2017 IEEE 37th International Conference on Distributed Computing Systems

RFIPad: Enabling Cost-efficient and Device-free In-air Handwriting using Passive Tags

Han Ding*, Chen Qian†, Jinsong Han*, Ge Wang*, Wei Xi*, Kun Zhao*, and Jizhong Zhao*
*School of Electronic and Information Engineering, Xi'an Jiaotong University
†Department of Computer Engineering, University of California Santa Cruz

***Abstract*—An important function of smart environments is the ubiquitous access of computing devices. In public areas such as hospitals, libraries, and airports, people may want to interact with nearby computing systems to get information, such as directions to a hospital room, locations of books, and flight departure/arrival information. Touch screen based displays and kiosks, which are commonly used today, may incur extra hardware cost or even possible germ and bacteria infection. This work provides a new solution: users can make queries and inputs by performing in-air handwriting to an array of passive RFID tags, named RFIPad. This input method does not require human hands to carry any device and hence is convenient for applications in public areas. Besides the mobile and contactless property, this system is a cost-efficient extension to current RFID systems: an existing reader can monitor multiple RFIPads while performing its regular applications such as identification and tracking. We implement a prototype of RFIPad using commercial off-the-shelf UHF RFID devices. Experimental results show that RFIPad achieves >91% accuracy in recognizing basic touch-screen operations and English letters.**

I. Introduction

Ubiquitous access and interaction to computing devices is a fundamental task provided by smart environments. In public areas such as clinics, hospitals, libraries, museums, parking garages, and airports, people may want to interact with computing systems in the environment to get important feedbacks, including directions to clinic and hospital rooms, locations of books, free slots in the garage, and flight departure/arrival information. Legacy input methods such as keyboards, mice, and touch screens are all contact-based: they require physical contact from users. Limitations of contact-based input methods in a public area include: 1) the devices are easily to be overused by large volume of people and become broken. Hence extra cost and timely device replacement are necessary; 2) they may incur possible germ and bacteria infection [18], especially for clinics and hospitals where health protection is in top priority.

Contactless input methods usually require users to perform hand or body motions and recognize these motions to be input information. One possible approach is the imagery-based devices, *e.g.*, Xbox Kinect [1]. However these devices incur non-trivial cost and could be easily damaged in public areas. In addition, they require line-of-sight and raise concerns on user privacy. Another approach, which is our focus in this paper, is to use radio frequency (RF) based sensing techniques of wireless devices [17], [22], [26], [33]. Prior works on RF-based motion estimation can be categorized into two groups, device-binding methods and device-free methods. Device-binding methods require every user to carry a sensing device, such as wearable sensors [22] or RFID tags [26]. These solutions are not feasible to applications in public areas because we shall not expect everyone visiting a public area to carry a designated tag or sensor. Furthermore, existing device-binding methods require rigorous constraints for better accuracy performance, yielding costly and inflexible deployment in real applications. For example, RF-IDraw [26] and Tagoram [33] should be equipped with 8 and 4 spatially-separated antennas for motion recognition. On the other hand, device-free methods become attractive due to releasing the limit of carrying special devices, as well as the line-of-sight constraint in imagery-based approaches. Existing device-free approaches [5], [17], [23], however, are mainly depends on dedicated devices and complex signal processing techniques, limiting the widely adoption in public areas.

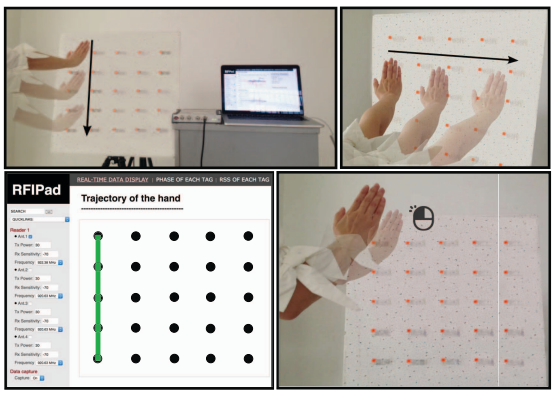


Fig. 1. Top: (Left) Writing stroke "|". (Right) Right turning motion. Bottom: (Left) Real-time stroke illustration when writing "|". (Right) Click motion.

In those public environments we mentioned, an ideal contactless input method should satisfy: 1) cost-efficient; 2) no device-binding; and 3) realtime reaction and no training period. *No prior solution satisfies all requirements.* This paper proposes a cost-efficient solution for real-time and ubiquitous human-computer interaction via RF based hand motion recognition techniques. Radio Frequency Identification (RFID) has been widely accepted and deployed in public areas for tracking and monitoring objects. The proposed method RFIPad is *a cost-efficient extension to existing RFID systems*. For public areas where an RFID reader is already deployed, the extra cost of RFIPad is only a few passive tags (in some cases an extra antenna may be needed), which are extremely cheap (50 cents per tag) compared to imagery and other RF devices. RFIPad is device-free and does not require training period.

1063-6927/17 $31.00 © 2017 IEEE
DOI 10.1109/ICDCS.2017.141

IEEE computer society

RFIPad is inspired by the following phenomena observed in our experiments. If a person moves a hand over an ultra high frequency (UHF) tag array, the signal backscattered from those tags is disturbed significantly. For different hand motions, the disturbance to those tags presents distinct profiles, in terms of changes on both the phase and radio signal strength (RSS). We realize that by combining with reported tag IDs and timestamps, such differentiation can be utilized to correlate to the moving trajectory of hands. In this way, we can identify fine-grained movements of hands, *i.e.*, in-air handwriting.

Our prototype of RFIPad is shown in Fig. 1. We deploy an array of tags as a sensing plate. The user just moves one of his/her hands above the plate to operate or write in air. The induced disturbance of RF signals can be collected and correlated to the motions/letters written by the user. RFIPad does not require training of user behaviors and can react to hand motions instantly. Another advantage of RFIPad is minimal deployment cost by utilizing existing RFID infrastructure and the full compatibility with industrial standards, *i.e.*, EPC Global C1G2 [11]. We implement a prototype of RFIPad using COTS RFID devices.

Contributions. We make the following contributions:

- RFIPad is a device-free in-air hand motion recognition scheme and a cost-efficient extension of existing RFID systems by only adding very cheap components. It exhibits several attractive merits, including the ability of estimating the motion direction and resiliency to environment changes.
- RFIPad achieves device-free fine-grained motion recognition of hands using passive tags, without prior training process. To our knowledge no other system can achieve this goal.
- We demonstrate the feasibility and effectiveness of our RFIPad via a prototype implementation of real COTS RFID devices. The results show that it can detect different motions with an accuracy of about 94% and recognize English letters with an accuracy of 91%.

II. Background and Overview

In this section, we first discuss the background and important observations in our preliminary experiments, and then present the overview of RFIPad.

A. Backscatter communication in RFID systems

Communication in passive RFID systems is based on backscatter radio links. An RFID reader communicates with passive tags in the full-duplex mode. It transmits continuous waves (CW) and keeps listening to the tag signals, *i.e.*, responses. The passive tags are usually not equipped with a battery or radio transmitter. Instead, it harvests power from the constant CWs and modulates its ID and other information for responses to the reader.

B. Observations

In a typical passive RFID system, as the reader continuously interrogates a set of tags, it can obtain the following information about each tag: the tag's ID and channel parameters. In particular, the channel parameters include Doppler Frequency Shift (Doppler), Phase, and Received Signal Strength (RSS). Doppler depicts the relative movement between the reader and the tag. Phase and RSS jointly reflect the RF situation around the tag, sometimes they are used as a signature of the ambient environment. We conduct preliminary experiments for observing the channel parameters of tags. In those experiments, we find that all parameters are nearly constant in the static environment, as the black lines shown in Fig. 2. And the most striking insight comes with above experiments is that, when there is a hand-movement around the tag, its disturbance triggers a significant variation in the tag's channel parameters, as shown by the red lines in Fig. 2. In particular, although the variation in Doppler is indistinguishable (due to the large noise) between the static and hand-movement cases, it shows distinct difference between the two cases in both Phase and RSS. With above insights, we attempt to utilize such variations in Phase and RSS to identify fine-grained hand motions.

C. Our idea

To allow inputs of touch screen operations and English letters, we define 7 basic hand motions, including "$\odot$", "$-$", "$|$", "$/$", "$\backslash$", "$\subset$", and "$\supset$" (Number #1 to #7). Three of them (*i.e.*, "$\odot$", "$-$", and "$|$") support the basic operations of touch screen. Specifically, $\odot$ denotes a "push" towards a certain tag, representing the "click" motion. Both "$-$" and "$|$" contain two directions (*e.g.*, "$\leftarrow$", "$\rightarrow$", "$\uparrow$", "$\downarrow$") and hence support page-swiping and scroll-bar controlling respectively. On the other hand, "$-$", "$|$", "$/$", "$\backslash$", "$\subset$", and "$\supset$" combine the basic strokes of English letters [6], supporting the in-air handwriting inputs. For example, the alphabet "H" is composed of 3 successive strokes, namely "$|$", "$-$", and "$|$". If these discrete strokes can be identified by recognizing hand-movements, it would be possible to infer the corresponding letter.

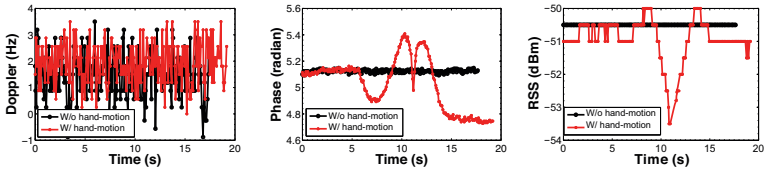
We deploy our RFIPad system over a group of tags, which are well-positioned to form an array. An RFID reader continuously collects the IDs and channel parameters of those tags by keeping interrogations on them. As aforementioned, when a human hand moves over the array, the movement will influence some tags in sequence. RFIPad correlates the hand positions to the tag positions, and identifies the corresponding motion by analyzing the variations in the channel parameters of those tags over time. Furthermore, after recognizing the sequential strokes (motions) written in-air, RFIPad is able to form the input letters accordingly.

III. RFIPad Design

In this section, we introduce the design of RFIPad. Our approach consists of three components: basic motion/stroke recognition, direction estimation, and English letter recognition.

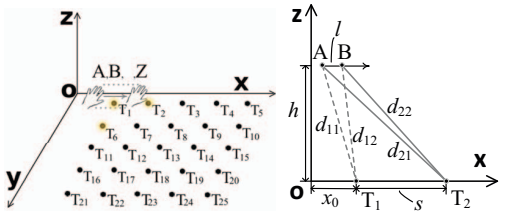
A. Phase-based motion/stroke recognition

RFIPad proposes to utilize phase to conduct motion recognition. The reasons are twofold. Firstly, the phase value reported by the reader has high resolution (*i.e.*, 0.0015 radians, which offers $\approx 320mm \times 0.0014/(4 * 3.14) = 0.038mm$ distance resolution). Secondly, we have the following theoretical analysis.



(a) (b) (c)

Fig. 2. Doppler, phase, and RSS values measured over time.



(a) (b)

Fig. 3. Theory model.

1) Theoretical analysis: In our system, the relative location of the reader antenna and tags are constant, hence the signal backscattered from each tag is nearly stable in the static environment. When a human hand moves above the tag array, certain RF signals are reflected from the moving hand to the tags. Since the hand is near the tag, the influence of reflections is a non-trivial contribution in the tag's received signal. Thus, the hand can be regarded as a powerful virtual transmitter that generates the reflected signals [23]. As shown in Fig. 3(a), suppose the hand moves over T_1 (*i.e.*, tag #1) along the direction of x-axis. Its trail is from location A to Z. The induced phase change of each individual tag during this procedure shall strongly correlate to the wave path-difference. We assume that at a certain time during above movement, the hand moves to position B. Theoretically, the phase difference of tag T_1 and its adjacent tag T_2 can be represented respectively as:

$$\Delta\theta_{T_1} = (\frac{2\pi}{\lambda} \times \Delta d_{T_1}) \ mod \ 2\pi \tag{1}$$

$$\Delta\theta_{T_2} = (\frac{2\pi}{\lambda} \times \Delta d_{T_2}) \ mod \ 2\pi \tag{2}$$

where $\Delta d_{T_1} = d_{12} - d_{11}$, $\Delta d_{T_2} = d_{22} - d_{21}$. As the relative positions shown in Fig. 3(b), we further have

$$\Delta d_{T_1} = \sqrt{h^2 + (x_0 - l)^2} - \sqrt{h^2 + x_0^2} \tag{3}$$

$$\Delta d_{T_2} = \sqrt{h^2 + (x_0 + s - l)^2} - \sqrt{h^2 + (x_0 + s)^2} \tag{4}$$

Since $y = \sqrt{h^2 + (x - l)^2} - \sqrt{h^2 + x^2}$ is a monotonous decreasing function, we have $\Delta d_{T_1} > \Delta d_{T_2}$. Hence $\Delta\theta_{T_1} > \Delta\theta_{T_2}$. (Note that the periodicity of phase value reported by the RFID reader may introduce error into the phase difference comparison. Hence, the phase we use in our system is the unwrapped one after de-periodicity operation, which is detailed in Section III-A3.)

Based on above analysis, considering the whole procedure of hand-movement from A to Z, it is conceivably hypothesised that $ideally$, the accumulative phase difference of T_1 shall be larger than T_2, *i.e.*, $\sum_A^Z \Delta\theta_{T_1} > \sum_A^Z \Delta\theta_{T_2}$. The principle also holds along y-axis, thus $\sum_A^Z \Delta\theta_{T_1} > \sum_A^Z \Delta\theta_{T_6}$. Following above analysis, we propose to detect which tag the hand passes through in a specific time duration by examining which tag is with the maximum accumulative phase difference among all

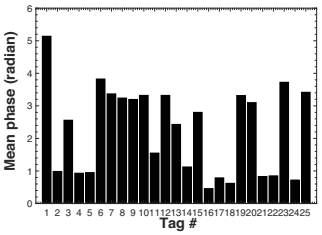


Fig. 4. Average phase value of different tags in the static scenario.

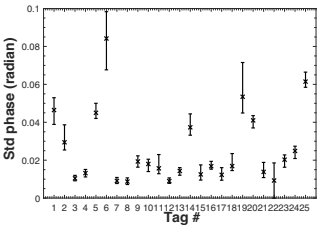


Fig. 5. Standard deviation of phase measurements of different tags.

tags. Suppose there are N tags, we formulate above principle as the following target function:

$$I_i = arg \ max_i \sum_{i=1}^{N} \Delta\theta_{T_i} \tag{5}$$

To validate above theoretical analysis, we deploy a 5×5 tag array on a white board, as illustrated in Fig. 1. We ask a volunteer to move his hand over the tags in the third column in sequence and collect the phase values of each tag. Unfortunately, the result calculated by Equation 5 is inconsistent with our ideal hypothesis. The phase value of each tag varies along different central values. Meanwhile, when the hand passes through a tag, it is possible that another tag's accumulative phase difference is larger than that of this tag. We conjecture two reasons emerged from above phenomena: *tag diversity* and *location diversity*. The tag diversity (*i.e.*, device diversity) is introduced during the manufacture [33], and the location diversity brings different multipath experiences to different tags [24]. Hence, these two diversities jointly affect the phase measurement of tags in the layout. We conduct experiments to demonstrate the influence. We collect the signal of tags in the 5×5 array. Each tag is interrogated for 100 times and there is no hand-movement during the collection. The average phase value is reported in Fig. 4. The result proves that the phase value of each tag irregularly distributes within the range of [0, 2π], *i.e.*, near different central values. We also test the stability of each tag's RF signal. Fig. 5 shows the standard deviation of phase measurement derived from multiple groups of experiments in static scenario. The result shows that the phase value vibrates in different levels, *i.e.*, the standard deviation of tags varies significantly (we call it *Deviation bias*). These results together indicate that it is necessary to suppress the diversities before conducting motion recognition using the proposed hypothesis.

2) Diversity suppression: Suppose the distance from the reader to a tag is d. Thus, the RF signals sent from the reader's antenna traverse $2d$ in the backscatter communication. The phase of those signals can be expressed as [16]: $\theta = (2\pi\frac{2d}{\lambda} + \theta_T + \theta_R + \theta_{tag})\ mod\ 2\pi$, where λ is the wavelength, θ_T, θ_R, and θ_{tag} are the phase rotations introduced by the reader's transmission circuits, the reader's receiver circuits, and the tag's reflection characteristic, respectively. Note that θ_{tag} varies among tags due to hardware imperfection (*i.e.*, tag diversity).

RFIPad leverages sequential samplings on a tag's RF signals to compensate the tag diversity. Let θ_{ij} denote the phase measurements of tag i during the procedure the hand moves, where j corresponds to the jth RF sample collection. Suppose $\widetilde{\theta}_i$ is the average phase measurement of tag i in the static environment, d_{ij} is the distance between the reader and tag i at the time point of collecting the jth sample. We have:

$$\widetilde{\theta}_i = (2\pi\frac{2\widetilde{d}_i}{\lambda} + \theta_T + \theta_R + \theta_{tag_i})\ mod\ 2\pi \tag{6}$$

$$\theta_{ij} = (2\pi\frac{2d_{ij}}{\lambda} + \theta_T + \theta_R + \theta_{tag_i})\ mod\ 2\pi \tag{7}$$

Subtracting Equation 7 from Equation 6, we have:

$$\theta'_{ij} = (4\pi\frac{d_{ij} - \widetilde{d}_i}{\lambda})\ mod\ 2\pi,\ \ j = 1, 2, 3, ... \tag{8}$$

From Equation 8, the parameters θ_T, θ_R, and θ_{tag} are naturally wiped out. As a result, the impact of tag diversity is suppressed. Meanwhile, by doing so, the phase value of each tag shall vibrate around the same central value (*i.e.*, around zero).

On the other hand, RFIPad suppresses the location diversity through two processes. We first estimate the *Deviation bias* of each tag i (b_i) in the array, and then formulate a weighting function:

$$w_i = \frac{E(b_i)}{\sum_{i=1}^{N} E(b_i)} \tag{9}$$

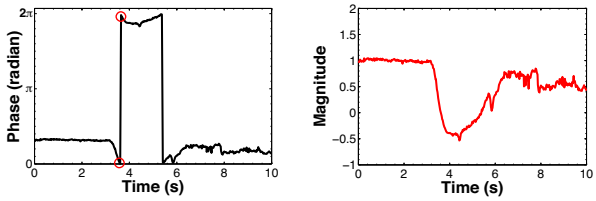
where $k = j + 1,\ j = 1, 2, .., M - 1$ (M is the number of phase samples), $E(b_i)$ is the expectation of standard deviation derived from each tag in static scenarios, and N is the number of tags. Then, the accumulative phase difference of tag i can be revised as:

$$I'_i = w_i^{-1} \sum (\theta'_{ik} - \theta'_{ij}) \tag{10}$$

The principle of above weighting function is simple: a tag with higher b_i indicates that it will be more sensitive to the ambient changes, and hence the phase measurements on this tag (or at its location) should be appropriately weakened. In contrast, a tag with lower b_i indicates the influence caused by the location diversity is lower, and the phase measurements on this tag deserves a higher weight. With such adaptive tuning, the location diversity can be effectively mitigated.

3) Motion recognition: An important step of RFIPad is to correlate the phase measurements to different hand motions/strokes. The process includes three components: phase de-periodicity, accumulative phase difference calculation (Equation 10), and image-assisted motion recognition.

Phase de-periodicity. The phase value reported by RFID readers variates periodically from 0 to 2π. As a result, it is possible that the phase might present a sudden change in its value from 0 to nearly 2π (Fig. 6(a)) or vice versa, which will cause errors when calculating Equation 10. To tackle this problem, we adopt the method proposed in [14] to unwrap the phase value. As the example shown in Fig. 6(a) and (b), the phase trend suffering a sudden phase change becomes smooth and continuous after the unwrapping process.



(a) phase trend before unwrapping (b) phase trend after unwrapping

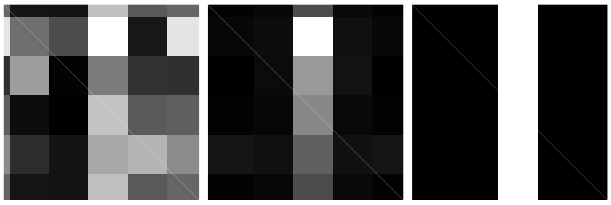
Fig. 6. Phase de-periodicity.

Image-assisted recognition. After above preprocessing, we are able to calculate I'_i for each tag i. For better illustration, we visualize I'_i values in a grey-scale image and use the OTSU's algorithm [21], which performs clustering-based image thresholding, to convert the image to a binary image. Specifically, after OTSU's algorithm the image only contains two classes of pixels, foreground and background pixels (*e.g.*, '1' and '0'). The '1' pixels implies that the hand just moves over the corresponding area. As a result, RFIPad can identify the motion by simply estimating the '1's in the tag (pixel) array. The detail of OTSU's algorithm can be referred in [21].

As an example, Fig. 7 shows the image generated when a volunteer moves his hand across the third column in the array. The whiter the pixel is, the larger I'_i value the tag bears. Fig. 7(a) and (b) are the results before and after the process of diversity suppression, respectively. These results show that the diversity interference is significantly weakened when adopting our suppression algorithm. Then, we conduct the OTSU's algorithm on the image in Fig. 7(b). The result is exhibited in Fig. 7(c). We can find that the hand-movement area is explicitly outlined.

B. RSS-based direction estimation

Above building blocks provide a geometric representation of the air-written stroke/motion. In practice, the direction can bear useful information. Generally, two motions in a same shape of trails but with opposite directions usually correspond to opposite operations, like *open* and *close*. In the following, we recognize the direction of the motion by estimating the sequence of tags that the hand passes. For instance, we distinguish whether the direction of hand-movement is $\downarrow$ or $\uparrow$ in Fig. 7(c). There are two options for achieving this objective, through phase or RSS. Phase measurement depends on both



(a) without diversity suppression (b) with diversity suppression (c) after OTSU algorithm

Fig. 7. Illustration of motion identification when a volunteer moves his hand across the third column in the array.

the location of the tag in the layout and the relative position of moving hand to it. Hence, when moving the hand over different tags, phase profiles could be totally different. As demonstrated in Fig. 8, the phase trend might be monotonous, axial symmetric, or circular symmetric. Thus, inconsistent pattern in phase trend makes it difficult to infer the sequence of tags. On the other hand, RSS profile is more distinguishable: there will always be a distinct $trough$ when the hand is moving perpendicularly over the tag. The principle of detecting the tag ordering is simple but effective. By observing the appearance order of the tags' troughs, we can acquire the sequence of tags influenced by the hand-movement. We obtain the order of the influenced tags by employing a two-staged RSS trough estimation solution. Due to space limit, we skip the technical details.

C. Composing English letters

RFIPad also provides the capability of recognizing English letters.

1) Segmentation: RFIPad first separates strokes from continuous phase streams. The stroke segmentation scheme is based on the fact that people tend to take a short pause after each stroke, in which the hand/arm will be raised/adjusted to the starting position of next stroke. We term such a procedure as the *adjustment interval*. RFIPad detects the adjustment interval to separate strokes apart according to two insights:

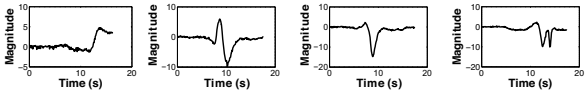
• When there is a hand moving over the tag array, each tag shall experience distinct variations in its phase values.

• In the adjustment interval, phase variations of all tags are relatively small and stable.

To mitigate the possible interference from unevenly distributed sampling along time, we segment the phase streams into non-overlapped frames. Each frame, denoted as f, is 100ms long. Then, we calculate the Root Mean Square (RMS) of each frame:

$$rms(f) = \sum_{i=1}^{M} \sqrt{\frac{\sum_{j=1}^{n} p_{ij}^2}{n}} \tag{11}$$

where M is the number of tags in the plane, n is the number of phase samples in current frame, and p_{ij} denotes the j-th phase sample of the i-th tag. Since a stroke (motion) may last for seconds and cover several frames, it is difficult to determine whether a stroke occurs or not merely based on the features of a single frame. Therefore, several successive frames are grouped as a window (w) and each window is treated as a unit for processing. The default window size in RFIPad is



(a) (b) (c) (d)

Fig. 8. Symmetrical characteristic of phase trends. (a) monotonous (b) circular symmetrical (c) axial symmetrical (d) axial symmetrical.

set to 0.5 second (*i.e.*, 5 frames). To determine whether the stroke occurs in (or during) a given window, we measure the standard deviation of the window's RMS, and see if it exceeds a threshold value:

$$std(rms(w)) > thre \tag{12}$$

where *thre* can be empirically determined. As an example, Fig. 9 shows the average phase, RMS and Std(RMS) when a volunteer writes letter 'H'. We can see that in the adjustment interval, the Std(RMS) is nearly 0, and during the stroke, Std(RMS) values are much larger (Fig. 9(c)). Therefore, we can effectively filter out the adjustment intervals and identify the stroke windows for stroke detection.

2) Deducing letters: After the stroke/motion detection, RFIPad composes identified strokes into letters. We adopt the tree-structure grammar (shown in Fig. 10) proposed in [6] to recognize letters. For instance, RFIPad observes two strokes " $-$ " and "|" in sequence. With the hint from tree-structure grammar, these motions are identified as letter "T". Interestingly, RFIPad can naturally eliminate such ambiguities as "D" and "P", or "O" and "S", although they are composed by the same sequence of strokes. This is because that according to the human writing habit and the property of letters, when writing "D", the last position of " $\supset$ " is usually overlapped with the bottom of stroke "|". Such a physical position information is useful for differentiating the letters with similar stroke sequence. Fortunately, RFIPad can easily obtain such position information of strokes by estimating the corresponding tag IDs. Thus, RFIPad is able to explicitly organize strokes into letters with no doubts, even for those with the similar stroke sequence. For current implementation we only focus on recognizing individual letter. We will leave the recognition of a succession of letters as our future work.

IV. Implementation

A. System setup

Hardware: we implement the prototype of RFIPad using COTS UHF RFID devices, including an Impinj Speedway R420 reader, a Laird antenna model A9028R30NF, and a set of passive tags (5 × 5). The size of antenna is 25.4cm × 25.4cm × 3.8cm and with 8dBi gain. The whole system operates on the frequency of 922.38MHz. The reader is connected to a backend PC via an Ethernet cable. It continuously reports the features of backscattered tag signal, such as ID, phase, and RSS.

Software: the software of RFIPad is implemented using C# and adopting the LLRP [12] protocol for communicating with the reader. We modify the Octane SDK (an extension of LLRP Toolkit) to enable the phase reporting. The software runs on

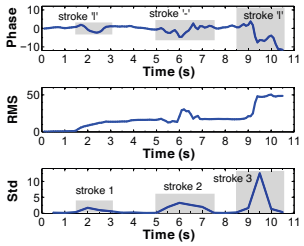


Fig. 9. The Phase, RMS, and Std(RMS) values when a volunteer writes a letter 'H'.

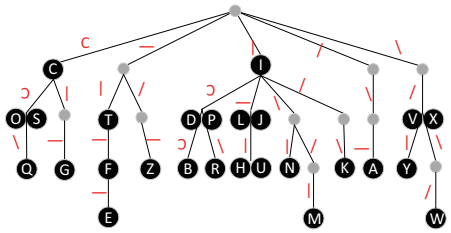


Fig. 10. Tree-structure grammar.

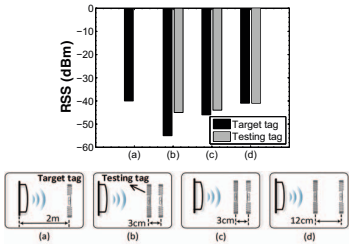


Fig. 11. Interference within a pair of tags.

a Lenovo PC, which is with an Intel Core i7-4600U 2.10GHz CPU and 8GB RAM.

B. Deployment issues

In the backscatter communication, the radio signal sent from the reader is converted by the tag antenna into currents. The current is used to power up the tag's IC to operate. Meanwhile, it also affects the electric fields of nearby tags [15]. Hence, before deploying the tag array, it is worth to study the interference among tags.

1) Interference between a pair of tags: To test the interference within two tags adjacent to each other, we conduct a series of experiments. As shown in Fig. IV(a), we put a single tag, called *target tag*, 2m away from the reader antenna. Its RSS is about -41bBm. When another tag, denoted as *testing tag*, approaches, we observe the RSS change of the target tag. In particular, we also check the situation when two tags are placed very close (*i.e.*, about 3cm, within the near-field region $\lambda/2\pi \approx 5.2cm$) and parallel to each other. With two tags' antennas toward a same direction (Fig. IV(b)), the target tag's RSS experiences a significant decrease. It indicates that the testing tag produces a shadow effect to the target tag, resulting in a significant suppression of the target tag's receiving power. Even worse, if the receiving power is weakened below the threshold that enables IC operations, the target tag will become unreadable. The results of our extensive experiments suggest two possible ways to mitigate above interference. One is to select a proper antenna directions. Instead of putting them in a same direction, we deploy the two tags opposite to each other, as shown in Fig. IV(c). We find that the target tag's RSS value will become similar to that of testing tag, implying that the receiving power of target tag is not suppressed too much. Another possible way is to enlarge the distance between the pair of tags. In particular, when the distance is larger than 12cm (*e.g.*, the far-field region $2\lambda/2\pi$), as shown in Fig. IV(d), the interference from the testing tag to the target tag becomes nearly negligible. Considering the constraint of the size of the deployed tag array, we suggest to deploy adjacnet tags with 6cm distance (*i.e.*, the transition region between the near-field and far-filed region) and with antennas facing opposite directions.

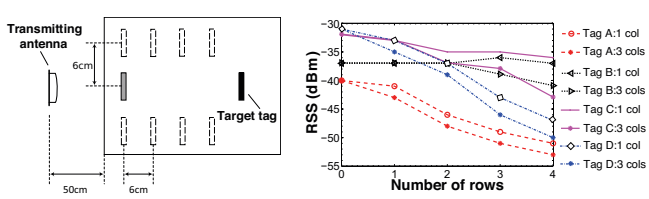
2) Interference within the tag array: The interference becomes more complicated within a tag array. Each individual tag will be severely affected by the shadow effect produced by the tags around it. We examine this interference here. As shown in Fig. 12(a), the reader antenna is 50cm from the plane on which the array is deployed, and the tags in this array are spaced 6cm lengthways and 6cm laterally. We vary the type and the number of tags that are placed on the plane. In particular, we test four kinds of commercial tags with different manufacturers (*e.g.*, Impinj and Alien) and different antenna designs (Fig. 12(c)). In each test, we use the same type of tags in the array. We treat another tag behind the array as the target tag and observe the interference coming from other tags. The results are plotted in Fig. 12(b). We have following observations:

- When there is only one column of tags in the array, with more tags in the column, the shadow effect becomes larger (*i.e.*, RSS of the target tag reduces gradually).
- When additional columns of tags are added in the array, the RSS of the target tag is further reduced.
- The interference appears diverse among variant types of tags. For example, three columns of Tag D reduces the received power of the target tag by 20dBm. While using Tag B in the three-columns array only reduces the receiving power by 2dBm at the target tag on average.

Former two observations are straightforward. The third observation can be explained by the finding proposed by [10]. That is, the unmodulated Radar Scattering Cross-section (RCS) determines the tag radiative efficiency as well as the backscattered power. RCS is defined as the ratio of power radiated by the tag to the power density incident on it. Generally, a smaller antenna corresponds to a smaller RCS, due to its higher reactive impedance and smaller radiation resistance. Intuitively, the less radiated power it yields, the less interference it injects to nearby tags. Hence, to mitigate the tag scattering effects, the simplest approach is to use those tags with a small RCS. This finding suggests that among these four kinds of tags, Tag B (Impinj AZ-E53) is the best choice for deploying the tag array.

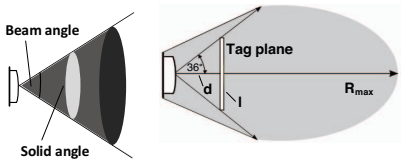
3) Distance between the plane and reader antenna: In real world implementation, it is necessary to set the minimum distance from the reader antenna to the tag plane, for ensuring that every tag is readable. It is known that typical passive RFID systems are forward-link limited, *i.e.*, tags require the incident power sent from the reader sufficiently high to operate their ICs. To this end, two essential factors deserve intensive selection, the *gain* and *read zone* of the reader's antenna. We investigate the effective solution for determining the distance from these two perspectives.

The *gain* of the directional antenna represents the maximized radiation intensity. Assume that the energy radiated by the antenna is uniformly distributed with a solid angle (Ω_s),



(a) (b) (c)

Fig. 12. Measured RSS with various number of rows and columns in the plane populated with tags, for four different commercial tag designs.



(a) (b)

Fig. 13. Idealized radiation pattern of a directional antenna.

as shown in Fig. 13(a). Then, the gain (G) of the antenna can be approximated as the ratio of the solid angle to the area of the unit sphere (*i.e.*, 4π):

$$G \approx \frac{4\pi}{\Omega_s} \tag{13}$$

Above equality holds only if the antenna radiates all the power it receives. In this case, the beam angle of the antenna can be approximated as:

$$\theta_{beam} \approx \sqrt{\Omega_s} = \sqrt{\frac{4\pi}{G}} \tag{14}$$

Hence, a higher gain of the reader antenna results in a narrower beam. According to Equation 14, the beam angle of reader antenna in RFIPad's prototype is about 72° ($\sqrt{4\pi/8} \approx 72°$).

Ideally, a directional antenna exhibits an ellipsoidal radiation pattern. In practice, the read zone of the reader's antenna, however, is hard to outlined, since it is influenced by many factors, such as the sidelobes and the reflections introduced by moving objects or environments. Fig. 13(b) shows an ideal read zone. The zone can reach its longest distance along the center beam of the antenna (*e.g.*, R_{max}), and become shorter toward the edges. Here R_{max} depends on the transmitted power of the reader.

To ensure tags in the plane have balanced coverage from the reader, we put the tag plane parallel to the antenna panel. In this part, we show how to set the required minimum distance (d) between the antenna and tag plane for RFIPad's prototype. According to the tag size (4.4cm) and interval between adjacent tags (6cm), the length of our tag plane (l) is about 46cm. Based on trigonometric formula, $d = \frac{l/2}{\tan 36°} \approx 31.7cm$, which is the minimum distance from the reader antenna to tag plane, given that all tags are with the 3dB beam coverage. Although RFIPad requires particular location and direction of a reader antenna. There is no constraint of the location of the reader, the most expensive component of an RFID system. A reader can carry multiple antennas and connect them via wires.

V. Evaluation

In this section, we evaluate the performance of RFIPad.

A. Experimental setups

We deploy a 5×5 tag array on a carton, with 6cm interval between adjacent tags. The reader antenna is mounted in two positions (as shown in Fig. 14). One is on the ceiling to simulate the LOS scenario (the left side of Fig. 14). In this case, the hand will move across the line-of-sight (LOS) paths between the antenna and certain tags. The other position is behind the board to simulate the non-line-of-sight (NLOS) scenario, in which the hand movement involves in the reflections. We conduct the experiments in four different locations in a typical office environment. Ten volunteers are invited to participate in the experiments, moving their hands over the tag array closely but without touch.

Ground truth collection. We use a Kinect to measure the user's motions. These measurements are actually the ground truth for comparison. The Kinect device is placed behind the user, facing the tag array plane so as to capture the infrared images of his/her motions. We manipulate the Kinect using a software development kit [2] and obtain the Kinect's skeletal output to track each volunteer's hand trajectory.

Metrics. We mainly evaluate the performance of RFIPad from the aspects of accuracy, False positive rate (FPR), and False negative rate (FNR). In particular, FPR is the percentage of falsely detected motions, and FNR is the percentage of undetected motions.

B. Motion detection

Since the motion/stroke is the basis of RFIPad, supporting both the touch screen like operations and English letters, we mainly focus on the motion/stoke detection performance in the evaluation part.

1) Reader antenna position: We first evaluate the influence of the reader antenna positions to the accuracy. We deploy the antenna at two positions (*i.e.*, the LOS and NLOS scenarios presented in Section V-A). The whole system operates on the frequency of 922.38MHz. The reader transmitting power is 30dBm. The distance from the reader to the tag plane is sbout 32cm in the NLOS scenario (We treat above parameters as default settings in the following). The volunteers are asked to perform 13 strokes (stroke 2∼7 with two directions), 20 times for each. We run 3 groups of above experiments and totally detect 780 hand motions. The results are shown in Table I. Surprisingly, RFIPad achieves the average detection accuracy of 94% for NLOS scenario, which is higher than that of LOS scenario (say 88%). A reasonable explanation on

TABLE I
Accuracy of Motion Identification

Case	Group 1	Group 2	Group 3	Average
LOS	0.88	0.86	0.91	0.88
NLOS	0.94	0.92	0.96	0.94

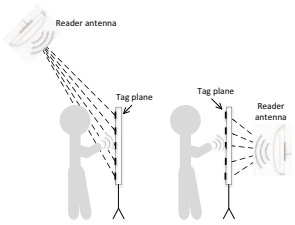


Fig. 14. Illustration of LOS and NLOS scenarios.

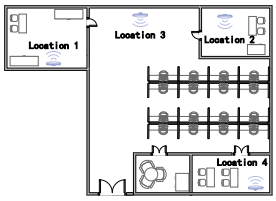


Fig. 15. Floorplan of the experimental lab locations.

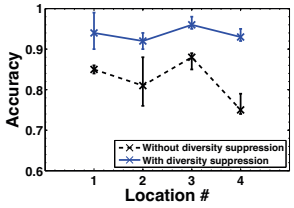


Fig. 16. Detection accuracy $v.s.$ different environments.

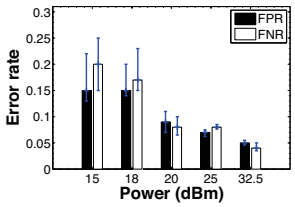


Fig. 17. Detection accuracy $v.s.$ different reader transmitting powers.

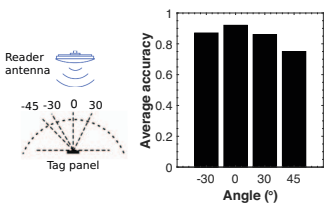


Fig. 18. Accuracy $v.s.$ reader-to-tag angles.

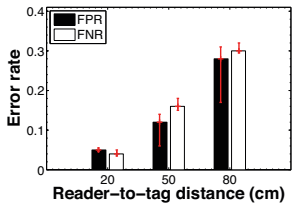


Fig. 19. Error rate $v.s.$ reader-to-tag distances.

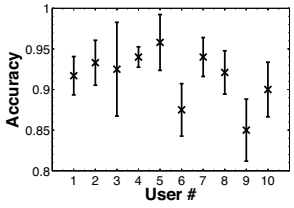


Fig. 20. Detection accuracy $v.s.$ different users.

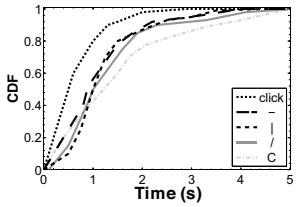


Fig. 21. Detection accuracy $v.s.$ the time to write the stroke.

this phenomenon is that the interference from human body, *e.g.*, arm, might introduce noises to the detection, because it is highly possible that the volunteer's arm also disrupts a number of tags' LOS paths to the reader. On the contrary, the arm's affect is much weaker in the NLOS scenario, since the hand is always closer to the tags than the arm. In what follows, we adopt the NLOS deployment mode in our experiments.

2) Evaluation in different environments: We run experiments in different environments (*i.e.*, four locations as shown in Fig. 15) in our lab to examine the effectiveness of the proposed diversity suppression algorithm. Similar to the previous experimental settings, each volunteers performs every stroke 30 times in each location. Fig. 16 shows the accuracy of RFIPad with and without the diversity suppression. The results indicate that after performing the suppression algorithm, the accuracy of RFIPad in all environments is improved to some extent. Specifically, we observe the largest accuracy improvement at location #4, *i.e.*, from 75% to 93%. Referring to Fig. 15, the tags at location #4 may experience the strongest multipath reflections from nearby objects, such as walls and tables, compared with the cases at other locations. Indeed, this result demonstrates the effectiveness of our diversity suppression algorithm.

3) Reader transmitting power: This group of experiments is to investigate the impact of the reader's transmitting power on detection accuracy. We vary the power from 15 to 32.5dBm. Note that the commercial reader power cannot exceed 32.5dBm according to the regulations. For each of the power levels, we have the volunteer conduct the 13 motions, 30 times for each. We plot the average false positive rate and false negative rate in Fig. 17. When the reader's power is 32.5dBm, the error rates are merely around 5%. The error rates gradually increase when reducing the power, and reach around 20% when the power becomes 15dBm. This result can be explained by the fact that the passive tags are battery-free. Since the energy is harvested from the reader, when the power level is high, the signal reflected from the hand shall have more significant influence on the approached tag. Hence, the data variation among tags tends to be more distinct. The enriched distinction facilitates distinguishing the motions that are unable to detect under a low reader's power setting. Therefore, we suggest to use larger reader transmitting power in real applications.

4) Reader-to-tag angle: We conduct experiments to vary the angle (including -30,0, 30,45) between the reader antenna plane and the tag panel. In this series of experiments, we ask a volunteer to perform the motions — and , over different columns and rows of the tag panel, each for 10 times. The top-view of the deployment and the average motion recognition results are shown in Fig. 18. The results show that the system achieves best performance at the angle of 0. As the angle increases, the motion recognition accuracy decreases. We recommend the 0 deployment for high accuracy. If the already-deployed reader has an angle to the tag panel, we may need preliminary signal processing or parameter tuning to ensure the system efficiency.

5) Reader-to-tag distance: We then check the impact of the reader-to-tag distance. In this trial of experiments, we vary the distance from the reader antenna to the tag plane from 20cm to 80cm. Other settings are consistent as default. The average detection error rates are shown in Fig. 19. As expected, shortening the distance can decrease the error rate. For example, FPR and FNR of RFIPad is only about 5% when the distance is 20cm. Further data analysis reveals that a larger distance may involve more complex environmental interference to the transmission from the reader and tags, which introduces irregular variation to the tag backscattered signals. Hence, we suggest that the reader-to-tag distance should not be longer than 50cm.

6) User diversity: Then we examine the usability of the system. We invite ten volunteers to stand at the front of our prototype to perform different strokes, 20 times for each. We

balance the diversity of volunteers in terms their gender (6 males and 4 females), age (ranging from 22 to 30), and other physical conditions (158 $\sim$ 183cm in height, 45 $\sim$ 80kg in weight, and 56 $\sim$ 70cm in the arm length, etc.). Note that they are naturally moving their hands during the operation or in-air writing. Fig. 20 shows the average accuracy and standard deviations for all motions. We see that most volunteers show comparable accuracy, and the median accuracy is above 90%. However, for two volunteers, *i.e.*, #6 and #9, the accuracy of RFIPad experiences a slight degradation, but still keeps at a high level (85%). We find that these two volunteers move their hands in a relatively fast speed. This motivates us to further investigate the impact from hand moving speed later. Overall, the result implies that RFIPad scales well cross diverse users.

7) Motion speed: The motion speed is an essential factor when performing the detection. Generally, different users move their hands at different speeds. For each volunteer, we record the time duration for successfully recognizing each stroke/motion across 300 rounds of experiments. Fig. 21 plots the cumulative distribution function (CDF) of the time used to correctly recognize a letter. Although RFIPad spends relatively a long time in recognizing a small portion of motions, 90 percent of recognitions can be completed within 2 seconds for motions/stokes "click", "$-$", "$|$", and "$/$". Specifically, stroke "$\subset$" takes a longer time than others, since the user needs to move a longer distance to fulfill it. Generally speaking, RFIPad prefers slow motions. This is consistent with the findings in Blink[34], which showed that the system suffers from undersampling when the hand moves at a high speed, resulting in ineffective detection on hand motions.

C. Letter recognition

We evaluate the accuracy of letter recognition in two phases: stroke segmentation and letter recognition. We first examine the stroke segmentation performance using two metrics: insertion rate and underfill rate. The insertion rate is the proportion of cases that our system detects a stroke within the repositioning period. It reflects how resilient RFIPad is when distinguishing the stroke data and the adjustment interval. The underfill rate is the proportion of the cases that the segmented stroke is incomplete. It indicates whether our method is capable to accurately yet completely excavate the entire data for a stroke. Fig. 22 shows the segmentation and letter deduction performance over 5 representative letters (L, T, Z, H, E), which incorporate 2 (L and T), 3 (Z and H), and 4 (E) strokes respectively. We find that the underfill rate is always lower than 0.07. The insertion rate is diverse for different letters. For example, a letter with a larger number of strokes experiences a larger insertion rate, which is consistent to our intuition, *i.e.*, recognizing it requires more times of segmentation. In a nutshell, the segmentation operation has an essential impact on letter deduction since it is the basis and first step for letter recognition. For achieving better segmentation performance, we suggest the user raises his/her arm when adjusting to the starting position of next stroke.

Then we focus on the accuracy of recognizing letters. The users are asked to write each of the 26 letters following the instruction of Fig. 10. We show the experiment results of recognizing letters in Fig. 23. They can be categorized into four groups. Each represents a typical type of letters in terms of the number of strokes. For instance, Group #1 includes "C" and "I" which have only one stroke. The others are Group #2 {D,J,L,O,P,S,T,V,X}, Group #3 {A,B,F,G,H,K,N,Q,R,U,Y,Z}, and Group #4 {E,M,W}. From Fig. 23, we can see that the average successful rate of recognizing a letter remains around 91%. This effectiveness of RFIPad's letter recognition is also confirmed by the ground truth data obtained from Kinect. For instance, one volunteer writes the letter "Z", and Fig. 25 shows the Kinect's skeletal points, the hand trajectories traced by Kinect and RFIPad. The two trajectories are very consistent.

D. Latency

Short latency is important for most interactive applications. We evaluate the latency of RFIPad through the response time of motion recognition, *i.e.*, the time between when a volunteer finishes one motion (stroke) and when the motion is correctly reported. For each motion, we randomly choose 50 records of motions from 10 users and calculate the response time. Fig. 24 shows the average response time for all the motions. Except two outliers for motion #5, the maximum response time across all motions is less than 0.1s. In particular, the largest difference of response time for each motion is less than 0.035s. This demonstrates the excellent capability of RFIPad in supporting online motion recognition.

VI. Limitation and Discussion

In the implementation of RFIPad, we encounter several challenging issues deserved discussions.

Distance from the tag plane: Although RFIPad realizes a "contactless" touch screen like interface, it still has a soft constrain on the distance from the user's hand to the tag plane. Our prototype achieves satisfactory accuracy if this distance is within 5cm. Beyond this range, the error would increase with enlarging the distance. The reason is that passive tags operate without an on-board power source, while they draw energy from the RF waves emitted from the reader's antenna. If the user is far from the tag plane, the interference introduced to the tags' signals might not be distinct enough for fine-grained gesture-sensing. As such, one of our future works is to loosen this constraint by advancing both the hardware and signal processing algorithm of RFIPad.

Low throughput: RFIPad prefers slow motion/hand-writing which would result in low throughput. The reason is that RFID system might suffer from under-sampling when the hand moves at a high speed. We envision that the problem could be mitigated by optimizing specific parameters involved in RFID systems, such as reducing the tag packet length (*e.g.*, reducing the communication time of each tag), avoiding the readings of tags out of our tag plane, and the like.

Compounding errors: For the current implementation, we adopt the tree-based grammar to recognize letters, *i.e.*, we deduce the letter after identifying each stroke sequentially. This would lead to compounding errors, *e.g.*, first in the segmentation, stroke recognition, then in the letter deduction. One possible direction to mitigate this interference is to treat a letter as a whole, and resort to image processing techniques for identifying the whole letter after RFIPad's OTSU operation.

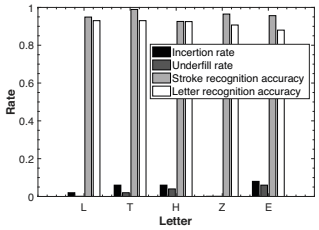


Fig. 22. The impact of stroke segmentation.

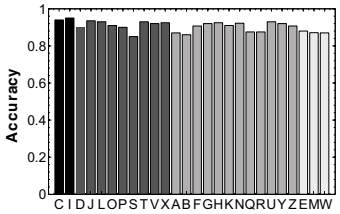


Fig. 23. The accuracy of letter recognition.

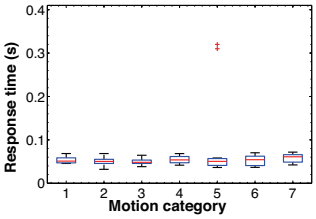
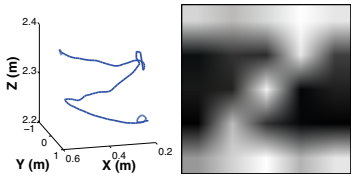


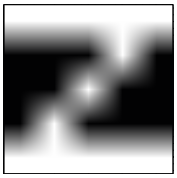
Fig. 24. Response time for various motions.

(a) Kinect skeletal tracking



(b) Hand trajectory of Kinect output

(c) Gray maps obtained by RFIPad



(d) Output after OTSU algorithm

Fig. 25. Kinect skeletal outputs and RFIPad's graymaps when a user writes "Z".

VII. Related Work

RFIPad bears similarity to prior works proposed for motion/gesture recognition and indoor localization.

A. Motion/Gesture recognition

Existing motion and gesture recognition related research can be classified into three categories.

Camera-based: Camera or vision based systems usually make use of depth sensors (*e.g.*, Wii [29]) or infrared cameras (*e.g.*, Kinect [1]) to trace fine-grained motions and enable 3-D in-air [3] human computer interactions. The basic requirement is the line-of-sight (LOS) guarantee between the device and user. Unlike these solutions, RFIPad yet does not require the dedicated hardware setup and LOS for operation.

Motion sensor-based: Sensor based systems rely on motion sensors (*e.g.*, accelerometer, gyroscope, and pressure sensor) in smartphones [6], smartwatches [8], [32], rings [13], or wristbands [22] to reveal human motions through machine learning techniques. However, they always require the instrumentation of users. In contrast, RFIPad achieves fine-grained motion tracing in a device-free manner.

RF-based: Recently, using RF signals [4], [5], [9], [17], [19], [20], [23], [28], [30] has shown the promising ability in both the localization and motion detection. Although some of those works provide attractive detection accuracy, most of them depend on either customized hardware or dedicated signal processing equipment, which are costly and inefficient in real world applications. For example, the works proposed in [4], [30] require GHz of bandwidth. In addition, several of them [5], [17], [35] require a priori learning on the signal patterns of certain predefined motions, which is the basis for later recognition on the those gestures. Sometimes, wearable devices are needed for supporting the accurate motion detection. For example, RF-IDraw [26] requires users to wear RFID tags on their fingers for tracking the trajectory of hand movements.

B. Indoor localization

There are also a large spectrum of related works focusing on the indoor localization [7], [26], [27], [31], [33], including the RSS based [7], phased based [33], and AOA based [26] techniques. The localization accuracy of those works ranges from a few decimeters to centimeters. Recently, quite a few approaches [25] adopt synthetic aperture radar (SAR) technique, which leverages moving antennas to form an antenna array for localization. RFIPad differs from those prior works in the aspect that it traces moving objects (hands) rather than localizing static objects. Moreover, it is device-free so that the traced object does not need to wear or carry any external devices. Another major difference from prior works is that RFIPad is fully compatible with current commercial RFID infrastructure, and does not require any hardware modification.

VIII. Conclusion

This paper presents RFIPad, which transforms a tag plane into a virtual touch screen, allowing a user to perform in-air handwriting and touch screen operations in public areas. RFIPad does not require the user to carry devices and introduces minimal extra cost. We implement a prototype RFIPad using COTS UHF RFID devices and conduct extensive experiments for evaluating its performance. The results show that RFIPad is fully compatible with existing industrial standard and exhibits high accuracy and reliability in motion detection and English letter recognition.

Acknowledgements

This work was supported by National Basic Research Program of China (973 Program) under Grant No.2015CB351705, NSFC Grant No.61572396, 61373175, 61402359, 61325013, China 863 Grant 2013AA014601, and National Science and Technology Major Project of the Ministry of Science and Technology of China JZ-20150910. Chen Qian was supported by National Science Foundation Grant CNS-1701681.

REFERENCES

[1] Microsoft kinect. http://www.microsoft.com/en-us/kinectforwindows/.

[2] Microsoft Releases Kinect for Windows SDK. http://latimesblogs.latimes.com/technology/2011/06/microsoft-releases-kinect-for-windows-sdk.html.

[3] Microsoft Research, Write in Air. *TechFest*, 2009.

[4] F. Adib, Z. Kabelac, D. Katabi, and R. C. Miller. 3D Tracking via Body Radio Reflections. In *USENIX NSDI*, 2014.

[5] F. Adib and D. Katabi. See Through Walls with WiFi. In *ACM SIGCOMM*, 2013.

[6] S. Agrawal, I. Constandache, S. Gaonkar, R. Roy Choudhury, K. Caves, and F. Deruyter. Using Mobile Phones to Write in Air. In *ACM MobiSys*, 2011.

[7] P. Bahl and V. Padmanabhan N. RADAR: An In-building RF-based User Location and Tracking System. In *IEEE INFOCOM*, 2000.

[8] S. Chen, H. Wang, and C. Romit Roy. I am a Smartwatch and I can Track my User's Arm. In *ACM MobiSys*, 2016.

[9] H. Ding, J. Han, L. Shangguan, W. Xi, Z. Jiang, Z. Yang, Z. Zhou, P. Yang, and J. Zhao. FEMO: A Platform for Free-Weight Exercise Monitoring with RFIDs. *Accepted to appear at IEEE Transactions on Mobile Computing*, PP, 2017.

[10] D. Dobkin. *The RF in RFID: Passive UHF RFID in Practice*. Elsevier Inc, 2008.

[11] EPCglobal. *Specification for RFID Air Interface EPC: Radio-Frequency Identity Protocols Class-1 Generation-2 UHF RFID Protocol for Communications at 860 MHz-960 MHz*, 2008.

[12] EPCglobal. *Low Level Reader Protocol (llrp)*, 2010.

[13] J. Gummeson, B. Priyantha, and J. Liu. An Energy Harvesting Wearable Ring Platform for Gestureinput on Surfaces. In *ACM MobiSys*, 2014.

[14] J. Han, H. Ding, C. Qian, W. Xi, Z. Wang, Z. Jiang, L. Shangguan, and J. Zhao. CBID: A Customer Behavior Identification System using Passive Tags. *IEEE/ACM Transactions on Networking*, 24(5):2885–2898, 2016.

[15] J. Han, C. Qian, X. Wang, D. Ma, J. Zhao, W. Xi, Z. Jiang, and Z. Wang. Twins: Device-free Object Tracking using Passive Tags. *IEEE/ACM Transactions on Networking*, 24(3):1605–1617, 2016.

[16] Impinj. *Speedway Revolution Reader Application Note: Low Level User Data Support*, 2010.

[17] B. Kellogg, V. Talla, and S. Gollakota. Bringing Gesture Recognition to All Devices. In *USENIX NSDI*, 2014.

[18] A. Khan, A. Rao, C. Reyes-Sacin, K. Hayakawa, S. Szpunar, K. Riederer, K. Kaye, J. T. Fishbain, and D. Levine. Use of portable electronic devices in a hospital setting and their potential for bacterial colonization. *American Journal of Infection Control*, 2015.

[19] J. Liu, M. Chen, S. Chen, Q. Pan, and L.-j. Chen. Tag-Compass: Determining the Spatial Direction of an Object with Small Dimensions. In *IEEE INFOCOM*, 2017.

[20] P. Melgarejo, X. Zhang, and P. Ramanathan. Leveraging Directional Antenna Capabilities for Fine-grained Gesture Recognition. In *ACM Ubicomp*, 2014.

[21] O. Noboyuki. A Threshold Selection Method from Gray Level Histogram. *IEEE Transactions on System, Man, and Cybernetics*, 9(1), 1979.

[22] A. Parate, M.-C. Chiu, C. Chadowiz, D. Ganesan, and E. Kalogerakis. RisQ: Recognizing Smoking Gestures with Inertial Sensors on Awristband. In *ACM MobiSys*, 2014.

[23] Q. Pu, S. Gupta, S. Gollakota, and S. Patel. Whole-home Gesture Recognition using Wireless Signals. In *ACM MobiCom*, 2013.

[24] L. Shangguan and K. Jamieson. The Disign and Implementation of a Mobile RFID Tag Sorting Robot. In *ACM MobiSys*, 2016.

[25] J. Wang and D. Katabi. Dude, where's my card? RFID Positioning that Works with Multipath and Non-line of Sight. In *ACM SIGCOMM*, 2013.

[26] J. Wang, D. Vasisht, and D. Katabi. RF-IDraw: Virtual Touch Screen in the Air using RF Signals. In *ACM SIGCOMM*, 2014.

[27] X. Wang, L. Gao, and S. Mao. PhaseFi: Phase Fingerprinting for Indoor Localization with a Deep Learning Approach. In *IEEE GLOBECOM*, 2015.

[28] B. Wei, W. Hu, M. Yang, and C. T. Chou. Radio-based Device-Free Activity Recognition with Radio Frequency interference. In *ACM IPSN*, 2015.

[29] Wii. *http://en.wikipedia.org/wiki/Wii.*

[30] J. Wilson and N. Patwari. See-Through Walls: Motion Tracking Using Variance-Based Radio Tomography Networks. *IEEE Transactions on Mobile Computing*, 10(5), 2011.

[31] Y. Xie, Z. Li, and M. Li. Precise Power Delay Profiling with Commodity wifi. In *ACM MobiCom*, 2015.

[32] C. Xu, P. Pathak H., and P. Mohapatra. Finger-writing with Smartwatch: A Case for Finger and Hand Gesture Recognition using Smartwatch. In *ACM HotMobile*, 2015.

[33] L. Yang, Y. Chen, X.-Y. Li, C. Xiao, M. Li, and Y. Liu. Tagoram: Real-Time Tracking of Mobile RFID Tags to High Precision Using COTS Devices. In *ACM MobiCom*, 2014.

[34] P. Zhang, J. Gummeson, and D. Ganesan. Blink: A High Throughput Link Layer for Backscatter Communication. In *ACM MobiSys*, 2012.

[35] Y. Zou, J. Xiao, j. Han, K. Wu, Y. Li, and L. M. Ni. GRfid: A Device-free RFID-based Gesture Recognition System. *IEEE Transactions on Mobile Computing*, pp(pp), 2016.