WILEY | Hindawi

Research Article

MoSeFi: Duration Estimation Robust Human Motion Sensing via Commodity WiFi Device

Xu Wang [ID],[1,2] **Linghua Zhang** [ID],[1,3] **Qin Cheng** [ID],[1] **and Feng Shu** [ID][1]

[1]*School of Communications and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China*
[2]*College of Mechanical Engineering, Optoelectronics and Physics, Huaihua University, Huaihua 418000, China*
[3]*Jiangsu Engineering Research Center of Communication and Network Technology, Nanjing 210003, China*

Correspondence should be addressed to Linghua Zhang; zhanglh@njupt.edu.cn

Accurate motion interval segmentation is the basic and crucial step in the advanced human perception based on WiFi signals. However, previous works have rarely considered motion duration, which is one of the important parameters for complete description of human motion. On this basis, we deeply investigate the properties of the CSI ratio from the perspective of Mobius transformation and construct a novel motion indicator using its complementary real and imaginary parts. The new indicator can attenuate the impact of motion fragmentation under short-window conditions and significantly reduce the duration error while ensuring detection accuracy. Moreover, we propose a universal subcarrier screening method based on response sensitivity and shape similarity, which provides more accurate information for perception. Furthermore, we present MoSeFi—a duration estimation robust human motion detection system using an existing commercial WiFi device. Detailed experimental results demonstrate that MoSeFi is lightweight yet effective compared to state-of-the-art systems.

1. Introduction

Perceiving human motion in the region of interest is the basic task of context awareness, which will enable various intelligent applications and services, including monitoring, control, and analysis. In recent years, with the widespread deployment of WiFi devices in daily life scenarios, ubiquitous awareness based on WiFi signals has attracted the attention of an increasing number of researchers. Compared with traditional video-based and device-based methods, reuse of the existing WiFi devices for passive human motion sensing has many advantages, such as no additional cost, no privacy disclosure, and no line-of-sight (LOS) path limitation. Thus, a large number of WiFi-based sensing technologies have emerged.

The wireless signal is affected by the surrounding environment in the process of propagation, causing variations in signal amplitude, phase, and frequency. By analysing the modulated signal, we can obtain relevant environmental information. Early wireless sensing used the received signal

strength indicator (RSSI) to implement application [1–3]. However, the coarse-grained RSSI measures the average receiving signal intensity over multiple propagation paths, which limits its stability and reliability. In recent years, researchers have turned their attention to fine-grained channel state information (CSI), which contains a richer environmental context. Compared with the RSSI, CSI describes the multipath propagation effects of wireless signals to a certain extent and provides more detailed and robust features for advanced environmental perception. Therefore, multiple CSI-based sensing methods have been proposed, including indoor location [4, 5], intrusion detection [6, 7], behavior classification [8, 9], and gesture recognition [10–12].

In these applications, distinguishing the static and moving states of the human body is a basis and crucial step for subsequent processing. A good system must not only be able to correctly detect the human motion but also accurately segment the motion interval. Here, we call the former detection accuracy, which is measured by false positive (FP) and false negative (FN), and the latter duration accuracy, which

can be represented by the difference between the detected duration and actual duration. Among them, the duration accuracy is crucial for applications such as identification and classification, because in these cases, we prefer to obtain pure CSI fragments that contain only complete motion information. However, most previous work focused on motion detection accuracy while ignoring duration accuracy, which results in an incomplete description of motion intervals. In addition, modern wireless communication systems provide us with multiple data streams for perception, but different data streams respond differently to environmental changes due to the influence of frequency differences, propagation paths, and system noise. Among these jagged data streams, how to select the ones with excellent environmental awareness is also an intractable problem.

To this end, we deeply analyse the characteristics of the derived CSI ratio from the perspective of the Mobius transform and find that its real and imaginary parts are sine-like and shape-complementary. Thus, we construct a novel indicator for motion sensing using both the real and imaginary parts of the CSI ratio, which can significantly improve the duration accuracy while ensuring the detection accuracy under short-window conditions.

Furthermore, we conduct proof-of-concept experiments and find that the nice data streams usually show appropriate environmental sensitivity and high shape similarity. Based on these observations, we propose a universal data stream screening algorithm driven by responsivity and similarity, which utilizes the sum of variance for rough selection and the curvature distance for fine selection. The proposed algorithm provides more accurate information for perception and fundamentally guarantees the motion detection accuracy.

Meanwhile, we prototype a passive human motion sensing system based on WiFi named MoSeFi. We deploy the system in two typical scenarios, and the results show that MoSeFi can achieve good performance in both detection accuracy and duration accuracy.

The rest of this paper is arranged according to the following structure. In Section 2, we review the related works on wireless sensing, and some preliminary analysis of CSI is introduced in Section 3. Then, we provide the detailed system design in Section 4, and the evaluation of the system performance is presented in Section 5. Finally, we summarize our work in Section 6.

2. Related Works

Driven by the needs of system applications, researchers have conducted extensive research on human motion sensing based on video, sound waves, and wearable devices. However, these device-dependent methods either require additional equipment such as cameras or destroy the normal state of the subject. In recent years, with the widespread deployment of WLAN equipment, WiFi-based passive human motion sensing has gradually become a research hotspot. According to the signal used, previous works can be divided into RSSI-based and CSI-based studies.

2.1. RSSI-Based Studies. Due to the advantages of universality and accessibility, RSSI was widely used in early wireless sensing works. RASID [13] utilized the variance of RSSIs to capture human motion, which operates in a short offline phase and a monitoring phase. Siebert et al. [14] provided a human motion detection and classification system based on the random forest algorithm. Time differencing transformation and empirical mode decomposition (EMD) are adopted in the system, and four statistical variables, including variance, mean, skewness, and kurtosis, are selected as the features. In another work [15], an occupancy estimation system using WiFi power measurements was proposed. Based on developing a simple motion model that characterizes the impact of blocking the LOS and scattering effects, the number of occupants was estimated using Kullback-Leibler divergence. WLID [16] expanded the detection area of human presence to the whole-home level by integrating WiFi-enabled Internet of Things (IoT) devices, and the system reached a 98% true-positive rate and a 3.8% false-positive rate by establishing a nonparametric algorithm. Furthermore, WiDet [17] captured human walking events by utilizing a deep convolutional neural network (CNN), where the wavelet coefficients and the raw RSSI signal were used as the input to the CNN. In general, the RSSI describes coarse-grained environmental variations, and the small-scale fading caused by the multipath effect limits the stability and reliability of motion sensing based on the RSSI.

2.2. CSI-Based Studies. Compared with the RSSI, CSI provides much more fine-grained information, such as the amplitude and phase of multiple subcarriers; therefore, an increasing number of CSI-based human motion detection studies have emerged since the CSI Tool [18] was released. FIMD [19] first leveraged the amplitude of CSI for distinguishing human motion, which slightly outperformed the RSSI-based system in [13]. To use the noisy phase of CSI, PADS [20] employed a linear transformation on the raw phase and computed the eigenvalues of the covariance matrix of both amplitude and phase sequences; furthermore, a threshold-based SVM classification was used for human motion detection. R-TTWD [21] considered the special case of through-the-wall detection of moving humans. The system took advantage of the correlated changes over different subcarriers and extracted the first-order difference of the eigenvector across different subcarriers for human detection. AR-Alarm [22] presented a real-time human intrusion detection system using the phase difference between different antennas. To achieve the purpose of environmental self-adaptation, the system utilized the normalized standard deviation as the motion detection feature and adopted a real-time static profile update scheme. WiSH [23] integrated correlations in both the time and frequency domains as a novel motion indicator and achieved a detection accuracy greater than 98% when a low sampling rate was used. Yang et al. [24] provided a device-free alarm system employing the CSI signal to identify human motion. In the active alarm module, specific help-seeking action combinations were captured through variance and Fourier transform, while in the passive alarm module, the foreground detection algorithm

was used to distinguish dangerous actions. WiMonitor [25] presented a room-level human vitality monitoring method. By eliminating auto gain control (AGC) noise, WiMonitor obtained a more stable sensing boundary parameter. Moreover, the system extracted the Doppler frequency shift (DFS) from the more robust CSI ratio signal and further constructed an activity intensity coefficient to distinguish between silence and different human activities.

Existing work has greatly promoted the development of passive wireless sensing technology based on WiFi. On this basis, this paper further proposes a subcarrier screening method based on environmental sensitivity and shape similarity and verifies the advantages of real and imaginary parts of the CSI ratio in improving the accuracy of motion duration detection.

3. Preliminaries

3.1. About CSI: Overview.
CSI describes the propagation characteristics of wireless signals in the form of the channel frequency response (CFR). In indoor scenarios, the signal received is the superposition of signals from different paths due to the multipath effect; thus, the CSI of the wireless system can be expressed as

$$H(f, t) = \sum_{k=1}^{K} A_k(t) \exp\left(-j2\pi f \frac{L_k(t)}{\lambda}\right), \qquad (1)$$

where A_k is the amplitude attenuation, K is the number of total propagation paths, L_k is the propagation length of the kth path, and f and λ represent the subcarrier center frequency and wavelength, respectively. A change in the environment, such as moving the transmitter or receiver, replacing the surrounding facilities, or walking along the propagation paths, brings a corresponding change in CSI, which makes it possible to use CSI for environmental sensing.

According to prior work [26], the wireless propagation paths can be divided into two parts: static paths and dynamic paths. The signals from the static path are coherent with each other, which can be regarded as a constant, and the dynamic component is incoherent with the static component, which causes the fluctuation of the CSI. Thus, we can rewrite the CSI as

$$H(f, t) = H_s(f, t) + \sum_{k=1}^{K_d} A_{kd}(t) \exp\left(-j2\pi f \frac{L_{kd}(t)}{\lambda}\right), \quad (2)$$

where $H_s(f, t)$ denotes the static component, K_d denotes the total dynamic reflection path, and A_{kd} and L_{kd} are the amplitude attenuation and propagation length of the k_d^{th} dynamic path, respectively. However, due to the imperfect hardware of the WiFi system, there are two major types of noise in the collected CSI data: amplitude noise, which is caused by the power amplifier uncertainly, and phase noise, which includes the packet detection delay (PDD), sampling fre-

quency offset (SFO), and carrier frequency offset (CFO) [27]. Thus, the polluted CSI can be expressed as

$$\tilde{H}(f, t) = \Delta_{\text{agc}} \exp\left(-j\theta_e\right) H(f, t), \qquad (3)$$

where Δ_{agc} is the amplifier noise and θ_e is the total phase offset.

Figure 1(a) shows the trajectory of the raw CSI on the complex plane. It can be seen that the CSI is distributed on a circle in the static state, while it is distributed on a ring in dynamic state. Due to the interference of phase noise, the phase is random in both two states; therefore, only its amplitude can be directly used for perception.

3.2. From CSI to the CSI Ratio.
In modern wireless networks, orthogonal frequency division multiplexing (OFDM) and multiple-input multiple-output (MIMO) technologies are used to improve the channel capacity and communication quality. Hence, the CSI evolves into a three-dimensional matrix, which characterizes the wireless channel variation simultaneously in the temporal, frequency, and spatial domains. Thus, we can define the CSI ratio as the quotient of raw CSI between two antennas [28], which can be expressed as

$$\text{CSI}_{\text{ratio}} = \frac{\Delta_{\text{agc}}^{(1)}\left(H_s^{(1)}(f, t)\right) + \sum_{k=1}^{K_d^{(1)}} A_{kd}^{(1)}(t) \exp\left(-j2\pi f\left(L_{kd}^{(1)}(t)/\lambda\right)\right)}{\Delta_{\text{agc}}^{(2)}\left(H_s^{(2)}(f, t)\right) + \sum_{k=1}^{K_d^{(2)}} A_{kd}^{(2)}(t) \exp\left(-j2\pi f\left(L_{kd}^{(2)}(t)/\lambda\right)\right)}, \qquad (4)$$

where the superscripts (1) and (2) are used to distinguish the parameters of the two antennas.

As different antennas share the same RF chain and clock in a commercial WiFi card, the power amplifier noise Δ_{agc} and the random phase offset θ_e are almost identical between different antennas [29]. Therefore, the calculation process of the CSI ratio can naturally eliminate amplitude and phase noise. Without loss of generality, we assume that there is only one dominating dynamic path corresponding to the moving object; then, the CSI ratio derived from Equation (4) can be written as

$$\text{CSI}_{\text{ratio}} = \frac{H_s^{(1)}(f, t) + A_d^{(1)} \exp\left(-j2\pi f\left(L_d^{(1)}/\lambda\right)\right)}{H_s^{(2)}(f, t) + A_d^{(2)} \exp\left(j2\pi f(\Delta L_d/\lambda)\right) \exp\left(-j2\pi f\left(L_d^{(1)}/\lambda\right)\right)}, \qquad (5)$$

where $\Delta L_d = L_d^{(1)} - L_d^{(2)}$ is the dynamic path length difference of two antennas. If we further employ $z = \exp\left(-j2\pi f(L_d^{(1)}/\lambda)\right)$ to represent the unit circle on the complex plane, it can be seen that the CSI ratio is the Mobius transformation of z, which includes the mapping of scaling, rotation, complex inversion, and translation. Since Mobius transformation is a conformal mapping, the trajectory of the CSI ratio during motion is still a circle on the complex plane.

As shown in Figure 1(b), the CSI ratio concentrates at a point in the static state, and its trajectory is a circle in the
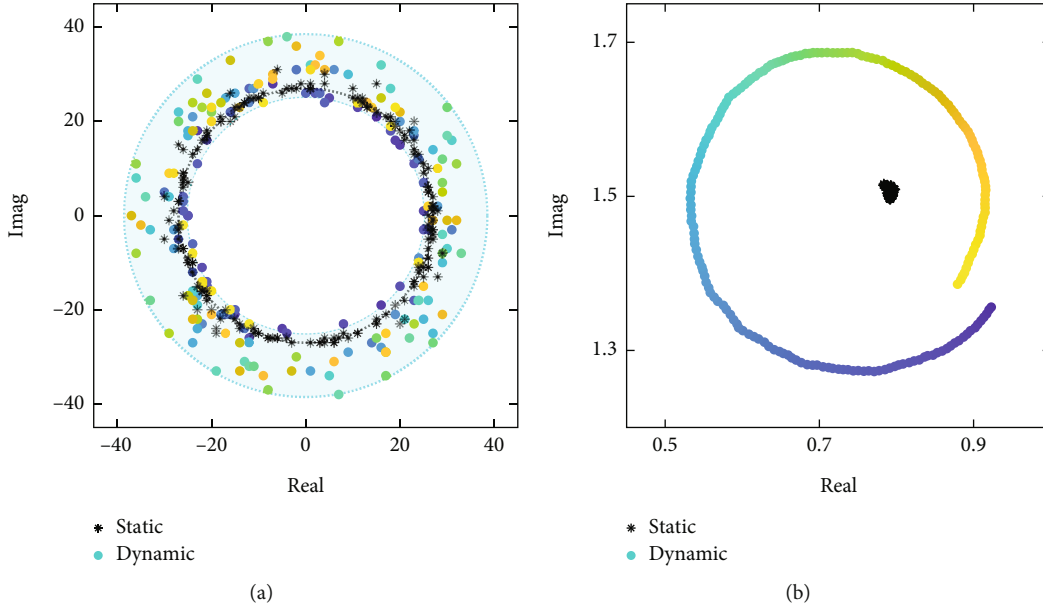
(a)

(b)

FIGURE 1: Trajectory on the complex plane. (a) CSI. (b) CSI ratio.

dynamic state. Since the impacts of amplitude and phase noise are eliminated, in addition to the amplitude, the phase of CSI ratio can also be used for sensing, which verifies the superiority of CSI ratio.

3.3. Key Observations of Motion Detection Using CSI Ratio. In this subsection, we conduct proof-of-concept experiments to further illustrate some key issues of motion sensing based on WiFi. We separate the transmitter and the receiver by 1.5 m at equal height and move a box with a size of $0.4 \times 0.3 \times 0.6$ m along the perpendicular bisector of the LOS; at the same time, the CSI data are collected.

3.3.1. Contradiction between Motion Integrity and Duration Accuracy. Figure 2(a) shows the amplitude of the CSI ratio, and it can be seen that there is a complete continuous motion interval. As mentioned above, the amplitude fluctuates differently in static and dynamic states. Previous works usually quantify this feature by variance or correlation coefficient and select the values of static state as a threshold. Since the above two parameters are calculated in a certain time window, next, we take the variance as an example to illustrate the impact of window size on detection result.

As shown in Figure 2(b), when the window is long (2 s), a complete motion interval with significantly large variance can be detected; however, the detected motion duration is significantly longer than the true value. It is reasonable because for the data points in the static interval that are close to the head or tail of the dynamic interval, their variance will be larger due to the influence of the dynamic data points contained in the window. These data points will be misjudged as dynamic state, making the detected motion duration longer than the true value. The longer the window, the more data points are affected around the junction.

Appropriately reducing the window length would alleviate the problem above, as Figure 2(c) shows. It can be seen

that the minimum variance of the dynamic state is just greater than the maximum variance of the static state, which allows us to obtain a complete motion interval that is closer to the true duration. Such a window seems like a suitable window.

If the window is shortened even further, such as 0.3 s, we found that the original continuous motion interval was divided into multiple short-term motion intervals, as shown in Figure 2(d). Here, we refer to these short-term motion intervals as motion fragments. The reason is that for a sinusoidal-like time series, the fluctuations at the position of the peak and trough in a small window are much smaller than those of the linear parts, which are comparable to the slight fluctuations in the static state. Meanwhile, the short window reduces the influence of the surrounding data points, and the boundary between static and dynamic states becomes clearer. If all these fragments are merged correctly, we will get a motion interval with a duration closer to the true value. However, the merging of fragments is not easy, especially when there are too many fragments; the false and missed merging not only reduces the accuracy of motion detection but also increases the error of motion duration.

From the foregoing discussion, we find that there is a contradiction between the motion integrity and the duration accuracy. From the perspective of motion integrity, a long window should be used, while for duration accuracy, a short window is better. Therefore, a compromise window size is usually selected in practice. Although the analysis above is conducted on the amplitude of the CSI ratio, the conclusions are also applicable to phase since they are all sine-like in shape.

3.3.2. Diversity of Subcarrier Response. For a MIMO-OFDM WiFi system, there are multiple data flows; however, the responses of different data flows are diverse due to the differences in frequency, propagation path, and noise level.
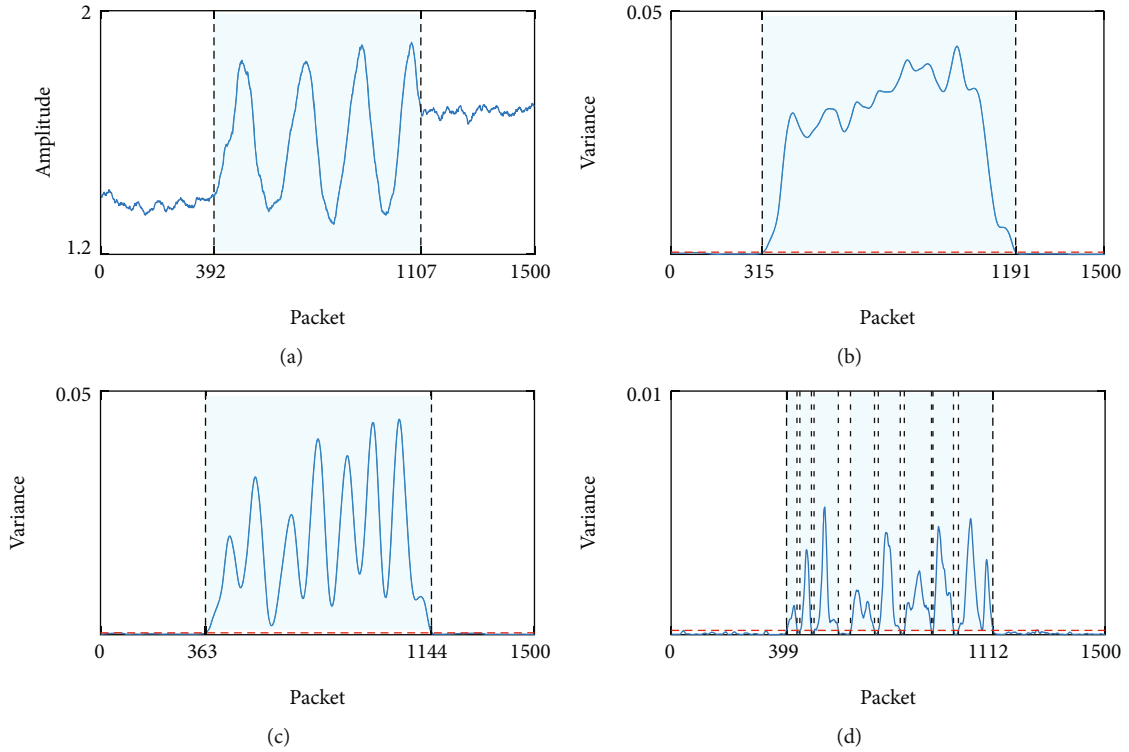
Figure 2: Detection results under different window sizes. (a) Amplitude. (b) Window size is 2 s. (c) Window size is 1 s. (d) Window size is 0.3 s.
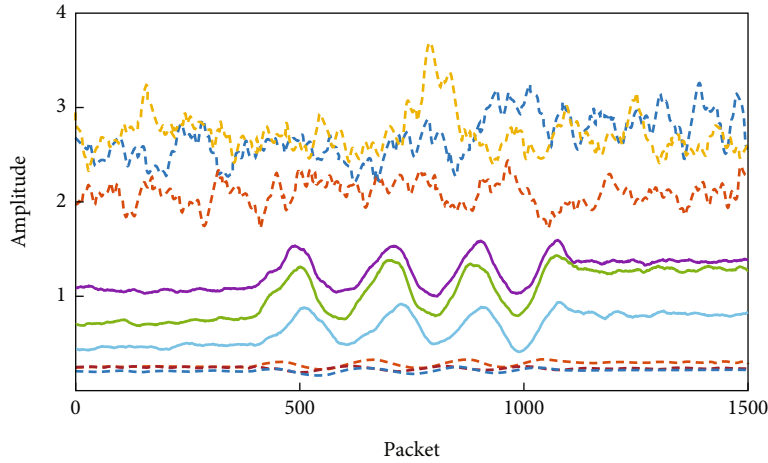

Figure 3: Diversity of subcarriers.

Figure 3 shows nine typical amplitude series of CSI ratio, and it can be seen that the data flows can be divided into three categories based on the noise level and response intensity.

(1) Noisy Subcarriers. These subcarriers contain a large amount of noise, such as those series plotted in dotted dashed lines. The fluctuations caused by motion are submerged in the noise and almost impossible to distinguish; thus, such subcarriers are shape-independent and useless for motion detection.

(2) Insensitive Subcarriers. We plot these subcarriers in dotted lines, and it is clear to see that they are stable when there is no moving object, but the fluctuations caused by motion are also very small. They are unresponsive to environmental changes regardless of whether they come from noise or human motion.

(3) Nice Subcarriers. As the solid line shows, these subcarriers show clear differences between two states, which remain stable in the static state but react significantly to the dynamic state. Moreover, they show strong correlation

and exhibit similar shapes during the motion period since they usually present approximate physical reality. Compared with the first two categories, these subcarriers are more suitable for motion detection.

From the analysis above, we find that avoiding the interference of poor subcarriers is necessary for motion detection, which is also the basis for determining the quality of the subsequent processing.

Via the case study, we find that although CSI ratio shows better attributes than CSI, there are still two key issues that need to be further solved for accurate and robust motion sensing:

(a) How to perceive the motion on the premise of both motion integrity and duration accuracy

(b) How to select the nice subcarriers with better environmental awareness

Aiming at the first problem, we construct a novel motion indicator by using both the real and imaginary parts of the CSI ratio; furthermore, we propose a subcarrier screening mechanism based on environmental sensitivity and shape similarity.

4. System Design

4.1. Data Preprocessing. Due to environmental noise and incomplete hardware, the collected CSI data cannot be directly used for motion detection. To solve the problem of data loss caused by sampling jitter, we conduct interpolation to obtain a uniform time interval. Then, the Savitzky-Golay filter is chosen to further smooth the CSI data.

4.2. Motion Detection Using New Feature. As described in Section 3, the CSI ratio of the dynamic state can be regarded as the Mobius transformation of the unit circle, and its trajectory is a circular on the complex plane when the dynamic path changes one wavelength. Thus, the real and imaginary parts of CSI ratio change periodically. Meanwhile, the CSI ratio centralizes on a point in the static state, which means that its real and imaginary parts are both stable at this time. Therefore, both parts can be used for motion sensing like amplitude and phase. At the same time, since the trajectory of CSI ratio on the complex plane is a circle, the extreme points of the real part appear in the direction of the real axis, and the extreme points of the imaginary part appear in the direction of the imaginary axis, and the extreme points of the two are staggered.

In Figure 4, we show the schematic illustration of Mobius transformation, and it can be seen that the real and imaginary parts of the CSI ratio are not standard sine curves due to the scaling, translation, and complex inversion operations in the transformation. The degree of shape deviation depends on the coefficients in the transformation. When $\rho = |H_s^{(2)}(f, t)/A_d^{(2)}|$ is far from 1, the real and imaginary parts are both sine-like. The condition above is easy to satisfy in practice, especially when the LOS exists. At the same time, since the trajectory of CSI ratio is a circle, the

extreme points of the real part appear in the direction of the real axis, and the extreme points of the imaginary part appear in the direction of the imaginary axis, and the extreme points of the two are staggered.

Figure 5(a) shows the real and imaginary parts of the CSI ratio collected in real scenario; it can be seen that their shape features are consistent with the previous analysis. Similar to magnitude and phase, when we use a short window to calculate the variance and select the maximum variance of static state as the threshold, there are multiple motion fragments in the raw detection results, as shown in Figures 5(b) and 5(c).

Recall that the root reason of motion fragments under short-window conditions is the relatively small fluctuation range at the extreme points. Since the real and imaginary curves of CSI ratio are complementary in shape, this inspires us to use both of them for motion sensing and construct a new motion indicator as

$$M_{\text{indi}} = \text{Var}_{\text{real}} + \text{Var}_{\text{imag}}, \tag{6}$$

where Var_{real} and Var_{imag} are the variance of the real part and imaginary part, respectively. It is clearly that the value of M_{indi} is small in the static state because both real and imaginary parts are stable. For the dynamic state, if there is a small value of Var_{real}, there must be a large value of Var_{imag}, and vice versa. Thus, M_{indi} is always greater than that of the static state during the whole motion period.

Therefore, we still use the largest M_{indi} of the static state as a benchmark, which is denoted as $\text{Th}_{M\_\text{indi}}$, and determine the data points whose M_{indi} is larger than $\text{Th}_{M\_\text{indi}}$ as dynamic. In order to initialize $\text{Th}_{M\_\text{indi}}$, we keep the environment static at the beginning of data collection and pick the largest M_{indi} of this static interval as the initial $\text{Th}_{M\_\text{indi}}$. As Figure 5(d) shows, a complete and continuous motion interval can be obtained under short-window conditions. Meanwhile, the shorter window reduces the number of affected points close to the actual motion interval, and the motion duration detected is closer to the true value.

Due to the complexity of human movement and environmental noise, there may still be a small amount of residual moving debris in the raw detection results, which is located in a complete motion interval or a stationary interval.

(1) The residual fragments within the motion interval are generally caused by the transient immobility of the human body in a complete motion, so the time interval between the adjacent motion fragments is relatively short

To this end, we define the threshold of static duration Th_s which characterizes the minimum time interval between two independent motions. When the time interval between two motion fragments is less than Th_s, the two motion fragments are regarded as being located in a complete motion interval and merged. Conversely, the two motion fragments are both treated as independent.
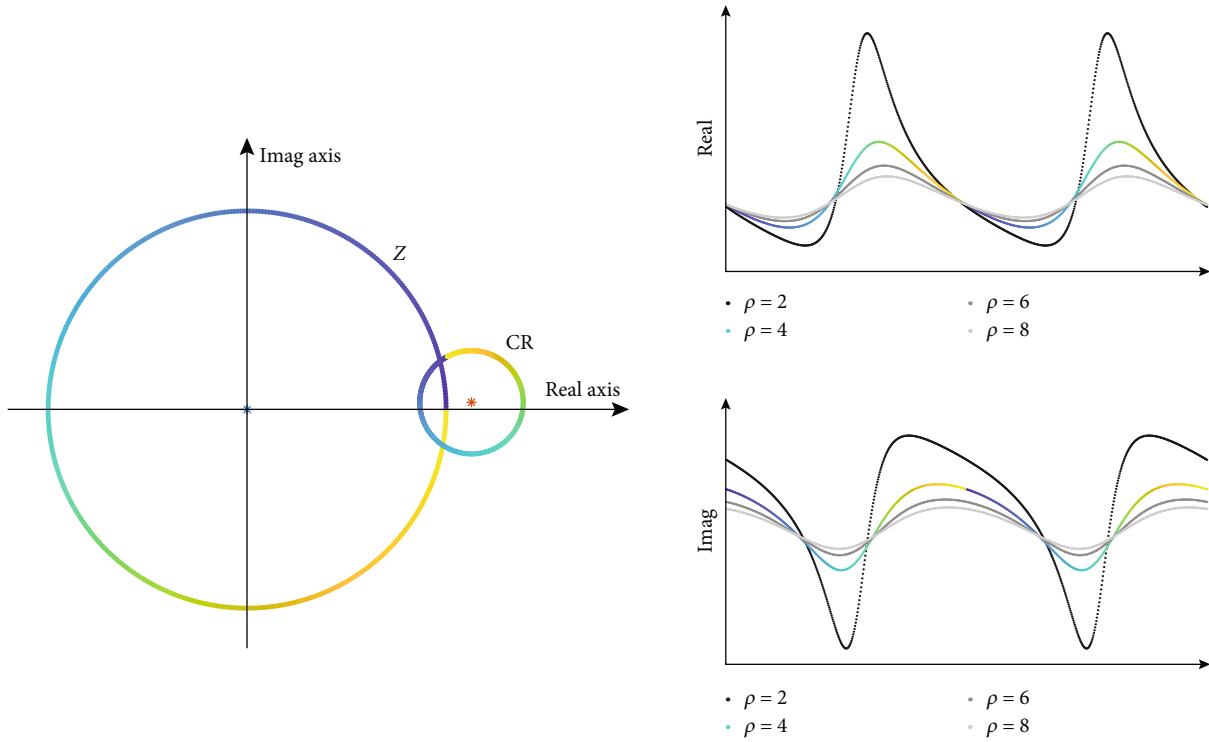
Figure 4: Schematic illustration of the Mobius transformation.
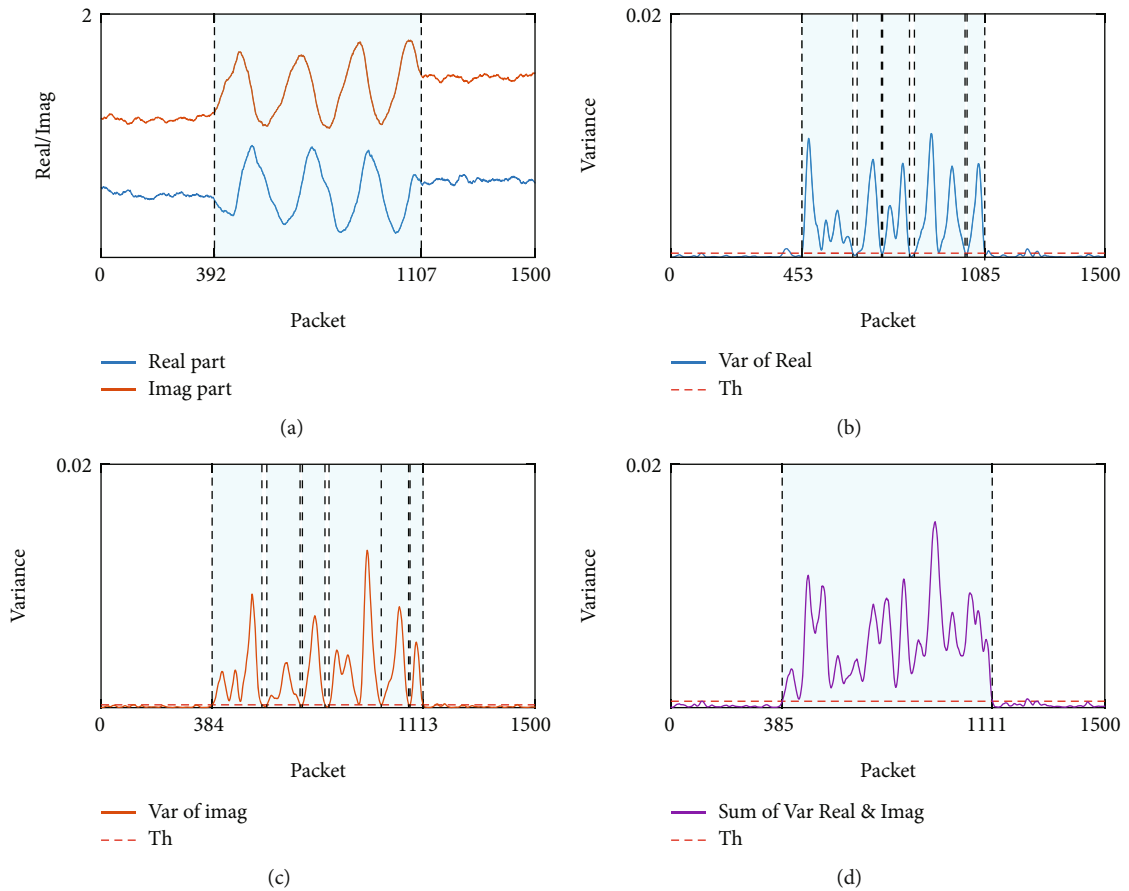

(a)

(b)

(c)

(d)

Figure 5: Motion sensing using different parts of CSI ratio. (a) Real and imaginary parts of the CSI ratio. (b) Detection result using the real part. (c) Detection result using the imaginary part. (c) Detection result using the real and imaginary parts.

(2) The pseudomotion fragments in the stationary region are generally caused by noise, either the duration is short or the fluctuation is small. Thus, we utilize two parameters to eliminate them: the threshold of motion duration $\mathrm{Th_m}$ and the threshold of average variance $\mathrm{Th_{av}}$

The $\mathrm{Th_m}$ is defined as the smallest possible duration of a true motion. When the duration of a motion fragment is less than $\mathrm{Th_m}$, it is discarded. In this way, the pseudo-short-duration motion fragments are eliminated.

The $\mathrm{Th_{av}}$ is defined as α times $\mathrm{Th}_{M\_\mathrm{indi}}$, which can be expressed as

$$\mathrm{Th_{av}} = \alpha \cdot \mathrm{Th}_{M\_\mathrm{indi}}. \tag{7}$$

For an independent motion segment, we calculate its average variances and then compare it with $\mathrm{Th_{av}}$. If the average value is greater than $\mathrm{Th_{av}}$, this motion segment is determined to be valid and retained; otherwise, the motion segment is considered to be caused by environment noise and discarded.

In the actual detection process, the above three parameters are fused together to remove the impact of motion fragments. Firstly, we calculate all the time intervals between adjacent motion fragments and merge the two fragments with time interval less than $\mathrm{Th_s}$. Then, we repeat the above operation until the intervals between all adjacent motion fragments are greater than $\mathrm{Th_s}$. For these new independent motion segments, we further utilize $\mathrm{Th_m}$ and $\mathrm{Th_{av}}$ to judge them one by one and remove the segments whose duration is less than $\mathrm{Th_m}$ or the average variance is smaller than $\mathrm{Th_{av}}$. At last, the remaining independent motion segments with long duration and large variance are picked as the final detection result.

4.3. Nice Subcarrier Screening.
According to the characteristics of the excellent subcarrier, we divide the subcarrier screening process into two steps: coarse selection driven by responsiveness and fine selection driven by similarity.

4.3.1. Coarse Selection.
From the discussion in Section 3, we know that the poor subcarriers are either disorganized or unresponsive; moreover, a nice subcarrier remains relatively stable in the static state and generates an appropriate response to the motion. Thus, we select the subcarriers approximately according to their volatility intensity. Here, the sum of variance (SV) is used to describe the volatility of the time series, which can be written as

$$\mathrm{SV}^{(k)} = \sum_{i=1}^{P} \mathrm{Var}_i^{(k)}, \tag{8}$$

where $\mathrm{Var}_i^{(k)}$ is the variance of the ith packet in the kth subcarrier and P is the length of the CSI series.

Figure 6(a) shows the SVs in gradient colors, and it can be seen that the SV values between subcarriers are significantly different. Here, we firstly sorted the SVs and divided the subcarriers into three groups equally according to values

of SV. The split benchmarks are shown as the dashed lines in Figure 6(a). Since the SV describes the magnitude of fluctuation, each group shows different responsiveness. Specifically, the first group contains the subcarriers with the 30 largest SV values, which are usually noisy due to being hyperreactive to environmental changes. The second group contains the subcarriers with the 30 smallest SV values, which are usually insensitive to environmental changes. Obviously, the above two groups are not suitable for motion detection, so they are discarded. Finally, the 30 subcarriers with the middle SV values are reserved as the result of coarse selection. These subcarriers usually keep stable in static state but fluctuate strongly enough in dynamic state. Figure 6(b) shows the coarse-selected subcarriers, and it can be seen that most of them meet the characteristics of the nice subcarrier.

4.3.2. Fine Selection.
Recall that the nice subcarriers appear in clusters and show a great deal of similarity in shape, which inspires us to design a similarity-driven selection mechanism to further pick the appropriate subcarriers.

Curvature is one of the important mathematical tools to describe the geometry of curves and is widely used in many shape processing applications [30]. Here, we utilize k-cosine functions [31] to calculate the curvature. Let

$$\begin{aligned} \mathrm{s}_{ik} &= (x_i - x_{i+k}, y_i - y_{i+k}), \\ \mathrm{t}_{ik} &= (x_i - x_{i-k}, y_i - y_{i-k}) \end{aligned} \tag{9}$$

be the k-vectors of point $p_i = (x_i, y_i)$ on the curve; the k-cosine is defined as

$$\cos p_i^k = \frac{\mathrm{s}_{ik} \cdot \mathrm{t}_{ik}}{|\mathrm{s}_{ik}||\mathrm{t}_{ik}|}. \tag{10}$$

Then, the curvature of point $p_i = (x_i, y_i)$ can be expressed as

$$c_i = \frac{p_i^k}{|\mathrm{s}_{ik}| + |\mathrm{t}_{ik}|}. \tag{11}$$

For any time series X and X', we can construct the corresponding curvature series $C = [c_1^k, c_2^k, \cdots, c_i^k, \cdots]$ and $C' = [c'^k_1, c'^k_2, \cdots, c'^k_i, \cdots]$. Since the curvature describes the shape of time series, we can leverage the curvature distance (CD) to measure the shape similarity. The curvature distance is defined as

$$D_c\left(X, X'\right) = \frac{\sum_{i=1}^{P} \left| c_i^k - c'^k_i \right|}{P - 1}, \tag{12}$$

where P is the length of the time series. The shorter the curvature distance is, the more similar the shape of the two time series.

Figure 7(a) shows the CDs of the coarse screened real parts, and it can be seen that the CDs between different subcarriers are distinct since they experience diverse variations. If we pick M groups with the smallest CD values, the shapes
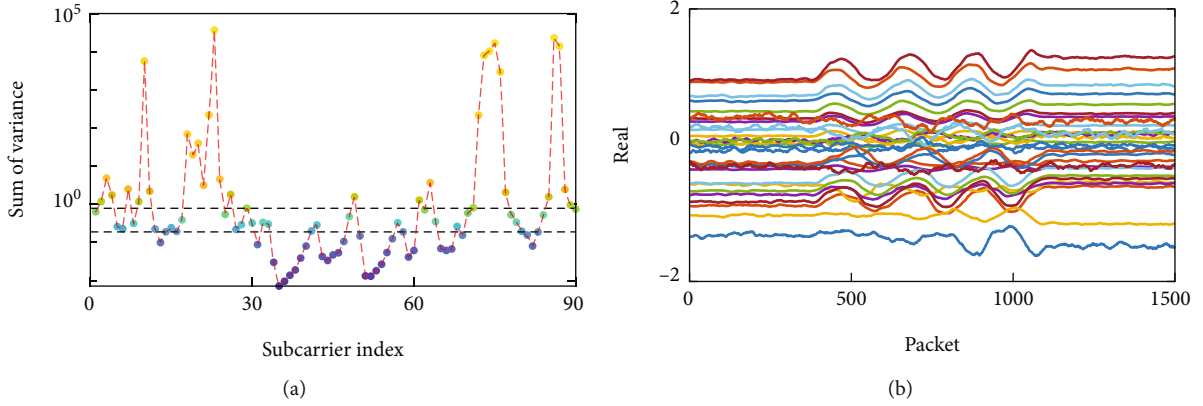
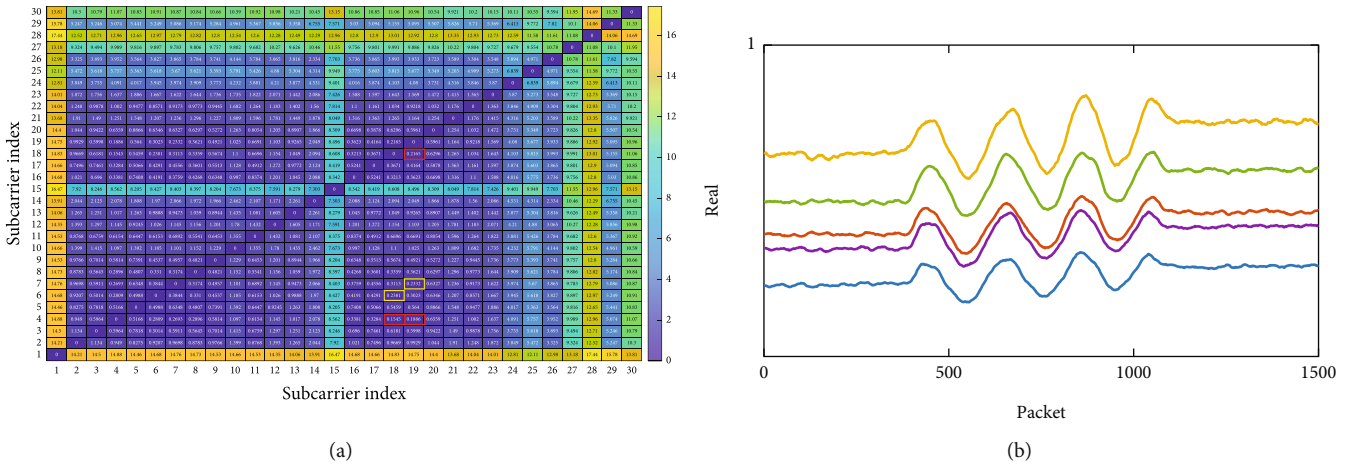FIGURE 6: Subcarrier coarse selection. (a) SVs of different subcarriers. (b) Coarse selected subcarriers.

FIGURE 7: Subcarrier fine selection. (a) Curvature distance. (b) Fine selected subcarriers.

of corresponding real parts are similar. It is worth noting that M should not be too large; otherwise, some subcarriers with poor shape similarity may be introduced. However, a small M sometimes results in too few subcarriers selected due to the overlap of subcarrier indexes, which is not conducive to the subsequent selection of imaginary parts. Therefore, we additionally specify the minimum number of selected real parts denotes as $\mathrm{Num}R_{\min}$, and once the number of subcarriers contained in the M groups is less than $\mathrm{Num}R_{\min}$, we continue to pick the group with the smallest CD value among the remaining ones until no less than $\mathrm{Num}R_{\min}$ subcarriers are selected.

Here, we take $M = 3$ and $\mathrm{Num}R_{\min} = 5$ as an example to illustrate the specific screening process. In Figure 7(a), we mark the three positions with the smallest CD values in red boxes, and it can be seen that only three subcarriers (#4, #18, and #19) are selected due to the overlapping of indexes. Thus, we successively pick two groups with the smallest CD values (yellow boxes) from the remaining subcarriers. Finally, the five subcarriers (#4, #6, #7, #18, and #19) are chosen as the result of fine selection. The screened real parts are shown in Figure 7(b) accordingly, and it can be seen that the shapes of them are exactly similar. Since

the curvature distance concentrates on the difference of shape, the nice subcarriers can still be picked out even if they lie far away in terms of Euclidean distance.

Since we use both real and imaginary parts of CSI ratio for motion detection, to ensure that both parts meet the requirements, we firstly utilize SV to roughly select the real part and obtain R_{coarse}; then, we pick M groups of real part series with the closest curvature distance to form R_{fine}. If R_{fine} contains less than $\mathrm{Num}R_{\min}$ subcarriers, we continue to pick the group with the smallest CD value among the remaining R_{coarse} until more than $\mathrm{Num}R_{\min}$ subcarriers are selected. Next, we select the corresponding imaginary parts I_{fine} according to the subcarrier indexes of R_{fine} and choose two pieces of imaginary part series with the smallest curvature distance in I_{fine}, which are denoted as I_{sel1} and I_{sel2}, respectively. According to the subcarrier indexes of I_{sel1} and I_{sel2}, we can find the corresponding two real part series R_{sel1} and R_{sel2}. The pseudocode of this module is shown in Algorithm 1.

At last, we obtained two subcarriers with similar real and imaginary parts. Since only one subcarrier is needed in the calculation of M_{indi}, we further calculate the average of the

---

Input: K: number of subcarriers; P: number of CSI packets; $R \in \mathbb{R}^{P \times K}$: real
parts of CSI ratio; $I \in \mathbb{R}^{P \times K}$: imaginary parts of CSI ratio;
Output: R_{sel}, I_{sel}
1: \\coarse selection of real part
2: **for** $k = 1$ to K **do**
3:          $SV[k] \longleftarrow \sum_{p=1}^{P} \text{movvar}(R[p, k])$
4: **end for**
5: $R_{coarse} \longleftarrow$ 30 pieces of R with middle SV
6: \\fine selection of real part
7: $CR \longleftarrow$ discrete curvature of R_{coarse}
8: **for** $i = 1$ to $30 - 1$ **do**
9:          **for** $j = 1 + 1$ to L **do**
10:                    $D_{CR}(i, j) \longleftarrow \sum_{p=1}^{P} |CR[p, i] - CR[p, j]|/(P - 1)$
11:          **end for**
12: **end for**
13: $R_{fine} \longleftarrow$ M groups of R with the shorts D_{CR} in R_{coarse}
14: $Count \longleftarrow$ number of subcarriers in R_{fine}
15: **while** $Count < \text{Num}R_{\min}$ **do**
16:          Add the group of R with the shortest D_{CR} in remained R_{coarse} to R_{fine}
17:          $Count \longleftarrow$ number of subcarriers in R_{fine}
18: **end while**
19: \\fine selection of imaginary part
20: $\text{ind}_{R-fine} \longleftarrow$ subcarrier indexes of R_{fine}
21: $I_{fine} \longleftarrow I[\text{ind}_{R-fine}]$
22: $CI \longleftarrow$ discrete curvature of I_{fine}
23: **for** $m = 1$ to $length(\text{ind}_{R-fine}) - 1$ **do**
24:          **for** $n = m+1$ to $length(\text{ind}_{R-fine})$ **do**
25:                    $D_{CI}(m, n) \longleftarrow \text{sum}_{p=1}^{P} |CI[p, m] - CI[p, n]|/(P - 1)$
26:          **end for**
27: **end for**
28: $[I_{sel1}, I_{sel2}] \longleftarrow$ two pieces of I with the shortest D_{CI} in I_{fine}
29: $[\text{ind}_{I-sel1}, \text{ind}_{I-sel2}] \longleftarrow$ subcarrier indexes of $[I_{sel1}, I_{sel2}]$
30: $[R_{sel1}, R_{sel1}] \longleftarrow R[\text{ind}_{I-sel1}, \text{ind}_{I-sel2}]$

ALGORITHM 1: Subcarrier screening module.

real and imaginary parts of the two subcarriers, respectively, which can be expressed as

$$R_{sel} = \frac{R_{sel1} + R_{sel2}}{2},$$
$$I_{sel} = \frac{I_{sel1} + I_{sel2}}{2}. \tag{13}$$

Here, we take R_{sel} as an example to illustrate the rationality of above operations. Since R_{sel1} and R_{sel2} are of equal length, we can regard R_{sel2} as the translation of R_{sel1} and denote the translation item as $\Delta R = R_{sel1} - R_{sel2}$. Thus, the variance of R_{sel} can be expressed as

$$\text{Var}(R_{sel}) = \frac{\text{Var}(R_{sel1})}{2} + \frac{\text{Var}(R_{sel2})}{2} - \frac{\text{Var}(\Delta R)}{4}. \tag{14}$$

Since R_{sel1} and R_{sel2} are similar in shape, we have the following: (1) the shape of R_{sel} is close to R_{sel1} and R_{sel2}. (2) The elements of translation item are approximately equal; thus, $\text{Var}(\Delta R)$ is small, and $\text{Var}(R_{sel})$ is mainly composed of the first two items. Since $\text{Var}(R_{sel1})$ of dynamic state is larger than that of static state and so does $\text{Var}(R_{sel2})$, R_{sel} can be used for motion detection.

Similarly, we can infer that I_{sel}, which is similar in shape to I_{sel1} and I_{sel2}, can also be used for motion detection. Furthermore, it can be seen that R_{sel} and I_{sel} are complementary in shape because R_{sel1} and I_{sel1} and R_{sel2} and I_{sel2} are complementary in shape. Thus, we can utilize R_{sel} and I_{sel} to calculate M_{indi} for motion detection.

4.4. Subcarrier and Benchmark Update. For long-term motion perception, subcarrier screening with excessively long CSI data incurs vast computational overhead. Meanwhile, it is not safe to use the fixed subcarrier for a long time because the nice subcarrier is not invariable. Therefore, we propose a motion-enhanced subcarrier updating mechanism.

When the system detects new motion, we reperform the subcarrier screening module using the data containing the motion interval. Furthermore, if the environment maintains a static state for a long time, such as 15 minutes, the system also conducts subcarrier reselection to avoid nice subcarrier

drift. Meanwhile, the benchmark used to distinguish between static and dynamic states is also updated. Specifically, we recalculate the M_{indi} of new subcarriers in the static state and pick the maximum value as the new Th_{M_indi} for subsequent motion detection. Correspondingly, the Th_{av}, which is α times Th_{M_indi}, will also be updated.

5. Performance Evaluation

5.1. Experiments Setup.
For MoSeFi, we utilized a TP-Link WDR5600 wireless router with two antennas working as the transmitter in AP mode at 2.4 GHz and a Lenovo X200 laptop equipped with an Intel 5300 NIC and three omnidirectional antennas working as the receiver. The laptop ran Ubuntu 12.04 OS with the modified firmware CSI Tool installed. We set the height of both transmitter and receiver to be 1.3 meters, and the distance between them is 3 meters. The sampling rate was set as 100 packets/s. We collected the CSI data in two typical scenarios: (1) a student dormitory containing only one performer during data collection and (2) a graduate studio containing 5 resident students and some random visitors. The floor plans of the two scenarios are shown in Figure 8. In order to obtain the ground truth of human motion duration, we used a camera to record the data collection process.

We adopted the false-positive rate (FPR) and false-negative rate (FNR) to characterize the motion detection accuracy, which can be depicted as follows:

$$
\begin{aligned}
\text{FPR} &= \frac{\text{False} \quad \text{positives}}{\text{True} \quad \text{positives} + \text{false} \quad \text{positives}}, \\
\text{FNR} &= \frac{\text{False} \quad \text{negatives}}{\text{True} \quad \text{positives} + \text{false} \quad \text{negatives}}.
\end{aligned}
\tag{15}
$$

For a motion detection system, FPR describes the proportion of false detections of motion when the environment is static, while FNR describes the proportion of missed detections of true motion. Compared to accuracy that calculates the proportion of the sum of false and missed detections, FPR and FNR more clearly depict the ability of the system to correctly distinguish between static and dynamic states.

Meanwhile, we recorded the difference between the detected motion duration and the true motion duration for all files under true-positive (TP) conditions and took their average as the motion duration error (MDE). In practice, the raw detection results are packet indexes, which can be converted to second according to the timestamp of the CSI data. Meanwhile, the actual motion duration can be obtained by video recording.

In addition, we used *tic* and *toc* commands in Matlab to record the time consumption of each file and took their average value as the running time (RT). The RT measures the real-time performance of the system.

We selected two typical systems for comparison: AR-Alarm uses the variance of the phase difference to characterize changes in the environment; WiSH utilizes the correlation of CSI in both time and frequency domains. Since the subcarrier selection mechanism is not clearly stated in AR-Alarm, we traverse all subcarriers and choose the best performing one as its final performance. And for WiSH, we select all subcarriers to calculate the correlation matrix.

5.2. Case Study of Single Motion Series.
In this subsection, we collected the CSI data in scenario (a), and five volunteers were recruited to complete six different actions, including walking, sitting, kicking, waving, stepping, and squatting. The volunteers were first asked to remain still, then complete a specified action under voice commands, and finally remain still for a period of time. For each type of motion, we asked each volunteer to complete 30 repetitions in three different positions. Specifically, as Figure 8(a) shows, the volunteers completed the walking action according to the gray lines ($A_1 \longrightarrow A_1{}'$, $B_1 \longrightarrow B_1{}'$, and $C_1 \longrightarrow C_1{}'$), respectively, and did the remaining five actions in the positions A_1, B_1, and C_1, respectively. Finally, a total of 2700 sets of CSI data were collected. Obviously, this type of data is often used in the applications such as classification and recognition.

5.2.1. Overall Performance.
In order to verify the influence of the window size on the detection results, we tested the performance of the three systems with window sizes of 0.5 s and 1 s, respectively, and the experimental results are shown in Table 1.

It can be seen that when the window size is 1 s, the detection accuracy of MoSeFi is significantly higher than that of the other two systems. This emphasizes the importance of subcarrier screening. On the one hand, it is not safe to always use a fixed subcarrier because the drift of nice subcarriers may cause the system to make wrong judgments. On the other hand, rough use of all subcarriers is not an optimal choice, and the processing of poor subcarriers is redundant or even harmful. Meanwhile, we find that the MDE of WiSH is smaller than that of the other two systems when the window is long. However, the real-time performance of WiSH is poor due to the calculation of complex correlation coefficient, and its RT is far greater than the other two, since AR_Alarm only uses the threshold of motion duration to judge each motion fragment and does not address the issue of incorrectly merging fragments. Meanwhile, we only counted the MDE under the true-positive conditions; the MDE of AR_Alarm is smaller although its FPR and FNR are significantly larger than those of MoSeFi. The above results also show that the merging of motion fragments is risky, and the wrong merging may lead to the increase of MDE.

When the window reduces to 0.5 s, our system can still maintain a relatively stable performance, while the performance of the other two systems is seriously degraded, especially AR_Alarm. It is because AR_Alarm employing variance of phase difference suffers from severe motion fragmentation when the window is short. Due to lack of fragment merging mechanism, if there are multiple fragments with durations greater than Th_m, each fragment will be misjudged as an independent motion, and the FPR increases accordingly. On the contrary, if the durations of all fragments are less than Th_m, they will all be

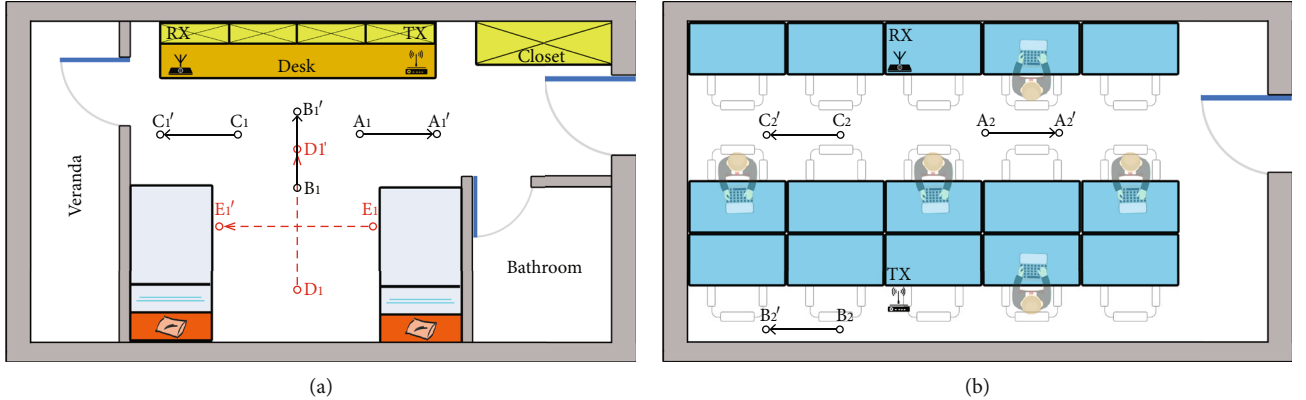(a)   (b)

FIGURE 8: Floor plan of experiment scenarios. (a) Student dormitory. (b) Graduate studio.

TABLE 1: System performance for single motion case.

System	Window (s)	FPR (%)	FNR (%)	MDE (s)	RT (s)
MoSeFi	1.0	0.33	0.89	1.22	0.14
	0.5	0.78	0.67	0.53	0.13
AR-Alarm	1.0	2.56	2.90	0.94	0.12
	0.5	12.42	10.46	0.76	0.11
WiSH	1.0	3.12	3.46	0.61	1.41
	0.5	4.49	8.06	0.66	1.33

discarded, resulting in the increase of FNR. By contrast, WiSH and MoSeFi, which adopt the fragment merging module, both outperform AR_Alarm under short-window conditions. Since the ingenious motion indicator of MoSeFi can significantly reduce the number of motion fragments, thereby reducing the probability of false merging and missed merging, its FPR and FNR are still low with short window. Benefiting from the shortening of the window and the limited increase of motion fragments, the MDE of MoSeFi declines significantly. On the contrary, severe motion fragmentation not only declines the motion detection accuracy but also offsets the benefits of window shortening in improving motion duration accuracy, resulting in a slight increase in MDE of WiSH.

5.2.2. Importance and Universality of Subcarrier Selection Module. In order to pick the nice subcarriers, we have carried out coarse selection and fine selection successively. To further illustrate the role of two steps, we tested the performance of MoSeFi when only coarse or fine selection was used. Here, we refer to the above two systems as MoSeFi$^{(C)}$ and MoSeFi$^{(F)}$, respectively. For MoSeFi$^{(C)}$, we directly calculated the mean of coarse selected 30 subcarriers, while for MoSeFi$^{(F)}$, we utilized the curvature distance to select two nice subcarriers among all subcarriers. Furthermore, aiming at verifying the universality of the proposed subcarrier selection method, we have also tested the performance of AR-Alarm and WiSH with the full subcarrier selection module, which are called AR-

Alarm$^{(CF)}$ and WiSH$^{(CF)}$, respectively. Table 2 shows the experimental results.

From Table 2, it can be seen that compared with MoSeFi using full subcarrier selection module, the detection accuracy of both MoSeFi$^{(C)}$ and MoSeFi$^{(F)}$ drops. As described in Section 4.3, there are still some insensitive or oversensitive subcarriers in the coarse selection results. Indiscriminately averaging these subcarriers reduces the discrimination between static and dynamic states, resulting in a large increase in the FPR and FNR of MoSeFi$^{(C)}$. On the other hand, when using the curvature distance to fine select across all subcarriers, the results may locate in the unresponsive or hyperreactive group, causing the variance of motion interval to be too small or the selected threshold to be too large. Thus, the performance of MoSeFi$^{(F)}$ deteriorated. The above results also show that the coarse selection and the fine selection play different roles, and one of them alone is not enough to always screen out the nice subcarriers; therefore, it is necessary to combine the above two steps in the actual human motion detection.

As for AR-Alarm$^{(CF)}$ and WiSH$^{(CF)}$, their motion detection accuracy is significantly improved in both window conditions, benefiting from the subcarriers that more accurately describe the environmental changes. At the same time, subcarrier screening reduces the data for calculating the correlation coefficient, so that the RT of WiSH$^{(CF)}$ considerably drops. The results further illustrate the influence of subcarrier quality on the motion detection

TABLE 2: System performance using different subcarrier selection modules.

System	Window (s)	FPR (%)	FNR (%)	MDE (s)	RT (s)
MoSeFi(C)	1.0	2.16	4.67	1.24	0.12
	0.5	3.75	6.02	0.66	0.12
MoSeFi(F)	1.0	1.34	1.60	1.20	0.14
	0.5	1.44	1.67	0.54	0.14
AR-Alarm(CF)	1.0	0.89	1.34	1.05	0.12
	0.5	3.50	1.86	0.65	0.12
WiSH(CF)	1.0	0.78	1.00	0.56	0.38
	0.5	2.38	4.54	0.49	0.31

accuracy and also show that our universal subcarrier screening method can combine with other works to improve system performance. However, we find that although the performance of AR_Alarm(CF) and WiSH(CF) has been greatly improved under short-window conditions compared with the native systems, their overall detection accuracy is still not high. This shows that subcarrier screening alone is not enough, and the impact of motion fragmentation cannot be ignored under short-window conditions.

5.2.3. Advantage of Proposed Motion Indicator. To further verify the superiority of using both the real and imaginary parts of CSI ratio for improving the duration accuracy, we tested the system performance with different window sizes. For the sake of fairness, we selected AR-Alarm(CF) and WiSH(CF) for comparison.

As shown in Figure 9, when the window is larger than 0.8 s, the motion detection accuracy of the three systems is high, but if the window continues to be shortened, their performance will decline to varying degrees. Take the window size of 0.3 s as an example. Benefiting from the complementarity between the real and imaginary parts of CSI ratio, the number of motion fragments in MoSeFi can be significantly reduced, and the FPR and FNR are only slightly increased to 1% and 0.89%, respectively. On the contrary, the serious fragmentation problem makes the FNR of AR_Alarm(CF) increase sharply to 7.58%. At the same time, the correlation coefficient calculated within a short window is difficult to accurately describe the environment state, resulting in a sharp increase in the FNR of WiSH(CF) to 9.62%. In addition, we found that the size of the window is not as long as possible. Although a long window improves the probability of detecting complete motion interval, it also blurs the boundaries between static and dynamic states, making the system performance worse.

Figure 10 shows the motion duration accuracy of different systems. It can be seen that the MDE of MoSeFi monotonically decreases as the window gets shorter. Since the new motion indicator effectively reduces the number of motion fragments, we can directly obtain a complete motion interval with a higher probability under short-window conditions. As discussed in Subsection 3.3.1, reducing the window size can weaken the influence of the surrounding data points, which makes the MDE of our system decrease as the window shortens. For AR-Alarm(CF) that also uses variance, when the

window length is shortened from 1 s to 0.6 s, its MDE is reduced from 1.06 s to 0.63 s. However, excessively shortening the window will result in severe motion fragmentation, which prevents the MDE of AR-Alarm(CF) from being further reduced. WiSH(CF) achieves the best duration accuracy when the window length is 0.7 s, and its MDE declines to 0.43 at the cost of increased FPR and FNR. When the window size is less than 0.7 s, its MDE gradually increases due to serious motion fragmentation.

5.2.4. Parameters Analysis

(1) Motion Type. Figure 11 shows the system performance of different types of motion. In general, MoSeFi is robust to human motion with different magnitudes and velocities. Specifically, the system achieved the highest motion detection accuracy and motion duration accuracy when multiple body parts move as a whole, such as sitting. Conversely, when multiple parts of the body move simultaneously but independently, the FPR and MDE of the system become larger, such as walking and stepping. Through further analysis, we find that the increase of the motion complexity makes the raw detection results more fragmented. For example, walking and stepping contain 40% and 37% more residual motion fragments than sitting, respectively. These more motion fragments increase the probability of false and missed merging, resulting in larger FPR and MDE. On the other hand, compared with the above large-scale motion, the CSI fluctuations caused by small-scale waving are more likely to be overwhelmed by noise, resulting in the highest FNR.

5.2.5. Threshold of Static/Motion Duration. As shown in Figures 12(a) and 12(b), both FPR and FNR decrease as Th_s increases. It is because for motion fragments located in a complete motion, if their durations are greater than Th_m, they are more likely to be merged into one as Th_s increases, rather than being misjudged as multiple independent motions, so FPR drops. Conversely, when the durations of these fragments are less than Th_m, they are more likely to be merged into one as Th_s increases, rather than being removed, so FNR declines. On the other hand, when Th_m becomes larger, more motions are discarded as noise, so that FPR decreases and FNR increases. For MDE, we find that it increases with Th_s and Th_m, as Figure 12(c) shows. In
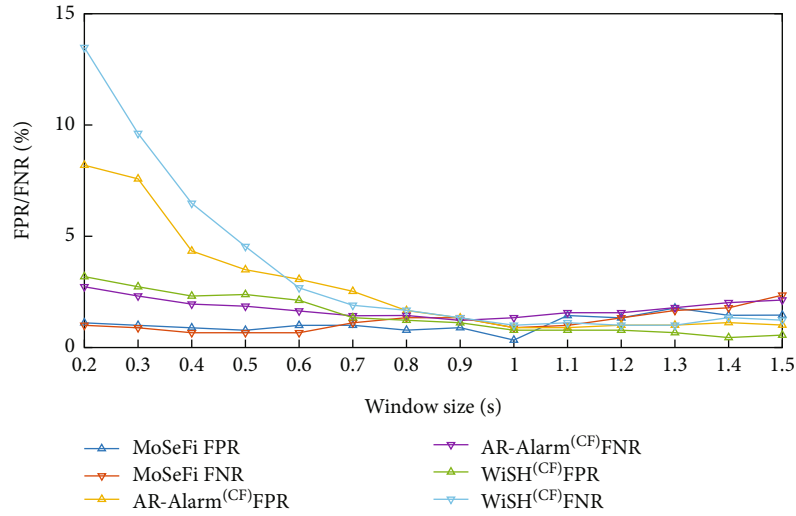
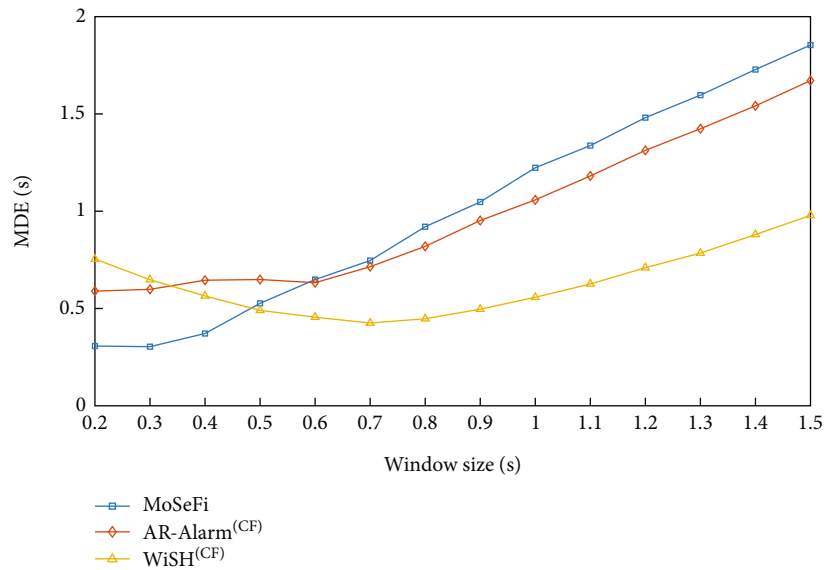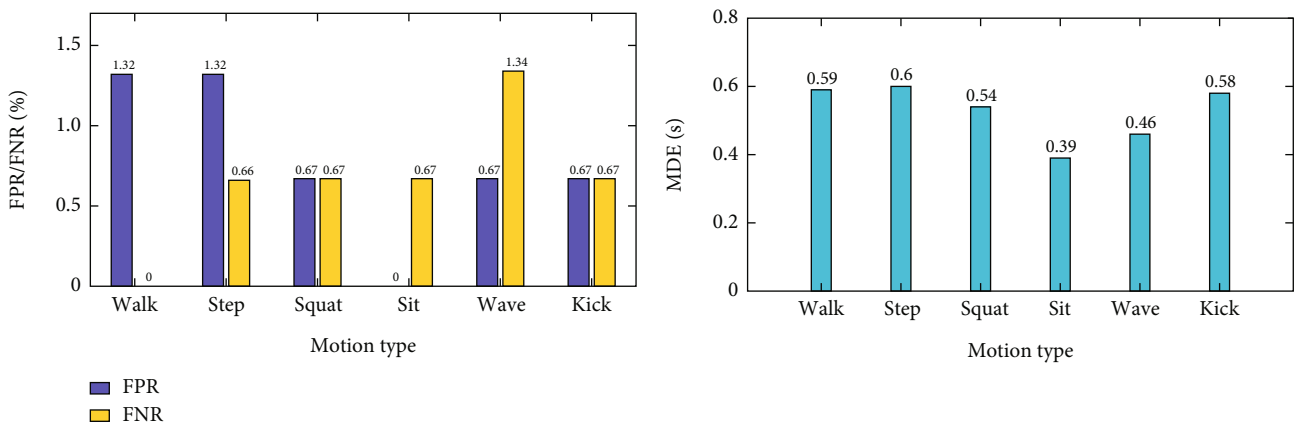Figure 9: Detection result with different window sizes.

Figure 10: Motion duration accuracy with different window sizes.

(a)

(b)

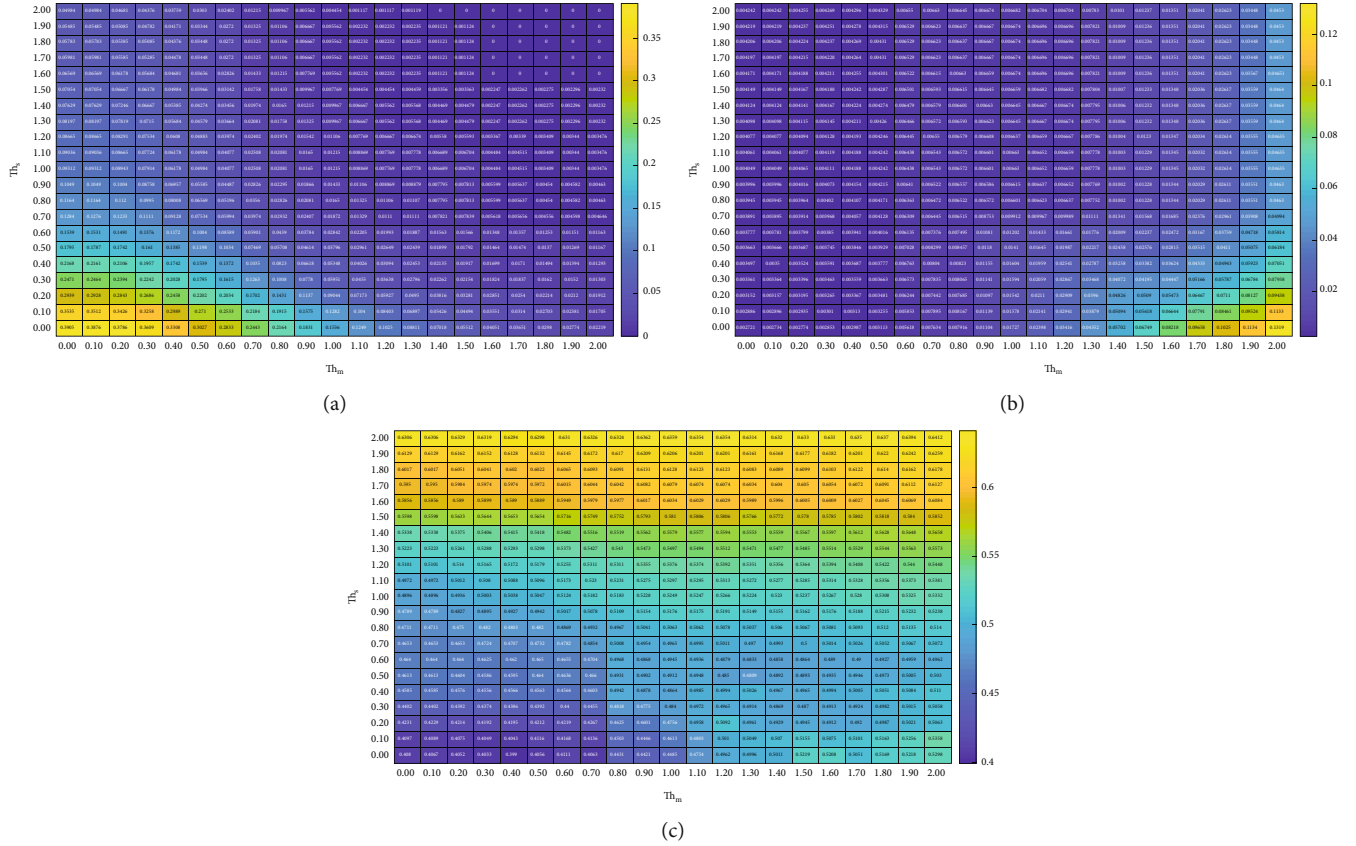Figure 11: Impact of motion type. (a) FPR and FNR. (b) MDE.

(a)

(b)

(c)

FIGURE 12: Impact of static and motion duration threshold. (a) FPR. (b) FNR. (c) MDE.
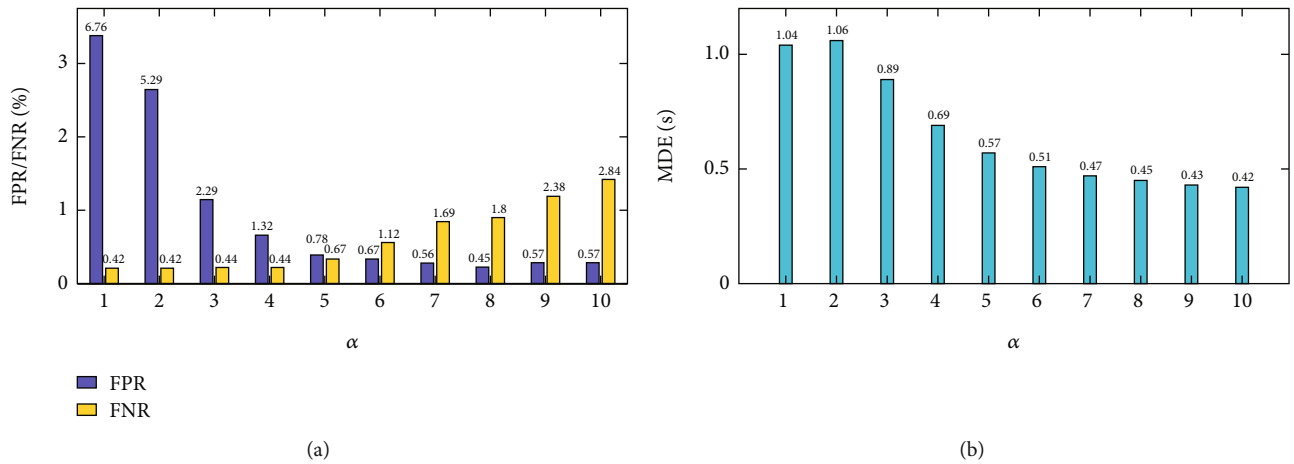

(a)

(b)

FIGURE 13: Impact of average variance threshold. (a) FPR and FNR. (b) MDE.

comparison, Th_s has a greater impact on MDE because it determines whether the fragments near the head or tail of the true motion interval can be merged into the detection result. Considering the motion detection and duration accuracy comprehensively, we choose Th_s as 1 s and Th_m as 1.2 s in this paper.

5.2.6. Threshold of Average Variance. Figure 13(a) shows the detection accuracy when α changes from 1 to 10. It can be seen that FPR decreases and FNR increases with α. This is

reasonable because we discard motion segments with average variance less than α times Th_{M_indi}; in other words, α determines the minimum average variance of the retained motion segments. When α gets larger, some true motion segments with small variance may be erroneously eliminated, resulting in an increase of FNR. On the contrary, once the value of α is small, it is impossible to remove all pseudomotion segments with small variance, so that the FPR increases. Meanwhile, if α is too small, the motion fragments caused by noise may be mistakenly merged into the real motion
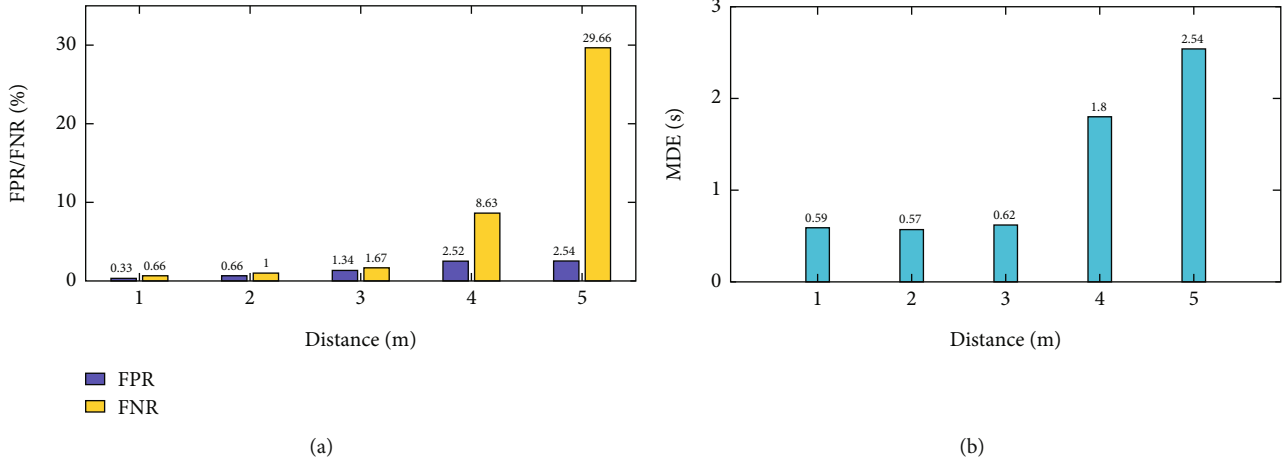
(a)

(b)

FIGURE 14: Impact of distance between human and LOS. (a) FPR and FNR. (b) MDE.
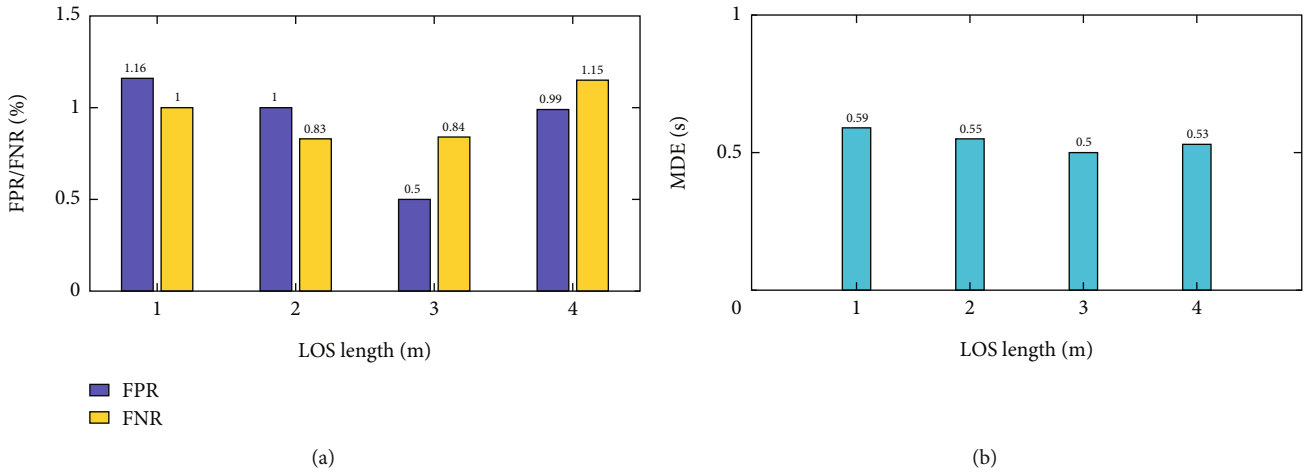

(a)

(b)

FIGURE 15: Impact of LOS length. (a) FPR and FNR. (b) MDE.

interval, thereby reducing the accuracy of motion duration, as Figure 13(b) shows. In this paper, we choose α as 5.

5.2.7. Distance between Human and LOS. To test the sensing range of MoSeFi, we selected 5 positions on the midperpendicular of LOS, which are 1 to 5 meters away from the midpoint of LOS. Then, the volunteers completed the above six actions in these positions, respectively, and the experimental results are shown in Figure 14. It can be seen that MoSeFi is very stable when motion occurs within 3 m from the LOS; the FPR, FNR, and MDE are all at a very low level. However, when the motion occurs beyond 3 m, the induced CSI fluctuation becomes small and the system performance degrades. Thus, equipment deployment should be reasonably designed to ensure that motion occurs within the effective detection range of the system. In addition, we find that small-scale motions, such as waving and kicking, are more likely to be missed as distance increased. The increase in distance further weakens the already small CSI fluctuations, which makes these small-scale motions more difficult to detect.

5.2.8. Length of LOS. We changed the distance between the transmitter and receiver from 1 m to 4 m and asked the volunteers to complete the aforementioned six actions within the effective detection range. Figure 15 shows the detection results, and it can be seen that the impact of LOS length is weaker than that of the distance between human and LOS, and the system can obtain relatively better performance with middle-length LOS. A short LOS means a large static signal strength, while a long LOS makes the dynamic reflection path longer, both of which make the proportion of the dynamic signal smaller, resulting in a decrease in system performance. Furthermore, we found that the relationship between system performance and LOS length is little affected by motion types; too long or too short LOS is unfavorable to the detection of all six actions.

5.2.9. Sampling Rate. We gradually reduced the sampling rate from 200 packets/s to 20 packets/s, and the detection results are shown in Figure 16. It can be seen that MoSeFi performs well when the sample rate is above 80 packets/s, and the motion detection and duration accuracy of the
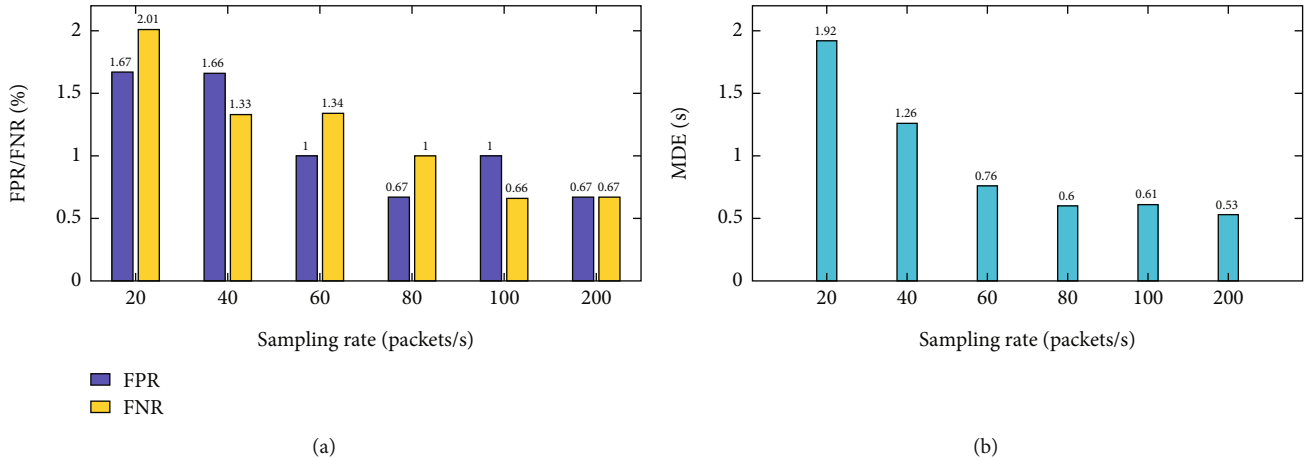
Figure 16: Impact of sampling rate. (a) FPR and FNR. (b) MDE.

system is maintained at a relatively high level. Besides, we find that the detection accuracy and duration accuracy are not significantly improved when the sampling rate increases to 200 packets/s, but the volume of data doubles. However, when the sample rate drops to 60 packets/s and below, the system performance gradually deteriorates. On the one hand, the decreasing of captured information reduces the motion detection accuracy. On the other hand, the reduction of the temporal resolution makes the MDE rise significantly.

5.2.10. Environment Change. In order to verify the system performance in different environments, we redeployed MoSeFi in scenario (b) and asked the volunteers to complete the above six actions in three positions which are shown in Figure 8(b). The experimental results show that our system can maintain a relatively stable performance in new scenario, and its FPR, FNR, and MDE are 0.70%, 0.83%, and 0.58 s, respectively. When the environment changes, the subcarrier screening module of our system can still pick excellent subcarriers and recompute suitable benchmarks to detect human motion. On the whole, our system is robust to the changes of environment.

5.3. Motion Sequence Case. The action sequence is an important form of human motion in real life; as Figure 8(a) shows, we designated two paths ($D_1 \longrightarrow D_1' \longrightarrow D_1$ and $E_1 \longrightarrow E_1' \longrightarrow E_1$) in scenario (a) and asked the volunteers to complete a series of actions, including standing up, walking, drinking, walking back, and sitting down, according to the voice commands, and each motion sequence was completed twenty times on each path. As discussed in the single motion case, the performance of AR_Alarm and WiSH is rather poor under short-window conditions; thus, we only tested their performance with 1 s window for comparison.

In Table 3, we show the system performance. On the whole, the detection accuracy of MoSeFi remains at a high level. The FPR and FNR of MoSeFi are 1.00% and 0.90%, respectively, which are the lowest among the three systems. Nevertheless, due to the lack of an effective subcarrier screening method and motion debris disposal

mechanism, the FPR and FNR of AR-Alarm are 4.70% and 4.80%, respectively. Compared to the single motion case, the detection accuracy of WiSH improves slightly. Through in-depth analysis, we found that the errors of WiSH are mainly concentrated in waving and kicking in the case of single motion; however, these small-amplitude motions are not included here, which makes the FPR and FNR WiSH drop slightly.

Benefit from the shorter window, the MDE of MoSeFi is 0.75 s, which is much smaller than the 1.25 s of AR-Alarm. Although WiSH performs slightly better in terms of the duration accuracy, its FPR and FNR are obviously larger than those of MoSeFi. Meanwhile, the RT of WiSH is 3.27 s, which is the longest of the three systems. It is obvious that the real-time performance of the variance-based method is much better than that of the correlation-based method.

Note that we have not updated the subcarrier or the threshold for the medium-length CSI data, and MoSeFi still achieves satisfactory performance. This verifies that the nice subcarriers can maintain fine environmental awareness for a period of time.

5.4. Evaluation of Long-Term Performance. To evaluate the stability of MoSeFi, we deployed the system in scenario (b) and collected ten hours of CSI data on a normal working day. During the data collection, the status of indoor personnel was not restricted; they can either stay still or perform different actions as usual.

Table 4 shows the overall performance of the three systems. Specifically, the FPR of MoSeFi is 1.71%, and the values of AR-Alarm and WiSH are 6.61% and 4.96%, respectively. We note that the FP of MoSeFi is significantly less than that of the other two systems. Since the noise level in the real environment for a long time is more complicated, the mean variance threshold adopted by MoSeFi can effectively eliminate the pseudomotion caused by noise, which is not available in the other two systems. Meanwhile, both MoSeFi and WiSH have FNRs of 2.54%, which are smaller than AR-Alarm. The nice subcarriers are more likely to degrade under long-term

TABLE 3: System performance of motion sequence case.

System	FPR (%)	FNR (%)	MDE (s)	RT (s)
MoSeFi	1.00	0.90	0.75	0.26
AR-Alarm	4.70	4.80	1.25	0.21
WiSH	2.91	3.31	0.74	3.27

TABLE 4: System performance for long-term case.

System	Time	8:00-10:00	10:00-12:00	12:00-14:00	14:00-16:00	16:00-18:00	Total	FPR/FNR (%)
	TP	27	11	37	15	25	115	
MoSeFi	FP	0	0	1	0	1	2	1.71
	FN	1	1	0	0	1	3	2.54
	TP	27	12	34	15	25	113	
AR-Alarm	FP	3	2	1	0	2	8	6.61
	FN	1	0	3	0	1	5	4.24
	TP	27	12	36	15	25	115	
WiSH	FP	1	0	2	2	1	6	4.96
	FN	1	0	1	0	1	3	2.54

conditions, resulting in an increase in the number of missed detection. These phenomena are consistent with the conclusions obtained in the previous experimental procedure. In addition, we find that our system can correctly capture the movement when a random visitor enters the room, which indicates that MoSeFi can also be applied in applications such as intrusion detection.

6. Conclusion

In this paper, we present the design and implementation of MoSeFi, a device-free and duration estimation robust human motion sensing system using ubiquitous WiFi signal. Based on the analysis of Mobius transform, we construct a novel indicator for motion detection using the shape-complementary real and imaginary parts of the CSI ratio, which can significantly reduce the motion duration error under short-window conditions. Furthermore, we propose a universal subcarrier screening method based on sensitivity and similarity and provide an update mechanism to attenuate the impacts of environmental variations. We conduct detailed experiments in real environments, and the results show that MoSeFi is lightweight yet efficient. We believe that this system enriches the practical solutions of passive human motion detection and facilitates the development of upper-level applications.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] M. Seifeldin, A. Saeed, A. E. Kosba, A. El-Keyi, and M. Youssef, "Nuzzer: a large-scale device-free passive localization system for wireless environments," *IEEE Transactions on Mobile Computing*, vol. 12, no. 7, pp. 1321–1334, 2012.

[2] J. Yang, Y. Ge, H. Xiong, Y. Chen, and H. Liu, "Performing joint learning for passive intrusion detection in pervasive wireless environments," in *2010 Proceedings IEEE INFOCOM*, pp. 1–9, San Diego, CA, USA, 2010.

[3] Y. Gu, L. Quan, and F. Ren, "WiFi-assisted human activity recognition," in *2014 IEEE Asia Pacific Conference on Wireless and Mobile*, pp. 60–65, Bali, Indonesia, 2014.

[4] K. Qian, C. Wu, Y. Zhang, G. Zhang, Z. Yang, and Y. Liu, "Widar2. 0: passive human tracking with a single Wi-Fi link," in *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, pp. 350–361, Munich, Germany, 2018.

[5] C. Han, W. Xun, L. Sun, Z. Lin, and J. Guo, "DSCP: depthwise separable convolution-based passive indoor localization using CSI fingerprint," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 8821129, 17 pages, 2021.

[6] J. Lv, D. Man, W. Yang, X. Du, and M. Yu, "Robust WLAN-based indoor intrusion detection using PHY layer information," *IEEE Access*, vol. 6, pp. 30117–30127, 2017.

[7] J. Wang, Z. Tian, M. Zhou, J. Wang, X. Yang, and X. Liu, "Leveraging hypothesis testing for CSI based passive human intrusion direction detection," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 8, pp. 7749–7763, 2021.

[8] J. Zhao, J. Zhou, Y. Yao, D. A. Li, and L. Gao, "RF-motion: a device-free RF-based human motion recognition system," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 1497503, 9 pages, 2021.

[9] Z. U. A. Akhtar, H. F. Rasool, M. Asif, W. U. Khan, and M. Ali, "Driver's face pose estimation using fine-grained Wi-Fi signals for next-generation Internet of Vehicles," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 7353080, 18 pages, 2022.

[10] X. Dang, Y. Liu, Z. Hao, X. Tang, and C. Shao, "Air gesture recognition using WLAN physical layer information," *Wireless Communications and Mobile Computing*, vol. 2020, Article ID 8546237, 14 pages, 2020.

[11] R. Gao, M. Zhang, J. Zhang et al., "Towards position-independent sensing for gesture recognition with Wi-Fi," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 5, no. 2, pp. 1–28, 2021.

[12] H. Yang, Z. Ji, J. Sun et al., "Recognition for human gestures based on convolutional neural network using the off-the-shelf Wi-Fi routers," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 782124, 12 pages, 2021.

[13] A. E. Kosba, A. Saeed, and M. Youssef, "RASID: a robust WLAN device-free passive motion detection system," in *2012 IEEE International Conference on Pervasive Computing and Communications*, pp. 180–189, Lugano, Switzerland, 2012.

[14] C. Siebert, M. Leng, S. G. Razul, C. M. S. See, and G. Wang, "Human motion detection and classification using ambient WiFi signals," in *2017 Sensor Signal Processing for Defence Conference (SSPD)*, pp. 1–5, London, UK, 2017.

[15] S. Depatla, A. Muralidharan, and Y. Mostofi, "Occupancy estimation using only WiFi power measurements," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 7, pp. 1381–1393, 2015.

[16] Y. Jin, Z. Tian, M. Zhou, Z. Li, and Z. Zhang, "A whole-home level intrusion detection system using WiFi-enabled IoT," in *2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC)*, pp. 494–499, Limassol, Cyprus, 2018.

[17] H. Huang and S. Lin, "WiDet: Wi-Fi based device-free passive person detection with deep convolutional neural networks," *Computer Communications*, vol. 150, pp. 357–366, 2020.

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

[19] J. Xiao, K. Wu, Y. Yi, L. Wang, and L. M. Ni, "FIMD: fine-grained device-free motion detection," in *2012 IEEE 18th International conference on parallel and distributed systems*, pp. 229–235, Singapore, 2012.

[20] K. Qian, C. Wu, Z. Yang, Y. Liu, and Z. Zhou, "PADS: passive detection of moving targets with dynamic speed using PHY layer information," in *2014 20th IEEE international conference on parallel and distributed systems (ICPADS)*, pp. 1–8, Hsinchu, Taiwan, 2014.

[21] H. Zhu, F. Xiao, L. Sun, R. Wang, and P. Yang, "R-TTWD: robust device-free through-the-wall detection of moving human with WiFi," *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 5, pp. 1090–1103, 2017.

[22] S. Li, X. Li, K. Niu, H. Wang, Y. Zhang, and D. Zhang, "AR-Alarm: an adaptive and robust intrusion detection system leveraging CSI from commodity Wi-Fi," in *International conference on smart homes and health telematics*, M. Mokhtari, B. Abdulrazak, and H. Aloulou, Eds., vol. 10461 of Lecture Notes in Computer Science, , pp. 211–223, Springer, Cham, 2017.

[23] T. Hang, Y. Zheng, K. Qian et al., "WiSH: WiFi-based real-time human detection," *Tsinghua Science and Technology*, vol. 24, no. 5, pp. 615–629, 2019.

[24] X. Yang, Y. Yin, P. Chen, and Q. Niu, "A device-free intelligent alarm system based on the channel state information," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 10, pp. 11419–11427, 2020.

[25] X. Niu, S. Li, Y. Zhang et al., "WiMonitor: continuous long-term human vitality monitoring using commodity Wi-Fi devices," *Sensors*, vol. 21, no. 3, p. 751, 2021.

[26] X. Li, S. Li, D. Zhang, J. Xiong, Y. Wang, and H. Mei, "Dynamic-music: accurate device-free indoor localization," in *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing*, pp. 196–207, Heidelberg, Germany, 2016.

[27] Y. Zhuo, H. Zhu, H. Xue, and S. Chang, "Perceiving accurate CSI phases with commodity WiFi devices," in *IEEE INFOCOM 2017-IEEE Conference on Computer Communications*, pp. 1–9, Atlanta, GA, USA, 2017.

[28] Y. Zeng, D. Wu, J. Xiong, E. Yi, R. Gao, and D. Zhang, "FarSense," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 3, no. 3, pp. 1–26, 2019.

[29] D. Wu, Y. Zeng, R. Gao et al., "WiTraj: robust indoor motion tracking with WiFi signals," *IEEE Transactions on Mobile Computing*, p. 1, 2021.

[30] L. Najman and P. Romon, *Modern Approaches to Discrete Curvature*, vol. 2184, Springer, Berlin, 2017.

[31] S. Zhang, B. Li, Z. Zhang, J. Ma, P. Li, and H. Wang, "Robust corner finding based on multi-scale K-cosine angle detection," *IEEE Access*, vol. 8, pp. 66741–66748, 2020.