Research Article

CSI-HC: A WiFi-Based Indoor Complex Human Motion Recognition Method

Zhanjun Hao [iD],[1,2] **Yu Duan** [iD],[1] **Xiaochao Dang** [iD],[1,2] and **Tong Zhang**[1]

[1]*College of Computer Science and Engineering, Northwest Normal University, Lanzhou, Gansu 730070, China*
[2]*Gansu Province Internet of Things Engineering Research Center, Northwest Normal University, Lanzhou, Gansu 730070, China*

Correspondence should be addressed to Zhanjun Hao; zhanjunhao@126.com

WiFi indoor personnel behavior recognition has become the core technology of wireless network perception. However, the existing human behavior recognition methods have great challenges in terms of detection accuracy, intrusion, and complexity of operations. In this paper, we firstly analyze and summarize the existing human motion recognition schemes, and due to the existence of the problems in them, we propose a noninvasive, highly robust complex human motion recognition scheme based on Channel State Information (CSI), that is, CSI-HC, and the traditional Chinese martial art XingYiQuan is verified as a complex motion background. CSI-HC is divided into two phases: offline and online. In the offline phase, the human motion data are collected on the commercial Atheros NIC and a powerful denoising method is constructed by using the Butterworth low-pass filter and wavelet function to filter the outliers in the motion data. Then, through Restricted Boltzmann Machine (RBM) training and classification, we establish offline fingerprint information. In the online phase, SoftMax regression is used to correct the RBM classification to process the motion data collected in real time and the processed real-time data are matched with the offline fingerprint information. On this basis, the recognition of a complex human motion is realized. Finally, through repeated experiments in three classical indoor scenes, the parameter setting and user diversity affecting the accuracy of motion recognition are analyzed and the robustness of CSI-HC is detected. In addition, the performance of the proposed method is compared with that of the existing motion recognition methods. The experimental results show that the average motion recognition rate of CSI-HC in three classic indoor scenes reaches 85.4%, in terms of motion complexity and indoor recognition accuracy. Compared with other algorithms, it has higher stability and robustness.

1. Introduction

Benefiting from the widespread deployment of the wireless communication infrastructure, human behavior recognition based on wireless communication network technology has become a core technology to promote various applications [1, 2]. Traditionally, in order to recognize human behavior, physical sensing devices (e.g., ultra-wideband (UWB), radio frequency identification (RFID), and acceleration sensors) are first required to be worn on the human body or deployed in the environment. On this basis, the information collected by these physical devices is read to facilitate the identification of the person's behavioral state. Although this traditional behavior recognition method has been widely used

and achieved good results, most of them require specific sensor equipment. In addition, WiFi-based behavior recognition overcomes the shortcomings of traditional methods; that is, it can automatically recognize human behavior without the user wearing a sensor or device and has been widely used in real life, including smart home, remote health care, campus security, severe illness patient care, and elderly activity detection [3].

Currently, WiFi-based human motion sensing technology has broad development prospects. Compared with the human behavior recognition of the motion sensor and optical camera, the advantage of WiFi-based human motion recognition is that the coverage of the WiFi signal is wide, there is no sense of dead angle, and there is no sense of light

requirement and obscurity. Not only can it run on the cheap commercial WiFi device, but also it has lower hardware cost and maintenance cost than the motion sensor and optical camera. The traditional indoor human behavior detection mainly depends on the received signal strength (RSS) method, but the experimental results show that the RSS method has poor motion recognition and low stability. Compared with RSS, CSI is a finer measure of the physical layer, which describes the amplitude attenuation and phase shift of a wireless signal, based on which it can effectively identify a variety of behaviors, from vital signs to basic behavior, as well as complex activities [4–7]. In reference [4], Liu developed a system to track vital signs such as the heart rate and respiration rate during sleep by analyzing CSI signals in real time. The Wi-Sleep system proposed in reference [5] can extract the respiratory information of users in various sleep positions by identifying the rhythm patterns associated with breathing. In reference [6], WiDraw is a hand motion tracking system, which uses the angle of arrival (AOA) value of the WiFi signal on the mobile device to track the hand motion trajectory. The WiWho framework proposed in reference [7] can use CSI for human gait recognition.

According to the different recognition algorithms and application scenes, a WiFi-based human motion recognition system can be divided into two categories: one is the model-based recognition system and the other is the fingerprint recognition system. The main difference between them is whether a priori learning is required [8]. The model-based recognition system can recognize human motion without training. For example, the Fresnel model of WiFi human recognition is established in reference [9]. However, the model-based recognition system needs to access a large number of access points (APs) to accurately identify human behavior, which leads to a significant increase in hardware cost and maintenance cost. Based on the fingerprint database recognition scheme, through offline phase training and online phase recognition for pattern matching, human motion recognition can be realized.

However, the existing WiFi-based human motion recognition scheme recognizes the action is relatively simple, the actual scene availability is not strong, or it is a daily behavior such as walking, opening the door, sleeping, or a single human body standing, picking up, and sitting down. Some complex motions cannot be accurately identified, and the application scenarios are relatively simple, and the existing application scenarios are diversified (special scene detection such as prison hospitals, wildlife behavior detection, indoor elderly activity detection, and construction site safety detection). Therefore, it is imperative to find a complex motion recognition scheme. Considering the application scene in real life, we use WiFi to identify the Chinese traditional martial art XingYiQuan motion in the indoor environment. We guide users to carry out correct fitness motions, so as to maintain human health.

The main contributions of our work are as follows:

(1) In this paper, a complex human motion recognition scheme CSI-HC based on WiFi is proposed and verified with the background of the Chinese traditional martial art XingYiQuan. Its advantage is that the detection motion is relatively complex and does not need human wearing equipment. It can work effectively on cheap commercial devices and is of great value in guiding and monitoring human health campaigns.

(2) CSI-HC uses the amplitude of the signal received by the receiver array antenna to establish a mapping relationship with the different actions of the human body. It uses the Butterworth low-pass filter and Sym8 wavelet function to construct a powerful denoising method to filter outliers in motion data. The offline fingerprint construction is completed through RBM training. In the online phase, SoftMax regression is used to modify the RBM classification to accurately sense the complex motions of different human bodies.

(3) We analyze the key factors that affect the perceptual effect, such as the distance between transceiver devices and transmitter contracting rate, find the appropriate parameter settings, and explore the impact of user diversity on system performance through experiments.

(4) In three scenarios where the multipath effect ranges from weak to strong (meeting room, corridor, and office), the performance of CSI-HC is tested. The experimental results show that CSI-HC has high robustness, and in three different scenes, the accuracy of the XingYiQuan motion can be more than 85%.

The main contents and organizational structure of this paper are as follows: We first introduce the related work in Section 2, and then we describe the proposed research methodology in Section 3. The specific implementation process of the method is provided in Section 4. Section 5 explores parameter settings and demonstrates performance. Finally, we summarize all the work of this paper in Section 6.

2. Related Works

In this section, we introduce the advantages and disadvantages of existing motion recognition works from two perspectives, that is, device-based and device-free.

2.1. Device-Based Human Motion Recognition. As we all know, most human perception systems need additional hardware support to complete the recognition of related motions. These hardware devices are divided into four categories: special sensors, infrared devices, optical cameras, and smartphones. The special sensor realizes the perception of different actions by collecting the relevant physical information of human body movements. Skinput [10] uses a wearable bioacoustic sensor array to analyze the mechanical vibrations that travel in the body to identify the movement of the arms and fingers. FEMO [11] is a human motion detection system based on RFID, which uses the backscattering signal of the passive RFID tag installed on the training

equipment to detect the motion state of the user. eFisio-Track [12] is a telemedicine assistance system that detects patients' rehabilitation training movements by using the accelerometer equipment. The GrandCare [13] system detects patient behavior by calling a motion sensor installed on the door. Although the special sensor can realize the high-precision perception of the fine-grained movement of the human body, users need to wear sensors or deploy special equipment, and it is difficult to carry and invade the privacy of users. The infrared device images the human body through infrared rays to realize the perception of an independent light source. The representative product is Microsoft's Kinect [14]. Infrared rays have a limited detection range due to problems such as their frequency bands and transmission distances and require expensive additional equipment, and it is difficult to achieve large-scale deployment. The optical camera captures the image sequence of the human motion through the camera, analyzes the human motion characteristics and the motion trajectory in the image sequence, and senses the state of the human body. The method based on optical cameras can be used in many scenes, such as gesture recognition [15], gait recognition [16], and target tracking [17]. However, this method still has shortcomings, and it cannot work in low-light conditions and where privacy is involved. Smartphones use the built-in sensors (accelerometer, gyroscope, magnetometer, and pressure gauge) to detect human activities. They have the advantages of being easy to use and not disturbing the user's normal activities. Smartphones also have some research results in detecting human motions. Gu et al. [18] combined the accelerometer and pressure gauge in a smartphone to monitor 7 different states of motion. PerFallD [19] uses the accelerometer in the smartphone to detect the fall of the human body. Unfortunately, although smartphones have become popular and have many advantages, they are not applicable in some scenarios. In particular, when it comes to detecting falls among the elderly, it is often impractical to have them carry around smartphones.

2.2. Device-Free Human Motion Recognition

2.2.1. RSS-Based Human Motion Recognition.
In the past few decades, due to the rapid development of sensing technology, device-based human perception has been widely used in daily life. However, device-based perception requires special equipment because of its high hardware and maintenance costs. It is difficult to deploy effectively on a large scale. To solve this problem, researchers have begun to focus on device-free human perception technology, that is, the detection of human behavior without the need for the human body to wear any physical device [20]. The widespread deployment of wireless networks makes it possible to realize device-free human perception based on WiFi signals. Seifeldin et al. [21] realized simple human motion detection by analyzing RSS changes in WiFi signals caused by the human motion, and Sigg et al. [22] used a software-defined radio to transmit RF signals and to determine the human motion based on changes in the received RSS. With the maturity of this technology, the human body motion that the WiFi signal can detect is more and more detailed. Wi-Vi [23] and WiSee [24] systems use the WiFi signal to realize gesture recognition. Although RSS-based technology has made great progress, the essence of RSS is to reflect the strength of signal reception, especially susceptible to multipath effects and narrowband interference, resulting in its own flaws with low accuracy. In order to overcome the inherent defects of RSS, a finer granularity WiFi channel feature CSI based on the physical layer is discovered. Compared with RSS, CSI can estimate the channel characteristic information on each subcarrier by orthogonal frequency-division multiplexing (OFDM) technology. The frequency attenuation variation of the WiFi channel can be better described to reduce the influence of multipath components and narrowband interference. In addition, CSI contains rich frequency-domain information such as amplitude and phase, which can better reflect the influence of the human body on the WiFi signal to achieve higher sensing accuracy.

2.2.2. CSI-Based Human Motion Recognition.
The research of human motion perception based on CSI originated from a tool, CSI Tool [25], which can obtain CSI information from the commercial WiFi network card. CSI Tool greatly facilitates the extraction of CSI sensing data so that the use of more fine-grained CSI signals for sensing has become a new trend. Since CSI has excellent characteristics of being relatively stable in an indoor environment and sensitive to human motions, CSI signals are widely used in various motion recognition scenarios, and FIMD [26] attempts to use CSI to detect user location changes without actively carrying any physical devices. DeMan [27] is a noninvasive detection scheme which can judge the motion and static state of the human body. The motion state is judged by amplitude combined with phase information, and the static state of the human body is judged by the periodic model caused by human respiration in the wireless signal. R-PMD [28] is a passive motion detection scheme that uses PCA to extract the covariance of CSI data as a motion feature, and the human motion state is judged by the mapping between covariance and human motion. CRAM [29] constructed a CSI speed model, which correlates the movement speed of different parts of the human body with the dynamic changes of CSI and detects specific human activities such as walking, running, and sitting down. RT-Fall [30] is a device-free fall detection system based on CSI data, which can effectively identify the normal walking state and abnormal fall state of the human body. Wi-Finger [31] recognizes finger gestures by establishing user gestures and CSI signal changes caused by different gestures (for example, numbers 1 to 9 in ASL). E-eye [32] is a device-independent activity detection system, which distinguishes a large number of daily activities by matching the measured values of CSI with known features and realizes the recognition of different actions such as sleeping, cooking, and watching TV.

3. Preliminaries and Program

In this section, we introduce the WiFi-based recognition model and the reasons why CSI can perform human motion perception and give the specific solution and the overview of this solution.

3.1. Research Theory

3.1.1. WiFi-Based Recognition Model. The principle of human motion perception of WiFi signals is depicted in Figure 1. When a person is in a signal link, the propagation of wireless signals will be reflected, scattered, and diffracted by the influence of the human body. The signal received at the receiving end is a composite signal that is propagated by the direct path and the human body reflection path as well as the reflection path of the floor ceiling. The influence of the human body on the propagation of the WiFi signal will be characterized by the wireless signal arriving at the receiving end. Assume that the line of sight (LOS) length from the transmitter to the receiver is d, then the distance between the reflection point of the ceiling and the floor and the LOS is R. Combining the Friis free-space propagation equations and signals with the reflection scattering generated by the human body, the impact of the human body on wireless signal propagation can be defined as [33]

$$P_{\text{rx}}(d) = \frac{P_{\text{tx}} G_{\text{tx}} G_{\text{rx}} \lambda^2}{(4\pi)^2 (d + 4R + \varepsilon)^2}, \qquad (1)$$

where P_{tx} is the transmitting power of the transmitting end, $P_{\text{rx}}(d)$ is the receiving power of the receiving end, G_{tx} is the transmitting gain, G_{rx} is the receiving gain, λ is the wavelength of the WiFi signal, and ε is the approximate change of the path length caused by the scattering of the signal by the human body. Because the signal scattering paths caused by different actions are different, according to the above formula, different human actions will cause the difference in receiver receiving power, and by establishing the mapping relationship between these differences and different human actions, it lays the basic idea of WiFi human motion perception.

3.1.2. Channel State Information. The CSI signal can achieve universal, low-cost, fine-grained human perception, and there are three reasons for this: (1) The maturity of WiFi technology and the widespread use of WiFi devices. (2) CSI data can be easily extracted from commercial WiFi devices using tools released by Halperin. (3) The WiFi signal transmits a modulation scheme using OFDM under the IEEE 802.11N protocol, and OFDM can encode the CSI data to a plurality of subcarriers of different frequencies. Therefore, the radio channel information at the subcarrier level can be obtained from the original CSI measured in the WiFi data link. In OFDM transmission systems, it is assumed that a general model of channel state information can be represented as

$$\overrightarrow{Y} = H\overrightarrow{X} + \text{noise}, \qquad (2)$$
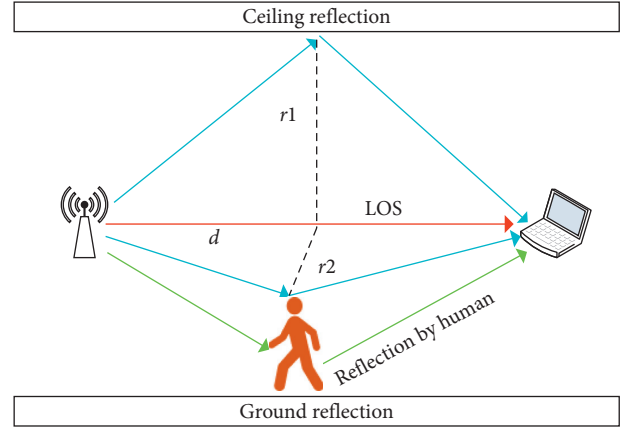
Figure 1: WiFi human motion recognition model.

where $\overrightarrow{Y}$ and $\overrightarrow{X}$ represent the received signal vector and the transmitted signal vector, respectively; noise is additive white Gaussian noise; and H is the channel impulse response (CIR) complex matrix in the CSI frequency domain, reflecting the channel gain information at the subcarrier level. Assuming that we obtain the measured CSI values H of the 2×3 frames received by the Atheros AR9380 NIC (that is, 2 transmit antennas and 3 receiving antennas) under the condition that the channel bandwidth is 20 MHz and the time is T, we can obtain the CSI values of 336 subcarriers:

$$H = [H(f_1), \ldots, H(f_N), \ldots, H(f_{336})]^T, \quad 1 \le N \le 336. \qquad (3)$$

Therefore, the CIR can characterize the frequency-domain information of each different subcarrier, and the i-th subcarrier can be defined as

$$H(i) = |H(i)|e^{j\sin \angle H(i)}, \qquad (4)$$

where $|H(i)|$ represents the amplitude value of the i-th subcarrier and $\angle H(i)$ represents the phase value of the i-th subcarrier.

3.2. Research Program

3.2.1. Specific Solution. CSI-HC proposed in this paper is a scheme to identify complex human motions. Specifically, CSI-HC uses CSI amplitude characteristics to identify human motions on commercial WiFi devices, filters environmental noise through low-pass filtering and wavelet function, and uses RBM and SoftMax machine learning classification algorithms to identify human motions. Taking the ancient Chinese martial art Xingyi boxing as the background of action recognition, XingYiQuan is a traditional Chinese fitness martial art, which is composed of six motions, such as QiShi, BengQuan, HuQuan, MaXingQuan, ZuanQuan, and ShouShi. Each action consists of different combinations of hands, arms, head, torso, and legs. We collect the data of six XingYiQuan motions, as shown in Figure 2, and show that these motions correspond to the characteristics of CSI signals generated in the frequency
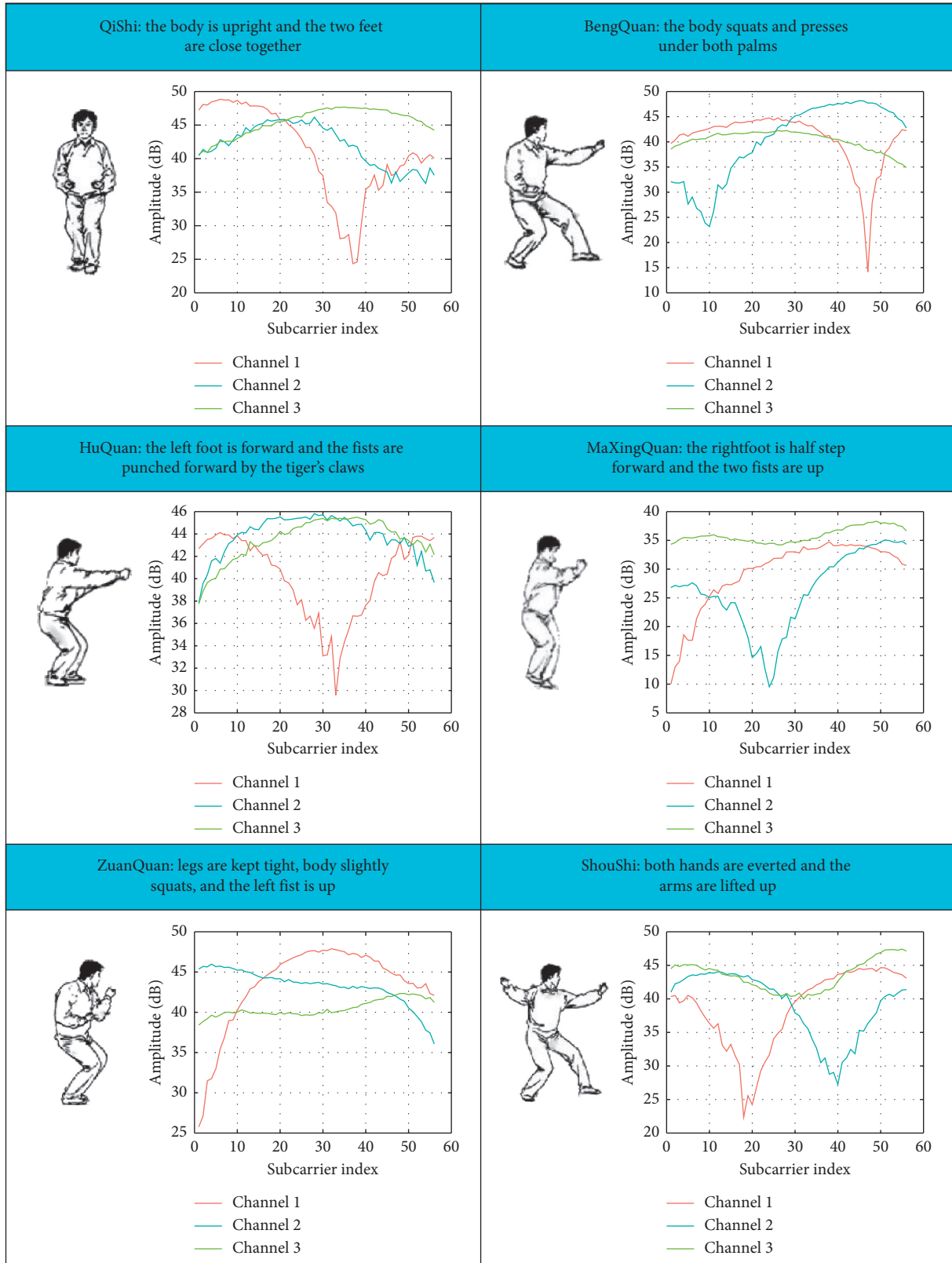
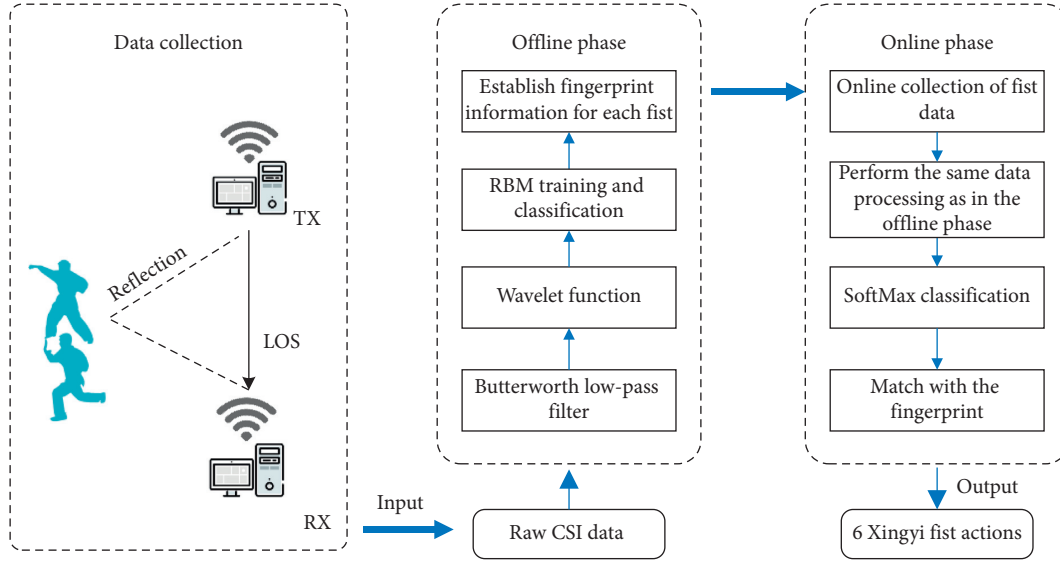FIGURE 2: XingYiQuan motion and the corresponding amplitude feature.

FIGURE 3: CSI-HC system flow chart.

domain, from which we can see that there are differences in CSI amplitude characteristics between different motions. By using these differences to establish the mapping with the action and carry on the pattern recognition, the concrete XingYiQuan motions can be recognized.

3.2.2. Overview. We describe in detail the overall architecture of the CSI-HC system, which consists of three phases: the data collection phase, the offline motion fingerprint establishment phase, and the online motion recognition phase, as shown in Figure 3. In the data collection phase, the WiFi device with the Atheros AR9380 NIC is used to collect CSI data of human motion perception in a variety of scenarios.

In order to remove the influence of the multipath effect and environmental noise, we filter the outliers of the data obtained in the data collection phase. The Butterworth low-pass filter is used to remove the high-frequency outliers and then combined with the wavelet function of the Sym8 wave base to filter the low-frequency outliers. Then, we get the CSI data which can better reflect the XingYiQuan motions, use these data as the sample set of RBM training, adjust the learning parameters, classify the data, and construct the standard fingerprint information of each XingYiQuan motion. In the online motion recognition phase, the same exception handling and machine learning classification methods as in the offline phase are adopted, the data consistency is maintained, and the RBM classification results are corrected using the SoftMax classifier. Then, it matches the data with the fingerprint database of the constructed motion and recognizes the specific XingYiQuan motion.

4. Methodology

In this section, we mainly describe in detail the two core data processing processes in the CSI-HC method: (1) Effective outlier filtering is performed on the collected CSI data using a Butterworth low-pass filter and wavelet function. (2) By constructing the RBM training model, the CSI data of complex motions are classified and the RBM classification result is modified by SoftMax to achieve higher precision motion recognition.

4.1. CSI Outlier Filtering. When using CSI data as a feature of motion perception, filtering outliers due to environmental noise interference and multipath components is critical to the accuracy of motion perception. In this paper, the amplitude of the acquired CSI data is taken as the feature value. Figure 4 shows the original CSI amplitude image of XingYiQuan, in which it can be seen that there are many abnormal values. These outliers may reduce the accuracy of the recognition of the motion method. For this reason, the Butterworth low-pass filter is used to filter the high-frequency interference in the outliers, and Sym8 is used as the wave-based wavelet function to filter the low-frequency interference in the outliers. The collected original CSI amplitude data are filtered by the outliers to preserve the signal feature integrity to the greatest extent, so as to establish the feature fingerprint information of each XingYiQuan motion.

4.1.1. Filtering Effect Evaluation. As shown in Figure 5, we selected four different filters to filter the abnormal data in Figure 4(a) to evaluate the performance of different filters. Here, Figure 5(a) shows abnormal data, which we use as the original input data of various filters, Figure 5(b) shows the data processed by the mean filter, Figure 5(c) shows the data processed by the Butterworth low-pass filter, Figure 5(d) shows the data processed by the median filter, and Figure 5(e) shows the data processed by the threshold filter.

By comparing the processing effects of four different filters, we can find that the processing effect of the mean filter
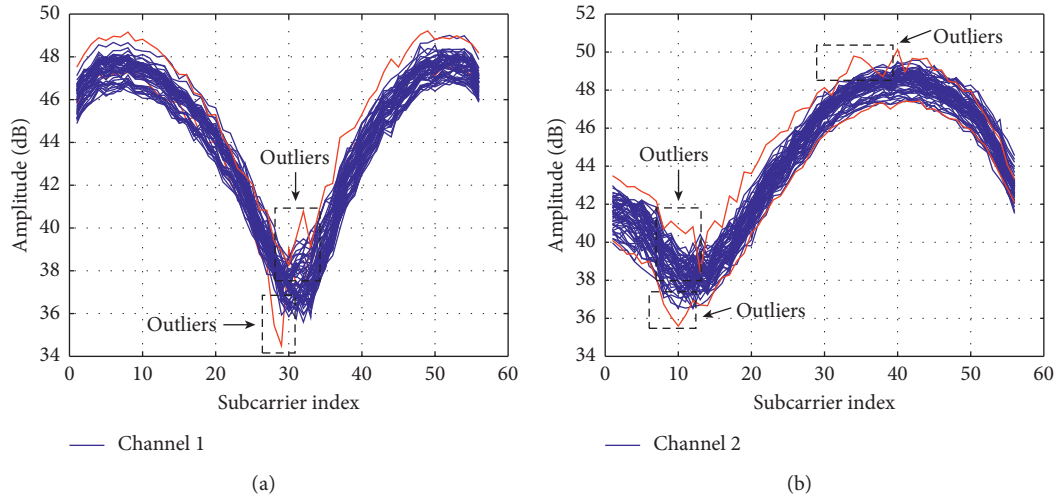
(a)

(b)

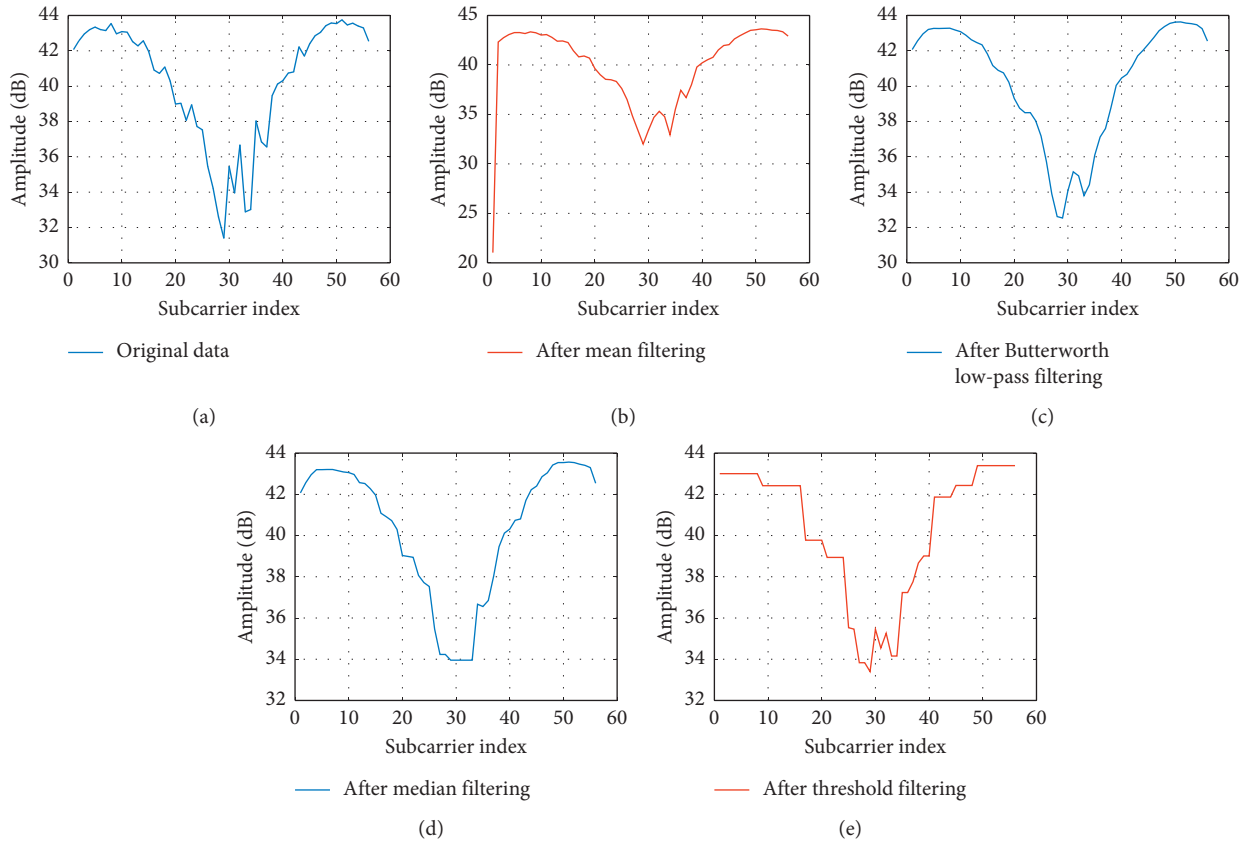FIGURE 4: Original CSI data.

(a)

(b)

(c)

(d)

(e)

FIGURE 5: Evaluation of four different filtering effects. (a) Original data. (b) After mean filtering. (c) After Butterworth low-pass filter. (d) After median filtering. (e) After threshold filtering.

is poor and it cannot filter the anomalies effectively. This is because the original data become smaller and the data waveform changes after averaging the CSI data. However, although the median filter and threshold filter can filter out the obvious noise, they ignore the detailed noise in the high-frequency part. It is obvious that the Butterworth low-pass filter has the best filtering effect on the data. It can preserve the integrity of the data to the greatest extent, that is, remove the detailed noise in the abnormal data without changing the data size and data waveform. By comprehensive comparison and consideration, we use the Butterworth low-pass filter in this paper to complete the data exception processing.
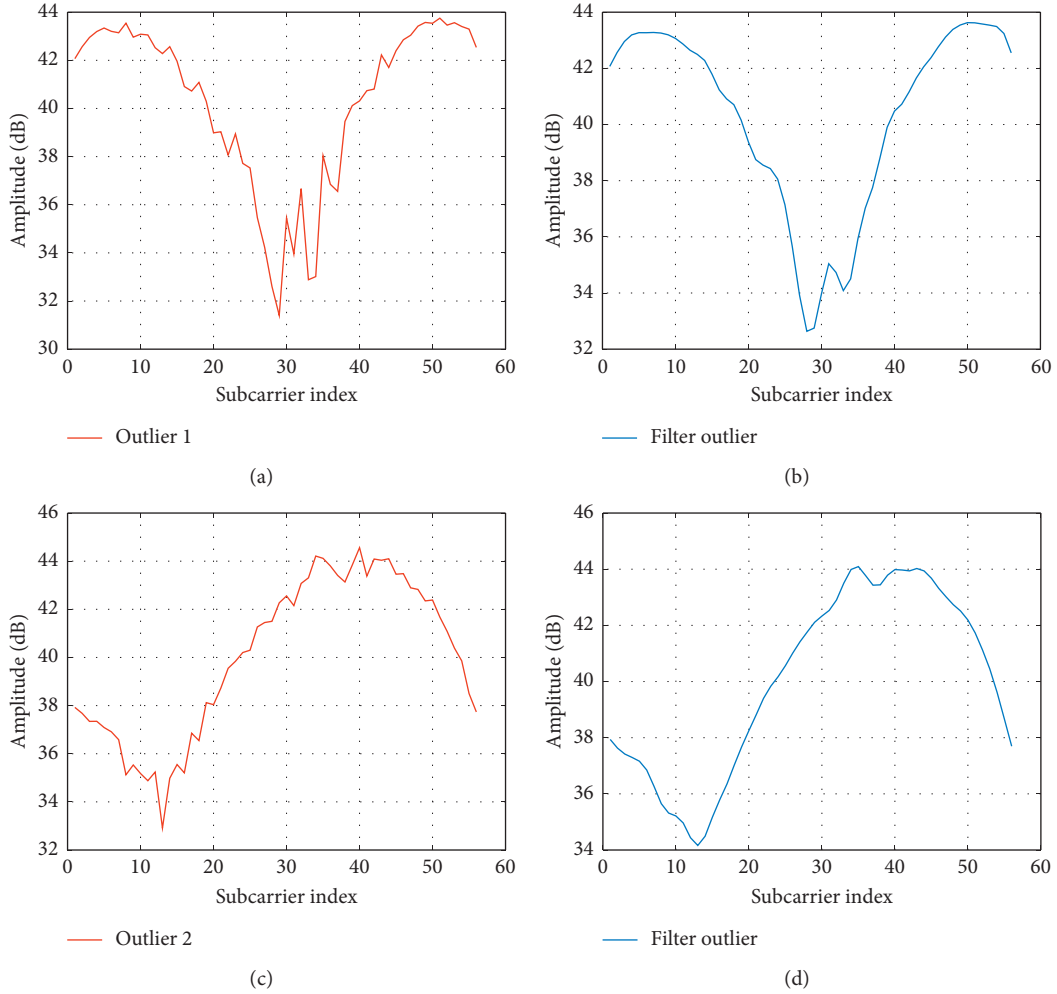
(a)

(b)

(c)

(d)

FIGURE 6: Data before and after comparison by the Butterworth low-pass filter: (a) channel 1 outlier 1; (b) after the Butterworth low-pass filtering of outlier 1; (c) channel 2 outlier 2; (d) after the Butterworth low-pass filtering of outlier 2.

4.1.2. High-Frequency Outlier Processing.

The high-frequency outliers in the CSI data of the XingYiQuan motion are filtered by the Butterworth low-pass filter. The processing steps are as follows:

Step 1: the amplitude-frequency transfer function $H(\varepsilon)$ of the Butterworth low-pass filter is designed as $|H(j\Omega)|^2 = (1/(1 + (\Omega/\Omega_C)^{2N}))$, where N is the filter order, Ω_C is the cutoff frequency (in rad/s), and j is the denominator coefficient of each order.

Step 2: the relevant parameters of the Butterworth low-pass filter $\eta(f_n, f_s, f_p, R_p, R_s)$ are set, where f_n is the sampling frequency, f_s is the stopband cutoff frequency, f_p is the passband cutoff frequency, R_p is the minimum attenuation of the fluctuation in the passband, and R_s is the minimum attenuation in the stopband. Using the influence frequency of the human body on the signal as the passband cutoff frequency, the part of the environmental noise higher than the average frequency of the human body is filtered out, generating relevant parameters suitable for processing amplitude data.

Step 3: the amplitude matrix of XingYiQuan is regarded as the original signal, and the Butterworth low-pass filter designed in Step 2 is used to filter out the high-frequency outliers in the amplitude information.

After the Butterworth low-pass filter filters out the abnormal XingYiQuan motion data before and after comparison, as shown in Figures 6(a)–6(d), it can be seen that the CSI data become smoother and high-frequency anomalies are filtered out.

4.1.3. Low-Frequency Outlier Processing.

The wavelet function is used to filter out low-frequency interference data in the CSI data. Taking the XingYiQuan motion feature information as the original input signal of the wavelet transform, and Sym8 wavelet as the wave base of wavelet transform, the CSI signal is decomposed by five layers and
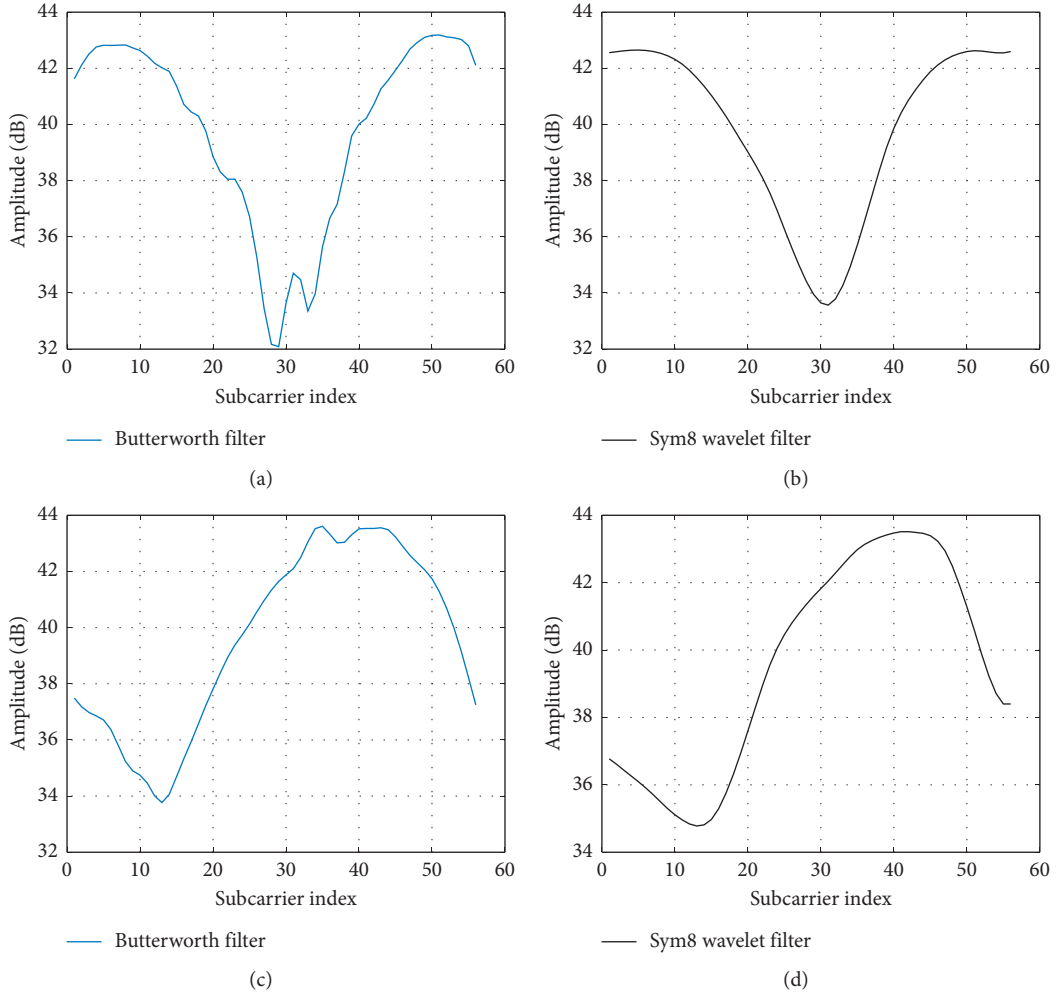
FIGURE 7: Sym8 wavelet function filters out low-frequency outliers. (a, c) After the Butterworth low-pass filter. (b, d) After the Sym8 wavelet filter.

the corresponding approximate part and detailed part are obtained, which approximately represent the low-frequency information. The detail represents the high-frequency information, and the sure threshold mode and scale noise are selected for the detail coefficient.

Figures 7(b) and 7(d) are CSI amplitude diagrams generated by Sym8 wavelet function. It can be seen that, after filtering out low-frequency anomalies, CSI data become more stable and retain the integrity to a large extent so that the features of each motion can be extracted more easily, and they can be used as the sample set of XingYiQuan data training, so as to better realize the classification of motion data in the next step.

4.2. XingYiQuan Motion Classification

4.2.1. Offline XingYiQuan Motion Fingerprint Establishment. The XingYiQuan data filtered by the outliers are used as the training sample. The RBM training requires the learning parameter θ'. It is assumed that a training set $\lambda = \left\{ \lambda^{(1)}, \ldots, \lambda^{(t)} \right\}$ is given, where t is the number of training

samples. The maximum likelihood method is used to solve the likelihood function $P(v)$ of RBM training:

$$
\begin{aligned}
\text{maximize } L(\theta') &= \text{maximize}_{\{W_{ij}, a_i, b_j\}} \frac{1}{t} \sum_{i=1}^{t} \ln \left(\sum_{h} P(v^l, h^l) \right) \\
&= \frac{1}{t} \sum_{l=1}^{t} \ln \left(P(v^{(l)}) \right),
\end{aligned}
$$

(5)

where v is the visible layer, h is the hidden layer, W_{ij} is the connection weight between the visible layer and the hidden layer, l is the number of visible and hidden layer neural units when solving the maximum likelihood function, and t is the number of training samples. In this paper, the k-step contrast divergence (k-CD) algorithm is used to train the RBM network, that is, it only needs M steps of Gibbs sampling to obtain an approximation suitable for the sample model to be trained. According to the establishment of the likelihood function, when the visible layer data are constant, the probability of the activation state of all hidden

layer neurons can be deduced, as shown in formula (6). Similarly, when the hidden layer data are constant, the activation state of the visible layer neurons can also be deduced from formula (7).

$$P\left(v_i = 1 \mid h; \theta'\right) = \sigma\left(\sum_j W_{ij} h_j + a_i\right), \qquad (6)$$

$$P\left(h_j = 1 \mid v; \theta'\right) = \sigma\left(\sum_i W_{ij} v_i + b_j\right), \qquad (7)$$

$$\sigma(x) = \text{sigmoid}(x) = \frac{1}{1 + e^{-x}}, \qquad (8)$$

where a_i is the bias of the visible layer, b_j is the bias of the hidden layer, and $\sigma(x)$ is a nonlinear sigmoid activation function of the neuron.

After collecting a large amount of action data of the XingYiQuan motion, the data are preprocessed. The data of each motion after processing are used as the fingerprint information $S = (S_1, S_2, S_3, \ldots, S_W)^T$, and S is used as the sample set to be trained, and the RBM is trained. The training process is as follows:

Step 1: we first initialize the RBM and assign initial values to the visible neurons, $v^0 = v$.

Step 2: Gibbs sampling on the training sample set S and the M-step Gibbs sampling are carried out when the time is $l = 1, 2, \ldots, M$.

Step 3: the Gibbs sampling method is used to sample the visible layer neurons and process them. By setting up a likelihood function to sample the visible layer from the hidden layer, we can use $P(h^{(l)} \mid v^{(l-1)})$ to sample the visible layer and calculate the corresponding v^l. Similarly, to Gibbs sampling of hidden layer neuron data, we can use $P(h^{l-1} \mid v^{l-1})$ to sample the hidden layer and calculate h^{l-1}.

Step 4: the corresponding expected values are calculated for the hidden layer and the visible layer to complete the training.

After the M-step Gibbs sampling of the training sample set, the system can obtain an approximate sampling value which is suitable for the training sample model. At the same time, the system can also determine the activation state of the neurons in the visible layer. In this paper, the fingerprint information after training is recorded as $S' = (S_1', S_2', S_3', \ldots, S_W')^{T'}$ to establish the characteristic fingerprint database of each XingYiQuan motion.

4.2.2. Online SoftMax-Modified RBM Classification.
In the online phase, the data are collected in real time in the experimental environment and the same abnormal filtering and RBM classification are carried out in the offline phase, and the RBM is trained to get $\theta' = (W, a, b)$ as the test input of the SoftMax function. The logistic regression cost function is used as the cost function of the SoftMax classifier, and the

SoftMax function is set as $U_\eta(\theta')$. The probability $P(y^i = j \mid \theta'^i)$ of each class of boxing samples is given to obtain the category label. If the input classification parameter is $\theta' = (\theta_1', \theta_2', \theta_3', \ldots, \theta_m')$, the model parameter is $\eta_1, \eta_2, \ldots, \eta_k$ and the sum of all probabilities is 1 by normalization.

$$U_\eta\left(\theta'^{(i)}\right) = \begin{bmatrix} P\left(y^{(i)} = 1 \mid \left(\theta'^{(i)}\right); \eta_1\right) \\ P\left(y^{(i)} = 2 \mid \left(\theta'^{(i)}\right); \eta_2\right) \\ \vdots \\ P\left(y^{(i)} = k \mid \left(\theta'^{(i)}\right); \eta_k\right) \end{bmatrix} = \frac{1}{\sum_{j=1}^k e^{\eta_j^T \theta'^{(i)}}} \begin{bmatrix} e^{\eta_1^T \theta'^{(i)}} \\ e^{\eta_2^T \theta'^{(i)}} \\ \vdots \\ e^{\eta_k^T \theta'^{(i)}} \end{bmatrix}. \qquad (9)$$

Next, the maximum likelihood estimation is used to obtain the cost function of SoftMax. In order to simplify the cost function, the indicator function $\text{ind}(\cdot)$ is introduced:

$$\text{ind}(y_i = j) = \begin{cases} 1 & \text{true} \\ 0 & \text{false} \end{cases}. \qquad (10)$$

Then, the cost function can be expressed as

$$J(\eta) = -\left(\sum_{i=1}^k \text{ind}(y = j_i) \ln \frac{e^{\eta_i^T \theta_I}}{\sum_{i=1}^k e^{\eta_i \theta_I}}\right). \qquad (11)$$

By constructing the SoftMax classification model, the data sample sets of different actions are input in turn, and each time, the data representing different actions are selected, a one-dimensional matrix containing six categories is returned, and select the category with the highest probability each time to correspond to the six motions to be classified. The six motions of the classification further correct the classification accuracy of the RBM through the SoftMax classification function and ensure that the CSI-HC method has higher motion recognition accuracy.

4.3. Online XingYiQuan Motion Recognition.
In the online recognition phase, the transmitter is used to collect the data of each XingYiQuan motion to be identified, the collected data are sent to the receiver, and the amplitude of the CSI data is selected as the feature. And the frequency in the signal is represented by the rate of multipath change caused by body motion, while the amplitude represents the energy of the signal. Therefore, by analyzing the amplitude information in the signal, we can obtain the amplitude characteristics of each XingYiQuan motion and then discriminate the different motions. The online phase recognition process is as follows:

Step 1: the real-time data of the XingYiQuan motion are collected between the receiver and the transmitter equipped with the Atheros network card.

Step 2: the amplitude in the CSI data is selected as a feature.

Step 3: the outliers are filtered out by data pre-processing, which is easy to match with the established standard fingerprint database.

Step 4: the RBM classification model is constructed by using the trained learning parameter $\theta' = (W, a, b)$, and the detection data are classified.

Step 5: the SoftMax classifier is used to modify the RBM classification results to improve the classification accuracy.

Step 6: the classified amplitude data are matched with the offline fingerprint database in real time.

Step 7: according to the above steps, the specific XingYiQuan motion performed by the tester is finally recognized.

5. Experiments and Evaluation

In this section, in order to evaluate the performance of CSI-HC, a large number of experiments have been carried out. First of all, we have introduced the experimental configuration in detail. Then, combined with three typical indoor scenarios, the key parameters that affect the recognition accuracy and the diversity of users are analyzed, and finally, the overall performance of CSI-HC is shown.

5.1. Experimental Design

5.1.1. Hardware Configuration. In order to verify the feasibility of the CSI-HC method in the actual scenario, this paper adopts the Atheros AR9380 network card solution based on the IEEE 802.11N protocol. The required equipment is two desktop computers equipped with the Atheros AR9380 network card, CPU model is Intel Core i3-4150, operating system is Ubuntu 16.04 LTS4.1.10 Linux kernel version, and two 1.5 m long 5 dB high-gain external antennas, one of which acts as the transmitting end and the other as the receiving end, which, respectively, connect the antenna contacts of the Atheros NIC of the transmitter and receiver with a 1.5 m external antenna. The Atheros AR9380 NIC and external antenna are shown in Figure 8.

5.1.2. Experimental Scenarios. We choose the testbed in three classical indoor scenarios, such as the meeting room (6.5 m × 10 m), corridor (2 m × 46 m), and office (6.5 m × 13 m), in order to correspond to the change of the multipath effect from low to high. The experimental scenarios and experimental scenarios' plane structure are shown in Figures 9 and 10. First of all, keeping the height and distance of the receiver and the transmitter, and the transmitter sending a certain rate, in the three scenarios, the testers are arranged to stand at the midpoint of the deployed transmitter and receiver to make a fixed motion of XingYiQuan. Each motion collects 10,000 packets of CSI data and saves them to the receiver PC. After processing by the CSI-HC method, the RBM is used to train and learn the data of each XingYiQuan motion, and the standard feature fingerprint information of each XingYiQuan motion is established.

In the online recognition phase, the testers stand in the middle of the transmitter and receiver in the deployed experimental scenario to do the XingYiQuan motion, collect, and process CSI data in real time and classify them using the constructed RBM model. Through the established standard feature fingerprint database for real-time matching, the specific motions of the testers are identified. In Table 1, we give the relevant steps for a XingYiQuan motion recognition, as well as the relevant hardware used in each step and the specific time it takes to process these operations.

5.2. Analysis of Influencing Factors of the Experiment

5.2.1. TX-RX Distance Analysis. In the process of establishing the offline motion fingerprint database, two important device parameters (TX contract rate and RX and TX distance) played a key role in the effect of online XingYiQuan motion recognition. In order to analyze the effect of the distance between TX and RX on the motion recognition method proposed in this paper, we take the method of controlling variables during the offline fingerprint database construction phase. For the same user, the sampling time is kept constant (100 s), the contract rate is 10 p/s, and the TX-RX distances of 0.6 m, 0.8 m, 1 m, 1.2 m, 1.4 m, 1.6 m, 1.8 m, and 2 m are set in three scenarios such as the office, corridor, and meeting room, to analyze the impact of different TX and RX distance settings on the effect of XingYiQuan motion recognition in the online phase during offline training, in order to find the most suitable distance setting for offline fingerprint training. Figure 11 shows how the recognition accuracy of six XingYiQuan motions varies with TX-RX distance in three scenarios.

As can be seen from Figure 11, with the increase of the distance between the transmitter and the receiver in three different scenarios, the accuracy of recognition of the six XingYiQuan motions generally shows an upward trend. The overall data except the MaXingQuan motion indicate when the distance between the transmitter and the receiver is 1.4 m, the accuracy of motion recognition reaches a peak, and when the distance is between 1.4 m and 2 m, the accuracy of motion recognition begins to decline slowly. After the distance exceeds 1.4 m, the accuracy of motion recognition begins to decrease slowly. Because of the complexity of the MaXingQuan motion and the ZuanQuan motion, the multipath interference and the interference within the environment are large, so there are fluctuations that are inconsistent with other XingYiQuan motions. And the experimental results also verify the previously set multipath effect-increasing scene, so the distance between the transmitter and the receiver is 1.4 m, which is most suitable for the XingYiQuan motion recognition in this question.

5.2.2. TX Contracting Rate Analysis. The transmitter's packet rate determines the number of samples collected during the same time period, as well as the sensitivity of the XingYiQuan motion to be acquired in the data link. Since

(a)

(b)

FIGURE 8: (a) WiFi commercial Atheros AR9380 NIC. (b) 5 dB high-gain external antenna.
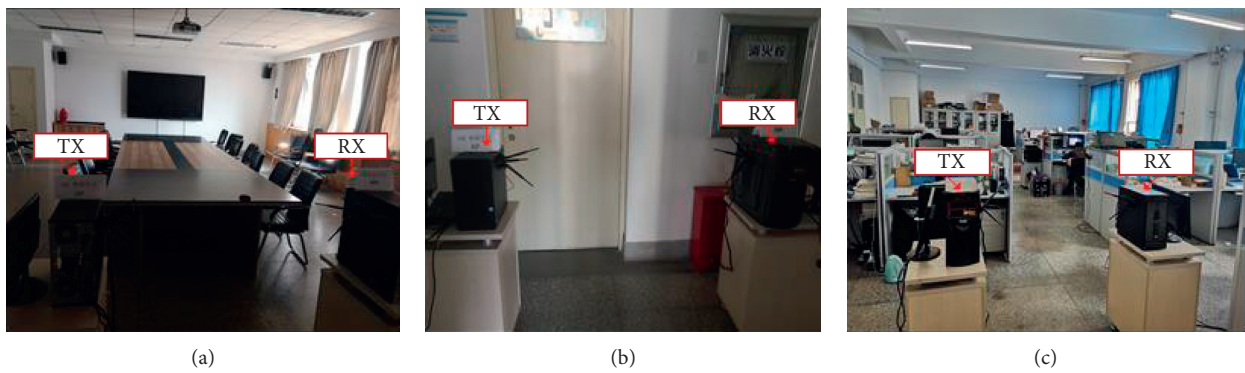

(a)

(b)

(c)

FIGURE 9: Experimental scenarios: (a) meeting room; (b) corridor; (c) office.
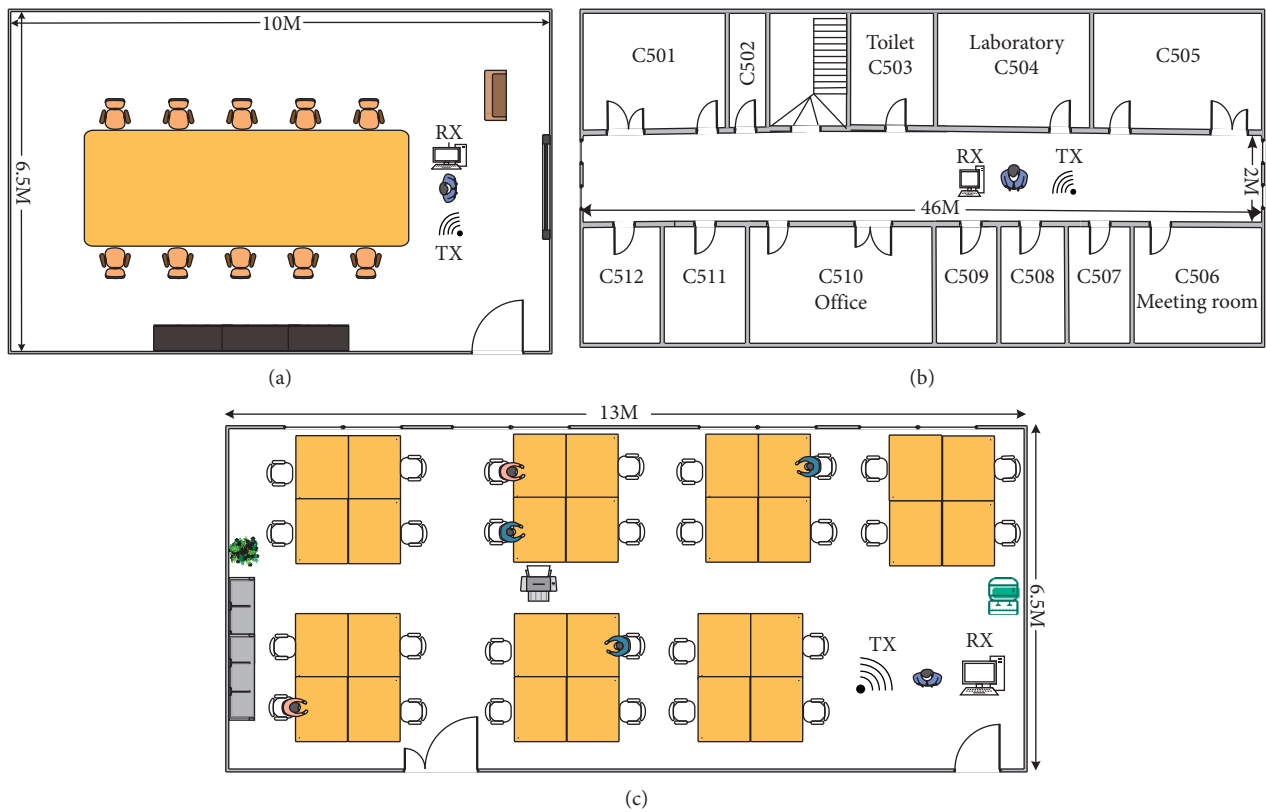

(a)

(b)

(c)

FIGURE 10: Floor plan of the experimental scenarios.

TABLE 1: Hardware processing time design.

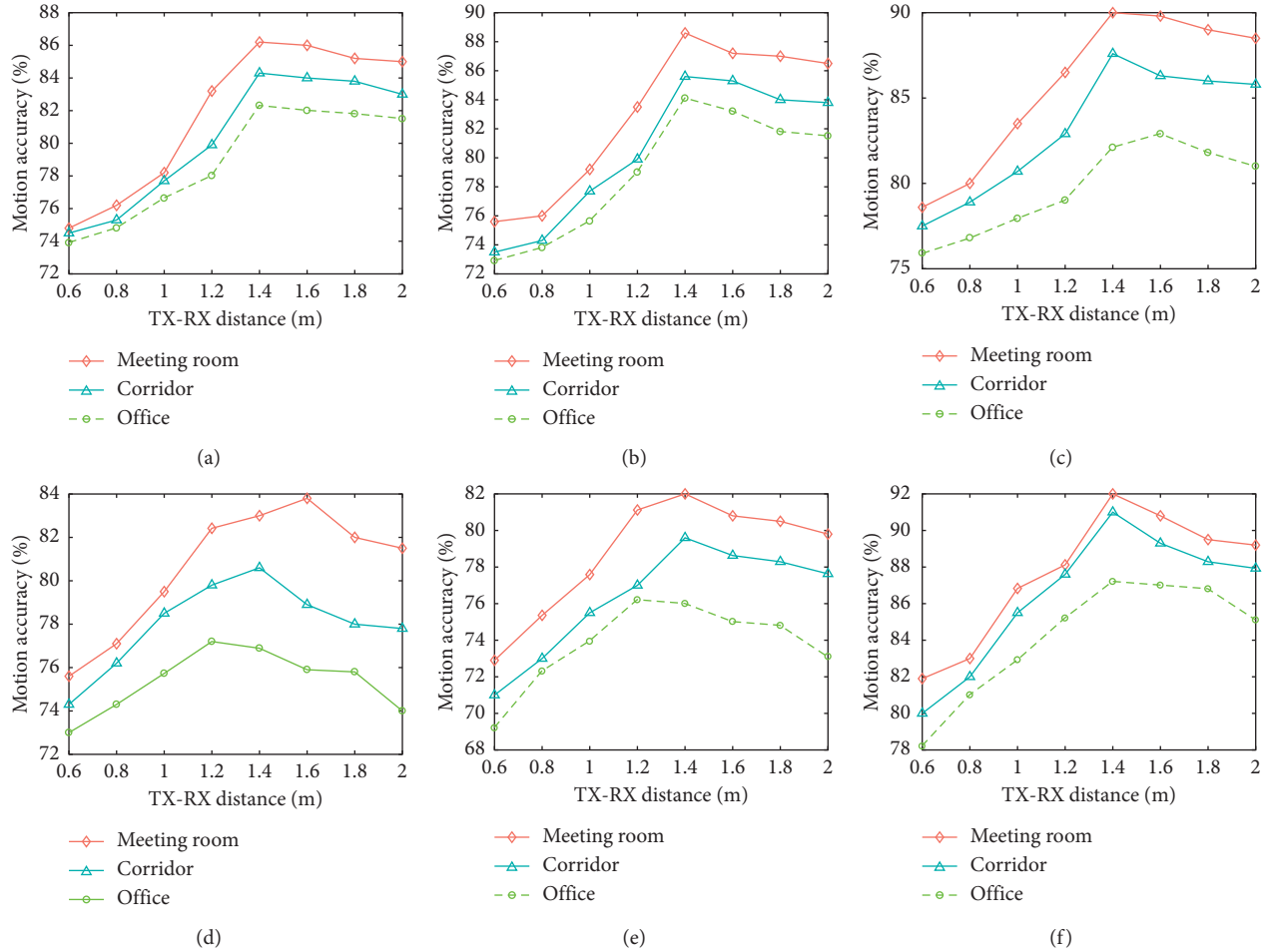| Related steps	Data collection	Offline fingerprint building	Online motion recognition
Hardware type	TP-Link WDR4310	PC	TP-Link WDR4310 + PC
Processing time	0.28 hours	0.11 hours	0.05 hours

FIGURE 11: Effect of TX-RX distance on the accuracy of motion: (a) QiShi motion; (b) BengQuan motion; (c) HuQuan motion; (d) MaXingQuan motion; (e) ZuanQuan motion; (f) ShouShi motion.

TABLE 2: Relationship between the transmitter contract rate and the packet loss rate.

Contract rate (p/s)	Estimated number of samples/p (10 s)	Actual number of samples/p (10 s)	Packet loss rate (%)
10	100	91	9
20	200	173	13.5
50	500	418	16.4
100	1000	813	18.7

the signal propagates in the air with a certain attenuation, the number of sample data collected at different packet rates is different at the same time. Assuming that q is the preset contract rate, T is the sampling time, and N is the actual number of samples, then the estimated number of sampling points $N1$ is $N1 = T * q$ and the packet loss rate loss can be defined as loss $= (N - T * q)/N$. According to the definition of the packet loss rate, we test the packet loss number of the transmitter under the 10 p/s, 20 p/s, 50 p/s, and 100 p/s

packet loss rates under the condition that the same user has a sampling time of 10 s and calculate the corresponding packet loss rate. Table 2 reflects the relationship between the contract rate and the packet loss rate of the transmitter.

What is obvious in Table 2 is that when the contract rate of the transmitter is 10 p/s, the actual number of packets collected accounts for the highest proportion of the estimated number of sampled packets, that is, the lowest packet loss rate. Because the minimum packet loss rate can get the

most true sampling number when the total collection number is fixed, and as many sampling points as possible can more truly reflect the motion state of the human body and then achieve a high-precision motion recognition, we choose the contract rate of 10 p/s in the construction phase of the offline fingerprint database.

In order to verify the effectiveness of the contract rate set during the offline phase, we tested different offline contract rate settings. Under the condition that the distance between TX and RX is set to 1.4 m and the total number of samples is 1000 packets, the training users in the offline phase conduct offline training on the contract rate settings of 10 p/s, 20 p/s, 50 p/s, and 100 p/s in three scenarios including the office, corridor, and meeting room and conduct the corresponding XingYiQuan motion recognition test.

By analyzing and comparing the contract rate and motion recognition accuracy data in Figure 12, we find that the accuracy of motion recognition is the highest when the packet rate is 10 p/s in the experimental scenario designed by this method and the action rate is 20 p/s. When the packet sending rate is 50 p/s and 100 p/s, the accuracy of recognition action is greatly reduced because of excessive packet loss. Therefore, the most suitable contract rate for the CSI-HC method is 10 p/s.

5.2.3. User Diversity Analysis.

5.2.3. User Diversity Analysis. In order to explore the influence of user diversity on the accuracy of motion recognition, two different testers were arranged in the experiment. Under the condition that the distance between TX and RX was set to 1.4 m, the total number of samples was 1000 packets, and the contract rate was 10 p/s, six XingYiQuan motions were performed many times to collect and process data. Then, the average recognition rates of the six XingYiQuan motions of the two testers were calculated. The result of the accuracy distribution of two users' motion recognition is shown in Figure 13.

As can be seen from Figure 13, the accuracy of the motion recognition of both testers was maintained above 80%, and the highest accuracy of the first tester's Xingyi action was 92.3%. This is because the first tester had a long period of XingYiQuan motion practice and performed the standard motion. In addition, the second tester, as he was a novice, had a low degree of proficiency in the XingYiQuan motion. The lower the accuracy of the motion due to the nonstandard motion (80.8%), the higher the standard of the motion performed by the tester in the box diagram and the shorter the length of the box, such as the BengQuan and ShouShi motions of tester 1. The lower the standard of the tester's motion, the longer the length of the box, such as the BengQuan and ShouShi motions of tester 2.

5.3. Overall Performance Evaluation

5.3.1. Robustness Testing. In order to test the robustness of the CSI-HC method, we cut through the environment and the user and arrange two different testers. First, let a tester do the same XingYiQuan motion in the meeting room, corridor, office, etc. The CSI data of the action are collected, the

data processing is trained, and the difference of CSI-HC recognition accuracy in different scenarios is analyzed, and then another tester is arranged to do the same motion in the three scenarios. The motion data are collected and processed and compared with the CSI characteristics of the previous person. Figure 14 shows the amplitude of the QiShi motion of the two testers in the meeting room, corridor, office, etc. Table 3 shows the specific accuracy of motion recognition.

The amplitude features of CSI motions of different testers in three scenarios (taking the QiShi motion as an example) are compared, as shown in Figure 14. On the one hand, according to the similar amplitude trend of (a), (c), and (e) in Figure 14, it is shown that CSI-HC is less affected by environmental factors and can have higher performance even in the case of serious multipath interference. On the other hand, by comparing (a) with (b), (c) with (d), and (e) with (f) in Figure 14, we can see that the trend of amplitude characteristics collected by different people in the same environment is approximately the same. Since the principle of motion recognition of CSI-HC is based on the motion and the corresponding amplitude characteristics, the amplitude change trend of different testers is consistent, which shows that CSI-HC is not sensitive to user differences and has high robustness.

Table 3 shows the accuracy of the CSI-HC detection of six motions of XingYiQuan in three different scenarios. The average detection accuracy in the meeting room is 89.33%. However, the average detection accuracy in the office is only 81.23%. In addition, the average detection rate in the corridor and meeting room is higher than that in the office because there are more multipath components and human interference in the office.

5.3.2. Overall Performance. In order to evaluate the overall performance of the CSI-HC proposed in this paper, we define the following three metrics, compare them with two classic human motion detection methods: R-PMD and FIMD, and compare their detection results in the meeting room, corridor, and office. The experimental results are shown in Figure 15 and Table 4.

   (i) Precision: precision = TP/(TP + FP) (also defined as sensitivity) refers to the probability of correctly recognizing the human motion.

  (ii) Recall: recall = TP/(TP + FN) (also defined as particularity) refers to the probability of correctly identifying nondetected human motions.

 (iii) $F1$ score: $F1$ score = $2 * \text{precision} * \text{recall}/(\text{precision} + \text{recall})$. The $F1$ indicator is a comprehensive evaluation standard combining precision and recall. By calculating the $F1$ score index of different methods, the stability of the method can be effectively evaluated.

Here, TP = true positives, FP = false positives, and FN = false negatives.

Figure 15 shows the comparison of CSI-HC, R-PMD, and FIMD methods in terms of precision, recall, and $F1$ score. From Figure 15, we can see that the precision, recall,
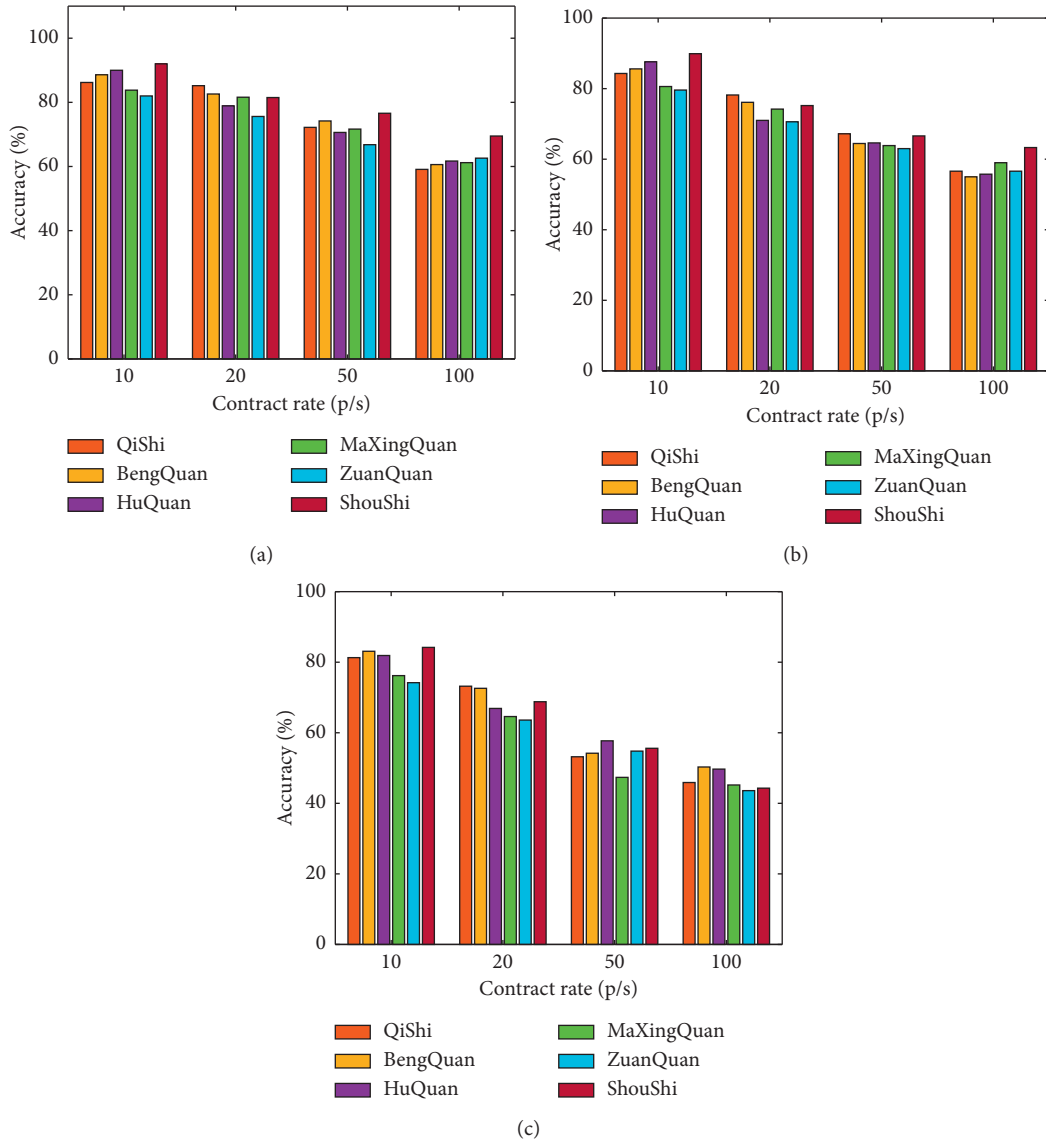
FIGURE 12: Effect of the contract rate on motion accuracy in three testbeds: (a) meeting room; (b) corridor; (c) office.
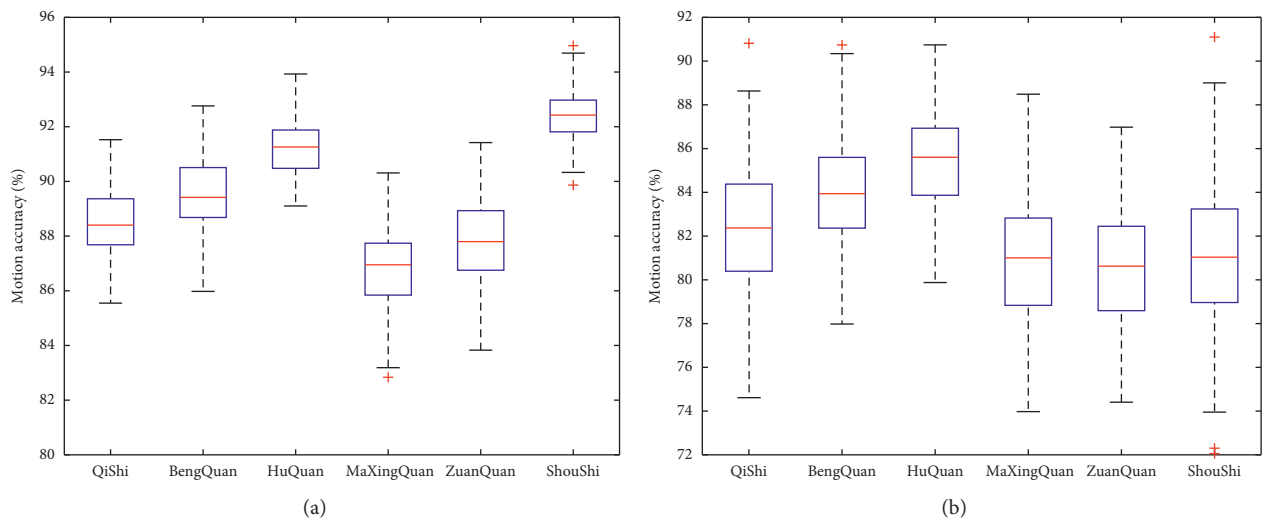

FIGURE 13: User diversity and recognition accuracy analysis. (a) Tester 1. (b) Tester 2.
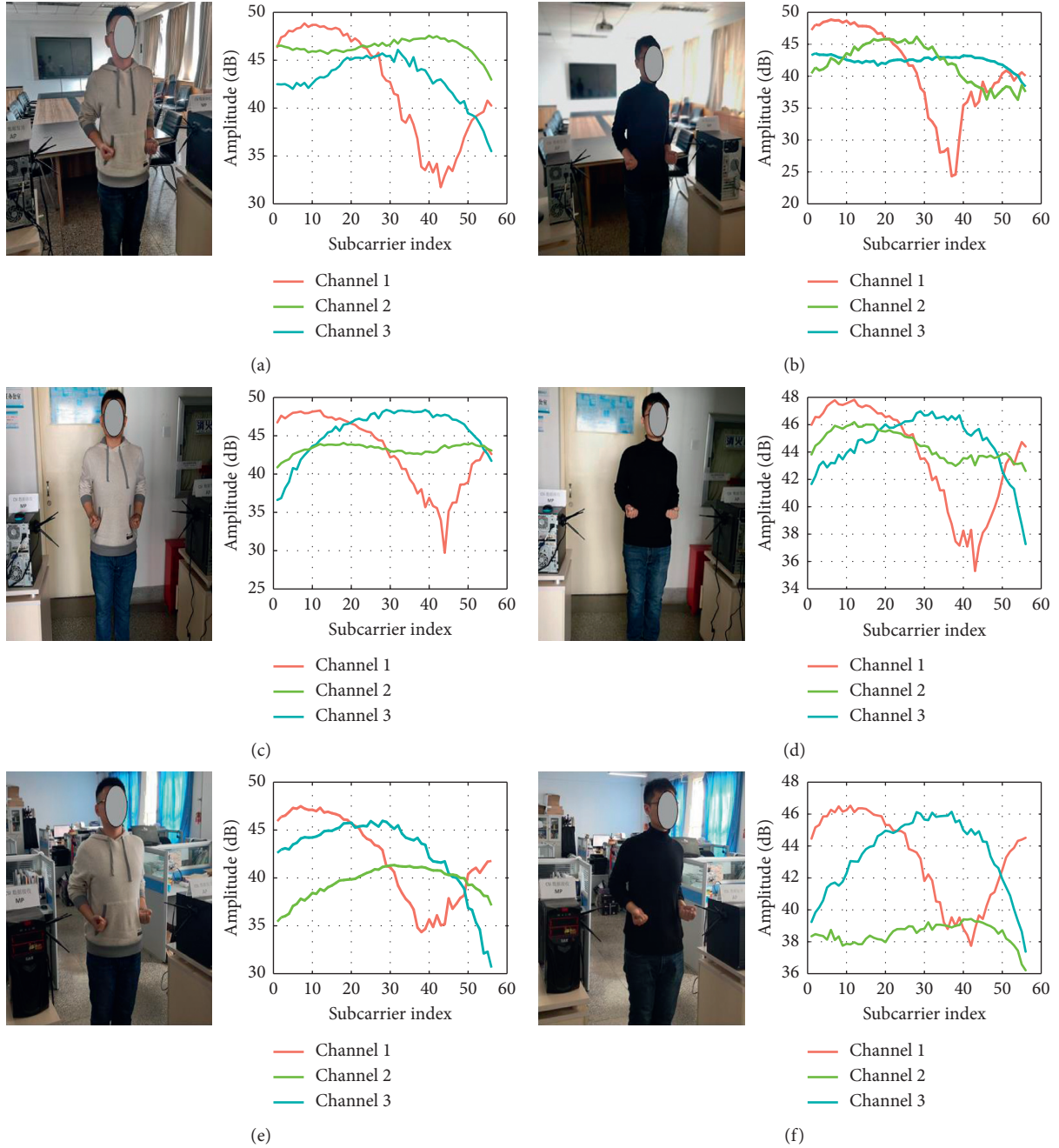
Figure 14: Amplitude characteristic maps of two testers performing the QiShi motion in three scenarios. (a) Meeting room (tester 1). (b) Meeting room (tester 2). (c) Corridor (tester 1). (d) Corridor (tester 2). (e) Office (tester 1). (f) Office (tester 2).

Table 3: Robustness evaluation of CSI-HC in three scenarios.

Different scenarios	Detection accuracy of different XingYiQuan actions (%)					
	QiShi	BengQuan	HuQuan	MaXingQuan	ZuanQuan	ShouShi
Meeting room	88.2	89.6	91.0	87.8	87.1	92.3
Corridor	84.3	85.9	86.6	83.6	85.5	88.0
Office	80.9	80.1	81.0	81.0	81.2	83.2

and $F1$ score index of CSI-HC are obviously better than those of the other two methods. This is because the denoising method of CSI-HC is a combination of low-pass filtering and

wavelet transform, which greatly filters multipath interference and compares the complete retention of the motion features. However, R-PMD uses low-pass filtering and
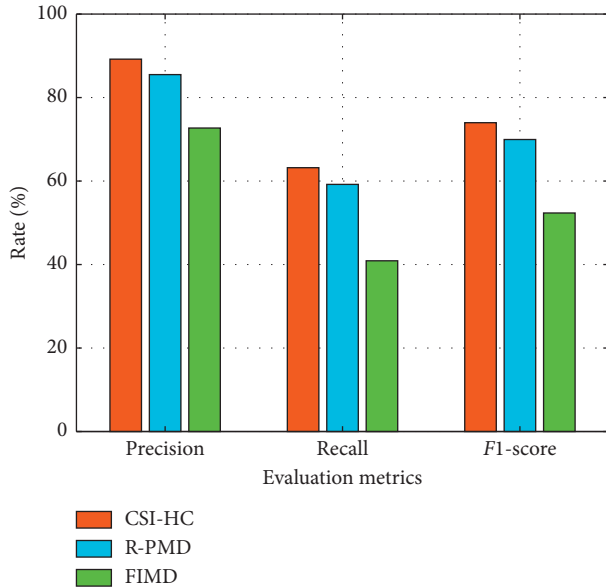
Figure 15: Comprehensive evaluation of three methods.

Figure 16: Impact of the number of samples.

Table 4: Motion detection method performance in three indoor environments.

Method	Meeting room (%)	Corridor (%)	Office (%)
CSI-HC	89.3	85.7	81.2
R-PMD	88.2	84.0	80.5
FIMD	76.1	71.8	70.6

principal component analysis (PCA), filtering out more motion features, retaining only a few representative features, and affecting the recognition accuracy. However, FIMD adopts the method of false alarm. This method can only eliminate the erroneous data with large deviation from the CSI feature and cannot filter out some interference data with small deviation.

Table 4 shows the motion detection accuracy of CSI-HC, R-PMD, and FIMD in the three scenarios of this paper. The experimental results show that the detection accuracy of CSI-HC in the meeting room reaches 89.3%. The approach has high environmental adaptability and robustness.

Figure 17: Impact of the number of features.

5.3.3. The Impact of Sample Size.

We find that the number of training samples affects the detection accuracy of the motion recognition method. To this end, we test different sample sizes in the real environment we have built, and the experimental results are shown in Figure 16.

Figure 16 reflects the relationship between the recognition accuracy of the three motion detection methods and the number of training samples. It can be seen that the detection accuracy of the motion recognition methods increases with the number of training samples. Compared with the other two methods, the CSI-HC method has a high motion recognition rate. The motion detection accuracy of the R-PMD method is slightly lower than that of CSI-HC, and FIMD has the worst effect. This is because CSI-HC uses the RBM training method based on contrast divergence.
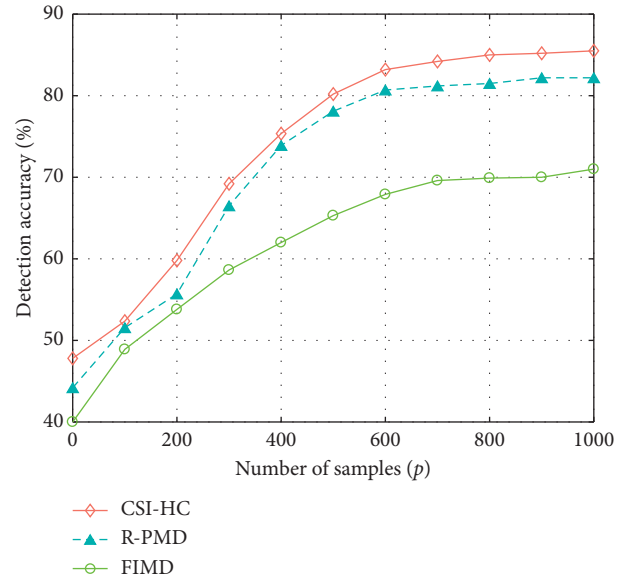
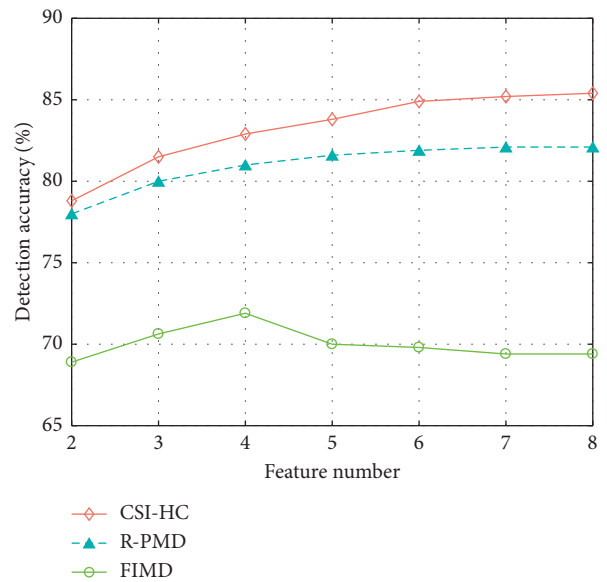After setting the training weight and learning rate, it can quickly fit with the sample by adjusting the weight parameters, reduce the system overhead by repeated reuse, improve the classification training efficiency of motions, and increase the accuracy of motion recognition. However, R-PMD uses the PCA approach. It means that, with the increase of data, principal component analysis should be carried out on all data to select characteristic data, which increases the system overhead. In contrast, FIMD uses the density-based spatial clustering of applications with noise (DBSCAN) method to determine the clustering center. By constantly calculating the Euclidean distance between sample points and updating clustering points, the time complexity of the algorithm is greatly increased, and with the increase of the number of training samples, the

TABLE 5: Comparison of parameters between CSI-HC and various human motion recognition methods.

Parameters	Methods			
	CSI-HC	Computer vision	Infrared	Dedicated sensor
Perceived distance	Non-line-of-sight	Line of sight	Non-line-of-sight	Short distance
Deployment costs	Low	Medium	High	High
Environmental limitations	No	Yes	No	Yes
Computational complexity	Low	High	High	Medium
Motion detection accuracy	Relatively high	High	High	High

computational burden of the system continues to increase, and the training and recognition effects on the motion data are not good.

5.3.4. The Impact of the Number of Features.

As shown in Figure 17, the motion detection accuracy increases as the number of features used increases. Specifically, when the number of features increases, the detection accuracy of the CSI-HC method proposed in this paper will also increase. When the number of feature values reaches 6, the detection accuracy reaches the highest and remains stable. In contrast, the accuracy change trend of the R-PMD method is roughly the same as that of CSI-HC, but it is significantly lower than that of CSI-HC. However, the turning point of FIMD detection accuracy is 5, and the detection rate decreases significantly when the number of features is 2–4.

The test results in Figure 17 show that the detection rate of our proposed CSI-HC method is higher and more stable. This is because the CSI-HC method uses the RBM network to train the model suitable for CSI data, and the characteristics are concentrated in the training data set, which preserves the data integrity to a large extent, so it has a high detection rate and stability. However, the R-PMD method uses the PCA method to extract the most representative data as features. The data integrity is not as high as that of CSI-HC, which causes the detection accuracy to decrease. The FIMD method uses the correlation matrix of CSI data as the feature and uses the DBSCAN method for clustering. The data are scattered in the feature, which makes it unstable, and the detection accuracy is low.

5.3.5. Comparison with the Optical Sensor Method.

We compare the CSI-HC method proposed in this paper with the optical sensor-based method. The CSI-HC method uses CSI data as a recognition feature, while the optical sensor-based method generally uses an optical signal composed of inherent characteristics of light as a feature for motion recognition. The average detection accuracy of our proposed CSI-HC for the motion is 85.4%, while the accuracy of motion detection based on optical sensor-based methods is as high as 98.9% [34]. Although in terms of recognition accuracy, the CSI-HC method is lower than the optical sensor-based method. However, in terms of application scenarios, the optical sensor-based method does not work properly in low-light and dark environments or in scenarios involving privacy. The CSI-HC method overcomes this defect and can be used in all indoor environments. In the future research, we will further improve the accuracy of the CSI-HC method to make up for its lack of recognition accuracy.

5.3.6. Comparison with Previous Motion Recognition Methods.

In order to reflect the unique advantages and comprehensive performance of the CSI-HC method proposed in this paper, we compare the CSI-HC method with the previous three human motion recognition methods (computer vision method, infrared method, and special sensor method) on multiple parameters [35]. The results are shown in Table 5.

As can be seen from Table 5, the CSI-HC method has the advantages of non-line-of-sight, low deployment cost, not limited by environmental conditions, and low algorithm complexity, while the other three methods have high deployment cost and computational complexity. Among them, computer vision methods cannot work effectively in dark and privacy-related scenes. However, although the infrared method and the dedicated sensor method have achieved high motion detection accuracy, the deployment cost is too high and they are not suitable for widespread deployment in indoor scenarios.

6. Conclusion

In this paper, we propose a new complex human motion recognition method, namely, CSI-HC, which is verified by the XingYiQuan motion which is a complex motion. The core part is the denoising of CSI signals and the classification of complex motions. We collect CSI amplitude data in the offline phase, use the Butterworth low-pass filter combined with the Sym8 wavelet function method to filter the outliers, and use the RBM to train and classify the XingYiQuan motion to establish standard fingerprint information for the XingYiQuan motion. Next, in the online phase, the XingYiQuan motion is collected in the experimental scenarios to collect real-time data, the abnormal value filtering and RBM training classification are performed, the SoftMax classification model is constructed to correct the RBM classification result, and the offline phase XingYiQuan motion fingerprint database is used for data matching to achieve recognition of different XingYiQuan motions. Through a large number of experiments, this paper explores the influence of parameter setting and user diversity on the accuracy of motion recognition. The experimental results show that the CSI-HC achieves an average motion recognition rate of 85.4% in three practical scenarios. It has good performance in robustness, motion recognition rate, practicability, and so on.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] Z. Wang, B. Guo, Z. Yu, and X. Zhou, "Wi-fi CSI based behavior recognition: from signals, actions to activities," *IEEE Communications Magazine*, vol. 56, no. 5, pp. 109–115, 2017.

[2] Y. Lu, S. Lv, X. Wang, X. Zhou et al., "A survey on WiFi based human BehaviorAnalysis technology," *Chinese Journal Of Computers*, vol. 1–222018, in Chinese.

[3] X. Dang, Y. Huang, Z. Hao, and X. Si, "PCA-Kalman: device-free indoor human behavior detection with commodity Wi-Fi," *EURASIP Journal on Wireless Communications and Networking*, vol. 2018, no. 1, 2018.

[4] J. Liu, Y. Wang, Y. Chen et al., "Tracking vital signs during sleep leveraging off-the-shelf WiFi," in *Proceedings of the ACM International Symposium on Mobile Ad Hoc Networking & Computing. ACM*, pp. 267–276, Hangzhou, China, June 2015.

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

[6] L. Sun, K. H. Kim, S. Sen et al., "WiDraw: enabling hands-free drawing in the air on commodity WiFi devices," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, pp. 77–89, Paris, France, September 2015.

[7] Y. Zeng, P. Patha, and P. Mohapatra, "WiWho: WiFi-based person identification in smart spaces," in *Proceedings of the 2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pp. 1–12, Vienna, Austria, April 2016.

[8] L. Chen, X. Chen, L. Ni, Y. Peng, and D. Fang, "Human behavior recognition using Wi-Fi CSI: challenges and opportunities," *IEEE Communications Magazine*, vol. 55, no. 10, pp. 112–117, 2017.

[9] D. Zhang, H. Wang, and D. Wu, "Toward centimeter-scale human activity sensing with Wi-Fi signals," *Computer*, vol. 50, no. 1, pp. 48–57, 2017.

[10] C. Harrison, D. Tan, and D. Morris, "Skinput: appropriating the body as an input surface," in *Proceedings of the Sigchi Conference on Human Factors in Computing Systems ACM*, pp. 453–462, Atlanta, GA, USA, April 2010.

[11] H. Ding, L. Shangguan, Z. Yang et al., "Femo: a platform for free-weight exercise monitoring with rfids," in *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems, ACM*, pp. 141–154, Seoul, Republic of Korea, November 2015.

[12] D. Ruiz-Fernandez, O. Marín-Alonso, A. Soriano-Paya, and J. D. García-Pérez, "eFisioTrack: a telerehabilitation environment based on motion recognition using accelerometry," *The Scientific World Journal*, vol. 2014, Article ID 495391, 11 pages, 2014.

[13] S. Herath, M. Harandi, and F. Porikli, "Going deeper into action recognition: a survey," *Image and Vision Computing*, vol. 60, pp. 4–21, 2017.

[14] Z. Zhang, "Microsoft kinect sensor and its effect," *IEEE Multimedia*, vol. 19, no. 2, pp. 4–10, 2012.

[15] H.-M. Zhu and C.-M. Pun, "An adaptive superpixel based hand gesture tracking and recognition system," *The Scientific World Journal*, vol. 201412 pages, 2014.

[16] D. Xu, S. Yan, D. Tao, S. Lin, and H.-J. Zhang, "Marginal fisher analysis and its variants for human gait recognition and content- based image retrieval," *IEEE Transactions on Image Processing*, vol. 16, no. 11, pp. 2811–2821, 2007.

[17] M. Fiaz and B. Ijaz, "Vision based human activity tracking using artificial neural networks," in *Proceedings of the International Conference on Intelligent and Advanced Systems (ICIAS)*, Kuala Lumpur, Malaysia, June 2010.

[18] F. Gu, A. Kealy, K. Khoshelham, and J. Shang, "User-independent motion state recognition using smartphone sensors," *Sensors*, vol. 15, no. 12, pp. 30636–30652, 2015.

[19] J. Dai, X. Bai, Z. Yang, Z. Shen, and D. Xuan, "PerFallD: a pervasive fall detection system using mobile phones," in *Proceedings of the 2010 8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, pp. 292–297, IEEE, Mannheim, Germany, March 2010.

[20] M. Youssef and M. Mah, "Challenges: device-free passive localization for wireless," in *Proceedings of the ACM Mobi-Com*, pp. 222–229, Montréal, Canada, September 2007.

[21] M. Seifeldin, A. Saeed, A. E. Kosba, A. El-Keyi, and M. Youssef, "Nuzzer: a large-scale device-free passive localization system for wireless environments," *IEEE Transactions on Mobile Computing*, vol. 12, no. 7, pp. 1321–1334, 2013.

[22] S. Sigg, S. Shi, F. Büsching, Y. Ji, and L. C. Wolf, "Leveraging rf-channel fluctuation for activity recognition: active and passive systems, continuous and rssi-based signal features," in *Proceedings of the 11th International Conference on Advances in Mobile Computing & Multimedia*, p. 43, Vienna, Austria, December 2013.

[23] F. Adib and D. Katabi, "See through walls with WiFi!" *ACM SIGCOMM Computer Communication Review*, vol. 43, no. 4, pp. 75–86, 2013.

[24] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, "Whole-home gesture recognition using wireless signals," in *Proceedings of the 19th Annual International Conference on Mobile Computing and Networking*, pp. 27–38, New York, NY, USA, December 2013.

[25] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: gathering 802.11n traces with channel state information," *ACM SIGCOMM Computer Communication Review*, vol. 41, no. 1, p. 53, 2011.

[26] J. Xiao, K. Wu, Y. Yi, L. Wang, and L. M. Ni, "FIMD: fine-grained device-free motion detection," in *Proceedings of the 2012 IEEE 18th International Conference on Parallel and Distributed Systems*, vol. 90, no. 1, pp. 229–235, Singapore, December 2012.

[27] C. Wu, S. Member, Z. Zhou, X. Liu, Y. Liu, and J. Cao, "Non-invasive detection of moving and stationary human with WiFi," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 11, pp. 2329–2342, 2015.

[28] H. Zhu, F. Xiao, L. Sun, X. Xie, P. Yang, and R. Wang, "Robust and passive motion detection with cots wifi devices," *Tsinghua Science and Technology*, vol. 22, no. 4, pp. 345–359, 2017.

[29] W. Wang, A. X. Liu, M. Shahzad et al., "Understanding and modeling of WiFi signal based human activity recognition," in *Proceedings of the 21st Annual International Conference on Mobile Computing andNetworking*, pp. 65–76, NewYork, NY, USA, September 2015.

[30] H. Wang, D. Zhang, Y. Wang, J. Ma, Y. Wang, and S. Li, "RT-fall: a real-time and contactless fall detection system with commodity WiFi devices," *IEEE Transactions on Mobile Computing*, vol. 16, no. 2, p. 1, 2016.

[31] H. Li, W. Yang, J. Wang, X. Yang, and L. Huang, "WiFinger: talk to your smart devices with finger-grained gesture," in *Proceedings of the ACM International Joint Conference on Pervasive & Ubiquitous Computing. ACM*, pp. 250–261, Heidelberg, Germany, September 2016.

[32] Y. Wang, J. Liu, Y. Chen, M. Gruteser, J. Yang, and H. Liu, "E-eyes: device-free location-oriented activity identification using fine-grained WiFi signatures," in *Proceedings of theACM MobiCom*, pp. 617–628, Maui, HI, USA, September 2014.

[33] C. Han, K. Wu, Y. Wang, and L. M. Ni, "WiFall: devicefree fall detection by wireless networks," in *Proceedings of the IEEE Conference on Computer Communications IEEE INFOCOM*, pp. 271–279, Toronto, ON, Canada, May 2014.

[34] M. Pierre, D. Yohan, B. Rémi, P. Vasseur, and X. Savatier, "A study of vicon system positioning performance," *Sensors*, vol. 17, no. 7, p. 1591, 2017.

[35] A. F. M. Saifuddin Saif, A. S. Prabuwono, and Z. R. Mahayuddin, "Moving object detection using dynamic motion modelling from UAV aerial images," *The Scientific World Journal*, vol. 2014, Article ID 890619, 12 pages, 2014.