not have any cardiac history, and more information about the testers are shown in Table II. The driving route is a cycle of 50.7 miles, including local routes and highway, where the road conditions can be referred to Maryland's GIS Data Set [54]. During the data collection, the driver is driving following their own habits with no further constraints, and a copilot is responsible for collecting data.¹

To further evaluate the performance of the proposed system, we compare it with the state-of-art work, V^2iFi [56], which estimates driver's vital signs using the CIR of a UWB radar. With the assumption that the distance change caused by vital signals are identical in different vital bins, V^2iFi estimates the respiration and heartbeat signal by multisequence variational mode decomposition (MS-VMD). Note that V^2iFi cannot estimate vital signs when drivers have body motion. For a fair comparison, motion compensation proposed in Section IV-C is also applied to V^2iFi to remove the motion artifacts before estimating the vital signals.

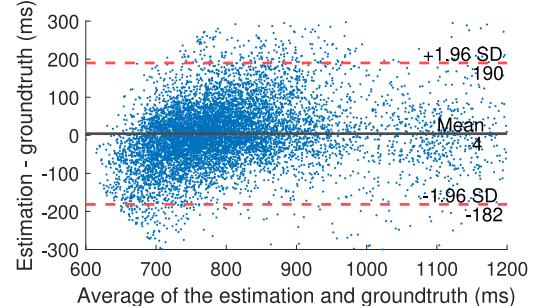


Fig. 13. Bland-Altman plot for the proposed method.

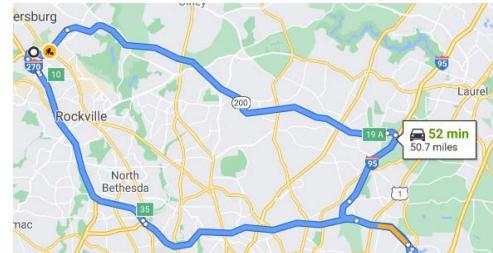


Fig. 14. Experiment path.

B. Overall Performance

Fig. 12 depicts the overall performance of the proposed system and V^2iFi . The experiments consist of road tests with different pavement conditions, device locations, as well as the controlled experiments with different motion types, including stationary, head motion, hand motion, and back-and-forth torso motion for four different users. Fig. 12(a) plots the empirical cumulative distribution function (CDF) of absolute RR estimation error, where the 90-percentile error for the proposed system and V^2iFi are 0.64 and 0.86 RPM, respectively. The performance improvement is more significant for HR estimation, where the proposed system achieves a median error of 0.82 BPM, and the median error of V^2iFi is 5.12 BPM, as shown in Fig. 12(b). Fig. 12(c) shows the performance of the IBI estimation for the two systems, where V^2iFi yields about 84 ms medium error, while the proposed system achieves a medium error of 46 ms, outperforming V^2iFi by about 45.2%. The Bland-Altman plot is shown in Fig. 13, where the solid line shows the mean of the difference between the estimation and the ground truth, and the dashed lines show the ± 1.96 times of standard deviation of the difference. It is obvious that

¹The data were obtained from the exclusive data from Origin Wireless [55].

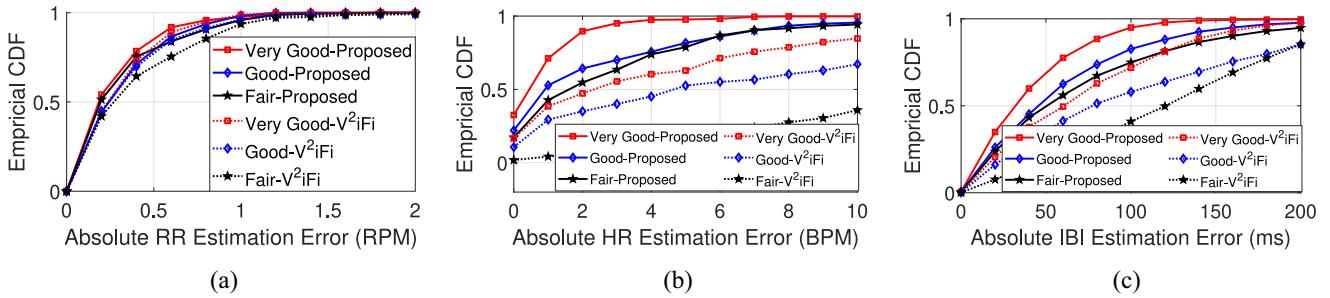


Fig. 15. Vital sign estimation performance versus pavement condition. (a) CDF of RR estimation error. (b) CDF of HR estimation error. (c) CDF of IBI estimation error.

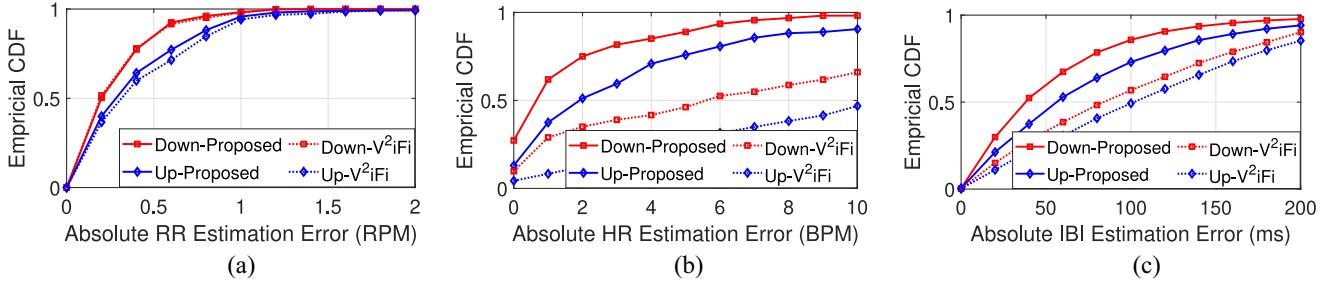


Fig. 16. Vital sign estimation performance versus device location. (a) CDF of RR estimation error. (b) CDF of HR estimation error. (c) CDF of IBI estimation error.

our estimation is nearly unbiased compared with the ground truth.

C. Impact of Road Condition

In this section, we investigate the effect of road condition on the estimation accuracy. Note that the road/pavement condition is assessed by several factors including rutting, friction, structural cracking density, etc. The better the pavement condition is, the smoother the road is, and less body motion induced by uneven road will be involved.² The test route shown in Fig. 14 roughly includes three different pavement conditions, i.e., *Very Good*, *Good*, and *Fair*, and the length of the road corresponding to *Very Good*, *Good*, and *Fair* condition is about 17, 16, and 18 mile, respectively, according to Maryland's GIS Data Set [54]. During the data collection, the copilot saves data every 2 min and records the corresponding pavement condition at the same time. On average, every round of road test contains about 18 sets of data.

As expected, the performance degrades with the deterioration of the pavement condition, as shown in Fig. 15. The median error of the proposed method in terms of RR, HR, and IBI estimations are 0.18 RPM, 0.45 BPM, and 32 ms, respectively, when the pavement condition is *Very Good*. The median error increase to 0.19 RPM, 1.64 BPM, and 50 ms for the RR, HR, and IBI estimations when the pavement condition is *Fair*. Since the distance change caused by respiration is larger than heartbeat, which means a higher signal-to-noise-ratio (SNR) of the RR estimation, we observe a slight degradation in RR estimation compared with HR and IBI estimations.

Besides, the proposed method outperforms V²iFi in all three vital signs measurement, and the performance gap increases

²The case study of the impact of the uneven road can be referred to the Appendix.

with the deterioration of the pavement condition. In specific, the median error of V²iFi in terms of RR, HR, and IBI estimations are 0.23 RPM, 2.34 BPM, and 61 ms, respectively, when the pavement condition is *Very Good*, which correspond to a relative increment of 17.8%, 420.0%, and 90.6% compared with the proposed method. The performance gap in terms of median error w.r.t. RR, HR, and IBI estimation increases to 42.1%, 620.7%, and 142.0% when the pavement condition is *Fair*. The main reason for the performance gap between V²iFi and the proposed method is that V²iFi assumes the same distance change of vital signals in different vital bins, which is hard to meet when the SNR of the signal is small. However, the proposed system only assumes the same periodicity of vital signals in different vital bins when reconstructing vital signals, and the distance change in different vital bins are jointly optimized, which is more robust to the noise.

D. Impact of device location

In this study, we investigate the impact of device location on the vital signs estimation. The radar is placed at the top of windshield (denoted as “up”), as shown in Fig. 11(b), and under the steering wheel (denoted as “down”), as shown in Fig. 11(c). Fig. 16 plots the CDF of the absolute error of RR, HR, and IBI estimations, where the red lines correspond to the “down” setting, and the blue lines correspond to the “up” setting.

It is shown that the “down” setting achieves better performance for all estimations. In specific, for the proposed system, the median error for RR, HR, and IBI estimation are 0.2 RPM, 0.65 BPM, and 38 ms, respectively, for the “down” setting. However, it increases to 0.28 RPM, 1.91 BPM, and 56 ms for the “up” setting, corresponding to 40%, 193.85%, and 47.37% performance degradation, respectively.

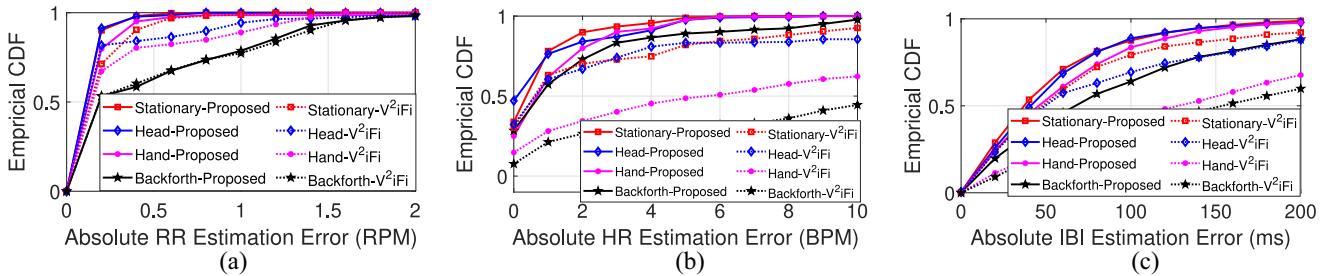


Fig. 17. Vital sign estimation performance versus motion type. (a) CDF of RR estimation error. (b) CDF of HR estimation error. (c) CDF of IBI estimation error.

We observe the similar phenomenon in V^2iFi , where the median error for all the three metrics increases when the device is placed as the “up” setting, as shown in dashed lines in Fig. 16. The reason is that when the device is mounted on the windshield, the vital bins mainly correspond to the chest, whereas, for the “down” setting, the vital bins mainly correspond to the lower chest and the abdomen. Note that for the same scenario (e.g., car decelerates due to brake), a severer motion will be involved in the upper chest than the abdomen, therefore, the SNR of vital signals for the “down” setting is larger than the “up” setting. However, comparing to the proposed system, V^2iFi yields a larger estimation error for all the three metrics, because it is less robust to noise as discussed in Section VI-C.

E. Impact of Motion Type

As driving involves different kinds of motion of head, hand, and body when looking at the side mirror, or controlling the steering wheel, etc., to better understand the impact of different motion types, we conduct controlled experiments and analyze their corresponding impact in this section, as shown in Fig. 17. During the experiment, drivers are asked to continuously perform a specific type of motion in a parked car, including sitting stationary, head motion to check the surroundings, hand motion to operate steering wheel, and randomly sway their body back-and-forth to emulate the body motion caused by acceleration and deceleration. Every data collection lasts for 2 min for both “up” setting and “down” setting as shown in Fig. 11. In total, we have 32 sets of data for analysis.

Fig. 17(a) shows the CDF of RR estimation error with different motion types, where we can see that the median estimation error when driver performs head motion is nearly the same as the stationary case. The performance slightly degrades when the driver performs hand motion, where the median error increases from 0.11 to 0.12 RPM compare to the stationary setting. However, for the large back-and-forth motion, we observe a severe performance degradation, and its median error of RR estimation is 0.19 RPM, 72.73% worse than the stationary setting. Similar performance degradation can be observed in terms of the HR and IBI estimation.

Fig. 17(b) shows that the median error of HR estimation increases from 0.35 BPM corresponding to the stationary setting to 0.68 and 0.75 BPM when the driver performs hand and back-and-forth motion, respectively. As for IBI estimation, the median error when the driver performs sitting stationary, head

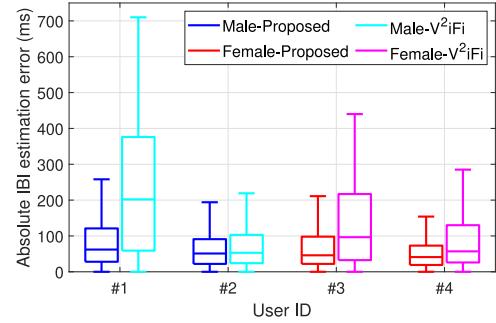


Fig. 18. Impact of user heterogeneity.

motion, hand motion, and random back-and-forth motion are 37, 41, 45, and 68 ms, respectively, as shown in Fig. 17(c).

We also plot the estimation performance of V^2iFi in dashed lines in Fig. 17, where the similar performance degradation can be observed. However, we can see that V^2iFi is more vulnerable to motion artifacts, and the performance degradation of hand and back-and-forth motion is more severe compared to the proposed system. In specific, we can see that the median error of HR estimation for the back-and-forth setting is larger than 10 BPM, which is almost useless for the driver’s HR estimation.

F. Impact of User Heterogeneity

In this part, we study the impact of the user heterogeneity on the performance. Fig. 18 summarizes the absolute IBI estimation error of four drivers using the data of all the settings above. Clearly, the proposed method demonstrates different IBI estimation errors for different users, where the median error varies from 41 to 62 ms, as shown in blue and red boxes in Fig. 18. The difference in error distribution can be caused by various factors, such as different driving habits and heart-beat strength over individuals. Besides, the performance of V^2iFi is shown in cyan and magenta boxes for comparison. It is obvious that the proposed method outperforms V^2iFi for all the four users, which is benefited from its dedicated design to resist motion artifacts.

G. Impact of Window Length on HRV Calculation

In this part, we investigate the impact of the window length on HRV calculation. Known that the HRV metrics can be

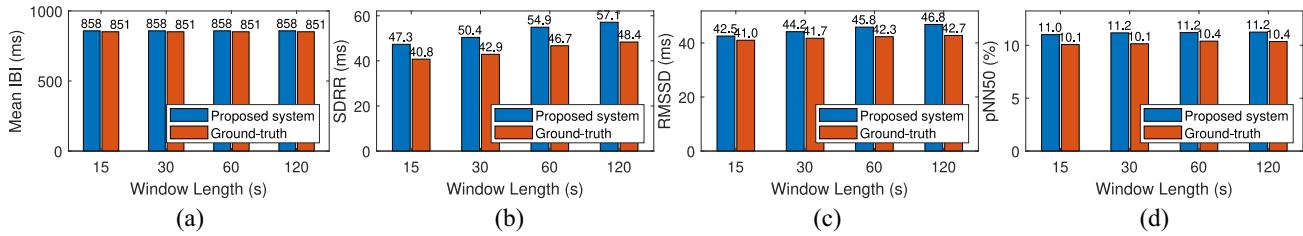


Fig. 19. HRV metrics with different time window. (a) Mean IBI. (b) SDRR. (c) RMSSD. (d) pNN50.

derived from the IBI sequence [57], Fig. 19 shows four different commonly used HRV metrics with window length ranging from 15 to 120 s. The mean of IBI and the standard deviation of the IBIs (SDRR) under different time window are shown in Fig. 19(a) and (b), respectively. The root-mean-square-of-successive differences (RMSSDs) is shown in Fig. 19(c), which can be calculated by

$$\text{RMSSD} = \sqrt{\frac{1}{N_{\text{IBI}} - 1} \sum_{i=2}^{N_{\text{IBI}}} (\text{IBI}(i) - \text{IBI}(i-1))^2} \quad (25)$$

where N_{IBI} is the total number of IBIs in the given time window. Fig. 19(d) shows the percentage of successive IBI that differ by more than 50 ms (pNN50), which can be calculated by

$$\text{pNN50} = \frac{\sum_{i=2}^{N_{\text{IBI}}} \mathbb{1}\{(\text{IBI}(i) - \text{IBI}(i-1)) > 50 \text{ ms}\}}{N_{\text{IBI}}} \quad (26)$$

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

As shown in Fig. 19(a), the mean of IBI barely changes over the window length. However, the other three metrics (i.e., SDRR, RMSSD, and pNN50) increase with the window length for both estimation and ground truth, as shown in Fig. 19(b)–(d). Furthermore, the estimation error of SDRR increases from 6.5 to 8.7 ms when the window length increases from 15 to 120 s. Similar performance/trend can be observed in RMSSD, where the estimation error increases from 1.5 to 4.1 ms when the window length increases from 15 to 120 s.

VII. DISCUSSION AND FUTURE WORK

With the miniaturization of antennas and chips, mmWave has been widely used in automotive radar system for ambient sensing, and our work shows that it can also be applied for in-vehicle sensing to detect driver's vital signals. In addition, existing work [58] has shown the feasibility of driver's drowsiness detection using vital signs, such as RR and HR estimation with the estimation errors of 5% and 2.82%, respectively. While in our system, the estimation error of RR and HR estimation are 0.83% and 1.06%, respectively, which is more accurate than [58]. Based on the estimated RR and HR, it is possible to detect the abnormal conditions of the driver, such as the tendency of road rage and driver's drowsiness, because it has been shown that the RMSSD increases by 10 ms when the human subject changes from neutral to angry [59], and the RR and HR decrease from wakefulness and extreme drowsiness by 3.5 [60] and 8 BPM [3], respectively. However, different applications may expect different acceptable error ranges, which is

also related to the post-processing of a particular application. It is an interesting topic to explore such tolerance levels for our future research.

VIII. CONCLUSION

In this article, we propose a novel system that can accurately detect driver's vital signs in the presence of practical driving motions using the reflections of RF signals off the human subject only. To locate the reflections from the driver, the system first performs conventional beamforming to get the CIR with different range-azimuth bins, followed by a clutter removal module to remove the reflection from the background. Then, the 2-D correlation between different CIR samples have been used to eliminate large displacement caused by body roaming. Finer motion artifacts are further removed by the smoothing spline, which can accurately estimate motion artifacts without dedicated choose of hyper-parameter as in polynomial fitting. The displacement caused by respiration and heartbeat are then estimated by jointly optimizing the decomposition of vital signals in all vital bins, and the RR, HR and IBI can be extracted from the reconstructed respiration and heartbeat wave. We prototype our system using a commercial mmWave radio, and conduct experiments in the real driving scenario to evaluate the performance. The experimental results show that the proposed system can estimate vital signs accurately with driving motion artifacts, outperforming the state-of-art works.

APPENDIX

In this part, we use a specific case to study the impact of the car vibration. During driving, the car vibrates due to the running engine, the friction between tires and roads, as well as bumps/pits in the road. To measure the vibration of the car, we fix an accelerometer at driver's seat during the driving test. Figs. 20–25 show the accelerometer readings and the procedure of vital sign extraction. During the data collection, the car is running on MD495, where the corresponding pavement condition is “Fair.”

In specific, Fig. 20(a) shows the accelerometer readings in the time domain, where Acc-X, Acc-Y, and Acc-Z denote the acceleration in range, azimuth, and elevation domain. The corresponding spectrum and the CDF of the spectrum energy are shown in Fig. 20(b). It is clear to see that the major frequency of the car vibration is much larger than the normal range of vital signs. However, it does not mean that the car vibration will not influence the vital signs estimation. From observation, the car vibration can be roughly classified into two types, i.e.,

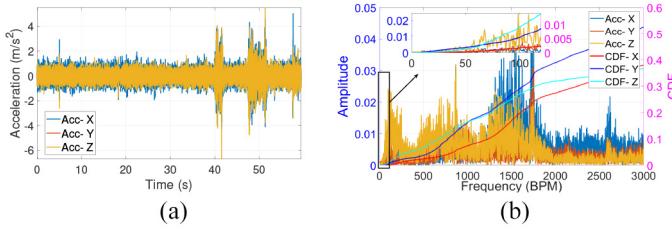


Fig. 20. Accelerometer signal at the driver's seat. (a) Time domain. (b) Frequency domain.

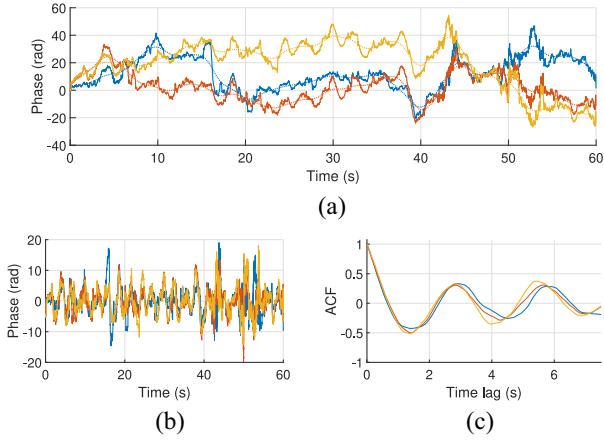


Fig. 21. Example of the motion compensation w.r.t. bins failing the periodic check. (a) Original phase. (b) Phase after motion compensation. (c) ACF of the phase after motion compensation.

the abrupt large acceleration caused by bumps/pits in the road, corresponding to the case when $t \in [40, 50]$ in Fig. 20(a), and the normal car vibration caused by running engine and friction between tires and roads, corresponding to the case when $t \in [0, 40]$ in Fig. 20(a).

Here, we need to clarify that the car vibration and human motion are two different notions. The reason that the car vibration may impact the estimation result is that the vibration may introduce body motion. When there is a pit/bump in the road (e.g., the accelerometer readings at around $t \in [40, 50]$), it may cause large body motion, which can corrupt the vital signals at some range-azimuth bins. However, note that different parts of human body have a different response to the abrupt large vibration, as shown in different lines in Fig. 22(a). Besides, the distance change caused by vital signals varies with different parts of the human body, as discussed in Section V. As a result, the SINR of the vital signals is different in distinct range-azimuth bins.

Fig. 21 shows the original phase and the signal after motion compensation in three different range-azimuth bins that fail the periodicity check. Meanwhile, Fig. 22 shows the original phase and the signal after motion compensation in three different range-azimuth bins that pass the periodicity check.³ We can see that the large abrupt acceleration change has caused large

³In our system, we use the property of the ACF peak, including peak height, peak width as well as peak prominence to check the periodicity, as discussed in Section IV-D. The peak that passes the periodicity check is marked as a triangle as shown in Fig. 22(c).

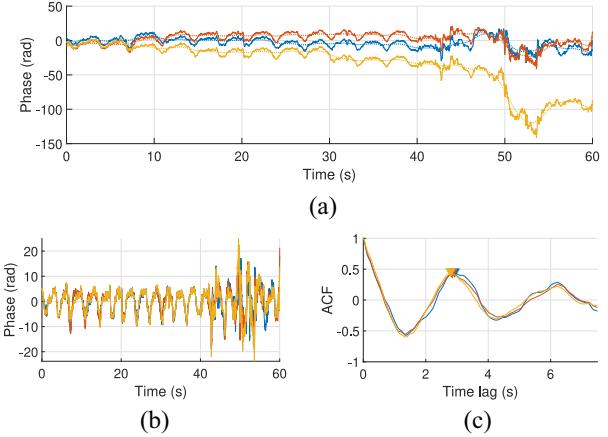


Fig. 22. Example of the motion compensation w.r.t. bins passing the periodic check. (a) Original phase. (b) Phase after motion compensation. (c) ACF of the phase after motion compensation.

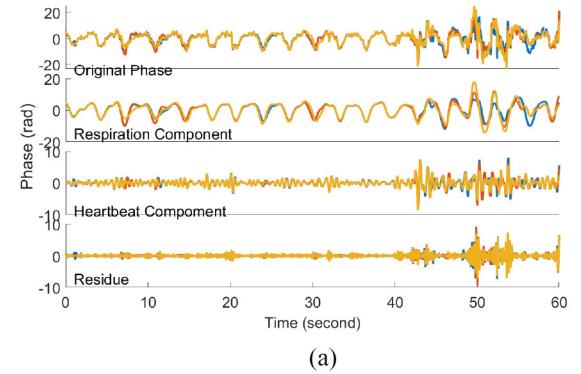


Fig. 23. Example of phase measurement decomposition. (a) Time domain. (b) Frequency domain.

body motion in all the six bins, as shown in Figs. 21(a) and 22(a). However, for the bins in Fig. 22, the condition of the phase change caused by vital signals is still good enough for vital signs extraction. The corresponding phase decomposition result is shown in Fig. 23. Fig. 24 shows the reconstructed respiration and heartbeat signals, where all the bins that pass the periodicity check (including but not limited to the three bins shown in Fig. 22) have been used, as discussed in Section V-C. Fig. 25 gives the estimated vital signs as well as their ground truths for the whole 2-min test, where we can see that even though there are large acceleration changes caused by the uneven road during the data collection, the estimation results match the ground truth.

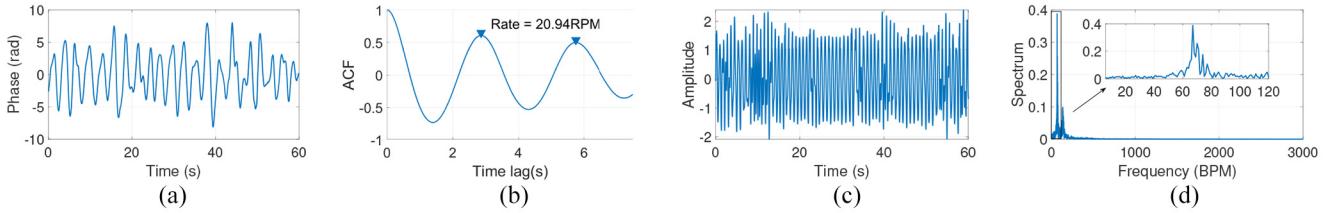


Fig. 24. Illustration of retrieved vital signals. (a) Estimated respiration signal. (b) ACF of the respiration signal. (c) Estimated heartbeat signal. (d) Spectrum of heartbeat signal.

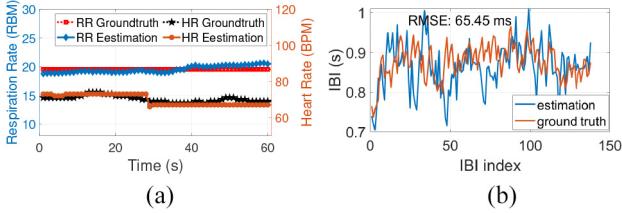


Fig. 25. Example of estimated result versus ground truth. (a) RR and HR estimation. (b) IBI estimation.

Besides the large abrupt acceleration change caused by the uneven road, the running engine as well as the friction between tires and roads causes continuous minute car vibrations, as shown in Fig. 20(a) when $t \in [0, 40]$. However, this kind of car vibration has a smaller impact on the vital signals, as shown in Fig. 22(a) when $t \in [0, 40]$. For the given testing route, we do not observe too much difference on this kind of vibration on three different pavement conditions. However, the large abrupt vibration happens more frequently on the road with “Fair” condition than the other two types. This matches the analysis of the impact of pavement condition on the system performance as discussed in Section VI-C.

ACKNOWLEDGMENT

K. J. Ray Liu was the founder of Origin Wireless which commercializes technology developed in part with collaborations with his labs.

REFERENCES

- [1] “World vehicles in use.” [Online]. Available: https://www.oica.net/wp-content/uploads//Total_in-use-All-Vehicles.pdf (Accessed: Mar. 1 2021).
- [2] “Road traffic injuries.” [Online]. Available: https://www.who.int/health-topics/road-safety#tab=tab_1 (Accessed: Mar. 1 2021).
- [3] S.-H. Jo, J.-M. Kim, and D. K. Kim, “Heart rate change while drowsy driving,” *J. Korean Med. Sci.*, vol. 34, no. 8, p. e56, Mar. 2019.
- [4] F. Zhang *et al.*, “SMARS: Sleep monitoring via ambient radio signals,” *IEEE Trans. Mobile Comput.*, vol. 20, no. 1, pp. 217–231, Jan. 2021.
- [5] U. R. Acharya, K. P. Joseph, N. Kannathal, C. M. Lim, and J. S. Suri, “Heart rate variability: A review,” *Med. Biol. Eng. Comput.*, vol. 44, no. 12, pp. 1031–1051, Nov. 2006.
- [6] 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 *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Berlin, Germany, Jul. 2019, pp. 6578–6582.
- [7] K. Fujiwara *et al.*, “Heart rate variability-based driver drowsiness detection and its validation with EEG,” *IEEE Trans. Biomed. Eng.*, vol. 66, no. 6, pp. 1769–1778, Jun. 2019.
- [8] S. Hu, R. L. Bowlds, Y. Gu, and X. Yu, “Pulse wave sensor for non-intrusive driver’s drowsiness detection,” in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Minneapolis, MN, USA, Sep. 2009, pp. 2312–2315.
- [9] M. Gromer, D. Salb, T. Walzer, N. M. Madrid, and R. Seepold, “ECG sensor for detection of driver’s drowsiness,” *Procedia Comput. Sci.*, vol. 159, pp. 1938–1946, Jan. 2019.
- [10] K. T. Chui, K. F. Tsang, H. R. Chi, C. K. Wu, and B. W.-K. Ling, “Electrocardiogram based classifier for driver drowsiness detection,” in *Proc. IEEE 13th Int. Conf. Ind. Informat. (INDIN)*, Cambridge, U.K., Jul. 2015, pp. 600–603.
- [11] P. C. Nissimagoudar and A. V. Nandi, “Precision enhancement of driver assistant system using EEG based driver consciousness analysis & classification,” in *Computational Network Application Tools for Performance Management*. Singapore: Springer, Oct. 2020, pp. 247–257.
- [12] C.-T. Lin, R.-C. Wu, S.-F Liang, W.-H. Chao, Y.-J. Chen, and T.-P. Jung, “EEG-based drowsiness estimation for safety driving using independent component analysis,” *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 52, no. 12, pp. 2726–2738, Dec. 2005.
- [13] J. Solaz *et al.*, “Drowsiness detection based on the analysis of breathing rate obtained from real-time image recognition,” *Transp. Res. Procedia*, vol. 14, pp. 3867–3876, Jan. 2016.
- [14] K. Lee, D. K. Han, and H. Ko, “Video analytic based health monitoring for driver in moving vehicle by extracting effective heart rate inducing features,” *J. Adv. Transp.*, vol. 2018, Dec. 2018, Art. no. 8513487.
- [15] A. Pai, A. Veeraraghavan, and A. Sabharwal, “HRVCam: Robust camera-based measurement of heart rate variability,” *J. Biomed. Opt.*, vol. 26, no. 2, Feb. 2021, Art. no. 022707.
- [16] E. M. Nowara, T. K. Marks, H. Mansour, and A. Veeraraghavan, “SparsePPG: Towards driver monitoring using camera-based vital signs estimation in near-infrared,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops*, Salt Lake City, UT, USA, Jun. 2018, pp. 1272–1281.
- [17] 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 Process. Mag.*, vol. 35, no. 3, pp. 59–80, May 2018.
- [18] K. J. R. Liu and B. Wang, *Wireless AI: Wireless Sensing, Positioning, IoT, and Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2019.
- [19] E. Cianca, M. De Sanctis, and S. Di Domenico, “Radios as sensors,” *IEEE Internet Things J.*, vol. 4, no. 2, pp. 363–373, Apr. 2017.
- [20] F. Zhang, C. Wu, B. Wang, H.-C. Lai, Y. Han, and K. J. R. Liu, “WiDetect: Robust motion detection with a statistical electromagnetic model,” *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 3, no. 3, pp. 1–24, Sep. 2019.
- [21] F. Zhang, C. Chen, B. Wang, and K. J. R. Liu, “WiSpeed: A statistical electromagnetic approach for device-free indoor speed estimation,” *IEEE Internet Things J.*, vol. 5, no. 3, pp. 2163–2177, Jun. 2018.
- [22] Q. Xu, B. Wang, F. Zhang, D. S. Regani, F. Wang, and K. J. R. Liu, “Wireless AI in smart car: How smart a car can be?” *IEEE Access*, vol. 8, pp. 55091–55112, 2020.
- [23] S. D. Regani, Q. Xu, B. Wang, M. Wu, and K. J. R. Liu, “Driver authentication for smart car using wireless sensing,” *IEEE Internet Things J.*, vol. 7, no. 3, pp. 2235–2246, Mar. 2020.
- [24] F. Wang, F. Zhang, C. Wu, B. Wang, and K. J. R. Liu, “Respiration tracking for people counting and recognition,” *IEEE Internet Things J.*, vol. 7, no. 6, pp. 5233–5245, Jun. 2020.
- [25] C. Chen *et al.*, “TR-BREATH: Time-reversal breathing rate estimation and detection,” *IEEE Trans. Biomed. Eng.*, vol. 65, no. 3, pp. 489–501, Mar. 2018.
- [26] X. Wang, C. Yang, and S. Mao, “TensorBeat: Tensor decomposition for monitoring multiperson breathing beats with commodity WiFi,” *ACM Trans. Intell. Syst. Technol.*, vol. 9, no. 1, pp. 1–27, Oct. 2017.
- [27] F. Wang, F. Zhang, C. Wu, B. Wang, and K. J. R. Liu, “ViMo: Multiperson vital sign monitoring using commodity millimeter-wave radio,” *IEEE Internet Things J.*, vol. 8, no. 3, pp. 1294–1307, Feb. 2021.

- [28] F. Adib, H. Mao, Z. Kabelac, D. Katabi, and R. C. Miller, "Smart homes that monitor breathing and heart rate," in *Proc. 33rd Annu. ACM Conf. Human Factors Comput. Syst.*, Apr. 2015, pp. 837–846.
- [29] Z. Yang, P. H. Pathak, Y. Zeng, X. Liran, and P. Mohapatra, "Monitoring vital signs using millimeter wave," in *Proc. 17th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, Jul. 2016, pp. 211–220.
- [30] M. Mercuri, I. R. Lorato, Y.-H. Liu, F. Wieringa, C. V. Hoof, and T. Torfs, "Vital-sign monitoring and spatial tracking of multiple people using a contactless radar-based sensor," *Nat. Electron.*, vol. 2, no. 6, pp. 252–262, Jun. 2019.
- [31] M. Zhao, F. Adib, and D. Katabi, "Emotion recognition using wireless signals," in *Proc. 22nd Annu. Int. Conf. Mobile Comput. Netw.*, Oct. 2016, pp. 95–108.
- [32] V. L. Petrović, M. M. Janković, A. V. Lupšić, V. R. Mihajlović, and J. S. Popović-Bozović, "High-accuracy real-time monitoring of heart rate variability using 24 GHz continuous-wave doppler radar," *IEEE Access*, vol. 7, pp. 74721–74733, 2019.
- [33] N. T. P. 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 *Proc. IEEE MTT-S Int. Microw. Biomed. Conf. (IMBioC)*, vol. 1. Nanjing, China, May 2019, pp. 1–4.
- [34] T. Ohtsuki and E. Mogi, "Heartbeat detection with Doppler radar based on estimation of average R-R interval using Viterbi algorithm," in *Proc. IEEE 27th Annu. Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC)*, Valencia, Spain, Dec. 2016, pp. 1–5.
- [35] F. Wang, X. Zeng, C. Wu, B. Wang, and K. J. R. Liu, "mmHRV: Contactless heart rate variability monitoring using millimeter-wave radio," *IEEE Internet Things J.*, vol. 8, no. 22, pp. 16623–16636, Nov. 2021.
- [36] S. Leonhardt, L. Leicht, and D. Teichmann, "Unobtrusive vital sign monitoring in automotive environments—A review," *Sensors*, vol. 18, no. 9, p. 3080, Sep. 2018.
- [37] Z. Yang, M. Bocca, V. Jain, and P. Mohapatra, "Contactless breathing rate monitoring in vehicle using UWB radar," in *Proc. 7th Int. Workshop Real-World Embedded Wireless Syst. Netw.*, Nov. 2018, pp. 13–18.
- [38] C. Li and J. Lin, "Random body movement cancellation in Doppler radar vital sign detection," *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 12, pp. 3143–3152, Dec. 2008.
- [39] M.-C. Tang, F.-K. Wang, and T.-S. Horng, "Single self-injection-locked radar with two antennas for monitoring vital signs with large body movement cancellation," *IEEE Trans. Microw. Theory Techn.*, vol. 65, no. 12, pp. 5324–5333, Dec. 2017.
- [40] A. Sharifi, M. Baboli, M. Eshghi, and A. Ahmadian, "Respiration-rate estimation of a moving target using impulse-based ultra wideband radars," *Aust. Phys. Eng. Sci. Med.*, vol. 35, no. 1, pp. 31–39, Dec. 2012.
- [41] Y. S. Koo, L. Ren, Y. Wang, and A. E. Fathy, "UWB microdoppler radar for human gait analysis, tracking more than one person, and vital sign detection of moving persons," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Seattle, WA, USA, Jan. 2013, pp. 1–4.
- [42] S. K. Leem, F. Khan, and S. H. Cho, "Vital sign monitoring and mobile phone usage detection using IR-UWB radar for intended use in car crash prevention," *Sensors*, vol. 17, no. 6, p. 1240, May 2017.
- [43] Z. Yang, H. Shi, S. Zhao, and X.-D. Huang, "Vital sign detection during large-scale and fast body movements based on an adaptive noise cancellation algorithm using a single doppler radar sensor," *Sensors*, vol. 20, no. 15, p. 4183, Jul. 2020.
- [44] C. S. Marino and P. M. Chau, "High-resolution DOA estimation from synthetic aperture beamforming," in *Proc. IEEE Antennas Propag. Soc. Int. Symp.*, Washington, DC, USA, Dec. 2005, pp. 279–282.
- [45] L. G. Brown, "A survey of image registration techniques," *ACM Comput. Surv.*, vol. 24, no. 4, pp. 325–376, Dec. 1992.
- [46] M. Richards, *Fundamentals of Radar Signal Processing*, 2nd ed. Chicago, IL, USA: McGraw-Hill Educ., 2014. [Online]. Available: <https://books.google.com/books?id=i5KAMAECAAJ>
- [47] I. J. Schoenberg, *Cardinal Spline Interpolation*. Philadelphia, PA, USA: Soc. Ind. Appl. Math., 1973. doi: [10.1137/1.978161970555](https://doi.org/10.1137/1.978161970555)
- [48] A. D. Groote, M. Wantier, G. Cheron, M. Estenne, and M. Paiva, "Chest wall motion during tidal breathing," *J. Appl. Physiol.*, vol. 83, no. 5, pp. 1531–1537, Nov. 1997.
- [49] N. U. Rehman and H. Aftab, "Multivariate variational mode decomposition," *IEEE Trans. Signal Process.*, vol. 67, no. 23, pp. 6039–6052, Dec. 2019.
- [50] "Single-chip 76-GHz to 81-GHz industrial radar sensor integrating DSP, MCU and radar accelerator," Texas Instrum., Dallas, TX, USA, Data Sheet IWR1843, 2020. Accessed: Aug. 23, 2020. [Online]. Available: <https://www.ti.com/product/IWR1843>
- [51] "MIMO radar," Texas Instrum., Dallas, TX, USA, Rep. SWRA554A, 2020. Accessed: Aug. 23, 2020. [Online]. Available: <https://www.ti.com/cn/cn/lit/an/swra554a/swra554a.pdf?ts=1598454063259>
- [52] "Bitalino revolution plugged kit BT," BITALINO Inc. 2020. [Online]. Available: <https://bitalino.com/en/plugged-kit-bt> (Accessed: Aug. 23, 2020).
- [53] "Go direct® respiration belt." [Online]. Available: <https://www.vernier.com/product/go-direct-respiration-belt/> (Accessed: Dec. 1, 2020).
- [54] "MDOT SHA pavement condition." 2021. [Online]. Available: <https://data-maryland.opendata.arcgis.com/datasets/mdot-sha-pavement-condition-1>
- [55] "Origin wireless." [Online]. Available: <https://www.originwirelessai.com> (Accessed: Dec. 1, 2020).
- [56] T. Zheng, Z. Chen, C. Cai, J. Luo, and X. Zhang, "V2iFi: In-vehicle vital sign monitoring via compact RF sensing," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 4, no. 2, pp. 1–27, Jun. 2020.
- [57] F. Shaffer and J. P. Ginsberg, "An overview of heart rate variability metrics and norms," *Front. Public Health*, vol. 5, p. 258, Sep. 2017.
- [58] Z. Dong *et al.*, "A fatigued driving detection method based on frequency modulated continuous wave radar," in *Proc. IEEE Int. Conf. Consum. Electron. Comput. Eng. (ICCECE)*, Guangzhou, China, Jan. 2021, pp. 670–675.
- [59] Y. Wu, R. Gu, Q. Yang, and Y.-J. Luo, "How do amusement, anger and fear influence heart rate and heart rate variability?" *Front. Neurosci.*, vol. 13, p. 1131, Oct. 2019.
- [60] S. E. H. Kiashari, A. Nahvi, A. Homayounfar, and H. Bakhoda, "Monitoring the variation in driver respiration rate from wakefulness to drowsiness: A non-intrusive method for drowsiness detection using thermal imaging," *J. Sleep Sci.*, vol. 3, nos. 1–2, pp. 1–9, Nov. 2018.



Fengyu Wang (Member, IEEE) received the 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, and the Ph.D. degree in electrical engineering from the University of Maryland at College Park, College Park, MD, USA, in 2021.

She is currently a Lecturer with the School of Artificial Intelligence, Beijing University of Posts and Telecommunications. Her current research

interests include Internet of Things, wireless sensing, statistical signal processing, and semantic communications.



Xiaolu Zeng received the B.S. degree from Harbin Institute of Technology, Harbin, China, in 2014, and the Ph.D. degree from the School of Electronic Engineering, Xidian University, Xi'an, China, in 2020.

He is currently a Postdoctoral Research Associate with the Department of Electrical and Computer Engineering, University of Maryland at College Park, College Park, MD, USA, where he was a visiting student with the Department of Electrical and Computer Engineering from September 2017 to June 2020. His research interests include target detection, tracking and localization, wireless sensing, 5G millimeter wave communication and applications, intelligent recognition and classification, and advanced driver assistance systems.



Chenshu Wu (Senior Member, IEEE) received the B.E. degree from the School of Software, Tsinghua University, Beijing, China, in 2010, and the Ph.D. degree in computer science from Tsinghua University in 2015.

He is currently an Assistant Professor with the Department of Computer Science, University of Hong Kong, Hong Kong. He is also the Chief Scientist with Origin Wireless Inc., Greenbelt, MD, USA, a spotlight startup commercializing WiFi sensing products. He has authored/coauthored three books, more than 70 papers, and more than 40 granted/filed patents. His research focuses on wireless IoT systems at the intersection of wireless sensing, ubiquitous computing, digital health, and the Internet of Things.

Dr. Wu has received four paper awards. He is a member of the ACM.



Beibei Wang (Senior Member, IEEE) received the B.S. degree (Hons.) in electrical engineering from the University of Science and Technology of China, Hefei, China, in 2004, and the Ph.D. degree in electrical engineering from the University of Maryland at College Park, College Park, MD, USA, in 2009.

She was with the University of Maryland as a Research Associate from 2009 to 2010 and with Qualcomm Research and Development, San Diego, CA, USA, from 2010 to 2014. Since 2015, she has been with Origin Wireless Inc., Greenbelt, MD, USA, where she is currently the Vice President of Research. She is also affiliated with the University of Maryland. She has coauthored *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). Her research interests include Internet of Things, mobile computing, wireless sensing and positioning, and communications and networking.

Dr. Wang was a recipient of the 2020 IEEE INTERNET OF THINGS JOURNAL Best Paper Award, the 2015 IEEE Signal Processing Society Overview Paper Award and several research and invention awards from the University of Maryland. She has served on the Editorial Board of IEEE SIGNAL PROCESSING LETTERS, IEEE INTERNET OF THINGS JOURNAL, and IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS.



K. J. Ray Liu (Fellow, IEEE) is a Distinguished University Professor and a Distinguished Scholar-Teach of University of Maryland at College Park, College Park, MD, USA, where he is also a 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. As the founder and the President of Origin Wireless, Greenbelt, MD, USA,

his inventions won the 2017 CEATEC Grand Prix and three CES Innovation Awards, including CES Best of Innovation in 2021.

Mr. Liu received the Teaching and Research Recognitions from the University of Maryland, including University-Level Invention of the Year Award and College-Level Poole and Kent Senior Faculty Teaching Award, the Outstanding Faculty Research Award, and the Outstanding Faculty Service Award, all from A. James Clark School of Engineering. He is the 2021 IEEE President Elect. He was the IEEE Vice President, a Technical Activities, and a member of IEEE Board of Director as Division IX Director. He has also served as the President of IEEE Signal Processing Society, where he was the Vice President—Publications and the Editor-in-Chief of *IEEE Signal Processing Magazine*.