Contactless Respiration Monitoring Using Ultrasound Signal With Off-the-Shelf Audio Devices

Tianben Wang, Daqing Zhang, *Senior Member, IEEE*, Leye Wang, Yuanqing Zheng, *Member, IEEE*, Tao Gu, *Senior Member, IEEE*, Bernadette Dorizzi, and Xingshe Zhou

***Abstract*—Recent years have witnessed advances of Internet of Things technologies and their applications to enable contactless sensing and elderly care in smart homes. Continuous and real-time respiration monitoring is one of the important applications to promote assistive living for elders during sleep and attracted wide attention in both academia and industry. Most of the existing respiration monitoring systems require expensive and specialized devices to sense chest displacement. However, chest displacement is not a direct indicator of breathing and thus false detection may often occur. In this paper, we design and implement a real-time and contactless respiration monitoring system by directly sensing the exhaled airflow from breathing using ultrasound signals with off-the-shelf speaker and microphone. Exhaled airflow from breathing can be regarded as air turbulence, which scatters the sound wave and results in Doppler effect. Our system works as an acoustic radar which transmits sound wave and detects the Doppler effect caused by breathing airflow. We mathematically model the relationship between the Doppler frequency change and the direction of breathing airflow. Based on this model, we design a minimum description length-based algorithm to effectively capture the Doppler effect caused by exhaled airflow. We conduct extensive experiments with 25 participants (7 elders, 2 young kids, and 16 adults, including 11 females and 14 males) in four different rooms. The participants take four different sleep postures (lying on one's back, on right/left side, and on one's stomach) in different positions of the bed. Experiment results show that our system achieves a median error lower than 0.3 breaths/min (2%) for respiration monitoring and can accurately identify Apnea. The results also demonstrate that the system is robust to different respiration styles (shallow, normal, and deep), respiration rate variation, ambient noise, sensing distance variation (within 0.7 m), and transmitted signal frequency variation.**

***Index Terms*—Acoustic sensing, contactless sensing, Doppler effect, respiration detection.**

I. Introduction

NON-INTRUSIVE vital signs monitoring is an important topic for smart home and smart healthcare [1]–[4]. Respiration rate is a vital sign that informs health conditions, indicates progression toward recovery, and tracks decline of illness. Abnormal respiratory events, such as obstructive or central sleep apnea-hypopnea are quite common in the elders [25]–[27]. These respiration disorders reduce sleep quality and even threaten one's life. In particular, chronic obstructive pulmonary disease is the third most common cause of death for people aged 65 and above [8]. Thus, it is crucial to monitor one's respiration continuously and accurately in real-time at home for elders, especially those living alone, with respiratory diseases.

The traditional way to monitor vital signs requires a person to visit hospitals or wear dedicated respiration monitoring devices, such as thoracic impedance pneumography [9] or capnography [10]. However, these methods are quite costly and also intrusive, preventing these systems from large scale deployment at home settings. In order to develop cost-effective and nonintrusive respiration monitoring systems during sleep, researchers turn their attention to contactless sensing [4]–[7]. The approaches based on laser [12], microwave [14], commodity Wi-Fi [15], [16], [32]–[38] or acoustic devices [13], [31], [44] to monitor respiration rate in a contact-free manner. These approaches typically measure respiration by detecting the displacement of human chest. However, the chest movement is hard to measure with current approaches when a user is covered by a thick blanket or quilt during sleep. In addition, for a user suffering from obstructive sleep Apnea (OSA), the respiration may stop (i.e., no exhaled airflow), but the chest can still move as if the user is breathing normally [28]. As such, current approaches based on chest movement detection cannot accurately monitor respirations or reliably detect abnormalities. Considering the cost and functional requirements of respiration monitoring in home settings, an ideal respiration monitoring system should: 1) directly sense breathing airflow, rather than chest

Manuscript received August 9, 2018; revised September 14, 2018; accepted October 19, 2018. Date of publication October 23, 2018; date of current version May 8, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 61332013 and Grant 61702437, in part by the Chinese Scholarship Council Program, and in part by the Hong Kong ECS under Grant PolyU 252053/15E. *(Corresponding author: Tianben Wang.)*

T. Wang and X. Zhou are with the School of Computer Science, Northwestern Polytechnical University, Xi'an 710072, China (e-mail: wangtianbengx@163.com; zhouxs@nwpu.edu.cn).

D. Zhang is with the Department of Réseaux et Services Multimédia Mobiles, Institut Mines-Télécom/Télécom SudParis, 91011 Évry, France (e-mail: daqing.zhang@telecom-sudparis.eu).

L. Wang is with the Department of Computer Science and Engineering, Hong Kong University of Science and Technology, Hong Kong (e-mail: wangleye@gmail.com).

Y. Zheng is with the Department of Computing, Hong Kong Polytechnic University, Hong Kong (e-mail: csyqzheng@comp.polyu.edu.hk).

T. Gu is with the School of Computer Science, RMIT University, Melbourne, VIC 3001, Australia (e-mail: tao.gu@rmit.edu.au).

B. Dorizzi is with the Department of Recherche et Formations Doctorales, Institut Mines-Télécom/Télécom SudParis, 91011 Evry, France (e-mail: bernadette.dorizzi@telecom-sudparis.eu).

Digital Object Identifier 10.1109/JIOT.2018.2877607

movement; 2) leverage the existing cheap commodity devices; and 3) be nonintrusive and ideally contact-free.

In this paper, we design and implement a contact-less human respiration monitoring system using commodity speaker and microphone. The system directly senses exhaled airflow with a pair of acoustic transceiver which consists of one speaker and one microphone. The speaker transmits inaudible sound waves (i.e., disturb-free design) to be received by the microphone. Our motivation is based on the observation that exhaled airflow from breathing causes signal changes in the received sound waves. Our aim is to detect such changes for respiration monitoring. Note that our system does not require users to wear any devices.

Building an acoustic-based respiration system entails many practical challenges. First, although exhaled airflow would cause changes in sound waves, it still remains elusive to reliably monitor respirations by analyzing the received sound waves. Besides, without a comprehensive theoretical model to capture the inherent influences of respirations to the received sound waves, it is hard to configure the monitoring system to accurately monitor and reliably detect the sound waves. Second, a user may change sleep postures and the direction of exhaled airflow may vary during sleep. It is difficult to quantify the influence of changing airflow directions. Third, many real-life factors (e.g., body movement and wind) in the environment may cause changes in the airflow from breathing, affecting detection accuracy.

This paper aims to overcome these challenges. We first conduct several experiments to: 1) study the feasibility of sensing exhaled airflow leveraging the off-the-shelf acoustic devices and 2) investigate the characteristics of sound wave changes during human breathing. Based on our empirical study, we then build a theoretical model to describe the relationship between the variation pattern of Doppler shifts and the direction of breathing airflow. Based on our theoretical model, we optimize the system parameters to effectively catch the Doppler shift to substantially enhance detection performance. We then profile the Doppler shift using power spectrum density (PSD) in a specific band derived from the model mentioned above. Afterwards, to meet the real-time requirement, PSD is compressed exploiting minimum description length (MDL) principle [24]. To reduce the dimension of PSD while keeping sensitivity for exhaled airflow, the compressing method segments PSD into several bands and ensures that the PSD segment in the same band has similar sensitivity for exhaled airflow. Finally, we leverage the periodicity of respiration to differentiate body movement or other noise factors to further improve the robustness and accuracy of our system in real practical scenarios. A demo video is provided at https://tinyurl.com/ybncm2jz, which verifies the feasibility of sensing exhaled airflow using commodity acoustic devices (00:00–03:47), illustrates the theoretical model described in Section III-A (03:48–04:57), and records one measurement study (04:58–08:22). The contributions of this paper can be summarized as follows.

1) We design and implement a respiration monitoring system which directly senses breathing airflow by leveraging commodity microphone and speaker.
2) We model the relationship between the exhaled airflow direction and the Doppler frequency change pattern. Based on the model, we design an MDL-based compressing algorithm to effectively capture the Doppler effect caused by exhaled airflow.
3) We design an auto-correlation-based method to characterize the periodicity of the Doppler effect and differentiate respirations from nonperiodic Apnea and body movement.
4) We conduct extensive experiments to evaluate our system in four different rooms with 25 participants. The participants take four different sleep postures (lying on one's back, on right/left side, and on one's stomach) in different positions of the bed. The experiment results show that our system achieves a median error lower than 0.3 breaths/min (2%) for respiration monitoring and can accurately identify Apnea. We also conduct extensive experiments to evaluate system robustness in various scenarios and the results show that the system is robust to different respiration styles (shallow, normal, and deep), respiration rate variation, ambient noise, sensing distance variation (within 0.7 m), and transmitted signal frequency variation (within the band [20 kHz, 21 kHz]).

II. Related Work

A. Contact-Based Methods

The traditional vital sign monitoring systems require hospital visits and contact-based monitoring devices. For instance, thoracic impedance pneumography [9] needs to attach electrodes on a subject's chest and measures the change of electrical impedance during the subject's respiration. Capnography [10] utilizes the partial pressure of carbon dioxide to monitor a subject's respiration. Both devices need to be operated by medical specialists in hospitals and clinics and incur high deployment costs, which are not affordable for large scale deployment in ordinary homes. Moreover, they require subjects' direct contact with the devices and cause inconvenience for everyday use. Other works [19] adopt wearable sensing systems and build bed sensing systems with pressure sensors [20]. However, they still need specialized devices and are not suitable for large-scale deployment [21].

B. Contactless Method

Compared to the contact-based approaches, the contactless methods do not require direct contact with monitoring devices. In the literature, most of the contactless solutions leverage various signals to detect chest movement during respiration, such as laser [12], ultrasonic sensors [13], and radio frequency technologies, like microwave [14], WiFi [15], [16], [32]–[38], and RFID [46]. All these works detect respiration by measuring the chest movement displacement during respiration. However, the chest movement is hard to measure with current approaches when a user is covered by thick blanket or quilt during sleeping. Moreover, these methods assume that chest movement is a reliable indicator of respiration. However, such an assumption may not always hold in practice. For instance, the users suffering OSA could stop breathing (i.e., no exhaled

airflow) but their chests may still move as if the users were breathing normally [28]. It can be life-threatening if a monitoring system considers such chest movements without exhaled airflows as normal.

Additionally, visual analysis-based methods have also been investigated. For instance, camera-based method [11] utilizes a time-of-flight camera to record subjects' daily activities and adopts computer vision algorithms to analyze subjects' respiration. The main problem is that the camera-based methods may raise users' privacy concerns. Besides, camera-based methods highly rely on good lighting conditions.

Acoustic-based approaches [53] have recently attracted wide attention. The method proposed in [30] detects respiration by recording breathing sound with earphone. However, users may not be willing to wear earphone when they sleep. The methods proposed in [31] and [44] detect respiration by measuring chest displacement during breathing with acoustic signals. As mentioned above, these methods cannot reliably monitor respiration if a user is covered by blanket or in case of OSA. The work [18] requires a specialized device to generate and receive high frequency (40 kHz) ultrasound signals. Commodity speakers however typically transmit in the spectrum between 20 Hz to 20 kHz [48]. Moreover, many existing respiration detection approaches are designed for a controlled sleep posture, and thus do not work well when users change their postures during sleep.

III. Acoustic Doppler Shift Caused by Exhaled Airflow

Theoretically, the basic principle supporting acoustic respiration detection is that the intermittent exhaled airflow of respiration can be seen as turbulence and cause the Doppler frequency shift. In this section we intend to answer the following questions: 1) why exhaled airflow incurs acoustic Doppler shift and 2) what is the relationship between the speed and direction of exhaled airflow and the Doppler shift. In order to answer the above questions, we first derive a mathematical model to quantify the Doppler frequency shift caused by exhaled airflow. To verify the derived model, we conduct real experiments using commodity microphone and speaker. Finally, we study various factors that may interfere the acoustic Doppler frequency shift, e.g., body movements, wind, etc.

A. Acoustic Doppler Shift Caused by Exhaled Airflow

The exhaled airflow can be regarded as turbulence, which is able to scatter ultrasound signals. As turbulence contains many unsteady vortices moving irregularly [29], the velocity of turbulence at time t, i.e., $\boldsymbol{u}(t)$ is generally composed of two parts: 1) average velocity $\bar{\boldsymbol{u}}$ and 2) fluctuating velocity $\boldsymbol{u}'(t)$, i.e., $\boldsymbol{u}(t) = \bar{\boldsymbol{u}} + \boldsymbol{u}'(t)$. During breathing, the average velocity of exhaled airflow mainly contributes to $\bar{\boldsymbol{u}}$, whose direction and norm are relatively steady, while the irregularly moving vortices mainly contribute to $\boldsymbol{u}'(t)$, whose direction and norm change over time. Projecting $\boldsymbol{u}(t)$ to the line between the scatterer and the device, suppose the angle of $\boldsymbol{u}(t)$, $\bar{\boldsymbol{u}}$ and $\boldsymbol{u}'(t)$ are $\gamma(t)$, α, and $\beta(t)$, respectively, the projection result of $\boldsymbol{u}(t)$ can be denoted as

$$|\boldsymbol{u}(t)|\cos(\gamma(t)) = |\bar{\boldsymbol{u}}|\cos(\alpha) + \left|\boldsymbol{u}'(t)\right|\cos(\beta(t)). \tag{1}$$

The traditional Doppler shift [47] is given by

$$\Delta f = \left(\frac{\pm 2v}{c}\right)\cdot f. \tag{2}$$

Replacing v with the projection of $\boldsymbol{u}(t)$, i.e., $v = |\boldsymbol{u_x}(t)|\cos(\gamma(t))$, we can finally derive the Doppler shift caused by exhaled airflow as

$$\begin{aligned}\Delta f &= \left(\frac{\pm 2|\boldsymbol{u_x}(t)|\cos(\gamma(t))}{c}\right)\cdot f \\ &= \pm 2f\frac{|\overline{\boldsymbol{u_x}}|\cdot\cos(\alpha) + \left|\boldsymbol{u}'_{\boldsymbol{x}}(t)\right|\cdot\cos(\beta(t))}{c}.\end{aligned} \tag{3}$$

Δf is positive if the scatterer moves toward the device. Otherwise, Δf is negative.

According to our empirical results, the value of $\max(|\boldsymbol{u}'(t)|)$ can be roughly estimated as 0.9 m/s by setting α in (3) as $\pi/2$. Then, the value of $|\bar{\boldsymbol{u}}|$ can be roughly estimated as 1.3 m/s by setting α in (3) as 0. Referring to the above model, we have the following observations. When angle α is close to 0, we have $|\bar{\boldsymbol{u}}|\cdot\cos(\alpha)+|\boldsymbol{u}'(t)|\cdot\cos(\beta(t)) \approx |\bar{\boldsymbol{u}}|+|\boldsymbol{u}'(t)|\cdot\cos(\beta(t)) > 0$, so $\Delta f_{\text{turb}} > 0$. So we can only observe the frequency shift above the transmit frequency f. When angle α changes from 0 to $\pi/2$ gradually, we will observe that the frequency shift above f decreases and the frequency shift below f starts occurring. When angle α stabilizes at around $\pi/2$, we have $|\bar{\boldsymbol{u}}|\cdot\cos(\alpha)+|\boldsymbol{u}'(t)|\cdot\cos(\beta(t)) \approx |\boldsymbol{u}'(t)|\cdot\cos(\beta(t))$, due to the randomness of $\boldsymbol{u}'(t)$ in time domain and space, we will observe symmetric frequency shifts below and above f simultaneously.

The Doppler frequency shift variation when α changes from $\pi/2$ to π and the Doppler frequency shift variation when α changes from 0 to $\pi/2$ are symmetric with respect to transmit frequency f.

B. Empirical Verification of Acoustic Doppler Shift Caused by Exhaled Airflow

In this section we conduct two experiments to verify: 1) the feasibility of sensing exhaled airflow using commodity acoustic devices and 2) the theoretical model described in Section III-A.

1) Experimental Settings: We bind a speaker (JBL Jembe, 6 Watt, 80 dB) and a microphone (SAMSON Meteor Mic, 16 bit, 48 kHz) as a simple acoustic radar, as shown in Fig. 1(a). The device is placed in front of a subject facing toward the effective sensing area where the exhaled airflow passes [as shown in Fig. 1(b)] at a distance of 50 cm. The speaker transmits inaudible ultrasound waves at $f = 20$ kHz continuously. The speaker sends ultrasound signals which are scattered in the effective sensing area due to the exhaled airflow of the subject. Meanwhile, the microphone records ultrasound signals (with the sampling rate of 48 kHz with 16 bits).

2) Experimental Protocol: The subject is asked to breathe naturally first, and then wears a face mask and keeps breathing. This process is repeated twice. The PSD of the reflected ultrasound signal is shown synchronously.

3) Experiment Results: The experiment process was recorded in the demo video provided at the end of Section I (the part 00:00–03:47). From this experiment we

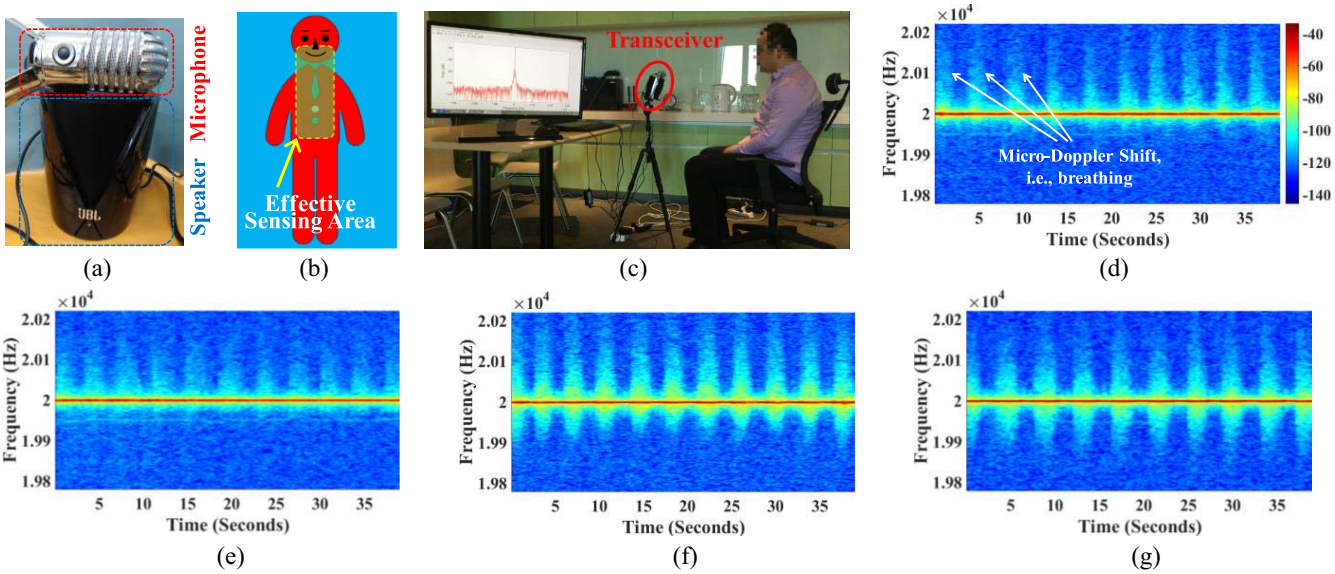

Fig. 1. Variation of Doppler frequency shift caused by exhaled airflow as radius angle varies. (a) Device we use. (b) Effective sensing area on the body of subject. (c) Panorama of experimental environment. (d) Radius angle $\alpha = 0$. (e) $\alpha = \pi/6$. (f) $\alpha = \pi/3$. (g) $\alpha = \pi/2$.

observe: when the subject breathes naturally without face mask, the echo PSD shows clear periodical amplitude variation, i.e., micro-Doppler shift around 20 kHz with respiration. Note that the ultrasound signal backward scattered by exhaled airflow is only a very small part of the echo and its power is low. The main part of the echo is the ultrasound signal reflected from the static room environment. This could explain the reason that even though the signal backward scattered by exhaled airflow embedded Doppler effect. The main frequency of the echo keeps the same as transmitted signal at a frequency of 20 kHz. While breathing with face mask, the exhaled airflow is blocked and the micro-Doppler effect disappears. The experimental results demonstrate that exhaled airflow can indeed cause Doppler shifts.

Next, we verify our theoretical model and conduct the following experiments.

4) Experimental Settings: To facilitate the adjustment of the acoustic beam direction, the acoustic transceiver [as shown in Fig. 1(a)] is fixed on a tripod facing toward the effective sensing area where the exhaled airflow passes [as shown in Fig. 1(c)].

5) Experimental Protocol: To verify our theoretical model, we first collect ultrasound echo signals in the following four scenarios: breathing at different angles $\alpha = 0$, $\pi/6$, $\pi/3$, and $\pi/2$, respectively, where α denotes the angle between the exhaled airflow direction and the acoustic beam direction. We then compute the PSD of the received ultrasound echo in each scenario to observe the micro-Doppler shift.

6) Experiment Results: As shown in Fig. 1(d)–(g), the results of the Doppler shift with different angle α match the results derived from our model. For $\alpha = 0$ [Fig. 1(d)], we only observe a relatively strong frequency shift above transmitting frequency 20 kHz. For $\alpha = \pi/6$ [Fig. 1(d)], the frequency shift above 20 kHz reduces. When $\alpha = \pi/3$ [Fig. 1(e)], the frequency shift below 20 kHz starts to appear. While $\alpha = \pi/2$ [Fig. 1(f)], we observe almost symmetrical frequency shifts on both sides of 20 kHz. The experiment process was recorded in the demo video provided at the end of Section I (the part 03:48–04:57)

Our experiments demonstrate that the Doppler frequency shift varies with the angle between the exhaled airflow direction and the acoustic beam direction. With the angle changes, the frequency shift may appear on one side or two sides of the transmitting frequency.

C. Other Factors Interfering Acoustic Doppler Shift

Other factors may interfere with the acoustic Doppler frequency shift caused by exhaled airflow. We now discuss two factors: 1) body movement and 2) wind.

1) Body Movement: The Doppler frequency shift caused by human body movement is much stronger than the Doppler frequency shift caused by exhaled airflow. When body movement exists, the Doppler frequency shift caused by exhaled airflow will be hard to detect.

2) Wind: The exhaled airflow is the direct detecting target in our system. When there exists wind in the effective sensing area [as shown in Fig. 1(b)], the exhaled airflow will be disturbed and influence the respiration monitoring accuracy.

In reality, when a person sleeps, most of these interference factors can be eliminated or well controlled.

1) A person will generally keep stable and her respiration rate can be roughly estimated while she is moving or turning for a short while during sleep. To mitigate the problem caused by body movement, we can suspend respiration detection when body movement occurs and activate it for detecting respiration after the body movement disappears.
2) People who need respiration monitoring could be asked to avoid fans or direct airflow blowing directly toward their bodies, especially for the elders and kids.

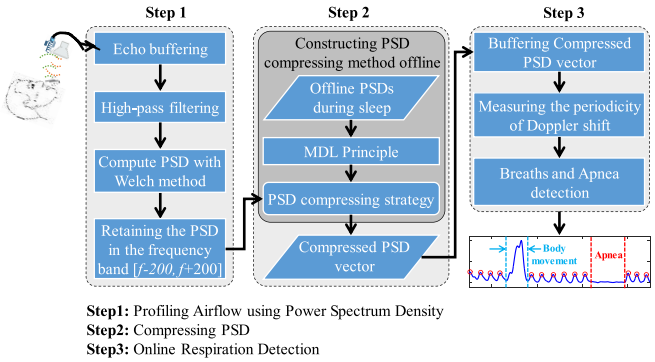

Fig. 2. Overview of the proposed framework.

IV. Respiration Detection Exploiting Periodical Acoustic Doppler Shift

As mentioned in Section III-C, to design a respiration detection system that is robust to breathing direction variation, we need to: 1) effectively capture the Doppler frequency shift regardless of the breathing direction and 2) identify the Doppler frequency shifts due to respiration from the received ultrasound signals. We employ PSD to capture the Doppler shift in a specific frequency band derived from the model in (3). We notice that human respiration is generally periodical, while Apnea and the interferences, such as body movement have no periodicity. Based on this observation, we differentiate respiration from other arhythmical factors by leveraging the rhythmicity of human respiration.

Based on the above idea, we propose a framework (as shown in Fig. 2) which consists of three key steps. In the first step, we profile the Doppler frequency shift using PSD. Next, in order to meet real-time requirement, we compress PSD into a short vector base on the MDL principle. Finally, we measure the periodicity of the elements in the short vector to detect whether the subject is breathing without body movement or Apnea. If the subject is detected as breathing without body movement or Apnea, we locate each breath as the peak of Doppler frequency shift. Otherwise, we use variance of Doppler shift to differentiate Apnea and body movement. We now describe the three key steps in details.

A. Profiling Doppler Shift Using PSD

PSD is an effective tool to profile the Doppler shift of the echo. We first construct a sliding window to buffer the echo. The sliding window abandons the obsolete echo sample and accepts the latest one continuously. Then we use a high-pass filter to filter out the ambient noise. Afterwards, we compute PSD to profile Doppler effect.

According to our empirical study, the value of average velocity $\overline{u_x}$ is about 1.3 m/s, the value of maximum of fluctuating velocity of exhaled airflow, i.e., $\max(|u'_x(t)|)$ is about 0.9 m/s. Based on the model described in (3), the theoretical upper bound of Doppler shift caused by exhaled airflow ($\alpha = 0$ and $\beta(t) = 0$) is 257 Hz. Due to the randomness of fluctuating velocity $u'_x(t)$ of exhaled airflow in time domain and space, the upper bound of Doppler shift during 1 s is about 200 Hz in practice. Therefore, we only retain the PSD in the frequency band $[f-200$ Hz, $f+200$ Hz$]$. As our emitted frequency is configured as an inaudible frequency of 20 kHz, this range is [19800 Hz, 20200 Hz].

B. Compressing PSD

For convenience in the following description, we define the part of echo PSD within the frequency band [19 800 Hz, 20 200 Hz] as effective PSD vector.

From the overview of the proposed system (as shown in Fig. 2), we can see that measuring the periodicity of Doppler shift is the key procedure in step 3. Our intuition to measure the periodicity of Doppler shift is to measure the periodicity of each frequency bin (i.e., one element in the PSD vector). However, the high dimension of PSD vector[1] leads to high computational overhead. We also observe that PSD values in some adjacent bins exhibit similar trends, meaning that the values in adjacent bins can be positively correlated. Based on these observations, we propose to group the PSD values in adjacent bins into a frequency band so as to enhance signal strength and reduce computational overhead. Ideally, the frequency bins grouped into a frequency bands should exhibit similar trends (named preciseness requirement), and meanwhile we want to minimize the total number of frequency bands (named conciseness requirement).

The sensitivity and informativeness of each frequency bin can be evaluated with its variance. For example, in our implementation, the iteration period of respiration detection system is 0.1 s, i.e., respiration detection system produces a 300-dimension[1] effective PSD vector per 0.1 s. If we collect the effective PSD vectors when monitoring respiration for 50 s, the system accumulates $50/0.1 = 500$ effective PSD vectors, i.e., a 500×300 matrix. The sensitivity of ith frequency bin can be measured as the variance of the ith column of 500×300 matrix. Larger variance indicates higher sensitivity and thus more informative in respiration monitoring. For example, Fig. 3(a) shows a set of effective PSD vectors during 5 min while breathing and Fig. 3(b) shows the variance (highlighted as the red line). Thus, the problem is transformed to segmenting the variance curve of effective PSD vectors with respect to the above two objectives, i.e., conciseness and preciseness.

As a matter of fact, the two requirements are contradictory to each other. For example, if each bin is placed into an independent band, preciseness is maximized but conciseness is minimized. In contrast, if all the bins are grouped into the same band, conciseness is maximized but preciseness is minimized. Therefore, we need to find an optimal tradeoff between conciseness and preciseness. To address this problem, we adopt the MDL [24] which allows us to strike a balance between the two requirements.

[1]The dimension of effective PSD vector $d = \text{NFFT} \times \Delta f / Fs \cdot Fs = 48$ kHz is sampling rate; $\Delta f = 400$ Hz is the width of frequency band $[f - 200$ Hz, $f + 200$ Hz$]$. NFFT is number of FFT points when computing PSD. In our implementation, we set NFFT as the length of data, i.e., $\text{NFFT} = \Delta t \times Fs$. $\Delta t = 0.75$ s is the buffer length in our system implementation. So $d = \text{NFFT} \times \Delta f / Fs = \Delta t \times Fs \times \Delta f / Fs = \Delta t \times \Delta f = 0.75 \times 400 = 300$.

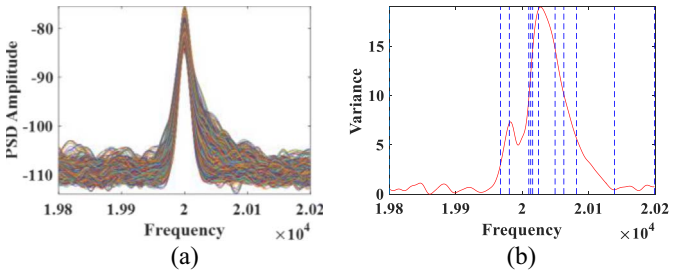

Fig. 3. Sensitivity of PSD for exhaling. (a) PSDs during breathing. (b) Variance (red line) as the sensitivity of PSD for exhaling and the blue dotted lines show an optimum partition employing the method based on MDL principle.

In particular, the MDL cost is defined as $L(H) + L(D|H)$, where H is the partition strategy, D denote the data; $L(H)$ is the cost of the partition strategy; $L(D|H)$ is the cost of the data description using partition strategy H [24]. The best partition strategy H is to minimize the MDL cost $L(H)+L(D|H)$.

Given:

1) The variance of all the frequency bins in effective PSD vector during the specific time slot, $\text{Vars} = \{v_{f_1}, v_{f_2}, \ldots, v_{f_m}\}$, where v_{f_i} is variance of the point at frequency f_i.
2) *Partition Strategy H*: Suppose Vars will be segmented into n bands, and the cut-points are $\text{CPs} = \{f_{c\_1}, f_{c\_2}, \ldots, f_{c_{n+1}}\}$, where $f_{c\_j} \in \{f_1, f_2, \ldots, f_m\}, j = 1, 2, \ldots, n$. CPs are in the ascending order, i.e., $f_{c\_1} = f_1 < f_{c\_2} < \ldots < f_{c_{n+1}} = f_m$. The elements of Vars in each band $[f_{c\_j}, f_{c_{j+1}}]$ are compressed as

$$H\left(\left[f_{c_j}, f_{c_j+1}\right]\right) = \frac{\max\left(\text{Vars}\left(\left[f_{c_j}, f_{c_j+1}\right]\right)\right) + \min\left(\text{Vars}\left(\left[f_{c_j}, f_{c_j+1}\right]\right)\right)}{2} \tag{4}$$

then, the cost of partition strategy H, $L(H)$ is defined as

$$L(H) = \sum_{i=1}^{n-1} \log 2\left(\text{abs}\left(f_{c_i} - f_{c_{i+1}}\right)\right). \tag{5}$$

The cost of the data description using partition strategy H, $L(D|H)$ is defined as

$$L(D|H) = \sum_{i=1}^{n-1} \sum_{j=c_i}^{c_{i+1}} \log 2\left(\text{abs}\left(v_{f_j} - H\left(\left[f_{c_j}, f_{c_j+1}\right]\right)\right)\right). \tag{6}$$

We can see that $L(H)$ measures the conciseness and $L(D|H)$ measures the preciseness. As mentioned previously, we need to find a segmenting strategy that can minimize the MDL cost [i.e., $L(H)+L(D|H)$]. In practice, the computation cost to find the optimal segmenting strategy is prohibitively high since we need to consider every subset of the points in the variance curve. Therefore, we adopt the approximate method [23] to get an approximate solution which provides a near-optimal grouping strategy (approximately 80% optimal) in much shorter time. In Fig. 3(b), the blue dotted lines show the near-optimal grouping result obtained using the approximate method.

We evaluate the performance of the grouping method with the four most common sleep postures (sleeping on one's back, on left side, on right side and on one's stomach). We collect four datasets, each of which contains data collected under each of the four sleep postures. We also combine the four datasets to form a mixed dataset. Then, we conduct the following experiments to validate the effectiveness of our segmenting method. The experiment settings in this section is the same as those in Section III.

TABLE I
RESPIRATION DETECTION ERROR USING DIFFERENT SEGMENT STRATEGY

	On one's back	Left side	Right side	On one's stomach
Seg_{back}	0 (0%)	0 (0%)	4.1 (27.3%)	0.6 (4%)
Seg_{left}	6.8 (45.3%)	0 (0%)	0.3 (2%)	0.3 (2%)
$\text{Seg}_{\text{right}}$	7.2 (48%)	0 (0%)	0 (0%)	0 (0%)
$\text{Seg}_{\text{stomach}}$	7.9 (52.7%)	0.3 (2%)	0.6 (4%)	0.3 (2%)
$\text{Seg}_{\text{mixed}}$	**0 (0%)**	**0 (0%)**	**0 (0%)**	**0.3 (2%)**
$\text{Seg}_{\text{uniform 8}}$	6.9 (46%)	7.4 (49.3%)	8.2 (54.7%)	7.9 (52.7%)
$\text{Seg}_{\text{uniform 10}}$	6.8 (45.3%)	9.2 (61.3%)	8.7 (58%)	8.5 (56.7%)
$\text{Seg}_{\text{uniform 12}}$	7.4 (49.3%)	7.9 (52.7%)	8.2 (54.7%)	8.1 (54%)

Error unit: breaths/min (percentage)

1) Experimental Protocol: First, we compute four segmenting strategies by applying our method to the breathing dataset of all four sleep postures (lying on one's back, on right/left side and on one's stomach) as well as the mix ed dataset, respectively, (as shown in Table I we denote these five segmenting strategies as Seg_{back}, Seg_{left}, $\text{Seg}_{\text{right}}$, $\text{Seg}_{\text{stomach}}$, and $\text{Seg}_{\text{mixed}}$). Second, we create three uniform segmenting strategies with the segment number 8, 10, 12 (as shown in Table I we denote them as $\text{Seg}_{\text{uniform-8}}$, $\text{Seg}_{\text{uniform-10}}$, and $\text{Seg}_{\text{uniform-12}}$). Finally, we replace the above eight segmenting strategies into our framework one by one, and test the framework performance with the four different sleep postures.

2) Experimental Results: Table I shows the respiration detection error using different segmenting strategies on different sleep posture datasets, i.e., $\text{Seg}_{\text{mixed}}$. We observe that the segmenting strategy computed by applying our segmenting method to the mixed breathing dataset achieves smallest error rate, while other seven segmenting strategies miss part of breaths in some datasets. The error rate 0.3 (2%) of $\text{Seg}_{\text{mixed}}$ when sleeping on one's stomach is caused by the fact that when sleeping on one's stomach, the Doppler shift is relatively weak than the Doppler shift when sleeping on left of right side. Weak Doppler shift will affect the accuracy of respiration monitoring. Nevertheless, our method is able to detect the Doppler shift in both sides of transmitted frequency and finally achieves error rate 2%, which is accurate enough for many respiration monitoring applications.

With the best segmenting strategy computed from the mixed dataset with our method, we compress effective PSD vector as a vector composed of the median values of all the segments.

C. Online Respiration Detection

It is a common sense that human respiration is usually rhythmical, i.e., relaxed periodicity, while body movement or other noise factors discussed in Section III-C are not rhythmical. Thus, the Doppler frequency shift embedded in echo

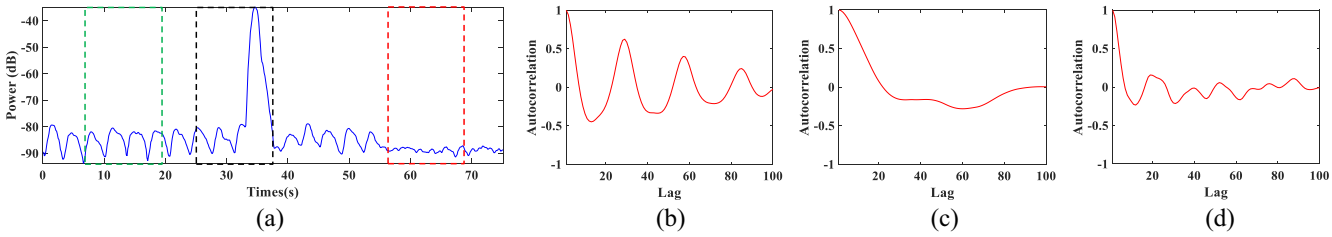

Fig. 4. Autocorrelation functions of the features corresponding to breathing, body movement, and environment noise. (a) One feature in the 75 s time window, the green, black, and red dotted rectangles correspond to the time windows when breathing without movement, body movement occurring, and only environment noise, respectively. (b)–(d) Autocorrelation function of the feature in green, black, and red time window, respectively.

PSD caused by respiration will inherit the relaxed periodicity. Hence, the periodicity of Doppler frequency shift can only occur during respiration. Therefore, we aim to exploit the periodicity of Doppler frequency shift to detect human respiration. In particular, we first measure the periodicity of Doppler frequency shift variations based on the compressed effective PSD vector. If the Doppler shift variation shows a strong periodicity in a specific time window, we can infer that the subject is breathing normally. Then we locate each breath as the peak of the Doppler shift. Otherwise, we can conclude that there is body movement or respiration arrest, i.e., Apnea occurring. In latter case, we use variance of Doppler shift to differentiate Apnea and body movement. By using a slide window, when body movement is detected, the system will automatically stop respiration detection, and when there is no movement occurring, the system detects the apnea event.

1) Measuring the Periodicity of Doppler Shift: The periodicity of Doppler frequency shift is reflected in the variation of the echo PSD. We construct a sliding window to buffer the compressed effective PSD vector. The sliding window abandons the obsolete compressed effective PSD vector and accepts the latest one continuously. Suppose the compressed effective PSD vector is M-dimension, and the length of the sliding widow is N. The system updates an $N \times M$ matrix in each iteration. We measure the periodicity of each column employing autocorrelation function. Given the jth ($j = 1, 2, \ldots, M$) column $\{x_{(1,j)}x_{(2,j)} \ldots x_{(3,j)}\}$, where $x_{(i,j)}$ is the element at ith row and jth column of the matrix, the autocorrelation function of jth column is defined as follows:

$$R_j(k) = \frac{\frac{1}{N}\sum_{t=1}^{N-k}\left(x_{(t,j)} - \mu\right)\left(x_{(t+k,j)} - \mu\right)}{\sigma^2} \tag{7}$$

where μ and σ are the expectation and standard deviation of the jth column, respectively. If the column is near periodical [shown as the green time window in Fig. 4(a), corresponding to breathing quietly], its autocorrelation function [as shown in Fig. 4(b)] looks like a sinusoid but its amplitude decreases gradually. If the column fluctuates randomly and significantly [shown as the black time window in Fig. 4(a), corresponding to body movement], its autocorrelation function [as shown in Fig. 4(c)] looks like an exponential function with a base smaller than 1. If the trend of the column is relatively stable [shown as the red time window in Fig. 4(a), corresponding to Apnea], its autocorrelation function [as shown in Fig. 4(d)] varies irregularly.

Based on the characteristics of the autocorrelation function and our observations, we construct a model to recognize whether one column of the matrix exhibit strong periodicity. It can be used to dynamically filter out weak or none periodical columns and also filter interferences due to body movement as well as environment noise. In particular, the model works as follows.

Given the autocorrelations, $R_j(k)$, of the jth column, indexes of peaks (i.e., local maximums, except for the first peak whose values is 1) of the $R_j(k)$ are denoted as $k_{c_1}, k_{c_2}, \ldots, k_{c_{\text{num}}}$.

1) If $T_{\text{PN1}} \leq \text{num} \leq T_{\text{PN2}}$.
2) If $R_j(k_{c_1}) > R_j(k_{c_2}) > \ldots > R_j(k_{c_{\text{num}}})$ then the jth column exhibits strong periodical. T_{PN1} and T_{PN2} are two thresholds determined as follows. As human respiration rate typically ranges from 12 breaths/min to 40 breaths/min (note that, after exercise, human respiration can reach 40 breaths/min), and the length of slide window is 10 s in our experiments. Thus, there should be 2∼6.7 breaths in one sliding window meaning that there are 2∼7 peaks in the autocorrelation results. Thus, in our experiments, we set $T_{\text{PN1}} = 2$ and $T_{\text{PN2}} = 7$ to filter out body movements and other noise.

If the number of strong periodical columns is larger than a specific threshold T_{SPA}, we can infer that the subject is breathing normally. Otherwise, there is body movement or respiration arrest, i.e., Apnea happens. In our experiment, $T_{\text{SPA}} = 3$ works well in distinguishing periodic respirations and nonperiodic body movement and Apnea. In the next section, we describe how to identify normal breathing and Apnea online.

2) Identifying Breaths and Apnea Online: As mentioned above, when we infer that the subject is breathing normally (i.e., the number of strong periodical columns is larger T_{SPA}), the next step is to identify each breathing. During exhaling, the Doppler frequency shift will first increase and then decrease. The variation of the Doppler frequency shift over time can be described as the sum of instances of the periodical columns (as described in Section IV-C1). Thus, the exhaling can be identified as the peak of the normalized Doppler frequency shift. When we infer that body movement or Apnea happens (i.e., the number of strong periodical columns is not larger T_{SPA}), we use the variance of Dopper shift to identify Apnea. Even though body movement and Apnea both affect the periodicity of Doppler shift, they result in very different Doppler shift variation. Body movement causes drastic and irregular Doppler shift and Apnea causes no Doppler

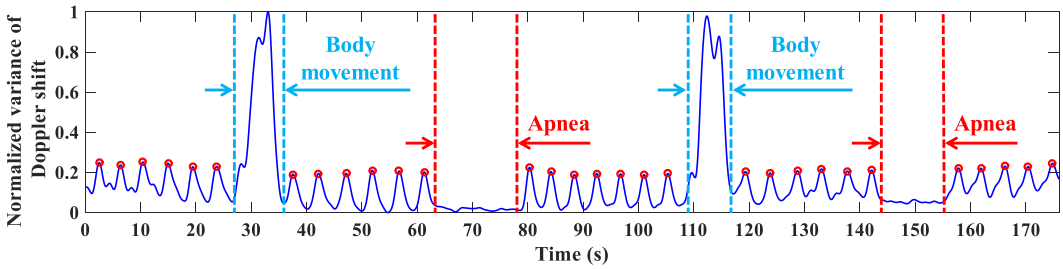

Fig. 5. Respiration detection results. The blue curve is the variation of the sum of instances of the periodical columns over time. The small red circles mark all respirations.

shift. So, we can use variance of normalized Doppler shift to identify Apnea. Specifically, if the variance of normalized Doppler shift in sliding window, whose length is 7 s, is larger than specific threshold T_{variance}, then body movement is detected. Otherwise, Apnea is detected. In our experiment, $T_{\text{variance}} = 0.01$ works well in differentiating Apnea and body movement.

Fig. 5 shows an example of the respiration detection results in 175 seconds' time frame. The blue curve is the variation of the sum of instances of the periodical columns over time. We can see that the system accurately identifies breaths, marked as small red circles, body movement (from 27 s to 36 s and from 109 s to 116 s) and Apnea (from 63 s to 78 s and from 143 s to 155 s). By using slide windows, when body movement is detected, the system will automatically stop respiration detection, and when movement stops, the system automatically resumes the detection.

V. Experimental Evaluation

In this section, we conduct comprehensive experiments to evaluate the proposed system. First, we introduce the system configuration and experiment settings. Then, we briefly describe the baseline method. We conduct experiments with 25 participants (7 elders, 2 young kids, and 16 adults, including 11 females and 14 males) in four different rooms. The participants take four different sleep postures (i.e., on one's back, on right/left side, and on one's stomach) in different positions of the bed. We compare our system with the baseline method in various experiment settings. In addition, we conduct experiments to test whether the system can identify Apnea. We also test the system robustness against body movement, wind, different respiration styles (shallow, normal, and deep), respiration rate variation, ambient noise, sensing distance variation, and transmitted signal frequency variation.

A. System Configuration and Experiment Settings

Theoretically, our design is not limited to COTS microphones and speakers and should be able to implement using smartphones. In practice, we face some technical challenges to implement using smartphone.

1) It is hard to ensure the that the transmitted acoustic beam passes through the exhaled airflow when the smartphone is placed on the nightstand or bed.

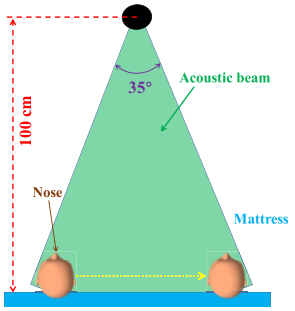

Fig. 6. Maximum detectable areas of one transceiver.

2) Even though the acoustic beam passes through the exhaled airflow, the relatively low power of the speaker on smartphone cannot ensure that the microphone receive sufficiently strong echo.

Therefore, currently we implement our respiration monitoring system with commodity microphones and speakers. Specifically, our system consists of a pair of commodity speaker and microphone [shown in Fig. 1(a)] which forms an acoustic transceiver. The speaker is programmed to transmit 20 kHz acoustic wave continuously, which is inaudible to users. Meanwhile, the microphone receives the echo at 48 kHz sampling rate and sends it to the connected laptop for data processing and respiration detection. For data collection and processing, the transceivers are connected to two laptops (Thinkpad T450 with Intel Core i5-5200 CPU, 8G RAM; Dell Latitude E6540 with Intel Core i7-4800MQ, 4GB RAM). The proposed respiration detection algorithms are implemented in MATLAB and run on each laptop in real-time.

First of all, we test the maximum detectable areas of one transceiver. To this end, we ask each participant to lie on a bed and place one transceiver above the subject's head for respiration detection. The experiment settings and corresponding results are illustrated in Fig. 6. We can see that the effective maximum detection distance is about 100 cm and the angle of the detectable area for sleeping on one's back is about 35°. We note that one transceiver cannot fully cover all possible facing directions of a subject. As such, we deploy two transceivers at both sides of a subject to fully cover different sleeping postures. As shown in Fig. 7, two transceivers are placed at the upper-left and upper-right of the head, respectively, facing the effective sensing area [as shown in Fig. 1(b)] with an angle of about 60°. The perpendicular distance between mattress and

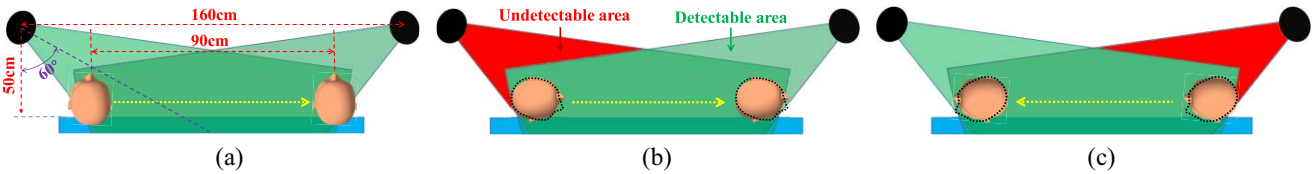

Fig. 7. Device layout and detectable facing in different sleep postures. (a) Sleep on one's back. (b) Sleep on right side and on one's stomach while the head (the head contour is highlighted as back dashed line) facing right. (c) Sleep on left side and on one's stomach while the head (the head contour is highlighted as back dashed line) facing left.

Fig. 8. System settings in real room environment.

device is about 50 cm and the distance between two pairs of devices is about 160 cm. Fig. 8 shows the real experiment environment.

According to the bias and limits of agreement of clinical respiration rate monitoring device [49], our targeted error for respiration monitoring should be smaller than 1 breaths/min. This accuracy is not only enough for general respiration monitoring applications but also can be used for the patients after surgery [49].

B. Baseline Approach

To the best of our knowledge, the paper in [18] is the only existing work that attempts to detect the airflow of respiration using ultrasound signals. Thus, we choose it as the baseline for comparison. However, the baseline method requires a specialized device to generate 40 kHz ultrasound signals. Moreover, the respiration detection approach was designed for a controlled sleep posture. For fair comparison: 1) the baseline approach is implemented using the same devices and deployment manner and 2) we adopt the same optimal parameter settings and configurations as specified in [18] and fine-tune the system.

C. System Performance Evaluation

We conduct comprehensive experiments to evaluate our system in four different rooms with 25 subjects, who have four different sleep postures in different positions of the bed and compare the system performance with the baseline approach. The experiment process was recorded in the demo video provided at the end of Section I (the part 04:58–08:22). In addition, we conduct experiments to test whether the system can identify Apnea.

1) Evaluation With Different Subjects: We recruit 25 participants (7 elders, 2 young kids, and 16 adults, including 11 females and 14 males) to evaluate the effectiveness of our system. To test the system usability for users, all participants are asked to set up the system and properly adjust facing direction of audio transceivers according to the requirements specified in Section V-A. We set aside 15 min for the participants to lie on his/her back so that the participants really fall asleep before measurements. We then detect the respiration with each subject for about 2 h. During the measurements, two subjects watch the video stream to record the ground truth manually. Fig. 9(a) shows the CDF of respiration detection error. We can see that the median respiration detection error of our approach is 0, while that of the baseline approach is around 0.9 breaths/min (6%). In addition, the max error of our system is about 0.6 breaths/min (4%), while that of the baseline is larger than 2.1 breaths/min (14%).

In addition, to further verify the proposed system, we use micro-movement sensitive mattress sleep monitoring system RS-611 (produced by Xinxingyangsheng Technology Company, Ltd., Bejing, China) to record the ground truth during the experiments. Two subjects are recruited to evaluate system performance for about 2 h. The experimental results show that the median error of the proposed system is 0.

This experimental results indicate: 1) our system is able to accurately detect human respiration when the participants sleep on his/her back, outperforming the baseline and 2) the system is easy to set up for users in practice.

2) Evaluation With Different Sleep Postures: Except for sleeping on one's back, lying on one side and sleeping on one's stomach are also common sleep postures. With the same experiment settings, all the 25 subjects are recruited to evaluate the performance of our system under the condition of subjects lying on one side and sleeping on one's stomach. Fig. 9(b)–(d) shows the CDF of respiration detection error when the subjects are sleeping on the left side, right side, and lying on one's stomach, respectively, (note that here lying on one's stomach requires the exhaled airflow not been blocked or covered by the pillow. Otherwise,

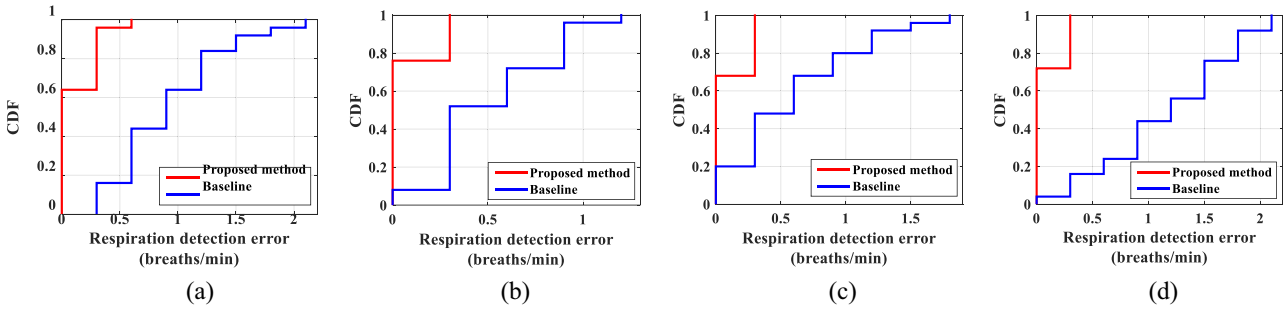

Fig. 9. Respiration detection error of 25 subjects with four different sleep postures. (a) Sleep on one's back. (b) Facing left. (c) Facing right. (d) Sleep on one's stomach.

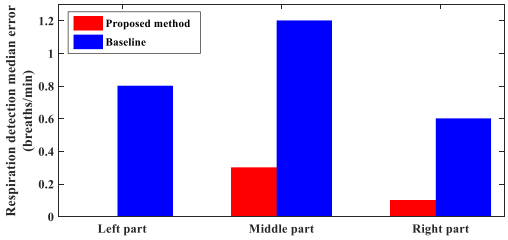

Fig. 10. Respiration detection median error in different positions in bed.

the system cannot detect respiration. In general, because of hypoxia the subject has to change sleep posture if the nose is covered by the pillow more than 10 s). The result shows that our system outperforms the baseline method with different sleep postures in terms of detection accuracy. In particular, the median errors of our system are all 0, while the median errors of the baseline are 0.3 breaths/min (2%), 0.6 breaths/min (4%), and 1.2 breaths/min (8%), respectively, for different sleep postures. In addition, the max error of our system can be controlled smaller than 0.6 breaths/min (4%), while the maximum error of the baseline reaches 2.2 breaths/min (14.7%). This is because when a subject sleeps on his/her back, the angle between airflow direction and acoustic beam direction is relatively small. Thus, the baseline method cannot reliably detect the Doppler frequency shift. In contrast, our method is able to detect the Doppler shift in both sides of transmitted frequency. Thus, the Doppler shift caused by breathing with all four sleep postures can be well captured.

3) Evaluation With Different Positions on Bed: In the experiment, the subjects are asked to sleep in different positions of the bed, i.e., left part of the bed, middle part of the bed, and right part of the bed. Fig. 10 shows the CDF of median respiration detection error when the subjects are sleeping in different positions of the bed. We can see that our method outperforms the baseline method and achieves a median respiration detection error lower than 0.3 breath/min.

4) Evaluation With Different Sleep Environments: We deploy the system in four rooms with different sizes and layouts. Fig. 11(a)–(d) shows the average respiration detection error CDF in the four test rooms, respectively. The median detection errors of our system for all four sleep postures in four test rooms are 0 breaths/min, while the median errors of the baseline are larger than 0.6 breaths/min (4%). There is no obvious difference in the four test rooms for both our method and the baseline. It indicates that our system is not sensitive to the experiment environment.

5) Apnea Detection Evaluation: Detecting Apnea is an important objective of monitoring respiration during sleep. Even though body movement and Apnea both affect the periodicity of Doppler shift, they result in very different Doppler shift variation. Body movement causes drastic and irregular Doppler shift and Apnea causes no Doppler shift. So, we can use the variance of Doppler shift to identify Apnea.

Constrained by legal issues, we could not test our system with real Apnea patients in hospitals for now. Instead, we simulate central Apnea and obstructive Apnea following the clinical symptom described as "hold breath for a while" [31], [39], and simulate Hypopnea event following the clinical symptom described as "breathing becomes shallow gradually and then recovers" [31]. We recruit 23 participants (13 male and 10 female) to test Apnea detection performance. Each participant is asked to simulate Apnea and generate body movement 10 times during the 30 minutes' respiration monitoring period. Fig. 12(a) and (b) show examples of Doppler shift variation when Apnea and Hypopnea happen, respectively. The experimental results show that the proposed system can accurately identify all the simulated Apnea.

D. System Robustness Testing

In this section, we conduct experiments to evaluate the robustness of our system. Specifically, we test various factors which may influence the performance of our system, including body movement, wind, different respiration styles (shallow, normal, and deep), respiration rate variation, ambient noise, sensing distance variation, and transmitted signal frequency variation.

1) Impact of Body Movement: Body movement generates strong but arhythmical Doppler frequency shift variation. The Doppler frequency shift caused by exhaled airflow would be submerged. Under this condition, it is difficult to detect breathing. To reduce the false alarm rate, our system is designed to suspend respiration detection once body movement is detected and recover for detecting respiration after the body movement disappears. Twenty two subjects are recruited to test whether our system can actually suspend when a body movement occurs. We detect subjects' respiration for about 30 min.

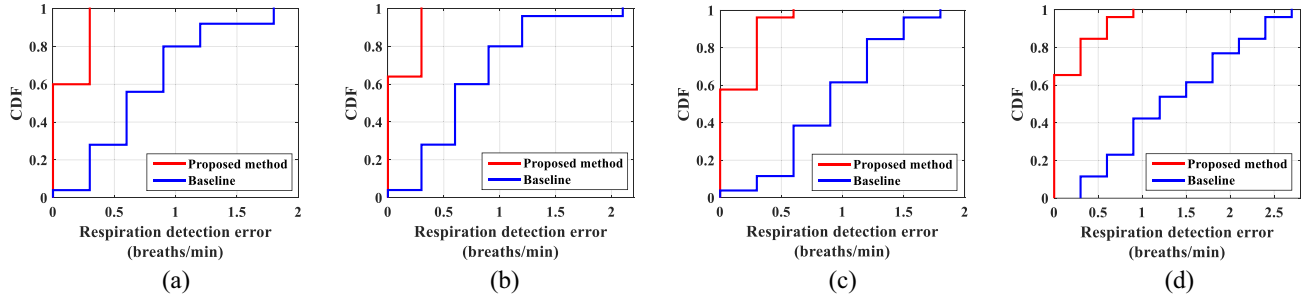

Fig. 11. Respiration detection error in four rooms of different sizes and layout.

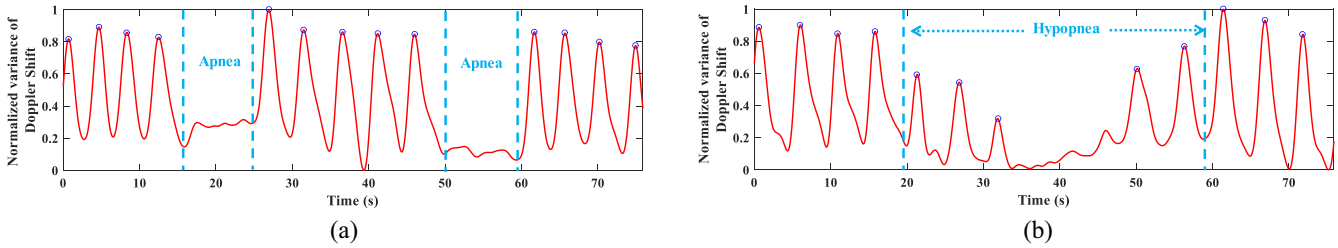

Fig. 12. Respiration detection when Apnea happens. (a) Apnea. (b) Hypopnea.

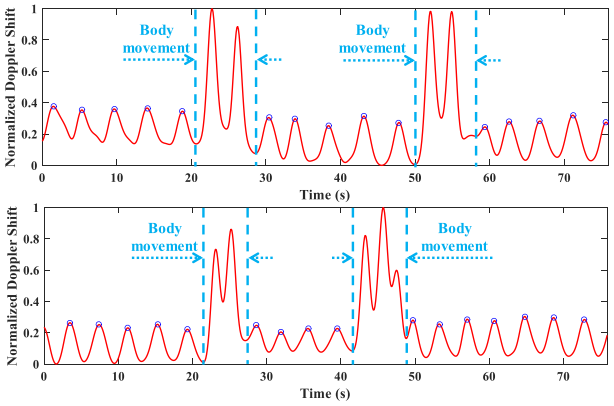

Fig. 13. Two examples of respiration detection process when body movement occurs.

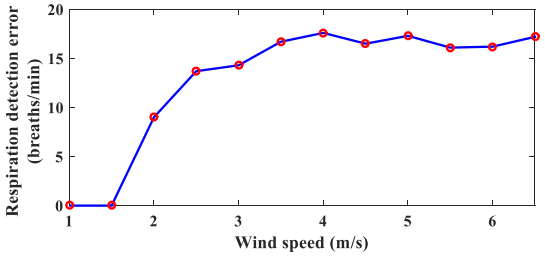

Fig. 14. Respiration detection error as the speed of interfering airflow varies.

During the detection process, the subjects change sleep postures or move limbs several times. Fig. 13 shows two examples that our system suspends the detection while body movement occurs, and resumes detection when movement stops.

2) Impact of Wind: Our system works by sensing the exhaled airflow. If the wind airflow in the effective sensing area [as shown in Fig. 1(b)] is large enough, the system cannot work well. To quantitatively test the impact of wind, we conduct an experiment using fan to generate airflow toward subject's body. The airflow speed is measured by a handheld anemometer (thermal anemometer testo 405-V1). We adjust the airflow speed by adjusting the distance between the fan and subject. The fan is placed at different places with different direction toward subject's head. Twenty subjects are recruited to test system performance under different airflow speed. Fig. 14 shows the respiration detection median errors with different airflow speed. We can see that when the indoor airflow speed is higher than 1.5 m/s regardless of the fan directions toward subject's head, the system cannot detect respiration rate accurately. The experiment results imply that our system is sensitive to indoor airflow. In reality, people under monitoring conditions always avoid fans or air conditions blowing directly toward their bodies, especially for elders and kids.

3) Impact of Different Respiration Styles: Breathing strength will affect the system performance. Generally, a deeper breath will lead to a lower respiration detection error. We recruit 21 subjects, including 4 young kids (8 years old on average), 12 adults (26 years old on average), and 5 elders (63 years old on average), to test the system performance when different subjects breathing with different respiration styles. In the first round, the participants breathe naturally. In the next two rounds, the subjects are asked to intentionally control their breath and take relatively shallow and deep breath, respectively. In each round, the system monitors their respiration for about 30 min. Fig. 15 shows the respiration detection median error for each subject category in each round. We can see that when breathing naturally and deeply, the respiration detection median error for three subject categories are all smaller than 0.5 breath/min. Even when the adults breathe gently, the median error of the proposed system is 0. When young kids and elders breathe gently, the respiration detection median error reaches 3.2 breaths/min (21.3%) and

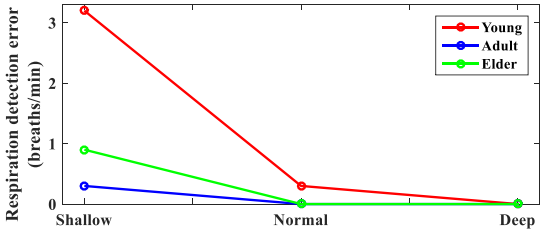

Fig. 15. Respiration detection error when subjects breathe with different respiration styles.

TABLE II
RESPIRATION DETECTION ERROR WITH DIFFERENT AMBIENT NOISE

Noise Source	Error (breaths/min, (percentage))
Talking in low voice	0 (0%)
Talking in normal voice	0.3 (2%)
Talking in loud voice	0 (0%)
Play music/video	0.1 (0.67%)
Noise from air condition	0 (0%)

0.9 breaths/min (6%), respectively. Too gent breaths not only mean low velocity, which results in weak Doppler shift (i.e., narrow frequency shift in PSD of echo) but also scatters backward weak ultrasound signal which results in low energy (i.e., low amplitude in PSD of echo). Low energy reflected signals and weak Doppler shifts will increase the respiration detection error rate. Fortunately, this problem can be mitigated by decreasing sensing distance. When we decrease the distance between the transceiver and subjects to 40 cm, the median error of all subject categories reduces to 0.

4) Impact of Respiration Rate Change: As our system detects the breathing rate by measuring the periodicity of Doppler shift, respiration rate change will weaken the periodicity and may influence our system. In this paper, we use a sliding window method which can discard the obsolete signals and adapt to the change of respiration rate. Twenty participants are asked to do high intensity exercises like push-up, which will increase the respiration rate significantly. Then, we let the participants lie on the bed and detect their respiration rates until the respiration rates gradually fall back to a normal level (10∼15 breathing counts per minute). Fig. 16 shows two examples of respiration detection results after high intensity exercise. The blue lines show the measured Doppler frequency shift and the pink lines track the instantaneous respiration rates. We can observe that the respiration rates change from about 40 breaths/min to about 15 breaths/min. Our system accurately tracks both rapid breathing and slow breathing throughout the process.

5) Impact of Ambient Noise: The proposed system senses human respiration using ultrasound signal. The ambient noise will also be received by the system. It is necessary to test whether the ambient noise has impact on respiration detection. We test the proposed system in several typical real scenes that continues to generate noise. Specifically, the real scenes include: 1) talking in low voice; 2) talking in normal voice; 3) talking in loud voice; 4) play music or video; and 5) noise from air condition. Twenty subjects are recruited to test system performance in the above five scenes. Table II shows the experimental results. We can see that

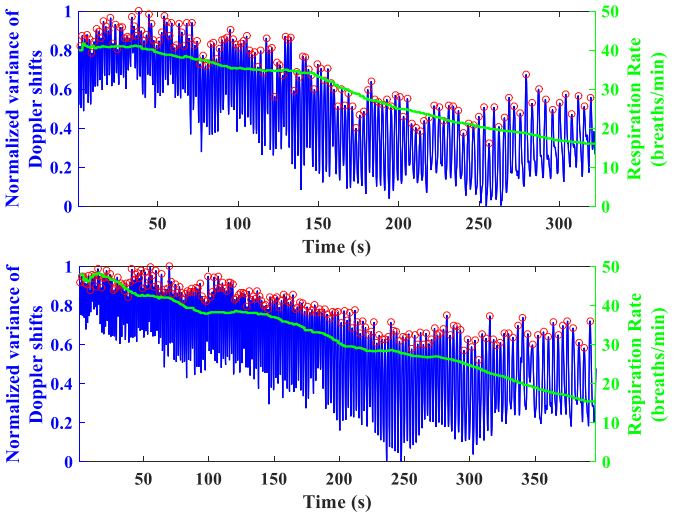

Fig. 16. Two examples of respiration detection process as respiration rate is changing.

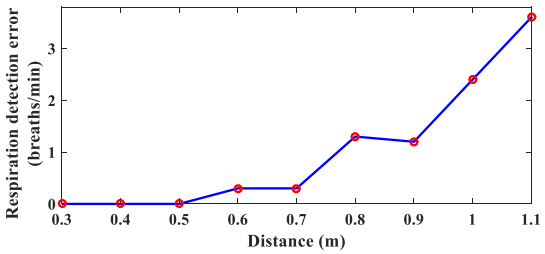

Fig. 17. Respiration detect error as sensing distance is varied.

the typical ambient noise has no impact on the proposed system. The proposed system senses exhaled airflow using 20 kHz ultrasound signal and capture the Doppler effect in the frequency band [19.8 kHz, 20.2 kHz]. In real scenes, there is hardly any ambient noise which can reach such a high frequency band. Studies show that the highest frequency of human voice is 3 kHz [40]; the highest frequency of music is 16 kHz [41]; the highest frequency of the noise produced by air condition or other household electric appliances is 8 kHz [42], [43]. The frequencies of all these ambient noises are far below the system working frequency band, hence, the ambient noise can be easily filtered out using lowpass filter.

6) Impact of Sensing Distance: We vary the distance between transceiver and subject, ranging from 0.3 m to 1.1 m with an interval of 0.1 m. 21 subjects are recruited to evaluate the performance of our system. At each position, we test for 30 min. Fig. 17 shows the median respiration detect error as distance varies. We can see that within 0.7 m, the system achieves respiration detection error smaller than 0.5 breaths/min (3.3%). Beyond 0.7 m, the error will increase with the distance mainly due to signal attenuation. We hence suggest setting the distance between transceiver and subject to a value smaller than 0.7 m. In practice user can also make a tradeoff between respiration detection error and sensing distance for a specific application environment.

7) Impact of Transmitted Signal Frequency: The frequency of the transmitted signal should be higher than the upper

TABLE III
EXPERIMENTAL RESULTS SUMMARY

Test Category	Test items	Results
Performance evaluation	Respiration monitoring with different subjects	Median error: 0; Max error: 0.6 (unit: breaths/min)
	Respiration monitoring with different sleep postures	Median error: 0; Max error: 0.6 (unit: breaths/min)
	Respiration monitoring with different positions on the bed	Median error of left part: 0, middle part: 0.3, right part: 0.1 (unit: breaths/min)
	Respiration monitoring with different sleeping environments	Median error: 0; Max error: 0.8 (unit: breaths/min)
	Respiration monitoring with Apnea	Accurately identifies all the simulated Apnea
Robustness testing	Effect of body movement	The system accurately identifies body movement; suspends the detection while body movement occurs, and resumes detection when movement stops.
	Effect of wind	If airflow speed is higher than 1.5 m/s, the system cannot monitor respiration accurately.
	Effect of different respiration styles	When breathing naturally and deeply, the respiration detection median error are smaller 0.5 breath/min; when young kids and elders breathe gently, median error reaches 3.2 breaths/min and 0.9 breaths/min, respectively.
	Effect of respiration rate change	The system is robust to respiration rate change.
	Effect of ambient noise	The system is robust to common ambient noise.
	Effect of sensing distance	Max error is smaller than 0.5 breaths/min (distance < 0.7 m)
	Effect of transmitting signal frequency	Max error is smaller than 1 breaths/min (frequency range [20kHz, 21kHz])

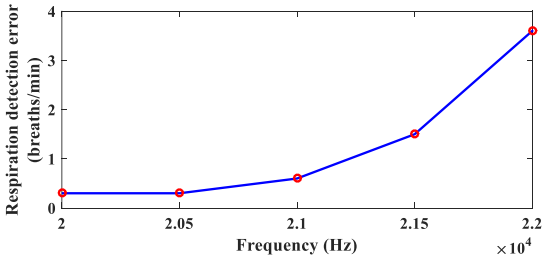

Fig. 18. Respiration detect error as the frequency of transmitted acoustic signal is varied.

bound of human audibility range, 20 kHz, and lower than the upper limit of frequency response of commodity acoustic device 22 kHz. We vary the transmitted signal frequency from 20 kHz to 22 kHz with an interval of 0.5 kHz. Twenty three subjects are recruited to performance of our system for about 10 min. Fig. 18 shows the median respiration detection error as the transmitted signal frequency varies. We can see that within 21 kHz, the system achieves respiration detection error smaller than 1 breaths/min (6.7%). Beyond 21 kHz, the error will increase with the transmitted signal frequency mainly due to the decrease of frequency response of commodity audio system. For commodity audio device, beyond 21 kHz, the system frequency response will decrease dramatically. In our system, we set the transmitted signal frequency to 20 kHz. In practice user can also make a tradeoff between respiration detection error and transmitted signal frequency for a specific application environment.

In summary, the proposed system is robust to different respiration styles (shallow, normal, and deep), respiration rate variation, ambient noise, sensing distance variation (within 0.7 m), and transmitted signal frequency variation (within the band [20 kHz, 21 kHz]), but sensitive to wind and body movement. The experimental results are summarized in Table III.

E. Discussion

The experiment results demonstrate that the proposed system can detect human respiration with four common sleep postures in different positions of the bed. The proposed system is robust to different respiration styles (shallow, normal, and deep), respiration rate variation, ambient noise, sensing distance variation (within 0.7 m), and the transmitted signal frequency variation (within the band [20 kHz, 21 kHz]). Yet, we note that current implementation can be improved in the following aspects.

1) Body Movement: As presented in Section V-D, the system is sensitive to sporadic body movement during sleep and the airflow around the subject. When body movement occurs, the weak Doppler shift caused by exhaled airflow would be submerged by the drastic and irregular Doppler shift caused by body movement; when interference airflow exists in the effective sensing area, the exhaled airflow will be disturbed. We plan to detect and filter out Doppler shifts caused by body movement in the future.

2) Multiple Users: Current system can only be used to monitor a single person lying on the bed at this moment. With multiple persons lying on the bed, the exhaled airflow may be blocked by other persons. In order to simultaneously monitor multiple people, we plan to study the feasibility of using motors to adjust acoustic transceivers in the future.

3) Apnea Detection: We note that due to legal issues, we could not evaluate our system with real Apnea patients in hospitals at this moment. The performance evaluation of Apnea detection was conducted by simulating the typical symptoms of Apnea. To better evaluate the efficacy of Apnea detection, we plan to conduct more extensive evaluations by inviting real Apnea patients to our lab in future work.

VI. CONCLUSION

This paper presents a continuous and real-time respiration monitoring system that is built purely using commodity audio devices. It utilizes the Doppler effect generated by the exhaled airflow of breath on the acoustic wave as the respiration indicator. We formally model the relationship between the exhaled airflow direction and the Doppler frequency change pattern. Based on this model, we design an MDL-based algorithm to effectively capture the Doppler effect caused by exhaled airflows. We implement a practical respiration monitoring system

with commodity microphone and speaker and tested the detection performance in various experiment settings. The extensive experiments demonstrate that 1) our proposed system achieves low respiration detection error [lower than 0.3 breaths/min (2%)] without assuming subject sleeping postures and positions in the bed and can accurately identify Apnea and 2) our proposed system is robust to different respiration styles (shallow, normal, and deep), respiration rate variation, ambient noise, sensing distance variation (within 0.7 m), and transmitted signal frequency variation. In order to further enhance the robustness of respiration monitoring performance, we plan to improve the system so that it can mitigate the influences caused by sporadic body movement during sleep. We also plan to further evaluate our system in larger scale deployment in ordinary homes.

ACKNOWLEDGMENT

The authors would like to express their special appreciation to the 25 volunteers for participating in their experiments. The study was approved by the medical ethics committee of NPU.

REFERENCES

[1] L. Catarinucci *et al.*, "An IoT-aware architecture for smart healthcare systems," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 515–526, Dec. 2015.

[2] S. Amendola, R. Lodato, S. Manzari, C. Occhiuzzi, and G. Marrocco, "RFID technology for IoT-based personal healthcare in smart spaces," *IEEE Internet Things J.*, vol. 2, no. 2, pp. 144–152, Apr. 2014.

[3] Y. Zhang, L. Sun, H. Song, and X. Cao, "Ubiquitous WSN for healthcare: Recent advances and future prospects," *IEEE Internet Things J.*, vol. 1, no. 4, pp. 311–318, Aug. 2014.

[4] D. Graham, G. Simmons, D. T. Nguyen, and G. Zhou, "A software-based sonar ranging sensor for smart phones," *IEEE Internet Things J.*, vol. 2, no. 6, pp. 479–489, Dec. 2015.

[5] Y. Gu, F. Ren, and J. Li, "PAWS: Passive human activity recognition based on WiFi ambient signals," *IEEE Internet Things J.*, vol. 3, no. 5, pp. 796–805, Dec. 2017.

[6] S. Mahmud, H. Wang, A. M. Esfar-E-Alam, and H. Fang, "A wireless health monitoring system using mobile phone accessories," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 2009–2018, Dec. 2017.

[7] Y. Gu *et al.*, "MoSense: A RF-based motion detection system via off-the-shelf WiFi devices," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 2326–2341, Dec. 2017.

[8] M. R. Vann. (2015). *The 15 Most Common Health Concerns for Seniors*. [Online]. Available: http://goo.gl/EQn2fn

[9] J. N. Wilkinson and V. U. Thanawala, "Thoracic impedance monitoring of respiratory rate during sedation—Is it safe?" *Anaesthesia*, vol. 64, no. 4, pp. 455–456, 2009.

[10] M. B. Jaffe, "Infrared measurement of carbon dioxide in the human breath:'Breathe-through' devices from Tyndall to the present day," *Anesthesia Analgesia*, vol. 107, no. 3, pp. 890–904, 2008.

[11] J. Penne, C. Schaller, J. Hornegger, and T. Kuwert, "Robust real-time 3D respiratory motion detection using time-of-flight cameras," *Int. J. Comput. Assisted Radiol. Surgery*, vol. 3, no. 5, pp. 427–431, 2008.

[12] T. Kondo, T. Uhlig, P. Pemberton, and P. D. Sly, "Laser monitoring of chest wall displacement," *Eur. Respiratory J.*, vol. 10, no. 8, pp. 1865–1869, 1997.

[13] S. D. Min *et al.*, "Noncontact respiration rate measurement system using an ultrasonic proximity sensor," *IEEE Sensors J.*, vol. 10, no. 11, pp. 1732–1739, Nov. 2010.

[14] M. Nowogrodzki, D. D. Mawhinney, and H. F. Milgazo, "Non-invasive microwave instruments for the measurement of respiration and heart rates," in *Proc. NAECON*, 1984, pp. 958–960.

[15] X. Liu, J. Cao, S. Tang, and J. Wen, "Wi-Sleep: Contactless sleep monitoring via WiFi signals," in *Proc. RTSS*, 2014, pp. 346–355.

[16] R. Ravichandran *et al.*, "WiBreathe: Estimating respiration rate using wireless signals in natural settings in the home," in *Proc. PerCom*, 2015, pp. 131–139.

[17] S. Venkatesh, C. R. Anderson, N. V. Rivera, and R. M. Buehrer, "Implementation and analysis of respiration-rate estimation using impulse-based UWB," in *Proc. IEEE Mil. Commun. Conf. (MILCOM)*, 2005, pp. 3314–3320.

[18] P. Arlotto, M. Grimaldi, R. Naeck, and J. M. Ginoux, "An ultrasonic contactless sensor for breathing monitoring," *Sensors*, vol. 14, no. 8, pp. 15371–15386, 2014.

[19] R. Paradiso, "Wearable health care system for vital signs monitoring," in *Proc. EMBS*, 2003, pp. 283–286.

[20] S. Nukaya, T. Shino, Y. Kurihara, K. Watanabe, and H. Tanaka, "Noninvasive bed sensing of human biosignals via piezoceramic devices sandwiched between the floor and bed," *IEEE Sensors J.*, vol. 12, no. 3, pp. 431–438, Mar. 2012.

[21] H. Gokalp and M. Clarke, "Monitoring activities of daily living of the elderly and the potential for its use in telecare and telehealth: A review," *Telemed. E-Health*, vol. 19, no. 12, pp. 910–923, 2013.

[22] P. D. Welch, "The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms," *IEEE Trans. Audio Electroacoust.*, vol. AE-15, no. 2, pp. 70–73, Jun. 1967.

[23] J.-G. Lee, H. Jiawei, and K.-Y. Whang, "Trajectory clustering: A partition-and-group framework," in *Proc. SIGMOD*, 2007, pp. 593–604.

[24] P. D. Grünwald, I. J. Myung, and M. A. Pitt, *Advances in Minimum Description Length: Theory and Applications*. Cambridge, MA, USA: MIT press, 2005.

[25] J. R. Cooke and S. Ancoli-Israel, "Normal and abnormal sleep in the elderly," *Handbook Clin. Neurol.*, vol. 98, no. 98, p. 653, 2011.

[26] D. Norman and J. S. Loredo, "Obstructive sleep apnea in older adults," *Clinics Geriatric Med.*, vol. 24, no. 1, pp. 151–165, 2008.

[27] T. L. Lee-Chiong and U. Magalang, "Monitoring respiration during sleep," *Clinics Chest Med.*, vol. 24, no. 2, pp. 297–306, 2003.

[28] *Wikipedia*. Accessed: May 5, 2018. [Online]. Available: https://en.wikipedia.org/wiki/Obstructive_sleep_apnea

[29] H. Tennekes and J. L. Lumley, *A First Course in Turbulence*. Cambridge, MA, USA: MIT Press, 1972.

[30] Y. Ren, C. Wang, J. Yang, and Y. Chen, "Fine-grained sleep monitoring: Hearing your breathing with smartphones," in *Proc. INFOCOM*, 2015, pp. 1102–1194.

[31] R. Nandakumar, S. Gollakota, and N. Watson, "Contactless sleep apnea detection on smartphones," in *Proc. Mobisys*, 2015, pp. 45–57.

[32] H. Wang *et al.*, "Human respiration detection with commodity WiFi devices: Do user location and body orientation matter?" in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2016, pp. 25–36.

[33] H. Abdelnasser, K. A. Harras, and M. Youssef, "UbiBreathe: A ubiquitous non-invasive WiFi-based breathing estimator," in *Proc. 16th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2015, pp. 277–286.

[34] O. Kaltiokallio, H. Yigitler, R. Jäntti, and N. Patwari, "Non-invasive respiration rate monitoring using a single COTS TX-RX pair," in *Proc. 13th Int. Symp. Inf. Process. Sensor Netw. (IPSN)*, 2014, pp. 59–69.

[35] X. Liu, J. Cao, S. Tang, J. Wen, and P. Guo, "Contactless respiration monitoring via off-the-shelf WiFi devices," *IEEE Trans. Mobile Comput.*, vol. 15, no. 10, pp. 2466–2479, Oct. 2016.

[36] N. Patwari, L. Brewer, Q. Tate, O. Kaltiokallio, and M. Bocca, "Breathfinding: A wireless network that monitors and locates breathing in a home," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 1, pp. 30–42, Feb. 2014.

[37] N. Patwari, J. Wilson, S. Ananthanarayanan, S. K. Kasera, and D. R. Westenskow, "Monitoring breathing via signal strength in wireless networks," *IEEE Trans. Mobile Comput.*, vol. 13, no. 8, pp. 1774–1786, Aug. 2014.

[38] C. Wu *et al.*, "Non-invasive detection of moving and stationary human with WiFi," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 11, pp. 2329–2342, Nov. 2015.

[39] *WebMD*. Accessed: Apr. 21, 2018. [Online]. Available: http://www.webmd.com/sleep-disorders/guide/sleep-disorders-symptoms-types

[40] *Human Voice Frequency Range*. Accessed: May 3, 2018. [Online]. Available: http://www.seaindia.in/blog/human-voice-frequency-range/

[41] *Table of Musical Notes and Their Frequencies and Wavelengths*. Accessed: May 5, 2018. [Online]. Available: http://www.liutaiomottola.com/formulae/freqtab.htm

[42] P. Susini *et al.*, "Characterizing the sound quality of air-conditioning noise," *Appl. Acoust.*, vol. 65, no. 8, pp. 763–790, 2004.

[43] S.-P. Huang and R.-P. Lai, "Frequency characteristics of interior noises in houses," in *Proc. World Sustain. Build. Conf.*, Sep. 2005, pp. 1503–1508.

[44] T. Wang *et al.*, "C-FMCW based contactless respiration detection using acoustic signal," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 1, no. 4, 2017, Art. no. 170.

[45] P. Nguyen, X. Zhang, A. Halbower, and T. Vu, "Continuous and fine-grained breathing volume monitoring from afar using wireless signals," in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, 2016, pp. 1–9.
[46] Y. Hou, Y. Wang, and Y. Zheng, "TagBreathe: Monitor breathing with commodity RFID systems," in *Proc. IEEE Int. Conf. Distrib. Comput. Syst.*, 2017, pp. 404–413.
[47] W. G. Carrara, R. S. Goodman, and R. M. Majewski, "Spotlight synthetic aperture radar: Signal processing algorithms," *J. Atmos. Solar Terrestrial Phys.*, vol. 59, no. 5, pp. 597–598, 1997.
[48] G. Zhang *et al.*, "DolphinAttack: Inaudible voice commands," in *Proc. ACM Conf. Comput. Commun. Security (CCS)*, 2017, pp. 103–117.
[49] A. Gaucher, D. Frasca, O. Mimoz, and B. Debaene, "Accuracy of respiratory rate monitoring by capnometry using the Capnomask(R) in extubated patients receiving supplemental oxygen after surgery," *Brit. J. Anaesthesia*, vol. 108, no. 2, pp. 316–320, 2012.
[50] A. G. Hagargund, R. Udayshankar, and N. Rashmi, "Radar based cost effective vehicle speed detection using zero cross detection," *Int. J. Elect. Electron. Data Commun.*, vol. 1, no. 9, Nov. 2013.
[51] J. Dybedal, "Doppler radar speed measurement based on a 24 GHz radar sensor," M.S. thesis, Norwegian Univ. Sci. Technol., Trondheim, Norway, 2013
[52] P. Yolanda, T. Guzmán, and J. T. Gonzales, "Development of a low-cost, short-range radar system to measure speed and distance," *Tecciencia*, vol. 12, no. 22, pp. 99–106, 2017.
[53] I. Bisio *et al.*, "Ultrasounds-based context sensing method and applications over the Internet of Things," *IEEE Internet Things J.*, to be published, doi: 10.1109/JIOT.2018.2845099.

Tianben Wang received the B.S. degree in computer science from Northwest A&F University, Yangling, China, in 2011, and the M.S. degree in computer application technology from Northwestern Polytechnical University, Xi'an, China, in 2013, where he is currently pursuing the Ph.D. degree.

His current research interests include ubiquitous computing, contactless behavior sensing, and intelligent elder assisting technology.

Daqing Zhang (M'11–SM'16) received the Ph.D. degree from the University of Rome La Sapienza, Rome, Italy, in 1996.

He is a Chair Professor with the School of Electrical Engineering and Computer Science, Peking University, Beijing, China. He has authored or co-authored over 200 technical papers in leading conferences and journals. His current research interests include context-aware computing, urban computing, mobile computing, big data analytics, and pervasive elderly care.

Dr. Zhang was a recipient of the ten-year CoMo Rea Impact Paper Award at IEEE Per-Com 2013, the Honorable Mention Award at ACM UbiComp 2015, the Best Paper Award at IEEE UIC 2015 and 2012, and the Best Paper Runner-Up Award at Mobiquitous 2011. He is an Associate Editor of the *ACM Transactions on Intelligent Systems and Technology* and the IEEE TRANSACTIONS ON BIG DATA. He served as the General or Program Chair for over ten international conferences and giving keynote talks at over 16 international conferences.

Leye Wang received the B.Sc. and M.Sc. degrees in computer science from Peking University, Beijing, China, and the Ph.D. degree from Institut Mines-Télécom/Télécom SudParis, Évry, France, and the Université Pierre et Marie Curie, Paris, France.

He is currently a Senior Researcher with the Department of Computer Science and Engineering, Hong Kong University of Science and Technology, Hong Kong. His current research interests include mobile crowdsensing and ubiquitous computing.

Yuanqing Zheng (S'11–GS'12–M'14) received the B.S. degree in electrical engineering and M.E. degree in communication and information system from Beijing Normal University, Beijing, China, in 2007 and 2010, respectively, and the Ph.D. degree from the School of Computer Engineering, Nanyang Technological University, Singapore, in 2014.

He is currently an Assistant Professor with the Department of Computing, Hong Kong Polytechnic University, Hong Kong. His current research interest includes mobile and wireless computing and RFID.

Dr. Zheng is a member of the ACM.

Tao Gu (S'03–M'07–SM'14) received the bachelor's degree from the Huazhong University of Science and Technology, Wuhan, China, the M.Sc. degree from Nanyang Technological University, Singapore, and the Ph.D. degree in computer science from the National University of Singapore, Singapore.

He is currently an Associate Professor in computer science with RMIT University, Melbourne, VIC, Australia. His current research interests include mobile computing, ubiquitous/pervasive computing, wireless sensor networks, distributed network systems, sensor data analytics, cyber physical system, Internet of Things, and online social networks.

Dr. Gu is a member of the ACM.

Bernadette Dorizzi received the Ph.D. (Thèse d'état) degree (with a focus on integrability of dynamical systems) theoretical physics from the University of Orsay (Paris XI-France), Orsay, France, in 1983.

She led the Electronics and Physics Department with Telecom SudParis (formerly, INT), Évry, France, from 1995 to 2009, where she has been a Professor since 1989, and the Dean of Research since 2013. She has coordinated the Biosecure Network of Excellence, and is currently the Chairwoman of the Biosecure Foundation. She is in charge of the Interaction for Multimedia Research Team. She has authored over 300 research papers and has supervised over 20 Ph.D. thesis. Her current research interests include pattern recognition and machine learning applied to activity detection, surveillance-video, and biometrics.

Xingshe Zhou received the M.S. degree from Northwestern Polytechnical University, Xi'an, China, in 1984.

He is a Professor with the School of Computer Science, Northwestern Polytechnical University. He is the Director with the Shaanxi Key Laboratory of Embedded System Technology, Xi'an. His current research interests include embedded computing and pervasive computing.