Assistive Intelligent Transportation Systems: The Need for User Localization and Anonymous Disability Identification


©ISTOCKPHOTO.COM/OLASER

David Fernández-Llorca, Raúl Quintero Mínguez, Ignacio Parra Alonso, Carlos Fernández López, Iván García-Daza, and Miguel Ángel Sotelo
Computer Engineering Department, Polytechnic School, University of Alcalá, Madrid, Spain, E-mail: llorca@aut.uah.es

Cristina Alén Cordero
Signal Theory Department, Polytechnic School, University of Alcalá, Madrid, Spain

Digital Object Identifier 10.1109/MITS.2017.2666579
Date of publication: 20 April 2017

Abstract—The main goal of Assistive Technology (AT) is to ensure the functional independence of disabled individuals. This paper proposes the definition of a new concept of AT within the context of the ITS, Assistive Intelligent Transportation System (AITS), analyzing its intrinsic requirements and providing a set of examples. We demonstrate that AITS must localize users with disabilities and identify their specific type of impairment in order to provide an efficient response, and we propose a specific procedure to guarantee anonymity while identifying the type of disability. Moreover, this new type of AT is illustrated by means of a new assistive intelligent pedestrian crossing application that is capable of localizing pedestrians with disabilities, identifying the specific type of impairment and providing an adaptive response to enhance functional capabilities of impaired pedestrians while crossing. By combining stereo-based object detection with radio-frequency identification technology (RFID and Bluetooth Low Energy), a specific solution to the problem of user localization and anonymous disability identification is proposed. Our approach has been validated in a real crosswalk scenario and it may be extended to other types of AITS, depending on the localization accuracy requirements and the range of operation of the specific application.

I. Introduction

Assistive technology (AT) is usually defined as any item, piece of equipment, software or product system used to increase, maintain or improve the functional capabilities of individuals with disabilities [1]. Different disabilities require different assistive technologies. Some examples of these include wheelchairs, walkers or power lifts for mobility impairments; screen readers or video magnifications for visual impairments; hearing aids and assistive listening for hearing impairments, memory aids and educational software for cognitive impairments, etc. Furthermore, AT may be applied to many different fields including sports, education, computer accessibility, ambient-assisted living [2], etc. In this paper, we propose the extension of the range of AT applications to the field of Intelligent Transportation Systems (ITS), in a so-called Assistive Intelligent Transportation System (AITS). Based on the aforementioned definition for AT and considering the definition of ITS provided by EU Directive 2010/40/EU [3], we offer the following definition for AITS:

- *Assistive Intelligent Transportation Systems* are ITS capable of interacting with users with disabilities and providing an adaptive response to each type of disability, to increase, maintain or improve their functional capabilities while making use of the transportation infrastructure.

This definition may be applied to the entire field of road transport, including intelligent infrastructure and intelligent vehicle applications, as well as traffic and mobility management. For example, an assistive intelligent vehicle could identify the drivers type of disability, providing an adaptive response that may involve fully autonomous navigation or different types of Advanced Driver Assistance Systems (ADAS); an assistive intelligent parking infrastructure would be able to identify a vehicle with an impaired passenger (driver or other) in order to guide the driver to a specific location, or allow the vehicle to park in a special spot; an assistive traffic control system (electronic tolls, dedicated lanes, access control systems, etc.) may be able to automatically permit the entrance of a vehicle with an impaired passenger, or to allow the use of a High Occupancy Vehicle (HOV) lane, etc. In all cases, there are three main requirements to be fulfilled by the AITS:

1) Identification of the specific type of disability: since different disabilities require different responses, the AITS must clearly identify the type of disability. In order to guarantee that no personal information is managed by the AITS, this identification process has to be anonymous.
2) Adaptive response to increase, maintain or improve his/her functional capabilities: this requirement involves a wide range of possible solutions that will be linked to the specific type of AITS.
3) Localization of the user with disabilities: in most cases the AITS would need to have an idea of the relative or global position of the impaired user in order to provide an efficient response that is adapted to his/her needs. The accuracy of the localization will depend on the specific type of application.

In order to gain insight into the needs of the proposed technology, this paper presents a specific type of AITS as an example: assistive intelligent pedestrian crossings. In this case, the pedestrian crossing will be able to localize pedestrians with disabilities, identify the specific type of disability, and provide different adaptive responses depending on the disability. Furthermore, the knowledge of the impaired pedestrians position allows the infrastructure to provide an oriented response to enhance assistance to the disabled pedestrian. Some examples include:

- Adaptive green phase for pedestrians with mobility or cognitive impairments.
- Variable audible messages beyond beeps or ticks, for pedestrians with visual or cognitive impairments. By knowing the position of the disabled pedestrian, the message content of the audible message may be adapted and its volume can be modulated depending on the closer waiting area.
- Variable visual messages for pedestrians with hearing or cognitive impairments. Again, the position of the impaired pedestrian permits the system to adapt the content of the visual message, and to provide visual aid as close as possible to the disabled pedestrian.
- More sophisticated approaches such as in-pavement flashing light systems to guide pedestrians with cognitive or hearing impairments.

However, since the adaptive response of the AITS will largely depend on the specific application and given that our aim is to illustrate the AITS concept, in this paper we are mainly concerned with transversal tasks, so we shall mainly focus on the first two requirements: type of disability identification and impaired pedestrian localization. A specific solution is presented for dealing with anonymous identification of the type of disability, which may be easily extended to other assistive systems, and shall therefore be presented as a general procedure. For this purpose, two wireless technologies are compared: passive Radio Frequency Identification (RFID) and active Bluetooth Low Energy (BLE). Pedestrians with disabilities are only required to wear a portable passive RFID tag or active BLE beacon[1], whereas the infrastructure is equipped with a RFID reader and two RFID antennas or two BLE adapters. Pedestrian (impaired or not) localization is implemented by means of a wide-angle stereo-based pedestrian detection and tracking system which besides the 3-D information, also exploits the fact that cameras are static so the background may be modeled and used to improve segmentation. The stereo-based object detection system allows us to implement a fast and automatic Receive Signal Strength Indicator (RSSI) to a distance (RSSI-distance) calibration procedure that will be used to obtain relative distances between the wireless tags and the antennas. Finally, in order to associate detected tags with detected pedestrians in the scene, a global nearest neighbor algorithm is applied, including a novel and robust distance metric that can handle noisy RSSI measurements.

> Accurate range measurements are still necessary in order to permit the application of safety measurements taken from the infrastructure.

Specifically, the main contributions of this work are:

- We define the new concept of Assistive Intelligent Transportation System (AITS), analyze its intrinsic requirements and provide a set of examples.
- We illustrate this new type of assistive technology by means of a new assistive intelligent pedestrian crossing application, capable of localizing pedestrians with disabilities, identifying the specific type of impairment and providing an adaptive response to enhance functional capabilities of impaired pedestrians while crossing.
- A specific procedure to ensure anonymity while identifying the type of disability is presented.
- A new RSSI-distance calibration procedure is proposed by combining stereo-based object detection with RFID/BLE identification technologies.
- A specific solution to the problem of user localization and anonymous disability identification is proposed by means of a new metric and a general nearest neighbor technique that associates pedestrians detected by the stereo system and the RSSI values given by the radio-frequency tag (passive RFID or active BLE) and at least two antennas. This approach may be extended to other types of AITS, depending on localization accuracy requirements and the range of operation of the specific application.

The remainder of the paper is structured as follows: Section II presents related works on pedestrian detection from the infrastructure and RFID/BLE-vision localization approaches. In Section III, the system layout is summarized. The general procedure proposed to ensure anonymous identification of type of disability is described in Section IV. In Section V, stereo-based pedestrian detection, RFID/BLE localization and stereo-RFID/BLE data association procedures are detailed. Experimental results are presented in Section VI, including a detailed comparison of both wireless technologies. Finally, Section VII offers a discussion, conclusions and future work possibilities.

II. Related Work

A. Infrastructure-Based Pedestrian Detection

Pedestrian detection is a well-known topic in the field of intelligent vehicles. Numerous surveys have been published over the past decade, including both monocular and stereo approaches [4]–[8]. Stereo cues are particularly relevant since they enhance both the region of interest selection [9] and the classification [10] stages, providing more accurate relative distance values than monocular approaches, which are essential for collision avoidance maneuvers such evasive steering or automatic braking [11]–[13]. On the other hand, in the context of infrastructure-based applications such as traffic surveillance, pedestrian-vehicle conflict or collision detection, pedestrian behavior modeling, etc., monocular approaches have been widely used given that the camera is static can be exploited by applying background subtraction, optical flow, motion history images, and other techniques, in order to segment pedestrians [14]–[21].

However, accurate range measurements are still necessary in order to permit the application of safety measurements taken from the infrastructure. Thus, in [22] a multi-sensor network to perceive the intersection environment has been proposed, including 14 laser scanners and 10 cameras. The proposed setup was then used to detect pedestrian intention at intersections [19]. Considering pedestrian detection at crosswalks, one of the first stereo approaches was provided by the SafeWalk commercial system [23]. However, its narrow field of view and limited range only allows the system to be used at pedestrian waiting areas on sidewalks. Thus, a multiple lane crosswalk would require a minimum of two SafeWalk systems for the pedestrian waiting areas and once C-Walk (monocular) for the crosswalk, and still, stereo measurements will only be

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

In most cases, the combination of wireless sensors and vision-based localization techniques is used to increase global localization accuracy.

available at pedestrian waiting areas. The feasibility of using only one stereo platform to monitor both the crossing area and the waiting zones at crosswalks was revealed in [24] where a wide-angle stereo system was used to detect pedestrians and nearby vehicles in a two-lane crosswalk. The system was also capable of providing relevant features regarding the pedestrian's intent to cross or wait. With the same goal of modeling pedestrian behavior at crosswalks, in [25] a 360 degree field of view, a Velodyne laser scanner was used for pedestrian and vehicle detection. However, in this case, manual labeling was used to obtain relevant features, so the approach may not be directly used to perform automatic pedestrian detection.

B. Localization Based on Radio Frequency (RF) and Computer Vision

Object localization based on radio frequency identification technology has been widely proposed to address numerous different applications [26], including different technologies such as RFID, Ultra-Wide Band (UWB), Bluetooth, BLE, ZigBee, Wi-Fi, etc. [27], and different RSSI-based localization approaches such as multilateration, Bayesian inference, nearest-neighbor and proximity [26]. By combining localization with its identification capability existing applications may be enhanced and new ones may be developed. Numerous works have been proposed for the localization of radio-frequency tags (objects) with fixed nodes (antennas or adapters), as well as the localization of moving nodes using a fixed set of tags [27]. 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.

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 [28], 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 [29], 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 [30]. Four different antennas are used to cover an indoor region of 4 × 4 meters. The RSSI-distance calibration procedure involves manual distance computation, and a linear-regression model is used to obtain the distance from RSSI measurements. Multilateration 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 [31] 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 [32] 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 [33], by means of a RSSI-room calibration process and a video tracking system that is able to detect an individual entering or leaving 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 [34] and [35], 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 [34], 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 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.

> We propose using the RFID/BLE system as an identification tool (type of disability), and using the vision system (stereo) for localization.

However, as suggested by several studies [36], [37], intrinsic limitations exist when using RSSI as a distance metric in terms of accuracy and stability for localization purposes. Thus, as in [38], 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 15 m, which is a clear contribution to the state of the art. Our previous study [35] was based on the use of RFID technology. However, BLE has been found to be a more suitable technology for indoor location tracking with respect to accuracy, stability and range [39]. Accordingly, we contribute to this topic by providing a specific comparison between RFID and BLE technologies in outdoor scenarios. Furthermore, a new RSSI-distance directional model is proposed. The presented solution mainly focuses on the type of disability identification in crosswalks, but it may also be extended to other types of scenarios or applications.

III. System Layout

A. Sensor Architecture

A global overview of the sensor architecture is depicted in Fig. 1. On the one hand, the stereo platform is composed of two CMOS USB cameras with VGA resolution and a baseline of 30 cm, with automatic gain control, synchronized with an IR illumination device Raymax 25 controlled by a photocell and having two wide angle optics with a focal length of 2.8 mm. 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 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.

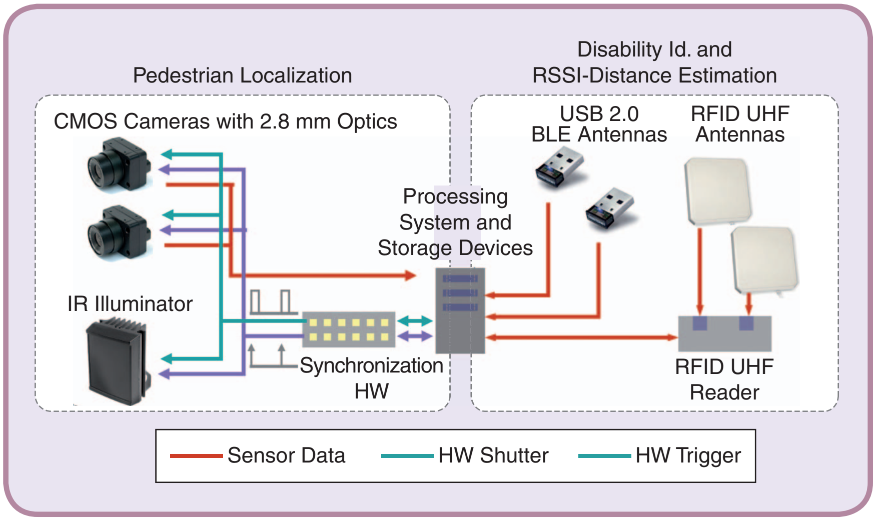


FIG 1 Global overview of the sensor architecture.

B. Scenario Description

In order to validate the proposed methodology to develop assistive intelligent pedestrian crossing systems, a two-lane crosswalk including the pedestrian waiting zones was selected. Note that in order to estimate distance measurements from wireless sensors, more than one antenna is needed. Thus, the sensor architecture depicted in Fig. 1 can be installed in various ways, depending on the location of the wireless antennas. Previous wireless localization studies

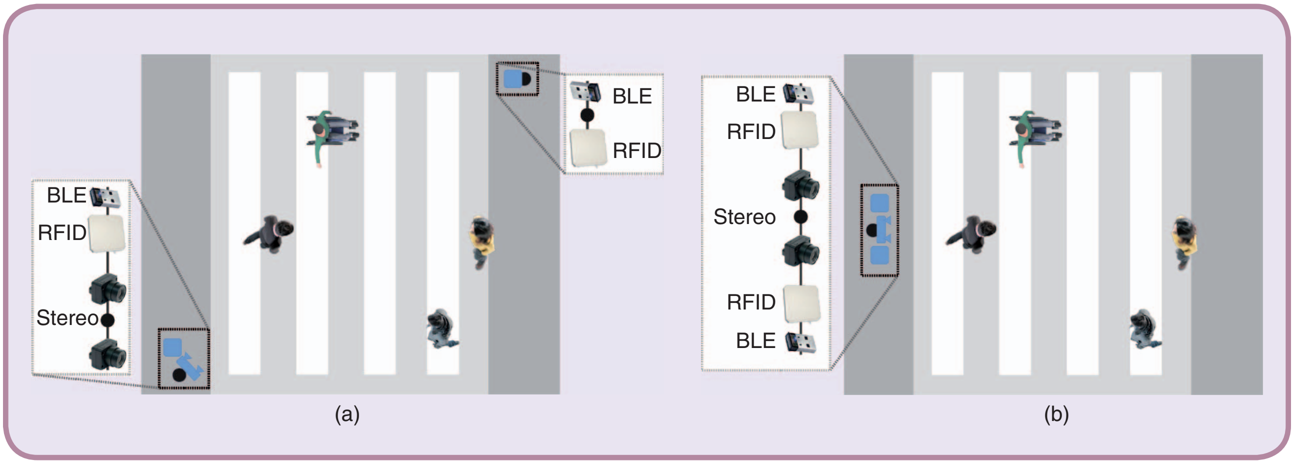


FIG 2 Scenario description. (a) Optimal sensor configuration; (b) Sensor configuration used due to deployment limitations.

[26], [27] suggested that the optimal relative position between the antennas should maximize the distance between them within the region of interest. Such is the case when each antenna is installed at each waiting region, as depicted in Fig. 2(a). However, in our case and due to implementation limitations, we have installed both antennas at the same waiting zone as depicted in Fig. 2(b). The baseline between both antennas (RFID or BLE) has been selected considering the maximum distance permitted by the length of the cable of the antennas.

IV. Anonymous Disability Identification

Although computer vision approaches have evolved dramatically over recent years, it is impossible to consider a potential vision-based solution to recognize different types of disabilities. The automatic recognition of wheelchair users, people with crutches, people with white canes, or even a rough estimation of the pedestrian's age may be possible in the near future. However, there are no visual evidence of individuals having visual (unless carrying a white stick), hearing, or cognitive impairments. Therefore, different type of technology is necessary.

The general approach proposed to maintain anonymity in the type of disability identification process performed by the infrastructure may be described as follows (see Fig. 3):

- The disabled user applies to a local government or administration for a specific disability identification device. The provided device must be easily wearable and inexpensive. A passive RFID tag or an active BLE beacon (with batteries) is proposed for use as disability identification devices.
- The local administration certifies the user's type of disability and requests an RFID tag or BLE beacon from the Central Management Unit (CMU). The CMU may be either publicly or privately managed.
- The CMU writes the code corresponding to the specific type of disability in the writable memory of the RFID tag or the BLE beacon (feature needed), and sends it to the local administration.
- The local government provides the user with the identification device. This device and its written codes does not contain personal information about the user.
- From this point on, the assistive infrastructure may interact with the disabled user, performing fully anonymous disability identification to adapt the infrastructure response.

Note that the proposed procedure has been updated with respect to our previous work [35], avoiding the stage in which all the infrastructure databases have to be updated with the corresponding RFID identifier. Thus, assistive infrastructures do not require remote updating each time a new user applies for his/her identification device.

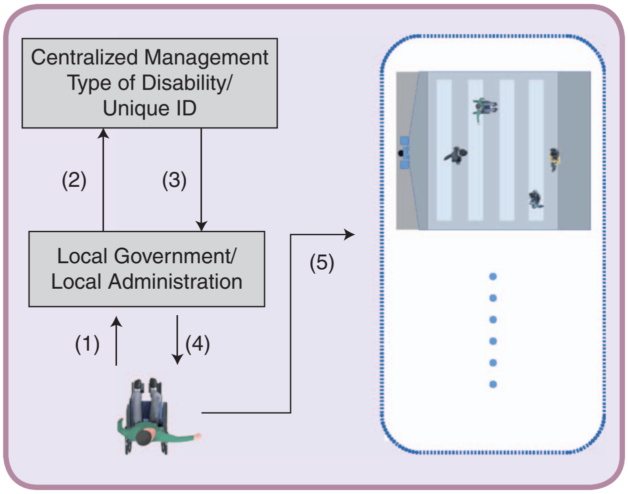


FIG 3 Anonymous disability identification procedure.

V. User Localization

A. Stereo-Based Pedestrian Detection

Our stereo-based pedestrian detection approach has been previously described in [24] and [35]. Here, a short summary is provided (see Fig. 4). Pedestrian localization is carried out using a temporal XZ density map that contains large values in regions with a high density of 3-D points (after removing points outside the range $0.2\,\mathrm{m} < Y < 2\,\mathrm{m}$ w.r.t. the road plane). This map includes information related to moving (pedestrians and vehicles) and static (poles, trees, etc.) objects. Static points are removed using a dynamical background subtraction algorithm [40] to mask the disparity map with the foreground objects. Region-growing is then applied over the masked temporal density map providing a list of potential candidates. Vehicle/pedestrian classification is carried out using features such as object velocity, size, and image location of its first appearance in the scene. In addition, an occlusion reasoning algorithm [21] is used to divide large objects that may correspond to multiple pedestrians. A Kalman-based filter with a constant velocity model is applied to track both pedestrians and vehicles. Data association problems are solved using the Hungarian assignment with a metric that combines the 3-D Mahalanobis distance between blobs and 2-D blob appearance (normalized cross-correlation) [41].

RSSI-based localization automatically provides thousands of measurements in a short period of time, avoiding manual intervention.

B. RSSI-Based Localization

In most RSSI-based localization approaches, the signal strength received from one sensor to another is considered as a monotonically decreasing function of their distance, including the reception and transmission antennas power and their gains. As described in [37], a simplified form of the relation between distance and receive power has been primarily used:

$$P_r(dBm) = P_{r1}(dBm) - K.\log_{10}(D(m)) \qquad (1)$$

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. The values of P_{r1} and K are 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, and considered to be one of the main contributions, the calibration data including thousands of RSSI, distance and angle measurements may be automatically obtained. Using a sequence of one person wearing one tag in a fixed position and orientation, and moving around the stereo region, the stereo-based pedestrian location system can be applied to obtain 3-D measurements w.r.t. one reference point (left camera in our 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 3-D 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 [36], [37], 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.

As discussed in Section III-B, due to implementation limitations, all of the sensors (RFID/BLE antennas and stereo cameras) are located at the same waiting region, integrated

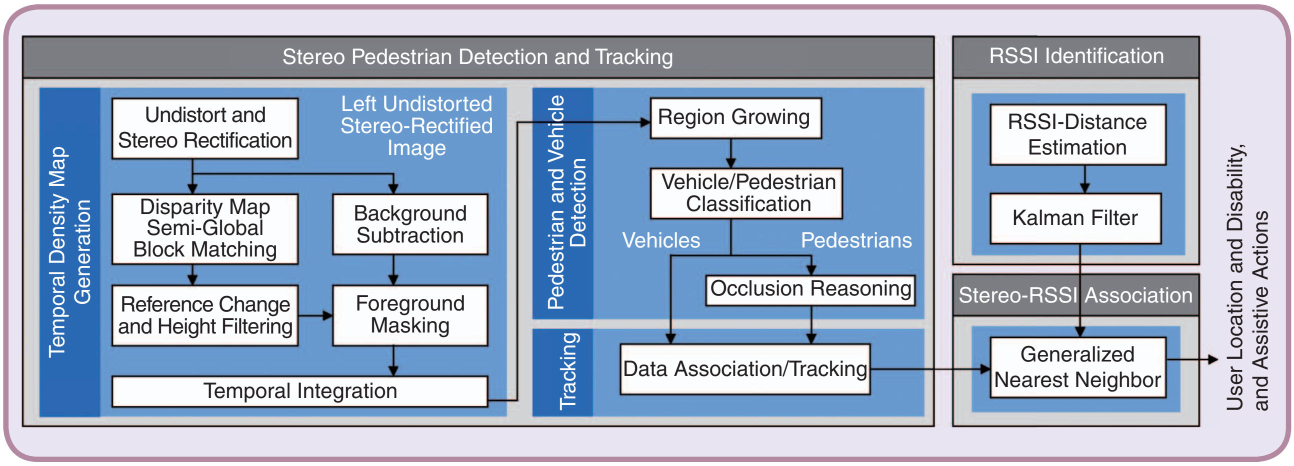


FIG 4 Overall block diagram: stereo-based pedestrian detection, and RSSI identification.

in the same stereo baseline (see Fig. 5). Stereo reconstruction provides 3-D 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. 6 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. We refer to this model as the standard RSSI-distance approach. 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. The possible location of the tag w.r.t. the antenna shall then be defined as a circumference centered at the antenna position with radius D and uncertainty σ^2_D.

However, the standard approach does not consider the directional (angular) dependence of the signal strength between the antenna and the tags. In order to take into account the radiation pattern of the wireless antennas, a more sophisticated model has been proposed, including the angle θ between the antenna and the user, that is, $P_r(dBm) = f(D,\theta)$. Although the signal strength can be considered as a logarithmically decreasing function of its distance, this is not the case w.r.t. the angle. After analyzing the calibration data (see Fig. 7) we concluded that the signal strength linearly decreases w.r.t. the angle, therefore we propose the use of a directional form of the relation between distance and power received as follows:

$$P_r(dBm) = P_{r1}(dBm)K_1.\log_{10}(D(m)) + K_2.\theta \qquad (2)$$

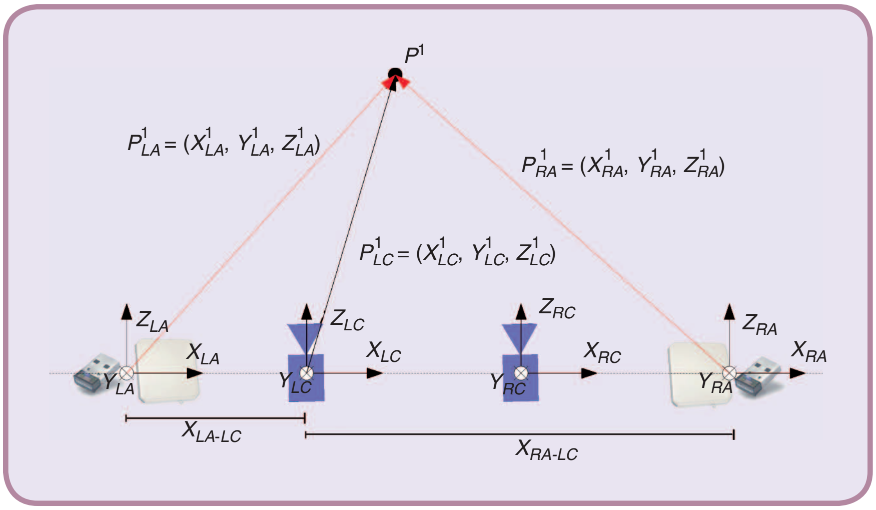


FIG 5 Relative position between cameras and wireless antennas.

where θ is the angle of the relative position between the tag (stereo-based) and its corresponding antenna. Thanks to the automatic calibration procedure, non-linear least squares fitting may be applied over data to obtain the parameters of the directional model (P_{r1}, K_1, K_2). For this case, we compute the variance as a function of both the distance and the angle $\sigma^2_{D,\theta}$. These models and their corresponding parameters are depicted in Fig. 8. Now, for a given RSSI measurement P_{ri} we compute the curve where Eq. (2) intersects the plane $P_r = P_{ri}$, which shall represent the potential location of the tag.

Finally, in both, standard and directional cases, 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.

C. Stereo-RSSI Data Association

In the standard approach (non-directional), 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 tag locations. In our case, a fixed and known tag height is assumed to reduce the 3-D sphere to a 2-D circumference. Then, the tag position may be determined by intersecting the circumferences generated by each antenna. For isotropic antennas with a 360° radiation pattern, a minimum of 3 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. A similar reasoning may be used for the directional case, in which the tag fixed height assumption provides 2-D curves that should intersect at a unique point.

However, as suggested by previous works [36], [37], and as supported by our data (see Figs. 6 and 7), 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 3-D object (stereo-based) and a detected tag (RSSI-based) has been proposed.

The distance, d^{ij}_k, between a 3-D object i and the tag j (assuming fixed height) detected by antenna k ($k = LA$ for left antenna and $k = RA$

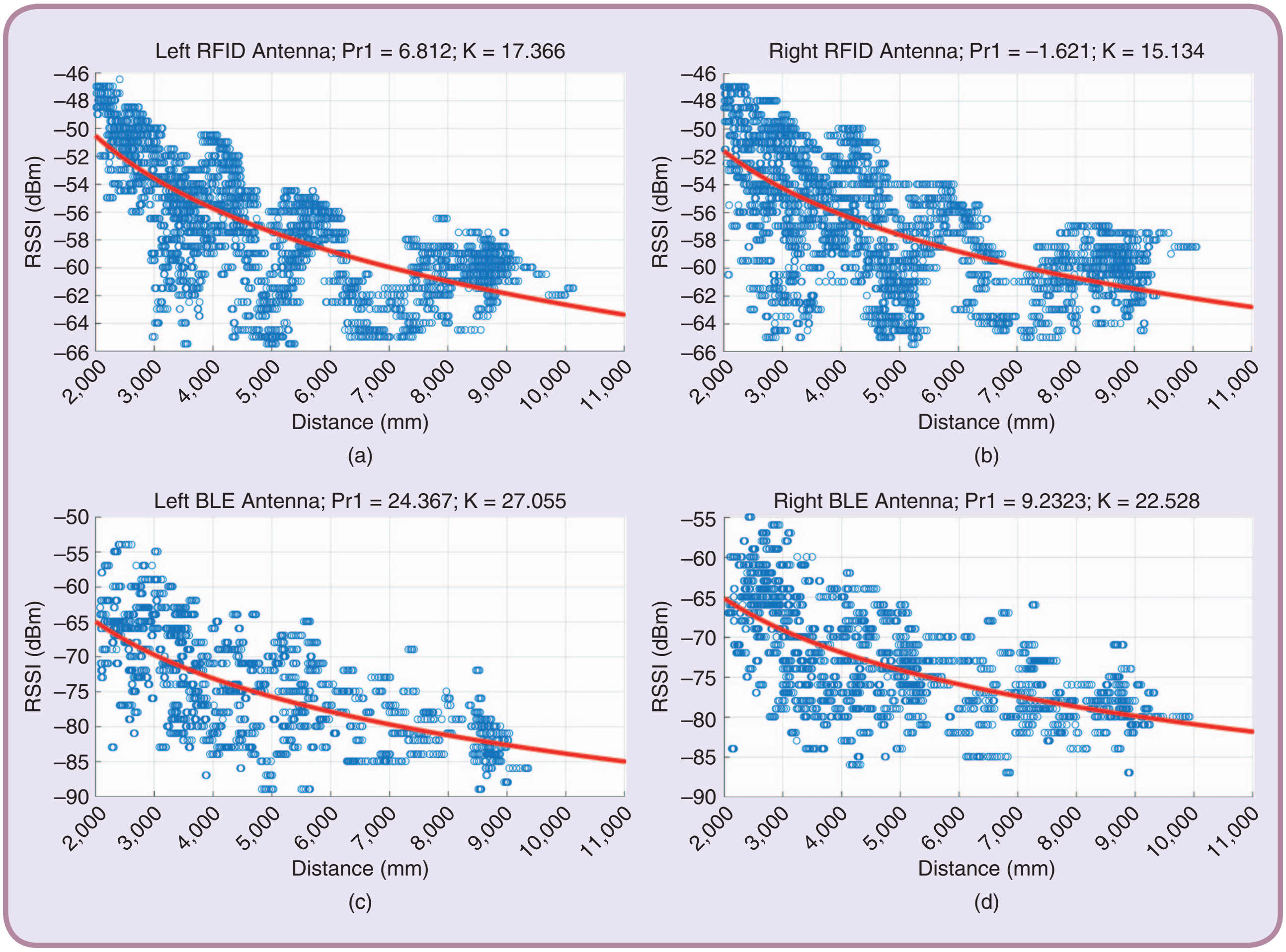


FIG 6 Standard RSSI-distance model. Upper row: RFID. Lower row: BLE. Left and right antennas respectively.

for right antenna) is modeled using a univariate normal distribution where the mean value is the RSSI-based computed distance d_k^j, the variance is that computed after the RSSI-distance calibration $\sigma^2_{d_k^j}$ (standard) or $\sigma^2_{d_k^j,\theta_k^j}$ (directional) and the independent variable is the 3-D object position w.r.t. the antenna $d^i_{\text{stereo},k}$:

$$d_k^{ij} = \frac{1}{\sigma_{D,\theta}\sqrt{2\pi}} e^{-\frac{(d_i^{\text{stereo},k} - d_k^j)^2}{2\sigma^2_{D,\theta}}} \tag{3}$$

where $D = d_k^j$ and $\theta = \theta_k^j$. Note that Eq. (3) may be valid for both the standard and the directional approach, assuming that $\theta = 0$ for the standard model. The graphical representation of this metric is depicted in Fig. 9 for the standard approach. For the directional case, the curves resulting from the intersection between the directional model and the RSSI plane shall be used instead of circumferences.

Eq. (3) is computed for both antennas. If one of these does not receive a signal, the metric shall be set to zero. In order to compute the global metric d_{ij} that represents the probability that tag j is being worn by person i, the following equation would be applied:

$$d^{ij} = d^{ij}_{LA} \cdot d^{ij}_{RA} \tag{4}$$

Eq. (4) can be easily extended to N antennas by applying the following expression:

$$d^{ij} = \prod_{k=1}^{N} d_k^{ij} \tag{5}$$

To achieve a reliable data association, a global nearest-neighbor (GNN) [42] algorithm is applied. The association probability between the predicted position of all pedestrians ($i = 1 \ldots P$) and all the detected tags ($j = 1 \ldots TB$) are computed at each time iteration t. The corresponding probability matrix $C_{P \times TB}$ is defined using the computed distances d_{ij}. The Hungarian or Munkres algorithm is then 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 3-D 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 3-D object i that has the maximum number of associations.

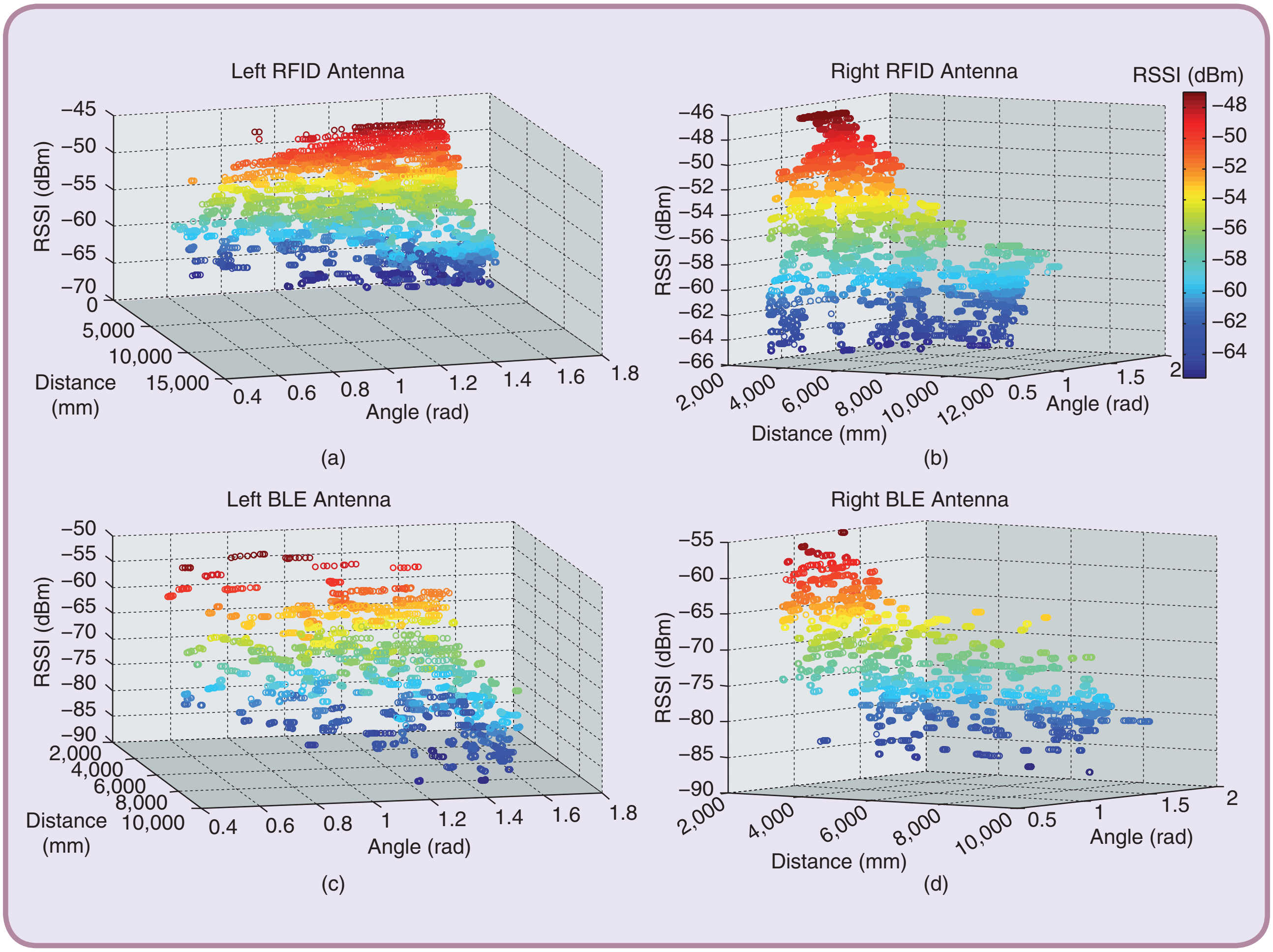


FIG 7 Calibration measurements. Upper row: RFID. Lower row: BLE. Left and right antennas respectively.

When this counter achieves a maximum threshold, the association is fixed until the tag or the 3-D object exits the detection area.

VI. Experimental Results

The stereo-based object detection system has been previously validated in different types of scenarios [24], [35] (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 seconds 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 pedestrians (users with disabilities), different types of sequences have been recorded in a crosswalk scenario, including different number of people, tags and trajectories (see Table 1). Some users were required to carry one tag at a fixed height and pointing to the antennas. Other users were only required to cross the road as usual. 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 pedestrian (CA, Correct Association) and percentage of time a tag has not been associated (NA, Not Associated). Due to the nature of our problem, a tag associated to an incorrect pedestrian for cases in which the pedestrian is very close to the tagged one may be considered to be correct associations since the infrastructure shall still be able to provide an effective response. Accordingly, we have also computed the percentage of time that the tag is correctly associated or associated to a near pedestrian 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 3-D object. Note that the system is currently running at 30 Hz, so we can easily convert D to time in seconds.

We provide results corresponding to the standard approach and the directional one for both RFID and BLE

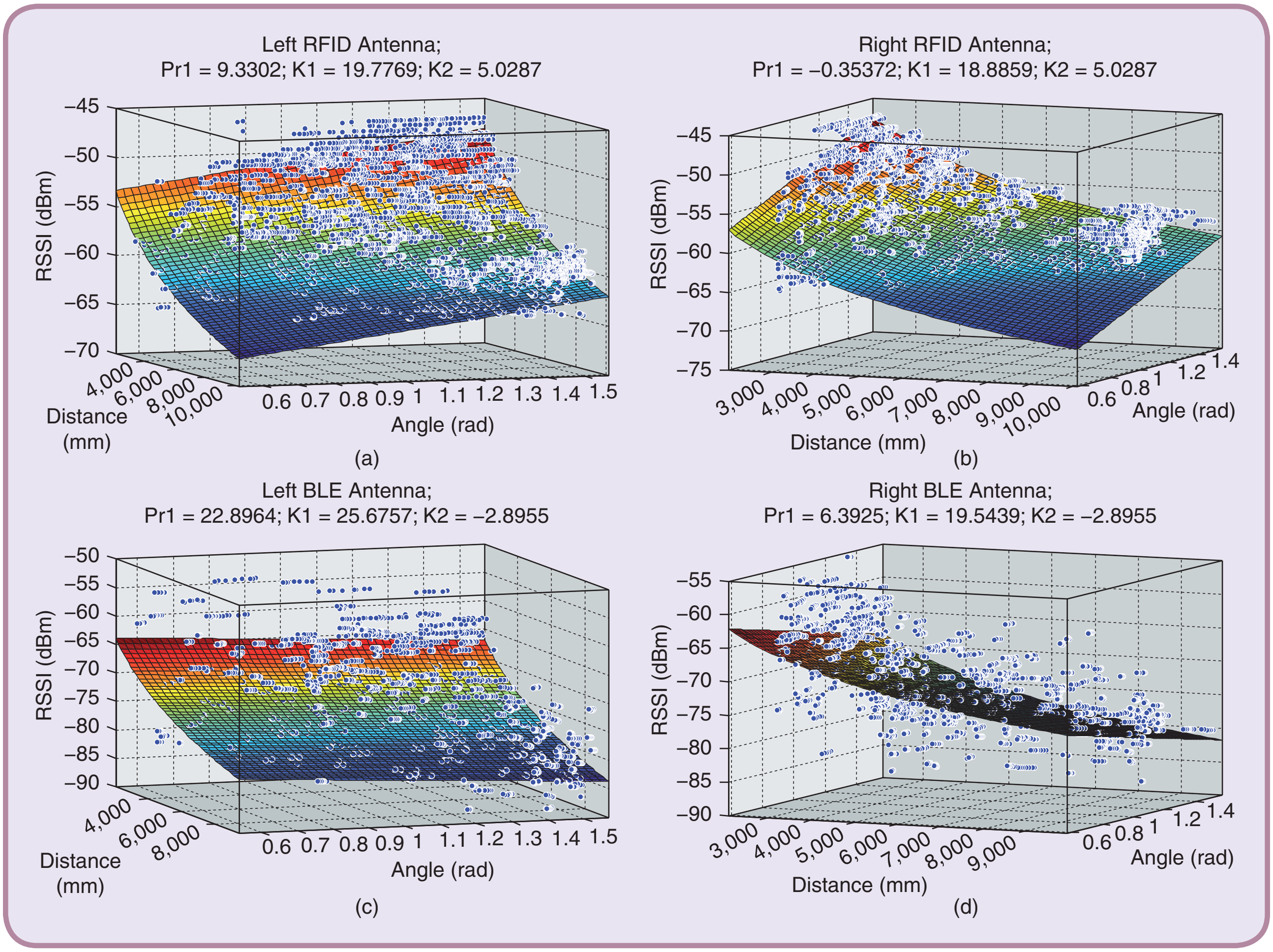


FIG 8 Directional RSSI-distance model. Upper row: RFID. Lower row: BLE. Left and right antennas respectively.

technologies in Table 2. On the one hand, comparing the RSSI-distance model, in most cases, the directional approach clearly outperforms the results provided by the standard one. Thus, *CA* increases by 6.9% (RFID) and 6.1% (BLE) for sequences in which one tagged and one non-tagged pedestrian cross in parallel, 8.4% (RFID) for cases where two tagged pedestrians cross in opposite directions, 5.5% (RFID) and 14.7% (BLE) for sequences where two tagged pedestrians cross in parallel, and 5.7% (RFID) and 4.82% (BLE) in cases where there is one tagged pedestrian among several non-tagged ones in mixed conditions. In addition, the delay considerably decreases for those cases which on average decrease by 5.0 frames (RFID) and 3.9 (BLE)

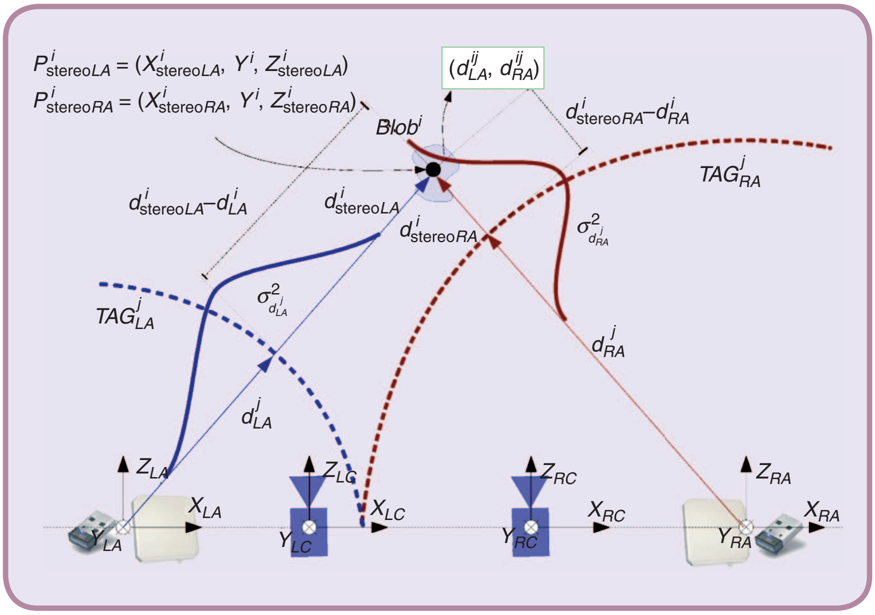


FIG 9 Graphical representation of the new metric defined between a 3-D object and the tag detected by both antennas.

Table 1. Description of the sequences, duration and identifier.

Identifier	Duration (Frames)	Sequence Description
1	8230	Calibration
2	3270	One Tagged Pedestrian Crossing
3	2710	One Tagged/One Non-tagged Pedestrians Opposite Crossing
4	2380	One Tagged/One Non-tagged Pedestrians Parallel Crossing
5	4740	One Tagged/Two Non-tagged Pedestrians Mixed
6	1270	Two Tagged Pedestrians Opposite Crossing
7	1250	Two Tagged Pedestrians Parallel Crossing
8	9180	One Tagged/Five Non-tagged Pedestrians Mixed

frames from the standard model to the directional one. The increase in the *CA* metric is mainly due to the superior performance of the directional model when associating the tag between close pedestrians crossing in parallel. On average *CA* increases by 3.3% (RFID) and 3.2% (BLE) for the directional approach. However, the greater lateral discrimination capacity of the directional model does not involve a considerable increase in the *CNA* metric, which on average is only 0.4% (RFID) and 1.7% better for the directional model than for the standard approach.

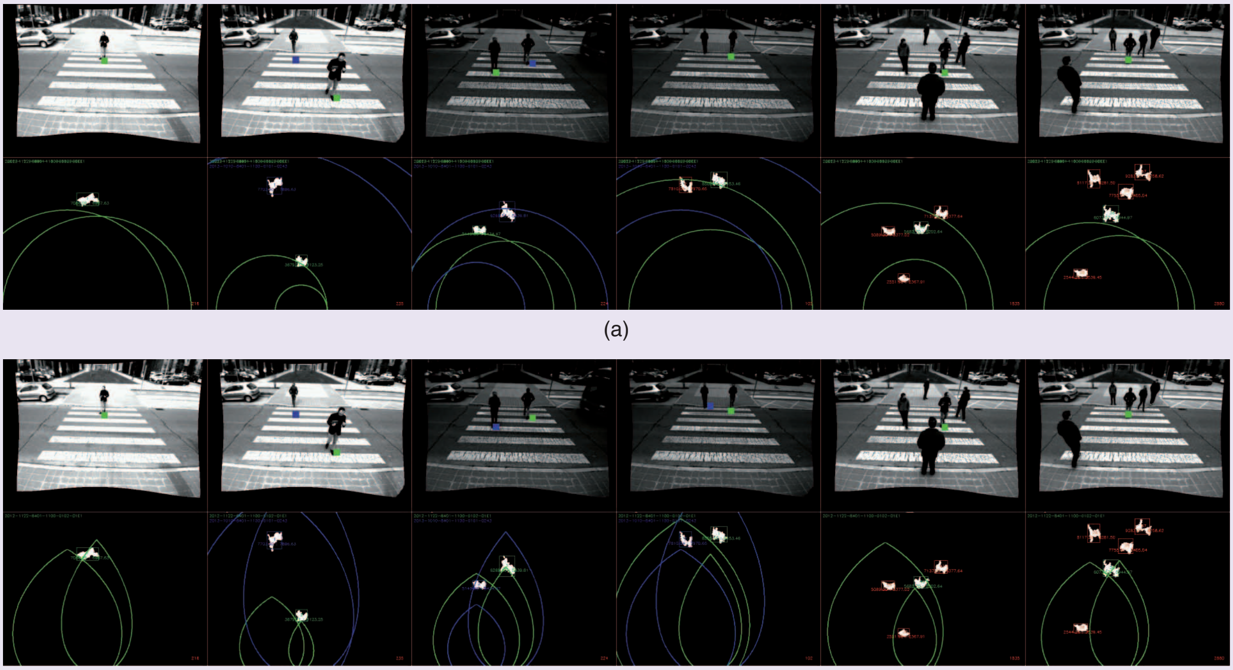
On the other hand, considering the radio frequency identification technology, we observe that RFID clearly outperforms BLE in its lateral discrimination capacity, since *CA* metric is 7.7% (standard) and 4.5% (directional) better for RFID than for BLE. However, when considering parallel and close pedestrians as correct associations, BLE technology provides a considerable improvement in *CNA* metric which increases by 8.3% for the standard model and 10.0% for the directional one, obtaining a $CNA = 91.5\%$ as the best result. This improvement is mainly due to the delay variable, which on average decreases 18.6 frames for the standard model and 22.5 frames for the directional one, and the *NA* metric, which on average decreases by 7.5% for the standard approach and 8.6% for the directional model.

In other words, if the association between the tag and the tagged pedestrian is a critical issue in order to allow the infrastructure to provide an effective assistive action, the RFID directional model provides the best results, correctly associating the tag to its corresponding pedestrian 78% of the time, with an average delay of 1.4 seconds. However, if tag associations to pedestrians near the tagged one, waiting or walking in parallel, may be considered to be correct given that the action taken by the infrastructure shall not suffer from it, then BLE technology with the directional model shall be the best solution, correctly associating the tag to its corresponding pedestrian 91.5% of the time, with an average delay of 0.7 seconds. It is important to highlight the fact that these results were obtained with a suboptimal antenna configuration as described in Section III-B and illustrated in Fig. 2.

Different examples are depicted for both the standard and the directional RFID approach in Figs. 10(a) and 10(b) respectively. The upper row shows the images of the left camera with a color-coded square that represents the associated tag next to the detected pedestrian. The lower row depicts the *XZ*-map (bird's eye view) without road points, including the detected blobs and the corresponding RSSI

Table 2. Stereo-RSSI data association results.

	RFID								BLE							
	Std.				Dir.				Std.				Dir.			
Seq. Id.	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	99.5	99.5	0.5	5.0	100	100	0.0	0.0	100	100	0.0	0.0
2	87.5	87.5	12.5	35.0	87.6	87.6	12.4	34.8	98.0	98.0	2.0	4.8	96.8	96.8	3.2	7.9
3	67.0	67.0	33.0	94.0	67.0	67.0	33.0	93.8	90.0	90.0	7.2	19.2	89.9	89.9	7.9	21.2
4	68.3	76.5	23.5	60.0	75.2	77.7	22.3	57.5	28.4	98.5	1.5	15.0	34.5	98.8	1.2	8.5
5	57.1	84.2	14.4	52.0	59.5	77.8	13.1	44.3	70.2	86.2	4.6	9.2	75.8	87.2	4.8	6.4
6	59.6	59.6	35.1	48.8	68.0	68.0	32.0	43.0	82.1	82.1	17.9	45.0	81.7	81.7	18.3	46.4
7	58.1	58.1	41.9	46.9	63.7	63.7	36.3	31.9	44.7	83.9	16.1	58.0	59.4	85.0	15.0	56.3
8	62.6	75.5	8.7	31.0	68.3	74.6	5.2	39.6	43.0	79.1	7.0	46.0	47.8	85.3	2.5	19.5
Avg.	74.8	81.1	14.7	48.7	78.0	81.5	12.9	43.7	70.3	89.8	5.4	25.1	73.5	91.5	4.3	21.3



(a)

(b)

FIG 10 (a) Standard and (b) directional results. 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).

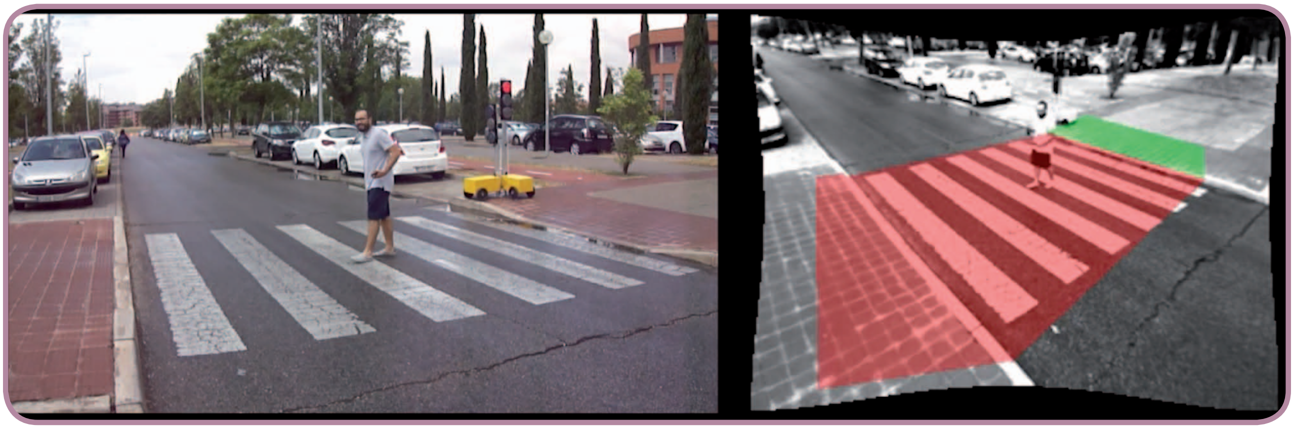


FIG 11 Adaptive green phase for pedestrians with mobility or cognitive disabilities.

circumferences or curves for each antenna depending on the model (standard or directional). Each tag is labeled with a different color (green or blue). It may be observed, in most cases that the RSSI curves of the directional model are closer to the tagged pedestrian than the RSSI circumferences of the standard approach. In the third example the standard approach incorrectly associates each tag for two tagged-pedestrians crossing and in the fourth example the model is unable to correctly associate one of the tags to its corresponding pedestrian.

As described in Section I, once the infrastructure is able to localize the user with disabilities, a set of different actions may be taken to enhance the functional capabilities of the disabled pedestrian while crossing. Some examples are the adaptive green phase for pedestrians with mobility or cognitive impairments (see Fig. 11), variable audible messages, variable visual messages to both pedestrians and drivers (see Fig. 12), etc. Note that more in-depth analysis of the performance and efficiency of these solutions extend beyond of the scope of this paper.

FIG 12 Variable visual messages provided to drivers to alert them about the presence of pedestrians with special needs.

VII. Conclusions

In this paper, for the first time, we have extended the range of applications of the Assistive Technology to the context of Intelligent Transportation Systems. This new concept, Assistive Intelligent Transportation System, involves the need for localizing disabled users and identifying their specific type of impairment. Thus, the AITS shall be able to develop different actions to enhance the functional capabilities of disabled users while interacting with the transportation infrastructure or vehicle. A specific procedure that ensures the individual's anonymity while identifying the type of disability was proposed and a set of distinct AITS examples have been described.

In order to illustrate the needs of AITS, a specific example, an assistive intelligent pedestrian crossing, was developed. The crosswalk was equipped with a stereo vision system to accurately localize all pedestrians in the waiting and crossing areas. Users with special needs shall carry a small, lightweight (passive -RFID- or active -BLE-) tag, that contains the identifier of their type of disability. The infrastructure was equipped with radio-frequency antennas (and a reader in the case of the RFID), and a processing unit shall be responsible for performing stereo-based object detection and RSSI-stereo data association. A new and automatic RSSI-distance calibration procedure was proposed by combining stereo vision with RFID/BLE identification technologies. Two different models, standard and directional, were defined and tested. A specific solution to the problem of user localization and anonymous disability identification was presented based on a new probabilistic metric and a general nearest neighbor technique that associates pedestrians detected by the stereo system and the distance values given by the radio-frequency tag RSSI and at least two antennas. Results were obtained in a real crosswalk scenario. The RFID tags were correctly associated to their corresponding pedestrians the 78% of the time, with an average delay of 1.4 seconds. Considering associations to close or parallel non-tagged pedestrians as correct, the BLE tags were correctly associated to their corresponding pedestrians the 91.5% of the time, with an average delay of 0.7 seconds. A set of assistive examples were presented in the context of adaptive pedestrian crossings. This approach may be easily extended to other types of AITS, depending on the localization accuracy requirements and the range of operation of the specific application.

Future works shall examine the use of more than two antennas located as far as possible from one another, so as to improve the association between tags and tagged pedestrians (users with special needs) or even combine both RFID and BLE technologies. The sensitivity of the tag position shall be also analyzed. Furthermore, user acceptance and other AITS shall be explored in order to extrapolate the insight derived from this experience, attempting to advance in the development of new assistive technologies so as to enhance the functional capabilities of transportation users with disabilities.

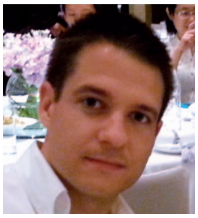
VIII. Acknowledgments

This work was supported by Research Grants DPI2014-59276-R (Spanish Ministry of Economy), SPIP2014-1300 and SPIP2015-01737 (General Traffic Division of Spain) and SEGVAUTO-TRIES-CM S2013/MIT-2713 (Community of Madrid).

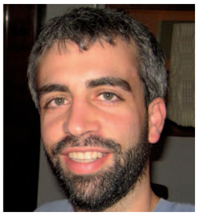
About the Authors

David Fernández-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. 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.



Raúl Quintero Mínguez 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.



Ignacio Parra Alonso 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.

Carlos Fernández López received the B.Sc. degree in computer science engineering in 2008 and the M.S. and Ph.D. degrees in Electrical Engineering in 2010 and 2016 respectively from the University of Alcalá (UAH). He is currently a research member at the Computer Engineering Department at the UAH. He has been research visitor at Griffith University, Australia, in 2011, and at Karlsruhe Institute for Technology, Germany, in 2014. His research interest include scene understanding, intelligent vehicles and intelligent transportation systems.

Iván García-Daza received the M.Sc. degree in Electrical Engineering in 2004 from the University of Alcalá (UAH), Madrid, Spain. He obtained his PhD from UAH in November 2011, which was focused on a drowsy driver detection system. His research interests include computer vision, pattern recognition, machine learning, stochastic processes optimization and control theory. He is currently an Assistant Professor at the Computer Engineering Department at the UAH and he collaborates with the company Escribano in developing high quality control systems and engineering solutions.

Miguel Ángel 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.

Cristina Alén Cordero received M.S. in Industrial Engineering (Mechanics and Machines) and Ph.D. in Industrial Engineering from the Polytechnic University of Madrid in 1996 and 2010 respectively. She is currently a research member at the Theory and Signal Department at the UAH. She has been research visitor at the Laboratorio di Robotica e Meccatronica de la Universita degli Studi di Cassino (Italy) in 2011, and at Kaiserslautern University, Germany, in 2016.Her research interest include safety in design of mechanical structures and analysis of deformation in mechanisms.

References

[1] J. Relton. (2004). The assistive technology act of 2004. *Access World Mag.* [Online] *6(1)*. Available: http://www.afb.org/afbpress/pub.asp?DocID=aw060109

[2] P. Rashidi and A. Mihailidis, "A survey on ambient-assisted living tools for older adults," *IEEE J. Biomed. Health Inform.*, vol. 17, no. 3, pp. 579–590, 2013.

[3] EC. (2010, July 7). Directive 2010/40/ey of the European Parliament and of the council [Online]. Available: eur-lex.europa.eu

[4] T. Gandhi and M. M. Trivedi, "Pedestrian protection systems: Issues, survey, and challenges," *IEEE Trans. Intell. Transport. Syst.*, vol. 8, no. 3, pp. 413–430, 2007.

[5] M. Enzweiler and D. M. Gavrila, "Monocular pedestrian detection: Survey and experiments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 12, pp. 2179–2195, 2009.

[6] D. Gerónimo, A. M. López, A. D. Sappa, and T. Graf, "Survey of pedestrian detection for advanced driver assistance systems," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 7, pp. 1239–1258, 2010.

[7] P. Dollár, C. Wojek, B. Schiele, and P. Perona, "Pedestrian detection: An evaluation of the state of the art," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 4, pp. 743–761, 2012.

[8] D. Gerónimo and A. M. López, *Vision-Based Pedestrian Protection Systems for Intelligent Vehicles*. New York: Springer-Verlag, 2014.

[9] D. F. Llorca, M. A. Sotelo, A. M. Hellín, M. Gavilán, I. G. Daza, and A. G. Lorente, "Stereo regions-of-interest selection for pedestrian protection: a survey," *Transport. Res. C*, vol. 25, pp. 226–237, 2012.

[10] C. G. Keller, M. Enzweiler, M. Rohrbach, D. F. Llorca, C. Schnorr, and D. M. Gavrila, "The benefits of dense stereo for pedestrian detection," *IEEE Trans. Intell. Transport. Syst.*, vol. 12, no. 4, pp. 1096–1106, 2011.

[11] D. F. Llorca, V. Milanés, I. Parra, M. Gavilán, I. G. Daza, J. Pérez, and M. A. Sotelo, "Autonomous pedestrian collision avoidance using a fuzzy steering controller" *IEEE Trans. Intell. Transport. Syst.*, vol. 12, no. 2, pp. 390–401, 2011.

[12] C. G. Keller, T. Dang, H. Fritz, A. Joos, C. Rabe, and D. M. Gavrila, "Active pedestrian safety by automatic braking and evasive steering," *IEEE Trans. Intell. Transport. Syst.*, vol. 12, no. 4, pp. 1292–1304, 2011.

[13] V. Milanes, D. F. Llorca, J. Villagrá, J. Pérez, I. Parra, C. González, and M. A. Sotelo, "Vision-based active safety system for automatic stopping," *Expert Syst. Appl.*, vol. 39, no. 12, pp. 11 234–11 242, 2012.

[14] Y. Malinovskiy, Y.-J. Wu, and Y. H. Wang, "Video-based monitoring of pedestrian movements at signalized intersections," *Transport. Res. Rec.*, vol. 2073, pp. 11–17, 2008.

[15] J. Versavel, "Road user protection via intelligent camera surveillance: Improving pedestrian safety via video image processing technology," in *Proc. 15th World Congress on Intelligent Transport Systems*, 2008.

[16] L. Boudet and S. Midenet, "Pedestrian crossing detection based on evidential fusion of video-sensors," *Transport. Res. C*, vol. 17, no. 5, pp. 484–497, 2009.

[17] K. Ismail, T. Sayed, N. Saunier, and C. Lim, "Automated analysis of pedestrian-vehicle conflicts using video data," *Transport. Res. Rec.*, vol. 2140, pp. 44–54, 2009.

[18] T. Morris, X. Li, V. Morellas, and N. Papanikolopoulos, "Video detection and classification of pedestrian events at roundabouts and crosswalks," Tech. Rep., 2013.

[19] S. Kohler, M. Goldhammer, S. Bauer, S. Zecha, K. Doll, U. Brunsmann, and K. Dietmayer, "Stationary detection of the pedestrian's intention at intersections," *IEEE Intell. Transport. Syst. Mag.*, vol. 5, no. 4, pp. 87–99, 2013.

[20] M. S. Shirazi and B. Morris, "Contextual combination of appearance and motion for intersection videos with vehicles and pedestrians," *Lecture Notes Comput. Sci.*, vol. 8887, pp. 708–717, 2014.

[21] S. Álvarez, D. F. Llorca, and M. A. Sotelo, "Hierarchical camera autocalibration for traffic surveillance systems," *Expert Syst. Appl.*, vol. 41, pp. 1532–1542, 2014.

[22] M. Goldhammer, E. Strigel, D. Meissner, U. Brunsmann, K. Doll, and K. Dietmayer, "Cooperative multi sensor network for traffic safety applications at intersections," in *Proc. 15th Int. IEEE Conf. Intelligent Transportation Systems*, 2012.

[23] W. Favoreel, "Pedestrian sensing for increased traffic safety and efficiency at signalized intersections," in *Proc. 8th IEEE Int. Conf. Advanced Video and Signal-Based Surveillance*, 2011.

[24] D. F. Llorca, I. Parra, R. Quintero, C. Fernández, R. Izquierdo, and M. A. Sotelo, "Stereo-based pedestrian detection in crosswalks for pedestrian behavioural modelling assessment," in *Proc. Int. Conf. Informatics in Control, Automation and Robotics*, 2014.

[25] B. Volz, H. Mielenz, G. Agamennoni, and R. Siegwart, "Feature relevance estimation for learning pedestrian behaviour at crosswalks," in *Proc. 18th Int. IEEE Conf. Intelligent Transportation Systems*, 2015.

[26] J. Zhou and J. Shi, "Rfid localization algorithms and applicationsa review," *J. Intell. Manuf.*, vol. 20, no. 6, pp. 695–707, 2009.

[27] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Trans. Syst. Man Cybern. B*, vol. 37, no. 6, pp. 1067–1080, 2007.

[28] T. Germa, F. Lerasle, N. Ouadah, and V. Cadenat, "Vision and rfid data fusion for tracking people in crows by a mobile robot," *Comput. Vis. Image Underst.*, vol. 114, pp. 641–651, 2010.

[29] T. Nick, S. Cordes, J. Gotze, and W. John, "Camera-assisted localization of passive rfid labels," in *Proc. Int. Conf. Indoor Positioning and Indoor Navigation*, 2012.

[30] F. Schwiegelshohn, T. Nick, and J. Gotze, "Localization based on fusion of rfid and stereo image data," in *Proc. 10th Workshop on Positioning Navigation and Communication*, 2013.

[31] T. Miyaki, T. Yamasaki, and K. Aizawa, "Tracking persons using particle filter fusing visual and wi-fi localizations for widely distributed camera," in *Proc. IEEE Int. Conf. Image Processing*, 2007.

[32] A. de San Bernabe, J. R. M. d. Dios, and A. Ollero, "Mechanisms for efficient integration of rssi in localization and tracking with wireless camera networks," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems*, 2013.

[33] L. Radaelli, Y. Moses, and C. S. Jensen, "Using cameras to improve wi-fi based indoor positioning," *Lecture Notes Comput. Sci.*, vol. 8470, pp. 166–183, 2014.

[34] M. Goller, C. Feichtenhofer, and A. Pinz, "Fusing rfid and computer vision for probabilistic tag localization," in *Proc. IEEE Int. Conf. RFID*, 2014.

[35] D. F. Llorca, R. Quintero, I. Parra, R. Izquierdo, C. Fernndez, and M. A. Sotelo, "Assistive pedestrian crossings by means of stereo localization and rfid anonymous disability identification," in *Proc. IEEE Intelligent Transportation Systems Conf.*, 2015.

[36] A. T. Parameswaran, M. I. Husa, and S. Upadhyaya, "Is rssi a reliable parameter in sensor localization algorithms an experimental study," in *Proc. Field Failure Data Analysis Workshop*, 2009.

[37] K. Heurtefeux and F. Valois, "Is rssi a good choice for localization in wireless sensor network?" in *Proc. IEEE 26th Int. Conf. Advanced Information Networking and Applications*, 2012.

[38] A. Isasi, S. Rodriguez, J. L. D. Armentia, and A. Villodas, "Location, tracking and identification with rfid and vision data fusion," in *Proc. European Workshop on Smart Objects: Systems, Technologies and Applications*, 2010.

[39] X. Zhao, Z. Xiao, and A. M. N. T. Y. Ren, "Does btle measure up against wifi? A comparison of indoor location performance," in *Proc. 20th European Wireless Conf.*, 2014.

[40] S. Álvarez, M. A. Sotelo, D. F. Llorca, R. Quintero, and O. Marcos, "Monocular vision-based target detection on dynamic transport infrastructures," *Lecture Notes Comput. Sci.*, vol. 6927, pp. 576–583, 2012.

[41] I. Parra, D. F. Llorca, M. A. Sotelo, L. M. Bergasa, P. R. de Toro, J. Nuevo, M. Ocana, and M. A. García-Garrido, "Combination of feature extraction methods for svm pedestrian detection," *IEEE Trans. Intell. Transport. Syst.*, vol. 8, no. 2, pp. 292–307, 2007.

[42] S. Blackman and R. Popoli, *Design and Analysis of Modern Tracking Systems*. Norwood, MA: Artech House, 1999.