From Lens to Prism: Device-Free Modeling and Recognition of Multi-Part Activities

LIU WENYUAN1,2, WANG SIYANG1, WANG LIN1,2, (Member, IEEE), LI BINBIN3, (Member, IEEE), SU XING1, AND JING NAN1

1School of Information and Engineering, Yanshan University, Qinhuangdao 066004, China
2Key Laboratory for Software Engineering of Hebei Province, Qinhuangdao 066004, China
3School of Economics and Management, Yanshan University, Qinhuangdao 066004, China

Corresponding author: Wang Lin (wlin@ysu.edu.cn)

This work was supported in part by the NSFC under Grant 61672448, Grant 61772453, and Grant 61702440, in part by the Natural Science Foundation of Hebei Province under Grant F2018203444 and Grant F2016203176, in part by the Overseas Students Science and Technology Activities Project Merit Funding of Hebei Province under Grant CL201625, and in part by the Youth Fundation in Basic Research of Yanshan University under Grant 16LGA009.

ABSTRACT Exploiting radio frequency signals with different physical characteristics of refraction and reflection of different parts and movements of the human body, most of the existing work focuses on the identification of single activity from a special part, and the elimination of mutual interference in different activities recognition of various parts of the body is deliberately diluted. A reliable, real-time, and non-invasive human body combination activities recognition system, namely, WiCoach, is proposed in this paper. WiCoach extends the possibility of utilizing the concept of sensorless sensing and the effective use of wireless signal side channels. Leveraging well-designed signal features and robust denoising methods, WiCoach models the speed and duration between channel state information and the body’s multiple concurrent activities, and then recognizes activity of different parts and guides the individual fitness. The experimental results show that WiCoach can effectively discriminate the activities of different parts of the human body, and no longer relies on modeling the signal characteristics of single part, compared with the current work.

INDEX TERMS Activity recognition, device-free sensing, channel state information, side-channel, eight section brocade.

I. INTRODUCTION

Benefiting from the widespread deployment of wireless infrastructure and the universal enhancement of privacy protection, the radio frequency-based activity recognition forms a wide applications spectrum with contactless, wireless, and non-intrusive characteristics. The basic assumption of previous work is that the wireless signal in space is affected by the physical effects of different activities. Harnessing the refraction and reflection of the signal transmission, the different signal features enable to be captured in the time domain and frequency domain.

According to recognition algorithm, recent work in the field can be classified into two categories: the model-based and the pattern-based methods [1]. The former utilize the RF signals features constituted by the motion of different parts of the human body to model physical attribute with CSI. Specialy, Fresnel zone theory, as a remarkable modeling tool, has been widely apply to analyze and understand the possibilities and limits of device-free sensing. The latter harnesses a large number of feature learning and parameter training in offline stage, and exploits pattern recognition to recognize the human activity in online stage.

It is worth noting that the related work has focused on recognizing the single activity, for example, falling [2], breathing [3], smoking [4], keystroke [5], and so on. Such activity/behavior is concerned only with a part of the body, such as arms, legs, fingers, chest, lips, or the whole body as a rigid body in falling determination. As a general knowledge, many activities of human body are performed by multiple parts of the body, such as traditional Chinese body-building technique, shadowboxing or eight section brocade.

Taking advantage of the built-in sensors of wearable device, the information of fine-grained motion state in different parts of human body can be obtained, such as accelerometer, gyroscope, etc. Through the smart wrist band to obtain the user arm movement data, FitCoach [6]monitors the user’s single fitness activity. Femo [7] makes use of the RFID tag installed on the fitness equipment to obtain the user’s
activity target. Myovibe [8] captures activity information by installing customized muscle vibration devices on the user. Though wearable devices can capture fine-grained data such as speed, angle, strength, and amplitude for body parts or muscle movements, the wearable method increases the burden of the body in the perception of systemic data, and the price and maintenance costs are not conducive to commercialization. Depth camera and machine vision techniques are often exploited to recognize complex movements of the human body, has been extensively used in human-computer interaction and virtual reality. However, such techniques often require deployment of specific devices with potential privacy leaks.

In this paper, we will answer the question: can we use the Wi-Fi link of Commercial Off-The-Shelf devices to model and recognize multi-part activities in sensorless sensing pattern? WiCoach is designed to take a typical and popular body fitness exercise into account, eight section brocade (ESB), which is a Chinese traditional fitness composed of eight sections of movements, each section of movements by the hands, arms, head, neck, torso, legs, ankles, feet of different combinations constitute a compound activity or posture. Fig. 1 shows the eight sections of movements and gestures in the ESB, and some of the features of the CSI signals generated by these activities in frequency domain and time domain. It can be seen that there are significant differences between CSI images. This observation reveals that the differences in CSI signals generated by different combinations of activities can be modeled and identified. However, to achieve a device-free complex activity recognition system that combines reliability, robustness and timeliness, we have to overcome the following three challenges:

1) Concurrent interference of internal and external noises of an activity. On the one hand, the influence of different parts of the body on the signal is concurrent. On the other hand, due to the subcarriers are sensitive to different parts of the reflection signal differences, the multipath interference from the subcarriers of CSI is concurrent. These concurrent noises are called the internal and the external noise of activity respectively.

2) Local similarity interference between different activities. Since each group of activities in an ESB consists of multiple parts of the body, the CSI signal characteristics produced by the body’s local motion can be merged into any given section.

3) Repetition of the same activity can lead to false negative interference. Since each group of ESB needs periodic repetition during the learning process, beginners may produce different signal characteristics, and WiCoach needs to evaluate these differences in robust pattern.

To address these challenges, WiCoach firstly calculates the instantaneous frequency and energy of CSI signal at different activities by Hilbert-Huang Transform (HHT), instead of modeling a single movement or activity. Secondly, WiCoach extracts the signal features reflecting different activities, and leverages the regular path distance in dynamic time warping (DTW) to replace the Euclidean distance in KNN for classifying sections. Thirdly, WiCoach eliminates the difference between activities that result from repetition of the same activity. Finally, according to the FITT principle [9], three metrics are utilized to evaluate the effectiveness of the individual activity, including the duration of activity, the motion cycle, and the degree of physical concurrency.

The main contributions of this paper can be summarized as follows:

- We propose a HHT-based method to extract the activity feature. The method has high robustness and reliability for the segmentation and recognition of different activity on different body parts. The differential interference caused by the repetition of the same activity is eliminated by combining the regular path distance in the dynamic time wrapped.

- We demonstrate an algorithm among the activity duration, the interval time and the CSI signal feature. The algorithm quantitates the activities of different parts of the body in real time, which is a solid progress on commercialization of device-free sensing.

- We perform this system by multi-part activities recognition in various indoor environments with LOS and NLOS path. By an inertial measurement unit acquisition system, Xsens MVN, we collect the real movement data of whole body as the ground truth to evaluate and analyze the recognition accuracy in ESB.

To our knowledge, WiCoach is the first system that demonstrates the practicality of recognize the combined activities with strong similarity. Taking advantage of modeling the influence of multiple parts of the body, CSI signal features are converted to the speed and duration of the whole set of activities, instead of a single activity.

II. RELATED WORK

Our work is closely related to two research topics. One is Human activity recognition based on CSI, and another is physical exercises monitoring.
A. HUMAN ACTIVITY RECOGNITION BASED ON CSI
In recent years, human activity recognition based on CSI mainly includes vital signs monitoring [3], [10] and fall detection [2], contextual awareness [4], gait recognition [11], gesture recognition [12]–[14], activity recognition on daily activities in smart homes, etc [15]. Liu et al. [10] develop a system to tracking vital signs during sleep leveraging off-the-shelf Wi-Fi. Smokey [4] takes the first attempt to build a ubiquitous passive smoking leaves on CSI signals to identify the smoking activity. In addition, WiFiU [11] uses COTS Wi-Fi devices to capture fine-grained gait patterns to recognize humans. These studies show that CSI-based sensing system have the capability to accurately recognize human activities. Furthermore, the Fresnel zone model [1] is introduced to analyze the human activities in indoor environment, leveraging the model to depict RF signal-propagation properties and further perfect the theory of human activity recognition based on CSI. Unfortunately, the existing work have tackled this problem in an overly simplistic setting by assuming that users often carry out single part of the body at a time or multiple activities consecutively, one after another.

WiCoach focuses on the thorough study on concurrent or interleaving activities in a real world scenario, in other words, multiple sets of activities involving multiple parts of the human body.

B. PHYSICAL EXERCISES MONITORING
Another closely related aspect focuses on the automatically monitoring physical exercises. Over the past years there have been a number of solutions based on mobile devices with sensors [6], [7], [16]–[19]. Ding et al. [7] propose to recognize free-weight activities by attaching passive RFID tags on the dumbbells. In addition, Guo et al. [6] develop a virtual fitness coach leveraging users’ wearable mobile devices to assess dynamic postures in workouts. Alone this line, Zhao et al. [17] use event data recorder to provide various types of real-time information to cyclists to help them achieve their desired exercise results. Most of these technologies rely on additional sensors or specific hardware. Most importantly, whether can give an exercise feedback or guidance to improve the performance of exercise is still an open question.

Along this trend, WiCoach takes one step forward by utilizing commercial Wi-Fi devices to monitor physical exercises in sensorless matter. In particular, WiCoach can distinguish the lower body activity and quantitative analysis of exercise quality by non-invasive way.

III. SYSTEM OVERVIEW
WiCoach recognizes and assesses individual activities by modeling the dynamics of CSI caused by the ESB. The system modules are shown in Fig. 2. The CSI signals is captured using the Intel 5300 NIC as raw data. Through three steps of data preprocessing, activity recognition and activity assessment, WiCoach outputs the recognition and assessment result of an individual activity.

- **Data preprocessing.** WiCoach first uses the Hampel filter [20] to remove the original signal outliers, and then processes the pass signal through the Butterworth low-pass filter, thus the high-frequency part of the data will be filtered out and leave the low-frequency part implying the information of individual activity. Finally, using the idea that most of the information in the signal is concentrated on certain characteristics, WiCoach applies the principal component analysis (PCA) to reduce the feature dimension, removes the interference of the sub-carrier correlation, and further reduces the noise of the signal.

- **Activity recognition.** ESB consists of eight routines, and each routine is composed of five to seven identical activities. Therefore, WiCoach needs to recognize the routine to which the activity belongs. To extract more fine-gained features, the data are sliced based on density clustering algorithm so that the data belonging to the time period of the activity enable to be segmented. Furthermore, HHT is leveraged into the segmented data to extract the features and create a feature matrix. Finally, the warping path distance in DTW is exploited to classify the user activity, instead of the Euclidean distance in KNN.

- **Activity Assessment.** We take assessing the effectiveness of the recognized activities into account. As ESB, too fast or too slow movement will lead to the decline of the
fitness effect. Therefore, the duration of each activity and the period between activities should be maintained in a stable and reasonable range. Moreover, the movement intensity is one of the important indexes for the physical activity assessment. Using these three indicators, we quantitatively evaluate the individual’s movement quality.

IV. SYSTEM DESIGN

A. PREPROCESSING

As the raw CSI data obtained from commercial Wi-Fi devices contain a lot of noise and cannot be extracted directly from the raw data into fine-grained activity information, the sensing system need preprocessing the data. The preprocessing part of the WiCoach system consists of three procedures: outlier removal, low-pass filtering and principal component analysis (PCA).

1) OUTLIER REMOVAL

Internal state transition such as transmission rate adaptation and transmission power changes introduces burst noise in the CSI streams [21]. Fig. 3(a) shows the original CSI data of a segment of routine 2 of ESB. Some mutations can be observed and they are CSI changes caused by burst noise. Since WiCoach recognizes and assesses user activities by analyzing CSI changes caused by activities, these changes are not caused by activities that greatly affect the performance of the system. WiCoach uses the Hampel identifier [22] to remove these outliers. Hampel identifier first uses a sliding window to detect all sample points one at a time. If the value of a point is out of the interval \([\mu - \gamma \times \sigma, \mu + \gamma \times \sigma]\), the point is treated as an outlier and replaced by the median of the current window, \(\mu\), \(\sigma\) is the median and absolute median of the current window, \(\gamma\) is a parameter that affects the detection sensitivity of outliers, and the size of the window in WiCoach is 100, \(\gamma\) is 2. Fig. 3(b) is the CSI data after the Hampel identifier removes the outliers.

2) LOW-PASS FILTERING

Carrier frequency offset (CFO) is another reason for CSI to contain a lot of noise. We found that CFO generally has higher frequency and that most of the changes in the CSI caused by activities are in the low-frequency portion of the signal. Therefore, WiCoach uses Butterworth low-pass filter to remove the high-frequency noise from CSI signal; the cutoff frequency is 20Hz. Fig. 3(c) shows the filtered CSI signal, we can observed that most high-frequency noise has been successfully filtered.

3) PRINCIPAL COMPONENT ANALYSIS (PCA)

After Hampel identifier and low-pass filtering, the CSI signals still cannot be used to recognize. It is mainly due to the following two reasons: First, there is high-bandwidth impulse noise in the CSI, which cannot be removed well by low-pass filters alone [23]. Second, on the 802.11n protocol, we can obtain 30 OFDM subcarriers from each Wi-Fi link that each with its own frequency. Based on the Fresnel zone model [24], it is impossible to know which subcarriers are sensitive or insensitive to activity when the relative position is uncertain between the target and the antennas. For the above two reasons, we apply principal component analysis (PCA) to extract the principal components that are primarily affected by human activity, based on the fact that most of the information in the data is focused on certain characteristics. Since the second principal component contained the least noise [25], we choose the second principal component as the analysis data. Fig. 3(d) demonstrates the second principal component after PCA.

B. ACTIVITY SEGMENTATION

For ESB, each routine is compose of 5 to 7 independent activities, and there is a time interval between both activities. Therefore, we first split the CSI data segments caused by human activities.

We determine the segmentation algorithm based on an insight. As the motion of the object changes the transmission path of the signal, the CSI value fluctuates significantly when the activity occurs (gain suddenly increases or weakens). On the contrary, the CSI value is in a relatively stable state. Therefore, we apply a dynamic threshold algorithm to realize activity segmentation. The segmentation algorithm mainly consists of the following four steps.

In first step, WiCoach selects the second principal component and computes the first order difference of it. As above mentioned, the human activity engenders a significant gain in the CSI signal. Therefore, by computing the first order difference, the wave will fluctuate around 0, and when an activity is performing, the amplitude of the wave will be very
large. Meanwhile, in the intermittent phase the amplitude will be almost 0. As a result, we can detect all activities with a single threshold.

In second step, WiCoach detects the increase and decrease in the data at the same time, and then computes the absolute value of the first order difference. The reason is that the sudden increase in the data will take on a larger positive value in the first order difference. Correspondingly, a sudden decrease in the data will take on a larger negative value in the first order difference.

In third step, WiCoach traverses the absolute value of the first order difference and locate the first point where the value is greater than the threshold $Thr$. This moment is recorded as $T_{s1}$. Moreover, starting from this point, WiCoach traverses the next sample and locating the first point whose value is less than $Thr$. The moment is recorded as $T_{e1}$, and we notice that $\omega_1 = (T_{s1} + T_{e1})/2$. Repeat the above steps until all of samples are traversed. As an output, WiCoach may get $n$ time groups: $\{T_{s1}, \omega_1, T_{e1}\}, \{T_{s2}, \omega_2, T_{e2}\}, \ldots \{T_{sn}, \omega_n, T_{en}\}$ that each group contains a waveform with a large amplitude.

Finally, since ESB is a complex multi-part activities of whole-body, and each activity may contain several waveforms, which means that each activity contains several groups. Therefore, we need to combine those groups that belong to the same activity. In this paper, we define the difference between $\omega$ in each group as the inter-distance distance. We then defined a threshold $\theta$, if the distance of inter-group is less than $\theta$, they are considered to belong to the same activity group. The benefit from this step is that there is only one time group for each activity. We treat $T_s$ of each group as the beginning of the activity and $T_e$ as the end of the activity.

Based on the ground truth that we get from MVNX, $\theta$ is set to 2. Fig. 4 shows the CSI data of the routine 1 after the activity segmentation. For each segment of data after the activity segmentation, we can then extract the fine-grained activity feature from it.

![FIGURE 4. Second PC of routine 1 after activity segmentation.](image)

### C. ACTIVITY CLASSIFICATION

1) CONSTRUCTING FEATURE MATRIX

WiCoach extracts the frequency and amplitude information of CSI from different activities to obtain the features of ESB. This is because there are significant differences in two aspects between the activities of different routines: amplitude and speed of movement of the body. For example, the routine 8 has a smaller amplitude of movement, but a faster movement speed. On the contrary, the routine 2 has a larger amplitude of movement and a slower movement speed. Moreover, the frequency in the signal represents the multipath change speed caused by the body movement, while the amplitude represents the energy of the signal. Therefore, by analyzing the frequency and amplitude information in the signal, we can get the speed and amplitude characteristics of the body movement related to them.

In this paper, we apply the Hilbert-Huang transform (HHT) to extract the frequency and amplitude information in the signal. The choice of HHT is mainly based on the following two reasons: First, HHT is a completely based on the signal itself time-frequency processing technology, without prior determination of the basis function or window function. Second, HHT can calculate the instantaneous frequency and amplitude of the signal, and it has better performance in detecting the similarity between complex activities and classifying activities with different duration and amplitude. HHT algorithm consists of two parts: empirical mode decomposition (EMD) and Hilbert transform (HT): First, EMD decomposes CSI into a set of $imfs$ according to the scale characteristics of CSI itself:

$$\xi_{csi}(t) = \sum_{j=1}^{n} imf_j(t) + r_n(t)$$  \hspace{1cm} (1)

where $r_n(t)$ is the remainder and $n$ is the number of $imf$, then the instantaneous frequency $f$ and the instantaneous amplitude $a$ at any moment can be calculated by applying HT to each $imf$.

$$[f_j(t), a_j(t)] = HT(imf_j(t))$$  \hspace{1cm} (2)

In this way, for each activity $i$, we can get a feature matrix about it through HHT, which is used to recognize and assess the activity. The feature matrix consists of a set of instantaneous frequencies and a set of instantaneous amplitudes.

$$H_i = [f_{i1}, f_{i2}, \ldots f_{in}, a_{i1}, a_{i2}, \ldots a_{in}]$$  \hspace{1cm} (3)

where $f$ is the instantaneous frequency, $a$ is the instantaneous amplitude, and $n$ is the number of $imf$. Because of space constraints, we do not give the theory of HHT and more details, interested readers can refer to [26].

2) ACTIVITY CLASSIFICATION

In this part, we will describe how to classify activities using the feature matrix. We note that two new challenges need to be solved in activity classification of ESB. One is that there is a large difference between the activities of different users. Another is that with the repetition of the activities in each routine, there is a large different between the former activities and the latter. For the above findings, we use the warping path distance to describe the difference between the routines. The warping path distance represents the distance of the shortest
warping path between two unequal signals in the dynamic time warping (DTW). The smaller the warping path distance, the greater the similarity between the signals. Conversely, the larger the warping path of the two signals, the less the similarity between the signals. Furthermore, we use the k-NN classifier and replace the Euclidean distance in the k-NN similarity between the signals. Furthermore, we use the k-NN classifier with the warping path distance. The detailed process is described as follows.

In first step, for each unknown activity $i$, we select $f_{1i}$, $a_{1i}$, $\xi_{csi}$ to constitute the classification matrix $C_i$ of activity $i$.

$$C_i = [f_{1i}, a_{1i}, \xi_{csi}]$$  \hspace{1cm} (4)

We chose $f_{1i}$ and $a_{1i}$ because they led to the highest accuracy in our experiment.

In second step, for each component in $C_i$, such as $f_{1i}$, we apply k-NN classifier, which essentially computes the warping path distance between the $f_{1i}$ and the $f_{j}$ of all the activities in the training set, and then define the category of unknown activity $i$ as the most frequent activity category among the k smallest distances. Similarly, the same is done for $a_{1i}$ and $\xi_{csi}$.

$$R_1 = \text{KNN}(f_{1i})$$  \hspace{1cm} (5)

$$R_2 = \text{KNN}(a_{1i})$$  \hspace{1cm} (6)

$$R_3 = \text{KNN}(\xi_{csi})$$  \hspace{1cm} (7)

where R represents the classification category after k-NN. In our experiment, k=4.

In final step, we apply the voting mechanism to determine the final classification result R:

$$R = \text{volt}(R_1, R_2, R_3)$$  \hspace{1cm} (8)

D. ACTIVITY ASSESSMENT

The purpose of activity assessment is to show the quality of the user’s activities and provide feedback to user. Unlike other fitness sports, such as free weight fitness, ESB as a complex sport of the whole body, moving too fast or too slowly will greatly reduce the effectiveness of the exercise, and the user must ensure the coordination of the whole body, so that the whole body is fully exercised. This requires the entire practice time to maintain a stable and appropriate range, while ensuring that the upper and lower body are fully exercised. Therefore, we adopt the user’s activity duration, activity period, and activity intensity to describe the effectiveness of user activity.

1) DURATION

WiCoach achieves the duration of each activity in each routine by calculating the length of the segment in Section III-B:

$$\delta = T_e - T_s$$  \hspace{1cm} (9)

2) PERIOD

We define the distance between two segments as an activity period:

$$P_{ij} = |\omega_{j} - \omega_{i}|$$  \hspace{1cm} (10)

3) INTENSITY

Intensity is another important indicator of the effectiveness of an activity. It reflects the amount of energy consumed by the user during the activity [3]. We can note that, as shown in Fig. 5, some routines need to be involved in the upper and lower halves (e.g., the routine 1, routine 2, routine 3 and routine 4). Some routines consist only of upper body movements (e.g., the routine 5, routine 6 and routine 7), while the routine 8 includes only the lower body movements. Therefore, WiCoach gives a specific assessment of the intensity according to the user’s activity category. Furthermore, as the whole body activity, the movement speed of the upper body is greater than the movement speed of the lower body. In the feature matrix $f_1$ is always greater than $f_2$. As a result, we select the energy vectors $a_1$ and $a_2$ corresponding to $f_1$ and $f_2$ to represent the intensity values of the upper and lower body, and for the routine involving only the upper or lower body, we extract only $a_1$ to represent the intensity values of the activity.

FIGURE 5. ESB Classification.

V. EVALUATION AND RESULTS

A. EXPERIMENTAL SETUP

1) PARTICIPANTS

We invited 6 volunteers including 2 females as the detecting target to collect the CSI data. These volunteers are university students and researchers, ranging in age from 23 to 40, average age 29, weight distribution of 49kg to 82kg, average weight of 74kg, and height distribution of 160cm to 175cm, average height of 170cm.

2) IMPLEMENTATION

We used TP-Link WDR5300 as AP to send data and the PC equipped with Intel 5300 NICs as MP to receive data. To evaluate the scalability of WiCoach, the experiments were conducted in two typical indoor environments: a lab and a hall. Experiments were first conducted in the laboratory’s typical multipath indoor environment. As shown in Fig. 6(a), the laboratory is 7.8 m in length and 7.2 m in width, and is furnished with office equipment such as computers, desks and chairs. The location of the AP and MP was shown in
FIGURE 6. Photographs and layouts of the two experiments. (a) Experimental scenario of the 1st experiment. (b) Layout of 1st experiment. (c) Experimental scenario of the 2nd experiment. (d) Layout of the 2nd experiment.

Moreover the second experiment was conducted in an empty hall with relatively few multipath. The hall is 13.6 m in length and 5 m in width. In this environment, the AP and MP are located at both ends of the hall respectively, as shown in Fig. 6(c) - (d).

3) GROUND TRUTH
We used the Xsens MVN to collect the ground truth. The Xsens MVN is a whole-body inertial motion capture system. It captured and inertial data of all parts of the body by wearing seventeen sensor units throughout the user’s body and simultaneously generates video. Fig. 7(a) shows the position change of the head sensing unit and CSI time amplitude image in routine 1, and Fig. 7(b) shows the video generated by Xsens MVN. In order to achieve the reasonable experimental results, during the experiment, we asked the volunteers wearing Xsens MVN to perform the ESB to collect ground truth, while using WiCoach to collect CSI data of the movements at that time. By comparing the ground truth, we made a comprehensive evaluation of the system performance of WiCoach.

FIGURE 7. Captured CSI feature v.s. ground truth captured by Xsens MVN. (a) CSI time amplitude image of routine 1. (b) The video from Xsens MVN.

4) DATASETS
We collected 1203 activity samples from 6 subjects to build our data set. Specifically, the environment of laboratory and the hall contain 739 samples and 464 samples, respectively. Then, we took 30% of the data from each of the two environments as two training sets, and the rest as two test sets. In addition, for different subjects, we trained our system with a single training set, and for different environments, we trained our system with two training sets separately.

B. EVALUATION ON ACTIVITY SEGMENTATION

1) EVALUATION METRIC
We use the following five metrics to evaluate our activity segmentation schemes: insertion rate, deletion rate, fragmentation rate, merge rate, and accuracy [14]. The former four metrics are used to evaluate the robustness of the activity segmentation schemes, while the last one shows the accuracy of the scheme as a whole. The detailed explanations for these metrics are as follows.

- **Insertion rate**: $\frac{N_I}{N_e}$, represents the probability that a period that is not an activity will be cut out as if an activity had occurred, where $N_I$ represents the number of time slices incorrectly identified as routine $e$. $N_e$ represents the number of activities in routine $e$ in ground truth. It demonstrates how resistant WiCoach is to outside interference.

- **Deletion rate**: $\frac{N_D}{N_e}$, represents the probability that WiCoach will lose an activity, where $N_D$ represents the number of activity that are not split in routine $e$. The deletion rate represents how sensitive WiCoach is to the occurrence of an activity.

- **Fragmentation rate**: $\frac{N_F}{N_e}$, represents the probability of WiCoach splitting an activity into multiple slices, where $N_F$ represents the number of fragmented activity of the routine $e$, a fragmentation rate that reflects WiCoach’s ability to handle complex, non-coherent activities.

- **Merge rate**: $\frac{N_M}{N_e}$, represents the probability that a segment contains multiple activities, where $N_M$ represents the number of combined activities of the routine $e$, and the merge rate reflects WiCoach’s ability to handle more continuous activities with shorter intervals.

- **Accuracy**: $\frac{#correctlydetectedactivity}{#activitysthatareperformed}$, represents the overall accuracy of the WiCoach activity segmentation scheme.

2) EVALUATION RESULT
Fig. 8(a) shows the fine-grained result of our activity segmentation scheme. For the insertion rate, the rate of routine 6 is zero, and the rate of the routine 1, routine 3, routine 5 and routine 7 distributed between 0.03 and 0.04. Finally the rate of the routine 2, routine 4 and routine 8 increases slightly to 0.06 $\sim$ 0.08. The average rate of the eight routines as a
whole is 0.04, which proved that our scheme has good anti-interference performance.

For the deletion rate, the rate is 0 for the routine 1, routine 2, routine 3, routine 4, routine 5 and routine 8, the deletion rate of the routine 6 and routine 7 is from 0.02 to 0.03. Finally, the average deletion rate is 0.006. The result indicates that, for all routines, WiCoach is sensitive to CSI changes incurred by the activity.

For the fragmentation rate, the rate of the routine 2, routine 3, routine 4, routine 5 and routine 8 is zero, it then slightly increases to 0.01 in routine 7, and for the routine 1 and routine 6, the fragmentation rate increases to about 0.08. This is because the routine 1 and routine 6 are composed of two stages, compared with other routines, their activity is less coherent. However, for all routines, the average fragmentation rate is 0.02. The result shows that WiCoach has a good performance at segmenting complex, incoherent activity.

Fig. 8(a) also suggests that the merge rate of the routine 1, routine 2, routine 6, and routine 7 is 0, then the merge rate increases for another routine. The average merge rate of eight routines is 0.04, which indicates that WiCoach has a desirable performance to segment the relatively continuous activities.

The overall accuracy of segmentation schemes of WiCoach compared with SEARE [27] is shown in Fig. 8(b). SEARE segments the entire CSI by applying First-order Difference (FOD) and Fast Fourier Transform (FFT) to identify the start and end of free weight fitness movements. We can see that the overall segmentation accuracy for WiCoach is very high (between 0.82 and 0.93), while that for SEARE is very low (between 0.03 and 0.16). This is because unlike free weight movements, which only cause sharp changes in the CSI at the beginning and end of the activity, the ESB movements are more complex, not only at the beginning and at end of the activity, but also in the course of the activity, the fluctuations of the CSI are very violent. Therefore, the way in which the moment of sharp change in CSI is considered the starting and ending point of an activity, as in the case of SEARE, is not a good way to segment the ESB activities. We can observed from the Fig. 8(b) that the average accuracy of all the routines for WiCoach is 0.88, which indicates that our segmentation scheme can segment most activities of ESB.

In addition, Fig. 8(c) shows the impact of different principal components on accuracy as the final output for the segmentation. In addition, we use the algorithm in WiKey [5] as the benchmark. WiKey applies an algorithm based on mean absolute deviations to integrate the second to fourth principal components as the final output for activity segmentation. The figure shows that when the first principal component as the output, the accuracy rate is lowest since the first principal component contains most of the noise. It also shows that the second principal component has the highest accuracy rate of 0.91, while the third and fourth principal components have a slight decrease of 0.82 and 0.83. This is because they contain less effective information than the second principal component. In particular, the accuracy of WiKey is 0.87. In order to ensure the real-time performance of the system, we directly choose the second principal component as the final output for activity segmentation.

C. EVALUATION ON ACTIVITY CLASSIFICATION

1) EVALUATION METRIC

We use the following three metrics to evaluate WiCoach’s classification scheme:

- **Precision**: \( \frac{N_e^T}{N_e^T + M_e^T} \), where \( N_e^T \) represents the number of all activities that are correctly classified as routine \( e \), and \( M_e^T \) represents the number of all activities of other routines that are wrongly classified as routine \( e \).

- **Recall**: \( \frac{N_e^T}{N_e} \); the recall for routine \( e \) is defined as the ratio of the number of correctly classified activities of routine \( e \) to the number of all activities of routine \( e \) in ground truth.

- **F1-score**: \( F_1^{(e)} = 2 \times \frac{\text{precision}_e \times \text{recall}_e}{\text{precision}_e + \text{recall}_e} \), which represents the harmonic average of precision and recall. It reaches its best value at 1 and worst at 0.

2) EVALUATION RESULT

Fig. 9(a) shows the confusion matrix of the eight routines of volunteers. Each row represents the true category of the activity, while each column represents the category of the activity predicted by WiCoach. Each element represents a subset of all the activities in this row, which are predicted to be
the activity category in this column. In particular, we can see that the average classification accuracy of the ESB is 0.92 and the standard deviation is 0.07. The highest classification rate is 0.94 (routine 1, routine 6) and the lowest is 0.8 (routine 3). This is due to the fact that the duration of the routine 3 is very similar to that of the routine 8, are 3-6 seconds. Moreover, the routine 3 and the routine 1 both have raised hands, so that some of the activities of routine 3 are mistakenly classified as the routine 8 and the routine 1. However, the results of all experiments show that WiCoach has a good and stable performance in activity classification.

In addition, Fig. 9(b) shows the rate of evaluation metrics of WiCoach and WIAG [10], we can see that the precision, recall, and F-1 score of WiCoach are about 0.91, significantly better than the WIAG experimental results. This is because, in WIAG, in order to apply wavelet transform (DWT), all the CSI data of the activity were unified by smoothing spline to 1024 points (sampling frequency is 100HZ). However, for ESB, this approach is not suitable because the duration of each activity varies greatly (from 3s to 12s), so that the difference of data between activities is reduced when all activity data are unified by smoothing spline to the same length. Moreover, WiCoach uses HHT to generate feature matrices, and uses regular path distance instead of Euclidean distance in k-NN for activity recognition. Compared with WIAG, which decomposes details coefficients into blocks, averages each block to generate feature vectors and puts them directly into k-NN for activity recognition, our classification algorithm not only extracts the frequency and energy information effectively, but also has better robustness to recognize the activity with unequal length. Therefore, our comparative experiments show that WiCoach has higher accuracy and better robustness in recognizing whole-body complex movements.

Fig. 9(c) compares the performance of HHT with common features including mean value, maximum value, minimum value and standard deviation. The results show that the precision, recall and F1-score based on HHT remain around 0.91, while those based on common features are around 0.57, 0.63, 0.31 and 0.66, respectively. The reason is that these common features are not effective in extracting the differences between different actions in both time and frequency domains.

### D. EVALUATION ON ACTIVITY ASSESSMENT

WiCoach is designed to present the user the detailed assessment information about duration, period and intensity of his/her activities. We apply \( \sigma = 1 - \frac{R_P}{R_g} \) to compare the results of the three metrics predicted by WiCoach with their ground truth, where \( R_P \) to represent the predicted results of WiCoach, and \( R_g \) to represent the true result of the metrics in the ground truth of the user’s activities, and a higher \( \sigma \) indicates that \( R_P \) is closer to \( R_g \). Fig. 10 shows the evaluation results of the performance of the activity assessment. It can be seen that all the \( \sigma \) of the three metrics are between 0.8 and 0.95, which shows that the assessment results of WiCoach are very close to the real results of users, which proves that WiCoach can accurately feedback the assessment of activities.

### E. RUNTIME PERFORMANCE

To evaluate the runtime performance of our system, we measure the average runtime of each stage of the radio signal processing pipeline for WiCoach. Specifically, we have the six volunteers perform a full set of ESB and collect their CSI data. We then record the average runtime of the program at each stage of processing the data and calculate the total run time of the entire processing pipeline. The results are recorded in Table 1. From the table, we note that the process of CSI extraction takes the most amount of the total processing time, followed by activity classification. Furthermore, it shows that the total run time of the processing pipeline is kept within 1.5s. This indicates that our system can achieve real-time performance.
TABLE 1. Runtime performance.

<table>
<thead>
<tr>
<th>Processing Pipeline</th>
<th>Processing Time (second)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSI Collection</td>
<td>0.683</td>
</tr>
<tr>
<td>Preprocessing</td>
<td>0.053</td>
</tr>
<tr>
<td>Activity Segmentation</td>
<td>0.065</td>
</tr>
<tr>
<td>Feature matrix Creation</td>
<td>0.284</td>
</tr>
<tr>
<td>Activity Classification</td>
<td>0.325</td>
</tr>
<tr>
<td>Activity Assessment</td>
<td>0.052</td>
</tr>
<tr>
<td>Total</td>
<td>1.482</td>
</tr>
</tbody>
</table>

F. IMPACT OF K ON ACTIVITY CLASSIFICATION

In this part of the experiment, we analyze the impact of the value of K in k-NN on the accuracy of activity classification. The results are shown in Fig. 11. We can see that in the laboratory and hall environments, the trend of the results is approximately the same. As K starts at 1 and increases gradually, the accuracy has a clear upward trend and reaches a peak of 0.91 (K=4), then the accuracy starts to decrease as K values continue to increase. This phenomenon is caused by the characteristics of the k-NN algorithm itself. The smaller the K value, the smaller the approximate error of classification, but prone to over-fitting phenomenon. Conversely, increasing the K value will increase the systematic deviation of k-NN. Therefore, according to the experimental results, we finally take the K of WiCoach as 4.

![FIGURE 11. Impact of K in two environments.](image_url)

G. IMPACT OF KEY PARAMETERS ON ACTIVITY SEGMENTATION

In this experiment, we analyze the impact of two key parameters of WiCoach’s activity segmentation scheme, Thr and θ. Thr is the threshold for detecting the fluctuation of CSI data. The smaller the Thr, the more sensitive the algorithm is to the fluctuation of CSI. θ is the threshold for distance between groups, which determines whether two adjacent groups are merged. According to our definition in 5.2, the value of Thr affects the insertion rate, the deletion rate, and θ affects the insertion rate, the deletion rate, both of Thr and θ affect the accuracy.

The results of the evaluation on Thr are shown in Fig. 12(a). We can see that when Thr increases from 0.06 to 0.10, the insertion rate becomes a significant downward trend, and after 0.1, it gradually flattens out, because the fluctuation range of CSI is basically less than 0.1 in the interval phase of the activity, and when Thr is greater than 0.1, the activity interval is not mistakenly segmented. Then as Thr continues to increase from 0.11 to 0.15, the deletion rate tends to increase significantly. This is because an increase in Thr means a decrease in sensitivity to fluctuations in CSI, and Thr is too large causes movements of little magnitude not to be segmented. Therefore, for accuracy, when Thr increased from 0 to 0.1, the accuracy increased to a peak of 0.9 due to a decrease in the insertion rate, and when Thr continued to increase to 0.15, the deletion rate increased sharply, resulting in a decrease in accuracy.

![Fig. 12(b) shows the impact of the size of θ on the experimental results. It can be seen that when θ is increased from 2.5 to 4.5, the merger rate tends to increase significantly. Because the larger θ, the greater the chance that two activities will be merged into one segment, this results in a high rate of merging. Conversely, if θ is too small, the greater the chance that an activity will be mistakenly identified into multiple segments, increasing the fragmentation rate. So for accuracy, theta increases from 0 to 2, the merger rate is low, the fragmentation rate drops sharply, resulting in a significant increase in accuracy to a peak of 0.91, then θ continues to increase, as the merger rate rises sharply, the accuracy rate begins to decrease significantly.](image_url)

H. IMPACT OF DIFFERENT ENVIRONMENTS ON RECOGNITION PERFORMANCE

In this experiment, we analyze the impact of different experimental environments on the performance of WiCoach activity classification. We performed this part of the experiment in the two experimental scenarios mentioned in V, and the results are shown in Fig. 13. It can be seen that more than 85% of the activities, whether in the laboratory with more paths or in the hall with less paths, have classification accuracy of more than 0.8. This indicates that WiCoach can achieve high accuracy in both experimental environments. In addition, it can be seen that more than 95% of the activities in the hall have recognition accuracy of more than 0.8, while in the laboratory is 85%, indicating that the results of the experiments in the hall are better than those of the laboratory. This is due to the relatively stable experimental environment in the hall.

I. IMPACT OF USER DIVERSITY

In this experiment, we evaluate the impact of user diversity on the accuracy of activity classification. For each subject, we calculated the accuracy of their different routine and showed the results of the distribution of accuracy in Fig. 14. From the figure, the accuracy of the six volunteers were kept at a high level. The average accuracy of the five volunteers was above 0.8, with the third one having the highest accuracy rate of 0.9. This was because the third subject had the longest time to practice the ESB and performed the most standardized movements. In addition, the accuracy of the second one was lowest (0.76), since the second volunteer was a novice. In particular, We further analyzed the detailed distribution of each subject’s accuracy, which showed that volunteers with more standard movements had shorter box...
lengths (e.g., volunteers 1 and 3) and volunteers with less standard movements had longer box lengths (e.g., volunteers 2 and 5), due to the fact that non-standard movements reduced WiCoach’s accuracy in recognizing certain activity.

J. IMPACT OF DIFFERENT Locations

In this experiment, we analyze the impact of different locations on the performance of WiCoach activity classification. We compared the performance of WiCoach’s activity classification in both LOS and NLOS, and the results are shown in Fig. 15. It can be seen that more than 93% of the activities, whether in the LOS or in the NLOS, have classification accuracy of more than 0.8. This indicates that our system is robust to location changes. In addition, we found that the results in NLOS were slightly better than those in LOS. The probable cause is that an individual hinders signals propagation.

VI. CONCLUSION

Compared with wearable devices and depth cameras with active sensing features, CSI-based passive sensing method has many advantages, such as non-intrusive, device-free, and so on. Due to the openness and uncertainty of the uncontrolled activity, most of the existing researches adopt the control experimental environment and individual performance evaluation. However, it is difficult for the system to ensure effective scalability. We try to identify multiple objects for a series of activities. Taking the concurrent, continuous and similar high characteristics into account, the proposed system provides new trail of thought for the design of real-world applications with sensorless sensing. In addition, the current system is designed for and tested with only a specific multi-part activity (ESB), but our denoising algorithm and HHT do not depend on this particular activity. This further extends the extensibility of our system. Therefore, although our system requires training to extract features, our current algorithm of denoising and feature extraction ensure the possibility of applying our system to other exercise recognition.

REFERENCES


WANG SIYANG received the B.S. degree in mechanical design, manufacturing and automation from the Hebei University of Engineering, Handan, China, in 2016. He is currently pursuing the M.S. degree in the computer science and technology with Yanshan University. His research interests include device free wireless localization and recognition.

WANG LIN (M’11) received the B.Ed. degree from Northwest Normal University in 2003, the M.E. and Ph.D. degrees in computer engineering from Yanshan University in 2009 and 2015, respectively. He was a Visiting Scholar with the Wireless Networking Laboratory, Illinois Institute of Technology. He has been a Visiting Student with the GreenOrbs Group since 2010. He is currently an Associate Professor with the School of Information Science and Engineering, Yanshan University. His research interests include mobile networks and RFID.

LI BINBIN (M’17) received the B.E. and M.E. degrees from the School of Computer and Technology, Harbin Institute of Technology, China. He is currently pursuing the Ph.D. degree with the School of Information Science and Engineering, Yanshan University, China. He is currently an Assistant Professor with the School of Economics and Management, Yanshan University. His research interests include mobile networks and RDF.

SU XING received the B.S. degree from Shijiazhuang Tiedao University, Shijiazhuang, China, in 2016. He is currently pursuing the M.S. degree in computer science and technology with Yanshan University. His research interests include device-free indoor localization and activity recognition, wireless networks, and mobile computing.

JING NAN received the B.S., M.S., and Ph.D. degrees in circuit and system from Yanshan University, Qinhuangdao, China, in 2001, 2004, and 2014, respectively. Since 2013, she has been an Associate Professor with the School of Information Science and Engineering, Yanshan University. Her research interests include sparse channel estimation for MIMO OFDM system, sparse signal recovery and its applications in communications, wireless sensor networks, and mobile networks.

LI WENYUAN received the B.S. and M.S. degrees from the Northeast Heavy Machinery Institute, China, and the Ph.D. degree from the Harbin Institute of Technology in 2000, all in computer science. Since 1996, he has been with the School of Information Science and Engineering, Yanshan University, Qinhuangdao, China, where he is currently a Chair Professor. He is also a special Invited Researcher at IoT, TNlist, Tsinghua University. His research interests include wireless sensor networks and mobile networks.