Supporting Autonomous Navigation of Visually Impaired People for Experiencing Cultural Heritage

Daniele Croce, Giovanni Galioto, Natale Galioto, Domenico Garlisi, Laura Giarré, Federica Inderst, Federica Pascucci and Ilenia Tinnirello

Abstract In this chapter, we present a system for indoor and outdoor localization and navigation to allow the low vision users in experiencing cultural heritage in autonomy. The system is based on the joint utilization of dead-reckoning and computer vision techniques on a smartphone-centric tracking system. The system is explicitly designed for visually impaired people, but it can be easily generalized to other users, and it is built under the assumption that special reference signals, such as colored tapes, painted lines, or tactile paving, are deployed in the environment for guiding visually impaired users along pre-defined paths. Differently from previous works on localization, which are focused only on the utilization of inertial sensors integrated into the smartphones, we exploit the smartphone camera as an additional sensor that, on one side, can help the visually impaired user to identify the paths and, on the other side, can provide direction estimates to the tracking system. The users with the help of the navigation system may experience the museum or the cultural site in auton-

D. Croce · G. Galioto · N. Galioto · I. Tinnirello (✉)
DEIM, Università degli Studi di Palermo, Viale delle Scienze, 9, 90128 Palermo, Italy
e-mail: ilenia.tinnirello@unipa.it

D. Croce
e-mail: daniele.croce@unipa.it

G. Galioto
e-mail: giovanni.galioto@unipa.it

N. Galioto
e-mail: natale.galioto@unipa.it

D. Garlisi
CNIT Consortium, Viale G.P. Usberti, 181/A, 43124 Parma, Italy
e-mail: domenico.garlisi@cnit.it

L. Giarré
DIEF, Università degli Studi di Modena e Reggio Emilia, Via P. Vivarelli, 10, 41125 Modena, Italy
e-mail: laura.giarre@unimore.it

F. Inderst · F. Pascucci
DING, Università degli Studi Roma Tre, Via della Vasca Navale, 79, 00146 Rome, Italy
e-mail: federica.pascucci@uniroma3.it

D. Seychell and A. Dingli (eds.), *Rediscovering Heritage Through Technology*, Studies in Computational Intelligence 859,
https://doi.org/10.1007/978-3-030-36107-5_2

461200_1_En_2_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:9/3/2020 Pages: 22 Layout: T1-Standard

omy, by following a path previously decided and by going from a location of interest to another, without any external personal assistant. We demonstrate the effectiveness of our approach, by means of experimental tests performed in a controlled indoor environment and in a real outdoor installation.

1 Introduction

In the last years we have assisted to many initiatives and experiments for strengthening the social role of museums, from a place mostly dedicated to the storage and exposition of art crafts, to a space of social dialogue and cultural activities, where museum interpretations and representations are often co-created by visitors and museum staff. In such a context, participation of vulnerable groups can be one of the key aspects to promote and reinforce their social inclusion. In this chapter we focus on people affected by visual problems. Visual impairment affects approximately 285 million people worldwide, of which 39 million are blind, and 246 million have moderate to severe visual impairment [1]. Estimates suggest that another person in the world goes blind every 5 s [2]. This is a large number of people who rely on a combination of their other senses—hearing, touch, and even smell—and tools like walking sticks and helpers. Improving museum accessibility for these groups of people is very relevant, especially taking into account that the impact of these disabilities is increasing with the aging of the world population.

Dealing with museum accessibility requires to face two different problems: extending the accessibility of the art experience, by finding innovative ways of perception beyond the sense of the sight; extending the physical accessibility of the spaces in which the art collections are placed. Most of the current research and experimentation efforts are focused on the first aspect. Indeed, people with visual impairment are excluded from directly experiencing the cultural heritage presented by museums in a traditional manner. A common approach for mitigating this problem is participating to the museum visits through audio descriptions and explanations of pieces, whether they are live (or recorded) commentaries provided by museums or provided from friends in real time. However, descriptions provided by others prevent one of the key aspect of art fruition, that is interpreting pieces of art by the users themselves. Innovative ways of perception are considering the exploitation of multimodal descriptions relying on multiple senses. Braille extensions and 3D reconstructions of paintings or building models, for example, have been proposed in many museums and exhibitions. Multimodal interaction allows visually impaired people to access cultural heritage involving large spatial information content, as described in [3] or [4], where the system makes it possible to interact with haptic/acoustic active objects and to select the information that must be shown on the basis of user requirements. But we really think this is only an aspect of the problem. As stated in a recent article appeared online [5]: *people are disabled more by inconsiderate design, inaccessible services, and other people's unaccommodating attitudes than by their own impairment*. Social inclusion of such large portion of population resides in

461200_1_En_2_Chapter ✓ TYPESET ☐ DISK ☐ LE ✓ CP Disp.:9/3/2020 Pages: 22 Layout: T1-Standard

letting them enjoy culture by a personal and autonomous experience, i.e. by allowing a low vision or blind person to enter in a museum, walk, visit and stop in front an artwork without a personal assistant.

The focus of the present chapter is concerning navigation systems for blind users, devised to encourage autonomous visits and improve physical accessibility of muse-

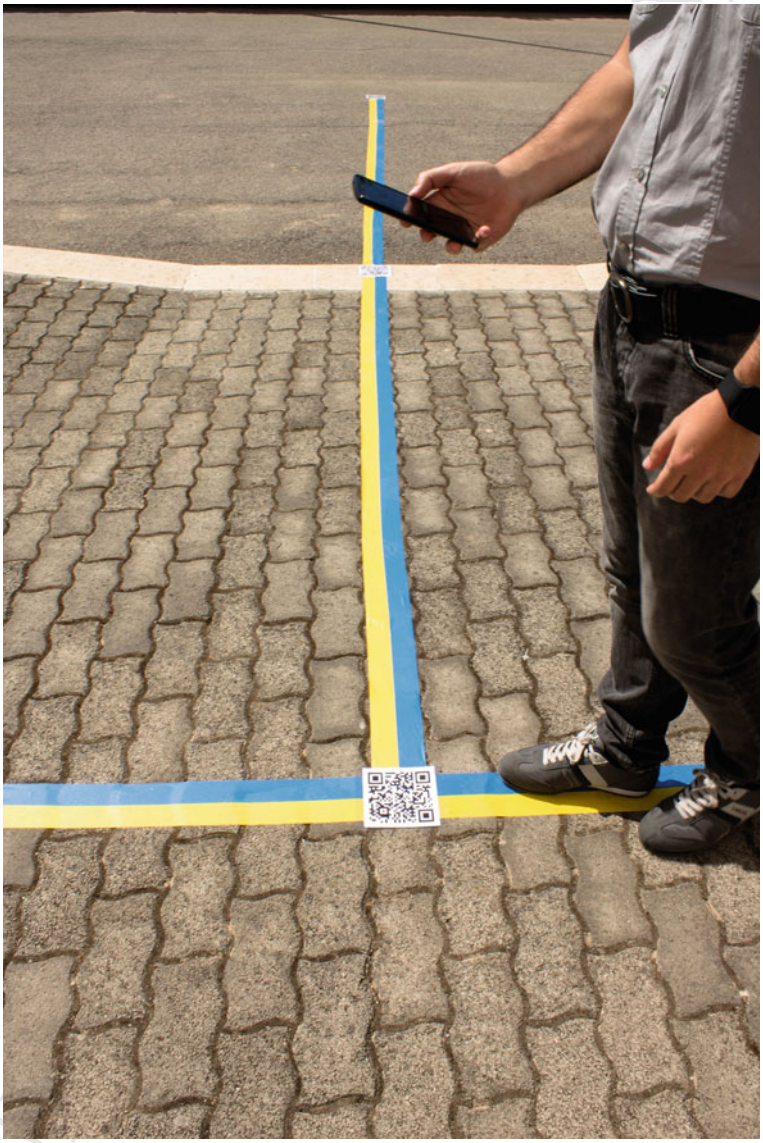

Fig. 1 An exemplary installation of the ARIANNA system

ums. Although indoor navigation systems are of wide interest for many different applications, it is worth noticing that solutions for blind people have stricter requirements than general systems, in terms of accuracy and reliability. For example, visually impaired people are used to a continuous reference signal, such as the one provided by tactile pavings, which guides the users along the path to the destination. This means that the navigation system cannot provide simple information on the directions towards a destination and new interfaces have to be considered. Our solution, called ARIANNA (pAth Recognition for Indoor Assisted NavigatioN with Augmented perception) [6–9], proposes to use the smartphone as a mediation instrument between the users and the reality. The main idea of the system is represented in Fig. 1: the smartphone camera detects the tapes in the environment using computer vision and provides user feedback in terms of vibration signals for identifying the paths. The system has been tested in permanent and temporary exhibitions in various cultural sites, such as the Farm Cultural Park (Favara, Italy), the GAM (modern art gallery) and the Branciforte Palace in Palermo (Italy), Expo 2015 in Milan, as well various other scientific events. The system has been also installed in the Blind institute of Palermo and tested by many low vision users. In the rest of the chapter, after a brief description of other possible solutions, we describe the technical features, localization performance and user experience results of ARIANNA.

2 Navigation Technologies for Blind People

Due to the pervasive availability of ICT technologies for localization, in recent years there was a proliferation of interesting solutions devised to improve the independence and social inclusion of visually impaired people [10]. In most cases, these solutions are focused on the design of specific user interfaces, rather than innovative localization solutions taking into account requirements for blind people. For example, Wayfinder offers an audio interface to a navigation system based on traditional GPS, providing directions and general descriptions about the outdoor places mapped by the application. Other systems working in indoor, where GPS cannot be used, are also based on general-purpose technological solutions, which that can be generally categorized into three approaches: triangulation of RF signals (mainly WiFi), direct sensing of reference points (implemented with RFIDs, ultrasound, bluetooth, etc.), ego-motion estimate provided by dead reckoning algorithms using *Inertial Measurement Units* (IMU, i.e., accelerometers, magnetometers, and gyroscopes). Examples of systems following these approaches and designed for blind people are: e-White, that uses WiFi or Bluetooth signals to coarsely locate the user's position and provide information on the nearest point of interest; RF-PATH-ID and Sesamonet, based on the detection of reference sensors whose presence is signaled by means of alert messages; Navatar [11], based on dead reckoning, where the user can provide references (such as a door in proximity) to improve the estimate of the position in a known map.

461200_1_En_2_Chapter ✓ TYPESET ☐ DISK ☐ LE ✓ CP Disp.:9/3/2020 Pages: 22 Layout: T1-Standard

All these solutions have well known accuracy and practical limitations, which may represent a significant barrier for their adoption in museum itineraries. On one side, distance estimates by means of WiFi signals, which are available in most indoor environments for offering wireless connectivity, are affected by intrinsic positioning errors due to multipath; on the other side, reference points such as RFID or iBeacons may suffer of collisions in case of dense deployments and require a rigid environment instrumentation [12], not suitable for temporary exhibitions and frequent itinerary updates. In some cases, reference points can be represented by WiFi radio maps (also called fingerprinting), which need long and periodic calibration phases [13, 14]. Pedestrian Dead Reckoning (PDR) solutions for tracking user positions between consecutive reference points suffer from drift due to noise [15] on long itineraries, which can be only partially mitigated by filtering techniques [16], exploiting activity recognition [17], or periodic resets at the detection of reference points [12]. In other words, no system is currently able to provide a reliable and continuous reference signal along a museum itinerary, which can be considered equivalent to the tactical pavings.

Assistive technologies may also include other types of services for blind people, such as scene analysis and detection of obstacles. Obviously, these services are completely different from providing orientation signals towards a desired destination. Examples of systems working on obstacle detection are SmartCane [18] and UltraCane/Batcane [19], which integrate sonars and cameras. Other systems, such as the ones described in [20], offer robot-assisted navigation, in which decision-making is not left to the user, but it is rather performed by an artificial intelligence agent. Researchers have also evaluated novel interfaces for enabling such a guidance of a human user. For example, in [21] it is proposed the use of a vibration belt with distinct vibration patterns to communicate directional and rotational commands for navigation.

Our contribution is focused on the guidance service towards a desired destination, without any additional service on obstacle detection. We also propose an innovative interface for providing orientation information and a continuous reference signal along the paths. However, decision making is left to blind users, which are free to plan their visit in the museum itineraries according to their needs and experience.

3 The ARIANNA Navigation System

The ARIANNA navigation system allows autonomous mobility of blind people in public spaces, adopting a solution that is based on ICT (robotic and vision) technologies to mediate between users and environment. The system is especially designed for indoor scenarios, where GPS-based solutions are unavailable, and exploits the visual sensor and vibration signals of commodity smartphones. The system permits to find some points of interests in an indoor environment by following a path painted or sticked on the floor. The path is detected by the camera of the smartphone which also generates a vibration signal providing a feedback to the user for correct-

461200_1_En_2_Chapter ✓ TYPESET ☐ DISK ☐ LE ✓ CP Disp.:9/3/2020 Pages: 22 Layout: T1-Standard

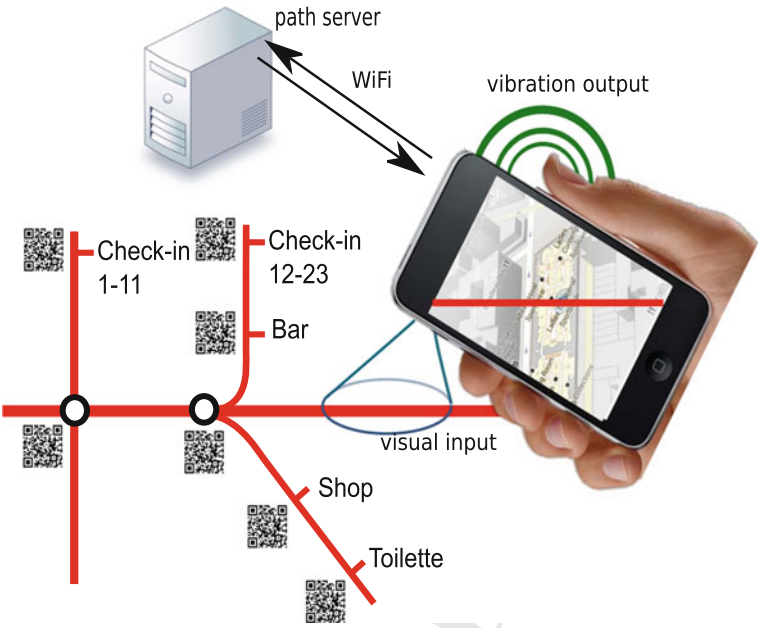

Fig. 2 ARIANNA navigation system description

ing his/her direction. Some special landmarks (e.g. QRcodes or iBeacons) can be deployed along the path for coding additional information detectable by the camera. All the computations and the computer vision algorithms run on the smartphone in real time.

Figure 2 highlights the different components of the system in an airport scenario. The paths of interest are marked with colored lines on the floor. This is the only dedicated instrumentation applied to the environment and is a quite simple and low cost solution; QRcodes are settled close to points of interest, on line intersections and are also used for landmarking. They provide information on the right line to follow in order to get to the desired destination. The user interface employs tactile stimuli to receive feedback on the heading corrections to be employed, as better described in the following. The systems itself is composed by five main components: (A) ambient instrumentation; (B) sensors; (C) data transport network; (D) path server; (E) tactile interface.

Ambient instrumentation. The ambient instrumentation is composed of colored tapes which can be easily sticked on the floor or carpets to define different paths. This is the only dedicated instrumentation applied to the environment. Paths can intersect each other forming a planar graph where intersections are nodes of the graph. To add information on the paths, any segment (the graph edges) may be deployed with two parallel strips with different colors, so the ordered couples (color1, color2) and (color2, color1) encode both direction and orientation. Additionally, using bar codes or QRcodes it is possible to encode relevant information regarding the edges (as

for example the distance from/to the extremes of the segment) and for landmarking. Another possible solution is to deploy iBeacons to provide proximity information close to intersections or points of interests. QRcodes are generally more precise in terms of positioning, although aesthetically more visible.

Sensors. The main sensor used in the ARIANNA system is the camera, which most smartphones on the marketplace are equipped of. The camera is used to reveal the presence of lanes on the floor and acts as a visual control for the haptic transducer. We also use the embedded compass and accelerometer sensors to help maintain or recover the visual control of the line. All these sensors are available on most commodity smartphones: this is a key aspect for keeping the system low-cost and affordable for a vast public.

Data transport network. We assume that a data network connection is available for downloading the ambient map (e.g. via a WiFi or cellular network). The data transport network does not require specific adaptations but is a facility that permits communication between the phone and the ARIANNA server. The server is used to provide localization information, correlation between paths and points of interest, routing towards the destination. The presence of the server and the wireless network is necessary only in case the application is unaware about the building topology and its deployed paths. On the contrary, if the application loaded on the phone has such information locally available, the presence of network and server is optional (even if flexibility is possible only with those elements, as explained below).

The path server. The path server stores and retrieves information from a path repository via the url printed into the QRcode or provided by the iBeacon. The content pointed out by the (fixed) url can be changed on the fly with a simple update on the server. Such flexibility permits path adaptation required by topological changes due to maintenance or load balancing. When the smartphone detects a QRcode/iBeacon on the path, it immediately runs an http request to the server using the detected url. The server knows the position of the user (because of its proximity to the tag position) and sends back to the smartphone the next edge to follow. In fact, among all paths deployed in the building, thanks to the indications provided by the path server, the smartphones provides haptic feedback only towards the "enabled" paths according to the server indication.

Tactile Interface. The tactile interface is a key point of the system. The behavior of the haptic feedback can be summarized as follows: the camera continuously grabs the scene in front of the person. The tracking system incorporates the information on the line (together with the compass and accelerometer data) and provides feedback with the phone vibration. The intensity and type of the vibration is based on the output of the EKF and is designed to keep the camera always in contact with the line or to bring back the visual contact when it is lost. Vibration is a native functionality of the phone obtained through a rotating eccentric mass. It has been shown that the current consumption of typical vibration motors has a limited impact on the battery life of commercial smartphones [22] and that the energy savings coming from switching off the screen are higher than the costs introduced by vibrational cues [23]. Unlike other approaches in haptic interfaces, our solution does not need a selective vibration

of the touched point (that is also difficult to obtain and requires special piezo-electric materials, etc.).

4 Computer Vision Algorithms

The main idea of ARIANNA is exploiting computing vision and vibration signals for replacing the special tiles deployed along the tactile pavings with easy to deploy colored tapes. There are many different computer vision functions that can be combined for the identification of a painted line, taking into account the constraints of our problem: (i) the path identification has to be prompt and reliable, without perceivable latencies for the users, which could correspond to discontinuous signals; (ii) the lifetime of the smartphone battery has to be compatible with the timing required for a complete museum visit, in order to guarantee the practical usage of the system. These constraints correspond to the identification of robust solutions, with limited complexity, able to work in real-time.

There are two main features that can be exploited for detecting the paths: the geometry of the tapes (which in the end are given by piecewise lines), and the colors of the tapes (which combine two different colors for representing a direction without ambiguity). We developed two different solutions, focused on both the geometric factors and the color identification, with different complexity and additional information that can be exploited by ARIANNA, as detailed in the following.

4.1 Geometry-Based Path Identification

An obvious solution for detecting a path is searching lines into the images, by using the well-known Canny algorithm, which is able to identify a set of edge points in an image delimiting areas with large luminance changes. This set of points can be associated to a path whenever they result aligned. Moreover, this kind of analysis allows to detect the slope of the identified path, which can be mapped into an heading measurement of the user movements along the path. In case a map of the paths is known and the user can be positioned (even roughly) on this map, the relative heading of the user can be converted into an absolute orientation.

To identify the line seen by the camera, we implemented three different steps: (i) filtering the image, for reducing the noise and the details of the image background; (ii) applying the Canny algorithm, for detecting the edges of the objects in the image; (iii) identifying the sub-set of edges which can be considered as a line using the Hough transform.

Eliminating image details. The first step is performed by using a Gaussian smoothing filter, whose main goal is defocusing the image for avoiding that some regular patterns of the floor (e.g., such as the edges of squared tiles) can be erroneously considered as a path trace. Since the lines deployed on the floor are very thick in

comparison with the tiles' edges, such a filtering operation does not affect the identification of the line edges. The filter is characterized by a parameter σ which represents the standard deviation of the Gaussian function used for smoothing the image details. Higher values of σ lead to a more evident loss of image details.

Detecting edges. The second step is given by the application of the well-known Canny scheme. The output is a binary matrix, whose dimension is equal to the original frame and whose values are set to 1 for the pixels corresponding to the detected edges. These pixels are identified by computing the maximum luminosity gradient (in each possible direction) for each pixel, and by selecting the pixels for which the gradient is higher than a threshold T. Higher values of T correspond to a lower number of detected edges.

Detecting lines and slopes. The last step works on the binary image found by the Canny scheme for transforming the line identification problem in a maximum search problem. The Hough transform is used for mapping each edge point in a set of parameters (ρ, θ) representing the bundle of lines passing through that point. When multiple edge points are aligned, there is a common (ρ, θ) value representing the line passing through all the points. Therefore, the scheme simply works by counting the maximum number of occurrences (i.e., votes) of quantized (ρ, θ) values computed for each edge point. The quantization allows compensating noise effects in the identification of the alignments. This step is critically affected by the resolution used for quantization and by the threshold used for assuming that a given (ρ, θ) value corresponds to a line in the image. We chose to only consider the global maximum in the parameter space, i.e., a single (ρ, θ) value receiving the greatest number of votes, because we assume that the greatest number of aligned points always correspond to the thickly painted line. However, since the painted line is actually implemented with a bi-colored stripe, such a choice implies that we are only able to detect one of three parallel linear edges along the stripe. In proximity of a turning point, such as L-like junction points, the number of path segments with a different orientation is at least equal to two. However, by always choosing the line with the highest number of votes, which usually correspond to the longest detected line, we avoid ambiguity in the reference orientation for the user.

The above scheme can be optimized by considering the effects of the image resolution, the smoothing factor of the Gaussian filter, and the threshold values of the Canny schemes on the accuracy of line detection. Figure 3 visualizes the effects of different smoothing factors on a real image of a path trace. The standard deviation is normalized as a function of the image resolution (for representing the portion of the area involved in the filtering operations). As the standard deviation increases, the image loses details as evident in the reduction of the number of edges identified by the Canny algorithm. This, in turns, corresponds to a lower number of detected lines (from 9 lines for $\sigma = 0.25$ to 3 lines for $\sigma \in [0.5, 0.625]$), which are likely to coincide with the path trace, rather than with the floor regular patterns. Image resolutions have a similar effect on line detection accuracy: too high resolutions (apart from complexity issues) include too many details, which can correspond to the detection of wrong lines; too low resolutions imply rough estimates of the line orientation. We also noticed that the threshold of the Canny scheme has a limited

461200_1_En_2_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:9/3/2020 Pages: 22 Layout: T1-Standard

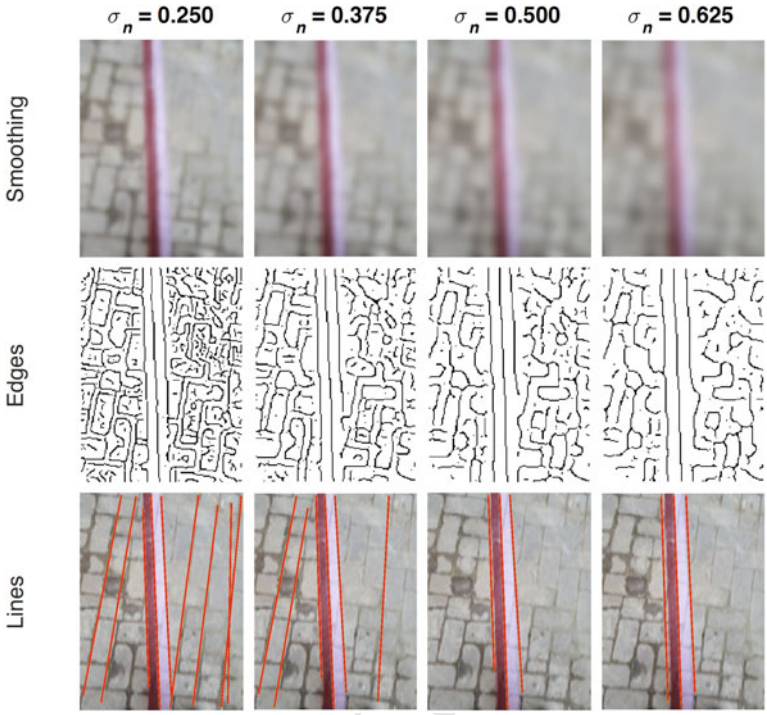

Fig. 3 Effects of the Gaussian smoothing filter on the capability of detecting the correct path trace

impact on the error metric. In the next experiments discussed in the paper, we used the optimal configurations found on the basis of these considerations.

4.2 Color-Based Path Identification

Another possible solution for identifying the path is to detect areas with the expected tape colors into the image. Image colors can have different representations, among which the HSV (hue, saturation, value) representation is more robust to luminance changes. In principle this approach can be quite simple, because it works by applying conditional statements on each image pixel, in order to verify if it belongs to the desired intervals of hue, saturation and values. The complexity required by this filtering operation is much lower than the complexity required by the previous approach, which requires many operations on the image matrices. Moreover, the direction information can be easily extracted from the blob of points identified as belonging to the path, by evaluating the second-order moments of the blob.

Despite of these considerations, the approach based on color searches has the drawback of being potentially affected by the light conditions in which the camera

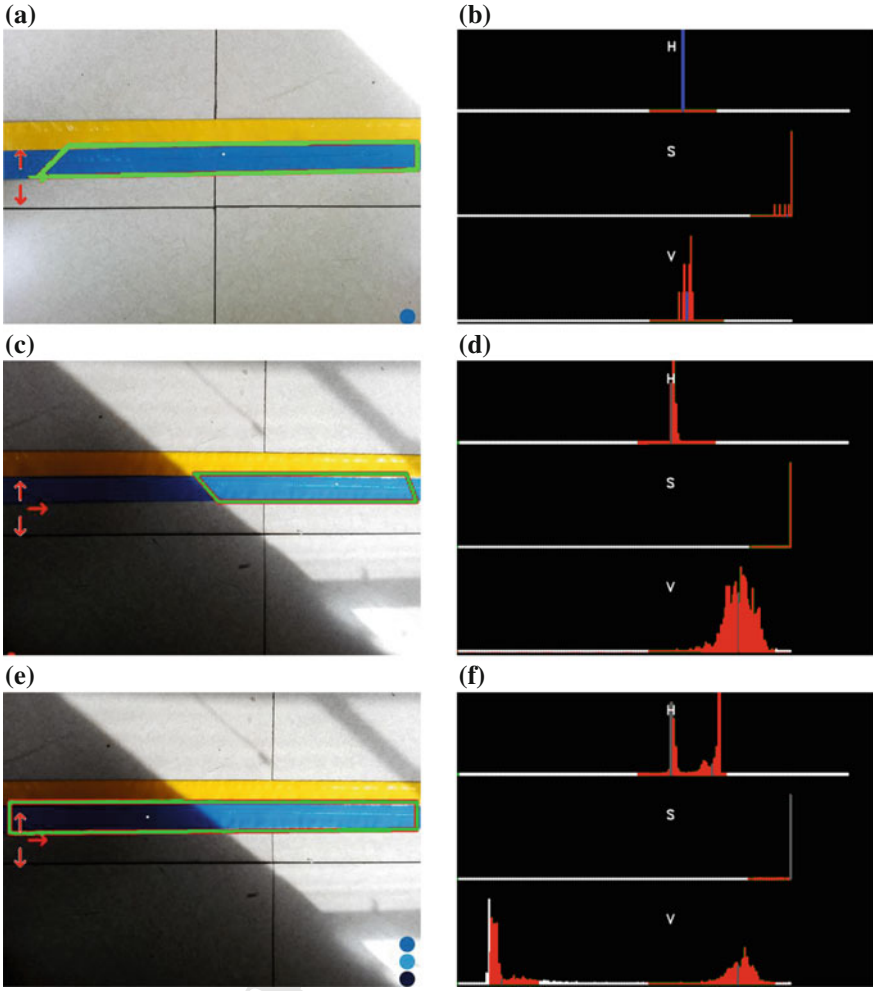

Fig. 4 Effects of the dynamic resampling of HSV filtering values

works. Indeed, it is not obvious how to define absolute HSV intervals for identifying the path under variable (even on a day-basis) light conditions. Schemes removing shadows on the images or correcting luminance changes can be more complex than edge-based solutions. We therefore implemented a mixed algorithm, in which the two schemes are opportunistically combined: ARIANNA normally works by applying a simple color-based detection, but at regular time intervals the geometry-based scheme is executed for delimiting the tapes and re-sampling the color intervals.

Figure 4 shows an example of dynamic resampling of the HSV values used by the color-based path identification approach. In the left part of the figure, we can see the bi-colored tape under three different light conditions. A red delimiter is used for segmenting the image area recognized as belonging to one color of the path. In the

right part of the figure, we also show the HSV distributions of the points belonging to the segmented area. In the top case, when the light conditions are almost uniform, we can see that the HSV distributions are very narrow and color recognition by means of HSV filtering works well. In the middle case, we see that the HSV filtering allows to recognize only one part of the path, not covered by the shadowed area. Only after the execution of the geometric-based path identification function, the whole path captured in the image is correctly identified and the HSV values can be resampled, leading to the bi-modal distributions shown in the bottom case of the figure. These distributions can be used for applying the color filters to the subsequent image frames, until the user moves completely to the shadowed area. At this point, the next execution of the geometric-based identification function will update the HSV values to new single-modal distributions typical of the shadowed image.

5 Extending ARIANNA with IMU-Based Tracking

A possible question that could arise from the description of ARIANNA is understanding what happens whenever the user loses the path. If the camera is no more able to capture the colored tape, in some cases it is enough to sweep the smartphone around for finding it again. However, we cannot exclude that users accidentally abandon the path. In this case, our idea is exploiting the inertial sensors of the smartphone, i.e. the IMU systems, for supporting PDR solutions, able to localize the user in the space and provide information for bringing him/her back to the path. Basically, the IMU systems are utilized at regular time intervals for providing a direction and velocity measurement, by reading the measurements provided by the gyroscope and accelerometers of the smartphone. It is also important to estimate the user activity, in order to stop the measurement acquisition whenever the user stops walking. Indeed, when the user stands in front of an artwork, for example experiencing a tactile fruition of a 3D model, IMU measurements could be very noisy and lead to positioning errors.

PDR solutions can be executed also when the camera correctly captures the colored tapes. In this case, the measurements provided by the IMU system can be aggregated to the heading information provided by the computer vision algorithms, in order to improve the accuracy of the estimated user position along the path.

5.1 Activity Recognition

The *Activity Recognition* is devoted to classify the human motion. This step is crucial, since according to the output of this subsystem, different models are applied to track the displacement of the user. This step is also difficult to be performed when the number of motions to be identified is high, however, in this set up only a limited number of activities are required.

Most of the cultural sites, indeed, are organized in planar environments connected by stairs and/or elevators. The user walks in this environment and stops to appreciate

461200_1_En_2_Chapter ✓ TYPESET ☐ DISK ☐ LE ✓ CP Disp.:9/3/2020 Pages: 22 Layout: T1-Standard

the artworks. Using markers to identify the floor or the area, the user moves in a two-dimensional space and only 2 activities, i.e., *standing still* and *walking* need to be recognized. These two activities can be easily detected by exploiting the accelerations recorded by the IMU, as proposed in [24–26].

5.2 Heading Estimation

The *Heading Estimation* aims at computing the direction of the user when visiting the cultural sites. It is related only to the *walking* activity since when the user stops is supposed standing still in front of an artwork.

The heading estimate is calculated with respect to the fixed Cartesian reference frame (i.e., the *Navigation Frame*, NF) and it is performed using data from gyroscopes and accelerometer merged with the information provided by the camera in a two-step procedure:

- The *Attitude Estimation*;
- The *Heading Refinement*.

The *Attitude Estimation* exploits the measurements collected from gyroscopes, accelerometers and magnetometers to provide the attitude of the smartphone reference frame (i.e., the Body frame) with respect to the NF. The attitude is computed as quaternions and an Extended Kalman Filter is applied to merge the data from the different sensors as detailed in [26, 27]. Once the quaternions are updated, both the rotation matrix $\mathbf{R}_b^n$ from the Body frame to the NF and the heading γ_k with its uncertainty Γ_k can be easily retrieved. The initial condition q_0 is obtained from the acceleration and the data provided by the camera considering the user standing still when the system is activated as in [26].

Data from gyroscopes and accelerometers are affected by bias that downgrade the accuracy of the estimate. To reduce this effect, the bias is recalculated when the *standing still* activity is detected. Furthermore, the *Heading Refinement* uses the data from the camera to bound the drift error. Specifically, the only the heading γ_k and the related uncertainty Γ_k feed the correction step and the following simple update is performed

$$\phi_k = (1 - W_k)\gamma_k + W_k \gamma_{C,k} \tag{1}$$

where W_k is a gain computed according to EKF equation as

$$W_k = \frac{\Gamma_k}{\Gamma_k + R}$$

where R is the uncertainty related to the heading measurement $\gamma_{C,k}$ performed by processing data from the camera and is supposed known and time-invariant.

461200_1_En_2_Chapter ✓ TYPESET DISK LE ✓ CP Disp.:9/3/2020 Pages: 22 Layout: T1-Standard

5.3 Position Estimation

The *Position Estimation* computes the position (x, y) of the user with respect to the *Navigation Frame* NF. To this end, the rotation matrix $\mathbf{R}_b^n$ provided by the Heading Estimation is used to project the accelerometer data in the the NF: in this way the acceleration along the z-axis lies on the sagittal plane of the user, perpendicular to the floor. This signal is fundamental during the *walking* activity, since it is analyzed to to identify the step event. Specifically, the local minima and the local maxima of the vertical acceleration signal is retrieved to perform both peak detection and zero crossing detection. The sharp changes to the vertical acceleration associated with the heel strike. These features are also exploited to compute the cardinality c_i of the set of samples to be processed to cope with different walking speed. During *walking* activity, the displacement l_i is estimated as proposed in [28].

The initial position is supposed known while the position (i.e., the position of the first marker met by the user when approaching an artwork) of the user during *walking* is recursively computed, by estimating the length of the stride on step event detection i

$$\begin{bmatrix} x_i \\ y_i \end{bmatrix} = \begin{bmatrix} x_{i-1} \\ y_{i-1} \end{bmatrix} + l_i \begin{bmatrix} \sin \bar{\phi}_i \\ \cos \bar{\phi}_i \end{bmatrix} \tag{2}$$

where l_i is the displacement of the user between two step events with respect to the NF and $\bar{\phi}_i$ is the average heading in the same time interval. Finally, the user displacement is $l_i = \begin{bmatrix} 0 & 0 \end{bmatrix}^T$ when the output of the classification phase is the *standing still*.

6 Experimental Validation

To evaluate the performance of the proposed system, several experiments have been carried out. Specifically, two different types of trials have been considered. The first set aims at assessing the performance of the system comparing the tracking results with a ground truth. The second set is devoted to prove the effectiveness of the approach in a real scenario. In both cases a Samsung Galaxy S6 (SM-G920F), running Android 6.0.1 is used: it is equipped with an IMU-MPU6500 by Invensense and an IMX240 camera by Sony. Data from IMU are available at sampling frequency 100 Hz, while the images from the camera are acquired at 20 Hz.

In Table 1, the parameters used to initialize Tracking System (TS) during the experiments are shown.

Table 1 Parameters initialization

Param init value	
β	0.48
α_x	0.02
α_y	0.01
α_z	0.03

461200_1_En_2_Chapter ✓ TYPESET ☐ DISK ☐ LE ✓ CP Disp.:9/3/2020 Pages: 22 Layout: T1-Standard

The parameter β is related to the human gait and has been calibrated on the user adopting the procedure introduced in [27]; the covariances associated to the accelerations $\{\alpha_x, \alpha_y, \alpha_z\}$ are used to discriminate between different human activities and are computed at the beginning of the experiment when the user is supposed standing still for 10 s.

To assess the performance of the system, the optical motion capture system OptiTrack has been used to compute the ground truth (GT). The system for motion capture exploits 10 infrared cameras to detect the position of markers in a limited area: it reach the accuracy of 10^{-4} m. To build the GT, the smartphone has been equipped with 4 markers: the CoM of the markers corresponds to the CoM of the smartphone. The accuracy of the proposed system is evaluated according to the estimated error on checkpoints. Specifically, the considered key performance indicator is represented by the Euclidean distance between the estimate (i.e., PDR or TS) and the corresponding points on the GT.

In the first trial, the user walks on an *S shaped* path, (10 m) long. Both the results obtained using PDR-only and the complete tracking system are compared with the GT and reported in Fig. 5. Considering PDR-only approach, the error accumulated is relevant, although the path is short and the opposite curves partially compensate the drift. Using the Tracking System, the error is reduced to 0.54 cm in the worst case.

The second experiment aims at evaluating the accuracy of the estimate when a closed loop is executed. To this end the user is required to repeat a square-path 5

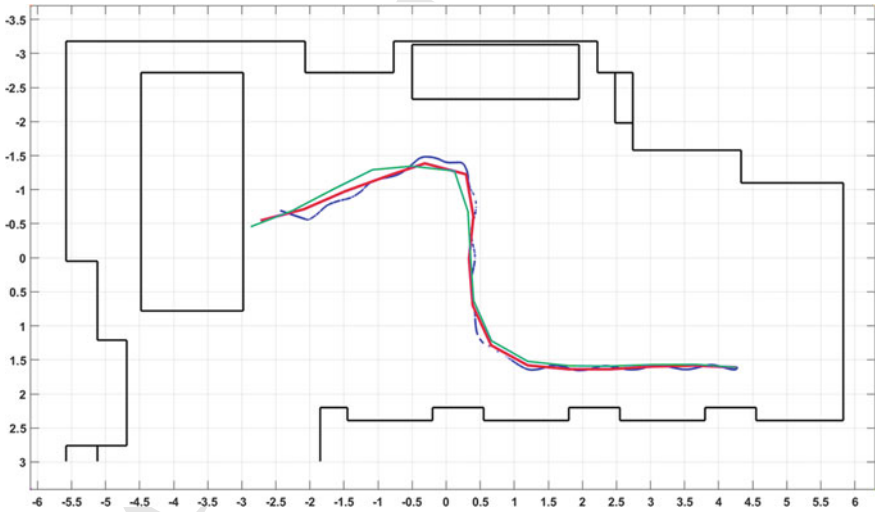

Fig. 5 Results of the S-path experiments: the $(x-y)$ axes represent the NF $[m]$

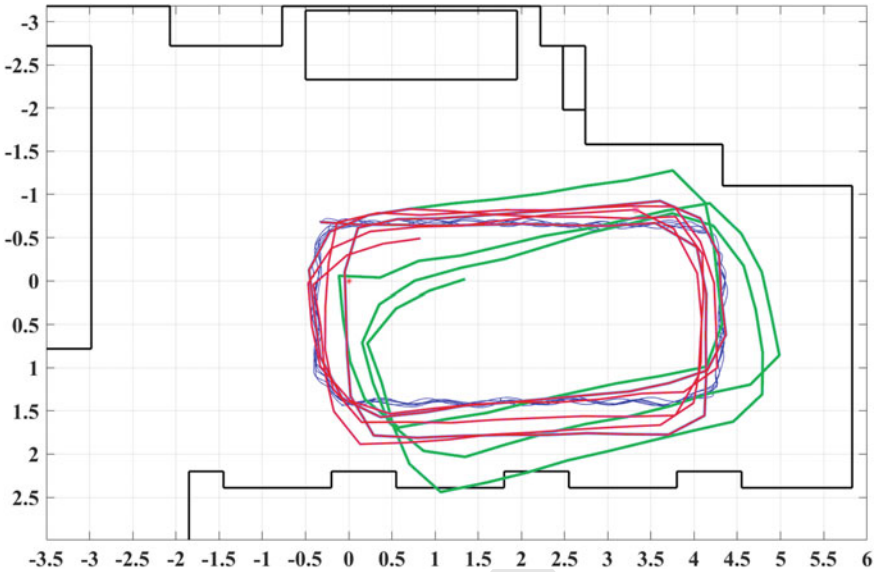

Fig. 6 Results of the Q-path experiments: the $(x-y)$ axes represent the NF $[m]$

Table 2 S and square-path: performance index

	Mean err $[m]$	Min err $[m]$	Max err $[m]$	Std dev $[m]$
S-path				
PDR	0.21	0.01	1.25	0.09
TS	0.10	0.001	0.54	0.02
Square-path				
PDR	0.66	0.15	1.77	0.22
TS	0.34	0.15	0.61	0.02

time without stops (see Fig. 6).The user walks for 130 m: since the user turns in the same direction, the shape of the path does not compensate the drift. The heading drift makes the PDR error unbounded, thus the covariance of the error is larger than the one obtained in the previous trial. When the heading is continuously corrected by the vision system the corresponding covariance does not change.

The values of the key performance indicator selected to asses the accuracy of the proposed system are collected in Table 2.

To prove the effectiveness of the approach in a real scenario, the ARIANNA system has been tested in the Farm Cultural Park, an open-air museum in city center of the Favara (AG), Italy. The project of the Farm Cultural Park has involved

Fig. 7 ARIANNA installation at Farm Cultural Park, Favara, Italy

several semi-abandoned buildings that were completely renovated. They host expositions, exhibitions by international and local artists. The system ARIANNA has been installed to help the visually impaired people visit the outdoor museum, as shown in Fig. 7.

In the trial, the user travels among the buildings, exploiting the path shown in Fig. 8. The path is about 76 m: in this case, the ground truth is not available, so the key performance indicator is represented by the error accumulated at the end of path. Considering PDR, the final positioning error is 3.1 m, that represents about 4% of the distance traveled. This error is limited, since the path is almost straight, however, using the correction provided by the camera, the error is reduced to 0.41 m, that represents less than the 1% of all the distance traveled.

461200_1_En_2_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:9/3/2020 Pages: 22 Layout: T1-Standard

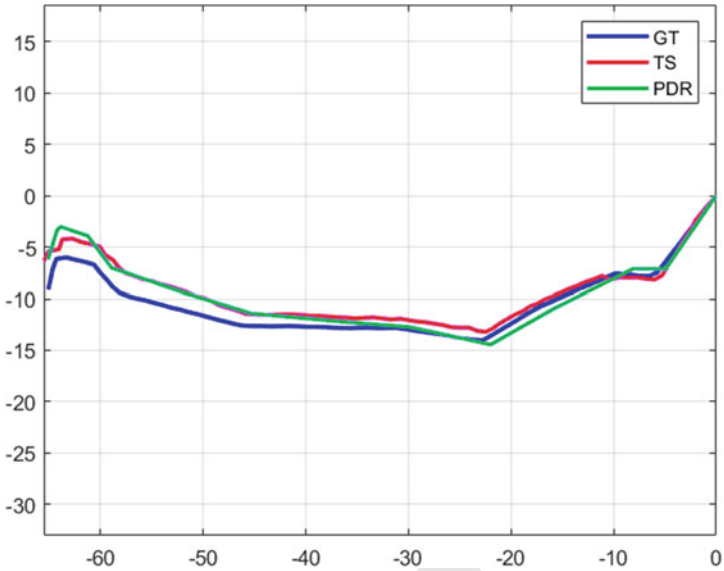

Fig. 8 Real and estimated path in a real installation: the $(x-y)$ axes represent the NF $[m]$, GT (blue line), PDR (green line), and TS (red line)

7 Conclusions

In this chapter we have presented ARIANNA, an innovative smartphone-centric tracking system for indoor and outdoor environments, based on the joint utilization of dead-reckoning and computer vision techniques. The system is explicitly designed for visually impaired people and has been tested in collaboration with key institutions such as the Italian Blind Association (Unione Italiana Ciechi). Many have been the installations (permanent or temporary) that have been provided to the users to be tested. In the occasion of these events, we interviewed 20 visual impaired people (10 low vision users and 10 blind users), asking them to answer to an assessment questionnaire. In both cases, the answers where very encouraging, with a satisfaction grade of 90%. A summary of the users' feedback is provided in Tables 3 and 4.

In real outdoor tests, such as the one carried in the installation present in the Farm Cultural Park, the system is able to suitably reduce the navigation error with respect to the results obtained using PDR only, as demonstrated by the trial carried out using the precise optical tracking system as GT. Future development of the system will include also a model of the hand movement and possibly the design of a vibration feedback to tell the user how to correct its location. Based on the presented tests and experiments, we have implemented and embedded all the algorithms into an app available for both Android and iOS platforms.

Table 3 Answers by low vision people

Question		Answer			
		Yes	No		
1	Have you understood the shape of the path?	9	1		
		None	Mild	Moderate	Very
2	Have you found difficult the change in direction?	2	2	6	0
3	How easy was to follow the path?	0	2	6	2
4	Do you need a learning phase?	1	3	4	2
5	The vibration is useful?	0	0	4	6
6	Have you noticed delays?	6	2	1	1
		Unusable	Not good	Satisfactory	Good
7	What's about the haptic interface operation?	0	3	4	3
		None	Mild	Moderate	Very
8	Do you think this app can substitute tactile paving?	0	1	2	6
9	Do you think interactive information are useful?	0	1	3	6
10	Do you think ARIANNA will increase your independence?	0	0	4	6
		With direction	Without direction	Both	
11	Which operaton mode do you prefer?	6	2	2	
		Indoor	Outdoor		
12	In which context is ARIANNA useful?	8	9		
		None	Mild	Moderate	Very
13	Have you found difficulties in the use of ARIANNA?	0	0	6	3

461200_1_En_2_Chapter ✓ TYPESET DISK LE ✓ CP Disp.:9/3/2020 Pages: 22 Layout: T1-Standard

Table 4 Answers by blind people

Question		Answer			
		Yes	No		
1	Have you understood the shape of the path?	6	4		
		None	Mild	Moderate	Very
2	Have you found difficult the change in direction?	3	6	1	0
3	How easy was to follow the path?	0	4	6	1
4	Do you need a learning phase?	0	3	4	2
5	The vibration is useful?	0	1	6	3
6	Have you noticed delays?	6	3	1	0
		Unusable	Not good	Satisfactory	Good
7	What's about the haptic interface operation?	0	1	7	2
		None	Mild	Moderate	Very
8	Do you think this app can substitute tactile paving?	0	0	2	8
9	Do you think interactive information are useful?	2	1	2	5
10	Do you think ARIANNA will increase your independence?	1	0	6	3
		With direction	Without direction	Both	
11	Which operaton mode do you prefer?	3	5	2	
		Indoor	Outdoor		
12	In which context is ARIANNA useful?	5	8		
		None	Mild	Moderate	Very
13	Have you found difficulties in the use of ARIANNA?	4	6	0	0

Acknowledgements This work was supported in part by the European Union (EU) Funded Research Projects INTERREG V-A Italia-Malta 2014–2020 "I-ACCESS: Implementing the Accessibility to Urban Historic Center's Use and Knowledge

References

1. WHO, Visual impairment and blindness (2014), http://www.who.int/mediacentre/factsheets/fs282/en/. Accessed Oct 2018
2. WHO, World sight day: 13 October 2011 (2011), http://applications.emro.who.int/dsaf/. Accessed Oct 2018
3. F. Felice, F. Renna, C. Attolico, A. Distante, Omero: a multimodal system that improves access to cultural heritage by visually impaired people. Archeologia e Calcolatori supplemento 1 (2007)
4. V. Cantoni et al., A multi-sensory approach to cultural heritage: the Battle of Pavia exhibition. IOP Conf. Ser.: Mater. Sci. Eng. **364**, 012039 (2018)
5. "Building tourism markets accessible to vision impaired-people", https://www2.ift.edu.mo/NewsPortal/building-tourism-markets-accessible-to-vision-impaired-people/. Accessed Oct 2018
6. D. Croce, P. Gallo, D. Garlisi, L. Giarré, S. Mangione, I. Tinnirello, ARIANNA: a smartphone-based navigation system with human in the loop, in *22nd Mediterranean Conference of Control and Automation (MED)* (2014), pp. 8–13
7. D. Croce, L. Giarré, F.G. La Rosa, E. Montana, I. Tinnirello, Enhancing tracking performance in a smartphone-based navigation system for visually impaired people, in *24th Mediterranean Conference of Control and Automation (MED)* (2016)
8. Italian Patent N. BG2014A000054, Sistema di navigazione per non vedenti, presented 2015, patented 2017
9. D. Croce, L. Giarré, F. Pascucci, I. Tinnirello, G. Galioto, D. Garlisi, A.L. Valvo, An indoor and outdoor navigation system for visually impaired people. IEEE Access **7**, 170406–170418 (2019)
10. R. Velázquez, Wearable assistive devices for the blind, *Smart Environment: Issues and Characterization*. LNEE, vol. 75 (Springer, Berlin, 2010), pp. 331–349
11. N. Fallah, I. Apostolopoulos, K. Bekris, E. Folmer, The user as a sensor: navigating users with visual impairments in indoor spaces using tactile landmarks, in *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems (CHI'12), Austin, Texas* (2012), pp. 425–432
12. B. Krach, P. Roberston, Cascaded estimation architecture for integration of foot-mounted inertial sensors, in *IEEE/ION Position, Location and Navigation Symposium, Monterey, CA* (2008), pp. 112–119
13. K. Kaemarungsi, P. Krishnamurthy, Properties of indoor received signal strength for WLAN location fingerprinting, in *Proceedings of 1st Annual International, (MobiQuitous04)* (2004)
14. J.A.M. Ladd, K.E. Bekris, A.P. Rudys, D.S. Wallach, L.E. Kavraki, On the feasibility of using wireless ethernet for indoor localization. IEEE Trans. Wirel. Commun. **5**(10), 555–559 (2006)
15. A.R. Jimènez, F. Seco, J.C. Prieto, J. Guevara Rosas, Indoor pedestrian navigation using an INS/EKF framework for yaw drift reduction and a foot-mounted IMU, *Workshop on Positioning Navigation and Communication, Dresden* (2010), pp. 135–143
16. A.R. Jimènez, F. Seco, F. Zampella, J.C. Prieto, J. Guevara Rosas, Improved heuristic drift elimination (iHDE) for pedestrian navigation in complex buildings, in *International Conference on Indoor Positioning and Indoor Navigation, Guimaraes* (2011), pp. 1–8
17. L. Filardo, F. Inderst, F. Pascucci, C-IPS: a smartphone based indoor positioning system, in *International Conference on Indoor Positioning and Indoor Navigation (IPIN), 4–7 October 2016, Alcala de Henares, Madrid, Spain* (2016)
18. V. Singh, R. Paul, D. Mehra, A. Gupta, V.D. Sharma, S. Jain, C. Agarwal, A. Garg, S.S. Gujral, M. Balakrishnan, K. Paul, P.V.M. Rao, D. Manocha, Smart'Cane for the visually impaired: design and controlled field testing of an affordable obstacle detection system, in *TRANSED 2010: 12th International Conference on Mobility and Transport for Elderly and Disabled Persons*, vol. 53 (2010), pp. 1689–1699
19. B. Hoyle, D. Waters, Mobility AT: the batcane (ultracane), in *Assistive Technology for Visually Impaired and Blind People* (Springer, Berlin, 2008), pp. 209–229

20. A. Wachaja, P. Agarwal, M. Zink, M.R. Adame, K. Möller, W. Burgard, Navigating blind people with a smart walker, in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg* (2015), pp. 6014–6019
21. A. Cosgun, E.A. Sisbot, H. Christensen, Guidance for human navigation using a vibro-tactile belt interface and robot-like motion planning, in *International Conference on Robotics and Automation* (2014)
22. M. Pielot, R. de Oliveira, *Peripheral Vibro-Tactile Displays, Mobile HCI 2013 - Tactile User Interfaces* (ACM Press, New York, 2013). https://doi.org/10.1145/2493190.2493197
23. M. Pielot, How the phone's vibration alarm can help to save battery, http://pielot.org/2012/12/11/how-the-phones-vibration-alarm-can-help-to-save-battery/. Accessed 12 May 2014
24. G. Galioto, I. Tinnirello, D. Croce, F. Inderst, F. Pascucci, L. Giarré, Demo: sensor fusion localization and navigation for visually impaired people, in *MobiCom* (2017)
25. G. Galioto, I. Tinnirello, D. Croce, F. Inderst, F. Pascucci, L. Giarré, Sensor fusion localization and navigation for visually impaired people, in *ECC* (2018)
26. F. De Cillis, L. Faramondi, F. Inderst, S. Marsella, M. Marzoli, F. Pascucci, S. Setola, Hybrid indoor positioning system for first responders, in *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50(2) (2020), pp. 468–479
27. L. Faramondi, F. Inderst, F. Pascucci, R. Setola, U. Delprato, An enhanced indoor positioning system for first responders, in *International Conference on Indoor Positioning and Indoor Navigation* (2013)
28. H. Wienberg, Using the ADXL202 in pedometer and personal navigation applications, Analog devices AN-602 application note (2002)

461200_1_En_2_Chapter ☑ TYPESET ☐ DISK ☐ LE ☑ CP Disp.:9/3/2020 Pages: 22 Layout: T1-Standard