Pers Ubiquit Comput (2012) 16:225–234
DOI 10.1007/s00779-011-0395-z

ORIGINAL ARTICLE

Gesture recognition using RFID technology

Parvin Asadzadeh · Lars Kulik · Egemen Tanin

Received: 15 October 2010 / Accepted: 4 April 2011 / Published online: 11 June 2011

Abstract We propose a gesture recognition technique based on RFID: cheap and unintrusive passive RFID tags can be easily attached to or interweaved into user clothes, which are then read by RFID antennas. These readings can be used to recognize hand gestures, which enable interaction with applications in an RFID-enabled environment. For instance, it allows people to interact with large displays in public collaboration spaces without the need to carry a dedicated device. We propose the use of multiple hypothesis tracking and the use of subtag count information to track the motion patterns of passive RFID tags. To the best of our knowledge, this work is the first on motion pattern tracking using passive RFID tags. Despite the reading uncertainties inherent in passive RFID technology, our experiments show that the proposed gesture recognition technique has an accuracy of up to 93%.

Keywords RFID · Gesture recognition · Space partitioning

1 Introduction

Ubiquitous computing enables omnipresent and intuitive interaction of people with their surrounding environment. The use of hand gestures for natural human computer interaction attracts great interest. Hand gesture recognition systems allow us to interact with and control computing devices and applications in intelligent environments [10].

Hand gesture recognition techniques can be mainly divided into vision-based and device-based techniques. Pavlovic et al. [17] give a comprehensive review of vision-based gesture recognition techniques. The challenges in such techniques are cluttered backgrounds and varying illuminations, especially in public places. Moreover, recording user movements using video is resource intensive.

Device-based hand gesture recognition techniques use glove-based equipments or accelerometer-enabled devices such as the Wiimote to measure user movements. Glove-based techniques [13, 19] are relatively intrusive for users. A less intrusive glove-based gesture recognition technique is proposed by Rahman et al. [24]: an infrared (IR) camera tracks an infrared emitter attached to a user's hand gloves and produces a sequence of motion points to recognize the intended hand gesture. Similarly, in the study by [26], on both user hands, thumb and index finger are equipped with retro-reflective markers that are tracked by a Wiimote. A finger pairing and pinch recognition method is then used to discriminate the hands and to initiate actions. The line-of-sight alignment requirement is the main drawback of systems based on IR technology. Accelerometer-enabled devices, on the other hand, while accurate and less intrusive, are not readily available in public places.

In this paper, we propose the use of passive radio-frequency identification (RFID) to facilitate natural user interaction. RFID is an effective automatic identification technology that allows for easy proximity sensing of tagged objects. Objects tagged with small inexpensive and unintrusive passive RFID tags can be sensed from a few centimeters up to several meters. Passive RFID tags

P. Asadzadeh (✉) · L. Kulik · E. Tanin
National ICT Australia (NICTA), Department of Computer Science and Software Engineering, University of Melbourne, Melbourne, Australia
e-mail: parvina@csse.unimelb.edu.au

L. Kulik
e-mail: lars@csse.unimelb.edu.au

E. Tanin
e-mail: egemen@csse.unimelb.edu.au

operate without a battery, and it is possible to tag large collections of objects with multiple tags. All RFID tags contain unique identification numbers along with other data to easily identify tagged objects.

Typically, RFID tags are used for supply chain management and automatic identification of objects [2, 4]. They also allow us to augment physical objects and the environment with digital information [25] and to monitor indoor human activities [8] or to detect people's interactions with RFID-tagged objects [11]. RFID tags have also been deployed as landmarks on the floor or within the environment to support navigation of mobile robots equipped with RFID readers [16, 20]. Moreover, many location-sensing techniques have been developed that use RFID technology, which are discussed in Sect. 2.

We propose the use of multiple hypothesis tracking to track the motion patterns of RFID tags and recognize a gesture. We use a combined tag consisting of multiple subtags to increase the readout reliability of the RFID readers. We also incorporate the subtag count information to improve the accuracy of our technique. Since there is no line-of-sight requirement, the user can easily draw the intended gesture. New gestures can also be easily added to the system as no training is required. Furthermore, multiple users can use the system without increasing the system complexity.

Because of the various error sources in passive RFID systems, reliable operation as the tag moves in the environment is inherently difficult and presents a significant challenge [15]. Particularly in a multi-tag and multi-reader configuration false-negative readings occur [12], i.e., a present tag is not detected. Researchers suggested different ways to deal with the resulting uncertainty in passive RFID systems, such as the use of the percentage of positive tag reads [28] or simple filtering mechanisms [7], which are only applicable if the object is stationary. To the best of our knowledge, we are the first to study the motion pattern tracking of passive RFID tags in a multi-reader, multi-tag environment.

We have implemented a prototype of our system and conducted a detailed performance evaluation. Our results show that the system can recognize hand gestures with up to 93% accuracy, without requiring any learning or training. We also show our initial findings for independent gesture recognition of two users.

Our system provides support for a variety of applications. Users could interact with large displays in public collaboration spaces without the need to carry a device, use hand gestures to interact with devices like turning on a projector in a lecture theater, moving an advertisement page on a public display, or control gaming devices. Tagged gloves, for example, can be designed to be worn by a user for the purpose of recognizing her interactions with objects in her physical surroundings.

In the remainder of this paper, we first discuss different location-sensing techniques using RFID technology (Sect. 2). Section 2 also describes space partitioning, which is the positioning technique that we use to track RFID tag motions. Section 3 details our setup. Our proposed tracking technique is described in Sects. 3.1 and 3.2. We then show the experimental results and evaluate the performance in Sect. 4. We conclude in Sect. 5.

2 Related work

There are a number of location-sensing techniques based on *active* RFID technology that measures the received signal strength (RSS) to estimate a tag's location [5]. RSS is usually severely affected by the propagation environment and the tagged object properties. Moreover, it cannot be universally approximated with a distance-dependent path loss model. Therefore, the use of RSS in tag localization is more accurate for active tags since they carry a power source and hence have more stable performance within crowded environments.

To localize *passive* RFID tags, some researchers use angulation technique to estimate the direction of arrival of a tag signal by measuring phases [23, 29] or the relative strength ratio [21] of the received signal at several receiving antennas. Nikitin et al. [23] also estimate distance to the tag by measuring tag phase at different frequencies. Furthermore, Wilson et al. [28] use the percentage of positive tag reads as an indication of distance and Chawla et al. [9] infer a tag's position based on the relative power level that is necessary for a reader to detect the tag.

A number of passive RFID-based location-sensing systems use only the presence information from RFID readers to localize a tag. RFID readers can only sense the presence of a tag within their detection fields, providing proximity information of the tag, but they cannot directly determine the tag's distance to the reader. However, one positive detection of a tag greatly reduces its possible locations, since it indicates that the tag is in the reader's detection field. However, the bigger the detection field of the reader is, the more uncertainty is in the localization of the tag.

The tag readings from a mobile RFID reader from different vantage points can be combined to reduce the uncertainty in a tag's location and to estimate the whereabouts of the tag more precisely. This technique has been employed in design of Ferret [18] and Sherlock [22] to estimate the position of tagged objects in a room swept by readers.

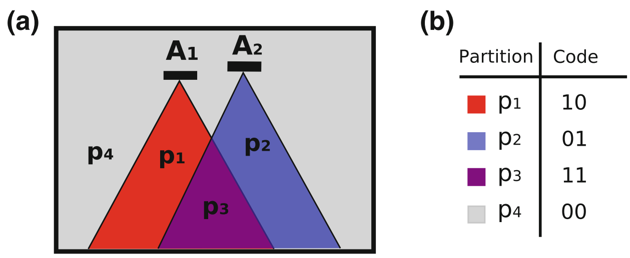


Fig. 1 Space partitioning using two antennas

The area of uncertainty can also be reduced if the outputs of several readers are combined and analyzed. Multiple stationary RFID readers/antennas attached to readers can be used to detect the presence or absence of a tag. Combining the detection results of all antennas provides a more accurate estimation of the location of a tag at a given time: the monitored area is divided into multiple partitions so that each partition is in detection fields of a particular set of antennas. We refer to the technique of using stationary RFID antennas to partition a space as *space partitioning*.

Figure 1 shows the partitioning of a rectangular space using two directional antennas A_1 and A_2. The monitored area is divided into four partitions p_1 to p_4 by the two antennas. These partitions are distinguishable in the sense that each partition is covered by detection fields of a different set of antennas. A tag detected by both antennas, for example, is estimated to be in partition p_3; whereas, a tag detected by A_1 only, is estimated to be in partition p_1.

We differentiate the created partitions by their assigned codes. In an arrangement of n antennas, an n-bit binary code (BC) of partition p_i is $BC[p_i] = bc_1, bc_2, \ldots, bc_n$. The kth bit is set if partition p_i is within the detection field of the kth antenna. The codes of partitions p_1 to p_4 in Fig. 1a are shown in Fig. 1b.

A basic implementation of this technique is presented in the study by [14], in which a table surface is equipped with an array of omni-directional RFID antennas and is hence divided into many distinguishable partitions. When a multi-tagged object is placed on the equipped surface, the partition each individual tag is within is determined based on the reading results gathered from readers. Bouet and Pujolle [6] have also used space partitioning to localize tagged objects in a room equipped with equally distanced arranged RFID readers on its floor and ceiling. A tag's position is estimated as the center of gravity of the different volumes formed by the intersection of the reader detection fields.

Existing passive RFID-based location-sensing techniques mainly focus on localization of stationary tagged objects [6, 9, 14, 21, 23, 28, 29]. A few of the proposed methods also try to localize and track moving objects [23, 28, 29] under particular conditions. However, none is capable of accurate online tracking of arbitrarily moving tags. In the study by [29], it is assumed that the tag is moving along a straight line, perpendicular to the reader at a fixed distance apart and at a known speed. Also, in the study by [23], the tagged object is moving along one of the several straight lines, perpendicular to the reader and it stops periodically to allow the reader to take measurements of the RSS and phase of the received tag signal on all frequency channels. Wilson et al. [28] try to estimate the speed of a mobile reader moving perpendicular to a set of tags at a known distance apart.

We propose the use of multiple hypothesis tracking along with subtag count information to track a tag's motion and recognize its gesture. Next section describes our experimental setup to simplify the description of our algorithm, described in Sects. 3.1 and 3.2.

3 Our RFID-based gesture recognition system

We have built an experimental system using the SkyeModule M9 UHF reader from SkyeTek [1] (Fig. 2a) and their linear broadband UHF antennas (Fig. 2b). We chose UHF tags ISO 18000-6C since they are small and compatible with the RFID readers we use (Fig. 2c).

The read range of a M9 SkyeTek reader interrogating ISO 18000-6C passive tags is around two and a half meters, when the transmitter's power output is set to the maximum of 27 dBm. A sample detection field of a UHF SkyeTek antenna interrogating ISO 18000-6C passive tags in a 140 cm by 130 cm area on a desk is shown in Fig. 3a. The area is divided into equally sized square cells, with the side equal to the width of the used tags, which is 10 cm. To measure the field, the tag is placed in parallel to the antenna in each cell and the transmitter's power output is set to 21 dBm.

The tag-antenna orientation plays an important role in determining whether the tag receives enough energy to be detected by the antenna. Particularly, when a tag's antenna is perpendicular to the reader's antenna, it is not detectable in a large area within the antenna's detection field (Fig. 3b), comparing with when the tag is in parallel to the antenna (Fig. 3a). To increase the reading reliability of a tag when it is close enough to the antenna, we use a combined tag instead of a single tag.

As shown in Fig. 2d, our combined tag consists of four individual colocated tags, where each tag is rotated 45 degrees to its neighbor tag. All single tags, or *subtags*, in a combined tag have different identifiers but they are combined to represent one "super" tag: if one of the subtags is oriented such that it is not detectable by the antenna, another subtag with a different orientation with respect to

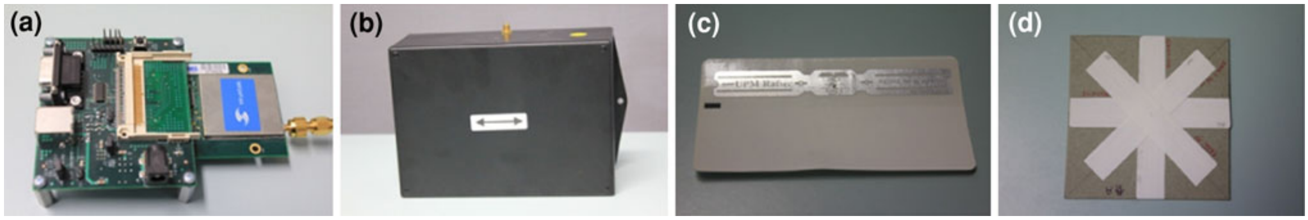


Fig. 2 Our **a** reader **b** antenna **c** single tag **d** combined tag

Fig. 3 Sample antenna's field, when the tag is **a** parallel **b** perpendicular to the antenna

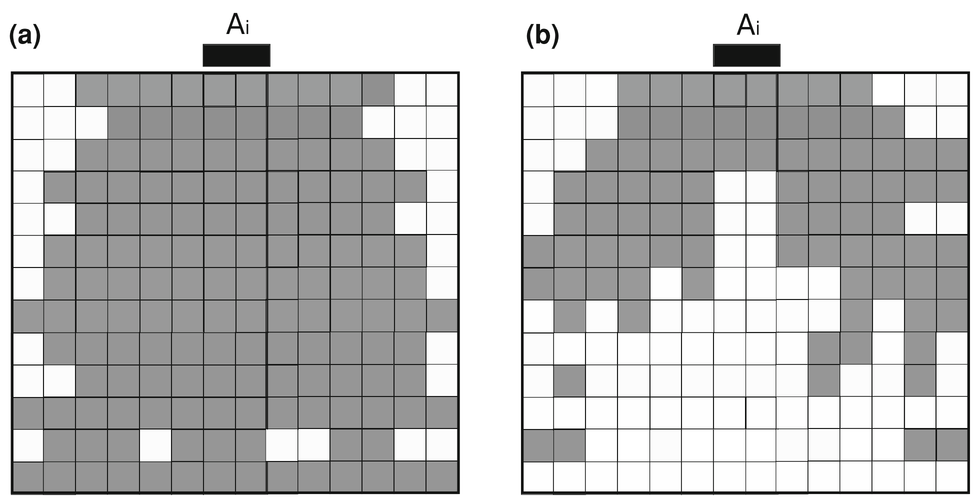
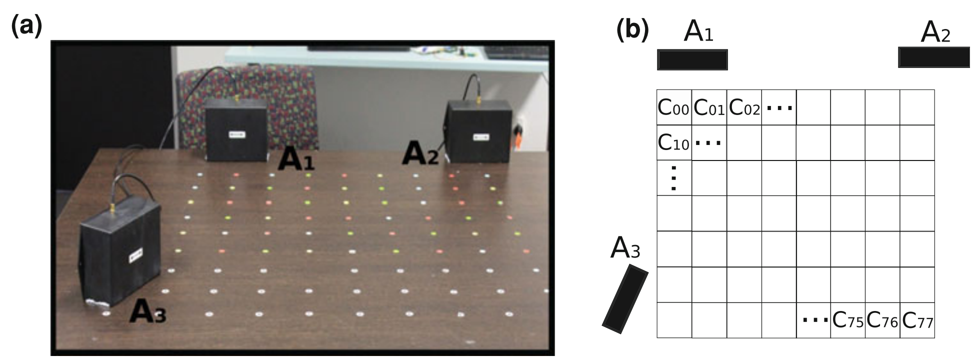


Fig. 4 Our monitored area



the antenna might be detectable. Our experiments showed that a combined tag of four subtags is optimal to provide a consistent antenna field for an arbitrary orientation while keeping the tag's reading time small. In the remainder of this paper, we simply use tag to refer to our combined supertag.

The tests were conducted in a busy student laboratory. We monitored a 80cm by 80cm square area on a desk in the laboratory (Fig. 4a). Similar to the fields shown in Fig. 3, the square monitored area is divided into 64 equally sized square cells $C_{00}, \ldots, C_{77}$, as shown in Fig. 4b, with the side equal to the width of the used tag, which is 10 cm.

The reader works in inventory mode, which runs an anti-collision protocol to read many tags simultaneously. It is connected to three antennas (A_1–A_3) via a multiplexer, which are placed just outside the monitored area (Fig. 4a). Time slicing is used to avoid an interference between the antennas. The three antennas are sequentially energized, which in turn return the tag identifiers in their detection fields.

The reader is connected to a laptop via a USB connection. Using the provided API, we developed a C program, which periodically returns an inventory of tags within the range of each antenna. RFID readings are then passed to

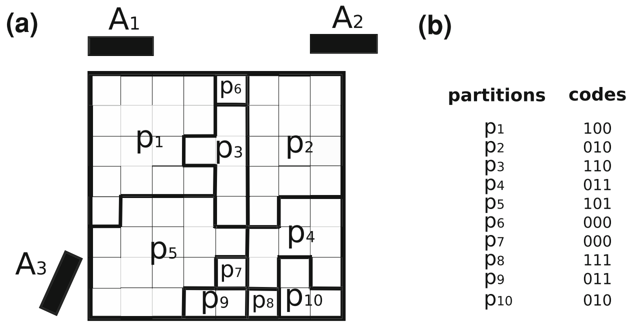


Fig. 5 **a** The partitioning of our monitored area **b** Partitions' codes

A_i

0	0	1	1	4	4	4	4	4	4	0	0	0	0
0	0	1	1	3	4	4	4	4	4	3	1	0	0
0	0	1	2	4	4	4	4	4	3	3	1	0	0
0	0	2	2	3	4	4	4	4	3	3	1	0	0
1	1	2	3	3	3	3	3	3	2	2	1	1	0
1	1	2	2	2	2	3	3	2	2	2	2	1	0
2	2	1	1	2	2	3	3	2	2	2	2	2	1
1	1	2	3	3	3	3	3	3	3	2	2	2	1
0	1	2	1	2	2	2	2	2	2	2	2	2	0
0	1	1	2	2	2	1	2	3	2	2	1	1	0
2	0	2	2	2	2	2	2	2	2	2	2	0	1
0	1	1	0	0	0	1	1	0	0	0	2	1	0
0	0	2	2	1	1	1	2	1	1	1	1	0	0

Fig. 6 Space partitioning based on subtag counts in the field of a single antenna

the gesture recognizer, which takes a multiple hypothesis approach to recognize the gestures as described in Sect. 3.2.

3.1 Space partitioning

In Sect. 2, we explained how a monitored area can be divided into different partitions by a number of antennas with overlapping fields. Figure 5a shows the partitioning of our monitored area into ten partitions, at a given time. The binary code assigned to each partition is shown in Fig. 5b. Partition p_1, for example, is assigned a binary code of 100, since it is only in the detection field of Antenna A_1.

The position and orientation of the antennas are determined based on the size of the monitored area and the shape of the antenna fields. They are placed such that they overlap as much as possible while minimizing the upper-bound on spatial resolution of the monitored area, such that the largest diameter of partitions is less than the minimum size of the intended gestures.

We propose the use of *subtag counts*, the number of detected subtags, as another measure to partition a space. Our experiments show that the subtag counts are an indication of the relative position of a supertag to the antenna, regardless of the tag-antenna orientation. Figure 6, for example, shows the subtag counts (of a supertag composed of four subtags) detectable by an antenna, in the same area as Fig. 3.

At any given time, in a system with n antennas, an n-bit *binary code*, $BC[C_k]$, is assigned to each cell C_k, the jth digit of which is set if any subtag of a supertag in cell C_k is detectable by antenna A_j.

To further improve the accuracy of our tracking method, we also use subtag count information. A *count code*, $CC[C_k]$, is also assigned to each cell C_k, at any given time. The jth digit of $CC[C_k]$ is set to m if m subtags of a supertag in cell C_k are detectable by antenna A_j.

Because of false-negative and false-positive readings, there are always unavoidable uncertainties about the presence of RFID tags. Tags outside the defined range of an antenna might be read at times, and similarly, there is no guarantee that all tags within the range of an antenna are read at all times. This is the biggest challenge in designing a passive RFID-based system, especially when antennas are in close proximity.

To cope with uncertainties in RFID readings, instead of assigning a fixed code to each cell of the monitored area, we consider a set of possible codes with their corresponding weights. The weight assigned to each code shows how probable the occurrence of that particular code is. Consequently, each cell C_k is assigned two sequences of pairs of possible codes, *BCSEQ* and *CCSEQ*, with their associated weights as shown below, where n is the number of antennas and t is the number of subtags in a supertag:

$$BCSEQ : \{(BC_1, W(C_k, BC_1)), \ldots, (BC_p, W(C_k, BC_p))\}; p_{max} = 2^n$$
$$CCSEQ : \{(CC_1, W(C_k, CC_1)), \ldots, (CC_q, W(C_k, CC_q))\}; q_{max} = (t+1)^n$$

The code weights are the number of times that a specific code is read by placing the tag in each cell and being read multiple times. To minimize the error in the tracking accuracy, the codes are measured multiple times at different dates and times of the day and with different tag-antenna orientations.

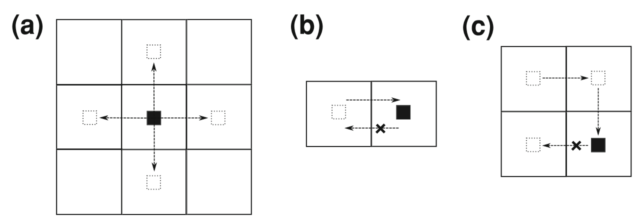


Fig. 7 **a** Possible local moves (**b**–**c**) Illegal local moves

On the other hand, at any given time, a total weight, $W(C_i, BC, CC)$ is assigned to each cell, which is a weighted sum of BC and CC as:

$$W(C_i, BC, CC) = W(C_i, BC) + \alpha \times W(C_i, CC); \alpha < 1$$

We give less weight to $W(C_i, CC)$, since it is more vulnerable to environmental changes. In other words, the movement pattern of the tag, the position of the user *etc*. have more impact on the number of detectable subtags; whereas BC shows if *any* of the subtags, no matter how many, are detectable and hence is less vulnerable to the environmental changes.

3.2 RFID-based gesture recognition

Whenever a user draws a gesture by moving a tag on the monitored area, a sequence of readings from RFID antennas is generated. The RFID readings are used to estimate the track of a tag T, as a sequence of traversed cells:

$$Track(T) = X_1 \cdot X_2 \ldots X_{k-1} \cdot X_k, X_i \in \{C_{00}, C_{01}, \ldots\},$$

which is later mapped to a gesture.

The basic approach used to track a tag is to generate a set of hypotheses to account for all possible tracks of the tag based on the received RFID readings. The key principle of this approach is that the track update decisions are deferred until more RFID readings are received. In our method, we also incorporate the tag movement pattern information to overcome the uncertainties in RFID readings and improve the results.

We assume that the tag is either moving horizontally or vertically, as shown in Fig. 7a. The antennas' reading speed is high enough to ensure that the tag does not move more than one cell away between any two consecutive readings. Furthermore, Fig. 7b, c show two illegal local movements, both assuming that the tag is always moving forward.

Multiple hypothesis tracking is generally used for solving the data association problem in multiple target tracking (MTT) systems [3]. However, in tracking multiple RFID tags, data association is not an issue, since the tag IDs are transferred to the readers at detection time, and hence, the readings can be associated with the tags without any ambiguity. Thus, the problem of tracking multiple tags can be decomposed into a set of smaller problems of tracking individual tags. Therefore, tracking of several tags can be done independently and hence simultaneously.

A flow diagram of our gesture recognition algorithm is shown in Fig. 8. On the receipt of new data (the kth set of antenna readings), BC_k and CC_k codes are generated in CODEGEN (code generator), as explained in Sect. 3.1. The INI (initiator) process creates a hypothesis tree once BC_0 is received, which includes as its children the cells C_i with $W(C_i, BC_0) > 0$.

On the receipt of new codes BC_k and CC_k at time k, HPGEN (hypotheses generator) expands each hypothesis into a set of new hypotheses by considering all possible new locations of the tag, which are determined by considering the possible movements of the tag. Each branch of the hypothesis tree represents a possible track of the tag, and nodes of the tree are the cells the tag has traversed. A hypothesis list is also created that contains all possible current locations of the tag along with the corresponding track weight, SW_i, which is the sum of the weights of all cells contained in that track. The track weight is later used to assess track validation as well as track selection.

Since the complete hypothesis tree for every gesture grows exponentially as more readings are processed, there is a clear potential explosion in the number of possible tracks (hypotheses) that our system can generate. Therefore, we use two pruning techniques to keep this potential growth in check, *track-based* and *weight-based* pruning methods, which both eliminate unlikely hypotheses. In the track-based pruning method, the number of hypotheses is reduced by eliminating the tracks that are not within the set of validated paths, following the rules illustrated in Fig. 7a–c. Moreover, it can be assumed that the tag does not remain in the same cell for long—the speed of the tag is greater than a threshold. In weight-based pruning, on the other hand, the tracks are evaluated based on their weight and tracks that are unlikely to reach the minimum weight requirement are removed from the hypothesis tree. This allows us to use a threshold to reject all tracks rather than picking the nearest matched track.

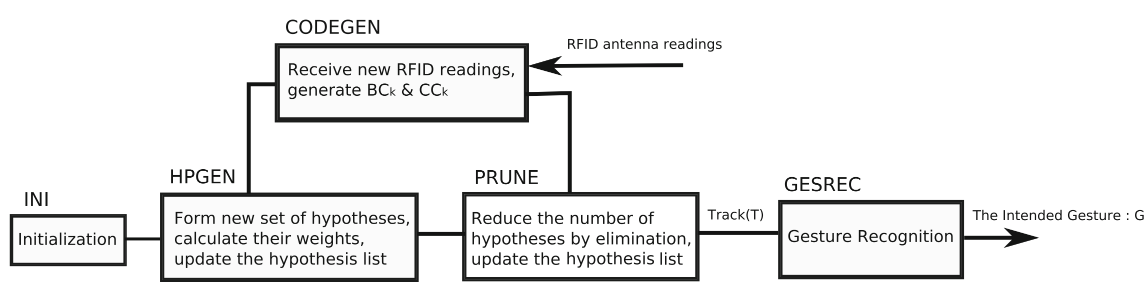
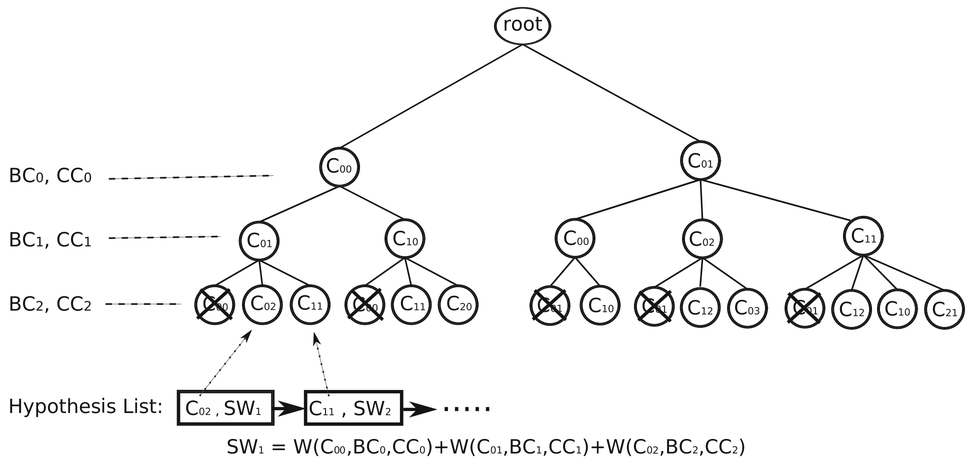


Fig. 8 Our gesture recognizer architecture

Fig. 9 The hypothesis tree



In a real-time variant of our system, the size of the hypothesis tree is significantly reduced by simply pruning the tracks with consecutive occurrences of the same cells. Details on run time of the system are included in Sect. 4. Figure 9 shows a sample hypothesis tree after three sets of readings are received for our real-time gesture recognition system. The crossed out nodes are the invalid new positions because they are a return to an immediate cell, as shown in Fig. 7b.

The tree expansion process continues until the end of the gesture, which is recognized by a long sequence of the same readings, indicating that the tag is stationary. After the last validation phase, the most likely tracks are the ones with $SW > \beta \times SW_{max}$ ($\beta = 0.95$, in our tests). For each likely track, the GESREC (gesture recognizer) process then finds the gesture that best matches that track, using Algorithm 1. The output gesture is the one with maximum probability of occurrence. If no valid gesture matches any of the likely tracks, recognition fails.

We use a strict matching algorithm, which first converts the sequence of traversed cells to a sequence of directional moves. Sequences of the same movements are recognized, which can only contain single movements of other directions. The sequences of different movements are then looked up in our defined gesture dictionary. Figures 10a–d, for example, show the matching process of four tracks to gestures, while Fig. 10b and c show two sample cases where no valid gesture matches the results.

4 Gesture recognition experiments

In this section, we will report and evaluate the results of our gesture recognition experiment. The experiments show that our system can recognize hand gestures with up to 93% accuracy, without requiring any learning or training. We will also present our initial findings for independent gesture recognition of two users (Sect. 4.2). We will first discuss the factors that influence the accuracy of this technique.

The accuracy of the proposed gesture recognition technique depends on the size of the partitions relative to the gestures. For best performance, the created partitions must be small enough to ensure that every *gesture element*—same direction movements—crosses more than one partition. Otherwise, without further information, inferring the direction of a tag's movement is not possible.

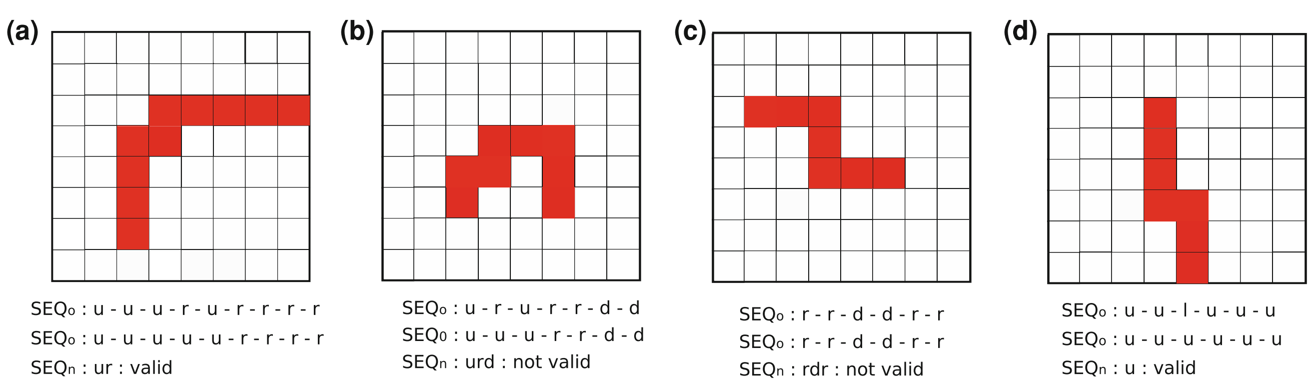


Fig. 10 Matcher

Algorithm 1 The matcher algorithm

Data: $Track(T) = \{X_1, X_2, \ldots, X_k\}$ where $X_i \in \{C_{00}, C_{01}, \ldots\}$
Result: The Intended Gesture : $G \in$ { possible gestures }
Convert *Track*(*T*) to a sequence of movements
$SEQ_o = \{m_1, m_2, \ldots, m_{k-1}\}$ where $m_i \in \{u, d, l, r\}$.
/** u:up ; d:down ; r:right ; l:left */.
foreach m_i in the SEQ_o **do**
if $m_i \neq m_{i-1}$ and $m_{i-1} = m_{i+1}$ **then**
$m_i \leftarrow m_{i-1}$
end
end
foreach *Consecutive occurrences of character 'c' in* SEQ_o **do**
/** More than one consecutive occurrence of 'c' */
Add a single 'c' to the SEQ_n.
end
if SEQ_n *is in the dictionary* **then** output G; **else** fail ;

The number of antennas, the shape of their detection fields, their relative positions and orientations determine the resolution of the space partitioning and, as a result, the accuracy of the gesture recognition technique. Therefore, the main challenge is to identify the required number of antennas, their placement and orientation such that the monitored area is divided into partitions with desired sizes. In particular, it is desirable to minimize the upper-bound on the spatial resolution of the monitored area, such that the largest diameter of partitions is less than the minimum size of the gesture elements.

In general, the accuracy can be increased by dividing the area into smaller partitions, which can be achieved by employing more antennas. As each antenna requires its own time slot to interrogate the tags, employing a larger number of antennas requires faster hardware (this applies to antennas as well as tags) and a faster singulation protocol. However, there is always a trade-off between cost and accuracy, as more antennas introduce higher costs and require more space for an actual deployment.

4.1 Single gestures test results

We have evaluated the performance of our gesture recognition technique by testing an alphabet of gestures. Any set of gestures can be defined for a service that are natural in the given context. Figure 11 shows our example gestures. Gesture $G1$, for example, is performed when the user moves the tag from left to right on a line, anywhere within the monitored area. Gestures $G5$ to $G12$ consist of two elements each.

Users perform gestures by moving the supertag on the surface of the desk. Since the antennas are working in different time slots, one reading cycle equals n times the interrogation time of one antenna. In our experiments, the reading cycle time is three times the interrogation time of an antenna, which equals 3×160 ms. Gestures have different sizes, and in average, they cross nine cells and are performed in an average of 4.5 s with an average speed of 0.2 m/s. They are also performed in different parts of the monitored area.

We collected quantitative data to determine the percentage of correctly recognized gestures. A total of 120 samples was collected: ten samples of each gesture. Table 1 shows the number of correctly recognized samples (CR) for each gesture in our alphabet.

A recognition error occurs when the system is unable to match the track to a unique gesture, or it outputs a gesture other than the drawn one. In the latter case, the output gesture (OG) is shown on the third row of Table 1. The average rate of correctly recognized gestures of all 120 (ten of each gesture shown in Fig. 11) gestures was 93%.

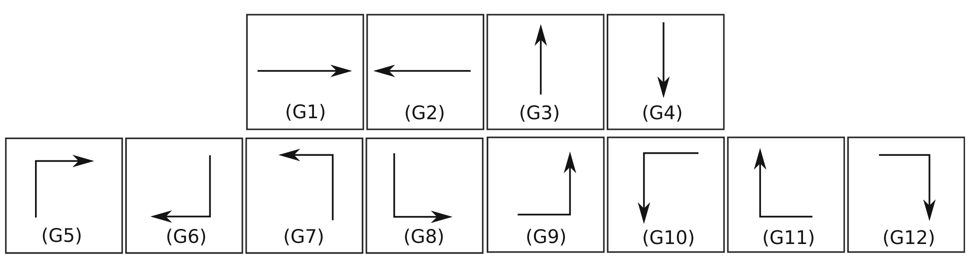


Fig. 11 Example gestures

Table 1 Recognition results

Gesture	G1	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12
CR	10	9	10	9	10	9	10	9	10	9	7	10
OG	–	–	–	G10	–	–	–	–	–	–	G7	–

4.2 Double gestures test results

Like all modern RFID systems, our used M9 SkyeTek readers employ an anti-collision and singulation technique that allows tags to be accessed in an ordered way. The read time of the singulation protocol is approximately linearly correlated with the number of tags. In practice, larger number of tags require longer read times, which means that in order to achieve a comparable recognition rate when more tags are in the field, the gesture speed must be decreased or faster singulation protocols and technologies have to be applied.

Our M9 SkyTek reader is relatively slow with a tag interrogation rate of 25 tags per second; in addition, it does not allow direct programmatic access to the singulation protocol for faster interrogation times. In fact, the passive tag read rate can be upto 500 tags per second [27], which is 20 times faster than our M9 SkyeTek reader. This means that a commercial system is likely to be able to recognize much faster gesture movements, by employing faster technologies and therefore acquiring a comparable recognition rate even when more supertags are employed.

To demonstrate that we can get comparable results when more gestures are to be recognized, we have used two supertags to perform two gestures simultaneously.

Using two supertags, the interrogation time of an antenna is increased to 270 ms, which makes the reading cycle time 3×270 ms. Gestures have different sizes and in average, they cross seven cells and are performed in an average of 7 s with an average speed of 0.12 m/s. They are also performed in different parts of the monitored area and with different relative distance to each other.

An alphabet of four double gestures, as shown in Fig. 12, is tested. We collected quantitative data to determine the percentage of correctly recognized double gestures. A total of 40 double samples were collected: ten samples of each double gesture.

Table 2 shows the number of correctly recognized samples (CR) for gestures of both users in the alphabet shown in Fig. 12. If the system outputs a gesture other than the drawn one, the output gesture (OG) is shown on the third and fifth row of Table 2 for the first and second user, respectively. The average rate of correctly recognized gestures of all 40 double gestures (ten of each double gesture shown in Fig. 12) was 85%.

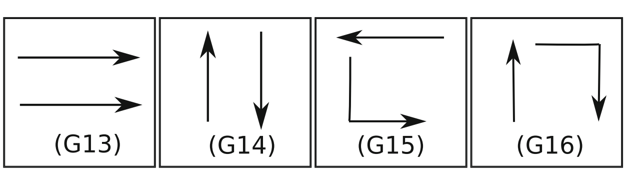


Fig. 12 Example double gestures

Table 2 Recognition results

Gesture	G13	G14	G15	G16
CR-User1	9	7	10	10
OG-User1	–	G1	–	–
CR-User2	8	7	8	9
OG-User2	G4	–	G1	G9

Our proposed system is capable of real-time gesture recognition. In fact, we have also evaluated a real-time variant of our system, which runs around 100 times faster than the original system with a very similar accuracy (still more than 90% recognition rate for our gesture alphabet). In this near real-time gesture recognition system, the size of the hypothesis tree is significantly reduced by simply pruning the tracks with consecutive occurrences of the same cells. The system can recognize the gestures of 6–11 cells long, in an average computation time of 13–430 ms. The original system is a more flexible algorithm that can recognize more general gestures but requires more time due to its larger search space. It can recognize gestures of 6–11 cells long in an average computation time of 50–45000 ms. It is important to mention that the gesture recognizer runs on a 2.16-GHz Intel Core 2 Duo with 2 GB of RAM laptop.

Although the proposed technique has been evaluated in a relatively small setup, our approach scales in several dimensions. Firstly, it can be extended to larger areas by using more powerful antennas or simply by employing more antennas. Secondly, it can be extended to recognize independent gestures by different users. Multiple users, for example, could interact with a large display in a public area by performing the gestures at different positions. Lastly, the partitioning resolution can be increased by using more antennas to make recognition of finer resolution gestures possible—at the cost of increased read times.

It is important to mention that our technique requires no learning or training. As RFID readers improve in range, cost and reliability, this technique will become more robust and viable.

5 Conclusions and future work

In this paper, we presented the design and evaluation of a hand gesture recognition technique based on RFID, which can be used to develop intuitive interfaces for ubiquitous applications. We proposed the use of multiple hypothesis tracking and the use of subtag count information to track the motion patterns of passive RFID tags and hence the hand gestures. Our gesture recognition technique was able to recognize gestures with up to 93%. We also demonstrated our initial findings to recognize independent gestures by different users. Due to the low cost of passive RFID tags and the fact that they operate without a battery, we believe that passive RFID technology is unintrusive and easy to use and, therefore, forms a promising solution for gesture recognition.

Acknowledgments This work was funded by National ICT Australia (NICTA). NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program.

References

1. Skyetek Company. http://www.skyetek.com
2. Angeles R (2005) RFID technologies: supply-chain applications and implementation issues. Inform Syst Manag 22:51–65
3. Blackman SS (2004) Multiple hypothesis tracking for multiple target tracking. IEEE Aerosp Electron Syst Mag 19:5–18
4. Borriello G, Brunette W, Hall M, Hartung C, Tangney C (2004) Reminding about tagged objects using passive RFIDs. In: UbiComp, pp 36–53
5. Bouet M, dos Santos AL (2008) RFID Tags: positioning principles and localization techniques. Wirel Days 1–5
6. Bouet M, Pujolle G (2008) 3-D localization schemes of RFID tags with static and mobile readers. In: Networking, pp 112–123
7. Brusey J, Harrison M, Floerkemeier C, Fletcher M (2003) Reasoning about uncertainty in location identification with RFID, In: Workshop on reasoning with uncertainty in robotics at IJCAI
8. Buettner M, Prasad R, Philipose M, Wetherall D (2009) Recognizing daily activities with RFID-based sensors. In: UbiComp, pp 51–60
9. Chawla K, Robins G, Zhang L(2010) Object localization using RFID. In: IEEE international symposium on wireless pervasive computing (ISWPC), pp 301–306
10. Coen MH (1998) Design principles for intelligent environments. In: AAAI, pp 547–554
11. Fishkin KP, Jiang B, Philipose M, Roy S (2004) I sense a disturbance in the force: Unobtrusive detection of interactions with RFID-tagged objects. In: Ubicomp, pp 268–282
12. Floerkemeier C, Lampe M (2004) Issues with RFID usage in ubiquitous computing applications. In: Pervasive, pp 188–193
13. Fröhlich M, Wachsmuth I (1998) Gesture recognition of the upper limbs—from signal to symbol. In: Gesture workshop on gesture and sign language in human-computer interaction, pp 173–184
14. Hinske S (2007) Determining the position and orientation of multi-tagged objects using RFID technology. In: IEEE international conference on pervasive computing and communications workshops (PerComW), pp 377–381
15. Hodges S, Thorne A, Mallinson H, Floerkemeier C (2007) Assessing and optimizing the range of UHF RFID to enable real-world pervasive computing applications. In: Pervasive, pp 280–297
16. Kubitz O, Berger MO, Perlick M, Dumoulin R (1997) Application of radio frequency identification devices to support navigation of autonomous mobile robots. In: IEEE vehicular technology conference, pp 126–130
17. Pavlovic Vl, Sharma R, Huang TS (1997) Visual interpretation of hand gestures for human-computer interaction: a review. IEEE Trans Pattern Anal Mach Intell 19:677–695
18. Liu X, Corner MD, Shenoy P (2006) Ferret: RFID localization for pervasive multimedia. In: UbiComp, pp 422–440
19. Bender M, Deller M, Ebert A, Hagen H (2006) Flexible gesture recognition for immersive virtual environments. In: Information visualization, pp 563–568
20. Mehmood MA, Kulik L, Tanin E (2008) Autonomous navigation of mobile agents using RFID-enabled space partitions. In: ACM GIS, pp 173–182
21. Myungsik K, Young CN (2006) Tracking an RFID transponder using a dual directional antenna. J Nippon Kikai Gakkai Robotikusu, Mekatoronikusu Koenkai Koen Ronbunshu 2006:2P1–E05
22. Nemmaluri A, Corner MD, Shenoy P (2008) Sherlock: automatically locating objects for humans. In: MobiSys, pp 187–198
23. Nikitin PV, Martinez R, Ramamurthy S, Leland H, Spiess G, Rao KVS (2010) Phase based spatial identification of UHF RFID Tags. In: IEEE RFID, pp 102–109
24. Mahfujur Rahman ASM, Hossain MA, Parra J, Saddik AEl (2009) Motion-path based gesture interaction with smart home services. In: MM, pp 761–764
25. Want R, Fishkin K, Gujar A, Harrison BL (1999) Bridging physical and virtual worlds with electronic tags. In: CHI, pp 370–377
26. Vlaming L, Smit J, Isenberg T (2008) Presenting using two-handed interaction in open space. In: IEEE international workshop on horizontal interactive human computer systems, pp 29–32
27. Want R (2006) An introduction to RFID technology. Pervasive 5:25–33
28. Wilson P, Prashanth D, Aghajan H (2007) Utilizing RFID signaling scheme for localization of stationary objects and speed estimation of mobile objects. In: IEEE international conference on RFID, pp 94–99
29. Zhang Y, Amin MG, Kaushik S (2007) Localization and tracking of passive RFID tags based on direction estimation. Int J Antennas Propag 2007:1–9