Cluster Comput
DOI 10.1007/s10586-017-0764-0

Recognizing individuals in groups in outdoor environments combining stereo vision, RFID and BLE

D. F. Llorca[1] · R. Quintero[1] · I. Parra[1] · M. A. Sotelo[1]

Received: 10 November 2016 / Accepted: 26 January 2017

Abstract Vision-based people localization systems in outdoor environments can be enhanced by means of radio frequency identification technologies. This combination has the potential to enable a wide range of new applications. When individuals wear a radio frequency tag, they may be both identified and localized. In this way, the technology may interact with individuals in a personalized way. In this paper, two radio frequency identification technologies, UHF Radio Frequency IDentification (RFID) and Bluetooth Low Energy (BLE), are combined with a stereo-based people detection system to recognize individuals in groups in complex outdoor scenarios in medium sized areas up to 20 m. The proposed approach is validated in crosswalks with pedestrians wearing portable RFID passive tags and active BLE beacons.

Keywords Individuals identification · Stereo · RFID · BLE · Data association · Sensor fusion · Outdoors

1 Introduction

The ability for intelligent systems to localize and identify individuals or specific objects is a very important feature that has the potential to enable a wide range of new applications. In this way, the system may interact with individuals in a personalized way or may be capable of getting very detailed information of objects such as vehicles, piled products, luggage, etc., or even animals, which, in parallel, are being localized in a specific scenario. The identification of individuals does not necessarily mean to manage personal information but any variable that may be of interest for the application, such as the type of disability for assistive intelligent transportation systems [11], the access level for access control in surveillance systems, etc. However, this becomes especially challenging in outdoor and dynamic environments that require fast and non-intrusive localization and identification of the users, animals or objects.

Object detection systems based on vision (monocular or stereo), structured light, radar or LIDAR, are capable of accurately determining where objects are in their field of view. However, to identify their specificity (e.g., who they are, what type of disability they have, what are their access level, what type of animal it is, what type of vehicle -make, model, emissions, etc.- it is, what destination has the suitcase, etc.) remains an open research challenge, specially for outdoor applications in medium-size areas.

The use of Radio Frequency-based (RF) wireless communication in combination with vision-, radar- or LIDAR-based object detection approaches may be the solution to provide the system with the ability to both localize and identify individuals or objects in groups. RF identification has achieved a widespread success in various applications ranging from asset tracking, highway toll collection, supply chain management, animal identification, surveillance systems, aerospace, etc. [13,15]. More specifically, passive Ultra High Frequency (UHF) RFID technology has attracted a great attention from both industry and academia due to the fact that a built-in power source in the tag is not needed. The passive tag can communicate with the reader thanks to the use of backscattered coupling from the tag to the reader. Active technologies, such as BLE (or Bluetooth Smart) provides a high communication range at a reduced power consumption and a minimum cost. Although powered active beacons[1] are needed in this

✉ D. F. Llorca
david.fernandezl@uah.es

1 Computer Engineering Department, University of Alcalá, Alcalá de Henares, Madrid, Spain

[1] From now on, BLE active beacons will be named as BLE tags.

Published online: 09 February 2017

case, they are small enough to be used in many different applications. Besides identification, RF-based technologies are capable of providing a rough estimation of the relative position between the tag an the antennas by modeling the Received Signal Strength Indicator (RSSI-based). When at least two (non-isotropic) or three (isotropic) antennas are available, different multilateration techniques can be applied to compute the global position of the tagged objects. However, the accuracy of RSSI-based localization systems is low due to intrinsic limitations and directional dependence when using RSSI as a distance metric [6,17]. When other sensors are available (e.g., vision or LIDAR), different fusion schemes can be applied to improve localization [9,21]. However, if the accuracy of the range measurements given by these sensors is greater than the one provided by RF-based systems, then RSSI-based localization is only performed to solve a data association problem, linking tags with objects, and considering the physical location of the tagged object as the one given by these sensors [7,8,11,12].

In [12] we presented an experimental comparison between UHF RFID and BLE for stereo-based tag association in outdoor scenarios. In this paper, the method presented in [12] is extended by combining both RF technologies to improve the accuracy of the system. A rough estimation of the location of tagged individuals is obtained by means of a RSSI-distance model with parameters that are automatically computed by applying an automatic stereo-RSSI calibration process. A robust data association method based on a global nearest neighbor (GNN) and a new distance metric is presented to deal with complex outdoor scenarios in medium sized areas with a measurement range up to 15 m. Two different methods are presented and analyzed to combine the measurements given by both RFID and BLE systems. The combination of RFID, BLE and stereo to deal with both individual localization and identification in groups in outdoor environments is validated in an intelligent pedestrian crossing scenario (see Fig. 1). A stereo-based pedestrian detection

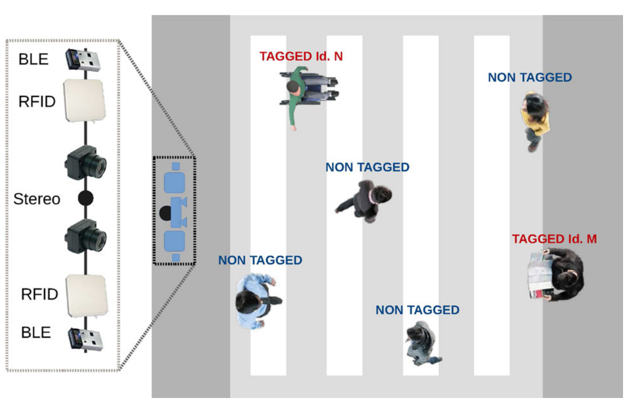


Fig. 1 Pedestrian crossing scenario. Up to two pedestrians are tagged. The system needs to associate the detected tag with the corresponding pedestrian

system [10] provides accurate locations of pedestrians that may wear portable RFID/BLE tags. The most typical scenario involves several pedestrians crossing, but only a few wearing a tag. The infrastructure has to estimate the tagged pedestrian among all the tracked pedestrians to efficiently provide an adaptive response to users with disabilities [11].

2 Related work

Object localization based on radio frequency identification technology has been widely proposed to address numerous different applications [21], including different technologies such as RFID, Ultra-Wide Band (UWB), Bluetooth, BLE, ZigBee, Wi-Fi, etc. [9], and different RSSI- and phase-based localization approaches such as multilateration, Bayesian inference, nearest-neighbor, proximity, etc. [21]. Numerous works have been proposed for the localization of RF tags (objects) with fixed nodes (antennas or adapters), as well as the localization of moving nodes using a fixed set of tags [9]. However, for the course of this work, we have focused on the localization of moving passive/active tags using fixed or moving nodes in combination with vision-based approaches (see Table 1).

In most cases, the combination of wireless sensors and vision-based localization techniques is used to increase global localization accuracy by means of some Bayesian filter (Kalman Filter -KF-, Extended KF -EKF-, Particle Filter -PF-, Unscented Kalman Filter -UKF-, etc.), that fuses the range measurements coming from the different sensors. Thus, in [4], eight directive RFID antennas, and one camera are embedded on a mobile robot to detect passive tags worn on the user's clothes, in indoor environments with a range of 5 m. Saliency maps are obtained for each antenna by counting occurrence frequencies and are translated to the image domain. These maps are used to filter particles on a PF applied over a skin probability image. In [16], RFID-based localization in a small indoor area of interest with a limited number of objects is carried out via RSSI measurements and combined with a camera-based localization system by means of an UKF. There is an obvious improvement in RFID-based localization accuracy thanks to the use of the monocular vision system. The formula between RSSI measurements and distance is adjusted using a manual calibration process. No data association is performed since results are provided with only one object that is directly associated with the detected tag. A similar fusion scheme using a PF to combine RSSI data from passive RFID tags with stereo measurements is proposed in [19]. Four different antennas are used to cover an indoor region of 4×4 m. The RSSI-distance calibration procedure involves manual distance computation, and a linear-regression model is used to obtain the distance from RSSI measurements. Multi-

Table 1 RF- and vision-based approaches for individuals localization and identification

Publication and year	Sensors configuration	Type of scenario and range (m)	RF-localization	# tagged/non-tagged individuals	RF-vision, data association
Miyaki et al. [14]	Wi-Fi, monocular	Outdoor, ≈20 m	RSSI-GPS calibration	1/1	PF RSSI-vision fusion, no data association
Germa et al. [4]	RFID, 8 antennas, monocular	Indoor, ≈5 m	Antenna level frequencies, counting	1/1	PF antennas-vision, fusion, no data association
Isasi et al. [7]	RFID, 3 antennas, monocular	Indoor, room level ≈5–10 m	Tag detection (no localization)	1/1	Vision localization, RFID detection, no data association
Nick et al. [16]	RFID, 1 antenna, monocular	Indoor, ≈3 m	RSSI-distance manual calibration	1/1	UKF RSSI-vision fusion, no data association
Bernabe et al. [3]	RFID, 1 antenna, monocular	Indoor, node level ≈2–4 m	RSSI-monocular automatic calibration	1/1	EIF RSSI-vision fusion, no data association
Cafaro et al. [2]	RFID, 2.5D depth, (Kinect)	Indoor, ≈ 1–2 m	RSSI-monocular automatic calibration	2/2	Probabilistic filtering RSSI-vision, data association (left-right)
Schwiegelshon et al. [19]	RFID, 4 antennas, stereo	Indoor, ≈ 4–6 m	RSSI-distance manual calibration, multilateration	1/1	PF RSSI-stereo fusion no data association
Radaelli et al. [18]	Wi-Fi, monocular	Indoor, room level, ≈4–6 m	RSSI-room supervised calib. trilateration	1/1	PF RSSI-stereo fusion no data association
Goller et al. [5]	RFID, monocular	Indoor, ≈2–4 m	RSSI-distance supervised calib.	1/1	Probabilistic RSSI-vision fusion, HMM data association
Li et al. [8]	RFID, 2.5D depth, (Kinect)	Indoor, ≈5 m	RSSI-phase distance SAR	5/5	SVM RSSI-phase-depth data association (motion needed)
Ours 2016	RFID, BLE, 4 antennas, stereo	Outdoor, ≈15–20 m	RSSI-distance automatic calibration	2/6	GNN RFID–BLE stereo data association, univariate metric

lateration is used to perform RSSI-based localization. Again, no data association is applied since only one object is taken into account. PF is also applied in [14] to fuse Wi-Fi and vision measurements in outdoor scenarios. The so-called fingerprints (SSID and RSSI of different nodes) and a GPS are used to perform RSSI-distance calibration. The GPS is only used for calibration, and its accuracy is limited when no differential corrections are available. RSSI-based localization is conducted using the centroid position for all access points. Data association is not applicable since results are obtained using only one person.

A dynamical RSSI-distance calibration process is proposed in [3] using linear local models around the target, combining RSSI and vision measurements using an Extended Information Filter (EIF) in indoor environments. Although the dynamic RSSI model increases localization accuracy, its use is limited to a one-object one-tag scenario. In real scenarios with multiple targets, perfect data association will be needed. A room-level accuracy system is proposed in [18], by means of a RSSI-room calibration process and a video tracking system that is able to detect an individual entering orleaving a room. Trilateration is then applied to solve the room-level localization problem. Results are provided with only one candidate; therefore no data association process is applied.

As we can observe, and as suggested by [5] and [11], the problem of data association between objects or blobs and tags has been somehow neglected in the literature, which limits the applicability to real scenarios. In [5], a probabilistic framework was proposed to combine RFID and monocular vision measurements for indoor scenarios in a limited range. A pre-defined and manual grid is used to perform RSSI-distance calibration, modeling each grid position with a Gaussian distribution. RSSI-based localization is solved by means of a Mixture of Gaussians, where each mode corresponds to one RFID antenna. A Hidden Markov Model

(HMM) is finally applied to handle the data association problem using a Gaussian distribution as the metric, and finally combining RSSI and vision measurements to compute the person/tag final position. In [2] a hybrid Kinect depth camera and RFID system that uses only one antenna is proposed to determine which of two people are standing on the left and right of an interactive display. An extended version was proposed in [8] using a reverse Synthetic Aperture Technique (SAR) to recognize individuals in groups in indoor scenarios up to 5 m. A Support Vector Machine (SVM) is trained to correlate changes in RSSI and Phase parameters as the tags are moving in space to the motion of the individuals as seen by the depth camera. The system is capable of determining the identity of moving individuals within 4 s and moving groups of five people in 7 s at an accuracy greater than 95%, but limited to indoor scenarios.

As suggested by several studies [6,17], intrinsic limitations exist when using RSSI as a distance metric in terms of accuracy and stability for localization purposes. Thus, as in [7], we propose using the RFID/BLE system as an identification tool (type of disability), and using the vision system (stereo) for localization. In this way, the data fusion problem may become simply a data association problem. A global nearest neighbor algorithm with a novel distance metric is proposed to link radio frequency tags with stereo objects (pedestrians). Our RSSI-distance calibration process is fully automatic. The system was devised for use in outdoor scenarios (crosswalks), in medium-sized areas with a measurement range of up to 20 m, which is a clear contribution to the state of the art. The goal of our previous study [12] was to provide a specific comparison between RFID and BLE technologies in outdoor scenarios. In this paper, we contribute to this topic by developing and analyzing two new methodologies to fuse both technologies in outdoor scenarios.

3 System layout

A global overview of the sensor architecture is depicted in Fig. 2. On the one hand, the stereo platform is composed of two CMOS USB cameras, with two wide angle lenses with

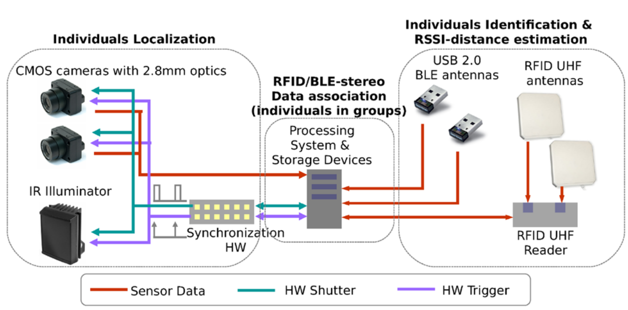


Fig. 2 Global overview of the sensor architecture (Color figure online)

a focal distance of 2.8 mm, with VGA resolution and a baseline of 30 cm, with automatic gain control. An infrared (IR) illumination device Raymax 25 is used for individuals localization in nighttime conditions. It is automatically turned on at night by means of a photocell. A specific synchronization HW controls both the external trigger and the shutter between the cameras and the IR illumination device.

On the other hand, a UHF Class 1 Gen 2 RFID Speedway Revolution R220 reader with two inputs is connected to the PC's Ethernet card. Two far field circularly polarized panel antennas within the 865–870 MHz band (Europe frequency allocation) are connected to the reader. Due to our outdoor scenario range needs, the Onmi-ID Dura 3000 RFID passive tags were selected, which have a theoretical read range of up to 20 m.

Finally, two Trendnet Class I micro Bluetooth 4.0 USB adapters with BLE protocol (2.4 GHz) are directly connected to the PC, which provides a theoretical wireless range up to 100 m at a power consumption of 100 mW. In this case, the active beacons used during our experiments are from Gelo Inc. due to their special features for outdoor scenarios. However, any other models can be used. Note that the synchronization between all sources of information (stereo, RFID and BLE) is carried out by retrieving the PC timestamp.

4 RSSI-based individuals localization

One of the main features of UHF RFID systems is that the passive tags do not actively transmit radio waves. Instead, the passive tags reflect the carrier wave of the RFID reader back to it. This feature attribute involves that the RFID reader can measure the RSSI of the back-scattered signal as well as the phase angle between its transmitted signal and the reflected one. The phase angle reported by the RFID reader can be only used to compute a relative distance due to the fact that the phase angle will rotate 2π radians for every λ wavelength (phase wrapping). Accordingly, in order to compute the absolute distance some initialization method is needed at the beginning. In addition, discontinuities in the phase angle have to be precisely detected to maintain a reliable estimation of the absolute distance. On the other hand, active BLE tags broadcast their identifier as well as information about its signal power in adjustable intervals between 100 and 2000 ms (950 ms by default).

In this way, RFID and BLE technologies are capable of obtaining the RSSI value for each tag. RSSI is represented as a scalar measurement of the tag's RF signal power received by the RFID reader or the BLE adapters. Among the different system parameters affecting the absolute value of the RSSI for a particular tag we have the transmit power, the antennas gain, the carrier wavelength, the tag characteristics, the angle of arrival and the distance [20]. In addition, other environmental factors such as occlusion, multi-path, interference, etc., may also be significant. In our case, besides angle of arrival and distance, the rest of the system parameters are fixed. In outdoor scenarios we can not expect to manage a precise and realistic model capable of taking into account all environmental factors. In addition, we are mainly concerned with RSSI variation with distance. In this way, we propose the use of a simplified form of the relation between distance and received power as in [6], which models the RSSI from one sensor to another as a monotonically decreasing function of their distance, including the main system parameters:

$$P_r(\mathrm{dBm}) = P_{r1}(\mathrm{dBm}) - K \cdot log_{10}(D\,(\mathrm{m})) \tag{1}$$

where P_{r1} is the received power in dBm at 1 m, K is the loss parameter and D is the distance between the receiver and the transmitter. Since accurate readings can be obtained in close proximity because of denser signal coverage, a precise estimation of P_{r1} is possible by computing the average RSSI received at $1m$ distance. However, in our case, both parameters P_{r1} and K are jointly determined by minimizing the root mean square error using calibration data, i. e., RSSI and ground-truth distance measurements.

Thanks to the stereo-based object detection system [10] the calibration data including thousands of RSSI and distance measurements may be automatically obtained. Using a sequence of one individual wearing one tag in a fixed position and orientation, and moving around the stereo region, the stereo-based individual location system [10] can be applied to obtain 3D measurements w.r.t. one reference point (left camera in this case). These measurements may be directly associated with the RSSI values provided by the antennas since data association is not necessary at this stage (one person-one tag). The 3D position of the tag w.r.t. the stereo system is approximated as the center of the blob in the XZ-map, assuming a fixed tag height w.r.t. the road plane. Although this approach provides distance measurements that suffer from both stereo inaccuracies and simplification (due to considering the tag at the center of the blob at a fixed height), its accuracy shall be much greater than that provided by the RSSI-based procedure in [17] or [6], therefore it can be perfectly used as ground truth. In addition, this process automatically provides thousands of measurements in a short period of time, avoiding manual intervention.

Due to implementation limitations, all of the sensors (RFID/BLE antennas and stereo cameras) are located at the same waiting region, integrated in the same stereo baseline (see Fig. 1). Stereo reconstruction provides 3D points P^1_{LC} referenced to the left camera (LC). The relative positions of both the left and the right antennas (LA; RA) w.r.t. the left camera are approximated using an identity rotation matrix and translation vectors containing only the X component. Thus, points P^1_{LA} and P^1_{RA} may be easily computed and associated with their corresponding RSSI values.

After applying the automatic calibration procedure, we obtain the parameters of Eq. (1) and the RSSI-distance curves depicted in Fig. 3 for both RFID and BLE, and the left and right antennas respectively. Furthermore, we compute the exact variance as a function of the RSSI-based distance, which shall be used later on. For a given RSSI value (P_{ri}), we compute the corresponding distance as $D = 10^{(P_{ri}-P_{r1})/-K}$, and we get the associated pre-computed variance σ^2_D. Finally, a Kalman filter is used to receive steadier distance estimations for each tag and antenna. A constant variation model is used. The state vector includes the RSSI value and its variation, whereas the measurement vector is defined by the RSSI value. RSSI variance is computed during the calibration process.

5 Stereo-RSSI data association

5.1 Antenna-level association metric

A single RSSI value yields a sphere with the antenna position at its center and a radius equal to the RSSI-based distance measurement as possible tagged individuals locations. In our case, a fixed and known tag height is assumed to reduce the 3D sphere to a 2D circumference. Then, the tagged individual position may be determined by intersecting the circumferences generated by each antenna. For isotropic antennas with a 360° radiation pattern, a minimum of three antennas are required to compute the tag location. However, in our case, directional 180° antennas are used and one of the intersection points may be discarded. Accordingly, two antennas are sufficient to provide a unique solution.

As suggested by previous works [6,12,17] the intrinsic limitations when using RSSI as a distance metric in terms of accuracy and stability, as well as, in our case, the suboptimal position of both antennas (at the same baseline) results in an intersection point or area (including the uncertainties) that is not a robust and accurate metric to be used for solving the data association problem. Therefore, a new distance metric that models the probability of association between a 3D object (stereo-based) and a detected tag (RSSI-based) has been proposed.

The distance, d^{ij}_k, between a 3D object i and the tag j (assuming fixed height) detected by antenna k ($k = LA_{RFID}$ for left RFID antenna, $k = LA_{BLE}$ for left BLE antenna, $k = RA_{RFID}$ for right RFID antenna and $k = RA_{BLE}$ for right BLE antenna) is modeled using a univariate normal distribution where the mean value is the RSSI-based computed distance d^j_k, the variance is that computed after the RSSI-distance calibration $\sigma^2_{d^j_k}$ and the independent variable is the 3D object position w.r.t. the antenna $d^i_{stereo,k}$:

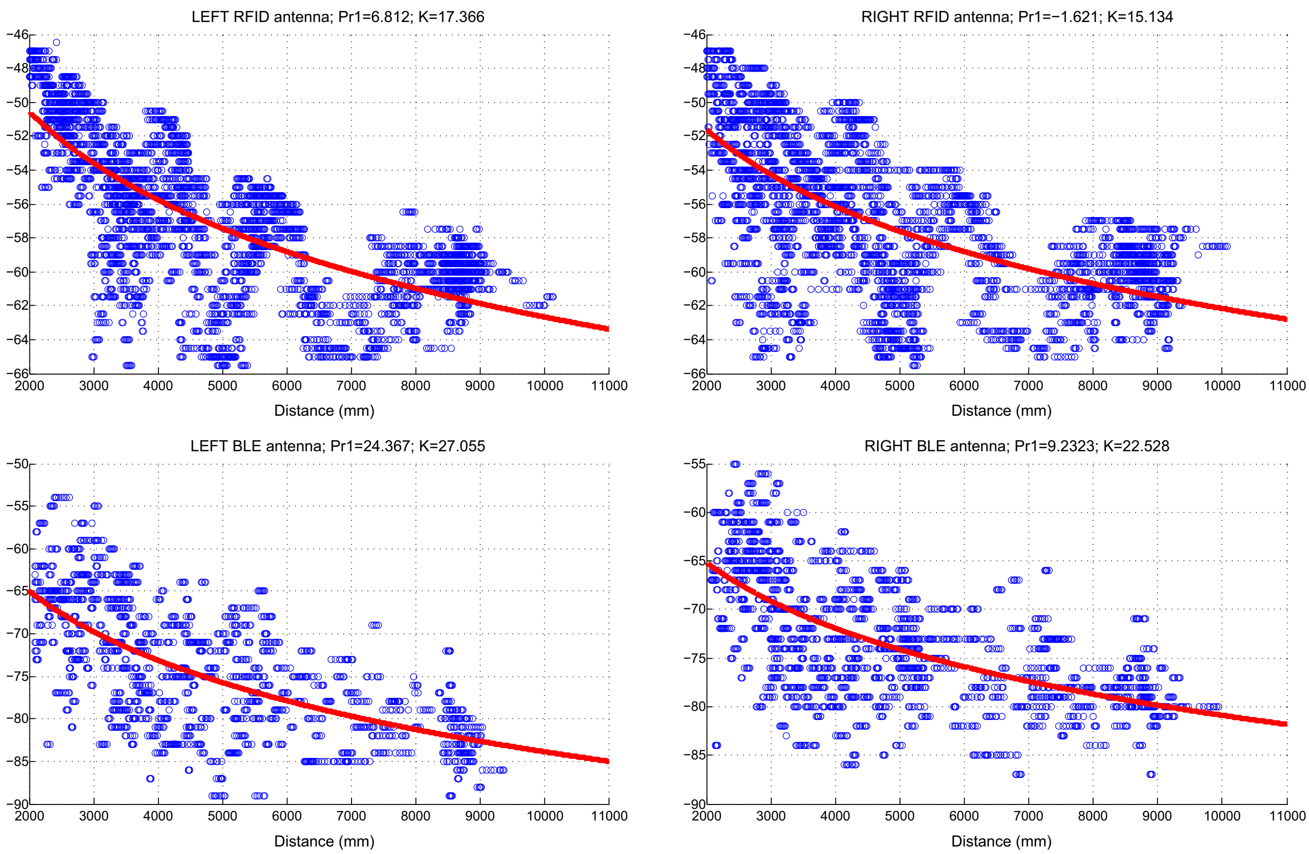


Fig. 3 RSSI-distance model. *Upper row* passive UHF RFID. *Lower row* BLE. *Left* and *right* antennas respectively

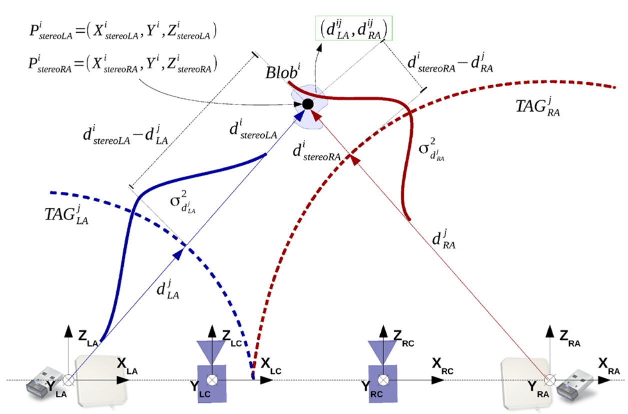


Fig. 4 Graphical representation of the new metric defined between a 3D object and the tag detected by both antennas

$$d_k^{ij} = \frac{1}{\sigma_{d_k^j}\sqrt{2\pi}} e^{-\frac{(d^i_{stereo,k} - d^j_k)^2}{2\sigma^2_{d^j_k}}} \quad (2)$$

A graphical representation of this metric is depicted in Fig. 4. Equation (2) is computed for both RFID and BLE antennas.

5.2 RFID and BLE combination

In [12] we presented an experimental comparison of both RFID and BLE technologies to deal with stereo-RSSI data association. In order to compute the global metric d^{ij} that represented the probability that tag j is being worn by person i, the following equation was applied:

$$d^{ij} = d^{ij}_{LA} . d^{ij}_{RA} \quad (3)$$

Equation (3) was independently used for RFID and BLE systems in [12]. If one of the antennas did not receive a signal, the metric given by Eq. (2) shall be set to zero, and so the final probability given by Eq. (3). We will refer to this approach as a RFID- or BLE-based approaches.

As a substantial contribution, in this paper we propose to combine both RFID and BLE technologies to improve the association capabilities of each one of the technologies by its own. In this case, the tagged individual will be wearing a tag composed of one passive RFID tag and one active BLE beacon, with two identifiers that are considered as a unique identifier. Two methodologies are proposed, so-called **mean-based** and **max-based** approaches. In the **mean-based** approach, the distances d^j_{LA} and d^j_{RA} are computed as follows:

$$d_k^j = \begin{cases} d_{k_{RFID}}^j & \text{if } d_{k_{BLE}}^j = 0 \\ d_{k_{BLE}}^j & \text{if } d_{k_{RFID}}^j = 0 \\ (d_{k_{RFID}}^j + d_{k_{BLE}}^j)/2 & \text{otherwise} \end{cases} \quad (4)$$

being $k = LA$ for left antennas and $k = RA$ for right ones. This approach computes the mean RSSI-distance value of both RFID and BLE antennas. In this case, one antenna (left or right) is considered as a group of two antennas: one RFID and other BLE. When one of the RFID or BLE antennas does not receive any signal, the distance of the mean-based approach will be fixed as the distance provided by the system (RFID or BLE) from which the signal is received. The value computed in Eq. (4) will be then used in Eq. (2), and the final result will be given by Eq. (3).

On the other hand, in the **max-based** approach, we firstly compute the probabilistic distance metric for all the antennas using Eq. (2). Then, the global metric d^{ij} that represented the probability that tag j is being worn by individual i is estimated as follows:

$$d^{ij} = \max(d_{LA_{RFID}}^{ij}, d_{LA_{BLE}}^{ij}) \cdot \max(d_{RA_{RFID}}^{ij}, d_{RA_{BLE}}^{ij}) \quad (5)$$

So in this case, for each pair of antennas (RFID and BLE) we select the maximum association probability between tag j and individual i. The result given by Eq. (5) will substitute, in this case, the one given by Eq. (3).

5.3 Data association

To achieve a reliable data association, a global nearest-neighbor (GNN) [1] algorithm is applied. The association probability between the predicted position of all pedestrians ($i = 1 \dots P$) and all the detected tags ($j = 1 \dots TB$) are computed at each time iteration t. The corresponding probability matrix $C_{P \times TB}$ is defined using the computed distances d^{ij}. Then, the Hungarian or Munkres algorithm is applied so that the global association probability is maximized, as long as the final assignment is always greater than 0.5 (higher thresholds may not be used due to unstable RSSI measurements).

In order to avoid oscillations between the associations, a variable c^{ij} is used for each 3D object i accounting for the number of times it has been associated with tag j. The final association at time t is given by the 3D object i that has the maximum number of associations. When this counter achieves a maximum threshold, the association is fixed until the tag or the 3D object exits the detection area.

Table 2 Description of the sequences, duration and identifier

Id.	Duration (frames)	Sequence description
1	8230	Calibration
2	3270	One tagged pedestrian crossing
3	2710	One tagged/one non-tagged individuals, opposite crossing
4	2380	One tagged/one non-tagged individuals, parallel crossing
5	4740	One tagged/two non-tagged individuals, mixed
6	1270	Two tagged individuals, opposite crossing
7	1250	Two tagged individuals, parallel crossing
8	9180	One tagged/five non-tagged individuals, mixed

Fig. 5 *Left* RFID tags (including Onmi-ID Dura 3000). *Right* BLE beacon from GeLo Inc

6 Experimental results

The stereo-based object detection system has been previously validated in different types of scenarios [10,11] (daytime and nighttime), with an average Detection Rate (DR) of 99% and a False Positive Rate (FPR) of 1.5%. In addition, 90% of the objects detected by the system were tracked in less than 10 frames once they were fully visible (0.33 s running at 30 Hz). Below, results concerning data association between tags and pedestrians are presented.

In order to validate the proposed methodology for localizing tagged individuals, different types of sequences have been recorded in a crosswalk scenario, including different number of people, tags and trajectories (see Table 2). Some users were required to carry one RFID tag and one BLE beacon (see Fig. 5) at a fixed height and pointing to the antennas

Fig. 6 *Left* UHF Class 1 Gen 2 RFID Speedway Revolution R220 reader. *Middle* far field circularly polarized panel antennas within the 865–870 MHz band (Europe frequency allocation). *Right* Trendnet Class I micro Bluetooth 4.0 USB adapter with BLE protocol (2.4 GHz)

Fig. 7 Sensor setup including stereo cameras (baseline of 30 cm), RFID antennas and BLE adapters (baseline of 3 m)

(see Fig. 6). Other users were only required to cross the road as usual. The sensor setup is depicted in Fig. 7.

In order to validate the system performance, the following metrics have been used: percentage of time that the tag is correctly associated to its corresponding tagged individual (CA, Correct Association) and percentage of time a tag has not been associated (NA, Not Associated). We have also computed the percentage of time that the tag is correctly associated or associated to a near individual who is walking or waiting in parallel (CNA, Correct-Near Association) to the tagged one. In addition, we have measured the average association delay (D, Delay), that is, the average number of frames that the system needs to correctly associate each detected tag with its corresponding 3D object. Note that the system is currently running at 30Hz on average, so we can easily convert D to time in seconds.

We provide results corresponding to RFID-based, BLE-based, RFID–BLE mean-based and RFID–BLE max-maxed approaches in Table 3. On the one hand, as detailed in our previous work [12], when using each technology separately, BLE outperforms the results given by RFID in most cases. It is faster and more stable. However RFID reports better CA performance when discriminating between parallel individuals. On the other hand, the combination of RFID and BLE technologies provides better metrics on average in all cases and for both mean and maximum methods than by using each technology separately. The best performance is achieved by means of the RFID–BLE mean-based approach, correctly associating the tag to its corresponding individual (CA) 78.6% of the time, with an average delay (D) of 0.31 s. If tag associations to individuals near the tagged one, waiting or walking in parallel, may be considered to be correct (CNA), this approach correctly associates the tag to its corresponding individual 93.7% of the time. The mean-based approach provides slightly better results than the max-based approach, since the max-based one suffers in cases where the tagged and non-tagged individuals are walking in parallel.

Some examples corresponding to the RFID–BLE mean-based approach are depicted in Fig. 8. The upper row shows the images of the left camera with a color-coded square that represents the associated tag next to the detected individual.

Table 3 Stereo-RSSI data association results.

Seq. Id.	RFID-based				BLE-based				RFID & BLE Mean-based				RFID & BLE Max-based			
	CA (%)	CNA (%)	NA (%)	D fr.	CA (%)	CNA (%)	NA (%)	D fr.	CA (%)	CNA (%)	NA (%)	D fr.	CA (%)	CNA (%)	NA (%)	D fr.
1	99.4	99.4	0.6	22.0	100	100	0.0	0.0	100	100	0.0	0.0	99.8	99.8	0.2	8.5
2	87.5	87.5	12.5	35.0	98.0	98.0	2.0	4.8	99.5	99.5	0.5	1.3	86.5	86.5	13.5	37.9
3	67.0	67.0	33.0	94.0	90.0	90.0	7.2	19.2	91.5	91.5	8.5	27.2	93.8	93.8	5.4	18.6
4	68.3	76.5	23.5	60.0	98.5	1.5	15.0	34.5	46.6	99.0	1.0	18.0	15.0	98.5	1.5	30.0
5	57.1	84.2	14.4	52.0	70.2	86.2	4.6	9.2	73.1	93.9	1.7	3.1	69.0	90.9	1.7	4.5
6	59.6	59.6	35.1	48.8	82.1	82.1	17.9	45.0	91.0	91.0	9.0	22.6	88.9	88.9	9.8	21.8
7	58.1	58.1	41.9	46.9	44.7	83.9	16.1	58.0	59.8	96.8	3.2	1.5	61.9	95.4	4.6	24.8
8	62.6	75.5	8.7	31.0	43.0	79.1	7.0	46.0	58.9	83.0	2.4	2.5	56.3	89.4	1.4	5.5
Avg.	74.8	81.1	14.7	48.7	70.3	89.8	5.4	25.1	78.6	93.7	2.3	9.5	74.9	93.8	3.0	18.9

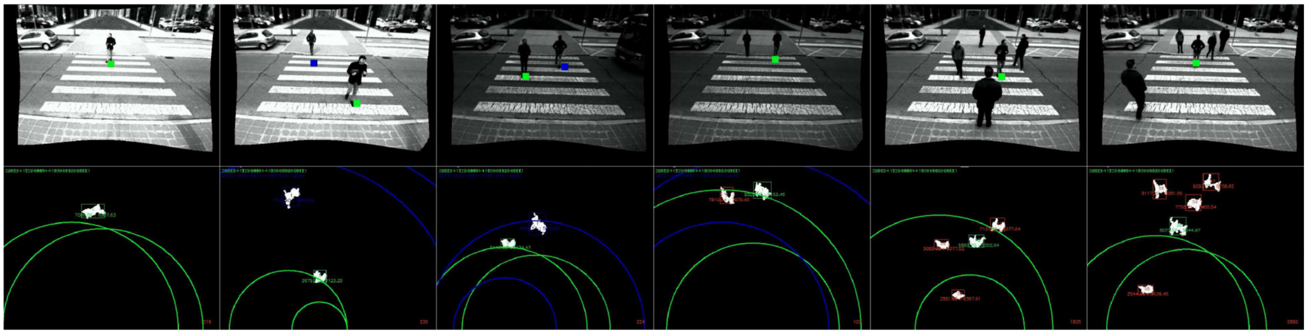


Fig. 8 RFID–BLE meam based examples. *Upper row left image* with color-coded identification (*squares*). *Lower row* XZ-map (*top-view* without road points), detected blobs and RSSI circumferences/curves. Each tag is labeled with a different color (*green* or *blue*) (Color figure online)

The lower row depicts the XZ-map (bird's eye view) without road points, including the detected blobs and the corresponding RSSI circumferences or curves for each antenna. In this case, the radii of the circumferences correspond to the mean value between RFID and BLE (using Eq. (4). Each tag is labeled with a different color (green or blue).

7 Conclusions

In this paper, we have extended the method and the results presented in [12]. For the first time, a combination between RFID and BLE technologies has been presented to recognize individual in groups in outdoor environments up to 20 m. A pedestrian crossing scenario has been selected to validate the proposed approach. It was equipped with a stereo vision system to accurately localize all individuals in the waiting and crossing areas. Some of the individuals carried a small, lightweight tag composed of one passive UHF RFID tag and one active BLE beacon. The sensor setup includes four RF antennas (two RFID including the reader, and two BLE). An automatic RSSI-distance calibration procedure was proposed by combining stereo vision with RFID/BLE identification technologies. Two different approaches has been developed and analyzed to combine the measurements provided by RFID and BLE antennas: mean-based and max-based approaches. Results were obtained in a real crosswalk scenario. The combination of RFID and BLE increases the percentage of time that a tag is correctly associated to the corresponding individual. The mean-based approach reported the best results, correctly associating the tags to their corresponding individuals the 78.6% of the time, with an average delay of 0.31 s. Considering associations to close or parallel non-tagged individuals as correct, RFID–BLE mean-based method correctly associated tags to their corresponding individuals the 93.7%.

Future works shall examine the use of more antennas located as far as possible from one another, so as to improve the association between tags and tagged individuals. The sensitivity of the tag position shall be also analyzed. The use of RF phase-based distance estimation will be explored to improve the tracking of tagged individuals. Finally, extensive validation in more complex scenarios with different lighting conditions shall be carried out.

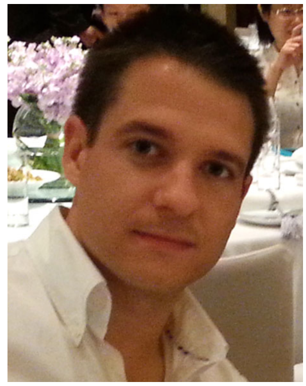
Acknowledgements This work was supported by Research Grants DPI2014-59276-R (Spanish Ministry of Economy), SPIP2015-01737 (General Traffic Division of Spain) and SEGVAUTO-TRIES-CM S2013/MIT-2713 (Community of Madrid).

References

1. Blackman, S., Popoli, R.: Design and Analysis of Modern Tracking Systems. Artech House, Norwood (1999)
2. Cafaro, F., Panella, A., Lyons, L., Roberts, J., Radinsky, J.: I see you there!: developing identity-preserving embodied interaction for museum exhibits. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '13, pp. 1911–1920. ACM (2013)
3. de San Bernabe, A., Martínez-de Dios, J.R., Ollero, A.: Mechanisms for efficient integration of rssi in localization and tracking with wireless camera networks. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (2013)
4. Germa, T., Lerasle, F., Ouadah, N., Cadenat, V.: Vision and rfid data fusion for tracking people in crows by a mobile robot. Comput. Vis. Image Underst. **114**, 641–651 (2010)
5. Goller, M., Feichtenhofer, C., Pinz, A.: Fusing rfid and computer vision for probabilistic tag localization. In: IEEE International Conference on RFID (IEEE RFID) (2014)
6. Heurtefeux, K., Valois, F.: Is RSSI a good choice for localization in wireless sensor network? In: IEEE 26th International Conference on Advanced Information Networking and Applications (AINA) (2012)
7. Isasi, A., Rodriguez, S., Armentia, J.L.D., Villodas, A.: Location, tracking and identification with rfid and vision data fusion. In: 2010 European Workshop on Smart Objects: Systems, Technologies and Applications (RFID Sys Tech) (2010)
8. Li, H., Zhang, P., Moubayed, S.A., Patel, S.N., Sample, A.P.: Id-match: a hybrid computer vision and rfid system for recognizing individuals in groups. In: CHI Interactivity (2016)
9. Liu, H., Darabi, H., Banerjee, P., Liu, J.: Survey of wireless indoor positioning techniques and systems. IEEE Trans. Syst. Man Cybern. C **37**(6), 1067–1080 (2007)

10. Llorca, D.F., Parra, I., Quintero, R., Fernández, C., Izquierdo, R., Sotelo, M.A.: Stereo-based pedestrian detection in crosswalks for pedestrian behavioural modelling assessment. In: ICINCO2014, International Conference on Informatics in Control, Automation and Robotics (2014)
11. Llorca, D.F., Quintero, R., Parra, I., Izquierdo, R., Fernndez, C., Sotelo, M.A.: Assistive pedestrian crossings by means of stereo localization and rfid anonymous disability identification. In: IEEE Intelligent Transportation Systems Conference (ITSC) (2015)
12. Llorca, D.F., Quintero, R., Parra, I., Jiménez, M., Fernández, C., Izquierdo, R., Sotelo, M.A.: Comparison between uhf rfid and ble for stereo-based tag association in outdoor scenarios. In: 6th International Conference on IT Convergence and Security (ICITCS16) (2016)
13. Miles, S.B., Sarma, S.E., Williams, J.R.: RFID Technology and Applications. Cambridge University Press, New York (2008)
14. Miyaki, T., Yamasaki, T., Aizawa, K.: Tracking persons using particle filter fusing visual and wi-fi localizations for widely distributed camera. In: IEEE International Conference on Image Processing (ICIP) (2007)
15. Nath, B., Reynolds, F., Want, R.: Rfid technology and applications. IEEE Pervasive Comput. **1**, 22–24 (2006)
16. Nick, T., Cordes, S., Gotze, J., John, W.: Camera-assisted localization of passive rfid labels. In: International Conference on Indoor Positioning and Indoor Navigation (2012)
17. Parameswaran, A.T., Husa, M.I., Upadhyaya, S.: Is rssi a reliable parameter in sensor localization algorithms an experimental study. In: Field Failure Data Analysis Workshop (2009)
18. Radaelli, L., Moses, Y., Jensen, C.S.: Using cameras to improve wi-fi based indoor positioning. In: Lecture Notes in Computer Science, vol. 8470, pp. 166–183 (2014)
19. Schwiegelshohn, F., Nick, T., Gotze, J.: Localization based on fusion of rfid and stereo image data. In: 10th Workshop on Positioning Navigation and Communication (WPNC) (2013)
20. Impinj. Speedway revolution reader application note. Low level user data support (2013). https://support.impinj.com/hc/en-us/article_attachments/200774268/SR_AN_IPJ_Speedway_Rev_Low_Level_Data_Support_20130911.pdf
21. Zhou, J., Shi, J.: Rfid localization algorithms and applications review. J. Intell. Manuf. **20**(6), 695–707 (2009)

D. F. Llorca received the M.S. and Ph.D. degrees in telecommunications engineering from the University of Alcalá (UAH), Madrid, Spain, in 2003 and 2008 respectively. He currently works as Associate Professor at UAH. He is the author of more than 90 refereed publications in international journals, book chapters and conference proceedings. His research interests are focused on computer vision and intelligent transportation systems. Dr. F.-Llorca is currently an Associate Editor of the IEEE Transactions on Intelligent Transportation Systems and the Journal of Advanced Transportation. He received the IEEE ITSS Outstanding Application Award 2013, the Best Young Researcher Award from the UAH in 2013, the Best PhD Award by the UAH in 2008, the Best Research Award in the domain of Automotive and Vehicle Applications in Spain in 2008, the 3M Foundation Awards under the category of eSafety in 2009, the MSc Thesis Award in eSafety from ADA Lectureship at the Technical University of Madrid in 2004, and the Best Telecommunication Engineering Student Award by IVECO in 2004.



R. Quintero received the M.S. degree in computer science engineering from the University of Alcalá (UAH), Madrid, Spain, in 2009. He is currently a research member at the Computer Engineering Department at the UAH and he is working towards his Ph.D. in pedestrian path, intention and pose prediction. He was the main author of the Best Workshop Paper at IEEE ITSC2015. His research interest include pedestrian detection, computer vision, intelligent transportation systems and intelligent vehicles.

I. Parra received the M.S. and Ph.D. degrees in telecommunications engineering from the University of Alcalá (UAH) in 2005 and 2010, respectively. He is currently working as an Assistant Professor at the Computer Engineering Department, UAH. His research interests include intelligent transportation systems, intelligent vehicles, artificial vision and operating systems. Dr. Parra Alonso received the Master Thesis Award in eSafety from the ADA Lectureship at the Technical University of Madrid, Spain, in 2006, and the 3M Foundation Award under the category of eSafety in 2009.

M. A. Sotelo received the degree in Electrical Engineering in 1996 from the Technical University of Madrid, the Ph.D. degree in Electrical Engineering in 2001 from the University of Alcalá (UAH), Madrid, Spain, and the Master in Business Administration (MBA) from the European Business School in 2008. He is currently a Full Professor at the Department of Computer Engineering at the UAH. His research interests include real-time computer vision and control systems for autonomous and assisted intelligent road vehicles. He is author of more than 200 publications in journals, conferences, and book chapters. He has been recipient of the Best Research Award in the domain of Automotive and Vehicle Applications in Spain in 2002 and 2009, and the 3M Foundation Awards in the category of eSafety in 2004 and 2009. He has served as Project Evaluator, Rapporteur, and Reviewer for the European Commission in the field of ICT for Intelligent Vehicles in FP6 and FP7. He was Director General of Guadalab Science & Technology Park (2011–2012). He is member of the IEEE ITSS Board of Governors and Executive Committee. He served as Editor-in-Chief of the ITSS Newsletter in the period Jan 2013 to Jan 2014. At present, he is Editor-in-Chief of the IEEE Intelligent Transportation Systems Magazine. He was recipient of the 2010 Outstanding Editorial Service Award for the IEEE Transactions on Intelligent Transportation Systems. He has served as General Chair of the 2012 IEEE Intelligent Vehicles Symposium (IV'2012) that was held in Alcalá de Henares (Spain) in June 2012. In 2013, he got the IEEE ITSS Outstanding Application Award.