Pattern Recognition Letters 137 (2020) 37–52

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Wearable assistive devices for visually impaired: A state of the art survey

Ruxandra Tapu[a,b,*], Bogdan Mocanu[a,b], Titus Zaharia[a]

[a] ARTEMIS Department, Institut Mines-Télécom / Télécom SudParis, UMR CNRS MAP5 8145, 9 rue Charles Fourier, 91000 Évry, France

[b] Telecommunication Department, Faculty of ETTI, University "Politehnica" of Bucharest, Splaiul Independentei 313, 060042 Bucharest, Romania

ARTICLE INFO

Article history:
Available online 29 October 2018

Keywords:
Assistive devices survey
Visually impaired people
Sensorial networks ETAs
Video camera based ETAs

ABSTRACT

Recent statistics of the World Health Organization (WHO), published in October 2017, estimate that more than 253 million people worldwide suffer from visual impairment (VI) with 36 million of blinds and 217 million people with low vision. In the last decade, there was a tremendous amount of work in developing wearable assistive devices dedicated to the visually impaired people, aiming at increasing the user cognition when navigating in known/unknown, indoor/outdoor environments, and designed to improve the VI quality of life. This paper presents a survey of wearable/assistive devices and provides a critical presentation of each system, while emphasizing related strengths and limitations. The paper is designed to inform the research community and the VI people about the capabilities of existing systems, the progress in assistive technologies and provide a glimpse in the possible short/medium term axes of research that can improve existing devices. The survey is based on various features and performance parameters, established with the help of the blind community that allows systems classification using both qualitative and quantitative measu.res of evaluation. This makes it possible to rank the analyzed systems based on their potential impact on the VI people life.

1. Introduction

Recent statistics of the World Health Organization (WHO), published in October 2017, estimate that more than 253 million people worldwide suffer from visual impairment (VI) with 36 million of blinds and 217 million people with low vision [1].

The visual sense plays a significant role in guiding humans to reach a desired destination, to find the correct path in known / unknown indoor environments or to navigate in outdoor scenes. Unfortunately, blinds or visually impaired (VI) people face a lot of challenges when performing such tasks and feel, most of the time, disoriented or even intimidated [2,3]. The lack of vision may affect personal, professional and environmental relations and can represent a brake in performing daily life routine [4].

Various studies [5] related to the mobility of VI people indicate that the support for acquiring spatial mapping and orientation skills should be provided at two levels, which are *perceptual* and *conceptual*. At the perceptual level, the deficiency from the visual channel should be replaced by other senses such as acoustic or haptic [6]. These sensorial substitution methods transform the real images and distances into acoustic signals or electrical stimulation [7].

At the conceptual level, any novel, VI-dedicated assistive device, should focus on the development of various orientation strategies, including spatial models / surface mapping, in order to adapt the user to unpredictable conditions that can arise during navigation [8]. However, in novel, unknown environments, the VI user is facing a set of difficulties and suffers from insecurity or anxiety [9].

In the last years, various research works have addressed such challenges in their attempt to gain a higher level of understanding of the surroundings, to increase cognition and to facilitate the navigation of the VI people in both indoor and outdoor environments. Most of the existing assistive technologies incorporate functions for obstacle avoidance and route selection. Because most partially sighted or blind people live in developing countries, the assistive device should be both relatively affordable, from a financial point of view, and easily available [3].

The state of the art offers a wide range of electronic traveling systems (ETA), dedicated to VI people, designed to provide a sensorial substitution to the human vision [10]. In this paper, we propose a comprehensive review of the most recent assistive devices systems, based on computer vision / machine learning methods and we try to explain why, despite impressive technological ad-

* Corresponding author at: ARTEMIS Department, Institut Mines-Télécom / Télécom SudParis, UMR CNRS MAP5 8145, 9 rue Charles Fourier, 91000 Évry, France.
E-mail address: ruxandra.tapu@telecom-sudparis.eu (R. Tapu).

https://doi.org/10.1016/j.patrec.2018.10.031
0167-8655/

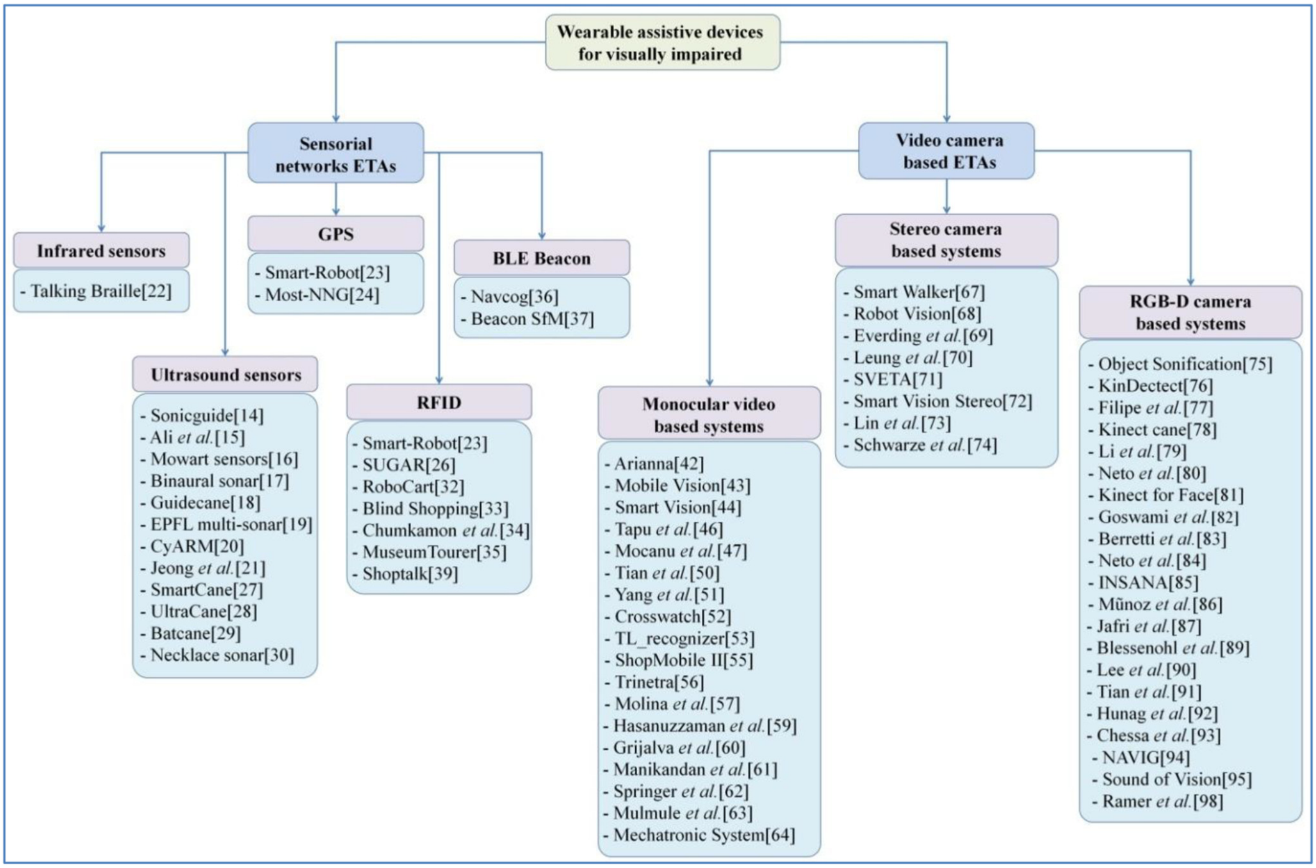

Fig. 1. Overview of wearable assistive devices dedicated to visually impaired users.

vances [11, 12] in the field, such devices are currently rarely used by VI people, which still prefer white canes or guiding dogs.

The rest of the paper is organized as follows. In Section 2 we review the state-of-the-art systems dedicated to the VI people. Section 3 presents and discusses the most relevant functional elements that represent key components in the computer vision-based assistive devices. In Section 4, we propose a set of both quantitative and qualitative criteria that we exploit for evaluating the most relevant video-based approaches. Finally, Section 5 concludes the paper and opens novel directions for further work and developments.

2. State of the art review

In the last couple of years, various ETAs technologies have been introduced. They aim at increasing the mobility of VI users and at providing additional knowledge about nearby surroundings [13]. Existing ETAs can be divided in two major families, including: (1) sensorial network systems (*i.e.*, active systems) and (2) video-based systems (*i.e.*, passive systems), as illustrated in Fig. 1.

2.1. Sensorial networks ETAs

The sensorial network ETAs have been designed to collect environmental information that is translated into acoustic / haptic signals transmitted to the VI user. In particular, sensorial-based assistive devices provide subject localization and object identification using: ultrasound [14–21], infrared [22], global positioning system (GPS) [23,24] and radio frequency identification (RFID) [23,25,26].

One of the pioneering systems proposed in the literature exploited Mowat sensors [16] that estimate the distance between the VI user and various obstacles existent in the scene. The system needs to be hand-held and requires an extensive training phase so that the VI user becomes familiar with the beeping patterns. Among such pioneering approaches, let us also mention the method introduced in [14], where two sonar-based environmental imaging sensors, called Sonicguide and Trisensor are proposed, and provide spatial information about the detected obstacles. However, both prototypes use low resolution sensors and the estimated distances are not very accurate in the 3D space.

More recently, the Talking Braille system introduced in [22] exploits a wireless ubiquitous computing network, designed to assist VI users to navigate safely in indoor public spaces. The system is limited to known buildings, on which various access / interest points are defined in advance. In addition, the infrared technology may not perform efficiently during daytime due to its high sensitivity to sunlight/illumination conditions. The binaural sonar system introduced in [17] (Fig. 2a) robustly detects objects situated at arbitrary locations on the left / right side of the VI user. The feedback is provided to the user as a set of vibrotactile simulations. The GuideCane (Fig. 2b) [18], SmartCane (Fig. 2c) [27], UltraCane [28], Batcane [29] (Fig. 2d) and Necklace cane (Fig. 2e) [30] systems exploit solely ultrasonic sensors networks that are aiming at improving / enriching the capabilities of a regular white cane.

Such frameworks attempt to estimate the optimal, obstacle free, walking path and inform the VI user through tactile simulation. However, when tested in real-life scenarios, they show quickly their limitations, notably when detecting objects situated above the user knee, overhanging obstacles or sidewalk borders.

The EPFL multi-sonar architecture (Fig. 2f) proposed in [19] is designed to function in indoor environments, in order to facilitate the user displacement through corridors by avoiding the possible

Fig. 2. Wearable assistive devices based on sensorial networks.

collision with obstacles situated at arbitrary levels of height. However, when tested in real-life conditions, the number of false positives returned by the system is extremely high, which makes the blind community reluctant to the acceptance of the system.

A haptic sensing device denoted by CyARM is proposed in [20] (Fig. 2g). CyARM allows the spatial localization of the VI user through ultrasonic transducers. When evaluated in real, outdoor scenarios the system obtains high detection rates for static objects, while correctly estimating the real distances between the users and obstacles that are present in the scene. However, for dynamic (moving) objects, the detection scores in terms of accuracy and recall rates significantly decrease (with more than 30%). In addition, the performance of ultrasonic sensors can be affected by the weather conditions.

A low-cost guidance system based on a microcontroller equipped with an ultrasonic sensor is introduced in [15]. The objective here is to detect objects situated in front of the walking stick. In [21], a similar electronic travel aid system is proposed (Fig. 2h), that performs obstacle detection and transmits the feedback through tactile stimulation. Such systems can be globally regarded as *smart-canes* that allow the localization of obstacles situated at arbitrary levels of height. However, they have never been tested and validated by real VI users or in crowded, outdoor scenes.

The Smart-Robot (Fig. 2i) device introduced in [23] is a VI navigation assistant that integrates RFID and GPS localization. The method is able to perform in both indoor and outdoor environments in order to assist the VI user in reaching a predefined destination. The user feedback is based on acoustic messages, together with vibrations patterns transmitted with the help of a hand-mounted glove. A similar approach, so-called MOST-NNG mobile phone navigation assistance is proposed in [24].

However, such GPS-based systems prove to be highly dependent on the signal loss and present a poor accuracy rate in estimating the user's position. This limitation notably appears in the case of urban environments, where the density of buildings is high [31].

Indoor and outdoor localization techniques that do not relay on GPS often involve structural modifications of the environment, that require the inclusion of additional technologies, such as those related to market access. Within this framework, the key solutions rely on the RFID technology and aim at facilitating the user navigation in indoor spaces [32,33]. In [32], a RFID system is developed (Fig. 2j), designed to help the VI users reach a path to a specified destination. In [33], the RFID is combined with Quick Response (QR) codes to locate products inside grocery stores (Fig. 2k). In [34,35] (Fig. 2l and m), the RFID technology is used in order to develop a tourist guide within museums. The major limitation of any RFID system is the necessity of disposing of some previous knowledge on the building architecture. Such systems cannot offer a wide range of capabilities, due to the necessity of installing tags everywhere within the considered environment. In addition, the related frameworks are considered as invasive, because the RFID readers need to be mounted on the subject's body.

Various authors [36,37] (Fig. 2n), propose using the BLE beacon, instead of RFID, because it can be temporally applied to the environment with velcro stickers and quickly removed without permanent alterations of the surroundings. However, as in the case of RFID frameworks, the BLE systems performance significantly depends on the location of the BLE beacon devices.

Independent and safe mobility is vital for independent shopping. The VI users have ranked the shopping centers as the most challenging environments to navigate within, while the overall shopping experience is considered as a major problem [38]. In [39], the ShopTalk buying assistant is introduced (Fig. 2o). The system is composed of regular helmets, barcode readers, numeric keypad and a processing unit. The experimental studies have shown that ShopTalk returns a high success in product retrieval. However, the system is difficult to carry on and requires accessing the store inventory control, which is intractable.

The SUGAR system proposed in [26] offers guiding services to the VI users when navigating in indoor spaces (Fig. 2p). SUGAR uses the UWB (*Ultra-WideBand*) technology for user positioning. The interaction between the VI user and the system is managed through acoustic signals and voice commands, played through headphones. As pointed out by authors, the system is limited to indoor scenarios and to environments containing UWB tags. Nothing is mentioned about the power consumption or the system's lifetime.

The timeline evolution of the sensorial assistive devices presented above is illustrated in Fig. 3. The sensorial networks ETAs

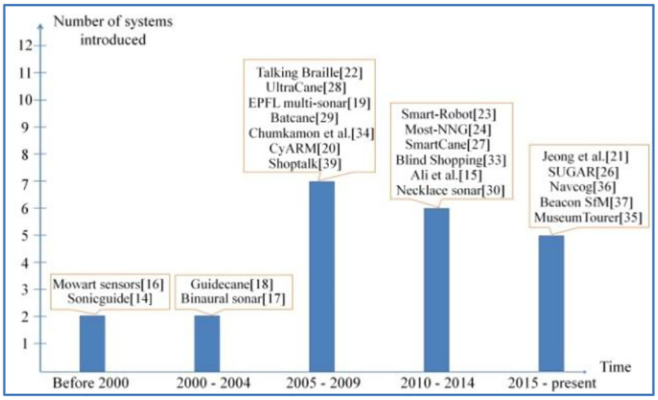

Fig. 3. The timeline evolution of the sensorial assistive devices.

are highly useful in assisting the visual impaired persons in navigation due to their high accuracy in establishing the actual distance between an obstacle and the VI user. However, none of the above-analyzed systems are able to locate, identify and recognize specific objects in an unknown environment, or to determine their relative degree of danger. In addition, ultrasonic sensors are highly sensitive and may suffer from interference problems with other sensors or signals from the environment. Although laser scanners are highly used in mobile robots navigation [40], due to their precision and resolution, they are expensive and heavy to wear, which makes them poorly-suited for inclusion in a VI-dedicated wearable assistive device.

On the contrary, video-based systems incorporating computer vision algorithms, offer a superior level of reproduction and interpretation of real scenes, at the price of a higher computational complexity [41]. Let us now analyze computer vision-based ETAs.

2.2. Computer vision-based ETAs

Due to the proliferation of hardware processing devices, computer vision algorithms and machine learning technologies, various systems designed to increase the mobility of VI users are based on artificial intelligence. Depending on the type of camera sensor considered, the video-based methods can be divided into the following three categories: monocular, stereoscopic and RGB-D. Let us now review the state-of-the-art approaches, and emphasize related strengths and limitations.

2.2.1. Monocular camera based systems

In [42], a low-cost augmented reality (AR) system for blind and partially sighted people is introduced (Fig. 4a). The framework, called Arianna, is designed to identify a safe walking path in indoor spaces. At the hardware level, the system is based on a video camera integrated on a smartphone device. The user feedback is transmitted with the help of a set of vibration patterns. The walking path is determined using a set of interests points, indicated by QR codes or by following a path painted on the floor.

Another marker-based approach, called Mobile Vision, is introduced in [43]. The system is integrated on a smartphone device and uses specific color markers in an indoor environment (Fig. 4b). By using red, green and blue color markers the user is directed to find some locations of interest such as restrooms, elevators or exits. Text-to-speech transcripts are used to provide feedback messages to the VI user. Both methods presented above require an initial knowledge of the environment and thus cannot be applied to arbitrary, unknown paths or scenes.

In [44], the SmartVision navigation framework is designed for outdoor scenarios and integrates GPS, Wi-Fi localization with GIS (*Geographic Information System*) [45], passive RFID tags, and computer vision techniques (Fig. 4c). The system does not have the ambition to replace the white cane, but to complement it, in order to alert the VI user of looming hazards. A database with potential objects of interest is here developed (*e.g.,* elevator, welcome desk, plants, cash machine, telephone booth). In the test phase, the reference images, stored *a priori*, are searched within the video frames acquired by the camera. However, the method is highly sensitive to camera motion and heavily depends on the size of the training dataset.

In addition, it suffers from scalability issues, since for a larger database with multiple objects of interest, the computational time increases significantly.

An obstacle detection and classification method completely integrated on a regular smartphone is introduced in [46] (Fig. 4d) and extended in [47] (Fig. 4e). The system is designed to facilitate the VI user navigation in both indoor and outdoor environments. In [46], authors propose to identify the location of an obstruction by extracting interest points that are tracked between successive frames using the traditional Lucas-Kanade algorithm [48]. The object's motion is distinguished from the camera movement with the help of a set of homographic transforms that are classified by applying the RANSAC algorithm [49]. The detected obstacles are further categorized by incorporating the HOG descriptor into a Bag of Visual Words (BOW) representation. Even though the system returns overall good results, it cannot detect large, flat structures or correctly estimate the distance between the VI user and an obstruction. In [47], the authors proposed to solve the above-mentioned limitations by integrating within the system ultrasonic sensors (Fig. 3e). The approach shows promising results, but proves to be sensitive when multiple moving obstacles are present in the scene.

A computer vision-based way-finding system that allows independent access to indoor, but unfamiliar environments is proposed in [50] (Fig. 4f). At the hardware level, the system is composed of a video camera, a microphone, a computer and Bluetooth earpieces. The framework detects doors, elevators and cabinets using the objects geometric shape combined with corners and edges identification algorithm. Then, by exploiting an optical character recognition approach, the system is able to distinguish between foreground and background objects.

An additional module has been added to the framework that allows cloth identification [51]. Using a novel Radon Signature descriptor, four types of clothes textures can be identified: plaid, striped, pattern less and irregular. Even though both modules have been developed for VI people, no studies / experiments have been performed up to now with real VI users. In addition, the framework cannot function in real-time or handle object occlusion.

The Crosswatch approach [52] exploits computer vision algorithms and GIS. Crosswatch is a navigation assistant based on a regular smartphone and is designed to provide guidance to the VI users when crossing intersections. However, because the Crosswatch transmits acoustic signals using the phone speakers, the warning messages are difficult to hear in noisy urban environments. A similar mobile traffic light mobile device has been introduced in [53] and extended in [54]. The system, called TL-recognizer, uses various color filters and contour detection methods to identify potential traffic lights. However, the approach is sensitive to camera movement and variability of illumination conditions.

In [55], the so-called ShopMobile shopping assistant, integrated on a regular smartphone, attempts to reduce the limitations of ShopTalk [39]. The smartphone's embedded video camera is here used in order to locate the products barcodes. Similarly, in [56] a smartphone grocery shopping assistant is proposed with the goal

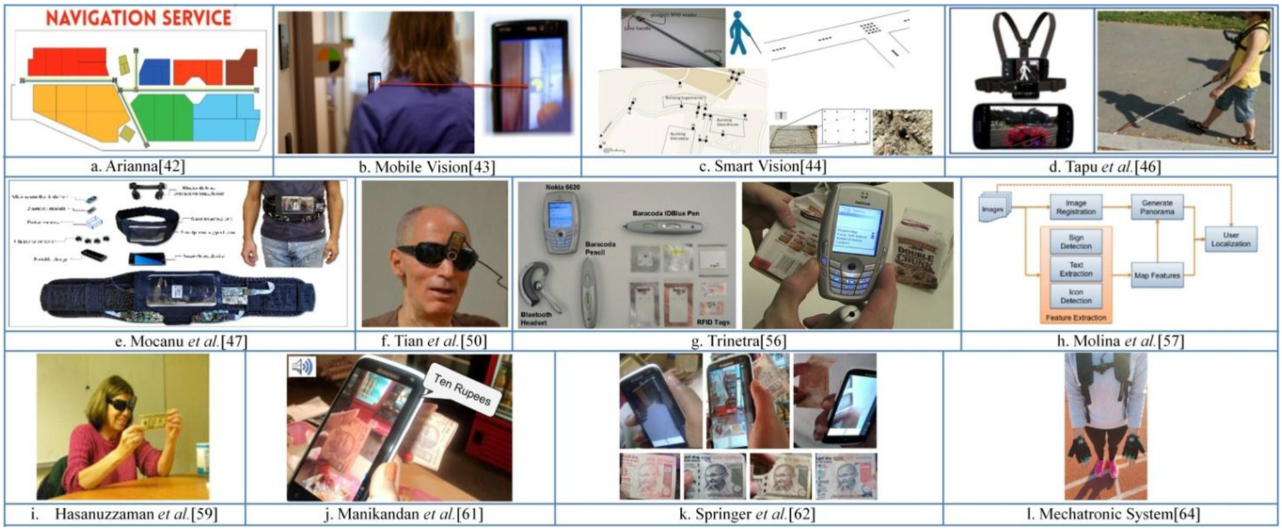

Fig. 4. Wearable assistive devices using monocular cameras.

of portability and cost effectiveness (Fig. 4g). All above systems use the video camera as a barcode scanner.

In [57], Molina et al. introduced the concept of visual nouns for VI user navigation in indoor and outdoor environments. The system creates mosaic images that are further used to facilitate the VI people navigation through streets and hallways. A visual noun is described by three types of features: text, visual icons and signs Fig. 4h). However, there are still a set of open requirements that need to be fulfilled in order to make the system useful from the VI people point of view: (1) the development of an appropriate human-machine interface; (2) the integration within a wearable assistive device and (3) the development of the acoustic or haptic interface.

The method is extended in [58], where a person localization and navigation system in indoor scenes is proposed. Compact and omni-directional video features are extracted using the video camera embedded on a smartphone that are further transmitted to a GPU server for processing. Even though the method shows promising results in the evaluation stage, the systems has been tested only on a limited number of indoor spaces. A large scale scene dataset is required that includes multiple buildings with various floors. Moreover, in order to increase the computational speed, hierarchical and context-based methods can be used in order to avoid searching within the entire database. Finally, the development of the user interface is mandatory to facilitate the communication.

Different systems [59] (Fig. 4i), [60,61] (Fig. 4j), [62] (Fig. 4k) and [63] based on regular monocular video cameras, embedded on various hardware platforms address the problem of real-time banknote recognition in the context of blind people. Even though the necessity of such an application appears as obvious, such systems still suffer from a set of limitations related to the high sensitivity to the illumination conditions or to occlusions. Thus, the banknote needs to face the direction of the video camera. In addition, low recognition and accuracy scores are obtained in the case of important camera/background motion.

Starting from the observation that in the Boston marathon more than 100 participants were blind, in [64] the Mechatronic system is introduced (Fig. 4l). The prototype is designed to facilitate jogging or running activities. At the hardware level, the system is composed of a regular monocular camera, a processing unit and a haptic device. At the software level, the authors propose detecting ground lanes and running corridors. However, the framework accuracy is highly dependent on the illumination conditions. Because of the vibrating gloves mounted on the user hands the system is considered as invasive. In addition, the battery life of the entire system is inferior to one hour.

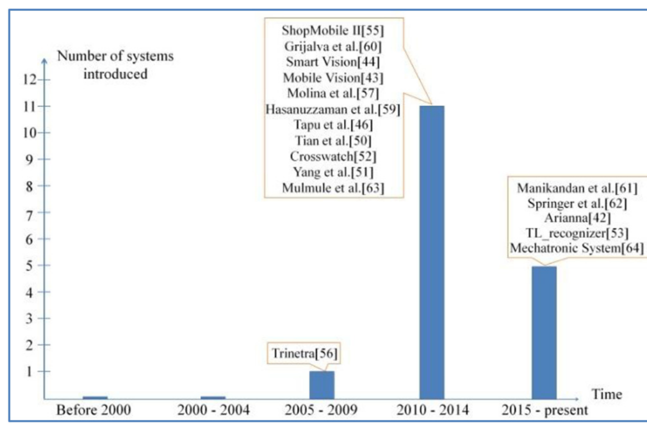

Fig. 5. The timeline evolution of the ETAs based on monocular cameras.

The timeline evolution of the electronic travel aids based on monocular cameras is illustrated in Fig. 5.

The assistive devices based on CCD camera systems are more compact and easy to maintain than sensorial frameworks. However, they have such a low accuracy when estimating the real distance between the VI user and a detected obstacle. A major disadvantage of any monocular system is that it cannot estimate the global object scale based on a single image [65].The problem gets aggravated when dealing with larger scenes because it is more likely to have scale drifts between map portions and their estimated motion vectors [66]. Moreover, the systems cannot differentiate between foreground and background objects. In addition, despite the efforts made to detect static objects using the camera apparent motion, such frameworks are valid only in limited scenarios with known obstacle appearance models.

The stereo-based systems, presented in the following section, aim at overcoming such limitations.

2.2.2. Stereo camera based systems

The Smart Walker system introduced in [67] aims to detect dangerous obstacles and hazardous ground changes, both outdoors

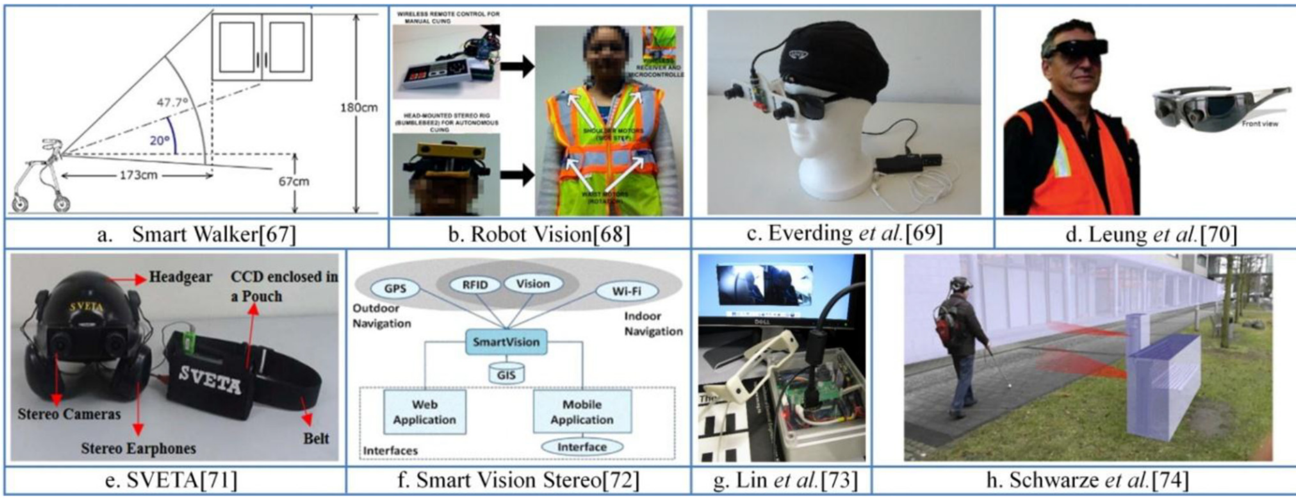

Fig. 6. Wearable assistive devices using stereoscopic cameras.

and indoors (Fig. 6a). Smart Walker is based on real-time stereo obstacle detection algorithm that uses the sparse 3D information. By working directly with acquired 3D points, the method is computationally more advantageous than constructing a 3D map of the environment. However, the system can operate at only 4 fps and returns a detection rate inferior to 65%, which is not sufficiently reliable in the context of VI-dedicated applications.

In [68], the authors introduce Robot Vision, a head-mounted stereo-vision navigational assistance device for VI users (Fig. 6b). In order to extract and maintain the orientation information and create the sense of egocentricity, the system incorporates the visual odometry and a feature-based metric topological SLAM (*Simultaneous Localization And Mapping*) [68]. The method is able to reconstruct the 3D scene through stereo triangulation and develop a vicinity map of the user's environment. This map serves to perform a 3D traversability analysis in order to guide the subject through a path with potential obstacles. From the user's perspective, the system is considered as invasive since it requires to be mounted on the head. In addition, the development of the 3D map requires a powerful processing unit that needs to be worn by the user during navigation.

The wearable, light-weighted navigation device proposed in [69] is based on dynamic vision sensors, which transmit, instead of the entire image at a fixed frame rate, only the local changes caused by movement in a scene (Fig. 6c). In this way, the visual information is represented as a stream of moving pixels that are asynchronously generated. The reported experimental results show that the system's main limitation concerns the detection of objects situated at the head level. In addition, the use of regular headphones is not optimal because they block the user's hearing.

A visual navigational assistant that estimates the egomotion in highly dynamic scenes is proposed in [70] (Fig. 6d). Using the 3D coordinates of the feature points obtained by standard triangulation, the system is able to determine the ground plain. Then, the ground plain normal is estimated using two different sources of information: stereo and inertial data. The experimental evaluation performed on synthetic and real scenes validates the approach. At the hardware level, the system is composed of a stereo camera mounted on sunglasses and a desktop computer. From the VI user point of view the system is not portable. In addition, nothing is said about on how to transmit the video steam to the processing unit or the warning signals to the VI users.

The stereo vision-based electronic travel aid proposed in [71] and denoted by SVETA consists of a helmet equipped with a stereo camera and earphones (Fig. 6e). The disparity map of the surrounding scene is obtained using a stereo matching algorithm. Then, the depth map is converted into musical sounds transmitted to the VI user.

The SmartVision Stereo navigation assistant dedicated to VI people introduced in [72] detects obstacles and provides navigational information in indoor and outdoor scenarios (Fig. 6f). As any stereoscopic-based method, both SVETA and SmartVision suffer from matching errors when estimating the disparity maps, especially in bright outdoor environments. In addition, such frameworks require an extensive training phase so that the VI user becomes familiar with the specifically-tuned sound patterns.

In [73], a wearable stereo vision system for visually impaired users is introduced (Fig. 6g).

The device is composed of eyeglasses and embedded processing device. In order to offer information about the VI user position, the video stream is transmitted through the 3 G network on a mobile device. The major strength of the system is its capacity to locally compute the disparity map and to detect various types of obstacles. However, some alerting functionalities are required while the processing speed needs to be increased.

A wearable device that allows the VI people to perceive the environment using a head-worn stereo camera is introduced in [74] (Fig. 6h). The method constructs a scene model and updates it as the user is moving through the environment. The foreground objects are differentiated from the background structure by using a dense disparity estimator that is combined with a visual odometry and an inertial measurement unit. The system is dedicated to navigation in outdoor scenes and transmits acoustic messages through bone conducting headphones. As pointed out by the authors, the system is not able to differentiate between various obstacles detected or to determine their degree of danger. In addition, if multiple objects are situated in the user's vicinity, a set of warning messages are launched and the VI user may become overwhelmed by the transmitted information.

The timeline evolution of the electronic travel aids based on stereoscopic video cameras is presented in Fig. 7.

Although numerous stereo-vision-based systems have been introduced, some inherent problems still need to be solved. First, because of the incorrect estimation of large depth cues the stereo-matching algorithms fail, especially in the case of poorly textured regions. Second, the quality and accuracy of the estimated depth map is sensitive to artifacts in the scene and abrupt changes in the illumination.

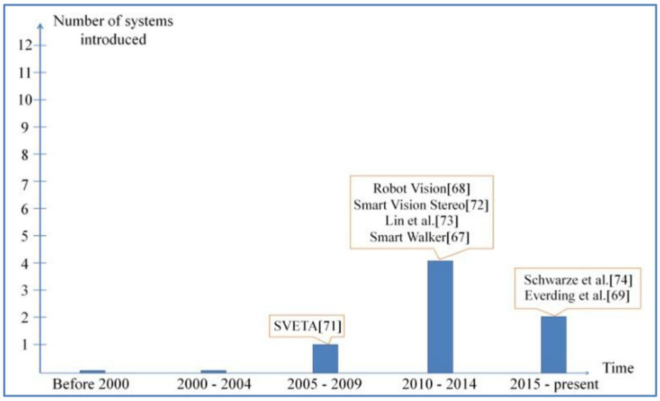

Fig. 7. The timeline evolution of the electronic travel aids based on stereoscopic video cameras.

More recently, the emergence of RGB-D cameras enabled the apparition of a new family of VI guidance systems exploiting such technologies. They are described in the following section.

2.2.3. RGB-D camera based systems

Several works [75] (Fig. 8a), [76] (Fig. 8b), [77] (Fig. 8c), [78] (Fig. 8d) have demonstrated the feasibility of using the RGB-D sensors, as the main hardware component of an wearable assistive device, due to its capability of sensing depth information along with RGB data. However, none of the above-mentioned systems has been evaluated in terms of recall and accuracy when performing object detection tasks.

In [79–81] (Fig. 8e), [82,83] various complex frameworks have been introduced, but they are not suitable for real-time systems integrated in low processing devices.

In contrast, recently, a real-time face recognition system, dedicated to blind and low-vision people, is proposed in [84] (Fig. 8f). The framework integrates wearable Kinect sensors, performs face detection, and uses a temporal coherence along with a simple biometric procedure to generate a specific sound that is associated with the identified person. The underlying computer vision algorithms are tuned in order to minimize the required computational resources (memory, processing power and battery life). From this point of view, they are overcoming most state-of the art techniques. However, the range of the Kinect sensors limits the applicability of the approach to solely indoor environments.

Another indoor navigation system, so-called INSANA (Fig. 8g), enriched with obstacle detection capabilities and integrated on a Tango device is proposed in [85]. The approach is specifically designed for indoor spaces and develops a semantic map of the environment by parsing information from a building architecture model. The detection module identifies the objects situated in front of the depth sensor. The safe, obstacle free walking path is determined based on ego-motion tracking and localization alignment. The guiding messages are transmitted to the VI user in real time, through a speech-audio interface. The experimental results performed with blind users validate the method.

However, the system can work only in simulated scenarios, for which a previous knowledge about the building structure is available.

A similar method is proposed in [86], where the authors introduce an indoor staircase detection algorithm based on the RGB-D camera integrated on a Tango Tablet (Fig. 8h).

The staircase candidates are firstly detected from RGB frames by extracting a set of concurrent parallel lines with the help of a Hough transform. Then, the depth frames are further analyzed in order to distinguish between candidate classes: upstairs, downstairs, and negatives (*i.e.,* corridors) using a SVM (*Support Vector Machine*) classifier trained with multiple categories. The detection rates demonstrate the effectiveness and efficiency of the algorithm. The Tango framework is also used in other systems dedicated to VI users [87] or [88]. However, none of such systems achieved a sufficient maturity to be tested by actual VI users in real-life conditions.

In [89], a head-worn depth camera system designed to help the VI move safely in an indoor environment is proposed (Fig. 8i). The system is designed to detect floors, walls or obstacles and to convert the information yielded by the detection module into a 3D sound map. The method runs in real-time and is useful in spotting both small and large objects situated on the floor.

In [90] (Fig. 8j), [91,92] various RBG-D camera-based indoor navigation systems for VI users are proposed. The frameworks perform real-time 6DOF ego-motion estimation using sparse visual features, dense point clouds and ground plain detection in order to obtain a safe, obstacle free walking path. The system in [90] is portable and at the hardware level is composed of a RGB-D head mounted camera, a laptop acting as a backpack processing unit and a haptic feedback vest (Fig. 8j). From the VI user point of view, the system is invasive because it requires being head-mounted and the subject's physical contact with the vibration mechanism from the vest. In addition, the 3D maps that are successively updated in time are likely to be affected by the accumulated errors and sensors noise.

Similarly, the 3D environment interpretation system for VI users proposed in [93] includes the following functionalities: object detection and recognition, face identification and optical character recognition (that helps identifying street names and food products). In addition, as in [52] (Fig. 8k), [53,54] a crossing detector is developed together with a pedestrian line detection [13]. Even though the entire framework looks quite promising and incorporates a highly complete software package, nothing is said about the hardware specifications. In addition, the information about the acoustic feedback is missing. Moreover, the most difficult part will be to prioritize all the warning messages coming from different software modules.

The NAVIG assistive device based on artificial vision, geo-located visual landmarks and GIS introduced in [94] (Fig. 8l) is composed of a RGB-D video camera, microphone, regular headphones and a backpack computer. In contrast with other devices, the object localization function is used not to avoid dangerous situations, but to guide the VI people reach the object / point of interest using also GPS, GIS positioning and 3D sound patterns. From the experimental evaluation, it can be observed that the system processing speed is significantly influenced by the object size and the maximum number of landmarks included in the database. In addition, it requires an extensive training phase.

In [95,96] the Sound of Vision (SoV) assistive device is introduced. The system creates a joint audio and tactile representation of the information acquired from the surrounding environment. Starting from the reconstructed 3D representation of the indoor/outdoor scene, an image segmentation is performed in order to extract objects of interest (Fig. 8m). The framework is quite complex and requires a training phase in order to understand the warning messages. Furthermore, experiments with actual VI user in real life, crowded urban scenes, are necessary in order to validate the approach.

The problem of VI people independent jogging is address in [97,98] using 3D video cameras (Fig. 8n). However, none of the systems have been tested with actual VI users. In addition, the proposed methods are highly sensitive to the camera motion despite the stabilization operations proposed.

The timeline evolution of the assistive devices using RGB-D cameras systems is presented in Fig. 9.

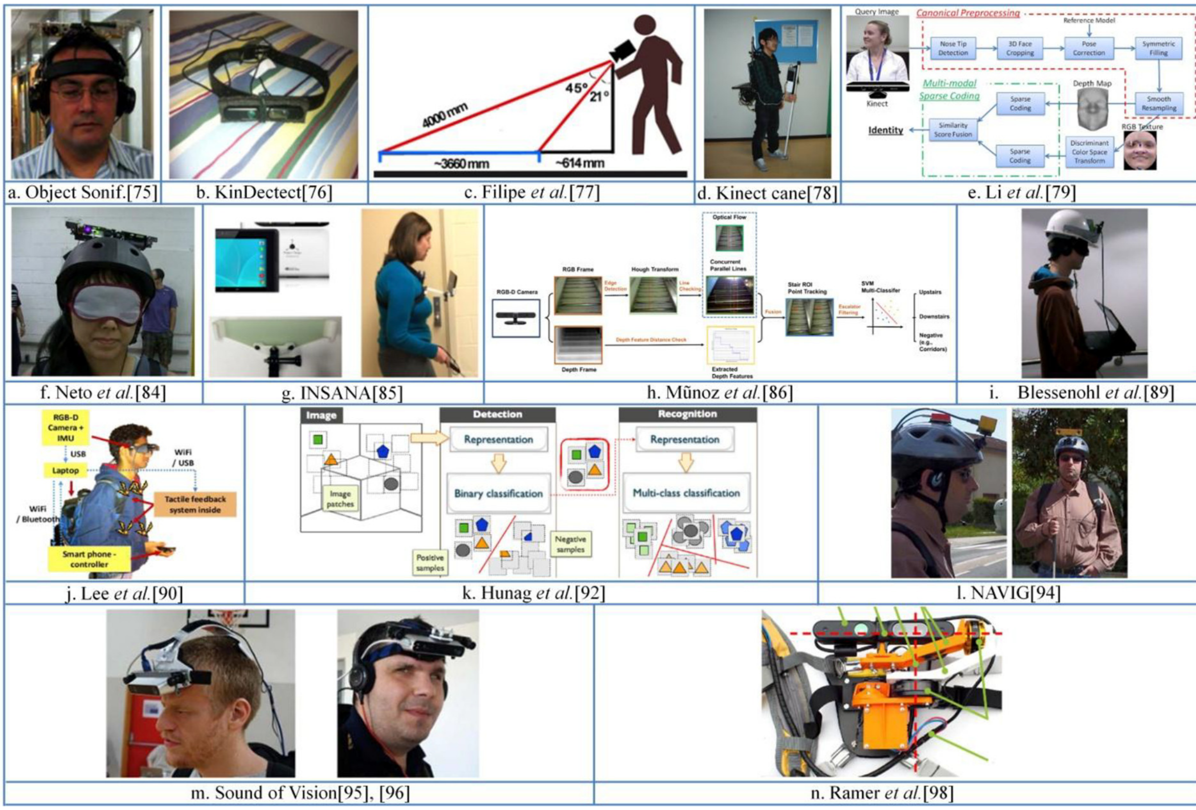

Fig. 8. Wearable assistive devices using RGB-D camera.

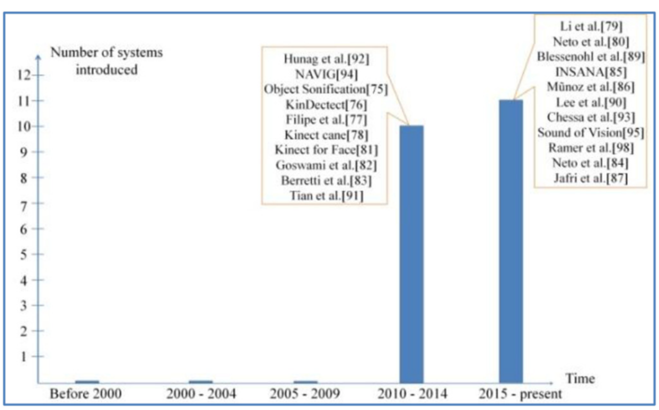

Fig. 9. The timeline evolution of the electronic travel aids based on RGB-D cameras.

Although various stereo and depth cameras systems have been introduced, some inherent drawbacks need to be overcome. First, the frameworks are highly dependent on the quality of the estimated disparity map, which, for poorly textured regions, is not sufficiently accurate. Second, when tested in real-life, outdoor scenarios that contain changes in the illumination, visual artifacts or various types of motion, the frameworks show quickly their limitations. Moreover, due to the limited computational resources, the development of accurate and dense depth maps is still expensive.

The global timeline evolution of the methods presented in Section 2 is illustrated in Fig. 10. As it can be observed before 2000 all the frameworks proposed were based exclusively on the sensorial substitution of the human vision using various types of sensors: ultrasonic, infrared or RFID. However, with the development of computer vision and machine learning algorithms, nowadays most of the ETAs are based on monocular, stereoscopic or 3D video cameras. When embedded on wearable devices, such approaches make it possible to obtain a higher level of understanding of the scene where the VI user is evolving in.

For this reason, in the following, we will focus our attention on computer vision-based systems, which seem to show the most promising results in terms of accuracy and repeatability. Let us analyze the functionalities proposed by such systems.

3. Computer vision based systems: functionality analysis

Most sensorial networks ETAs are designed to identity the safe walking path (*cf.* Section 2.1), but none of the frameworks can include additional, high level, functionalities. In the last couple of years, the assistive devices integrating computer vision algorithms have captured the attention of the scientific community. Compared with sensorial networks ETAs, the camera based systems offer a superior level of reproduction and interpretation of real scenes, at the price of a higher computational complexity. Based on this observation, in this section we focus our attention on computer vision-based systems designed to perform a semantic analysis of the surrounding scene in order to improve VI user cognition over the environment.

The main functionalities that an assistive device dedicated to VI people should satisfy, in order to make it acceptable by the blind community, are the following:

- static/moving obstacle detection,
- capability to provide information about the distance between the user and the detected obstruction,
- ability to provide directional and navigational messages so that the VI user reach a desired destination or to identify the safe walking path.

Table 1 presents a comparison of the most relevant features introduced in ETAs based on computer vision algorithms. As

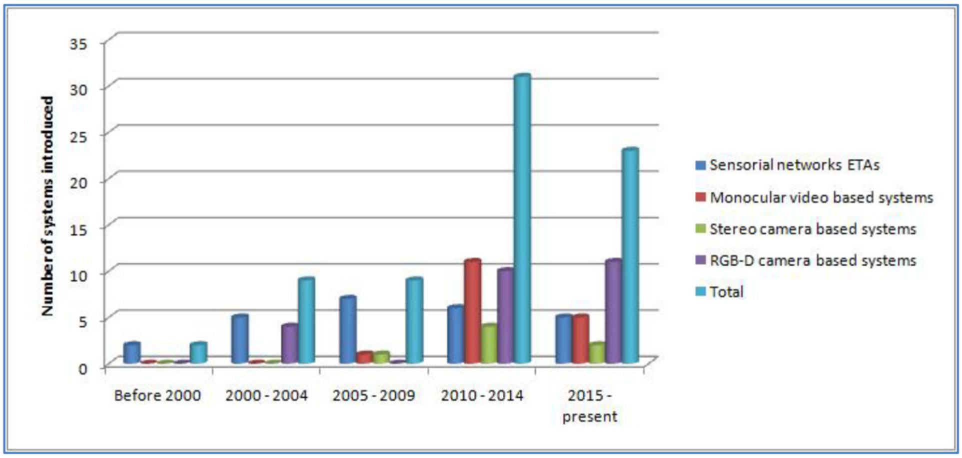

Fig. 10. Timeline evolution of the analyzed systems.

it can be observed, most of the systems have been developed in order to facilitate the safe navigation of the VI user in an unknown indoor/outdoor environment. The majority of devices [42,64,68,70,74,75,88,89,90,98] include a walking path identification and a dedicated module that performs object detection able to identify various types of hazards or dangerous obstacles.

In addition, most systems include high level capabilities. For example in [42,57,90] the VI user position is estimated based on computer vision algorithms and GPS/GIS signals. In [46,47,50,55,67,72,76,78,93], object classification modules are proposed in order to help establishing a semantic interpretation of the detected objects. Furthermore, in [46,47] the obstacles are prioritized based on their degree of danger and relative position with respect to the VI user.

The authors in [43,44] designed the ETAs with a landmark detection module (that requires a training stage), while in [77, 86, 91] a stairs identification algorithm is introduced. In order to facilitate the VI people displacement in crowded urban scenes various authors proposed a traffic light detection module [53,54] or a zebra-crossing identification [52,53,54,72,91]. The independent shopping problem of the VI user is addressed in [55,56] where a mobile grocery assistant is proposed. The assistant can be further extended with a banknote recognizer as in [59,60,93]. Finally, the semantic information about the presence of a familiar face in the VI people near surrounding is addressed in [84,93] where face recognition algorithms designed for wearable devices are proposed.

As a general conclusion of the above state of the art methods we can say that the difficulty is not developing a system that has all the “bells and whistles” but to conceive the technology that can last in time and be useful.

Regarding the user interface, most state of the art systems attempt to replace vision by other senses such as acoustic or haptic. The best asset of haptic feedback is its discretion (only the user feels the effect of touch). For example, in [44,55,64,68,72,90,78.98] various vibration actuators are proposed, located either on a waist belt, a vest or gloves. However, from the VI user’s perspective, all tactile display frameworks are considered invasive because they require an actual physical contact with the subject. In addition, the main limitation of haptic devices is related to its low resolution capability. For this reason, the tactile interfaces are only suitable for limited information feedback. As indicated in [99], the majority of blind and visually impaired people prefer the speech interface when using a navigation assistant.

Based on this observation, various systems adopt the acoustic feedback in order to transmit the warning messages to the VI user. In the state of the art various strategies are proposed, including beeping sequences [43], pre-recorded verbal messages [50,51,56,60], text to speech transcripts [57,83,95,93], musical sounds [71], remote voice commands [73] or binaural acoustic feedback [74]. However, the VI people showed a set of concerns when using regular headphones because they block other auditory cues and warnings from the environment. In [46,47,88] bone conducting headphones are adopted, since they make it possible to avoid any obstruction with the sense of hearing.

The analysis of the state of the art also shows that relatively few frameworks have considered the issue of prioritizing the warning messages that need to be transmitted to the user. However, such an issue is essential if we want to avoid overwhelming the user with useless information, and needs to be carefully treated. For experimental testing/validation purposes, some systems employ normal or blindfolded persons to complete a navigational task. However, we consider that the experimental results are not relevant and some additional test with real VI user in real life scenarios needs to be performed. Moreover, the frameworks performing the test on prerecorded videos should extend the evaluation stage and present the feedback of real VI users after using the proposed prototype.

4. Qualitative evaluation

The requirements, needs and wishes of end users are the backbone of any assistive device designed to facilitate the live and inclusion of any person suffering from a disability. In order to offer a qualitative evaluation of the systems from the state of the art we developed a structured interview with closed questions.

We consider the interview the most appropriate element for gathering user requirements. Structured interviews have some positive features: they allow a rapid analysis of the questions asked, are consistent across participants and are more informative than unstructured interviews.

After an extensive consultation/discussion with several groups/associations of visually impaired users, researchers and software developers, the requirements were collected, compiled and analyzed. In this way, we have established a set of features that a wearable assistive device should present in order to be accepted by the blind community. The results of our analysis are summarized in Table 2.

The identified structural elements can be considered as basic, operational features that any assistive device must incorporate in order to be accepted/used by VI/blind users during navigation in indoor/outdoor unknown environments.

Table 1
Structural features existent in computer vision based assistive device.

	System name / Year of publication	Main Functionality	High Level Features	Feedback	Participants	Test - Task & Environment
Monocular camera based systems	Trinetra[56] / 2007	NO ETAs functions	Mobile shopping assistant / Barecode scanner	Pre-recorded audible messages	Blinds	Identify grocery in campus store
	ShopMobile[55] / 2010	NO ETAs functions	Mobile shopping assistant / Barecode scanner	Haptic and audio	Blindfolded subjects	Laboratory study
	Grijalva et al. [60] / 2010	NO ETAs functions	Banknote recognizer	Acoustic messages	Blindfolded subjects	218 banknotes of different denominations
	Smart Vision[44] / 2011	Navigational assistant	Known landmark detection / Object identification in pantry or refrigerator	Vibration actuator and speech synthesis	Blinds	Walking on predefined paths marked RFID tags
	Mobile Vision[43] / 2012	Sign-based wayfinding system	Detect specific color markers	Acoustic signals - beep sequences	Blinds	Wayfinding in known buildings based on landmarks
	Molina et al. [57] / 2012	Local orientation and navigation	User localization / Context augmentation / OCR and signs identification	Text-to-Speech	Not specified	5 s videos of known indoor scenes
	Hasanuzzaman et al. [59] / 2012	NO ETAs functions	Banknote recognizer	Acoustic messages	Blinds	579 positive images with US dollars
	Tapu et al. [46] / 2013	Navigational assistant	Obstacles detection and classification (pedestrians, cars, bikes, obstructions)	Acoustic warnings - boneconduction headphones	Blinds	Walking in real life crowded urban scenes
	Tian et al. [50] / 2013	Navigational assistant	Doors, elevators, and cabinets detection / Optical character recognition	Verbal messages	Blinds	Dataset of 221 images acquired with the help of blinds
	Crosswatch[52] / 2013	Guidance in traffic intersections	Crosswalk detection / Walk light status identification	Audio tones	Blinds	Crossing two intersections
	Yang et al. [51] / 2014	NO ETAs functions	Clothes pattern recognition	Verbal messages	Blinds	Texture classification
	Arianna[42] / 2016	Navigational assistant	Walking path identification / User position estimation	Smartphone vibrations	Blindfolded subjects	Indoor navigation scenarios / Simulated environments
	Mocanu et al. [47] / 2016	Navigational assistant	Obstacles detection and classification (pedestrians, cars, bikes, obstructions)	Acoustic warnings - boneconduction headphones	Blinds	Walking in real life crowded urban scenes
	TL_recognizer[54] / 2016	NO ETAs functions	Traffic lights detection / Crosswalk detection	Speech synthesis / Mobile vibration patterns	Blinds and low vision subjects	Traffic light detection in real urban intersection
	Mechatronic[64] / 2018	Navigational assistant during walking and running	Line or lane detection on the runway	Haptic device - two gloves with vibration motors	Normal users	Running on a known marked path
Stereo camera based systems	SVETA[71] / 2007	Navigational assistant	Obstacle detection and distance estimation	Stereo musical sounds and voice commands	Blinds	Walking in unknown environments
	Robot Vision[68] / 2010	Navigational assistant	Ground plane estimation / Obstacle detection	Haptic feedback vest	Blindfolded subjects	Walking in unknown environments
	SmartVision[72] / 2010	Navigational assistant	Obstacle detection / Zebra-crossing identification / Building entrance detection	Haptic device - vibration actuators	Normal users	Recorded videos captured within the campus
	Lin et al. [73] / 2012	Navigational assistant	Obstacle detection / Video streaming through the 3 G network	Audio feedback and remote guidance from a sighted person	Normal users	Simulated environment
	Smart Walker [67] / 2014	Navigational assistant	Detect obstacles and hazardous ground changes / Cabinets detection	Not specified	Not specified	Static stereo pictures acquired with the hardware device
	Leung et al. [70] / 2014	Navigational assistant	Ground plane estimation / Egomotion estimation	No feedback	Not tested with users	Crowded known routes with an average length of 250 m

(*continued on next page*)

Table 1 (*continued*)

	System name / Year of publication	Main Functionality	High Level Features	Feedback	Participants	Test - Task & Environment
	Schmarze et al. [74] / 2015	Navigational assistant	Obstacle detection / Plane estimation / Egopose estimation	Binaural acoustic feedback	Visually impaired	Navigate towards the obstacles and pass between them to validate the sound localization concept.
	Everdinget al. [69] / 2016	Navigational assistant	Obstacle detection	Spatial auditory signals	Normal users	20 trails performed by 11 subjects to detect and localize obstacle and determine their size
RGB-D camera based systems	NAVIG [94] / 2011	Navigational assistant	Landmark detection	Not specified	Normal users	Outdoor navigation with geolocated landmarks
	Object sonification[75] / 2012	Navigational assistant	Auditory augmented reality / Ground plane estimation	Sonify recognized objects	Normal users	Sighted users listened to an auditory rendering of various scenes
	KinDetect[76] / 2012	Navigational assistant	Obstacle and human detection	Text-to-Speech	Blindfolded subjects	Walking in known indoor environments
	Filipe et al. [77] / 2012	Navigational assistant	Obstacle detection / Stairs detection	Not specified	Normal users	1200 images recorded in indoor environment
	Kinect cane[78] / 2013	NO ETAs functions	Object recognition based on user demand	Tactile device	Blindfolded subjects	Recognize floors, chairs, upward and downward stairs
	Tian et al. [91] / 2014	Navigational assistant	Detection of stairs and pedestrian crosswalks	Acoustic messages	Not specified	A dataset with 106 images of stairs and 52 crosswalks
	Neto et al. [84] / 2015	NO ETAs functions	Face recognition	3D audio messages, pre-recorded	Blindfolded and blind subjects	Testing on 600 video of 30 peoples
	Blessenohl et al. [89] / 2015	Navigational assistant	Walking path identification Floor and side walls detection	Binaural acoustic feedback	Blindfolded subjects	Participants were asked to navigate along routes in real floor layout
	INSANA[85] / 2016	Navigational assistant	High level semantic map of the building architecture / Obstacle detection	Speech - audio messages	Blindfolded and blind subjects	Indoor navigation within buildings with previous known architecture
	Munoz et al. [86] / 2016	Navigational assistant	Staircase and corridor detection	Text to speech	Normal users	A database of 115 upstairs, 111 downstairs and 120 negative data
	Lee *at al* [90] / 2016	Navigational assistant	Egomotion estimation, mappingand path planning / Obstacle avoidance	Haptic feedback vest	Visually impaired /Blindfolded subjects	Navigation in buildings guided by sighted people / Pose estimation while walking from point A to B
	Chessa et al. [93] / 2016	Navigational assistant	Semantic annotation and interpretation of the 3D scene / Face recognition / Text detection and recognition / Banknote recognition	Text to speech	Not specified	Various datasets depending on the module
	Sound of Vision[95] / 2016	3D representation of theenvironment	Obstacle detection	Hearing and tactil-erepresentation	Blinds	Test in virtual environment and real world setups
	Ramer et al. [98] / 2016	Navigational assistant during walking and running	Course/lane detection / Collision avoidance	Haptic feedback vest	Blindfolded subjects	People were asked to walk and jog for one lap of a stadium
	Jafri et al. [88] / 2017	Navigational assistant	Ground plane detection / Obstacle detection	Acoustic warnings - boneconduction headphones	Blindfolded subjects	Simulated scenarios at various illumination conditions with different obstacles

Table 2
Operational features of wearable assistive devices.

Abbr.	Characteristics	Description
F1	*Processing speed*	The system should function in real-time and to alert immediately the user about an obstacle situated at a distance inferior to 1.5 m
F2	*Usability*	The system should function in both, indoor and outdoor scenes
F3	*Robustness*	The system should not by influenced about the scene dynamics or lighting conditions
F4	*Coverage distance*	The maximum distance between the user and an object for which detection can be performed
F5	*Obstacle detection*	The system should be able to detect any type of object regardless its position, shape, size or the object dynamics
F6	*Portability*	The system should be light, ergonomic and easy to wear
F7	*Friendliness*	The system should be easy to learn without any medium or long-term dedicated training

Table 3
Comparison of the assistive devices dedicated to visually impaired users that incorporate computer vision algorithms.

	System name / Year of publication	Processing speed	Usability	Robustness	Coverage distance	Obstacle detection (Shape / Size / Static / Dynamic / Position)	Portability	Friendly
Monocular camera based systems	Smart Vision[44] / 2011	5 fps	Indoor / Outdoor	Low	2 m	Any / Any / Static & Dynamic / Ground level	Yes	Moderate
	MobileVision[43] / 2012	8 fps	Indoor	High	3.5 m	None	Yes	Moderate
	Molina et al. [57] / 2012	Not specified	Indoor	Low	3 m	None	No	N/A
	Tapu et al. [46] / 2013	7 fps	Indoor / Outdoor	High	2 m	Any / Any / Static &Dynamic / Any	Yes	High
	Tian et al. [50] / 2013	Not specified	Indoor	Moderate	Not specified	Geometric specific / Large / Static / Any	Yes	Low
	Crosswatch[52] / 2013	Not specified	Outdoor	Low	1 m	None	Yes	Low
	Arianna[42] / 2016	Not specified	Indoor	High	5 m	None	Yes	Moderate
	Mocanu et al. [47] / 2016	10 fps	Indoor / Outdoor	High	5 m	Any / Any / Static &Dynamic / Any	Yes	Moderate
	Mechatronic[64] / 2018	20 fps	Outdoor	Moderate	10 m	None	Yes	Low
Stereo camera based systems	SVETA[71] / 2007	0.8 fps	Indoor	Moderate	3 m	Any / Any / Static &Dynamic / Any	Yes	Moderate
	Robot Vision[68] / 2010	10 fps	Indoor / Outdoor	Moderate	3 m	Any / Any / Static / Any	No	N/A
	SmartVision[72] / 2010	Not specified	Indoor / Outdoor	Low	2 m	Geometric specific / Small / Static /Ground level	No	N/A
	Lin et al. [73] / 2012	0.5 fps	Indoor / Outdoor	Low	Not specified	Any / Any / Static &Dynamic / Any	Yes	Low
	Smart Walker [67] / 2014	4 fps	Indoor / Outdoor	Moderate	2 m	Specific (Known) / Any / Static / Any	No	N/A
	Leung et al. [70] / 2014	1 fps	Outdoor	Moderate	Not specified	None	No	N/A
	Schmarze et al. [74] / 2015	15 fps	Outdoor	High	10 m	Any / Any / Static &Dynamic / Any	Yes	Moderate
	Everding et al. [69] / 2016	20 fps	Indoor	Moderate	6 m	Any / Any / Static &Dynamic / Any	Yes	Moderate
RGB-D camera based systems	NAVIG[94] / 2011	5 fps	Outdoor	Low	1.5 m	None	Yes	Moderate
	Object sonification[75] / 2012	Not specified	Indoor	Moderate	5 m	Any / Any / Static / Ground level	Yes	Moderate
	KinDetect[76] / 2012	Not specified	Indoor	Low	1.5 m	Any / Medium to large / Static / Any	Yes	Moderate
	Filipe et al. [77] / 2012	Not specified	Indoor	High	2 m	Specific / Large / Static / Any	No	N/A
	Tian et al. [91] / 2014	5 fps	Indoor / Outdoor	High	4.7 m	None	No	Low
	Blessenohl et al. [89] / 2015	15 fps	Indoor	Medium	2.5 m	Any / Any / Static &Dynamic / Any	Yes	Low
	INSANA[85] / 2016	5 fps	Indoor	High	3 m	Any / Any / Static / Any	Yes	High
	Munoz et al. [86] / 2016	2 fps	Indoor	Medium	3.5 m	None	No	N/A
	Lee et al. [90] / 2016	28.6 fps	Indoor	High	5 m	Any / Medium to large / Static / Any	Yes	Low
	Chessa et al. [93] / 2016	Not specified	Indoor / Outdoor	High	1.5 m	Geometric specific / Any / Static /Any	No	Low
	Sound of Vision[95] / 2016	10 fps	Indoor / Outdoor	High	5–10 m	Any / Any / Static &Dynamic / Any	Yes	Low
	Ramer et al. [98] / 2016	11 fps	Outdoor	Low	1.5 m	None	Yes	Low
	Jafri et al. [88] / 2017	Not specified	Indoor	Medium	2 m	Any / Any / Static / Any	Yes	Moderate

Table 4
Scores and experimental evaluation for each system considered.

	Feature	Processing speed	Usability	Robustness	Coverage distance	Obstacle detection	Portability	Friendly	Global score (after weighting)
	Weight	**0.18**	**0.12**	**0.21**	**0.16**	**0.14**	**0.10**	**0.09**	
	System name / Year of publication	**Scores**							
Monocular camera based systems	Smart Vision[44] / 2011	5	10	3	5	10	8	8	6.45
	MobileVision[43] / 2012	7	5	8	8	1	9	6	6.4
	Molina et al. [57]. / 2012	3	5	3	7	1	1	1	3.22
	Tapu et al. [46] / 2013	7	10	8	5	10	9	9	8.05
	Tian et al. [50]. / 2013	3	5	6	5	5	6	3	4.77
	Crosswatch[52] / 2013	7	5	2	3	1	3	3	3.47
	Arianna[42] / 2016	8	5	9	10	1	8	7	7.1
	Mocanu et al. [47]. / 2016	8	10	8	10	10	7	8	**8.74**
	Mechatronic[64] / 2018	10	5	5	10	1	4	3	5.86
Stereo camera based systems	SVETA[71] / 2007	1	5	4	7	10	7	7	5.47
	Robot Vision[68] / 2010	8	10	6	7	8	1	1	6.33
	SmartVision[72] / 2010	8	10	3	5	3	1	1	4.68
	Lin et al. [73]. / 2012	1	10	3	7	10	5	4	5.39
	Smart Walker [67] / 2014	4	10	6	5	6	1	1	5.01
	Leung et al. [70]. / 2014	2	5	5	8	1	1	1	3.62
	Schmarze et al. [74]. / 2015	10	5	8	10	10	7	6	**8.32**
	Everdinget al. [69]. / 2016	10	5	6	10	10	8	6	8
RGB-D camera based systems	NAVIG[94] / 2011	5	5	3	4	1	8	7	4.34
	Object sonification[75] / 2012	4	5	7	10	7	7	5	6.52
	KinDetect[76] / 2012	5	5	3	4	7	8	8	5.27
	Filipe et al. [77]. / 2012	8	5	9	5	5	1	1	5.62
	Tian et al. [91]. / 2014	5	10	7	9	1	1	1	5.34
	Blessenohl et al. [89] / 2015	10	5	5	6	10	5	3	6.58
	INSANA[85] / 2016	5	5	8	7	8	10	9	7.23
	Munoz et al. [86]. / 2016	2	5	5	8	1	1	1	3.62
	Lee et al. [90]. / 2016	10	5	9	10	7	5	3	7.64
	Chessaet al. [93]. / 2016	2	10	8	4	6	1	1	4.91
	Sound of Vision[96] / 2016	8	10	8	10	10	6	3	**8.19**
	Ramer et al. [98]. / 2016	8	5	4	4	1	5	3	4.43
	Jafri et al. [88]. / 2017	6	5	4	5	8	8	5	5.69

Based on the set of criteria presented in Table 2, we introduce in Table 3 a comparative overview of the state of the art systems focusing our attention on computer vision-based approaches. The study provides a systematic evaluation of the ETAs designed specifically as navigational assistants and fulfilling, up to a certain degree, the operational features collected after the interview phase. Using the proposed evaluation parameters, we can estimate the degree of acceptance of any assistive device. In order to globally rate the systems effectiveness we have first allocated a qualitative score (V_i), ranging from 0 to 10, to each feature element (F_i) retained.

A maximum value of 10 is assigned to a fully satisfactory degree of a specific feature. However, some intermediary scores have been used for systems accomplishing only in a partial manner the considered operational feature. Regarding the *processing speed* criterion we have assigned a 10 for systems performing the computation sufficiently fast so that the user can walk normally. In the case of the *usability* feature, we have allocated a value of 5 to a system designed to function solely in indoor or outdoor environments and a 10 for both. Regarding the *robustness* parameter, a value of 10 is proposed only if the accuracy and recall rates are both superior to 99%. Otherwise, the score is correlated with the number of false positives and negatives. Thus, systems with precision and recall scores inferior to 50% have been considered as highly sensitive and have been consequently assigned a *low* value for the robustness parameter. Frameworks with the accuracy scores inferior to 75% are considered with *moderate* robustness, while a F1-norm score superior to 75% is considered for *high* robust systems.

For the *coverage distance* we have assigned a value of 10 only for systems designed to inform the VI user about any type of static/dynamic obstacle situated at a distance superior to 5 m. In the case of *obstacle identification* only the systems detecting any type of obstacle, situated at arbitrary levels of height, without any *a priori* information received a maximum score.

For the *portability* parameter, we have assigned a 10 for wearable and light weighted frameworks.

Finally, for the *friendliness* degree, solely prototypes that can be directly used without any additional training have received a maximum score of 10. In Table 3, systems with the friendliness score inferior to 4 are considered as *low* because they require an extensive training phase with dedicated personnel. Prototypes with scores superior to 8 were assigned a *high* degree of friendliness because they can be used directly by the VI users without any training.

In our case, in order to assign a mark to the relevance of each considered feature, we have used in the evaluation the VI people feedback, provided by authors when testing the device. The results obtained are summarized in Table 4.

In order to obtain a global rate, we have used the opinions of a group of 42 VI users in order to assign a weight to each of the considered feature, based on its perceived impact on the overall system performance. The weighting mechanism has been developed based on the subjective perception of the VI users. Thus, each user was asked to prioritize the 7 features presented in Table 2 in descending order of importance. Then, we summed the scores and normalized them (using the number of VI user involved and the

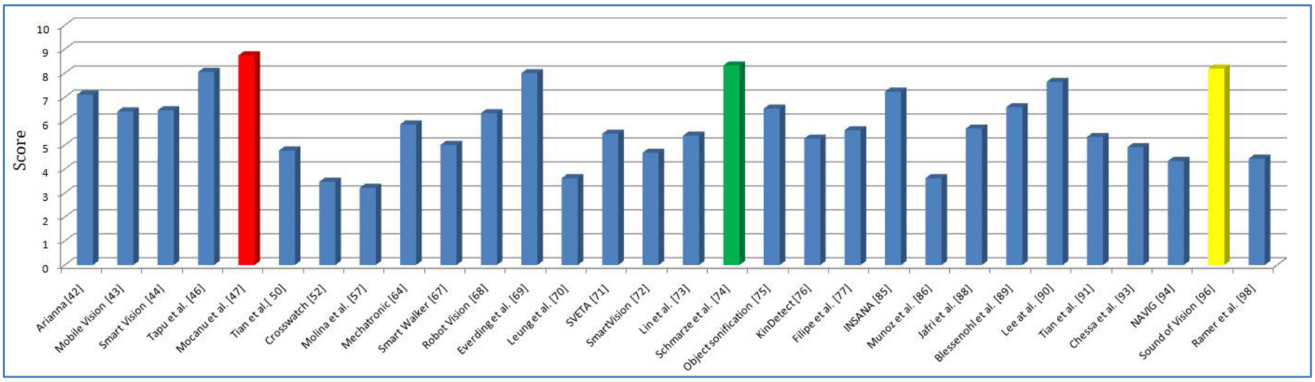

Fig. 11. The performance scores obtained for each system considered in the evaluation.

total number of features). The resulting weights are also reported in Table 4.

Finally, the global rate associated with a given system is defined as described in Eq. (1):

$$Global\ Score = \frac{\sum_{i=1}^{N} w_i \cdot F_i}{N}, \quad (1)$$

where F_i represents the score assigned to the i^{th} feature, N is the total number of characteristics used in the evaluation and w_i is the weight assigned to each feature. Fig. 11 presents an overall picture of the performances achieved by the considered assistive devices. Systems with higher scores such as Mocanu et al. [47] and Schmarze et al. [74] and Sound of Vision [96] propose state of the art, solid and robust performances.

From the results presented in Fig. 11 we can observe that none of the systems are able to completely satisfy all the criteria imposed by the blind community and return a global score superior to 9.5. The highest performance of 8.74 (the red bar) is obtained by the Mocanu et al. [47] that proposed a system based on two independent sources of information: ultrasonic sensors and the video camera embedded on a regular smartphone device. The second performance score (8.32 the yellow bar) is obtained by the Sound of Vision [96] that uses a RGB-D camera and provides an audio and tactile representation of the surrounding scene. While, the third score (8.19 the green bar) is obtained by Schmarze et al. [74] that introduced a framework based on dense disparity maps combined with visual odometry and inertial measurements units.

5. Conclusions and discussion

In this paper, we have proposed a survey of wearable and portable assistive devices dedicated to blind and visually impaired users, while providing a deep and critical analysis of related strengths and limitations.

After analyzing the state of the art developments in the area of assistive devices dedicated to blind/VI users, we have observed that despite promising potential and many years of research, sensorial/camera based substitution devices have not been widely adopted [100]. A problem for VI people using such systems is the lack of ability to developing cognitive maps of the environment [101]. In order to tackle this issue, several research works recently reported aim at recognizing places [102], in order to connect habitual actions / locations [103] for assistive purposes (*e.g.*, visually impaired, memory diseases…). Solely a few wearable assistive devices have been used outdoors in real crowded scenes and not in simulated environments (*i.e*., in laboratory conditions), controlled by researchers. To the best of our knowledge none of the state of the art devices have been adopted by the blind community so far [3,104]. The underlying problem is two-folded. First, there are some limitations regarding the accuracy and precision of the proposed architectures [105]. Moreover, there is an important factor constraining their development and acceptance, that is related to the limitation of the visual rehabilitation programs in general.

In the past, the development of wearable assistive devices faced significant stumbling blocks: the devices were difficult to operate by the VI users, expensive and not sufficiently effective to be tested in real world. In addition, socio-psychological factors such as the reluctance to novel technologies have also hampered their progress and adaptation. However, from our perspective the biggest drawback of the proposed technologies is the lack of organized training procedures. Within this context, the VI users have to train themselves without a trainer or a clear set of instruction to follow.

In the last couple of years, the researches dedicated to the development of a sensorial substitution to the human vision have flourished [106,107]. It has been demonstrated that commonly-used mobile platforms offer a large variety of assistive applications by using the built-in sensors of the mobile devices, and combining this sensory information with cloud resources in real-time [108]. Unfortunately, such devices contribute only partially to the VI people integration into the society, since most systems were tested solely in a laboratory, controlled environment, under specific training experimental conditions, while not addressing the problem of closing the sensorial-motor loop.

Approaches based on computer vision and machine learning technologies have evolved swiftly, but are still far from approaching the regular human level of semantic content understanding of the environment. We argue that a carefully designed hybridization between various technologies such as: computer vision, GPS, Wi-Fi, GIS integrated in a wearable device (*e.g.,* a regular smartphone) with voice recognition capabilities can provide a single, versatile and hands-free assistive device that can be accepted by the blind community.

References

[1] World Health Organization (WHO), Visual impairment and blindness, accessed on (23.05.2018) http://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment.
[2] R. Imrie, P. Hall, Inclusive Design: Designing and Developing Accessible Environments, Spon Press, London, 2001.
[3] S. Maidenbaum, S. Abboud, A. Amedi, Sensory substitution: closing the gap between basic research and widespread practical visual rehabilitation, NeurosciBiobehav Rev 41 (2014) 3–15.
[4] D. Dakopoulos, N.G. Bourbakis, Wearable obstacle avoidance electronic travel aids for blind: a survey, IEEE Trans. Syst. Man Cybern. Part C (Appl. Rev.) 40 (1) (2010) 25–35.
[5] O. Lahav, D.W. Schloerb, S. Kumar1, M.A. Srinivasan, A virtual environment for people who are blind—a usability study, J. Assist Technol 6 (1) (2012) 1–21.
[6] M. Leo, G. Medioni, M. Trivedi, T. Kanade, G. Farinella, Computer vision for assistive technologies, Comput. Vision Image Underst. 148 (2016) 1–5.

[7] A. Csapo, G. Wersenyi, M. Jeon, A survey on hardware and software solutions for multimodal wearable assistive devices targeting the visually impaired, Acta Polytechnica Hungarica 13 (5) (2016) 39–63.
[8] W. Elmannai, K. Elleithy, Sensor-based assistive devices for visually-impaired people: current status, challenges and future directions, Sensors 17 (3) (2017) 1–42 565.
[9] G.F. Lourenço, A.C. Honório, A.O. Figueiredo, Satisfaction of use of assistive devices for orientation and mobility of adults with visual impairment, J. Occup. Therapy Univ. São Paulo 28 (3) (2017) 340–348.
[10] M. Salerno, M. Re, A. Cristini, G. Susi, M. Bertola, E. Daddario, F. Capobianco, AudiNect: an aid for the autonomous navigation of visually impaired people, based on virtual interface, Int. J. Human Comput. Interact. 4 (1) (2013) 25–33.
[11] V.R. Schinazi, T. Thrash, D.R. Chebat, Spatial navigation by congenitally blind individuals, Wiley Interdiscip. Rev. Cogn. Sci. 7 (2016) 37–58.
[12] J.M. Loomis, R.L. Klatzky, N.A. Giudice, Sensory substitution of vision: importance of perceptual and cognitive processing, in: R Manduchi, S Kurniawan (Eds.), Assistive Technology for Blindness and Low Vision, CRC Press, Boca Raton, FL, 2012, pp. 162–191.
[13] S.L. Phung, M.C. Le, A. Bouzerdoum, Pedestrian lane detection in unstructured scenes for assistive navigation, Comput. Vision Image Underst. 149 (2016) 186–196.
[14] L. KeyR.-G. Busnel, J.F. Fish (Eds.), Air sonars with acoustic display of spatial information, Animal Sonar Systems (1980) 769–816.
[15] A.J. Ali, A.H. Sankar, Artificial guidance system for visually impaired people, Bonfring Int. J. Man Mach. Interface 2 (1) (2012) 10–12.
[16] N. Pressey, Mowat sensor, Focus 11 (3) (1977) 35–39.
[17] R. Kuc, Binaural sonar electronic travel aid provides vibrotactile cues for landmark, reflector motion and surface texture classification, IEEE Trans. Biomed. Eng. 49 (10) (2002) 1173–1180.
[18] I. Ulrich, J. Borenstein, The guidecane: applying mobile robot technologies to assist the visually impaired, IEEE Trans. Syst. Man, Cybern. A Syst. Hum. 31 (2) (2001) 131–136.
[19] S. Cardin, D. Thalmann, F. Vexo, A wearable system for mobility improvement of visually-impaired people, Vis. Comput. 23 (2007) 109–118.
[20] L. Akita, T. Komatsu, K. Ito, T. Ono, M. Okamoto, CyARM: haptic sensing device for spatial localization on basis of exploration by arms, Adv. Human-Comp. Interact. 2009 (2009) 1–6.
[21] G.-Y. Jeong, K.H. Yu, Multi-section sensing and vibrotactile perception for walking guide of visually impaired person, Sensors 16 (7) (2016) 1070.
[22] D. Ross, A. Lightman, Talking braille: a wireless ubiquitous computing network for orientation and wayfinding, in: Proceedings of the Seventh International ACM SIGACCESS Conference on Computers and Accessibility, 2005, pp. 98–105.
[23] K. Yelamarthi, D. Haas, D. Nielsen, S. Motherse, RFID and GPS integrated navigation system for the visually impaired, in: Proceedings of the 53rd IEEE International Midwest Symposium on Circuits and Systems, 2010, pp. 1149–1152.
[24] N. Márkus, A. Arató, Z. Juhász, G. Bognár, L. Késmárki. MOST-NNG: an accessible GPS navigation application integrated into the MObile Slate Talker (MOST) for the blind. In Computers Helping People with Special Needs, Lect. Notes Computer Science, 2010, pp. 247–254.
[25] M. Bessho, S. Kobayashi, N. Koshizuka, K. Sakamura, Assisting mobility of the disabled using space-identifying ubiquitous infrastructure, in: Proceedings of the Tenth International ACM SIGACCESS Conference on Computer and Accessibility, 2008, pp. 283–284.
[26] A.S. Martines-Sala, F. Losilla, J.C. Sanchez-Aarnoutse, J. Garcia-Haro, Design, implementation and evaluation of an indoor navigation system for visually impaired people, Sensors 15 (2015) (2015) 32168–32187.
[27] A. Wahab, M. Helmy, A.A. Talib, H.A. Kadir, A. Johari, A. Noraziah, R.M. Sidek, A.A. Mutalib, SmartCane: assistive cane for visually-impaired people, Int. J. Comput. Sci. 8 (4) (2011) 21–27.
[28] W. Penrod, M.D. Corbett, B. Blasch, A master trainer class for professionals in teaching the UltraCane electronic travel device, J. Visual Impairment Blindness 99 (11) (2005) 711–714.
[29] B. Hoyle, D. Waters, Mobility at: the batcane (ultracane), in: Assistive Technology for Visually Impaired and Blind People, Springer, London, UK, 2008, pp. 209–229.
[30] L.H. Villamizar, M. Gualdron, F. Gonzalez, J. Aceros, C.V. Rizzo-Sierra, A necklace sonar with adjustable scope range for assisting the visually impaired, in: Proceedings of the 35th Annual International Conference IEEE Engineering Medical Biology Society, 2013, pp. 1450–1453.
[31] R. Tapu, B. Mocanu, T.B. Zaharia, ALICE: a smartphone assistant used to increase the mobility of visual impaired people, J. Ambient Intell. Smart Environ. 7 (5) (2015) 659–678.
[32] V. Kulyukin, C. Gharpure, J. Nicholson, RoboCart: toward robot-assisted navigation of grocery stores by the visually impaired, in: Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, Edmonton, AB, Canada, 2005, pp. 2845–2850.
[33] D. Lopez-de-Ipina, T. Lorido, U. Lopez, Indoor navigation and product recognition for blind people assisted shopping, in: Ambient Assisted Living, 6693, Springer, Berlin/Heidelberg: Málaga, Spain, 2011, pp. 33–40.
[34] S. Chumkamon, P. Tuvaphanthaphiphat, P. Keeratiwintakorn, A blind navigation system using RFID for indoor environments, in: Proceedings of the 5th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, 2, „ KrabiThailand, 2008, pp. 765–768.
[35] A. Meliones, D. Sampson, Blind MuseumTourer: a system for self-guided tours in museums and blind indoor navigation, Technologies 6 (1) (2018) 1–31.
[36] D. Ahmetovic, C. Gleason, C. Ruan, K. Kitani, H. Takagi, C. Asakawa, Navcog: A navigational cognitive assistant for the blind, in: Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services, 2016, pp. 90–99.
[37] T. Ishihara, J. Vongkulbhisal, K.M. Kitani, C. Asakawa, Beacon-guided structure from motion for smartphone-based navigation, in: IEEE Winter Conference on Applications of Computer Vision (WACV), Santa„ RosaCA, 2017, pp. 769–777.
[38] L. Hakobyan, J. Lumsden, D. O'Sullivan, H. Bartlett, Mobile assistive technologies for the visually impaired, Surv. Ophthalmol. 58 (6) (2013) 513–528.
[39] J. Nicholson, V. Kulyukin, D. Coster, Shoptalk: independent blind shopping through verbal route directions and barcode scans, Open Rehabil. J. 2 (2009) 11–23.
[40] J. Bai, S. Lian, Z. Liu, K. Wang, D. Liu, Smart guiding glasses for visually impaired people in indoor environment, IEEE Trans. Consum. Electron. 63 (3) (2017) 258–266.
[41] R. Tapu, B. Mocanu, T. Zaharia, DEEP-SEE: joint object detection, tracking and recognition with application to visually impaired navigational assistance. Sensors, vol. 17(11), pp. 2473.
[42] D. Croce, L. Giarré, F.G. Rosa, E. Montana, I. Tinnirello, Enhancing tracking performance in a smartphone-based navigation system for visually impaired people, in: Proceedings of the 24th Mediterranean Conference on Control and Automation (MED), Athens, 2016, pp. 1355–1360.
[43] R. Manduchi, Mobile vision as assistive technology for the blind: An experimental study, in: Proceedings of the 13th International Conference on Computers Helping People with Special Needs, 2012, pp. 9–16.
[44] J.M.H. Buf, J. Barroso, J.M.F. Rodrigues, H. Paredes, M. Farrajota, H. Fernandes, J. Jose, V. Teixeira, T. Saleiro, The smartvision navigation prototype for blind users, Int. J. Digit. Content Technol. Appl. (2011) 361–372.
[45] S. Kammoun, M. Mac, B. Oriola, C. Jouffrais, Towards a geographic information system facilitating navigation of visually impaired users. In LNCS, vol. 7383 LNCS, in: Lecture Notes in Computer Science, Part 2, IEEE, Berlin, Germany, 2012, pp. 521–528.
[46] R. Tapu, B. Mocanu, A. Bursuc, T. Zaharia, A smartphone-based obstacle detection and classification system for assisting visually impaired people, in: Proceedings of the IEEE International Conference on Computer Vision Workshops, Sydney, 2013, pp. 444–451.
[47] B. Mocanu, R. Tapu, T. Zaharia, When ultrasonic sensors and computer vision join forces for efficient obstacle detection and recognition, Sensors 16 (2016) 1807.
[48] B. Lucas, T. Kanade, An iterative technique of image registration and its application to stereo, in: Proceedings of the IJCAI'81 7th International Joint Conference on Artificial Intelligence, 2, Vancouver, BC, USA, 1981, pp. 674–679.
[49] J.J. Lee, G. Kim, Robust estimation of camera homography using fuzzy RANSAC, in: Proceedings of the International Conference on Computational Science and its Applications, „ Kuala Lumpur, Malaysia, 2007, pp. 992–1002.
[50] Y.L. Tian, X. Yang, C. Yi, A. Arditi, Toward a computer vision-based way finding aid for blind persons to access unfamiliar indoor environments, Mach. Vis. Appl. 24 (2013) 521–535.
[51] X. Yang, S. Yuan, Y. Tian, Assistive clothing pattern recognition for visually impaired people, IEEE Trans. Human–Mach. Syst. 44 (2) (2014) 234–243.
[52] M. Coughlan, H. Shen, Crosswatch: a system for providing guidance to visually impaired travelers at traffic intersections, J. Assist Technol. 7 (2) (2013) 131–142.
[53] S. Mascetti, D. Ahmetovic, A. Gerino, C. Bernareggi, M. Busso, A. Rizzi, Supporting pedestrians with visual impairment during road crossing: a mobile application for traffic lights detection, in: Proceedings of the 15th International Conference On Computers Helping People With Special Needs, Springer International Publishing, Cham, 2016, pp. 198–201.
[54] S. Mascetti, D. Ahmetovic, A. Gerino, C. Bernareggi, M. Busso, A. Rizzi, Robust traffic lights detection on mobile devices for pedestrians with visual impairment, Comput. Vision Image Underst. 148 (2016) 123–135.
[55] V.A. Kulyukin, A. Kutiyanawala, Demo: ShopMobile II: Eyes-free supermarket grocery shopping for visually impaired mobile phone users, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, „ San Francisco, CA, 2010, pp. 31–32.
[56] P.E. Lanigan, A.M. Paulos, A.W. Williams, D. Rossi, P. Narasimhan, Trinetra: Assistive technologies for grocery shopping for the blind, in: Proceedings of the IEEE-BAIS Symposium Research on Assistive Technologies, 2007, pp. 147–148.
[57] E. Molina, Z. Zhu, Y. Tian, Visual nouns for indoor/outdoor navigation, in: Lecture Notes in Computer Science, 7383, Springer, Berlin Heidelberg, 2012, pp. 33–40.
[58] F. Hu, Z. Zhu, J. Zhang, Mobile panoramic vision for assisting the blind via indexing and localization, in: Proceedings of the 2nd Workshop Assistive Computer Vision Robotics Conjunction (ECCV), 2014, pp. 600–614.
[59] F.M. Hasanuzzaman, X. Yang, Y. Tian, Robust and effective component-based banknote recognition for the blind, IEEE Trans. Syst. Man Cybern. C Appl. Rev. 42 (2012) 1021–1030.
[60] F Grijalva, J.C. Rodriguez, J. Larco, L. Orozco, Smartphone recognition of the U.S. banknotes' denomination, for visually impaired people, in: Proceedings of the IEEE ANDESCON Conference, „ Bogota, Colombia, 2010, pp. 1–6.
[61] K. Manikandan, T. Sumithra, Currency recognition in mobile application for visually challenged, Discovery 30 (2015) 245–248.
[62] K. Springer, P. Subramanian, T. Turton, Australian banknotes: Assisting people with vision impairment, Bull. Reserve Bank (2015) 1–12.

[63] D. Mulmule, A. Dravid, A study of computer vision techniques for currency recognition on mobile phone for the visually impaired, Int. J. Adv. Res. Comput. Sci. Softw. Eng. 4 (2014) 160–165.
[64] A Mancini, E. Frontoni, P. Zingaretti, in: Mechatronic system to help visually impaired users during walking and running, 19, 2018, pp. 649–660.
[65] A. Aladrén, G. López-Nicolás, L. Puig, J.J. Guerrero, Navigation assistance for the visually impaired using RGB-D sensor with range expansion, IEEE Syst. J. 10 (3) (2016) 922–932.
[66] D. Gutiérrez-Gómez, L. Puig, J.J. Guerrero, Full scaled 3D visual odometry from a single wearable omnidirectional camera, in: Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Vilamoura, 2012, pp. 4276–4281.
[67] S. Cloix, V. Weiss, G. Bologna, T. Pun, D. Hasler, Obstacle and planar object detection using sparse 3D information for a smart walker, in: Proceedings of the International Conference on Computer Vision Theory and Applications (VISAPP), „ Lisbon, Portugal, 2014, pp. 292–298.
[68] V. Pradeep, G. Medioni, J. Weiland, Robot vision for the visually impaired, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Workshops, San Francisco, CA, 2010, pp. 15–22.
[69] L. Everding, L. Walger, V.S. Ghaderi, J. Conradt, A mobility device for the blind with improved vertical resolution using dynamic vision sensors, in: Proceedings of the IEEE 18th International Conference on e-Health Networking, Applications and Services (Healthcom), Munich, 2016, pp. 1–5.
[70] T.S. Leung, G. Medioni, Visual navigation aid for the blind in dynamic environments, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2014, pp. 579–586.
[71] G. Balakrishnan, G. Sainarayanan, R. Nagarajan, S. Yaacob, Wearable real-time stereo vision for the visually impaired, Eng. Lett. 14 (2) (2007) 6–14.
[72] H. Fernandes, P. Costa, V. Filipe, L. Hadjileontiadis, J. Barroso, Stereo vision in blind navigation assistance, in: Proceedings of the World Automation Congress, 2010, pp. 1–6.
[73] K. Lin, T. Lau, C. Cheuk, Y. Liu, A wearable stereo vision system for visually impaired, in: Proceedings of the IEEE International Conference Mechatronics Automation, 2012, pp. 1423–1428.
[74] T Schwarze, M. Lauer, M. Schwaab, M. Romanovas, S. Böhm, T. Jürgensohn, An intuitive mobility aid for visually impaired people based on stereo vision, in: Proceedings of the IEEE International Conference on Computer Vision Workshop (ICCVW), Santiago, 2015, pp. 409–417.
[75] F. Ribeiro, D. Florencio, P. Chou, Z. Zhang, Auditory augmented reality: Object sonification for the visually impaired, in: Proceedings of the IEEE 14th International Workshop Multimedia Signal Process, 2012, pp. 319–324.
[76] A. Khan, F. Moideen, J. Lopez, W. Khoo, Z. Zhu, KinDectect: Kinect detecting objects, in: Proceedings of the 13th International Conference Computer Helping People Special Needs, 2012, pp. 588–595.
[77] V. Filipe, F. Fernandes, H. Fernandes, A. Sousa, H. Paredes, J. Barroso, Blind navigation support system based on Microsoft Kinect, in: Proceedings of the 4th International Conference Software Development Enhancing Accessibility Fighting Info-exclusion, 2012, pp. 94–101.
[78] H. Takizawa, S. Yamaguchi, M. Aoyagi, N. Ezaki, S. Mizuno, Proceedings of the 6th International Conference on Human System Interactions (HSI), Sopot (2013) 473–478.
[79] B. Li, A. Mian, W. Liu, A. Krishna, Face recognition based on Kinect., Pattern Anal. Appl. (2015) 1–11.
[80] J.B. Cardia Neto, A. Marana, 3DLBP and HAOG fusion for face recognition utilizing Kinect as a 3D scanner, in: Proceedings of the 30th Annual ACM Symposium Application Computer, 2015, pp. 66–73.
[81] W. Li, A. Mian, W. Liu, A. Krishna, Using Kinect for face recognition under varying poses, expressions, illumination and disguise, in: Proceedings of the IEEE Workshop Application Computer Vision, 2013, pp. 186–192.
[82] G. Goswami, S. Bharadwaj, M. Vatsa, R. Singh, On RGB-D face recognition using Kinect, in: Proceedings of the IEEE 6th International Conference Biometrics: Theory, Application System, 2013, pp. 1–6.
[83] S Berretti, N. Werghi, A. del Bimbo, P. Pala, Selecting stable key-points and local descriptors for person identification using 3D face scans, Vis. Comput. 30 (11) (2014) 1275–1292.
[84] L.B. Neto, F. Grijalva, V. Maike, L. César Martini, D. Florencio, M.C. Baranauskas, R. Anderson, G. Siome, A kinect-based wearable face recognition system to aid visually impaired users, IEEE Trans. Human-Mach. Syst. 47 (2017) 52–64.
[85] B. Li, J.P. Mũnoz, X. Rong, J. Xiao, Y. Tian, A. Arditi, ISANA: wearable context-aware indoor assistive navigation with obstacle avoidance for the blind, in: Proceedings of the ECCV - Workshops, 9914, 2016, pp. 448–462.
[86] R. Munoz, X. Rong, Y. Tian, Depth-aware indoor staircase detection and recognition for the visually impaired, in: Proceedings of the IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Seattle, WA, 2016, pp. 1–6.
[87] R. Jafri, A GPU-accelerated real-time contextual awareness application for the visually impaired on Google's project tango device, J. Supercomput. 73 (2) (2017) 887–889.
[88] R. Jafri, R.L. Campos, S.A. Ali, H.R. Arabnia, Visual and infrared sensor data-based obstacle detection for the visually impaired using the Google Project Tango Tablet Development Kit and the Unity Engine, IEEE Access 6 (2018) 443–454.
[89] S. Blessenohl, C. Morrison, A. Criminisi, J. Shotton, Improving indoor mobility of the visually impaired with depth-based spatial sound, in: Proceedings of the IEEE International Conference on Computer Vision Workshop (ICCVW), Santiago, 2015, pp. 418–426.
[90] Y.H. Lee, G. Medioni, RGB-D camera based wearable navigation system for the visually impaired, Comput. Vision Image Underst. 149 (2016) 3–20.
[91] Y. Tian, RGB-D sensor-based computer vision assistive technology for visually impaired persons, in: Computer Vision and Machine Learning With RGB-D Sensors, Springer, New York, NY, USA, 2014, pp. 173–194.
[92] A.S. Huang, A. Bachrach, P. Henry, M. Krainin, D. Maturana, D. Fox, N. Roy, Visual odometry and mapping for autonomous flight using an RGB-D camera, in: Proceedings of the International Symposium of Robotics Research (ISRR), 2011, pp. 235–252.
[93] M. Chessa, N. Noceti, F. Odone, F. Solari, J. Sosa-García, L. Zini, An integrated artificial vision framework for assisting visually impaired users, Comput. Vision Image Underst. 149 (2016) 209–228.
[94] A. Brilhault, S. Kammoun, O. Gutierrez, P. Truillet, C. Jouffrais, Fusion of artificial vision and GPS to improve blind pedestrian positioning, in: Proceedings of the 4th IFIP International Conference on New Technologies, Mobility and Security (NTMS), Paris, France, 2011, pp. 1–5.
[95] A. Burlacu, S. Bostaca, I. Hector, P. Herghelegiu, G. Ivanica, A. Moldoveanul, S. Caraiman, Obstacle detection in stereo sequences using multiple representations of the disparity map, in: Proceedings of the 20th International Conference on System Theory Control and Computing (ICSTCC), 2016, pp. 854–859.
[96] S. Caraiman, A. Morar, M. Owczarek, A. Burlacu, D. Rzeszotarski, N. Botezatu, P. Herghelegiu, F. Moldoveanu, P. Strumillo, A. Moldoveanu, Computer vision for the visually impaired: the sound of vision system, in: Proceedings of the IEEE International Conference on Computer Vision Workshops (ICCVW), Venice, 2017, pp. 1480–1489.
[97] C. Ramer, C. Ziegler, S. Reitelshöfer, J. Franke, Sensor-guided jogging for visually impaired, in: Proceedings of the 5th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics, Sao Paulo, 2014, pp. 467–472.
[98] C. Ramer, T. Lichtenegger, J. Sessner, M. Landgraf, J. Franke, An adaptive, color based lane detection of a wearable jogging navigation system for visually impaired on less structured paths, in: Proceedings of the 6th IEEE International Conference on Biomedical Robotics and Biomechatronics (BioRob), 2016, pp. 741–746.
[99] R.G. Golledge, J.R. Marston, J.M. Loomis, R.L. Klatzky, Stated preferences for components of a personal guidance system for nonvisual navigation, J. Visual Impair. Blind. 98 (3) (2004) 135–147.
[100] R. Velázquez, Wearable Assistive Devices for the Blind, in: Wearable and Autonomous Biomedical Devices and Systems for Smart Environment, 75, Springer, Berlin Heidelberg, 2010, pp. 331–349.
[101] L. Ran, S. Helal, S. Moore, Drishti: an integrated indoor/outdoor blind navigation system and service, in: Proceedings of the Second IEEE Annual Conference on Pervasive Computing and Communications, Orlando, FL, USA, 2004, pp. 23–30.
[102] A. Furnari, S. Battiato, G.M. Farinella, Personal-location-based temporal segmentation of egocentric videos for lifelog-ging applications, J. Visual Commun. Image Represent. 52 (2018) 1–12.
[103] A. Ortis, G.M. Farinella, V. DAmico, L. Addesso, G. Torrisi, S. Battiato, Organizing egocentric videos of daily living activities, Pattern Recognit. 72 (2017) 207–218.
[104] M.C. Murphy, A.C. Nau, C. Fisher, S.G. Kim, J.S. Schuman, K.C. Chan, Top-down influence on the visual cortex of the blind during sensory substitution, Neuroimage 125 (2016) 932–940.
[105] H. Ye, U. M. Malu, L. Oh, Findlater "current and future mobile and wearable device use by people with visual impairments", in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems ser. CHI'14, 2014, pp. 3123–3132.
[106] Z. Reynolds, B. Glenney, When sensory substitution devices strike back: an interactive training paradigm, Philos. Study: 432 (2012).
[107] S. Maidenbaum, G. Buchs, S. Abboud, O. Lavi-Rotbain, A. Amedi, Perception of graphical virtual environments by blind users via sensory substitution, PLoSONE 11 (2) (2016).
[108] G. A. Csapó, H. Wersényi, T. Nagy, Stockman "A survey of assistive technologies and applications for blind users on mobile platforms: a review and foundation for research", J. Multimodal User Interfaces 9 (2015) 275–286.