Human-centric Computing
and Information Sciences

　　　　　　　　　　　　　　　　　　　　　　　　Open Access

CrossMark

User-centered design of a depth data based obstacle detection and avoidance system for the visually impaired

Rabia Jafri[*] and Marwa Mahmoud Khan

*Correspondence:
rabia.ksu@gmail.com
Department of Information
Technology, College
of Computer and Information
Sciences, King Saud
University, Riyadh, Saudi
Arabia

Abstract

The development of a novel depth-data based real-time obstacle detection and avoidance application for visually impaired (VI) individuals to assist them in navigating independently in indoors environments is presented in this paper. The application utilizes a mainstream, computationally efficient mobile device as the development platform in order to create a solution which not only is aesthetically appealing, cost-effective, lightweight and portable but also provides real-time performance and freedom from network connectivity constraints. To alleviate usability problems, a user-centered design approach has been adopted wherein semi-structured interviews with VI individuals in the local context were conducted to understand their micro-navigation practices, challenges and needs. The invaluable insights gained from these interviews have not only informed the design of our system but would also benefit other researchers developing similar applications. The resulting system design along with a detailed description of its obstacle detection and unique multimodal feedback generation modules has been provided. We plan to iteratively develop and test the initial prototype of the system with the end users to resolve any usability issues and better adapt it to their needs.

Keywords: Visual impairment, Blindness, Navigation, Obstacle detection, Obstacle avoidance, Google Project Tango, Assistive technologies

Introduction

According to the World Health Organization (WHO), 285 million people are estimated to be visually impaired (VI) worldwide, of which 39 million are blind and 246 million have low vision [1]; these numbers are projected to steadily increase in the coming years [2, 3].

Navigating independently in an indoors environment is a major challenge for VI individuals—the inability to do so causes them frustration, undermines their confidence and autonomy and poses a serious risk to their physical safety [4–7]. In particular, lack of access to visual information about obstacles and temporary barriers in their paths significantly hinders their capacity to navigate not only in unfamiliar surroundings but even in environments which they traverse on a regular basis [8–10]. White canes, guide dogs and human companions are typically utilized to assist in detecting and avoiding obstacles. However, the white cane is highly conspicuous, has limited reach dependent on its

Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 2 of 30

length, cannot sense obstacles above the waist level, requires contact with obstacles in order to sense them (which may be impractical or undesirable—e.g., for detecting people or fragile objects) [4, 11, 12] and may not be usable by VI individuals suffering from additional disabilities which deplete their strength and impair their motor skills [13]; also, individuals with low vision tend not to use a cane for detecting nearby obstacles [14]. Guide dogs are costly, require protracted training and need proper care which VI individuals, especially aged ones, may find hard to provide [11]. A sighted companion may give extraneous, ambiguous or incorrect directions because of a lack of knowledge and understanding of how VI people navigate [15] and it may not be possible or desirable for such a human aide to be available at all times (indeed, according to a recent report, 26% of blind adults in the United States live alone [16, 17]).

Several assistive systems based on various sensor technologies, have, therefore, been proposed in recent years to support VI people in obstacle detection and avoidance during navigation. Among these, infrared-enabled depth sensor-based and computer vision-based systems on mobile devices have emerged as some of the most promising solutions [18–23]; however, both these kinds of solutions incur high computational costs. Given the current limitations of mobile platforms in terms of both the processing power and the battery life required to support real-time execution of computationally intensive code for extended periods, most of these solutions use the mobile device simply as a input/output frontend while all the actual processing of the data is done on a remote server equipped with powerful processors. This, in turn, leads to additional constraints and challenges: the user's device has to be connected to a remote system, the inevitable communication overhead may negatively affect the real-time performance and the absence of or failure of a network connection would render the system useless (network connectivity is an even more pertinent issue in developing countries where 90% of the VI population resides [1, 24, 25]). Moreover, the sensor modules utilized by many of these systems are not designed to be wearable or handheld; employing makeshift methods to affix these to various body locations has resulted in awkward bulky contraptions unlikely to be practically adopted by VI individuals who prefer assistive technologies to be provided on mainstream, or at least unobtrusive, devices to alleviate physical inconvenience [10] and facilitate social acceptability [26–29]. Last, but not least, instead of directly consulting with VI individuals to assess their navigation support needs prior to system design, the developers of most of these systems have simply relied on common perceptions about these needs held by sighted people [9] which may be erroneous and insufficient due to stark differences between the cognition and mental models of the sighted and VI [15, 30, 31]. Consequently, these solutions suffer from myriad usability issues as their interface designs and functionalities do not meet the unique needs and expectations of the target users [32]. Thus—to summarize—currently such systems fall short in terms of network independence, local computational efficiency necessary for real-time (and thus, practicable) feedback about approaching obstacles, physical convenience, social acceptability and ease of use.

The Project Tango Tablet Development Kit [33] is a 7" Android tablet, equipped with various sensors which allow it to track its motion, determine its position and orientation and learn the areas it is travelling through. Its integrated infrared based depth sensors enable it to measure the distance from the device to objects in the real world providing

Jafri and Khan *Hum. Cent. Comput. Inf. Sci. (2018) 8:14*

Page 3 of 30

depth data about the objects in the form of point clouds [34, 35] while its powerful NVIDIA Tegra K1 processor can perform expensive computations in real-time on the device itself without the need to connect to an external server or rapidly draining the battery. Also, being compact, lightweight, relatively discreet and affordable renders it aesthetically appealing, socially acceptable and accessible for VI users [24]. These considerations have motivated us to use this platform to design and develop an application to assist VI users in detecting and avoiding obstacles in their path during navigation in an indoors environment.

To ensure that our solution would meet the needs of real users and to minimize usability problems, we decided to adopt a user-centered design (UCD) approach, which is characterized by involving the users in every phase of the development process [36]. This paper, therefore, presents an exploratory study based on semi-structured interviews conducted with VI individuals in Riyadh, Saudi Arabia to uncover the challenges faced by them in detecting and avoiding obstacles and to identify the requirements for a mobile application that would assist them in micro-navigation indoors. The study results have been used to inform the design of a micro-navigation application for the VI based on the Tango platform. The initial design for the system, detailing the obstacle detection and feedback mechanisms, has also been presented in this paper. We plan to continue involving the target users in the iterative design, prototyping and evaluation process to ascertain that the final product is usable and meets their needs and expectations.

The research contributions of our work are threefold: the invaluable insights into the micro-navigation practices, challenges and preferences of VI individuals gained from interviews conducted with the actual targeted users, which have not only informed the design of our system but would also greatly benefit other researchers developing similar assistive micro-navigation applications for VI people; the design of a novel real-time assistive stand-alone application on a cutting-edge mainstream (thus, socially acceptable) high-performance power-efficient mobile device, which utilizes depth sensor data in an innovative manner to allow VI users to detect and avoid obstacles independently during navigation in possibly unfamiliar indoor environments; the creation of a unique multimodal user interface, developed in accordance with VI users' preferences, which employs audio and vibration cues to convey obstacle warnings and navigation instructions to the user.

The rest of the paper is organized as follows: The "Related work" section provides an overview of existing infrared-enabled depth sensor-based micro-navigation systems for the VI as well as previous user requirements collection studies for navigation systems for the VI. The "User requirements elicitation study" section describes the exploratory study conducted with VI users in the local context. The "System design" section presents the assistive obstacle detection and avoidance system designed in accordance with the preferences expressed by the users in the exploratory study. The "Conclusion and future work" section specifies some directions for future work and concludes the paper.

Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 4 of 30

Related work

To concretize our system concept and to formulate our exploratory study, we first researched existing work on infrared-enabled depth sensor-based micro-navigation systems for the VI as well as previous user requirements collection studies for navigation systems for the VI. An overview of the literature in these two areas is provided below.

Infrared-enabled depth sensor-based micro-navigation systems for the VI

In recent years, several technologies embedded with infrared-enabled depth sensors capable of extracting 3D information about the environment (e.g., Microsoft's Kinect [37], Occipital's Structure Sensor [38], and, most recently, Google's Project Tango Tablet Development Kit [33] and Tango-enabled smartphones [39, 40]) have become available at increasingly affordable prices; this has led to their being utilized for the development of various assistive micro-navigation systems for the VI. These solutions offer the advantages of being affordable, discreet and unobtrusive (since infrared light is invisible to the naked eye), and being able to operate in low illumination environments [13, 41]. However, such systems can detect objects only within a certain range and their performance may be negatively impacted by interference from other infrared sources, such as sunlight or fluorescent light [13].

Recent development work on obstacle detection has specially focused on Kinect, either utilizing the data from its depth sensor alone [42–45] or from both its RGB and depth sensors [12, 46]: for example, Khan et al. [47] divide the depth image of the scene obtained from a waist-mounted Kinect sensor into 5×3 regions, calculate a depth metric for each region and instruct the user to go in the direction with the smallest probability of an obstacle. Filipe et al. [43] extract six vertical line profiles at pre-defined locations from depth images acquired from a chest-mounted Kinect sensor, classify them using a feedforward neural network with back propagation and inform the user about the location of any obstacles found in terms of right, left and center. Huang et al. [45] determine the ground height threshold from the depth image obtained from a Kinect sensor affixed on the user's helmet, chest or waist, utilize it to detect descending stairs, then remove the ground plane, apply a region growing approach to label different objects and analyze the labelled objects to determine if any of them are ascending stairs; the user is informed via audio about how far he is from any obstacles in his path; if stairs are detected, the direction and distance to the stairs is conveyed. Liu et al. [48] apply a multiscale voxel plane segmentation method on the 3D point cloud data obtained from a chest-mounted Kinect sensor to extract planar structures, then remove the ground plane and use an area growing algorithm to segment all the non-ground regions into independent clouds which are regarded as generalized obstacles. The area in front of the user is then divided into three cuboids (left, center and right) and a multi-level voice feedback strategy is employed to generate warnings and avoidance instructions for any detected obstacles. Brock et al. [44] down-sample the depth data from a Kinect sensor hung from the user's neck and split it into isolated structures, representing obstacles, at different depth levels using the marching cubes algorithm. The 3D location of the obstacle is then sonified so that the horizontal position, vertical position and distance from the user are encoded by the panning position, pitch and volume of the sound, respectively. Bernabei et al. [49] detect the floor based on the depth data acquired from a waist-mounted Kinect sensor

Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 5 of 30

and then simultaneously analyze the volume in front of the user to determine if there is sufficient room for him to move without colliding with an obstacle and the output from a wearable accelerometer to establish if the user is walking and if so, at what speed. Speech-based instructions are given for obstacle avoidance and sonification is used to convey the obstacle's location and distance from the user. The data processing for the above systems is done on portable devices such as laptop computers [44, 45, 47], mini-PC processor [48] and smartphones [49] while the audio feedback is provided via Bluetooth headphones [44, 47, 48], laptop speakers [45] or smartphone speakers [49].

Some Kinect-based solutions have opted for tactile—instead of audio—feedback: Zöllner et al. [42] mount the Kinect sensor on the user's head and put vibe boards on the left, right and center of the user's waist to indicate the direction in which an obstacle was detected; if the pixel area of the depth window moved over the depth histogram of the current frame exceeds a certain threshold area, the average depth value of the window is mapped to the pulse of the appropriate vibe board. Mann et al. [50] affix an array of six vibrating actuators inside a helmet, mount a Kinect sensor on top, divide its depth sensing region into six zones, calculate a distance map of the depth image and make the vibration of each actuator inversely proportional to the distance of its corresponding zone to create the sensation of objects in the visual field pressing against the forehead before collision occurs—the sensation increases in strength as collision becomes more imminent.

All the above systems detect obstacles in general. However, a few other Kinect-based solutions focus on detecting specific obstacles such as staircases and traffic [51–53]. Data from Kinect has also been combined with other modalities such as ultrasonic and sonar for obstacle detection [54, 55].

Since the Kinect sensor module is not designed to be wearable or handheld and the data captured by it needs to be sent to an external server for processing, solutions based on it suffer from the lack of aesthetic appeal, network connectivity issues and real-time performance challenges mentioned previously in the "Introduction" section.

The Project Tango tablet appears to have a distinct advantage over Kinect in terms of aesthetic appeal, on-board handling of significant compute workloads, and being equipped with several additional embedded sensors and in-built functionalities, which can be utilized for extending and improving the obstacle detection application in the future, thus, making it an attractive platform for developing assistive obstacle detection and avoidance solutions for the VI. It should be noted that a few preliminary applications for the Tango tablet have already been proposed for this purpose: The system presented by Anderson [56] collects depth information about the environment, saves it in a chunk-based voxel representation, and generates 3D audio for sonification relayed to the VI user via headphones to alert him to the presence of obstacles. Wang et al. [57] cluster depth readings of the immediate physical space around the users into different sectors and analyze the relative and absolute depth of different sectors to establish thresholds to differentiate among obstacles, walls and corners, and ascending and descending staircases; users are then given navigation directions and information about objects using Android's text-to-speech (TTS) feature. Lock et al.'s [58] navigation system includes an obstacle detection component which utilizes the depth sensor data to detect any approaching obstacle (details of the detection process have not been provided) and

Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 6 of 30

communicates its distance to the user by adjusting the vibration intensity; a co-adaptation module will eventually be added to the system to adjust the feedback parameters according to the user's skill levels. Munoz et al. [41] focus on detecting staircases: their system preprocesses RGB images and depth data from a chest-mounted Tango tablet, applies additional geometric constraints to identify staircase candidates, classifies them using an SVM-based classifier as upstairs, downstairs or negatives and provides feedback such as directions and distance to the staircase via TTS (voice commands are received from the user via a speech-to-text (STT) module). Jafri et al. [59] utilize the Tango Unity SDK to create a 3D-reconstruction of the surrounding environment based on the tablet's depth data and associate a Unity collider box with the user; audio alerts are relayed to the user via bone conduction earphones whenever the collider box comes in contact with the reconstructed mesh (steps are taken to ensure that the ground does not trigger obstacle warnings). All these applications need further development and are yet to be tested with the target users.

Li et al.'s [60] indoor assistive navigation system has an obstacle detection component that de-noises the point cloud data, de-skews it to align it with the horizontal floor plane, and projects it in both the horizontal and vertical directions. The vertical projection detects in-front and head-height obstacles locally while the horizontal projection is used to update obstacle information in a global 2D grid map maintained by the application. The Android TTS module is used to convey the obstacle detection results and navigation directions while a beeping alert sound signals the distance to the obstacle. The system is reported to have been tested with blindfolded and blind subjects. However, details about the number of subjects and the data collection procedure and an in-depth analysis of the users' performance have not been provided.

None of the Tango tablet-based systems mentioned above has been developed by following a user-centered design approach consistently from the early product development phases, i.e., VI users have not been involved in the requirements collection, design and initial implementation of the system. Also, unlike our system, none of them appear to combine audio and vibration to provide feedback about the obstacles and directions to avoid them.

User requirements collection studies for navigation systems for the VI

Most of the systems mentioned above focus on technical aspects and were designed without prior consultation with the end users themselves to understand their perspective which caused them to suffer from several usability and interaction issues [32]. Since it is becoming increasingly apparent that user involvement in assistive technology design and development is imperative for ensuring its usability and eventual acceptance by the target users, recently, a handful of studies have been conducted in which VI individuals were directly approached to elicit their navigation needs and identify any major challenges that may be alleviated by electronic technology-based solutions as a prelude to developing assistive navigation solutions. For example, Sánchez and Elías [61] interviewed blind students and mobility and orientation experts in Santiago, Chile about the challenges faced by blind children when navigating in indoors environments to derive some guidelines for developing navigation software for blind individuals in general and blind children in particular. Quinones et al. [10] interviewed blind and low vision

Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 7 of 30

individuals both in the USA and South Korea about their way finding techniques and challenges and use of assistive tools to gather their information and technology needs in order to inform the design of a system for providing landmark information to people with visual disabilities. Miao et al. [62] interviewed blind users, a mobility coach and a sighted person (who performed indoors navigation tasks while blindfolded) to elicit user requirements for the functionality, usability and route descriptions of an indoor navigation application. Burigat and Chittaro [9] interviewed disabled students (mostly with motor or visual impairments) at a university in Italy about their navigation challenges, information needs and mobile technology use within university buildings in order to derive design implications for mobile navigation and information services for disabled students. Paredes et al. [8] interviewed low vision and blind individuals about their assistive technology use for mobility and the obstacles and barriers frequently encountered by them to gather requirements for an assistive navigation system. Williams et al. [15] conducted focus groups with VI individuals, observed VI participants during independent navigation tasks and observed blind participants navigate real environments with sighted companions; this enabled them to identify navigation challenges as well as common misperceptions held by sighted people about the guidance information needed by the VI and to infer several design considerations for future navigation technology. We came across only one study [63] in the local context, i.e., in Saudi Arabia, in which an online survey was conducted with VI users from the online Twitter community followed by semi-structured interviews with participants from a local organization for the VI to elicit several requirements and design implications for the development of indoor mobile navigation systems for the VI.

Some other studies with VI users, though exploring a broader range of issues, have also elicited some requirements related to their navigation needs and interface preferences. For example, Hamilton-Fletcher et al. [64] briefly trained low vision and blind users to use three sensory substitution devices and then interviewed them to gain some insight into the limitations and benefits of such devices for improving their everyday lives; identifying a variety of obstacles was highlighted as one of the most desired practical functions.

All of the above studies concentrated on navigation in general rather than obstacle detection and avoidance in particular. Also, several of them limited their focus to very specific contexts (e.g., a particular building type [9]). This highlights the need to conduct further studies with VI users which focus on thoroughly investigating issues related to obstacle detection and avoidance in particular in a variety of environments. Furthermore, the dearth of such studies in Saudi Arabia (we were able to find only one such study [63]), especially given the lack of assistive tools and adapted environments supporting the VI in this region [63], delineates the necessity for more research with VI users in the local context pertaining to this task.

The usability issues plaguing existing systems, the valuable insights into the navigation needs and interface design preferences of VI individuals provided by the above studies and the lack of such studies focusing on micro-navigation in the local context convinced us of the necessity of formulating our own exploratory study to gather feedback from local users about their micro-navigation practices and challenges to inform the design of our proposed system. We will refer back to the insights of the previous studies when

reporting the results of our own exploratory study with VI users accentuating the similarities as well as any differences among them. Furthermore, the appropriateness of the Tango tablet for assistive technology deployment (as explained in the "Infrared-enabled depth sensor-based micro-navigation systems for the VI" section) and the limited utilization of this platform so far for the obstacle detection and avoidance task has motivated us to select this platform for our development work.

We proceed to describe the exploratory study in the next section. The design of the proposed micro-navigation system based on the Tango tablet will be presented in the subsequent section.

User requirements elicitation study

A study was conducted with VI individuals from the local community in order to gain some insight into their electronic device use and current micro-navigation practices and challenges as well as to obtain their feedback about the usefulness, desired functionalities and interaction design of the proposed system. A semi-structured interview method was selected to conduct the study since it allows direct contact and interaction with the users, is broadly replicable but yields rich and complex information related to individual users' perceptions and opinions and can be accomplished within a reasonable amount of time with limited resources [8, 36].

The interview was structured around the following six main topics: (1) basic demographic information and visual impairment history of the participants, (2) their electronic device use, (3) their electronic and non-electronic navigational aid use, (4) their obstacle detection challenges, (5) their preferences for the interface design of a mobile micro-navigation application and (6) their impressions and suggestions for the proposed system.

The following sub-sections explain the methodology for conducting the study, report the results of the interviews, discuss the findings and recommendations generated based on the results and explicate how these were used to inform the design of the proposed system.

Methodology

Participant recruitment and interview sessions

Semi-structured interviews were scheduled with ten VI participants who were recruited via the Disability Center located in the College of Education, Girls' Campus, King Saud University (KSU), Riyadh, Saudi Arabia.

The interviews were conducted in-person and on-site at the Disability Center to minimize any inconvenience to the participants in terms of time and transportation and to provide a familiar environment in which they would feel comfortable and at ease. For each participant, the goals of the study were explained and formal consent for participating in the study was obtained by having her sign a consent form.

Pre-planned questions about the six topics mentioned previously were asked in an informal manner allowing the participants to express their opinions and volunteer further information if they so desired. Any dubious points or unanticipated responses were further examined by additional follow-up questions. The interview session lasted for about an hour. The interviews were conducted over a period of 2 weeks.

Table 1 Demographic information of participants

ID	Age range (years)	Year in Bachelors' program	Degree of visual impairment	Congenitally VI?	Financial support from government?
LV1	18–25	Senior	LV	Yes	Yes
LV2	18–25	Junior	LV	No	Yes
LV3	18–25	Sophomore	LV	Yes	Yes
LV4	18–25	Junior	LV	Yes	Yes
LV5	18–25	Senior	LV	Yes	Yes
LV6	26–35	Sophomore	LV	Yes	Yes
BL1	18–25	Sophomore	BL	No	Yes
BL2	18–25	Sophomore	BL	No	Yes
BL3	26–35	Junior	BL	Yes	Yes
BL4	18–25	Sophomore	BL	Yes	No

LV extreme low vision with some color and light perception, *BL* blind

Data collection, processing and analysis

Though it would have been desirable to video record the sessions facilitating facial expression analysis later if needed, deference to cultural norms made this infeasible. All participants also declined having the sessions audio-recorded.

The participants' responses were, thus, recorded manually by taking notes. For closed questions with a limited range of response options, the answers were aggregated across all participants and converted into percentages using the SurveyGizmo tool [65]. For open-ended questions, the answers were examined manually, positive and negative comments were separated, and similarities as well as disparities among the opinions and suggestions about the system were noted. The aggregated results were examined to discover any emerging patterns and/or correlations among various factors. The results were used to develop findings and recommendations to inform the design of the proposed system.

Results

The responses of the participants for the six main interview topics are compiled and summarized below.

Demographic information

All ten participants were female, undergraduate students at KSU between the ages of 18–35 years. All of them did not have any disabilities other than visual impairment. Table 1 shows their ages, undergraduate study year, degree of visual impairment, whether they were congenitally VI and if they receive any government support for their disability. Four were blind (BL) while the remaining six had extreme low vision (LV) with some light and color perception. The LV participants reported varying levels of residual visual perception: LV2 had no vision in one eye while the other eye was normal; LV5 could not decipher text such as room numbers or building names; LV1, LV3, LV4 and LV6 could not discern details such as facial expressions; LV3 mentioned not being able to see in environments where the intensity of light is very high or very low.

Seven of the participants were VI from birth while others became VI later in life. The causes of visual impairment included genetic factors, various diseases and medical errors.

All participants except BL4 received financial support from the government mainly through the Comprehensive Rehabilitation Center under the Ministry of Social Affairs [66] in Saudi Arabia in the form of cash or assistive devices. Two of the participants mentioned that another well-established regional hospital, King Khaled University Hospital [67] could also provide special devices upon request.

Electronic device use

The purpose of this part of the interview was to determine the participants' familiarity and level of expertise with electronic devices, the type and cost of devices currently being used by them, the most frequently performed tasks on these devices and the factors that they deem the most important for selecting and adopting a device.

All participants used smartphones daily while seven of them had also either tried out a tablet device or owned one. Only BL3 mentioned that she owned additional devices, namely, a laptop computer and an electronic note-taker (a portable device that uses braille or standard keyboard to help take notes in a class [68]). All smartphones and tablets were Apple iOS-based [69]—only LV2's smartphone was Android-based [70]. The cost of the devices ranged from 650 to 800 USD for all participants excepting three for whom this exceeded 800 USD.

The most frequent tasks performed by all participants using these devices were making phone calls, browsing the web and social networking. Three participants (2 BL, 1 LV) also used their smartphones for navigation in outdoor environments using the embedded global positioning system (GPS) receiver whereas one LV participant used the GPS only for locating places without navigation.

The most frequently mentioned important factors which influence the preference of one device over another were: availability of an embedded screen reader service (90%), availability of accessible utilities to interact with the device (60%), ease of learning to use the device (60%), and ease of carrying the device (50%).

Navigation aid usage

This part of the interview explored which traditional and electronic navigation aids, if any, were currently being used by the participants in indoor environments and what problems were experienced using these aids. It also probed into how often they ventured outside their homes unaccompanied, how willing they were to adopt a new electronic travel aid and what their projections were about the impact of such an aid on their independent navigation practices.

Regarding traditional aids, two LV participants did not use any such aids for indoor navigation. Of the remaining eight, the majority used a human guide while only two used a white cane (Fig. 1). The main shortcomings mentioned for the white cane were its inability to warn about obstacles which are distant or above the waist-level and its requiring contact with an obstacle in order to sense it. LV6 also mentioned that it limits her walking speed and produces irritating sounds when walking on a rough surface thus drawing unwanted attention from passersby. The main limitations mentioned for a human guide were dependency on others and the guide's failure to warn about approaching obstacles; a couple of participants (BL3 and LV6) also mentioned that this limits their walking speed and draws attention to their being VI.
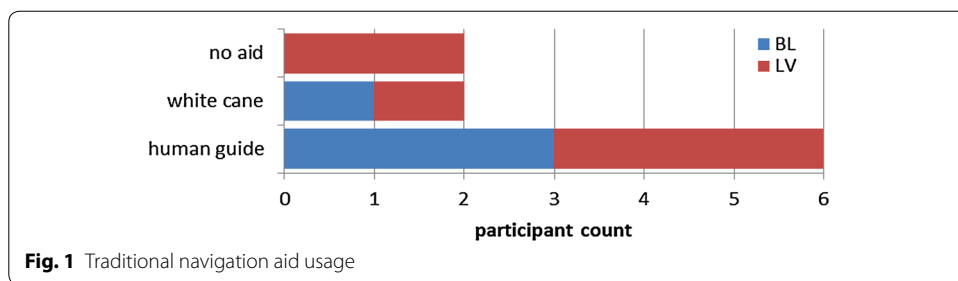
Fig. 1 Traditional navigation aid usage

Table 2 Aggregated count of participants' ratings of the difficulty of detecting various obstacles

Obstacle	Difficulty of detection				
	Very easy	**Easy**	**Neutral**	**Difficult**	**Very difficult**
Ascending stairs	5	1	0	2	2
Descending stairs	2	1	4	1	2
Moving obstacles	5	0	1	0	4
Very small obstacles	1	0	1	0	8
Walls	7	0	2	0	1
Chairs/tables	3	1	6	0	0
People	5	3	0	1	1

When asked about their degree of satisfaction with their current navigation aid, for both the white cane and the human guide, half of the participants indicated that they were satisfied while half of them remained neutral.

None of the participants were currently using any electronic aids for indoor navigation. However, two participants, LV1 and BL3, mentioned that they tested two smartphone navigation applications—for obstacle detection and navigation, respectively—developed by students at the Computer Science department at KSU. LV1 also participated in the one-time testing of some navigation and localization aids available in the College of Education at KSU. BL3 mentioned that she was aware of some navigation aids available in the Apple store [71] but that they were too expensive for her to purchase.

All participants went out alone regularly in familiar environments but always asked a human guide to accompany them in new/unfamiliar environments.

When asked if they would be likely to go out alone more often if they had an effective application/electronic device which would assist them in obstacle detection and navigation, nine of the participants replied in the affirmative while only one (LV5) expressed some reservations indicating that she would consider this if and only if she was certain about the accuracy of the electronic aid's functions.

Challenging obstacles

This part of the interview aimed to discover which obstacles the participants found the most difficult to detect and avoid. Table 2 shows how the participants rated various kinds of obstacles. Very small obstacles, moving obstacles and descending stairs were rated as the most challenging to detect overall while walls and people were rated as the easiest. Some differences were observed based on whether the participant used a navigation aid

Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 12 of 30

or not and also the type of the aid. Participants using the white cane rated moving obstacles and people as the most challenging and descending stairs and walls as the easiest to detect. Those using a human guide ranked very small obstacles and descending stairs while those who did not use a navigation aid rated very small obstacles and sill on the floor as the most difficult to detect.

Some additional challenging obstacles mentioned by the participants were poles, recycling bins, sill on the floor, and children.

User interface preferences for a mobile micro-navigation application

This part of the interview aimed to investigate the input and output modalities the participants would prefer for interacting with a mobile micro-navigation application and sought their opinions about the format, extent and frequency of the feedback provided by such a system.

Input preferences   When presented with three options to provide input to the system—speech, touch or a combination of both (with the possibility to select either one)—, 80% of the participants chose the combined option. The reasons given were that such a model would balance the tradeoffs for both modalities—touch would be preferable for commands that need to be executed quickly while speech would be preferable to give more precise commands that do not require instant execution; touch is more discreet and would be preferable in public places while speech would be more convenient in private spaces.

Participants were further probed about their past experience and satisfaction with both these modalities as well as format preferences for the speech input. Their responses are summarized below:

A. Speech input All participants had used the embedded input speech services in their devices. The most popular services were Apple's Siri [72], iPhone keyboard voice dictation [73] and WhatsApp [74] speech-to-text services. The level of visual impairment seems to affect the frequency of use of these services: the majority of the LV participants rarely used them while half the blind participants used these often. Most participants utilized the speech input for text messaging and interacting with recreational apps while a couple of them employed it for web searching and one blind participant used it for initiating phone calls.

As shown in Fig. 2, only 40% of the participants were satisfied with the speech input services, 40% were neutral and 20% were dissatisfied. The dissatisfaction arose from the software not recognizing the Arabic language input correctly, the requirement of enunciating the commands slowly and clearly, the interference of environmental sounds with the speech and the incompatibility of some apps with the speech services.

When asked if they would prefer to input configuration commands to the assistive mobile application using short sentences (e.g., "Change to vibration mode") or single words ("Vibration"), 60% chose the former claiming this format would be less ambiguous while 40% opted for the latter opining that it would enable faster delivery and interpretation of commands.
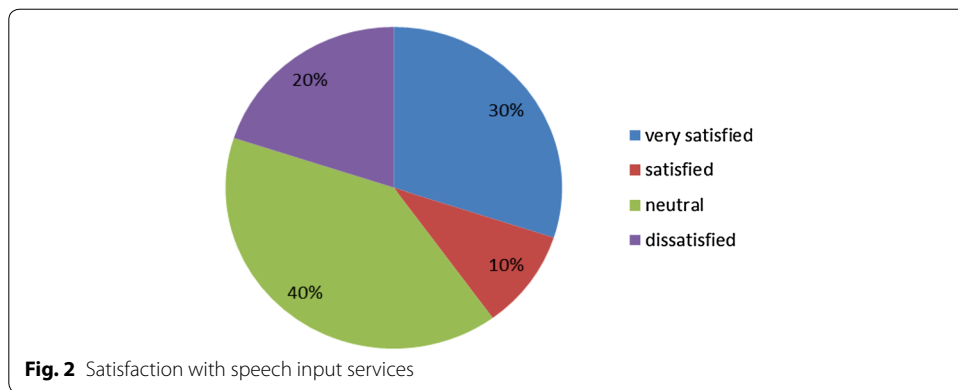
Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 13 of 30

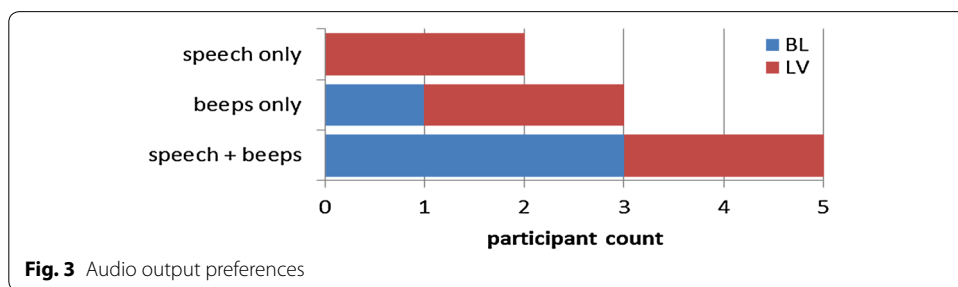**Fig. 2** Satisfaction with speech input services

Fig. 3 Audio output preferences

B. Touch input All participants were familiar with touch input with 90% of them using it frequently. Tasks included writing, searching, making calls, browsing the internet and opening apps.

All participants were satisfied in general with the touch input though some complained that it occasionally causes their devices to slow down or stop responding. BL1 and LV4 also mentioned that it is difficult to distinguish among buttons while BL3 commented that a touch model requires a training phase especially for first-time VI users.

Output preferences   Similar to the input, most of the participants (70%) preferred a combined audio and touch model—as opposed to audio-only or touch-only—for the output.

Participants were further probed about their past experience, satisfaction and format preferences for both these modalities as well as their combination. LV users were also asked if they would like to receive some coarse visual feedback. The responses are summarized below:

A. Audio output Since audio output can be provided in two forms, speech or beeps/tones, the participants were asked about which format they would prefer. As shown in Fig. 3, the majority opted for a combination of speech and beeps. However, the majority of the blind (3 out of 4) chose the combination while the LV ones were equally split among the three options (i.e., speech-only, beeps-only, speech and beeps). Further responses regarding these two audio forms are reported below.

Speech-only output All the participants had experience with text-to-speech technology through iPhone VoiceOver service [75]. They were questioned about their preferences for the speech type (automated or human), language (Arabic, English or both)
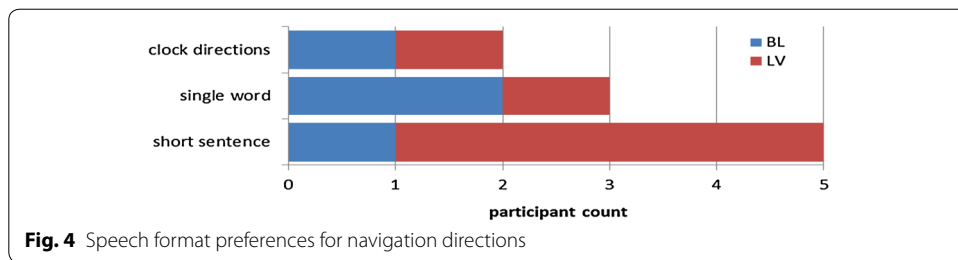
Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 14 of 30

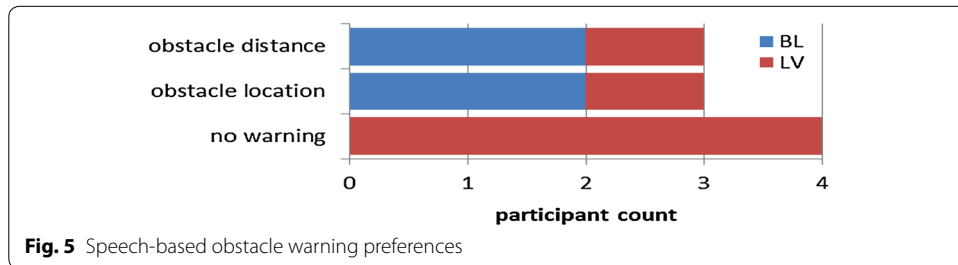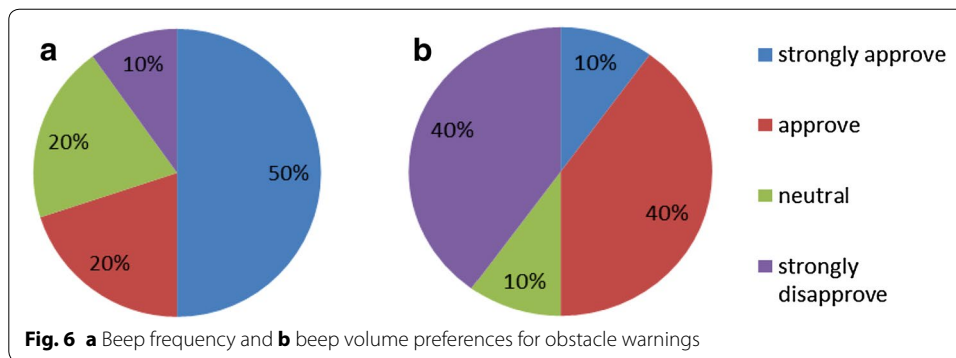**Fig. 4** Speech format preferences for navigation directions

Fig. 5 Speech-based obstacle warning preferences

and navigation directions' format (short sentence (e.g., "Turn right"), single word (e.g., "Right"), or clock directions (e.g., "1 o'clock")). The participants were split equally in their speech type preferences with half choosing automated speech and the other half favoring a human voice; however, this preference appeared to be affected by the level of visual impairment with most of the LV participants choosing automated and most of blind participants choosing human speech. Half the participants preferred Arabic-only while the other half favored both Arabic and English. Among the speech formats, as shown in Fig. 4, short sentences were preferred by most followed by single words and finally, clock directions (the participants cited novelty and more precise description of directions than only the cardinal ones as reasons for choosing this format).

When asked if they would like to receive speech-based obstacle warnings in addition to navigation directions, the majority responded positively (Fig. 5). Among these, half preferred an obstacle warning consisting of a short sentence indicating the location of the obstacle relative to the user (e.g., "obstacle in front") while the rest wanted the distance of the obstacle to be reported in addition to this warning (e.g., "Obstacle in front at 1 m/4 steps"). Furthermore, all participants indicated that if the distance is reported by the system, they preferred it to be specified in steps, as opposed to meters (this is in accordance with the findings from some other studies [32, 63]). They all also agreed that they would like the system to explicate the type of the obstacle (e.g., chair, pole, bin, etc.).

Note that all blind participants wanted to receive speech-based obstacle warnings in addition to navigation directions while most of the LV participants (66%) believed that navigation directions were sufficient.

Beeps-only output The following feedback scheme based on beeps was described to the participants to clarify how such output may be utilized for obstacle warnings and navigation: beeps in the right ear, left ear and both ears indicate an obstacle on the left, right and center, respectively, of the area in front of the user; no beep in both ears indicates that the path directly in front of the user is clear (i.e., no obstacles).

Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 15 of 30

Fig. 6 **a** Beep frequency and **b** beep volume preferences for obstacle warnings

Two techniques were suggested to warn a user that he/she was getting closer to a detected obstacle: (1) increasing the beep frequency, (2) increasing the beep volume. As shown in Fig. 6, 70% of the participants approved of the frequency method and only 10% strongly disapproved of it. On the other hand, though 50% approved of the volume method, 40% strongly disapproved of it asserting that the increased volume would draw attention in public places causing embarrassment and would also be irritating for the ears since the user would be wearing headphones.

In the suggested feedback scheme, the beep constitutes an obstacle warning. However, the participants were asked if they thought it should indicate navigation directions to avoid obstacles instead (e.g., beep in the right ear should mean "Bear towards the right"). 60% (3 BL, 3 LV) replied in the affirmative while the remaining 40% (1 BL, 3LV) did not (note the differences between the blind and LV participants: most of the blind favored this option while the LV were split equally on this issue).

Two participants also pointed out such a feedback scheme is not acceptable in general since it requires the user to wear headphones which would block out environmental sounds which constitute essential contextual/situational awareness cues for the VI.

B. Touch (vibration-based) output 90% of the participants indicated that they would like to receive a vibration-based warning for detected obstacles. They all concurred that a high vibration frequency should indicate nearby obstacles while a low frequency should denote far-off obstacles.

C. Hybrid (audio + touch) output Participants were then asked to rate different combinations of audio and vibration-based outputs for conveying navigation directions and the distance to the obstacle. The highest ranked combination, approved by 90% of the participants, utilizes speech for navigation directions and vibrations for the distance to the obstacles while the lowest rated one uses speech and beep volume, respectively, for these tasks.

D. Visual output (for LV users) Since mobile devices are usually equipped with a visual display screen, participants with some residual vision were asked if they would like to receive coarse visual feedback on the screen depicting the locations of detected obstacles and their distances from the user. Distances could be represented either by varying the light intensities (the closer the obstacle, the darker it will appear) or by different colors (e.g., from closest to farthest: red, orange, yellow, green, blue). Three of the LV participants selected the light intensities option and two chose the colors' one. However, the

Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 16 of 30

Fig. 7 Frequency of receiving feedback preferences
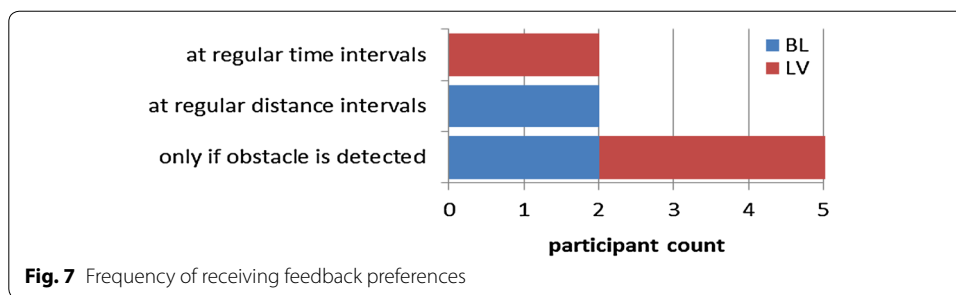
remaining LV participant rejected the visual feedback option asserting that even though she could differentiate among various colors, she did not want any visual feedback since it is very difficult to distinguish among colors or intensities on a mobile device screen. One other participant, though she had chosen the intensities option, also stated that she would rather have vibration-based—instead of visual—feedback.

E. Frequency of receiving feedback Since one major concern when designing an assistive device for the VI is not to overwhelm the user by providing too much information, the participants were asked if they wanted to receive feedback from the system only if an obstacle was detected or at regular intervals regardless of whether an obstacle is detected, where the intervals are determined by either time (every few seconds or minutes) or distance (every few meters or steps). The majority indicated that they would want feedback only if an obstacle was detected; of the remaining 4 (2 BL, 2LV), the two blind participants chose the distance intervals while the two LV ones chose the time intervals for receiving feedback (Fig. 7).
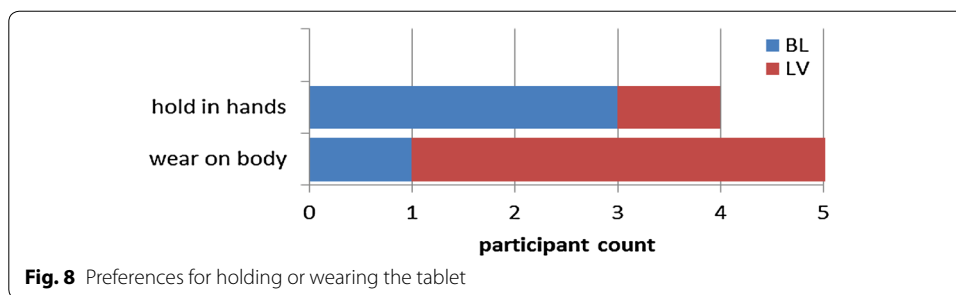
When asked at what distance from an obstacle should the feedback about it start being provided, 40% of the participants chose 1 m while the remaining were split equally among the following three options: < 1 m, 2 m and 3 m. Note that most of the blind participants (3 out of 4) selected 1 m.

Impressions and suggestions for the proposed system

This part of the interview aimed to gather the participants' initial impressions of the Tango tablet, their preferences for carrying it (held in their hands or worn on their bodies), and their opinions about the extent to which a device's dependency on an internet connection affects their acceptance and adoption of it.

A brief description of the Tango tablet and the proposed obstacle detection and avoidance application was provided to the participants and they were allowed to handle the device to get an idea of its weight and physical dimensions.

Four participants preferred to hold it in their hands while the remaining six favored wearing it on their body (Fig. 8). There is a clear divide between the blind and LV participants here with most of the blind participants gravitating towards the hand-held option and most of the LV ones towards the wearable one. One participant mentioned that she would actually like to have both options available so she can select the most appropriate one according to her situation and environmental context. When those choosing the wearable option were questioned further about which body location they would prefer, four opted for the waist, one indicated the chest while one chose the head. Some

Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 17 of 30

Fig. 8 Preferences for holding or wearing the tablet

justifications provided for these choices were that on the waist, the tablet's camera will be more accurately aligned with the direction in which the user would be facing; furthermore, some participants expressed concerns that placing the tablet on the chest area may have harmful effects on their health.

Seven of the participants deemed it very important for the application to run entirely on the device without having to connect to the internet, one stayed neutral while the remaining two remarked that connecting to services on the internet would be acceptable only if it would significantly improve the application's results.

Though all the participants were satisfied with the tablet's weight purporting that it was light enough to carry around easily, almost all of them preferred its size to be smaller. Since the Tango platform is being made available on other mobile devices with smaller form factors (e.g., the recently introduced Lenovo Phab 2 Pro [39] and Asus Zenfore AR [40] smartphones), moving our application to these smaller devices in the future may address this concern.

Discussion of results and design recommendations

Observations about electronic device and navigation aid use

Some general observations based on the interview results for electronic devices and navigation aid use, along with their implications for our system, are highlighted below:

The results for the electronic device use indicate that all participants use smartphones and most use tablets frequently revealing their familiarity with mobiles devices, even though the use is mostly limited to communication tasks (the prevalence of mobile phone usage by the VI has been reported by several other studies as well, both locally and internationally [8, 9, 63]). However, the majority does not own or use dedicated assistive electronic devices. This finding is consistent with previous studies which reported that VI individuals prefer not to carry additional devices (because this is physically inconvenient [10]) or use special devices (since such tools are usually costly, conspicuous and draw unwanted attention [28, 61]). This reinforces the need to develop assistive applications on mainstream devices—as is the case with our system which will be launched on a consumer-level tablet device; since the Tango platform is rapidly being integrated into other typical mobile devices (e.g., smartphones), the system could eventually be ported to these if the user so prefers. The results also show that the participants already own electronic devices in a price range of 650 USD-800 USD and most of them also receive financial support from the government for purchasing assistive tools implying that assistive solutions developed for mobile devices in this range, such as the Tango tablet, would

not be beyond their financial reach. Most of the devices currently being used by the participants are iOS-based and it would be recommendable to use the same platform for further development in order to save the users the time and effort to familiarize themselves with a new operating system. However, in our case, our system is constrained to Android devices since this is the platform on which Tango is currently available. It should be noted that Android is reported to have the largest mobile operating system market share in Saudi Arabia [76], so the preference for iOS appears to be limited to the local VI community. Nevertheless, given the popularity of Android locally as well as worldwide [77, 78], we are optimistic that VI users switching to this operating system will not experience major platform-dependent usability issues. From a development and dissemination point of view, too, Android appears to be a more attractive choice since its open source nature allows free distribution of software and facilitates updates. The interview results show that the three most important factors which influence the VI user's preferring one device over another are the availability of an embedded screen reader service, availability of accessible utilities to interact with the device, and ease of learning to use the device. In the case of the Tango tablet, there are several Android screen readers available [79, 80]; usability testing conducted with the target users will eventually help to determine the learnability and ease of interaction with the device.

There appears to be a general reluctance to use navigation aids: 80% of the participants do not even use a white cane instead opting either not to use an aid at all or to rely on the relatively discreet assistance of a friend or family member. This is consistent with the results obtained in some previous studies in the local context [27, 63]. The major shortcomings explicated for both the white cane (i.e., failure to detect distant and above waist-level obstacles, requiring contact with obstacles, conspicuousness) and the human guide (i.e., dependency on others, failure to warn about approaching obstacles, conspicuousness) confirm and extend upon the limitations for these aids reported by previous studies [15, 63] (some additional limitations mentioned by other studies include inaccurate distance estimations to obstacles and ill-timed, redundant, ambiguous and erroneous directions given by sighted companions [9, 15, 63], high fees and inappropriate treatment meted out by paid human guides [63], and the long training time required for learning to effectively use the white cane [11]); all these shortcomings can be alleviated by using an electronic aid on a mainstream device. However, though a small number of participants use the GPS service on their smartphones for localization, none of them currently use an electronic aid for navigation in general or obstacle detection in particular (this finding is also consistent with the low assistive technology use for navigation in the local context reported in [63]). Most of them seemed unaware of the availability of such aids and were surprised when asked about their use of them. However, almost all of them seemed eager to try out such an aid and appeared excited about the prospect of independent navigation in unfamiliar surroundings that this offered, thus, reaffirming the need for developing such aids.

Requirements and recommendations for a micro-navigation application for the VI

Some requirements and design recommendations for an assistive application for micro-navigation for the VI emerged from the responses about challenging obstacles, proposed

Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 19 of 30

system suggestions and user interface preferences. A few important ones have been listed below and the ones especially applicable to our system have been highlighted.

Requirements   Since human guides frequently fail to provide warnings about approaching obstacles and white canes cannot detect overhanging and distant obstacles, the system should be able to detect such obstacles. Our system will be designed to take these cases into account.

Very small obstacles, moving obstacles and descending stairs were rated as the most challenging obstacles by the majority of the participants. Poles, recycling bins, and children were additionally mentioned as difficult to detect. It should be noted that many of these obstacles have also been mentioned by VI users in other studies [8, 9, 61, 63].

However, when participants were divided into groups based on whether they used a white cane, human guide or no aid at all, then within each group, the following obstacles were also rated the most difficult to detect: people, sill on the floor, stairs going up. Hence, the results imply that the traditional navigation aid being used by a participant alters her perception of which obstacles are more challenging to detect. Since potential users of the proposed system would probably use the system to complement their current aid rather than to replace it, this has the following implications: The system should be able to deal with all obstacles specified as challenging by the different categories of users (white cane users, human guide users, users not using traditional navigation aids); The user should be given the option to receive warnings only about those obstacles which she finds difficult to detect with her traditional navigation aid (if using one).

Our system's initial prototype will attempt to detect all obstacles mentioned above other than descending stairs (since our current obstacle detection method searches for objects above the floor level). The custom warnings option would be included in future iterations since this requires recognizing and classifying obstacles after detecting them which is intended as future work.

Requiring the device to connect to the internet was strongly disfavored and was deemed tolerable only if this provided some substantial improvement in the application's performance (this supports the findings of the previous local study [63]). Our application will, therefore, run entirely on the device itself and not require internet connectivity.

Blind and LV participants favored holding the device in their hands and wearing it on their body, respectively. The initial prototype of the system will support the handheld option only. However, a 3D-printed holder for the tablet (as suggested in [59]) to be mounted on the user's waist (the body location preferred by most participants) would be designed in the future.

Design recommendations   *Input interface* The results indicate that all participants were familiar with both speech and touch but were more dissatisfied with speech mainly due to the unreliability of the speech recognition software. Touch also has the added advantage of being discreet. This implies that speech input should be used sparingly with most of the input being provided via touch. However, since blind participants showed a marked preference for using speech despite its unreliability, speech input alternatives should still be provided.

Since the initial prototype of our system will simply provide feedback about detected obstacles, it will not require any input from the user. However, when customization options which would require users to input their preferences are added in future iterations, the above recommendations would be taken into consideration.

Output interface The majority of the participants preferred speech for navigation directions and vibrations for the distance to the obstacles. Our system will, thus, adopt these modalities for these feedbacks.

Some differences based on the level of visual impairment were apparent for some of the output preferences such as speech type, speech format, obstacle warnings, etc. However, since there was general agreement on most other issues, this suggests that one interface should be provided for both blind and LV users but with customization options for the aspects in which differences were observed. Our system's initial prototype will, thus, provide a single interface for all users with short sentences for navigation directions and higher vibration frequency for indicating closeness to obstacles, as per the majority opinion. For speech, synthetic speech in English would be utilized as the default option, since half of the participants endorsed this and text-to-speech software for generating this is readily available. However, since human voice recordings and Arabic speech were also highly recommended, these options would be provided in future iterations.

Regarding visual output for LV users, varying light intensities—rather than colors—to indicate the distance to the obstacle was the preferred option. However, one-third of the LV participants dismissed the need for such feedback and the rest did not seem strongly convinced of its utility, mainly based on the difficulty they have experienced in viewing information on their smartphone screens (this corroborates the results reported by a previous study [14]). The initial prototype of our system will, therefore, not provide visual output but this option will be considered for inclusion in future iterations of the system subject to applying enhancements to the visual information, such as those suggested in [14, 81], to facilitate its perception.

The results indicate that most of the participants want to receive feedback from the system only when an obstacle is detected which is consistent with previous findings [8, 15, 61] indicating that the amount of information being provided should be minimized to avoid cognitive overload. Most of the participants—and especially, the majority of the blind participants—specified a distance of 1 m from an obstacle to start receiving feedback about it. It should be noted that this is contrary to the findings from previous studies [82] which recommend a distance of 2 m or more to give the user some time to stop or change direction to avoid the obstacle. We will defer to our participants' preferences in the design of the initial system prototype and issue speech-based instructions to avoid the obstacle at a distance of 1 m. However, to reconciliate the previous studies' recommendation, a vibration-based alert for a detected obstacle will start being issued at a distance of 2 m. If the testing reveals that either of these distances is impractical, they would be modified accordingly.

Though all participants favored footsteps as the distance measurement unit, this option does not appear to be practical since step length varies among individuals and even the same individual may adjust his gait according to the environment (e.g., walking slowly with smaller steps in a crowded setting vs. an uncrowded one); therefore, the standard unit of meters would be adopted in our system for specifying the distance.

Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 21 of 30

It should be noted that due to time constraints and the challenges of accessing and recruiting VI individuals for user research in the local context, the study was conducted with a relatively small sample of potential users all of whom were female and had quite similar demographic characteristics in terms of age, education, economic background and experience with technology. However, the invaluable insights which were obtained from these target users would be instrumental in ensuring that the design of the initial prototype of the system would meet basic user expectations and avoid some major usability issues which would otherwise not be uncovered until much further in the development process and, that too, after incurring a much higher cost if the system were developed without any prior consultation with the target users. For future iterations of the system, we plan to conduct user research and usability tests with a larger and more diverse sample to further refine the system design so that it would accommodate the needs of a broader range of potential users.

System design

The initial design of a system to assist VI users in detecting and avoiding obstacles while navigating indoors is presented in this section. The solution is being developed for Google's Project Tango Tablet Development Kit, a 7" Android tablet equipped with a 2.3 GHz NVIDIA Tegra K1 processor with 192 CUDA cores running on the Android 4.4 KitKat operating system. It has 4 GB RAM and 128 GB flash memory (expandable via microSD), a depth-sensing array (an infrared projector, 4 MP 2 μm rear-facing RGB/IR camera and 180° field of view fisheye rear-facing camera), several other sensors (a 120° front-facing camera, accelerometer, ambient light, barometer, compass, GPS, gyroscope), accurate sensor time stamping, and a software stack that enables application developers to use motion tracking, area learning and depth sensing [34, 59, 83].

The Java API provided by Tango [84], the Android NDK [85] and the Point Cloud Library [86] would be used for accessing and processing the data and developing the application.

An overview of the system is shown in Fig. 9. The user holds the tablet in his hands roughly at waist level slightly tilted forward with the screen facing towards him in landscape orientation. This ensures that the field of view of the cameras, which are located along the upper edge at the back of the tablet in this orientation, covers the walking area in front of the user in which obstacles are to be detected. As the user walks, depth data of the scene in front of him as well as his location and orientation information is acquired based on the input received from the various inbuilt sensors of the tablet. This data is then processed and analyzed by an obstacle detection module to detect any obstacles in front of the user. If any obstacle is detected, a feedback module then generates warnings and directions to avoid the obstacle(s) which are conveyed to the user via audio and vibrotactile feedback.

Since the data is acquired via the tablet's inbuilt sensors and the obstacle detection and feedback modules run on the tablet itself, this ensures that the system is self-contained and does not require any external hardware components for the data acquisition and processing. This also incorporates our interviewees' preference for no internet connectivity.
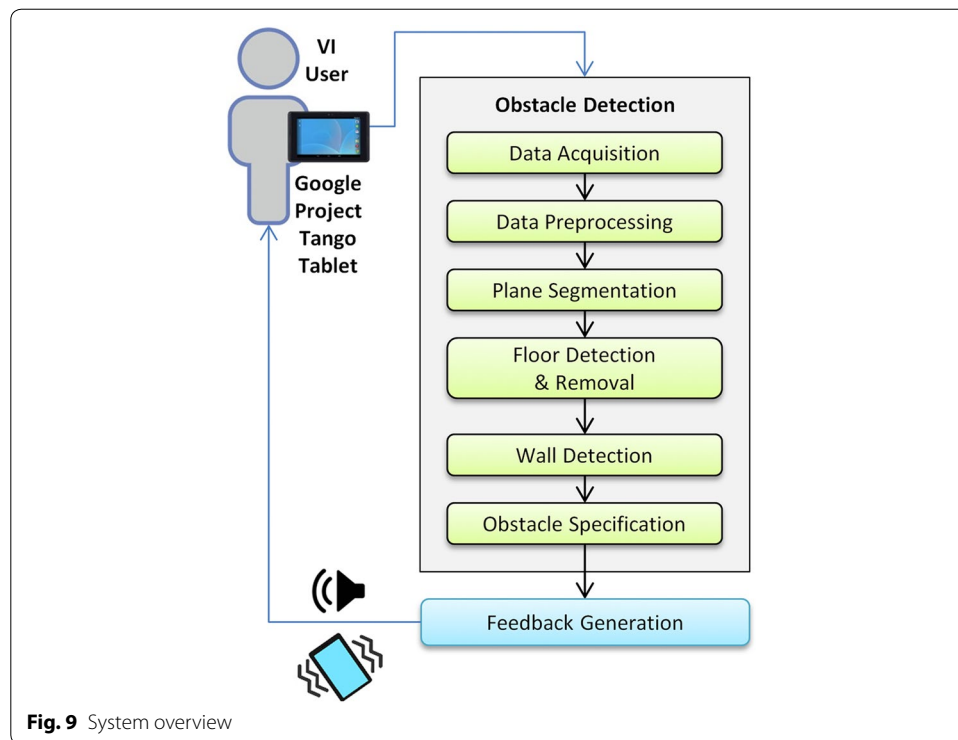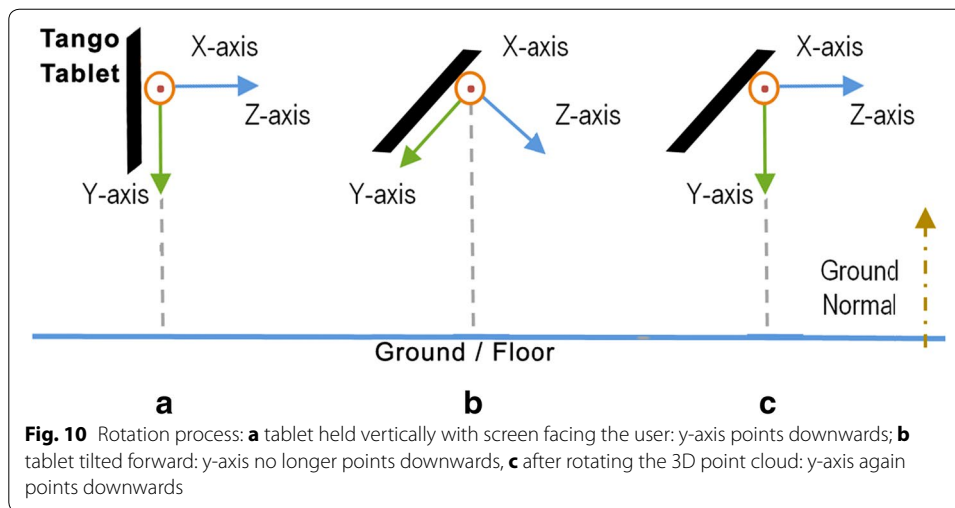
Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 22 of 30

Fig. 9 System overview

The design of the system—especially that of the feedback generation module—is informed by the results of the semi-structured interviews conducted with the VI users. The detailed design of the two main components of the system—the obstacle detection and the feedback generation modules—is presented below.

Obstacle detection

The obstacle detection component is designed to include six steps (as illustrated in Fig. 9):

1. *Data acquisition* Depth data of the scene in front of the user is acquired as a 3D point cloud using the Tango tablet's RGB-D camera; the data is captured at 320×180 resolution with adjusted frame rate (one frame per 800 ms) [87]. Working with the 3D point cloud data requires two processes to be executed: (1) The workable coordinate system should be set to the depth camera coordinate system, in which the y-axis is pointed down toward the gravity, the x-axis is pointed to the right and the z-axis is pointed forward away from the user [88] (Fig. 10a). However, when the user holds the tablet tilted at an angle, the y-axis is no longer pointed downwards (Fig. 10b). (2) The 3D point cloud data should, thus, be rotated around the x-axis utilizing the pose data provided by the Tango API which causes the y-axis to become parallel to the ground normal and the XZ plane to become parallel to the ground plane as shown in Fig. 10c.

2. *Data preprocessing* Two processes are executed in this step: (1) Noise removal: This involves discarding 3D points with low confidence values (lower than 0.8—note that the range of confidence values is [0-1] [33]) as well as those that lie beyond a 2 m range (according to the recommendation discussed in the "Design recommendations" section). (2) Downsampling: The noise-free point cloud is then down-sampled

Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 23 of 30

Fig. 10 Rotation process: **a** tablet held vertically with screen facing the user: y-axis points downwards; **b** tablet tilted forward: y-axis no longer points downwards, **c** after rotating the 3D point cloud: y-axis again points downwards

using the Voxel grid filtering algorithm [89] in order to reduce its complexity and speed up the execution time.

3. *Segmentation* Since indoor environments consist mainly of man-made objects with planar surfaces, this step combines all points that belong to the same plane together as one segment using the Random Sample Consensus (RANSAC) algorithm [90, 91], thus, dividing the point cloud into several planes.

4. *Floor plane extraction* Since the floor should not be considered as an obstacle, hence, this step identifies and removes the floor plane, which is defined as the plane that satisfies the following conditions: (1) Its surface normal is parallel to that of the XZ-plane (as shown in Fig. 10c); (2) The angle between it and the XZ plane is 0°; (3) It is at a distance of at least 1 m from the origin (the tablet's camera)—this constraint is based on the observation that an adult of average height [92] would hold the tablet at an elevation of at least 1 m from the ground.

5. *Wall detection* Since walls are frequently encountered in indoors environments, hence, this step detects any planes consisting of walls, where a plane is considered a wall plane if it meets the following two conditions, as specified in [93]: (1) its surface normal is perpendicular to the floor; (2) its size is larger than a minimum pre-specified wall plane size.

6. *Obstacle specification* In this step, the Euclidian Cluster Extraction algorithm is used, as in [93], which clusters points whose Euclidean distance is below a certain threshold; the resulting clusters are considered as obstacles. To specify the position of an obstacle, the scene is divided into three parts (similar to the approach followed in [94]): left, middle, and right, as shown in Fig. 11. The middle division represents the area directly in front of the user with an exact width of 1 m (this is considered to be enough space for a person to walk through), whereas the left and right divisions have equal width.

If the middle division does not contain any obstacles within a range of 1 m (note that the 1 m range is based on the interview results in the "User requirements elicitation study" section), the walking path directly in front of the user is clear of obstacles and safe
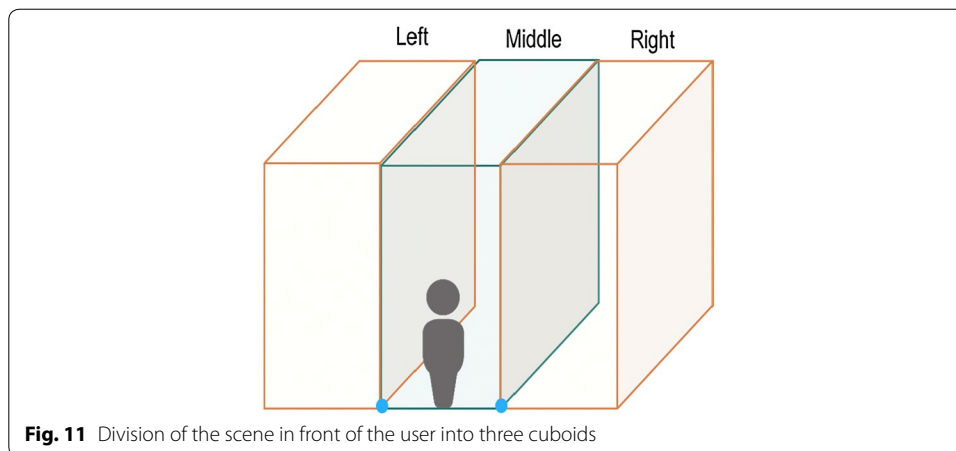
Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 24 of 30

Fig. 11 Division of the scene in front of the user into three cuboids

Table 3 Vibration frequencies and their meanings

Vibration frequency (vibrations/s)	Distance to obstacle (m)	Implication
1–3	>1 m	There is some space to continue walking forward
4–6	≤1 m	Walking forward will result in collision with an obstacle

to navigate. Otherwise, an obstacle warning is issued and the left and right divisions are checked; if either of them does not contain any obstacle within a 1 m range, the user is directed to bear in that direction. However, if both of them contain obstacles but beyond a range of 1 m, then the one with the obstacle(s) farthest away from the user is selected. In the worst case scenario, if all divisions contain obstacles within a 1 m range, then no safe path exists and the user is informed about this.

Feedback generation

In accordance with the design recommendations generated from the exploratory study, a multimodal user interface has been designed for our system which incorporates two modalities—touch (vibrations) and speech—to provide alerts and avoidance directions, if necessary, for any detected obstacles.

If no obstacles are detected in front of the user, no feedback is given in accordance with the preferences of the VI users. However, if an obstacle is detected in the path directly in front of the user, feedback is generated as follows. (1) The tablet starts vibrating; the frequency of the vibrations indicates the distance to the obstacle (see Table 3). (2) Speech-based feedback, generated using Android TTS, is provided to the user via bone conduction headphones to indicate the safest path for him to take to avoid colliding with the obstacle; this feedback is given only when the distance to the obstacle becomes less than 1 m (since the VI users we interviewed specified 1 m as the distance at which they would prefer to receive a warning about an obstacle). The speech messages consist of short sentences (again, according to the user preferences). Table 4 shows the various speech messages generated, along with their meanings. Note that "free path"

Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 25 of 30

Table 4 Speech messages and their meanings

Speech message	Meaning
Left and right are free	Free paths exist on both the left and the right in front of the user
Left is free	Free path exists only on the left in front of the user
Right is free	Free path exists only on the right in front of the user
Stop, no free path	No free path on the left, right, or middle in front of the user (dead end)

refers to a path that does not contain any obstacles at a distance of less than 1 m, and is, thus, safe to take.

One of the design recommendations from the exploratory study was that the identity of frequently appearing obstacles should be conveyed to the user. To keep the computational cost low in our initial prototype, we have decided to defer the incorporation an object recognition module to future iterations; however, since walls occur very frequently in indoors environment and help users in orienting themselves [8] and also, since wall planes can easily be identified based on some geometric constraints (as explained in step 5 in the "Obstacle detection" section), a wall detection step has been added to the obstacle detection component (see step 5 in the "Obstacle detection" section). The last message in Table 4 is then modified to "Wall: Stop, no free path" if the obstacle blocking the path is a wall.

The vibrations are generated by utilizing the inbuilt functions in the tablet itself and can be felt by the user via his hands which are in contact with the tablet. Since the participants expressed some concerns about audio feedback drawing unwanted attention as well as blocking out environmental sounds which are crucial for them to discern their context and orient themselves (note that similar concerns have also been voiced by VI users in other studies [8, 15, 61]), we have chosen open-ear Bluetooth-enabled bone conduction earphones to convey the audio feedback since the sounds from these are audible only to the wearer ensuring discreetness and they also do not occlude environmental sounds.

Conclusion and future work

The development of a novel depth-data based real-time obstacle detection and avoidance application for VI users to assist them in navigating independently in indoors environments has been presented. The application utilizes the capabilities of the Project Tango Tablet Development Kit to provide an aesthetically acceptable, cost-effective, portable, stand-alone solution for this purpose. A user-centered approach has been adopted throughout the development process in an effort to ensure that the resulting product would meet the target users' actual needs bolstering its potential for eventual acceptance and adoption by the end users. Semi-structured interviews with VI individuals in the local context were, thus, conducted to acquire a better understanding of their current micro-navigation practices and challenges and to gather their recommendations for an electronic assistive aid for this task. The invaluable insights gained from the interviews have not only informed the design of our system but would also benefit other assistive technology designers building similar applications. Following the exploratory study, an initial prototype of a micro-navigation system for the Tango platform was designed in

Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 26 of 30

accordance with the users' preferences; an overview of this system along with a detailed description of the obstacle detection and multimodal feedback generation modules has been provided. We plan to iteratively develop and test the initial prototype of the system with the end users to resolve any usability issues and better adapt it to their needs.

Subsequent to the validation of the initial prototype, the following work directions will be pursued to incorporate some additional features which were highly recommended by the interviewed participants: Since most interviewees requested that the identities of the obstacles be provided to them, a module that employs computer vision based methods to recognize commonly encountered obstacles and conveys their identities to the user will be added. Also, the ability to customize the feedback frequency and format would be included; since certain types of visual impairment are degenerative [61] and some clear differences were found among the interface preferences of LV and BL users, such customization would also facilitate accommodating the user in case of further vision loss. For speech feedback, Arabic language and human voice recording options may be provided in view of the interviewees' preferences; a visual output option may also be added for LV users, subject to addressing the concerns expressed by the participants about the perception of the visual information on the tablet; an option for speech-based obstacle warnings specifying the distance to the obstacles in steps may also be added for BL users (though this would require a training phase to calculate the user's average step length). Since obstacles such as descending stairs and distant ones (beyond a 2 m range) are not currently detected by our system but have been rated challenging by the interviewees, the obstacle detection mechanism would be updated to detect these objects as well. Furthermore, since the definition of what constitutes an obstacle differs from one user to the other [9, 15], and some users may not want to be alerted to obstacles which they can easily detect with their primary navigation aid (e.g., ground-level obstacles within the reach of the white cane) [15], an option to define what an obstacle is would be provided to the user. Also, a 3D printed holder which allows the tablet to be mounted on the waist would be designed to accommodate the preference regarding this expressed by the majority of the LV participants.

We are optimistic that our decision to adopt a user-centered approach and to use a mainstream standalone mobile device for the development would result in a usable solution widely accepted by the target users that would significantly assist them in autonomously detecting and avoiding obstacles in indoors environments. We hope that our project would inspire further research into user-centered design and development of micro-navigation systems which exploit the capabilities of the new generation of high-performance sensor-rich mobile devices to produce practical, user-friendly assistive solutions.

Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 27 of 30

References

1. Visual impairment and blindness: fact sheet number 282 (2014) http://www.who.int/mediacentre/factsheets/fs282/en/. WHO media center. http://www.who.int/mediacentre/factsheets/fs282/en/
2. Schwab L (2007) Eye care in developing nations. CRC Press, Boca Raton
3. Pascolini D, Mariotti SP (2012) Global estimates of visual impairment: 2010. Br J Ophthalmol 96(5):614–618
4. Manduchi R, Kurniawan S (2011) Mobility-related accidents experienced by people with visual impairment. AER J 4(2):44–54
5. Sharma S (2002) How does visual impairment affect performance on tasks of everyday life?: the SEE project. Evid Based Ophthalmol 3(4):218–219
6. Karlsson JS (1998) Self-reports of psychological distress in connection with various degrees of visual impairment. J Vis Impair Blindness 92:483–490
7. Wilson M (2015) "Who put that there!" The barriers to blind and partially sighted people getting out and about. https://www.rnib.org.uk/sites/default/files/Who%20put%20that%20there!%20Report%20February%202015.pdf. Accessed 3 Apr 2017
8. Paredes H, Fernandes H, Martins P, Barroso J (2013) Gathering the users' needs in the development of assistive technology: a blind navigation system use case. In: Stephanidis C, Antona M (eds) Universal access in human-computer interaction. Applications and services for quality of life: 7th international conference, UAHCI 2013, held as part of HCI international 2013, Las Vegas, NV, USA, July 21–26, 2013, Proceedings, Part III. Springer, Berlin, Heidelberg, pp 79–88. https://doi.org/10.1007/978-3-642-39194-1_10
9. Burigat S, Chittaro L (2012) Mobile navigation and information services for disabled students in university buildings: a needs assessment investigation. In: Proceedings of the 2nd workshop on mobile accessibility, MOBILE HCI 2012 conference, 2012
10. Quinones P-A, Greene T, Yang R, Newman M (2011) Supporting visually impaired navigation: a needs-finding study. In: CHI '11 extended abstracts on human factors in computing systems, Vancouver, BC, Canada, 2011. ACM, 1979822, pp 1645–1650. https://doi.org/10.1145/1979742.1979822
11. Shoval S, Ulrich I, Borenstein J (2003) NavBelt and the guide-cane [obstacle-avoidance systems for the blind and visually impaired]. Robot Automat Magazine IEEE 10(1):9–20
12. Lee YH, Medioni G (2014) Wearable RGBD indoor navigation system for the blind. In: Computer vision-ECCV 2014 workshops. Lecture Notes in Computer Science vol 8927. Springer, Berlin, pp 493–508
13. Jafri R, Khan MM (2016) Obstacle detection and avoidance for the visually impaired in indoors environments using Google's Project Tango device. In: Miesenberger K, Bühler C, Penaz P (eds) Computers helping people with special needs. Lecture notes in computer science, vol 9759. Springer, Cham, pp 179–185. https://doi.org/10.1007/978-3-319-41267-2_24
14. Szpiro S, Zhao Y, Azenkot S (2016) Finding a store, searching for a product: a study of daily challenges of low vision people. In: Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing, Heidelberg, Germany, 2016. ACM, 2971723, pp 61–72. https://doi.org/10.1145/2971648.2971723
15. Williams MA, Galbraith C, Kane SK, Hurst A (2014) "Just let the cane hit it": how the blind and sighted see navigation differently. In: Proceedings of the 16th international ACM SIGACCESS conference on computers & accessibility, Rochester, New York, USA, 2014. ACM, 2661380, pp 217–224. https://doi.org/10.1145/2661334.2661380
16. Zuckerman DM (2004) Blind adults in America: their lives and challenges. National Center for Policy Research for Women & Families, Washington, DC. http://www.center4research.org/wp/wp-content/uploads/2010/05/blind02041.pdf
17. Jafri R, Ali SA (2014) A multimodal tablet-based application for the visually impaired for detecting and recognizing objects in a home environment. In: Miesenberger K, Fels D, Archambault D, Peňáz P, Zagler W (eds) Computers helping people with special needs. Lecture notes in computer science, vol 8547. Springer, Cham, pp 356–359. https://doi.org/10.1007/978-3-319-08596-8_55
18. Jafri R, Ali SA, Arabnia HR, Fatima S (2014) Computer vision-based object recognition for the visually impaired in an indoors environment: a survey. TVC 30(11):1197–1222. https://doi.org/10.1007/s00371-013-0886-1
19. Coughlan J, Manduchi R (2007) Color targets: fiducials to help visually impaired people find their way by camera phone. J Image Video Process 1:096357. https://doi.org/10.1155/2007/96357

Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 28 of 30

20. Mascetti S, Ahmetovic D, Gerino A, Bernareggi C (2016) ZebraRecognizer: pedestrian crossing recognition for people with visual impairment or blindness. Pattern Recogn 60:405–419. https://doi.org/10.1016/j.patcog.2016.05.002

21. Takizawa H, Yamaguchi S, Aoyagi M, Ezaki N, Mizuno S (2015) Kinect cane: an assistive system for the visually impaired based on the concept of object recognition aid. Pers Ubiquit Comput 19(5):955–965. https://doi.org/10.1007/s00779-015-0841-4

22. Wang S, Tian Y Detecting stairs and pedestrian crosswalks for the blind by RGBD camera. In: bioinformatics and biomedicine workshops (BIBMW), 2012 IEEE international conference on 4–7 Oct 2012, pp 732–739. https://doi.org/10.1109/bibmw.2012.6470227

23. Mu K, Hui F, Zhao X (2016) Multiple vehicle detection and tracking in highway traffic surveillance video based on SIFT feature matching. J Inf Process Syst 12(2):183

24. Jafri R (2017) A GPU-accelerated real-time contextual awareness application for the visually impaired on Google's Project Tango device. J Supercomput 73(2):887–889. https://doi.org/10.1007/s11227-016-1891-8

25. Gilani SMM, Hong T, Cai Q, Zhao G (2017) Mobility scenarios into future wireless access network. J Inf Process Syst 13(2):236–255

26. Golledge R, Klatzky R, Loomis J, Marston J (2004) Stated preferences for components of a personal guidance system for nonvisual navigation. J Visual Impairment Blindness 98(3):135–147

27. Jafri R, Ali SA (2014) Exploring the potential of eyewear-based wearable display devices for use by the visually impaired. In: Proceedings of the 3rd international conference on user science and engineering (i-USEr 2014), Shah Alam, Malaysia, September 2–5 2014. IEEE Computer Society, pp 119–124

28. Shinohara K, Wobbrock JO (2011) In the shadow of misperception: assistive technology use and social interactions. In: Proceedings of the SIGCHI conference on human factors in computing systems, Vancouver, BC, Canada, 2011. ACM, 1979044, pp 705–714. https://doi.org/10.1145/1978942.1979044

29. Wang H-C, Katzschmann RK, Teng S, Araki B, Giarré L, Rus D (2017) Enabling independent navigation for visually impaired people through a wearable vision-based feedback system. In: 2017 IEEE international conference on robotics and automation (ICRA), pp 6533–6540

30. Golledge RG (1993) Geography and the disabled: a survey with special reference to vision impaired and blind populations. Trans Instit Br Geograph 18:63–85

31. Sánchez J, Sáenz M (2006) 3D sound interactive environments for blind children problem solving skills. Behav Inf Technol 25(4):367–378

32. Wise E, Li B, Gallagher T, Dempster AG, Rizos C, Ramsey-Stewart E, Woo D (2012) Indoor navigation for the blind and vision impaired: where are we and where are we going? In: 2012 international conference on Indoor positioning and indoor navigation (IPIN), pp 1–7

33. Google Project Tango. https://www.google.com/atap/project-tango/. Accessed 5 Feb 2016

34. Tango Developer Overview. https://developers.google.com/tango/developer-overview. Accessed 16 Dec 2016

35. Jafri R, Campos RL, Ali SA, Arabnia HR (2016) Utilizing the Google Project Tango tablet development kit and the unity engine for image and infrared data-based obstacle detection for the visually impaired. In: Proceedings of the 2016 international conference on health informatics and medical systems (HIMS'15), Las Vegas, Nevada, USA, July 27–30, 2016, pp 163–164

36. Yvonne R, Helen S, Preece J (2015) Interaction design: beyond human—computer interaction, 4th edn. Wiley, New York

37. Microsoft Kinect. https://dev.windows.com/en-us/kinect. Accessed 23 Dec 2015

38. Structure Sensor, Occipital Inc. http://structure.io/. Accessed 5 Feb 2016

39. Lenovo Phab 2 PRO. http://shop.lenovo.com/us/en/tango/. Accessed 17 Dec 2016

40. Asus Zenfone AR (ZS571KL). https://www.asus.com/Phone/ZenFone-AR-ZS571KL/. Accessed 8 Jan 2017

41. Munoz R, Rong X, Tian Y (2016) Depth-aware indoor staircase detection and recognition for the visually impaired. In: 2016 IEEE international conference on multimedia & expo workshops (ICMEW), pp 1–6

42. Zöllner M, Huber S, Jetter H-C, Reiterer H (2011) NAVI—A proof-of-concept of a mobile navigational aid for visually impaired based on the Microsoft Kinect. In: Campos P, Graham N, Jorge J, Nunes N, Palanque P, Winckler M (eds) Human–computer interaction—INTERACT 2011, vol 6949. Lecture Notes in Computer Science. Springer, Berlin, pp 584–587. https://doi.org/10.1007/978-3-642-23768-3_88

43. Filipe V, Fernandes F, Fernandes H, Sousa A, Paredes H, Barroso J (2012) Blind navigation support system based on Microsoft Kinect. Procedia Comput Sci 14:94–101. https://doi.org/10.1016/j.procs.2012.10.011

44. Brock M, Kristensson PO (2013) Supporting blind navigation using depth sensing and sonification. In: Proceedings of the 2013 ACM conference on pervasive and ubiquitous computing adjunct publication, Zurich, Switzerland. ACM, New York, pp 255–258. https://doi.org/10.1145/2494091.2494173

45. Huang H-C, Hsieh C-T, Yeh C-H (2015) An indoor obstacle detection system using depth information and region growth. Sensors 15(10):27116–27141

46. Ribeiro F, Florencio D, Chou PA, Zhengyou Z (2012) Auditory augmented reality: object sonification for the visually impaired. In: 2012 IEEE 14th international workshop on multimedia signal processing (MMSP), 17–19 Sept. 2012, pp 319–324. https://doi.org/10.1109/mmsp.2012.6343462

47. Khan A, Moideen F, Lopez J, Khoo W, Zhu Z (2012) KinDectect: Kinect detecting objects. In: Miesenberger K, Karshmer A, Penaz P, Zagler W (eds) Computers helping people with special needs, vol 7383. Lecture Notes in Computer Science. Springer, Heidelberg, pp 588–595. https://doi.org/10.1007/978-3-642-31534-3_86

48. Liu H, Wang J, Wang X, Qian Y (2015) iSee: obstacle detection and feedback system for the blind. In: Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing and proceedings of the 2015 ACM international symposium on wearable computers, Osaka, Japan, 2015. ACM, 2800917, pp 197–200. https://doi.org/10.1145/2800835.2800917

49. H.R. 1902—108th Congress: Medicare Vision Rehabilitation Services Act of 2003 (2003) https://www.govtrack.us/congress/bills/108/hr1902. Accessed 13 Sept 2016

Jafri and Khan *Hum. Cent. Comput. Inf. Sci.* (2018) 8:14

Page 29 of 30

50. Mann S, Huang J, Janzen R, Lo R, Rampersad V, Chen A, Doha T (2011) Blind navigation with a wearable range camera and vibrotactile helmet. In: Proceedings of the 19th ACM international conference on multimedia, Scottsdale, Arizona, USA, 2011. ACM, 2072005, pp 1325–1328. https://doi.org/10.1145/2072298.2072005

51. Wang S, Pan H, Zhang C, Tian Y (2014) RGB-D image-based detection of stairs, pedestrian crosswalks and traffic signs. J Vis Commun Image Represent 25(2):263–272

52. Tang TJJ, Lui WLD, Li WH (2012) Plane-based detection of staircases using inverse depth. In: Australasian conference on robotics and automation (ACRA)

53. Pérez-Yus A, López-Nicolás G, Guerrero JJ (2014) Detection and modelling of staircases using a wearable depth sensor. In: Computer vision-ECCV 2014 workshops. Springer, Berlin, pp 449–463

54. Simpson R, LoPresti E, Hayashi S, Guo S, Ding D, Ammer W, Sharma V, Cooper R (2005) A prototype power assist wheelchair that provides for obstacle detection and avoidance for those with visual impairments. J Neuro Eng Rehabil 2:30

55. Yelamarthi K, Haas D, Nielsen D, Mothersell S RFID and GPS integrated navigation system for the visually impaired. In: 2010 53rd IEEE international midwest symposium on circuits and systems (MWSCAS), 1–4 Aug. 2010, pp 1149–1152. https://doi.org/10.1109/mwscas.2010.5548863

56. Anderson D (2015) Navigation for the visually impaired using a Google Tango RGB-D tablet

57. Tiresias (2015) An app to help visually impaired people navigate easily through unfamiliar buildings. HackDuke 2015. http://devpost.com/software/tiresias-ie0vum

58. Lock J, Cielniak G, Bellotto N (2017) Portable navigations system with adaptive multimodal interface for the blind. In: AAAI 2017 spring symposium—designing the user experience of machine learning systems, Stanford, CA, 27th–29th March 2017. AAAI

59. Jafri R, Campos RL, Ali SA, Arabnia HR (2017) 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. https://doi.org/10.1109/access.2017.2766579

60. Li B, Muñoz JP, Rong X, Xiao J, Tian Y, Arditi A (2016) ISANA: wearable context-aware indoor assistive navigation with obstacle avoidance for the blind. In: European conference on computer vision. Springer, Berlin, pp 448–462

61. Sánchez J, Elías M (2007) Guidelines for designing mobility and orientation software for blind children. In: Baranauskas C, Palanque P, Abascal J, Barbosa SDJ (eds) Human–computer interaction—INTERACT 2007: 11th IFIP TC 13 international conference, Rio de Janeiro, Brazil, September 10-14, 2007, proceedings, Part I. Springer, Berlin, pp 375–388. https://doi.org/10.1007/978-3-540-74796-3_35

62. Miao M, Spindler M, Weber G (2011) Requirements of indoor navigation system from blind users. In: Holzinger A, Simonic K-M (eds) Information quality in e-Health: 7th conference of the workgroup human-computer interaction and usability Engineering of the Austrian Computer Society, USAB 2011, Graz, Austria, November 25–26, 2011. Springer, Berlin, Heidelberg, pp 673–679. https://doi.org/10.1007/978-3-642-25364-5_48

63. Benabid A, AlZuhair M (2014) User involvement in the development of indoor navigation system for the visually impaired: a needs-finding study. In: 2014 3rd international conference on user science and engineering (i-USEr), 2–5 Sept. 2014, pp 97–102. https://doi.org/10.1109/iuser.2014.7002684

64. Hamilton-Fletcher G, Obrist M, Watten P, Mengucci M, Ward J (2016) I always wanted to see the night sky: blind user preferences for sensory substitution devices. In: Proceedings of the 2016 CHI conference on human factors in computing systems. ACM, New York, pp 2162–2174

65. "SurveyGizmo| Professional Online Survey Software & Form Builder", Surveygizmo.com. https://www.surveygizmo.com/. Accessed 07- Apr- 2016

66. Ministry of Social Affairs. http://www.mosa.gov.sa. Accessed 9 June 2017

67. King Khalid University Hospital, Riyadh, Saudi Arabia. https://medicine.ksu.edu.sa/en/node/371. Accessed 26 Dec 2017

68. Willings C Notetaker instruction. Teaching students with visual impairments

69. Apple Inc. http://www.apple.com/iphone/

70. Android Developers. https://developer.android.com/index.html

71. Apple Store. https://www.apple.com/retail/. Accessed 27 Dec 2017

72. Siri, Apple Inc. https://www.apple.com/ios/siri/. Accessed 1 May 2016

73. Vincent J (2017) Google upgrades its iPhone keyboard with voice dictation and 15 new languages. The Verge, New York

74. WhatsApp Inc. https://www.whatsapp.com/. Accessed 1 May 2016

75. Apple accessibility: VoiceOver, Apple.com. http://www.apple.com/sa/accessibility/ios/voiceover/. Accessed 1 May 2016

76. Mobile Operating System Market Share Saudi Arabia (November 2016–November 2017). StatCounter. http://gs.statcounter.com/os-market-share/mobile/saudi-arabia. Accessed 27 Dec 2017

77. Android, the world's most popular mobile platform. https://developer.android.com/about/android.html Accessed 26- Dec- 2017

78. Kobie N (2017) Android overtakes windows to become the internet's most popular OS. Wired, New York

79. Google Talkback. https://play.google.com/store/apps/details?id=com.google.android.marvin.talkback&hl=en. Accessed 26 Dec 2017

80. Android access: getting started with Android. http://www.androidaccess.net/intro.php. Accessed 26 Dec 2017

81. Zhao Y, Szpiro S, Knighten J, Azenkot S (2016) CueSee: exploring visual cues for people with low vision to facilitate a visual search task. In: Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing. ACM, New York, pp 73–84

82. Kim SY, Cho K (2013) Usability and design guidelines of smart canes for users with visual impairments. Int J Design. 7:1

83. Project Tango Tablet Teardown. https://www.ifixit.com/Teardown/Project+Tango+Tablet+Teardown/28148. Accessed 6 Apr 2017

84. Tango Java API. https://developers.google.com/tango/apis/java/. Accessed 27 Dec 2017

Jafri and Khan *Hum. Cent. Comput. Inf. Sci. (2018) 8:14*

Page 30 of 30

85. Getting Started with the NDK| Android Developers, Developer.android.com. https://developer.android.com/ndk/guides/index.html. Accessed 4 Dec 2016
86. Point Cloud Library (PCL), Pointclouds.org. http://pointclouds.org/about/. Accessed 4 Dec 2016
87. What is the resolution of the Project Tango tablet?. Stackoverflow.com (2016) https://stackoverflow.com/questions/37490158/what-is-the-resolution-of-the-project-tango-tablet. Accessed 9 Jun 2016
88. Tango developers overview: frames of reference. https://developers.google.com/tango/overview/frames-of-reference. Accessed 26 Aug 2017
89. PCL/OpenNI tutorial 2: Cloud processing (basic)-Robótica-ULE. http://robotica.unileon.es/mediawiki/index.php/PCL/OpenNI_tutorial_2:_Cloud_processing_(basic)#Outlier_removal. Accessed 17 Mar 2016
90. Kurban R, Skuka F, Bozpolat H (2015) Plane segmentation of kinect point clouds using RANSAC. In: The 7th international conference on information technology, pp 545–551
91. Derpanis KG (2010) Overview of the RANSAC algorithm. Image Rochester NY 4(1):2–3
92. Wikipedia: the free encyclopedia. Wikimedia Foundation Inc. "List of Average Human Height Worldwide". https://en.wikipedia.org/w/index.php?title=List_of_average_human_height_worldwide&oldid=806422555. Accessed 3 Apr 2017
93. Pham H-H, Le T-L, Vuillerme N (2016) Real-time obstacle detection system in indoor environment for the visually impaired using Microsoft Kinect Sensor. J Sens. https://doi.org/10.1155/2016/3754918
94. Gupta S, Girshick R, Arbeláez P, Malik J (2014) Learning rich features from RGB-D images for object detection and segmentation. In: European conference on computer vision. Springer, Berlin, pp 345–360