

RF-Kinect: A Wearable RFID-based Approach Towards 3D Body Movement Tracking

CHUYU WANG, Nanjing University, CHN

JIAN LIU, Rutgers University, USA

YINGYING CHEN*, Rutgers University, USA

LEI XIE*, Nanjing University, CHN

HONGBO LIU, Indiana University-Purdue University Indianapolis, USA

SANGLU LU, Nanjing University, CHN

The rising popularity of electronic devices with gesture recognition capabilities makes the gesture-based human-computer interaction more attractive. Along this direction, tracking the body movement in 3D space is desirable to further facilitate behavior recognition in various scenarios. Existing solutions attempt to track the body movement based on computer version or wearable sensors, but they are either dependent on the light or incurring high energy consumption. This paper presents *RF-Kinect*, a training-free system which tracks the body movement in 3D space by analyzing the phase information of wearable RFID tags attached on the limb. Instead of locating each tag independently in 3D space to recover the body postures, RF-Kinect treats each limb as a whole, and estimates the corresponding orientations through extracting two types of phase features, *Phase Difference between Tags* (PDT) on the same part of a limb and *Phase Difference between Antennas* (PDA) of the same tag. It then reconstructs the body posture based on the determined orientation of limbs grounded on the human body geometric model, and exploits Kalman filter to smooth the body movement results, which is the temporal sequence of the body postures. The real experiments with 5 volunteers show that RF-Kinect achieves 8.7° angle error for determining the orientation of limbs and 4.4cm relative position error for the position estimation of joints compared with Kinect 2.0 testbed.

CCS Concepts: • **Networks** → Sensor networks; **Mobile networks**; • **Human-centered computing** → **Mobile devices**;

Additional Key Words and Phrases: RFID; Body movement tracking

ACM Reference Format:

Chuyu Wang, Jian Liu, Yingying Chen, Lei Xie, Hongbo Liu, and Sanglu Lu. 2018. RF-Kinect: A Wearable RFID-based Approach Towards 3D Body Movement Tracking. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 2, 1, Article 41 (March 2018), 28 pages. <https://doi.org/10.1145/3191773>

*Yingying Chen and Lei Xie are the co-corresponding authors, Email: yingche@scarletmail.rutgers.edu, lxie@nju.edu.cn.

Authors' addresses: Chuyu Wang, Nanjing University, State Key Laboratory for Novel Software Technology, 163 Xianlin Ave, Nanjing, 210046, CHN; Jian Liu, Rutgers University, Department of Electrical and Computer Engineering, North Brunswick, NJ, 08902, USA; Yingying Chen, Rutgers University, Department of Electrical and Computer Engineering, North Brunswick, NJ, 08902, USA; Lei Xie, Nanjing University, State Key Laboratory for Novel Software Technology, 163 Xianlin Ave, Nanjing, 210046, CHN; Hongbo Liu, Indiana University-Purdue University Indianapolis, Department of Computer, Information and Technology, Indianapolis, IN, 46202, USA; Sanglu Lu, Nanjing University, State Key Laboratory for Novel Software Technology, 163 Xianlin Ave, Nanjing, 210046, CHN.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Association for Computing Machinery.

2474-9567/2018/3-ART41 \$15.00

<https://doi.org/10.1145/3191773>

Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, Vol. 2, No. 1, Article 41. Publication date: March 2018.

1 INTRODUCTION

The gesture-based Human-Computer Interaction (HCI) embraces an increasing number of practical usage enabled by the growing popularity of electronic devices with gesture recognition capabilities. Recent survey reveals that the global gesture recognition market is anticipated to reach USD 48.56 billion by 2024 [4]. In particular, the success of Microsoft Kinect [8] in tracking human gestures in gaming consoles has induced many emerging applications to adopt gesture recognition solutions in the fields like healthcare, smart homes, mobile robot control, *etc.* For example, numerous applications are developed to monitor human's well-being based on their activities (such as fitness, drinking, sleeping, *etc.*) with either wearable devices or smartphones. The success of gesture and activity recognition leads to a growing interest in developing new approaches and technologies to track the body movement in 3D space, which can further facilitate behavior recognition in various occasions, such as VR gaming, mobile healthcare, and user access control.

Existing solutions for body movement recognition fall into three main categories: (i) Computer vision-based solutions, such as Kinect and LeapMotion [5, 8], leverage the depth sensors or infrared cameras to recognize body gestures and allow the user to interact with machines in a natural way. However, these methods suffer from several inherent disadvantages of computer vision including light dependence, dead corner, high computational cost, and ambiguity of multi-people. (ii) Sensor-based solutions, such as the smartwatch and wristband [3], are designed to track the movement of the limbs based on the accelerator or gyroscope readings. But these systems usually require the user to wear different kinds of sensing devices, which present short life cycles due to the high energy consumption. Further, there are also some products (*i.e.*, Vicon [6]) integrating the information from both cameras and wearable sensors to accurately track the body movement, however the high price of the infrastructure is not affordable for many systems. (iii) Wireless signal-based solutions [17, 25] capture the specific gestures based on the changes of some wireless signal features, such as the Doppler frequency shift and signal amplitude fluctuation. But only a limited number of gestures could be correctly identified due to the high cost of training data collection and the lack of capabilities for multi-user identification.

With the rapid development of RFID techniques [34, 41], RFID tag now is not only an identification device, but also a low power battery-free wireless sensor serving for various applications, such as the localization and motion tracking. Previous studies, such as Tagoram [41] and RF-IDraw [35], could achieve cm-level accuracy on tracking an individual RFID tag in 2D space (*i.e.*, a tagged object or finger). Further, Tagyro [39] could accurately track the 3D orientation of objects attached with an array of RFID tags, but it only works for objects with the fixed geometry and rotation center. However, due to the complicated body movement involving multiple degrees of freedom, the problem of 3D RFID tag tracking associated with the human body movement, including both limb orientation and joint displacement (*e.g.*, elbow displacement), remains elusive.

Inspired by these advanced schemes, we explore the possibility of tracking the *human body movement* in 3D space via RFID system. In particular, we propose a *wearable RFID-based approach* as shown in Figure 1, which investigates new opportunities for tracking the body movement by attaching the lightweight RFID tags onto the human body. *Wearable RFID* refers to the gesture recognition towards the human body wearing multiple RFID tags on different parts of the limbs and torso. In actual applications, these tags can be easily embedded into the fabric [11], *e.g.*, T-shirts, with fixed positions to avoid the complicated configurations. During the process of the human motion, we are able to track the human body movement, including both the rigid body [7] movement (*e.g.*, the torso movement) and non-rigid body movement (*e.g.*, the arm/leg movement), by analyzing the relationship between these movements and the RF-signals from the corresponding tag sets. Due to the inherent identification function, *wearable RFID* solves the distinguishing problem of tracking multiple subjects in most device-free sensing schemes. For example, in regard to tracking the body movement of multiple human subjects, different human subjects or even different arms/legs can be easily distinguished according to the tag ID, which is usually difficult to achieve in the computer vision or wireless-based sensing schemes. Even RF-IDraw [35] makes the first

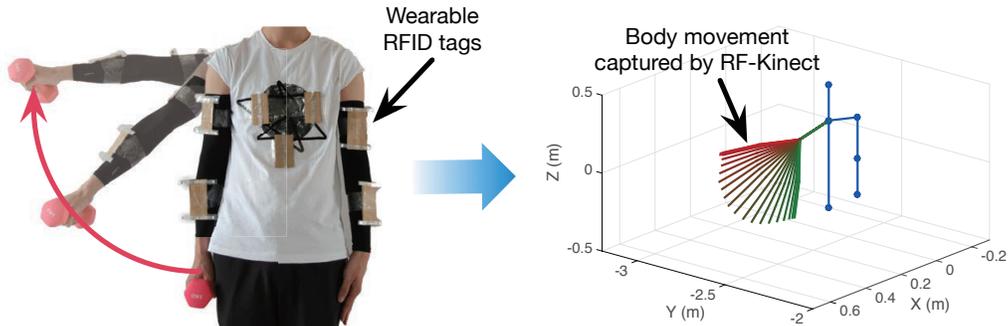


Fig. 1. RF-Kinect: Tracking the body movement based on wearable RFID tags.

attempt to track the finger by wearing one RFID tag on the finger, we are the first to systematically explore the usage of wearable RFID on tracking the whole body movement (*i.e.*, the limb orientation and joint displacement) in 3D space, which is more complicated and challenging.

In order to investigate the applicability of wearable RFID, we present *RF-Kinect* which consists of multiple wearable RFID tags and one dual-antenna RFID reader measuring the RF signal variations from these tags. In particular, RF-Kinect focuses on tracking the complicated 3D limb orientation and movement with multi-degree of freedom other than the simple trajectory of the finger or hand, which has been well studied in previous work [27, 31, 35]. The key novelty of RF-Kinect lies in (i) *training-free* and (ii) *minimum hardware requirements*. First, it is impractical to traverse numerous body movements (*e.g.*, [13, 17]) to build a complete training dataset. Thus, we build a geometric model of the human body to assist the body movement tracking with little efforts on training data collection. Second, it is also not applicable to place a large number of antennas on RFID readers around the user, making the tracking of the body movement cumbersome. The existing RFID-based localization systems require either at least three static antennas or a moving antenna [34–36, 41] to accomplish the task, posing a big challenge on the hardware design for the body movement tracking. Therefore, we aim to design RF-Kinect with the minimum hardware requirement by leveraging a single dual-antenna RFID reader.

The basic idea of RF-Kinect is to derive the limb orientation by leveraging the phase information of RF signals collected from multiple wearable RFID tags, and then construct the body movement, which is represented as a temporal sequence of the limb orientation estimations, grounded on the predefined human body geometric model. Specifically, RF-Kinect extracts two types of phase features to perform the limb orientation estimation: (i) *Phase Difference between any two Tags (PDT)* attached to the same part of a limb (*e.g.*, the upper arm), and (ii) *Phase Difference between the two Antennas (PDA)* of the same tag. By regarding the two tags on one skeleton as a rigid body, we can formulate PDT as a function of the skeleton orientation with respect to the transmitting direction of the antenna. The possible orientations derived from a single antenna thus form a conical surface in 3D space, where the apex of the cone is the rotation center of the limb [2]. When two antennas are employed to perform the orientation estimation, the possible range of orientations can be largely reduced by examining the overlapping range of two conical surfaces. However, two antennas still lead to two ambiguity orientations on the mirroring sides. Therefore, we further model the relationship between two rigid bodies to filter out the ambiguity. Particularly, we calculate the relative distance between the tags on different skeletons to describe the relationship. Since the relative distance describes the relative postures between two skeletons, the ambiguity orientation on the mirroring side can be thus filtered out due to the unmatched relative distances. Finally, we can correctly estimate the orientations of each skeleton of the limb.

The key contributions in this work are summarized as follows: 1) To the best of our knowledge, we are the first to propose the wearable RFID research and systematically investigate the applicability of it by presenting

RF-Kinect. It is the first training-free and low-cost human body movement tracking system, including both the *limb orientation* and *joint displacement*, by leveraging multiple wearable RFID tags, and it overcomes many drawbacks on existing light-dependent works. 2) We demonstrate that RF-Kinect could accurately track the 3D body movement other than simply tracking one joint on the body, with the minimum hardware requirements involving only a dual-antenna RFID reader and several low-cost wearable RFID tags. 3) Instead of locating the absolute position of each joint for tracking, we regard the human body as the combination of several rigid bodies (*i.e.*, skeletons) and use a kinematic method to connect each skeleton as the human body model. Then, we exploit the features PDT and PDA to estimate the orientations of each skeleton and use the relative distances to measure the relationship between different skeletons for tracking. 4) The fast adoption and low-cost deployment of RF-Kinect are also validated through our prototype implementation. Given the groundtruth from the Kinect 2.0 testbed, our systematic evaluation shows that RF-Kinect could achieve the average angle and position error as low as 8.7° and 4.4cm for the limb orientation and joints' position estimation, respectively.

2 RELATED WORK

Existing studies on the gesture/posture recognition can be classified into three main categories:

Computer Vision-based. The images and videos captured by the camera could truthfully record the human body movement in different levels of granularity, so there have been active studies on tracking and analyzing the human motion based on the computer vision. For example, Microsoft Kinect [8] provides the fine-grained body movement tracking by fusing the RGB and depth image. Other works try to communicate or sense the human location and activities based on the visible light [15, 16, 18]. LiSense [20] reconstructs the human skeleton in real-time by analyzing the shadows produced by the human body blockage on the encoded visible light sources. It is obvious that the computer vision-based methods are highly light-dependent, so they could fail in tracking the body movement if the line-of-sight (LOS) light channel is unavailable. Besides, the videos may incur the privacy problem of the users in some sensitive scenarios. Unlike the computer vision-based approaches, RF-Kinect relies on the RF device, which can work well in most Non-line-of-sight (NLOS) channel environments. Moreover, given the unique ID for each tag, it can also be easily extended to the body movement tracking scenario involving multiple users.

Motion Sensor-based. Previous research has shown that the built-in motion sensors on wearable devices can also be utilized for the body movement recognition [19, 44]. Wearable devices such as the smartwatch and wristband can detect a variety of body movements, including walking, running, jumping, arm movement etc., based on the accelerometer and gyroscope readings [22, 23, 33, 40, 45]. For example, ArmTrack [30] proposes to track the posture of the entire arm solely relying on the smartwatch. However, the motion sensors in wearable devices are only able to track the movement of a particular part of the human body, and more importantly, their availability is highly limited by the battery life. Some academic studies [32] and commercial products (e.g., Vicon [6]) have the whole human body attached with the special sensors, and then rely on the high-speed cameras to capture the motion of different sensors for the accurate gesture recognition. Nevertheless, the high-speed cameras are usually so expensive that are not affordable by everyone, and the tracking process with camera is also highly light-dependent. Different from the above motion sensor-based systems, RF-Kinect aims to track the body movement with RFID tags, which are battery-free and more low-cost. Moreover, since each RFID tag only costs from 5 to 15 U.S. cents today, such price is affordable for almost everyone, even if the tags are embedded into clothes.

Wireless Signal-based. More recently, several studies propose to utilize wireless signals to sense human gestures [10, 17, 25, 35, 37, 38, 42, 43]. Pu *et al.* [25] leverage the Doppler influence from Wi-Fi signals caused by body gestures to recognize several pre-defined gestures; Kellogg *et al.* [17] recognize a set of gestures by analyzing the amplitude changes of RF signals without wearing any device; Adib *et al.* [9] propose to reconstruct a human

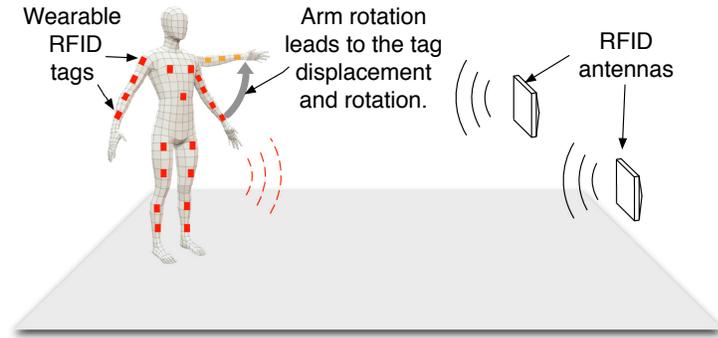


Fig. 2. RF-Kinect illustration of the RFID-based human body movements tracking.

figure by analyzing the RF signals' reflections through walls and occlusions, thereby accurately locating each part of the human body. Nowadays, as the rapid development of RFID-based localization techniques [28, 34, 41], more systems are developed to sense human activities based on RFID. Wang *et al.* [35] recover the moving trace of the tagged finger on a surface plane based on the AoA model. Shanguan *et al.* [29] tracks the tagged object in the 2D plane for user feedbacks based on only one antenna. But these methods only work in 2D space by tracking a rigid body, and thus are not suitable for tracking the complicated movement of the human body. Lin *et al.* [21] track the motion status of the tagball based on the phase variation read from the attached tags. Tagyro [39] estimates the orientation of passive objects that have the constant geometric by analyzing the Phase Difference of Arrival of attached tags. Ding *et al.* [13] aim to detect the fitness gestures leveraging the Doppler profile extracted from the phase trend of RF signals. These RFID-based methods mainly focus on estimating the position/orientation of one single passive rigid body or recognizing the gestures via pattern matching. However, RF-Kinect is designed to track the *whole body movement* through a model-based approach, which involves several related rigid bodies (*i.e.*, skeletons) and thus is more challenging for the design of the model-based approach.

3 APPLICATIONS & CHALLENGES

In this section, we first present the application scenario of RF-Kinect, and introduce the preliminaries of tracking human body movements using RF signals. We then describe the main challenges of the proposed RF-Kinect.

3.1 RF-Kinect Application Scenario

The wireless information gathered from wearable RFID tags opens a new research opportunity for developing gesture recognition systems and supporting related applications. RF-Kinect is such a system aiming to track *human body movements* based on the RF signals emitted from the wearable RFID tags attached to the human body. Taking the *Virtual reality (VR)* gaming as one example, we can utilize RF-Kinect to recognize the user gestures during the game, in the meanwhile, RF-Kinect can also identify specific users based on the wearable tag IDs at any time. Therefore, RF-Kinect can easily support multi-player games by identifying the users from the tag IDs and automatically reload the gaming process for each user from the tag ID. In the contrast, even traditional vision-based approaches can provide good accuracy in the games, they usually need to manually configure for different users and may also suffer from the interference of surrounding people, leading to bad user experience. *Personal fitness*, as another example, could also rely on RF-Kinect to associate the recognized activities with the subject for the fitness monitoring. Due to the energy harvesting capability of the wearable RFID tags from backscattered signal, RF-Kinect could operate for a long term without the battery supply issues

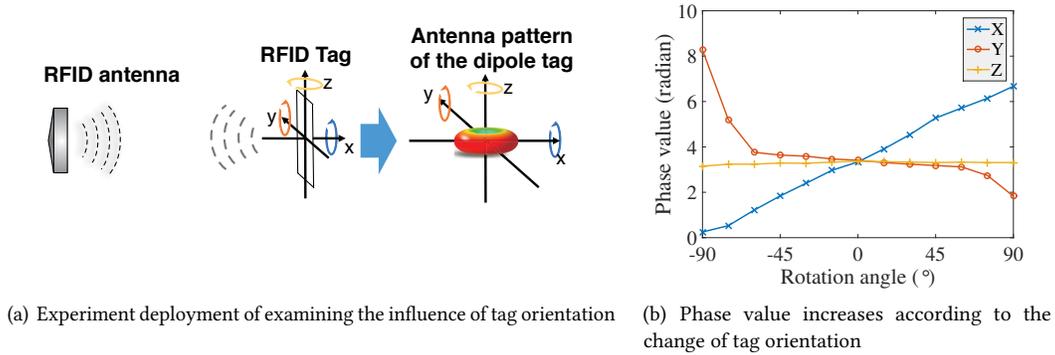


Fig. 3. Phase VS tag orientation: phase value is related to the tag orientation.

like wearable sensors. Compared with the vision-based approaches, RF-Kinect can efficiently filter out the other users from the tag IDs and usually work well when some objects block the line of sight path.

As shown in Figure 2, RF-Kinect utilizes a dual-antenna RFID reader to continuously scan the wearable RFID tags attached on the user (e.g., on the clothes) for the body movement tracking. In actual applications, we embed the tags into the T-shirts with fixed positions to avoid the complicated configurations. The changes of user's posture (e.g., the arm rotation) lead to the displacement and rotation of the wearable tags accordingly, thus producing unique RF signal patterns. Specifically, during each scan of RFID reader, the RF signals collected from multiple wearable RFID tags are combined to estimate the orientation of each limb (e.g., the upper arm, lower arm) and the position of each joint (e.g., the elbow, wrist), and finally the body posture will be successfully reconstructed. By concatenating the body postures derived from multiple scans, the entire body movement will be uniquely determined. To get rid of the exhaustive training efforts on covering all body movements, a *training-free* framework thus is essentially needed to reduce the complexity of the body movement tracking.

3.2 Preliminaries

In order to track the body movement, we need to identify some reliable RF signal features on distinguishing different postures and corresponding changes. There are several RF signal features, such as phase, RSSI and reading rate, available from the RFID system. According to the recent studies [21, 28, 35, 41], the phase information is proved to be an optimal choice than other features for the localization and many other sensing applications. In particular, the phase indicates the offset degree of the received signal from the original transmitting signal, ranging from 0 to 360 degrees. Assuming the distance between the antenna and tag is d , the phase θ can be represented as:

$$\theta = (2\pi \frac{2d}{\lambda} + \theta_{dev}) \mod 2\pi, \quad (1)$$

where λ is the wavelength, and θ_{dev} is the system noise caused by factorial imperfection of tags.

Since the body movement can unavoidably lead to the change of the tag orientation in 3D space, we first conduct controlled experiments to study the influence of the tag orientation on the phase as illustrated in Figure 3(a). The RFID tag spins 180° on a fixed spot along three different axes before an RFID reader, and the corresponding phase change is presented in Figure 3(b). We find that the phase changes linearly as the tag rotating along the X-axis, and remains stable along the Z-axis. When rotating along the Y-axis, most phase are similar except the perpendicular direction (i.e., -90° and 90°). We observe the similar phase variation trend when conducting the same experiment with different tags placed at different locations. The reason behind such phenomenon is that the RFID tags are commonly equipped with the linear-polarized dipole antenna, while the reader antenna works

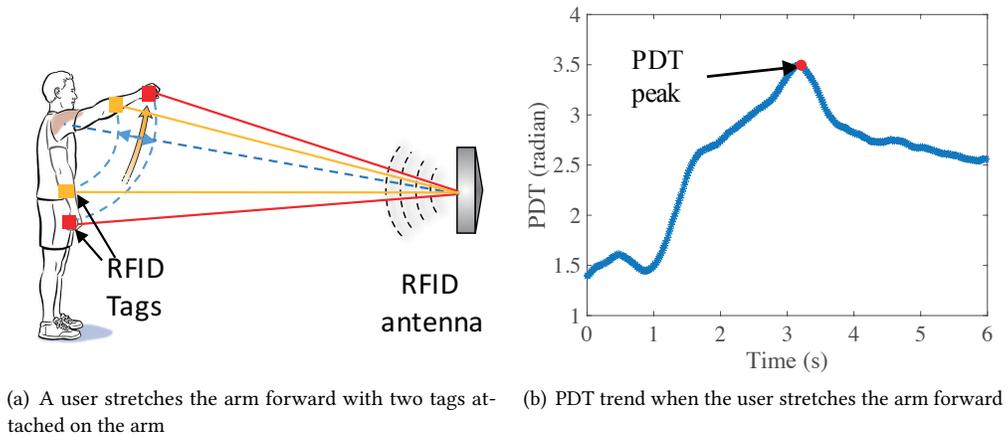


Fig. 4. Preliminary study of PDT trend when the tag-attached user stretches the arm forward.

at the circular polarization mode, which is compatible with two perpendicular polarizations with 90° phase difference, to identify the tags of different orientations. The rotation of RFID tag along the X-axis will change the polarization of the dipole antenna, and thereby affect the phase measurements due to the changes on the electromagnetic coupling between the tag and reader antenna. As a result, rotating 180° along the X-axis leads to 2π changes on the phase measurement, whereas rotating along the Y or Z-axis leads to no phase change due to the constant electromagnetic coupling. Note that when the tag rotates 90° along Y-axis, the tag faces the direction of the minimum gain on the RFID reader antenna, thus producing erroneous phase measurements due to multi-path effects. The above observation implies that the tag orientation has a significant impact on the extracted phase, and results in the ambiguity on body movement tracking.

In order to eliminate such negative impacts from tag orientation changes, we utilize the Phase Difference between Tags (PDT) to track the body movement. Specifically, we deploy multiple RFID tags in parallel and then measure the phase difference between these tags. Since all the tags have the same phase offset due to the consistent orientation, we can cancel such phase offset via the phase difference between different tags. Moreover, even though the tags have slight different orientations due to deployment or manufacturing error, the phase offset due to the rotation can still be canceled, because these tags are under the same rotation and hence have the same phase offset. To validate its effectiveness, we further conduct another experiment with the posture where the user stretches his/her arm forward as shown in Figure 4(a). Two tags with the same orientation are attached on the lower arm with fixed distance and the corresponding phase difference (PDT) is presented in Figure 4(b). We find that the PDT first increases and then slightly decreases. It coincides with the varying trend of distance difference (the difference between the red and yellow line) with respect to the two tags for one antenna, which is small at first and increases to the maximum value as the arm points to the antenna (e.g., the blue dash line). Therefore, it is feasible to track the body movement based on the PDT.

3.3 Challenges

To achieve the goal of the accurate and training-free body movement tracking with minimum hardware support, we identify three key challenges as follows:

Tracking with a Dual-antenna RFID Reader. Given the dual-antenna RFID reader to facilitate the minimum hardware requirement, it is a challenging task to track the human body movement. Existing RFID-based localization methods (e.g., [26, 41]) are not applicable for our problem, since they require at least three antennas

or moving antennas to locate a target tag in 2D environment. Other related studies such as [27] and [29] can locate a tagged object with two antennas or only one antenna, but they only work in 2D plane and they only track one object by attaching one or more tags on it. For the complex body movement in 3D space, it is not applicable to the motion tracking in our application scenario. Thus, a dual-antenna-based solution needs to be proposed to facilitate the 3D body movement tracking.

Imperfect Phase Measurements. Unlike previous RFID-based localization studies [35, 41] that track the tag movement in 2D space, our work aims to achieve a more challenging goal, i.e., tracking the movement in 3D space. So it poses even higher requirements on the phase measurements. There are multiple factors that may affect the uniqueness and accuracy of phase measurements related to the body movement. According to our preliminary study, the phase change of the RF signal is determined by both the tag-antenna distance and the tag orientation. Moreover, both the water-rich human body and the muscle deformation during the body movement may also affect the phase measurements from RF signals. All the above factors together make it much harder to track the human body movement in 3D space leveraging the phase information in RF-signals.

Training-free Body Movement Tracking. Existing studies on gesture tracking usually spend significant efforts on training the classification model by asking the users to perform each specific gesture multiple times [13, 25]. However, the number of gestures that can be recognized highly relies on the size of the training set, so the scalability to unknown gestures is greatly thwarted. Some other methods [29, 31, 35] are designed to recover the trace of a rigid body (e.g., the finger and box) from the signal models, but they are not suitable for the complex human body, which consists of several rigid bodies. In order to identify diverse gestures or postures flexibly of the complex human body, it is critical to develop a body movement tracking system that does not rely on any training dataset.

4 SYSTEM DESIGN

In this section, we first introduce the architecture of our RF-Kinect system, and then present the building modules of RF-Kinect for tracking the 3D body movement.

4.1 System Architecture

The basic idea of RF-Kinect is to derive the body posture in each scanning round by analyzing the RF signals from the wearable RFID tags attached on the limbs and chest, and then reconstruct the body movement from a series of body postures in consecutive scans. Figure 5 illustrates the architecture of RF-Kinect. We first extract the phase information of M RFID tags from two antennas in consecutive scanning rounds as *Phase Stream*, where all the attached tags are read in each scanning round. Then the system is initialized by requiring the user to stand still with his/her arms hanging down naturally. As a perfect rigid object, the tags on the chest enable *Body Position/Orientation Estimation* module to determine the position and facing orientation of the user relative to the antennas based on a model-based approach in the previous work (e.g., [21]). Then, *Coordinate Transformation* module converts the relative positions of the antennas into the Skeleton Coordinate System (SCS), which is defined based on the human body geometric structure in Section 4.2, so that the coordinates of both the tags and antennas could be expressed properly. Based on the coordinates of the antennas and tags attached on the user body when the user stands still, the theoretical phase value of each tag is calculated from Eq. (1). *Phase Deviation Elimination* module then computes the phase offset between the theoretical and the measured phase value, which is used to eliminate the phase deviation in the following biased phase stream.

After the above preprocessing, *Phase Difference Extraction* module extracts two phase related features from the RF signal measurements in each scanning round: (i) Phase Difference between any two Tags (PDT) attached to the same part of a limb (e.g., the upper arm), and (ii) Phase Difference between the two Antennas (PDA) of the same tag. The two phase related features are then utilized to estimate the limb postures based on the *3D Limb*

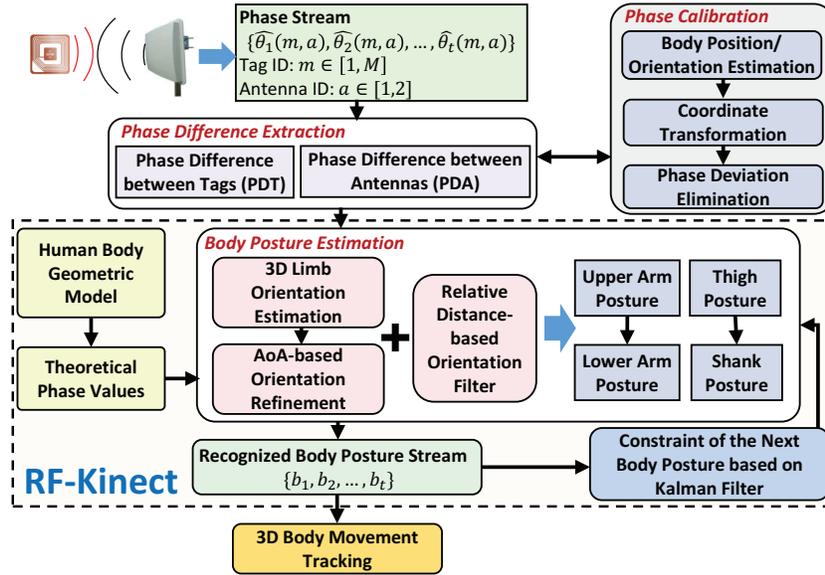


Fig. 5. System architecture of training-free RF-Kinect.

Orientation estimation, AoA-based Orientation Refinement and Relative Distance-based Orientation Calibration methods in the *Body Posture Estimation*. The first two methods determine the limb postures by comparing the extracted PDT/PDA with the theoretical PDT/PDA derived from the *Human Body Geometric Model*. Moreover, *Relative Distance-based Orientation Filter* removes the impossible orientations by measuring the relationship between different skeletons, which shrinks the searching range in the orientation estimation. Specifically, the arm posture estimation starts from the upper arm to the lower arm, while the leg posture estimation follows the order from the thigh to the shank. Then, the individual postures estimated from multiple scanning rounds construct a *Recognized Body Posture Stream* to represent the Body Movement, which is further smoothed with Kalman Filter. After smoothing, the positions of each tag can be calculated from the estimated body posture. Then we compute the theoretical phase values of each tag and extract the theoretical PDA as the constraint condition to calibrate the next body posture estimation. Finally, the 3D body movement is reconstructed accordingly, and can be applied to many interesting applications, such as gaming, healthcare, etc.

4.2 Human Body Geometric Model

Before diving into the details of the body movement tracking, we first introduce the human body geometric model in RF-Kinect. Inspired by the robotics studies, which model the human arm into 7 rotational degrees of freedom (DoF) [24], we use 4 of 7 DoFs to model the joints of a single arm by ignoring the other 3 DoFs on the wrist, which beyond the detection capability of our system. Similarly, we extend the method to model the joints of the leg with 3 DoFs.

Figure 6 illustrates the model of the right side of human body, while the left side follows a similar model. We use red and blue arrows to indicate the DoFs on the arm and leg, respectively. Specifically, ϕ_1, ϕ_2 and ϕ_3 are the 3 DoFs on the shoulder joint, which correspond to flexion/extension, abduction/adduction and internal/external rotation of the shoulder, respectively, and ϕ_4 represents flexion/extension on the elbow joint. Here, $\phi_1 = \phi_2 = \phi_3 = \phi_4 = 0^\circ$ refers to the posture where the arm is naturally hanging down with the palm facing forward. When the user lifts up one arm with bent elbow as shown in Figure 6, both ϕ_1 and ϕ_4 will change accordingly. Specifically,

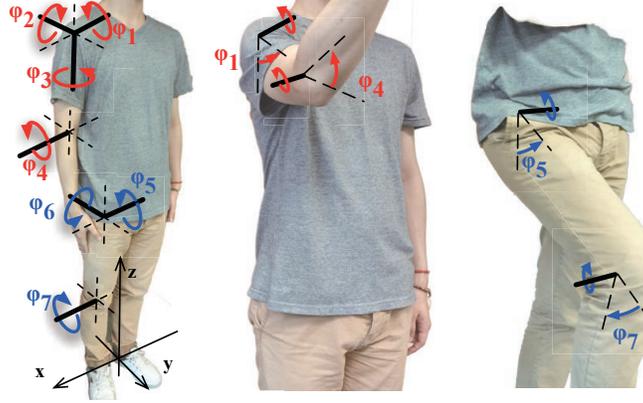


Fig. 6. Human body geometric model with rotation angle on each joint.

ϕ_1 represents the lift angle of the upper arm and ϕ_4 represents the bent angle of the elbow. In fact, ϕ_1 and ϕ_2 together determine the orientation of upper arm in 3D space, while ϕ_3 and ϕ_4 are for the lower arm orientation. It is interesting that ϕ_3 , the DoF on the shoulder, does not affect the orientation of the upper arm, but the lower arm instead, because ϕ_3 measures the angle of internal/external rotation around the arm. Further, ϕ_5 , ϕ_6 and ϕ_7 are the 3 DoFs on the leg. Since the lower body can be modeled in a similar way as the upper body, we will focus on the upper body to demonstrate the human body geometric model.

Given the length of upper arm l_u and lower arm l_l , the positions of elbow and wrist are determined by the rotation values of ϕ_1 , ϕ_2 , ϕ_3 and ϕ_4 in the Skeleton Coordinate System (SCS). Figure 6 illustrates the SCS in our system - the plane where the user's torso locates (i.e., the chest) serves as the XZ plane, and the line emanating from the shoulder in the frontward direction indicates the Y axis in the SCS. The midpoint between two feet is the origin of the SCS. Therefore, according to the mechanism model with Denavit-Hartenberg transformation [12], we can express the posture of the arm by calculating the position of elbow and wrist with ϕ_1 , ϕ_2 , ϕ_3 and ϕ_4 . For example, since the orientation of upper arm is determined by ϕ_1 and ϕ_2 , the position of elbow p_e is a function of ϕ_1 , ϕ_2 and l_u as follows:

$$p_e = p_s + f(\phi_1, \phi_2, l_u) = p_s + l_u \begin{pmatrix} \sin \phi_2 \\ \cos \phi_2 \times \sin \phi_1 \\ -\cos \phi_1 \times \cos \phi_2 \end{pmatrix}, \quad (2)$$

where p_s represents the position of shoulder and function $f(\cdot)$ calculates the vector pointing from the shoulder to the elbow. Similarly, the position of wrist p_w can be represented as:

$$p_w = p_e + g(\phi_1, \phi_2, \phi_3, \phi_4, l_l), \quad (3)$$

where $g(\cdot)$ computes the vector pointing from the elbow to the wrist.¹

4.3 Body Posture Estimation

As the core module of RF-Kinect system, three key techniques for estimating the body posture, *3D Limb Orientation Estimation*, *AoA-based Orientation Refinement* and *Relative Distance-based Orientation Calibration*, are proposed in this subsection.

¹The details of function $g()$ can be found in [12].

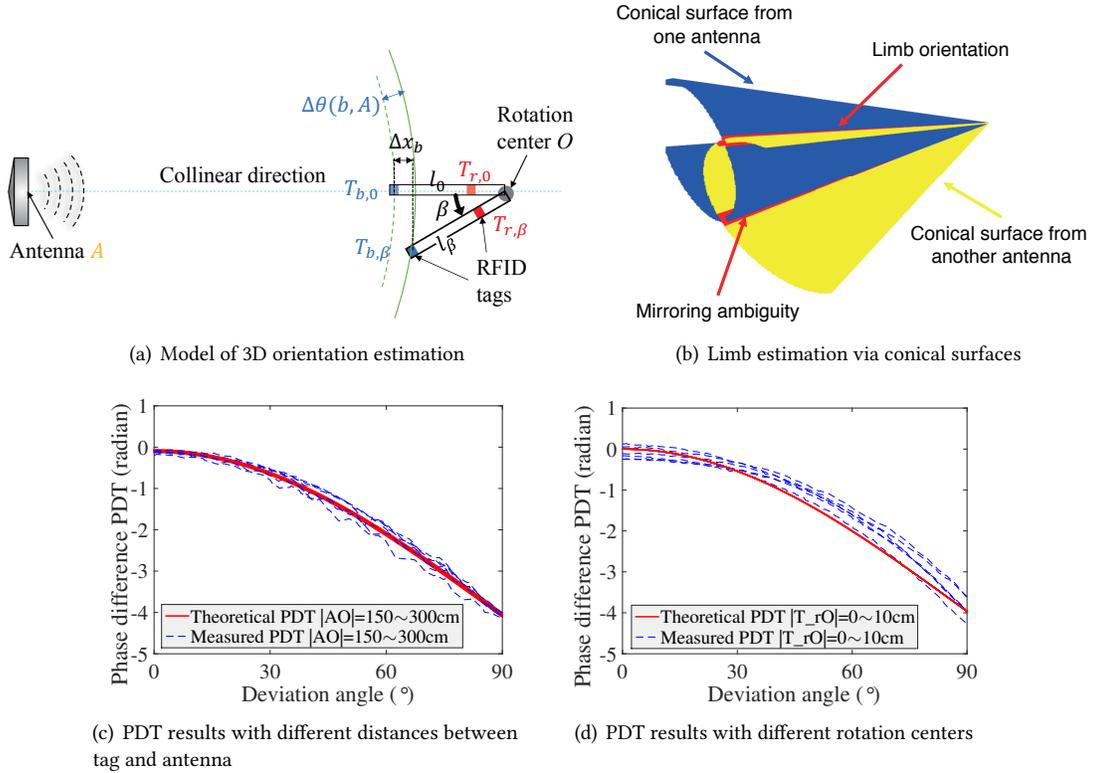


Fig. 7. Basic idea of estimating 3D orientation and its experimental results.

4.3.1 3D Limb Orientation Estimation. Given the human body geometric model, we do not need to locate each tag individually to construct the limb posture, which is also difficult for a dual-antenna RFID system. Instead, we treat the wearable tags on the upper arm or lower arm as an integrity for the orientation estimation. Specifically, the limb is viewed as a straight line attached with several wearable RFID tags. Assuming the position of user shoulder is known to RF-Kinect, we can recover the posture of the whole arm if the 4 DoFs of the arm are known (i.e., $\phi_1, \phi_2, \phi_3, \phi_4$) by examining the RF signals emitted from all the wearable tags.

Specifically, RF-Kinect utilizes *Phase Difference between any two Tags (PDT)* attached on the same part of a limb (e.g., the upper arm) to estimate the 3D limb orientation. A simple example of the limb rotation in 2D space is given in Figure 7(a) and Figure 7(b) to demonstrate our basic idea. A limb attached with two tags $T_{b,0}$ and $T_{r,0}$ is modeled as a straight line l_0 , while the corresponding joint is modeled as the rotation center O of the line. Define the direction between RFID antenna A and rotation center O as the *collinear direction*, which means the limb and antenna are collinear along this direction, the limb orientation thus can be represented as a *deviation angle* with respect to the collinear direction. For example, the position of limb l_β can be represented as the deviation angle β . Suppose the initial posture of limb is l_0 , and it rotates around center O to l_β . The deviation angle changes from 0 to β accordingly, leading to the displacements of the two tag, $T_{b,0} \rightarrow T_{b,\beta}$ and $T_{r,0} \rightarrow T_{r,\beta}$. Hence, the phase

difference between the two positions (i.e., l_0 and l_β) at antenna A is derived based on Eq. (1) as:

$$\begin{aligned}\Delta\theta(b, A) &= \theta_\beta(b, A) - \theta_0(b, A) = 4\pi \frac{|AT_{b,\beta}| - |AT_{b,0}|}{\lambda} + \omega(\beta) \pmod{2\pi}, \\ \Delta\theta(r, A) &= \theta_\beta(r, A) - \theta_0(r, A) = 4\pi \frac{|AT_{r,\beta}| - |AT_{r,0}|}{\lambda} + \omega(\beta) \pmod{2\pi},\end{aligned}\quad (4)$$

where $|\cdot|$ measures the Euclidean distance and $\omega(\beta)$ calculates the phase offset due to the change of the tag orientation. The phase deviation θ_{dev} of each tag could be canceled during the above calculation. Theoretically, if the distance between rotation center O and RFID reader antenna A , $|AO|$, is much larger than $|T_{b,0}O|$, $|AT_{b,\beta}| - |AT_{b,0}|$ is approximated to the distance change Δx_b along the collinear direction as follows:

$$\Delta x_b = |T_{b,0}O|(1 - \cos \beta) \approx |AT_{b,\beta}| - |AT_{b,0}| = \left(\frac{\Delta\theta(b, A) - \omega(\beta)}{2\pi} + k \right) \frac{\lambda}{2} \text{ where } k = 0, \pm 1, \dots \quad (5)$$

Similarly, we also derive the change on the distance between the red tag Δx_r and antenna A as:

$$\Delta x_r = |T_{r,0}O|(1 - \cos \beta) \approx |AT_{r,\beta}| - |AT_{r,0}| = \left(\frac{\Delta\theta(r, A) - \omega(\beta)}{2\pi} + k \right) \frac{\lambda}{2} \text{ where } k = 0, \pm 1, \dots \quad (6)$$

To remove the phase offset resulted from the changes of tag orientation $\omega(\beta)$, we combine the phase changes with respect to the two tags as:

$$|T_{b,0}T_{r,0}|(1 - \cos \beta) \approx \left(\frac{\Delta\theta(b, A) - \Delta\theta(r, A)}{2\pi} + k \right) \frac{\lambda}{2} \text{ where } k = 0, \pm 1, \dots \quad (7)$$

Therefore, given the distance between the two tags $|T_{b,0}T_{r,0}|$, deviation angle β could be obtained based on the PDT $\Delta\theta(b, A) - \Delta\theta(r, A)$.

Once the deviation angle β with respect to one RFID antenna is obtained, the possible postures of the limb form a conical surface around the collinear direction, where the apex of the cone is rotation center O . Figure 7(b) presents a sample for the estimation. Given the deviation angles with respect to two antennas located at two different positions, we can determine the posture of the limbs by finding the common places of the two conical surfaces. Since we use only two antennas, there is an ambiguity on the mirror side with respect to the line between the two antennas.

We validate the feasibility of using PDT for the deviation angle estimation via the following two basic experiments. 1) In the first experiment, we vary the distance of $|AO|$ from 150cm to 300cm, and in the meanwhile spin the line l_0 from 0° to 90° as shown Figure 7(a). The red tag is always fixed at O ($|T_{r,0}O| = 0$) during the movement. The actual and theoretical PDT variations of different AO distances presented in Figure 7(c) are closely overlapping. It indicates that the estimation of the deviation angle via PDT is independent of the distance between the tag and antenna. 2) For the second experiment, before spinning line l_0 around O in a similar way as above, we change the location of rotation center O by varying the distance $|T_{r,0}O|$ from 0cm to 10cm. As shown in Figure 7(d), the actual and theoretical PDT variations are still similar to each other, so the effectiveness of using PDT for the limb orientation estimation is verified. In addition, since all the tags on the same part of limb (i.e., the upper arm or lower arm) have the same orientation, the phase offset due to the tag orientation is canceled through the PDT calculation and will not affect the final estimation of the deviation angle.

4.3.2 AoA-based Orientation Refinement. Since the resolution of the 3D Limb Orientation Estimation, reflected as the range of possible PDT, is determined by the relative distance between two tags, it is difficult to achieve higher resolution by increasing the distance between tags, due to the limited length of the human arm. Besides, the PDT is capable of estimating the deviation angle related to the transmitting direction of the antenna, but can hardly differentiate orientations with the same deviation angle. Therefore, we utilize an AoA-based method [35] to refine the orientation estimation. As shown in Figure 8, we use a simple case to show the basic idea of using

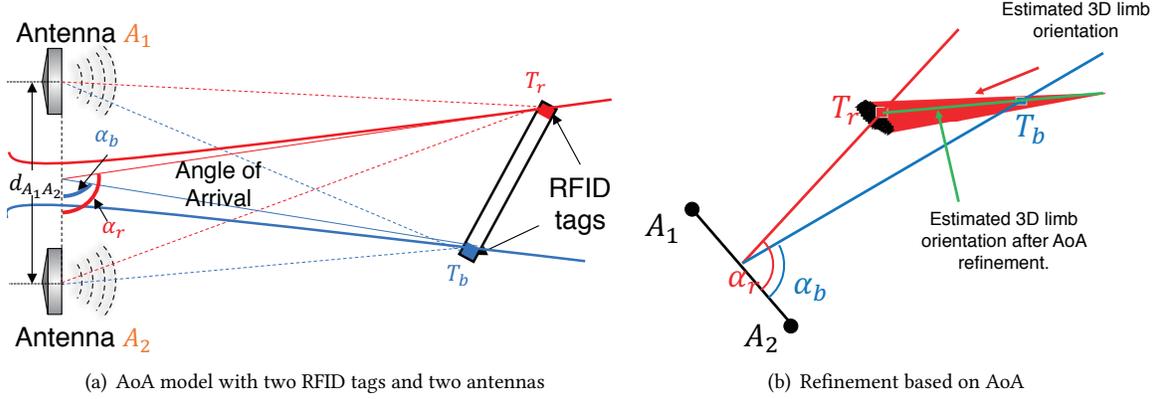


Fig. 8. AoA-based Orientation Refinement.

the *Phase Difference between the two Antennas (PDA)* of the same tag to determine the Angle of Arrival (AoA) of each tag and then refine the limb orientation. The limb is also modeled as a line segment with two tags T_b and T_r indicating two end points, and the corresponding AoAs of the received signals from the two tags are α_b and α_r , respectively. Taking blue tag T_b as an example, when two antennas A_1 and A_2 on RFID reader interrogate the tag, we could obtain two phase values $\theta(b, A_1)$ and $\theta(b, A_2)$ from tag T_b . So the PDA is represented as:

$$\Delta\theta(b) = \theta(b, A_1) - \theta(b, A_2). \quad (8)$$

Since the phase value of the RFID tag is linearly proportional to the distance between the antenna and tag, the distance difference, Δd_b , from tag T_b to two antennas on RFID reader could be obtained from the PDA as:

$$\Delta d_b = |A_1 T_b| - |A_2 T_b| = \frac{\lambda}{2} \left(\frac{\Delta\theta(b)}{2\pi} + k \right), \text{ where } k = 0, \pm 1, \dots \quad (9)$$

Then a half hyperbolas is generated based on Δd_b , which indicates the AoA of tag T_b . According to the analysis in [35], the distance difference Δd_b with respect to two antennas can be approximated as $|A_1 A_2| \cos \alpha_b$, so we can further derive angle α_b as:

$$\cos \alpha_b = \frac{\lambda}{2|A_1 A_2|} \left(\frac{\Delta\theta(b)}{2\pi} + k \right), \text{ where } k = 0, \pm 1, \dots, \quad (10)$$

where α_b indicates the angle of the asymptotic line of the hyperbolas. Similarly, AoA α_r for tag T_r can be obtained in the same way as described above, and obviously α_r is larger than α_b . Given the AoA estimation of the attached tags, we can refine the estimated 3D limb orientation by searching for the intersection between the AoA directions (i.e., α_r and α_b) and the estimated 3D limb orientation (i.e., β). Figure 8(b) illustrates the refined 3D limb orientation estimation based on the example given in Figure 7(b). The green line shown in Figure 8(b) indicates the refined limb orientation, which is determined by the intersections.

4.3.3 Relative Distance-based Orientation Filter. Both the 3D Limb Orientation Estimation and AoA-based Orientation Refinement are designed to measure the orientation of each skeleton (e.g., the upper arm or lower arm), separately. Hence, the relationship and constraints between different skeletons are not sufficiently considered. To improve the measurement accuracy, we further use the relative distance between different skeletons as constraints to filter out the impossible orientations. As shown in Figure 9(a), we present a simple case of bending the arm to show the basic idea of calibrating the orientation from the relative distance. Two RFID tags are attached on the arm, i.e., a yellow tag on the upper arm and a red tag on the lower arm. As the arm bends over in the figure, the joint angle of elbow decreases, accordingly. As a result, the relative distance between the two tags also reduces.

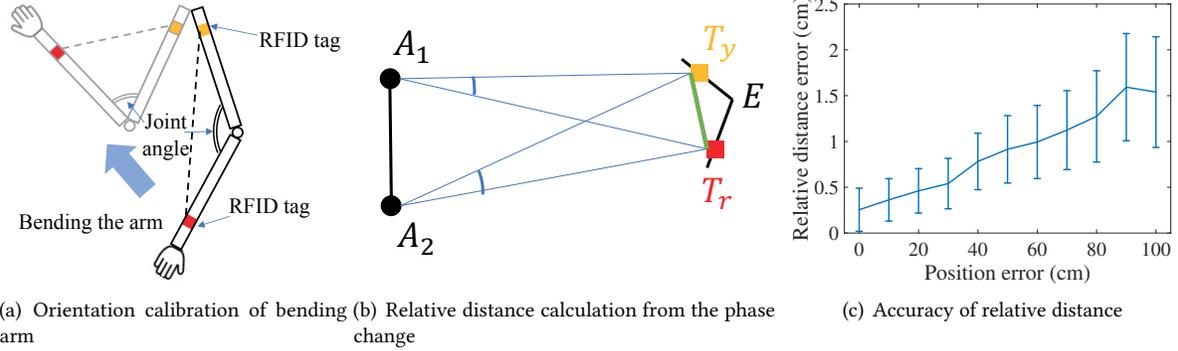


Fig. 9. Relative Distance-based Orientation Calibration.

Therefore, the relative distance can be regarded as a constraint, which connects the two skeletons and determines the joint angle of elbow in return. Suppose we have measured the relative distance between the two tags, we can recalculate the joint angle by leveraging the law of cosines [14] and then filter out the impossible orientations of the limbs based on the joint angle.

Next, we demonstrate to calculate the relative distance from the phase values. We first simplify it to a geometric problem as shown in Figure 9(b). A_1 and A_2 are the positions of antennas, and T_y and T_r are the two tags. E represents the location of the elbow. Basically, according to the law of cosines, the relative distance $|T_y T_r|$ can be calculated as:

$$\begin{aligned} |T_y T_r| &= \sqrt{|A_1 T_y|^2 + |A_1 T_r|^2 - 2|A_1 T_y| |A_1 T_r| \cos \angle T_y A_1 T_r} \\ |T_y T_r| &= \sqrt{|A_2 T_y|^2 + |A_2 T_r|^2 - 2|A_2 T_y| |A_2 T_r| \cos \angle T_y A_2 T_r}. \end{aligned} \quad (11)$$

In this equation, only the *tag-antenna distance*, i.e., $|A_1 T_y|$, and the *included angle of tags*, i.e., $\angle T_y A_1 T_r$, are still unknown. In regard to the tag-antenna distance, the variance of tag-antenna distance is represented as the phase change of tag in RFID system. Therefore, we can update the tag-antenna distance from the phase change and the tag-antenna distance in the previous timestamp. For example, if we use the superscript $[t]$ to represent the time, $|A_1 T_y|^{[t]}$ of time t can be calculated as:

$$|A_1 T_y|^{[t]} = |A_1 T_y|^{[t-1]} + \lambda \frac{\theta^{[t]}(y, A_1) - \theta^{[t-1]}(y, A_1)}{4\pi}, \quad (12)$$

where $\theta^{[t]}(y, A_1)$ represents the phase value received by antenna A_1 from tag T_y at time t . It is worth noting that the initial positions of each tag can be estimated in the *Body Position/Orientation Estimation* module and then the following tag-antenna distance can be updated from the phase changes accordingly. In regard to the included angle $\angle T_y A_1 T_r$, since the movement of tag between adjacent timestamps is much smaller than the tag-antenna distance, the included angle hence varies slightly across different timestamps. Therefore, we can calculate the angle value of last timestamp from the estimated posture of the arm in the SCS, and use it to approximate current included angle. Finally, combining the updated tag-antenna distance and include angle, we can calculate the relative distance from Eq. (11) and use the average value of different equations as our result.

Since RF-Kinect focuses on estimating the relative rotation of each skeleton instead of the accurate locations of each tag, the tag-antenna distance may have some error due to the inaccurate locations of each tag. Therefore, we validate the accuracy of calculating relative distance from the inaccurate locations of each tag. Particularly, we vary the position error of the joint from 10cm to 100cm, and calculate the relative distance based on the

incorrect position when the arm is bending over with random orientations. For each position error, we repeat the experiments 10,000 times, and plot the statistic result in Figure 9(c). We find that the error of relative distance increases slightly with the position error. However, even though the joint is located 100cm away from its actual position, the relative distance error is only about 1.5cm. The reason is that when the user is about 2 ~ 3m away from the antennas, the inaccurate locations of each joint only leads to very small error ($< 2^\circ$) of the included angle of tags. Hence, we can still efficiently measure the relative distance from the inaccurate locations of each joint.

5 3D BODY MOVEMENT TRACKING

In this section, we introduce the workflow for the 3D body movement tracking by leveraging both the proposed 3D Limb Orientation Estimation method, AoA-based Orientation Refinement method and Relative Distance-based Orientation Filter method. In particular, we estimate the orientations of all the limbs in each scanning round, which is represented as the joint rotation angles in the body geometric model. We take the right arm as an example for illustration unless mentioned otherwise, and it can be easily extended to other limbs for the complete body posture reconstruction. Once the orientation of each limb is determined, they will be concatenated to complete the body movement tracking.

5.1 Estimating Orientation of Limbs

Due to the physical structural constraints on human body, most of the joints have limited flexibility to behave. For example, there is only one Degree of Freedom (DoF) on the elbow (i.e., flexion/extension) and the joint rotation angle ϕ_4 must be positive in Figure 6. To quantify the physical structural constraints on joints, we integrate the Range of Motion (RoM), which is presented by ArmTrack [30] based on some medical studies, into the body geometric model.

We first estimate the orientation of each limb, and decompose the process into two steps: 1) deriving the orientation of the upper part of one particular limb (e.g., the right upper arm); and 2) deriving the orientation of the lower part (e.g., the right lower arm) with the knowledge of the upper part's orientation. Without loss of generality, we focus on the upper arm to look into the algorithm flow. The only parameter in estimating the posture of upper arm is the position of shoulder p_s , which is supposed to be fixed when the user moves his/her limbs. Assume we have m tags ($m \geq 2$) attached on the upper arm with known distances from the shoulder, i.e., the distances $d_{i,s}$ between the shoulder and the tag i on the upper arm can be measured in advance. Therefore, given the joint rotation angles $\Phi = \langle \phi_1, \phi_2 \rangle$ regulated by RoM, the position $p_{i,\Phi}$ of tag i can be derived from Equation (2) as:

$$p_{i,\Phi} = p_s + f(\phi_1, \phi_2, d_{i,s}), i = 1, \dots, m. \quad (13)$$

Assuming p_{A_1} and p_{A_2} are the positions of antennas on an RFID reader, the theoretical phase captured by antenna A_j from tag i is presented as:

$$\theta_\Phi(i, j) = (2\pi \frac{2\|p_{i,\Phi} - p_{A_j}\|}{\lambda} + \omega(\Phi)) \mod 2\pi, j = 1, 2, \quad (14)$$

where $\|\cdot\|$ denotes the Euclidean distance between the two given positions and $\omega(\cdot)$ estimates the phase offset caused by the tag orientation. Note we do not consider phase deviations θ_{dev} of each tag, since we have removed them in *Phase Deviation Elimination* module in Section 4.1. Due to the impact of tag orientation, the phase measurements can not be directly applied to determine the tag's location. Therefore, we calculate the phase difference between tags (PDT) based on the model in Section 4.3.1 for the estimation of limb orientation as:

$$\Delta\theta_\Phi(i, j) = \theta_\Phi(i, j) - \theta_\Phi(1, j). \quad (15)$$

Given the measured phases $(\hat{\theta}^{[t]}(1, 1), \hat{\theta}^{[t]}(1, 2), \dots, \hat{\theta}^{[t]}(m, 1), \hat{\theta}^{[t]}(m, 2))$ obtained in the round t , we also calculate the PDT as:

$$\Delta\hat{\theta}^{[t]}(i, j) = \hat{\theta}^{[t]}(i, j) - \hat{\theta}^{[t]}(1, j). \quad (16)$$

Therefore, our mission is to search for the joint rotation angle Φ to minimize the difference between the measured PDT $\Delta\hat{\theta}^{[t]}(i, j)$ and theoretical PDT $\Delta\theta_\Phi(i, j)$. We follow Tagoram [41] to define the likelihood function to evaluate the closeness the two PDTs as:

$$Prob_\Phi = \sum_{i=1}^m \sum_{j=1}^2 |e^{J(\Delta\theta_\Phi(i, j) - \Delta\hat{\theta}^{[t]}(i, j))}| / 2m. \quad (17)$$

The key idea is that if the joint rotation angles Φ are close to the groundtruth, the theoretical PDT $\Delta\theta_\Phi(i, j)$ will be close to the measured PDT $\Delta\hat{\theta}^{[t]}(i, j)$, meaning $\Delta\theta_\Phi(i, j) - \Delta\hat{\theta}^{[t]}(i, j) \approx 0$. Then, the vector $e^{J(\Delta\theta_\Phi(i, j) - \Delta\hat{\theta}^{[t]}(i, j))}$ will reach its maximum value, and we can obtain higher probability value on $Prob_\Phi$. Otherwise, $\Delta\theta_\Phi(i, j) - \Delta\hat{\theta}^{[t]}(i, j)$ is a random value within $[0, 2\pi]$, resulted from the differences between the joint angles Φ and groundtruth. The vector $e^{J(\Delta\theta_\Phi(i, j) - \Delta\hat{\theta}^{[t]}(i, j))}$ is thus smaller than the maximum value and the calculated probability $Prob_\Phi$ has a relatively lower value. Finally, we can correctly estimate the joint rotation angles Φ with the largest likelihood at round t , which indicates the orientation of the upper arm.

5.2 Refining Orientation of Limbs Based on AoA Model

Theoretically the joint rotation angle can be accurately estimated by following the above procedures, however, the existence of phase measurement errors may produce large errors on determining the limb orientation. The intuitive way to reduce such error is to increase the resolution of PDT, which is closely related to the Euclidean distance between tags. However, due to limited length of human arm, it is difficult to increase the distance between tags for more accurate joint rotation angle estimation. So we adopt the AoA model used in RF-IDraw [35] to calibrate the joint rotation angle estimation. The idea is to calculate the AoA of each tag and choose the the appropriate joint rotation angles Φ that ensure the theoretical AoAs of related tags are the most similar as the measured AoAs.

In particular, we first give the theoretical and estimated phase difference between antennas (PDA) of tag i as follows:

$$\begin{aligned} \Delta\theta_\Phi(i) &= \theta_\Phi(i, 2) - \theta_\Phi(i, 1); \\ \Delta\hat{\theta}^{[t]}(i) &= \hat{\theta}^{[t]}(i, 2) - \hat{\theta}^{[t]}(i, 1), \end{aligned} \quad (18)$$

where $\Delta\theta_\Phi(i)$ represents the theoretical PDA related to the limb orientation Φ and $\Delta\hat{\theta}^{[t]}(i)$ is the measured PDA in round t . Then the likelihood between $\Delta\theta_\Phi(i)$ and $\Delta\hat{\theta}^{[t]}(i)$ is given as follows:

$$Prob'_\Phi = \left| \sum_{i=1}^m e^{J(\Delta\theta_\Phi(i) - \Delta\hat{\theta}^{[t]}(i))} \right| / m. \quad (19)$$

Therefore, the calibrated estimation of joint rotation angles can be obtained by summing $Prob$ and $Prob'$ as follows:

$$\arg \max_{\Phi} (S), \text{ where } S = (Prob_\Phi + Prob'_\Phi) / 2. \quad (20)$$

Once ϕ_1 and ϕ_2 are obtained, we calculate the position of elbow p_e and estimate the posture of lower arm following the similar way as that for upper arm. The only difference is that the position derivation of each tag i in Eq. (13) is replaced as follows for lower arm orientation estimation:

$$p_{i, \Phi} = p_e + g(\phi_1, \phi_2, \phi_3, \phi_4, d_{i, e}), i = 1, \dots, m, \quad (21)$$

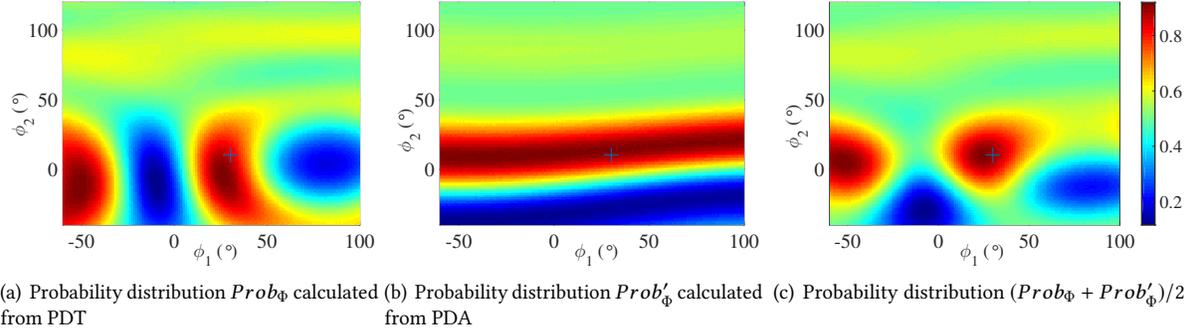


Fig. 10. An example of probability distribution for the orientation estimation of upper arm.

where $d_{i,e}$ indicates the distance between tag i and the elbow. The arm posture thus can be reconstructed by integrating both the upper and lower arm. Finally, given a series of arm postures over time, Kalman filter is adopted to smooth the estimated joint rotation angles to produce the complete arm movement.

Figure 10 presents an example of the probability distribution $Prob_{\Phi}$ and $Prob'_{\Phi}$ for the estimated orientation of the upper arm, where the groundtruth in this example is $\Phi = \langle \phi_1, \phi_2 \rangle = \langle 30^\circ, 10^\circ \rangle$. Here, $Prob_{\Phi}$ is obtained in Eq. (17) based on PDT and $Prob'_{\Phi}$ is obtained in Eq. (19) based on PDA. The larger the value of $Prob_{\Phi}$ or $Prob'_{\Phi}$, the more likely the estimated limb orientation Φ is close to the groundtruth. We find that the two features, PDT and PDA, have different focuses on the limb orientation estimation. PDT can almost uniquely determine Φ from $Prob_{\Phi}$, but it cannot provide accurate estimation on ϕ_2 . On contrary, PDA gives no knowledge about ϕ_1 , but accurately estimates ϕ_2 , which could compensate the error in PDT. Combining $Prob_{\Phi}$ and $Prob'_{\Phi}$ in Figure 10(c), we find two regions with high probabilities: $\langle 30^\circ, 10^\circ \rangle$ and $\langle -50^\circ, 10^\circ \rangle$, which represent the two possible orientations of the limb. The region around $\langle 30^\circ, 10^\circ \rangle$ is the actual orientation of the limb, where the other region around $\langle -50^\circ, 10^\circ \rangle$ is a mirroring ambiguity of the actual limb orientation with respect to the line between the two antennas as illustrated in Figure 7(b). Such mirroring ambiguity can be easily eliminated based on the relative distance, which is demonstrated in the next subsection. Finally, we can uniquely determine the orientation of the limb from the above two distributions.

5.3 Filtering Orientations Based on Relative Distance

Even though we can estimate the orientation of the arm from the largest likelihood, it is inefficient to calculate likelihoods of all the possible orientations. Therefore, before estimating the orientation of arm from PDT and PDA, we first use the relative distance between tags to filter out the impossible orientations. First, in each timestamp when we receive the phase values of each tag, we calculate the actual relative distances between tags according to the model in Section 4.3.3. Then for each possible orientation, we examine its effectiveness based on the relative distances. Particularly, we calculate the theoretical relative distances for each possible orientation and compare them with the actual relative distances. If all the theoretical relative distances are close to the actual relative distances within a difference threshold (*i.e.*, 2cm), then the corresponding orientation will be regarded as a validated possible orientation. Otherwise, it will be regarded as impossible orientation and be filtered out. Finally, we can only focus on the validated possible orientations for estimation from the PDT and PDA.

5.4 Putting Things Together

The discussion so far focuses on demonstrating each module, separately. Now we put all the pieces together and sketch the algorithm as shown in Algorithm 1. Specifically, it presents the pseudo code for single arm posture

ALGORITHM 1: Tracking the arm posture

Require: Initial body position/orientation and arm length l_e
Set joint rotation angles $\phi_1, \phi_2, \phi_3, \phi_4$ to 0° ;
Initialize the range of joint rotation $\Delta\phi = [-7.65^\circ, 7.65^\circ]$;
for Each scanning round t **do**
 Filter $\Delta\phi$ based on the relative distances and get possible angles for round t as $\Delta\phi^{[t]}$
 Calculate the PDT/PDA as $\Delta\theta^{[t]}(i, j)$ and $\Delta\theta^{[t]}(i)$ for tag i and antenna j ;
 for $\Phi = (\phi_1, \phi_2) + \Delta\phi^{[t]}$ is inside RoM **do**
 Calculate the theoretical tag position $p_{i,\Phi} = f(\phi_1, \phi_2, d_i)$;
 Calculate $\Delta\theta_\Phi(i, j)$ and $\Delta\theta_\Phi(i)$ and the corresponding likelihood S ;
 end for
 Select $\hat{\phi}_1, \hat{\phi}_2$ with the maximum likelihood S ;
 for $\Phi = (\phi_3, \phi_4) + \Delta\phi^{[t]}$ is inside RoM **do**
 Calculate the theoretical tag position $p_{i,\Phi} = g(\hat{\phi}_1, \hat{\phi}_2, \phi_3, \phi_4, d_i)$;
 Calculate $\Delta\theta_\Phi(i, j)$ and $\Delta\theta_\Phi(i)$ and the corresponding likelihood S ;
 end for
 Select $\hat{\phi}_3, \hat{\phi}_4$ with the maximum likelihood S ;
 Use Kalman filter to smooth $\hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3, \hat{\phi}_4$;
 Update $(\phi_1, \phi_2, \phi_3, \phi_4)$ with $\hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3, \hat{\phi}_4$;
end for

estimation, *i.e.*, estimates the joint rotation angles ϕ_1, ϕ_2, ϕ_3 and ϕ_4 for one scanning round t . In order to avoid the 2π jump of phase measurements due to periodical feature, the displacement of a tag between every two consecutive scans must be less than $\lambda/2$, which limits the moving speed of a tag less than $5.33m/s$ given the $30Hz$ scanning rate of RF-Kinect. Since the average human arm length is $63.5cm$ [1], the rotation angle limitation between two consecutive scans (*i.e.*, $33ms$) can be calculated as $\Delta\phi = [-7.65^\circ, 7.65^\circ]$ based on the maximum speed. Even though the instantaneous speed of some activities (*e.g.*, swinging a racquet or a golf club) may exceed such maximum speed, the instantaneous speed usually will reduce largely after the maximum speed. As a result, the movement between two consecutive scans (*i.e.*, $33ms$) is usually smaller than our limitation (*i.e.*, $\lambda/2$). Therefore, $\Delta\phi = [-7.65^\circ, 7.65^\circ]$ describes the searching range of the limb orientations in each iteration.

During one iteration of a scanning round, we first use the relative distance to filter the possible orientations inside $[-7.65^\circ, 7.65^\circ]$, which further reduces the searching range. Moreover, the mirroring ambiguity is solved since the mirroring postures have different relative distances. Next, we in turn search for the postures of the upper arm and lower arm based on PDT and PDA, and obtain the estimated rotation angles $\hat{\phi}_1, \hat{\phi}_2, \hat{\phi}_3, \hat{\phi}_4$ from the likelihood S . Finally, we use Kalman filter to smooth the rotation angles and update them accordingly.

6 SYSTEM EVALUATION

In this section, we present the implementation of our system and evaluate the performance with several gestures and users.

6.1 Experimental Setup & Methodology

Experimental Setup. The experimental setup of our RF-Kinect system, consisting of one dual-antenna RFID reader and multiple wearable RFID tags, is presented in Figure 11. The Impinj Speedway modeled $R420$ RFID reader integrated with two Laird $S9028PCL$ directional antennas, which is compatible with EPC $C1G2$ standard, is placed on a $80cm$ -high desk. The two antennas on the RFID reader are separated by $150cm$ from each other. The RFID reader works in channel hopping mode over the frequency band $920 \sim 926$ MHz, and scans the RFID tags with both antennas. The user who wears RFID tags stands $200cm$ in front of the RFID reader while performing

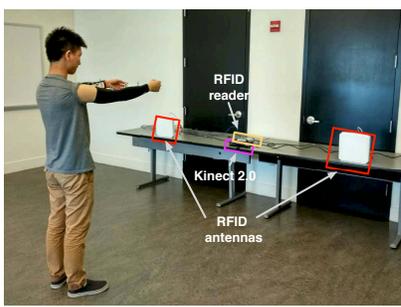


Fig. 11. Experimental setup.

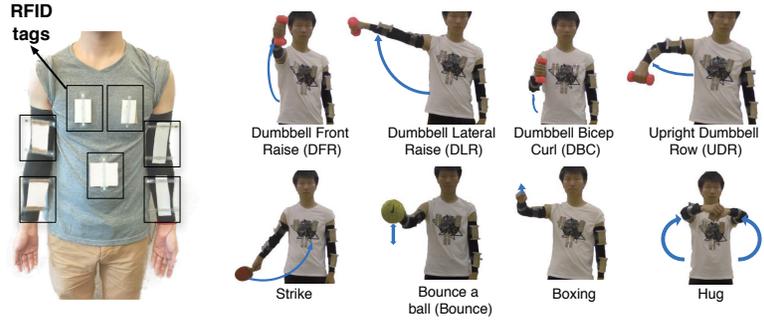


Fig. 12. Body gestures.

different postures. Specifically, each part of the limb (i.e., upper arm, lower arm, thigh and shank) has attached with a pair of RFID tags, while the distance between tags is set to 10cm as default. Another three tags, forming a triangle, are attached on the user's chest for the body position and facing direction estimation, and also serve as the reference points for the whole body postures. To maintain the relative geometric relationship among these tags, we attach the tags together with cardboard to the clothes of user.

Experimental Methodology. The experiments are conducted in a typical indoor environment involving 5 participants in total (4 males and 1 female) to validate the effectiveness of RF-Kinect on body movement tracking. Each participant is asked to wear the sport sleeves and T-shirt attached with RFID tags, and perform both pre-defined static postures and dynamic gestures. Specifically, the static postures include pointing frontward/rightward/front-rightward with one arm raising up to 30°/60°/90°, while three different types of dynamic gestures are adopted as: (i) *Exercise gestures*: Dumbbell Front Raise (DFR), Dumbbell Lateral Raise (DLR), Dumbbell Bicep Curl (DBC), Upright Dumbbell Row (UDR); (ii) *Functional gestures*: Strike, Bounce a Ball (Bounce), Boxing, Hug; and (iii) *Leg gestures*: Knee Kick (KK) and Front Kick (FK). Figure 12 presents the body movement of exercise gestures and functional gestures. All the above postures or dynamic gestures will be repeated 10 times by each participant. In addition, we also evaluate the robustness of our RF-Kinect system on tracking different practical gestures such as drinking and hand shaking, etc., while the users move around in a complicated indoor environment.

In order to record the groundtruth of body movement, we place a Microsoft Kinect 2.0 [8] in the middle of two RFID antennas for videotaping. The movements of irrelevant people in the video will be filtered out to avoid the visual interference on the ground truth. A series of estimated postures, involving both limb (solid line) orientations and joint (circle marker) positions, consist of a complete body movement, and then compare with the groundtruth for performance evaluation. Figure 13 shows an illustrative example of DLR gesture recorded by RF-Kinect and Kinect 2.0, respectively.

6.2 Metrics

To evaluate the performance of RF-Kinect on body movement tracking, we define the following two metrics.

Angle Error. The angle error is defined as the difference between estimated orientation vector of each limb (e.g., the estimated orientation vector of upper arm is obtained as $\mathbf{V}_{\text{upper}} = p_e - p_s$) and its actual orientation vector obtained by Microsoft Kinect 2.0.

Position Error. The position error is defined as the relative Euclidean distance between the estimated position of each joint (e.g., the elbow) with its actual position. For example, given the estimated position of elbow p_e and the groundtruth \hat{p}_e obtained by Kinect, the position error is $\|p_e - \hat{p}_e\|$.

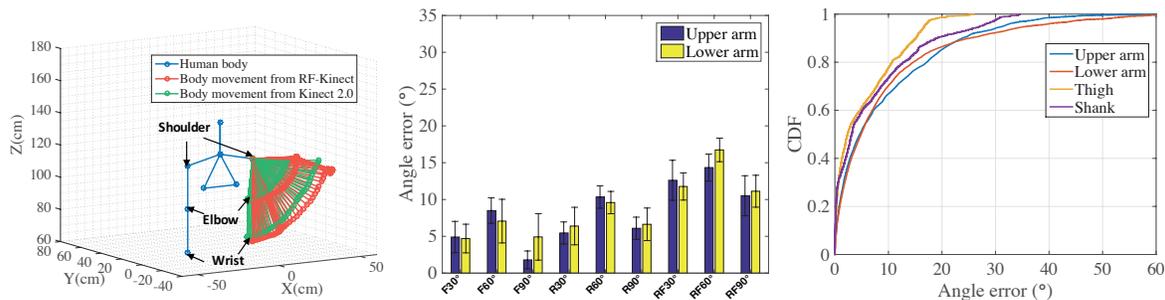


Fig. 13. Illustration of body movement DLR.

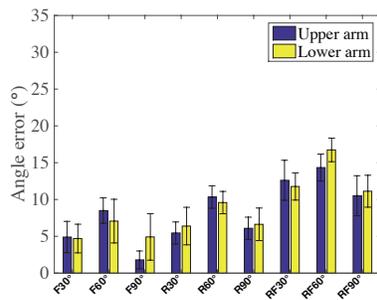


Fig. 14. Angle error of static pointing postures.

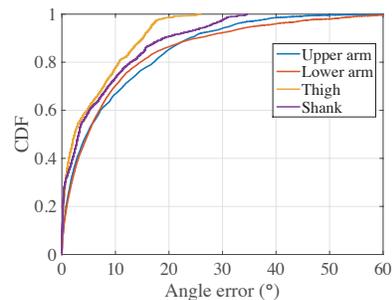


Fig. 15. Overall angle error of dynamic gestures.

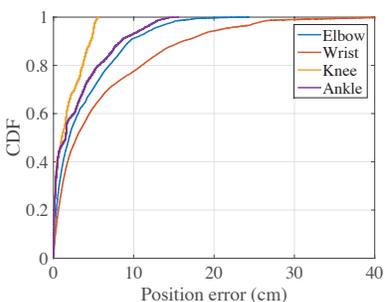


Fig. 16. Overall position error of dynamic gestures.

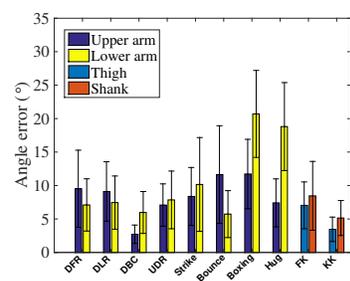


Fig. 17. Angle error of different dynamic gestures.

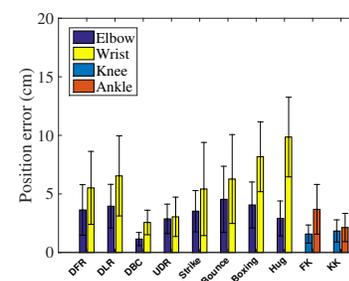


Fig. 18. Position error of different dynamic gestures.

6.3 Tracking of Static Postures

Our system can accurately estimate the pointing direction of the arm. We first evaluate the orientation estimation for both upper arm and lower arm while performing static postures (i.e., pointing in a direction with their hands). Figure 14 presents the mean and standard deviation of angle errors for nine postures involving both upper and lower arm. We use F/R/RF to represent the hand pointing frontward/rightward/front-rightward, followed with the degrees of arm raising angle. The mean angle errors across all the nine postures are 8.2° and 8.7° for the orientation estimation of upper and lower arm, respectively. Specifically, the posture of pointing frontward has the lowest angle error 5.4° while pointing front-rightward performs the worst with the angle error of 12.7° . The reason behind is that the arm moving frontward does not introduce any orientation changes for the tags on the arm, whereas moving front-rightward results in the phase deviation due to inconsistent orientation changes of different tags. We also find that the angle error reaches the maximum when the arm is raised to 60° , because PDT value is close to the peak value (Figure 4) at 60° , leading to ambiguity orientation estimations. Overall, the above results demonstrate that our proposed RF-Kinect can provide satisfactory accuracy to track static postures.

6.4 Tracking of Dynamic Gestures

Our experimental results show that our system can achieve good performance in tracking different dynamic gestures. We then evaluate the performance of RF-Kinect when tracking dynamic gestures. Figure 15 and Figure 16 present the cumulative distribution function (CDF) of the angle and position errors with respect to 4 different limb movements (i.e., upper arm, lower arm, thigh and shank) across different gestures for all the participants. Specifically, we can achieve comparable orientation estimation accuracy with average angle errors, 9° , 9.3° , 5.3° ,

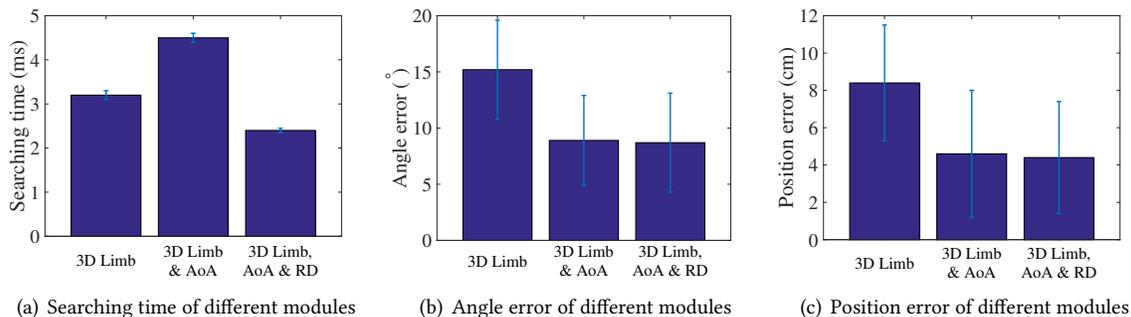


Fig. 19. Performance evaluation results with different modules.

and 6.8° , when tracking upper arm, lower arm, thigh and shank movement, and the average position errors for locating each body joint (i.e., elbow, wrist, knee and ankle) are 3.7, 6.1, 1.7 and 3.1cm, respectively. Even the 80 percentile angle errors for all the limb movements are always limited within 20° , while the corresponding 80 percentile position errors are less than 12cm. These promising results imply that RF-Kinect could accurately track 3D human body movements with complex gestures (e.g., bounce a ball, boxing), which can be found in many motion sensing applications such as virtual reality gaming. We also observe that both the angle error and position error of the leg-involved gestures are smaller than those of the arm-involved gestures, since the motion range of the leg is usually much smaller than that of the arm. The high accuracy of RF-Kinect provides comparable results with sensor-based approaches [30], which locates the elbow and wrist with 4.5cm and 5.7cm error, respectively. Compared to the volunteers' average arm length of 49.8cm, we believe the performance of RF-Kinect makes it amenable to most gesture recognition applications, including both the VR gaming and fitness monitoring.

We further compare the tracking accuracy across different dynamic gestures. Specifically, for each dynamic gesture, we present the angle error for both upper and lower arm orientation estimation, and the position error of both elbow and wrist tracking. Similarly, the angle error for both thigh and shank and the position error of knee and ankle are also provided. Figure 17 and Figure 18 present the mean and standard deviation of the angle and position errors for each individual dynamic gesture. It is encouraging to find that 8 out of 10 gestures achieve the angle error less than 15° , and the corresponding position errors of those 8 gestures are as low as 7cm. The average angle and position errors of the other two gestures (i.e., boxing and hug) are lower than 25° and 10cm, respectively due to the possible signal attenuation and multipath effects caused by heavy involvement of torso and interference from bent/crossed arm. In addition, we have a consistent observation that the accuracy for tracking leg-involved gestures is higher than that of arm-involved gestures.

6.5 Impact of Different Modules

Next, we evaluate the impact of the three core modules in RF-Kinect, i.e., 3D Limb Orientation Estimation (3D Limb), AoA-based Orientation Refinement (AoA) and Relative Distance-based Orientation Filter (RD). Particularly, we evaluate the impact by comparing the searching time during each loop and the tracking accuracy as shown in Figure 19. From the three figures, we find that all the three methods can be efficiently finished within 5ms for each searching loop, while the average errors of angle and position are all below 16° and 10cm, respectively. This means RF-Kinect can correctly estimate the trace of the human body in a real-time application. Moreover, when we compare among the three methods, we find that AoA-based Orientation Refinement slightly increases the searching time for refinement, but reduces the tracking error. On the other hand, even though Relative Distance-based Orientation Filter does not improve the tracking accuracy compared with 3D limb & AoA, it instead reduces the searching time to 2.5ms by filtering the impossible orientations. Therefore, by combining

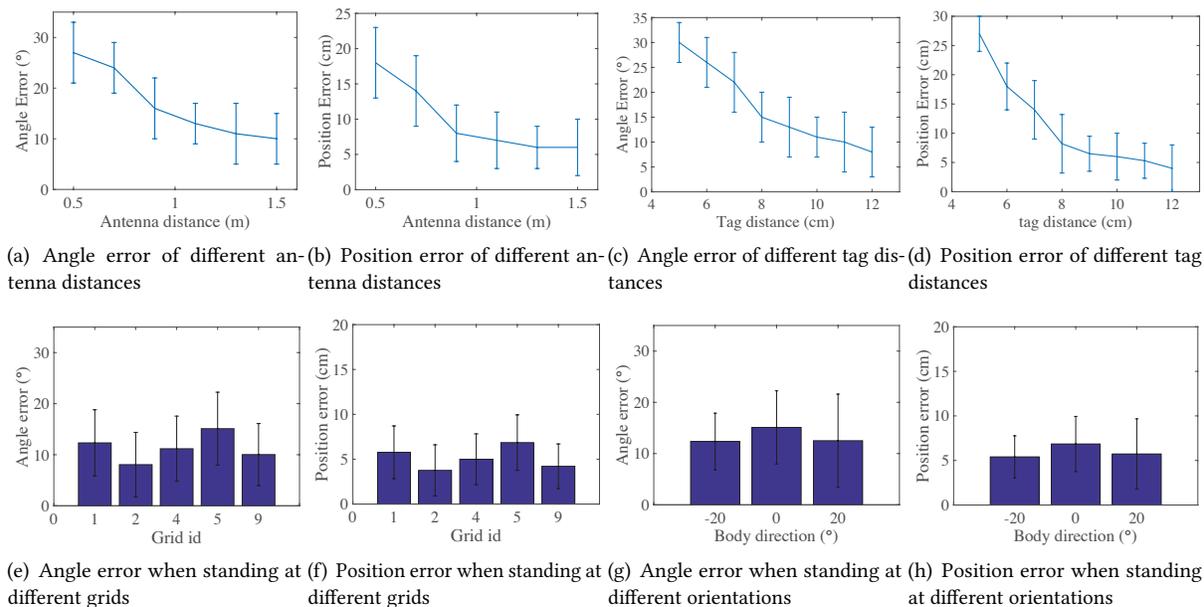


Fig. 20. Performance evaluation results with different system configurations and user locations/orientations.

the three methods, RF-Kinect can efficiently and accurately track the human movements, which can be used in different kinds of gesture recognition applications.

6.6 Impact of Different Distances between Antennas

We next evaluate the impact of the distance between antennas on the accuracy of body movement tracking. Specifically, we vary the distance between two antennas from $0.5m$ to $1.5m$ with the incremental step $0.2m$. Given a specific distance between two antennas, the user is asked to perform 8 different dynamic gestures, each repeating 10 times, at around $200cm$ in front of the antennas. As shown in Figure 20(a) and Figure 20(b), both the angle error and position error decreases as the distance between two antennas increases. The rationale behind the above results is that larger distance can achieve higher resolution on AoA estimation, and thereby boost the tracking accuracy. In particular, when the distance between two antennas is larger than $1m$, the angle error and position error are less than 13.3° and $7.4cm$, respectively, which are accurate enough to perform body movement tracking. It is worth noting that the increasing distance between antennas will also create more spatial ambiguity due to the AoA model [35]. However, since RF-Kinect tracks the relative rotation of each limb from the phase changes instead of the absolute positions, the spatial ambiguity can be eliminated from the continuous movement of each tag. Moreover, 3D Orientation Estimation and Relative Distance calculation is insensitive to the actual position as shown in Figure 7(c) and Figure 9(c), so we can still accurately track the gestures even AoA model leads to some spatial error at the first timestamp.

6.7 Impact of Different Distances between Tags

We next evaluate the impact of the distance between tags on the accuracy of body movement tracking. Specifically, we vary the distance between two tags on the limb from $5cm$ to $12cm$ with the incremental step $1cm$. Given a specific distance between tags, the user is asked to perform 8 different dynamic gestures, each repeating 10 times, at around $200cm$ in front of the antennas. As shown in Figure 20(c) and Figure 20(d), we find both the angle

error and position error decreases as the distance between tags increases. It is reasonable since the increment of tag distance will increase the resolution of the model of 3D orientation estimation as shown in Figure 7(a). Particularly, when the distance is larger than $8cm$, the angle error and position error reduce to 14.7° and $8.2cm$, respectively.

6.8 Impact of Different User Locations/Orientations

Our experimental results show that the tracking accuracy is stable as the location changes. We next demonstrate the scalability of RF-Kinect by evaluating the body movement tracking performance when the user locate at different distance from the RFID reader. We partition the main scanning region of RF-Kinect, a $150cm \times 150cm$ square area, into 3 grids labeled from 1 to 9. The grids with the labels $1 \sim 3/4 \sim 6/7 \sim 9$ correspond to three different levels of distances from RFID reader, i.e., close/medium/far, respectively. Without loss of generality, we only focus on the grids 1/2/4/5/9, since other grids have similar distances. In the experiments, the users always face the same direction when standing at different grids. As shown in Figure 20(e) and Figure 20(f), the average limb orientation estimation errors at different locations are always less than 16° , and the average position errors are limited within $8cm$. Specifically, we can achieve the best accuracy at the grid 2 due to its stable RFID readings at the closest distance from RFID reader.

Our experimental results show that the tracking accuracy is stable as the user orientations changes. Beside the impact of user locations, we also study the tracking accuracy with respect to user body orientation as shown in Figure 20(g) and Figure 20(h). To evaluate the performance experimentally, we ask the user to stand a fixed location and perform the gestures on three different body orientations $-30^\circ/0^\circ/30^\circ$, where 0° represents the user is right in front of the antennas. We find that the tracking accuracy changes little as the user changes his/her body orientation. In particular, the average angle error and position errors still maintain within 16° and $8cm$, respectively. In addition, among all the participants, RF-Kinect has a lower tracking accuracy for the participant 5 due to the stronger mutual RFID interference when facing the midpoint of two separated antennas.

6.9 Tracking Performance across Multiple Users

Our system can achieve comparable tracking accuracy across different users, even when they are performing simultaneously. We next evaluate the impact of different users on the accuracy of body movement tracking in RF-Kinect. The tracking accuracy for five different participants are presented Figure 21(a) and Figure 21(b). Both the orientation and position estimation maintain consistently low errors across different participants. It denotes that different users do not have a significant impact on the tracking performance of RF-Kinect. We also observe a relatively lower tracking accuracy for the user 5 with the estimation errors on limb orientation as 11.3° and joint position as $8.5cm$. This is because the intensive Doppler effects on the RF signals resulted from a faster movement of this user.

Furthermore, we also study the performance on tracking the body movement of multiple users simultaneously. Due to the unique ID on each RFID tag, it is easily for RF-Kinect to differentiate the gestures of each individual person. Particularly, since RFID system uses the Frame-Slotted ALOHA protocol for identification, we first study the reading rate as we increase the number of users. As shown in Figure 21(c), when the number of users increases, the reading rate of each user will decrease accordingly. It is reasonable because RFID system need to identify more tags, which consumes more time for a reading cycle. However, the reading rate is not linearly decreasing with the number of users. When the number of users increases to 4, the reading rate is still more than 10Hz, which is about 1/3 of the reading rate for one user. Even though 10Hz reading rate is not sufficient track the faster gestures, we can still track some fitness gestures with modest speed, which may usually involves multiple users. When the number of users is 1 or 2, we can track the faster gestures in applications such as VR gaming.

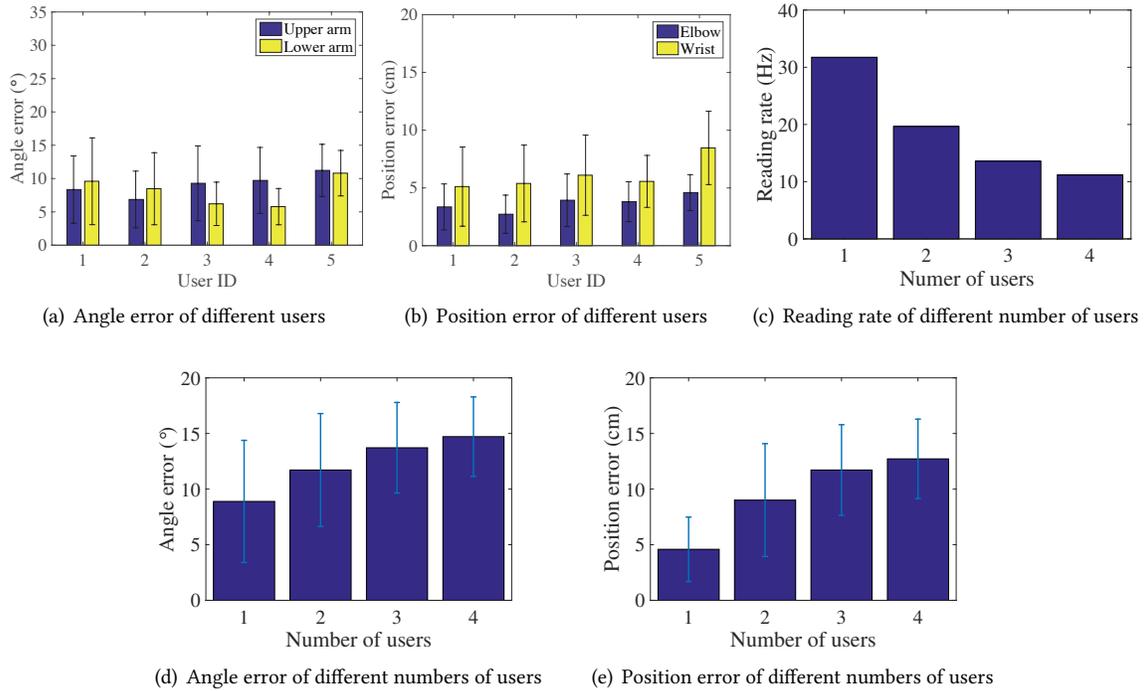


Fig. 21. Performance evaluation across different users and multiple users.

Next, we evaluate the tracking accuracy of different number of users. As shown in Figure 21(d) and Figure 21(e), the tracking error may slightly increase with the increment of number of users. This is because the increment of user number will reduce the reading rate and finally affect the tracking accuracy. Specifically, when 4 users are performing gestures simultaneously, RF-Kinect can still achieve less than 15° average angle error and 13cm position error, respectively. The above results confirm that RF-Kinect is capable to accurately track the body movement of multiple users simultaneously.

6.10 Robustness Study

Our system can accurately track both the practical gestures and whole body movements of multiple users in complicated environments. In order to evaluate the performance of RF-Kinect in more practical scenarios, RF-Kinect is then set up in a typical office environment with complex layout and a lot of furniture/appliances. The RFID signals also experience heavy interferences from other wireless devices in such an environment. In the experiments, we ask multiple participants move around in the office and behave as usual. Specifically, the participants perform either practical gestures (i.e., turn on the light, open the fridge, shake the hand and drink water) or whole body movement (i.e., the rotation of human body and walking) in the absence of any instructions.

Figure 22 presents the average errors of limb orientation and position estimation for both the practical gestures and whole body movements. We can find that RF-Kinect can still limit the limb orientation and position estimation error within 20° and 11cm , respectively. Considering the random movement pattern of practical gestures, RF-Kinect can still achieve satisfactory tracking accuracy for the body movement, and the performance is not significantly degraded when comparing with the pre-defined gestures in Section 6.4. We further evaluate the performance on tracking the body orientation and position for whole body movement, and the results are

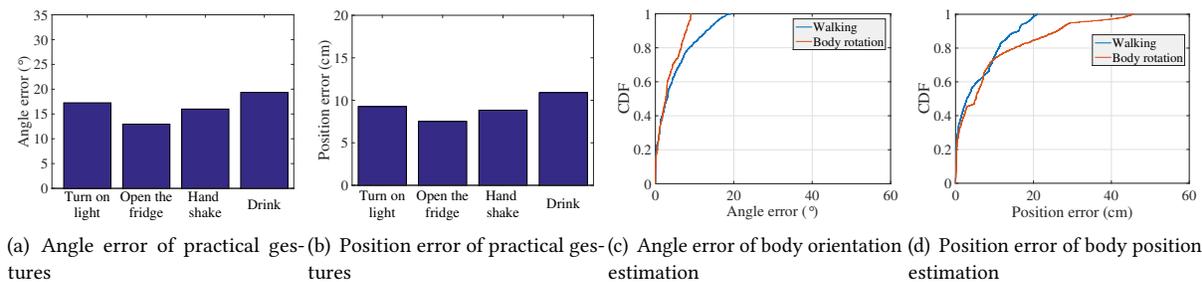


Fig. 22. Performance evaluation results of practical gestures and whole body gestures.

presented in Figure 22(c) and Figure 22(d). As depicted in the experimental setup, we have the three tags attached on user's chest for body orientation and position tracking. Specifically, the existing method [21] is deployed to locate the human body. It is obvious that RF-Kinect achieves higher tracking accuracy for whole body movements than practical gestures. This is because whole body movements have less involvement of limb movement that may produce strong interference on the RF signal propagation.

6.11 Impact of Blocks in NLOS Environment

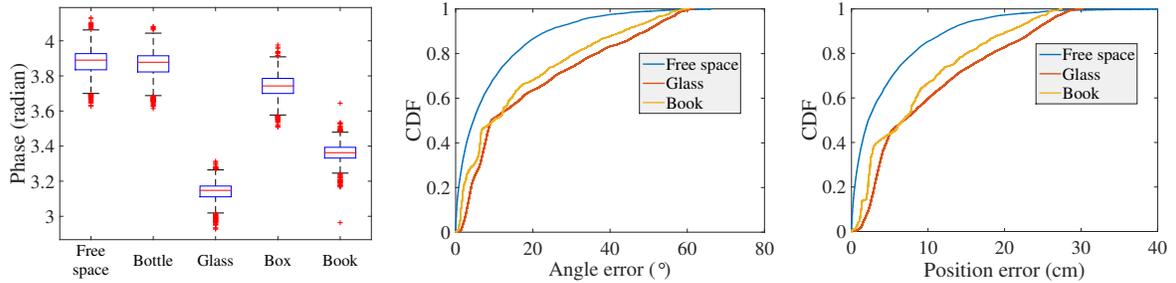
RF-Kinect can still accurately track the gestures in NLOS environment. In real application scenarios, there may exist different kinds of objects that block the line-of-sight path between the user and RFID antennas. Therefore, it is essential to study the performance of RF-Kinect in these NLOS environment. Particularly, we first study the change of phase measurement when different kinds of objects block the LOS path (*i.e.*, Bottles, Glass, Box and Books). As shown in Figure 23(a), we find while the glass and books reduce the phase about 1 radian, both the bottle and box almost do not change the phase value. It means RF-Kinect can work similar to the free space even if the bottle and box block the LOS path, but vision-based approaches cannot work in such situations.

Moreover, we evaluate the tracking performance of dynamic gestures when the glass or books block the LOS path. As shown in Figure 23(b) and Figure 23(c), we find the accuracy of both situations is slightly smaller than the free space, due to the multi-path effect in NLOS environment. However, such accuracy (*i.e.*, about 20° angle error and 10cm position error) is still sufficient for gesture recognition compared with some sensor-based approaches [30]. The reason of such accurate results in NLOS environment may be that the tags attached on one person may suffer from similar multi-path effect. Therefore, when we calculate the phase differences between tags to estimate the orientations, the multi-path effect is canceled by each other. Finally, we can still accurately track the gestures in the NLOS environment.

7 DISCUSSION

Number of Wearable Tags on Each Individual Limb. The limited limb length only allows two wearable RFID tags to be attached, but we can still achieve good orientation estimation accuracy. To evaluate the performance with more wearable tags, we also emulate the limb rotation with one stick that is longer than any limb. We find that as the tag number increases from 2 to 4, the orientation estimation accuracy only has 3° improvement. It implies that two tags on each limb have already ensured sufficiently high accuracy on tracking body movement. Nevertheless, as the advancement of RFID techniques, it is possible to have a dense deployment of wearable RFID tags with reduced size in the future.

Tracking Accuracy and Missing Tags. Currently, some tracking error is still as large as 30°, which is not accurate enough compared with real Kinect. By in deep analyzing the data, we find the tracking trace is exactly the same as shown in Figure 13, but the speed of each gesture can not be correctly estimated, leading to large



(a) Phase error of different NLOS environments (b) Angle error of different NLOS environments (c) Position error of different NLOS environments

Fig. 23. Performance evaluation results of different NLOS environments.

error at some timestamps. Therefore, the correct trace of each gesture can still satisfy the requirements in applications such as VR gaming and fitness monitoring. The rationale behind the unsynchronized error may be the non-uniform sampling for each gesture due to the missing tags in RFID system. One possible way is to increase the number of wearable tags and use the redundant tags to compensate the missing tags. Based on the redundant tags, even though several tags are missing in some cycles, we can still estimate the gestures from other tags.

Number of Antennas. Even though we can indeed increase the number of antennas to improve the tracking accuracy of RF-Kinect, it is not convenient to deploy more than two antennas in a real application. When we setup RF-Kinect with two antennas, we only need to separate the two antennas with a certain distance similar to the deployment of loudspeaker box, which is easy for the common user. However, when we deploy multiple antennas (*i.e.*, three antennas), we usually do not deploy them linearly. On the contrary, we need to carefully deploy them in 3D space, which provides the distinctive information of different dimensions. Therefore, the positions and relative distances of the antennas need to be configured with specialized knowledge, which is more professional and not suitable for a ubiquitous application. Moreover, since an RFID antenna is about \$100, the increment on the number of antennas will also increase the cost and make RF-Kinect intolerable.

Initial Postures and Body Parameters of A Specific User. Before launching RF-Kinect, we ask the user's hand to maintain in the vertical free-fall posture, and manually measure the positions of each tag. In a real application, we can embed the wearable RFID tags into the T-shirt with fixed positions to avoid the manual measurement. When the user starts RF-Kinect, we can require the user to raise the arm, which can be the gesture for starting the tracking application and also be used to calibrate the actual positions of each tag. After the calibration, the user can use RF-Kinect normally without worrying about the positions of tags.

8 CONCLUSION

In this paper, we propose RF-Kinect, a training-free and low-cost RFID-based body movement tracking system leveraging RFID tags attached on the human body. RF-Kinect provides a practical solution to capture various human body postures without the dependance of light conditions and with minimum hardware requirements by only using a dual-antenna RFID reader and a few RFID tags. A new 3D limb orientation estimation method is developed to determine limb orientation and joint position, which are then calibrated with the proposed Angle of Arrival (AoA) based orientation refinement scheme. Specifically, two phase related features, *phase difference between tags* (PDT) attached to the same part of a limb and *phase difference between antennas* (PDA) on the same tag, are composed to facilitate the posture estimation. Given a series of estimated body postures from multiple rounds of tag scanning, the body movement is accurately determined with Kalman filter-based

smoothing. Extensive experiments involving 5 volunteers are conducted in a typical indoor environment. The experimental results confirm the effectiveness on utilizing attached RFID tags for body movement tracking. In particular, RF-Kinect could achieve as low as 8.7° average angle error for limb orientation estimation and 4.4cm position error for joint position estimation in comparison with Microsoft Kinect 2.0.

ACKNOWLEDGMENTS

This work is partially supported by National Natural Science Foundation of China under Grant Nos. 61472185, 61321491, 61502224; Jiangsu Natural Science Foundation, No. BK20151390. This work is partially supported by Collaborative Innovation Center of Novel Software Technology and Industrialization. This work is partially supported by the program A for Outstanding PhD candidate of Nanjing University. This work is partially supported by the US National Science Foundation Grants CNS-1514436, CNS-1820624, CNS-1717356 and Army Office Research Grant W911NF-17-1-0467.

REFERENCES

- [1] 2017. Average human arm length. <https://www.reference.com/science/long-average-human-arm-62c7536c5e56f385..> (2017).
- [2] 2017. Cone - Wikipedia. <https://en.wikipedia.org/wiki/Cone>. (2017).
- [3] 2017. Fitbit. <http://www.fitbit.com/>. (2017).
- [4] 2017. Gesture Recognition Market-Global Industry Analysis Size Share Growth Trends and Forecast 2016-2024. <http://www.transparencymarketresearch.com/gesture-recognition-market.html..> (2017).
- [5] 2017. Leap Motion. <https://www.leapmotion.com>. (2017).
- [6] 2017. Motion capture systems | Vicon. <https://www.vicon.com>. (2017).
- [7] 2017. Rigid body - Wikipedia. <https://en.wikipedia.org/wiki/RigidBody..> (2017).
- [8] 2017. X-box Kinect. <http://www.xbox.com/>. (2017).
- [9] Fadel Adib, Chen-Yu Hsu, Hongzi Mao, Dina Katabi, and Frédo Durand. 2015. Capturing the human figure through a wall. *ACM Transactions on Graphics (TOG)* 34, 6 (2015), 219.
- [10] Fadel Adib, Zachary Kabelac, and Dina Katabi. 2015. Multi-person motion tracking via RF body reflections. In *Proc. of USENIX NSDI*.
- [11] Sena Agezo, Yuxiang Zhang, Ziyu Ye, Somesh Chopra, Shrenik Vora, and Timothy Kurzweg. 2016. Battery-free RFID heart rate monitoring system. In *Proc. of IEEE Wireless Health (WH)*. IEEE.
- [12] Jacques Denavit. 1955. A kinematic notation for lower-pair mechanisms based on matrices. *Trans. of the ASME. Journal of Applied Mechanics* (1955).
- [13] Han Ding, Longfei Shangguan, Zheng Yang, Jinsong Han, Zimu Zhou, Panlong Yang, Wei Xi, and Jizhong Zhao. 2015. Femo: A platform for free-weight exercise monitoring with rfids. In *Proc. of ACM SenSys*.
- [14] M. Hazewinkel. 1987. *Encyclopaedia of Mathematics*. Springer. <https://books.google.com/books?id=WzbzxytdAuUC>
- [15] Wenjun Hu, Hao Gu, and Qifan Pu. 2013. LightSync: unsynchronized visual communication over screen-camera links. In *Proceedings of the 19th annual international conference on Mobile computing & networking*. ACM, 15–26.
- [16] Mostafa Izz, Zhongyuan Li, Hongbo Liu, Yingying Chen, and Feng Li. [n. d.]. Uber-in-Light: Unobtrusive Visible Light Communication Leveraging Complementary Color Channel.
- [17] Bryce Kellogg, Vamsi Talla, and Shyamnath Gollakota. 2014. Bringing gesture recognition to all devices. In *Proc. of USENIX NSDI*.
- [18] Ye-Sheng Kuo, Pat Pannuto, Ko-Jen Hsiao, and Prabal Dutta. 2014. Luxapose: Indoor positioning with mobile phones and visible light. In *Proc. of ACM Mobicom*.
- [19] Oscar D Lara and Miguel A Labrador. 2013. A survey on human activity recognition using wearable sensors. *IEEE Communications Surveys & Tutorials* (2013).
- [20] Tianxing Li, Chuankai An, Zhao Tian, Andrew T Campbell, and Xia Zhou. 2015. Human sensing using visible light communication. In *Proc. of ACM MobiCom*.
- [21] Qiongzhen Lin, Lei Yang, Yuxin Sun, Tianci Liu, Xiang-Yang Li, and Yunhao Liu. 2015. Beyond one-dollar mouse: A battery-free device for 3d human-computer interaction via rfid tags. In *Proc. of IEEE INFOCOM*.
- [22] Chang Liu, Lei Xie, Chuyu Wang, J Wu, and Sanglu Lu. 2016. Track Your Foot Step: Anchor-free Indoor Localization based on Sensing Users' Foot Steps. In *Proc. of IEEE MASS*.
- [23] Abhinav Parate, Meng-Chieh Chiu, Chaniel Chadowitz, Deepak Ganesan, and Evangelos Kalogerakis. 2014. RisQ: recognizing smoking gestures with inertial sensors on a wristband. In *Proc. of ACM MobiSys*.
- [24] Joel C Perry and Jacob Rosen. 2006. Design of a 7 degree-of-freedom upper-limb powered exoskeleton. In *Proc. of IEEE BioRob*.

- [25] Qifan Pu, Sidhant Gupta, Shyamnath Gollakota, and Shwetak Patel. 2013. Whole-home gesture recognition using wireless signals. In *Proc. of ACM MobiCom*.
- [26] Longfei Shangguan and Kyle Jamieson. 2016. The Design and Implementation of a Mobile RFID Tag Sorting Robot. In *Proc. of ACM MobiSys*.
- [27] Longfei Shangguan and Kyle Jamieson. 2016. Leveraging Electromagnetic Polarization in a Two-Antenna Whiteboard in the Air. In *Proceedings of the 12th International on Conference on emerging Networking EXperiments and Technologies*. ACM, 443–456.
- [28] Longfei Shangguan, Zheng Yang, Alex X Liu, Zimu Zhou, and Yunhao Liu. 2015. Relative localization of rfid tags using spatial-temporal phase profiling. In *Proc. of USENIX NSDI*.
- [29] Longfei Shangguan, Zimu Zhou, and Kyle Jamieson. [n. d.]. Enabling Gesture-based Interactions with Objects. In *Proc. of ACM Mobisys*.
- [30] Sheng Shen, He Wang, and Romit Roy Choudhury. 2016. I am a Smartwatch and I can Track my User's Arm. In *Proc. of ACM MobiSys*.
- [31] Li Sun, Souvik Sen, Dimitrios Koutsonikolas, and Kyu-Han Kim. 2015. Withdraw: Enabling hands-free drawing in the air on commodity wifi devices. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. ACM, 77–89.
- [32] Daniel Vlasic, Rolf Adelsberger, Giovanni Vannucci, John Barnwell, Markus Gross, Wojciech Matusik, and Jovan Popović. 2007. Practical motion capture in everyday surroundings. In *ACM transactions on graphics (TOG)*, Vol. 26. Acm, 35.
- [33] He Wang, Ted Tsung-Te Lai, and Romit Roy Choudhury. 2015. MoLe: Motion Leaks Through Smartwatch Sensors. In *Proc. of ACM MobiCom*.
- [34] Jue Wang and Dina Katabi. 2013. Dude, where's my card?: RFID positioning that works with multipath and NOLS. In *Proc. of ACM SIGCOMM*.
- [35] Jue Wang, Deepak Vasisht, and Dina Katabi. 2015. RF-IDraw: virtual touch screen in the air using RF signals. In *Proc. of ACM SIGCOMM*.
- [36] Ju Wang, Jie Xiong, Hongbo Jiang, Xiaojiang Chen, and Dingyi Fang. 2016. D-watch: Embracing "bad" multipaths for device-free localization with COTS RFID devices. (2016), 253–266.
- [37] Wei Wang, Alex X Liu, Muhammad Shahzad, Kang Ling, and Sanglu Lu. 2015. Understanding and modeling of wifi signal based human activity recognition. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. ACM, 65–76.
- [38] Wei Wang, Alex X Liu, and Ke Sun. 2016. Device-free gesture tracking using acoustic signals. In *Proceedings of the 22nd Annual International Conference on Mobile Computing and Networking*. ACM, 82–94.
- [39] Teng Wei and Xinyu Zhang. 2016. Gyro in the air: tracking 3D orientation of batteryless internet-of-things. In *Proc. of ACM Mobicom*. 55–68.
- [40] Chao Xu, Parth H Pathak, and Prasant Mohapatra. 2015. Finger-writing with Smartwatch: A Case for Finger and Hand Gesture Recognition using Smartwatch. In *Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications (ACM HotMobile)*. ACM, 9–14.
- [41] Lei Yang, Yekui Chen, Xiang-Yang Li, Chaowei Xiao, Mo Li, and Yunhao Liu. 2014. Tagoram: Real-time tracking of mobile RFID tags to high precision using COTS devices. In *Proc. of ACM MobiCom*.
- [42] Lei Yang, Yao Li, Qiongzhen Lin, Huanyu Jia, Xiang-Yang Li, and Yunhao Liu. 2017. Tagbeat: Sensing Mechanical Vibration Period With COTS RFID Systems. *IEEE/ACM Transactions on Networking* 25 (2017), 3823–3835.
- [43] Lei Yang, Qiongzhen Lin, Xiangyang Li, Tianci Liu, and Yunhao Liu. 2015. See through walls with cots rfid system!. In *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. ACM, 487–499.
- [44] Zheng Yang, Chenshu Wu, Zimu Zhou, Xinglin Zhang, Xu Wang, and Yunhao Liu. 2015. Mobility increases localizability: A survey on wireless indoor localization using inertial sensors. *ACM Computing Surveys (CSUR)* 47, 3 (2015), 54.
- [45] Xiaojun Zhu, Qun Li, and Guihai Chen. 2013. APT: Accurate outdoor pedestrian tracking with smartphones. In *Proc. of IEEE INFOCOM*.

Received May 2017; revised November 2017; accepted January 2018