InDexMo: Exploring Finger-Worn RFID Motion Tracking for Activity Recognition on Tagged Objects

Rong-Hao Liang  
Eindhoven University of Technology  
rliang@tue.nl

Shun-Yao Yang  
National Taiwan University  
georgeyoung1231@gmail.com

Bing-Yu Chen  
National Taiwan University  
robin@ntu.edu.tw

ABSTRACT
This work explores and evaluates the designs of finger-worn radio-frequency identification (RFID) motion tracking for activity recognition on tagged objects. We propose an index-finger-worn device that consists of a short-range (~2cm) RFID reader and a pair of two inertial measurement units (IMUs), which are mounted at the locations where an artificial nail and a ring are worn. The short-range RFID reader recognizes the tagged object on finger touch, and then the IMU data are used for activity recognition. Data collected from the user of this device allows for a post-hoc analysis, which informs the activity recognition performance in various RFID+IMU and IMU-only configurations on the same task. The results of a ten-participant user study show that when the objects have similar physical form factors, the hybrid RFID motion tracking significantly outperforms the IMU-only tracking, especially in a larger-number set of objects. In our test, three IMU configurations (i.e., NailOnly, RingOnly, and Nail+Ring) achieved comparable action recognition performances, i.e., ≥90% accuracy, with 500 ms recognition time, though the NailOnly RFID+IMU configuration provided the highest wearability. The practical challenges toward a real-world deployment of a finger-worn RFID motion tracking system are also discussed.

CCS CONCEPTS
• Human-centered computing → Ubiquitous and mobile computing systems and tools.

KEYWORDS
Finger-worn device, nail-mounted device, RFID, motion tracking, activity recognition.

ACM Reference Format:

1 INTRODUCTION
Humans interact with the physical world using their hands; therefore, researchers on wearable interfaces have sought solutions that maximize the availability of human-computer interaction by appropriating the everyday surfaces near one’s hand as computer interfaces. Camera-based solutions [1, 4, 8, 14, 15, 24, 39] have been used for hand and object tracking; however, the line-of-sight occlusion problem limits the possible sensor location, which if inappropriate, may raise issues regarding technical validity, privacy [17], and social acceptance [18]. Data gloves [11] integrate occlusion-free low-level sensors, such as motion and bend sensors, into a socially acceptable glove form, but they also impede casual daily uses because the fabrics block the user’s haptic sensation and are hard to wear.

To facilitate everyday uses, several works have proposed incorporating low-level sensing with other conventional hand-worn forms. Devices have been developed in forms of wrist wears (e.g., watches [25, 26, 28, 41] and wristbands [9, 16, 32, 33]) and finger wears (e.g., rings [4, 20, 23, 31, 34, 38, 40], nails [5, 10, 21, 35], or nail+ring devices [7, 19]) to recognize user’s gestures as an alternative solution to the cameras and data gloves. Wrist-worn devices are effective in detecting macro hand and wrist movements, whereas finger-worn devices are generally more effective in detecting subtle movements of the fingertip(s) [5].

As the technical validity of activity (i.e., operation) recognition has been demonstrated and evaluated, the next prominent challenge would be recognizing the object (i.e., operand). Although previous works have shown the extra capability of object recognition-based machine learning [25, 26, 33], they may fail in distinguishing different objects that are of similar physical properties (e.g., form and material), which usually require tags for object identification. Instead, optical tags (such as barcodes), which usually need to be applied on an object surface, passive radio-frequency identification (RFID) tags are preferred, are used as they support embedded uses and straightforward maintenance [30]. Previous works have investigated the wearable uses of passive RFID tags (e.g., tattoo stickers [22], artificial nails [36]), and wrist-worn RFID motion sensors (e.g., gloves and bracelets [3, 12, 13]) that further recognize macro hand and wrist gestures on a tagged object. Nonetheless, the design space of RFID motion finger-worn tracking devices such as rings and nails, which are potentially more sensitive to subtle inputs, was not sufficiently explored.

In this paper, we research the design space through an exploratory study. We aim at designing a finger-worn RFID motion tracking device that not only identifies a tagged object but also recognizes the user’s activity performed on the object. Regarding object identification, we place a short-range tag reader on the fingernail so that it can read a tag under the fingertip. Regarding activity recognition,
we let the fingertip to be exposed and placed all sensors to the back.

we use inertial measurement units (IMUs) to track the index finger
posture as the main features of actions. From the index fingertip to
the root, the bones are interconnected by a one-degree-of-freedom
(DoF) distal interphalangeal (DIP) joint and a one-DoF proximal
interphalangeal (PIP) joint [2]. Such a constrained one-DoF posture
can be approximated using one or two IMUs. Figure 1 shows the
plausible realization in the nail or nail+ring forms. These forms have
different levels of wearing ease, but once worn, they are available
for mobile interaction.

The goal of this study is to understand the activity recognition
through users wearing finger-worn data collection devices. To make
the results meaningfully comparable, we propose an experimental
data collection device that can simulate the three forms shown in
Figure 1 so that we can evaluate the user’s performance on
the same task through a post hoc analysis. Figure 2 shows our
implementation, which is modified from a standard Proxinc PX572
fingerstall\(^1\) that can be easily and comfortably worn. The device
consists of a short-range (~2cm) RFID reader and a pair of two
IMUs, which are mounted at the locations where an artificial nail
and a ring are worn. To retain the user’s native haptic sensation,
we let the fingertip to be exposed and placed all sensors to the back
of the index finger. Two LEDs were mounted on the backhand to
provide simple visual feedback when a tag is detected during user
interactions. All the components were connected to an ARM Cortex
M4 microcontroller that is connected to a computer via a USB wire
or a Bluetooth low energy (BLE) module for data collection.

We conducted a 10-participant user study to understand the activity
recognition performance between the three IMU configurations in
both RFID-IMU and IMU-only conditions. The results show that
when the objects have similar physical form factors, the hybrid
RFID motion tracking significantly outperforms the IMU-only
tracking, especially in a larger-number set of objects. Furthermore,
the three IMU configurations (i.e., NailOnly, RingOnly, Nail+Ring)
achieved comparable action recognition performances: i.e., \(\geq 90\%\)
accuracy with 500 ms recognition time for the eight tasks in our
testing, although the NailOnly RFID-IMU configuration provided
the highest wearability (Figure 1a). Practical challenges such as
power management, adoption thresholds, physical object designs,
and technical limitations are also discussed to inform future real-
world deployment of such an activity recognition system based on
the proposed finger-worn RFID motion-tracking techniques.

This paper is organized as follows. First, we explain the design
and implementation of our system. Subsequently, we explain the
user study design and the results. Finally, we discuss the practical
issues and limitations for future research and draw the conclusion.

\section{SYSTEM DESIGN AND IMPLEMENTATION}

\subsection{Object Identification Through Touch}

When using an object by hand, the index fingertip is often the first
contact location. Therefore, at the location of the index fingernail,
we mounted a 15×20 mm\(^2\) 2 cm-range antenna to the the device and
then connected it to a Mifare 13.56MHz RFID reader. The 10-turn,
0.7 mm pitch inward spiral antenna, which is made by 0.29 mm
enameled wire, identifies an RFID tag under the fingertip like a
touch sensor.

The dimension of the antenna was determined through a formal
measurement on its tag reading distance. We fixed an RFID tag on
an elevated platform of a 3D printer, which is of 0.1 mm precision,
and fixed the customized RFID antenna above the platform at a
particular tilt angle using physical support. We measured the tag
reading distance by the following procedures: 1) Move the tag to
50 mm below the antenna. 2) Move the tag 0.1 mm closer to the
antenna until it identifies the tag. 3) Record the stopping height
and then return to step 1 until all measurements are made. Five tilt
angles (0°, 15°, 30°, 45°, and 60°) were tested, and each of them was
tested 50 times. Overall, a total of 250 maximum heights (5 [angles] \(\times\) 50 [measurements] = 250) were collected. Figure 3 shows the
results of the mean distance between 0° to 45°, which ranged from
18.8 mm to 22.4 mm (M = 19.86 mm; SD = 1.71 mm). The reader
did not read the tag at 60° at any distance, suggesting that ±45°
is a rough limit of tilt. The distance and operating angle suffice our
application.

\subsection{Activity Recognition Using Finger Posture}

To simulate the nail and ring motion-tracking devices, we mounted
a nine-axis MPU9250 IMU at both the fingernail and root locations.
Each IMU outputted its absolute orientations as a quaternion \(q\)
representing a 3D Euler orientation (\(\theta, \phi, \psi\)), which we considered
as reliable features for activity recognition. The motion processor
built in the IMU calculates the absolute orientation vector from the

\(^1\)http://www.proxinc.co.jp/
fusion of the accelerometer and gyroscope at 200 Hz and calibrates its yaw angle ($\psi$) using the magnetometer at 8 Hz.

For activity recognition, we collected the orientations of each of two IMUs for $N$ samples at a consistent rate of 20 Hz; thus, a gesture took 0.05s×$N$ second to be recognized. In each collection, we used the starting posture vector $P_1 = (\theta_1, \phi_1, \psi_1) = q_1$, the mean $\bar{P} = \sum_{i=1}^{N} P_i / N$ and the standard deviation $\rho$ as the statistic features of the trend and the intensity of movement. The feature vector $[\bar{P}, \bar{\theta}, \rho]$ of each collection was fed to an RBF-kernel support vector classifier (SVC) implemented with LibSVM [6], where $\gamma = 1.0$ and $c = 64$ were chosen through an 8×8 grid search, to build up a prediction model without overfitting.

3 USER STUDY
A user study was conducted to understand the activity recognition performance between the three IMU configurations in both RFID+IMU and IMU-only conditions.

3.1 Experiment Design

Participant and Apparatus. Ten users (7 males, 3 females) ranging from 22 to 25 years old ($M = 23.4; SD = 1.35$) were recruited in the study. All users were right-handed. To sensitize them to perform the gesture in a natural way, each user was requested to wear the sensing glove and perform given tasks using handy utensils in a simulated kitchen consisting of several tagged objects with similar material properties. Figure 4 shows the example uses of six handy kitchen utensils, which are a cup, plate, teapot, fork, knife, and spoon (Figure 4). The first three are graspable objects that requires a user to use it with a power grasp [29], and the latter three are pinchable objects that can be manipulated with a precision grasp [29].

Physical mediums such as water and clay were provided in some tasks to provide tactile feedback.

Tasks and Procedures. Each user was asked to wear the device on their dominant hand to perform eight tasks (Figure 4), including two static postures in different angles of holding an object and six dynamic gestures in the different wrist or wrist-and-finger movement. The tasks were touching the cup lid (T1), holding the cup (T2), lifting the plate (T3), pouring out water from the teapot (T4), using the fork (T5), using the knife (T6), spooning up water from the cup (T7), and stirring water using the spoon (T8). Two tasks were performed on the cup (T1-2) and the spoon (T7-8), whereas only one task was performed on the other four objects. These eight kitchen-relevant tasks were deliberately chosen to cover some common properties such as the type of postures and objects so that the subset of the tasks could be comparable as task groups in the later data analysis.

Each participant received an introduction before the formal session started. The participants were asked to demonstrate how they performed the task based on their individual preferences and then to attach an RFID tag to the position where he or she might first touch when performing the task. For each task of data collection in the formal session, a user followed on-screen prompt to perform each task. The user first moved a hand from the home position toward the objects, touched the RFID tag using the index finger, performed the task and placed the object back on the table surface, removed the finger from the object, and then moved the hand back to the home position to complete the task. The system recorded the gesture in a natural way, each user was requested to wear the sensing glove and perform given tasks using handy utensils in a simulated kitchen consisting of several tagged objects with similar material properties. Figure 4 shows the example uses of six handy kitchen utensils, which are a cup, plate, teapot, fork, knife, and spoon (Figure 4).

Figure 3: Understanding the tag reading range of the customized antenna: (a) experimental apparatus; (b) results.

Figure 4: Tasks: (T1) touching the cup lid; (T2) holding the cup; (T3) lifting the plate; (T4) pouring out water from the teapot; (T5) using the fork; (T6) using the knife; (T7) spooning up water from the cup; (T8) stirring water using the spoon.

Data Analysis and Variables. We collected a total of 4000 IMU data series (10 [participants]×8 [actions]×50 [repetitions] = 4000). Each data series consisted of $M$ samples, where $M > 20$ in every data collected, as the time for each data collection was longer than 1 s. In a post-hoc analysis, we formed 20 subseries of IMU data. Each subseries $S_N$, where $1 \leq N \leq 20$, selected the first $N$ samples from the $M$ to understand how the sample size affects the recognition accuracy regarding recognition time, which ranged from 50ms to 1000ms. The features $[\bar{P}, \bar{\theta}, \rho, q_1]$, where $P_i = (\theta_i, \phi_i, \psi_i)$ and $q_1$ were collected from each subseries and independently fed to the activity classifier (RBF-kernel SVC, $\gamma = 1.0$ and $c = 64$).

In addition to recognition time, IMU configuration and task group were also set as independent variables. Three IMU configurations were used in the data analysis:

- **NailOnly** (Figure 1a): Only the IMU mounted on the nail was used for activity recognition.
- **RingOnly** (Figure 1b): Only the IMU mounted on the finger root was used for activity recognition.
- **Nail+Ring** (Figure 1c): Both IMUs were used for activity recognition.

Four task groups were used in the data analysis:
Two-way ANOVA results also showed that recognition time (win-
accuracy
The
form factors. In the RFID+IMU case, the three
IMU configurations
more objects especially when some of them had similar physical
and the IMU(s) activity recognition performance decreased with
significantly outperformed IMU-only cases in all three
IMU configurations.
recognition time showed that for every
ables
and the eight actions performed on
the two objects of significantly different physical form factors.
shows that the IMU(s) still reasonably handled activity recognition
recognition time, no significance or borderline significance (p < .05)
dow size) had a significant (p < .05) effect on accuracy in all three
IMU configurations, showing that the activity recognition requires
several readings to obtain reasonable accuracy.

For the four actions in Cup&Spoo, the leave-one-user-out action
recognition accuracy among the three IMU configurations reached
M = 92.1% (SD = 6.5%) in RFID+IMU and M = 87.4% (SD = 7.0%)
in IMU-only, with a 500 ms recognition time, whereas the IMU-only
case only achieved M = 87.5% (SD = 8.9%) on Graspables, M = 76.8%
(SD = 9.4%) on Pinchables, and M = 75.5% (SD = 8.3%) on AllObjects
with the same 500 ms recognition time, as shown in Figure 5b-d.

The IMU-only results in Graspables and Pinchables can be consi-
dered as the generalization of the RFID+IMU results of Cup and
Spoon, respectively, as more actions can be trained within the same
object or transferred from other similar-form objects. Overall, all
the three IMU configurations had (significantly) higher accuracy on
Graspables than Pinchables with a < 850-ms recognition time. We
reviewed the records and found that the lower accuracy in Pinch-
ables was due to the individual ways of holding and using these
objects (spoon, fork, and knife), and the position of tag placement
was different among individuals. These individual differences
confused the cross-user activity recognizer. Therefore, we trained
another personalized recognizer based on the data collected from
each user, and the 10-fold cross-validation accuracy of all Gras-
Spoon, Pinchables, and AllObjects reached > 99% with a 500 ms
recognition time. This shows how well our approach can perform
if a user is willing to train the system.

In most cases, the three IMU configurations had no significance
in activity recognition among the four task groups, showing that the
NailOnly, RingOnly, and Nail+Ring configurations have comparable
performances on activity recognition; therefore, we recommend the
NailOnly RFID+IMU configuration (Figure 1a) due to its wear-
ing ease. Nonetheless, an exception is found that the RingOnly
significantly (p < .05) outperformed NailOnly in Pinchables in the
IMU-only case when the recognition time was ≥ 600 ms, which
shows the influence of sensor location can be further investigated.

3.2 Results
Figure 5 shows the results of leave-one-user-out activity recogni-
tion. Among all four task groups, student’s t-test showed that no
significant difference existed between any two of the three IMU
configurations when the recognition time was > 50 ms (N > 1).
Two-way ANOVA results also showed that recognition time (win-
dow size) had a significant (p < .05) effect on accuracy in all three
IMU configurations, showing that the activity recognition requires
several readings to obtain reasonable accuracy.

For the four actions in Cup&Spoon, the leave-one-user-out action
recognition accuracy among the three IMU configurations reached
M = 92.1% (SD = 6.5%) in RFID+IMU and M = 87.4% (SD = 7.0%)
in IMU-only, with a 500 ms recognition time (Figure 5a). For every
recognition time, no significance or borderline significance (p < .05)
was found between the RFID+IMU and IMU-only. This shows that the IMU(s)
still reasonably handled activity recognition on the two objects of significantly different physical form factors.

Nonetheless, for the four actions performed Graspables, Pinch-
ables and the eight actions performed on AllObjects, student’s t-test
showed that for every recognition time, the RFID+IMU (p < .05)
significantly outperformed IMU-only cases in all three IMU config-
trations. This shows that even the IMU data were well segmented
and the IMU(s) activity recognition performance decreased with
more objects especially when some of them had similar physical
form factors. In the RFID+IMU case, the three IMU configurations
reached a mean accuracy of M = 97.3% (SD = 5.0%) on Graspables,
M = 94.8% (SD = 6.5%) on Pinchables, and M = 96.0% (SD = 3.2%)
on AllObjects with 500 ms recognition time, whereas the IMU-only
case only achieved M = 87.5% (SD = 8.9%) on Graspables, M = 76.8%
(SD = 9.4%) on Pinchables, and M = 75.5% (SD = 8.3%) on AllObjects
with the same 500 ms recognition time, as shown in Figure 5b-d.

4 DISCUSSION
The current work still has practical limitations that should be con-
sidered in future real-world deployment.

Power Management. The RFID motion tracking is implemented
using power electronics, and such devices require maintenance (e.g.,
charging) in wireless uses. The maintenance efforts can be reduced
in ultra-high-frequency (UHF) RFID systems with the batteryless
wireless pair identification technique in RFIMatch [27] and the
batteryless RFID motion sensors (e.g., Farsens EVAL01-Kineo-RM)2
that are based on RF energy harvesting. Nonetheless, the availability
will also be constrained by the UHF RF signal coverage.

• Cup&Spoo (Figure 4 T1-2, T7-8): The four actions performed
on the two objects, which are of different physical form
factors, were recognized with or without RFID information.
• Graspables (Figure 4 T1-4): The four actions performed on
the graspable objects were recognized with or without RFID
information.
• Pinchables (Figure 4 T5-8): The four actions performed on
the pinchable objects were recognized with or without RFID
information.
• AllObjects (Figure 4 T1-8): All eight actions performed on
the objects were recognized with or without RFID information.

For clarity, we used RFID+IMU and IMU-only to annotate the
recognition with RFID and without RFID results, respectively.

The accuracy was measured as a dependent variable. The accu-
curacy was analyzed by a 10-fold leave-one-user-out cross-validation.
The accuracy was obtained from an SVM trained on nine users and
tested on the tenth user. This shows how well the approach works
when there is no training data from the current user.

Figure 5: Accuracy of leave-one-user-out activity recognition.

- For the four actions in Cup&Spoo, the leave-one-user-out action
recognition accuracy among the three IMU configurations reached
M = 92.1% (SD = 6.5%) in RFID+IMU and M = 87.4% (SD = 7.0%)
in IMU-only, with a 500 ms recognition time (Figure 5a). For every
recognition time, no significance or borderline significance (p < .05)
was found between the RFID+IMU and IMU-only. This shows that the IMU(s)
still reasonably handled activity recognition on the two objects of significantly different physical form factors.

Nonetheless, for the four actions performed Graspables, Pinch-
ables and the eight actions performed on AllObjects, student’s t-test
showed that for every recognition time, the RFID+IMU (p < .05)
significantly outperformed IMU-only cases in all three IMU config-
trations. This shows that even the IMU data were well segmented
and the IMU(s) activity recognition performance decreased with
more objects especially when some of them had similar physical
form factors. In the RFID+IMU case, the three IMU configurations
reached a mean accuracy of M = 97.3% (SD = 5.0%) on Graspables,
M = 94.8% (SD = 6.5%) on Pinchables, and M = 96.0% (SD = 3.2%)
on AllObjects with 500 ms recognition time, whereas the IMU-only
case only achieved M = 87.5% (SD = 8.9%) on Graspables, M = 76.8%
(SD = 9.4%) on Pinchables, and M = 75.5% (SD = 8.3%) on AllObjects
with the same 500 ms recognition time, as shown in Figure 5b-d.

The IMU-only results in Graspables and Pinchables can be consi-
dered as the generalization of the RFID+IMU results of Cup and
Spoon, respectively, as more actions can be trained within the same
object or transferred from other similar-form objects. Overall, all
the three IMU configurations had (significantly) higher accuracy on
Graspables than Pinchables with a < 850-ms recognition time. We
reviewed the records and found that the lower accuracy in Pinch-
ables was due to the individual ways of holding and using these
objects (spoon, fork, and knife), and the position of tag placement
was different among individuals. These individual differences
confused the cross-user activity recognizer. Therefore, we trained
another personalized recognizer based on the data collected from
each user, and the 10-fold cross-validation accuracy of all Gras-
pables, Pinchables, and AllObjects reached > 99% with a 500 ms
recognition time. This shows how well our approach can perform
if a user is willing to train the system.

In most cases, the three IMU configurations had no significance
in activity recognition among the four task groups, showing that the
NailOnly, RingOnly, and Nail+Ring configurations have comparable
performances on activity recognition; therefore, we recommend the
NailOnly RFID+IMU configuration (Figure 1a) due to its wear-
ing ease. Nonetheless, an exception is found that the RingOnly
significantly (p < .05) outperformed NailOnly in Pinchables in the
IMU-only case when the recognition time was ≥ 600 ms, which
shows the influence of sensor location can be further investigated.

4 DISCUSSION
The current work still has practical limitations that should be con-
sidered in future real-world deployment.

Power Management. The RFID motion tracking is implemented
using power electronics, and such devices require maintenance (e.g.,
charging) in wireless uses. The maintenance efforts can be reduced
in ultra-high-frequency (UHF) RFID systems with the batteryless
wireless pair identification technique in RFIMatch [27] and the
batteryless RFID motion sensors (e.g., Farsens EVAL01-Kineo-RM)2
that are based on RF energy harvesting. Nonetheless, the availability
will also be constrained by the UHF RF signal coverage.

2http://www.farsens.com/
Adaption Thresholds. Inherited from the previous wrist-worn RFID motion sensing systems, the proposed finger-worn system also requires the object to be tagged and the users to wear the device. To bypass these thresholds, application designers may first apply this system to the objects already embedded with RFID tags (e.g., keychains or bonus cards). Wearable device designers may corporate with fashion designers and fabric industries to make the devices easy-to-use and comfortably worn in everyday life.

Physical Object Designs. In our study, the RFID tags were manually placed at the position where users first touch. This could be impractical for some physical objects that allow many ways of interaction, such as symmetric objects or objects that do not necessarily involve the index’s fingertip in its use, such as chopsticks. Instead of adding more sensors to the device or adding more tags to the object to include more cases, we suggest solving the limitation through thoughtful form design that can clearly indicate the entry point of interaction through its affordance such as the push button on a spray can (Figure 1) or the physical landmark on a spoon (Figure 2d).

Technical Limitations. Regarding object identification, the presented system could be insensitive to a tag that is attached to a metal object. False negatives may happen when the fingertip is very close to the tag; the tag can be disambiguated by adding a touch sensor to the device, repurposing the IMUs for bio-acoustic touch sensing [25] and/or antenna optimization. Regarding activity recognition, although the proposed gesture classifier was robust enough to handle the slight drifting of the IMU’s yaw angle, the effects were unclear in an extended period of time. The activity segmentation can be further improved by using the IMU features, such as movement intensity. Macro-scale features, such as body postures and orientation can be featured by additional hardware and advanced signal processing, which will be considered in a future work.

5 CONCLUSION
We have evaluated finger-worn RFID motion-tracking devices for activity recognition on tagged objects. The results of a ten-participant user study revealed the activity recognition performance and several design implications. Based on the results, we recommend the NailOnly RFID+IMU configuration, which provides both wearing ease and accuracy of activity recognition. We also discuss the practical limitations of real-world deployment and the potential solutions for future endeavors in ubiquitous computing applications [37].

ACKNOWLEDGMENTS
This research was supported in part by the Ministry of Science and Technology of Taiwan (MOST 108-2633-E-002-001), National Taiwan University (NTU-108104039), and Intel Corporation.

REFERENCES


