

# Toward Low-Cost Passive Motion Tracking With One Pair of Commodity Wi-Fi Devices

Wei Guo , Graduate Student Member, IEEE, and Lei Jing , Member, IEEE

**Abstract**—With the popularity of Wi-Fi devices and the development of the Internet of Things (IoT), Wi-Fi-based passive motion tracking has attracted significant attention. Most existing works utilize the Angle of Arrival (AoA), Time of Flight (ToF), and Doppler Frequency Shift (DFS) of the Channel State Information (CSI) to track human motions. However, they usually require multiple pairs of Wi-Fi devices and extensive data training to achieve accurate results, which is unrealistic in practical applications. In this article, we propose Wi-Fi Motion Tracking (WiMT), a low-cost passive motion tracking system based on a single pair of commodity Wi-Fi devices. WiMT calculates the Doppler velocity and phase difference using the CSI obtained from the transmitter with one antenna and the receiver with three antennas. The Zero Velocity Identification and Calibration (ZVIC) algorithm is proposed to remove the random noise of Doppler velocity when the target is stationary. We take the Doppler velocity as the measurement and employ a particle filter to estimate the motion trajectory. A particle weight update method based on phase difference information is developed to eliminate particles with low confidence. Experimental results in real indoor environment show that WiMT achieves great performance with a motion tracking median error of 7.28 cm and a nonmoving recognition accuracy of 92.6%.

**Index Terms**—CSI ratio, channel state information, Doppler velocity, particle filter, passive tracking, phase difference, Wi-Fi sensing.

## I. INTRODUCTION

INDOOR motion tracking plays a crucial role in many intelligent applications, such as smart home, elderly care, indoor navigation, etc. Although satellite positioning technology has achieved high accuracy, satellite signals cannot be received indoors, unlike the outdoor environment. Therefore, the issue of how to implement indoor positioning has attracted the interest of researchers.

Recently, tracking solutions based on various kinds of devices have been proposed. Vision-based approaches [1], [2] can achieve high accuracy, but the high cost, limited viewing angle, and privacy issues have limited their popularity. Acoustic-based solutions [3] do not have the above disadvantages, but the coverage area is too small. Inertial measurement unit (IMU)-based methods [4] are not disturbed by the external environment and are easy to carry, but the tracking error accumulates over time, thus they need to be integrated with other sensors to achieve promising performance. Due to the ubiquity of radio frequency (RF) devices in daily life, RF-based solutions such as RF identification (RFID) [5], [6], ultra-wideband (UWB) [7], [8], Bluetooth [9], [10], and Wi-Fi [11], [12], [13], [14], [15],

[16], [17], [18] have attracted growing attention. In addition, Wi-Fi-based methods are more attractive due to their low cost and ubiquity.

Wi-Fi-based motion tracking can be divided into two categories: device-based [11], [12] and device-free [13], [14]. Device-based methods require the user to carry the device, which is impractical for applications, such as elderly care and intrusion detection. In contrast, device-free motion tracking offers more user-friendly experience, requiring only passive cooperation from users without the need to carry any devices. It utilizes the received signal strength indicator (RSSI) [19], [20], [21] and CSI of Wi-Fi signals to locate and track the target. Compared to RSSI, CSI is a fine-grained measurement that can achieve higher estimation accuracy. The existing CSI-based motion tracking solutions are classified into three categories: the Time of Flight (ToF) based [22], [23], the Angle of Arrival (AoA) based [11], [13], [24], and the DFS based [14], [24], [25], [26], [27]. In reality, CSI is polluted by ambient noise and hardware devices, and accurate ToF and AoA require large-scale antenna arrays to obtain, which is not feasible in commercial Wi-Fi. To overcome the influence of noise, Ding et al. [28] have presented a 3-D indoor localization and tracking system that exploits deep learning method to train offline CSI data collected from various positions and motions. While this method is limited to certain scenes and predefined motions, CSI data has to be collected and trained again when the room is changed or motions are added. IndoTrack [13] and Widar [14] employ DFS to estimate the location of target based on a model-based algorithm without training samples, but they require multiple Wi-Fi devices, which limits their application in practical scenarios. Widar2.0 [25] first proposed a motion tracking method with single Wi-Fi link, which employs multidimensional parameters including AoA, ToF, DFS, and the attenuation. With a pair of transceivers, there

Manuscript received 7 April 2023; revised 23 April 2023 and 12 June 2023; accepted 13 June 2023. Date of publication 20 June 2023; date of current version 7 July 2023. This work was supported in part by JSPS KAKENHI under Grant 22K12114, in part by JKA Foundation, and in part by NEDO Younger Research Support Project under Grant JPNP20004. (Corresponding author: Lei Jing.)

Wei Guo is with the Graduate School of Computer Science and Engineering, University of Aizu, Aizuwakamatsu, Japan (e-mail: d8232103@u-aizu.ac.jp).

Lei Jing is with the School of Computer Science and Engineering, University of Aizu, Aizuwakamatsu 965-0006, Japan (e-mail: lei-jing@u-aizu.ac.jp).

Digital Object Identifier 10.1109/JISPIN.2023.3287508

is no redundancy in the system to ensure robust tracking across the entire motion trajectory [27], especially when the target is moving along the tangent direction of the Fresnel zone, as the DFS cannot be accurately captured. WiTraj [27] proposed a robust motion tracking system that uses three pairs of Wi-Fi devices and calculates the Doppler velocity using the CSI quotient model [29], [30] on two adjacent antennas at each receiver. The number of transceivers and specific placement ensure robust performance.

From the comparison of the above studies, it is obvious that the balance between the number of devices, the robustness of the system, and the algorithm complexity is quite crucial. More accurate parameter information of the moving target can be obtained by using more Wi-Fi links. In practical applications, only one Wi-Fi link is usually available indoors. Therefore, it is vital to utilize a limited number of devices to obtain fine-grained parameters for passive tracking. In this article, we propose WiMT, a low-cost indoor passive motion tracking system based on a single pair of commodity Wi-Fi devices. The system overview is shown in Fig. 2. WiMT utilizes DFS and the phase difference of two adjacent antennas at the receiver to track the target's motion trajectory. The key insight is that the target's motions not only introduce the change in the path length of the reflected signals, but also change the phase difference between the antennas. WiMT employs DFS and phase difference information to calculate the Doppler velocity, and then combines the Doppler velocity and phase difference to estimate the moving trajectory based on a particle filter. However, there are three challenges that need to be solved.

First, it is hard to obtain DFS over a single Wi-Fi link when the target is moving along the tangent direction of the Fresnel zone. To the best of our knowledge, the DFS is caused by the change of the path length of the reflected signals, but the movement in the tangent direction cannot change the path length, and the DFS is zero even though the target is moving. Unlike WiTraj [27] which ensures that there are at least two receivers that get a good estimate of the DFS at any time.

Second, WiMT employs the DFS acquisition method [27] to obtain Doppler velocity. Theoretically, the antennas of a receiver share the same RF chain and clock, the amplitude noise and random phase shifts are almost identical for each antenna, so the CSI quotient cancels out both the amplitude noise and the phase noise [27]. However, the noise on two adjacent antennas at the same receiver is not identical, and the CSI quotient still contains noise, especially if the target is stationary. As a result, the Doppler velocity, which should be zero when the target is stationary, is randomly distributed. This leads to significant errors in motion tracking.

Third, the orientation and velocity information obtained from a single Wi-Fi link is limited compared to multiple Wi-Fi links, making it difficult to maintain stable performance during motion tracking.

To address the above challenges, this article makes the following contributions.

- 1) We propose a **Z**ero **V**elocity **I**dentification and **C**alibration (**ZVIC**) algorithm. It uses the phase difference information to accurately identify the stationary and

moving states of the target over a pair of Wi-Fi devices. In addition, it calibrates the random noise of the Doppler velocity when the target is stationary, addressing the shortcomings of the DFS acquisition method.

- 2) Applying the particle filter, we propose a trajectory estimation method based on Doppler velocity and phase difference. We use the Doppler velocity as a measure of the filter and use the phase difference information to remove particles with low confidence.
- 3) We propose a novel Wi-Fi-based motion tracking system, WiMT, which harnesses a single pair of unmodified Wi-Fi devices. In evaluations conducted within real indoor environments, the ZVIC accuracy is 92.6% for nonmoving and 85.2% for moving, median tracking error is 7.28 cm.

The rest of this article is organized as follows. In Section II, we introduce the related work. The preliminary is presented in Section III. Section IV is dedicated to the method. The WiMT is evaluated in Section V. Section VI covers the discussion. Finally, Section VII concludes this article.

## II. RELATED WORK

### A. Wi-Fi-Based Indoor Localization and Tracking

In recent years, Wi-Fi-based indoor localization has attracted the attention of researchers. As an early attempt, RADAR [19] utilized RSSI to locate and track the user in-building, but only meter-level accuracy is achieved due to the multipath effect. ViVi [20] employed fingerprint spatial gradient of RSSI to locate target and achieves great performance. Subsequently, researchers turn their attention to CSI, which can describe the amplitude and phase of Wi-Fi signals, enabling decimeter or even centimeter-level localization and tracking accuracy. ArrayTrack [12] expands the number of antennas on receiver and proposes multipath suppression algorithm to effectively remove the reflection paths between transmitter and receiver to obtain AoA. In applications such as elderly care and intrusion detection, it is inconvenient for users to carry Wi-Fi devices, so device-free localization and tracking solutions are proposed. xD-Track [17] is the first practical passive human localization system, which combines ToF, AoA, Angle of Departure (AoD), and DFS to fully characterize the wireless channel between transceivers. WiDir [31] innovatively analyzed the phase change dynamics from multiple Wi-Fi subcarriers based on Fresnel zone model and infers the target's moving directions. Widar [14] and IndoTrack [13] estimated target's moving speed, direction, and location at a decimeter level. In practical applications, it is inconvenient to deploy multiple Wi-Fi links. Hence, Widar2.0 [25] first implements target tracking via one Wi-Fi link. It combines the four parameters of AoA, ToF, DFS, and signal attenuation, but it cannot be robust in the whole motion tracking. PITrack [32] designed a scheme to dynamically select the best receivers among multiple Wi-Fi devices to maximize velocity estimation accuracy for moving targets, and achieve position independent target tracking. The mathematical model is very dependent on the angle of departure and angle of arrival of the signal. WiTraj [27] proposed a solution of three Wi-Fi link based on the CSI quotient model, no matter how the target moves, it can

ensure that there are two links in the system to achieve motion tracking. However, the transceiver equipment requires a special placement, which is inconvenient in practical applications. In addition, the Doppler velocity calculated by WiTraj is randomly distributed when the target is stationary, which causes a large error in trajectory estimation. In order to overcome the above shortcomings, WiMT proposed the ZVIC algorithm to remove the Doppler velocity noise when the target is stationary, and only use one Wi-Fi link to realize target tracking.

### B. Wi-Fi-Based Gesture and Activity Recognition

Wi-Fi-based human gesture and activity, as a novel communication approach, are widely adopted in human-computer interfaces (HCI) for its natural and straightforward properties [33]. WiGest [34] utilized variations in the RSSI of Wi-Fi signals to estimate human hand gestures. In order to overcome the shortcoming that coarse-grained RSSI is easily affected by the environment. CARM [18] proposed CSI-speed model and CSI-activity model, and uses these two models to establish the relationship between CSI value dynamics and human activities. WiDance [35] presented a Wi-Fi-based contactless dance-pad exergame based on DFS. However, it can only recognize eight predefined directions and cannot recognize complex movements. RT-Fall [36] exploited phase difference and amplitude of the fine-grained CSI to detect target's body falling. QGesture [37] used CSI values provided by COTS Wi-Fi devices to measure the movement distance and direction of human hands. SignFi [38] utilized CSI as the input and a convolutional neural network (CNN) as the classification algorithm to recognize 276 sign gestures, which involve the head, arm, hand, and finger gestures, with high accuracy. In [39] and [40], Wi-Fi-based human pose tracking solutions are proposed. These works are based on deep learning methods, CSI data needs to be recollected and trained when adding actions and changing the environment. Widar3.0 [26], WiGesture [30], DPSense [41], and [42] proposed position-independent gesture recognition methods, solve the problem that the recognition accuracy drops when changing scenes and locations. However, the above works depend on the number of devices and special placement. WiMT does not depend on the placement of devices, and only uses a pair of Wi-Fi devices to realize motion tracking.

## III. PRELIMINARY

In this section, we first introduce the basics of CSI, and then explain the principle of the CSI quotient.

### A. Channel State Information

In a narrow-band flat fading channel, the Wi-Fi orthogonal frequency division multiplexing (OFDM) system viewed in the frequency domain can be defined as [43]

$$\vec{Y} = H\vec{X} + \vec{N} \quad (1)$$

where  $\vec{Y}$  and  $\vec{X}$  represent the received and transmitted signal vectors, respectively.  $H$  denotes the channel frequency response (CFR) and  $\vec{N}$  is the additive white Gaussian noise (AWGN).

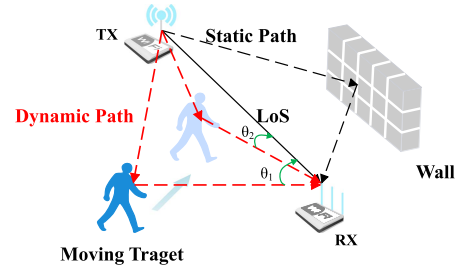


Fig. 1. Signal propagation paths indoors. The path between TX and RX is Line of Sight (LoS).

In an indoor environment, wireless signals propagate from TX to RX through multiple paths, i.e., one Line of Sight (LoS) path and multiple reflection paths from objects (such as walls, furniture, and the moving target) [44], as shown in Fig. 1. Hence, the CFR in time and frequency as amplitude and phase in the format of CSI is a superposition of signals from all the paths. Mathematically, the CSI can be represented as [44]

$$H(f_i, t) = \sum_{k=1}^L A_{i,k} e^{-j2\pi f_i \tau_k(t)} \quad (2)$$

where  $f_i$  is the carrier frequency of the  $i$ th subcarrier,  $i$  is the index of the OFDM subcarrier,  $i \in [1, 30]$ .  $L$  is the number of paths,  $A_{i,k}$  denotes the amplitude, and  $\tau_k(t)$  represents the propagation time of the  $k$ th path. Moreover, the phase of CSI at carrier frequency  $f_i$  propagating along the  $k$ th path can be written as  $\varphi_{i,k} = 2\pi f_i \tau_k(t)$ ,  $\tau_k(t)$  can be calculated by  $d_k(t)/c$ , where  $d_k(t)$  is the length of  $k$ th path, and  $c$  is the speed of light. In Fig. 1, we can see that the propagation paths are divided into static paths and dynamic paths. We assume there is only one dynamic path reflected by moving target and the static paths are composed of the LoS propagation and other reflection paths from static objects in the environment. Considering the case of one subcarrier, the CSI can be rewritten as [27]

$$H(f, t) = A_{\text{noise}}(f, t) e^{-j\varphi_{\text{offset}}(f, t)} (H_s(f) + H_d(f, t)) \quad (3)$$

where  $A_{\text{noise}}$  is the amplitude noise,  $\varphi_{\text{offset}}$  is the random phase offset caused by hardware imperfections,  $H_s(f)$  is the static component, and  $H_d(f, t)$  is the dynamic component.

### B. CSI Quotient

The CSI-quotient model takes the ratio of CSI readings of two antennas on the same RX as a new base signal, which can be expressed as follows [27]:

$$H_q(f, t) = \frac{CSI_1(f, t)}{CSI_2(f, t)} = \frac{H_{s1}(f) + H_{d1}(f, t)}{H_{s2}(f) + H_{d2}(f, t)} \quad (4)$$

where  $H_q(f, t)$  denotes the CSI quotient,  $H_{s1}(f)$  and  $H_{s2}(f)$  represent the static components, while  $H_{d1}(f, t)$  and  $H_{d2}(f, t)$  represent the dynamic components.

The DFS can be derived by employing the CSI-quotient model. The rationale for this method is that the CSI quotient between two antennas is a Möbius transform, the length of the reflection path changes by one wavelength, and the CSI quotient

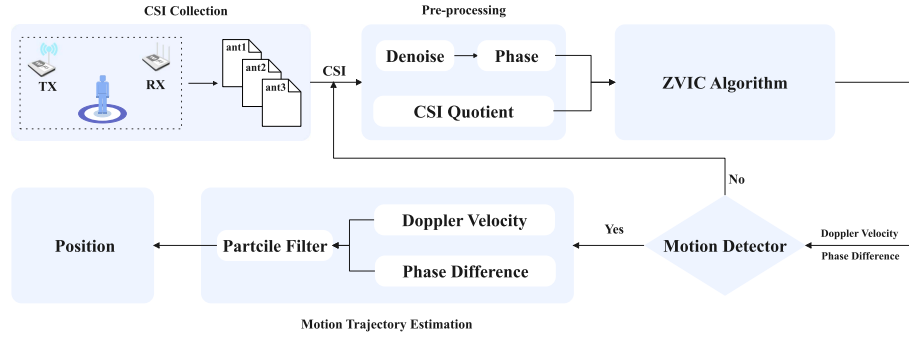


Fig. 2. WiMT System Overview: CSI data are collected by single pair of Wi-Fi devices and preprocessed to calculate the CSI quotient and phase difference. The random noise in the Doppler velocity is identified and removed by the ZVIC algorithm. The motion detector is used to detect the moving state of the target. When the target is moving, the trajectory of the target is estimated using a particle filter.

rotates into a perfect circle in the complex plane accordingly. Therefore, the DFS can be calculated by how many complete circles are rotated in the complex plane by the CSI quotient:

$$f_D = \frac{\Delta\rho}{2\pi\Delta t} \quad (5)$$

where  $f_D$  denotes the Doppler frequency shift,  $\Delta\rho$  is the phase change of the rotating circles in the  $H_q(f, t)$ , and  $\Delta t$  represents the sampling intervals. Accordingly, the Doppler velocity  $v_D$  can be obtained as follows:

$$v_D = f_D\lambda \quad (6)$$

where  $\lambda$  is the signal wavelength. The Doppler velocity indirectly reflects the speed of a person's movement, providing velocity information for motion tracking.

#### IV. METHOD

In this section, we demonstrate how WiMT achieves motion tracking with a single WiFi link. The system framework is illustrated in Fig. 2. First, the CSI is acquired by the RX equipped with three antennas. Then, the CSI data undergoes preprocessing to remove noise from the amplitude and phase while calculating the CSI-quotient to obtain the Doppler velocity. Subsequently, the denoised phase and Doppler velocity are fed into the ZVIC algorithm to derive fine-grained Doppler velocity. Finally, the Doppler velocity and phase difference are used as observation inputs for the particle filter, which calculates the target's position.

##### A. CSI Collection and Preprocessing

1) *CSI Collection*: We use Intel 5300 NIC to receive CSI data, TX is equipped with one antenna, and RX is equipped with three antennas. Therefore, the dimension of the CSI data matrix is  $1 \times 3 \times 30$ .

2) *CSI Denoise*: Due to the imperfections in hardware devices, the phase of CSI contains phase noise introduced by carrier frequency offset (CFO), sampling frequency offset (SFO), and packet detection delay (PDD), while the amplitude of CSI is polluted by environmental noise. We employ the Savitzky–Golay filter and linear transform method [16], [36], [45] to remove noise from the amplitude and phase, respectively. Fig. 3 shows the comparison of the raw CSI and the denoised CSI,

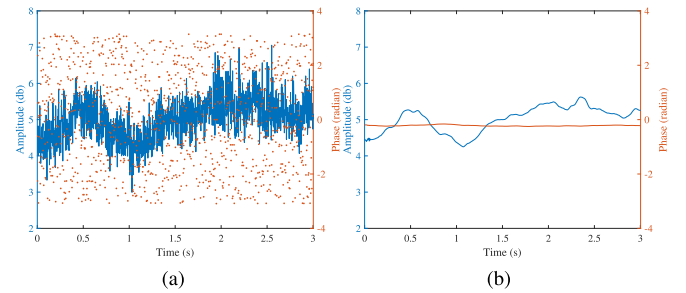


Fig. 3. Comparison of raw and denoised CSI. The blue lines are amplitude and orange dots are phase. The noise in amplitude and phase are effectively removed. (a) Raw CSI. (b) Denoised CSI.

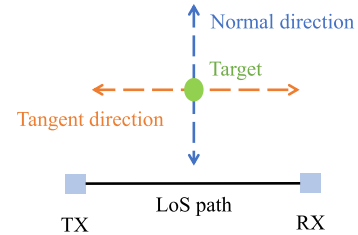


Fig. 4. Normal and tangent motion directions.

where blue line is the amplitude, and the orange dots represent the phase. Obviously, the amplitude noise and the randomly distributed phase offset noise in Fig. 3(a) are effectively removed by the above methods, and the results are shown in Fig. 3(b). We can calculate the phase difference with the denoised phase.

3) *Coarse-Grained Doppler Velocity Calculation*: As previously stated, we use the CSI-quotient model to compute the Doppler velocity. First, we calculate the CSI quotient between two antennas using (4), and then we compute the Doppler velocity using (5) and (6). In an indoor environment, we collect CSI data of the target moving in both the normal and tangential directions, as shown in Fig. 4, and calculate the CSI quotient and Doppler velocity. In Fig. 5(b) and (d), the CSI quotient in the complex plane is plotted for two traces, with trace A moving along the normal direction and trace B along the tangent direction. For clarity, Fig. 5(a) and (c) displays the

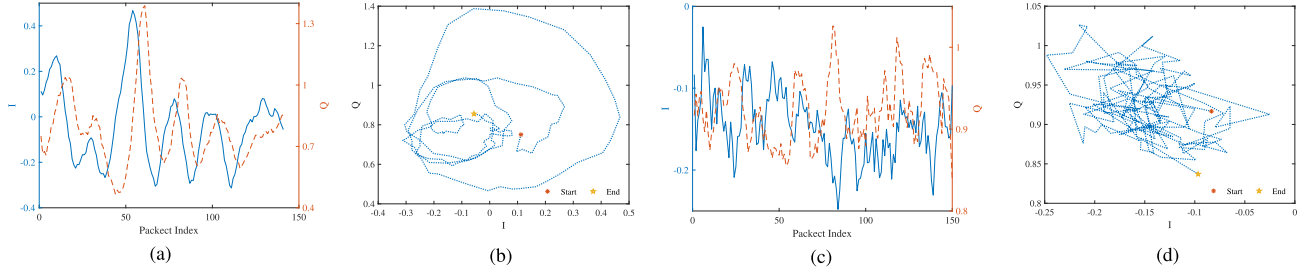


Fig. 5. CSI quotient of traces A and B. Trace A is along normal direction and trace B is along tangent direction. I and Q are the real and imaginary parts of CSI quotient. The CSI quotient of trace A form complete circles on the complex plane, but trace B cannot form. (a) I/Q wave of trace A. (b) CSI quotient of trace A. (c) I/Q wave of trace B. (d) CSI quotient of trace B.

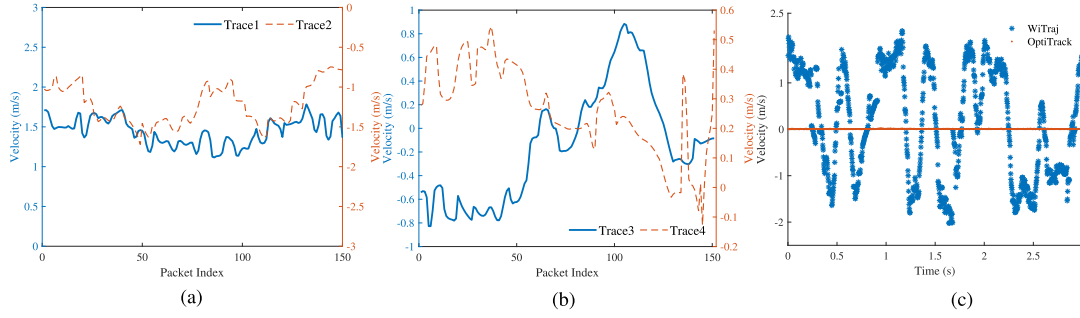


Fig. 6. Doppler velocity calculated by CSI quotient method. (a) is the velocity of normal direction motion, has a stable and regular trend. (b) is the velocity of tangent direction motion, has an irregular trend. (c) is the velocity of nonmoving, should be equal to zero, but randomly distributed. (a) Along normal direction. (b) Along tangent direction. (c) Non-moving.

real and imaginary parts of the CSI quotient. The CSI quotient of trace A forms multiple complete circles, whereas that of trace B is irregular and lacks a circular shape. Similarly, as shown in Fig. 6, the Doppler velocity exhibits a stable and regular trend in the normal direction, while it is irregular in the tangent direction and during nonmotion. Notably, in Fig. 6(c), the target is stationary, yet the velocity is nonzero and randomly distributed.

From the results, it is evident that the CSI-quotient model accurately estimates Doppler velocity for motion directions with significant changes in reflection path length (e.g., normal direction). However, for directions with constant or insignificant changes in path length (e.g., stationary and tangent directions), the estimated Doppler velocity contains randomly distributed noise. This issue makes it difficult to determine whether the target is stationary or moving along the tangential direction. To eliminate random noise in Doppler velocity, we input the coarse-grained Doppler velocity obtained from the CSI-quotient model into the ZVIC algorithm.

## B. ZVIC Algorithm

The ZVIC algorithm consists of two parts. First, the phase difference between the antennas is calculated. Second, the phase difference information is used to estimate the motion state and direction of the target, identify the Doppler velocity noise, and calibrate according to the motion state.

1) *Phase Difference Calculation*: The AoA of the reflected signal varies with the target motion, as shown in Fig. 1. AoA

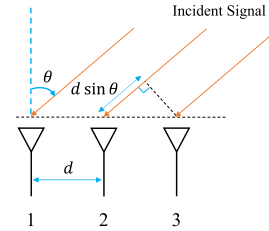


Fig. 7. Uniform linear antenna array.

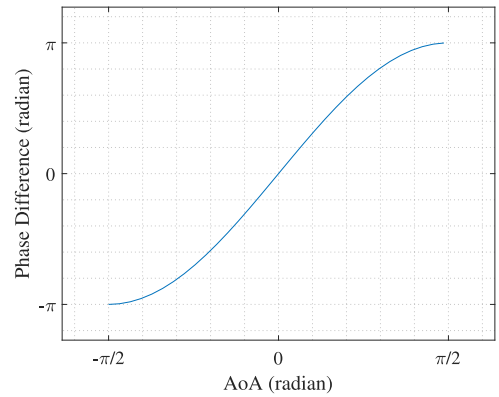


Fig. 8. Phase difference.

changes from  $\theta_1$  to  $\theta_2$  as the person moves. Moreover, the variation of AoA can be estimated by the phase difference between the two antennas. According to (2), the CSI phase of

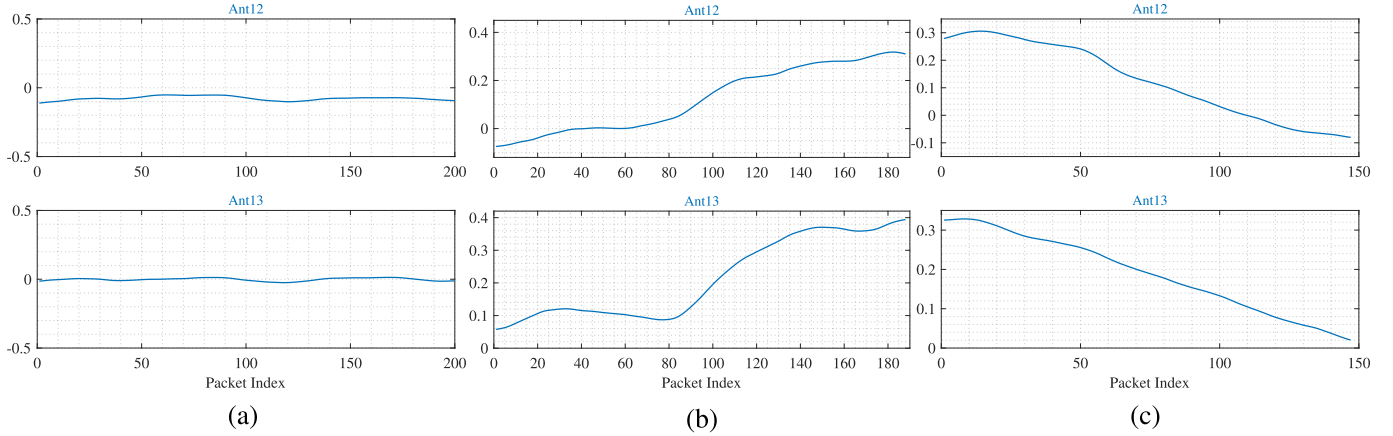


Fig. 9. Phase difference of distinct antenna pairs. The  $y$ -axis represents the phase difference and the unit is radian. Ant12 represents the antenna combination of first and second, and so on. (a) Static. (b) Move away from LoS. (c) Move close to LoS path.

the  $i$ th subcarrier propagated by the  $k$ th path to the  $n$ th antenna can be expressed as

$$\varphi_{i,k,n} = 2\pi f_i \tau_k + 2\pi f_i \Delta t (n-1) \quad (7)$$

where  $\Delta t$  represents the time difference of signal arriving at the adjacent antenna on RX. Assuming that the signal is on the same subcarrier via the same path, the CSI phase on different antennas can be expressed as

$$\varphi_n = 2\pi f \tau_n(t). \quad (8)$$

As shown in Fig. 7, the wave paths of the signals reflected by the target arriving the RX are different, and the signal wave path difference between adjacent antennas is  $d \sin \theta$ , where  $d$  is the antenna spacing  $\lambda/2$ , and  $\theta$  is the AoA. Taking antennas 1 and 2 as examples, the phase difference of the signals received by the two antennas can be expressed as

$$\Delta \varphi = \varphi_2 - \varphi_1 = 2\pi f (\tau_2(t) - \tau_1(t)) = \pi \sin \theta. \quad (9)$$

We define the variation range of AoA as  $-\frac{\pi}{2}$  to  $\frac{\pi}{2}$ , and the phase difference is monotonically increasing in the domain of definition as shown in Fig. 8, which allows us to accurately capture the variation of AoA.

In Fig. 9, we plot the phase difference when the target is stationary and moving. The  $x$ -axis is the packet index, and the  $y$ -axis is the phase difference between the antennas. ‘‘Ant12’’ represents the phase difference of the CSI received by antennas 1 and 2, and so on. In Fig. 9(a), it can be observed that the phase difference is a straight line when the target is stationary. As shown in Fig. 9(b), when the target moves away from the LoS on the right side of the LoS, the AoA gradually increases, and the phase difference also increases accordingly. In Fig. 9(c), when the target is close to the LoS, the AoA gradually decreases, and the phase difference decreases accordingly. Based on the variation of the phase difference, we can accurately distinguish the motion state and direction of the target, and then obtain a fine-grained Doppler velocity.

2) *Extracting Fine-Grained Doppler Velocity*: The phase difference exhibits minor fluctuations around a stable value during stationary periods, while it increases or decreases with the motion direction when the target is in motion. Observing this continuously, we notice peaks and troughs in the phase difference waveform, corresponding to changes in motion direction. The second part of the ZVIC algorithm identifies points with abrupt changes in slope and mean within the phase difference, with the peak-trough pairs containing these points representing the beginning and end of motion. We retain the Doppler velocities between these start and end points, setting the rest to zero, resulting in fine-grained Doppler velocities without random noise. The algorithm process is outlined in Algorithm 1. The input consists of phase differences  $P$  and coarse-grained Doppler velocities  $x_D$ , while the output is the fine-grained Doppler velocities  $v_D$ . The index of peaks and troughs in phase difference is stored in  $p$ , with the index of abrupt changes in slope and mean stored in  $p_1$  and  $p_2$ , respectively.  $M$  holds the index of identified motion start and end points, and  $C$  represents the confidence of Doppler velocities, with confidence values of 1 between the start and end points in  $M$  and zero elsewhere.

To test the performance of the ZVIC, we collected 20 s of CSI data at a sampling rate of 400 Hz while the target was moving and stationary, and use the data collected by the OptiTrack as ground truth. In Fig. 10, we employ the phase difference to identify the start, end, and direction of each piece of motion. Each pair of circles represents a start point and an end point. Moreover, the phase difference changes as the direction of motion changes. Based on above information, each Doppler velocity is given a confidence level of 0 for stationary and 1 for moving, and a fine-grained Doppler velocity is obtained. Notably, the ground truth represents the hand motion speed, and we utilize the metric of zero speed to differentiate between stationary and moving states of the target. In Fig. 11, the black line represents the ground truth, while the blue and orange lines correspond to the Doppler velocity obtained from WiTraj and ZVIC algorithms, respectively. The  $x$ -axis denotes time, and the  $y$ -axis represents Doppler velocity or hand movement speed. By comparison, it can be clearly seen that ZVIC accurately identifies the stationary

---

**Algorithm 1: ZVIC Algorithm.**


---

**Input:**  $P, \hat{x}_D$ 
**Output:**  $v_D$ 

- 1: Find the index of all peaks and troughs in  $P$  and store them in  $p$ .
  - 2: Compute the index of the values of the slope and mean abrupt changes in  $P$ , store them in  $p_1$  and  $p_2$  respectively.
  - 3:  $m_1 \leftarrow \text{sortasc}(p \cup p_1)$
  - 4:  $m_2 \leftarrow \text{sortasc}(p \cup p_2)$
  - 5: // Identify abrupt changes in the slope of the phase difference
  - 6: **for**  $i \leftarrow 1$  **to**  $\text{length}(p_1)$
  - 7:  $j \leftarrow \text{findindex}(m_1 == p_1[i])$
  - 8:  $M_1[2i - 1] \leftarrow m_1[j - 1]$
  - 9:  $M_1[2i] \leftarrow m_1[j + 1]$
  - 10: **end for**
  - 11: // Identify abrupt changes in the mean of the phase difference
  - 12: **for**  $i \leftarrow 1$  **to**  $\text{length}(p_2)$
  - 13:  $j \leftarrow \text{findindex}(m_2 == p_2[i])$
  - 14:  $M_2[2i - 1] \leftarrow m_2[j - 1]$
  - 15:  $M_2[2i] \leftarrow m_2[j + 1]$
  - 16: **end for**
  - 17:  $M \leftarrow \text{Merge } M_1 \text{ and } M_2, \text{ remove duplicates}$
  - 18:  $C \leftarrow \text{zeros}(\text{length}(P), 1)$
  - 19: // Find the start and end of motion
  - 20: **for**  $i \leftarrow 1$  **to**  $\text{length}(M)/2$
  - 21:  $C(M(2i - 1) : M(2i)) \leftarrow 1$
  - 22: **end for**
  - 23: // Remove the Doppler velocity noise
  - 24:  $v_D \leftarrow \hat{x}_D \cdot C$
- 

and moving state of the target, and filters out Doppler velocity noise when stationary.

### C. Motion Trajectory Estimation

In the Fresnel zone, a geometric relationship exists between the target's movement velocity and Doppler velocity. Leveraging this relationship, we utilize the fine-grained Doppler velocity to estimate the target's movement velocity. By establishing an observation equation based on the geometric relationship, we employ the fine-grained Doppler velocity and phase differences as inputs for the particle filter, enabling the estimation of the target's position coordinates.

1) *Geometric Relationship:* There is a geometric relationship between Doppler velocity and target moving velocity as shown in Fig. 12. The ellipse is the Fresnel zone with TX and RX are the focus,  $\alpha_T$  and  $\alpha_R$  are the Angle of Departure (AoD) and AoA of the signal, respectively.  $\vec{v}_H$  is the moving velocity of the target, the speed is  $v_H$ , and the direction is  $\phi$ .  $v_n$  represents the normal component of  $\vec{v}_H$ , and the reverse extension of the normal velocity intersects the  $x$ -axis at point  $F$ . We assume that the target is at point  $H$ , and TX and RX are at points  $T$  and  $R$ , respectively.  $HF$  is the angular bisector of  $\angle THR$ [13], [32].

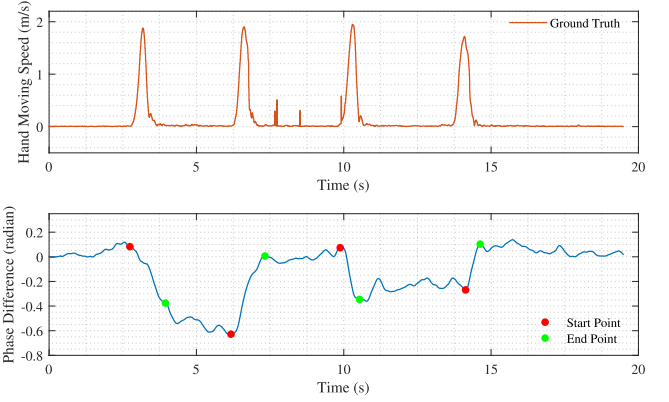


Fig. 10. Action start and end point. The orange line represents the ground truth, while the blue line corresponds to the phase difference. Red circles are the start points of the action, and green circles are the end points. The  $x$ -axis represents time, and the  $y$ -axis signifies either hand movement speed or phase difference. When the target moves, the phase difference will abruptly change from zero to the corresponding value, and show different trends with different moving directions.

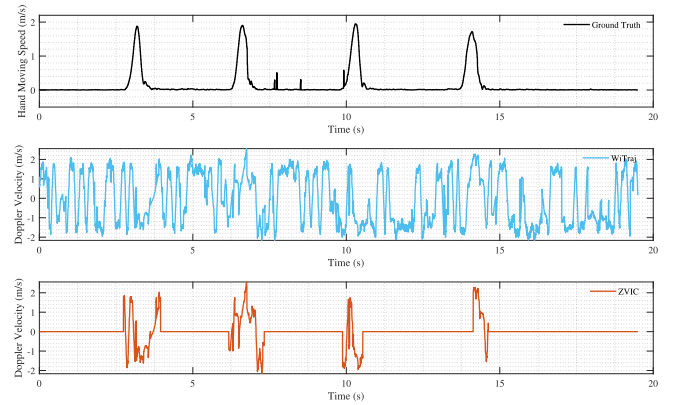


Fig. 11. Comparison of WiTraj and ZVIC. The black line represents the ground truth, while the blue and orange lines correspond to the Doppler velocity obtained from WiTraj and ZVIC algorithms, respectively. The  $x$ -axis denotes time, and the  $y$ -axis represents Doppler velocity or hand movement speed. ZVIC accurately identifies the stationary and moving state of the target, and filters out Doppler velocity noise when stationary.

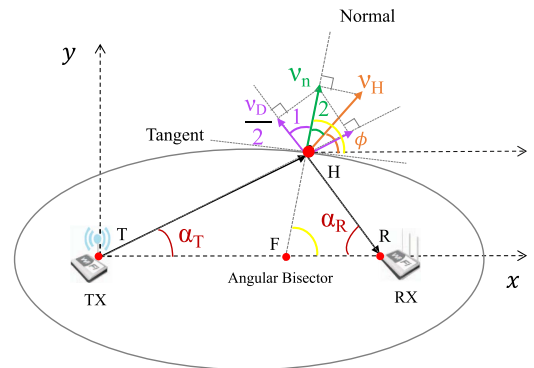


Fig. 12. Geometric relationship between human velocity and Doppler velocity.

The projection of  $v_n$  on the extension line of  $RH$  is  $\frac{v_D}{2}$ , where  $v_D$  is the Doppler velocity.

Based on the above geometric relationship, we can obtain the following equations:

$$\begin{cases} \angle 1 = \frac{\pi}{2} - \frac{\alpha_T + \alpha_R}{2} \\ \angle 2 = \frac{\pi}{2} + \frac{\alpha_T - \alpha_R}{2} + \phi \\ v_n = v_H \cos \angle 2 \\ v_D = 2v_n \cos \angle 1. \end{cases} \quad (10)$$

Simplifying (10) gives the velocity of the target

$$v_H = -\frac{v_D}{\sin(\phi + \frac{\alpha_T - \alpha_R}{2}) \cdot 2\sin(\frac{\alpha_T + \alpha_R}{2})}. \quad (11)$$

2) *Particle Filter*: We use particle filter to update the target's position, correspondingly, the state model and observation model can be represented as [46]

$$\begin{cases} X_t = f(X_{t-1}, u_t, w_t) \\ z_t = h(X_t, v_t) \end{cases} \quad (12)$$

where  $X_t$  denotes the state,  $u_t$  is the control input,  $w_t$  is the state noise,  $z_t$  represents the observation, and  $v_t$  is the observation noise at time  $t$ . The state transition and observation functions are represented by  $f(\cdot)$  and  $h(\cdot)$ , respectively.

We define the process state  $X_t$  and observation state  $z_t$  are as follows, respectively:

$$X_t = [x, y, v_x, v_y, \phi]^T \quad (13)$$

$$z_t = [v_D, \hat{\phi}_t, \Delta\varphi]^T \quad (14)$$

where  $x$  and  $y$  are the coordinates of the target,  $v_x$  and  $v_y$  are the velocity components of  $v_H$  along the  $x$ -axis and  $y$ -axis,  $\phi$  is the moving direction of target.  $v_{Dt}$  denotes the Doppler velocity at time  $t$ ,  $\hat{\phi}_t$  is the estimated heading angle, derived from the phase difference and the variance of the Doppler velocity,  $\Delta\varphi$  is the phase difference. We assign the known initial position coordinates to the state variables  $x$  and  $y$ , the initial heading angle is determined by  $\hat{\phi}_1$ , and the initial values of the two velocities depend on the heading angle. During each iterative update, the system state  $X_t$  is initialized with 1000 particles, which are sampled from a Gaussian distribution. Each particle is assigned a weight of  $1/1000$ . The observation equations are

$$\begin{cases} v_{Dt} = \sqrt{v_{xt}^2 + v_{yt}^2} \cdot \sin(\phi_t + \frac{\alpha_T - \alpha_R}{2}) \cdot 2\sin(\frac{\alpha_T + \alpha_R}{2}) \\ \phi_t = \hat{\phi}_t \end{cases} \quad (15)$$

where  $\sqrt{v_{xt}^2 + v_{yt}^2}$  is equal to target moving speed  $v_{Ht}$ .

Particle weights are updated based on new observations. Particles closer to the new observation values are more important, with larger weights. We use Gaussian distributions to update the weights. Accordingly, we employ low-variance resampling to select particles, and obtain the system state value for the current iteration by taking the mean of the resampled particles.

Notably, the phase difference  $\Delta\varphi$  is an observation state, but it cannot be used to estimate system state  $X_t$  directly. Although the phase difference is a fine-grained parameter to capture the motion state and direction of the target, it deviates from the

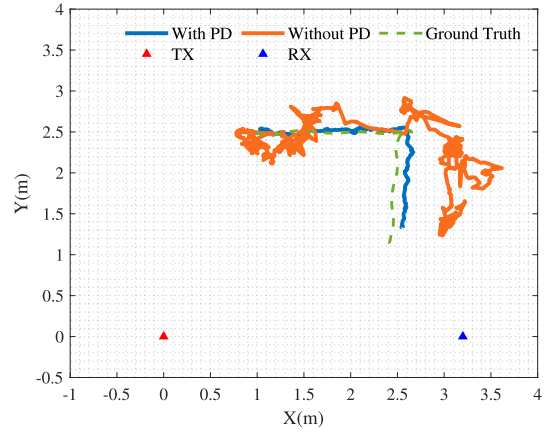


Fig. 13. Motion tracking result. The red and blue triangles are transceivers. Blue solid line is using phase difference information. Orange dashed line is without using phase difference information.

actual value. If it is directly involved in the operation, a large error will be introduced.

Fortunately, we found that the phase difference can be employed to remove low-confidence system state particles. The method flow is described as follows:

- 1) Calculate the slope  $slope_t$  of the phase difference from  $t-1$  to  $t$ , and then use the particle to calculate the slope  $slope'_t$  of the phase difference  $\Delta\varphi'_t$  in the same way.
- 2) If the product of the two slopes is greater than or equal to zero, the confidence of the particle is high, otherwise the confidence is low.

According to Fig. 12, we can derive the relationship between the AoA of the reflected signal and the target's position

$$\sin \alpha_R = \frac{y_t}{\sqrt{(x_t - x_r)^2 + (y_t - y_r)^2}} \quad (16)$$

based on (10) and (16), the formula for calculating the phase difference with particles can be written as

$$\Delta\varphi'_t = \frac{\pi \cdot y_t}{\sqrt{(x_t - x_r)^2 + (y_t - y_r)^2}} \quad (17)$$

where the  $x_r$  and  $y_r$  are the coordinates of RX.

The motion tracking result is shown in Fig. 13. The green solid line in the figure is the ground truth of the trajectory, the red and blue triangles represent the transmitter and receiver, respectively. The blue solid line and the orange dashed line are the trajectory estimation results based on particle filter, where the blue solid line represents the result using phase difference, and the orange dashed line represents the result without using phase difference. By comparison, it can be seen that the estimation result with the participation of phase difference has a smaller error, and without the participation of the phase difference, there is a large error in the estimation of the motion direction.

Ultimately, the coordinates of the target are updated with the particle filter, and the particles with low confidence are eliminated by the phase difference, which improves the estimation accuracy and realizes low-cost passive motion tracking based on one pair of Wi-Fi devices.



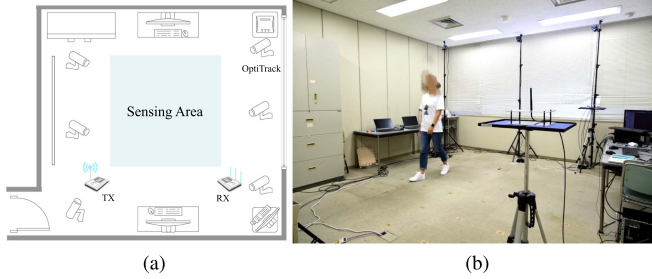


Fig. 14. Experimental environment setup. (a) The layout of the experimental room. (b) The experimental scene. The indoor environment.

## V. EVALUATION

In this section, we validate WiMT through a series of experiments. We first describe the experiment methodology, followed by assessing the Doppler velocity denoising and ZVIC identification accuracy in both stationary and moving states. Subsequently, we examine WiMT’s motion tracking performance under various experimental conditions.

### A. Experiment Methodology

1) *Implementation*: WiMT is evaluated in an indoor scenario depicted in Fig. 14, with a room size of approximately 5 m × 6 m. There are furniture and computer equipment in the room, which is a typical multipath propagation environment. Two PCs equipped with Intel 5300 network cards serve as TX and RX, where TX has a single antenna and RX has three antennas forming a uniform linear array. The spacing between RX’s antenna array is half the wavelength of the signal, and the placement angle between TX and RX is such that the line connecting TX’s antenna and RX’s second antenna is perpendicular to the RX array. The distance between RX and TX is set to 3.2 m, and the height is set to 1.3 m. Both TX and RX PCs run on Ubuntu 14.06, with CSI Tool [47] installed for transmitting and recording CSI. The devices operate in monitor mode, choosing channel 64 at 5.32 GHz, with a bandwidth of 20 MHz. TX sends 200, 400, and 1000 packets per second, that is, the packet sampling rate is 200 Hz, 400 Hz, 1000 Hz. The CSI data are processed offline using MATLAB 2022a, rather than in real time.

2) *Data Collection*: We collect three CSI datasets to separately evaluate the performance of the ZVIC algorithm in denoising, motion/stationary state identification, and the motion tracking accuracy of the WiMT system. The quantity of CSI samples included in each dataset is distinct, and due to variances in time and sampling rates, the number of packets in each sample also differs. For ease of distinction, we label them as dataset1, dataset2, and dataset3.

Dataset1 contains 50 CSI samples, each with a duration of 10 s. We ask volunteer to remain stationary at three different locations while collecting CSI data at three sampling rates. This dataset is used to assess the performance of the ZVIC algorithm in eliminating random noise from Doppler velocities.

Dataset2 consists of 20 CSI samples, with the first seven groups having a 40 s duration, while the remaining groups have a 20 s duration, all sampled at 400 Hz. We ask volunteer to sit

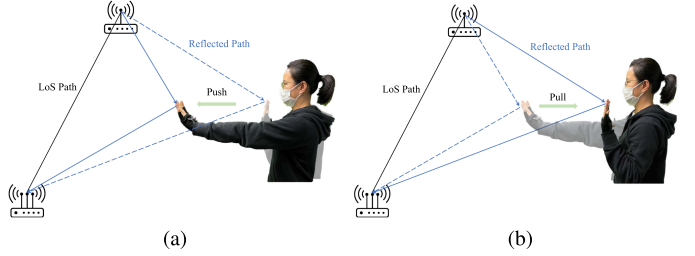


Fig. 15. Push-pull action. (a) Push. (b) Pull.

along the extended line of the LoS path midpoint and perform push-pull actions along this extension. The push involves moving the hand toward the LoS path, while the pull is the opposite, as shown in Fig. 15. In the 40 s data, 6–8 push-pull actions are executed, and in the 20 second data, 2–4 push-pull actions are performed. After each push or pull, volunteer pauses for 1–3 s, with the time ratio of moving to stationary states approximately 1:2.5. This dataset evaluates the ZVIC algorithm’s ability to identify moving and stationary states.

Dataset3 comprises 203 CSI samples, documenting volunteers walking along three indoor trajectories (line-shaped, L-shaped, and rectangle-shaped). We place labels on the ground, depicting walking trajectories, where the starting point of each trajectory is fixed and known in the trace estimation. When the target walks along L-shaped and rectangle-shaped trajectories, she/he pauses for 1–3 s at each turning point, where the direction changes, then continues walking. We collect 189 line-shaped trajectory data from four volunteers at three different sampling rates and nine L-shaped and five rectangular trajectory data from one volunteer at a 400 Hz sampling rate. This dataset is employed to evaluate the motion tracking accuracy of the WiMT system.

3) *Ground Truth*: We utilize the millimeter-level precision OptiTrack system to record the ground truth. When collecting dataset1 and dataset3, markers are placed on the target’s shoulder, and when collecting dataset2, markers are placed on the gloves worn by the target. The OptiTrack system records the marker coordinates at a 100 Hz sampling rate. To calculate the accuracy, we employ interpolation to align the lengths of the ground truth and the WiMT system outputs.

### B. Performance of ZVIC Algorithm

1) *Comparison with state-of-the-art*: As mentioned earlier, there is random noise in the Doppler velocity calculated by the WiTraj method when the target is stationary. In order to evaluate the performance of the ZVIC algorithm in removing noise, we utilize data from dataset1 to compute Doppler velocities using both WiTraj and ZVIC algorithms. We calculate the mean, standard deviation, and mean square error of the Doppler velocities. The result is shown in Table I. Although the mean Doppler velocity calculated by the WiTraj method is small, which is  $-0.312$  m/s, the standard deviation is 2.503 m/s, and the mean square error is 6.360. The Doppler velocity calculated by the ZVIC algorithm has a mean of 0 m/s, a standard deviation of 0.039 m/s, and a mean square error of 0.002. The comparison

TABLE I  
COMPARING WITRAJ AND ZVIC DOPPLER VELOCITIES

| Method | Mean   | Standard Deviation | Mean Square Error |
|--------|--------|--------------------|-------------------|
| WiTraj | -0.312 | 2.503              | 6.360             |
| ZVIC   | 0.000  | 0.039              | 0.002             |

<sup>1</sup> The unit of mean and standard deviation is  $m/s$ .

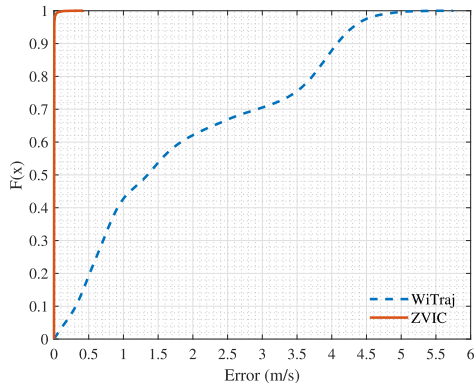


Fig. 16. Distribution on CDF of target velocity estimation error. The target is stationary at three different positions and 50 sets of CSI data are recorded at three different sampling rates.

shows that ZVIC effectively removes the noise in the Doppler velocity when the target is stationary.

In addition, in order to evaluate the influence of the estimation error of the Doppler velocity  $v_D$  on the target velocity  $v_H$ , we use the above CSI data to calculate the target velocity. As we know, the velocity should be zero when the target is stationary. We put  $v_D$  calculated by WiTraj and ZVIC into (10), respectively, to obtain  $v_H$  of target, where  $\phi$  equals to  $\pi/2$ , the coordinates of TX and RX are (0,0) and (3.2,0), the initial position of target is (1.6,1). We take the absolute value of all calculation results and make a difference with the ground truth, calculate the target velocity estimation error, and show the cumulative distribution function of the error in Fig. 16.

As shown in Fig. 16, the error of target velocity calculated by ZVIC is smaller than WiTraj. The median error of ZVIC is about 0 m/s and the WiTraj is 1.3476 m/s, the mean error of ZVIC and WiTraj are 0.0011 m/s and 1.8682 m/s, respectively. It can be seen from the results that the noise in the Doppler velocity calculated by the WiTraj method is not negligible, which introduces a large error to the estimation of the moving velocity and the ZVIC algorithm can effectively remove these noises.

2) *Moving and Nonmoving State Identification Accuracy:* We evaluate the recognition accuracy of ZVIC algorithm for nonmoving and moving states of target using dataset2. Using the ZVIC algorithm, we identify the target's moving and stationary states in each CSI sample. Then, we calculate the velocities by differentiating the target positions recorded by Optitrack, obtaining the ground truth for motion and stationary states, and compute the recognition accuracy of the ZVIC algorithm. The confusion matrix of nonmoving and moving state is shown in Fig. 17. The recognition accuracy of nonmoving and moving states are 92.6% and 85.2%, respectively.

|        |            |            |        |
|--------|------------|------------|--------|
| Actual | Non-moving | 0.926      | 0.074  |
|        | Moving     | 0.148      | 0.852  |
|        |            | Non-moving | Moving |
|        |            | Predicted  |        |

Fig. 17. Confusion matrix of moving and nonmoving state identification.

3) *Impact of Subcarrier:* We test the accuracy of ZVIC in removing Doppler velocity noise on different subcarriers. In Fig. 18(a), the  $x$ -axis represents the phase difference index, where "1" is the phase difference of CSI between the first antenna and third antenna, "2" denotes the first antenna and second antenna, "3" indicates the second antenna and third antenna. The  $y$ -axis is the error rate, and the  $z$ -axis is the subcarrier index. It can be seen that the identification error rates are distinct using different subcarrier and phase difference combinations. The performance of the 14th to 21st subcarriers is the most stable, and the error rate is zero in different phase difference combinations. The error rates of the first to seventh subcarriers are relatively large when PDI are 2 and 3, and the largest error rate is the sixth subcarrier when PDI is 3, which is 57.6%.

4) *Impact of Moving Direction:* We employ ten data samples with a sampling rate of 200 Hz in dataset1 to analyze the velocity estimation errors of ZVIC method in five moving directions, and compare with the WiTraj method. In Fig. 18(b), the  $x$ -axis is the moving angle, and the  $y$ -axis is the mean square error of the velocity  $v_D$ . It can be seen that the error gradually decreases with the increase of the moving angle, the error of ZVIC are distributed between 0–0.03, and the error of WiTraj are distributed between 0.5–2. The ZVIC has stable denoise performance in different moving directions.

5) *Impact of Sampling Rates:* In order to appraise the denoise performance of ZVIC under different sampling frequencies, we utilize 30 data samples, calculate the mean square error of the velocity  $v_D$  at 200 Hz, 400 Hz, and 1000 Hz, and compare with WiTraj. As shown in Fig. 18(c), the  $x$ -axis is the sampling rate, and the  $y$ -axis is the mean square error. As the sampling rate increases from 200 to 1000 Hz, the error of WiTraj also floats from 0.3 to 20 accordingly. In contrast, the error of ZVIC does not change with the sampling rate, and fluctuates stably between 0 and 0.03. Similarly, the ZVIC has stable denoise performance in different sampling rates.

### C. Performance of WiMT

We evaluate WiMT's motion tracking performance using the line-shaped trajectory data from dataset3. The line-shaped

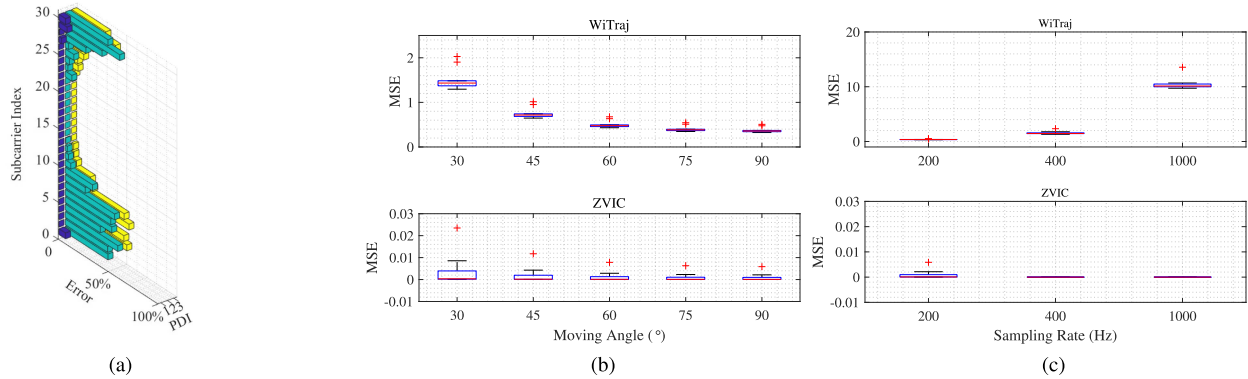


Fig. 18. Error distribution under different experimental settings. (a) x-axis represents the phase difference index, y-axis is the error rate, and z-axis is the subcarrier index. (b) x-axis is the moving angle and y-axis is the mean square error of Doppler velocity. (c) x-axis is the sampling rate and y-axis is the mean square error of Doppler velocity. (a) Different subcarriers. (b) Different moving directions. (c) Different sampling rates.

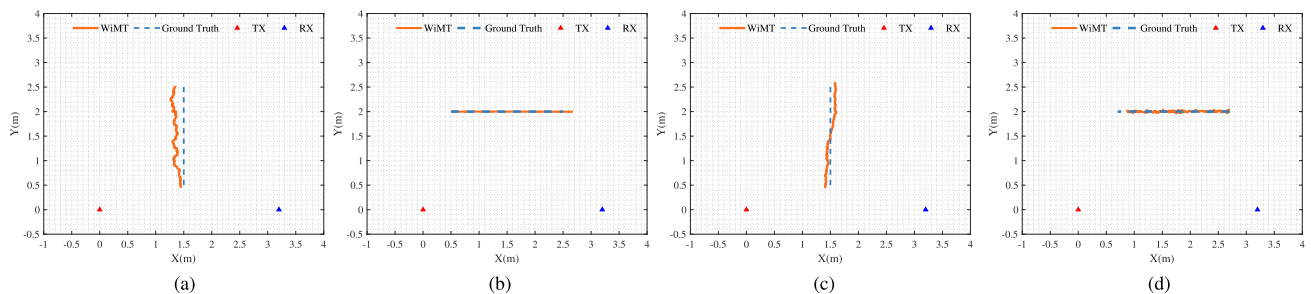


Fig. 19. Line-shaped trajectories. (a) and (c) are trajectories along the normal direction. (b) and (d) are trajectories along tangent direction. (a) Away LoS. (b) Towards RX. (c) Close LoS. (d) Towards TX.

trajectory includes four walking directions, namely away LoS, close LoS, toward RX, and toward TX. Close LoS and away LoS are directions perpendicular to the LoS path between RX and TX, while toward RX and toward TX are directions parallel to the LoS path. WiMT processes each CSI sample and outputs the target's position coordinates accordingly. The tracking error of WiMT is calculated as the Euclidean distance between the output position coordinates and the ground truth. The trajectory estimation results are shown in Fig. 19. The red and blue triangles in the figure represent TX and RX, respectively. The orange solid lines are the results of WiMT estimation, and the blue dotted lines are the ground truth. Fig. 19(a) and (c) are walking along the normal direction, Fig. 19(b) and (d) are walking along the tangent direction. To evaluate the tracking accuracy, we compute the Euclidean distance between each point estimated by WiMT and the ground truth.

We evaluated the overall performance of WiMT with median error, mean error, and 90th error of 7.28 cm, 9.25 cm, and 19.92 cm, respectively, the percentage of median error to walking distance is 3.64%. Then, we evaluate the four walking directions separately, as shown in Fig. 20 and Table II. The median error of the four directions is 6.02–11.52 cm, the 90th error is 16.66–22.90 cm, and the percentage of the median error to the walking distance is  $< 6.0\%$ .

1) *Impact of Moving Distance*: Since the line-shaped trajectory has a length of only 2 m, we aimed to assess WiMT's

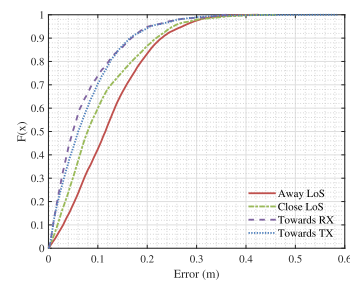


Fig. 20. CDF of line-shaped.

TABLE II  
TRACKING ERROR OF DIFFERENT MOVING DIRECTIONS

| Direction  | Median Error | Mean Error | 90 <sup>th</sup> Error | Percentage |
|------------|--------------|------------|------------------------|------------|
| Away LoS   | 11.52        | 12.35      | 22.90                  | 5.76%      |
| Close LoS  | 7.92         | 10.17      | 21.81                  | 3.96%      |
| Towards RX | 4.93         | 7.19       | 16.79                  | 2.47%      |
| Towards TX | 6.02         | 7.75       | 16.55                  | 3.01%      |

<sup>2</sup> The unit of error is *cm*.

performance on longer walking distances and more complex trajectories by collecting L-shaped and rectangle-shaped trajectories, with respective lengths of 3 and 6 m. WiMT's estimation results for both trajectories, as well as the ground truth, are presented in the Figs. 21 and 22. The red and blue triangles in the figure represent TX and RX, respectively. The orange solid lines

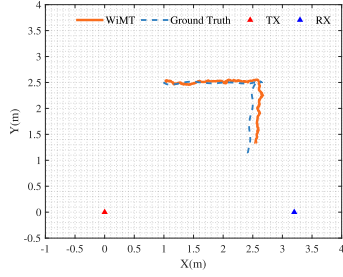


Fig. 21. L-shaped trajectory.

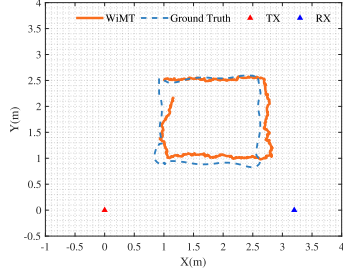


Fig. 22. Rectangular trajectory

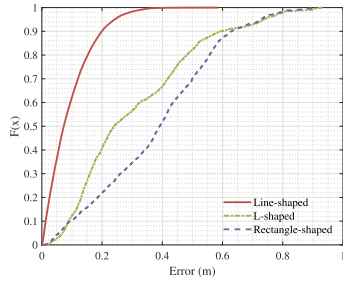


Fig. 23. CDF of three trajectories.

are the results of WiMT estimation, and the blue dotted lines are the ground truth. We have plotted the CDF graph of the WiMT tracking errors for three different types of trajectories, as shown in Fig. 23. The red line represents the line-shaped trajectory, the green line represents the L-shaped trajectory, and the purple line represents the rectangle-shaped trajectory. The median tracking errors for these three trajectories are 7.28 cm, 24.01 cm, and 39.14 cm, respectively.

2) *Impact of Sampling Rate:* We collect CSI with sampling rates of 200, 400, and 1000 Hz. Then, we calculate the tracking error of WiMT at different rates. In Fig. 24(a), the boxplot of median error at three sampling rate are printed. It can be seen that the performance of WiMT at different rates is stable, and the mean errors are distributed between 0.05 and 0.1 cm.

3) *Impact of Subject:* In order to test the tracking accuracy of WiMT for different walking targets, four volunteers were invited to participate in the experiments, including two males and two females, with an age distribution between 24 and 29, and a weight and height of 50 to 65 kg and 165 to 180 cm, respectively. From Fig. 24(b), we can learn that the error distribution of WiMT is basically the same, WiMT is barely affected by moving targets.

4) *Impact of Moving Direction:* In Fig. 24(c), 1 and 2 represent normal direction, 3 and 4 represent tangent direction, and WiMT performs better than normal direction in tangent direction. However, the errors in four directions are stably distributed between 1 and 25 cm.

## VI. DISCUSSION

In this section, we discuss the limitations of the proposed method and future work.

### A. Single Target Tracking

WiMT employs a pair of Wi-Fi devices to collect CSI data, proposes the ZVIC algorithm to obtain fine-grained Doppler velocities, and subsequently utilizes a particle filter to estimate target coordinates, achieving single-target motion tracking. Employing a pair of Wi-Fi devices to gather CSI and utilizing model-based approaches to analyze data for multitarget detection and tracking proves challenging. However, this will be the focus of our future work. We will attempt to leverage deep learning method to extract features from CSI phase, Doppler velocities, and phase differences to train models, ultimately achieving multitarget detection. In addition, we will explore the fusion of Wi-Fi with other sensors (e.g., smart insoles) to facilitate multitarget detection and tracking.

### B. Assumption of Known Initial Position

WiMT operates under the assumption that the initial position is known, initializing particles at the precise location. Accurately obtaining a target's starting position in real time using a pair of Wi-Fi devices is challenging in practical applications. In real-world applications, we will initially measure the coordinates of commonly used starting points such as beds, sofas, and doors. Before applying the particle filter, we will perform a coarse-grained position search, matching the position coordinates with parameters such as DFS, phase difference, and AoA to pinpoint the initial position of human movement. This location will then be fed into the particle filter for trajectory tracking. Moreover, if the filter does not achieve convergence, we will restart the process of coarse-grained position search.

### C. Coarse-Grained Direction Recognition

WiMT identify motion directions parallel and perpendicular to the LoS path with a pair of Wi-Fi devices collecting CSI. In practical applications, this can be used to sense forklift directions at intersections in warehouse aisles for collision prevention. Additionally, in gesture recognition, multiple gestures can be formed using four directions, eliminating the need for retraining when adding new gestures compared to deep learning methods. However, WiMT's performance significantly declines when tracking targets moving along circular or zigzag trajectories. Consequently, fine-grained motion direction recognition will be a focus of our future work. We intend to leverage the phase difference information in time and space, as well as Doppler shifts, to perceive changes in the direction of human movement

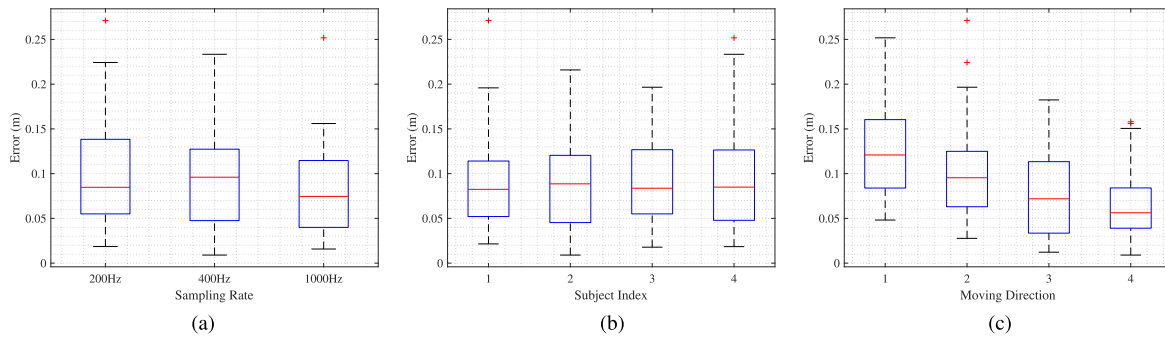


Fig. 24. Impact of different experimental setups. (b) Subjects are two males and two females, with age distribution 24 to 29, weight and height of 50 to 65 kg and 165 to 180 cm. (c) 1 and 2 represent normal direction, 3 and 4 represent tangent direction. (a) Impact of sampling rate. (b) Impact of subject. (c) Impact of moving direction.

in a more detailed manner, thereby achieving tracking of more complex trajectories.

#### D. Compared With Existing Work

In a nonmoving state, the recognition accuracy of WiMT is 92.6%, while WiTraj achieves 98.2%. For moving states, WiMT has an accuracy of 85.2%, with WiTraj at 87.8%. In terms of median tracking error for rectangular trajectories, WiMT is 39.13 cm, while WiTraj has an error of 40 cm. The errors as percentages of the total walking distance are 6.5% and 2.5% for WiMT and WiTraj, respectively. WiTraj requires three Wi-Fi links for implementation, with one link needing a special placement, whereas WiMT only utilizes one link. Although WiMT's precision is slightly inferior to that of WiTraj, it enhances usability and reduces deployment costs.

### VII. CONCLUSION

In this article, we present WiMT, a low-cost passive motion tracking system using a single pair of off-the-shelf Wi-Fi devices. We propose the ZVIC algorithm to remove the Doppler velocity noise, realize nonmoving and moving recognition with an accuracy of 92.6% and 85.2%, respectively. Moreover, we utilize the particle filter to track target's motion. To improve the tracking accuracy, phase difference information is employed to eliminate particles with low confidence. Thus, we achieve the trajectory tracking with median error of 7.28 cm for line-shaped trajectory, and the percentage of the median error to the walking distance is 3.64%. Furthermore, we set various experimental conditions to verify the robustness of WiMT. The results show that WiMT has stable recognition and tracking accuracy under different experimental conditions.

### REFERENCES

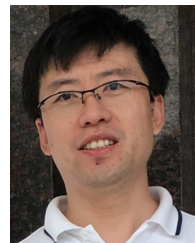
- [1] J. Berclaz, F. Fleuret, and P. V. Fua, "Robust people tracking with global trajectory optimization," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2006, vol. 1, pp. 744–750.
- [2] M. Yasir, S.-W. Ho, and B. N. Vellambi, "Indoor position tracking using multiple optical receivers," *J. Lightw. Technol.*, vol. 34, no. 4, pp. 1166–1176, Feb. 2016.
- [3] D. J. Schott et al., "Comparison of direct intersection and sonogram methods for acoustic indoor localization of persons," *Sensors*, vol. 21, 2021, Art. no. 4465.
- [4] W. T. Faulkner, R. Alwood, D. W. A. Taylor, and J. Bohlin, "GPS-denied pedestrian tracking in indoor environments using an IMU and magnetic compass," in *Proc. Int. Tech. Meeting Inst. Navigation*, 2010, pp. 198–204.
- [5] J. Wang, J. Xiong, H. Jiang, X. Chen, and D. Fang, "D-watch: Embracing "bad" multipaths for device-free localization with COTS RFID devices," *IEEE/ACM Trans. Netw.*, vol. 25, no. 6, pp. 3559–3572, Dec. 2017.
- [6] M. Merenda, L. Catarinucci, R. Colella, D. Iero, F. G. D. Corte, and R. Carotenuto, "RFID-based indoor positioning using edge machine learning," *IEEE J. Radio Freq. Identif.*, vol. 6, pp. 573–582, 2022.
- [7] M. Zhao et al., "Uloc: Low-power, scalable and cm-accurate uwb-tag localization and tracking for indoor applications," in *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2021, vol. 5, pp. 140:1–140:31.
- [8] M. Ridolfi, A. Kaya, R. Berkvens, M. Weyn, W. Joseph, and E. D. Poorter, "Self-calibration and collaborative localization for UWB positioning systems: A survey and future research directions," *ACM Comput. Surv.*, vol. 54, no. 4, pp. 1–27, 2021.
- [9] D. Oosterlinck, D. F. Benoit, P. Baecke, and N. V. de Weghe, "Bluetooth tracking of humans in an indoor environment: An application to shopping mall visits," *Appl. Geogr.*, vol. 78, pp. 55–65, 2017.
- [10] F. S. Danics, A. T. Cemgil, and C. Ersoy, "Adaptive sequential Monte Carlo filter for indoor positioning and tracking with Bluetooth low energy beacons," *IEEE Access*, vol. 9, pp. 37022–37038, 2021.
- [11] M. Kotaru, K. R. Joshi, D. Bharadia, and S. Katti, "Spotfi: Decimeter level localization using WiFi," in *Proc. ACM Conf. Special Int. Group Data Commun.*, 2015, pp. 269–282.
- [12] J. Xiong and K. Jamieson, "Arraytrack: A fine-grained indoor location system," in *Proc. Netw. Syst. Des. Implementation*, 2013, pp. 71–84.
- [13] X. Li et al., "Indotrack: Device-free indoor human tracking with commodity Wi-Fi," in *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2017, vol. 1, pp. 72:1–72:22.
- [14] K. Qian, C. Wu, Z. Yang, Y. Liu, and K. Jamieson, "Widar: Decimeter-level passive tracking via velocity monitoring with commodity Wi-Fi," in *Proc. 18th ACM Int. Symp. Mobile Ad Hoc Netw. Comput.*, 2017, pp. 1–10.
- [15] X. Li, S. Li, D. Zhang, J. Xiong, Y. Wang, and H. Mei, "Dynamic-music: Accurate device-free indoor localization," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2016, pp. 196–207.
- [16] J. Nan, G. D. Wei, Y. Yuling, and W. Lin, "Leveraging TPDV to direction estimation with a single pair of commodity devices," *IEEE Access*, vol. 8, pp. 62248–62260, 2020.
- [17] Y. Xie, J. Xiong, M. Li, and K. Jamieson, "xD-track: Leveraging multi-dimensional information for passive Wi-Fi tracking," in *Proc. 3rd Workshop Hot Topics Wireless*, 2016, pp. 39–43.
- [18] W. Wang, A. X. Liu, M. Shahzad, K. Ling, and S. Lu, "Understanding and modeling of WiFi signal based human activity recognition," in *Proc. 21st Annu. Int. Conf. Mobile Comput. Netw.*, 2015, pp. 65–76.
- [19] P. Bahl and V. N. Padmanabhan, "Radar: An in-building RF-based user location and tracking system," in *Proc. IEEE Conf. Comput. Commun. 19th Annu. Joint Conf. IEEE Comput. Commun. Soc.*, 2000, vol. 2, pp. 775–784.
- [20] C. Wu, J. Xu, Z. Yang, N. D. Lane, and Z. Yin, "Gain without pain: Accurate WiFi-based localization using fingerprint spatial gradient," *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 1, no. 2, pp. 1–19, 2017.
- [21] N. Singh, S. Choe, and R. Punmiya, "Machine learning based indoor localization using Wi-Fi RSSI fingerprints: An overview," *IEEE Access*, vol. 9, pp. 127150–127174, 2021.

- [22] A. T. Mariakakis, S. Sen, J. Lee, and K.-H. Kim, "Sail: Single access point-based indoor localization," in *Proc. 12th Annu. Int. Conf. Mobile Syst., Appl., Serv.*, 2014, pp. 315–328.
- [23] Y. Xie, Z. Li, and M. Li, "Precise power delay profiling with commodity Wi-Fi," *IEEE Trans. Mobile Comput.*, vol. 18, no. 6, pp. 1342–1355, Jun. 2019.
- [24] Z. Wang, J. A. Zhang, M. Xu, and J. Guo, "Single-target real-time passive WiFi tracking," *IEEE Trans. Mobile Comput.*, vol. 22, no. 6, pp. 3724–3742, Jun. 2023.
- [25] 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 *Proc. 16th Annu. Int. Conf. Mobile Syst., Appl., Serv.*, 2018, pp. 350–361.
- [26] Y. Zheng et al., "Zero-effort cross-domain gesture recognition with Wi-Fi," in *Proc. 17th Annu. Int. Conf. Mobile Syst., Appl., Serv.*, 2019, pp. 313–325.
- [27] D. Wu et al., "WiTraj: Robust indoor motion tracking with WiFi signals," *IEEE Trans. Mobile Comput.*, vol. 22, no. 5, pp. 3062–3078, May 2023.
- [28] J. Ding, Y. Wang, H. Si, S. Gao, and J. Xing, "3-D indoor localization and tracking for mobile target based on WiFi sensing," *IEEE Internet Things J.*, vol. 9, no. 21, pp. 21687–21701, Nov. 2022.
- [29] Y. Zeng, E. Yi, D. Wu, R. Gao, and D. Zhang, "A CSI-ratio model based house-level respiration monitoring system using cots WiFi devices," in *Proc. Adjunct Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput. Proc. ACM Int. Symp. Wearable Comput.*, 2019, pp. 354–357.
- [30] R. Gao et al., "Towards position-independent sensing for gesture recognition with Wi-Fi," *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 5, pp. 1–28, 2021.
- [31] D. Wu, D. Zhang, C. Xu, Y. Wang, and H. Wang, "Widir: Walking direction estimation using wireless signals," in *Proc. ACM Int. Joint Conf. Pervasive Ubiquitous Comput.*, 2016, pp. 351–362.
- [32] K. Niu, X. Wang, F. Zhang, R. Zheng, Z. Yao, and D. Zhang, "Rethinking doppler effect for accurate velocity estimation with commodity WiFi devices," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2164–2178, Jul. 2022.
- [33] Q. Zhou, J. Xing, W. Chen, X. Zhang, and Q. Yang, "From signal to image: Enabling fine-grained gesture recognition with commercial Wi-Fi devices," *Sensors*, vol. 18, 2018, Art. no. 3142.
- [34] H. Abdelnasser, M. Youssef, and K. A. Harras, "Wigest: A ubiquitous Wi-Fi-based gesture recognition system," in *Proc. IEEE Conf. Comput. Commun.*, 2015, pp. 1472–1480.
- [35] K. Qian, C. Wu, Z. Zhou, Y. Zheng, Z. Yang, and Y. Liu, "Inferring motion direction using commodity Wi-Fi for interactive exergames," in *Proc. CHI Conf. Hum. Factors Comput. Syst.*, 2017, pp. 1961–1972.
- [36] H. Wang, D. Zhang, Y. Wang, J. Ma, Y. Wang, and S. Li, "RT-Fall: A real-time and contactless fall detection system with commodity WiFi devices," *IEEE Trans. Mobile Comput.*, vol. 16, no. 2, pp. 511–526, Feb. 2017.
- [37] N. Yu, W. Wang, A. X. Liu, and L. Kong, "Qgesture: Quantifying gesture distance and direction with WiFi signals," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, pp. 51:1–51:23, 2018.
- [38] Y. Ma, G. Zhou, S. Wang, H. Zhao, and W. Jung, "Signfi: Sign language recognition using WiFi," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 2, pp. 23:1–23:21, 2018.
- [39] F. Wang, S. Panev, Z. Dai, J. Han, and D. Huang, "Can WiFi estimate person pose?," 2019, *arXiv:1904.00277*.
- [40] Y. Ren, Z. Wang, S. Tan, Y. Chen, and J. Yang, "Winect: 3D human pose tracking for free-form activity using commodity WiFi," *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 5, pp. 176:1–176:29, 2021.
- [41] R. Gao et al., "Towards robust gesture recognition by characterizing the sensing quality of WiFi signals," *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 6, pp. 1–26, 2022.
- [42] K. Niu, F. Zhang, X. Wang, Q. Lv, H. Luo, and D. Zhang, "Understanding WiFi signal frequency features for position-independent gesture sensing," *IEEE Trans. Mobile Comput.*, vol. 21, no. 11, pp. 4156–4171, Nov. 2022.
- [43] Q. Zhou, J. Xing, J. Li, and Q. Yang, "A device-free number gesture recognition approach based on deep learning," in *Proc. 12th Int. Conf. Comput. Intell. Secur.*, 2016, pp. 57–63.
- [44] Y. Zeng, D. Wu, J. Xiong, E. Yi, R. Gao, and D. Zhang, "Farsense: Pushing the range limit of WiFi-based respiration sensing with CSI ratio of two antennas," *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 3, no. 3, pp. 1–26, 2019.
- [45] C. Wu, Z. Yang, Z. Zhou, K. Qian, Y. Liu, and M. Liu, "Phaseu: Real-time LOS identification with WiFi," in *Proc. IEEE Conf. Comput. Commun.*, 2015, pp. 2038–2046.
- [46] S. Thrun, W. Burgard, and D. Fox, *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. Cambridge, MA, USA: MIT Press, 2005.
- [47] D. Halperin, W. Hu, A. Sheth, and D. Wetherall, "Tool release: Gathering 802.11 n traces with channel state information," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 41, no. 1, pp. 53–53, 2011.



**Wei Guo** (Graduate Student Member, IEEE) received the B.S. degree in communication engineering and the M.S. degree in electronic science and technology from Yanshan University, in 2016 and 2019, respectively. She is currently working toward the Ph.D. degree in computer and information systems with the Graduate School of Computer Science and Engineering, University of Aizu, Japan.

Her research interests include intelligent sensing systems, signal processing, and sensor fusion for mobile computing.



**Lei Jing** (Member, IEEE) received the Ph.D degree in computer science and engineering from University of Aizu, Aizuwakamatsu, Japan, in 2008.

He is a Senior Associate Professor at the University of Aizu, where he is leading the Human Sensing Laboratory, and the team dedicated to human motion capture and recognition. He is particularly interested in hand motion sensing with multimodal sensor fusion methods, and the applications include activity abnormal detection, sign language recognition, and smart agriculture.

He has published over 100 papers and holds six patents in related areas. He is a member of IEEE since 2010.