



2018

Literature Review on Wireless Sensing—Wi-Fi Signal-Based Recognition of Human Activities

Chao Wang

Key laboratory of Specialty Fiber Optics and Optical Access Networks, Joint International Research Laboratory of Specialty Fiber Optics and Advanced Communication, Shanghai Institute for Advanced Communication and Data Science, Shanghai University, Shanghai 200444, China.

Siwen Chen

Key laboratory of Specialty Fiber Optics and Optical Access Networks, Joint International Research Laboratory of Specialty Fiber Optics and Advanced Communication, Shanghai Institute for Advanced Communication and Data Science, Shanghai University, Shanghai 200444, China.

Yanwei Yang

Key laboratory of Specialty Fiber Optics and Optical Access Networks, Joint International Research Laboratory of Specialty Fiber Optics and Advanced Communication, Shanghai Institute for Advanced Communication and Data Science, Shanghai University, Shanghai 200444, China.

Feng Hu

Key laboratory of Specialty Fiber Optics and Optical Access Networks, Joint International Research Laboratory of Specialty Fiber Optics and Advanced Communication, Shanghai Institute for Advanced Communication and Data Science, Shanghai University, Shanghai 200444, China.

Follow this and additional works at: <https://tsinghuauniversitypress.researchcommons.org/tsinghua-science-and-technology>, Heilongjiang University, Harbin 150022, China.



Part of the [Computer Sciences Commons](#), and the [Electrical and Computer Engineering Commons](#)

See next page for additional authors

Recommended Citation

Chao Wang, Siwen Chen, Yanwei Yang et al. Literature Review on Wireless Sensing—Wi-Fi Signal-Based Recognition of Human Activities. *Tsinghua Science and Technology* 2018, 23(2): 203-222.

This Research Article is brought to you for free and open access by Tsinghua University Press: Journals Publishing. It has been accepted for inclusion in *Tsinghua Science and Technology* by an authorized editor of Tsinghua University Press: Journals Publishing.

Literature Review on Wireless Sensing—Wi-Fi Signal-Based Recognition of Human Activities

Authors

Chao Wang, Siwen Chen, Yanwei Yang, Feng Hu, Fugang Liu, and Jie Wu

Literature Review on Wireless Sensing—Wi-Fi Signal-Based Recognition of Human Activities

Chao Wang*, Siwen Chen, Yanwei Yang, Feng Hu, Fugang Liu, and Jie Wu

Abstract: With the rapid development and wide deployment of wireless technology, Wi-Fi signals have no longer been confined to the Internet as a communication medium. Wi-Fi signals will be modulated again by human actions when propagating indoors, carrying rich human body state information. Therefore, a novel wireless sensing technology is gradually emerging that can realize gesture recognition, human daily activity detection, identification, indoor localization and human body tracking, vital signs detection, imaging, and emotional recognition by extracting effective feature information about human actions from Wi-Fi signals. Researchers mainly use channel state information or frequency modulated carrier wave in their current implementation schemes of wireless sensing technology, called “Walls have eyes”, and these schemes cover radio-frequency technology, signal processing technology, and machine learning. These available wireless sensing systems can be used in many applications such as smart home, medical health care, search-and-rescue, security, and with the high precision and passively device-free through-wall detection function. This paper elaborates the research actuality and summarizes each system structure and the basic principles of various wireless sensing applications in detail. Meanwhile, two popular implementation schemes are analyzed. In addition, the future diversely application prospects of wireless sensing systems are presented.

Key words: channel state information; frequency modulated carrier wave; human activities recognition; wireless sensing applications; Wi-Fi

1 Introduction

Wireless sensing technologies based on common commercial off-the-shelf wireless devices are core

-
- Chao Wang, Siwen Chen, Yanwei Yang, and Feng Hu are with Key laboratory of Specialty Fiber Optics and Optical Access Networks, Joint International Research Laboratory of Specialty Fiber Optics and Advanced Communication, Shanghai Institute for Advanced Communication and Data Science, Shanghai University, Shanghai 200444, China. E-mail: wangchao@staff.shu.edu.cn.
 - Fugang Liu is with the Science and Technology, Heilongjiang University, Harbin 150022, China.
 - Jie Wu is with the Department of Computer and Information Sciences, Temple University, PA 19122-6096, USA.

* To whom correspondence should be addressed.

Manuscript received: 2017-12-28; accepted: 2018-01-22

technologies of many applications. The technology does not require users to wear any hardware devices; instead the devices only perform the detection functions with common commercial Wi-Fi devices in a passive manner. As computer technology has developed, while the mode that people perceive and identify human activities in the external environment has no longer been confined to their own sensory systems, the novel human-computer interaction devices such as the camera, RFID, thermal imaging, ultrasonic imaging, and radar system^[1] have been able to extend our horizons and help develop the human intellect and experiences through modes of communication. These devices are more common in the current identification system of human activities, but they have some limitations and disadvantages on the detection range and working environment. The camera is easily

affected by the light and horizon, limiting the application scope and scene of the camera. The thermal imaging system can easily be disturbed by various heat sources, and it has small coverage and poor penetrability. In addition, picking up the infrared radiation of the human body by the detector is difficult, because the radiation is so easily blocked. The radar system with the high-precision performance characteristic in the non-line-of-sight (NLOS) environment is suitable for human activities identification, but this system is too expensive and sophisticated to be handled by the average civilian rather than a trained professional. Current commercial gesture recognition systems such as Leap and X-box Kinect use deep perception and computer vision to make interaction between humans and computers based on the possible gestures^[2]. However, based on user experience, the systems require dedicated hardware settings and visibility ranges.

In recent years, as wireless network equipment and technology has rapidly been developed, researchers can use wireless signals to perceive and identify changes in the environment, thus promoting the research progress of many fields with wireless perception being the core technology. In contrast to traditional techniques, researchers have proposed some device-free wireless sensing schemes based on wireless signals, namely, the wireless sensing system called “Walls have eyes”. The general idea of wireless sensing systems is that the wireless signal is multipath propagation in an indoor space, including direct radiation, reflection, diffraction, and scattering propagation, due to the obstruction of the indoor walls, floors, furniture, and other obstacles. Human actions are always changing the propagation path of wireless signals, changing the path fading rule. In particular, the human actions cause multipath fading and a Doppler shift, indicating that the received signal contains abundant environmental information about human actions. In short, it is possible to realize wireless sensing systems with common wireless signals such as Wi-Fi. The wireless sensing system combines three technologies, radio frequency technology, digital signal processing technology, and machine learning technology, to form a novel wireless sensing technology, achieving the passive presentation detection including the position, gait, and gestures of a person. The wireless system also subverts the traditional sensing technology, requiring users to wear special sensors or special hardware devices to work in the space of the experimental infrastructure. In addition, the wireless system does not require users to wear any

sensor devices and can normally work in the line-of-sight (LOS) and NLOS environment, making wireless sensing systems many advantages such as non-invasive, robustness, universality, convenience, and low price.

There are two main kinds of current research focuses in the field: one focus is on using the Channel State Information (CSI) of Wi-Fi signals, using Orthogonal Frequency Division Multiplexing (OFDM) modulation based on the IEEE 802.11n standard for analysis and research at the Intel 5300 wireless network interface card platform^[3]; while the other focus is on using the Frequency Modulated Carrier Wave (FMCW) radio signal to achieve the wireless sensing system based on the software defined radio platform^[4].

Future wireless sensing systems will be based on ubiquitous advanced commercial Wi-Fi signals to provide all kinds of high accuracy, high reliability, high security, and convenient application services of which the core is human actions identification technology available for public use. The application scenario is confined to smart homes, hospitals, office buildings, shopping malls, and other daily life places, and the application scenario can be applied to detect vital signs in the disaster rescue, monitor human activities in tunnels, mines, and factories, and use the actions identification function of wireless sensing systems.

The remainder of this paper is organized as follows: An analysis about the basic principle of the wireless sensing system is first presented in Section 2. Then, Section 3 summarizes several typical cases of each application function. Afterwards, Section 4 reports four general application areas of the wireless sensing system. Finally, several limitations and potential alternative technical solutions for existing problems in the current wireless sensing technology are generalized in Section 5.

2 Basic Principle of the Wireless Sensing System

Wireless sensing systems combine the radio-frequency technology, the digital signal processing technology, and the machine learning technology to form a novel environmental perception technology with wireless signals. This technology collects environmental information among modulated wireless signals determined by the environmental state and uses digital signal processing technology to obtain current interest environmental states to achieve the purpose of perceiving the environment. The general schematic diagram of the

wireless sensing system model is shown in Fig. 1.

2.1 Feasibility analysis of the wireless sensing system

The main way of performing wireless signal transmission is using space wave, that is, direct wave, reflected wave, scattered wave, refracted wave, and diffraction wave. In the typical indoor environment with human bodies and obstacles present, wireless signals usually propagate to the receiver through multiple paths such as reflection, diffraction, and scattering. The received signals of the paths are affected by each respective propagation path and exhibit different physical characteristics at the receiver terminal such as the received signal strength and phase.

In the wireless communication system, the wireless signals in the free space have propagation loss. According to free space propagation model, the distance between the transmitting antenna and the receiving antenna is represented by d , and the received signal power obtained by the Friis transmission formula is as follows:

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi d)^2} \quad (1)$$

where P_t represents the transmitted power, $P_r(d)$ represents the received signal power that is a function of the distance d , G_r represents the receiver antenna gain, G_t represents the transmitter antenna gain, λ represents the wavelength in meters and d represents the distance from transmitter to receiver in meters. Considering the reflection paths by the ceiling and floor, the power received is represented by

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 (d^2 + 4h^2)} \quad (2)$$

where h represents the distance from reflection point on ceiling or floor to LOS path. When a person exists in the indoor environment, several scattered paths are produced by the human body. Those scattered power should also be added in the final received power denoted as follows:

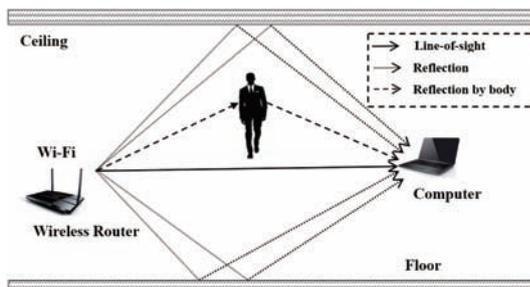


Fig. 1 Schematic diagram of the wireless sensing system model.

$$P_r(d) = \frac{P_t G_t G_r \lambda^2}{(4\pi)^2 (d^2 + 4h^2 + \Delta^2)} \quad (3)$$

where Δ represents a brief representation of path length caused by human body. Accordingly, the human body actions in the radio-frequency coverage change the signal propagation path, causing the received signals power $P_r(d)$ to significantly fluctuate. In the absence of motion interference, $P_r(d)$ basically remains stable^[5], and the total received power by the receiver is the sum of received power from multiple propagation paths. If the intensity magnitudes of the LOS path, the roof reflection path, the ground reflection path, and human scattering paths are E_{los} , E_{ref1} , E_{ref2} , E_{sca} , respectively, then the total received power $P_r(d)$ can be expressed by

$$P \propto |E_{los} + E_{ref1} + E_{ref2} + E_{sca}|^2 \quad (4)$$

where E_{los} , E_{ref1} , E_{ref2} , and E_{sca} are nearly constant in a static environment^[6]. While signal transmission is disturbed by human body actions, E_{sca} temporarily changes. According to the principle of signal propagation, the signal phase is a linear function of the propagation path distance^[7]. Therefore, information changes of the propagation path caused by human actions are also reflected in the received signal phase. In addition, human actions modify the magnitude and phase of the received signal^[5]. The wireless sensing system is sufficient because wireless signals carry enough physical information that represents the environment state.

The key to the wireless sensing system is determining how to obtain environment information from wireless signals. There are two main research technique categories: one category uses the CSI provided by common commercial Wi-Fi devices; while the other category uses the FMCW radio signal generated by FMCW signal generation designed by Katabi et al.^[4,7], and the MIT based software defined radio device USRP N210 collects the environmental information.

In addition, as early as 2013, WiSee designed by Pu et al.^[8] can enable whole-home gesture recognition using the Doppler shifts effect. In their context, the torso reflecting the signals can be thought of as a virtual transmitter. The set of Doppler shifts generated by the relative motion between the virtual transmitter and the receiver could be used to classify different gestures^[8]. In addition, in the same year, MIT Adib and Katabi^[9] could use inverse synthetic aperture radar technology to obtain environmental information. However, these two technologies have not been developed further in wireless sensing applications for the past few years; thus, the rest

of this work only focuses on the further discussing CSI and FMCW.

2.2 CSI

Modern Wi-Fi devices that support IEEE 802.11n/ac standard adopt OFDM multi-carrier modulation technology and typically consist of multiple transmitting and multiple receiving antennas and thus support MIMO. In the wireless communication system, the wireless channel performance is affected by time delay, amplitude attenuation, and phase shift^[6]. Consequently, the wireless channel state needs to be continuously monitored by Wi-Fi devices to effectively perform transmit-power allocations and rate adaptations for each individual MIMO stream such that the available capacity of the wireless channel is maximally utilized^[10].

In the frequency domain, the narrowband flat fading channel model of MIMO system is given by

$$Y = \mathbf{H} \times X + N \tag{5}$$

where Y represents the received signal vector, X represents the transmitted signal vector, N represents the noise signal vector and \mathbf{H} represents the channel state matrix. The CSI represents the estimation plural matrix of the wireless channel, that is, the estimation of the channel state matrix \mathbf{H} , and the CSI essentially describes the channel frequency response of each subcarrier^[3]. In the MIMO system, the CSI matrix is the $N_T \times N_R \times N_C$ dimension matrix, and each element provides the amplitude and phase information of the corresponding subcarrier channel, as follows:

$$\mathbf{H} = \begin{bmatrix} H_{11} & H_{12} & \dots & H_{1N_R} \\ H_{21} & H_{22} & \dots & H_{2N_R} \\ \vdots & \vdots & \ddots & \vdots \\ H_{N_T1} & H_{N_T2} & \dots & H_{N_TN_R} \end{bmatrix} \tag{6}$$

where $H_{ij}=(h_1, h_2, \dots, h_{N_C})$ represents the communication channel state matrix between transmitting antenna i and receiving antenna j , h_k denotes the channel state of the k -th subcarrier, N_T represents the number of transmitting antennas, N_R represents the number of receiving antennas, and N_C represents the number of OFDM subcarriers. In the OFDM system, each subcarrier is an independent fading channel, and CSI provides finer-grained channel description of each subcarrier, that is, the amplitude and phase of the received signal^[11]. In addition, h_k can be expressed by

$$h_k = |h_k|e^{j \sin \theta} \tag{7}$$

where $|h_k|$ represents the amplitude, and θ represents the phase.

The previous section shows that the CSI of Wi-Fi signals can carry enough information about human actions in the special environment to enable wireless sensing systems to be implemented. In practice, however, identical motion causes different extents of signal perturbation on different subcarriers due to frequency-selective fading. Specifically, the same motion does not consistently increase or decrease the received signal power because of constructive and destructive phaser superposition^[5].

The current research method of studying wireless sensing systems based on the CSI is usually built using the CSI tool designed by Halperin et al.^[10] based on Linux and the Intel 5300 wireless network interface card with three omni-directional antennas that provide the CSI. That means their standard hardware configuration uses just a computer as the receiver and a wireless access point, e.g., a common commercial router or a laptop serving as a hot spot as the transmitter^[12]. An $N_T \times N_R \times 30$ CSI matrix can be obtained from each packet, that is, the state information data of each of the 30 subcarriers of each communication link between each antenna pair. This means that the device can collect information on the environment that affects only the channel state of the part subcarrier^[13]; therefore, the CSI is more refined than the Received Signal Strength Information (RSSI). Current research results show that the wireless signal is sensitive enough to actions in the environment so that macro-scale body movements, such as walking, can be detected and that micro-scale motions, such as gestures, keystroke, breathing, and heartbeats, can be extracted and recognized.

2.3 Frequency modulated carrier wave

Frequency modulated carrier wave technology can produce narrowband signals with only a few thousand Hertz bandwidth, and FMCW technology makes the carrier frequency sweep in a linear fashion in time within a certain frequency range, as shown in Fig. 2.

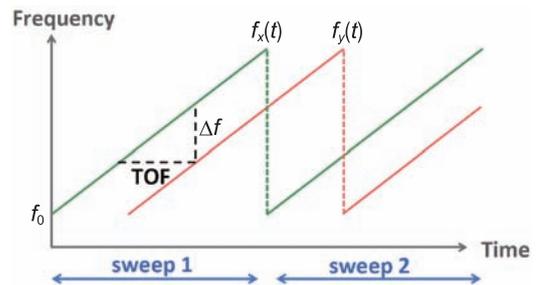


Fig. 2 FMCW provides TOF^[4].

The technique to use a linear relationship between frequency and time can measure the depth of different reflecting objects according to the time of flight (TOF) of the wireless signal as follows^[14]:

$$r = c \times \text{TOF} = c \times \frac{\Delta f}{k} \quad (8)$$

where r represents the round-trip distance of signals, c represents the speed of light, Δf represents the frequency shift between the transmitted and received signals, and k represents the slope that is equal to the total swept bandwidth divided by the sweep time^[14]. The sampled signal s_t can be represented as a complex discrete function of time t as follows^[7]:

$$s_t = A_t e^{-j2\pi \frac{r}{\lambda} t} \quad (9)$$

where r represents the distance traveled by the signal, λ represents its wavelength, and A_t represents its amplitude. Mathematically, a frequency chirp of slope k can be used to compute the signal power $P(r)$ emanating from a particular depth r as follows^[7]:

$$P(r) = \left| \sum_{t=1}^T s_t e^{j2\pi \frac{kr}{c} t} \right| \quad (10)$$

The received signal phase $\phi(t)$ is represented by^[15]

$$\phi(t) = 2\pi \frac{r}{\lambda} t \quad (11)$$

Naturally, due to the depth of the body relative to the detection device and constant changes of human actions in an environment, FMCW signals can accurately describe different depths of the reflection objects; thus, sufficient information about the environment can be captured from the depth information.

Several wireless sensing systems prototypes based on FMCW signals have been built by Katabi et al.^[7] These prototypes generate FMCW signals that sweep from 5.46 GHz to 7.25 GHz using directional antennas with a few milliwatts of transmitted power and can work in both LOS and through-wall scenarios. Katabi et al.^[7] have achieved multiple functions including indoor multi-location and tracking, fall detection, gesture recognition, emotion recognition, vital signs detection, and human body imaging.

2.4 Performance analysis of wireless sensing applications

In recent years, with wide deployment of Wi-Fi infrastructure worldwide, many applications using RSSI to perceive environmental characteristics are emerging.

In the wireless perception field, the universality of RSSI makes it widely used in mobile computing applications such as indoor wireless localization and passive human detection. In essence, RSSI depicts the total received power across all subcarriers, restricting the stability and reliability of RSSI^[16]. As wireless technology has been developed, more advanced wireless technology and signal processing capabilities enable researchers to break through traditional technical bottlenecks and use wireless signals to realize the perception of environmental information. The detection process characteristic of the wireless sensing system is device-free, non-invasive, passive, and through-wall.

2.4.1 Wireless sensing applications based on the CSI

Since the IEEE formally released the 802.11n standard in 2009, commercial Wi-Fi equipment has been able to provide fine-grained CSI, improving the functional and accuracy of wireless sensing systems based on the CSI. Therefore, since Halperin et al.^[10] developed the CSI tool that can extract the CSI from Wi-Fi with the IEEE 802.11n standard, many researchers have studied wireless sensing systems based on the CSI involved areas such as gesture recognition, vital signs, indoor human activity recognition, identification, and intrusion detection. In particular, in recent years, the main reasons why such systems are growing constantly are as follows:

- (1) Universality: This scheme only requires common wireless routers that support the IEEE 802.11n standard and computers with the Intel 5300 wireless network interface card whose firmware is modified according to the CSI tool provided by Halperin et al.^[10] The hardware configuration scheme completes the environmental awareness function only through the ubiquitous Wi-Fi, making wireless sensing systems cheaper and easier and have better prospects and universality than systems that use specialized hardware devices such as USRP N210.
- (2) Fine-grained: The CSI describes the channel state of the subcarrier level. Due to the mutually independent narrowband subcarriers, the multipath effect and shadow fading at different subcarriers result in significant differences in the observed amplitudes or phases, meaning that a small movement in the physical environment may lead to the change of the CSI at some subcarriers, whereas such variations may be canceled if the signal strength over the whole channel bandwidth is examined, and this is why the CSI of the PHY layer is more fine-grained. It can provide great

opportunity to capture the minute movements such as gestures, breathing, and heartbeats in a diverse range of areas^[13].

- (3) High-precision: Many proposed systems can complete the environmental sensing function with high accuracy in the range of controlled experimental space. For example, CARM achieves an average accuracy of greater than 96% for some specified indoor human activities detection^[1]. For example, the Wi-Finger system has achieved up to 90.4% average classification accuracy for recognizing 9 digits finger-grained gestures from American Sign Language (ASL)^[12].

Compared with wireless sensing systems based on FMCW signals, there are some disadvantages in such system schemes. Theoretically, the detection range of wireless sensing systems based on the CSI is just Wi-Fi coverage and there are walls, furniture, and other obstacles in the environment similar to an ordinary apartment or office, providing the function of the through-wall perception. However, since prototype systems are limited by factors such as the Wi-Fi signal frequency, bandwidth, transmitted power, and antenna polarity, they are only evaluated to verify the feasibility in a simple and controllable environment. Although a few systems such as CARM^[1] have achieved system testing with walls in the experimental environment, the accuracy also drops as the number of walls increases. Furthermore, mostly systems are based on the hypothesis that there are only specifically motions designed in advance in the environment and that they cannot detect multiple targets at the same time, and only a few detection systems can achieve this with high accuracy.

2.4.2 Wireless sensing applications based on frequency modulated carrier wave

In military purposes, FMCW radar technology has been used to detect the presence of obstructions such as walls. However, due to the unique characteristics of the military field, for instance, the large system power, the frequency band for the military level, expensive and cumbersome equipment, confidentiality technical, and many other factors make it not suitable for the civilian areas^[4,17].

In 2013, Adib and Katabi^[9] proposed the use of inverse synthetic aperture radar technology to achieve wireless sensing systems from MIT. In 2014, they proposed a solution of wireless sensing systems based on commercial universal software radio peripheral devices by using light-weight low-power FMCW radar technology and conforming to specifications of Federal Communications

Commission on consumer electronics equipment. This process has continually improved system accuracy and reliability involving multi-domain applications to extend radar technology from military to civilian areas successfully^[4]. Thus far, a series of wireless sensing systems based on FMCW have been achieved including Wi-Vi, Wi-Track, Wi-Track 2.0, Vital-Radio, RF-Capture, and EQ-Radio, and these components are used in human location tracking, gesture recognition, vital signs detection, human imaging, and emotional recognition. Furthermore, these systems have the following advantages:

- (1) Stability: The universal software radio peripheral enables computers to work as high-bandwidth software radios, providing a complete and stable development environment that enables researchers to generate radio signals that meet their requirements. The FMCW generation developed by Katabi et al.^[4] in 2014 can produce stable and adjustable FMCW signals to enable their subsequent researches on wireless sensing systems to work in a reliable and stable manner.
- (2) Through-wall detection: The frequency range of FMCW signal adopted by Katabi et al.^[4,7,14,17] is between 5.46 to 7.25 GHz, and the signal in this range can penetrate the walls of conventional urban buildings. Consequently, this system can achieve through-wall detection and not be affected by indoor walls, furniture, and other obstacles while ensuring accuracy.
- (3) Low-power and high-precision: It is an important factor for the detection range and accuracy of the wireless sensing system to increase the power of the transmission. However, as a future consumer electronics product, high-power wireless devices are not allowed. Consequently, determining how to perceive efficient environmental states with low-power wireless signals becomes a problem the researchers must solve. Relying on the directivity of directional antenna, Katabi et al.^[4,7,14,15,17] achieved centimeter-level through-wall detection in an NLOS environment by just transmitting the sub-milliwatt power radio frequency signal.
- (4) Multi-person detection: In the actual multi-person scenario, researchers will introduce more challenges and increase the complexity of effective signal analysis and processing, meaning that the reliability, robustness, and accuracy of wireless sensing systems will also be challenged. In this research, the three-dimensional positioning function of Wi-

Track based on the TOF of FMCW signals provides effective solutions in 2014 on which subsequent systems such as Wi-Track 2.0, Vital-Radio, and EQ-Radio focus on using the reflected signal from objects of different positions to detect multiple people simultaneously. For example, Wi-Track 2.0 can track four mobile people or five stationary people simultaneously. Vital-Radio can then monitor three individual vital signs simultaneously with the same accuracy for each user. EQ-Radio can also recognize the emotions of a specific user in a multi-person environment.

Compared with the current wireless sensing systems based on Wi-Fi, more widely functional, higher accuracy, increased stability, and increased robustness are advantages of these schemes. However, the requirement for special hardware equipment makes them not available to be widely deployed at this stage. With the development of wireless technology and devices, it is foreseeable that this problem can be solved in the future, and this is why many researchers still engage in this research.

3 Wireless Sensing Applications Samples

Since 2013, research on the wireless sensing system has been sharply emerging and has attracted extensive attentions from many research institutions. For example, since 2013, MIT Katabi et al.^[4,7,9,14,15,17] have proposed a series of prototypes on wireless sensing. In addition, Wi-See^[8] proposed by Pu et al. from the University of Washington has been able to control smart devices by gestures since 2013. Also, in China, the number of research institutions is also gradually growing such as the Yunhao Liu group from Tsinghua University, the Wei Wang group from Nanjing University, the Daqing Zhang group from Peking University, the Liusheng Huang group, and the Xiangyang Li group from the University of Science and Technology of China. At the same time, under the guidance of Professor Jie Wu, the Chao Wang group from ShangHai University has cooperated with Temple University to design a wireless sensing system that can detect in real time whether there is an illegal invasion in a specific indoor space.

Various conceptual verification prototypes of different applications have been also implemented in the past few years, mainly including micro-scale motion recognition, human activity recognition, identification, vital sign detection, indoor localization, imaging, and emotional recognition. Experimental results have shown that they can achieve relatively ideal effects. Because the number of

the wireless sensing system rapidly increases, the details of some representative prototypes in each application scenario are explained in the following sections.

3.1 Micro-scale motion recognition

3.1.1 Wi-Key^[18]

In 2015, Wi-Key, a keystroke recognition system proposed by Ali et al.^[18], showed for the first time that Wi-Fi signals can be exploited to recognize keystrokes of 37 keys (26 alphabets, 10 digits and 1 space bar), as shown in Fig. 3.

The intuition is that the hand and finger of a user move in a unique formation and direction while typing a certain key, generating a unique pattern in the time-series of CSI values, and the user calls the CSI-waveform for that key. Wi-Key uses shapes of each extracted keystroke CSI-waveforms as their features, because the shapes retain both time and frequency domain information of the waveforms and are thus more suited for use in classification. Due to the high data rates supported by modern Wi-Fi devices, enough CSI values can be provided within the duration of a keystroke to construct high resolution CSI-waveforms for each keystroke. To obtain clear CSI-waveforms of each keystroke, Principal Component Analysis (PCA) helps remove the uncorrelated noisy components from the signals by taking advantage of correlated variations in the CSI time series of different subcarriers, and these variations cannot be completely removed through traditional low-pass filtering. To lower computational costs in the classification process, Discrete Wavelet Transform (DWT) is first used to compress each keystroke CSI-waveforms as features while preserving most of the time and frequency domain information. Wi-Key trains a k-Nearest Neighbor classifier using features from each transmit-receive antenna pair for each keystroke with Dynamic Time Warping (DTW) distance as the comparison metric between keystroke shape features. The result is determined through majority voting from each classifier.

In real-world experiments, Wi-Key can recognize keystrokes in a continuously typed specific prepared

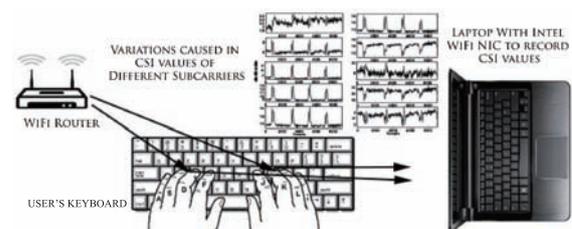


Fig. 3 Wi-Key system^[18].

sentence with an accuracy of 93.5%. Although it works well under relatively stable and controlled environments, there are still many limitations, because the accuracy of the current scheme is susceptible to interference-free surroundings, device positioning, the CSI sampling rate, and controlled typing.

3.1.2 Wi-Finger^[12]

In 2016, Wi-Finger, a finger-grained gesture recognition system proposed by Li et al.^[12] achieved number text input of 9 digits finger-grained gestures from ASL by using ubiquitous Wi-Fi signals, and the number text input has been applied to the human-computer interaction field.

To achieve higher Signal Noise Ratio (SNR) and higher data rate, the transmitter of the Wi-Finger is a wireless router with one DB-Link directional antenna operating in IEEE 802.11n AP mode at 5.745 GHz, because Li et al.^[12] think 2.4 GHz frequency band is crowded and more vulnerable.

In implementing the system, there are three steps in data preprocessing: First, Wi-Finger uses the Hampel identifier^[19] to eliminate outliers namely all points out of the interval $[\mu - \gamma \times \sigma, \mu + \gamma \times \sigma]$, where μ and σ are the median and the median absolute deviation of the data sequence respectively, γ varies in different situations and the most widely used value is 3. Second, the Butterworth low-pass filter is applied to eliminate the out-of-band interference of finger gestures frequencies within 1 Hz to 60 Hz. Finally, the weighted moving average method used in Wi-Fall^[6] is introduced to further clear the filtered CSI stream. Similar to Wi-Key, the CSI stream for each finger gestures needs to be split into segments to extract features as the training sample and identification unit for classification. To do that, Li et al.^[12] refer to CARM^[1] and design a similar adaptive algorithm to detect the starting and finishing points of finger gestures that define $E\{h_2^2/\delta_{q_2}\}$ as a sign indicator to indicate the occurrence of finger motions. Afterwards, 30 subcarriers are combined by averaging every 6 subcarriers and concatenating them to form a synthetic waveform for each gesture, and this is mentioned as a feature vector F . Figure 4 shows the compressed synthetic feature vector for finger gestures No. 1, No. 4, and No. 7, respectively. Similarly, Wi-Finger also uses DWT to compress feature vectors for less computational costs in the classification process and structures a k-Nearest Neighbor classifier with DTW distance as the comparison metric between feature vectors such as Wi-Key.

Wi-Finger runs well for single individuals under a relatively stable environment such as laboratories and

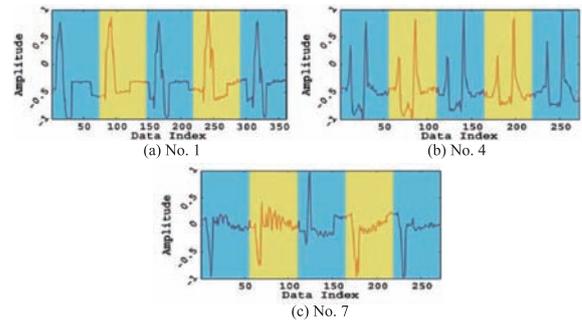


Fig. 4 Compressed features for gestures No. 1, No. 4, and No. 7^[12].

dormitories with only two occupants, a target user and a system controller. It can achieve up to 90.4% average classification accuracy per user and its average accuracy of continuous number text input for single individuals in a desktop reaches 82.67% within 90 digits.

3.2 Human activity recognition

3.2.1 E-eyes^[2]

In 2014, E-eyes, an indoor device-free location-oriented activity identification system proposed by Wang et al.^[2], achieved the human daily activities recognition based on the CSI of the indoor rich wireless network link between Wi-Fi devices. It is a solution that can provide sufficiently accurate tracking and recognition with minimal infrastructure requirements and without the need to carry a dedicated device, as shown in Fig. 5. Different from gesture recognition, activity identification needs to identify a more loosely defined series of motions over a period of time rather than a single well-defined body movement.

For the data preprocessing, E-eyes adopts the dynamic exponential smoothing filter^[20] to remove high frequency noises, which is an exponential smoother that changes its smoothing factor dynamically according to previous samples, and only retains CSI measurements with the modulation and coding scheme index value more than 263 for the activity identification, which will not severely

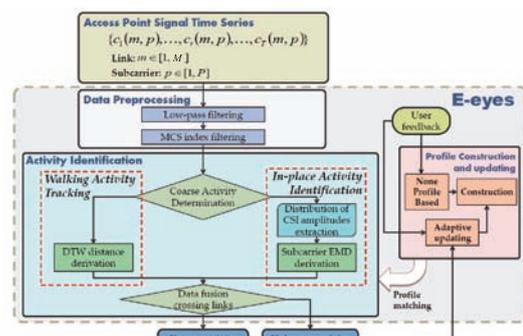


Fig. 5 Framework of E-eyes^[2].

influence the amplitude of the CSI due to the unstable wireless channel.

E-eyes mainly includes two mechanisms: activity identification and profile construction and updating. In activity identification, clear CSI measurements are divided into the corresponding activity category, including in-place activity and walking activity according to their moving variance. For in-place activity identification, E-eyes exploits the distribution of the CSI amplitude as the feature to identify different in-place activities by comparing against known profiles because different daily in-place activities respectively result in different relatively stable distributions of the CSI amplitude. For walking activity tracking, the principle that E-eyes identifies is walking activity according to the CSI measurement profile that when the person passes through doorways, a logical inference is made that allows activities to be strongly tied to which doorway an individual passes. For example, passing through a kitchen doorway at noon is very likely followed by cooking or eating. According to the above rules, E-eyes uses a semi-supervised mode to construct CSI profiles using the collected CSI measurements of typical in-place activity and walking activity by utilizing a distance based non-profiling clustering that can adaptively update activity profiles in time. They adopt the K-Means clustering technique to discriminate different activity instances in terms of the distance based on earth mover distance^[21] for in-place activity and multi-dimensional dynamic time warping^[22] distance for walking activity between the testing measurements and known CSI profiles.

The experimental evaluation in two apartments of different size shown in Fig. 6 demonstrates that E-eyes

can achieve an average true positive rate above 96% and an average false positive rate of 1% to distinguish a set of in-place and walking activities listed in Table 1 and Table 2, respectively, with only a single Wi-Fi access point. However, there must be only one person in the relatively stable experimental environment. In addition, the actual results show that the recognition accuracy of E-eyes can be improved by widening the signal bandwidth, increasing the transmission rate, having more Wi-Fi equipment, and optimizing the algorithm parameter selection.

3.2.2 CARM^[1]

In 2015, CARM, a CSI based human activity recognition and monitoring system proposed by Wang et al.^[1] can quantitatively build the correlation between CSI value dynamics and a specific human activity as the profiling mechanism and recognize a given activity shown in Table 3 by matching the activity to the best-fit profile.

Our previous work has shown that the traditional low pass and median filter cannot effectively eliminate the pulse and burst noise of the CSI stream^[18]. To obtain the optimal results through CSI streams, PCA needs be performed according to the correlation between the CSI data of different subcarrier channels for data fusion and the uncorrelated noisy components such as Wi-Key need to be removed.

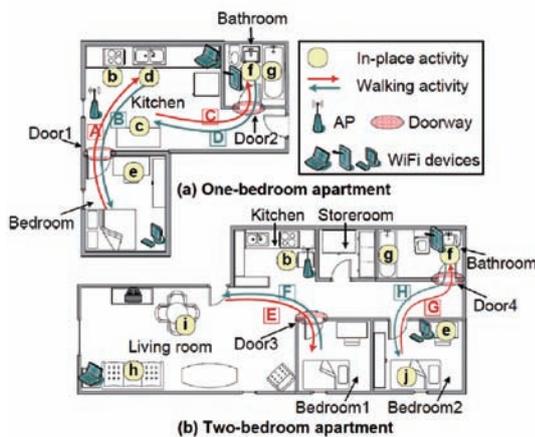


Fig. 6 Experimental setups and the illustration of activities^[2]. (a) One-bedroom apartment. (b) Two-bedroom apartment.

Table 1 In-place activity.

	1-bedroom apt.	2-bedroom apt.
a	Empty	Empty
b	Cooking	Cooking
c	Eating	
d	Washing dishes	
e	Studying	
f	Brushing	Brushing
g	Bathing	Bathing
h		Watching TV
i		Gaming
j		Sleeping
o	Others	Others

Table 2 Walking activity.

	1-bedroom apt.	2-bedroom apt.
A	Bedroom to Kitchen	
B	Kitchen to Bedroom	
C	Bathroom to Kitchen	
D	Kitchen to Bathroom	
E		Outside to Bedroom1
F		Bedroom1 to Outside
G		Bedroom2 to Bathroom
H		Bathroom to Bedroom2

Table 3 Summary of activity dataset^[1].

Activity	Number of samples	Training time (s)
Running	205	16.38
Walking	315	26.84
Sitting down	266	14.49
Opening refrigerator	213	13.49
Falling	98	5.02
Boxing	75	4.88
Pushing one hand	72	7.00
Brushing the teeth	96	7.35
Empty (i.e., no activity)	60	5.10

Then, to automatically detect and extract CSI segments that represent the occurrence of human activities, CARM designs an adaptive detection algorithm that defines the activity indicator as $E\{h_2^2/\delta_{q_2}\}$ compared with a dynamically detection threshold updated by an exponential moving average algorithm. For the signal analysis and feature extraction of each segment, CARM needs to extract frequencies at multiple resolutions on multiple time scales due to human activities associated with duration and frequency. Therefore, CARM applies DWT to decompose the PCA components into 12 levels that span the frequency range from 0.15 Hz to 300 Hz. In addition, a 27-dimensional feature vector is extracted from the averaged DWT results of the five PCA components with the movement information present on each 200 ms interval. To classify and recognize human activities, CARM constructs a Hidden Markov Model (HMM) using the feature vector of training samples for each activity. To estimate mean vectors and covariance matrixes corresponding to each state and the transition probabilities for the HMM, the well-known Baum-Welch algorithm^[23] is used.

In the indoor evaluation scenarios shown in Fig. 7, CARM can recognize the activities shown in Table 3 with an average accuracy of greater than 96%. However, CARM does not support the simultaneous detection of multiple human activities and proper through-wall detection.

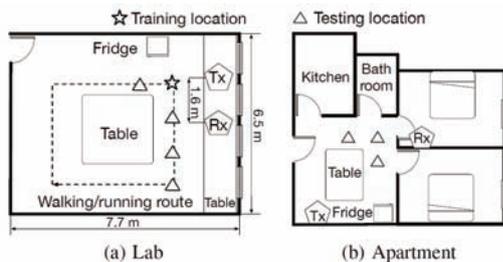


Fig. 7 Experimental scenarios^[1]. (a) Evaluation scenarios in the laboratory. (b) Evaluation scenarios in the apartment.

3.3 Identification

3.3.1 Wi-Who^[24]

In 2016, Wi-Who, a Wi-Fi based person identification system proposed by Zeng et al.^[24], achieved identification, a remaining fundamental question, by using gait characteristics extracted from the CSI of Wi-Fi signals in a device-free manner when a user walked in the smart indoor space. The person identification system supports the system feasibility that recent works have demonstrated that gait can be used as a biometric signature for person identification^[25,26].

While the high-frequency noise and DC component are removed by a Butterworth bandpass filter with cutoff frequency of 0.3 Hz to 2 Hz, distant multipath interference signals, a result of the reception of a strong signal due to reflection from a distant object or person, need be also eliminated in terms of delay characterization. Wi-Who removes the distant multipath components that have a delay longer than 0.5 microseconds of the channel impulse response that delays the profile of signal reception^[27].

To obtain the start indication of the identification, a metric referred as motion energy is introduced in the frequency domain as follows^[24]:

$$Energy = \sum_{i=1}^{window_length/2} magnitude^2 \quad (12)$$

where *magnitude* is the normalized Fast Fourier Transform (FFT) coefficients calculated over the time window and *window_length* is the window length. Wi-Who uses a decision tree-based machine learning classifier that selects information gain as the feature selection criteria to extract gait features of the time and frequency domain people for training and testing based on the step analysis and walk analysis. In step analysis, the system relies on a peak-valley detection algorithm instead of DTW to detect the local minimum and maximum of the time-series data to calculate the step cycle construction in which time-domain gait features of the CSI such as mean, variance, inter-quartile range, mean and crossing rate can be calculated. In walk analysis, frequency domain features such as energy, entropy, and FFT-peaks for the entire walk segment are calculated^[24].

In the evaluation, Wi-Who recognized a specific person's gait and identification with an average accuracy of 92% to 80% from a group of 2 to 6 people, respectively, while a person walked as few as 2–3 meters in most cases. Compared with mature biometric identification technology such as fingerprint, iris, and face recognition with high

accuracy and reliability, Wi-Who has the advantages of passive detection and low price for civilians while filling in the vacancy of identification in the wireless sensing field. The proposal of Wi-Who mainly clarifies the feasibility for the identification using gait features extracted from the CSI of common commercial Wi-Fi signals. However, the system implemented is still built on some assumptions. For example, the user is asked to walk along with a path a meter long parallel to the LOS path between the Wi-Fi sender and receiver. Thus, the system remains to be optimized.

3.3.2 Wi-Fi-U^[28]

In 2016, Wi-Fi-U, a gait detection system for identification proposed by Wang et al.^[28], captured fine-grained gait patterns to recognize identity by using the better signal processing techniques than previous schemes based on specifically designed Doppler radars. Gait features that can be considered adequate unique biometrics information for distinguishing humans include walking speed, gait cycle time, footstep length, movement speeds of torso and legs, and spectrogram signatures extracted from the CSI spectrogram.

As described by Wang et al.^[18] in 2015 regarding Wi-Key, PCA is applied to remove the uncorrelated noisy components from the signals by taking advantage of the important correlation between CSI measurements of different subcarriers in the data preprocessing phase. However, the signal reflections of different body parts are still mixed in the waveform, preventing the extraction of human gait information. To address this problem, Wi-Fi-U uses the short-time Fourier transform technique to transform the waveforms to spectrograms so that CSI waveforms of different body parts can be separated in the time-frequency domain, because the reflection frequencies of body parts of different moving speeds are different. Meanwhile, Wang et al.^[28] applied spectrogram enhancement techniques^[29] to further reduce the noise to obtain a high-quality spectrogram with detailed information for the walking, as shown in Fig. 8.

The system considers three times noise level variance

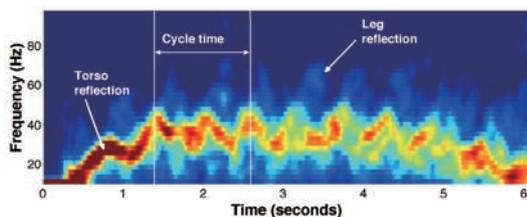


Fig. 8 CSI spectrogram with human walking^[28].

estimation updated by a dynamic thresholding algorithm based on exponential moving average algorithm as the starting detection threshold of walking. The 170 gait features are extracted on the period of steady walking that is defined as the optimal detection period where the torso and leg speed estimated by the percentile method developed for Doppler radars^[30] is no less than 80% of the maximum speed. According to these features, Wi-Fi-U trains a two-class classifier for each person using support vector machine with the radial basis function kernel whose optimal parameters are selected through grid search to build a gait recognition system.

The implementation of Wi-Fi-U demonstrates the feasibility of using COTS Wi-Fi devices for identification through gait patterns in a room with an area of 50 square meters, but there are still some limitations. For example, the user must walk 5.5 meters away on a predefined straight path in a predefined walking direction, and this is suitable for confined spaces such as a corridor or a narrow entrance. In addition, Wi-Fi-U devices cannot work while multiple people walk at the same time.

3.4 Vital signs detection

3.4.1 Vital-Radio^[17]

In 2015, Vital-Radio, an FMCW based device-free wireless sensing technology proposed by Adib et al.^[17], can monitor vital signs including breathing and heart rate remotely without body contact for smart environments. Vital-Radio can perceive periodic chest movements caused by inhaling and exhaling and skin vibrations due to heartbeats using low-power FMCW signals, as shown in Fig. 9. While combining with the location function of Wi-Track, Vital-Radio can simultaneously monitor the vital signs of up to three individuals.

Vital-Radio enables the reflections from different objects to be separated into buckets based on their reflection times, i.e., the TOF by using FMCW signals, which eliminates the mutual interference between the reflections of objects at different distance positions relative to the receiver. Because the resolution of FMCW buckets

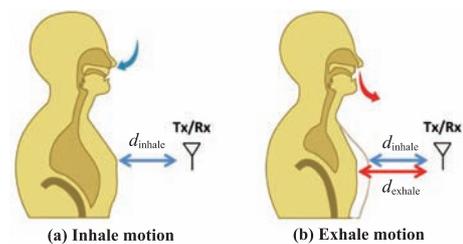


Fig. 9 Chest motion changes the signal reflection time^[17].

is about 8 cm, as shown in Fig. 10, reflections of any objects or motions further than 8 cm from the chest would naturally fall into different buckets, enabling Vital-Radio to focus on the extraction of breathing and heart rate without overmuch interference.

Regarding the interference from static objects such as walls and furniture that submerge useful reflections, Vital-Radio can eliminate them by subtracting consecutive time measurements since their reflections do not change over time. Finally, there are only the reflections of moving objects separated into buckets.

Periodicity is the basis for extracting the reflections from the chest movements and skin vibrations, and the system performs FFT on each 30 second window of the phase signal, and the value of the peak, as shown in Fig. 11, is higher than at least five times the average power in the remaining frequencies. In addition, it has been determined that dominant motions are the breathing and heart beats. However, the corresponding frequency of the peak is an initial estimate of the breathing rate of the individual. To more precisely measure breathing, the peak and its two adjacent bins are retained to be performed an inverse FFT to obtain a complex time-domain signal $s'(t)$. The phase of $s'(t)$ is linear and its slope corresponds to the breathing frequency; thus, an accurate estimate of the

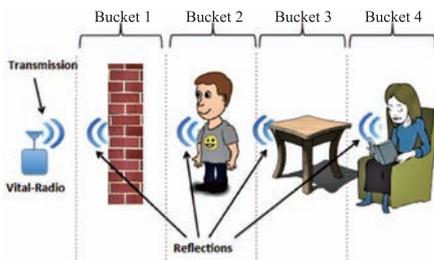


Fig. 10 Separating reflectors into different buckets^[17].

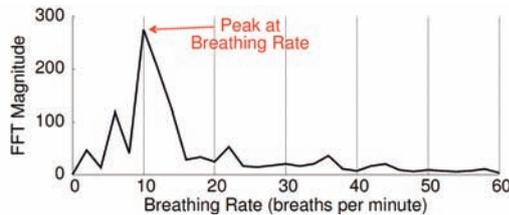


Fig. 11 Output of Fourier transform for breathing^[17].

breathing rate can be calculated by

$$Estimate = 60 \times \frac{slope\{\angle s'(t)\}}{2\pi} \quad (13)$$

This equation is based on a well-known property in signal processing that states that if the signal contains a single dominant frequency, then that frequency can be

accurately measured by performing a linear regression on the phase of the complex time-domain signal.

The extraction of the heart rate reduces the FFT window to 10 seconds after breathing and high frequency noises are both filtered by applying a Hanning window. The corresponding frequency of the maximum peak is a coarse estimate of the heart rate in the heartbeat frequency range between 40 and 200 beats per minute as shown in Fig. 12. Similar to breathing, the more precise estimate of the heart rate can be obtained by regressing on the phase of this signal using Eq. (13).

In the evaluation, Vital-Radio can track the breathing and heart rates of users with a median accuracy of 99% even when users are 8 meters away from the device or in a different room.

3.4.2 Vital signs detection system based on the CSI

Also in 2015, Liu et al.^[17] proposed a vital sign detection system that provides non-intrusive and continuous precise detection based on the CSI for the sleep scenario by using the fine-grained CSI of off-the-shelf Wi-Fi without any wearable or dedicated devices. The motivation of tracking human vital signs during sleep is that they help to assess the general physical health of a person and provide useful clues for diagnosing possible diseases. The data processing frame of the system is shown in Fig. 13.

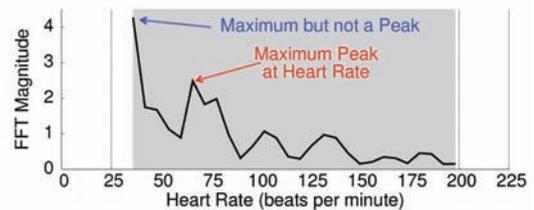


Fig. 12 Output of fourier transform for heart rate^[17].

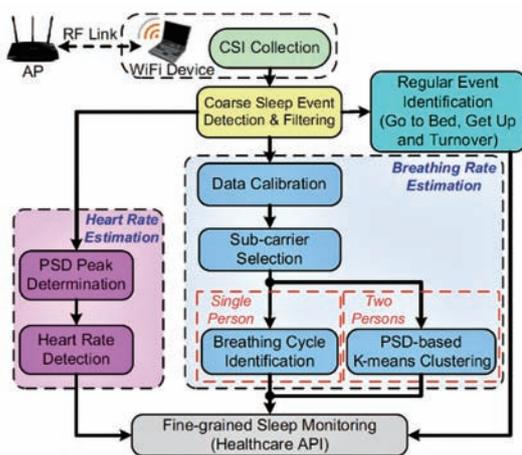


Fig. 13 System frame^[13].

The kind of motions, regular events such as going to bed or vital signs such as breathing and heart beats, contained in the segment of CSI measurements must first be determined using a threshold-based approach in terms of the short-time energy of the moving variance of CSI measurements, which is feasible.

The detection of breathing and heart beats is the core function of the system, and this is a CSI data process except large-scale body movements. For breathing detection, the Hampel filter is used to filter the outliers for each subcarrier and a moving average filter removes high-frequency noise to complete the data preprocessing. Based on the observation that CSI amplitudes of different subcarriers have different sensitivities to minute movements, the system uses a threshold based method that uses the variance of the CSI amplitude in a moving time window to quantify the sensitivity of a subcarrier to minute movements to select sensitive subcarriers in terms of threshold, which can decrease computational cost and noise interference. The feasibility for this method is based on subcarriers with higher variance being more sensitive to minute movements. In addition, a sinusoidal-like periodic changing pattern over time due to breathing indeed presents on CSI measurements. Thus, the breathing cycle E can be identified by a weighted mean of the peak-to-peak interval measurements from multiple subcarriers. Consequently, the breathing rate can be identified as $60/E$ beat per minute. For the detection of heart beats, after bandpass filtering with frequency range from 1 Hz to 1.33 Hz to eliminate breathing interference, the corresponding frequency of the maximum power, a noticeable peak of the average Power Spectral Density (PSD) of all selected subcarriers, is simply the heart rate.

For detecting the breathing patterns of two people in bed at the same time, two strong peaks present in the power spectral density of each subcarrier, and the system then uses a K-means clustering method to classify all peaks into two clusters based on two-dimensional features, including PSD amplitude and corresponding frequency. The average values of the frequencies in two clusters are identified as the breathing rates of these two people. In addition, to detect the heart rates of two people, the same method is used. Also, the system maps the detected breathing or heart rates to the corresponding individual, because the person close to the wireless link has larger impact on CSI changes and the stronger peak in PSD.

In the breathing rate estimation, the evaluation result for single person shows that more than 80% estimation errors are less than 0.5 bpm for all those three setups in

two apartments shown in Fig. 14. For the case of studying two persons in a bed using setup i in Bedroom 1, over 90% of estimation errors were less than 1 bpm. In heart rate estimation, about 57% of estimation errors are less than 2 bpm and over 90% of estimation errors are less than 4 bpm using setup i in Bedroom 1.

3.5 Indoor localization and tracking

In 2014, Wi-Track, designed by MIT Katabi et al.^[4] based on FMCW signal generation at the USRP N210 platform, can track the 3D motion of one person including 3D localization, coarse tracking of body parts, and identification of the direction of a pointing hand through walls and in the occlusive room.

By using the linear relationship characteristic among the FMCW carrier frequency, signal phase, time of flight, and the signal propagation distance between the antenna and the target object, as shown in Fig. 2, Wi-Track can map the information of the varying distance of the target relative to receiving antennas in the physical space to the signal phase for further data processing. In this scheme, the TOF that takes for the round trip of FMCW signal between the antenna and the reflecting body is the only measurement variable, and it is difficult to directly measure because of the speed of light transmission of wireless signals. Nonetheless, the TOF can be obtained by the FMCW technique because the frequency shift Δf between the transmitted and received signals is a function of both the slope of the sweep and the TOF, as follows:

$$\text{TOF} = \Delta f / \text{slope} \quad (14)$$

Wi-Track leverages the knowledge of the placement of the antennas placed in a “T” shape, as shown in Fig. 15a, where three directional antennas are placed horizontally for receiving and one directional antenna is placed vertically for transmitting. In this setting, where each transmit-receive antenna pair defines an ellipsoid

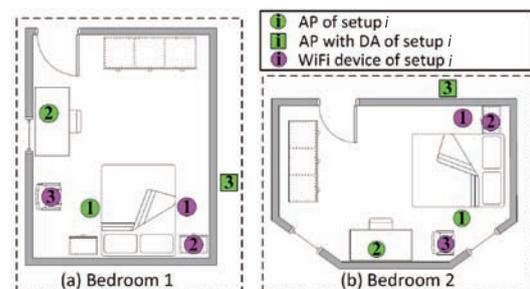


Fig. 14 Two apartments setup^[13]. (a) Evaluation scenarios in the Bedroom 1. (b) Evaluation scenarios in the Bedroom 2.

based on the round trip calculated by TOF, the intersection of three ellipsoids is insufficient to pinpoint the 3D location of a person, as shown in Fig. 15b.

In the experimental evaluation, the result of 3D Tracking shows that the median location errors along the x , y , and z dimensions are 9.9 cm, 8.6 cm, and 17.7 cm, respectively, for the LOS experiments and are 13.1 cm, 10.25 cm, and 21.0 cm in the through-wall experiments, respectively.

In 2015, Wi-Track 2.0, an updated version of Wi-Track also proposed by Katabi et al., was a multi-person localization system that can localize up to five people simultaneously with a median accuracy of 11.7 cm in each of the x and y dimensions and even localize static users by detecting the minute movements due to their breathing. Algorithmically, Wi-Track2.0 has two main components: (a) multi-shift FMCW, a technique that allows the management of multipath effects, (b) successive silhouette cancellation, an algorithm that allows Wi-Track2.0 to overcome the near-far problem for multi-person scenario.

However, there are still some limitations for these

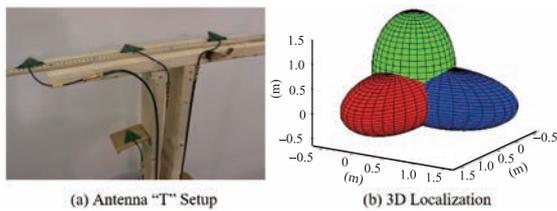


Fig. 15 Wi-Track's localization^[4].

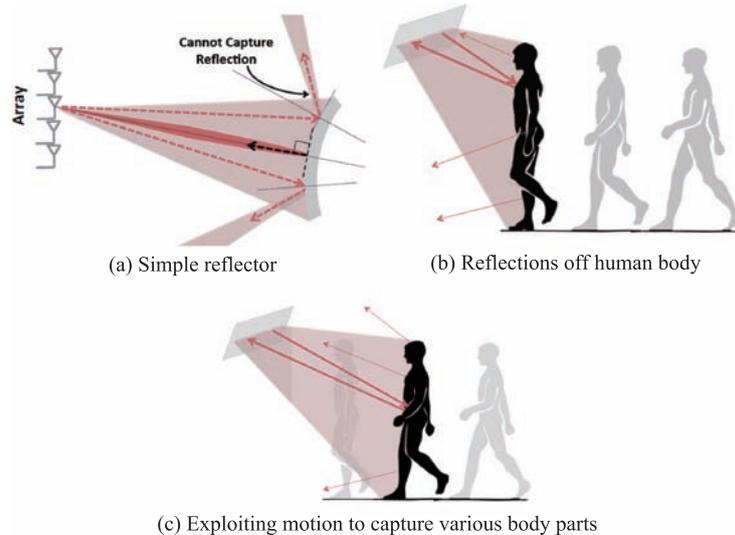


Fig. 16 RF reflections. (a) Only signals that fall along the normal to the surface are reflected toward the device. (b) The human body has a complex surface, but at any point in time only signals close to the normal to the surface are reflected toward the device. (c) As the person walks, different body parts reflect signals toward the device and become visible to the device^[7].

schemes. Wi-Track2.0 can only accurately track up to four moving users and five static users at a ten square meters coverage area, which has been better relative to Wi-Track. In addition, these schemes cannot yet identify the detected users.

3.6 Imaging

In 2015, RF-Capture, an imaging system proposed by Adib et al.^[7], can capture the human figure, i.e., a coarse skeleton extracted from the FMCW reflection signal in a through-wall scenario. In contrast to typical techniques for imaging humans such as visible light, x-ray, terahertz, and millimeter-wave, RF-Capture operates at lower frequencies from 5.46 GHz to 7.24 GHz and can work well in the NLOS scenario. The human body acts as a reflector rather than a scatterer at this frequency range. Although the antenna array receives reflections only from very few points on the surface of the user, these points vary among the entire body surface as the person moves, enabling RF-Capture to capture the human figure by combining the instantaneous RF reflections over consecutive time frames from various body parts. Figure 16 illustrates the discussed concept.

The core components of RF-Capture are the coarse-to-fine 3D scan and the motion-based figure capture.

(1) Coarse-to-fine 3D scan

With the combination of a 2D antenna array and FMCW chirps, RF-Capture can calculate the power of each voxel with spherical coordinate (r, θ, ϕ) in 3D space as shown in Eq. (15). N and M , respectively, represent

the number of receive antennas and transmit antennas, and $s_{n,m,t}$ represents the signal received by receive antenna n from transmit antenna m at time t . Thus, the direction of the high reflection power regions where the object exists will be recursively refined more and more precisely as the number of antennas increases up to the limit, which exploits an intrinsic property of antenna arrays, namely, the larger an array is, the narrower its beam, and the finer its spatial resolution. Similarly, the space depth of reflection source can be recursively refined by gradually increasing the amount of bandwidth which utilizes the principle that the depth resolution of FMCW is inversely proportional to the bandwidth of the signal. Finally, RF-Capture can generate one 3D snapshot of the reflection power from body parts every 75 ms according to the above method with low computational complexity.

(2) Motion-based figure capture

Each 3D snapshot needs to be calibrated through compensation to combine the information across consecutive snapshots to capture a clear human figure. Due to the conical beam making the image of the distant object blurrier, RF-Capture compensates for depth-related distortion by deconvolving the power in each snapshot using the Lucy-Richardson method^[31] for the compensation for depth. For the compensation for swaying, RF-Capture aligns each snapshot from the phase where the user walks straight by considering the chest position with the highest reflection power in the snapshot as the pivot and then automatically segments the remainder of the snapshot into eight regions corresponding

to different body parts respectively, as shown in Fig. 17a.

Because the structures of the torso and head do not suffer deformations significantly as the subject moves after compensating for calibration, the system calculates the sum of each torso and head regions correspondingly and selects the highest SNR segments of limb regions across the consecutive snapshots to generate a more complete human figure. Finally, the Alpha-Blending algorithm is performed to obtain a smooth figure as shown in Fig. 17b.

RF-Capture has realized two sets of functions, and it can capture the human figure through walls and distinguish various users. The other function is to identify and track the trajectory of certain body parts through walls.

3.7 Emotional recognition

In 2016, EQ-Radio, an emotion recognition system designed by Zhao et al.^[15], can infer the emotions of a person by extracting the individual heartbeats from the FMCW signal on the same hardware platform with Wi-Track. In addition, the minute variations in each individual beat length are of more relevance to emotion recognition than the periodicity^[32–34]. Therefore, the heartbeat extraction algorithm is the core of the system and includes the removal of the large order of magnitude of signal interference from breathing, the definition of the boundary of the single-cycle heartbeat, and the difference of tens of milliseconds between the heartbeat waves in each cycle.

According to the intuition that the acceleration of breathing is smaller than that of heartbeats, EQ-Radio can

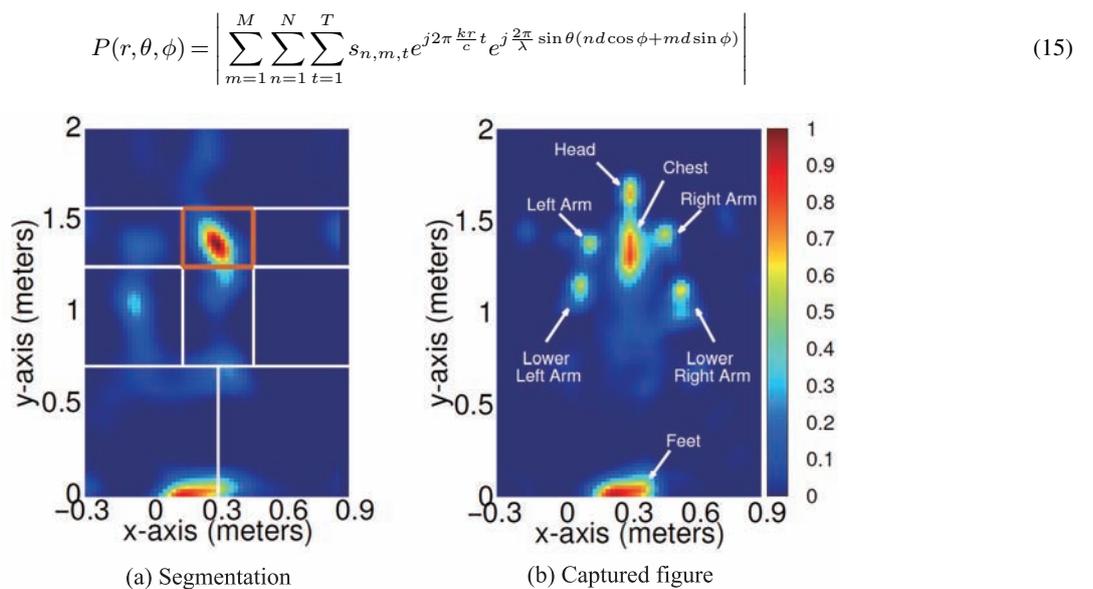


Fig. 17 Body part segmentation and captured figure. (a) The different regions used to identify body parts. (b) The captured figure^[7].

simply operate on the second derivative of the phase signal that is equivalent to the compound displacement caused by chest expansion and contraction due to breathing and body vibration due to heartbeats to mitigate the impact of breathing. Figure 18 demonstrates that although breathing is more pronounced than heart beats in the RF phase, there is a periodic pattern corresponding to each heartbeat cycle, while the breathing effect is negligible in the acceleration signal.

Based on the heartbeat segmentation algorithm designed by Zhao et al.^[15], EQ-Radio extracts individual beat segments to obtain the minute variations between the individual heartbeats based on the situation that the morphology of the heartbeat in the acceleration signal that is used to bootstrap this segmentation process is unknown. The algorithm can achieve global optimization in linear time complexity and return the optimal heartbeat template μ and segmentation set S containing all heartbeat segments, whose average error is only 3.2 ms in inter-beat-intervals with less than 0.4% of the average heartbeat cycle. Figure 19 shows the final beat segmentation for the data of Fig. 18 and the contrast with ECG.

In all optional 27 features^[15, 35, 36], EQ-Radio selects emotion-dependent features in each heartbeat cycle by using an embedded method that a class of feature selection mechanisms can pick out features that best contribute to the accuracy of the model while training the model^[37].

Then EQ-Radio uses l_1 -SVM to train a classifier based on selected features. In the evaluation, EQ-Radio, a commercial ECG monitor, and the Microsoft image-based emotion recognition system are simultaneously used to

monitor the emotion of each participant who elicits a certain emotional state using the prepared stimuli, e.g., personal memories, music, photos, and videos. The result shows that if EQ-Radio is separately trained on each subject, the accuracy of emotion classification is 87%, and if one classifier is used for all subjects, the accuracy is 72.3%. EQ-Radio's emotion recognition is on par with the state-of-the-art ECG-based system.

4 General Application Areas

The future development of the wireless sensing system truly makes us see the world through wireless signals. As early as 2015, the RF-Capture system, which was proposed by the MIT Katabi et al., was able to image the human body moving behind the wall. In addition, in early May 2017, Holl and Reinhard^[38] at the Technical University of Munich, Germany, demonstrated a holographic scheme to acquire three-dimensional images of building interiors from the radiation of an unmodified commercial narrow band Wi-Fi router that numerically simulates the hologram of a 10-m-sized building, finding that both localization of emitters and 3D tomography of absorptive objects could be feasible by this technique. Although these technologies are still not mature with the preliminary stage, it is sufficient to confirm that the application perspective of wireless sensing systems will be more widespread in the future. The current applications mainly include smart home, medical health care, search-and-rescue, and security.

4.1 Smart home

The concept of smart home has gradually become the indispensable vision of the future life, and the essence of which is a novel application scenario of human-computer interaction technology. The wireless sensing system that has been proposed thus far provides us an alternative implementation scheme in smart homes such as gesture recognition systems^[8, 12, 39, 40], gait recognition systems^[24, 28], emotion recognition systems^[15], and vital signs detection systems^[13, 17], all of which can be achieved through the ubiquitous Wi-Fi in the home without additional special equipment cost. The different components can still work properly with passive identification and non-invasive monitoring even in the NLOS and dark condition, and these benefits are advantages of wireless sensing systems for smart home.

Under the help of wireless sensing systems, any smart device such as TV, sound system, lighting system, and temperature adjustment system can be controlled by our gestures anywhere or anytime in a smart home. For

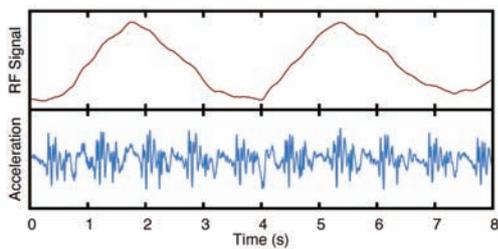


Fig. 18 RF signal and estimated acceleration^[15].

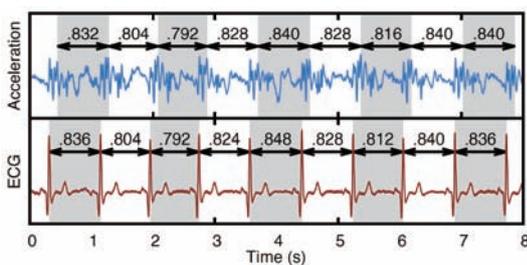


Fig. 19 Segmentation result compared to ECG^[15].

example, the ambience is automatically adjusted by the emotions of the user by easing his or her mood by playing soothing music or using soft lighting. Whether a home invasion happens can be also monitored through relying on the identification function of gait recognition systems. The wireless sensing system is the sensory system of smart home, making its interaction with human beings more intelligent and convenient.

4.2 Medical health care

Medical health monitoring is no longer confined to public medical institutions such as hospitals. Heart rate and blood pressure monitoring devices have become the daily necessities for many families, meaning that health awareness of people is gradually increasing. The monitoring devices will be possibly replaced by wireless sensing systems to make medical health monitoring more convenient for the public in the future. For families, the wireless sensing system can provide all family members vital signs monitoring such as breathing and heart rate within real-time, passive, non-invasive, and high precision. The collected data is important, as it can help to assess the general physical health of a person and provide useful clues for diagnosing possible diseases. In addition, falling is one of the major health threats and obstacles to independent living of elders, and falling aggravates the global pressure in the health care and injury rescue of elders^[6]. Studies have shown that the medical outcome of falling is largely dependent on the response and rescue time. The delay of medical treatment after a fall can increase the mortality risk in some clinical conditions, especially for the elders who live alone. For example, falls among older adults costed the U.S. health care system over \$19 billion dollars in 2000, and the number increased to \$30 billion dollars in 2010^[6]. This is sufficient to demonstrate the far-reaching implications of widely deployed wireless sensing systems for medical health monitoring.

For public medical institutions, wireless sensing systems can achieve widespread deployment, and the construction of a medical network based on these systems will improve the real-time monitoring on residents health status and enable rapid medical treatment response to emergencies. The wireless sensing technique is suitable for the recent emergence concept of the intelligent endowment system whether applied to detect falling or monitoring vital signs for old people in China. In particular, it is easier in some places to offer both real-time monitoring while ensuring that privacy is not violated.

4.3 Search-and-rescue

After a natural disaster such as a fire or earthquake, the disaster scene environment is so extremely complex to severely obstruct the progress of search and rescue. Consequently, quickly and accurately exploring the life signs and identifying the location of the wounded can save more lives in the ruins. The accurate vital signs monitoring and positioning function of the wireless sensing system can help rescuers find survivors more efficiently perform rescue work. For example, Vital-Radio, as the representative of the vital signs monitoring system proposed by Adib et al.^[17], can help search and rescue survivors in the ruins be found by their ability to search through walls. The system can also help searchers quickly determine the rubble area where some people are still alive, taking much less time than blindly searching.

However, the presented systems are just prototypes for search-and-rescue in experimental stages. To achieve this goal, the system must have stronger robustness to a complex open space environment while ensuring that the detection accuracy is determined..

4.4 Security

For the monitoring protection of office buildings and homes, the camera is a universal monitoring equipment. However, the quality of monitoring videos is affected by many environmental factors such as light intensity, distance of sight, and the blind spot area, while the wireless signal does not have these environmental factor restrictions. Current home camera surveillance systems do not have enough safety, because they can be hacked and cause private surveillance video data leakage. In contrast, the wireless sensing system can also realize identification and human activities recognition to perform security detection functions such as detecting illegal invasions in the home; consequently, it can work in some private spaces to accommodate for the camera deficiencies in the future. It plays an important role in large-scale dark environments such as subway tunnels and an underground pipe gallery to constantly monitor whether there is an uninvited guest.

RF-Capture can imagine the human body behind the wall using RF signals. Specially, it provides a more effective technology method that helps the police understand the distribution of criminals behind the wall to rescue hostages in anti-terrorism actions.

5 Conclusion

Based on the analysis in this work, all implementation schemes of the wireless sensing technology are based

on the basic principle that the wireless signal carries rich information about human activities due to signal propagation that is affected by human actions in the wireless signal coverage space. Therefore, how the wireless sensing technology could bring effective technical solutions to many fields in which there are key common points that are related to human action recognition was determined. However, all prototypes and results are only in the laboratory test stage with only the initial completion of the conceptual model in a small-scale controllable space. To put it into a complex practical application environment to use, the following categories of problems must be addressed:

- (1) Although they have been proven to have satisfactory accuracy and robustness, the current wireless sensing systems are only directed at experimental results in a limited controlled space. Experimental results demonstrate that the placement scheme of detection equipment, the relative distance between the detected users and devices, the multi-person complex environment, and the through-wall detection all seriously impact the accuracy of the system. Therefore, the first issue that must be solved is that wireless sensing systems are required to have the robustness and reliability in complex real-world environments while ensuring accuracy.
- (2) The innovation of the wireless sensing system based on the CSI helps realize the recognition perception function using the ubiquitous Wi-Fi signal. However, the IEEE 802.11n standard Wi-Fi signal is the only current option for these schemes. As mentioned, more advanced wireless technology will optimize and improve the performance of wireless sensing systems. Therefore, the second issue that must be solved is that this technology must consider the compatibility issues of protocol standards for Wi-Fi signals such as the IEEE 802.11ac and 802.11ah standards to make fully use their performance advantages.

With the rapid development of wireless technology, the ubiquitous Wi-Fi is quietly changing. Both transmission rate and signal bandwidth of modern mainstream Wi-Fi devices that comply with the IEEE 802.11n standard in 2.4-GHz and 5-GHz bands with MIMO and OFDM technology have been greatly improved. As a successor to the IEEE 802.11n standard, the IEEE 802.11ac standard in the 5-GHz band is emerging with its higher transmission speed and wider bandwidth. Meanwhile, in early 2016, the Wi-Fi Alliance released the IEEE 802.11 ah standard,

known as HaLow, and the better diffraction capability makes the transmission distance twice the current standard with the lower power consumption and the better through-wall ability in the 900 MHz band. Therefore, to improve the advanced wireless technology, the performance of the wireless sensing system will be further improved in the next research phase.

Although wireless sensing technology is still in the laboratory research stage, in terms of the results obtained in this field, the wireless sensing system has a great potential application in the context of artificial intelligence. As wireless technology, digital signal processing technology, and machine learning technology are continuously improved, more novel and valuable applications in many commercial areas will be created.

Acknowledgment

This work was supported in part by the National Natural Science Foundation of China under Key Program of NSFC (No. 61332019), NSFC (Nos. 61572304 and 61272056); Shanghai Key Laboratory of Specialty Fiber Optics and Optical Access Networks (No. SKLSFO2014-06).

References

- [1] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, Understanding and modeling of wi-fi signal based human activity recognition, in *Proc. ACM Mobi Com*, Paris, France, 2015, pp. 65–76.
- [2] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu, E-eyes: Device-free location-oriented activity identification using fine-grained wi-fi signatures, in *Proc. ACM Mobi Com*, Maui, HI, USA, 2014, pp. 617–628.
- [3] Enhancements for Higher Throughput, IEEE Standard 802.11n, 2009.
- [4] F. Adib, Z. Kabelac, D. Katabi, and R. C. Miller, 3D tracking via body radio reflections, in *Proc. Usenix NSDI*, Seattle, WA, USA, 2014, pp. 317–329.
- [5] C. S. Wu, Z. Yang, Z. Zhou, X. F. Liu, Y. H. Liu, and J. N. Cao, Non-invasive detection of moving and stationary human with wi-fi, *IEEE J. Sel. Area. Comm.*, vol. 33, no. 11, pp. 2329–2342, 2015.
- [6] C. M. Han, K. S. Wu, Y. X. Wang, and L. M. Ni, WiFall: Device-free fall detection by wireless networks, in *Proc. IEEE INFOCOM*, Toronto, Canada, 2014, pp. 271–279.
- [7] F. Adib, C. Y. Hsu, H. Z. Mao, D. Katabi, and F. Durand, Capturing the human figure through a wall, in *Proc. ACM MobiCom*, Paris, France, 2015, pp. 1–13.
- [8] Q. F. Pu, S. Gupta, S. Gollakota, and S. Patel, Whole-home gesture recognition using wireless signals, in *Proc. ACM SIGCOMM*, Hong Kong, China, 2013, pp. 485–486.

- [9] F. Adib and D. Katabi, See through walls with wi-fi, in *Proc. ACM SIGCOMM*, Hong Kong, China, 2013, pp. 75–86.
- [10] D. Halperin, W. J. Hu, A. Sheth, and D. Wetherall, 802.11 with multiple antennas for dummies, in *Proc. ACM SIGCOMM*, New Delhi, India, 2010, pp. 19–25.
- [11] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, Predictable 802.11 packet delivery from wireless channel measurements, in *Proc. ACM SIGCOMM*, New Delhi, India, 2010, pp. 159–170.
- [12] H. Li, W. Yang, J. X. Wang, Y. Xu, and L. S. Huang, Wi-Finger: Talk to your smart devices with finger-grained gesture, in *Proc. ACM UbiComp*, Heidelberg, Germany, Sep. 2016, pp. 250–261.
- [13] J. Liu, Y. Wang, Y. Y. Chen, J. Yang, X. Chen, and J. Cheng, Tracking vital signs during sleep leveraging off-the-shelf wi-fi, in *Proc. ACM MobiHoc*, Hangzhou, China, 2015, pp. 267–276.
- [14] F. Adib, Z. Kabelac, and D. Katabi, Multi-person localization via RF body reflections, in *Proc. Usenix NSDI*, Oakland, CA, USA, 2015, pp. 279–292.
- [15] M. M. Zhao, F. Adib, and D. Katabi, Emotion recognition using wireless signals, in *Proc. ACM MobiCom*, New York, NY, USA, 2016, pp. 95–108.
- [16] Z. Yang and Y. H. Liu, Wi-Fi radar: From RSSI to CSI, *CCF*, vol. 10, no. 11, pp. 55–60, 2014.
- [17] F. Adib, H. Z. Mao, Z. Kabelac, D. Katabi, and R. C. Miller, Smart homes that monitor breathing and heart rate, in *Proc. ACM CHI*, Seoul, Korea, 2015, pp. 837–846.
- [18] K. Ali, A. X. Liu, W. Wang, and M. Shahzad, Keystroke recognition using wi-fi signals, in *Proc. ACM MobiCom*, Paris, France, 2015, pp. 90–102.
- [19] L. Davies and U. Gather, The identification of multiple outliers, *J. Am. Stat. Assoc.*, vol. 88, no. 423, pp. 782–792, 1993.
- [20] E. S. Gardner, Exponential smoothing: The state of the art, *J. Forecasting*, vol. 4, no. 1, pp. 1–28, 1985.
- [21] Y. Rubner, Perceptual metrics for image database navigation, Ph.D. dissertation, Dept. S. U. C. S., Stanford Univ., Stanford, CA, USA, 1999.
- [22] G. T. Holt, M. Reinders, and E. Hendriks, Multi-dimensional dynamic time warping for gesture recognition, in *Proc. ASCI*, Heijen, Holland, Jun. 2007, pp. 1–8.
- [23] L. R. Welch, Hidden Markov models and the Baum-welch algorithm, *IEEE Information Theory Society Newsletter*, vol. 53, no. 4, pp. 10–13, 2003.
- [24] Y. Z. Zeng, P. H. Pathak, and P. Mohapatra, WiWho: wi-fi-based person identification in smart spaces, in *Proc. ACM IPSN*, Vienna, Austria, 2016.
- [25] J. Unar, W. C. Seng, and A. Abbasi, A review of biometric technology along with trends and prospects, *Pattern Recogn.*, vol. 47, no. 8, pp. 2673–2688, 2014.
- [26] T. T. Ngo, Y. Makihara, H. Nagahara, Y. Mukaigawa, and Y. Yagi, The largest inertial sensor-based gait database and performance evaluation of gait-based personal authentication, *Pattern Recogn.*, vol. 47, no. 1, pp. 228–237, 2014.
- [27] Y. Jin, W. S. Soh, and W. C. Wong, Indoor localization with channel impulse response based fingerprint and nonparametric regression, *IEEE Trans. Wirel. Commun.*, vol. 9, no. 3, pp. 1120–1127, 2010.
- [28] W. Wang, A. X. Liu, and M. Shahzad, Gait recognition using wi-fi signals, in *Proc. ACM UbiComp*, Heidelberg, Germany, Sep. 2016, pp. 363–373.
- [29] M. W. Whittle, Gait analysis: An introduction, *Physiotherapy*, vol. 77, no. 11, p. 786, 2003.
- [30] P. V. Dorp and F. Groen, Feature-based human motion parameter estimation with radar, *IET Radar Sonar Nav.*, vol. 2, no. 2, pp. 135–145, 2008.
- [31] L. B. Lucy, An iterative technique for the rectification of observed distributions, *Astron. J.*, vol. 79, no. 6, p. 745, 1974.
- [32] D. S. Quintana, A. J. Guastella, T. Outhred, I. B. Hickie, and A. H. Kemp, Heart rate variability is associated with emotion recognition: Direct evidence for a relationship between the autonomic nervous system and social cognition, *Int. J. Psychophysiol.*, vol. 86, no. 2, pp. 168–172, 2012.
- [33] R. D. Lane, K. McRae, E. M. Reiman, K. Chen, G. L. Ahern, and J. F. Thayer, Neural correlates of heart rate variability during emotion, *Neuroimage*, vol. 44, no. 1, pp. 213–222, 2009.
- [34] R. A. Calvo and S. D. Mello, Affect detection: An interdisciplinary review of models, methods, and their applications, *IEEE Trans. Affect. Comput.*, vol. 1, no. 1, pp. 18–37, 2010.
- [35] J. Kim and E. Andre, Emotion recognition based on physiological changes in music listening, *IEEE Trans. Pattern Anal.*, vol. 30, no. 12, pp. 2067–2083, 2008.
- [36] U. R. Acharya, K. P. Joseph, N. Kannathal, C. M. Lim, and J. S. Suri, Heart rate variability: A review, *Med. Biol. Eng. Comput.*, vol. 44, no. 12, pp. 1031–1051, 2006.
- [37] I. Guyon and A. Elisseeff, An introduction to variable and feature selection, *J. Mach. Learn. Res.*, vol. 3, no. 6, pp. 1157–1182, 2002.
- [38] P. M. Holl and F. Reinhard, Holography of wi-fi radiation, *Phys. Rev. Lett.*, vol. 118, no. 18, p. 183901, 2017.
- [39] H. Abdelnasser, M. Youssef, and K. A. Harras, Wigest: A ubiquitous wi-fi-based gesture recognition system, in *Proc. IEEE INFOCOM*, Kowloon, Hong Kong, China, 2015, pp. 1472–1480.
- [40] J. C. Shang and J. Wu, A robust sign language recognition system with sparsely labeled instances using Wi-Fi signals, in *Proceedings of IEEE 14th International Conference on Mobile Ad Hoc and Sensor Systems*, Orlando, FL, USA, 2017, pp. 99–107.



Chao Wang received the PhD degree from Tongji University in 1999. Currently, he is an IEEE senior member, council member of China Association of AI, Information Security Committee Vice Chair of China Electronics Institute, Committeeman of the Sixth Shanghai Expert Committee for Informatization,

Directorate of China Artificial Intelligence Institute, Committeeman of CCF, IEEE Shanghai Section Secretary, IEEE Shanghai CAS Chapter Vice Chair, and IEEE Shanghai Computer Chapter Vice Chair. His research interests include wireless sensor network, network information security, and ECC quantum computing cryptography.

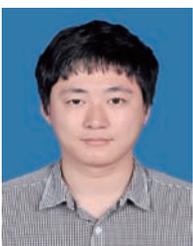


Siwen Chen received the B.S. degree from Shanghai University, Shanghai, in 2015. He is currently pursuing the M.S. degree in the School of Communication and Information Engineering from Shanghai University. His research interest is about wireless sensing technology.



Yanwei Yang received the B.S. degree in Communication Engineering from Huaihai Institute of Technology, China, in 2015 and the M.S. degree in Electronics and Communication Engineering from Shanghai University, Shanghai, China, in 2017. From 2015 to 2017, she was studying in the professor

Wang Chao's laboratory. Her researches are about the development of wireless and the emergency communications for small groups. She received the second-class graduate student scholarship.



Feng Hu received the B.S. degree from the Xi'an University of Science and Technology. Currently, he is a PhD student at Electronic and Information Engineering Dept. of Shanghai University. His research interests include information security and quantum computing cryptography.



Fugang Liu received the B.S. degree in computer science and technology from the Heilongjiang University of Science and Technology in 2004. He received the M.S. and PhD degrees in Communication and Information System in 2009 and 2013, respectively, from Harbin Engineering University. From 2004 to 2009, he was

an Assistant with the Heilongjiang University of Science and Technology, and from 2009 to 2013, he was a Lecture with there. Since 2013, he has been an Assistant Professor with the University. He is the author of three books, 12 articles and 15 inventions. His research interests include DOA Estimation of wideband Signals, D-InSAR technique, and array signals processing. Dr. Liu was awarded Science and Technology Prize of Coal Industry Association of China in 2014.



Jie Wu is the Chair and the Laura H. Carnell Professor at the Department of Computer and Information Sciences at Temple University. Prior to joining Temple University, USA, he was a program director at the National Science Foundation and a distinguished professor at Florida Atlantic University. He

received the PhD degree from Florida Atlantic University in 1989. His current research interests include mobile computing and wireless networks, routing protocols, cloud and green computing, network trust and security, and social network applications. He regularly publishes in scholarly journals, conference proceedings, and books. He serves on several editorial boards, including *IEEE Transactions on Computers*, *IEEE Transactions on Service Computing*, and *Journal of Parallel and Distributed Computing*. He was the general co-chair/chair for IEEE MASS 2006 and IEEE IPDPS 2008 and program co-chair for IEEE INFOCOM 2011. Currently, he is serving as a general chair for IEEE ICDCS 2013 and ACM Mobi Hoc 2014, and program chair for CCF CNCC 2013. He was an IEEE Computer Society Distinguished Visitor, ACM Distinguished Speaker, and chair for the IEEE Technical Committee on Distributed Processing (TCDP). He is a CCF Distinguished Speaker and a fellow of the IEEE. He is the recipient of the 2011 China Computer Federation (CCF) Overseas Outstanding Achievement Award.