Spin-Antenna: 3D Motion Tracking for Tag Array
Labeled Objects via Spinning Antenna

Chuyu Wang, Lei Xie, Keyan Zhang, Wei Wang, Yanling Bu, Sanglu Lu
State Key Laboratory for Novel Software Technology, Nanjing University, China
Email: {chuyu, lxie, ww, sanglu}@nju.edu.cn, {yanling, koye}@smail.nju.edu.cn

Abstract—Nowadays, the growing demand for the 3D human-computer interaction (HCI) has brought about a number of novel approaches, which achieve the HCI by tracking the motion of different devices, including the translation and the rotation. In this paper, we propose to use a spinning linearly polarized antenna to track the 3D motion of a specified object attached with the passive RFID tag array. Different from the fixed antenna-based solutions, which suffer from the unavoidable signal interferences at some specific positions/orientations, and only achieve the good performance in some feasible sensing conditions, our spinning antenna-based solution seeks to sufficiently suppress the ambient signal interferences and extracts the most distinctive features, by actively spinning the antenna to create the optimal sensing condition. Moreover, by leveraging the matching/mismatching property of the linearly polarized antenna, i.e., in comparison to the circularly polarized antenna, the phase variation around the matching direction is more stable, and the RSSI variation in the mismatching direction is more distinctive, we are able to find more distinctive features to estimate the position and the orientation. We build a model to investigate the RSSI and the phase variation of the RFID tag along with the spinning of the antenna, and further extend the model from a single RFID tag to an RFID tag array. Furthermore, we design corresponding solutions to extract the distinctive RSSI and phase values from the RF-signal variation. Our solution tracks the translation of the tag array based on the phase features, and the rotation of the tag array based on the RSSI variation. The experimental results show that our system can achieve an average error of 13.6cm in the translation tracking, and an average error of 8.3° in the rotation tracking in the 3D space.

I. INTRODUCTION

Nowadays, the 3D human-computer interaction (HCI) has become a brand-new approach to allow the users to interact with the computer via natural 3D gestures. Conventionally, the users perform the gestures with a specific device, while both the position and the orientation of this device are continuously tracked to generate the corresponding gestures for the interaction. Such interaction style can get rid of the constraint from traditional interaction devices, e.g., the keyboard and the mouse, which only work in the 2D space, and provide more comfortable user experience and accurate operation, when interacting in the 3D space.

The state-of-art HCI approaches mainly fall into three categories, i.e., the computer vision (CV)-based approaches, sensor-based approaches, and sensorless approaches. The CV-based approaches, e.g., Microsoft Kinect, use the features from captured images to perform the motion tracking and gesture recognition, which are easy to be affected by either the poor light condition or the object occlusion in the non-line-of-sight condition. Sensor-based approaches, e.g., IMU in the smartphone, reconstruct the gesture trace based on the inertial measurement data from the accelerometer and the gyroscope, which suffer from the limited battery life and the high cost of hardwares. Moreover, some specific sensors with the accurate sensing capability are relatively heavy, leading to uncomfortable user experience. Recently, the sensorless methods, such as WiFi and RFID-based sensing, have been widely investigated to design novel HCI schemes. In particular, due to the low-cost and the light-weight properties, the RFID system has enabled brand-new solutions for the motion tracking and the gesture recognition [4, 8, 9, 11, 15], by attaching the RFID tags on the HCI devices. However, most of the previous work tracks the motion of the tagged object only in the 2D space, which do not consider the orientation variation of the tagged object in the 3D space. Tagyro [13] tracks the orientation of a tagged object via multiple antennas, however, it is not able to track the absolute translation of the object simultaneously. Tag-compass [7] estimates the orientation of one tag based on multiple spinning antennas, however, it is based on the precondition that the tagged object is deployed in a specified 2D plane. Different from these work, we need to track the motion of the tag array in the 3D space with six degrees of freedom, including the translation and rotation, which has remained unresolved so far.

In this paper, we propose a novel RFID-based approach to track the 3D motion of the tagged object by continuously spinning the linearly polarized antenna. Specifically, we attach a tag array onto the specified object with multiple tags in different orientations, and extract the distinctive signal features based on the spinning antenna to track the 3D motion of the tagged object, including the translation and the rotation. Realizing that the fixed antenna-based solutions in the previous work usually suffer from the unavoidable signal interferences at some specific positions/orientations, and only achieve good...
In this paper, we make three main contributions as follows. 1) To the best of our knowledge, we are the first to thoroughly investigate the characteristics of the spinning linearly polarized antenna in the RFID system, with empirical studies and building models, which further facilitate the motion tracking in the 3D space. 2) We propose to leverage the tag array with different tag orientations for the 3D motion tracking, and build corresponding models to derive the translation and the rotation based on the RSSI and the phase extracted from the spinning linearly polarized antenna. 3) We implement a prototype system of a spinning antenna with the COTS RFID and evaluate its performance in the real environment. The experiments show that our system can achieve an average accuracy of 13.6cm in the translation tracking, and an average accuracy of 8.3° in the rotation tracking in the 3D space.

II. RELATED WORK

Computer Vision-based Approach. Based on the accurate captured images and videos, the CV-based approaches are widely used to track the motion or recognize the gesture of either objects or human subjects [3, 5]. However, these approaches are easily affected by the poor light condition or the object occlusion in the non-line-of-sight (NLOS) condition. In contrast, we leverage the RFID technology, which uses the backscatter communication to read the passive RFID tags, so that our system has no requirements of the light condition.

Sensor-based Approach. Built-in sensors in wearable devices, e.g., the accelerometer and the gyroscope, can be utilized to reconstruct the gesture trace [12, 14, 16]. For example, ArmTrack [10] proposes to track the posture of the entire arm solely relying on the smartwatch. However, the motion sensors in wearable devices have limited battery life and high cost, and the tracking accuracy is relatively low due to the noise of the measurement data. Some specific sensors with high accuracy sensing capability are too heavy to provide comfortable user experience. In contrast, we attach a paper-like passive tag array onto the target to track the 3D motion, which is battery-free, low-cost, light-weight and portable.

RFID-based Approach. Recently, several studies have proposed to utilize the RFID technique to track the motion of tagged objects [1, 6, 9, 11, 15, 17]. Wang et al. [11] track the moving tagged finger in a 2D plane with multiple fixed antennas. Shangguan et al. [9] track the 2D moving trace of the tagged object to obtain user feedbacks with only one fixed antenna. These solutions focus on tracking the moving trace in the 2D space, but they are not suitable for the orientation tracking in the 3D space. Tagyro [13] tracks the 3D orientation of a tagged objects based on the phase differences of the tags via multiple antennas. Liu et al. [7] leverage multiple spinning linearly polarized antenna to estimate the 3D orientation of one tag in a specified 2D plane. However, these methods only can tracks the 3D orientation and are unable to track the absolute translation of the object in the 3D space simultaneously. Different from the prior work, we focus on tracking the rigid motion of the tag array in the 3D space via only one spinning linear polarized antenna, including both the translation and the rotation, which has remained unresolved so far.
III. Empirical Study

We conduct empirical studies to investigate the RF-signal features of the spinning antenna, by using an RFID reader, a circularly polarized antenna, a linearly polarized antenna and a passive UHF RFID tag. The experiment setup is shown in Fig. 2. For the global coordinate system (GCS), the origin $O$ is set to the center of the antenna, the $Z$-axis is vertically straight up, the $Y$-axis is along the spin axis, and the GCS is a right-hand system. Specifically, for the linearly polarized antenna, the polarization angle, denoted as $\phi$, is defined as the incline angle between the $X$-axis and the linear polarization plane. The polarization direction is defined as the direction of intersection between the linear polarization plane and the antenna plane. For simplicity, we use the polarization angle to indicate the spin angle of the antenna during the spinning, i.e., when the linear polarization plane of the antenna is vertical as the blue plane in Fig. 2, the polarization angle is $90^\circ$, and the spin angle equals $90^\circ$. As a dipole, the tag is regarded as a line and depicted with the 3D orientation of the line.

A. Signal Features between Different Antennas

Observation 1. During the spinning process, in comparison to the circularly polarized antenna, the phase variation of the linearly polarized antenna is more stable, and the RSSI variation of the linearly polarized antenna is more distinctive. As shown in Fig. 2, we put a tag at position $A (0, 200, 0)$, which is parallel to the $XZ$-plane with the included angle against the $X$-axis (or $Z$-axis) of $60^\circ$ (or $30^\circ$), then spin the circular/linearly polarized antenna around the $Y$-axis to interrogate the tag, separately. We compare the RF-signal variation of the tag from the circularly polarized antenna with that from the linearly polarized antenna, and plot the results in Fig. 3(a) and Fig. 3(b). It is found that the phase value from the circularly polarized antenna is linearly changing with the antenna spinning, while that from the linearly polarized antenna almost keeps unchanged. For the circularly polarized antenna, the polarization direction rotates electronically, which can be regarded as the summation of multiple RF-signals with different polarization directions and initial phase values. The tag matches the RF-signal with the same polarization direction to harvest the maximum energy, thus, the phase value of the tag changes during the spinning. For the linearly polarized antenna, it has only one polarization direction with one initial phase value, so the phase value of the tag is independent of the spinning angle and keeps stable. Besides, the RSSI values from the two types of antennas are both changing periodically with the spinning antenna, but the RSSI variation range from the linearly polarized antenna is much larger, i.e., $23\text{dB}$ larger than that from the circularly polarized antenna. For the circularly polarized antenna, the RSSI reduction is due to the different antenna gains along different directions. For the linearly polarized antenna, the RSSI reduction is due to the mismatch of the linear polarization direction with the tag orientation, which leads to large energy loss and causes more distinctive RSSI variation. Therefore, compared to the circularly polarized antenna, the linearly polarized antenna can provide more reliable position information based on the stable phase values, and more sensitive orientation information based on the distinctive RSSI variation. Thus, we prefer the linearly polarized antenna in our work.

B. RSSI Variation Pattern of Linearly Polarized Antenna

Observation 2. For the linearly polarized antenna, the mismatching direction, corresponding to the minimum RSSI value, is more distinctive for the estimation of the tag orientation than the matching direction, corresponding to the maximum RSSI value. We move the tag from position $A (0, 200, 0)$ to position $B (−100, 200, 0)$ with the same orientation, and spin the linearly polarized antenna to interrogate the tag, the results are plotted in Fig. 3(c). By comparing Fig. 3(c) with Fig. 3(b), it is found that the two RSSI variation patterns of the tag at different positions share the similar shape, consisting of two arches, but they have quite different absolute values. The peak of the arch refers to the perfectly matching polarization direction with the tag orientation, called the matching direction, and the valley of the arch refers to the mismatching polarization direction with the tag orientation, called the mismatching direction. The mismatching direction in Fig. 3(b) is $153^\circ$, which is almost orthogonal to the tag orientation when the tag is at position $A$. But the mismatching direction changes to $166^\circ$ in Fig. 3(c). It indicates that even if the tag orientation is unchanged, the mismatching direction will change according to the position of the tag. Moreover, as shown in Fig. 3(c), the RSSI values are similar around the matching direction, while they are more distinctive around the mismatching direction. Hence, we use the mismatching direction to estimate the tag orientation instead of the matching direction, which is relatively difficult to deduce.

C. Phase Variation Pattern of Linearly Polarized Antenna

Observation 3. For the linearly polarized antenna, the phase value keeps stable when the polarization direction of the antenna matches the tag orientation perfectly; and the phase value fluctuates when the polarization direction mismatches the tag orientation due to the multi-path effect. We place the tag at position $A$ vertically or horizontally, i.e., the tag orientation is along the $Z$-axis or the $X$-axis, and spin the linearly polarized antenna to interrogate the tags with different orientations, respectively. We conduct the experiments in an open lobby, where the antenna is at the height of $150\text{cm}$ and $500\text{cm}$ away from the surrounding walls, so the horizontal multi-path effect is relatively smaller than the vertical one. As shown in Fig. 3(d), the phase value of the vertical tag almost keeps unchanged, while that of the horizontal tag changes to a certain extent, especially when the tag orientation
mismatches the polarization direction of the antenna. When the tag orientation mismatches the polarization direction of the antenna, the received signal power of the tag is relatively small, thus, the multi-path signal will affect the received signal and change the phase value. For the horizontal tag, the mismatching direction is corresponding to the vertical polarization direction, thus, the tag suffers from the larger vertical multi-path effect from the ground. For the vertical tag, the mismatching direction of polarization is horizontal. Since the horizontal multi-path effect is relatively smaller, the corresponding phase change is not distinct. As a result, the horizontal tag suffers larger multi-path effect than the vertical tag according to the phase measurement. It indicates that the phase values around the matching direction are more accurate.

D. Summary

Above all, we get the following three key findings. First, the linearly polarized antenna can capture the more stable phase value and distinctive RSSI variation compared to the circularly polarized antenna. Thus, it is more suitable to use linearly polarized antenna to estimate the position with the phase value and estimate the orientation with the RSSI variation. Second, for the linearly polarized antenna, the mismatching direction based on the RSSI variation is more distinctive to estimate the tag orientation compared to the matching direction. Third, for the linearly polarized antenna, the phase around the matching direction is more stable than the phase around the mismatching direction, indicating that we can calibrate the phase value by removing the noisy phase around the mismatching direction.

IV. Modeling

In this section, we build the models to describe the relationships between the signal features and the position/orientation of the tag, which are further used to depict the translation and the rotation of the object.

A. Modeling RSSI Pattern of a Single Tag with Spin-antenna

As shown in Fig. 4, in the Global Coordinate System (GCS), the linearly polarized antenna is located at the origin $O$, and spins around the $Y$-axis. We define the polarization angle as $\phi$, i.e., the polarization direction is offset the $X$-axis by $\phi$ in the $XZ$-plane. The polarization vector is defined as $V_\phi = [\cos \phi, 0, \sin \phi]^T$, which is the unit vector pointing to the polarization angle $\phi$. The tag is deployed at a distance of $d_t$ away from the antenna, and the $\text{antenna-tag direction}$ is $V_d = [x_d, y_d, z_d]^T$, which is a unit vector pointing from the antenna to the tag. The $\text{tag orientation}$ is defined as the orientation of the long side of the tag, which is represented by the elevation angle $\beta$ and the deflection angle $\gamma$ as shown in Fig. 4. Hence, the unit vector of the tag orientation is $V_t = [x_t, y_t, z_t]^T = [\cos \beta \cos \gamma, \cos \beta \sin \gamma, \sin \beta]^T$.

Polarization loss factor calculation. In the communication system, the Received Signal Strength Indicator (RSSI) is calculated as: $R = 10 \times \log(P_{RX})$. Theoretically, the received power $P_{RX}$ is calculated as: $P_{RX} = P_{TX} h(d_t) G_p$, where $P_{TX}$ is the transmit power. $h(d_t)$ represents the path loss factor in the communication, which can be calculated based on the Friis equation [2, 7]. $G_p$ presents the polarization loss factor (PLF) [7], which calculates the matching coefficient between the polarization direction of the antenna and the tag orientation in the RFID system. Since the change of the PLF value $G_p$ is caused by the linear polarization change when the antenna is spinning, we focus on investigating the variation of PLF with different linear polarizations. In theory, $G_p$ includes the signal loss of both the uplink $G_{p_u}$ and the downlink $G_{p_d}$ in RFID communications, i.e., $G_p = G_{p_u} \cdot G_{p_d}$. According to the radar principle, the amount of the energy received by the object, e.g., the tag, is dependent on the reflective area of the object, which is the Radar Cross Section (RCS) [2]. Similarly in the RFID system, we can project the polarized signal on the tag to calculate the reflective area in both the uplink and the downlink, which is the corresponding PLF.

For the uplink transmission, the linearly polarized antenna scatters the signal to the spherical surface in the ambient space as shown in Fig. 4(a). The projection of the linear polarization at the tag position with the polarization angle $\phi$ is:

$$ P_\phi = MV_\phi. $$

Here, $M$ is the transforming matrix to project the vector $V_\phi$ onto the tangent plane of the spherical surface. Since $V_d$ is the normal vector of the tangent plane, $M$ is calculated as:

$$ M = \begin{bmatrix} 1 & -x_d^2 & -x_d y_d & -x_d z_d \\ -x_d y_d & 1 & -y_d^2 & -y_d z_d \\ -x_d z_d & -y_d z_d & 1 & -z_d^2 \end{bmatrix}. $$
Then, based on the vector projection $P_{\phi}$ and the tag orientation $V_t$, the PLF in the uplink is calculated as $G_{p,u} = |P_{\phi} \cdot V_t|^2$.

Actually, the PLF represents the magnitude of the projection of the polarization vector $P_{\phi}$ along the tag orientation $V_t$, which reflects the reflective area of the tag [2, 7].

For the downlink transmission, given the orientation of the tag $V_t$ as shown in Fig. 4(b), which is the polarization direction of the dipole tag, the projected polarization vector $P_{\phi}$ at the antenna position is calculated similar to Eq. (1) as:

$$P_{\phi} = (x_{t,p}, y_{t,p}, z_{t,p}) = MV_{t}.$$  \(3\)

Since the linearly polarized antenna is a patch antenna, it can be regarded as a panel to receive the signal backscattered from the tag [7], while the tag only receives the matching signal of its orientation. Therefore, different from the uplink PLF, the downlink PLF is the projection of $P_{\phi}$ on the antenna plane, i.e., the $XZ$-plane, which is $G_{p,d} = x_{t,p}^2 + z_{t,p}^2$. Based on the PLF in the uplink and the downlink, we can calculate the theoretical RSSI change based on the PLF, and use it to calculate the mismatching directions in Section V-B.

Mismatching direction. As investigated in the experimental study, we try to use the mismatching direction to estimate the tag orientation. The mismatching direction, denoted as $\phi_m$, is defined as the spin angle of the antenna, when the projection of antenna polarization is orthogonal to the tag direction, i.e., $P_{\phi} \perp V_t$, and the RSSI value reaches its minimum value. Based on the orthogonal relationship, $P_{\phi} \cdot V_t = 0$, which is expanded as:

$$(MV_{\phi}) \cdot V_t = 0.$$  \(4\)

Then, we can deduce the mismatching direction $\phi_m$ as:

$$\tan \phi_m = \frac{(1 - z_{t}^2)x_{t} - x_{d}y_{t} - x_{d}z_{d}z_{t}}{x_{d}z_{d}x_{t} + y_{d}z_{d}y_{t} - (1 - z_{d}^2)z_{t}},$$  \(5\)

where $V_t = [x_t, y_t, z_t]^T$ is the tag orientation vector. According to Eq. (5), both the tag orientation $V_t$ and the antenna-tag direction $V_d$ are unknown, meaning that it is difficult to deduce the tag orientation $V_t$ based on the mismatching direction $\phi_m$ of only one tag. Besides, this equation also explains the difference of mismatching directions in Fig. 3(b) and Fig. 3(c), which is caused by the change of the antenna-tag direction $V_d$.

B. Modeling RSSI Pattern of Tag Array with Spin-antenna

Tag array deployment. Since it is difficult to estimate the tag orientation $V_t$ based on one tag, the key idea is to deploy multiple tags as a tag array to estimate the orientation of the tag array. However, the tag array is relatively small compared with the distance between the tag array and the antenna, thus, the antenna-tag directions $V_d$ of different tags are almost the same. Hence, the traditional tag array deployment with the same tag orientation cannot provide the more discrimination of tag orientation compared with one tag. Based on the understanding, we vary the orientations of tags in the tag array. Hence, even though one tag is misread due to mismatching, we can still read other tags with different orientations. Without loss of generality, in our system we attach $N$ tags on the plane as shown in Fig. 5. The orientation difference between adjacent tags is set to $\alpha = 2\pi/N$, which maximizes the orientation difference. Therefore, the differences of $\phi_m$ between adjacent tags are also maximized, which provides the maximum discrimination. Since the antenna-tag directions $V_d$ are similar for different tags in the tag array, we consider all tags are deployed at the same position with different orientations.

Definition of the tag array orientation. Basically, the 3D orientation is usually expressed with three angles, e.g., the Euler angles. For simplicity, we decompose the three angles of the orientation of the tag array into two parts. 1) The incline orientation represents the orientation of the tag array plane, which is denoted as the unit normal vector of the plane $N_i = [x_n, y_n, z_n]^T$. It is determined by two angles. 2) The rotation offset represents the rotation angle of the tag array inside the plane. Without loss of generality, we randomly choose a labeled tag $T_0$ in the tag array and use the angle difference $\delta$ from the reference orientation to represent it. The reference orientation is defined from the unit normal vector $N_i$ as:

$$V_r = \left[ \frac{y_n}{\sqrt{x_n^2 + y_n^2}}, \frac{-x_n}{\sqrt{x_n^2 + y_n^2}}, 0 \right]^T,$$  \(6\)

which is a unit vector along the intersecting line between the $XY$ plane and the tag plane as shown in Fig. 5. Then, we can use the tuple $\{N_i, \delta\}$ to define the orientation of the tag array.

Mismatching direction calculation. Then given the tuple $\{N_i, \delta\}$, the orientation of tag $T_i$ in the tag array can be calculated as:

$$V_i = R_i V_r.$$  \(7\)

Here, $R_i$ is the rotation matrix of tag $T_i$, which rotates around the normal vector $N_i$ with the rotation offset $\delta_i$. Particularly, $|\delta_i - \delta_0|$ is the angle difference between tag $T_i$ and $T_0$, whose value is the multiple of $\alpha$. Finally, based on the calculated orientation of tag $T_i$, we can rewrite Eq. (4) as:

$$(MV_{\phi_m}) \cdot (R_i V_r) = 0.$$  \(8\)

Therefore, we can calculate the mismatching direction $\phi_m$ from $V_{\phi_m}$, accordingly.

The mismatching direction of each tag in the tag array is more sensitive to the rotation offset $\delta$ compared with the tag plane orientation $N_i$ or the antenna-tag direction $V_d$. Since the mismatching directions in Eq. (5) are not linearly related to the tag array orientation and the antenna-tag direction, we use several experiments to examine the discrimination of mismatching directions in the orientation estimation of the tag array. By changing the tag plane orientation, i.e., $N_i$ and the antenna-tag direction, i.e., $V_d$, we vary the rotation offset $\delta$ from 0 to $\pi$ and calculate the corresponding $\phi_m$ based on Eq.
Mismatching directions estimation (radian)

Translation/Rotation estimation

V. System Design

A. Overview

The major objective of our work is to track the 3D motion of the tag array labeled objects with the new antenna deployment, i.e., Spin-antenna. By spinning the linearly polarized antenna, we can not only extract the most distinctive signal features due to the linear polarization; but also sufficiently suppress the ambient signal interference. Fig. 8 shows the system overview of Spin-antenna. We take as input both the RSSI and the phase stream. First, Preprocessing segments the signals into separated windows, calculates the mismatching directions, and calibrates the phase based on the RSSI variation, which produces the distinctive signal features. Then, Relative Direction Estimation uses the phase model in Section IV-C to estimate the relative direction of the antenna based on the calibrated phase values. After that, Coordinate Transformation transforms the relative direction of antenna in the local coordinate system to the global coordinate system based on the positions of the tag array in the previous window. Based on the transformed positions of the tag array, 3D Orientation Estimation finally estimates the rotation of the tag array based on the mismatching directions. Therefore, with the position and the orientation of the tag array in the consecutive window, we can estimate the corresponding translation and the rotation of the tag array, which are used to calibrate the 3D motion tracking in the following windows.

B. Preprocessing

In this section, we preprocess the raw phase/RSSI signals to extract the most distinctive signal features. Particularly, we segment the signals into separated windows based on the cycle of the spinning framework, then estimate the mismatching directions from the RSSI variation, and calibrate the phase
values based on the RSSI variation. We first segment the signals based on the spinning cycle and resample the raw signals. Since the linearly polarized antenna is symmetric based on the empirical study in Fig. 3, the spin angle of one cycle is 180° instead of 360°. Therefore, we segment the signals into windows, where the spin angles in each window range from 0° to 180°. Then, we resample the signals in each window with the cubic spline interpolation to tackle the random sampling problem in the RFID system. Specifically, based on the stable RSSI variation pattern studied in Section III-C, we set the RSSI value to −80dBm for the misreading tags. After the resampling, we have the RSSI vector \( S_{i,j} = \{s_1, \ldots, s_w\} \) and phase vector \( \Theta_{i,j} = \{\theta_1, \ldots, \theta_w\} \) of tag \( T_i \) in the \( j \)-th window, where \( w \) is the resampling window size.

Then, we estimate the mismatching directions from the RSSI variation by comparing the collected RSSI vector \( S_{i,j} \) with the theoretical RSSI trace. Formally, the mismatching direction \( \hat{\phi}_{m,i,j} \) of tag \( T_i \) can be calculated as:

\[
\hat{\phi}_{m,i,j} = \arg \min_{\phi_m, \in [0, \pi]} \sum_{k=1}^w (s_k - s_p,k)^2.
\]

(11)

Here, \( s_{p,k} \) is the \( k \)-th theoretical RSSI value calculated from \( \hat{\phi}_{m,j} \) based on the theoretical phase shift. The initial \( \phi_m \) equals to 90°, while the theoretical RSSI trace (red line) is totally different from the collected RSSI. By sliding the trace towards the left, we find the difference between the collected RSSI vector and the theoretical RSSI (original line) is minimum when \( \phi_m = 10^\circ \). Then, we get the estimated mismatching direction \( 10^\circ \).

After that, we calibrate the phase in each window according to the observation in Section III-C, the phase values around the mismatching direction are noisy due to the multipath effect. The basic idea is to calibrate the phase values based on the reliable RSSI collections around the matching direction. Fig. 9(b) shows the calibration flow with an example. We first select the phase collections where the collected RSSI value is close to the matching RSSI value. Then we use a weighted average algorithm to calibrate the phase in each window, where the weights are the corresponding RSSI values. In Fig. 9(b), we use the matching phase and the matching RSSI to calibrate the phase via weighted averaging. We use \( \theta_{i,j} \) to represent the calibrated phase of tag \( T_j \) in the \( j \)-th window.

### C. Relative Direction Estimation based on RF Phase

After the signal preprocessing, we use the calibrated phase values \( \theta_{i,j} \) of the tag array to estimate the direction of the antenna relative to the tag array. The relative topology is shown in Fig. 7, where the relative direction is based on the Local Coordinate System (LCS) of the tag array. According to the phase model in Section IV-C, the basic idea is to derive the direction of the antenna with respect to the tag array by comparing the collected phase difference with the theoretical phase difference. Particularly, given the calibrated phase difference \( \Delta \theta_{i,0,j} = \theta_{i,j} - \theta_{0,j} \) between the tag \( T_i \) and \( T_0 \) in the \( j \)-th window, we search for the antenna direction in the LCS, that maximizes the similarity between the computed phase difference \( \Delta \hat{\theta}_{i,0,j} \) and the collected phase difference \( \Delta \theta_{i,0,j} \). This antenna direction, denoted as \( \hat{\mathbf{V}}_d \) based on the phase model, is our estimate for \( \mathbf{V}_d \). The maximization formulation calculates the similarity of the phase differences based on the imaginary number to remove the periodical ambiguity of 2\( \pi \) as:

\[
\hat{\mathbf{V}}_d = \arg \max_{\mathbf{v} \in S^3} \sum_{i=1}^N \left| e^{j \Delta \hat{\theta}_{i,0,j}} - e^{j \Delta \theta_{i,0,j}} \right|^2,
\]

(12)

where \( J = \sqrt{-1} \) and \( S^3 \) represents the possible unit direction vectors in the 3D space. In regard to \( \Delta \hat{\theta}_{i,0,j} \), it can be calculated from Eq. (9) and Eq. (10), based on the preset tag array deployment. Fig. 10 uses an example to show the results of the maximization formulation, where \( N = 5 \). We find our method can uniquely determine the relation direction of the antenna based on the maximization formulation.

### D. Coordinate System Transformation

We then transform the estimated antenna direction \( \hat{\mathbf{V}}_d \) in the Local Coordinate System (LCS) to the global coordinate system (GCS), which is defined in Section III, and then estimate the position of tag array in the GCS. The basic idea is to estimate the possible position of tag array in the GCS based on the position of the tag array in the previous window and the phase change of each tag between consecutive windows. Specifically, given the phase change \( \Delta \theta_{i,j} = \theta_{i,j} - \theta_{i,j-1} \) of the tag \( T_i \), we search for the position of the tag array, that maximizes the similarity between the computed phase change \( \Delta \hat{\theta}_{i,j} \) and the collected phase change \( \Delta \theta_{i,j} \). This position of the tag array, denoted as \( \hat{C}_j \), is our estimate for \( C_j \). The optimization formulation can be written as:

\[
\hat{C}_j = \arg \min_{C_j \in A} \sum_{i=1}^N \left| e^{j \Delta \hat{\theta}_{i,j}} - e^{j \Delta \theta_{i,j}} \right|^2.
\]

(13)

Here, \( A \) is the candidate positions of the tag array based on \( C_{j-1} \) and the moving speed between the consecutive windows. To calculate the phase change \( \Delta \theta_{i,j} \), we first transform the position of the tag array in the LCS to the GCS. Then
we can calculate the theoretical phase values based on the actual tag positions in the GCS, which are further used to calculate the phase change $\Delta \hat{\phi}_{m,i}$. In regard to the coordinate transformation, since the direction from the antenna to the tag array is calculated as $\hat{V}_d = \hat{C}_f/|\hat{C}_f|$ in the GCS, which should represent the same vector with the estimated direction vector $\hat{V}_d$ in the LCS, we simply transform the coordinate of the tag array from the LCS to the GCS based on $\hat{V}_d$ and $\hat{V}_d^\prime$.

**E. 3D Orientation Estimation**

After determining the position of the tag array, we estimate the rotation angle of the tag array, which corresponds to the rotating angle around the estimated direction vector $\hat{V}_d$ in the GCS. The basic idea is to compare the calculated mismatching directions $\hat{\phi}_{m,i}$ with the theoretical mismatching directions $\phi_{m,i}$, which are calculated based on the model in Section IV-B. Specifically, we can calculate the rotation angle $\Delta \phi$ as:

$$
\Delta \phi = \frac{1}{N} \sum_{i=1}^{N} \left( \phi_{m,i} - \hat{\phi}_{m,i} \right).
$$

Then, according to the rotation angle $\Delta \phi$ and the estimated antenna-tag direction $\hat{V}_d$, we rotate the tag array around $\hat{V}_d$ with the angle of $\Delta \phi$. After the rotation, we can finally get the position and orientation of the tag array. Therefore, by connecting the consecutive windows, we can track the translation and the rotation of the tag array in the 3D space.

**VI. PERFORMANCE EVALUATION**

**A. Experimental Setup**

We have implemented a system prototype with the ImpinJ R420 Speedway RFID reader and the Laird PA912 linearly polarized antenna. As shown in Fig. 11, we design a spin framework, which can continuously spin the antenna around the spin axis and interrogate the tags simultaneously. The antenna spins 4 rounds per second in our system. We design three kinds of the tag array deployment with different numbers of tags, i.e., 3, 4 and 5, which separates the endpoints between adjacent tags to reduce the mutual interference. During the experiments, we vary the number of tags, the distance between the tag array and the antenna, and the multi-path environment to evaluate the performance of our system in the 3D motion tracking. The initial posture is known by default. For each specific setting, we move the tag array along one coordinate axis with the distance of 50cm to evaluate the tracking accuracy of the translation, and rotate the tag array around one coordinate axis with the angle of 90° to evaluate the tracking accuracy of the rotation. Particularly, we use two metrics to judge the accuracy: the translation error refers to the difference between the ground-truth translation and the estimated translation, and the rotation error refers to the angle difference between the ground-truth rotation and the estimated rotation. We use the OptiTrack system to capture the ground-truth of the translation and rotation with the high-speed camera.

**B. Overall Performance of 3D Motion Tracking**

Our solution can accurately track the translation with the average error of 13.6cm and track the rotation with the average error of 8.3°. We first show the overall tracking accuracy of our system with the CDF in Fig. 12(a) and Fig. 12(b). For the translation error in Fig. 12(a), the $Y$-axis outperforms the other two axes in tracking the movement of the tag array, because the translation along the $Y$-axis leads to more distinctive phase change compared with the translation along the $X$-axis or $Z$-axis. Overall, more than 80% experiment results achieve the translation error within 14.5cm along the $X$-axis, 5.8cm along the $Y$-axis and 13.4cm along the $Z$-axis. For the rotation error in Fig. 12(b), the tag array rotation achieves less than 4.8° error for 80% results, which is 6.2° smaller than the rotation error of the tag plane. The accuracy of the tag array rotation is based on the RSSI.
variation, which is caused by the polarization direction change of the spin-antenna. Overall, our solution can accurately track the translation and the rotation in the 3D space.

C. Impact of Number of Tags

The tracking error of both the translation and the rotation decreases when the number of tags increases. Then we compare the tracking accuracy with the three tag arrays to evaluate the impact of the number of tags. Fig. 12(c) shows the translation error with the number of tags. We find that as the number of tags increases from 3 to 5, the translation error along the X-axis decreases from 10.1cm to 6.6cm, and the translation error along the Z-axis decreases from 6.8cm to 5.7cm. The translation error along the Y-axis does not largely reduce, because all the tags share the similar phase change due to the translation along the Y-axis. For the rotation error in Fig. 12(d), we find that the error of the tag plane orientation decreases dramatically as we increase the number of tags. It is because the increment of the number of tags will increase phase pairs exponentially based on the phase model in Section IV-C, providing more phase differences for the estimation. Besides, the error of the tag array rotation maintains small when we increase the number of tags, indicating the RSSI variation is stable to estimate the orientation of each tag.

D. Impact of Distance between Tag Array and Antenna

The tracking error slightly increases as the distance between the tag array and the antenna increases. We evaluate the impact of the distance between the tag array and the antenna by conducting the experiments at the distance of 150cm, 250cm and 350cm away from the antenna, respectively. Fig. 12(e) shows the translation error along with the distance between the tag array and the antenna. It is found that the translation error along the Y-axis is as low as 2.2cm, when the distance is 150cm, and then increases to 4.1cm, when the distance increases to 350cm. Therefore, even if the distance between the tag array and the antenna indeed affects the translation error, our system can always keep the high tracking accuracy as the distance increases. For the rotation error in Fig. 12(f), the rotation errors are always below 15°, even if the actual error increases along with the increment of the distance between the tag array and the antenna.

E. Impact of Multi-path Effect

Our solution can accurately track the 3D motion in the heavy multi-path effect environment. Finally, to evaluate the robustness of the spin-antenna in the 3D motion tracking, we deploy multiple iron plates around the spin-antenna to generate the heavy multi-path effect. Fig. 12(g) and Fig. 12(h) show the corresponding translation error and the rotation error of the heavy and the light multi-path effect. It is found that even if the heavy multi-path effect leads to about 16cm translation error for both the X-axis and the Y-axis, we can still achieve the high accuracy for the Y-axis. Besides, we also find that the rotation error of the tag array rotation is also as small as 3.1° with the heavy multi-path effect. It indicates that the relative RSSI variation is stable enough to resist the interference of the multi-path effect. Therefore, the results prove that our system can still accurately track the 3D motion of the relatively complex environment with the heavy multi-path effect.

VII. CONCLUSION

In this paper, we propose to track the 3D motion of a tag array labeled objects via a spinning linearly polarized antenna. By modeling the relationship between the tag array and the spinning antenna, our system is able to sufficiently suppress the ambient signal interference and extract the most distinctive features. Based on the calibrated features, we design corresponding solutions to track the translation based on the phase values, and estimate the rotation of the tag array from the RSSI variation. The extensive experiment results on the real testbed show that our solution can achieve an average translation error of 13.6cm and an average rotation error of 8.3° in the 3D space.

ACKNOWLEDGMENTS

This work is supported in part by National Natural Science Foundation of China under Grant Nos. 61872174, 61832008, 61872173, 61802169; JiangSu Natural Science Foundation under Grant No. BK20180325. This work is partially supported by Collaborative Innovation Center of Novel Software Technology and Industrialization. Lei Xie is the corresponding author.

REFERENCES