A Robust and Device-Free Daily Activities Recognition System using Wi-Fi Signals

Enjie Ding[1], Yue Zhang[1,3], Yun Xin[1,3], Lei Zhang[2*], Yu Huo[1], and Yafeng Liu[1]
[1]IoT Perception Mine Research Center, China University of Mining and Technology
Xuzhou, 221000, China
[2]School of Information and Electrical Engineering, Xuzhou University of Technology
Xuzhou, 221000, China
[e-mail: lei@cumt.edu.cn]
[3]School of Information and Control Engineering, China University of Mining and Technology
Xuzhou, 221000, China
*Corresponding author: Lei Zhang

Abstract

Human activity recognition is widely used in smart homes, health care and indoor monitor. Traditional approaches all need hardware installation or wearable sensors, which incurs additional costs and imposes many restrictions on usage. Therefore, this paper presents a novel device-free activities recognition system based on the advanced wireless technologies. The fine-grained information channel state information (CSI) in the wireless channel is employed as the indicator of human activities. To improve accuracy, both amplitude and phase information of CSI are extracted and shaped into feature vectors for activities recognition. In addition, we discuss the classification accuracy of different features and select the most stable features for feature matrix. Our experimental evaluation in two laboratories of different size demonstrates that the proposed scheme can achieve an average accuracy over 95% and 90% in different scenarios.

Keywords: CSI, human activities recognition, phase transformation, device-free system, classification algorithms

1. Introduction

With the rapid development of wireless sensing technology, a series of Wi-Fi based human activities recognition systems are proposed [13-16]. Compared to traditional indoor human activities detection systems, which use smartphones [1, 2], cameras [3, 4], accelerometer [5-7] or wearable sensors [8-10], Wi-Fi signal based systems have some advantages. First, Wi-Fi based solutions are device-free and users do not need to wear any sensors. Second, compared to cameras, Wi-Fi signals have better coverage and can still work in poor lighting conditions. Third, Wi-Fi based solutions will not cause privacy intrusion on users.
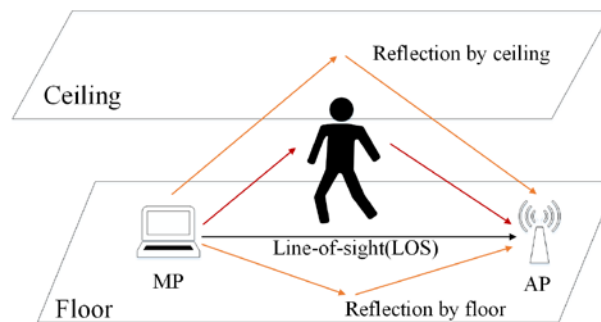

Fig. 1. Wi-Fi signal propagation in indoor environment.

The basis of Wi-Fi signal perception is the disturbance of human body on wireless signals. As illustrated in **Fig. 1**, the signal transmission paths change significantly when a human activity happens, so we can observe the impact of these multipath propagation by analyzing the received signals. Moreover, different human activities introduce different multi-path distortions on Wi-Fi signals, which can be used to distinguish corresponding activities. Received signal strength (RSS) of media access control (MAC) layer is utilized as a monitoring indicator in the early effort. However, it experiences severe performance degradation in complex environments due to temporal variations and signal reflections [11], and this will limit RSS to only be applied for recognizing coarse-grained activities. Fortunately, the channel state information (CSI) which can be extracted from off-the-shelf Wi-Fi devices provides fine-grained information [12]. Compared with RSS, each orthogonal frequency division multiplexing (OFDM) subcarrier of CSI consists of amplitude and phase information which is more fine-grained and can be used as the indicator of multi-path changes caused by human body.

In this paper, we propose a fine-grained activity detection system to implement the recognition of daily activities. Activities of daily living (ADL) refers to a series of basic activities that are necessary for people to take care of themselves in daily life. The identification of ADL is of great significance for health care and indoor personnel monitoring. As a result, our experiments take a number of ADL and divide them into two main categories: upper body movements (i.e. drinking water) and full-body movements (i.e. walking). In addition, since our system is designed for patients in hospital wards, prisoners in prisons and elderly people living alone, our experiments are mainly based on the scenarios in which no one moves except the target person. To achieve this, we focus on how different features describe human activities and select more sensitive features to construct feature matrix. We firstly pass the raw CSI data to the pre-processing module to eliminate significant random noise and

obtain useful phase information. Afterwards, we utilize principal component analysis (PCA) method on 30 subcarriers to extract principal components for feature extraction. Finally, we introduce machine-learning algorithms including the support vector machine (SVM), random forest (RF), and K-nearest neighbor (KNN) for activity classification.

In summary, the main contributions of our work are as follows:

(1) In addition to amplitude information, we derive the phase information from raw CSI through a linear transformation.
(2) Principal component analysis method is applied to solve the problem of subcarrier selection for correlated activity recognition. The basis of our selection is the contribution rate of current subcarriers.
(3) We extract multiple features from CSI amplitude and phase to discuss the performance of different features on human activities, and then select features that best describe human activities for feature matrix.
(4) We analyze some key factors such as the kind of features and the number of principal components to summarize how they affect the detection accuracy.
(5) The experiments are carried out in two different scenarios and the performance of the system in different scenarios is verified in section 6.

The rest of this paper is organized as follows. We review the related works in the field of human activities recognition in Section 2. In Section 3, we briefly introduce some preliminaries about CSI. Section 4 presents an overview of the system architecture. Implementation of our system and experiment evaluation are presented in Section 5 and Section 6. Section 7 discusses the application of the system. Finally, Section 8 concludes the paper.

2. Related Work

Wi-Fi based activities detection systems have made a lot of progress in recent years, which is mainly divided into two types: RSS-based systems and CSI-based systems.

RSS-based systems: The RSS-based systems mainly take advantage of the changes of received signal strength (RSS) caused by human activities. Sigg S et al. propose the software-defined-radio (SDR)-based device-free activity recognition (DFAR) system [13]. This system leverages the localization of activities performed, the characteristics of environment and the difference of walking speeds to realize activity recognition. [14] explores the influence of human activities on RSS, and achieves the average recognition ratio (from 75% to 92.58%) with proposed fusion algorithm. RSS can also be used for gesture recognition. [15] uses the RSS at a mobile phone for recognition and achieves an accuracy of 51% when distinguishing 11 gestures and can reach 72% for four gestures. Wi-Gest can extract the three changes of rise, fall and pause according to the influence of different gestures on RSS waveform to build different gesture clusters, and map these gesture clusters to different application operations to achieve gesture recognition [16]. However, RSS is easily interfered by the environment, which is more suitable for recognizing some coarse-grained activities (such as sitting, walking, etc.). Compared to RSS-based systems, our system can identify more fine-grained activities and achieve higher recognition accuracy.

CSI-based systems: Researchers have been seeking for more stable information that can be used for passive human detection due to the unsatisfactory application of RSS. Since [12] is published, a series of CSI-based activities detection systems have been proposed [17, 19-21, 26, 31]. Compared with RSS, CSI can be a more fine-grained signal feature, which characterizes the multipath effect at the granularity of the OFDM subcarrier in the frequency

domain [18]. CSI-based activity recognition systems also leverage variations in signal measurement which is similar to RSS-based systems. Wi-Fall describes the indoor radio signal transmission model under the interference of human activities, verifying the feasibility of CSI for activities detection [17]. In addition, it utilizes the time stability and frequency diversity of CSI for fall detection. E-eyes extracts CSI amplitude information to detect walking and in-place activities, and uses CSI histogram as fingerprint information to detect daily human activities [19]. To analyze the relationship between human activities and CSI changes, CARM proposes CSI-speed and CSI-activity models, and then extracts feature information for classification using discrete wavelet transform (DWT) method [26]. In these studies, only amplitude information provided by CSI is used instead of considering the phase information. Feng et al. propose Wi-chase, which utilizes all 30 subcarriers to realize activities classification [21]. However, Wi-chase only classifies three simple activities, such as walking, rather than multiple fine-grained activities. In [31], a device-free system is proposed to identify seven activities of older persons. Unlike Wi-chase and the system in [31], our dataset contains more human activities.

Moreover, since these studies neglect to discuss how the extracted features represent human activities, in this paper, we leverage both CSI amplitude and phase information for feature extraction and discuss the performance of different features.

3. Channel State Information

CSI is fine-grained information in the physical (PHY) layer, which describes the channel properties of the communication link, including scatter, environmental attenuation, distance attenuation and other information. In the Wi-Fi OFDM system, the effects of multipath can be effectively reduced, and channel impulse response (CIR) between the transmitter and receiver is parsed from the PHY layer in the form of CSI. In a narrow-band flat fading channel, CSI can be expressed in the frequency domain as

$$y = Hx + n \qquad (1)$$

Where y, x, n and H represent the received vector, the transmission vector, the noise matrix and the channel matrix respectively.

In OFDM system, 56 subcarriers can be used for data transmission with a 20 MHz bandwidth channel. However, the Intel 5300 NIC can only capture 30 subcarriers [12], which can be expressed as

$$H = [H_1, H_2, ..., H_i, ..., H_N]$$
$$(2)$$

Where N is the number of subcarriers, and each H_i describes the amplitude and phase information of one OFDM subcarrier.

4. Overview

Our proposed system leverages CSI to indicate the human activities in indoor environments and realize different daily activities recognition. **Fig. 2** gives an overview of the system, which is consisted of four modules: sensing, data pre-processing, subcarriers selection and activities recognition. First, the raw CSI data is extracted from commercial Wi-Fi devices in the sensing module, and then we feed raw CSI streams into data pre-processing module to filter the noise and outliers. In the activity recognition module, we extract features from the phase and amplitude information respectively. Finally, these features are passed to the machine-learning algorithms to realize activities classification.

The core of our system is the construction of features matrix. The system distinguishes between full-body activities and upper body activities and the performance of extracted features differ in different type of activities. In general, a full-body activity (such as walking and running) causes significant pattern changes of the CSI amplitude over time because of its significant body movements. An upper body activity (such as throwing debris) only involves upper body movements, which will not cause significant amplitude changes. Moreover, we use the phase and amplitude information as indicators, and the performance of features is also different under different indicators. Therefore, to improve the identification accuracy and efficiency of the system, we select multiple features to construct feature matrix through multi-angle analysis and extensive experimental evaluation.
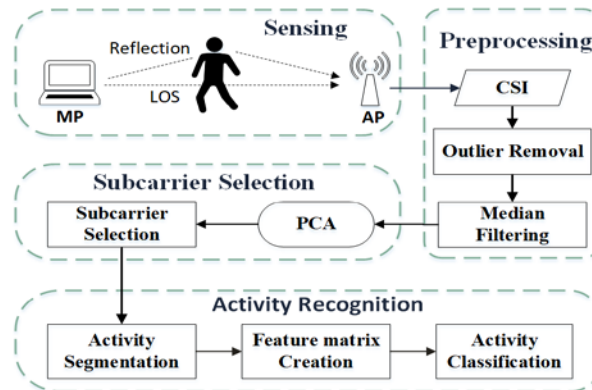

Fig. 2. Overview of System Architecture.

5. Methodology

In this section, we will elaborate on the overall design of the system. First, in the sensing module, we collect the raw CSI data in thirty subcarriers and nine streams. Next, in the data pre-processing module, the CSI phase information is extracted through a liner transformer. Meanwhile, the noise and outliers of CSI amplitude information are filtered by Hampel filter and median filter [24]. Since the CSI contains thirty subcarriers and these subcarriers are relatively similar, we apply PCA method on thirty subcarriers to obtain the most meaningful subcarriers in subcarriers selection module. Finally, in activity recognition module, we extract different features as input of machine-learning algorithm.

5.1 Data Pre-processing

Since the collected raw CSI data contains a lot of noise, in this part, raw CSI will be preprocessed for next work. According to Section 3, the raw CSI data includes phase and amplitude information. We extract the phase and amplitude information respectively, and then remove the noise using a series of filters.

5.1.1 Phase Sanitization

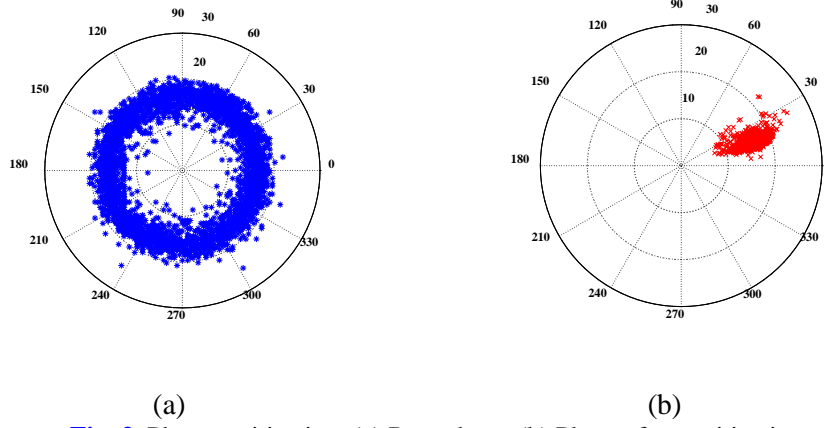(a)   (b)

Fig. 3. Phase sanitization. (a) Raw phase, (b) Phase after sanitization

Although CSI has been widely utilized for various applications, most of them only consider the amplitudes of CSI, and phase information does not attract enough attention, the reason of which is the synchronization error between the transmitter and receiver cannot be completely eliminated. The CSI information received at the receiver contains both the residual error of the carrier frequency synchronization and the clock synchronization error, which will have great influence on the phase information. Thus, in this section, we employ a linear transformation on the raw phase readings to obtain available phase information for activities detection [22]. The measured CSI phase $\hat{\phi}_i$ in the i_{th} subcarrier can be expressed as follows

$$\hat{\phi}_i = \phi_i - 2\pi \frac{k_i}{N} \delta + \beta + Z \tag{3}$$

Where $\hat{\phi}_i$ and ϕ_i represent the measured CSI phase and the true phase values of the i_{th} subcarrier, respectively. δ represents the time difference between receiver and transmitter. β is the phase deviation caused by the unsynchronization of the center frequency at the receiver and transmitter, and Z is Gaussian noise. k_i ranging from -28 to 28 denotes the index of the 30 subcarriers. N is the fast Fourier transform (FFT) size. As shown in **Fig. 3(a)**, the raw CSI phase information is disorganized and unusable. To reduce the influence of β and δ, a linear transformation recommended in [23] is used on the raw phase information. The slope and intercept of the linear transformation equation are defined as a and b respectively.

$$a = \frac{\hat{\phi}_n - \hat{\phi}_1}{k_n - k_1} = \frac{\phi_n - \phi_1}{k_n - k_1} - \frac{2\pi}{N} \delta \tag{4}$$

$$b = \frac{1}{n} \sum_{j=1}^{n} \hat{\phi}_j = \frac{1}{n} \sum_{j=1}^{n} \phi_j - \frac{2\pi\delta}{nN} \sum_{j=1}^{n} k_j + \beta \tag{5}$$

Since the frequency of subcarriers is symmetric [12], $\sum_{j=1}^{n} k_j = 0$ and b can be expressed as $b = \frac{1}{n} \sum_{j=1}^{n} \phi_j + \beta$. Noise Z is caused by measurement, which can be ignored when multiple measurements are taken. And then we can calculate $\hat{\phi}_i - ak_i - b$ to get $\tilde{\phi}_i$.

$$\tilde{\phi}_i = \hat{\phi}_i - ak_i - b = \phi_i - \frac{\phi_n - \phi_1}{k_n - k_1} k_i - \frac{1}{n} \sum_{j=1}^{n} \phi_j \tag{6}$$

Where $\hat{\phi}_i$ and $\tilde{\phi}_i$ denote the measured CSI phase and the approximate true phase. It is found that β and δ have been eliminated from the equation after using a linear transformation. Compared to **Fig. 3(a)**, the phase in **Fig. 3(b)** distributes relatively stable as expected.

5.1.2 Outlier Filter

It is found that there are always some outliers in the collected CSI data, which are obviously not caused by human activities. As shown in **Fig. 4(a)**, the amplitude information of raw CSI data extracted from the static environmen contains many outliers. Even in the non-human scenario, the collected signals contain a lot of ambient noise. We apply a Hampel filter to remove these deviated outliers, which classifies all points outside the closed interval [μ-γ*σ, μ+γ*σ] as outliers [24]. Where μ and σ are the median and the median absolute deviations (MAD) respectively. γ is set to 3 in this paper according to the actual application setting. The amplitude of the CSI after removing outliers is shown in **Fig. 4(b)**. It can be seen that the Hampel filter can effectively remove these outliers. To avoid signal fluctuations at the beginning and end of data collection, and fully capture the motion information, including the preparation and end of the motion, we set the size of the time window to 20 seconds.
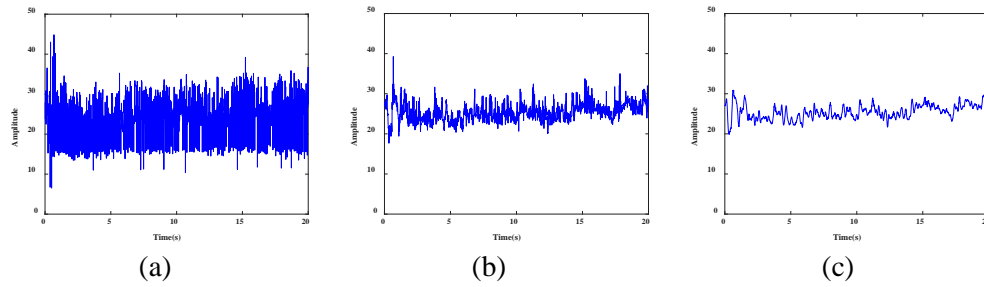

(a)   (b)   (c)

Fig. 4. Noise filtering for CSI amplitude. (a) Raw CSI amplitude, (b) CSI after Hampel filtering, (c) 10-poingts after median filtering.

5.1.3 Noise Filtering

After removing outliers, ambient noise will also be removed. Frequency domain approaches like low-pass filter (e.g. Butterworth) is employed at first time. However, we find that there is still some noise when the cut-off frequency of low-pass filters (LPF) is not small enough, and some meaningful information will be filtered out when the cut-off frequency is too small. As a result, in this paper, we apply a median filter to remove the environment noise. Median filter (MF) uses the means as the filtered value after removing a maximum and a minimum in a sliding window. Therefore, compared with LPF, MF can avoid noise residue and information loss generated from improper cut-off frequency setting. We apply 10-poingts median filter on CSI amplitude after Hampel filtering in **Fig. 4(b)**. The result after median filtering is illustrated in **Fig. 4(c)**.

5.2 Principal Component Analysis (PCA) and Subcarriers Selection

Principal Component Analysis (PCA) method is a common method of data dimension reduction [25]. Dimensionality reduction is a kind of dataset pre-processing technology, which can make data processing simpler and more efficient while minimizing the loss of information. Moreover, since random noise will not cause the change of signal principal components (PCs), PCA also can remove signal noise at the same time.

There are 30 subcarriers in each CSI stream, and most of subcarriers are correlated to some extent [26]. If they are all used for feature extraction, the calculation will be too heavy. As a result, PCA is used to select the principal components containing the most activity information.

In PCA, the filtered CSI information is taken as input to calculate the covariance matrix of the centralized data [27]. Then the matrix is decomposed into eigenvectors and the eigenvectors are sorted in descending order so that the larger eigenvectors come first. After obtaining the eigenvectors, we determine the first few principal components by calculating their contribution rates (*CRate*).

$$CRate = \frac{\text{Sum of selected eigenvectors}}{\text{Sum of all eigenvectors}} \tag{7}$$

The eigenvectors whose *CRate* are more than *C*% can be screened out by the above equation, and then the corresponding principal components can be obtained. **Fig. 5(b)** shows the CSI fluctuation of first 2 principal components (*CRate* is set as 95%), which removes a lot of redundant information compared to **Fig. 5(a)**. In our experiment, we apply PCA on all 30 subcarriers and select the subcarriers corresponding to eigenvectors with large variance according to *CRate*. In our later experimental evaluation, it is found that the best result is achieved when we select the first five subcarriers (i.e., when the *CRate* is set to about 90%).
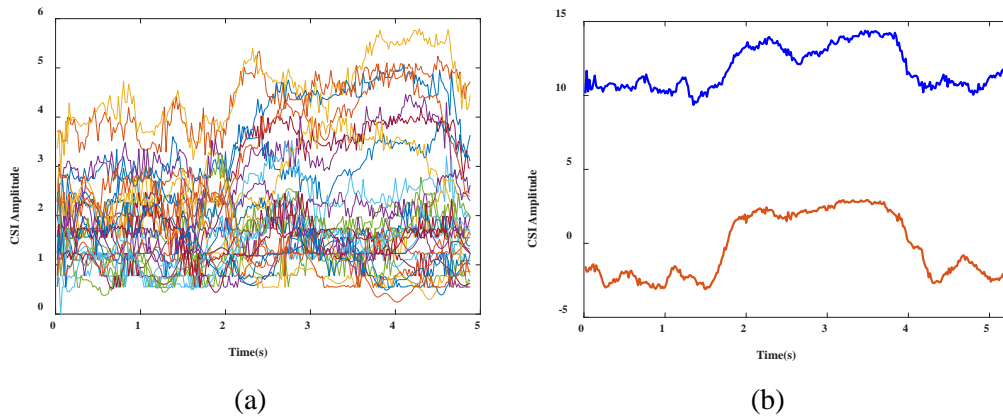

| (a) | (b) |

Fig. 5. Principle Components Analysis (PCA) on CSI amplitude.
(a) Raw CSI amplitude, (b) The first two PCs.

5.3 Activity Recognition

This section focuses on human activities recognition algorithm in the system. We will explain the construction of features matrix and the machine learning algorithms used for activities classification.

5.3.1 Constructing Feature Matrix

In this part, we will elaborate on the construction process of feature matrix. The feature matrix determines the final performance of activity recognition. Selecting suitable features can ensure the accuracy of classification and avoid excessive data processing time. In the previous work, the amplitude and phase information of filtered CSI have been obtained. We extract features from the amplitude and phase information respectively, and combine the features that

are significantly different into eigenvectors by analyzing the similarity between these features. In addition, the differences of amplitude and phase features are not exactly the same, so we analyze these features separately. A sliding window method is applied across CSI to calculate the following features.

1) Variance: One feature is variance $\sigma^2(k)$ that is a measure of the dispersion of a set of data. The variance of CSI value x_i in the kth sliding window of length n can reflect the degree of change of CSI caused by human activities. Mathematically it can be expressed as

$$\sigma^2(k) = \frac{1}{n}\sum_{i}^{n}(x_i - \mu)^2 \tag{8}$$

Where μ is the mean value of in the kth sliding window and $k \in (1, N)$ represents the number of sliding windows.

2) Normalized standard deviation (STD): The second feature is STD $\sigma(k)$. Compared with variance, STD is more intuitive to describe the degree of deviation between data and mean value. We can formula STD as

$$\sigma(k) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2} \tag{9}$$

3) Median absolute deviation (MAD): The third feature in our system is median absolute deviation $MAD(k)$. MAD represents the variability in human activity. It can be expressed as

$$MAD(k) = \frac{1}{n}\sum_{i=1}^{n}|x_i - \mu| \tag{10}$$

4) Interquartile range (IR): The fourth feature is the interquartile range (IR) $P_{25}(k)$. IR is the value at third segmentation points after CSI in the window are arranged from small to large and divided into four equal parts. Mathematically, it can be calculated by

$$P_{25}(k) = x(\frac{25}{100} \cdot n) \tag{11}$$

5) The offset of signal strength (Offset): The fifth feature is the offset of signal strength $offset(k)$. Human activities also can cause changes in signal strength, and the offset can reflect these changes. It can be expressed as following.

$$offset(k) = |\frac{1}{n}\sum_{i}^{n}x_i| \tag{12}$$

6) Signal entropy (SE): The last feature is signal entropy $SE(k)$. This serve as an important feature to distinguish between different types of human activities. We first figure out the probability distribution of the signal and then we calculate the signal entropy. The mathematical equation is as follows.

$$SE(k) = -\sum_{i=1}^{n}p(x_i) \cdot \log_2 p(x_i) \tag{13}$$

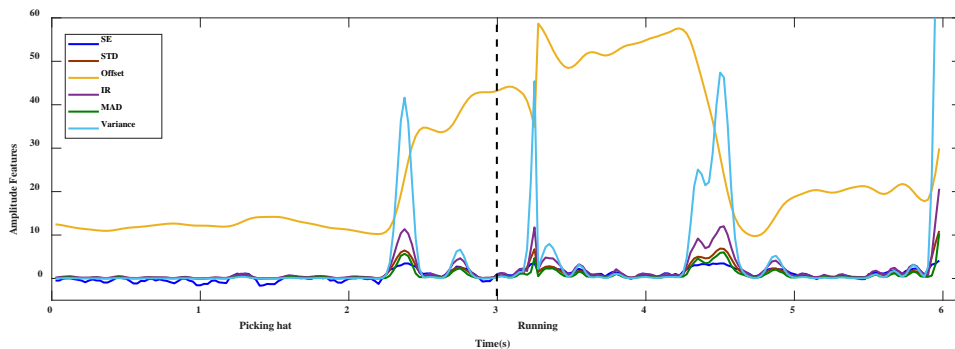Where $p(x_i)$ is the probability distribution of CSI.

Fig. 6. Features extracted from amplitude of picking hat and running.

Our experiment mainly includes the full-body movements (i.e. running) and upper body movements (i.e. picking hat). To select more representative features, we analyze the performance of all the features from full-body movement and upper body movement at the same time. **Fig. 6** shows the amplitude features extracted from running and picking hat. It can be seen that the most of the features in **Fig. 6** fluctuate slightly around 0. Among them, the offset fluctuates more significantly and has completely different patterns of fluctuation compared with other features. Therefore, offset is taken as one of our amplitude features. In addition, the remaining five features change similarly with the implementation of human activity. We select the signal entropy which includes more human activities information. Then we extract phase features to observe the performance of these features, which is shown in **Fig. 7**. It presents that the signal entropy, offset and interquartile range fluctuate significantly than other features. Meanwhile, there is little correlation between their fluctuations. Therefore, we finally select SE and offset extracted from amplitude information and IR, SE, and offset extracted from phase information. In addition, we also used a useful feature, the duration of human moving (duration), which illustrates the duration of a human activity.
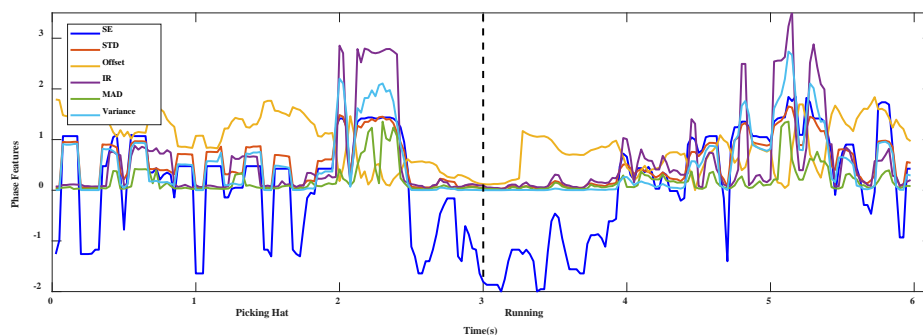

Fig. 7. Features extracted from phase of picking hat and running.

Meanwhile, our system is not limited to one person, but to different subjects for activity identification. How to distinguish similar actions made by different subjects has also become a problem to be studied. Since the experimental subjects have different heights and volumes, their influence on the line-of-sight (LOS) path is also quite different, and each subject will cause different reflection paths, which also leads to differences in the collected CSI data. Therefore, even if the subjects take similar actions, it can still be distinguished. At the same

time, in order to verify that the features we extract can show the differences in the actions of different people, we compared the picking hat action of three subjects, as shown in **Fig. 8**. **Fig. 8(a)** and **Fig. 8(b)** represent the amplitude features that we selected and the rest is the phase feature. **Fig. 8(c)**, **Fig. 8(d)** and **Fig. 8(e)** in the figure represent different subjects. It is illustrated from the figure that even if the experimental subjects take same actions, the features we extract from the action also have certain differences, which helps us to distinguish similar actions of different subjects.
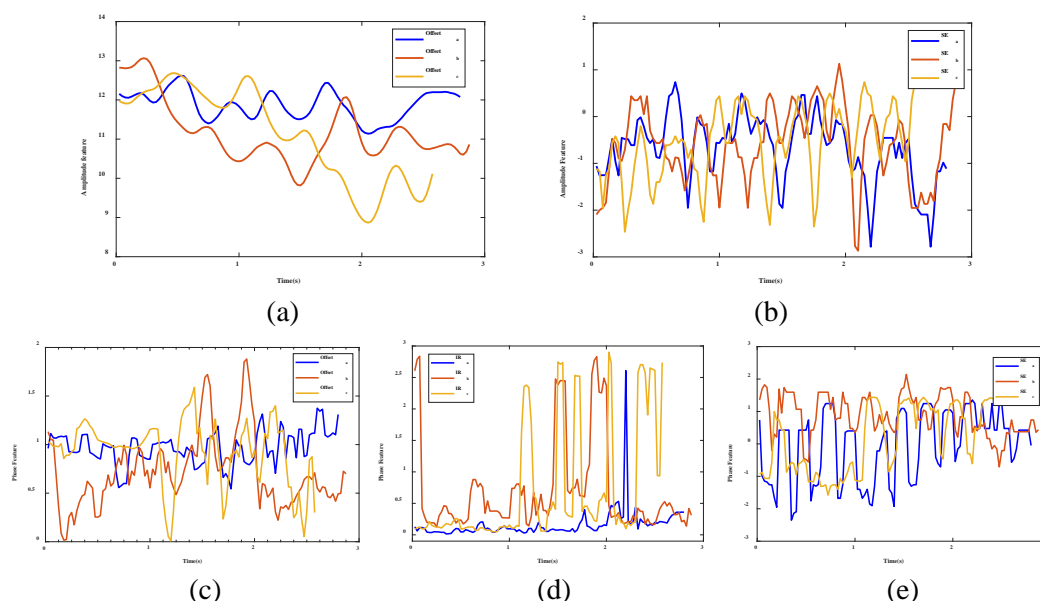

(a)   (b)

(c)   (d)   (e)

Fig. 8. Features extracted from picking hat activity. (a) Offset from amplitude, (b) SE from amplitude, (c) Offset from phase, (d) IR from phase, (e) SE from phase.

5.3.2 Activity Classification

To realize activity classification, we use support vector machine (SVM) as classification algorithm [28]. SVM is a useful technique for data classification, which maps input vector nonlinear to a very high-dimensional feature space, and constructs a linear decision surface. The special properties of this surface can ensure that the learning machine has high generalization ability. The kernel function is used to map the input vector to the high-dimensional feature space and then the maximum edge hyperplane in the transformed feature space can be found.

We utilize libsvm toolbox to build our SVM classification model [29]. The extracted features and corresponding labels are used as input of SVM classifier to construct the SVM classification model. In addition, Gaussian kernel function is used to solve nonlinear classification problem [30].

Meanwhile, two other classification algorithms (KNN and RF) are compared with SVM for better classification effect. KNN is a common algorithm in wireless signal location, and we consider using it in activities recognition to see the classification effect. The core idea of KNN is that if most of the K most adjacent samples of a sample in the feature space belong to a certain category, then the sample also belongs to this category and has the characteristics of samples in this category. In [31], the author uses three different machine-learning algorithms for activity recognition, including RF, SVM and Naive Bayes. Among them, RF achieves

higher recognition accuracy. Therefore, RF is also applied in our experiment. The basic unit of RF is decision tree, each decision tree is a classifier. Different decision trees may have different classification results, and RF algorithm integrates these classification results to specify the category with the most votes as the final output.

6. Evaluation

In this section, we present the implementation and evaluation results of our human activity detection system using commercial off-the-shelf Wi-Fi devices.
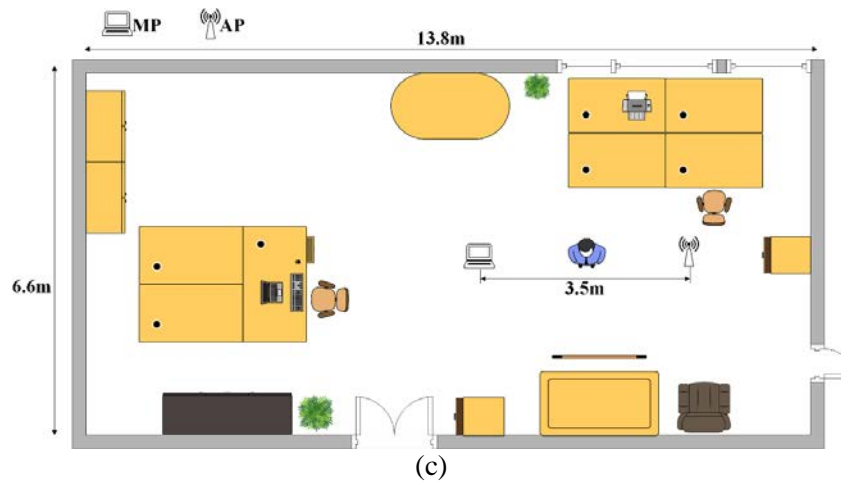
6.1 Experimental Setups

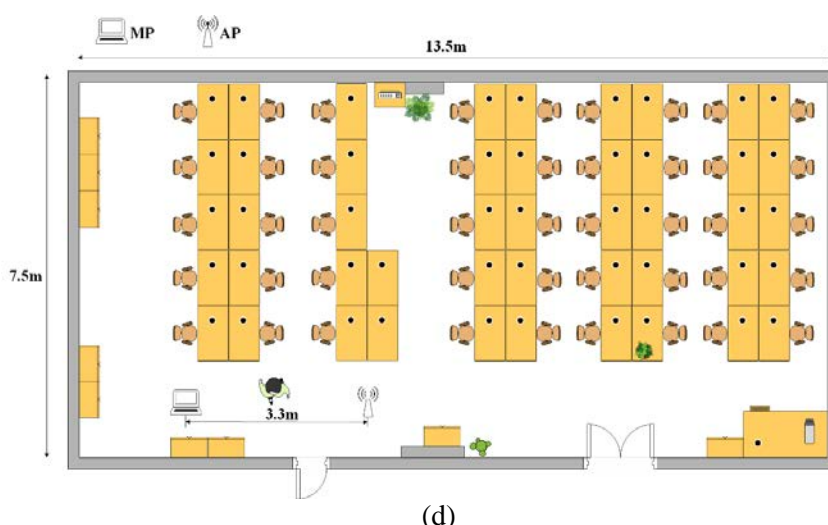(a)   (b)

(c)

(d)

Fig. 9. Photographs and layouts of the two experiment. (a) Experimental scenario of the 1st experiment, (b) Experimental scenario of the 2nd experiment, (c) Layout of 1st experiment, (d) Layout of the 2nd experiment.

6.1.1 Experimental Environments

We perform our detection system using an 802.11n Wi-Fi network consisting of an off-the-shelf Wi-Fi devices (laptop) and a commodity wireless access point (i.e., D-Link DIR-859 Router with three antennas). The laptop is installed Ubuntu 14.04 system with a modified Intel NIC driver.

The software that is used in our experiments is the open source CSI-Tool presented by Halperin et al. [12], and Matlab is used to analyze the collected data. Since there exist other wireless devices working in 2.4GHz in the laboratory, we run our experiment in 5GHz frequency to avoid interference from other devices. Moreover, we set the sampling frequency to 100Hz in order to fully capture CSI information caused by human activities.

To verify the scalability of our system in different environments, we conduct our experiments in two different laboratories. As shown in **Fig. 9(c)**, the length and width of the first laboratory is given, which is surrounded by many experimental equipment and desks. The location of AP and MP is shown in **Fig. 9(a)**. Moreover, the second laboratory is 13.8m in length and 6.6m in width, and is more crowded with relatively more multipath. The AP and MP are located between the cabinets and the desks, as shown in **Fig. 9(d)**.

6.1.2 Dataset

We recruit five volunteers (ages 22 to 32; two women and three men; weight: 48-105kg; height: 1.62-1.85m) to perform various daily activities in the two test environments over one months. The collected data includes upper body activities and full-body activities. Finally, we get 1600 samples listed in **Table 1**.

Table 1. Summary for activity dataset.

Activity	Samples
Walk (W)	200
Pick hat (P)	200
Throw debris (T)	200
Sit down (SD)	200

Drink water (D)	200
Wave arm (W)	200
Run (R)	200
Sleep (SP)	200

6.2 Performance Evaluation

In this section, we first evaluate the overall performance of our system and its robustness under different experimental conditions.
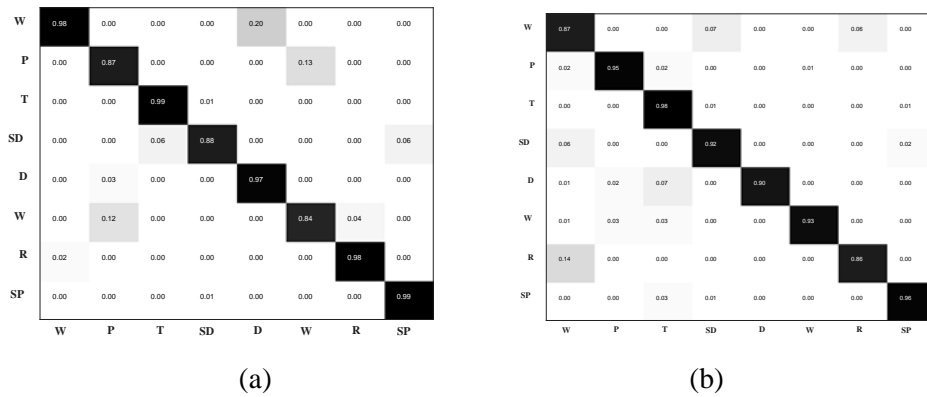
6.2.1 Evaluation Metrics

Fig. 10. (a) A confusion matrix for the classification of a person's different activities, (b) The confusion matrix for the classification of similar activities of five people.

Fig. 10 shows the confusion matrix of the eight activities. Each row in the matrix represents the true category of the activity, and each column represents the predicted category by our system. **Fig. 10(a)** shows the result in the case of a single person. It can be seen that the average classification accuracy of our system is 0.95. The high classification rate is 0.99 (sleeping and throwing debris), the reason of which is that these two activities have special modes of execution. In contrast, the lowest rate is 0.84 (walking), because walking may be incorrectly identified as running. Meanwhile, our system is also applied to all the data collected from the five objects. As shown in **Fig. 10(b)**, we classify eight similar activities of five people and achieve a certain degree of accuracy.

Moreover, we have calculated the three parameters of Precision, Recall and F-score to evaluate the performance of the system. These three parameters are calculated as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{14}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{15}$$

$$\text{F-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{16}$$

Where TP, FP and FN are the true positive, false positive and false negative respectively. As shown in **Fig. 11**, we compare the result of the precision, recall and F1 score of the eight experimental activities with Wi-chase [21] and Wi-Fall [17], which shows that the three rate of

our system are about 0.95 while the results of Wi-chase and Wi-Fall are not good enough. This is because that Wi-chase only uses a few mathematical features that inadequately represent movement changes and it may be better suited to identifying a small number of simple activities than to our data set. As for Wi-Fall, it neglects to take full advantage of the phase information. Compared with Wi-chase and Wi-Fall, in our system, features that can better describe human activities are selected to form the feature matrix and we can get more useful information for classification.
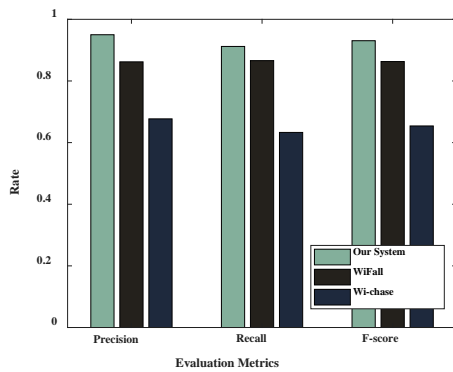

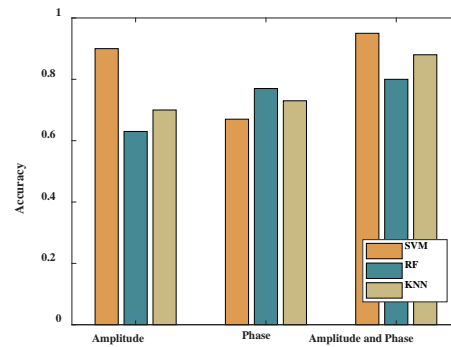**Fig. 11.** Precision, Recall and F-measure results.

Fig. 12. Detection accuracy with different basic Signals

To verify the validity of phase information, the classification accuracy is compared in three cases: only using amplitude, only using phase, using both amplitude and phase in **Table 2**. Moreover, we also test the performance of three classification algorithms, SVM, random forest (RF) and K-nearest neighbor (KNN) on our system. It is shown in **Fig. 12** that SVM algorithm performs better than other algorithms. Meanwhile, when we use SVM for classification, the accuracy achieves 0.9 and 0.82 with only amplitude or phase while it can achieve 0.95 with both amplitude and phase. Therefore, it can be seen that the addition of the phase can improve the accuracy of our system.

Table 2. Detection accuracy with three classification algorithms.

Algorithms	Amplitude	Phase	Amplitude and Phase
SVM	0.90	0.67	0.95
RF	0.63	0.77	0.80
KNN	0.70	0.73	0.88

6.2.2 Impact of Principal Components

In our system, the principal components of CSI are obtained for subcarriers selection by utilizing PCA. To show the validity of the Principal Component Analysis (PCA) method, we test the accuracy of each implemented activity with PCA and without PCA. **Fig. 13(a)** shows that the system accuracy decreased without using PCA algorithm. In addition, in this section, we also discuss the impact of the number of selected components. As shown in **Fig. 13(b)**, the system achieves better accuracy when the first five principal components are selected. However, increasing the number of principal components will not significantly improve the accuracy, which also leads to slower computing speed.
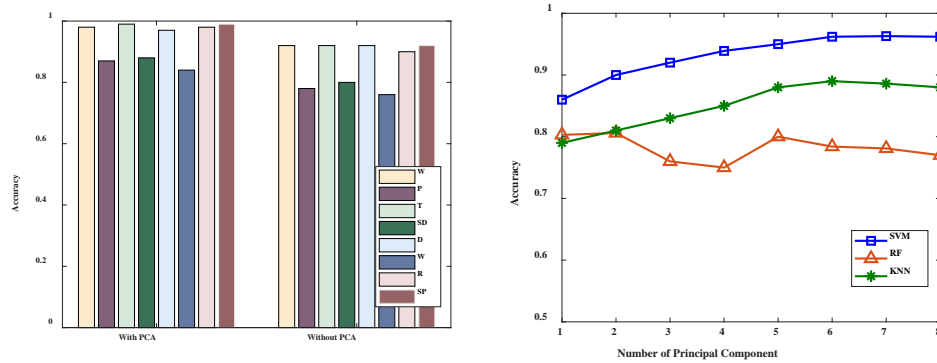
Fig. 13. The impact of PCA. (a) Comparison of system precision with PCA and without PCA, (b) Detection accuracy with PCs.

6.2.3 Impact of Features

In this experiment, there are multiple activities containing fine-grained activities to be classified, and our system cannot achieve expected accuracy with a single statistical feature. As a result, we extract multiple features from amplitude and phase and then select more useful features to construct the feature matrix according to the performance and the correlation between features. In this section, we analyze the impact of different feature and the kind of feature on the performance of our system. In addition, we compare the performance of different features in two different environments to test the feasibility of features.

1) Performance of Different Features:

Fig. 14(a) compares the performance of different features extracted from the amplitude of CSI, including variance (Variance), signal entropy (SE), offset of signal strength (Offset), interquartile range (IR), median absolute deviation (MAD), normalized standard deviation (STD) and duration of human moving (duration). The duration of human moving is a fixed value, so we put it in the figure of amplitude features. The result shows that the classification accuracy of single feature is around 0.3, and the offset of signal strength can achieve the accuracy of about 0.6. Moreover, the performance of single feature extracted from the phase of CSI is compared in **Fig. 14(b)**, which shows that the classification accuracy of phase feature is slightly lower than the amplitude feature, among which the offset and signal entropy are relatively stable in different environments.
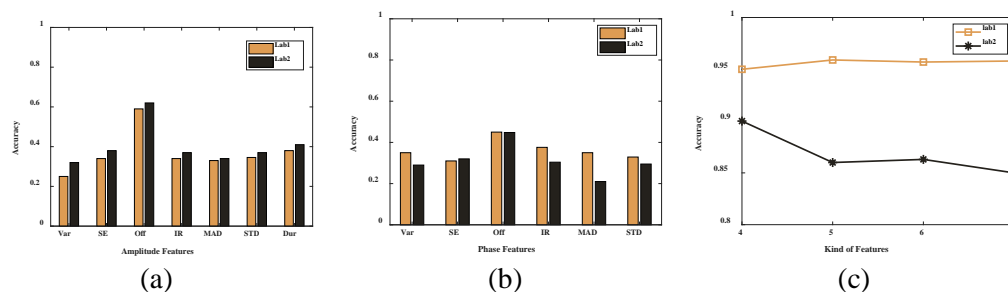

| (a) | (b) | (c) |

Fig. 14. Detection accuracy with different features. (a) Amplitude features, (b) Phase Amplitude features, (c) Kind of features.

Therefore, offset, signal entropy, IR and duration are selected to form the feature matrix for the requirement of accuracy in the experiment. We also discuss the impact of the increase of the kind of the features, as shown in **Fig. 14(c)**, and it presents that accuracy is not improved significantly and even slips a bit in lab2 with the increase of the kind of features. This is because not all features can describe human activities in different ways, and different features may have similar performance. In addition, since the SVM algorithm is targeted at small samples, if the number of samples is large, overfitting may occur.

2) Impacts of Sliding Window Size

Fig. 15 illustrates the performance of different sliding window sizes. The performance of our system is compared in two different environments. As shown in **Fig. 15**, the detection accuracy increases with the increase of window size in two schemes. With a larger window size, we get more feature samples in the time dimension that might mitigate the impact of temporal variance, but the rate might will keep stable when the size reaches some threshold.
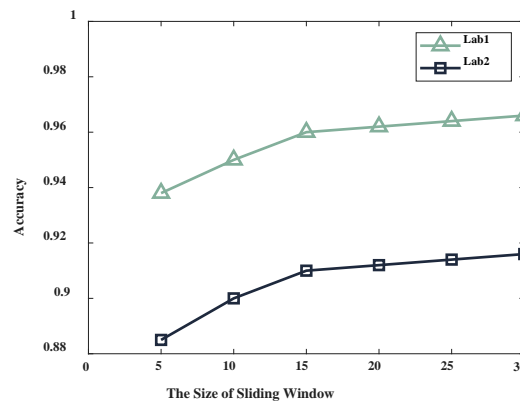

Fig. 15. The impact of different size of sliding window.

6.2.4 Impact of Different Environment

In this section, the impacts of different experimental environments on the performance of our system are analyzed. We performed the experiment in two laboratories mentioned in section 6.1. As shown in **Fig. 16**, it can be seen that the most of the activities, whether in the lab1 with less multipath or in the lab2 with more multipath, can achieve the accuracy of more than 0.8, which prove the robustness of our system. In addition, the accuracy of running activity is just over 0.7, which is due to the scenario limitations of lab2. In this environment, the object can only move in a limited area and cannot perform normal running activities because MP and AP are located between desks and cabinets. These limitations lead to the possibility that running can be misclassified as an activity such as walking. Meanwhile, sitting activity will be identified as sleeping both in two environments because they all have a still period.
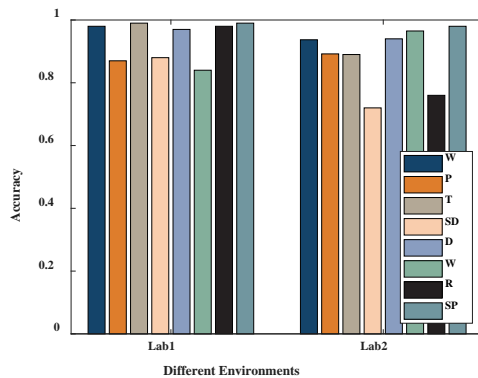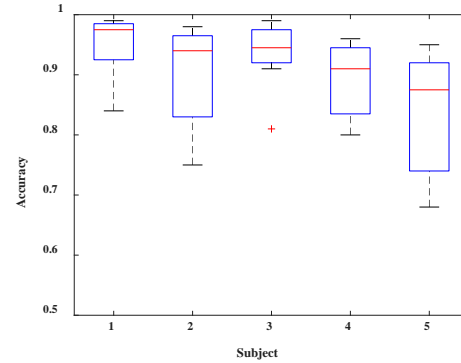
Fig. 16. The impact of different environments.

Fig. 17. The impact of different people.

6.2.5 Impact of Different People

The impacts of user diversity on the accuracy of activity recognition in our experiment are also evaluated. Five different volunteers varying in age, height, weight and gender are recruited. As shown in **Fig. 17**, the activity classification accuracy of them reach a high level and the average accuracy is above 0.85, which proves that this system still performs well to different users.

7. Discussion

In this section, We will discuss the limitations of the system and some problems in application from the following three aspects

1) Multi-person activities recognition: Our system is designed for recognizing the activities of a single person. We find that if only one person moves in the room and the others are still, our system can still work. These non-moving people can be considered part of the surrounding. However, if the target and other people move at the same time, the reflected signals they produce affect each other, making it difficult to distinguish the target's movements. Therefore, we consider the single person scenario when applying our system, or the case where only one person moves.

2) Impact of environmental changes: For our system, the premise of activity identification is to build a training set of a certain size, which is easily influenced by the environment. If you change to a new environment, the system will often require extra effort to retrain, which will result in additional burden on the system. For an untrained environment, the ability of our system to recognize also changes.

3) Location and orientation of users: In addition to the above two aspects, the location and orientation of target personnel also have an impact on the recognition. For the human activities, the differences on the location and orientation of the user could induce differesnt reflection paths, resulting in changes in CSI measurements.

8. Conclusion

In this paper, a novel human activity detection system based on CSI has been proposed. The system overcomes the limitations of traditional systems and achieves better accuracy. In addition to the amplitude information, we also extract the phase information using a linear transformation. And a series of filters are used to remove the noise of CSI. To solve the

problem of subcarriers selection, we employ PCA method and keep the first five principal components. Compared to other work, this system focuses on the performance of different features on human activity. We discuss the sensitivity of these features to human activity and the correlation between them, and finally select several unrelated sensitive features for feature matrix. Moreover, we apply three machine learning algorithms to realize activity classification, among which SVM has the best performance. The experiment is carried out in two different laboratories, and the results confirmed the robustness of our system.

References

[1] Reyes-Ortiz, J. L., Oneto, L., Samà, A., Parra, X., and Anguita, D., "Transition-aware human activity recognition using smartphones," *Neurocomputing*, vol. 171, pp. 754-767, January, 2016. Article (CrossRef Link)

[2] Shoaib, M., Bosch, S., Incel, O. D., Scholten, H., and Havinga, P. J., "A Survey of Online Activity Recognition Using Mobile Phones," *Sensors*, vol. 15, no. 1, pp. 2059-2085, 2015. Article (CrossRef Link)

[3] Donahue, J., Anne Hendricks, L., Guadarrama, S., Rohrbach, M., Venugopalan, S., Saenko, K., and Darrell, T, "Long-term recurrent convolutional networks for visual recognition and description," in *Proc. of the IEEE conference on computer vision and pattern recognition*, pp. 2625-2634, 2015. Article (CrossRef Link)

[4] Xia, L., Gori, I., Aggarwal, J. K., and Ryoo, M. S, "Robot-centric activity recognition from first-person rgb-d videos," in *Proc. of 2015 IEEE Winter Conference on Applications of Computer Vision*, pp. 357-364, January, 2015. Article (CrossRef Link)

[5] Kwapisz, J. R., Weiss, G. M., and Moore, S. A, "Activity recognition using cell phone accelerometers," *ACM SigKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74-82, 2011. Article (CrossRef Link)

[6] Alsheikh, M. A., Selim, A., Niyato, D., Doyle, L., Lin, S., and Tan, H. P, "Deep activity recognition models with triaxial accelerometers," in *Proc. of Workshops at the Thirtieth AAAI Conference on Artificial Intelligence*, March, 2016. Article (CrossRef Link)

[7] Catal, C., Tufekci, S., Pirmit, E., and Kocabag, G., "On the use of ensemble of classifiers for accelerometer-based activity recognition," *Applied Soft Computing*, vol.37, pp. 1018-1022, 2015. Article (CrossRef Link)

[8] Lara, O. D., and Labrador, M. A., "A survey on human activity recognition using wearable sensors," *IEEE communications surveys & tutorials*, vol. 15, no. 3, pp. 1192-1209, 2012. Article (CrossRef Link)

[9] Ordóñez, F., and Roggen, D., "Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition," *Sensors*, vol.16, no. 1, pp. 115, 2016. Article (CrossRef Link)

[10] Jiang, W., and Yin, Z., "Human activity recognition using wearable sensors by deep convolutional neural networks," in *Proc. of the 23rd ACM international conference on Multimedia*, pp. 1307-1310, October, 2015. Article (CrossRef Link)

[11] Yang, Zheng, Zimu Zhou, and Yunhao Liu, "From RSSI to CSI: Indoor localization via channel response," *ACM Computing Surveys (CSUR)*, vol. 46, no. 2, 2013. Article (CrossRef Link)

[12] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "802.11 with multiple antennas for dummies," *ACM SIGCOMM Computer Communication Review*, vol. 40, no. 1, pp. 19–25, 2010. Article (CrossRef Link)

[13] Sigg, S., Shi, S., Buesching, F., Ji, Y., and Wolf, L., "Leveraging RF-channel fluctuation for activity recognition: Active and passive systems, continuous and RSSI-based signal features," in *Proc. of International Conference on Advances in Mobile Computing and Multimedia*, December, 2013. Article (CrossRef Link)

[14] Gu, Y., Quan, L., and Ren, F, "Wi-Fi-assisted human activity recognition," in *Proc. of 2014 IEEE Asia Pacific Conference on Wireless and Mobile*, pp. 60-65, August, 2014. Article (CrossRef Link)

[15] Sigg, S., Blanke, U. and Tröster, G., "The telepathic phone: Frictionless activity recognition from Wi-Fi-rssi," in *Proc. of 2014 IEEE international conference on pervasive computing and communications (PerCom)*, pp. 148-155, March, 2014. Article (CrossRef Link)

[16] Abdelnasser, H., Youssef, M. and Harras, K. A., "Wigest: A ubiquitous Wi-Fi-based gesture recognition system," in *Proc. of 2015 IEEE Conference on Computer Communications (INFOCOM)*, pp. 1472-1480, April, 2015. Article (CrossRef Link)

[17] Wang, Y., Wu, K., and Ni, L. M., "Wifall: Device-free fall detection by wireless networks," *IEEE Transactions on Mobile Computing*, vol.16, no.2, pp. 581-594, 2016. Article (CrossRef Link)

[18] Bhartia, A., Chen, Y. C., Rallapalli, S. and Qiu, L., "Harnessing frequency diversity in wi-fi networks," in *Proc. of the 17th annual international conference on Mobile computing and networking*, pp. 253-264, September, 2011. Article (CrossRef Link)

[19] Wang, Y., Liu, J., Chen, Y., Gruteser, M., Yang, J. and Liu, H., "E-eyes: device-free location-oriented activity identification using fine-grained Wi-Fi signatures," in *Proc. of the 20th annual international conference on Mobile computing and networking*, pp. 617-628, September, 2014. Article (CrossRef Link)

[20] Palipana, S., Rojas, D., Agrawal, P., and Pesch, D., "FallDeFi: Ubiquitous fall detection using commodity Wi-Fi devices," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol.1, no.4, pp.155, 2018. Article (CrossRef Link)

[21] Arshad, S., Feng, C., Liu, Y., Hu, Y., Yu, R., Zhou, S., and Li, H., "Wi-chase: A Wi-Fi based human activity recognition system for sensorless environments," in *Proc. of 2017 IEEE 18th International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM)*, pp. 1-6, June, 2017. Article (CrossRef Link)

[22] K. Qian, C. Wu, Z. Yang, Y. Liu, and Z. Zhou, "PADS: passive detection of moving targets with dynamic speed using PHY layer information," in *Proc. of 2014 20th IEEE International Conference on Parallel and Distributed Systems (ICPADS)*, pp. 1–8, 2014. Article (CrossRef Link)

[23] S. Sen, B. Radunovic, R. R. Choudhury, and T. Minka, "You are facing the Mona Lisa: spot localization using PHY layer information," in *Proc. of the 10th International Conference on Mobile Systems, Applications, and Services - MobiSys '12, Low Wood Bay, Lake District, UK*, pp. 183–196, 2012. Article (CrossRef Link)

[24] Davies, L., and Gather, U., "The identification of multiple outliers," *Journal of the American Statistical Association*, vol. 88, no. 423, pp. 797-801, 1993. Article (CrossRef Link)

[25] Ringnér, M., "What is principal component analysis?," *Nature biotechnology*, vol. 26, no. 3, pp. 303-304, 2008. Article (CrossRef Link)

[26] Wang, W., Liu, A. X., Shahzad, M., Ling, K., and Lu, S., "Understanding and modeling of Wi-Fi signal based human activity recognition," in *Proc. of the 21st annual international conference on mobile computing and networking*, pp. 65-76, September, 2015. Article (CrossRef Link)

[27] Smith, L. I., "A tutorial on principal components analysis," 2002. Article (CrossRef Link)

[28] Cortes, C., and Vapnik, V., "Support-vector networks," *Machine learning*, vol. 20, no.3, pp. 273-297, 1995. Article (CrossRef Link)

[29] Chang, C. C., and Lin, C. J., "LIBSVM: A library for support vector machines," *ACM transactions on intelligent systems and technology (TIST)*, vol. 2, no. 3, pp. 27, 2011. Article (CrossRef Link)

[30] Hsu, C. W., Chang, C. C., and Lin, C. J., "A practical guide to support vector classification," 2003.

[31] Li, F., Al-Qaness, M. A. A., Zhang, Y., Zhao, B., and Luan, X., "A robust and device-free system for the recognition and classification of elderly activities," *Sensors*, vol. 16, no. 12, pp. 2043, 2016. Article (CrossRef Link)

Enjie Ding is a Professor in Perception Mine Research Center at China University of Mining and Technology. He received his Ph.D. degree in control theory and control engineering from China University of Mining and Technology and the masters and bachelor's degrees respectively from China University of Mining and Technology and University of Electronic Science and Technology of China. Dr. Ding has several years of experience working in the academia, teaching, and   consulting. He is leading the Mine Research Center. He has published more 100 papers. His research interests   include signal processing, wireless communications and applications in coal mine. Mine Wireless Sensor Network and mine things network.

Yue Zhang received the B.S. degree in Electronic and Information Engineering from Changzhou Institute of Technology, Changzhou, Jiangsu, China, in 2018. He is currently working toward the M.S. degree at China University of Mining and Technology, Xuzhou, Jiangsu, China. His research interests include device free wireless recognition, wireless networks, and machine learning.

Yun Xin is a master in Perception Mine Research Center at China University of Mining and Technology. She received her bachelor's degree from Harbin Engineering University. She major in Information and Communication Engineering. Her main research is Mine Wireless Sensor Network.

Lei Zhang is a PhD student in Perception Mine Research Center at China University of Mining and Technology. He received his Master's degree from the School of Information and Control Engineering in 2012 at China University of Mining and Technology. His research interests are in machine learning, mobile computing. His research focus is on Personnel positioning.

Yu Huo was born in Tianshui, China, in1983. She received the Ph.D. degree in communication and information system from China University of Mining and Technology, Xuzhou, China, in 2013.Currently, she is Research Assistant at the National Joint Engineering Laboratory of Internet Applied Technology of Mines, China University of Mining and Technology. Her main fields of interest are antenna systems and propagation models, wireless underground network, and emergency rescue communication.
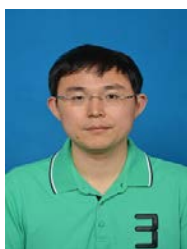
Yafeng Liu received his Ph.D degree in optical and optical engineering and B.S. degree in modern physical from University of Science and Technology of China, in 2004 and 2008, respectively. He is currently a research associate in the Perception Mine Research Center, China University of Mining and Technology. His research interests are in the areas of quantum-inspired evolutionary computation, computer vision, and image classification techniques. Dr. Liu is good at solving math problems