



# Interactive Travel Aid for the Visually Impaired: from Depth Maps to Sonic Patterns and Verbal Messages

Piotr Skulimowski<sup>(✉)</sup>  and Pawel Strumillo 

Institute of Electronics, Lodz University of Technology, Lodz, Poland  
piotr.skulimowski@p.lodz.pl

**Abstract.** This paper presents user trials of a prototype micro-navigation aid for the visually impaired. The main advantage of the system is its small form factor. The device consists of a Structure Sensor depth camera, a smartphone, a remote controller and a pair of headphones. An original feature of the system is its interactivity. The user can activate different space scanning modes and different sound presentation schemes for 3D scenes on demand. The results of the trials are documented by timeline logs recording the activation of different interactive modes. The aim of the first trial was to test system capability for aiding the visually impaired to avoid obstacles. The second tested system efficiency at detecting open spaces. The two visually impaired testers performed the trials successfully, although the times required to complete the tasks seem rather long. Nevertheless, the trials show the potential usefulness of the system as a navigational aid and have enabled us to introduce numerous improvements to the tested prototype.

**Keywords:** Assistive technology · Electronics travel aids · Visual impairment · Depth maps · Scene segmentation · Sonification

## 1 Introduction

Building an electronic travel aid (ETA) for the visually impaired, whether dedicated to micro- or macro-navigation tasks, has proven itself to be a difficult interdisciplinary challenge. Research efforts to build an out-of-laboratory ETA device date back to 1889, when the first attempts were undertaken by a Polish scientist named Kazimierz Noiszewski. Noiszewski built the “Electroftalm”, a device that converted visual signals to auditory stimulation and utilized the photoelectric properties of Selenium cells [1]. However, the device was heavy and inconvenient to use, so did not find any practical applications. Today, there are small form factor devices available that enable reconstruction of the 3D structure of the environment, such as structural light and stereo-vision. Such techniques have been applied in practice, e.g. in prototype electronic travel aids (ETAs) assisting the blind in mobility [2]. However, analysis of the operation of ETA solutions, and the fact that they have not been accepted or widely used by blind people, points to the need to develop more efficient, user-centered methods of presenting information about the three-dimensional environment to the blind.

Major barriers faced by blind people are spatial orientation and independent mobility. This is because the visually impaired cannot build accurate and dynamic internal cognitive models of the environment using their non-visual senses. The primary navigational aid for the blind is a white cane, which increases the range of haptic perception of the environment but only to a range of about 1.5 m. The authors of a very recent study [3] conducted a state of the art survey of wearable assistive devices for the blind and visually impaired. They investigated more than 70 different solutions, of which 29 used scene depth reconstructed from time of flight (ToF) cameras or stereovision. Based on this review, it can be concluded that the existing solutions do not enable the visually impaired to use interactive techniques to explore the environment.

In this work, we describe novel non-visual techniques which engage the senses of touch and hearing to give the visually impaired an impression of their environment. The novelty of the proposed solution is the interactive mode of audio-haptic presentation, whereby the user can flexibly control the amount of incoming information. The blind user can acquire a capacity for selective attention, similar to that for the sense of vision. This work is a continuation of our earlier studies, reported in [4], where we proposed a scheme for sonification of U-depth maps (i.e. histograms of the consecutive columns of the depth map) of the environment. The sound was generated as a sum of sinusoidal sounds, each one dedicated to the predefined depth range. The user interactively selected the depth map region for sonification by using touch gestures on the mobile device screen.

## 2 Related Work

Recently, there have been many scientific efforts aimed at building assistive devices for the visually impaired. Comprehensive reviews of research into technological aids for the visually impaired are given in [3, 5]. Electronic travel aids are devices or systems that aid mobility and navigation for the visually impaired [6]. These personal systems convert data about the environment into auditory or haptic stimulation [7]. Such use of healthy sensory modality to substitute a lost sense is termed sensory substitution [8]. Some approaches attempt to use sound to encode spatial data about the environment [9]. Others use haptic or tactile displays to present the geometric layout of the environment [10]. Text-to-speech (TTS) technologies have also been developed, which enable the visually impaired to access the written word. Notable examples are the optophone and Kurtzweil's "reading machine" [11]. Simpler technologies convert visual images into haptic or auditory modalities. Audio description is a method of explaining graphics or images in the form of verbal narration, commenting on the visual information. However, this requires human participation in interpreting the imaged scene content. Other solutions are based on automatic conversion of visual data into auditory signals [6] or haptic stimulation [12].

Unfortunately, few of the technical solutions to aid the visually impaired present non-visual information about the environment in an interactive way [13]. By interactivity, we mean functionality that would enable the blind user to take over control of the incoming stream of non-visual information presented by the device or computer application about the environment. Only a few interactive assistive approaches for the

visually impaired are reported in literature. For example, in [14] line elements of images were sonified, while in [15] Nintendo's Wiimote controller was utilized to provide the user with combined sonic and haptic feedback. The image was pre-segmented and the detected regions were assigned descriptors that were further mapped to sounds. However, the mapping method was not explained. A multilevel sonification approach has been proposed in [16], whereby low-level image features such as edges, colors, and textures are sonified in parallel, and machine learning algorithms are applied to recognize and verbally describe objects to the user. Finally, a system was reported in [17], which integrates a geographic information system of a building with computer vision for navigating a blind individual. Its advantage is that only one camera, e.g. smartphone camera can be used without any need for any additional hardware.

### 3 Materials and Methods

#### 3.1 System Hardware Platform

Our ETA device consists of a depth camera connected to a smartphone with Android OS, a pair of headphones and a remote controller. Using the remote controller, the blind user can set non-visual presentation parameters and the operating mode of the device. Figure 1 shows the system mounted on a mannequin head.



**Fig. 1.** Components of the ETA system, consisting of a depth camera, an Android OS smartphone, a pair of headphones and a remote controller.

The depth camera is a Structure Sensor (SS). It is a lightweight device (65 g) that uses an infrared structured light pattern projector and a low-range infrared CMOS camera to reconstruct 3D scene geometry, with a range of depth reconstruction of 0.4–5 m.

The SS generates depth images with a rate of 30 fps with  $640 \times 480$  pixel spatial resolution (or 60 fps and  $320 \times 240$  pixels), both with depth relative accuracy of less 1%. We selected the 60 fps mode. The horizontal field of view (FoV) of the SS is  $58^\circ$  and the vertical FoV is  $45^\circ$ . The SS device is mounted on the head using rubber bands (see Fig. 1). A special rig is used to set various camera tilt angles. The chosen camera does not work outdoor in a direct sunlight, but it can be replaced in the future with a new class of depth sensors equipped with stereo infrared cameras and laser projectors. Although, wearing a headset is not very comfortable, it allows to scan the scene by moving user's head. The Android OS smartphone is a Samsung Note 3 equipped with a Qualcomm Snapdragon 800 processor. The image processing software is written in Java and uses OpenNI2 drivers. The remote controller is a small form factor device to be held in the user's hand, with a set of convex keys that "click" when pressed. The controller communicates with the smartphone via a Bluetooth link.

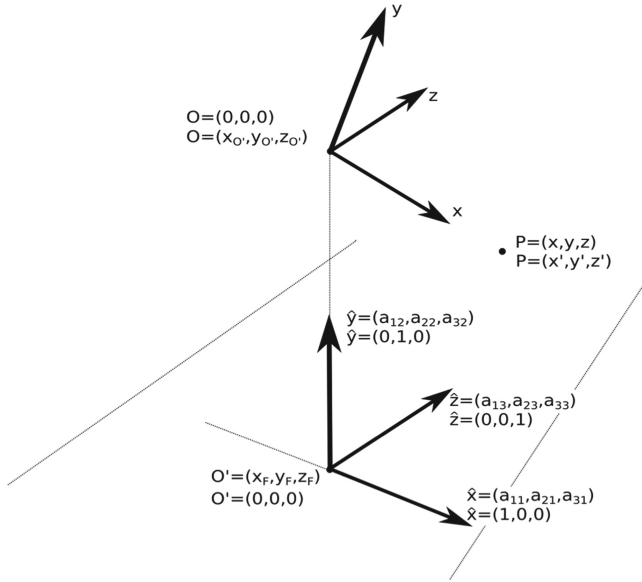
### 3.2 Image Processing

Our approach to designing an ETA is that the blind person should construct a higher level understanding of the environment from low-level features such as edges, fragments of planes or open space devoid of obstacles. This approach is strongly supported by earlier studies on visual rehabilitation and by in-depth understanding of the neural mechanisms behind nonvisual multisensory presentation of space [8, 18].

Reliable estimation of the position of the ground plane in a scene is a very important element of any navigation system for the blind. The ground plane should not be identified as an obstacle and should not be presented to the user. However, the accuracy of its detection determines other important system capabilities, such as identification of small but important obstacles or defects in the ground. Many automatic methods of detecting ground planes have been proposed [19, 20], especially for vehicle-related applications. These solutions, however, fail in the case of scenes with complicated structures, such as where the ground plane is visible only in a small part of the image, or in imaging systems where 6DoF (Degrees of Freedom) camera motion is enabled.

In the research described in this work, we applied the method of determining the ground plane described in our previous study [4]. This method is based on the assumptions that the system is placed on the user's forehead and the distance from the camera to the ground is within the presumed range. Knowing the orientation and location of the ground plane versus the camera optical axis, it is possible to determine the area of the plane in the image even if it is visible in only a small region of the image. In the present study, we limit the search area of the ground plane to the bottom part of the depth image [4]. Our earlier experiments showed that the tilt (pitch angle) of the camera towards the ground plane should be slightly less than half of the vertical field of view of the camera, which is  $22^\circ$ . This makes it possible to observe both the ground plane and the objects located slightly above the user's head. With these assumptions, the distance to the obstacle should be defined in a new coordinate system, which is related to the ground plane camera  $O'$  rather than to the camera coordinate system  $O$  (see Fig. 2).

Let the ground plane equation  $Ax + By + Cz + D = 0$  and the origin of the coordinate system associated with the ground  $O'$  be the point that is the projection of the origin of the coordinate system associated with the camera  $O$  on the ground plane. The new origin of the  $O'$  coordinate system in the  $O$  coordinate system has coordinates  $[x_F, y_F, z_F]$ , which are determined by assuming that the origin of the camera coordinate system is the closest point to the ground plane.



**Fig. 2.** Geometric configuration of the camera coordinate system ( $O$ ) and the ground related coordinate system ( $O'$ ). Coordinate systems  $O$ ,  $O'$ , point coordinates  $P$ , and the basis vectors are defined in the camera coordinate system (top line) and in the ground related coordinate system (bottom line), respectively.

The basis unit vectors of the  $O'$  coordinate system in the camera coordinate system  $O$  are  $\hat{x} = (a_{11}, a_{21}, a_{31})$ ,  $\hat{y} = (a_{12}, a_{22}, a_{32})$ ,  $\hat{z} = (a_{13}, a_{23}, a_{33})$ . The basis vector  $\hat{y}$  has the same direction as the normal vector to the ground plane  $[A, B, C]$ . The basis vector  $\hat{z} = [x_F, y_F, z_F]$  is oriented along the line for which the scene depth is defined. Finally,  $\hat{x}$  can be computed from the outer product  $\hat{x} = \hat{z} \times \hat{y}$ . It can be proven [21] that any point  $(x', y', z')$  defined in ground-related coordinate system  $O'$  has the following  $(x, y, z)$  coordinates in the camera coordinate system  $O$ :

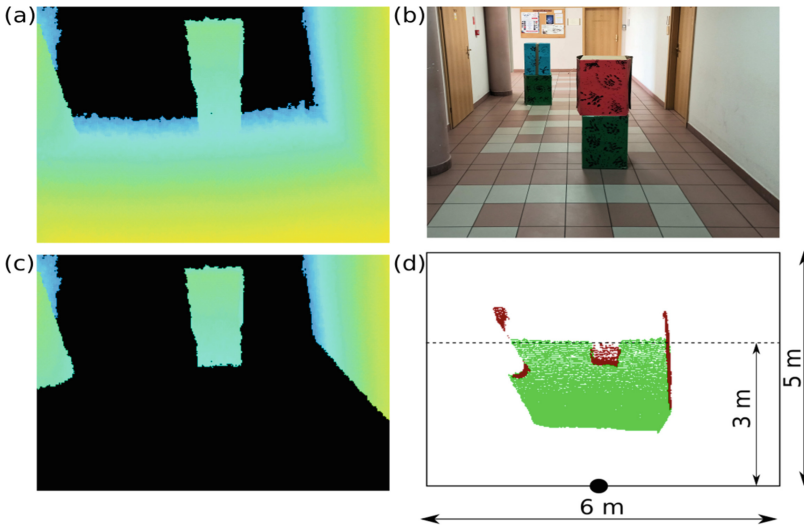
$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x_F \\ y_F \\ z_F \end{bmatrix} + \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \quad (1)$$

Similarly, an arbitrary point  $(x, y, z)$  in the camera coordinate system  $O$  has  $(x', y', z')$  coordinates in the new  $O'$  ground-related coordinate system:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x_{O'} \\ y_{O'} \\ z_{O'} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \tag{2}$$

where  $[x_{O'}, y_{O'}, z_{O'}]^T$  are the coordinates of the origin of the camera coordinate system given in the  $O'$  coordinate system, which can be calculated by solving Eq. 1 for the pair of points  $x = [0, 0, 0]^T$  and  $x' = [x_{O'}, y_{O'}, z_{O'}]^T$ .

Using Eq. 2, one can find the coordinates of each point in the scene defined in the  $O'$  coordinate system associated with the ground plane. Given these coordinates, an obstacle map (termed ‘‘occupancy grid’’) can be built. Figure 3 shows an example of an indoor scene, the result of detection and removal of the ground plane, and the occupancy grid of the scene geometric content for an area of 6 m × 5 m in front of the user.

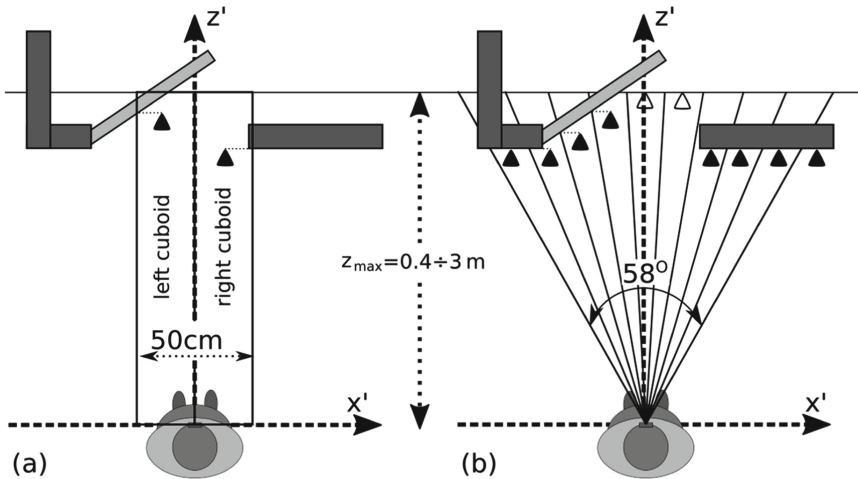


**Fig. 3.** Depth map captured by the camera shown in pseudo-colors (a), the corresponding indoor 3D scene image (b), the depth map with ground plane removed (c), and the occupancy grid (top view) generated for the depth map (d) (the green region represents the detected ground plane, brown regions represent detected obstacles, and the black dot denotes the position of the user in the scene). (Color figure online)

### 3.3 Presentation Methods

Two obstacle presentation methods are proposed: one based on parking sensors, the other on horizontal space scanning in the entire angular range of the SS depth camera (see Fig. 4). In the first method, which is the default mode, obstacles are detected in front of

the system user in two adjacent cuboids, each measuring the selected depth range, the user's height (estimated automatically by the system), and half the width of the user's shoulders (Fig. 4a). The user is informed by audio signals resembling a car parking sensor the nearest obstacle(s) in each cuboid. Sounds with a constant frequency of 457 Hz are generated for 30 ms. The interval between successive sounds  $dT$  (milliseconds) depends on the distance to an obstacle according to the heuristic formula:



**Fig. 4.** Comparison of the proposed sound modes for presenting obstacles: a) in the *default mode* obstacles are detected in front of the user only, b) the *scanning mode* presents the obstacle by horizontal scanning of the of the field of view. The black and white small triangles indicate depth values communicated to the user. Black triangles indicate sectors in which scene objects are within system depth range, whereas white triangles denote sectors in which scene objects are further than 3 m away from the camera.

$$dT = 0.2d + \frac{d^2}{40000} \quad (3)$$

where  $d$  is given in millimeters. The sound volume depends on the distance: the closer the obstacle, the louder the sound. In the absence of obstacles a low, quiet and cyclic sound with a frequency of 300 Hz is generated, confirming that the system is operating correctly.

If the user is not moving (such inactivity is detected by analyzing signals form the accelerometers integrated into the mobile device carried by the user) then the “verbal mode” of the system is automatically activated. In this mode, the system informs the user about the nearest distance (in meters) to any obstacle by short verbal commands, e.g. “zero three” means 0.3 m, “one two” means 1.2 m. There is a distinction between distances from the obstacles in each cuboid: if the difference between the distances is less than 10 cm then the smallest distance value is read centrally in the headphones. If the difference is larger than 10 cm, then two values are read sequentially, for the left

and right cuboid separately. The voice commands are panned slightly left or right for the corresponding cuboid. To avoid a cluttered sequence of verbal cues, the next distance is communicated to the user after not less than 2 s.

The second mode of system operation, *scanning mode*, is activated on demand by pressing the appropriate button on the remote control. In this presentation mode, the entire space in the depth camera field of view is divided into 9 vertical sectors, each with an angular range of approx.  $6.5^\circ$ . The scanning mode features two options (selected by the user): *scanning sonification mode* and the *scanning verbal mode*. In *scanning sonification mode*, the nearest obstacle in each sector is found and a sound with a frequency dependent on the distance is generated. A sound of 2000 Hz is generated for the nearest obstacles (0.4 m), then the further away the obstacle the lower the sound frequency, with the lowest frequency of 370 Hz corresponding to a depth of 3 m. Nine different sound frequencies are applied, covering in equal steps a depth range of  $0.4 \div 3$  m. These frequencies assume the following values given in Hz: 2000, 1619, 1312, 1062, 860, 697, 564, 457, and 370. These were computed by applying a logarithmic scale for mapping depth to frequencies. The duration of the sound is 100 ms, followed by 150 ms of silence. The sounds are generated in the headphones with panning between the left and right channels corresponding to the direction of the obstacle. To avoid ambiguity in direction detection, for sectors where there are no obstacles a muted sound with a frequency of 300 Hz is generated, indicating no obstacles. The entire scanning cycle lasts 2.4 s. In *scanning verbal mode*, the user can select verbal messages for reading the distance to any obstacles in consecutive sectors. As with the sounds, the spoken messages are presented using volume panning between the left and right headphones. If no obstacle is detected in a given sector, the message “empty” is spoken.

## 4 Mobility Tests of the System and Discussion

Mobility tests of a prototype version of the system were conducted with two blind people who were also asked for their opinions and ideas for improvements. The tests were approved by the Bioethics Commission at the Medical University of Lodz, Poland. Prior to the mobility trails, the users were acquainted with the system features, given time to test the system, and asked to share their impression of its usability. Table 1 summarizes user comments and steps taken to improve system functionality.

The mobility tests of the system were carried out with participation of two visually impaired individuals, a woman aged 34 (Tester 1) and a man aged 37 (Tester 2). Tester 1 belongs to visual impairment category 4 as defined by the World Health Organization [22], whereas Tester 2 belongs to category 3. Tester 1 is aided by a guide dog and Tester 2 uses a white cane as a primary mobility aid. During the trials, the testers did not use their primary travel aids.



Two indoor mobility testing scenarios were carried out:

- 1) a walk down a corridor with cardboard boxes positioned at random locations to evaluate the ability of the system to detect obstacles and aid the user to maintain the direction of travel along the corridor,
- 2) a walk down an empty corridor and finding a door opening, to evaluate the ability of the system to locate and present off-path empty spaces.

**Table 1.** Main user comments addressing system usability features and modifications of the system in response to user comments

No.	User comment	System modification
1	Increased speed of response to changes in obstacle location	In the <i>default mode</i> the system response to changing position of obstacle(s) was increased to a rate close to 60 fps
2	Frequencies used in the <i>scanning sonification mode</i> are too high and considered annoying	Frequencies reduced to the range of 370 Hz ÷ 2000 Hz
3	100% left/right panning of the sound in the <i>default mode</i> is misleading because the perceived obstacle(s) were well off the collision track	Sound panning was shifted to a more central location to position the sounds within the virtual cuboids in front of the users (see Fig. 4a)
4	No user interaction for the presentation mode. Interaction control of system functions and parameters was strongly advised by the users	Interactive presentation modes added on demand, i.e. <i>scanning sonification mode</i> and <i>scanning verbal mode</i> . This involved adding a remote control
5	The system depth detection range of 5 m is too large. Distant obstacles are detected which are irrelevant to the current mobility task of the blind user	The default obstacle detection range was reduced to 3 m (approx. twice the range of a white cane)

Figure 5 shows photographs taken during mobility testing scenarios 1) and 2), which were carried out indoors on the university campus. The testers were informed that they could freely use the different operating modes of the system, should the need arise. We implemented system capability to record time-stamped logs, noting which system operating modes were activated during the testing trials. A video recorded during one of the trials containing on-line segmentation results of the scene and the generated sound messages is available at: <https://youtu.be/dAYiw-OhjEw>.

#### 4.1 Trial 1: A Walk Down a Corridor with Cardboard Box Obstacles

The task of the user in this trial was to walk along a corridor approx. 3 m wide and avoid two obstacles randomly positioned on the walking path before stopping in front of a wall (see Fig. 5). The length of the path was approx. 12 m. The users kept the smartphone securely in their pocket and held the remote controller that served to activate the different operating modes of the system in their hand.

The data recorded by the smartphone application during this test is illustrated in Fig. 6. The upper panel shows the result for Tester 1 and the bottom panel the results for Tester 2. Note that Tester 1 needed  $T = 87$  s to walk the path, whereas Tester 2 walked the same path in less than  $T = 70$  s. Tester 1 frequently stopped to listen to voiced messages giving information on the distance to the obstacles (indicated by arrows on a timeline). Tester 1 preferred this presentation mode to the “parking sensor” mode, which is the default. Also note that Tester 1 did not switch on any of the scanning modes. Tester 2, on the other hand, used the scanning verbal mode twice and the scanning sonification mode 7 times.



Fig. 5. The two proposed testing scenarios: a walk down a corridor with cardboard boxes simulating obstacles (a), finding a door opening on an empty corridor (b).

Moreover, he stopped 12 times during trail, while Tester 1 stopped 27 times, which is one stop every 3 s (on average) while walking. These times seem long (in particular for Tester 1). However, we observed that both testers were excessively testing the capabilities of the system, rather than concentrating on achieving the task quickly (the results were obtained for one preliminary trial).

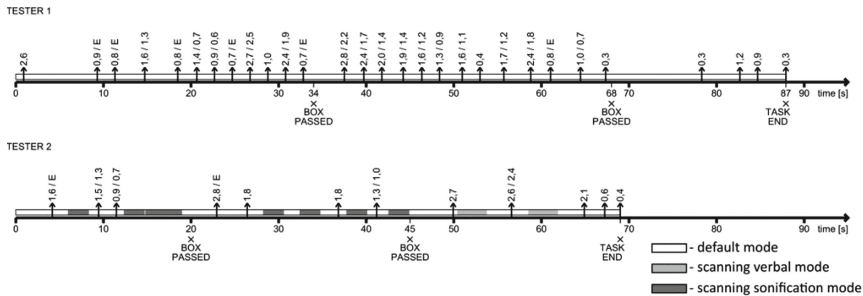
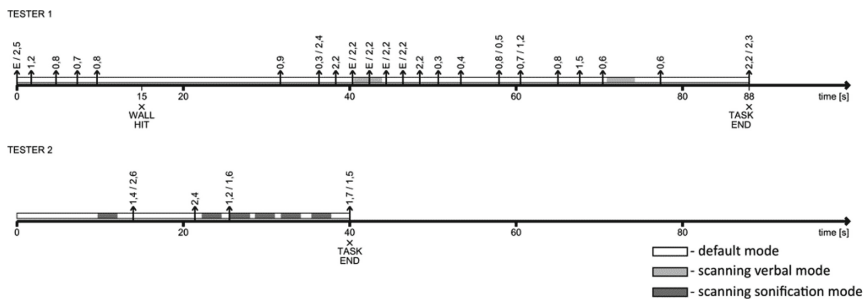


Fig. 6. Timelines showing how Testers 1 and 2 performed the task of walking along a corridor with obstacles. Vertical arrows indicate when the testers stopped and voice messages informed them of the distance to obstacle(s) in meters (E – denotes empty space). The crosses below the timelines indicate when the testers avoided the obstacles (no obstacle collisions were noted).

## 4.2 Trial 2: A Walk Down an Empty Corridor and Finding a Door Opening

The aim of the second test scenario was to evaluate how efficiently the system can aid the blind user in finding an open space, in this case an opened door. The task was to walk along an empty corridor and find a door opening. The users were told on which side of the corridor they should expect the door opening, but were not given any indication as to the location of the door along the corridor. The distance to the door opening from the starting point was in fact approx. 10 m.

The timelines showing how the testers performed in this trial are shown in Fig. 7. Note again that Tester 2 completed the task in a much shorter time ( $T = 40$  s) than Tester 1 ( $T = 85$  s). As in the first trial, Tester 1 was using mainly the default system mode and only twice switched on the scanning verbal mode. Tester 1 preferred to stop and by head turns obtain confirmation of her position relative to the corridor wall. This strategy was not very efficient, because it did not exploit the scanning modes designed to solve this type of task. Note also that Tester 1 touched the wall in spite of having been warned that she was approaching an obstacle (see the upper panel in Fig. 7). Tester 2, on the other hand, showed much greater skill in efficiently using the system. In the second part of the test, he frequently used the scanning sonification mode to search for the door opening (see the lower panel in Fig. 7) and stopped only three times while walking along the path.



**Fig. 7.** Timelines showing how Testers 1 and 2 performed the task of walking along an empty corridor and locating a door opening (explanations of the symbols are given in the caption in Fig. 6).

## 5 Conclusions

In this study, we presented user trials of a prototype ETA system aimed at aiding the visually impaired to complete micro-navigation tasks. The system has a small form factor and lightweight. An important feature of the system is its interactivity. The user can activate different space scanning methods and sound presentation modes on demand (see Fig. 4). Both verbal messages and sonification schemes for presenting 3D scenes are available. An important property of the tested prototype is that all computations related to processing depth images (e.g. ground plane detection) and auditory

presentations are performed on-line, on the smartphone platform. Initial trials of the system, although carried out with participation of only two visually impaired users, have shown the potential viability of the system as a micro-navigation aid. User feedback about the system was used to modify the sound presentation parameters and improve system functionality (see Table 1).

Comparing to our previous work [4] and added value is the way of interacting with the remote control, which is much more convenient than using the touch screen of the phone. Moreover, presenting data based on the 3D scene information instead of its U-depth representation is more detailed and more intuitive according to the comments of the trial blind participants.

Finally, we are aware that more rigorous trials are needed with a larger group of visually impaired users. In particular the testers should be given more time to attain automatism for efficient use of the system. Further research should also test how helpful the system is in combination with the users' primary travel aids (a white cane or a guide dog).

## References

1. Starkiewicz, W., Kuliszewski, T.: Progress report on the elektroftalm mobility aid. In: Proceedings of the Rotterdam Mobility Research Conference, pp. 27–38. American Foundation for the Blind, New York (1965)
2. Strumillo, P., et al.: Different approaches to aiding blind persons in mobility and navigation in the “Naviton” and “Sound of Vision” projects. In: Pissaloux, E., Velázquez, R. (eds.) *Mobility of Visually Impaired People*, pp. 435–468. Springer, Cham (2018). [https://doi.org/10.1007/978-3-319-54446-5\\_15](https://doi.org/10.1007/978-3-319-54446-5_15)
3. Tapu, R., Mocanu, B., Zaharia, T.: Wearable assistive devices for visually impaired: a state of the art survey. *Pattern Recogn. Lett.* (2018). <https://doi.org/10.1016/j.patrec.2018.10.0311>
4. Skulimowski, P., Owczarek, M., Radecki, A., Bujacz, M., Rzeszotarski, D., Strumillo, P.: Interactive sonification of U-depth images in a navigation aid for the visually impaired. *J. Multimodal User Interfaces* **13**(3), 219–230 (2018). <https://doi.org/10.1007/s12193-018-0281-3>
5. Hersh, M., Johnson, M. (eds.): *Assistive Technology for Visually Impaired and Blind People*. Springer, London (2008). <https://doi.org/10.1007/978-1-84628-867-8>
6. Bujacz, M., Strumillo, P.: Sonification: review of auditory display solutions in electronic travel aids for the blind. *Arch. Acoust.* **41**(3), 401–414 (2016)
7. Dakopoulos, D., Bourbakis, N.G.: Wearable obstacle avoidance electronic travel aids for blind: a survey. *IEEE Trans. Man Cybern. Part C Appl. Rev. Syst.* **40**(1), 25–35 (2010)
8. Maidenbaum, S., Abboud, S., Amedi, A.: Sensory substitution: closing the gap between basic research and widespread practical visual rehabilitation. *Neurosci. Biobehav. Rev.* **41**, 3–15 (2014)
9. Dobrucki, A., Plaskota, P., Pruchnicki, P., Pec, M., Bujacz, M., Strumillo, P.: Measurement system for personalized head-related transfer functions and its verification by virtual source localization trials with visually impaired and sighted individuals. *J. Audio Eng. Soc.* **58**(9), 724–738 (2010)
10. Brock, A., Jouffrais, Ch.: Interactive audio-tactile maps for visually impaired people. *ACM SIGACCESS Access. Comput. (ACM Digit. Libr.)* **113**, 3–12 (2015). Association for Computing Machinery (ACM)

11. Kurzweil, R.: The Kurzweil reading machine: a technical overview. *Science, Technology and the Handicapped* (1976)
12. Pissaloux, E., Maingreud, F., Velazquez, R., Fontaine, E.: Concept of the walking cognitive assistance: experimental validation. *ASME Int. J. Adv. Model. C* **67**, 75–86 (2007)
13. Pissaloux, E.E., Velazquez, R. (eds.): *Mobility in Visually Impaired People - Fundamentals and ICT Assistive Technologies*. Springer, Cham (2018). <https://doi.org/10.1007/978-3-319-54446-5>
14. Ramloll, R., Yu.W., Brewster, S., Riedel, B., Burton, M., Dimihgen, R.: Constructing sonified haptic line graphs for the blind student: first steps. In *Proceedings of the 4th International ACM Conference on Assistive Technologies*, pp. 17–25 (2000)
15. O'Neill, Ch., Ng, K.: Hearing images: interactive sonification interface for images. In: *EVA 2008 London Conference 22–24 July*, pp. 188–195 (2008)
16. Banf, M., Blanz, V.: A modular computer vision sonification model for the visually impaired. In: *Proceedings of International Conference of Auditory Display* (2012)
17. Serrão, M., et al.: Computer vision and GIS for the navigation of blind persons in buildings. *Univ. Access Inf. Soc.* **14**(1), 67–80 (2015)
18. Zhang, M., Weisser, V.D., Stilla, R., Prather, S., Sathian, K.: Multisensory cortical processing of object shape and its relation to mental imagery. *Cogn. Affect. Behav. Neurosci.* **4**(2), 251–259 (2004)
19. Kwon, J., Dragon, R., Gool, L.V.: Joint tracking and ground plane estimation. *IEEE Signal Process. Lett.* **23**(11), 1514–1517 (2016)
20. Lin, Y., Guo, F., Li, S.: Road obstacle detection in stereo vision based on UV-disparity. *J. Inf. Comput. Sci.* **11**(4), 1137–1144 (2014)
21. Stark, M.: *Analytical geometry*. Mathematical Monographies 26. Mathematical Institute, Polish Academy of Sciences (1951). (in Polish)
22. WHO webpage. International classification of diseases for mortality and morbidity statistics. <http://www.who.int/classifications/icd/en/>. Accessed 04 Jan 2020