

## Regression Analysis for Gesture Recognition Using Passive RFID Technology in Smart Home Environments

Kevin Bouchard<sup>1</sup>, Sylvain Giroux<sup>1</sup>, Bruno Bouchard<sup>2</sup> and Abdenour Bouzouane<sup>2</sup>

<sup>1</sup>DOMUS Laboratory, Universite de Sherbrooke, Canada

<sup>2</sup>LIARA laboratory, Universite du Quebec a Chicoutimi, Canada

<sup>1</sup>{Kevin.bouchard, Sylvain.Giroux}@USherbrooke.ca

<sup>2</sup>{Bruno.Bouchard, Abdenour.Bouzouane}@uqac.ca

### Abstract

*Gesture recognition is a field of research that consists in recognizing patterns of movement performed by a human using his body parts with or without the help of a comprehensive device (a mouse, a laser, etc.). This particular area has attracted a number of researchers over the years that applied such algorithms in a broad range of disciplines. In particular, it was exploited on early research initiatives with pervasive environments to enable simple communication with automation systems. Nowadays, those environments are used for more than automation. Many researchers, in fact, believe it is one of the most promising solutions to the problems related to ageing of the population. Smart homes are seen as an alternative to the full-time support of a semi-autonomous person by healthcare professional and thus also a potentially economically viable solution to the rising cost of such support. However, researchers are still facing many challenges in that regards, such as the comprehension of the context and of the ongoing activity of daily living. In that equation, gesture recognition could help extract more information from the collected data and thus reinforce the context modeling. The knowledge extracted could even help with monitoring of more fine-grained activities and with the understanding of normal or abnormal behaviors. Gesture recognition is often considered as a solved problem since the techniques to perform it work well as soon as we can track accurately. However, in smart environment, tracking is very imprecise and hard to achieve with limited technology (i.e. noninvasive sensors). In this paper, we present a novel gesture recognition algorithm that works under uncertainty, and that is scalable to the precision of the tracking system. The algorithm is based on the tracking of passive RFID tags installed on all everyday life objects in a smart environment. A set of experimentation in simulation and in a real smart home environment is presented. The results are very encouraging despite the very low accuracy of the passive RFID tracking system.*

**Keywords:** Linear Regression; Smart Home; Activity Recognition; Passive RFID

### 1. Introduction

The occidental societies of the world are currently facing ageing of their population [1]. The effects of the ageing have been well studied in the literature, and the various consequences are well-known to researchers in social sciences. Among them, there are the decrease, not to say a shortage, in qualified healthcare professional and the costs of the system which take an increasing percentage of gross domestic product (GDP) expenditure. Many researchers are now believing that the solution to these difficult societal transformations lie partially in the technological advances. The so call ambient intelligence combined to the miniaturization and improvement of sensing technologies are now exploited

everywhere in the world to create smart homes prototypes toward that very goal of addressing these issues. In particular, smart environments could be used to reduce the burden of families of persons with decreasing autonomy such as a cognitive impairment due to Alzheimer disease. However, to reach this goal, many difficulties must be overcome first (opportunistic networking, context modeling, human activity recognition, etc.). Context modeling [2] and human activity recognition (HAR) [3] are closely tied together as the first one leads the second to more precision and enables better delivery of assistive services. Extensive research has been conducted on both [4] but there is still a need for improvement. In particular, many researchers think that we need to better exploit the raw information extracted from the sensors installed in smart environments. For example, Jakkula & Cook [5] exploited the temporal relationships between events created by the trigger of sensors and Augusto & al. [6] designed a spatiotemporal inference engine for the smart home. We believe that being able to extract current and past gestures from the data would be another step toward a smarter use of information.

Gesture recognition is an old problem that has particularly attracted researchers on Human-Computer Interfaces (HCI) [7]. Many algorithms are used for natural and efficient design in video games, software engineering and even in smart home [8]. For that last field, cameras are usually the privileged sensors. In the context of smart home assistance, however, many researchers believe invasiveness should be limited for the technology to be accepted [9], and cameras are generally considered as one of the most invasive technologies. Moreover, in the case of some cognitive diseases, such as Alzheimer, it can even worsen the state of the resident. Accelerometers can also be exploited for gesture recognition [10]. These sensors provide a good tradeoff between price and accuracy. They enable strong gesture recognition with a good amount of precision. Nevertheless, they need to be mounted on body parts of the resident and are often also considered invasive. Moreover, there is no guarantee the resident wears the equipment at all times for the system to be able to extract the gestures. One of the most promising technologies for smart home is the Radio-Frequency IDentification (RFID). This technology possesses many advantages for our context. It is cheap (a passive tag cost only few pennies), but also non intrusive. Passive tags can basically be placed on any object of a smart home due to their small size. This technology has been ignored for gesture recognition, until recently [11], because of its inherent imprecision. However, the advances in localization and tracking system now enable us to foresee gesture recognition as something feasible in near future. In this paper, we propose to exploit our recent passive RFID tracking system [12] to recognize ongoing gestures from the manipulation of daily usage objects tagged in the smart home. To do so, we exploit linear regression combined with the correlation coefficient to identify the direction of movement and estimate the segmentation [13]. A large set of experiments was conducted with a random simulator using a normal distribution. Additionally, another set of experiments was conducted in realistic conditions inside a smart home infrastructure, and those results are compared with the literature.

The remainder of this paper is divided as follows. Section 2 is a brief overview of the literature on RFID localization and on the recognition of gestures. Section 3 describes the tracking method that was implemented to conduct the experiments in the smart home. Thereafter, section 4 relates how the regression analysis was exploited to determine basic direction of movement and to perform the task of gesture recognition. Section 5 details the experimental protocol that was followed in simulation and in the smart home infrastructure of the LIARA laboratory. Finally, section 6 concludes with an assessment of our new method and opens on the future work to be accomplished.

## 2. Related Work

Gesture recognition is a well-established field of research that traditionally focuses on HCI [14]. A gesture is widely described and recognized as an expressive and meaningful body motion (hand, face, arms, *etc.*) that conveys a message or more generally, embeds important information of spatiotemporal nature. Gestures are ambiguous and incompletely specified, since a multitude of conceptual information can be mapped to one gesture. The usual steps to perform gesture recognition from spatiotemporal data series are the following:

1. Segmentation
2. Filtering of the data
3. Limiting directions
4. Matching

The segmentation is the step that divides the data series into different part to find the beginning and the ending of the gestures. In many cases, the segmentation step is ignored because the user specifies the start and the end of a gesture with a device or simply because it is assumed that it is known by the computer system (*e.g.*: starts when the object is moving, end when it becomes idle). A difficult challenge of segmentation is the support of gestures of varying length interleaved with small to big inactive time. As you will see through this paper, segmentation is, in fact, the biggest challenge in our context. The filtering is a straightforward step consisting in standardizing the data (time, format, *etc.*) and compensating for missing data. Interpolation is often used to compensate for the missing data. In our case, this part of the gesture recognition is done by the tracking system which provides coordinates (position) at a fixed rate. The third step, limiting the directions, is used to reduce the quantitative data into a finite set of qualitative directions. The step of the limitation of directions is generally very different from one application to another. That is because it really depends on the granularity of gestures that need to be recognized. Finally, the last step consists in using those sequences of limited directions to match them with gestures that are defined or learned in a knowledge base. The research has always focused on that step [15].

### 2.1. Main Gesture Recognition Models

Many gesture recognition approaches are based on statistical modeling such as the Hidden Markov Machine (HMMs), the Kalman filtering or other particles filtering [7]. For instance, Samaria & Young [16] exploit HMMs to extract efficiently facial expressions from a single camera. The reasoning corresponds to the process of finding the HMM with the highest probability of explaining that set of observations. It is generally required to design and train one HMM per gesture that we desire to recognize [8]. Methods exploiting particle filters are also very popular [14]. For instance, Shan & al. [7] combined the technique with Mean shift to perform real time hand tracking. Their algorithm, named Mean Shift Embedded Particle Filter (MSEPF), was tested on a 12fps camera stream with a 240x180 pixels resolution. They showed that their method could robustly track a hand to recognize gestures. Particle filters are very effective in estimating the state of dynamic systems from sensors information. Finally, a large number of gesture recognition approaches effectively exploited Finite State Machines (FSMs). For instance, Hong & al. [17] exploited spatial clustering to learn a set of FSMs corresponding to gestures. They tested their approach using four sample gestures performed in front of a video camera. They achieved a hundred percent recognition rate, but admitted that with a very noisy data sample, the recognition would fail.

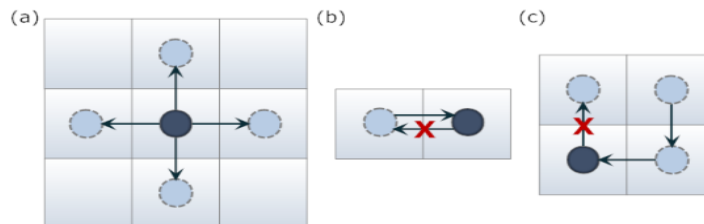
The main problem with the models of the literature for our specific context with passive RFID in smart home is the assumptions that are made. First of all, it is often assumed that obtaining the basic directions of the movement is straightforward. It is not the case with passive RFID tracking. Secondly, it is assumed that the amount of noise is not a problem (or that there is simply no noise). Thirdly, segmentation is often not an issue within an HCI context; therefore, few models address this issue. Finally, they generally suppose that the user is cooperative; an intended recognition context [18]. In our case, the recognition is done unbeknownst to the user (keyhole context) which means fewer definite gestures.

## 2.2. RFID Tracking

As we have seen, most approaches that perform gesture recognition exploit cameras. Such sensing devices possess the advantage of being extremely informative enabling either hand or objects tracking and even recognition of facial expressions. However, cameras are difficult to install, require very stable lighting condition, are invasive and the processing of their output data is highly complex. It would also be possible to use laser scan or ultrasonic methods, but these are limited by line of sight constraints. Many researchers, including our team, believe that passive RFID technology is the best option in the context of smart home. For gesture recognition, we could put passive tags on any objects and track their movement. A lot of research has been conducted on the tracking of RFID tags [12]. Despite this, most solutions are unusable for our context. The major part is based on reference tags [19, 20] which imply the deployment of a large quantity of tags fixed at known positions in order to compare the received signal strength indication values with the tracked objects. It is hardly feasible to do this for assistive smart homes. Others use hybrid technologies such as laser [21] or LEDs [22], but these technologies cannot be mounted directly on daily life objects. Finally, few approaches exploit the trilateration technique, similarly to the tracking system used in this paper. The problem is that they have only very limited accuracy [23] which does not permit gesture recognition. To overcome the challenges of passive RFID gesture recognition, specialized tracking systems need to be developed.

## 2.3. Gesture Recognition using RFID

The team of Asadzadeh & al. [11] is, to the best of our knowledge, the only one that investigated the problem of gesture recognition with passive RFID technology. With three antennas on a desk, they monitored an 80cm by 80cm area, which was divided into 64 equally sized square cells (10cm by 10 cm) and localized using reference tags (a technique that is not usable in our context). Their system is based on two important assumptions. First, the tracking system is considered as fast enough to never miss any cell in a sequence; that is, the tracked object cannot move farther than one cell away in between two readings. Second, only forward local moves are possible. Figure 1 below shows legal (a) and illegal moves (b-c) as defined in their system.



**Figure 1. (a) Legal Move, (b) (c) Illegal Moves**

Their algorithm was tested with a dictionary of twelve gestures and obtained a 93% recognition rate. However, their system is not able to perform segmentation and thus, two consecutive gestures cannot be recognized. Additionally, their assumptions made it difficult to apply their system in a smart home which requires more flexibility. Furthermore, they worked in an intended context, where the user purposely performs gestures accurately to help the system, while in our case, we perform in a keyhole recognition context. Nevertheless, their work showed that there is potential for gesture recognition with passive RFID.

### 3. Tracking Objects

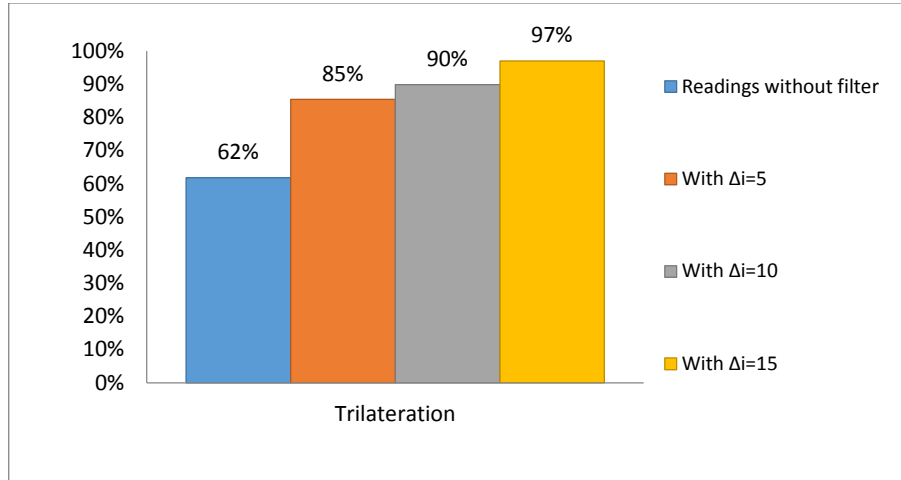
In order to recognize the gesture in our smart home, the first step that needs to be accomplished is the tracking of the hand of the human resident. However, since we prefer noninvasive methods, we want to avoid the use of bracelets. Therefore, we decided to track objects of the smart home with passive RFID tags. To do so, we adapted a version of our research project on RFID localization that has been published in [12]. As we explained in the previous version, we need a trilateration based algorithm and few solutions exist in the literature. In this section, we describe the different filters that we use to enhance the trilateration and adapt it to the particularities of the technology. We also describe how the algorithm was built to work with an elliptical model instead of a circular one. Finally, we present partial results from many of our experiments with passive RFID to demonstrate the contribution of each portion of the localization algorithm. Each of these experiments was conducted at the LIARA's smart home that is described in the section 5.

#### 3.1. Addressing the false reading

One of the difficulties that arises when developing tracking algorithms based on passive RFID technology is the false-negative reading (FNR) challenge. That problem occurs when a tag is in the antenna coverage area but is not detected during a period of time. This can happen for more than one reason. The reader can have an internal problem limiting its capacity for a period; the radio waves emitted from many tags can collide or environmental noise (metal, person, *etc.*) can create interference [24]. In our experience, FNRs are unavoidable and exist in all RFID systems. However, it is possible to limit the impact of this problem with simple methods. In our case, we use a function called  $f_{ite}(i)$  that exploit more than one reading cycle (equation 1).

$$f_{ite}(i, O_s) = \begin{cases} !O_s & |i_c - i_d| \geq \Delta i \\ O_s & otherwise. \end{cases} \quad (1)$$

That function uses the Boolean state of a tag ( $O_s$ ) and modifies it in function of iteration value. The state changes if the current iteration number ( $i_c$ ) minus the first detection iteration ( $i_d$ ) of a sequence of the opposite state is higher than a constant  $\Delta i$ . The  $\Delta i$  constant was selected by measuring the average false reading length experimentally. Of course, the higher it is, the less FNRs will occur. We tested this method at a 200ms per iteration in our smart home (see section 5) and various values for  $\Delta i$ . The results are compiled in Figure 2.



**Figure 2. False Readings Rate with and Without the Filter**

### 3.2. Stabilization of the RSSI

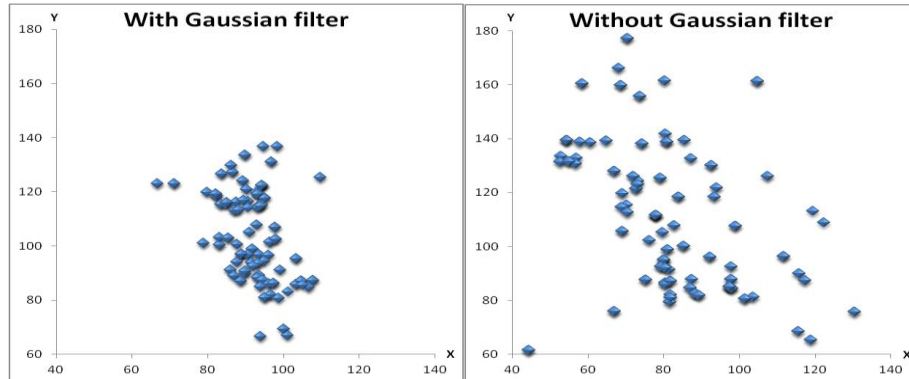
The second problem that needs to be addressed in order to perform trilateration is the flickering of the received signal strength indication (RSSI) over a period of time. To do so, the signal is smoothed by using a weighted average. The first part uses a Gaussian curve (equation 2) which is centered on the current iteration number  $i_c$ .

$$f_{Gaussian}(i) = e^{-\frac{1}{2}\left(\frac{i_c-i}{\sigma}\right)^2} \quad (2)$$

The parameter  $i$  is the iteration number associated with the RSSI record that we are weighting. The  $\sigma$  constant can be determined automatically corresponding to the number of iterations you want to weight (it is proportional to the length). Our general rule is to divide the number of iteration to weight by two to give the value of  $\sigma$ . For instance, at  $\sigma = 2$  the fifth iteration is weighted at 13.53% and the eleventh at 0.67%. The best value of this parameter could also be easily learned in a supervised manner. Thereafter, the mean weighted RSSI of a tag is computed by making use of the equation (3).

$$f_{strength}(t[i_c]) = \frac{\sum_{i=i_c-\Delta i}^{i_c} t[i]_{rssi} * f_{Gaussian}(i)}{\sum_{i=i_c-\Delta i}^{i_c} f_{Gaussian}(i)} \quad (3)$$

The function takes as an entry the array  $(t[\cdot])$  containing the RSSI readings. The  $t[i]_{rssi} * f_{Gaussian}(i)$  denotes the weighted value for the  $i^{\text{th}}$  iteration. The constant  $\Delta i$  is the number of iterations considered for the average RSSI calculation and is necessary only to limit the computation (optimization). This filter was tested in our previous publications [12] and Figure 3 shows the improvement of the basic circular trilateration of a tag with that filter (estimations are closer).



**Figure 3. Concentration of the Approximate Positions of an Object**

### 3.3. Elliptical Trilateration

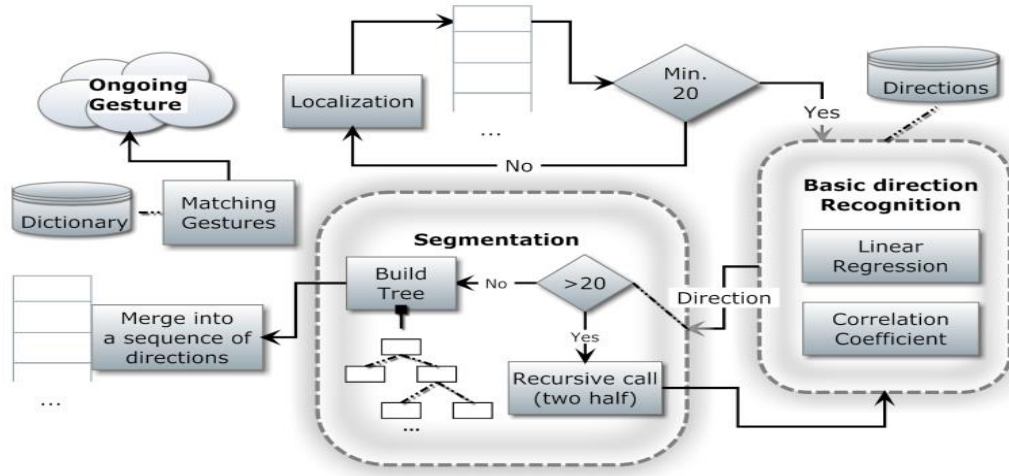
The last step consists of performing the trilateration. Traditionally, it is done by converting the RSSI to a distance from each antenna. Then, with at least three antennas, the values are used to create circle equation. The intersection of these circles is the estimated position of the tracked entity. However, the wave emissions of indoor RFID antennas are more alike to ellipses (directional). For that reason, our tracking system is based on elliptical trilateration. To transform the basic RSSI into ellipse equation, the first step was to record RSSI values of a tag placed at different distances in front and on the side of an antenna. With that dataset, we performed a polynomial regression that gave us one equation for each axis of the ellipse. With those two equations, the ellipses can be determined in function of the current RSSI. Thereafter, the idea is the same as usual. For the trilateration to work, we need at least two antennas on the same wall or at least three on different walls. We can find intersection points between a pair of antennas on the same wall easily. When on different walls, we use the method of Ferrari, which gives between 0 to four points. In the case that the ellipses do not intersect, the equations are progressively modified in proportion to the RSSI value. The final position is found by removing the intersection points that are outliers and doing a weighted average on the remaining ones.

We implemented and tested the elliptical trilateration in our smart home (section 5) [12]. The system was set to 200ms per iteration, and four tags (whenever possible) were installed on the tracked objects. The average accuracy obtained was 14.12cm. For the purpose of this research on gesture recognition, we increased the speed to 20ms per iteration. Gesture recognition is highly dependent on the amount of data. With limited data, we might otherwise miss fast movements, and thus the recognition will also be limited.

## 4. Gesture Recognition

In this section, we present the new gesture recognition method that takes as input the coordinates extracted from the localization algorithm described in Section 3. The method depends directly on the accuracy of the localization algorithm and takes it as a parameter ( $\epsilon = 14cm$ ). That is, a valid direction cannot be less than the average error and is probably significantly longer. This is the limitation of the granularity of our gestures. However, by performing the gesture recognition with the average error as a parameter, the model is more scalable since it adapts itself automatically to any tracking method. The new method also depends heavily on the reading rate of the tracking algorithm. In our case, as stated before, we obtain a new position for an object every 20ms or to put it simply; we have 50 positions per

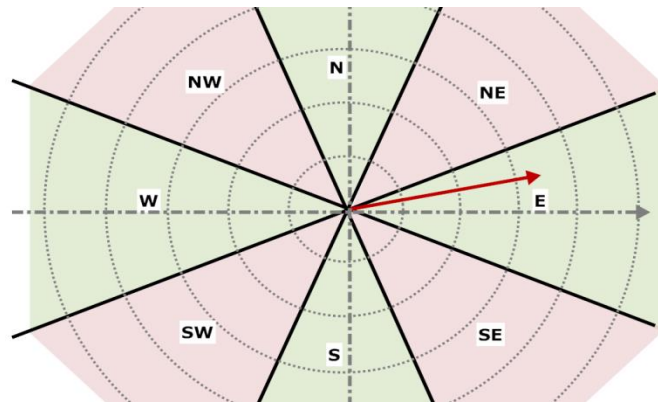
second. The main contribution of this new method is on the step of identification of the basic directions that are generally assumed as known in the literature and on the segmentation of these directions which is particularly arduous with passive RFID technology. The Figure 4 depicts the overall method that is presented through this section.



**Figure 4. The Overall Gesture Recognition Method**

#### 4.1. Data Processing

The first step toward gesture recognition is to ensure that the flow of incoming data is stable. The synchronization part is done at the step of localization. We receive the data from all the objects of the smart home simultaneously every 20ms. Thereafter, we have to convert the set of points that we obtain to basic directions. There are two essential elements to this step. First of all, the possible directions have to come in a limited set. Since our data is noisy, we think that a limit to eight qualitative directions is a good tradeoff between precision and accuracy. These eight directions are shown on the Figure 5 below:



**Figure 5. The Eight Direction Exploited in our Model**

In that qualitative model, a quantitative direction is converted by using its angle with the abscissa axis. For instance, on Figure 5 the red arrow representing a quantitative direction would be converted to a qualitative East. The use of qualitative direction is not only mandatory; it also provides many advantages. For example, it would be straightforward to



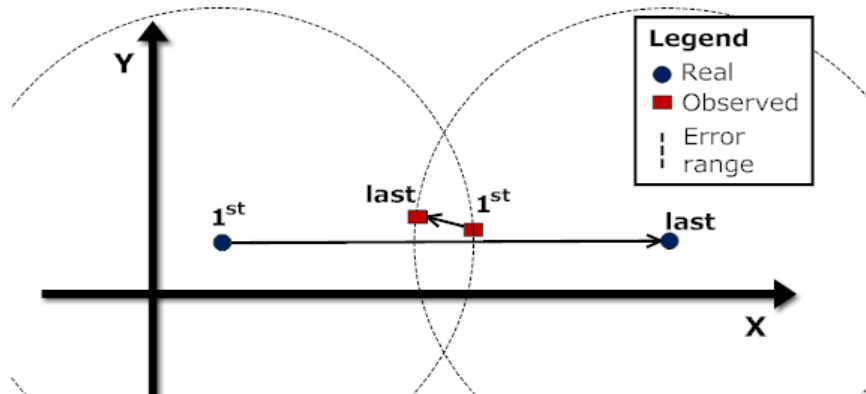
extend our model with fuzzy sets. These fuzzy sets could then be exploited by the Gesture recognition module when no direct matches are found in the gestures' dictionary. As discussed at the end of this paper, this is part of our future work on gesture recognition.

The second part of this step is to infer the said direction from the points. To do so, we perform linear regressions. At this step of the process, we suppose that the set of data correspond only to one of the qualitative directions. Later in this paper, we explain how to perform the important but difficult step of segmentation. From a set of positions, the linear regression gives a linear function of the form  $y = ax + b$  using equation 4 and 5.

$$a = \frac{(n(\sum_{i=0}^n x_i y_i) - (\sum_{i=0}^n x_i) (\sum_{i=0}^n y_i))}{(n(\sum_{i=0}^n x_i^2) - (\sum_{i=0}^n x_i)^2)} \quad (4)$$

$$b = \frac{(\sum_{i=0}^n y_i)}{n} - a \frac{(\sum_{i=0}^n x_i)}{n} \quad (5)$$

From the resulting equation, we can compute directly the angle from the x-axis by doing the  $\arctan(|a|)$  of the slope. The result of this operation leads to two opposite possible directions (from the qualitative framework). We chose between both directions simply by doing an average of the starting points and comparing it with the average of the ending points. To be clear, the set is divided into two equal part (if possible) and both averages are computed. We cannot simply compare the first and the last point because it could mislead us in unlucky situations. The Figure 6 depicts the situation that is generally avoided with a simple average:



**Figure 6. In Unlucky Situations, Taking the First and the Last Position can be Highly Misleading. It Doubles the Average Error obtained from the Localization**

The last but nonetheless, important aspect of the direction identification is the speed at which the system goes. We could let the system execute as fast as possible, but with many objects and a lot of incoming data, the performance could decrease significantly. Moreover, it appears to be clear that for a direction to be detected under uncertainty, a minimum of half a second of data is needed. Additionally, humans usually do not perform gestures shorter than few seconds. For those reasons, in our system, the recognition is only executed every 10 iterations (200ms) and when we have at least 20 positions. Besides, we must also eliminate old data over time to limit escalation of computation time required as a final step. More details are given later about how we do this.

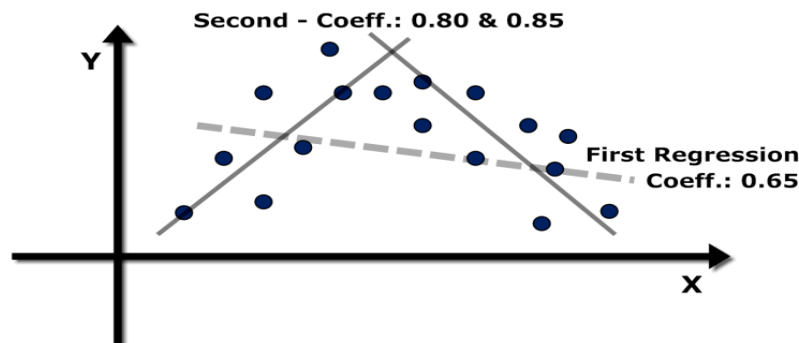
## 4.2. Segmentation

Being able to identify the current direction is not the only challenge toward gesture recognition with RFID. Another very difficult part is to be able to understand where the individual directions are ending or if the object is actually moving or not. This part of our new algorithm is based on the correlation coefficient which can be computed using equation 6.

$$\varphi = \frac{(n(\sum_{i=0}^n x_i y_i) - (\sum_{i=0}^n x_i)(\sum_{i=0}^n y_i))}{\sqrt{n(\sum_{i=0}^n x_i^2) - (\sum_{i=0}^n x_i)^2} * \sqrt{n(\sum_{i=0}^n y_i^2) - (\sum_{i=0}^n y_i)^2}} \quad (6)$$

From that equation, we always obtain a value of  $\varphi$  comprised between -1 and 1. If the value is far from 0, the correlation is high between the data points. Our hypothesis was that this information could be exploited to distinguish between an idle object (with high localization error) and a moving one. Knowing that, we recorded the values of several subsets of data during when an object was idle to learn the correct threshold of the correlation coefficient to exploit in our model. We found out that on average, when the object was idle,  $\varphi < 0.4$ . However, assuming that an idle object is moving when it is not can be very damaging for the algorithms using the data (keeps in mind that we work on gesture recognition in the goal of exploiting the knowledge for assistive smart homes). Consequently, we used a slightly higher value (0.5) in our implementation to decide whether the object was idle over the period evaluated or moving in a certain direction.

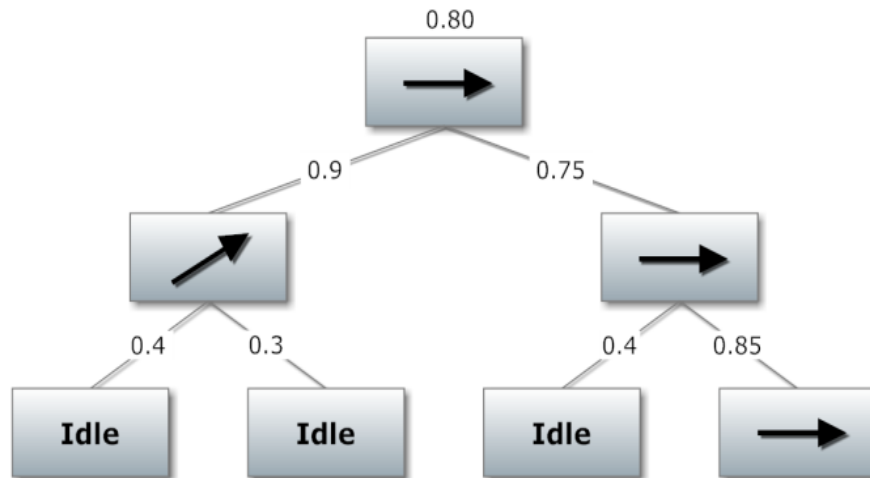
To perform the segmentation, we exploit a divide and conquer method. The idea is simply to divide the data in half and redo the previous steps until stop conditions are met. The stop conditions are either when we have 20 positions or less or when the correlation coefficient drop under idle while resulting in the same direction. The idea is that by doing that, we should be able to identify that a data sequence, even with a good correlation coefficient value, is, in fact, composed of multiple different directions. The Figure 7 shows an example of such a situation (the coefficients are given as an example) where the data is the result of two atomic directions.



**Figure 7. An Example of how the Correlation Coefficient is used to Perform the Segmentation**

**4.2.1. Decision Process:** The previous step gives a tree structure where each level corresponds to a potential sequence of directions. To choose the sequence of direction (0 to n), we need to merge the tree hypothesis by comparing the various correlation coefficients at each level. There are three possible situations. If two leaves have the same direction, they are merged to the superior level. If one of the two leaves is idle but not the other, the algorithm

takes the decision in function of the value of  $\varphi$ . If the direction of the lower-level leaf has a higher coefficient value, the conclusion is that the sequence is comprised of those two leaves. If it is lower, the leaves are deleted and the higher level is considered correct. The last possibility is that the two leaves are of different direction (but not idle). In that situation, to decide if they are kept, the algorithm computes the average value of both coefficient and compares it with one of the superior levels. At the end of the decision process, the sequence of directions is updated along with the approximate iteration each began and ended. The Figure 8 depicts an example tree structure that could be created during the step of segmentation.



**Figure 8. An Example of Tree Resulting from the Multiple Regressions. The Numbers are the Correlation Coefficients for each Part of the Dataset**

In that case, the conclusion would be that the first half of the data was NorthEast with a very high correlation (0.9). The second half would be composed of a short idle time with an East direction. However, in the recognition, the idle could be ignored if no gesture matches that sequence.

**4.2.2. Limiting the Data Growth:** As we mentioned before, it is important to limit the growth of the dataset if we do not want the computation time to explode. To do so, we designed a simple solution that exploits the growing certainty of our conclusions over time. That is, we only reevaluate the data points of the two last direction excluding the idle time in between. For example, it supposes that the first three recognition iterations give respectively:  $\{(1; N), (2; NE, E), (3; N, NE, E)\}$ . Then, for the fourth recognition iteration, the first direction (N) will be locked, and the data associated with it will no more be evaluated. The data associated is the set of positions that was used to infer that direction (in the tree structure).

### 4.3. Matching the Gestures

The final part of our method consists of matching the list of identified directions to the gestures in the dictionary. For this part, the literature proposes a variety of methods developed through years of research [14, 17]. For this work, we preferred to keep that part simple as it is not the main challenge to gesture recognition from RFID. Once we are able to find basic directions and to perform segmentation, we can rely on standard methods. Our gesture

dictionary is a set of finite state machines representing each gesture. The selected ongoing gesture is the state machine that matches the sequence of atomic directions identified. However, the matching module is not strict. If a sequence comprising of small idle moment does not match any gesture, they are progressively eliminated until either the sequence match or until no more remain. Contrary to most work in the literature, we do not assume that a gesture was intended. That is because in our context, the user is a normal resident or a resident with a cognitive deficit that does not purposely intend to perform a gesture with the objects he moves. Finally, note that for our experiments, each FSM was directly designed by us since they are fairly small. However, they could certainly be learned automatically with knowledge discovery techniques.

## 5. Experiments

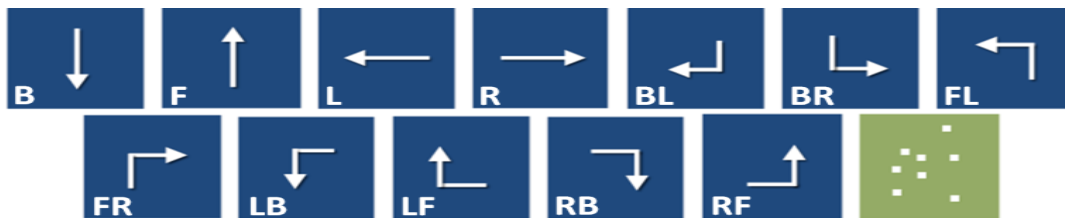
To validate the potential of our new method, we implemented it at the LIARA's smart home laboratory. Our smart home infrastructure is composed of eight RFID antennas and a large number of various sensors (ultrasonic, infrared, load cells, *etc.*) and effectors. However, to track objects in real-time, we need at least two antennas on the same wall or three on different walls. Only the kitchen zone is equipped for that purpose with four antennas that can detect up to three meters in front of them. The kitchen is the perfect place to track accurately everyday life objects since it is the place where the most complex activities of daily living (ADLs) take place. In fact, we think that it is not necessarily useful to have a very high level of information everywhere in the smart home since it would raise the cost of the technology without necessarily improving the quality of services. For instance, in the bedroom, most ADLs are very simple and easily identifiable. As it is not the purpose of this paper, reader should look into [25] which justify some of the technical choices made for our infrastructure.

The antennas we are currently using are A-PATCH-0025 working on the 860 - 960MHZ band and are circularly polarized for a better indoor GSM coverage. The zone we want them to cover is approximately 9m<sup>2</sup>. We set the parameter of the RFID system at high sensitivity (almost max), and the emission rate was configured at one reading per 20ms. Each object of the smart home has 2 to 4 class 3 RFID tags attached to it. It is important in order to limit the bad angle of arrival, and thus we always select the highest RSSI from the tags. Another point worth mentioning is that cabinets are shielded (partially) in order for the objects inside to remain mostly invisible to the antennas. The shielding is done in order to reduce the amount of collisions from the propagation of the radio-frequency waves. The RFID system can support a lot of objects before the localization accuracy significantly decreases, but further experiments would be needed to determine exactly how many.

### 5.1. Generation of Random Gestures

As a first set of experiment, we implemented a simulator that generated gestures to be recognized by our new method. The simulator enabled us to do an extensive amount of tests in a short time interval that the complex protocol needed with human subjects would never allow us to do. The generator works simply by randomly selecting a FSM corresponding to a gesture in the dictionary and computing the next position using user-specified parameters. These parameters comprise the generation speed (ms), the object speed (cm/s), the gesture length (seconds) and the average positioning error (cm). The error is used to generate noise. We suppose with the generator that the noise follows a normal distribution. For example, if the object should be at (10, 0) and the average error specified is  $\pm 14$ cm, then the generated position would be  $(-4 \leq x \leq 24, -14 \leq y \leq 14)$ .

Our algorithm is able to detect gestures composed of any sequence of basic directions. However, in a first time, we wanted to compare our performance with the only other RFID based gesture recognition model, so we decided to reproduce their experimental setting. In their experiment, Asadzadeh *et al.*, [11] used only four basic directions. On average, their gestures lasted 4.5 seconds at 20cm/sec. Their localization algorithm was, although unusable in our context, more accurate than our method with an average error of approximately 10cm. We used that error for the simulation but reduced the length of the gestures to approximately 40cm per basic directions. Moreover, to test our segmentation, we added a random variable to the distance (-10 to +20 cm). We also added the *idle* gesture in the dictionary because in a realistic context, most of the time, objects are idle. The Figure 9 shows the set of gestures.



**Figure 9. Example Gestures used for the Experiments. Eight are Composed of Two Directions, Four of Only One. The Last on the Picture is *Idle***

We let our generator work for about 2000 gestures generated randomly, and we obtained positive results (86% success). Table 1 details the results that were obtained from this set of tests. The most important thing to understand is that recognizing the directions was not difficult; most of the errors were due to the process of segmentation. It means that with the same assumption (no need for segmentation) of the team of Asadzadeh *et al.*, [11] our new method would have performed better. The other errors are mostly misclassification between the idle and real directions which is also significant since Asadzadeh *et al.*, [11]. did not include that possibility.

**Table 1. The Results Obtained from the Simulation**

Gesture	True Positive	False Negative	False Positive
Idle	157	14	<b>116</b>
Front (F)	149	26	8
Back (B)	142	23	11
Left (L)	145	14	13
Right (R)	151	20	16
Left Front (LF)	156	21	18
Left Back (LB)	163	12	12
Back Right (BR)	139	25	13
Right Back (RB)	126	29	9
Front Left (FL)	158	21	17
Front Right (FR)	155	13	14
Back Left (BL)	137	29	11
Right Front (RF)	148	27	16
	<b>1926</b>	<b>274</b>	<b>274</b>

## 5.2. Experiments with a Human Subject

Since the results obtained from the gesture generator were good, we decided to conduct a set of experiments directly in the smart home infrastructure. For that purpose, a human subject was asked to perform each gesture a total of ten times. The protocol of Asadzadeh et al. [11] was exactly reproduced this time. The human was using a standard cup of coffee with four RFID tags on it and the system used the position of the cup to infer the gesture. The cup was initially put on the kitchen counter for the tests. A physical guideline was put beside the cup to show the human subject the average distance and direction that should compose the gesture. Therefore, he had only to approximately move the cup following that guideline. The appropriate guideline was installed before each series of tests (for a gesture). No particular instructions were given to the human for the initial position he should begin the gesture. Sometime the human has held the cup in his hand before start and other times he just performed the gesture right away on the kitchen counter. The results of the experiments can be seen on Table 2.

**Table 2. The Results from the Set of Gestures Simulated by a Human Subject**

R	L	F	B	FR	BL	FL	BR	RF	LB	LF	RB	Idle	Total
9	8	7	9	7	8	8	7	7	8	7	7	9	77%

As shown on the table, the results are slightly inferior to those obtained with the generator. There are two explanations to this. First, the noise in the data obtained from the RFID system is actually not following a normal distribution. It means that often, when the data begin to be inaccurate, it moves in a distinguishable direction which can be recognized and thus, it creates some issues with the segmentation process. Secondly, in a realistic environment, there are unpredictable interferences that lead to recognize directions that never happened. For example, if the human is hiding one or many antennas for a certain time, this might lead to a significant modification of the estimated position and thus to identify a movement that is not real. We think that, in the future, these patterns could be learned and recognized and thus removed from the dataset whenever necessary.

## 6. Conclusion

In this paper, we presented our first attempt to address the problem of gesture recognition from passive RFID technology in smart environments. Before us, this technology was avoided in that field of research since it is considered too imprecise. The model we developed deal with the imprecision at many levels. First, it tries to improve the data directly collected from the RFID. Second, the localization method is adapted to the precise context of smart home. Third, the recognition method tackles the challenge of recognizing basic direction which few papers discuss in the literature. Finally, our model uses a novel method based on the correlation coefficient to perform the segmentation of a continuous stream of data.

The recognition of gesture would be very useful toward the implementation of the assistive smart home vision. In particular, it would enhance context modeling and help with human activity recognition. However, there are still many challenges toward the real-world applications of passive RFID gesture recognition. One limit of this work is the dependence on the absolute Cartesian position. This limits the scalability of such system to other smart environments. In future work, we aim to address this issue by exploiting relative positions of objects in the environment. We also aim to conduct further experiments with a bigger dictionary of gestures. We are working to obtain an approval certificate from the ethical

committee in order to conduct larger experiments with normal and cognitively impaired subjects. We also aim to improve the model by adding fuzzy inference to the recognition and by learning the gesture models automatically. Finally, we constantly work on the localization algorithms in parallel in order to improve the data. As the reader probably noticed, gesture recognition with RFID is highly dependent on both the improvement of the technology and of the localization algorithm.

## Acknowledgements

Financial sponsors: the Natural Sciences and Engineering Research Council of Canada (NSERC), the Quebec Research Fund on Nature and Technologies (FRQNT), the Canadian Foundation for Innovation (CFI).

## References

- [1] U. Nations, World population ageing 2009: United Nations, Dept. of Economic and Social Affairs, Population Division, (2010).
- [2] C. Phua, J. Biswas, A. Tolstikov, V. Foo, W. Huang, M. Jayachandran, A. P. W. Aung, P. C. Roy, H. Aloulou, M. A. Feki, S. Giroux, A. Bouzouane and B. Bouchard, "Plan Recognition based on Sensor Produced Micro-Context for Eldercare", Proceedings of The First International Workshop on Context--Awareness in Smart Environments: Background, Achievements and Challenges (CASEbac 2009), Toyama, Japan, (2009), pp. 39-46.
- [3] T. Gu, L. Wang, Z. Wu, X. Tao and J. Lu, "A Pattern Mining Approach to Sensor-Based Human Activity Recognition", IEEE Trans. on Knowl. and Data Eng., vol. 23, (2011), pp. 1359-1372.
- [4] C. Ramos, J. C. Augusto and D. Shapiro, "Ambient Intelligence: the Next Step for Artificial Intelligence", IEEE Intelligent Systems, vol. 23, (2008), pp. 15-18.
- [5] V. R. Jakkula and D. J. Cook, "Enhancing Smart Home Algorithms Using Temporal Relations", Technology and Aging, A. Mihailidis, J. Boger, H. Kautz, and L. Normie, Eds., ed Amsterdam: IOS Press, vol. 21, (2008), pp. 3-10.
- [6] J. C. Augusto, J. Liu, P. McCullagh and H. Wang, "Management of uncertainty and spatio-temporal aspects for monitoring and diagnosis in a Smart Home", International Journal of Computational Intelligence Systems vol. 1, (2008), pp. 361-378.
- [7] C. Shan, Y. Wei, T. Tan, and F. Ojardias, "Real time hand tracking by combining particle filtering and mean shift", Automatic Face and Gesture Recognition, 2004. Proceedings. Sixth IEEE International Conference, (2004), pp. 669-674.
- [8] T. Westeyn, H. Brashear, A. Atrash and T. Starner, "Georgia tech gesture toolkit: supporting experiments in gesture recognition", Proceedings of the 5th international conference on Multimodal interfaces, (2003), pp. 85-92.
- [9] K. Mäkelä, S. Belt, D. Greenblatt and J. Häkkinen, "Mobile interaction with visual and RFID tags: a field study on user perceptions", presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, San Jose, California, USA, (2007).
- [10] J. Liu, L. Zhong, J. Wickramasuriya and V. Vasudevan, "uWave: Accelerometer-based personalized gesture recognition and its applications", Pervasive and Mobile Computing, vol. 5, (2009), pp. 657-675.
- [11] P. Asadzadeh, L. Kulik and E. Tanin, "Gesture recognition using RFID technology", Personal and Ubiquitous Computing, vol. 16, 2012/03/01 (2012), pp. 225-234.
- [12] D. Fortin-Simard, K. Bouchard, S. Gaboury, B. Bouchard, and A. Bouzouane, "Accurate Passive RFID Localization System for Smart Homes", presented at the 3th IEEE International Conference on Networked Embedded Systems for Every Application, Liverpool, UK, (2012).
- [13] S. Brandt, "Data Analysis: Statistical and computational methods for scientists and engineers with 273 illustrations", includes CD-Rom: Springer, (1999).
- [14] S. Mitra and T. Acharya, "Gesture recognition: A survey", Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions, vol. 37, (2007), pp. 311-324.
- [15] L. Bretzner, I. Laptev and T. Lindeberg, "Hand gesture recognition using multi-scale colour features, hierarchical models and particle filtering", Automatic Face and Gesture Recognition, 2002. Proceedings. Fifth IEEE International Conference, (2002), pp. 423-428.
- [16] F. Samaria and S. Young, "HMM-based architecture for face identification", Image and vision computing, vol. 12, (1994), pp. 537-543.

- [17] P. Hong, M. Turk, and T. S. Huang, "Gesture modeling and recognition using finite state machines", Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on, (2000), pp. 410-415.
- [18] P. R. Cohen, C. R. Perrault, J. F. Allen, B. Bolt, i. Newman and U. S. A. R. P. Agency, Beyond question-answering: Bolt Beranek and Newman Inc., (1981).
- [19] P. Vorst, S. Schneegans, Y. Bin and A. Zell, "Self-Localization with RFID snapshots in densely tagged environments", Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on, (2008), pp. 1353-1358.
- [20] J. L. Brchan, Z. Lianlin, W. Jiaqing, R. E. Williams and L. C. Perez, "A real-time RFID localization experiment using propagation models", RFID (RFID), 2012 IEEE International Conference, (2012), pp. 141-148.
- [21] A. Milella, D. Di Paola, G. Cicirelli and T. D'Orazio, "RFID tag bearing estimation for mobile robot localization", Advanced Robotics, 2009. ICAR 2009. International Conference, (2009), pp. 1-6.
- [22] A. P. Sample, C. Macomber, J. Liang-Ting and J. R. Smith, "Optical localization of passive UHF RFID tags with integrated LEDs", RFID (RFID), 2012 IEEE International Conference, (2012), pp. 116-123.
- [23] C. Y. Chen, J. P. Yang, G. J. Tseng, Y. H. Wu and R. C. Hwang, "An Indoor positioning technique based on fuzzy logic", MultiConference of Engineers and Computer Scientists, Hong Kong, (2010).
- [24] J. Brusey, C. Floerkemeier, M. Fletcher and M. Lane, "Reasoning about Uncertainty in Location Identification with RFID", Security, (2003).
- [25] K. Bouchard, B. Bouchard and A. Bouzouane, "Guideline to Efficient Smart Home Design for Rapid AI Prototyping: A Case Study", International Conference on Pervasive Technologies Related to Assistive Environments, Crete Island, Greece, (2012).