

C-FMCW Based Contactless Respiration Detection Using Acoustic Signal

TIANBEN WANG, Northwestern Polytechnical University
 DAQING ZHANG, Institut Mines-Télécom/Télécom SudParis
 YUANQING ZHENG, The Hong Kong Polytechnic University
 TAO GU, RMIT University, Australia
 XINGSHE ZHOU, Northwestern Polytechnical University
 BERNADETTE DORIZZI, Institut Mines-Télécom/Télécom SudParis

Recent advances in ubiquitous sensing technologies have exploited various approaches to monitoring vital signs. One of the vital signs is human respiration which typically requires reliable monitoring with low error rate in practice. Previous works in respiration monitoring however either incur high cost or suffer from poor error rate. In this paper, we propose a Correlation based Frequency Modulated Continuous Wave method (C-FMCW) which is able to achieve high ranging resolution. Based on C-FMCW, we present the design and implementation of an audio-based highly-accurate system for human respiration monitoring, leveraging on commodity speaker and microphone widely available in home environments. The basic idea behind the audio-based method is that when a user is close to a pair of speaker and microphone, body movement during respiration causes periodic audio signal changes, which can be extracted to obtain the respiration rate. However, several technical challenges exist when applying C-FMCW to detect respiration with commodity acoustic devices. First, the sampling frequency offset between speakers and microphones if not being corrected properly would cause high ranging errors. Second, the uncertain starting time difference between the speaker and microphone varies over time. Moreover, due to multipath effect, weak periodic components due to respiration can easily be overwhelmed by strong static components in practice. To address those challenges, we 1) propose an algorithm to compensate dynamically acoustic signal and counteract the offset between speaker and microphone; 2) co-locate speaker and microphone and use the received signal without reflection (self-interference) as a reference to eliminate the starting time difference; and 3) leverage the periodicity of respiration to extract weak periodic components with autocorrelation. Extensive experimental results show that our system detects respiration in real environments with the median error lower than 0.35 breaths/min, outperforming the state-of-the-arts.

CCS Concepts: • **Human-centered computing** → **Ubiquitous and mobile computing systems and tools**

KEYWORDS

Health Monitoring, Respiration Detection, Contactless Sensing, Acoustic sensing, Device-free Ranging,

ACM Reference format:

Tianben Wang, Daqing Zhang, Yuanqing Zheng, Tao Gu, Xingshe Zhou, and Bernadett Dorizzi. 2017. C-FMCW Based Contactless Respiration Detection Using Acoustic Signal. *PACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 4, Article 170 (December 2017), 20 pages. <https://doi.org/10.1145/3161188>¹

This work was supported by National Natural Science Foundation of China (No. 61332013), Chinese Scholarship Council Program.

Author's addresses: Tianben Wang, Xingshe Zhou, School of Computer Science, Northwestern Polytechnical University, Xi'an Shaanxi, China; Daqing Zhang, Bernadette Dorizzi, Institut Mines-Télécom/Télécom SudParis 9, rue Charles Fourier 91011 Evry Cedex, France; Yuanqing Zheng, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong; Tao Gu, School of Computer Science and IT, RMIT University, Australia. Daqing Zhang is the corresponding author (daqing.zhang@telecom-sudparis.eu).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page.

Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2017 Association for Computing Machinery.

2474-9567/2017/12-ART170 \$15.00

<https://doi.org/10.1145/3161188>

1 INTRODUCTION

1.1 Motivation

Recent years have witnessed the surge of ubiquitous sensing technologies and their applications in vital sign monitoring [1, 2]. Respiration rate is one of the important vital signs, often used to assess and track human health condition. For instance, research studies show that a deep breath helps reduce blood pressure while irregular respiration patterns may indicate chronic stress even progression of illness and decline in health [3, 4, 5]. Respiration monitoring in bed also helps track sleep quality and allow timely intervention in case of respiration abnormality such as obstructive or central sleep apnea-hypopnea which have been found quite common among the elderly [6, 7, 8]. In particular, chronic obstructive pulmonary disease (COPD) is the third most common causes of death for people aged 65 and above [9]. As such, respiration monitoring is essential to enable various ubiquitous healthcare applications. Since these systems often target patients or the elderly with chronic disease, a convenient and non-intrusive system is highly desirable.

1.2 Limitations of Prior Work

Several respiration monitoring systems have been proposed in clinical setting in early days such as thoracic impedance pneumography [10] and capnography [11], however these systems are inconvenient and intrusive. For instance, these systems typically require users to attach/wear special devices such as nasal probes or chest bands thus restricting body movements and causing inconvenience in their daily lives. Moreover, clinical respiration monitoring systems typically require well-trained professionals to set up the systems and assist users to attach/wear respiration monitoring devices properly. To minimize the discomfort brought by the invasive respiration systems, several attempts have been made for long-term respiration monitoring leveraging on wearable devices [12, 13]. While these systems are made more tolerable for the elderly, additional issues may arise such as user acceptance and usability [14].

Contactless sensing technologies explore a new possibility of monitoring vital signs without direct and physical contact to users. For instance, Penne et al. use a Time-of-Flight camera and apply advanced image processing algorithms to estimate human respiration rate in a home [15]. Unfortunately, such camera-based approaches require a subject to face the camera closely besides having privacy concerns and being affected by the lighting conditions. Kondo et al. deploy a laser sensor to measure the chest wall motion during respiration [16]. RF based methods have been also widely studied, ranging from Doppler radar [17], UWB radar [18], FMCW radar [19] to USRP (Universal Software Radio Peripheral) based solutions [20, 21]. The basic idea of these methods is to detect the chest movement displacement during breathing. Though these methods are able to accurately detect human respiration, they incur prohibitive cost, preventing these systems from large-scale deployment in typical home settings. To discover cost-effective solutions, researchers have recently turned their attentions to widely available commodity WiFi or Zigbee devices. These works can be further classified into two sub-categories: channel state information (CSI) based method [22-26] and received signal strength (RSS) based method [27-30]. Limited by the granularity of RSS, the RSS based approach suffers from low resolution, and it is subjected to interference in the environment. The CSI based method typically yields a higher resolution due to fine-grained channel information used. However, the study in [22] shows that the minute body movement may not generate detectable CSI changes. As such, the CSI based approach typically calls for careful selection of device locations and subcarriers to achieve robustness.

In recent years, acoustic signal has been used to enable accurate affordable contactless respiration monitoring in home environments. The study in [31] uses 40KHz ultrasonic signal to detect human respiration by sensing the exhaled airflow. The Doppler effect introduced by the velocity of exhaled airflow can be used to detect respiration. While the idea of using acoustic signal to detect respiration is promising, these systems are vulnerable to ambient airflow, such as wind or even the airflow introduced by body movement around it. Frequency Modulated Continuous Wave (FMCW) has been first used to detect human respiration using acoustic signal by measuring the chest movement displacement during breathing [32]. It achieves about 0.7 cm distance

estimation resolution and can detect the respiration with chest movement displacement of about 2 cm. However, the average respiration amplitude, i.e., the chest movement displacement, is only about 1cm. Especially for Asian people whose body size is relatively thin, new-borns or the elderly, the chest movement displacement is usually within 1cm [22], and this is beyond the detection capability of the FMCW based systems [32].

1.3 Proposed Approach

Due to the fact that the distance estimation resolution of FMCW is restricted by the sweep band width, it is difficult to achieve very high distance estimation resolution (i.e., detecting human respiration with chest movement displacement of less than 1cm.) with narrow-band commodity acoustic device. In this paper, we propose a Correlation based FMCW (C-FMCW) approach to reliably monitor respiration with commodity audio devices. Unlike traditional FMCW, C-FMCW estimates the round-trip propagation time of acoustic signals by discovering the maximum correlation between transmitted signal and received signal. For digital FMCW signal, the round-trip propagation time can be measured by detecting the number of samples corresponding to the maximum correlation. As such, the distance estimation resolution of C-FMCW is limited only by the acoustic signal sampling rate. With current prevailing audio systems of 48KHz, our approach can achieve a ranging resolution of around 0.4 cm. The ranging resolution can be further improved if higher sampling rates (e.g., 96Khz) are supported by audio systems. In C-FMCW, linear chirps facilitate the correlation detection, and also tolerate the frequency selective fading of audio signals due to multipath effect.

Implementing such an audio based respiration system however entails several practical challenges. First, the sampling frequency offset between commodity speakers and microphones can seriously affect the correlation based detection. In practice, the sampling frequency offset squeezes or stretches audio signals such that the received signals and the transmitted signals cannot be aligned during the correlation process. Such an offset if not corrected properly may incur high ranging error. Second, in practice, even though we send a starting command to speaker and microphone simultaneously, the actual start time of the speaker is essentially different from that of the microphone. What's worse, the starting time difference varies every time we send a starting command to them, which will introduce intolerable absolute distance estimation errors. Third, due to multipath effect, weak reflected signals capturing user's respiration can be overwhelmed by stronger signals reflected from the static body parts of the user. It still remains elusive to extract the reflected signals capturing respiration.

We hence propose three novel techniques to address the above challenges. First, we propose an algorithm to compensate dynamically to the transmitted signal and counteract the signal shift caused by sampling frequency offset between speaker and microphone. Second, we co-locate speaker and microphone, thus, part of the transmitted signal will be directly received by the microphone without reflection (self-interference). We regard this part of received signal as a reference to estimate the absolute distance. In this way, we can counteract the starting time difference between speaker and microphone since the starting time difference remains the same for the signal without reflection and the signal reflected from abdomen. Third, based on the fact that respiration is almost periodic in nature, we extract periodic components by using autocorrelation to capture respiration while filtering out static signals.

1.4 Contribution

The main contributions of this paper can be summarized as follows:

1) We propose a novel C-FMCW method for human respiration monitoring, which is able to achieve sub-centimeter ranging accuracy. C-FMCW adopts the correlation based detection method to accurately measure the propagation time of audio signals with high resolution audio samples.

2) We propose several novel ideas to address practical issues, such as developing a compensation algorithm to dynamically add artificial shift to the transmitted signal to counteract the signal shift caused by sampling frequency offset between speaker and microphone; leveraging the received signal without reflection to eliminate the absolute distance estimation error introduced by the starting time difference between speaker and microphone; leveraging the prior knowledge of respiration periodicity to enhance and extract weak periodic

signals which can be otherwise overwhelmed by stronger reflection signals.

3) We conduct extensive experiments to evaluate our system with 22 subjects. The experimental settings vary in different sleep postures, different rooms, and different scenarios. We also verify the system robustness with respect to different sensing distance, different scenarios, different respiration rates, different kinds of simulated apnea, and body movement. The results show that our system achieves a median error of 0.35 breaths/min.

2 Related Work

In this section, we briefly review the existing contactless respiration detection work, which can be roughly grouped into two categories:

2.1 Customized Device based Methods

Recently, customized device based technologies ranging from laser [16], infrared [33], Doppler radar [17], UWB radar [18], FMCW radar [19], USRP [20, 21], ultrasonic sensor [31] have been proposed to detect human respiration. The basic idea is to measure the chest movement displacement during respiration. Although these methods achieve accurate respiration detection, they typically require costly signal transceivers or customized radio frontends. As such, these methods are less appealing than the commodity device based methods for large-scale deployment in home settings.

2.2 Commodity Device based Methods

In recent years, researchers have turned their attentions to commodity devices since they are readily available in home environments. The most widely studied methods are WiFi based and audio based.

2.2.1 WiFi based methods. Ubibreathe [27] measures the received signal strength (RSS) to extract the respiration rate of a person in vicinity. RSS-based methods, however, require the subject in the line-of-sight path. WiSleep [24] and the studies in [22, 23, 25, 26] use the fine-grained channel state information (CSI) in commodity WiFi cards to detect respiration rate. Wang et al. [22] apply the Fresnel zone model to characterize the radio propagation in indoor environments and reveal the principle and limitations of CSI based human respiration sensing with WiFi signals. According to their findings, CSI based approaches cannot always reliably track human respiration and require careful selection of relative distance between human and WiFi devices.

2.2.2 Audio based methods. ApneaApp [32] is a mobile application that allows smartphone users to monitor respiration rate. ApneaApp sends audio chirps and measures frequency shifts to monitor the breathing patterns so as to detect Apnea. The ranging resolution of ApneaApp is fundamentally limited by the narrow sweeping band of the device. As the ranging resolution of ApneaApp is only 0.7 cm, it can detect a deep breath with chest displacement of around 2 cm. Our system offers a more fine-grained ranging resolution of 0.4 cm within the 80cm range. Moreover, the maximum ranging error is less than 0.2 cm when the distance between the acoustic devices and subject is from 40 cm to 60 cm, which can reliably detect a normal breath of young kids or elders who typically have a chest displacement of less than 1 cm. iSleep [34] uses microphone on a smartphone to detect the events that are closely related to sleep quality such as body movement and snore. However, iSleep cannot detect respiration. BeepBeep [35] detects the roundtrip time of flight at the resolution of audio sampling rate and achieves high ranging accuracy between two smartphones. BeepBeep requires both smartphones to actively generate audio chirps, while our work detects the time of flight of reflected audio chirps. RunBuddy [36] measures the breathing pattern of a runner with Bluetooth headset and suggests a proper running rhythm. The headset is not intrusive for a runner but may cause inconvenience to a user during sleep.

In our work, we propose a novel acoustic respiration detection method leveraging commodity acoustic devices, which achieves a ranging resolution of around 0.4 cm. Based on the high ranging estimation resolution, we manage to detect human respiration with the chest movement displacement of less than 1cm.

3 C-FMCW: A HIGH RESOLUTION DISTANCE ESTIMATION METHOD

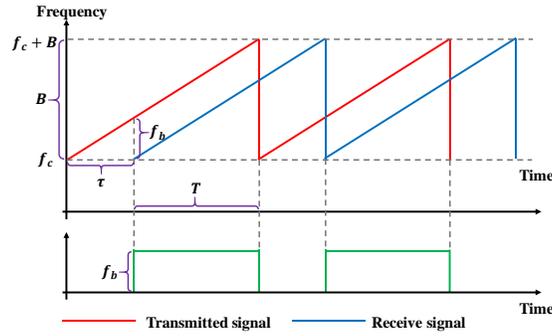


Fig 1. Chirp signal in FMCW

3.1 FMCW and its Limitation

3.1.1 FMCW. The naïve way to estimate distance d to an object using wireless signal is to transmit and receive a sharp pulse and compute the time delay τ , then estimate the distance using their relationship $d = \tau \cdot v$, where v denotes the speed of wireless signal. However, sending a sharp pulse requires a large bandwidth. Instead, FMCW [37] indirectly estimates the propagation delay based on the frequency shift of the chirp signal as follows.

As shown in Fig. 1, the red line denotes the transmitted signal, whose frequency linearly increases over time. The frequency of the transmitted signal at time t is given by: $f(t) = f_c + \frac{Bt}{T}$, where f_c , B and T denote carrier frequency, sweep bandwidth and sweep period, respectively. The phase of the transmitted signal is calculated by integrating $f(t)$ over time:

$$u(t) = \int_0^t f(t') dt' = 2\pi \left(f_c t + B \frac{t^2}{2T} \right)$$

the transmitted signal is then presented as $v_{tx}(t) = \cos(u(t))$. For simplicity, the amplitude is assumed to be 1. The transmitted signal is reflected by the target with the time delay τ . If amplitude attenuation is ignored, the received signal at time t can be represented as: $v_{rx}(t) = \cos(u(t - \tau))$ (the blue line in Fig. 1). Consider the transmitter and receiver are co-located. τ is calculated as:

$$\tau = 2(R + vt)/C \quad (1)$$

where R denotes the distance between transceiver and target, v denotes the moving speed of the target, and C denotes the propagation speed. The receiver multiplies the transmitted signal with the received signal as $v_m = v_{tx} \cdot v_{rx}$. By using $\cos A \cdot \cos B = (\cos(A - B) + \cos(A + B))/2$ and filtering out the high frequency component $\cos(A + B)$, v_m is then simplified as $v_m = \cos\left(2\pi\left(f_c \tau - \frac{B(\tau^2 - 2t\tau)}{2T}\right)\right)$ (the green line in Fig. 1).

Plugging Eq. (1), the frequency of v_m can be represented as: $f_b = \frac{1}{2\pi} \frac{\delta \text{phase}(v_m)}{\delta t} = \frac{2f_c v}{c} + \frac{2BR}{CT} + \frac{4Bvt}{CT} - \frac{4Bv^2 t + 4BRv}{C^2 T}$. The terms with respect to $1/C^2$ is too small and can be ignored, thus, Δf can be then simplified as: $f_b = \frac{2f_c v}{c} + \frac{2BR}{CT} + \frac{4Bvt}{CT}$. For the static or very slowly moving target, i.e., v is close to 0, we can further simplify the frequency of mixed signal, and thus obtain the distance between transceiver and target as follows.

$$R = \frac{CT}{2B} f_b \quad (2)$$

3.1.2 Theoretical Range Resolution Upper Bound of FMCW. The transmitted signal v_{tx} has the

fundamental frequencies at all multiples of the frequency $1/T$ [38, 39]. If ignoring the amplitude attenuation and assuming that the target is static or moving slowly, the received signal v_{rx} is a time shifted version of transmitted signal v_{tx} , so v_{tx} and v_{rx} have the same fundamental frequencies. Thus, $v_m = v_{tx} \cdot v_{rx}$ has the fundamental frequencies at all multiples of the frequency $1/T$, since the periodicity is preserved when two same period signals are multiplied [38]. It means that in the frequency domain, v_m has the spectral points only at $1/T$ Hz. In other words, the frequency resolution of mixed signal, $\delta f_b \geq 1/T$.

Based on Eq. (2), we can infer the theoretical range resolution of FMCW as:

$$\delta R = \frac{CT}{2B} \delta f_b \quad (3)$$

Eq. (3) indicates that the range resolution, i.e., fixed error, δR depends on frequency resolution of the mixed signal δf_b . However, as mentioned above, δf_b is bounded by the chirp frequency $1/T$, i.e., $\delta f_b \geq 1/T$. which means that in order to distinguish targets in two different distances, the frequency of the mixed signal cannot be smaller than the chirp frequency [39]. Thus, the theoretical range resolution upper bound of FMCW is computed as follows:

$$\delta R \geq \frac{CT}{2B} \cdot \frac{1}{T} = \frac{C}{2B} \quad (4)$$

Eq. (4) indicates that the range resolution of linear modulated FMCW depends on sweep bandwidth B . Narrower bandwidth results in lower range resolution.

3.1.3 Highest Range Resolution of FMCW Using Commodity Audio Devices. There are two practical frequency bounds that restrict the available sweeping bandwidth B (as shown in Fig. 1). On one hand, the upper bound of human audible range is about 18KHz. To be inaudible to users, the starting frequency of chirp signal f_c (as shown in Fig. 1) should be set higher than 18KHz. On the other hand, for commodity audio devices, the frequency response starts to decrease rapidly in the band higher than 23KHz. Thus, the available sweeping bandwidth $B = 23\text{KHz} - 18\text{KHz} = 5\text{KHz}$. According to Eq. (4), the highest range resolution of FMCW using commodity audio devices is represented as follows:

$$\delta R \geq \frac{C}{2B} = \frac{343}{2 \times 5000} = 0.0343 \text{ m} = 3.43 \text{ cm}$$

in which 3.43 cm is much larger than the typical chest displacement of 1 cm during breathing. It implies that directly implementing FMCW using commodity audio devices cannot reliably detect human respiration.

3.2 Cross-correlation function based FMCW (C-FMCW)

3.2.1 Basic Idea of C-FMCW. The objective of distance estimation is to measure delay τ . For transmitted signal with a modulation period T , the delay within $T/2$ reflected from target can be measured using the main peak of cross-correlation of transmitted signal and received signal. In this paper, we use FMCW wave as transmitted signal, and exploit cross-correlation to measure the delay. We call this method Cross-correlation function based FMCW (C-FMCW).

The cross-correlation function is defined as:

$$R(n) = \begin{cases} \frac{1}{N-n} \sum_{m=0}^{N-n-1} v_{tx}(m) \cdot v_{rx}(m+n), & n \geq 0 \\ \frac{1}{N-|n|} \sum_{m=0}^{N-|n|-1} v_{rx}(m) \cdot v_{tx}(m+n), & n < 0 \end{cases} \quad (5)$$

where N is the number of samples of transmitted signal in one modulation period T , $n = -N + 1, -N + 2, \dots, N - 1$. $R(n)$ measures the similarity of v_{tx} and n -sample shifted version of v_{rx} . Suppose $R(n)$ reaches its peak $R(\text{Lag})$ with the Lag -sample shift. Since v_{rx} is a time shifted version of v_{tx} with amplitude attenuation, it

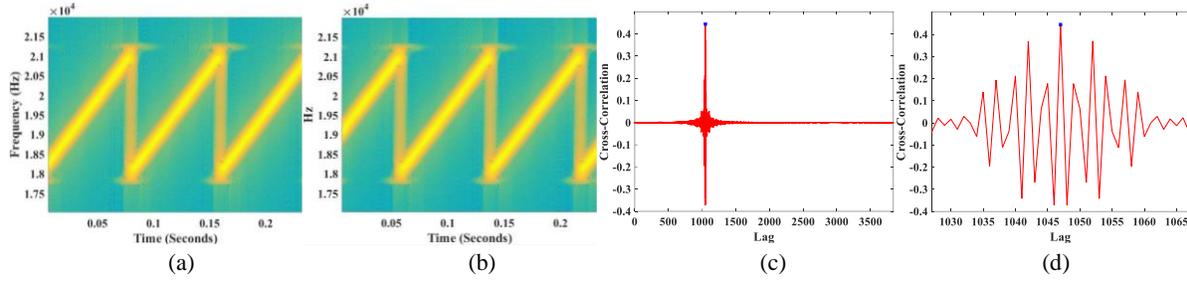


Fig. 2. Cross-correlation function of transmitted and received chirp signals with delay. (a) transmitted signal. (b) received signal with delay. (c) cross-correlation function of transmitted and received signals. (d) enlarged view round the peak of (c)

means that v_{rx} overlaps v_{tx} with the Lag -sample shift. Thus, we can measure the time of flight by counting the number of audio samples. For example, Fig. 2(a) and Fig.2(b) show the transmitted signal and received signal with 3 modulation periods (sampling frequency $F_s = 48$ KHz, $f_c = 18$ KHz, $B = 3$ KHz, $T = 0.08$ s), respectively. The received signal is set to $Lag = 1047$ samples shifted version of transmitted signal. Fig. 2(c) shows the cross-correlation $R(n)$ of transmitted signal and received signal. We can see the peak in $R(n)$. According to the definition of $R(n)$, the peak should be $R(Lag) = R(1047)$. In Fig 2(d), we can see that peak appears indeed at $n = 1047$. Thus, the delay τ can be calculated as:

$$\tau = \frac{Lag}{F_s} \quad (6)$$

where F_s is the sampling frequency. Combining with Eq. (1), we measure the distance R between transceiver and target as:

$$\begin{cases} \tau = Lag/F_s \\ \tau = 2(R + vt)/C \end{cases} \quad \text{and} \quad R = \frac{C \cdot Lag}{2F_s} - vt \quad (7)$$

For the static or very slowly moving target, i.e., v is close to 0, we can simplify R as:

$$R = \frac{C \cdot Lag}{2F_s} \quad (8)$$

3.2.2 Theoretical Range Resolution Upper Bound of C-FMCW. Based on Eq. (8), the theoretical range resolution of C-FMCW can be represented as:

$$\delta R = \frac{C \cdot \delta Lag}{2F_s} \quad (9)$$

According to the definition of cross-correlation function in Eq. (5), the theoretical resolution of Lag is $\delta Lag = 1$, since the cross-correlation is calculated at the resolution of each audio sample. Thus, the theoretical range resolution of C-FMCW can be calculated as:

$$\delta R = \frac{C \cdot \delta Lag}{2F_s} = \frac{C}{2F_s} \quad (10)$$

From Eq. (10) we can see that the range resolution of C-FMCW is limited only by the sampling frequency F_s and independent of sweeping bandwidth B . Higher sampling frequency results in higher ranging resolution. In other words, even if available sweeping bandwidth is narrow, C-FMCW can still achieve very high ranging resolution.

3.2.3 Range Distance Resolution of C-FMCW Using Commodity Audio Devices. Generally, commodity audio systems support 22.05KHz, 44.1KHz and 48KHz sampling frequencies. In order to achieve high ranging resolution, we set the sampling frequency to $F_s = 48$ KHz, and the ranging resolution of C-FMCW is calculated as:

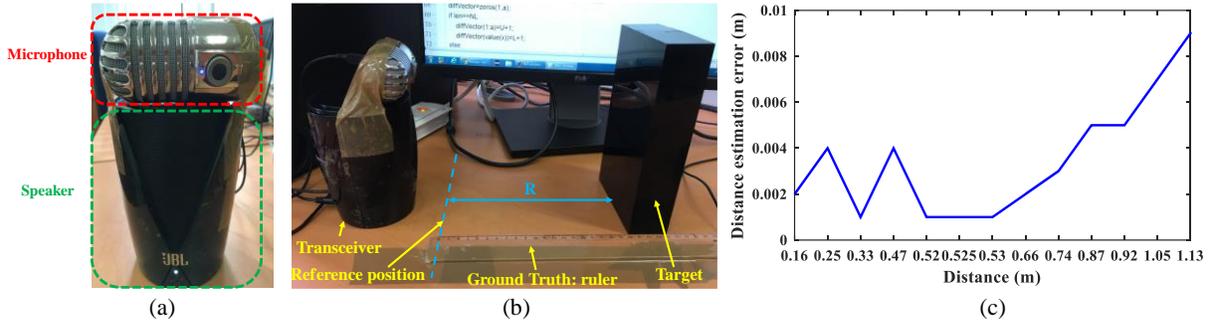


Fig. 3. C-FMCW verification. (a) transceiver, (b) full view of experiment, (c) distance estimation error as ground truth is varied

$$\delta R = \frac{C}{2F_s} = \frac{343}{2 \times 48000} = 0.00357 \text{ m} = 0.357 \text{ cm}$$

Such a ranging resolution suffices to reliably detect human respiration. In fact, by combining appropriate interpolation techniques, the resolution can be further improved. In this work, we have adopted 2x interpolation method to improve the resolution to

$$\delta R = \frac{C}{2 \times 2F_s} = \frac{343}{2 \times 2 \times 48000} = 0.179 \text{ cm}$$

Note that excessive interpolation has a limit in improving the resolution.

3.3 C-FMCW Verification

In this section, we conduct real experiments to verify Eq. (8) and Eq. (10) in C-FMCW. The experimental settings and results are reported as follows.

We bind a general speaker (JBL Jembe, 6 Watt, 80 dB) and microphone (SAMSON MeteorMic, 16 bit, 48 KHz) together (as shown in Fig. 3(a)), they work as a simple acoustic radar (transceiver). The transceiver illuminates target directly (as shown in Fig. 3(b)). The speaker transmits FMCW signal ($F_s = 48 \text{ KHz}$, $f_c = 18 \text{ KHz}$, $B = 3 \text{ KHz}$, $T = 0.02 \text{ s}$) continuously, which is beyond human audibility range. Meanwhile, the microphone receives the signal reflected by the target at the frequency of 48 KHz with 16 bits.

A ruler is placed on the desk as the ground truth. The position, where the reading of the ruler is 0, is regarded as the reference position (highlighted with blue dashed line in Fig. 3(b)). The distance between transceiver and reference position is set to 0. To verify the ranging resolution of C-FMCW, we move the target to 13 different positions ranging from 0.1 m to 1.2 m. The distance estimation results are shown in Fig. 3(c). We observe that within 80 cm, the distance estimated error can be controlled below 0.4 cm. Moreover, the maximum ranging error is less than 0.2 cm when the distance between the acoustic devices and subject is from 40 cm to 60 cm, meaning that chest displacement which is less than 0.5 cm can be detected using C-FMCW. As distance increases, we observe that error also increases due to degradation of reflected acoustic signal. For more details, readers are encouraged to watch the video at this link: <http://dwz.cn/6EE2Dx>

4 CONTACTLESS RESPIRATION DETECTION USING C-FMCW WITH COMMODITY ACOUSTIC DEVICES

In this section, we present the detailed design of C-FMCW based contactless respiration detection system using commodity acoustic devices. We first identify the technical challenges and our solutions. We then give an overview of the system architecture and describe each major function component in detail.

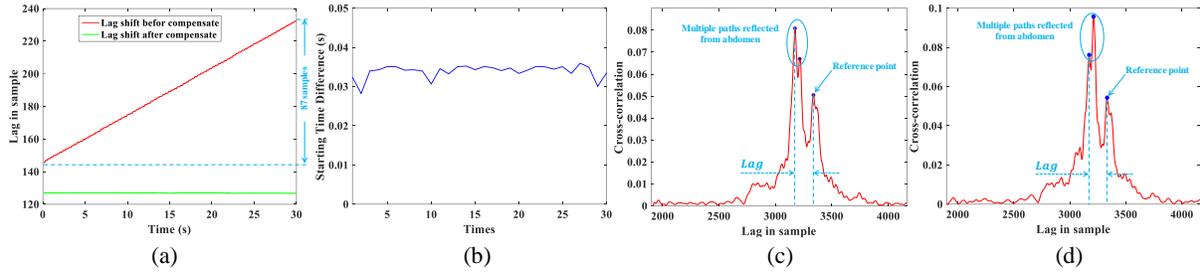


Fig. 4. Practical issues using C-FMCW to detect respiration with commodity acoustic devices. (a) lag shift over time caused by sampling frequency offset between transmitter and receiver. The red line indicates the delay shift over time before correction, while, the blue line indicates the delay shift over time after correction. (b) starting time difference between speaker and microphone. (c) multiple reflection path from abdomen, and (d) cross-correlation function peak amplitude.

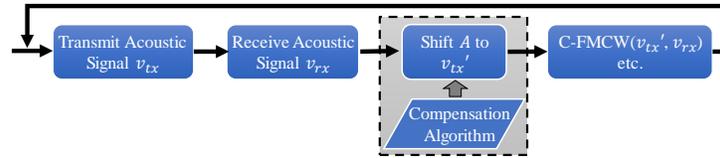


Fig. 5. Solution for sampling frequency offset between transmitter and receiver

4.1 Practical Challenges and Solutions using C-FMCW to Detect Respiration with Commodity Acoustic Devices

Even though C-FMCW is promising to achieve high ranging resolution, a number of practical challenges exist when applied to detect respiration with commodity acoustic devices such as sampling frequency offset between transmitter and receiver, the starting time offset between transmitter and receiver, as well as major reflection path selection problems. In this section, we present these challenges and our solutions.

4.1.1 Challenge 1: Sampling Frequency Offset Between Transmitter and Receiver. Existing audio based ranging approaches typically assume that the sampling frequencies of transmitter and receiver are exactly the same. Unfortunately, due to imperfect clocks, the frequency offset between transmitter and receiver introduces error in calculating *Lag* (in Eq. (8)). Moreover, the error will accumulate over time. In Fig. 4(a), the red line highlights the *Lag* shift caused by the sampling frequency offset between transmitter and receiver. We find that within 30 seconds (about 10 breathes), the *Lag* shift accumulates up to 87 samples, which leads to a distance measurement error of 31.1cm according to Eq. (8). In other words, during each breathing, the ranging error caused by the sampling frequency offset can exceed 3 cm, which is much larger than the chest displacement of 1 cm during breathing. In order to reliably detect respiration, we must eliminate the errors caused by the sampling frequency offset between transmitter and receiver.

4.1.2 Solution 1: Sampling Frequency Offset Calibration Algorithm. As shown in Fig. 5, the basic idea of our solution is to add a shift into transmitted signal to compensate the *Lag* shift caused by sampling frequency offset between transmitter and receiver. Specifically, given the *Lag* shift threshold Td and time duration D , our compensation algorithm will shift audio samples so that the *Lag* shift in duration D is smaller than Td by adapting two parameters compensation rate r and compensation unit $\Delta Shift$. Compensation rate $r = M/N$ ($M < N$ and both of them are positive integers) means that in every N iterations (as shown in Fig. 5), we will add M times shifts into transmitted signal. Compensation unit $\Delta Shift$ means that we add shift in unit $\Delta Shift$ samples. Our compensation algorithm is summarized in ALGORITHM 1. The algorithm firstly finds shift unit $\Delta Shift$ to make the symbols of system shift when $r = 0$ and $r = 1$ are different. Then, the algorithm uses the binary search method to find appropriate compensation rate r that makes system shift during period D smaller

ALGORITHM 1 Compensation Algorithm**Input:** Duration D , and offset threshold OT **Output:** compensation rate r and compensation unit $\Delta Shift$ $r_L \leftarrow 0; r_U \leftarrow 1;$ find the smallest $\Delta Shift$, which makes $getShift(r_L, \Delta Shift, D)$ and $getShift(r_U, \Delta Shift, D)$ with different symbol;% $getShift(r, \Delta Shift, D)$ gets the Lag shift in duration D with the compensation rate r and compensation unit $\Delta Shift$ **while** $|shift| > OT$ **do** % Binary search method to find appropriate r $shift \leftarrow getShift(r \leftarrow (r_L + r_U)/2, \Delta Shift, D);$ **if** $shift < 0$ % update the upper bound or lower bound according to the symbol of current system shift **do** $r_U \leftarrow r;$ **else** $r_L \leftarrow r;$ **end**

than shift threshold OT . Note that during the execution of the ALGORITHM 1, the target should be static. The target movement will introduce shift and influence compensation performance. In practice, we only need to use a static object to get the compensation parameters for once during system calibration before starting respiration detection. The time cost incurred by ALGORITHM 1 depends on the input parameters D and Td . Larger D and smaller Td will result in longer running time but more accurate compensation rate r and compensation unit $\Delta Shift$.

The blue line in Fig. 4(a) highlights the system shift after applying our compensation algorithm. We observe that there is almost no *Lag* shift after applying the compensation. In summary, with this compensation algorithm, the effect caused by sampling frequency offset between transmitter and receiver can be well controlled.

4.1.3 Challenge 2: Starting Time Offset Between Transmitter and Receiver. Ideally, if speaker and microphone start to work simultaneously, we can directly estimate the absolute distance from transceiver to abdomen. Unfortunately, even though we send the starting commands to speaker and microphone simultaneously, the starting time of speaker and microphone can be different in practice. Moreover, the starting time difference varies over time. Fig. 4(b) shows 30 times starting time difference between speaker and microphone in our system. The varied starting time difference will introduce a large error in initial distance estimation. In the example shown in Fig. 4(b) the starting time difference (average 0.034s) will introduce $0.034 \times 48000 = 1632$ samples *Lag* shift. According to C-FMCW distance estimation equation Eq. (8), it will introduce an absolute distance error of $343 \times 1632 / (2 \times 48000) = 5.83$ m.

4.1.4 Solution 2: Cancelling Starting Time Offset Using Self-interference Effect. As shown in Fig. 3(a), in our system, speaker and microphone are co-located. As a result, part of the transmitted signal will be received by the microphone without reflection (i.e. self-interference). This part of received signal will generate a special peak in the cross-correlation function of transmitted signal and received signal (the reference point shown in Fig. 4(c) and Fig. 4(d)). Due to no reflection, this part of received signal will not be affected by the target. Thus, both the amplitude and location of this special peak will keep relatively static. It is easy to use variation of amplitude and location to distinguish this special peak from others. We name the special peak *reference point*. As shown in Fig. 4(c) and Fig. 4(d), with the reference point, we can cancel the starting time offset between transmitter and receiver by using the location difference between reference point and the peak generated by the receive signal reflected from the target. Then, by using C-FMCW distance estimation equation Eq. (8), we can accurately get the distance. This solution can completely avoid the interference from the starting time difference between speaker and microphone, since the *Lag* shift caused by the starting time difference is the same for the reference point and the peak generated by the receive signal reflected from the target. Thus, when we calculate the *Lag* by

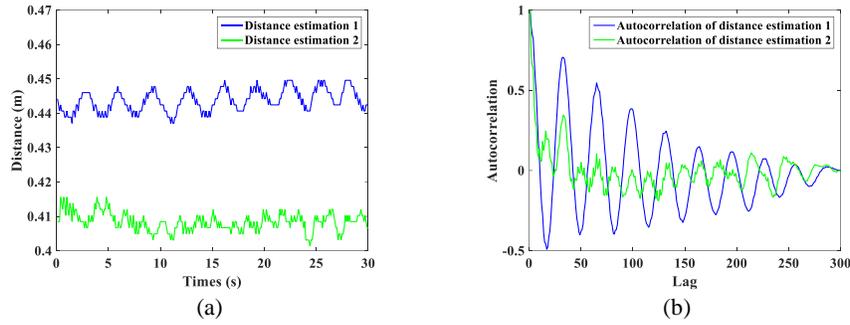


Fig. 6. Autocorrelations of two different distance estimation results. (a) two distance estimation sequences. (b) the autocorrelation function of the two distance estimation sequences in (a)

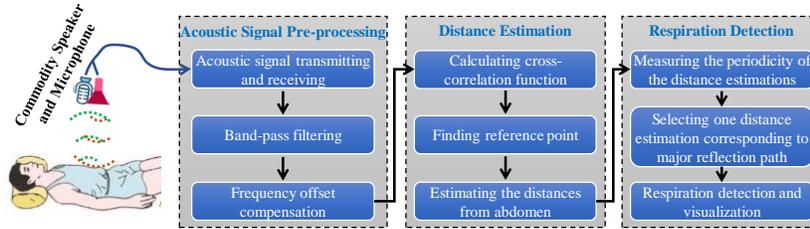


Fig. 7. Overall framework of C-FMCW based contactless respiration detection system using acoustic signal

using the location difference between the reference point and the peak generated by the receive signal reflected from abdomen, the *Lag* shift caused by the starting time difference can be cancelled.

4.1.5 Challenge 3: Multiple Reflection Paths from Different Parts of Body. In practice, there are multiple reflection paths from different parts of abdomen. As shown in both Fig. 4(c) and Fig. 4(d), there are two peaks (in the blue circle) corresponding to two different reflection paths (note that in practice, there may be more than 2 peaks). Different parts of abdomen have different movement displacements during breathing. The reflection paths from static parts cannot be used for respiration detection. Only the reflection paths from the moving parts during breathing can contribute to respiration detection. Moreover, the power of received signal from different reflection path fluctuates dynamically, which will result in amplitude fluctuation of cross-correlation peaks. As shown in Fig. 4(c) and Fig. 4(d), the two peaks in blue circle alternatively become the largest peak. It is challenging to identify the representative peak that can effectively reflect human respiration among the multiple peaks.

4.1.6 Solution 3: Selecting Major Reflection Path Using the Periodicity of Respiration. Even though there may exist multiple reflection paths from different parts of abdomen, we know that human respiration is almost periodical in nature. It means that the distances from the moving parts of the target during breathing will vary periodically, while other distances from static reflection paths are not periodical. Thus, we can select the effective reflection path by measuring the periodicity. The periodicity can be easily measured by using an autocorrelation function. For the sequence $X = x_1 x_2 \dots x_M$, autocorrelation function is defined as follows:

$$R(k) = \frac{E[(x_i - \mu)(x_{i+k} - \mu)]}{\sigma^2}, k = 1, 2, \dots, M - 1 \quad (11)$$

Fig. 6(a) shows two different distance estimations from two different reflection paths over time. Fig. 6(b) shows the autocorrelation functions of the two distance estimations. In Fig. 6(b), we see that the peaks of the autocorrelation function is higher if the distance estimation is more cyclical.

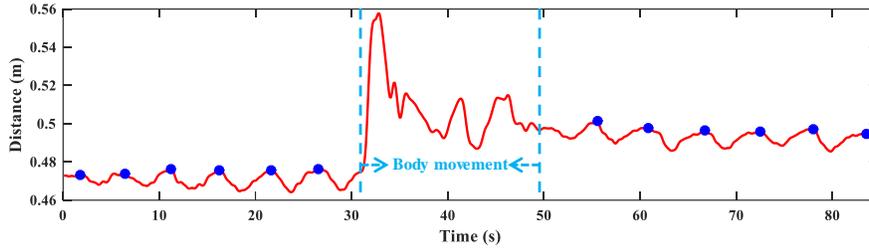


Fig. 8. Respiration identification with body movement

4.2 C-FMCW Based Respiration Detection System Framework

In this section, we present the key workflow of C-FMCW based contactless respiration detection system using acoustic signal. We first give an overview of the system architecture, and then describe each key functional module in detail.

4.2.1 System Architecture. As shown in Fig. 7, our system is composed of three major components, namely acoustic signal pre-processing, distance estimation, and respiration detection. The acoustic signal pre-processing module filters out noise and compensates the *Lag* shift. The distance estimation module estimates the distance of each reflection path from the target. The respiration detection module detects the respiration from the distance estimation results. In the following, we summarize the key functional modules of C-FMCW.

4.2.2 Acoustic Signal Pre-processing. The objectives of this module are to: 1) ensure real-time continuous transmission and reception of acoustic signals; 2) denoise received signals; and 3) compensate the *Lag* shift caused by sampling frequency offset between transmitter and receiver.

In order to filter background noise, we apply a band-pass filter (passband $[f_c, f_c + B]$), where f_c denotes carrier frequency and B denotes sweep band width.

To compensate the *Lag* shift caused by sampling frequency offset between transmitter and receiver, we firstly get the compensation rate r and compensation unit $\Delta Shift$ by using our composed compensation algorithm in ALGORITHM 1. Then, we distribute M times compensation (each time $\Delta Shift$ samples) to transmitted signal (as shown in Fig. 5) during N iterations uniformly to counteract the shift caused by the sampling frequency offset between transmitter and receiver. M and N have the relation $M/N = r$.

4.2.3 Distance Estimation. The key function of this module is to estimate the distance of each reflection path from abdomen using our proposed C-FMCW. Specifically, we first calculate the cross-correlation function of compensated transmitted signal and received signal using Eq. (5), and find all the peaks of cross-correlation function. Then, we identify the reference point by finding the peak whose variation of the amplitude and location are smallest. After the reference point is located, we estimate the absolute distances (here the distance means half of the length of reflection path) of all reflection path by using the solution presented in Section 4.2.2. We calculate *Lag* by using the location difference between reference point and the peak corresponding to current reflection path. Then, we estimate the distance by using C-FMCW distance estimation equation Eq. (8). The distance estimation result of each reflection path over time is saved as a distance sequence. These distance sequences of all the reflection paths will be regarded as the input of next module.

4.2.4 Respiration Detection. The major work of this module is to detect respiration based on the distance estimation results. As mentioned in Section 4.1.3, not all the distance estimation results can be used for respiration detection. We first calculate the periodicity of all the distance sequences by using autocorrelation function. Then, we select the distance sequence corresponding to the major reflection path which can reflect the respiration well if the periodicities of all the distance sequences are lower than a given threshold. Thereby, the interference factors (e.g., body movement, sleep posture change) can be detected. Fig. 8 shows the respiration detection result while the user changes sleep posture. The system will stop identifying breathing when the user changes sleep posture, since during this period the distance estimation fluctuation cannot reliably capture the chest movement. The system will recover to identify breathing once at least one distance sequence's periodicity

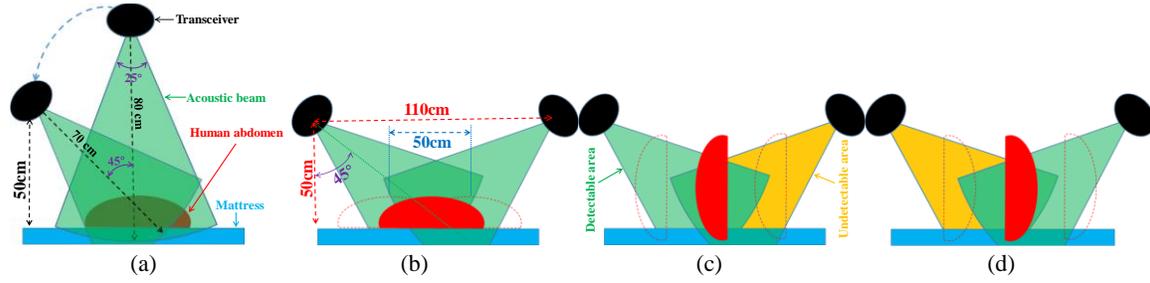


Fig. 9. The detectable area of transceiver and its layout facing in different sleep postures. (a) the range, covering area and max slant angle of the transceiver. (b) ~ (d) the device layout and detectable area facing in different sleep postures.

becomes higher than a threshold.

5 EVALUATION

In this section, we conduct comprehensive experiments to evaluate the proposed system. First, we describe our experiment settings. We then briefly describe the baseline approach. We conduct comprehensive experiments to evaluate our system in four different rooms with 22 subjects, who have three different sleep postures and compare the performance with that of the baseline approach. We also evaluate the system robustness with different scenarios, different sensing distance, different respiration rates, different kinds of simulated apnea, and body movement. Finally, we discuss the limitations of the system.

5.1 Experiment Setting and Set-ups

To set up our respiration monitoring system, we use two laptops (Thinkpad T450 with Intel Core i5-5200 CPU, 8G RAM; Dell Latitude E6540 with Intel Core i7-4800MQ, 4GB RAM) and connect them with a pair of commodity speaker (JBL Jembe, 6 Watt, 80 dB) and microphone (SAMSON MeteorMic, 16 bit, 48 KHz) which forms an audio transceiver in our system, as shown in Fig 3(a). In practice, we use two transceivers to increase the coverage of monitoring area. The proposed respiration detection algorithm implemented in Matlab keeps monitoring user's respiration in real-time. The speakers are programmed to transmit FMCW signal continuously. Meanwhile, the microphones receive the echo with 48 KHz sampling rate and send the received audio signal to the connected laptop for data processing and respiration detection. In order to avoid frequency band interference, the speaker in the first transceiver is programmed to transmit FMCW signal with the parameters: $f_c = 18\text{KHz}$, $B = 2.3\text{KHz}$, $T = 0.02\text{ second}$, $D = 0.1\text{ second}$, $F_s = 48\text{KHz}$, while the speaker in the second transceiver transmits FMCW signal with the parameters: $f_c = 20.7\text{KHz}$, $B = 2.3\text{KHz}$, $T = 0.02\text{ second}$, $D = 0.1\text{ second}$, $F_s = 48\text{KHz}$. f_c , B and T denote the carrier frequency, sweep band width and sweep period, respectively. D is FMCW signal length of time, i.e., the system processes D second signal for one time. In other words, our system processes $D/T = 5$ sweep periods FMCW signal in one system iteration. F_s is sampling frequency of speakers. The microphones in two transceivers receive the reflected signal with 48KHz and filter the received signal with band-pass filters. In particular, the pass bands of the two transceivers are set to [18KHz, 20.3KHz] and [20.7KHz, 23KHz], respectively.

Due to clothes absorption and attenuation in the air, the received signal reflected from abdomen is very weak. Thus, we place the transceiver toward subject's abdomen. However, one transceiver will fail to detect respiration if the subject changes sleep posture. We notice that subjects may change their sleeping posture from lying on one's back to lying on right side or left side. For system configuration, we firstly detect the effective range, covering area and max slant angle of the transceiver. To this end, we ask each subject to lie on a bed and place a transceiver above the subject's abdomen for respiration detection. As shown in Fig. 9(a), the acoustic beam width is about 25° and the effective detection distance is about 80cm, while the effective detection



Fig. 10. System settings in test rooms. The red circles mark the transceivers.

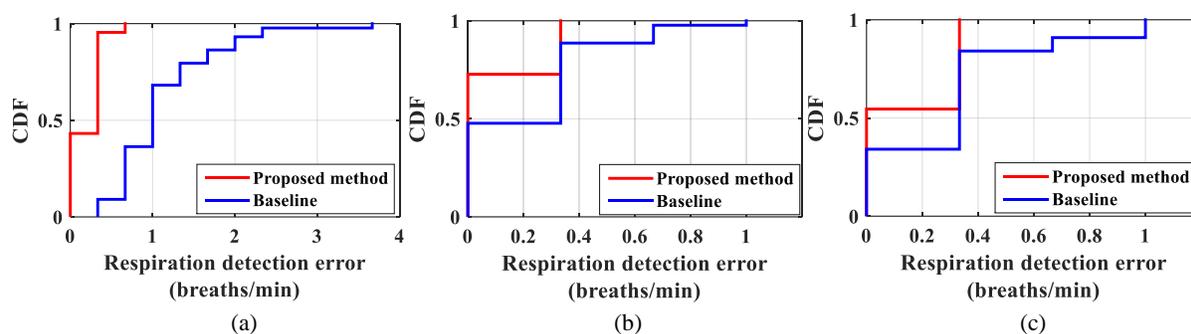


Fig. 11. Respiration detection results of 22 subjects with three different sleep postures. (a) sleep on one's back, (b) facing left, (c) facing right

distance is about 70cm when the slant angle is 45° . In practice, if the slant angle is set as larger than 45° , the transceiver cannot receive the effective reflection signal.

Based on the above results, we deploy two transceivers as shown in Fig. 9(b) -Fig. 9(d). The transceivers are placed at the upper-left and upper-right of the subject's abdomen, respectively with the angle about 45° pointing to the user's abdomen. The perpendicular distance between mattress and device is about 50 cm and the distance between two transceivers is about 110 cm. We can see that, theoretically this setup covers all the facings of a subject, no matter where the subject faces during sleep.

5.2 Baseline Approach

We compare C-FMCW with two representative audio based approaches [31, 32]. The theoretical ranging resolution of App [32] is around 0.7 cm which can be used to detect respiration with chest movement displacement of about 2 cm. Different from the subjects selected in [32] who are mainly western adults, people in Asia, young kids and elders have relatively mild chest movement during normal breathing, i.e. with displacement of less than 1 cm [22], which is beyond the detection capability of the ApneaApp system. Based on this fact and extensive verification experiments, we find that [31] is the only existing work that can detect human respiration with small chest movement using acoustic signal. Instead of directly monitoring the chest movement, the work [31] measures the airflow of respiration and monitor respiration. Specifically, the velocity of exhaled airflow will introduce the Doppler effect to the reflected acoustic signal. Thus, we choose it as the baseline to verify whether our system is also able to detect respiration accurately.

Yet, the work [31] needs a customized device to generate 40 KHz acoustic signals. As the off-the-shelf speaker can only generate acoustic signals with maximum frequency of 23 KHz, their proposed algorithm

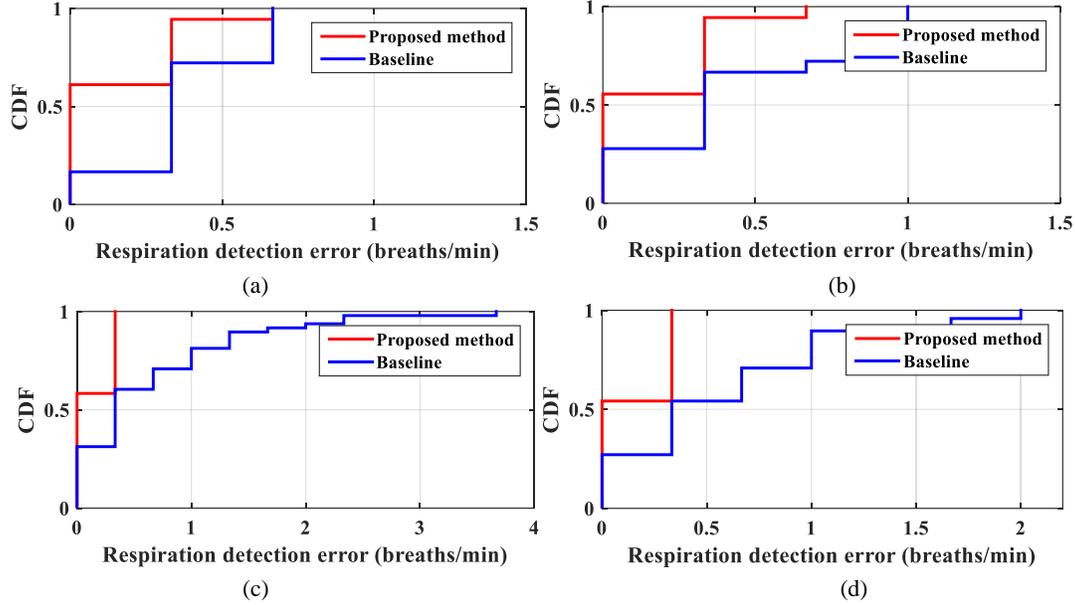


Fig. 12 Respiration detection error in four different rooms.

cannot be applied directly. We thus implement it with a 23 KHz acoustic signal, setting the key parameters accordingly to make the system work well. The key difference between two systems is that the transceiver faces different part of human body. Specifically, the system in [31] monitors the area around nose, while our system monitors the area around chest and abdomen.

5.3 System Performance Evaluation

We conduct comprehensive experiments to evaluate our system in four different rooms with 22 subjects, who have three different sleep postures. We compare the performance with baseline approaches in four different practical scenarios. We also evaluate the system robustness with different scenarios, different sensing distance, different respiration rates, different kinds of simulated apnea, and body movement.

5.3.1 Evaluation with Different Subjects. We recruit 22 participants (7 elders, 2 young kids and 13 adults, including 10 females and 12 males) to evaluate the effectiveness of our system. The participants are asked to set up the system and properly adjust facing direction of audio transceivers according to the requirements specified in Section 5.1. The experiment environment is shown in Fig. 10. We set aside 15 minutes 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 posture for about 200 seconds. During the measurements, two subjects watch the video stream to record the ground truth manually. Fig. 11(a) shows the respiration detection error CDF. We can see that the median respiration detection error of our system is lower than 0.35 breaths/min. In most cases ($p > 0.95$), our system achieves an error lower than 0.33 breaths/min. The median error of the baseline approach is around 1 breaths/min. In addition, the max error of our system is 0.6 breaths/min, while the maximum error of the baseline reaches 3.7 breaths/min. Note that, based on our measurements, even the highest and fattest subject (male, height 1.85m, weight 95 kg), his chest movement displacement is about 1cm when he breathes naturally. This experimental result indicates: 1) the proposed C-FMCW achieves very high distance estimation resolution; 2) our system is able to accurately detect human respiration when the participants sleep on his/her back; 3) the system is easy to set up in practice.

5.3.2 Evaluation with Different Sleep Postures. Except for sleeping on one's back, lying on one side is also

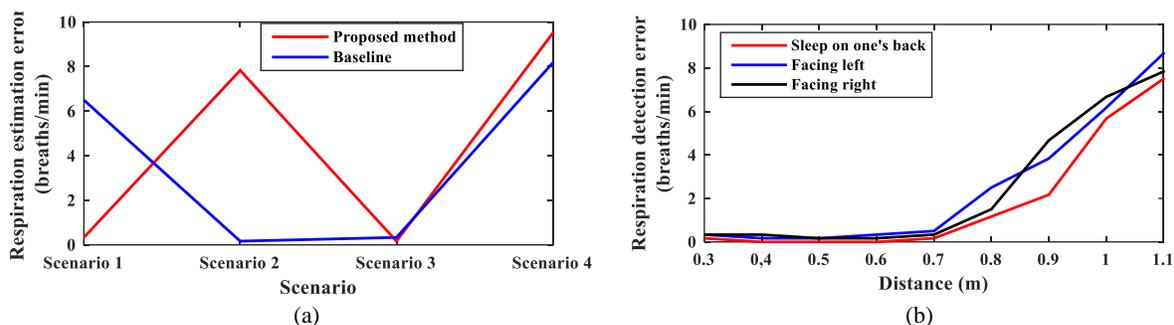


Fig. 13 System robustness to (a) different test scenarios and (b) different sensing distance

one common sleep posture. With the same experiment settings, all the subjects are recruited to evaluate the performance of our system when subjects lying on one side. Fig. 11(b) and Fig. 11(c) show the respiration detection error CDF when subjects sleeping on the left and right side, respectively. The result shows that the median error of our system is 0 breaths/min, while the median error of the baseline is 0.35 breaths/min. In addition, the max error of our system can be controlled smaller than 0.35 breaths/min, while the maximum error of the baseline is 1 breaths/min.

The experimental results indicate that our system is able to accurately detect human respiration for all three common sleep postures in real environment and achieve better performance.

In addition, we observe an increased error of the baseline when a user lies on his/her back. With this sleeping posture, the angle between the acoustic signal and the exhaled airflow is closed to 90° , which will result in relatively small effective velocity component of exhaled airflow. Thus, the Doppler effect will decrease and result in larger detection error.

5.3.3 Evaluation in Different Environments. We conduct the experiments in four different rooms with different size and layout. Fig.12(a) ~ (d) show the average respiration detection error CDF in the four test rooms, respectively. The median detection errors of our system for all three sleep postures in four test rooms are 0 breaths/min, while the median errors of the baseline are 0.35 breaths/min. 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. The real demo video is provided at link below: <http://dwz.cn/6EDZnK>

5.3.4 Evaluation with Different Scenarios. In order to test the applicability in real bed rooms, we recruit two subjects to evaluate our system and compare with the baseline in the following four practical scenarios.

- 1) A subject wears thin cloth or is covered by thin quilt with windows opened or air condition turned on (let breeze blow into the room, e.g., in summer scenario)
- 2) A subject wears thick cloth or is covered with quilt (e.g., in winter scenario)
- 3) A subject wears thin cloth or is covered by thin quilt with windows closed and air condition turned off (e.g., in spring and autumn scenarios)
- 4) A subject sleeps with quilt covering his/her head or sleep prostrate.

Fig. 13 (a) shows the respiration detection error in the above four scenarios. We can see that in the first scenario, our system works well and the error is smaller than 0.5 breaths/min, while the baseline fails to detect respiration. This is due to the fact that the baseline is to detect respiration by sensing exhaled airflow which can be easily disturbed by ambient airflow such as wind or even people passing by. In contrast, our system still works since our distance measurement between transceiver and the chest of subject will not be affected by airflow. In the second scenario, the baseline works well while our system fails to detect respiration. This is due to the fact that in this scenario, the chest movement displacement during breathing will be buried since the displacement of thick cloth or quilt is very small. The sensing target of baseline is the airflow around nose, so it will not be affected by thick cloth or quilt. In the third scenario, both systems work well, since there is no interference for both systems. In the fourth scenario, both systems fail to detect respiration, since the exhaled

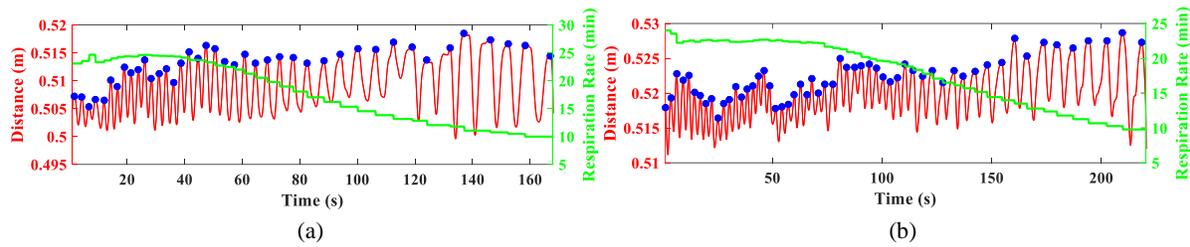


Fig. 14. Respiration detection results of two participants while respiration rate is changing over time.

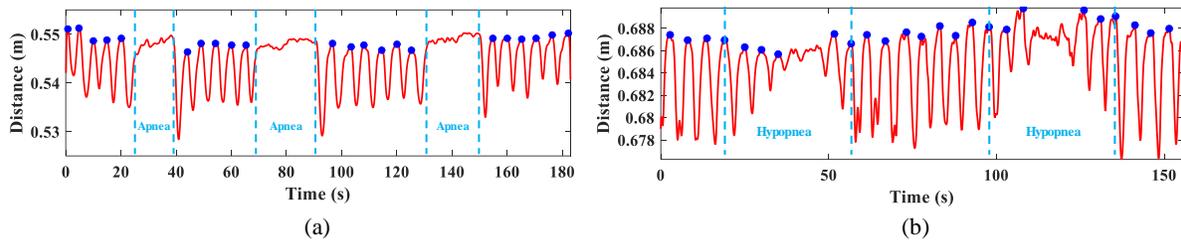


Fig. 15. Partial apnea identification results of two subjects.

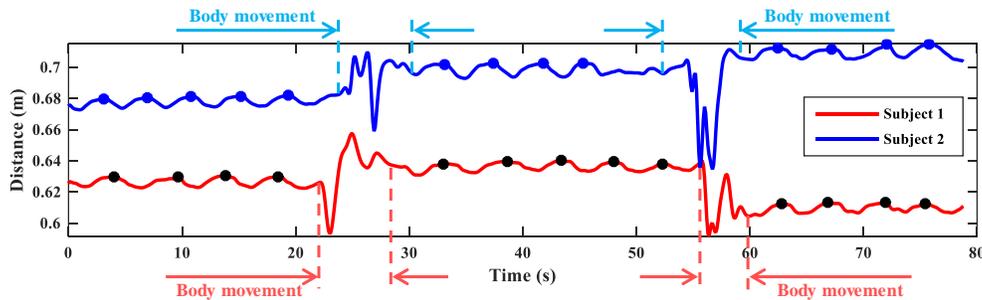


Fig. 16. Respiration detection with body movement

airflow around the nose are blocked and displacement reflected on her/his back is too small to be detected directly.

5.3.5 Evaluation with Different Sensing Distance. We vary the distance between the transceiver and the body of subject ranging from 0.3 m to 1.1 m with an interval of 0.1 m. Two subjects are recruited to evaluate the performance of our system. At each position, we test for 6 minutes. Fig. 13 (b) shows the average respiration detect error as distance is varied when subjects sleeping with three different postures. We can see that within 0.7 m, the system achieves respiration detection error smaller than 0.5 breaths/min. Beyond 0.7 m, the error will increase with the distance between the transceiver and the subject mainly because of the signal attenuation. The device with higher output power and higher frequency response in effective frequency band (18KHz~23KHz) will result in larger effective sensing range.

5.3.6 Evaluation with Respiration Rate Change. As our system detects the breathing rate by measuring the periodicity of distance variation, respiration rate change will weaken the periodicity and may influence our system. In this work, we use a sliding window method which can discard the obsolete signals and adapt to the change of respiration rate.

Two participants are asked to do high-intensity exercises like push-ups, 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. 14 shows the respiration detection results of the two participants after high-intensity exercises. The red lines show the measured displacements and the green lines track the instantaneous respiration rates. We can observe that the respiration rates of two participants change from about 20 breathing counts per minute to about 10 breathing counts per minute. Our system provides sufficient ranging resolutions to accurately track both rapid shallow breathing and slow deep breathing throughout the process.

5.3.7 Detection of Apnea. Detecting apnea is an important objective of monitoring respiration during sleep. When apnea happens, the periodicity of distance variation will be absent. Benefitting from the usage of sliding window to detect the periodicity of distance variation, after the subject recovers to breathe normally, the system will recover to detect breathing.

Limited by legal constraints, we could not test our system with real apnea patients in hospitals for now. Instead, we simulate Central Apnea (CA) and Obstructive Apnea (OA) following the clinical symptom described as “hold breath for a while” [32, 40], and simulate Hypopnea event following the clinical symptom described as “breathing becomes shallow” [32]. We recruit two participants (one male and one female) to test apnea detection performance. We repeat the experiment for 10 times each of which lasts for 2 to 3 minutes. Fig. 15 (a) shows that our system accurately detects CA and OA. Fig. 15(b) shows that our system also accurately detects the Hypopnea event.

5.3.8 Evaluation with body movement. Body movement (e.g., sleep posture change) generates strong but un-rhythmical distance displacement variation. The displacement caused by chest movement would be submerged. Under this condition, it is difficult to detect breathing. Our system is designed to suspend respiration detection once body movement is detected. Two subjects are recruited to test whether our system can actually suspend when a body movement occurs. We detect subjects’ respiration for about 15 minutes. During respiration detection process, the subjects change sleep postures or move limbs several times. Fig. 16 shows that our system suspends respiration detection while body movement occurs and recovers for detecting respiration after the movement stops.

5.4 Discussion

5.4.1 Limitations. Even though the experimental results demonstrate that the proposed system can detect human respiration with three common sleep postures in real rooms and can adapt to different real scenarios, different sensing distances, changing respiration rate, different kinds of simulated apnea and body movement, the current implementation has the following two limitations. First, the transceivers must face the abdomen directly to ensure that the transceivers can receive reflected acoustic signals. As mentioned in Section 5.3.4, wearing thick cloth or covering with quilt and sleeping prostrate may influence accuracy. This is due to the fact that when sleeping on one’s chest, the displacement reflected on her/his back is too small to be detected. Second, so far, we have not tested our system with subjects suffering sleep apnea yet due to users’ privacy and legal issues. We note that the apnea detection study is based on simulated apnea, where volunteers try their best to simulate the apnea events according to the clinical symptoms. Yet, we believe the experiments can at least allow us to evaluate whether our system can detect some types of apnea events that would lead to breath stopping or other irregular breathing patterns.

5.4.2 Potential applications. C-FMCW achieves 0.4 cm ranging resolution within 80 cm using commodity audio devices. The ranging resolution can be further improved if higher sampling rates (e.g., 96Khz) are supported by audio systems. Beneficial from the high ranging resolution, C-FMCW can be potentially applied to detect other types of subtle movements like heartbeat, finger gestures and even mechanical oscillations if the sampling rate is further increased.

6 CONCLUSION

In this work, we study contactless human respiration detection using acoustic signal. By exploring the traditional

FMCW principle, we proposed a novel correlation based distance estimation method C-FMCW. Based on C-FMCW, we develop a contactless respiration detection system using commodity audio devices. We develop and integrate several elaborated acoustic signal processing methods to address some practical challenges. We conduct comprehensive experiments to evaluate the proposed respiration detection system. We implement the proposed C-FMCW with commodity audio devices and demonstrate that our system can achieve a ranging resolution of less than 0.4 cm. We evaluate the performance of respiration monitoring in four different rooms with 22 subjects who have three different sleep postures. Experimental results show that our system accurately detects subjects' respiration with the median error lower than 0.35 breaths/min in different experiment settings.

ACKNOWLEDGMENTS

The authors would like to express their special appreciation to the 22 volunteers for participating in our experiments. The study was approved by the Medical Ethics Committee (MEC) of NPU. This work was supported in part by a grant from National Natural Science Foundation of China (No. 61332013), Chinese Scholarship Council Program, National Natural Science Foundation of China (No. 61702437), Hong Kong ECS under Grant PolyU 252053/15E, and Hong Kong PolyU under Grant G-YBMT.

REFERENCES

- [1] Adib Fadel, Hongzi Mao, Zachary Kabelac, Dina Katabi, and Robert C Miller. 2015. Smart homes that monitor breathing and heart rate. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 837 – 846.
- [2] Healthcare. 2012. F. M. Market for embedded health monitoring-gadgets to hit 170M devices by 2017. <http://www.fiercemobilehealthcare.com/story/market-embedded-health-monitoring-gadgets-hit-170m-devices-2017/2012-08-03>. (2012).
- [3] Michelle A Cretikos, Rinaldo Bellomo, Ken Hillman, Jack Chen, Simon Finfer and Arthas Flabouris. 2008. Respiratory rate: the neglected vital sign. *Medical Journal of Australia* 188, 11 (2008): 657.
- [4] J. Brian North and Sheila Jennett. Abnormal breathing patterns associated with acute brain damage. 1974. *Archives of neurology* 31, 5 (1974), 338-344.
- [5] Yuan, George, Nicole A. Drost, and R. Andrew McIvor. 2013. Respiratory rate and breathing pattern. *McMaster University Medical Journal* 10, 1 (2013), 23-25.
- [6] Cooke, Jana R., and Sonia Ancoli-Israel. Normal and abnormal sleep in the elderly. 2011. *Handbook of clinical neurology*/edited by PJ Vinken and GW Bruyn 98 (2011), 653.
- [7] Norman, Daniel, and José S. Loredó. Obstructive sleep apnea in older adults. 2008. *Clinics in geriatric medicine* 24.1 (2008), 151-165.
- [8] Lee-Chiong, L. Teofilo. 2003. Monitoring respiration during sleep. *Clinics in chest medicine* 24, 2 (2003), 297-306.
- [9] R. Madeline, MPH. Vann. 2015. The 15 Most Common Health Concerns for Seniors. URL: <http://goo.gl/EQn2fi>, 2015
- [10] J. N. Wilkinson, and V. U. Thanawala. 2009. Thoracic impedance monitoring of respiratory rate during sedation – is it safe?. *Anaesthesia*, 64 (2009), 455–456.
- [11] Jaffe, B. Michael. 2008. Infrared measurement of carbon dioxide in the human breath: “breathe-through” devices from Tyndall to the present day. *Anesthesia & Analgesia* 107, 3 (2008), 890-904.
- [12] Rita Paradiso. 2003. Wearable health care system for vital signs monitoring. In *Information Technology Applications in Biomedicine, 2003. 4th International IEEE EMBS Special Topic Conference on*. IEEE, 283–286.
- [13] Shoko Nukaya, Toshihiro Shino, Yosuke Kurihara, Kajiro Watanabe, Hiroshi Tanaka. 2012. Noninvasive bed sensing of human biosignals via piezoceramic devices sandwiched between the floor and bed. *IEEE Sensors journal* 12, 3 (2012), 431-438.
- [14] Hulya Gokalp and Malcolm Clarke. 2013. Monitoring activities of daily living of the elderly and the potential for its use in telecare and telehealth: a review. *TELEMEDICINE and e-HEALTH* 19, 12 (2013), 910–923.
- [15] Jochen Penne, Christian Schaller, Joachim Hornegger, and Torsten Kuwert. 2008. Robust real-time 3D respiratory motion detection using time-of-flight cameras. *International Journal of Computer Assisted Radiology and Surgery* 3, 5 (2008), 427–431.
- [16] T Kondo, T Uhlig, P Pemberton, and PD Sly. 1997. Laser monitoring of chest wall displacement. *European Respiratory Journal* 10, 8 (1997), 1865–1869.
- [17] M Nowogrodzki, DD Mawhinney, and HF Milgazo. 1984. Non-invasive microwave instruments for the measurement of respiration and heart rates. *NAECON 1984* (1984), 958–960.
- [18] Svetha Venkatesh, Christopher R Anderson, Natalia V Rivera, and R Michael Buehrer. 2005. Implementation and analysis of respiration-rate estimation using impulse-based UWB. In *Military Communications Conference, 2005. MILCOM 2005. IEEE*. IEEE, 3314–3320.
- [19] Fadel Adib, Hongzi Mao, Zachary Kabelac, Dina Katabi, and Robert C Miller. 2015. Smart homes that monitor breathing and heart rate. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 837–846.
- [20] Ruth Ravichandran, Elliot Saba, Ke-Yu Chen, Mayank Goel, Sidhant Gupta, and Shwetak N Patel. 2015. WiBreathe: Estimating respiration rate using wireless signals in natural settings in the home. In *Pervasive Computing and Communications (PerCom), 2015 IEEE International Conference on*. IEEE, 131–139.
- [21] Fadel Adib, Zachary Kabelac, Dina Katabi, Robert C. Miller. 2013. 3D Tracking via Body Radio Reflections. In *Proceedings of the*

- 11th USENIX Conference on Networked Systems Design and Implementation*. USENIX Association, 317-329.
- [22] Hao Wang, Daqing Zhang, Junyi Ma, Yasha Wang, Yuxiang Wang, Dan WU, Tao Gu, Bing Xie. 2016. Human respiration detection with commodity wifi devices: do user location and body orientation matter?. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 25-36.
- [23] Jian Liu, Yan Wang, Yingying Chen, Jie Yang, Xu Chen, and Jerry Cheng. 2015. Tracking Vital Signs During Sleep Leveraging Off-the-shelf WiFi. In *Proceedings of the 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing*. ACM, 267–276.
- [24] Xuefeng Liu, Jiannong Cao, Shaojie Tang, and Jiaqi Wen. 2014. Wi-Sleep: Contactless sleep monitoring via WiFi signals. In *Real-Time Systems Symposium (RTSS), 2014 IEEE*. IEEE, 346–355.
- [25] Xuefeng Liu, Jiannong Cao, Shaojie Tang, Jiaqi Wen, and Peng Guo. 2016a. Contactless Respiration Monitoring via WiFi Signals. *Mobile Computing, IEEE Transactions on* (2016).
- [26] Chenshu Wu, Zheng Yang, Zimu Zhou, Xuefeng Liu, Yunhao Liu, and Jiannong Cao. 2015. Non-Invasive Detection of Moving and Stationary Human With WiFi. *Selected Areas in Communications, IEEE Journal on* 33, 11 (2015), 2329–2342.
- [27] Heba Abdelnasser, Khaled A Harras, and Moustafa Youssef. 2015. Ubibreathe: A ubiquitous non-invasive wifi-based breathing estimator. *arXiv preprint arXiv:1505.02388* (2015).
- [28] Ossi Kaltiokallio, Huseyin Yigitler, Riku Jantti, and Neal Patwari. 2014. Non-invasive respiration rate monitoring using a single COTS TX-RX pair. In *Information Processing in Sensor Networks, IPSN-14 Proceedings of the 13th International Symposium on*. IEEE, 59–69.
- [29] Neal Patwari, Lara Brewer, Quinn Tate, Ossi Kaltiokallio, and Maurizio Bocca. 2014a. Breathfinding: A wireless network that monitors and locates breathing in a home. *Selected Topics in Signal Processing, IEEE Journal of* 8, 1 (2014), 30–42.
- [30] Neal Patwari, James Wilson, Sundaram Ananthanarayanan, Sneha Kumar Kasera, and Dwayne R Westenskow. 2014b. Monitoring breathing via signal strength in wireless networks. *Mobile Computing, IEEE Transactions on* 13, 8 (2014), 1774–1786.
- [31] Philippe Arlotto, Michel Grimaldi, Roomila Naeck and Jean-Marc Ginoux. 2014. An ultrasonic contactless sensor for breathing monitoring. *Sensors* 14.8 (2014), 15371-86.
- [32] Rajalakshmi Nandakumar, Shyamnath Gollakota, Nathaniel Watson M.D.. 2015. Contactless Sleep Apnea Detection on Smartphones. In *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, 45-57.
- [33] Carina Barbosa Pereira, Xinchu Yu, Michael Czaplak, Rolf Rossaint, Vladimir Blazek, and Steffen Leonhardt. 2015. Remote monitoring of breathing dynamics using infrared thermography. *Biomedical optics express*, 6, 11 (2015), 4378-4394.
- [34] Hao Tian, Guoliang Xing, and Gang Zhou. 2013. iSleep: unobtrusive sleep quality monitoring using smartphones. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*. ACM, 1-14.
- [35] Chunyi Peng, Guobin Shen, Yongguang Zhang, Yanlin Li, Kun Tan (2007). BeepBeep: a high accuracy acoustic ranging system using COTS mobile devices. In *Proceedings of the 5th international conference on Embedded networked sensor systems*. ACM, 1-14
- [36] Tian Hao, Guoliang Xing, Gang Zhou. 2015. RunBuddy: A Smartphone system for running rhythm monitoring. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 133-144
- [37] A. G. Stove. (1992). Linear FMCW radar techniques. *Radar & Signal Processing Iee Proceedings F*, 139, 5(1992):343-350.
- [38] Oppenheim A V, Willsky A S, Nawab S H. 1996. *Signals & systems (2nd ed.)* Prentice-Hall, Inc. 1996.
- [39] S. Suleymanov. (2016). *Design and Implementation of an FMCW Radar Signal Processing Module for Automotive Applications* (Master's thesis, University of Twente).
- [40] WebMD. <http://www.webmd.com/sleep-disorders/guide/sleep-disorders-symptoms-types>.

Received: May 2017; revised: September 2017; accepted: October 2017.