# Respiration Tracking for People Counting and Recognition

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

***Abstract*—Wireless detection of respiration rates is crucial for many applications. Most of the state-of-the-art solutions estimate breathing rates with the prior knowledge of crowd numbers as well as assuming the distinct breathing rates of different users, which is neither natural nor realistic. However, few of them can leverage the estimated breathing rates to recognize human subjects (also known as identity matching). In this article, using the channel state information (CSI) of a single pair of commercial WiFi devices, a novel system is proposed to continuously track the breathing rates of multiple persons without such impractical assumptions. The proposed solution includes an adaptive subcarrier combination method that boosts the signal-to-noise ratio (SNR) of breathing signals, and iterative dynamic programming and a trace concatenating algorithm that continuously tracks the breathing rates of multiple users. By leveraging both the spectrum and time diversity of the CSI, our system can correctly extract the breathing rate traces even if some of them merge together for a short time period. Furthermore, by utilizing the breathing traces obtained, our system can do people counting and recognition simultaneously. Extensive experiments are conducted in two environments (an on-campus lab and a car). The results show that 86% of average accuracy can be achieved for people counting up to four people for both cases. For 97.9% out of all the testing cases, the absolute error of crowd number estimates is within 1. The system achieves an average accuracy of 85.78% for people recognition in a smart home case.**

***Index Terms*—Crowd counting, identity matching, multipeople breathing estimation, people recognition, wireless sensing.**

## I. Introduction

HUMAN-CENTRIC sensing via wireless radio frequency (RF) has attracted an increasing interest for a range of Internet-of-Things (IoT) applications [1], [2]. Demands of accurate and passive awareness of the environment surge for many applications [3], [4]. For instance, a smart home can adjust the light and ventilation system based on the occupancy level to improve energy efficiency [5], [6]. People recognition in smart homes enables user authentication for home security and privacy protection [7]–[11]. Besides understanding the environment, monitoring the status of the human in the environment also has received great attention, and respiration/breathing rate serving as a significant vital sign has been an important topic for RF sensing.

Comparing with conventional methods that use dedicated sensors to monitor breathing rates, RF sensing provides contact-free solutions. The RF-based solutions can be classified into two categories, i.e., radar-based methods and WiFi-based methods. For radar-based methods, previous works have shown the potential of using millimeter wave (mmWave) [12], [13] or ultrawideband (UWB) [14], [15] to monitor respiration rate as well as the heart rate. Although these systems can achieve high accuracy, dedicated devices hinder their deployment. WiFi-based methods, on the contrary, can be easily deployed and have been studied in the past decade [16]–[21]. Received signal strength (RSS) measured by a WiFi device has been used in [16] to measure the chest movement during breathing. However, the accuracy of respiration rate estimation degrades when the test subjects do not hold the device. Fine-grained channel state information (CSI) is more sensitive to the environment changes, which has been utilized to capture minute movements caused by respiration in [17]–[21]. However, due to the omnidirectional propagation and narrow bandwidth of commonly used 2.4/5-GHz WiFi, the received signal can be reflected from multiple humans in an indoor space. This makes it difficult to extract the vital signs of multiple humans from the reflected signal. Most of the previous works assume that there is only one person in the observation area [17] or assume the respiration rates of different people are distinct and the number of people is known in advance [18]–[21].

In this article, we propose a solution to continuously track the human respiration rate without any prior knowledge of the crowd number or assuming that the breathing rates of different users are distinct. Different from the previous works, we are particularly interested in matching the breathing rates estimated in different time instances to different users, i.e., which breathing rate corresponds to which person. By utilizing the estimated breathing rate traces, our system can achieve people counting as well as recognition at the same time. Achieving identity matching based on multiperson breathing rates for both purposes, however, entails several unique challenges.

First, the subtle changes caused by human breathing could be easily undermined by measurement noises in WiFi CSI. The situation becomes even worse in multiperson scenarios. To overcome the problem, we leverage the diversity residing in subcarriers (SCs) and multiple links (antenna pairs) to reduce

Manuscript received December 16, 2019; revised February 19, 2020; accepted February 22, 2020. Date of publication February 28, 2020; date of current version June 12, 2020. *(Corresponding author: Fengyu Wang.)*

The authors are with the Department of Electrical and Computer Engineering, University of Maryland at College Park, College Park, MD 20742 USA, and also with the Department of Research and Development, Origin Wireless, Inc., Greenbelt, MD 20770 USA (e-mail: fywang@umd.edu; fzhang15@umd.edu; cswu@umd.edu; bebewang@umd.edu; kjrliu@umd.edu).

Digital Object Identifier 10.1109/JIOT.2020.2977254

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

Authorized licensed use limited to: University of Maryland College Park. Downloaded on November 30,2020 at 19:35:07 UTC from IEEE Xplore. Restrictions apply.

noises while boosting the breathing signals. We show that, by properly combining different SCs and links, the breathing signals can be considerably better enhanced than what could be achieved by any individual or the average of them.

Second, a person's breathing rate varies over time making it nontrivial to associate the successive breathing rate estimates to the corresponding persons. Considering that one's breathing rate will not fluctuate within a short time period, we build a Markov chain model for the natural breathing signals and further employ an iterative dynamic programming (IDP) algorithm to continuously track multiple breathing traces (i.e., sequences of breathing rates of different users).

Third, the number of users may not be fixed for an area of interest, since users might come and go. In order to output real-time estimates of the occupancy level, we propose a trace concatenating algorithm to maintain the traces of existing users, which concatenates the latest estimates with the presented traces to determine existing users, newly arriving users, and leaving users.

Finally, although the human respiration rate varies in a small range, the breathing pattern of different individuals is distinct. Based on this observation, we build a database of breathing traces and apply hypothesis testing to do human recognition for the smart home scenario.

We prototype our system using a pair of commodity off-the-shelf (COTS) WiFi devices and conduct experiments in two typical targeted environments (i.e., a lab office and a car), where 12 users are involved in the evaluation. The average accuracy of people counting up to four people is more than 86% for both environments. For 97.9% out of all the testing cases, the absolute error of crowd number estimates is within 1. Note that our system achieves similar performance for both environments, demonstrating the independence on environment, which outperforms the training-based method. Finally, we investigate the performance of people recognition using breathing traces for a 4-people family. The system can achieve 85.78% accuracy. We believe the proposed approach makes an important step toward smart home applications in terms of breathing rate estimation, passive people counting, and human recognition.

The main contributions of this article are as follows.

1) We devise a pipeline of novel techniques to estimate the breathing traces of multiple users, including an SC combination method to enhance multiple breathing signals, an IDP and trace concatenating algorithms to continuously track successive breathing rates of multiple users using the CSI of COTS WiFi devices.
2) We propose a nonintrusive training-free method to estimate the crowd number using the extracted breathing traces. Furthermore, a hypothesis-testing model is proposed to leverage the breathing traces to do human recognition.
3) We prototype and evaluate the proposed system on COTS WiFi devices. The results in indoor spaces and in a car demonstrate promising performance.

The remainder of this article is organized as follows. We review the related works in Section II. We overview the system in Section III, followed by multiperson breathing spectrum generation in Section IV and breathing rates tracking in Section V. The cases of people counting and recognition are studied in Section VI and the performance is evaluated in Section VII. This article is concluded in Section VIII.

## II. Related Works

The design of our system is closely related to the following categories of research.

*RF-Based Breathing Detection:* Respiration is an important vital indicator of health status and medical diagnosis. Conventional methods require a user to wear dedicated sensors, which are intrusive and costly [22]. RF-based monitoring solutions can estimate breathing rates contactlessly since chest motions can be captured by RF signals. mmWave [13] and UWB [14] radars can be used to estimate breathing as well as the heart rate. Although these systems can achieve high accuracy, dedicated systems hinder their wide deployment in practice. Some works have shown the viability of detecting breathing using commodity wireless devices [16]–[21]. Fine-grained CSI is utilized to capture minute movements caused by respiration [18]–[21]. However, most of these works assume that the number of people in the observation area is known in advance and do not address the occupancy level. The most relevant work [21] shows the potential of crowd counting by utilizing the affinity propagation on the estimated breathing rates, however, the accuracy of people counting is not analyzed.

Unlike [19] and [23], in this article, we do not assume the prior knowledge of the number of people and also can handle the case when the number of people changes.

*Wireless-Based Passive People Counting:* The occupancy level plays an important role in many modern applications. For example, a smart building can understand its utilization and perform adaptive crowd control based on the occupancy level. Previous works on wireless nonintrusive human detection [24]–[27] can determine whether the observation area is empty or not, but cannot determine the exact number of people in the target area. Since human activities can influence the reflection path of wireless signals, RF signals can be further utilized to determine the occupancy level. Compared with the device-based approaches [28]–[31], passive (or device-free) estimation algorithms do not require people's active participation and are easier to be deployed. Some device-free approaches have been proposed recently based on WiFi devices. The training-based method is used in [32]–[35] for people counting. However, most of them need a labor-intensive training phase and the prediction model is highly dependent on the environment. A few other works focus on the training-free method. For example, a renewal-type process model is built in [6] to estimate the occupancy level corresponding to the dip events of the received signals. The relationship between the number of moving people and the variation of CSI is studied in [36]. However, all of these works are based on the assumption that all human subjects would keep moving all the time, which is impractical, especially for the indoor scenario.

Different from previous works, we are particularly interested in estimating the occupancy level in quasistatic scenarios, such as attendees in a meeting or staffs in an office. Detecting and counting static users, however, is nontrivial.

Authorized licensed use limited to: University of Maryland College Park. Downloaded on November 30,2020 at 19:35:07 UTC from IEEE Xplore. Restrictions apply.

*RF-Based Human Recognition:* User identification has become a vital issue due to the increasing concerns about privacy. Traditional authentication systems, such as RFID-based systems [37] or video-based systems [38] require dedicated devices, which is not convenient for daily use in smart homes. With the ubiquity of WiFi infrastructure, the WiFi-based authentication system has been studied. Leveraging gait information (i.e., natural walking style) as unique human biometrics, indoor human identification systems were built [8]–[11]. However, the aforementioned systems require users to walk along the predefined paths, which is labor intensive. The existence of human radio biometrics has been shown in [39] and a system using the time-reversal technique has been built. However, their performance would degrade significantly due to environmental changes. In this article, we consider using the breathing rate traces to do human recognition based on a hypothesis-testing model. The proposed human recognition system, which is a byproduct of the respiration rate tracking system, has been proved to work well for a smart home scenario.

## III. System Overview

The computational pipeline underlying the proposed system is shown in Fig. 1. Different from many previous works aiming at estimating independent breathing rates at certain time instances, this article focuses on utilizing the frequency- as well as time-domain information to do identity matching. The core idea is to estimate the breathing rate sequences along the time (also known as breathing rate traces) of different individuals. Furthermore, utilizing the estimated breathing traces, we can estimate the occupancy level as well as predict the user identity in the observation area. This idea immediately leads to three stages of the proposed system: 1) multiuser breathing spectrum generation; 2) breathing rate trace tracking; and 3) people counting and recognition.

In the first stage, the proposed system first performs a short-term Fourier transform (STFT) on CSI measurements to extract the periodic breathing signals. As long as the breathing rates of different individuals are different, multiple frequency components would be observed in the frequency response. The extracted breathing signals are typically fairly weak on a single SC, which are further boosted by a novel adaptive SC combining method. Stage 1 finally outputs a spectrogram of the estimated breathing rates over time.

In Stage 2, the goal is to track the breathing rate traces (i.e., breathing sources) from the spectragram obtained from Stage 1. However, there is a significant gap between breathing rates to breathing rate traces because of two reasons. First, different individuals may have the same breathing rates that overlap with each other. Second, one's breathing rate varies over time. To address the challenges, a Markov chain model is introduced to handle dynamics in natural breathing. We propose a successive cancellation scheme that resolves each individual's breathing trace one by one by a novel algorithm of IDP. Thereafter, we concatenate the identified traces of breathing rates in adjacent time windows to further identify the arriving and leaving time of human subjects.

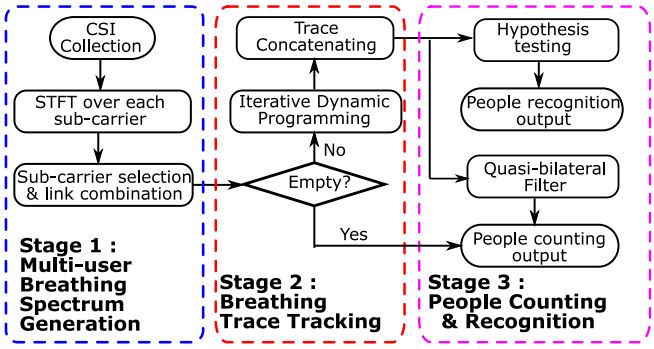



Fig. 1. Processing flow of the system.

In Stage 3, we leverage the estimated breathing rate traces given by Stage 2 to do people counting and recognition. By further utilizing the time-domain information and removing outliers of the estimated breathing rate traces by a novel quasi-bilateral filter, the system gives an estimate of the crowd number. Then by hypothesis testing, we can do people identity recognition according to the extracted breathing rate traces.

## IV. Multiuser Breathing Spectrum Generation

### A. CSI Model With Breathing Impact

The CSI depicts how radio signals propagate from a transmitter (Tx) to a receiver (Rx). In the presence of human beings, one or more paths of signal propagation will be altered due to human motion. Given a pair of Tx and Rx with multiple omni-directional antennas, the CSI of link $m$ at time $t$ and frequency $f_k$ is modeled as

$$h^{(m)}(t,f_k)=\sum_{l=1}^{L}a_l(t)\exp\left(-j2\pi f_k\frac{d_l(t)}{c}\right)+n(t,f_k) \tag{1}$$

where $k\in\mathcal{V}$ is the SC index with center frequency $f_k$ in the set of usable SCs $\mathcal{V}$. $L$ is the total number of multipath components (MPCs), while $a_l(t)$ and $d_l(t)$ denote the complex gain and propagation length of MPC $l$. $n(t,f_k)$ is the additive white noise, and $c$ is the speed of light.

In the presence of breathing, (1) can be rewritten as

$$h^{(m)}(t,f_k)=\sum_{i\in\mathcal{I}}\sum_{l\in\Omega_{d_i}(t)}a_l(t)\exp\left(-j2\pi f_k\frac{d_l(t)}{c}\right)+\sum_{l\in\Omega_s}a_l\exp\left(-j2\pi f\frac{d_l}{c}\right)+n(t,f_k) \tag{2}$$

where $\mathcal{I}$ denotes the set of human subjects. $\Omega_{d_i}$ denotes the MPCs scattered by human being $i$, resulting in time-variant complex gain and delay, $\Omega_s$ denotes the MPCs that are not affected by people's breathing, whose complex gain and delay keep time invariant. The gain of MPCs in $\Omega_{d_i}$ could be modeled as [40]

$$a_l(t)=a_l\times\left(1+\frac{\Delta d_l}{d_l}\sin\theta\sin\left(\frac{2\pi t}{T_{b_i}}+\phi\right)\right)^{-\Psi} \tag{3}$$

where $a_l$ and $d_l$ are gain and path length in a static environment, $\Delta d_l$ is the difference of propagation length caused by chest movement, $\theta$ is the angle between human scatter and the

Authorized licensed use limited to: University of Maryland College Park. Downloaded on November 30,2020 at 19:35:07 UTC from IEEE Xplore. Restrictions apply.

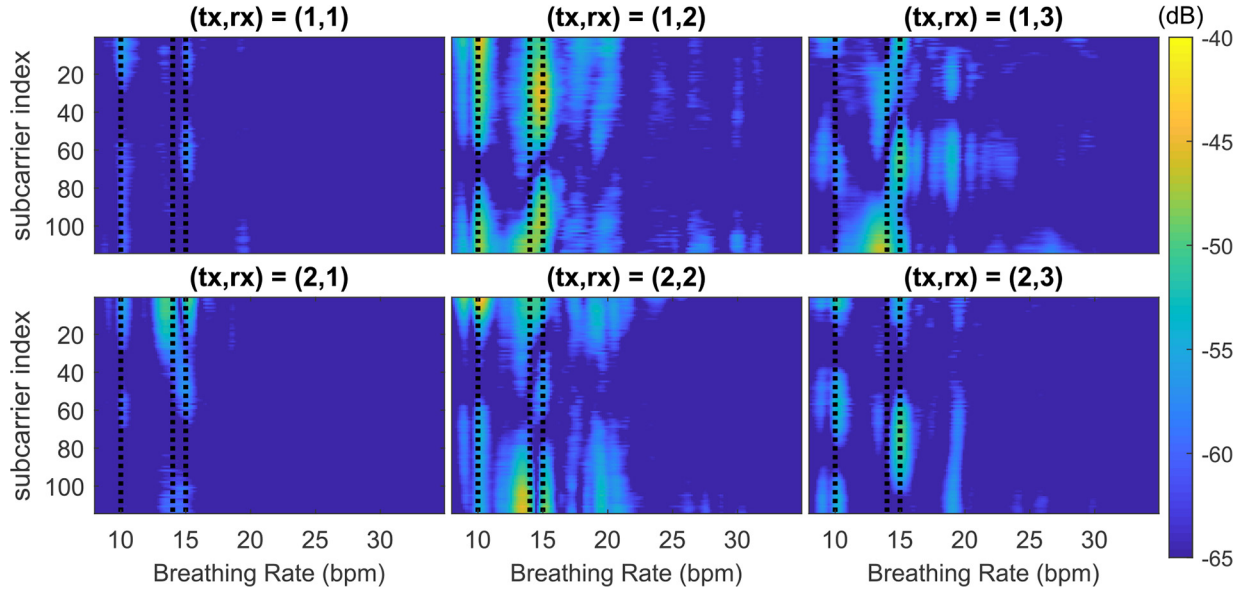



Fig. 2. PSD of different links. Ground truth: three people sitting in car with the breathing rate [10 14 15] bpm (marked with dashed lines).

EM wave, and $\phi$ is the initial phase. $\Psi$ is the path-loss exponent. Since the chest movement is much smaller than the path length, i.e., $\Delta d_l \ll d_l$, the amplitude of MPC in both $\Omega_s$ and $\Omega_{d_i}$ can be assumed to be time invariant, e.g., $a_l(t) \approx a_l$, and the set of MPCs are assumed to be the same, i.e., $\Omega_{d_i}(t) \approx \Omega_{d_i}$.

### B. Boosting SNR of Breathing Signals

It is noted that for each MPC subset $\Omega_{d_i}$, the delay is periodic due to the periodic chest movement, i.e., $d_l(t + T_{b_i}) = d_l(t) \quad \forall l \in \Omega_{d_i}$. Hence, we would be able to see multiple frequency components of the measured CSI in the frequency domain, each corresponding to a distinct breathing signal.

The breathing signals can be extracted by applying STFT to the CSI measurement. In specific, we first apply a sliding window of length $W$ to the captured CSI time series of each SC in every link, and then obtain the frequency spectrum by performing fast Fourier transform (FFT) over each time window. We then employ a band-pass filter on the spectrum to consider only the normal range of human breathing frequencies $[b_{\min}, b_{\max}]$. The FFT is performed on every SC to obtain the individual spectrum for all the $N_{\text{Tx}} \times N_{\text{Rx}} \times N_{\text{sc}}$ SCs, where $N_{\text{Tx}}$, $N_{\text{Rx}}$, and $N_{\text{sc}}$ are the number of Tx antennas, Rx antennas, and usable SCs on each Tx–Rx link, respectively.

As shown in Fig. 2, each breathing signal from one person contributes to one evident peak in the obtained power spectrum density (PSD). Note that different SCs experience diverse sensitivity levels to the identical breathing motion. Previous approaches attempt to select a set of best SCs based on variance, amplitude, or ensemble average of CSI among all SCs to improve the signal-to-noise ratio (SNR). However, the following observations show the flaws of these approaches.

1) The response power of different SCs to the same breathing source is different (see columns in Fig. 2).
2) For the same SC, the response power to different breathing sources is different (see rows in Fig. 2).
3) The response power of different links is distinct (different figures in Fig. 2).

Therefore, there is no single SC that is universally sensitive to all breathing sources. Using the same subset of SCs for different frequency components may not produce equally good SNR for all breathing signals. Furthermore, using a universal threshold for all links may lose information from links with low-response power.

Inspired by these observations, we first propose a novel adaptive SC combining criteria to boost the SNR of breathing signal of each link. For link $m$, the selected SCs for a given frequency component $q$ satisfy the condition that

$$E_k^{(m)}(q) \geq \alpha \max_{q \in Q, i \in \mathcal{V}} \left\{ E^{(m)}(q, i) \right\} \quad \forall k \in \mathcal{V} \tag{4}$$

where $Q$ is the set of frequency components in the range of $[b_{\min}, b_{\max}]$. $E_k^{(m)}(q)$ denotes the power of the $k$th SC over link $m$ at frequency component $q$ and $\max_{q \in Q, i \in \mathcal{V}} \{E^{(m)}(q, i)\}$ denotes the maximum power of link $m$ over all frequency components and SCs. $\alpha$ is a hyperparameter which determines a relative threshold $th^{(m)} = \alpha \max_{q \in Q, i \in \mathcal{V}} \{E^{(m)}(q, i)\}$ for SC selection. Note that $th^{(m)}$ is adaptive to individual link quality, as inspired by the third observation above. Thus, the enhanced power of frequency component $q$ in link $m$ is

$$E^{(m)}(q) \leftarrow \sum_{k \in \mathcal{V}} E_k^{(m)}(q) \mathbb{1}\left( E_k^{(m)}(q) \geq th^{(m)} \right). \tag{5}$$

To further incorporate diverse link quality, we normalize the power for each link and then combine them together to further improve the SNR

$$E^{(m)}(q) \leftarrow \frac{E^{(m)}(q)}{\sum_{i \in Q} E^{(m)}(i)} \quad \forall q \in Q \tag{6}$$

$$E(q) \leftarrow \sum_{m=1}^{M} E^{(m)}(q) \quad \forall q \in Q \tag{7}$$

where $E(q)$ is the power of frequency component $q$ after link combination and $M = N_{\text{Tx}} \times N_{\text{Rx}}$ is the total number of links.

Authorized licensed use limited to: University of Maryland College Park. Downloaded on November 30,2020 at 19:35:07 UTC from IEEE Xplore. Restrictions apply.

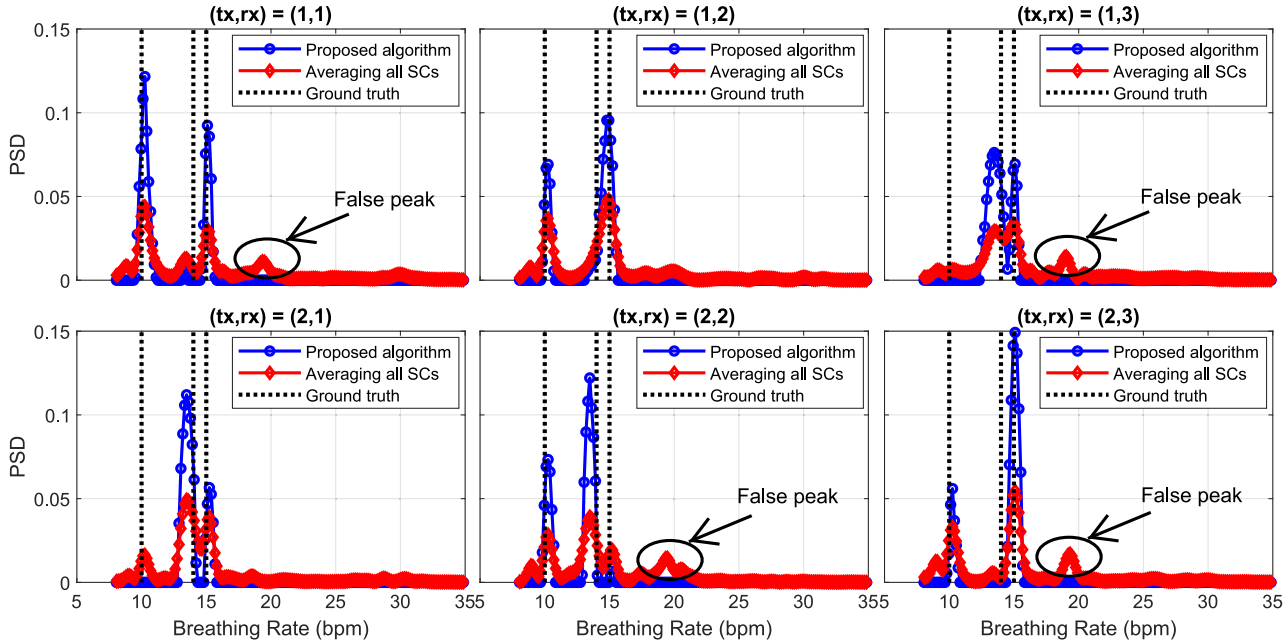



Fig. 3. Normalized PSD of different links after SC combination.

Fig. 3 shows the effect of SC combination for several exemplary links. As seen, the proposed SC selection and combination scheme (shown in blue curves) remarkably improve the SNR for the frequency components of interests, outperforming the simple average scheme (shown in red curves). Fig. 4 further depicts the PSD after the combination of all nine links, which demonstrates that noise and interference have been effectively suppressed. The ground truth of breathing rates are marked with the black dashed lines. As a comparison, a simple average of all SCs suffers from less dominant peaks for the desired breathing signals and false peaks.

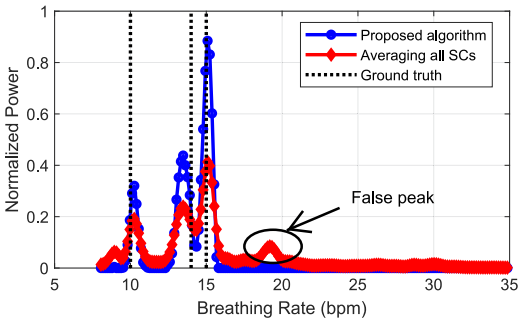



Fig. 4. PSD after link combination.

## V. Breathing Rate Trace Tracking

### A. From Breathing Rates to Breathing Rate Traces

Previous works estimate the number of people by the number of candidate breathing rates [21]. However, they have several limitations. First, the breathing rate estimation may not be accurate enough for a single time instance. Second, different users may have close breathing rates that are indistinguishable from the frequency spectrum, resulting in potential underestimation. Third, the number of people could vary over time as people may come and go. The accompanying motion will also corrupt the breathing signals.

To map the breathing rates to individuals and thus further estimate the accurate crowd number, we utilize the diversity in the time series of the breathing rate estimates for reliable estimation. We first model the dynamic breathing rates as a Markov process. Noting that the breathing signals are periodic where breathing frequency can smoothly change over time, the variation of breathing rate between two adjacent time bins is assumed to follow a normal distribution $\mathcal{N}(0, \sigma^2)$, with the probability density function (PDF) $p(f)$. Since the operation of STFT automatically discretizes the continuous frequency in the range of $[b_{\min}, b_{\max}]$ into $|Q|$ frequency components, where $|Q|$ means the cardinality of set $Q$, the natural breath can be modeled as a Markov chain, and the transition probability matrix is denoted as $\mathbf{P} \in \mathbb{R}^{|Q|} \times \mathbb{R}^{|Q|}$, which is defined as

$$\begin{aligned}\mathbf{P}(q, q') &= \mathbf{P}\big(g(i) = q' | g(i-1) = q\big) \\ &= \int_{\left(q'-q-\frac{1}{2}\right)*\Delta f}^{\left(q'-q+\frac{1}{2}\right)*\Delta f} p(f)\mathbf{d}f \end{aligned} \tag{8}$$

where $\forall q, q' \in Q$ and $g$ is a mapping indicating the frequency component of the breathing rate at given time slots.

To estimate the breathing rate trace in a given time slot $t$, our system leverages the spectrum in $[t-W, t]$, where $W$ is the window length. An output is produced every $W_s$ seconds, and the spectrum is updated at the same time. Thus, to estimate the breathing traces at time $t$, a spectrum $S \in \mathbb{R}_+^I \times \mathbb{R}_+^{|Q|}$ is leveraged, where $I = (W/W_s)$.

In principle, the breathing signal is more periodic than noise and other motion interference. Thus, it is more likely to be observed as peaks in most of the time, and thus the breathing signal will form a trace in the given spectrum along the time with the frequency changing slightly, as shown in Fig. 5. In the following, we first extract the traces of successive breathing

Authorized licensed use limited to: University of Maryland College Park. Downloaded on November 30,2020 at 19:35:07 UTC from IEEE Xplore. Restrictions apply.

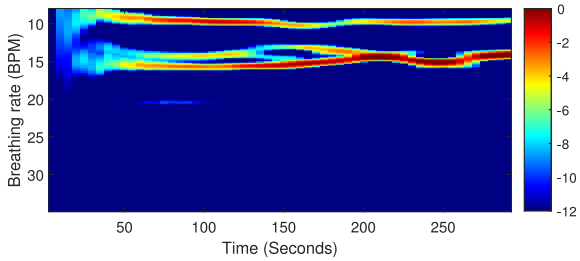



Fig. 5. Spectrogram after link combination.

rates in the given window, and then concatenate them over time.

### B. Extracting Breathing Rate Traces

*1) Theoretical Model:* For a given spectrum $S$, a reasonable estimate of the breathing trace can be obtained by

$$\mathbf{g}^* = \arg\max_{\mathbf{g}} \; E(\mathbf{g}) \tag{9}$$

where $\mathbf{g}$ indicates the breathing trace, denoted as

$$\mathbf{g} = (g(n), n)_{n=1}^{I}. \tag{10}$$

Here, $g : [1, I] \rightarrow Q$ is a mapping indicating the frequency component of the trace at the given time. $E(\mathbf{g})$ is the power of a trace, defined as

$$E(\mathbf{g}) = \sum_{i=i}^{I} S(i, g(i)) \tag{11}$$

where $S(i, j)$ denotes the power at time bin $i$ and frequency component $j$.

Considering that one's breathing rate will not fluctuate a lot within a short period, a regularization term is added to penalize sudden changes in frequencies of interests. A breathing trace is then a series of the breathing rate estimates that achieve a good balance between frequency power and temporal smoothness. The smoothness of a trace can be evaluated by a cost function $C(\mathbf{g})$, defined as

$$\mathbf{C}(\mathbf{g}) \triangleq -\log \mathbf{P}(g(1)) - \sum_{i=2}^{I} \log \mathbf{P}(g(i-1), g(i)) \tag{12}$$

where the frequency transition probability $\mathbf{P}(g(i-1), g(i))$ can be calculated by (8). Without loss of generality, we assume a uniform prior distribution, i.e., $\mathbf{P}(g(1)) = (1/|Q|)$. The cost function $C(\mathbf{g})$ is the negative of the log likelihood for a given trace. The smoother a trace is, the larger its probability is, and the smaller the cost it incurs.

The most probable breathing trace can be found by solving

$$\mathbf{g}^* = \arg\max_{\mathbf{g}} \; E(\mathbf{g}) - \lambda C(\mathbf{g}) \tag{13}$$

where $\lambda$ is a regularization factor. Here, we denote $E(\mathbf{g}) - \lambda C(\mathbf{g})$ as the regularized energy of trace $\mathbf{g}$. By properly choosing the hyperparameter $\lambda$, the system can ensure that the regularized energy of a true breathing trace is positive, while when the observation area is empty, the regularized energy for any trace candidate in the given spectrum is negative.

*2) Iterative Dynamic Programming:* The problem in (13) can be solved by dynamic programming. However, dynamic programming typically can only find the trace with the maximum regularized energy and cannot deal with multiple breathing traces. We propose a successive cancellation scheme to find multiple traces one by one via a novel method of IDP.

The principle idea of the IDP is intuitive. For a given spectrum $S$, the most probable breathing trace is first found by dynamic programming. To further determine if there are any other breathing traces, the identified trace will be erased from the spectrum, and then a new round of dynamic programming is performed to find another candidate trace. This successive cancellation procedure will be run iteratively until there is no more effective traces in the spectrum.

For clarity, $(i, q)$ denotes the bin index with timestamp $i$ and frequency component $q$. We want to find the best trace of frequency peaks from timestamp $i$ to $j$, which is denoted as $\mathbf{g}_i \rightsquigarrow \mathbf{g}_j$. Define the regularized energy of trace $\mathbf{g}_i \rightsquigarrow \mathbf{g}_j$ that ends at point $(j, n)$ as $s(\mathbf{g}_i \rightsquigarrow (j, n))$. Our approach is to search all possible traces $\mathbf{g}_i \rightsquigarrow (j, n)$ that end at frequency point $n$ and select the best one among them. This can be achieved by finding the optimal traces for all the bins along with the adjacent timestamps. For simplicity, we denote the regularized energy at each bin as its score given by

$$s(i, q) = S(i, q) + \max_{\forall q' \in Q} \left\{ s\big(i-1, q'\big) + \lambda \log \mathbf{P}\big(q', q\big) \right\} \quad i = 2, 3, \ldots, I \quad \forall\, q, q' \in Q \tag{14}$$

where $s(1, q) = S(1, q) + \lambda \log \mathbf{P}(g(1) = q)$. The score of a given bin is the maximum achievable regularized energy that it can obtain. In other words, it determines the optimal paths that pass through bin $(i, q)$.

The entire optimal breathing trace can be found by backtracking the bins that contribute to the maximum score $g^*(I)$ of the last timestamp. For the rest of the breathing trace in the observation window, i.e., $\forall\; i = I-1, I-2, \ldots, 1$, we have

$$g^*(i) = \arg\max_{\forall q \in Q} \; s(i, q) + \lambda \; \log \mathbf{P}\big(q, g^*(i+1)\big). \tag{15}$$

The backtracking procedure in (15) gives the optimal trace $\mathbf{g}^*$ for a given spectrum, which is the optimal solution for (13).

To further check if there are any other candidate breathing signals in the given spectrum, the trace $\mathbf{g}^*$ should be removed. For the ideal case, we only need to remove the bins along $\mathbf{g}^*$. However, since the number of FFT points are limited, the energy of the breathing signal is diffused around the center of breathing trace, which forms an energy strip in the given spectrum as shown in Fig. 5. Thus, if we only remove the energy along the optimal trace $\mathbf{g}^*$ and consecutively execute dynamic programming in (14) and (15), we will get a group of traces inside one energy strip. Therefore, IDP applies a windowing module on the optimal trace $\mathbf{g}^*$ to emulate the diffusing effect of FFT to get an energy strip. The updated spectrum after we erase the optimal energy strip is

$$\mathbf{S}(i) \leftarrow \mathbf{S}(i) - S\big(i, g^*(i)\big) * \mathbf{w} \quad \forall i = 1, 2, \ldots, I \tag{16}$$

where $\mathbf{S}(i)$ denotes the energy of spectrum at timestamp $i$, and $\mathbf{w}$ is the frequency response of the windowing module.

Authorized licensed use limited to: University of Maryland College Park. Downloaded on November 30,2020 at 19:35:07 UTC from IEEE Xplore. Restrictions apply.

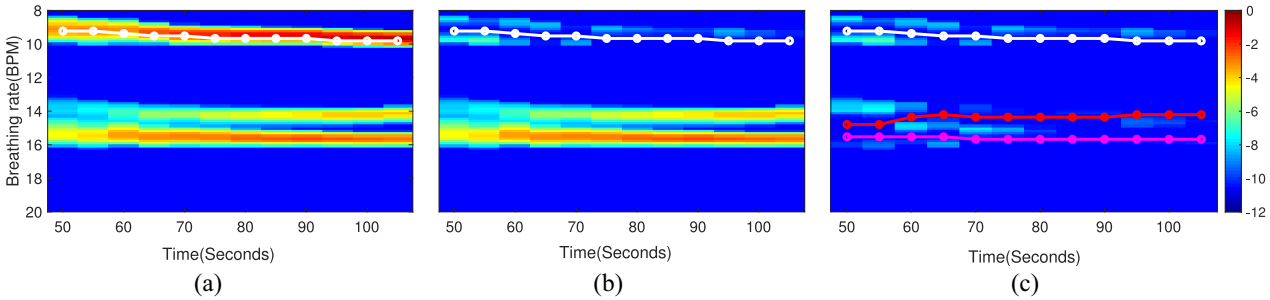



Fig. 6. Successive cancellation procedure of IDP. (a) First trace found by IDP. (b) Spectrogram after erasing the energy stripe of first trace. (c) Spectrogram after erasing all energy stripes found by IDP.

Operator $*$ denotes convolution operation, which can emulate the energy stripe caused by the diffusing effect of FFT.

We recursively perform the above-dynamic programming and spectrum cancelation to find multiple traces. The algorithm terminates when the score of the found trace is negative, indicating an empty spectrum without any effective traces.

The procedure of IDP is summarized in Algorithm 1. Fig. 6 illustrates the details of this finding-then-erasing procedure. In Fig. 6(a), the trace found by DP is marked by the line, and the energy stripe of this trace is removed as shown in Fig. 6(b). The spectrogram, when IDP terminates, is shown in Fig. 6(c), and lines in the figure indicate the breathing traces. It is clear to see that although there is still some residual energy not perfectly removed, IDP terminates properly since there are no traces satisfying the constraint of nonnegative regularized energy.

*3) Detecting Empty Case:* Ideally, when there is no person present in the monitoring area, no breathing trace would be picked up since the spectrum would be random due to the normal distribution of the thermal noise. In reality, however, false traces could be detected since some noise might be boosted in the empty case. To avoid this effect, we employ motion detection to determine empty cases. If no motion (not even chest movement) is detected, the system will directly claim empty; otherwise, the above steps are performed to find a potential breathing rate trace. Here the motion detector needs to be sensitive and robust enough to detect breathing motion. In this article, we employ the state-of-the-art approach proposed in [41] for this purpose, which achieves almost zero false alarm.

**Algorithm 1** IDP

Calculate regularized energy map $s(i, j)$ by (14)
Initialize trace number $t \leftarrow 0$, frequency response of rectangular window $\mathbf{w}$
**while** $\max_q s(I, q) > 0$ **do**
  $t \leftarrow t + 1$
  $g_t(I) \leftarrow \arg\max_q g(I, q)$
  $i \leftarrow I - 1$
  **while** $i \neq 0$ **do**
    $g_t^*(i) \leftarrow \arg\max_q \; s(i, q) + \lambda \, \mathbf{P}(q, g_t^*(i+1))$
    $i \leftarrow i - 1$
  **end while**
  update spectrum $\mathbf{S}(i) = \mathbf{S}(i) - S(i, g^*(i)) * \mathbf{w}, \forall i = 1, 2, \ldots, I$
  Calculate regularized energy map $s(i, j)$ by (14)
**end while**

## C. Trace Concatenating

IDP provides the breathing rate traces for each time window. In practice, a continuous monitoring system, however, would operate for a much longer time than a time window, posing extra information gains to enhance the trace extraction. In this part, we propose a novel trace concatenating algorithm to concatenate trace segments belonging to the same breathing signal in different time windows, which not only improves the trace segments but also enables detection of the start and end time of each trace (or equivalently, the entering and leaving time of a specific user).

For clarity, we store all presented traces in a database. The $j$th trace found previously is denoted as $\mathbf{g}_j^{\text{pre}}(t_{\text{st}} : t_{\text{end}})$, where $j = 1, \ldots, J$, and $t_{\text{st}}$ and $t_{\text{end}}$ denote the start and end time of the trace. The $k$th traces found in the current time window $[t - W, t]$ is denoted as $\mathbf{g}_k(t - W : t)$, where $k = 1, \ldots, K$. Furthermore, the similarity between two traces is defined as the ratio between the overlapped time in the window and the window length, which is expressed as

$$f\left(\mathbf{g}_j^{\text{pre}}, \mathbf{g}_k\right) = \frac{|\mathbb{1}\left(\mathbf{g}_j^{\text{pre}}(t_{\text{st}} : t_{\text{end}}) = \mathbf{g}_k(t - W : t)\right)|}{I - 1} \quad (17)$$

where $f(\mathbf{g}_j^{\text{pre}}, \mathbf{g}_k) \in [0, 1]$. A similarity matrix $F \in \mathbb{R}^J \times \mathbb{R}^K$ can be calculated according to (17) to show the similarity between all the traces in the current window and those in the database. In order to find the previous part for $\mathbf{g}_k(t-W : t)$, we only need to find the maximum item of $\mathbf{f}(\mathbf{k})$, which is the $k$th column of $F$. The row index of the maximum similarity indicates the index of the previous trace if the maximum similarity is above a predefined threshold.

If there exists a previous trace with a high enough similarity, it means that the corresponding breathing rate trace has been detected before. Then, the endpoint of the corresponding trace should be updated. We let the endpoint be the current time and update the corresponding frequency component accordingly. In case a new user arrives, there will be no existing traces that have a similarity larger than the threshold and thus a new trace is created with the corresponding timestamps and frequency

Authorized licensed use limited to: University of Maryland College Park. Downloaded on November 30,2020 at 19:35:07 UTC from IEEE Xplore. Restrictions apply.

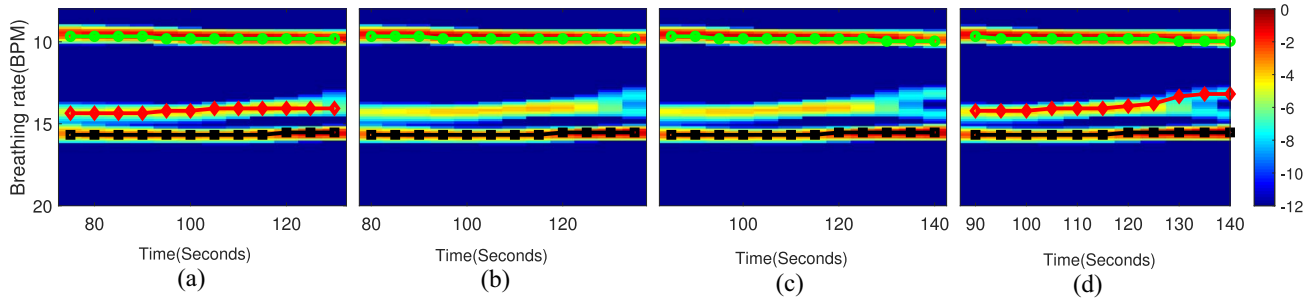



Fig. 7. Traces found by IDP in four adjacent time windows.

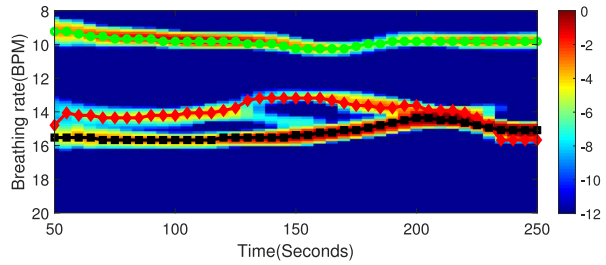



Fig. 8. Trace concatenating result of windows in Fig. 7.

components. Similarly, no trace in the current window being similar to the past traces corresponds to a user that has left, and thus the trace would be terminated. The procedure of trace concatenating is summarized in Algorithm 2.

Figs. 7 and 8 show the effect of the trace concatenating algorithm. Four adjacent time windows are shown in Fig. 7, and traces found by IDP are marked by lines. We can see that although the breathing trace in the middle of the spectrogram is not detected in the second and third window (due to body motion as well as the breathing rate change of the subject), since the trace found in the fourth window still achieves high similarity with the trace found in the first window, it still can be concatenated as shown in Fig. 8.

## VI. People Counting and Recognition

### A. People Counting

IDP and trace concatenation provides the estimation of breathing rate traces, and the trace number would be the estimate of the occupancy level. Although the IDP and trace concatenating have considered the stability of human breath in the observation area, the estimation result may still suffer from large noise and have some false alarms or underestimations for a real-time output as shown in Fig. 7(b) and (c). To eliminate/mitigate these outliers for a real-time system, we design a quasi-bilateral filter to explore the information contained in the previous estimations. Similar to the bilateral filter [42], the designed filter considers the distance in time and estimation domain simultaneously, but we make some improvements according to our system. First, for a real-time system, the filter can only use the past and current estimation result. Furthermore, since IDP and trace concatenation leverage time as well as frequency information to get continuous breathing rate traces, the preliminary estimations are consistent in a short period. Thus, if we directly use a bilateral filter,

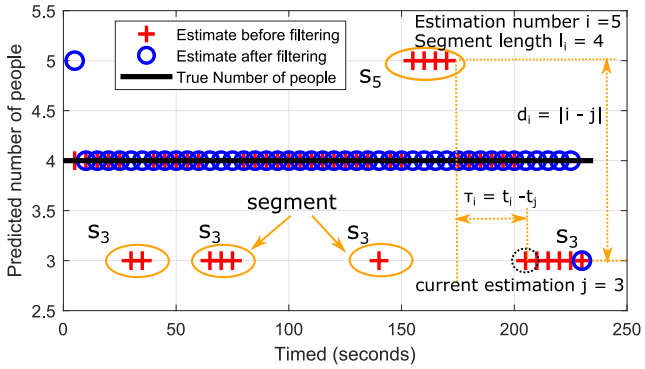



Fig. 9. Example and illustration of quasi-bilateral filter.

only the first incorrect output will be rectified. Given these two constraints, we develop a segment-based filter, where each segment is a consistent preliminary estimation sequence.

Specifically, the output is determined by the current estimation and the previous segments. We denote the weight of segment $s$ as $W_{\text{seg}}(s)$ expressed as

$$W_{\text{seg}}(s) = w(l_s) * w(\tau_s) * w(d_s) \tag{18}$$

where $l_s$ is the length of segment $s$, and $\tau_s$ is the time difference between segment $s$ and current timestamp. $d_s$ is the estimation difference between current estimation and segment $s$ as shown in Fig. 9. Intuitively, the longer the segment is, the greater weight will be assigned. In contrary, the larger the distance is, no matter in time or the estimate of the crowd number, the smaller the influence of this segment imposing on the current result. For clarity, the set of segments with $i$ estimated people is denoted as $S_i$, and the current estimated number as $j$. The weight that the currently estimated people is $i$ after quasi-bilateral filter can be calculated by

$$p(i) \leftarrow \frac{1}{N}\left(\sum_{s\in S_i}\frac{l_s}{\tau_s}\right)e^{-d_s} \tag{19}$$

where $N$ is the total number of segments, the estimation difference is $d_s = |i-j|$, and $W_{\text{seg}}(s)$ in (18) is designed as

$$W_{\text{seg}}(s) \leftarrow \frac{l_s}{\tau_s}e^{-d_s}. \tag{20}$$

The eventual result after filtering is $j'$, given by

$$j' = \arg\max_i p(i). \tag{21}$$

Authorized licensed use limited to: University of Maryland College Park. Downloaded on November 30,2020 at 19:35:07 UTC from IEEE Xplore. Restrictions apply.

**Algorithm 2** Trace Concatenating

Input: Trace num. $J$ and trace matrix $\mathbf{TR}$ in the database, trace num. $K$ and trace matrix $\mathbf{tr}$ in the current window, start and end timestamp of current window $t_{\text{st}}$, $t_{\text{end}}$
Initialization: Indicate vector flag $\leftarrow$ $ones(1, K)$
**if** $J == 0$ **then**
  $\mathbf{TR} \leftarrow \mathbf{tr}$
  $J \leftarrow K$
  flag $\leftarrow$ $zeros(1, K)$
**else**
  **for** $j = 1 : J$ **do**
    $\mathbf{f} = \frac{|\mathbb{1}(\mathbf{TR}(j)=\mathbf{tr})|}{I-1}$
    $f \leftarrow max\,\mathbf{f}$, $k \leftarrow \arg\max \mathbf{f}$
    **if** $f < th$ **or** sum(flag) == 0 **then**
      Continue
    **else**
      Update the end point and the location of trace:
      $\mathbf{TR}(j)$.end $\leftarrow$ $t_{\text{end}}$, $\mathbf{TR}(j)$.Loc($\mathbf{tr}(k, :)$, $t_{\text{end}}$)
      flag$(k) \leftarrow 0$
    **end if**
  **end for**
**end if**
**if** sum(flag) $> 0$ **then**
  index = find(flag == 1)
  $\mathbf{TR} \leftarrow [\mathbf{TR}, \mathbf{tr}(\text{index})]$
  $J \leftarrow J + \text{sum(flag)}$
**end if**

Fig. 9 shows the estimation results before and after quasi-bilateral filtering. Clearly, the novel quasi-bilateral filter can remove the estimation outliers effectively, and thus improve the performance of the people counting system.

### *B. Human Recognition in Smart Homes*

In this part, we apply our respiration tracking result to an interesting smart home application, where we want to recognize human identity in the observation area. Based on our observations, although at some time instances, the breathing rate of different people can be the same, for a specific person, his or her breathing rate tends to locate in a certain range, which is a quite unique feature for human/user recognition. In other words, the breathing distribution of different people tends to be different. Motivated by this observation, we utilize hypothesis testing to do human recognition leveraging the breathing traces we obtained.

The PDF of each person is assumed to follow a Gaussian distribution, i.e., $p_k(\theta) \sim \mathcal{N}(\mu_k, \sigma_k)$, where $k$ indicates the label of the testing subject. To obtain the breathing rate distribution of different people, we first get the rough PDF estimation from the histogram in the training data set. This histogram will be fitted into a Gaussian model to get PDF for each testing subject. Fig. 10 shows exemplary breathing PDFs for four subjects. Based on the PDF distribution, for a certain observation of breathing trace $g$, the log likelihood of the trace

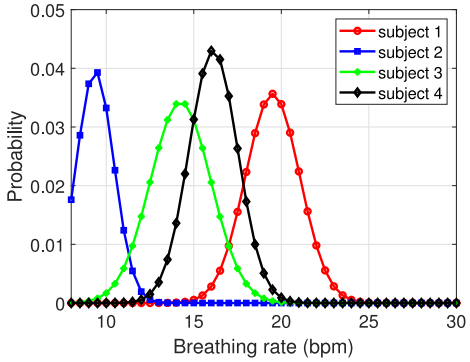
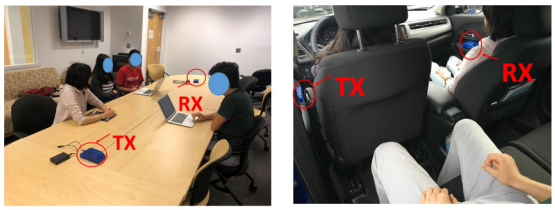



Fig. 10. Estimated PDF of four testing subjects.




(a) (b)

Fig. 11. Experiment setup. (a) LAB. (b) Car.

belonging to subject $k$ can be calculated by

$$\mathcal{L}(k) = \sum_{n=1}^{N} \log(p_k(g(n))) \tag{22}$$

where $N$ is the number of time instances of a given trace. The predicted label should be the one that achieves the maximum likelihood, i.e.,

$$K = \arg\max_k \mathcal{L}(k). \tag{23}$$

## VII. Experiments and Evaluation

In this section, we conduct extensive experiments to evaluate the performance of the proposed approach. Specifically, we first introduce the experimental setup and then the results corresponding to two different scenarios. The discussion on the impact of distinct modules is proposed in Section V.

### *A. Device and Methodology*

We conduct experiments using a pair of commodity WiFi devices, one as Tx and the other as Rx. The channel is set to 5.765 GHz with a bandwidth of 40 MHz. Both Tx and Rx are equipped with three omnidirectional antennas. Each link between a Tx antenna and an Rx antenna has a total of 114 SCs. Considering for practical long-term monitoring, we use a very low sampling rate of 10 Hz.

All the data in our experiments are collected in an on-campus lab and a car over two months with 12 participants. Fig. 11(a) shows the layout of the LAB in which two devices (Tx and Rx) are put on two different sides of a round desk, and the distance between the Tx and Rx is 3.5 m. Participants are invited to sit in chairs as if they were attending a meeting. During the experiments, the participants randomly choose their

Authorized licensed use limited to: University of Maryland College Park. Downloaded on November 30,2020 at 19:35:07 UTC from IEEE Xplore. Restrictions apply.

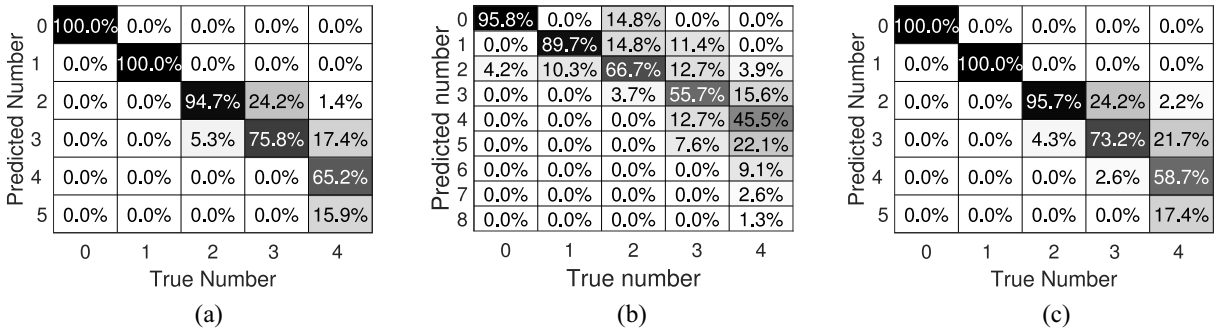



Fig. 12. Confusion matrix of people counting in LAB. (a) Proposed system. (b) TR-BREATH system. (c) Without quasi-bilateral filter.

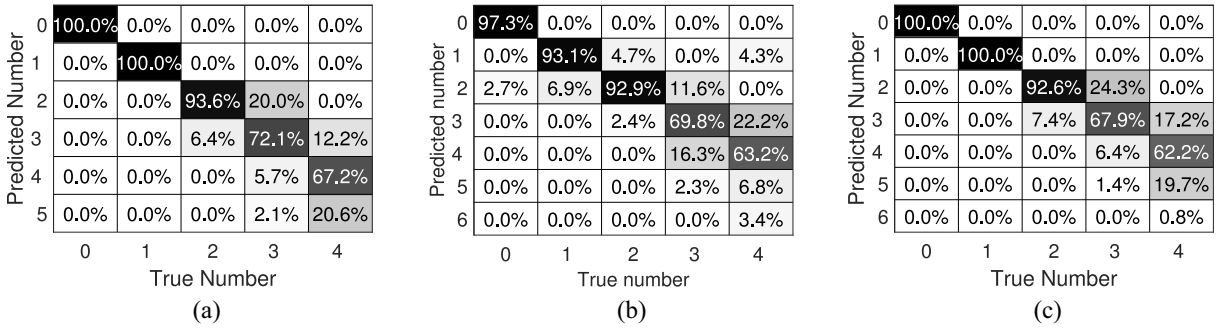



Fig. 13. Confusion matrix of people counting in car. (a) Proposed system. (b) TR-BREATH system. (c) Without quasi-bilateral filter.

seats and slight movements are allowed. To further verify that the proposed system is independent of the environment, we also conduct experiments in a car, which is an extreme case for indoor scenario, where there is limited space as well as strong reflection. For the car scenario, the Tx and Rx are put at the front door on the driver and passenger side, respectively, as shown in Fig. 11(b).

### *B. Overall Performance*

Fig. 12(a) shows the confusion matrix of our method in the LAB, and the overall accuracy is 87.14%, with the accuracy defined as

$$\text{Acc} = \frac{\text{\# of predicted label equals true label}}{\text{total \# of samples}}. \tag{24}$$

The counting error is within one person for 98.6% of the testing cases. Additionally, the proposed system can perfectly detect whether the monitoring area is occupied or not. The accuracy, however, decreases with more people present. This is as expected since the more people there are, the more likely their breathing traces may merge together and the more likely occasional motion may happen, both leading to counting errors. Fig. 13(a) shows that our testing result in the car can achieve a comparable performance with that in the LAB, which demonstrates the independence of our system on the environment.

To further evaluate our system, we compare it with the most relevant TR-BREATH [21] which also estimates multi-person breathing rates using commercial WiFi. TR-BREATH employs root-MUSIC for breathing rate estimation and uses the affinity propagation algorithm to estimate crowd number. In order to make a fair comparison, the quasi-bilateral filter is used to filter out the outliers of original TR-BREATH estimations. The estimation accuracy of TR-BREATH [21] in LAB and car are shown in Figs. 12(b) and 13(b), respectively. As seen, TR-BREATH shows a comparable performance in the car testing scenarios. The performance in the LAB environments is much worse, with an overall accuracy of 70.68%. The proposed approach improves the overall performance by 16.46% and 3.32% for LAB and car testing scenario, respectively, attributed to its three core techniques: 1) adaptive SC combination; 2) IDP; and 3) trace concatenation.

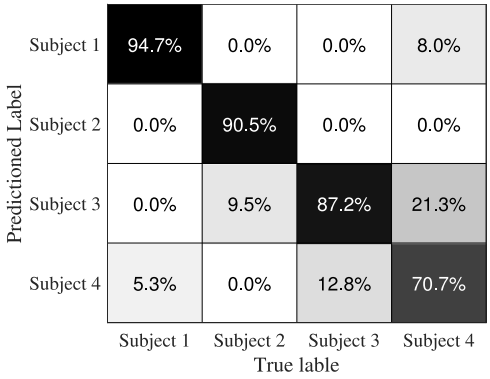



Fig. 14. Confusion matrix of people recognition.

Fig. 14 shows the confusion matrix of people recognition, where the PDFs of each individual are shown in Fig. 10. We can see that when the subject has a distinct breathing rate, it corresponds to a high recognition rate (i.e., subject 1 and subject 2). But for the case that the two subjects have very similar breathing rate ranges, it is hard to distinguish from

Authorized licensed use limited to: University of Maryland College Park. Downloaded on November 30,2020 at 19:35:07 UTC from IEEE Xplore. Restrictions apply.

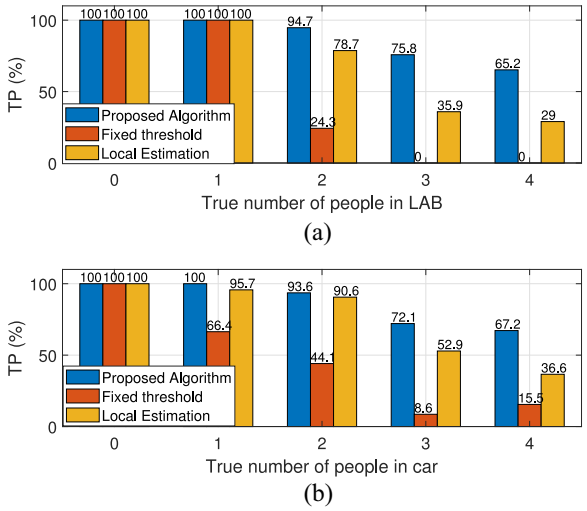



Fig. 15. TP comparison of different algorithms.

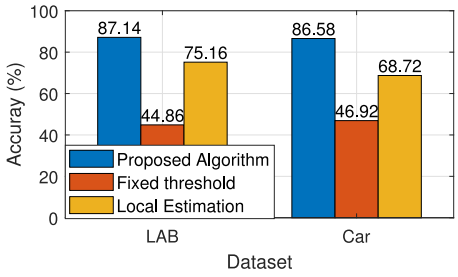



Fig. 16. Accuracy of different algorithms.

each other (i.e., subject 3 and subject 4), and this is also the reason for underestimations in the people counting system.

### C. Performance Gain of Individual Modules

In this section, we discuss how each independent module improves the performance of our system. Apart from the confusion matrix and accuracy, here we additionally adopt the true positive (TP) rate, which is calculated as

$$\mathrm{TP}_i = \frac{\#\text{ of samples that predicted label is } i}{\text{total }\#\text{ of samples that true label is } i}. \tag{25}$$

*1) Impact of SNR Boosting Algorithm:* Here, we compare the proposed SNR boosting algorithm with the commonly used one, i.e., selecting the SCs whose maximum energy is above a certain threshold (hereafter, it is called *fixed threshold* algorithm). For a fair comparison, we choose the 30% of the maximum link energy as the threshold for both methods. Furthermore, the energy of each link is normalized before link combination, thus the parameters used in the later process for both methods are also the same. Fig. 15 compares the TP of the two SC selection algorithms. It can be easily seen that our proposed algorithm shows better performance. This is benefited from our observation that SC has different sensitivity on distinct breathing signals, i.e., for each SC, we only choose the part that is most likely to be a signal, thus compressing noise and interference well. Fig. 16 also shows the superiority of our method in the accuracy aspect. Besides, the fixed threshold algorithm degrades significantly when there are more than three people because they cannot effectively suppress noise and interference.

*2) Impact of the IDP Estimation Algorithm:* In this experiment, we show the benefits of the proposed IDP. We compare the performance with a *local estimation* algorithm that estimates the number of people based on the spectrum at the current timestamp only.

For a fair comparison, the quasi-bilateral filter is also applied to the local estimation algorithm. The comparisons of TP and accuracy for the two algorithms are portrayed in Figs. 15 and 16, respectively. The results show that IDP considerably improves the performance for both data sets, which demonstrates the gains contributed by leveraging time diversity in counting.

*3) Impact of the Quasi-Bilateral Filter:* In this experiment, we show the effect of the designed quasi-bilateral filter on the performance of the people counting system. Figs. 12(c) and 13(c) show the confusion matrix of the people counting system without filtering on data sets collected in LAB and car, respectively. By comparing the result with Figs. 12(a) and 13(a), we can see that the quasi-bilateral filter can improve the performance in most cases, especially when the number of people is larger than 3 in the observation area. The reason is that when the number of subjects increases, more motion interference will be introduced. Furthermore, it is more likely that different breathing traces will merge. Even though we utilize the time domain as well as frequency diversity by IDP, estimation error still can occur. The quasi-bilateral filter is a post-processing method that will further utilize the diversity in the time domain and thus correct the estimate outliers.

### D. Discussion

*1) Resolution:* To further investigate the impact of spatial separation as well as respiration rate difference of human subjects, we perform experiments with three participants sitting with different spatial separations, as shown in Fig. 17. Considering the volume of a human subject, the minimum distance is set as 70 cm. The distance between Tx and Rx is 3.2 m. To ensure a constant breathing rate separation during the experiments, each of the subjects performs controlled breathing according to a metronome. The breathing rate separations of [0.5, 1, 1.5, 2] BPM are evaluated, respectively. Table I shows the performance of our system, where the 4-tuple $(*;*;*;*)$ denotes the detection accuracy and the relative breathing rate accuracy with the three users at location $a$, $b$, and $c$, respectively. We can see that the breathing rate separation has a significant impact on the performance, while the impact of the spatial separation is negligible. The detection rate raises more than 30% when the breathing rate separation increases from 0.5 to 1 BPM. The system achieves 100% detection rate once the breathing rate separation is above 1.5 BPM. Besides, as long as the breathing rate trace has been detected, our system can accurately estimate the breathing rate, and the breathing estimation accuracy is above 94.7% for all of the test cases.

*2) Motion:* Random body motion is one of the most challenging problems for wireless vital sign monitoring. To combat

Authorized licensed use limited to: University of Maryland College Park. Downloaded on November 30,2020 at 19:35:07 UTC from IEEE Xplore. Restrictions apply.

TABLE I
ACCURACY WITH DIFFERENT FREQUENCY AND SPATIAL SEPARATION

| Frequency separation / Spatial separation | 0.5 BPM | 1 BPM | 1.5 BPM | 2 BPM |
|---|---|---|---|---|
| 70 cm | $(65.5; 98.7; 98.0; 98.7)$[1] | $(96.5; 98.9; 96.5; 97.3)$ | $(100; 97.3; 98.8; 97.7)$ | $(100; 98.1; 98.9; 94.7)$ |
| 100 cm | $(62.1; 98.6; 99.4; 99.2)$ | $(96.6; 98.6; 98.4; 98.9)$ | $(100; 97.8; 99.0; 98.9)$ | $(100; 98.2; 98.5; 99.0)$ |
| 130 cm | $(65.0; 98.1; 98.5; 99.7)$ | $(96.6; 98.9; 98.5; 97.7)$ | $(100; 98.4; 99.1; 97.1)$ | $(100; 98.1; 99.3; 99.4)$ |

[1] The 4-tuple $(*; *; *; *)$ denotes the detection accuracy and the relative breathing rate accuracy with 3 participants at location $a$, $b$, and $c$ respectively. For example, $(65.5; 98.7; 98.0; 98.7)$ denotes the detection accuracy is $65.5\%$, and the breathing rate accuracy at location $a$, $b$, and $c$ are $98.7\%$, $98.0\%$, and $98.7\%$ respectively.

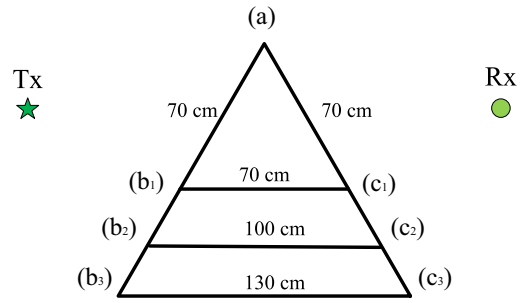



Fig. 17. Experiment setup for resolution investigation.

the impact of slight motion [e.g., typing a keyboard or slight limb motion as shown in Fig. 11(a)], we propose an adaptive SC selection algorithm as well as IDP to extract the breathing rate trace. However, it is hard to accurately extract the breathing signal when continuous large motion such as walking is present due to the inherent limitation of wireless sensing. As long as there is no continuous large motion, which is usually the case for most wireless monitoring applications, the proposed system can correctly pick up the breathing rate traces.

There are indeed dedicated systems [6] designed to detect the dip events of RSS of WiFi systems when humans are walking and estimate the crowd number. Gait information can also be extracted from WiFi systems to do human recognition [11]. However, those systems are specifically designed for the case when the subjects need to keep walking. Our system can perform crowd counting and recognition when the subjects are static.

Recall the experiments performed in a car to verify that the proposed system is independent of the environment, during which the car engine is on but the car is not moving. The proposed respiration tracking method performs well with the slight vibration introduced by the engine, however, it may not work well in a moving car since the vibration from a running car is large enough to overwhelm the minute vital signals. However, once the car is static, e.g., waiting before the traffic light, the proposed system can track the respiration rate traces for crowd counting and human recognition. Handling car vibration during driving is a great common challenge faced by various sensing techniques not only RF-based but also vision-based approaches. We keep this as a future direction to explore.

## VIII. CONCLUSION

This article presents a breathing rate trace tracking system to achieve people counting as well as recognition using commercial WiFi. The proposed system enables static crowd counting and human recognition by multiperson breathing rate trace tracking with three key components: 1) an adaptive SC combination method that boosts breathing signals; 2) an IDP algorithm to extract the successive breathing traces from different individuals; and 3) a trace concatenating algorithm that splices consecutive breathing trace segments. The experimental results show a respective average accuracy of 87.14% and 86.58% for an indoor office and car environments for people counting. Additionally, the counting error is within one person for 97.9% of the time. The average accuracy of people recognition is 85.78% for 4-people smart home scenario.

## REFERENCES

[1] B. Wang, Q. Xu, C. Chen, F. Zhang, and K. J. R. Liu, "The promise of radio analytics: A future paradigm of wireless positioning, tracking, and sensing," *IEEE Signal Process. Mag.*, vol. 35, no. 3, pp. 59–80, May 2018.

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

[3] E. Cianca, M. De Sanctis, and S. Di Domenico, "Radios as sensors," *IEEE Internet Things J.*, vol. 4, no. 2, pp. 363–373, Apr. 2017.

[4] H. Jiang, C. Cai, X. Ma, Y. Yang, and J. Liu, "Smart home based on WiFi sensing: A survey," *IEEE Access*, vol. 6, pp. 13317–13325, 2018.

[5] S. Depatla, A. Muralidharan, and Y. Mostofi, "Occupancy estimation using only WiFi power measurements," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 7, pp. 1381–1393, Jul. 2015.

[6] S. Depatla and Y. Mostofi, "Crowd counting through walls using WiFi," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun.*, 2018, pp. 1–10.

[7] Q. Xu, Y. Chen, B. Wang, and K. J. R. Liu, "Radio shot: Through-the-wall human recognition," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, 2016, pp. 1067–1071.

[8] J. Zhang, B. Wei, W. Hu, and S. S. Kanhere, "WiFi-ID: Human identification using WiFi signal," in *Proc. Int. Conf. Distrib. Comput. Sensor Syst. (DCOSS)*, 2016, pp. 75–82.

[9] M. N. A. Nipu, S. Talukder, M. S. Islam, and A. Chakrabarty, "Human identification using WiFi signal," in *Proc. Joint 7th Int. Conf. Informat. Electron. Vis. (ICIEV) 2nd Int. Conf. Imag. Vis. Pattern Recognit. (icIVPR)*, 2018, pp. 1–2.

[10] H. Zou, Y. Zhou, J. Yang, W. Gu, L. Xie, and C. J. Spanos, "WiFi-based human identification via convex tensor ShapeLet learning," in *Proc. 32nd AAAI Conf. Artif. Intell.*, 2018, pp. 1711–1719.

[11] W. Wang, A. X. Liu, and M. Shahzad, "Gait recognition using WiFi signals," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput. (UbiComp)*, 2016, pp. 363–373.

[12] S. Bakhtiari *et al.*, "Compact millimeter-wave sensor for remote monitoring of vital signs," *IEEE Trans. Instrum. Meas.*, vol. 61, no. 3, pp. 830–841, Mar. 2012.

Authorized licensed use limited to: University of Maryland College Park. Downloaded on November 30,2020 at 19:35:07 UTC from IEEE Xplore. Restrictions apply.

[13] Z. Yang, P. H. Pathak, Y. Zeng, X. Liran, and P. Mohapatra, "Monitoring vital signs using millimeter wave," in *Proc. 17th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2016, pp. 211–220.
[14] F. Adib, H. Mao, Z. Kabelac, D. Katabi, and R. C. Miller, "Smart homes that monitor breathing and heart rate," in *Proc. 33rd Annu. ACM Conf. Human Factors Comput. Syst.*, 2015, pp. 837–846.
[15] L. Ren, H. Wang, K. Naishadham, O. Kilic, and A. E. Fathy, "Phase-based methods for heart rate detection using UWB impulse doppler radar," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 10, pp. 3319–3331, Oct. 2016.
[16] H. Abdelnasser, K. A. Harras, and M. Youssef, "UbiBreathe: A ubiquitous non-invasive WiFi-based breathing estimator," in *Proc. ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2015, pp. 277–286.
[17] F. Zhang *et al.*, "SMARS: Sleep monitoring via ambient radio signals," *IEEE Trans. Mobile Comput.*, early access, doi: 10.1109/TMC.2019.2939791.
[18] J. Liu, Y. Wang, Y. Chen, J. Yang, X. Chen, and J. Cheng, "Tracking vital signs during sleep leveraging off-the-shelf WiFi," in *Proc. ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2015, pp. 267–276.
[19] D. Zhang, Y. Hu, Y. Chen, and B. Zeng, "BreathTrack: Tracking indoor human breath status via commodity WiFi," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 3899–3911, Apr. 2019.
[20] C. Chen, Y. Han, Y. Chen, and K. J. R. Liu, "Multi-person breathing rate estimation using time-reversal on WiFi platforms," in *Proc. IEEE Global Conf. Signal Inf. Process. (GlobalSIP)*, 2016, pp. 1059–1063.
[21] C. Chen *et al.*, "TR-BREATH: Time-reversal breathing rate estimation and detection," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 3, pp. 489–501, Mar. 2018.
[22] F. Q. AL-Khalidi, R. Saatchi, D. Burke, H. Elphick, and S. Tan, "Respiration rate monitoring methods: A review," *Pediatric Pulmonol.*, vol. 46, no. 6, pp. 523–529, 2011.
[23] Q. Zhu, M. Chen, C. Wong, and M. Wu, "Adaptive multi-trace carving based on dynamic programming," in *Proc. Asilomar Conf. Signals Syst. Comput.*, 2018, pp. 1716–1720.
[24] J. Yang, Y. Ge, H. Xiong, Y. Chen, and H. Liu, "Performing joint learning for passive intrusion detection in pervasive wireless environments," in *Proc. IEEE Int. Conf. Comput. Commun.*, 2010, pp. 1–9.
[25] K. Qian, C. Wu, Z. Yang, Y. Liu, and Z. Zhou, "PADS: Passive detection of moving targets with dynamic speed using PHY layer information," in *Proc. IEEE Int. Conf. Parallel Distrib. Syst.*, 2014, pp. 1–8.
[26] Y. Zhao, N. Patwari, J. M. Phillips, and S. Venkatasubramanian, "Radio tomographic imaging and tracking of stationary and moving people via kernel distance," in *Proc. ACM/IEEE Int. Conf. Inf. Process. Sensor Netw.*, 2013, pp. 229–240.
[27] C. Wu, Z. Yang, Z. Zhou, X. Liu, Y. Liu, and J. Cao, "Non-invasive detection of moving and stationary human with WiFi," *IEEE J. Sel. Areas Commun.*, vol. 33, no. 11, pp. 2329–2342, Nov. 2015.
[28] M. Soyturk, M. C. Bodur, A. B. Bakkal, and S. Öztürk, "Estimating the number of people in a particular area using WiFi," in *Proc. Signal Process. Commun. Appl. Conf.*, 2015, pp. 2541–2544.
[29] M. Wirz, T. Franke, D. Roggen, E. Mitleton-Kelly, P. Lukowicz, and G. Tröster, "Probing crowd density through smartphones in city-scale mass gatherings," *EPJ Data Sci.*, vol. 2, no. 1, p. 5, 2013.
[30] J. Weppner and P. Lukowicz, "Bluetooth based collaborative crowd density estimation with mobile phones," in *Proc. IEEE Int. Conf. Pervasive Comput. Commun.*, 2013, pp. 193–200.
[31] J. Weppner, B. Bischke, and P. Lukowicz, "Monitoring crowd condition in public spaces by tracking mobile consumer devices with WiFi interface," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2016, pp. 1363–1371.
[32] T. Yoshida and Y. Taniguchi, "Estimating the number of people using existing WiFi access point in indoor environment," in *Proc. Eur. Conf. Comput. Sci.*, 2015, pp. 46–53.
[33] H. Zou, Y. Zhou, J. Yang, W. Gu, L. Xie, and C. Spanos, "FreeCount: Device-free crowd counting with commodity WiFi," in *Proc. IEEE Global Commun. Conf.*, 2017, pp. 1–6.
[34] S. Liu, Y. Zhao, and B. Chen, "WiCount: A deep learning approach for crowd counting using WiFi signals," in *Proc. IEEE Int. Symp. Parallel Distrib. Process. Appl. IEEE Int. Conf. Ubiquitous Comput. Commun.*, 2017, pp. 967–974.
[35] S. Di Domenico, G. Pecoraro, E. Cianca, and M. De Sanctis, "Trained-once device-free crowd counting and occupancy estimation using WiFi: A doppler spectrum based approach," in *Proc. IEEE Int. Conf. Wireless Mobile Comput. Netw. Commun.*, 2016, pp. 1–8.
[36] W. Xi *et al.*, "Electronic frog eye: Counting crowd using WiFi," in *Proc. IEEE Conf. Comput. Commun.*, 2014, pp. 361–369.
[37] A. Juels, "RFID security and privacy: A research survey," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 2, pp. 381–394, Feb. 2006.
[38] L. Wolf, T. Hassner, and I. Maoz, "Face recognition in unconstrained videos with matched background similarity," in *Proc. CVPR*, 2011, pp. 529–534.
[39] Q. Xu, Y. Chen, B. Wang, and K. J. R. Liu, "Radio biometrics: Human recognition through a wall," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 5, pp. 1141–1155, May 2017.
[40] A. Goldsmith, *Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2005.
[41] F. Zhang, C. Wu, B. Wang, H.-Q. Lai, Y. Han, and K. J. R. Liu, "WiDetect: Robust motion detection with a statistical electromagnetic model," *Proc. ACM Interact. Mobile Wearable Ubiquitous Technol.*, vol. 3, no. 3, p. 122, 2019.
[42] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *Proc. Int. Conf. Comput. Vis.*, 1998, pp. 839–846.

Authorized licensed use limited to: University of Maryland College Park. Downloaded on November 30,2020 at 19:35:07 UTC from IEEE Xplore. Restrictions apply.