

RESEARCH

Open Access



Gesture recognition method based on misalignment mean absolute deviation and KL divergence

Yong Tian^{1*} , Chuanzhen Zhuang¹, Jiadong Cui¹, Runjie Qiao¹ and Xuejun Ding^{2*}

*Correspondence:
tianyong@lnnu.edu.cn;
dingxuejun@dufe.edu.cn

¹ School of Physics and Electronic
Technology, Liaoning Normal
University, Dalian, China

² School of Management
Science and Engineering,
Dongbei University of Finance
and Economics, Dalian, China

Abstract

At present, it has become very convenient to collect channel state information (CSI) from ubiquitous commercial WiFi network cards, and the location or activity of a human who affects the CSI can be recognized by analyzing the change of the CSI. Therefore, wireless sensing technology based on the CSI has received widespread attention. However, the existing CSI-based gesture recognition methods still have some problems, which include that subcarrier selection is not optimized and motion interval extraction is not accurate enough, so the accuracy of gesture recognition methods still needs to be further improved. In response to the above problems, a gesture recognition method based on misalignment mean absolute deviation (MMAD) and KL divergence is proposed in the paper, which is called MMAD-KL-GR method. This method uses the proposed MMAD algorithm to extract the CSI amplitude intervals containing gesture information, then selects subcarriers by comparing the KL divergence of the CSI amplitude, and finally uses the subspace K-nearest neighbor (KNN) algorithm to recognize the gestures. Several experiments show that the MMAD-KL-GR method can effectively improve the accuracy of the gesture recognition.

Keywords: CSI, Gesture recognition, KL divergence, Misalignment mean absolute deviation

1 Introduction

Since the 21st century, the rapid developments of big data, cloud computing and Internet of Things have promoted the developments of various intelligent technologies, such as smart homes, smart schools, smart cars, and robots. Human–computer interaction technology is the basis for realizing the above-mentioned intelligent technologies. In the human–computer interaction technology, a user should give instructions to a device through different gestures. Therefore, how the device recognizes the gestures is a key technology of the human–computer interaction.

At present, the gesture recognition technology mainly includes the following four categories: Video-based gesture recognition technology captures videos through cameras and then recognizes gestures by extracting motion features in the videos. The advantage of this technology is that it can detect tiny gestures with high recognition

accuracy, but the disadvantages are that the technology is sensitive to light in the environment, cannot recognize gestures in non-line-of-sight situations, and cannot protect user privacy [1, 2].

Gesture recognition technology based on infrared light uses the principle of infrared radiation to recognize gestures. The disadvantages are that the gestures cannot be recognized in non-line-of-sight conditions, the equipment costs are high, and the large-scale deployment is difficult [3].

Gesture recognition technology based on wearable devices requires a user to carry wearable devices which integrate a variety of sensors, such as accelerometers and gyroscopes, and these sensors can record the data about user gestures [4]. The disadvantage of this technology is that it is inconvenient to carry the device around. If a user forgets to wear the device, the gesture recognition will stop.

Commercial WiFi devices are ubiquitous and WiFi signals cover almost every corner of people's lives. In 2011, a tool that can obtain channel state information (CSI) data from Intel 5300 wireless network cards was released, which makes it very convenient to obtain the CSI data of each communication link by using the commercial WiFi devices [5]. In the communication process of WiFi devices, each communication link contains multiple subcarriers, and each subcarrier is composed of amplitude and phase information of the CSI data, so the CSI data can stably reflect the signal changes caused by reflection and diffraction of a human. Based on the above principles, many scholars use the CSI data to recognize locations and activities of a human [6, 7], including gesture recognition technology. Compared with traditional gesture recognition technologies, this technology based on CSI has these advantages, which include no special equipment deployment, nor the need to wear additional sensors, no privacy leakage, and no sensitivity to light intensity and line-of-sight (LOS) [8, 9]. In the early days of this technology, many scholars mostly used received signal strength (RSS) to recognize gestures. RSS is a coarse-grained measurement value that is greatly affected by the environment. For example, when someone moves in a monitoring area, the RSS may increase, decrease, or even remain unchanged. In recent years, many scholars have devoted themselves to exploring CSI of wireless signals. Compared with RSS, CSI is a fine-grained measurement value, which is less affected by the environment and has better stability, so the CSI-based recognition technology has received extensive attention from academic community [6–12].

Although the existing methods have good performance in their respective experimental environments, there are still some problems, which include that subcarrier selection is not optimized and motion interval extraction is not accurate, so the accuracy of gesture recognition methods needs to be further improved. In response to the above problems, we propose a gesture recognition method based on misalignment mean absolute deviation (MMAD) and KL divergence, which is called MMAD-KL-GR method. The main contributions are as follows:

1. Before recognizing gestures, it is necessary to extract CSI amplitude interval affected by the gestures in order to improve the accuracy of gesture recognition. For this reason, we propose a MMAD motion interval extraction algorithm. Based on the MAD algorithm, this algorithm considers the various situations of the starting point and

the ending point of the motion interval and improves the accuracy of the motion interval extraction.

2. Different subcarriers are affected by gestures to different degrees, so it is necessary to select subcarriers that are more affected by gestures and less interfered by noise in order to obtain higher gesture recognition accuracy. For this reason, we propose a subcarrier selection algorithm based on KL divergence. This algorithm compares the KL divergences of the CSI amplitudes to select subcarriers and obtains better results.
3. To further improve the accuracy of gesture recognition based on CSI, we propose a MMAD-KL-GR method based on subspace K-nearest neighbor (KNN) for classification. Through several experiments, we verified the good performance of the proposed MMAD-KL-GR method and analyzed the influence of training samples, transceiver spacing, human body position, indoor environment and features on the proposed method.

The rest of the paper is organized as follows. Section 2 introduces the related works of the paper. Section 3 describes in detail the MMAD-KL-GR method proposed in this paper. Section 4 gives a detailed analysis of the experimental results. Section 5 summarizes this paper.

2 Related works

In recent years, a large amount of CSI-based wireless sensing methods have emerged, such as activity recognition and gesture recognition methods.

Dang et al. used principal component analysis (PCA) algorithm to build a fingerprint database of CSI amplitude data and used Kalman filter algorithm to obtain data for classification and then used support vector machine (SVM) algorithm and fingerprint database matching for activity recognition [6]. Cheng et al. proposed a CSI-based human continuous activity recognition system. This system uses the CSI phase difference matrix and a method based on threshold and label to segment the continuous activities and then uses Gaussian mixture model–hidden Markov model (GMM-HMM) to recognize activities [10]. The above methods use the time-domain features of CSI to recognize activities, and some methods also use the frequency-domain features. Wavelet transform can locally analyze time and frequency and is an ideal tool for time–frequency analysis and processing of signals [13–19]. Therefore, Wang et al. used a wavelet transform to extract the features of CSI and designed a two-stage recognition method to jointly recognize the locations and activities of multiple targets [11]. Tian et al. constructed a time–frequency matrix by using signal preprocessing and wavelet transform and then extracted multi-dimensional features in time domain and frequency domain as the input feature vector of bidirectional long short-term memory (BLSTM) network [12].

The above methods based on CSI can recognize the activities with large amplitude. However, the gestures with small amplitude can also be recognized by using CSI.

Tian et al. proposed a CSI-based device-free gesture recognition system, namely the WiCatch system. Firstly, a new interference cancellation algorithm based on data fusion is proposed to capture weak reflected signals. Secondly, the motion trajectories of gestures are reconstructed by constructing a virtual antenna array of time-domain signal samples. Finally, SVM algorithm is used to complete the classification [20]. Zhang

et al. proposed a gesture recognition system called the WiGrus system. The system first uses PCA method and the first-order difference method to denoise CSI data, and then extracts multiple features that can characterize the gestures, and finally proposes a two-stage radio frequency algorithm to classify the gestures [21]. Thariq et al. proposed a sign language recognition system, namely the DF-WiSLR system. The system can be used to recognize 30 static gestures and 19 dynamic gestures and can obtain better recognition accuracy for dynamic gestures composed of compound word symbols [22]. Jiang et al. proposed a WiGAN gesture recognition system, which not only solves the problem of performance degradation caused by small samples and strong environmental dependence, but also incorporates more diverse features to improve the accuracy of gesture recognition [23]. Hao et al. proposed a fine-grained sign language recognition method, which first filters out environmental interference in the frequency domain through a Butterworth filter, and then uses wavelet transform to smooth the CSI data, and finally builds a low-complexity KSB classification model [24]. Tan et al. proposed a finger gesture recognition algorithm, which effectively removes environmental noise and develops a measure that can recognize gestures by dealing with individual diversity and gesture inconsistency. Experimental results show that the algorithm has good recognition accuracy and robustness in a changing environment [25]. Zhang et al. proposed a WiFi-based cross-domain gesture recognition system WiGr. The system proposes a dual-path network composed of a depth feature extractor and a dual-path recognizer, which can extract domain-independent gesture features, so that good gesture recognition accuracy can be obtained without retraining in a new domain [26]. Gu et al. proposed a gesture recognition system WiGRUNT based on dual-attention network for cross domain recognition. The system dynamically extracts the domain-independent features of CSI by using a spatial-temporal dual-attention mechanism and then recognizes the fine-grained gestures by using a depth residual network [27].

3 MMAD-KL-GR method

The framework of the proposed MMAD-KL-GR method is shown in Fig. 1. Firstly, the MMAD-KL-GR method needs to deploy a transmitter and a receiver with WiFi devices. Volunteers complete the required motions between the transceivers. The receiver collects CSI data affected by the gestures and stores the collected data in the computer for training and testing. Then, the MMAD-KL-GR method needs to perform data preprocessing on the collected CSI data. The processing process includes using a Hampel identifier to remove the abnormal values in the CSI data and using a Gaussian filter to remove the high-frequency noise in the CSI data for retaining the low-frequency features caused by the gestures. Next, the MMAD algorithm is used to detect the time starting point and end point of a gesture in order to extract the motion interval of the CSI amplitude, and the motion interval data is interpolated into a sequence of 50 data points by cubic spline, and then is normalized. The proposed subcarrier selection algorithm based on KL divergence is used to select the better subcarriers conducive to gesture recognition. Finally, the mean, median, upper quartile, lower quartile, variance, root-mean-square and skewness coefficient of the normalized data are calculated. These features are constructed as a feature matrix together with the normalized data for training and testing the subspace KNN algorithm.

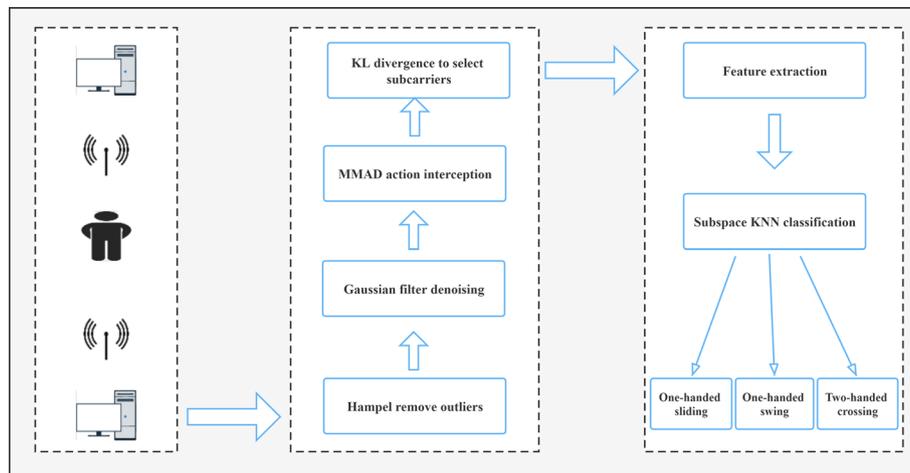


Fig. 1 MMAD-KL-GR method framework

3.1 Data preprocessing

Due to the interferences including WiFi device itself, complex indoor environment and various electromagnetic signals in space, there are outliers and high-frequency noise in the original CSI data [28]. These outliers and high-frequency noise can reduce the accuracy of gesture recognition. In order to eliminate the influence of the outliers, we use the Hampel identifier algorithm [29] to remove the outliers in the paper, and the specific method is as follows:

A sliding window with a length of $2h + 1$ is defined on the CSI sequence, and the amplitude of the window midpoint is x_i . The median m_i and the median absolute deviation MAD_i of the window are calculated as follows:

$$m_i = \text{median}(x_{i-h}, x_{i-h+1}, \dots, x_i, \dots, x_{i+h-1}, x_{i+h}), \tag{1}$$

$$MAD_i = \text{median}(|x_{i-h} - m_i|, \dots, |x_{i+h} - m_i|), \tag{2}$$

where $\text{median}()$ represents the function of the median. If x_i satisfies $|x_i - m_i| > nMAD_i$, the Hampel identifier algorithm determines that x_i is an outlier and replaces x_i with the median m_i , where n is a positive integer. Through some experiments, we have verified that the Hampel identifier algorithm has a good effect on removing outliers when $h = 5$ and $n = 3$.

Figure 2 shows the CSI amplitude curve before and after removing outliers by using the Hampel identifier algorithm, where the blue dashed line and the red solid line represent the CSI amplitude curve before and after removing outliers and the black boxes represent the outliers. From Fig. 2, it can be seen that the outliers in the CSI amplitudes have been effectively removed.

After using the Hampel identifier algorithm to remove the outliers, there is still a lot of high-frequency noise in the CSI amplitudes. To eliminate the influence of the high-frequency noise and retain the low-frequency fluctuations caused by the gestures, we use one-dimensional Gaussian filter to eliminate the high-frequency noise in this paper. The specific process is as follows:

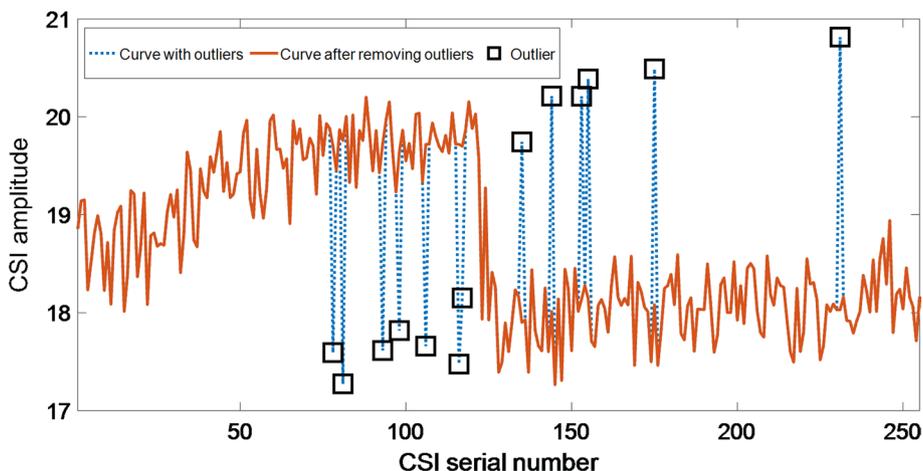


Fig. 2 Comparison before and after removing outliers

A sliding window with a length of $2k + 1$ is defined on the CSI sequence, the amplitude of the window midpoint is x_i , and the weighted normal distribution function $Q(x_j)$ is calculated as follows:

$$Q(x_j) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(j-i)^2}{2\sigma^2}}, j \in (i - k, \dots, i, \dots, i + k), \tag{3}$$

Then, the Gaussian filter function $G(x_i)$ is:

$$G(x_i) = \frac{1}{2k + 1} \sum_{j=i-k}^{i+k} x_j Q(x_j). \tag{4}$$

The characteristic of Gaussian filter is that x_i is the center, and the weights are symmetrically distributed. The closer the amplitude to x_i , the greater the influence on x_i , so the weight is also greater. Conversely, the smaller the influence, so the weight is also smaller. The parameters affecting the weighted normal distribution function include k and σ^2 , where the larger the k , the larger the range of the CSI amplitudes that affects x_i , and the larger the variance σ^2 of the normal distribution, the more concentrated the weight is in the center. Through some experiments, we verify that the Gaussian filter achieves good performance when $k = 30$ and $\sigma^2 = 20$.

Figure 3 is the comparison of the effects before and after using Gaussian filter, where the blue curve is the CSI amplitudes before filtering, and the red curve is the CSI amplitudes after filtering. As shown in Fig. 3, the Gaussian filter removes the high-frequency noise of the CSI amplitudes and turns them into a smooth curve. The low-frequency variation of the CSI amplitudes affected by the gestures is mainly concentrated in the L_1 interval, which is well-preserved to ensure the accuracy of the gesture recognition.

3.2 Motion interval extraction

In general, the CSI samples not only include the motion interval affected by the gestures, but also the no motion interval that is not affected by the gestures. The no motion interval is invalid information for the gesture recognition. If the filtered CSI

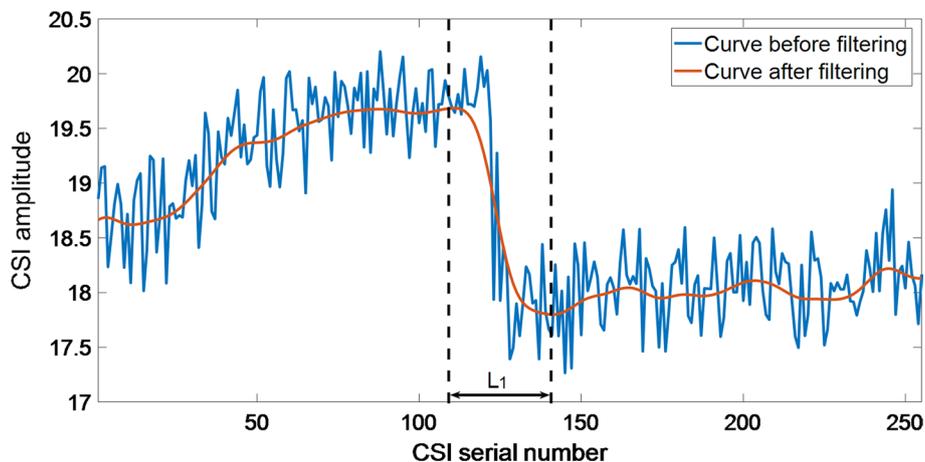


Fig. 3 Comparison of the effects before and after using Gaussian filter

amplitudes are directly used as the input of a machine learning algorithm, the accuracy of the gesture recognition will decrease. Therefore, we need to accurately extract the CSI motion interval. The MAD threshold method is a commonly used motion interval extraction method [30, 31]. This method needs to calculate the MAD value of the data sequence and compare it with the threshold to determine the starting and ending points of the motion interval. However, there are some problems in this method, such as inaccurate judgment of the starting point and possible misjudgment of the ending point. To solve these problems, we improve the MAD threshold method and propose the MMAD algorithm as follows.

A sliding window with a length of $2c + 1$ is defined on the CSI sequence, and the amplitude of the window midpoint is x_d , and the MAD and MMAD value of each point in the sliding window is calculated as follows:

$$\bar{x}_{MAD}(d) = \frac{1}{2c + 1} \sum_{i=d-c}^{d+c} x_i, \tag{5}$$

$$\bar{x}_{MMAD}(d) = \frac{1}{2c + 1} \sum_{i=d-2c}^d x_i, \tag{6}$$

$$MAD(d) = \frac{1}{2c + 1} \sum_{i=d-c}^{d+c} |x_i - \bar{x}_{MAD}(d)|, \tag{7}$$

$$MMAD(d) = \frac{1}{2c + 1} \sum_{i=d}^{d+2c} |x_i - \bar{x}_{MMAD}(d)|, \tag{8}$$

where $\bar{x}_{MAD}(d)$ is the mean value of the CSI amplitudes of $2c + 1$ points centered at point d , and $\bar{x}_{MMAD}(d)$ is the mean value of the CSI amplitudes of $2c + 1$ points that are at the left of point d (including point d), and $MAD(d)$ and $MMAD(d)$ are the MAD value

and MMAD value of the point d , respectively, and d is the integer changing from $2c + 1$ to $D - 2c$, and D is the total number of data points.

To illustrate the effectiveness of the MMAD algorithm, we have selected two typical CSI samples, and calculated the MAD values and the MMAD values of the two sample sequences, and then drawn the MAD and the MMAD curves, respectively, as shown in Fig. 4. In Fig. 4a, when the MMAD value is greater than the threshold T for the first time, the corresponding CSI data point S_1 is the starting point of the motion interval. When the MMAD value is less than the threshold T for the first time after S_1 , the corresponding CSI data point S_3 is the ending point of the motion interval. The interval S_1S_3 is the extracted motion interval. However, as shown in Fig. 4a, the starting point of the motion interval obtained by the MAD algorithm is S_2 , but the interval S_1S_2 contains the part information of the gesture. Figure 4b shows that the MAD algorithm incorrectly divides the interval S_4S_5 into the no motion interval. Therefore, the MMAD algorithm is better than the MAD algorithm in the judgment of the starting point and the ending point of the motion interval. Through some experiments, we verify that the MMAD algorithm has good performance when $c = 10$.

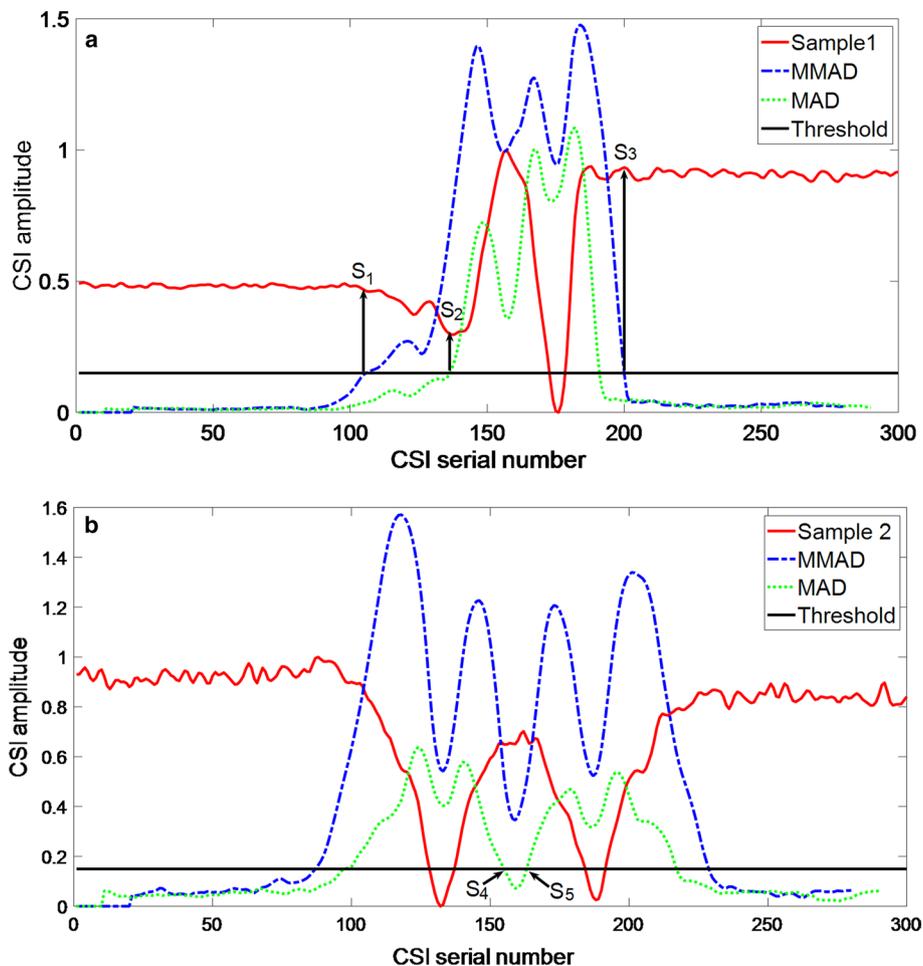


Fig. 4 Comparison of MMAD and MAD algorithms

Since the motion interval data length of each sample is different, and the subsequent classification algorithm requires that the data length of each sample must be the same, we use cubic spline interpolation method to interpolate the extracted motion interval data, so as to obtain a unified motion interval data length.

3.3 Subcarrier selection

In order to ensure the stability of data transmission, a commercial WiFi network card uses one or more antennas and each communication link contains multiple subcarriers when sending and receiving signals. Therefore, CSI data of each sample collected at the receiver contain multiple subcarriers. Because communication link, transmission frequency and multipath effect may be different and the CSI amplitudes of each subcarrier are also different [32], it is important to select a better subcarrier from a sample to improve the accuracy of gesture recognition. For this reason, we propose an algorithm for selecting subcarriers based on KL divergence.

KL divergence is an asymmetry measure of the difference between two probability distributions [33]. In the field of communication, the KL divergence can be calculated as the difference between the information entropy of two sets of data, where the information entropy is related to the appearance probability of data and is a measure of the time series complexity. Let $U(y)$ and $V(y)$ be the two probability distributions of the random variable y . When y is a discrete random variable, the KL divergence can be defined as follows:

$$KL(U \parallel V) = \sum U(y) \log \frac{U(y)}{V(y)}. \tag{9}$$

The properties of the KL divergence are: (i) KL divergence is always non-negative, that is, $KL(U \parallel V) \geq 0$. (ii) KL divergence is an asymmetry measure of two probability distributions, namely $KL(U \parallel V) \neq KL(V \parallel U)$.

Using the properties of the KL divergence, we calculate $KL(U \parallel V)$ by using the motion interval sequence of subcarrier as the probability distribution $V(y)$ and no motion interval sequence as the probability distribution $U(y)$, as shown in Fig. 5. The

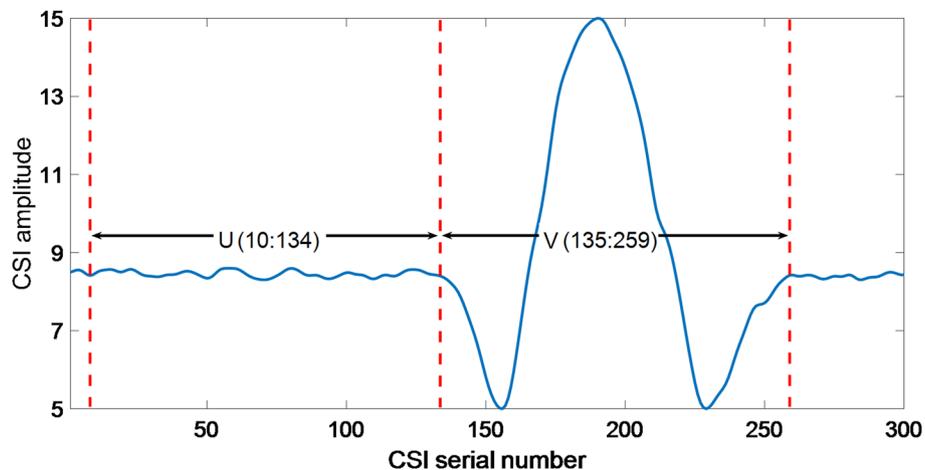


Fig. 5 Schematic of KL divergence extraction

larger the $KL(U \parallel V)$, the greater the difference between the motion interval and no motion interval of the subcarrier, and the greater the change of the CSI amplitudes caused by the gestures. Therefore, we can select the CSI amplitudes of the subcarrier whose $KL(U \parallel V)$ is the largest to recognize the gestures. To ensure good stability of the selected subcarriers, we calculate the subcarriers of all samples as follows:

$$Sum_a = \sum_{b=1}^B KL(U_{ab} \parallel V_{ab}), \quad (10)$$

where $a = 1, \dots, A$, and A is the number of subcarriers in a CSI sample, and B is the total number of CSI samples. We select the data of the subcarrier corresponding to the largest Sum_a for gesture recognition.

3.4 Feature extraction

Extracting features that are highly relevant to the gestures from the motion intervals is an important part of improving the accuracy of gesture recognition. In this paper, we use the mean, median, upper quartile, lower quartile, variance, root-mean-square, skewness factor and the CSI amplitudes to construct the feature vectors of samples which can represent the statistical characteristics and change trend of CSI amplitudes [34].

3.5 Subspace KNN algorithm

The subspace KNN algorithm is an improved KNN algorithm. The basic principle of the KNN algorithm is shown in Fig. 6. The algorithm assumes that all existing samples have a definite classification. When it is necessary to determine the category of a new sample, the KNN algorithm calculates the distance between each sample in the existing sample set and the new sample (in this paper, Euclidean distance is used) and finds the K samples with the smallest distance. In the above K samples, the number of samples belonging to which classification is the largest, and the new sample belongs to the classification [35].

Assuming that the feature matrix is R rows and C columns, the steps of the subspace KNN algorithm are as follows:

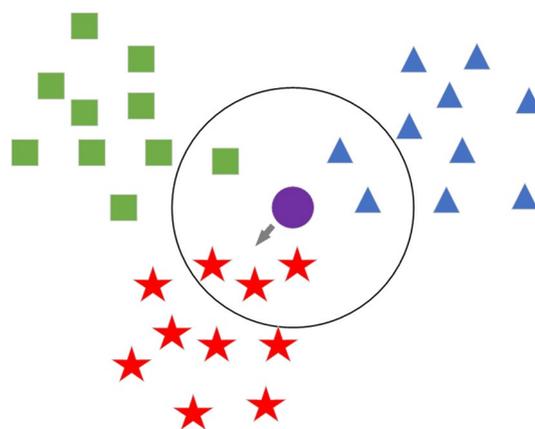


Fig. 6 KNN principle diagram

1. From the C columns of the feature matrix, M columns are randomly selected to construct a sub-feature matrix, and the step is repeated N times to obtain N sub-feature matrices.
2. The N sub-feature matrices are used to train the KNN algorithm, and N sub-classification models are obtained.
3. The N sub-classification models are used to classify a new sample. Then, we use the majority principle to determine the category of new samples.

The subspace KNN algorithm samples the feature matrix to form multiple sub-feature matrices and then trains the KNN algorithm multiple times, thereby improving the classification accuracy and stability of the KNN algorithm.

4 Experiment evaluation

4.1 Experimental setup and data collection

This paper conducted experiments in a laboratory with an area of $11.1 \text{ m} \times 9.6 \text{ m}$. The layout of the laboratory is shown in Fig. 7. In the experiment, we use two desktop computers with Intel 5300 wireless network card as the transmitter and the receiver, and both computers are equipped with Ubuntu 12.04 operating system. The transmitter sends signals through one antenna and the receiver receives signals through three antennas. The working frequency of the wireless network card is 2.4 GHz, and the channel bandwidth is 20 MHz. There are 30 subcarriers in each communication link.

During the experiment, the volunteers always sit on the chair at the designated position. When starting to collect the CSI data, the volunteers first remained still, then completed the prescribed motions, and then remained still again. The process lasted 4 seconds in total. In each experiment, the volunteers carried out three motions of

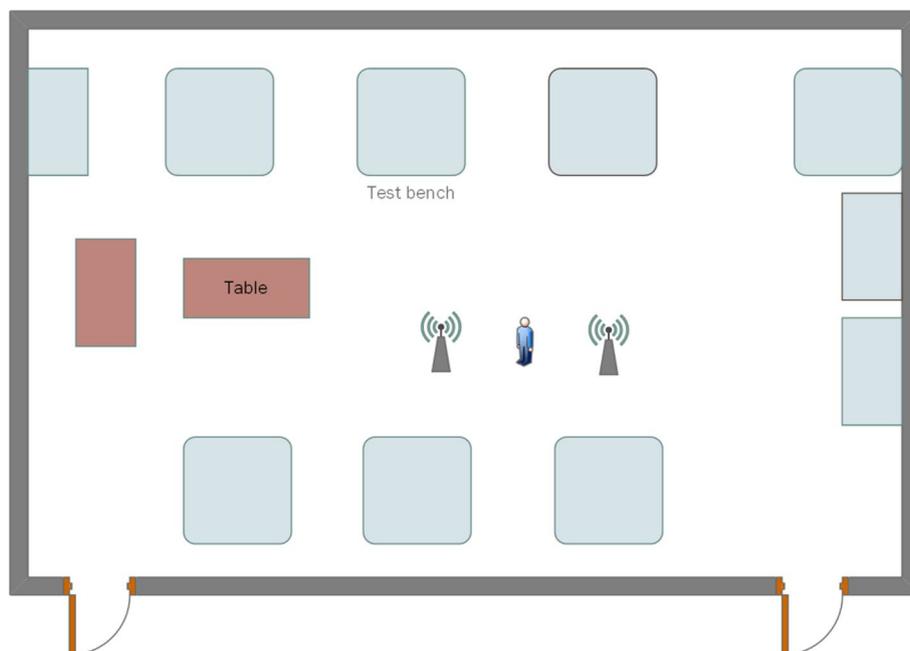
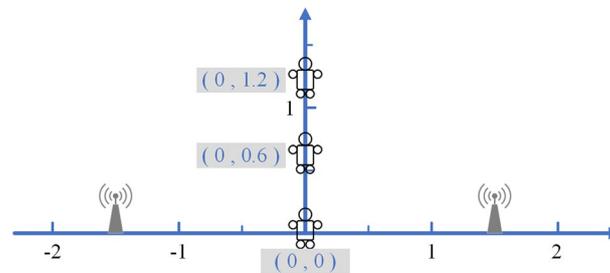


Fig. 7 Layout of the laboratory

Table 1 Information of experimental samples

Group	Transceiver distance (m)	Location	Interference
Group1	1.5	(0, 0)	N/A
Group2	2.0	(0, 0)	N/A
Group3	2.5	(0, 0)	N/A
Group4	3.0	(0, 0)	N/A
Group5	2.0	(0, 0)	Bluetooth
Group6	2.0	(0, 0)	Computer case
Group7	2.0	(0, 0)	Blackboard
Group8	2.0	(0, 0.6)	N/A
Group9	2.0	(0, 1.2)	N/A

**Fig. 8** Position of the transceiver and the human body

two-handed crossing, one-handed sliding and one-handed swing, and 130 samples were collected for each motion. To analyze and verify the MMAD-KL-GR method proposed in this paper, we conducted nine sets of experiments with different transceiver distances, volunteer positions, and interference factor combinations. The information of experimental samples is shown in Table 1. The positions of the transceiver and the human body are shown in Fig. 8. From each set of motion samples, we randomly select 60 samples as the training set and the remaining 70 samples as the testing set. Therefore, in each experiment, the training set contains 180 samples and the testing set contains 210 samples. We use four machine learning algorithms: the bagging tree, the subspace KNN, the linear SVM and the medium decision tree to evaluate the performance of the MMAD-KL-GR method.

4.2 Motion interval extraction and subcarrier selection algorithm evaluation

4.2.1 Analysis of MMAD motion interval extraction algorithm

To verify the effectiveness of the MMAD algorithm, we randomly selected 165 samples from the training set of the second group of experimental data to train the four machine learning algorithms, and tested them with 210 samples of the test set. The MMAD and MAD algorithms were used to extract the motion interval respectively. The experimental results are shown in Fig. 9. Figure 9 shows that the accuracy of the MMAD algorithm is better than that of the MAD algorithm. This is because the MMAD algorithm is more accurate in the judgment of the starting point and the MMAD algorithm can better avoid the truncation of the motion interval in the determination of the ending point, so the gesture recognition accuracy of the MMAD algorithm is higher. Figure 9 shows that

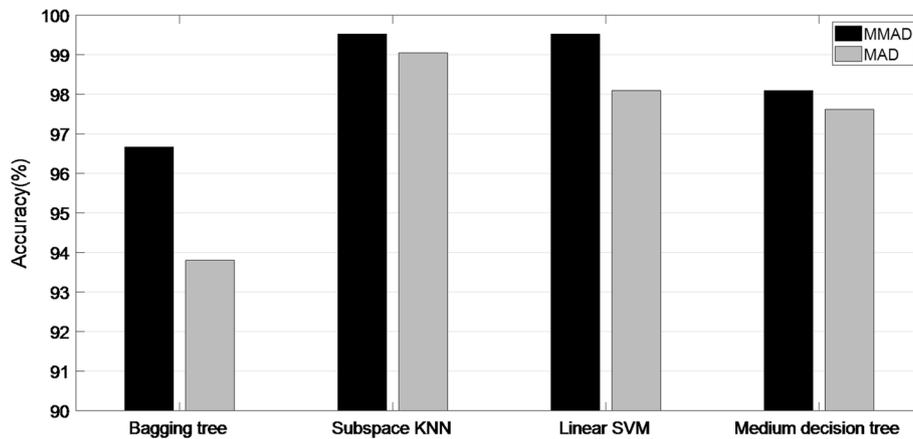


Fig. 9 Performance comparison of MMAD and MAD algorithm

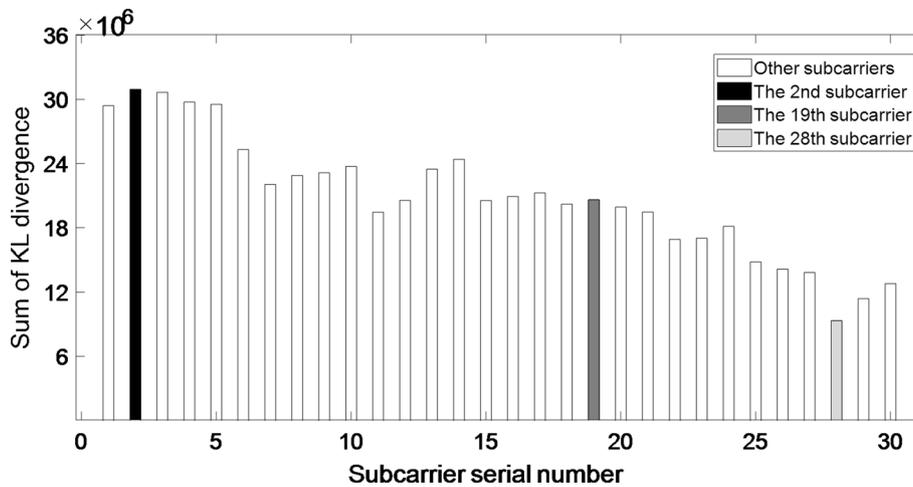


Fig. 10 Sum of KL divergence of 30 subcarriers

the subspace KNN algorithm can all achieve higher gesture recognition accuracy when using the MMAD algorithm and the MAD algorithm.

4.2.2 Analysis of KL divergence selection subcarrier algorithm

To verify the effectiveness of KL divergence selection subcarrier algorithm, we calculate the KL divergence of 30x3510 subcarriers (390x9=3510 samples in Table 1) and then calculate the sum of KL divergence of the subcarrier of 3510 samples, where $a = 1, 2, \dots, 30$. The experimental results are shown in Fig. 10. Figure 10 shows that the sum of the KL divergence of the second subcarrier is the largest. Therefore, we select the data of the second subcarrier for the gesture recognition.

To verify whether the data of the second subcarrier is better than other subcarriers for the gesture recognition, we use the same data as Sect. 4.2.1 for experiments. In the experiment, we only use the data of the 2th, 19th, and 28th subcarriers ranked 1, 15 and 30 in the sum of the above KL divergence to perform the gesture recognition. The experimental results are shown in Fig. 11. Figure 11 shows that the accuracy of the gesture

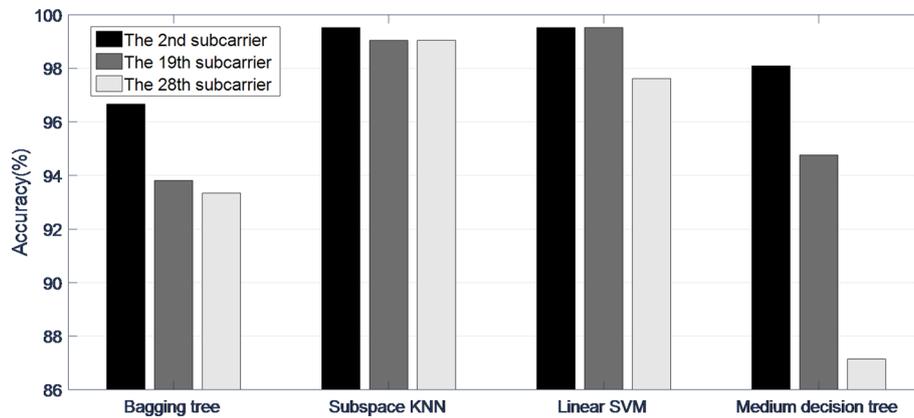


Fig. 11 Accuracy of the gesture recognition using the data of the 2nd, 19th and 28th subcarriers

recognition using the data of the second subcarrier is all the highest, 96.67%, 99.52%, 99.52%, and 98.10%, respectively. The accuracy of gesture recognition using the data of the 28th subcarrier is the lowest. This is because the sum of KL divergence of the second subcarrier is the largest, indicating that the CSI amplitudes caused by the gestures change greatly. Therefore, the extracted features can more accurately characterize the corresponding gestures, thus improving the accuracy of the gesture recognition. The experiment verifies the reliability and effectiveness of the proposed KL divergence selection subcarrier algorithm.

4.3 Evaluation of MMAD-KL-GR method

4.3.1 Impact of training samples

When training a machine learning algorithm, the number of training samples is an important factor affecting the accuracy of the machine learning algorithm. To evaluate the impact of training samples, we use the second group of experimental data, randomly select 15 to 180 samples with the step size 15 to train the four machine learning algorithms, and test these algorithms by using the remaining 210 samples. The experimental results are shown in Fig. 12. From 15 to 180 training samples, the gesture recognition accuracy of the subspace KNN algorithm is all the highest, and the stability is also the best. The advantage is especially obvious when the number of training samples is small. When the number of training samples reaches 165, the accuracy stabilizes at 99.52%. The reason is that the subspace KNN does not directly use the feature matrix for training, but instead samples the feature matrix to form multiple sub-training sets before training. Although the number of features of training samples is reduced for each sub-training set, the number of features is increased for the classification algorithm as a whole. Therefore, the subspace KNN algorithm effectively improves the accuracy of the gesture recognition and has better stability. According to the above experimental results, the gesture recognition accuracy increases with the increase in the number of training samples. When the number of training samples reaches a certain value, the gesture recognition accuracy remains stable. Therefore, the number of training samples does not need to be too large,

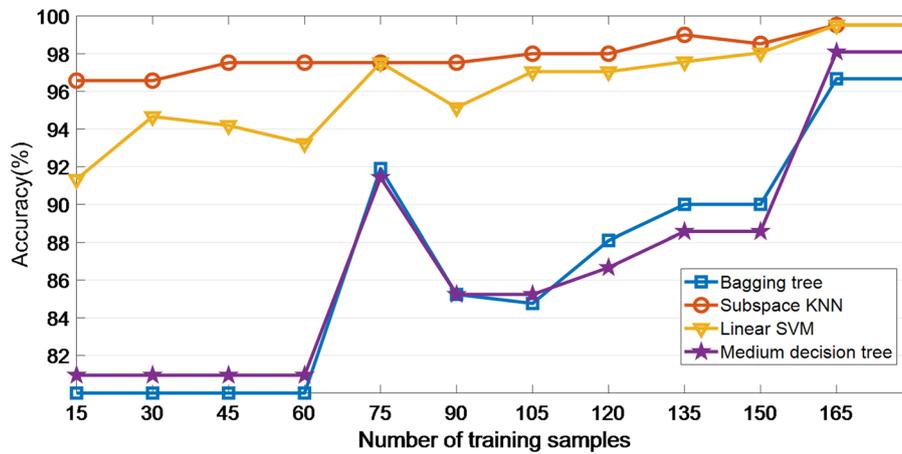


Fig. 12 Impact of the number of training samples

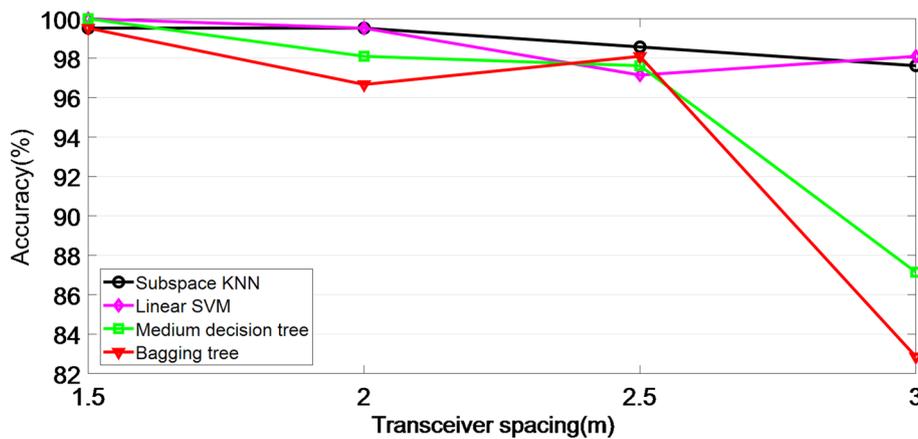


Fig. 13 Impact of transceiver spacing

because too many training samples will greatly increase the workload of sample collection and make the training time of classification algorithm too long. In the paper, 165 samples are randomly selected from 180 training samples for training.

4.3.2 Impact of transceiver spacing

To verify the performance of the MMAD-KL-GR method in the case of different transceiver spacing, we use the first, second, third, and fourth set of data in Table 1 to conduct experiments, and the experimental results are shown in Fig. 13. Figure 13 shows that when the distance between the transceivers is 1.5 meters, the gesture recognition accuracies of the four machine learning algorithms are all greater than 99%. As the distance between the transceivers increases, the accuracy of the gesture recognition begins to gradually decrease. However, the accuracy of the subspace KNN algorithm has also declined, but it is much higher than other algorithms. This shows that the subspace KNN algorithm still has a high gesture recognition accuracy and good robustness for a large transceiver spacing.

4.3.3 Impact of human body position

To verify the impact of human body position on the MMAD-KL-GR method, we set the human body at the center of transceiver connection, 0.6 meters away from the center vertically, and 1.2 meters away from the center vertically, as shown in Fig. 8. The experiment is carried out by using the data of groups 2, 8 and 9 in Table 1. The experimental results are shown in Fig. 14. Figure 14 shows that the gesture recognition accuracy of the four machine learning algorithms is high when the human body is at the center of the transceiver connection, and the accuracy of the other two experiments is low. The experimental results show that the accuracy of the gesture recognition decreases rapidly when the human body gradually moves away from the center of the transceiver connection. Therefore, to improve the accuracy of the gesture recognition, it is better for the human body to be at the center of the LOS path of the transceiver.

4.3.4 Impact of indoor environment

The application scenario of the MMAD-KL-GR method is indoors, and there are usually many environmental changes in the indoor environment. Therefore, we use the second, fifth, sixth, and seventh groups of data in Table 1 for experiments, and the experimental results are shown in Fig. 15. On LOS path of the transceiver, the second, fifth, sixth and seventh groups of data are respectively collected in the following four cases: no obstacle and no interference (referred to as no interference), no obstacle but Bluetooth headset interference (referred to as Bluetooth), computer case but no interference (referred to as computer case) and blackboard (2 m×1.2 m) but no interference (referred to as blackboard). As shown in Fig. 15, the gesture recognition accuracy of the subspace KNN algorithm is the highest in the three cases of no interference, Bluetooth and blackboard, and the recognition accuracy in the case of computer case is slightly lower than that of bagging tree and linear SVM algorithm. The interference of Bluetooth headset has little impact on the accuracy of the gesture recognition. This is because the transmission distance of Bluetooth is short and the power is small, so the interference to the CSI is also small. When the computer case blocks the LOS path of the transceiver, according to the multipath effect theory, the WiFi signal can also be transmitted to the receiver through

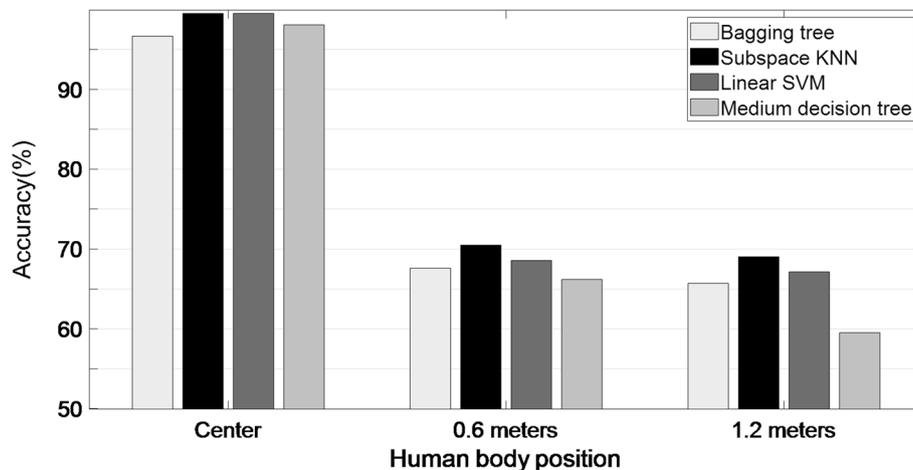


Fig. 14 Impact of human body position

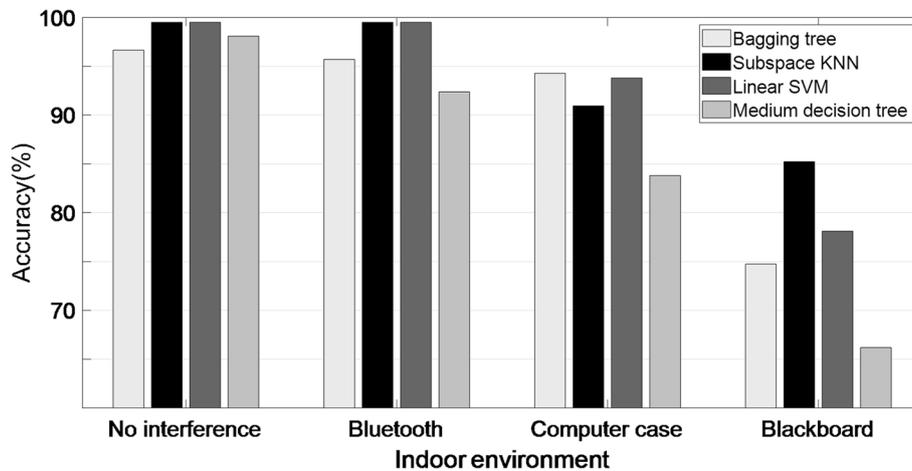


Fig. 15 Impact of the indoor environment

other reflection paths in the surrounding environment. However, when the blackboard blocks the LOS path of the transceiver, the power reduction in the received signals is very large, and the accuracy of the gesture recognition is very low because the blackboard blocks too many transmission paths of signals. Therefore, in the application environment of the MMAD-KL-GR method, it is better not to place any obstacles on the LOS path of the transceiver.

4.3.5 Impact of features

In the MMAD-KL-GR method, the features are the important factors that determine the recognition accuracy of the method. In the paper, we use mean, median, variance, root-mean-square, upper quartile, lower quartile, skewness factor and CSI amplitude as the features of the gesture recognition. However, these features are common statistical features in time domain. At present, other features such as energy, zero-crossing rate and entropy are used in some state-of-the-art methods [12, 36–39]. To verify the performance of different combined features and classifiers, we use the seven features given in Sect. 3.4, together with the sample entropy, time-domain energy and frequency-domain energy proposed in our previous work [12], to carry out some experiments. In the experiments, we use the different combinations of the above features and the four classifiers used in this paper. The experimental results are shown in Fig. 16, where the meanings of the feature combinations are shown in Table 2. From Fig. 16, it can be show that each classifier can obtain similar recognition accuracy when using Comb1 features and Comb2 features, while the recognition accuracies of all classifiers are improved when using Comb4 features, and the recognition accuracy of the subspace KNN is always the highest. When Comb3 features are used, the recognition accuracies of the four classifiers are not significantly different, while they are also improved when Comb5 features are used. However, compared with that when using Comb4 features, the recognition accuracy of the subspace KNN is reduced when using Comb5 features, while the recognition accuracies of the other three classifiers are improved, and the recognition accuracy of the bagging tree is higher than that of the subspace KNN. This shows that we need to

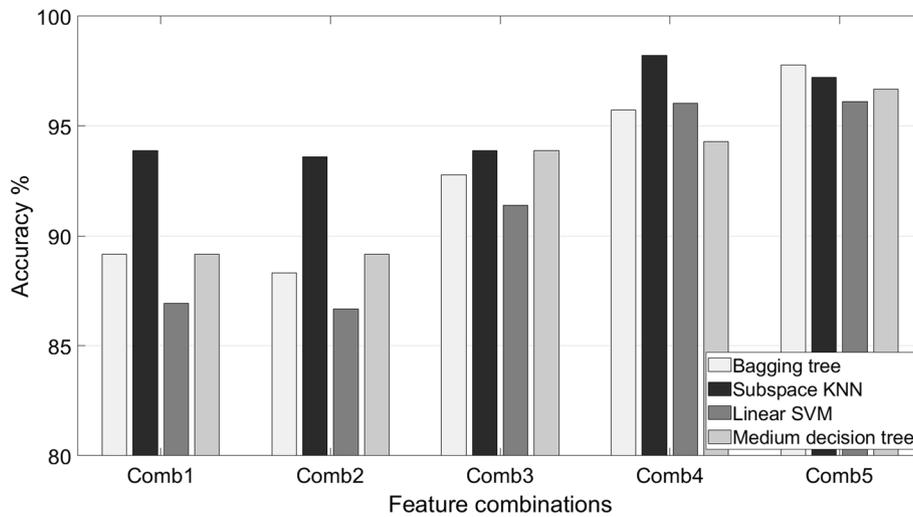


Fig. 16 Impact of the features

Table 2 Meanings of the feature combinations

Name	Feature combination
Comb1	Mean, Median, Variance and Root mean square
Comb2	Upper quartile, Lower quartile and Skewness factor
Comb3	Sample entropy, Time-domain energy and Frequency-domain energy
Comb4	Comb1 and Comb2
Comb5	Comb1, Comb2 and Comb3

select different classifiers according to different feature combinations when designing a gesture recognition algorithm.

4.3.6 Discussion and limitation

To verify the generalization of the MMAD-KL-GR method, we analyzed the impact of training samples, transceiver spacing, human body position and indoor environment on the performance of the method. From the experimental results, it can be seen that the recognition accuracy of the MMAD-KL-GR method can meet the needs of most applications and is very high even when the number of training samples is small. As the transceiver spacing increases, the recognition accuracy of the MMAD-KL-GR method gradually decreases, so the transceiver spacing should not be too large when using the method. The human body position has a great impact on the recognition accuracy of the MMAD-KL-GR method. When the human body gradually moves away from the LOS path, the recognition accuracy of the method decreases rapidly. Therefore, the human body should be on the LOS path when using the method, otherwise the high recognition accuracy cannot be guaranteed. Indoor environment also has a great impact on the recognition accuracy of the MMAD-KL-GR method. If there are large obstacles or walls between transceivers, the recognition accuracy is low, but Bluetooth interference

or small obstacles have little impact on the accuracy. In summary, the MMAD-KL-GR method can obtain high recognition accuracy by using few training samples when the transceiver spacing is small, there are no large obstacles on the LOS path, and the human body is located on the LOS path. If the above conditions cannot be met, the recognition accuracy of the method will reduce. It is necessary to comprehensively judge the application possibility of the method according to the actual situation and the degree of accuracy reduction.

5 Conclusion

For the problems of subcarrier selection and motion interval extraction in the existing gesture recognition methods based on CSI, we propose a gesture recognition method based on the MMAD and KL divergence, which is called the MMAD-KL-GR method. This method uses the MMAD algorithm to extract the motion interval of the CSI data, uses the properties of KL divergence to select subcarriers, and uses some extracted features to recognize the gestures through the subspace KNN algorithm. Through experimental comparison, we analyze the proposed MMAD algorithm and KL divergence subcarrier selection algorithm. The experimental results show that the proposed interval extraction and subcarrier selection algorithms can effectively improve the accuracy of the gesture recognition. To comprehensively evaluate the MMAD-KL-GR method, we also analyzed the impact of the number of training samples, the transceiver spacing, the human body position, the indoor environment and the features on the proposed method through a large number of experiments and gave better application parameters of the method. In future work, we will test more gestures to further expand the application range of the MMAD-KL-GR method, and study the impact of different gestures and application scenarios on the selection of features and classifiers.

Abbreviations

CSI	Channel state information
MMAD	Misalignment mean absolute deviation
GR	Gesture recognition
KNN	K-nearest neighbor
RSS	Received signal strength
PCA	Principal component analysis
MAD	Mean absolute deviation
SVM	Support vector machine
LOS	Line of sight

Acknowledgements

The authors would like to thank the anonymous reviewers for their valuable comments and suggestions that helped to improve the quality of this manuscript.

Author contributions

All authors have contributed equally. All authors have read and approved the final manuscript.

Funding

This work is supported by the Natural Science Foundation of China under Grant 62076114 and Grant 71874025, and the Humanities and Social Sciences Research Planning Foundation of the Ministry of Education of China under Grant 20YJA630058.

Data availability

Not available online. Please contact the author for data requests.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 29 November 2021 Accepted: 19 September 2022

Published online: 30 September 2022

References

1. S.S. Rautaray, A. Agrawal, Vision based hand gesture recognition for human computer interaction: a survey. *Artif. Intell. Rev.* **43**, 1–54 (2015)
2. S. Herath, M. Harandi, F. Porikli, Going deeper into action recognition: a survey. *Image Vis. Comput.* **60**, 4–21 (2017)
3. J. Wang, T. Liu, X. Wang, Infrared hand gesture recognition with convolutional neural networks in double-teachers instruction mode classroom. *Infrared Phys. Technol.* **111**, 103464 (2020)
4. C. Shen, Y. Chen, G. Yang et al., Toward hand-dominated activity recognition systems with wristband-interaction behavior analysis. *IEEE Trans. Syst. Man Cybern. Syst.* **50**, 2501–2511 (2020)
5. D. Halperin, W. Hu, A. Sheth et al., Tool release: Gathering 802.11n traces with channel state information. *ACM SIGCOMM Comput. Commun. Rev.* **41**, 53 (2011)
6. X. Dang, Y. Huang, Z. Hao et al., PCA-Kalman: device-free indoor human behavior detection with commodity Wi-Fi. *EURASIP J. Wirel. Commun. Netw.* **2018**, 214 (2018)
7. L. Zhang, E. Ding, Y. Hu et al., A novel CSI-based fingerprinting for localization with a single AP. *EURASIP J. Wirel. Commun. Netw.* **2019**, 51 (2019)
8. J. Wang, L. Zhang, C. Wang et al., Device-free human gesture recognition with generative adversarial networks. *IEEE Internet Things J.* **7**, 7678–7688 (2020)
9. X. Shen, Z. Ni, L. Liu et al., WiPass: 1D-CNN-based smartphone keystroke recognition using WiFi signals. *Pervasive Mob. Comput.* **73**, 101393 (2021)
10. X. Cheng, B. Huang, CSI-based human continuous activity recognition using GMM-HMM. *IEEE Sens. J.* (2022). <https://doi.org/10.1109/JSEN.2022.3198248>
11. J. Wang, X. Zhang, Q. Gao et al., Device-free simultaneous wireless localization and activity recognition with wavelet feature. *IEEE Trans. Veh. Technol.* **66**, 1659–1669 (2017)
12. Y. Tian, S. Li, C. Chen et al., Small CSI samples-based activity recognition: a deep learning approach using multidimensional features. *Secur. Commun. Netw.* **2021**, 5632298 (2021)
13. L. Yang, H. Su, C. Zhong et al., Hyperspectral image classification using wavelet transform-based smooth ordering. *Int. J. Wavelets Multiresolut. Inf. Process.* **17**, 1950050 (2019)
14. E. Guariglia, Primality, fractality and image analysis. *Entropy* **21**, 304 (2019)
15. X. Zheng, Y. Tang, J. Zhou, A framework of adaptive multiscale wavelet decomposition for signals on undirected graphs. *IEEE Trans. Signal Process.* **67**, 1696–1711 (2019)
16. X. Liu, H. Zhang, Y. Cheung et al., Efficient single image dehazing and denoising: An efficient multi-scale correlated wavelet approach. *Comput. Vis. Image Underst.* **162**, 23–33 (2017)
17. E. Guariglia, S. Silvestrov, Fractional-wavelet analysis of positive definite distributions and wavelets on $d'(c)$, in *Engineering Mathematics II, Springer Proceedings in Mathematics and Statistics*, pp. 337–353 (2017)
18. Y.Y. Tang, *Document Analysis and Recognition by Wavelet And Fractal Theories* (The World Scientific Publishing Co, Singapore, 2012)
19. E. Guariglia, Harmonic Sierpinski gasket and applications. *Entropy* **20**, 714 (2018)
20. Z. Tian, J. Wang, X. Yang et al., WiCatch: A Wi-Fi based hand gesture recognition system. *IEEE Access* **6**, 16911–16923 (2018)
21. T. Zhang, T. Song, D. Chen et al., WiGrus: a WiFi-based gesture recognition system using software defined radio. *IEEE Access* **7**, 131102–131113 (2019)
22. H. Thariq, H. Ahmad, K. Narasingamurthi et al., DF-WiSLR: device-free Wi-Fi-based sign language recognition. *Pervasive Mob. Comput.* **69**, 101289 (2020)
23. D. Jiang, M. Li, C. Xu, WiGAN: a WiFi based gesture recognition system with GANs. *Sensors* **20**, 4757 (2020)
24. Z. Hao, Y. Duan, X. Dang et al., Wi-SL: contactless fine-grained gesture recognition uses channel state information. *Sensors* **20**, 4025 (2020)
25. S. Tan, J. Yang, Y. Chen, Enabling fine-grained finger gesture recognition on commodity WiFi devices. *IEEE Trans. Mob. Comput.* **21**, 2789–2802 (2022)
26. X. Zhang, C. Tang, K. Yin et al., Wifi-based cross-domain gesture recognition via modified prototypical networks. *IEEE Internet Things J.* **9**, 8584–8596 (2022)
27. Y. Gu, X. Zhang, Y. Wang et al., WiGRUNT: WiFi-enabled gesture recognition using dual-attention network. *IEEE Trans. Hum. Mach. Syst.* **52**, 736–746 (2022)
28. L. Davies, U. Gather, The identification of multiple outliers. *Publ. Am. Stat. Assoc.* **88**, 782–792 (1993)
29. F.R. Hampel, The influence curve and its role in robust estimation. *J. Am. Stat. Assoc.* **69**, 383–393 (1974)
30. K. Ali, A.X. Liu, W. Wei et al., Keystroke recognition using WiFi signals, in *The 21st Annual International Conference on Mobile Computing and Networking*, 7–11 September 2015, Paris, France, pp. 90–102 (2015)
31. W. Wang, A.X. Liu, M. Shahzad et al., Device-free human activity recognition using commercial WiFi devices. *IEEE J. Sel. Areas Commun.* **35**, 1118–1131 (2017)
32. J. Liu, Y. Wang, Y. Chen et al., Tracking vital signs during sleep leveraging off-the-shelf WiFi. In *The 16th ACM International Symposium on Mobile Ad Hoc Networking and Computing*, 22–25 June 2015, Hangzhou, China, pp. 267–276 (2015)
33. S. Kullback, R.A. Leibler, On information and sufficiency. *Inst. Math. Stat.* **22**, 79–86 (1951)
34. Z. Akhtar, H. Wang, WiFi-based gesture recognition for vehicular infotainment system—an integrated approach. *Appl. Sci.* **9**, 5268 (2019)
35. Z. Chikr-Elmezouar, I.M. Almanjahie, A. Laksaci et al., FDA: strong consistency of the KNN local linear estimation of the functional conditional density and mode. *J. Nonparametr. Stat.* **31**, 175–195 (2019)

36. R.C. Guido, A tutorial on signal energy and its applications. *Neurocomputing* **179**, 264–282 (2016)
37. R.C. Guido, ZCR-aided neurocomputing: a study with applications. *Knowl.-Based Syst.* **105**, 248–269 (2016)
38. R.C. Guido, A tutorial-review on entropy-based handcrafted feature extraction for information fusion. *Inf. Fus.* **41**, 161–175 (2018)
39. R.C. Guido, Enhancing teager energy operator based on a novel and appealing concept: signal mass. *J. Franklin Inst.* **356**, 2346–2352 (2019)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

Submit your next manuscript at ▶ [springeropen.com](https://www.springeropen.com)
