Analysis of Non-Intrusive Hand Trajectory Tracking by Utilizing Micro-Doppler Signature Obtained From Wi-Fi Channel State Information

NOPPHON KEERATIVORANAN, (Member, IEEE), PANAWIT HANPINITSAK, (Member, IEEE), KENTARO SAITO, (Member, IEEE), AND JUN-ICHI TAKADA, (Member, IEEE)
Department of Transdisciplinary Science and Engineering, Tokyo Institute of Technology, Tokyo 152-8552, Japan
Corresponding author: Nopphon Keerativoranan (nopphon.keerativoranan@ap.ide.titech.ac.jp)

ABSTRACT Envision a ubiquitously device-free motion sensing, this work focuses on the analysis of Wi-Fi-based hand gesture trajectory tracking by utilizing the Doppler frequency obtained from channel state information (CSI). Because human limb movement generates variant micro-Doppler signatures produced by the contribution of different parts of the hand and arm surfaces to the spectrum, an estimation technique is proposed to extract the temporal profile of the hand-only Doppler signature. With a set of Doppler profiles from different pairs of Wi-Fi antennas, hand trajectory can be traced by exploiting the multi-static Doppler radar model. The Kalman filter (KF) was applied to mitigate the accumulated noise from the recursive process of trajectory estimation. To validate the proposed method, a human limb model was developed to simulate the deterministic movement of a hand gesture by exploiting the non-rigid motion of the robotic arm. The electromagnetic wave scattered from the human limb model was computed using a physical optics (PO) approximation to simulate time-variant CSI. In the experiment, the hand Doppler signature could be successfully extracted from the spectrum with an error of less than 4 Hz at the 90th percentile of the CDF. With the extracted profiles, the trajectories of a square and M-shaped gesture were successfully traced, albeit with moderate trajectory offset of $10\text{°} - 20\text{°}$. Measurement conducted in a meeting room with commodity Wi-Fi devices installed on laptops also confirmed the tracking capability of the proposed framework.

INDEX TERMS Channel state information, micro-Doppler, motion sensing, non-rigid motion, trajectory tracking, Wi-Fi.

I. INTRODUCTION
The advances in communication technology in recent years have drastically influenced the mean of interaction between humans and computing devices. Motion sensing is one of the emerging technologies that enables communication through certain gestures or actions. Various commercial non-intrusive motion sensing sensors such as Kinect, Orbbec, and Leap Motion can achieve promising accuracy and precision by utilizing vision sensors. However, their performance could sharply deteriorate in the absence of line of sight (LoS), and when the user located farther away from the sensor, due to the high sensitivity to obstacles and the limited size of the vision sensor’s lens respectively [1]–[4]. In addition, a privacy concern due to images captured by the vision sensor would limit the usage of motion sensing in certain private locations such as restrooms or office meeting rooms [5], [6]. To avoid these restrictions, there are applications such as the virtual reality gear controller, which utilizes many body-equipped sensors to capture human motions via an inertial measurement unit sensor. However, employing such intrusive wearable sensors places an extra burden on users, especially in the context of healthcare-related motion sensing applications [1], [7], [8]. To enable non-intrusive motion sensing in these scenarios, many studies have investigated the applicability of the radio frequency (RF) signal for motion sensing [1], [9]–[11]. This is because RF-based motion sensing could theoretically operate in a Non-LoS (NLoS) scenario, has fewer privacy concerns, and can operate in a Non-LoS (NLoS) scenario, has fewer privacy concerns.
of the CSI, the performance should be more robust to the trajectory tracking does not rely on a unique temporal pattern division multiplexing (OFDM) radar [26]–[28]. Because the concept of Wi-Fi-based hand gesture trajectory estimation. FIGURE 1. Concept of Wi-Fi-based hand gesture trajectory estimation.

Concerns, and also has a larger sensing area owing to its longer wavelength. Due to its low cost and ubiquity, Wi-Fi has been widely studied for motion sensing applications [8], [12], [13]. According to the literature, one of the typical methods of detecting movement is through an analysis of the Doppler frequency shift extracted from the channel state information (CSI) of Wi-Fi chips.

In the context of Wi-Fi-based hand motion sensing, most studies in the literature generally apply various pattern matching techniques to recognize a certain set of hand gestures by analyzing the temporal patterns of the Doppler frequency. For example, [3], [14]–[16] were able to distinguish a set of hand gestures moving in different directions, which could enable the performance of simple tasks in contactless interaction with computing devices. Similarly, [17], [18] could successfully classify a set of finger gestures [9], [19] by recognizing a distinct pattern of keystrokes while typing on keyboard and touching a smartphone, respectively. Although their works have empirically demonstrated the feasibility of motion analysis applied to even a small movement of a human limb, they rely on the strong assumption that each gesture is always associated with a unique pattern in the CSI streams. Unfortunately, the temporal pattern of CSI produced by the Doppler frequency depends on the position of the Wi-Fi transmitter (Tx) and receiver (Rx) relative to the user location as opposed to the recognition approach. Typically, a moving target object is supposedly far away from sensors and thus can be assumed as a point object in the radar tracking problem [20]. Therefore, the entire object is assumed to produce only a single Doppler component at a particular time window. However, the point object assumption is not always valid for human motion sensing, especially hand motion, because multiple Doppler components could be observed from these movements as shown in previous works on CSI-based human motion sensing [29]–[31]. In the case of hand motion, various Doppler components are generated because each part of the human limb is moving at a different speed due to non-rigid motion [32]. Accumulated in CSI, these components are expressed as a micro-Doppler effect in the Doppler power spectrum [31]–[33]. Instead of the single peak spectrum as in the point object assumption, the Doppler spectrum spreads out and contains multiple peaks owing to the presence of multiple Doppler components. However, this effect has not been addressed in previous works on device-free hand motion tracking with Wi-Fi CSI [34], [35]. Despite their useful results in trajectory estimation, the micro-Doppler effect was not taken into account in their systems. Therefore, it was not possible to explain whether the selected Doppler profile utilized for tracking corresponded to the hand segment or not. Without the support of a theoretical model, the usability and performance of the tracking framework in different setups and scenarios are difficult to justify. For example, an indoor human tracking system in [29], [36] utilized the highest peak in the Doppler power spectrum, which represents the movement of a human torso [32], [37]. Therefore, it supports the usability of the human tracking system.

Recognizing the lack of a supporting model in hand gesture trajectory tracking, we aim to develop a passive hand trajectory tracking framework based on the micro-Doppler characteristics of the hand gesture. To fulfill this goal, our work has led to the following three contributions. First, a deterministic Wi-Fi CSI model that can simulate the micro-Doppler effect caused by the physical movement of a hand gesture has been developed by utilizing the dynamic model of a robot arm and electromagnetic (EM) scattering theory. This model allows us to study the temporal pattern of Doppler components produced by different segments of our limb while performing hand gestures. The findings led to the second contribution, an algorithm to pinpoint the micro-Doppler signature of the hand segment from the Doppler spectrum. The last contribution is the formulation of hand gesture trajectory tracking by utilizing a multi-static Doppler radar system. The capability of our CSI model to simulate the micro-Doppler effect will be verified with the measurement result. The trajectory estimation performance will be discussed and compared with similar tracking techniques used in CSI-based motion tracking in the context of gesture distortion. The results from the measurements using commodity Wi-Fi devices will be further discussed to confirm the applicability of the hand gesture tracking framework.
structure consists of three joints located at the wrist, elbow, and shoulder. These joints then govern the position and orientation of the hand, forearm, and upper arm segments. Considering the human limb structure, the 6R-PUMA robot arm model depicted in Fig. 2 could be utilized to simulate human limb movement owing to its simplicity and configuration similarity.

A few constraints must be imposed to apply the 6R-PUMA model to the simulation framework of the human limb movement. Suppose a particular column vector \( \mathbf{p}_k \) represents a 3D temporal position at timestamp \( k \) in Cartesian coordinates where the subscript indicates either a human limb segment or a joint. First, the position of the shoulder joint, defined as \( \mathbf{p}_s \), is supposedly fixed and can rotate in both azimuth \( \phi_s(k) \) and co-elevation \( \theta_s(k) \) angles, which are equivalent to the rotation of the first and second joint in 6R-PUMA. With this setup, the length between these two joints \( l_0 \) is reduced to zero. Second, the elbow joint located at \( \mathbf{p}_e(k) \) is movable but can rotate vertically by \( \gamma_e(k) \). Finally, the movement reference or the end-effector represents the temporal position of the wrist \( \mathbf{p}_w(k) \), which is supposedly a non-rotatable part in this model. This last joint controls the overall hand gesture trajectory, which can be recursively determined by

\[
\begin{align*}
\mathbf{p}_w(k + 1) &= \mathbf{p}_w(k) + \mathbf{v}_w(k)T_s \\
\end{align*}
\]

where \( T_s \) and \( v_w(k) \) are a time sampling between consecutive snapshots and the wrist velocity, respectively. Here, the wrist velocity follows the Gaussian velocity profile described previously. \( \mathbf{p}_e(k), \gamma_e(k), \phi_\text{elbow}(k), \) and \( \theta_\text{elbow}(k) \) can then be obtained from the inverse kinematic solution of 6R-PUMA described in [43]. With this model, any position on the upper arm and forearm with lengths \( l_u \) and \( l_f \) can accordingly be easily computed.

\section*{B. SIMPLIFIED MODEL OF HUMAN LIMB SURFACE}

By employing the velocity profile and the 6R-PUMA robot arm, the general non-rigid body motion of the human limb can be simulated. However, the shape of the human limb has not yet been included in the model. Because the characteristics of the micro-Doppler effect are also dependent on the shapes of scattered objects, as mentioned in [32], it is necessary to provide this non-rigid skeleton motion with an appropriate surface shape such that the micro-Doppler pattern can be deterministically simulated. Although modeling the shape of the human limb appears complicated at first glance, it can be simplified by regarding the entire structure as a composition of simple geometrical shapes. Hence, in our model, a limb is broken down into an upperarm (ua), elbow (eb), forearm (fa), and hand (hd) segments. Each segment is represented by a cylinder, sphere, truncated cone, and cuboid, respectively, as shown in Fig 3(a). The sizes of the arm segments and elbow are defined by radii \( r_1 \) and \( r_2 \), and the lengths \( l_u \) and \( l_f \). Similarly, the \( l_{hd}, w' \), and \( w'' \) parameters define the size of the hand, as shown in Fig. 3(a). Moreover, the lengths of the limb segments can be expressed in terms of the human height.
and forearm utilize the joint movement can refer to the joints partially relax the hand orientation from the forearm segment. This operation is introduced to maintain a smooth alignment. This operation is introduced to does not change the azimuth angle from that of the forearm to the forearm segment, respectively. The rotation matrix, however, the difference between the co-elevation angle of the hand and L as stated in [44]. To be specific, \(l_{up}, l_{fa}, \) and \(l_{hd}\) can be determined as 0.186L, 0.146L, and 0.108L, respectively.

Each limb segment requires reference vectors to orientate and translate them into the aligned position, as shown in Fig. 3(b). They can be computed based on the direction of the upper arm, forearm, and hand segments at a timestamp \(k\), and are defined in terms of the following unit vectors:

\[
\hat{a}_{fa}(k) = \frac{p_{fa}(k) - p_{hd}(k)}{\|p_{fa}(k) - p_{hd}(k)\|} \quad (2)
\]

\[
\hat{a}_{fa}(k) = \frac{p_{fa}(k) - p_{sh}(k)}{\|p_{fa}(k) - p_{sh}(k)\|} \quad (3)
\]

\[
\hat{a}_{hd}(k) = \hat{a}_{hd}(k) \hat{a}_{fa}(k) \quad (4)
\]

where \(\|\cdot\|\) and \(\gamma_{hd}(k)\) represent the Euclidean norm and the rotation matrix that vertically orients \(\hat{a}_{hd}(k)\) away from the \(\hat{a}_{fa}(k)\) direction by \(\gamma_{off}(k)\), respectively. This angle defines the difference between the co-elevation angle of the hand and forearm segment, respectively. The rotation matrix, however, does not change the azimuth angle from that of the forearm to maintain a smooth alignment. This operation is introduced to partially relax the hand orientation from the forearm segment. For the translation correction, the upper arm and hand segment can refer to the joints \(p_{sh}\) and \(p_{hd}\). Similarly, the elbow and forearm utilize the joint \(p_{fa}\). Finally, the human limb represented by a set of nearly uniform points on the surface can be visualized as depicted in Fig. 4.

III. TIME-VARIANT CSI OF MOVING LIMB MODEL

CSI has been measured in various wireless communication systems, including Wi-Fi, to assess the spatial and temporal conditions of the wireless channel. Multiple-input multiple-output (MIMO) transmission allows the Wi-Fi system to utilize multiple channels, thus effectively improving the quality of data communication through the use of beamforming, spatial multiplexing, and diversity techniques [23], [45]. In this study, the usage of CSI is further explored for motion detection and tracking through an analysis of the Doppler frequency. To be specific, this section describes the development of a deterministic CSI model that can physically simulate the micro-Doppler effect produced during a hand gesture. The model will be used to study the characteristics of hand Doppler components while performing hand gestures.

A. CSI MODEL BASED ON KINEMATIC MODEL OF HUMAN LIMB

In the multipath propagation scenario, CSI can be physically described as a combination of the wireless channel gain from multiple propagation paths where RF signals have propagated between Tx and Rx [45]. The RF signal may interact with either moving or static objects along their propagation paths, imparting the time-variant (TV) and time-invariant (TIV) properties to the channel. Considering a Wi-Fi system with \(M\) number of Tx and \(N\) number of Rx antennas, the CSI corresponding to the channel between the \(m\)-th Tx and the \(n\)-th Rx antenna can be modeled according to [30], [46] as

\[
H_{nm}(k, f) = H_{nm}^{TV}(f) + \mathcal{N}_{nm}(k, f) + \mathcal{N}_{nm}(k, f) \quad (5)
\]

where \(\mathcal{N}_{nm}(k, f)\) and \(f\) represent additive white Gaussian noise (AWGN) and the subcarrier frequency, respectively. For the sake of simplicity, the propagation path with respect to the TIV component is restricted to only the dominated component of the LoS static path visualized as a sine wave in Fig. 4, and can be written as

\[
H_{nm}^{TV}(f) = \frac{\lambda}{4\pi d_{LoS}} e^{-j\frac{2\pi}{\lambda} d_{LoS}} \quad (6)
\]
where $d_{\text{LoS}}$ and $\lambda$ are the distances between the Rx and Tx antennas located at $p_R$ and $p_T$, respectively, followed by a subcarrier wavelength. The TV component, on the other hand, should contain all the propagation paths scattered from the surface points of the moving human limb model described in Section II. Considering the geometry of the limb object with respect to the antenna positions, there exists a shadowed region on the surface whose LoS of those surface points are blocked. Depicted as black points in Fig. 4, they are assumed to generate no Doppler frequency regardless of how fast the surface points were moving. Suppose a set $\Omega(k)$ consists of surface points located only in the illuminated region (non-shadowed) at timestamp $k$ that are visualized as colored points in Fig. 4, the dynamic component of CSI can be expressed as

$$H_{nm}^{\text{TV}}(k, f) = \sum_{i \in \Omega(k)} a_{nm}^{(i)}(k, f) e^{-j \frac{2\pi}{\lambda} d_{nm}^{(i)}(k)}$$

(7)

where $d_{nm}^{(i)}(k)$ are the distances from the $i$-th surface point to the positions of the Tx and Rx antenna, respectively. The complex attenuation $a_{nm}^{(i)}(k, f)$ represents the propagation loss between the Tx antenna and the $i$-th surface point, the RF scattering loss at the surface, and the propagation loss from the surface point back to the Rx antenna. Because the RF scattered wave can be calculated by using the solution from EM theory, a physical optics (PO) approximation [47] has been applied to estimate the TV component of CSI. To apply our limb model consisting of surface points to the PO framework, a small surface area surrounding the $i$-th surface point has been constructed. Here, each surface area is represented by a circular mesh to avoid the need for a surface reconstruction algorithm. The detailed solution and validation of this PO based on a circular mesh can be found in the report in [48]. Therefore, the term $H_{nm}^{\text{TV}}(k, f)$ is equivalent to the total scattered electric field at the Rx obtained from the PO solution normalized by the emitted electric field from the Tx antenna.

By utilizing the PO, the CSI model could deterministically simulate the wireless channel and micro-Doppler effect based on EM scattering theory, allowing us to fully investigate the behavior of the hand Doppler signature that will be further discussed in the next section.

**B. MICRO-DOPPLER EFFECT PRODUCED BY HAND GESTURE**

In general, the temporal change in the propagation distance of an RF signal when it interacts with a moving point object is proportional to the Doppler frequency shift [1], [20], [32]. In the case of a non-rigid body, each body segment is likely moving with various speeds and therefore generates different Doppler shifts. Considering the human limb model described in Section II, the micro-Doppler frequency $f_{d(nm)}^{(i)}$ with respect to the $i$-th surface point can be expressed as the phase term of (7) according to the bistatic Doppler radar [32]:

$$f_{d(nm)}^{(i)}(k) = \frac{1}{\lambda} \left( \frac{d}{dt} d_{nm}^{(i)}(k) + \frac{d}{dt} d_{Rn}^{(i)}(k) \right)$$

(8)

For clarification, the distribution of the micro-Doppler frequency on the surface of a human limb with size parameters summarized in Table 1 was simulated assuming propagation of the RF signal at 5.31 GHz. The continuous variation in the Doppler frequency distribution can be observed in Fig. 5.

| Parameter                     | Values       |
|-------------------------------|--------------|
| $\ell_1, \ell_2$              | 180, 7, 3.5 cm |
| $w_{x(h)}, w_{y(d)}$          | 9, 18 cm     |
| Hand co-elevation angle       | 30°          |
| Average $p_{\ell}(k)$ resolution | 5 mm         |
| Snapshot sampling period      | 0.4 ms       |
| Motion segment period         | 1.2 s        |

**FIGURE 5.** Spatial distribution of bistatic Doppler frequency along the surface of a moving limb.

A simple way of obtaining the instantaneous Doppler profile is to take the first derivative of the phase term at each timestamp. However, this method works only when there is a single component of the Doppler frequency, which is not the case in this scenario. Hence, a short-time Fourier transform (STFT) with Gaussian window $g(k)$ is applied to decompose the micro-Doppler components within a short-time window. Hence, the Doppler-variant transfer function is computed by

$$B(\delta, f_0, f) = \sum_{k=-\infty}^{\infty} H_{nm}^{\text{TV}}(k, f) g(k-C\delta) e^{-j2\pi f_0 k T_s}$$

(9)

where $\delta$ and $C$ are the index of the short-time window and the hop size between consecutive STFTs, respectively. For brevity, the $nm$ subscript in the above equation and from here onward is omitted in this section. It should be noted that the
TV component can be easily obtained from (5) if the TIV component is given. Because this static component can be assumed to be a constant within a short period, it can be estimated by taking the average of CSI weighted by \( g(k) \). Consequently, we can obtain the Doppler spectral density from (9) with the following transformation, according to [45]:

\[
P_B(\delta, f_d) = \sum_{\tau \in \mathcal{T}} \left| \sum_{f \in \mathcal{F}} B(\delta, f_d, f)e^{j2\pi f \tau} \right|^2
\]

where \( \tau, \mathcal{T}, \) and \( \mathcal{F} \) are a propagation delay bin, a set of propagation delay bins within the range of the TV propagation paths, and a set of subcarrier frequencies, respectively. The benefit of restricting the computation to a certain delay window is to exclude the undesired noise power outside the range of RF propagation. Here, the areas with relatively higher power correspond to various components of the Doppler frequency produced by the movement of each surface point on the limb. According to the closed-form solution of the STFT with the Gaussian window provided in [49], the Doppler power is proportional to the complex attenuation of the TV propagation in (7).

The time-Doppler power spectrum in (10) will be applied to CSI to study the behavior of the hand Doppler signature in the presence of the micro-Doppler effect and thereby develop a method to extract the target hand Doppler profile from the spectrum. A detailed explanation of this will be given in the next section.

### IV. EXTRACTION OF HAND DOPPLER PROFILE

Because this work focuses on the tracking of hand gestures, the unique Doppler signature produced by the hand segment is essential for tracing the hand trajectory. To fulfill this goal, the characteristics of the hand Doppler signature will be examined based on the distribution of the TV path gain derived from the complex attenuation and Doppler frequency produced by CSI. Based on the micro-Doppler analysis findings, the criteria for extracting the profile of the hand Doppler signature will be established.

#### A. CHARACTERISTICS OF HAND DOPPLER SIGNATURE DURING GESTURE

The influence of the micro-Doppler effect on the shape of the Doppler power spectrum is analyzed in the following. First, the set of surface points \( \Omega(k) \) is decomposed into subsets \( \Omega_{hd}(k) \) and \( \Omega_{arm}(k) \) containing surface points only on the hand and arm segments, respectively. Therefore, the Doppler profiles and the total path gain scattered from the hand and arm segments can be separately computed. The temporal distribution of the Doppler frequency for hand motion in Fig. 5 is visualized in Fig. 6(a). The hand segment produced a larger Doppler shift than the arm segment most of the time due to the higher acceleration of the hand. On the other hand, the temporal path gain from the hand segment depicted in Fig. 6(b) depends on the orientation of the human limb relative to both the Tx and Rx positions, which explains the lower path gain on the hand segment during the periods 0.71–0.9 s and 2.31–2.5 s. Consequently, it is challenging to pinpoint the micro-Doppler signature of the hand from the path gain without knowledge of angular information.

#### B. PROPOSED HAND DOPPLER EXTRACTION TECHNIQUE BASED ON PEAK DETECTION

According to the discussion in Section IV-A, micro-Doppler components corresponding to the hand segment should be likely situated near the farthest peak in the spectrum due to its faster motion. Regarding the time-Doppler frequency uncertainty, multiple Doppler components are indistinguishable within the time-Doppler uncertainty area [50]. This prevents us from deterministically extracting a representation of the hand Doppler component from the spectrum.

Therefore, a heuristic approach is chosen for the proposed estimation technique of the hand Doppler signature. In a nutshell, the Doppler frequency produced by the hand surface is defined as the expectation of the Doppler spectrum over the region that likely contains the hand Doppler component. Defining this region as the hand Doppler window \( S_h \), we can compute the temporal profile of the hand Doppler by,

\[
\hat{f}_d(\delta) = \frac{\sum_{f_d \in S_h} f_d P_B(\delta, f_d)}{\sum_{f_d \in S_h} P_B(\delta, f_d)}
\]

Through a heuristic investigation, the criteria to determine \( S_h \) can be broken down into four scenarios (S1–S4) with different spectral shapes.

Let us first consider the spectrum with a single peak. This scenario occurs when a dominant Doppler component with a distinct high power is surrounded by other components with a weaker Doppler power. Here, the difference among the
Doppler frequencies is shorter than the uncertainty Doppler deviation $\sigma_{fd}$, which is inversely proportional to $\sigma_t$ [50], they become indistinguishable. As a result, all the Doppler components are superimposed within a single peak.

- **S1: Single peak**
  
  Suppose the dominant peak, depicted as the green cross-marker in Fig. 7(a), is located at $\rho_d = \rho_d^{(1)}$ on the spectrum whose power is greater than the noise floor (black line). The simple approach to estimating the noise threshold is from the peak of the Doppler power during an absence of motion. In this case, the region on the spectrum spanning $\rho_d^{(1)}$ to the first minimal point on both sides of the peak is regarded as the selected region. This Doppler interval is labeled with the magenta line in the figure. The portion on the spectrum below the noise floor is excluded from the selected region.

  In the case of the multiple-peak Doppler spectrum, $L_\rho$ peaks are present in the spectrum with the power exceeding the noise floor. The Doppler frequencies corresponding to these peaks can be denoted in descending order as $\rho_d^{(1)} > \rho_d^{(2)} > \ldots > \rho_d^{(L_\rho)}$. Hence, the regions over the noise level that span the first minimal points of these peaks are defined by the following intervals of the Doppler frequency $S_\delta^{(q)}$, where $q = 1, 2, \ldots, L_\rho$.

- **S2: $\rho_d^{(q)}$ values have the same sign**
  
  As illustrated in Fig. 7(b), multiple peaks are situated in the same (negative) Doppler plane where $L_\rho = 2$. In comparison, it is obvious that the farthest peak $\rho_d^{(L_\rho)}$ contains the component with the larger Doppler shift. Hence, the selected region denoted by the magenta line should be $S_\delta^{(L_\rho)}$. Otherwise, the $S_\delta^{(1)}$ region will be chosen when all $\rho_d^{(q)}$ values are positive.

- **S3: $\rho_d^{(q)}$ values have different signs and $|\rho_d^{(q)}| \geq \sigma_{fd}$**
  
  In this scenario, the multiple peaks are distributed in both planes on the Doppler spectrum. Therefore, two candidate peaks are present that contribute to the largest Doppler shift in negative and positive planes represented by $\rho_d^{(L_\rho)}$ and $\rho_d^{(1)}$, respectively. Suppose a majority of limb movements produced the same sign of the Doppler frequency. The selected region can then be decided based on $\mu_\delta$, which is the expected value of the entire spectrum. The example in Fig. 7(c) shows the $L_\rho = 3$ peaks spectrum, where $S_\delta^{(L_\rho)}$ is chosen because $\mu_\delta$ is negative. By contrast, the selected region becomes $S_\delta^{(1)}$ when the sign of $\mu_\delta$ becomes positive.

- **S4: $\rho_d^{(q)}$ values have different signs and $|\rho_d^{(q)}| \leq \sigma_{fd}$**
  
  In this scenario, most of the Doppler components produced by the hand and arm are relatively small and...
indistinguishable from the spectrum due to coherency. Therefore, the selected region is determined by including all intervals associated with \( L_p \) peaks, as depicted in Fig. 7(d).

This temporal profile computed from (11) is supposed to approximately represent the Doppler of the entire hand, although the actual Doppler shifts produced by each hand surface point have slightly variation in magnitude, as seen from the simulation results in Fig. 6(a). Nevertheless, the reference Doppler profile of the hand gesture motion \( \hat{d}_d^\text{ref}(\delta) \) is established for validation of the proposed extraction method and is given by

\[
f_d^\text{ref}(\delta) = \frac{1}{2} \left[ \max_{\ell \in \Omega_d(k)} f_d^{(\ell)}(\delta) + \min_{\ell \in \Omega_d(k)} f_d^{(\ell)}(\delta) \right]
\]

Visually, the reference hand Doppler of the gesture in Fig 5 is depicted as the solid line in Fig 6(a). The reliability of the reference profile has been heuristically validated to reconstruct a fine-grained trajectory.

V. HAND GESTURE TRAJECTORY ESTIMATION

The Doppler profile from (11) provides only the information of the temporal speed in a radial direction relative to the position of Tx and Rx as detailed in [22], [51]. To trace the hand trajectory in 3D space, one must obtain the temporal position in Cartesian coordinates, which represents the movement of the hand segment. In this work, the trajectory is recursively estimated using a set of Doppler frequency information collected from \( NM \) pairs of Rx and Tx antennas. Because this is a dynamic model, estimation of the initial value will also be discussed.

A. TRAJECTORY ESTIMATION MODEL

Suppose the temporal position and velocity of the entire hand within a short-time period \( \delta \) are interpreted by the 3D column vectors \( v_{hd}(\delta) \) and \( p_{hd}(\delta) \), respectively. The bistatic Doppler radar in (8) can be written in Cartesian coordinates as

\[
\hat{f}_{d,(nm)}(\delta) = u_{nm}(\delta) \nu(\delta)
\]

where the vector \( u_{nm}(\delta) \) represents the relative direction from the hand position to the \( n \)-th Tx and \( m \)-th Rx, respectively, which can be expressed by

\[
u_{nm}(\delta) = \frac{1}{\lambda} \left[ \frac{p_{T_m} - p_{hd}(\delta)}{d_{T_m}(\delta)} + \frac{p_{R_n} - p_{hd}(\delta)}{d_{R_n}(\delta)} \right]
\]

These variables and vectors can be stacked up as the Doppler vector and the relative direction matrix in a multi-static radar system with multiple Tx and Rx antennas, respectively,

\[
f_d(\delta) = \left[ \hat{f}_{d,(11)}(\delta) \hat{f}_{d,(21)}(\delta) \cdots \hat{f}_{d,(NM)}(\delta) \right]^T \quad U(\delta) = \left[ u_{11}(\delta) u_{21}(\delta) \cdots u_{NM}(\delta) \right]^T
\]

Therefore, the least squares (LS) solution of the multi-static Doppler radar yields the estimated velocity:

\[
v_{hd}(\delta) = U^T(\delta)f_d(\delta)
\]

where \([\cdot]^T\) represents the Moore–Penrose pseudoinverse operation. However, the LS solution in (17) assumes a given \( U(\delta) \), which is unfortunately impractical because it consists of the unknown temporal position \( p_{hd}(\delta) \). Therefore, a simple approach to predict the hand trajectory, the path described by the movement of the temporal position vector, is to recursively apply the kinematic equation

\[
p_{hd}(\delta + \Delta \delta) = p_{hd}(\delta) + \nu_{hd}(\delta)\Delta \delta \quad (18)
\]

where \( \Delta \delta = CT_s \) represents a time sampling between consecutive short-time windows. Likewise, the predicted trajectory \( p_{hd}(\delta + \Delta \delta) \) is substituted into \( U(\delta) \) to recursively estimate \( \nu(\delta + \Delta \delta) \) given the input \( f_d(\delta) \). Depending on the quality of the Doppler profile estimated from (11) and the initial position \( p_{hd}(0) \), the trajectory performance may deteriorate due the error accumulated from this recursive process, which could be mitigated by applying a KF to the system [52]. Assuming a constant velocity within a short-time window, (18) is treated as the process model of the KF. The estimated velocity in (17) is the measurement variable in the KF framework to predict the state estimates \( p_{hd}(\delta) \) and \( \nu_{hd}(\delta) \) at the next snapshot.

In regard to the antenna separation, it is evident that all the Tx and Rx antennas should be widely separated relative to the distance from the hand position. This condition can maintain the spatial diversity between Doppler measurements such that the rank of \( U(\delta) \) remains equal to or greater than three for the estimation of the velocity components. Direction ambiguity is another problem in this trajectory tracking framework. Because the Doppler represents only the 1D movement characteristics with respect to the positions of the Tx and Rx antennas, they should be distributed along the three main axes to resolve the temporal position along the three components. It should be noted that the estimated trajectory does not physically trace any particular movement of points of the hand surface, because it is estimated based on the extracted Doppler profile from (11). However, it should statistically produce a similar trajectory as those from the hand surface.

B. INITIAL POSITION ESTIMATION PROBLEM

Doppler profile information alone may not be sufficient to obtain a precise absolute initial position. The localization scheme generally requires either fine-grained angular or delay information [45], which is obtained by utilizing an array antenna or a wideband signal, respectively. Unfortunately, this requirement may be impractical to achieve with a typical narrowband Wi-Fi system. Because trajectory estimation is the subject of this study, an arbitrary initial position should be acceptable as long as the hand trajectory is correctly estimated. Although the trajectory is a temporal sequence of positions, as mentioned previously, the velocity itself dictates the direction of motion. In other words, the estimated velocity in (17) that yields the smallest residual during the entire gesture given an arbitrary initial position could implicitly represent the existence of an overall trajectory. Hence, this
condition can be expressed in terms of the optimization problem as

$$\arg\min_{p_{hd}(0)} \sum_{\delta \in \mathcal{L}} \epsilon(\delta)$$

subject to

$$\epsilon(\delta) = \left\| f_{d}(\delta) - U(\delta)p_{hd}(\delta) \right\|,$$

$$p_{a} \leq p_{hd}(0) \leq p_{b}$$

where $\mathcal{L}$, $\epsilon(\delta)$, $p_{a}$, and $p_{b}$ are a set of the short-time windows within a single gesture motion, the LS residual at each windowed snapshot, and the lower and upper search boundaries of the initial position, respectively. In practice, a set of $x(\delta)$ with different initial positions are generated in parallel using the KF model. Only the trajectory satisfying the above optimization problem is selected as the final estimated trajectory.

VI. EXPERIMENT AND DISCUSSION

A. EXPERIMENTAL SETUP AND SCENARIO

In this section, three experiments were designed to verify and evaluate the performance of the CSI model, hand Doppler profile extraction techniques, and hand trajectory estimation framework. These experiments were primarily conducted in a simulation environment using MATLAB, based on the CSI model and proposed techniques described in Sections III–V. AWGN noise was supplied to the simulated CSI with a signal-to-noise ratio (SNR) that was heuristically set to 50 dB. Except for the second experiment, the measurements were also performed to verify the applicability of the simulation results. The hardware configuration was similar to the measurement setup in [23]. Linux 802.11n CSI Tool [53] was utilized to obtain CSI from each Wi-Fi packet propagating between two commercial laptops at 5.31 GHz center frequency and 40 MHz bandwidth. The b2b calibration proposed in [23] was applied to mitigate the effect of the undesired phase rotation. As a result, the calibrated CSI has the recovered phase component, thus allowing a full analysis of the micro-Doppler frequency. Tx and Rx antennas were externally extended from laptops and placed according to the simulation scenarios described in the following.

The results of the first experiment are discussed in Section VI-B. The capability of the CSI model to simulate the micro-Doppler effect was compared with the measurement results in terms of the characteristics of the time-Doppler power spectrum. The bistatic radar system and hand waving depicted in Fig. 5 were applied in this investigation. The characteristics of the Doppler spectrum are discussed based on the corresponding temporal Doppler and the path gain characteristics shown in Fig. 6.

The following experiment described in Section VI-C was designed to evaluate the performance of the hand Doppler extraction technique. The effect of the moving average (MA) on the quality of the extracted profiles was examined by applying various MA window sizes. In addition, the Doppler extraction performance was compared with other techniques used in CSI-based human tracking [29], [36] and 2D hand gesture tracking [34]. Because these Doppler profiles will be applied to estimate hand trajectories, the simulation scenario depicted in Fig. 8(a) was established considering the tracking framework. Multi-static radar systems with three Tx and two Rx widely-separated antennas were set up to capture the Doppler frequency produced by the hand gestures depicted in Fig. 8(b). To simplify the discussion, the movement along the $y$-direction of a square gesture is referred to as the horizontal motion segment. Similarly, the $z$-direction movement in both gestures is regarded as a vertical motion segment. On the other hand, the diagonal motion segment represents the second and third segments of the M-shaped gesture. The absolute Doppler error term is introduced to measure the discrepancy of the extracted Doppler profiles from the reference defined in (12).

The last experiment, described in Section VI-D, evaluated the performance of hand trajectory estimation that utilized Doppler profiles obtained from the previous experiment. The trajectory framework described in Section V-A was applied to estimate the gesture trajectory. Because a motion path is governed by the velocity, two velocity-related quantities were introduced to quantify the trajectory error relative to the reference trajectory, which was estimated by applying the reference Doppler profiles in (12) to the same trajectory framework. The temporal magnitude error measures the speed offset at each snapshot, whereas the temporal direction error describes the motion direction offset from the reference.
Both quantities can be mathematically expressed as

\[
\psi_{v, \text{error}}(\delta) = \cos^{-1}\left( \frac{\mathbf{v}_{\text{hd}}(\delta) \cdot \mathbf{v}_{\text{ref}}(\delta)}{\|\mathbf{v}_{\text{hd}}(\delta)\| \|\mathbf{v}_{\text{ref}}(\delta)\|} \right)
\]

\[
\psi_{v} = \psi_{v, \text{error}}(\delta)
\]

where \(\mathbf{v}_{\text{hd}}(\delta)\) denotes the velocity of the reference trajectory. In the first evaluation of trajectory tracking, the initial point was given to avoid the impact of the initial position estimation. The effect of the KF on the trajectory was investigated by comparing it with the effect observed without applying the KF as well as by comparing the performance with Doppler profiles extracted from the other techniques mentioned previously. After that, the trajectory performance with an estimated initial position was examined to clarify the performance of the algorithm described in Section V-B.

B. VALIDATION OF CSI MODEL FOR MICRO-DOPPLER EFFECT PREDICTION

The Doppler power spectra computed from the simulated and measured CSI using (10) with 16 ms standard deviation of the Gaussian window are illustrated in Fig. 9.

Comparatively, it was found that the overall micro-Doppler effect characteristics computed from the simulated CSI in Fig. 9(a) showed a certain degree of agreement with the measurement depicted in Fig. 9(c). At approximately 0.5–1 s, both results exhibited a small Doppler power spreading out in the negative Doppler plane. On the other hand, a relative higher power was observed at 1–1.5 s in all cases due to the rapid change of the path gain as shown in Fig. 6(c). In the vertical motion, the Doppler frequency shift was comparatively smaller due to the orthogonal direction of motion to the LoS direction between Tx and Rx [23], [51]. Although the simulated Doppler spectrum also successfully generated a similar micro-Doppler pattern, the measurement exhibited a fairly larger Doppler shift within ±20 Hz in comparison with the ±15 Hz spread in the simulated Doppler. This is possibly due to the inaccuracy of the kinematic model. To simulate the noisy Doppler power spectrum, the simulated result with 40 dB SNR was also compared in Fig. 9b. The result could balance the simulated spectrum shape closer to the measurement in terms of the noise level.

In regard to the level of Doppler power, the result from the measurement had a shorter dynamic range of approximately 30 dB in comparison with the simulated Doppler of 50 dB. Because the measured CSI was normalized by the internal reference of the Wi-Fi chip [54], the temporal fluctuation in CSI may also have been normalized which affected the dynamic range of the Doppler power spectrum. Although a further detailed examination is needed to clarify this effect, it may not have a strong impact on the motion tracking because the spectrum shapes are the most important factor for extracting the hand Doppler profile, as discussed in Section IV.

This experiment has confirmed the usability of the simulated CSI model to deterministically reproduce the micro-Doppler effect due to the motion of hand gestures in terms of the joint time-Doppler spectrum. It is worth noting that the proposed CSI model could alternatively be utilized to generate the trained dataset, which is one of the problems in the CSI-based gesture recognition system as stated in [25].

C. VALIDATION OF DOPPLER PROFILE ESTIMATION

The estimation technique of Section IV-B was applied to extract the hand-only Doppler profile from the Doppler spectrum. The evaluation is depicted in Fig. 10 in terms of the CDF function of the absolute Doppler error representing the overall discrepancy of six estimated Doppler profiles. An MA filter was applied to smooth the estimated profile. Although the performance does not show a significant difference at the median level of 1 Hz error, the MA filter began to gradually take effect at the 90th percentile with a 1–1.5 Hz Doppler error reduction with 3- and 5-point MA window filters shown in Fig. 10 for both gestures. It should be noted that there was less improvement in the estimated Doppler profile with MA window sizes larger than the 5-point MA window. As a result, it could be assessed that the proposed profile estimation technique can extract the hand Doppler profile with an error of less than 5 Hz in most of the cases, given the 50 dB SNR, which could be regarded as a noise-free environment.

Finally, the performance of the proposed Doppler profile extraction was compared with the other two techniques used in CSI-based motion tracking. The first method pinpointed the strongest power in the spectrum as the human temporal

FIGURE 9. Doppler spectrum comparison of the hand gestures between the simulated and measured CSI.
position in [29], [36]. Another extraction method in [34] utilized the phase derivative to obtain the Doppler frequency in terms of the path length change to iteratively trace the 2D hand gesture. Overall, the proposed method outperformed both methods, as depicted in Fig. 10. In fact, a significantly large error of > 25 Hz is observed from the square gesture, as shown in Fig. 10(a), whereas the proposed method with the 5-point window MA filter maintains the error within 10 Hz. Because these methods did not consider the micro-Doppler effect, they could not handle Doppler frequency incoherency, which manifests in the Doppler spectrum as multi-peak spectral shapes, as depicted in Figs. 7(b)–7(d). In the case of the M-shaped gesture, on the other hand, the proposed method exhibits less error than the other methods, as shown in Fig. 10(b). Interestingly, the derivative-based method performed slightly better until 0.8 CDF, before it rapidly drifted out to approximately 17 Hz in error. This is because the M-shaped gesture consists of vertical and diagonal trajectories, which generally produced less Doppler variation due to the smaller radial velocities. Therefore, the estimated profiles of the M-shaped gesture did not vary much from the reference profile.

This comparison confirmed the robustness of our method in addressing with the micro-Doppler effect for hand Doppler profile extraction, as it produced a smaller error. Although the Doppler extraction techniques used in the literature may be applied for capturing the hand Doppler component, their performance clearly depends on the direction of motion and the position of the antennas.

D. EVALUATION OF HAND TRAJECTORY ESTIMATION

1) IMPACT OF THE KF ON THE ESTIMATED TRAJECTORY
Six estimated Doppler profiles with a 5-point MA filter were chosen for estimating the trajectory of hand gestures. Two sets of trajectories—one with and one without the KF—were plotted in Fig. 11 to validate the effect of the KF. The former trajectory was depicted with an orange line in the figures, whereas the latter trajectory was labeled with a yellow line. The initial position of the wrist, $p_w(0)$, was given in this analysis.

Although the estimated trajectories could be traced without applying the KF as illustrated in Figs. 11(a) and 11(d) for the square and M-shaped gestures, respectively, the estimated trajectories experienced a scaling offset from the actual length depicted with the blue line, especially in vertical motion. However, the excess length of the estimated trajectories reduced significantly after applying the KF. This improvement can be explained by comparing the magnitude error of the trajectory. Let us consider the magnitude errors during 2–3 s of Fig. 11(b) and during 0.5–1.5 s of Fig. 11(e), which correspond to the vertical motion of the square and M-shaped gestures, respectively. As the magnitude error was positive most of the time in the case of the trajectory without the KF, the cumulative sum, which is proportional to the length of trajectory, contributed to the larger scaling offset. Similarly, a larger negative error within 3.5–4.5 s of Fig. 11(e) produced a shorter length offset for the M-shaped gesture. On the other hand, the estimated trajectory after applying the KF produced a similar magnitude error characteristics that fluctuated between positive and negative values within each motion segment. As a result, the KF effectively mitigated the cumulative magnitude error, thus reducing the scaling offset.

In regard to the direction error, the KF clearly smoothed the temporal orientation offset of the estimated trajectory, as shown in Figs. 11(c) and 11(f). However, a distinct pattern of a larger offset with more than 90° at the beginning and end of each motion segment still persisted through all the segments. Fortunately, the corresponding magnitude offset within this period has been vastly mitigated, thus minimizing trajectory distortion. Therefore, the effective orientation offset primarily contributed to a moderate direction error at the center of motion segments that varied between 10° and 30°.

These findings imply that the shape of the hand gesture may not be significantly improved with the KF because the direction offset is slightly reduced. Nevertheless, the impact of the KF is significantly manifested in terms of scaling offset mitigation.

2) IMPACT OF THE ESTIMATED INITIAL POSITION ON THE TRAJECTORY
At this point, the trajectory estimation was computed by assuming a known initial hand position, which is usually impractical. The optimization problem presented in Section V-B was applied to estimate the initial position based
on the cumulative LS residual. The search boundary could be easily determined by analyzing the joint delay-Doppler spectrum [22] from the six streams of CSI. Although the position uncertainty could span a magnitude of a few meters due to a coarse-grained delay resolution in the narrowband signal, it was sufficient to determine the search region. Given the 8-m uncertainty derived from the bandwidth, $p_a$ and $p_b$ have been set to (-4, 4, 0) and (4, 4, 2), respectively. A 2 m distance in the z-direction was arbitrary chosen based on the human height indicated in Table 1.

By way of comparison, the hand trajectory computed using the estimated initial position is depicted with a magenta line in Figs. 11(a) and 11(d) for both gestures. Here, it was apparent that these trajectories were still preserved despite a meter-level difference in the initial state in comparison with the reference position. This is because the objective behind was not to find the absolute initial state but rather an arbitrary position that could sustain the overall trajectory. In terms of the magnitude error, depicted with magenta lines in Figs. 11(b) and 11(e) for the square and M-shaped gesture, respectively, the performance is similar to those with the given initial position even though a slightly larger error could be observed. On the other hand, the direction error, on average, showed a constant discrepancy of approximately $15^\circ$–$20^\circ$ degrees from the trajectory with the given initial position, according to the results in Figs. 11(c) and 11(f) for both gestures.

The effect of the initial position on the orientation offset can be explained by considering the multi-static Doppler radar in Section V-A. Given the estimated Doppler profiles, the proportion of velocity components would be altered to accommodate the change in $U^T(\delta)$ because of the different initial positions. The change in the velocity vector is the cause of the offset. In summary, this experiment has confirmed the robustness of the trajectory to the initial position. In contrast, the quality of the Doppler profiles derived from the spectrum strongly influences the trajectory performance.

3) TRAJECTORY TRACKING WITH THE ACTUAL WI-FI CSI

This proof of concept has also been tested with measurements conducted in the meeting room. In the measurement, each gesture was performed twice at the same position while CSI was being captured. The trajectory and initial position were estimated offline using the same proposed technique. As depicted in Fig. 12, the estimated trajectories resembled the square and M-shaped gestures in both trials, although both of them experienced a rotation offset relative to the reference trajectory as well as a slight scale and shape distortion. This was likely due to a higher noise power that partially contaminated the Doppler power in the spectrum, thus resulting in inaccurate extraction of the Doppler profiles.

For clarification, an additional experiment was conducted by contaminating the simulated CSI with a higher AWGN. Two samples of estimated trajectories with 40 dB SNR were
selected because it exhibited a similar spectrum distortion due to the noise contamination, as shown in Fig. 9. Both trajectories from the simulated Doppler produced a similar distortion and initial position offset compared to those from the measurement, as illustrated in Fig. 12 for both gestures. The quality of the Doppler profiles could be the cause of the initial position difference, as explained previously. In addition, the distortion could be due to various uncertainty factors during the measurement, such as the difference in the speed profiles and the hand position while performing gestures. More experiment with reproducible instruments such as a robot arm are needed to further clarify the effect of measurement uncertainty. Nevertheless, the measurements successfully confirmed the usability of 3D hand trajectory tracking from CSI measured by commodity Wi-Fi devices given a sufficient SNR level.

VII. CONCLUSION
RF-based motion sensing has been investigated to enable sensing capability in scenarios where commercial vision-based sensing systems find it difficult to operate. This work analyzes the non-intrusive trajectory tracking of hand gestures based on the CSI Doppler frequency. To understand deterministically how the movement of a hand gesture generates a Doppler frequency in CSI, a surface-point-based human limb model was developed by simulating the non-rigid motion model of the robot arm. In this work, simulated CSI was computed based on EM scattering from surface points on a human limb by using PO. This approach enables us to obtain micro-Doppler signatures produced by each segment on the human limb through an analysis of the TV Doppler power spectrum. A Doppler profile estimation technique based on peak detection was established to extract the portion of hand micro-Doppler signatures from the Doppler power spectrum. The trajectory of the hand gesture was reconstructed by recursively estimating the velocity and temporal position from a set of extracted Doppler profiles.

An experiment was conducted to validate the performance of the CSI model, Doppler profile extraction, and trajectory estimation from hand movement. The micro-Doppler effect produced from the simulated CSI showed a similar pattern to those obtained from the off-the-shelf Wi-Fi CSI. It was revealed that the error from the extracted Doppler profile was rather small, with a less than 4-Hz offset at the 90th percentile of the CDF. However, the discrepancy could increase to 10 Hz when the movement direction was approaching orthogonality with the relative direction from the hand to the Tx and Rx antennas. Despite the profile error, the trajectories of both square and M-shape gestures were successfully traced albeit a moderate scale and orientation distortion. In addition, it was found that the quality of the overall trajectory is robust to the initial position but highly sensitive to the extracted Doppler profile. Measurements were conducted to test the applicability of this trajectory tracking framework. The trajectory of hand gestures was successfully reconstructed with promising results in all the measurement trials.

REFERENCES
[1] Q. Pu, S. Gupta, S. Gollakota, and S. Patel, “Whole-home gesture recognition using wireless signals,” in Proc. 19th Annu. Int. Conf. Mobile Comput. Netw., 2013, pp. 27–38.
[2] 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., Services, New York, NY, USA, Jun. 2018, p. 350.
[3] M. A. Aide Al-Qaness and F. Li, “WiGeR: Wifi-based gesture recognition system,” ISPRS Int. J. Geo-Inf., vol. 5, no. 6, pp. 1–17, 2016.
[4] G. Wang, Y. Zou, Z. Zhou, K. Wu, and L. M. Ni, “We can hear you with Wi-Fi,” IEEE Trans. Mobile Comput., vol. 15, no. 11, pp. 2907–2920, Nov. 2016.
[5] X. Liu, J. Cao, S. Tang, and J. Wen, “Wi-sleep: Contactless sleep monitoring via WiFi signals,” in Proc. Real-Time Syst. Symp., vol. 2014, pp. 346–355.
[6] H. Wang, D. Zhang, J. Ma, Y. Wang, Y. Wang, D. Wu, T. Gu, and B. Xie, “Human respiration detection with commodity WiFi devices: Do user location and body orientation matter?”, in Proc. ACM Int. Joint Conf. Pervas. Ubiquitous Comput., New York, NY, USA, Sep. 2016, p. 25.

[7] X. Wang, C. Yang, and S. Mao, “PhaseBeat: Exploiting CSI phase data for vital sign monitoring with commodity WiFi devices.” in Proc. IEEE 27th Int. Conf. Distrib. Comput. Syst. (ICDCS), Jun. 2017, pp. 1230–1239.

[8] D. Wu, D. Zhang, C. Xu, H. Wang, and X. Li, “Device-free WiFi human sensing: From pattern-based to model-based approaches,” IEEE Commun. Mag., vol. 55, no. 10, pp. 91–104, Oct. 2017.

[9] Y. Ma, G. Zhou, and S. Wang, “WiFi sensing with channel state information: A survey,” ACM Comput. Surv., vol. 52, no. 3, p. 36, 2019.

[10] F. Adith, Z. Kabelac, and D. Katabi, “Multi-person localization via RF body reflections,” in Proc. 12th USENIX Symp. Netw. Syst. Des. Implement. (NSDI), 2015, pp. 279–292.

[11] F. Adith, H. Mao, Z. Kabelac, D. Katabi, and R. C. Miller, “Smart homes that monitor breathing and heart rate,” in Proc. 33rd Annu. ACM Conf. Hum. Factors Comput. Syst., 2015, pp. 837–846.

[12] S. Yousef, H. Narui, S. Dayal, S. Ermon, and S. Valaee, “A survey on behavior recognition using WiFi channel state information,” IEEE Commun. Mag., vol. 55, no. 10, pp. 98–104, Oct. 2017.

[13] D. Zhang, H. Wang, and D. Wu, “Toward centimeter-scale human activity sensing with Wi-Fi signals,” Computer, vol. 50, no. 1, pp. 48–57, Jan. 2017.

[14] 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’16 Ext. Abstr. Hum. Factors Comput. Syst., May 2016, pp. 1961–1972.

[15] Z. Tan, J. Wang, X. Yang, and M. Zhou, “Wicatch: A Wi-Fi based hand gesture recognition system,” IEEE Access, vol. 6, pp. 16911–16923, 2018.

[16] W. He, K. Wu, Y. Zou, and Z. Ming, “WiG: WiFi-based gesture recognition system,” in Proc. 24th Int. Conf. Comput. Commun. Netw. (ICCCCN), Aug. 2017, pp. 1–7.

[17] S. Tan and J. Yang, “WiFinger: Leveraging commodity WiFi for fine-grained finger gesture recognition,” in Proc. 17th ACM Int. Symp. Mobile Ad Hoc Netw. Commun., Paderborn, Germany, 2016, pp. 201–210.

[18] H. Li, W. Yang, J. Wang, Y. Xu, and L. Huang, “WiFinger: Talk to your smart devices with finger-grained gesture,” in Proc. ACM Int. Joint Conf. Pervas. Ubiquitous Comput., Sep. 2016, pp. 250–261.

[19] K. Ali, A. X. Liu, W. Wang, and M. Shahzad, “Keystroke recognition using WiFi signals,” in Proc. 21st Annu. Int. Conf. Mobile Comput. Netw., Paris, France, 2015, pp. 90–102.

[20] M. Skolnik, Radar Handbook (Electrical Engineering), New York, NY, USA: McGraw-Hill, 2008.

[21] Y.-C. Xiao, P. Wei, and T. Yuan, “Observability and performance analysis of Bi/Multi-static Doppler-only radar,” IEEE Trans. Aerosp. Electron. Syst., vol. 46, no. 4, pp. 1654–1667, Oct. 2010.

[22] Z. Solatizadeh and S. Azimi-Shahbazi, “Accelerating target detection in passive radar sensors: Delay-Doppler-Acceleration estimation,” IEEE Sensors J., vol. 18, no. 13, pp. 5445–5454, Jul. 2018.

[23] N. Keerativoranan, A. Haniz, K. Saito, and J.-I. Takada, “Mitigation of CSI temporal phase rotation with B2B calibration method for fine-grained motion detection analysis on commodity Wi-Fi devices,” Sensors, vol. 18, no. 11, p. 1–18, Nov. 2018.

[24] M. A. Al-Qaness, M. A. Alzaid, S. Kim, A. A. Ewees, A. A. Abbasi, Y. A. Alhaj, and A. Hawbani, “Channel state information from pure works with effective SNR,” Ph.D. dissertation, Dept. Comput. Sci. Eng., College Eng., Univ. Washington, Washington, DC, USA, 2012.
NOPPHON KEERATIVORANAN (Member, IEEE) was born in 1989. He received the B.E. degree (Hons.) in electrical engineering from Thammasat University, Thailand, in 2012, the M.S. degree from Seoul National University, South Korea, in 2015, and the D.E. degree from the Tokyo Institute of Technology, Japan.

In 2012, he was a Telecommunication Engineer with Advance Info Service (AIS), Thailand. From 2015 to 2016, he was a Research Assistant with the National Electronics and Computer Technology Center (NECTEC), Thailand. He is currently a Researcher with the Tokyo Institute of Technology. His research interests include RF-based localization and tracking, radio propagation channel modeling, and measurements for wireless communication system and application.

PANAWIT HANPINITSAK (Member, IEEE) was born in 1991. He received the B.E. degree (Hons.) in electronics and communications engineering from the Sirindhorn International Institute of Technology, Thammasat University, Thailand, in 2013, and the M.E. and D.E. degrees from the Tokyo Institute of Technology, Japan, in 2016 and 2019, respectively.

He was a Guest Ph.D. Researcher with Aalborg University, Denmark, and the Ilmenau University of Technology, Germany, in 2016 and 2018, respectively. He is currently a Researcher with the Tokyo Institute of Technology. His research interests include radio propagation channel modeling, measurement, and prediction at millimeter waves, localization, and spectrum sensing in cognitive radio. He is also a member of the IEICE. He also received the Best Student Presentation Award from the IEICE Short Range Wireless (SRW) conference, and the IEEE Antenna and Propagation Society (APS) Best Student Paper Award in 2017 and 2019, respectively.

KENTARO SAITO (Member, IEEE) was born in Kanagawa, Japan, in 1977. He received the B.S. and Ph.D. degrees from The University of Tokyo, Japan, in 2002 and 2008, respectively.

He joined NTT DOCOMO, Kanagawa, in 2002. Since then, he has been involved in the research and development of mobile communication systems and radio propagation. He joined the Tokyo Institute of Technology, Japan, in 2015. Since then, he has also been involved in research on radio propagation for mobile communication systems. He is a member of the IEICE.

JUN-ICHI TAKADA (Member, IEEE) received the B.E., M.E., and D.E. degrees in electrical and electronic engineering from the Tokyo Institute of Technology, Tokyo, Japan, in 1987, 1989, and 1992, respectively.

He was a Research Associate with Chiba University, Chiba, Japan, from 1992 to 1994. He was an Associate Professor with the Tokyo Institute of Technology, from 1994 to 2006. He was also a Researcher with the National Institute of Information and Communications Technology, Kanagawa, Japan, from 2003 to 2007. He has been a Professor with the Tokyo Institute of Technology, since 2006. His research interests include radio-wave propagation and channel modeling for mobile and short-range wireless systems, regulatory issues of spectrum sharing, and ICT applications for international development. He is a Fellow of the IEICE, Japan.