3D Gesture Recognition Using Passive RFID Tags

Krigslund, Rasmus; Popovski, Petar; Pedersen, Gert Frølund

Published in:
In proceeding; IEEE International Symposium on Antennas and Propagation and USNC-URSI National Radio Science

DOI (link to publication from Publisher):
10.1109/APS.2013.6711812

Publication date:
2013

Document Version
Early version, also known as pre-print

Link to publication from Aalborg University

Citation for published version (APA):
Abstract—This paper investigates the application of passive UHF Radio Frequency IDentification (RFID) tags for recognizing gestures in three dimensions. We experimentally assess the possibility to differentiate among predefined motions with a tag or object. We propose a novel method estimating and tracking the tag orientation in 3D based solely on the physical characteristic of the tag reply, using multiple reader antennas distributed around the interrogation zone. The results demonstrate a good potential, as even with simple data processing differentiation of simple gestures is possible. Our investigation therefore shows that the orientation information is available, and can be extracted using suitable data and signal processing techniques.

I. INTRODUCTION

Radio Frequency IDentification (RFID) technology has been widely used for identification of objects. But recently, research has focused on using passive RFID tags as dedicated sensors, collecting information about the object they are attached to or the environment around them, such as humidity [1] or body monitoring [2].

In this work we propose a novel usage of standard passive RFID tags for orientation sensing in three dimensions, in order to enable recognition of gestures, i.e. predefined motions with the tag or tagged object. This is achieved by illuminating the interrogation zone with multiple reader antennas. Each antenna collects information about the polarization characteristic in the tag reply, from which the orientation is estimated and tracked.

The related work treats RFID motion detection [3], and [4] presents small low power sensor devices for tracking the orientation of human body limbs. However, computing the orientation from noisy velocity/acceleration samples are subject to several sources of error, e.g. integration drift and collision accelerations [5].

In this work, the tag motion information is collected by multiple antennas continuously collecting tag replies and estimating their orientation. We use dual linearly polarized antennas that are capable of to decompose the tag Received Signal Strength (RSS) into horizontal/vertical dimensions and thus obtain the tag polarization. This method was initially utilized for motion capture in two dimensions [6]. The proposed method is investigated and evaluated through experimental evaluation and measurements.

II. PROPOSED METHOD

The tag orientation is determined based on the method from [7]: the reader is interrogating the tag using an antenna with two orthogonal and linearly polarized antenna elements, such that the tag reply is a two-dimensional RSS vector with the horizontal, $H$, and vertical, $V$, RSS components.

To estimate the orientation in three dimensions we use $M$ antennas distributed around the tag. The directions from which they each observe the tag therefore differs significantly, hence the orientation of the tag appears different depending on which antenna is collecting the RSS samples. This is referred to as a projective transformation between coordinate spaces. As an example consider the illustration in Fig. 2. The tag orientation is given by a three dimensional orientation vector, $t_0$, in the reference, world coordinate space. The two orthogonal antenna elements spans the $xz$-plane of the antenna coordinate space, and the $y$-dimension is parallel to the observation direction of the antenna, see Fig. 2. The orientation vector observed by antenna $A_i$, $t_i$ where $i \in \{1, \ldots, M\}$, can then be described as a three dimensional rotation of $t_0$ from the world coordinate space to the coordinate space of $A_i$, followed by a projection to the $xz$-plane $[t_{i,x} \ t_{i,y} \ t_{i,z}]^T = R_i \cdot t_0$. In practice the projection means that $A_i$ only observes $t_{i,x}$ and $t_{i,z}$, i.e. $t_i = [t_{i,x} \ t_{i,z}]^T$. The rotation matrix $R_i$ represents a rotation around three axes, $x$, $y$ and $z$ in that order, according to the three rotation angles $\theta_{x,i}$, $\theta_{y,i}$ and $\theta_{z,i}$. For convenience we write $R_i$ as a function of the three rotation angles, i.e. $R_i(\theta_{x,i}, \theta_{y,i}, \theta_{z,i})$. We assume that $R_i$ is known in advance due to the known position/orientation of all $M$ antennas with respect to the world coordinate space. The rotation from the antenna coordinate space back to the world coordinate space is given by the inverse rotation matrix, $R_i^{-1}$. Since the $y$-component is not obtained, $t_i$ is two dimensional and 3D back rotation cannot be used to estimate $t_0$ from known $R_i$ and $t_i$.

The vector $t_i$ obtained by $A_i$ represents the orientation of the tag from the observed direction. A Kalman filter estimates the tag orientation vector in the world coordinate space. At a given time $n$ the orientation vector element $t_{i,x}$ is $t_{i,x}[n] = t_{i,x}[n-1] + v_{i,x}[n-1] \cdot \Delta$, where $t_{i,x}[n-1]$ is the orientation, and $v_{i,x}[n-1]$ the angular velocity, in the $x$-dimension for the previous time instance. $\Delta$ refers to the time interval from $n-1$ to $n$. The angular velocity is $v_{i,x}[n] = v_{i,x}[n-1] + u_x[n]$, where $u_x[n] \sim \mathcal{N}(0, \sigma_u^2)$ and represents the random changes in the motion velocity. The magnitude of $\sigma_u^2$ depends of the type of motion to be captured,
but it also represents the certainty of the model. For large $\sigma^2$, the filter relies more on the data directly, and less on the data model. We then write the orientation vector to time $n$ as:

$$t_i[n] = t_i[n - 1] + w_i[n - 1] \cdot \Delta$$  \hspace{1cm} (1)

The problem of capturing the motion of a tag can thus be described as updating the estimates of the three dimensional orientation and angular velocity, through Kalman filtering.

III. EXPERIMENTS AND RESULTS

In order to capture the motion of an RFID tag and recognize gestures, we have designed a test scenario with a well defined motion, using a step motor for rotating a tag in a single plane. We use $M = 3$ antennas, placed so their direction of observation is parallel with one of the world coordinate axes: $A_1$ is parallel with the $y$-axis, $A_2$ with the $x$-axis and $A_3$ with the $z$-axis. This gives the following rotation matrices: $R_1(0, 0, 0)$, $R_2(0, 0, 90)$ and $R_3(90, 0, 0)$. The distance to the tag, an Alien 9640 “squiggle” tag [8], is 130 cm for all three antennas. The setup uses an Impinj Speedway Revolution Reader [9], at an interrogation power of 19 dBm. The dual linearly polarized reader antennas are custom made horn antennas, matched to the RFID frequency band around 900 MHz. When a tag is completely aligned to one antenna element, the orthogonal element experiences a 100 % polarization loss. The method for estimating the tag orientation relies on two dimensional RSS samples from each of the dual polarized reader antennas, and we therefore introduce artificial samples whenever an interrogation round for a given antenna element returns empty. The RSS samples are processed with second order Butterworth low pass filter. A circular and a semicircular gesture are used, with rotations in each of the three planes, $xy$, $xz$ and $yz$. The estimated orientation vector for both gestures in the $xy$, $xz$ and $yz$ planes are plotted in 3D, in Fig. 1. We see that the Kalman Filter is able to reconstruct some periodic behaviour, as the estimated orientation vector from the circular gesture forms crude circular shapes, while the semicircular gesture forms lunar-shapes. In order to differentiate between the circular and half circular motions, and enable automatic classification of these gestures we define two heuristic metrics: 1) From the standard deviation of the orientation vector components we are able to correctly identify the plane of motion in the given examples. 2) By calculating the Euclidian distance between two points on the motion trajectory, we are able to determine the correct gesture in five out of six of the given examples. This means that even with the relatively simple methods used in this work for data processing and classification we are able to differentiate between two types of gestures. Hence, the orientation information is available in the physical characteristic of the tag reply, and using more advanced methods the reliability of the classification can be improved.

REFERENCES


