



Article

Spot-Presentation of Stereophonic Earcons to Assist Navigation for the Visually Impaired

Yuichi Mashiba ¹, Ryunosuke Iwaoka ¹, Hisham E. Bilal Salih ², Masayuki Kawamoto ³, Naoto Wakatsuki ², Koichi Mizutani ² and Keiichi Zempo ^{2,*} 

- ¹ Graduate School of Systems and Information Engineering, University of Tsukuba, Tsukuba, Ibaraki 305-8573, Japan; mashiba@aclab.esys.tsukuba.ac.jp (Y.M.); iwaoka@aclab.esys.tsukuba.ac.jp (R.I.)
- ² Faculty of Engineering, Information and Systems, University of Tsukuba, Tsukuba, Ibaraki 305-8573, Japan; hisham@aclab.esys.tsukuba.ac.jp (H.E.B.S.); wakatsuki@iit.tsukuba.ac.jp (N.W.); mizutani@iit.tsukuba.ac.jp (K.M.)
- ³ Headquarters for International Industry-University Collaboration, University of Tsukuba, Tsukuba, Ibaraki 305-8575, Japan; kawamoto.masayuki.gn@un.tsukuba.ac.jp
- * Correspondence: zempo@iit.tsukuba.ac.jp

Received: 12 June 2020; Accepted: 16 July 2020; Published: 20 July 2020



Abstract: This study seeks to demonstrate that a navigation system using stereophonic sound technology is effective in supporting visually impaired people in public spaces. In the proposed method, stereophonic sound is produced by a pair of parametric speakers for a person who comes to a specific position, detected by an RGB-D sensor. The sound is a stereophonic earcon representing the target facility. The recipient can intuitively understand the direction of the target facility. The sound is not audible for anyone except for the person being supported and is not noisy. This system is constructed in a shopping mall, and an experiment is conducted, in which the proposed system and guidance by a tactile map lead to a designated facility. As a result, it is confirmed, that the execution time of the proposed method is reduced. It is also confirmed that the proposed method shows higher performance in terms of the average time required to grasp the direction than the tactile map approach. In the actual environment where this system is supposed to be used, the correct answer rate is over 80%. These results suggest that the proposed method can replace the conventional tactile map as a guidance system.

Keywords: earcon; parametric speaker; visually impaired people; assistive technology; sonification

1. Introduction

As of October 2017, it has been estimated that 130 million people suffer from visual impairment globally, and this number is expected to rise to more than 550 million by 2050 owing to the increasing prevalence of disabilities in the growing older population arising from increased lifespan [1]. Despite the large number of people affected by visual impairments, visually impaired people still face various obstacles in daily life, such as lack of accessibility to information [2]. For example, information about many architectural environments and traffic is not accessible to the visually impaired, making it difficult for them to use various facilities in public spaces and in particular creating problems when using public transportation. This greatly restricts the movement of the visually impaired, so that they often require assistance with going out. Therefore, the visually impaired need a system that can smoothly transmit information about the map of their surroundings.

Previously, Braille, speakers, and braille block were used to provide information to the visually impaired, but these methods have several disadvantages when used as a means of presenting the surrounding map information. For tactile maps, Braille usage is limited by the fact that only less than

10% of visually impaired people can read Braille. In addition, it is difficult for people who have recently become visually impaired to learn Braille. When speakers are used, the audio broadcast by the speaker can be heard by people other than the person needing support, thus increasing noise. Additionally, for some cases of the language used, support cannot be provided. Finally, by nature, braille blocks cannot convey the direction of the facility. Owing to the shortcomings of the existing approaches, there is a need for a system that can easily provide visually impaired people with information and can transmit peripheral map information, regardless of the presence or absence of visual impairment or differences in the language used.

Therefore, in this study, we seek to construct an acoustic navigation system for the use in a noisy environment based on stereophonic sound technology and earcons with reduced noise for the environment. The earcon to be presented is information for the direction of the facility relative to the person needing support (Figure 1).

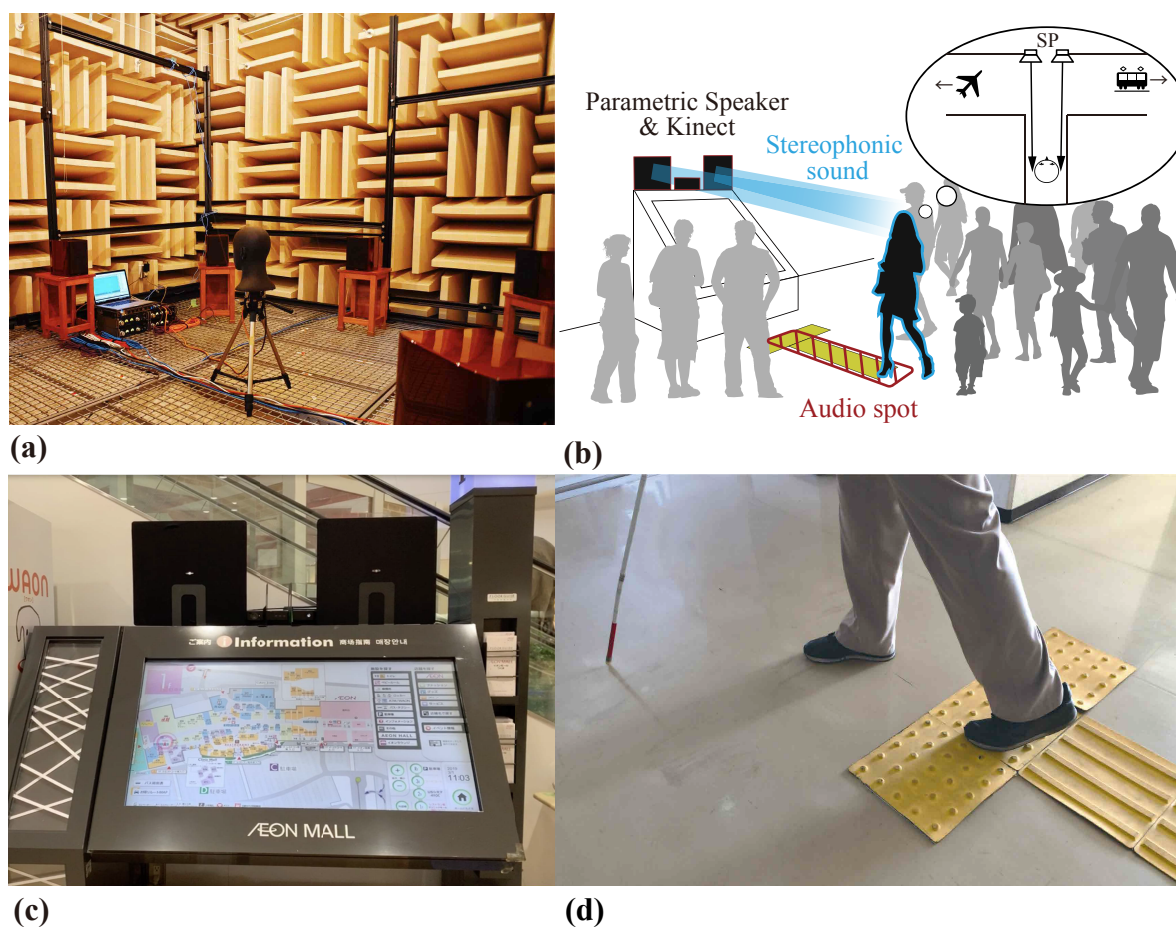


Figure 1. (a) Head transfer function (HRTF) is measured using a dummy head. A stereophonic sound can be produced by the convolutional integration of an HRTF to an arbitrary sound source in time domain. (b) Installation image of the proposed method. Using the stereophonic sound, it is possible to present an earcon bearing facility information only to a person who needs support among a crowd. Using the proposed method, the direction of the facility can be grasped intuitively. (c) Image of an actual experiment in a shopping mall. A parametric speaker and Kinect are installed on the digital signage. This provides a mechanism for presenting information when a recipient approaches. (d) Using this system, the range of movement of the visually impaired is extended to areas without a braille block.

2. Method

2.1. Related Work

A white cane is widely used by people who are visually impaired for independent travel. Other than that, a braille block is used in some countries (Figure 2). Although it is easy to learn how to use these tools, they cannot provide textual data such as facility information. Tactile maps are used to compensate for this, but learning Braille is indispensable for using tactile maps. The acquisition rate of Braille is less than 10% among visually impaired people [3], and it is clear that many visually impaired persons cannot use it. For this reason, studies have been conducted to make tactile maps audible and provide information by voice [4–6]. However, it is necessary to search for the presence of a tactile map by fumbling, which is a burden for the visually impaired.

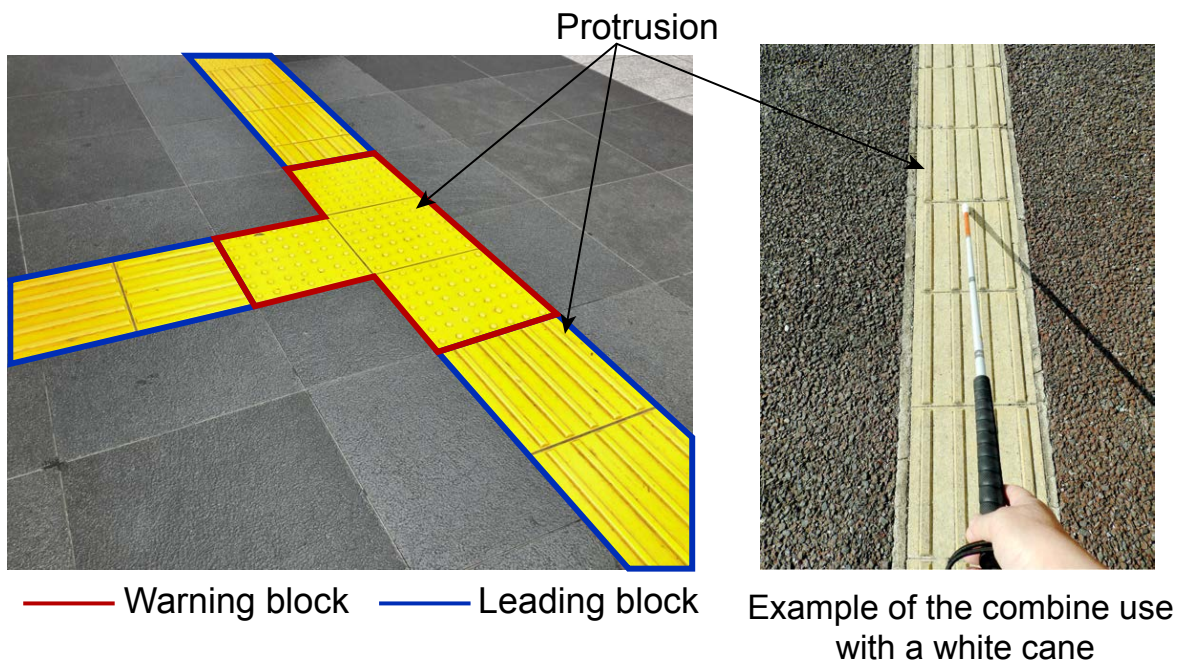


Figure 2. There are two types of braille blocks: leading blocks and warning blocks. The leading block is installed so that a visually impaired people can follow the direction indicated by the protrusion while checking them with the sole of the foot or a white cane. The warning block is a block indicating a position to be noted. It is installed in front of stairs, in front of pedestrian crossings, in front of information boards, in front of obstacles, etc.

Auditory displays can easily provide a large amount of information, as compared to tactile displays. In recent years, there has been research on indoor localization, route planning, and navigation using a smartphone or tag. With GPS, it is difficult to estimate self-position in insufficient radio wave environments such as indoors. For this reason, self-localization using tags installed in the environment is mainly studied. Tag readers are mainly implemented on smartphones and white papers. The method is divided according to the type of tag to be installed, and the first is RFID tag [7–11]. The RFID tag can be installed inexpensively, but the readable distance is short, and a large amount of them should be installed. The IR signal may not be readable by the receiver because of obstacles. It can also be affected by natural or artificial light [12]. In addition, there is an obstacle problem in self-localization using the QR code [13,14]. However, there are studies that use Bluetooth [15,16] or perform indoor LED lighting and visible light communication with smartphones to estimate self-position [17]. These navigation systems can present information to the user anytime and can provide detailed guidance to the destination.

On the other hand, there is conventionally a guidance method using voice from a speaker installed in the environment. Because there is no need to have a receiver and information can be obtained just by listening to the sound, there is no burden on the visually impaired and installation can be found in various places such as stations. For this reason, guidance signs of public facilities have been standardized internationally [18].

The disadvantages of using a speaker are sound diffusion and a possible mismatch between the languages of the user and the guidance voice from the speaker. The approaches to these problems are listed below.

2.1.1. Parametric Speaker

In Miyauchi's study, a parametric speaker is used, and guidance can be performed on the opposite bank on a pedestrian crossing without braille blocks [19,20]. However, it is desirable to guide the visually impaired in multiple directions. Therefore, Aoki et al. showed that multidirectional presentation is possible by presenting stereophonic sound from parametric speakers [21].

However, these methods have the following problems in stereophonic presentation and guide speech. For stereophonic sound presentation, it was pointed out that the direction of the sound image depends on the shape of the head and the auricle, and the effectiveness of this approach has been verified in a noiseless environment, but not in a noisy environment where the system must actually be used. As mentioned above, if the language of the guidance voice is different from that of the person needing support, the person cannot receive the necessary guidance.

2.1.2. Auditory Display

The term auditory display refers to the presentation of information through the auditory sense. Users do not have to move their bodies to receive information and do not have to focus their visual attention. It is also easy to attract the user's attention. Transmitting some events by sound has been studied using a computer interface since the 1980's, and several representative examples are described below.

1. Auditory icons

Auditory Icons are sounds associated with objects in a non-natural language. Gaver first proposed the use of auditory icons as a computer interface in 1986 [22–25]. The sound of scrapping wastepaper heard when a file is dropped into the trash is an example of an auditory icon, in which the sound can be intuitively associated with an object without being remembered.

2. Earcons

Earcons were proposed by Blattner et al. in 1989 [26]. Each sound represents a different event. For example, file deletion is indicated by a sequence of a file and an erase command. Western music is adopted in this approach to express the information with the hierarchical sound structure. Earcons have been proposed as a non-GUI interface for navigational cues in the menu hierarchy [27,28] and mobile phones [29,30].

Vice versa, a guidance method that uses earcons does not rely on the knowledge of a specific language. However, currently, there is no unified standard for earcons, so it is necessary to memorize earcons for each facility. Therefore, by creating a unified standard for earcon generation, the burden of memorizing the earcons can be reduced.

2.2. Proposed Method

2.2.1. System Overview

The proposed system consists of an RGB-D sensor (Kinect), a parametric speaker, and a PC (Figure 3). When the person needing support reaches a specific position, a stereophonic earcon is presented to both ears. The RGB-D sensor (Kinect) detects that the recipient is in a specific position,

and stereophonic sound is generated by a parametric speaker. The presented sound signal is called an earcon, and is a stereophonic sound signal containing map information, where each sound signal is a melody with a short phrase. The acoustic signal is stereophonic, and the recipient perceives the earcon from multiple directions. Thus, the map information is presented only to the person needing support, and the direction of the destination can be recognized by simply listening to the sound generated by stereophonic sound presentation, so that the problems of noise, the direction of the facility, and the language of the presentation voice can be solved. The next section describes the creation of parametric speakers, stereophonic sound, and design indicators for earcons.

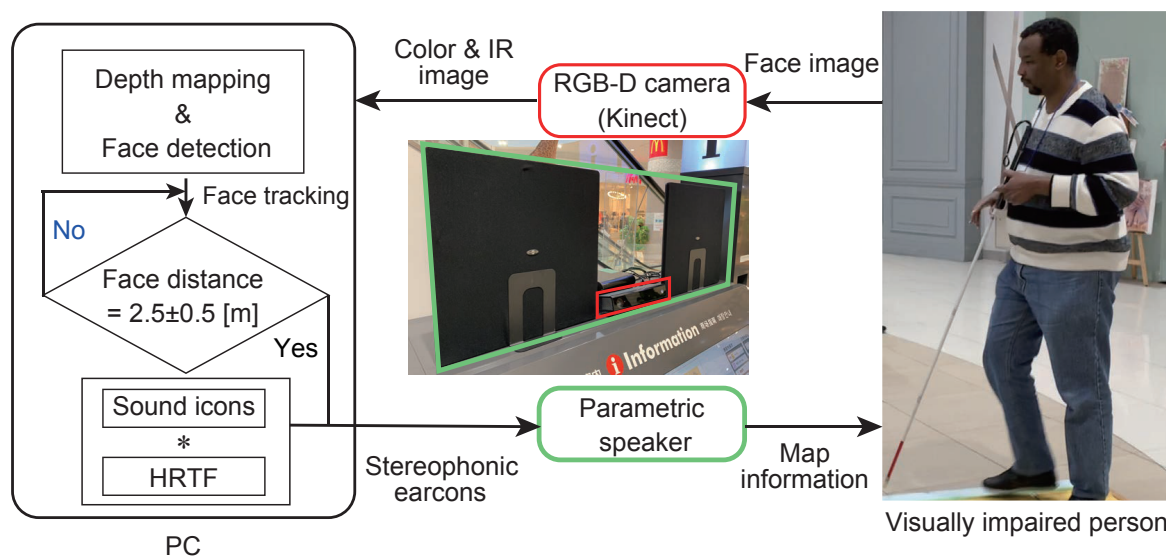


Figure 3. Block diagram of the proposed method: Face recognition is performed using images from an RGB-D camera. When a face approaches a certain distance, a stereophonic sound including map information is presented by a parametric speaker.

2.2.2. Parametric Speaker

A traditional loudspeaker emits an audible sound that diffuses and propagates through diffraction. However, parametric speakers emit ultrasonic waves as carrier waves. Ultrasound is an inaudible sound with high directivity and a frequency of greater than 20 kHz. The ultrasonic wave radiated from the parametric speaker is demodulated into audible sound by the nonlinearity of the medium, and the language problem of the presentation voice can be solved.

2.2.3. Stereophonic Sound

Humans perceive the sound direction from a sound source when it reaches the eardrum after it is reflected and diffracted on the floor, shoulder, head, pinna, etc. [31]. Physical characteristics including differences in the sound volume, time, and frequency from the sound source to the eardrum are called the head-related transfer function (HRTF). The HRTF mimics the shape of a human head, and is measured using a device called a dummy head with a microphone in the ear canal or by inserting a microphone into the human ear canal. By convolving the HRTFs of the right and left ears in time domain measured as shown in Figure 4 into an audio signal and then playing back the respective sounds in the left and right ears, the audio signal can be localized. However, because individual HRTFs are different, the use of a commercially available dummy head, voice recorded with another person's head, or convolutional voice of the HRTF measured with the dummy head may lead to the user hearing the voice in a direction that deviates from the original audio signal direction. Therefore, Perrott et al. proposed the use of dynamic binaural signals in the presentation of stereo sound images [32]. The dynamic binaural signal is an audio signal that convolves the HRTF in time domain

following the direction of the head of the listener. Dynamic binaural signals increase the information content of the sound image, making it easier to sense the direction. However, the dynamic binaural signal must be listened only using wearable sound devices (earphone and headphone), and it must be appropriately measured in front of the listener by attaching the sensor to the listener’s head. In this study, we implemented a recording method using a dummy head, in which a dynamic binaural signal can be received without using a headphone.

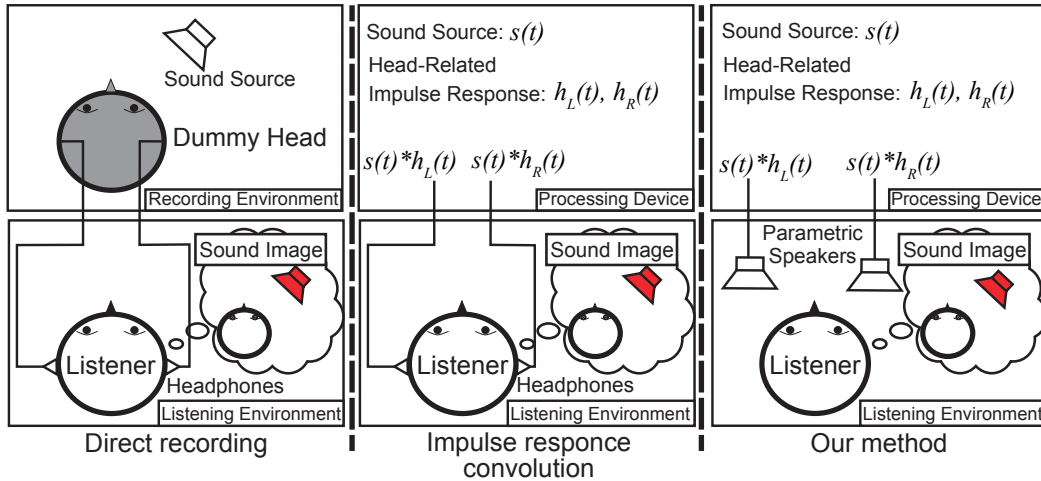


Figure 4. Method for providing localization to a sound image.

In our proposed system, information about facilities is transmitted by earcons. Auditory icons are superior mobile service notifications, as compared to earcons [33]. However, it is difficult to create sounds that can intuitively remind the user of the facility, so that sound icons are not intuitively symbolic, but abstract. Therefore, the sounds used in our system are classified as earcons. The user must remember the facilities associated with sounds, but can learn quickly [34]. Moreover, the recall rates of earcons are affected by learning techniques [34]. Various studies have been conducted on the design index of earcons [35,36]. The original developers of earcon proponent, Blattner et al., have proposed that earcons should have common features based on similar characteristics of the objects that they are referring to [26]. This study adopts this proposal, and earcons are generated according to a unified standard. By unifying the design rules of earcons, the burden on the user is reduced because the user does not have to memorize the earcons repeatedly. It is also possible to create new earcons by combining the features of the existing earcons.

In addition, the design index of earcons is described. Earcons must be designed so as not to give an unpleasant feeling when listening. Roberts’s subjective experiments showed that, among the major, minor, diminished, and augmented chords, major chords were identified as the most harmonious sounds [37]. Therefore, all earcons in this work were created with melodies using major chords. The facilities that are audibly represented by the earcons in our work are often located in public spaces; more specifically, we focus on elevators, escalators, stairs, and toilets. Each earcon represents the facility using the difference in the tone color, and is designed based on the constraint that expresses the attribute of the facility by the changes in the tone, such as up and down. Objects that are related conceptually to the directions of up and down had earcons with the common feature of the melody going up and down on the scale (Figure 5). The musical scale of the men’s room is lowered by an octave, and that of the women’s room is raised by an octave. Figure 6 shows the score of the created earcons. As described in the next section, using the earcons designed based on the design index above, we verify that the proposed method can guide visually impaired and sighted people to the desired destination in a noisy environment.

scale note	Up	Down
staccato	Up stairs	Down stairs
note 2 minutes	Up ESL	Down ESL
slur	Up ELV	Down ELV
Whole note	Men`s toilet	Women`s toilet

ESL = Escalator, ELV =Elevator

Figure 5. Design indicator of the earcons.

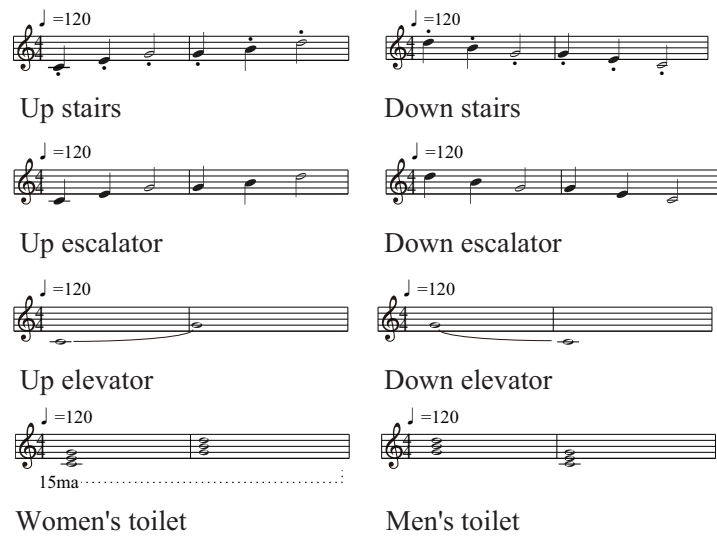


Figure 6. Notes for each earcon: each facility has its own tone color, with the scale going up and down. For the toilet, the musical scale changes according to the gender.

2.3. Contribution

In this paper, we propose a nonwearable stereophonic presentation system for visually impaired people that can be used for map presentation. Presentation of stereophonic sound consists of a pair of stereo parametric speakers, and by delivering stereophonic sound only to the user, it is possible to present a sound image centered on the user’s position regardless of the position of the speaker. Intuitive understanding is encouraged. In the next section, we describe experiments, whose results confirm that the proposed method can be a substitute for conventional tactile maps, which need to be touched, and map presentation by announcement that provides the user with surrounding information in natural language.

3. Experiment

3.1. Overview

It was verified that the proposed method can guide the user to the destination in a noisy environment. The proposed method was compared experimentally to the guidance to the destination using a tactile map. Experiments were carried out by installing parametric speakers, tactile maps, and braille blocks around the digital signage in a shopping mall. The warning block of the braille block was set to be the point where the stereophonic sound could be heard (Figure 7).

In addition to the proposed method, a tactile map was installed around the bifurcation point of the braille block, and similar experiments were conducted (Figure 7). Also, we conducted a questionnaire about the proposed system.

3.2. Environment

The experiment was conducted in a section of a shopping mall (AEON Mall Tsukuba). In the experiment, a braille block sheet was placed around the digital signage. There are two types of braille blocks: leading and warning blocks. The leading block is installed so that visually impaired people can follow the direction indicated by the protrusion while checking them with the sole of the foot or a white cane. The warning block is a block indicating a position to be noted. It is installed in front of stairs, pedestrian crossings, information boards, obstacles, etc.

For the braille block, a leading block and a warning block of 0.3 m per side were used. Additionally, a parametric speaker (Holosonics AS-16B) was installed on the digital signage (Figure 7). When a subject reaches a branch point (warning block) on the braille block, a stereophonic acoustic signal including map information is transmitted.

The noise level (A-weighted sound pressure level (A-weighting is a frequency dependent curve which reflects the response of the human ear to noise.)) of the environment was measured by a noise meter (RION NL-31). Measured for 5 min. The average A-weighted sound pressure level was as follows: $L_{pa} = 73.4$ dB.

3.3. Procedure

The subjects walked to the warning block located at a distance of 2.4 m from the starting point, where they were guided by the proposed method. The results obtained using the new approach were compared to those obtained using the tactile map method.

In the experiment, the subjects were instructed to go to the directed destination depending on the presented map information. First, the subjects walked on the leading block from the starting point to the warning block. They could not access the map before reaching the warning block. Second, the subjects stopped there, and a stereophonic sound was generated by a parametric speaker. After obtaining the map information, the subjects were asked to head in the direction of the target facility.

The destination point was located either forward, backward, right, or left relative to the direction of the travel. The subjects were required to respond as soon as they could after listening to the sounds.

Next, an experiment using a tactile map (Figure 8) was carried out in a similar manner. A tactile map was placed near the warning block, and then the subjects were asked to read the tactile map and head in the direction of the target facility. The visually impaired performed the operations of the proposed method and the tactile map twice. In the aforementioned experiments, the sighted people were blinded by wearing an eye mask. The presented map information was different from the actual map information, it was not possible to know the answer. After each measurement, an interview questionnaire was conducted for the visually impaired. The contents of the questionnaire were as follows.

- Impression, feeling, and suggestions for improving the proposed system and the tactile map.

3.4. Subjects

The subjects were 35 people (27 males and 8 females) aged from 19 to 42 years old. Among the subjects, there were 18 people with clear vision (13 males and 4 females) and 17 people with visual impairment (14 males and 4 females). The subjects understand how to use braille blocks. All visually impaired subjects have a disability certificate and can read Braille. Also, all sighted subjects cannot read Braille.

3.5. Evaluate Function

The time taken by the subjects to understand the direction and the direction of the travel (front and back, right and left) was recorded and analyzed. The time taken to grasp the direction is given as the time from the moment when the sound starts to emanate from the parametric speaker to the moment when the subject actually starts walking in some direction. To determine if there was a significant difference between the experimental results, Student's t-test was performed. The significance level was $p < 0.01$. Similarly, for a tactile map, the time is measured from the moment when the tactile map is touched. Additionally, the hearing questionnaire after the experiment was used as a reference for the evaluation of the system.

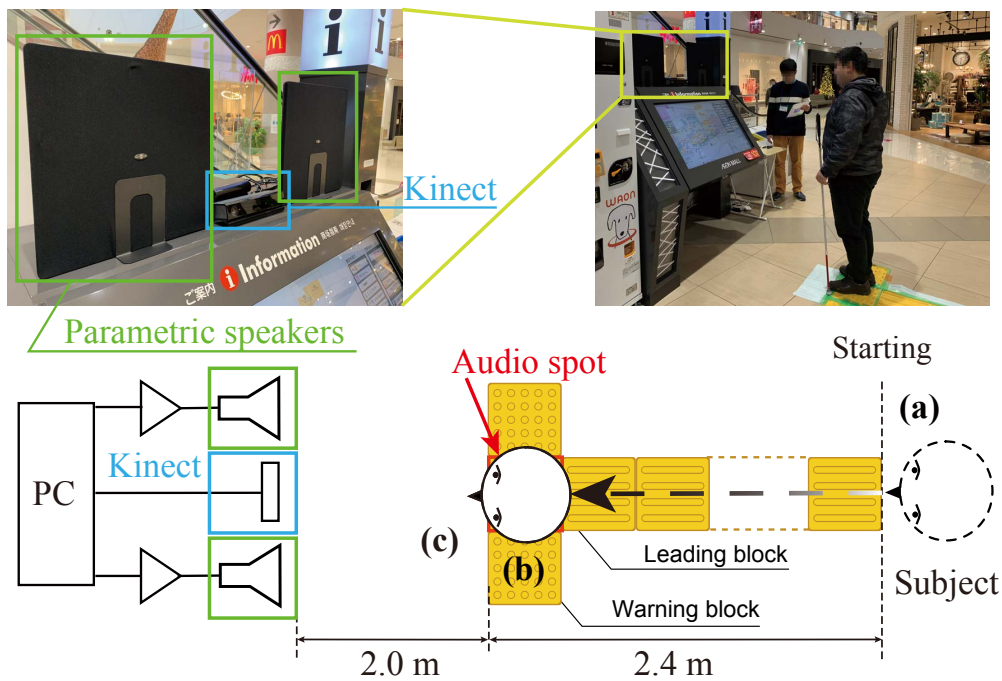


Figure 7. Experiment overview: (a) Subject walks from the beginning of the braille block. (b) When the subject reaches the warning block, a stereophonic sound is generated by a parametric speaker. (c) After hearing the sound, the subject moves in the direction of the target facility indicated by the earcon.

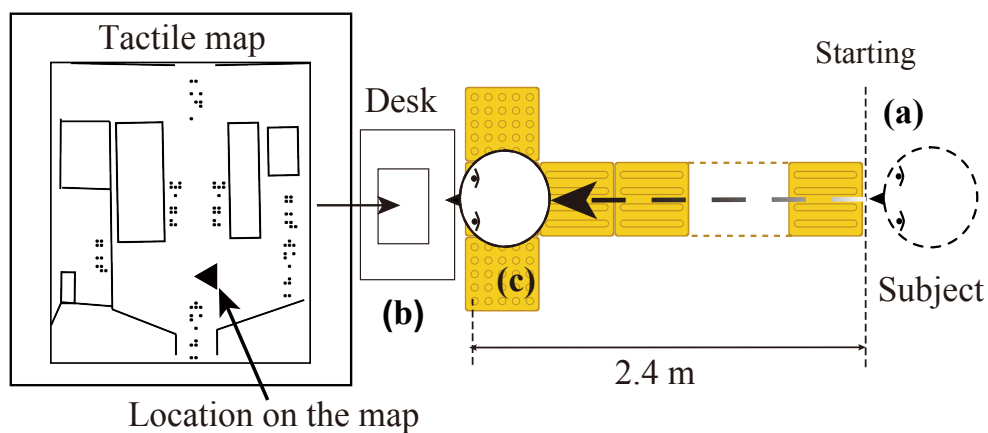


Figure 8. Experiment overview: (a) Subject walks from the beginning of the braille block. (b) The tactile map on the desk is read. (c) Subject moves in the direction of the target facility.

4. Results

4.1. Elapsed Time

Figure 9 shows the times taken to grasp the direction for the sighted and the visually impaired subjects. First, the results obtained for the sighted subjects are described. The average time taken to grasp the direction decreased, but there was no statistically significant difference. Next, the results for the visually impaired, there was no significant difference between the average time using the proposed method and the tactile map. Large differences were observed in the results and population maps, and the sighted subjects required much more time for using the tactile map than the visually impaired subjects. By contrast, for the proposed method, there was no significant difference between the time required for the sighted and the visually impaired subjects.

4.2. Accuracy

Figure 10 shows the percentages of the correct answers for the sighted and visually impaired subjects. First, the results for the sighted subjects are described. The accuracy of the proposed method was higher than that of the tactile map method, but there was no statistically significant difference. In addition, it was reduced for the sighted subjects compared to that obtained using the tactile map. In the proposed method, there was no significant difference between the results obtained with or without a blindfold. The accuracy of the proposed method with the blindfold was as low as that of the tactile map.

Next, the results for the visually impaired, the accuracy using the tactile map was 100%. Thus, the accuracy of the proposed method was lower than that of the tactile map, but there was no statistically significant difference. By comparing the results of the sighted and visually impaired subjects, it was observed that the accuracy of the sighted subjects without blindness was almost the same as that of the visually impaired subjects.

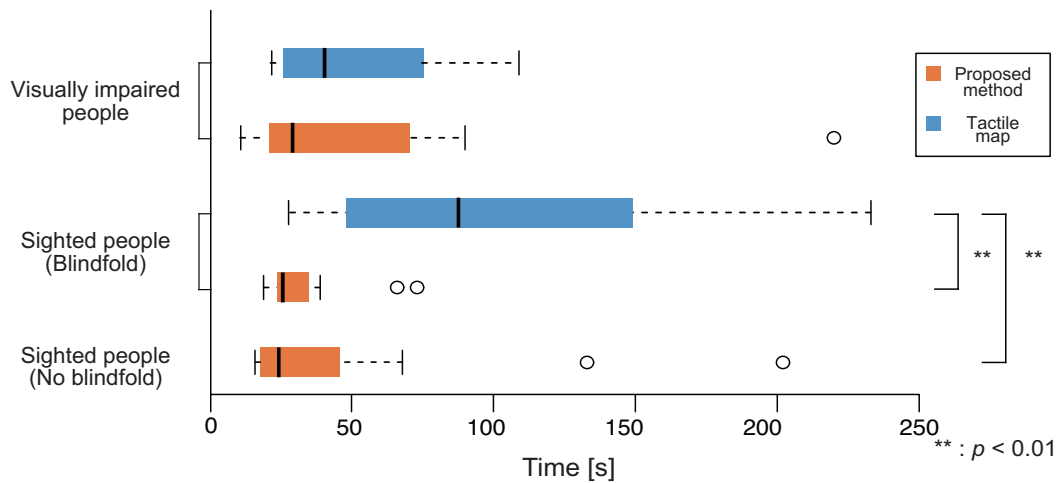


Figure 9. Time taken to determine the direction: 1 Sighted subjects-The average time using the proposed method was shorter than that for the tactile map. The proposed method took less time with the second blindfold. 2 Visually impaired subjects-There was no significant difference in the time between the tactile map and the proposed method.

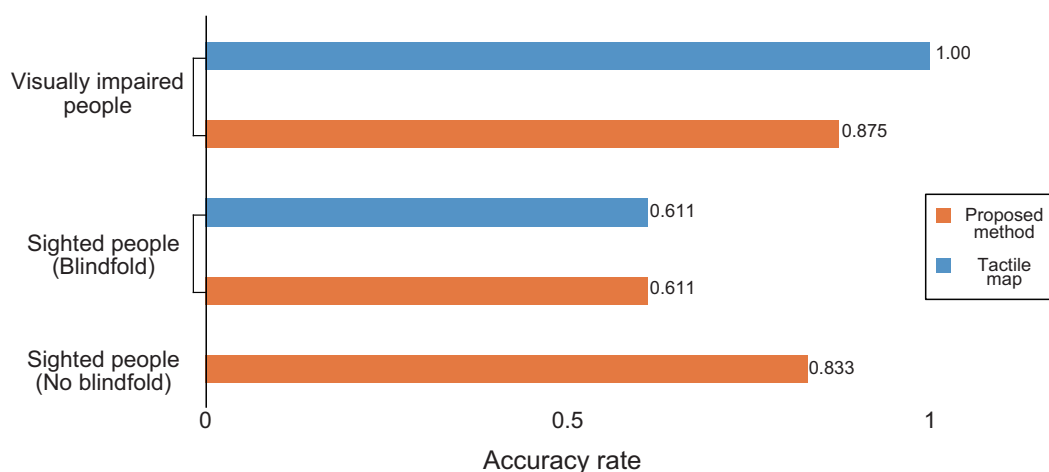


Figure 10. Accuracy rate: In the proposed method, the accuracy was almost the same between those without a blindfold and those with visual impairment. For the tactile map, the accuracy of the visually impaired subjects was 100%, which was different from that of the sighted subjects.

5. Discussion

The accuracy of the proposed method for the sighted subjects did not differ from that of the tactile map. Then, in the proposed method, an experiment was carried out without blindfolding a sighted person. As a result, the percentage of the correct answers for the sighted subjects without a blindfold increased significantly (Figure 10). This is attributed to the fact that the sighted subjects performed the task after the visual information was not available, unlike in the usual case. In the previous experiment, a sighted person walked to the warning block blindfolded. Compared with the visually impaired, the sighted are not used for walking relying on the braille block. As a result, they could not walk straight, and they heard the presentation sound at a location slightly away from the audio spot. Therefore, we considered that the stereophonic sense of the stereophonic sound was impaired, and the accuracy rate deteriorated.

Based on our previous study, we investigated the extent to which dislodgement from the front of the speaker impaired the orientation.

We experimentally confirmed that auditory localization can be maintained when moving 0.4 m in parallel from 2.0 m in front of a parametric speaker (Figure 11). The localization angle of the sound image was randomly changed, and the sound image was presented and answered to each of three subjects 40 times for each sound source. As a result, the sense of localization was impaired instead of hearing sounds in front of the parametric speaker (Figure 12). Based on this experiment, we considered that the correct answer rate deteriorated in the experiments owing to a loss of orientation sensation due to the displacements in the opposite directions. Moreover, we confirmed the following from the experimental results.

Moreover, we confirmed the following from the experimental results.

1. The elapsed time was reduced using the proposed method for the sighted, which was approximately the same as that of the visually impaired.
2. There was no significant difference in the accuracy rate between the visually impaired and the sighted without blindfolding for the proposed method.
3. There was no significant difference in the accuracy rate between the proposed method and the touch map method for the visually impaired.

From 1 and 2, it is suggested that the proposed method is useful for people who cannot read tactile maps or Braille (less than 10% of visually impaired people can read Braille [3]).

From 3, it is suggested that the proposed method can be used as a substitute for tactile maps even for those who can read Braille. Thus, it was confirmed that the peripheral map information

could be presented to the recipient regardless of the existence of the visual impairment in a noisy environment (the average of A-weighted Sound Pressure Level: $L_{pa} = 73.4$ dB). In particular, it can be said that the proposed method is effective for the visually impaired who cannot read Braille.

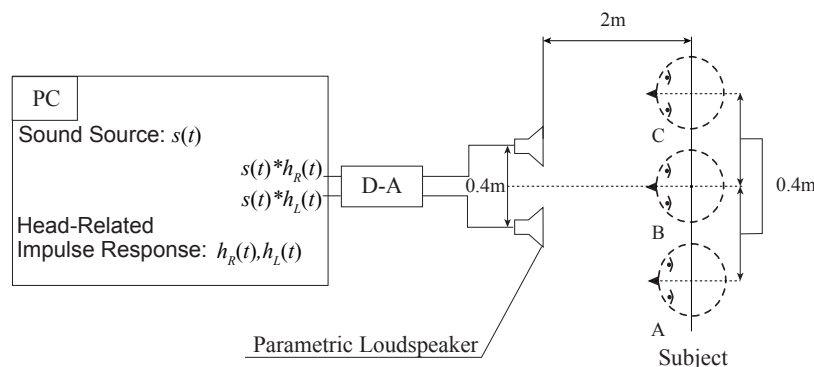


Figure 11. Experimental process: Subjects listened to a stereophonic sound at A, B and C. We then asked from which direction the sound was heard.

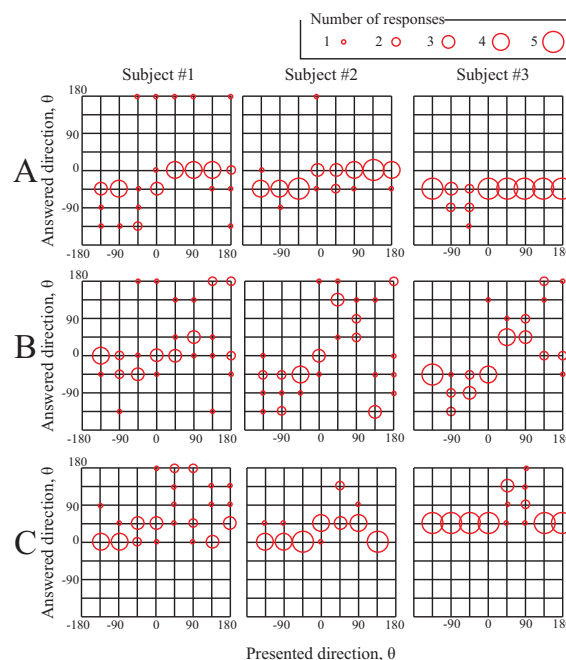


Figure 12. Results: The subjects listened to a stereophonic sound and answered from which direction the sound was heard. A, B and C are the same as in Figure 11.

From the descriptive questionnaire, about the earcons

- I think it would be better to make the icons easier to remember.
- It is difficult to understand because it sounds on the same scale.

Almost all of the visually impaired subjects said that the earcons were difficult to understand. Thus far, all earcons are generated at the same scale, making it difficult to understand the result. One of the subjects suggested changing the elevator earcon into a motor sound, using the sound actually emitted by the subject as an earcon. It is necessary to improve the design index of earcons for perception.

6. Future Work and Application

Our proposed system uses parametric speakers to present the location information of the facility to the recipient using stereophonic sound. This system has two main improvements.

The first is a speaker. Currently, the supported person needs to stop at a specific position in order to perceive the sound. When the position is shifted, a sense of localization may not be obtained. Therefore, in the next study, it is necessary to be able to provide the supported person with stereophonic presentation even if he/she is not at the specific position. For that purpose, it is necessary to change the direction of the directivity according to the position of the face of the supported person. It is believed that this can be achieved by controlling the direction of the speaker or by implementing the spatial acoustic technology using an ultrasonic phased array [38]. By utilizing this technology, we considered that the supported person can make a system capable of presenting information in real time even when walking without standing still.

The second is a detection system. Currently, Kinect (RGB-D sensor) is used to detect the face of the supported recipient and measure the distance to the speaker. We considered that by combining this with image recognition, it is possible to select the target person automatically to perform sound presentation.

It is believed that the improvement in these two points will make it possible to identify the supported person automatically by image recognition in the future, and to present the sound automatically to the walking target. For example, a system is proposed in which only a pedestrian with a white cane is detected and map information is dynamically presented according to the walking position.

7. Conclusions

This study sought to demonstrate that a navigation system using stereophonic sound technology is effective for the support of the visual impaired in a noisy environment. Pinpoint presentation of stereophonic sound consists of a pair of stereo parametric speakers, and by delivering stereophonic sound only to the user, it is possible to present a sound image centered on the user's position regardless of the position of the speaker. The presented sound is an earcon representing the target facility that was created in accordance with the design index. The recipient can intuitively understand the direction of the target facility by listening to the earcon with a stereophonic sound. The sound is presented by a parametric speaker, and the sound is not noisy and is not audible for everyone except for the user.

This system was constructed in a shopping mall, and an experiment was carried out in which the proposed system and guidance by a tactile map lead to a target facility. The subjects included the visually impaired and sighted with a blindfold. The system was evaluated using the time required to grasp the direction and the accuracy rate.

As a result, it was confirmed that the time using the proposed method was reduced for the visually impaired and sighted. It was confirmed that the percentage of the correct answers decreased because the sighted were blindfolded. This is because the sighted subjects deviated from the audible region of the stereophonic sound while walking on a braille block when blindfolded.

On the one hand, the accuracy of the sighted subjects without a blindfold was almost the same as that of the visually impaired subjects. Therefore, the proposed method presented map information regardless of the presence or absence of visual impairment in a noisy environment. Moreover, in the actual environment where this system is supposed to be used, the correct answer rate was over 80%. These results suggest that non-wearable stereophonic presentation for the visually impaired can be used for map presentation and can replace the conventional tactile map, which require touch, and the map presentation which cannot confirm the languages of the user.

However, some of the subjects complained that the sound localization was poor, and it was difficult to distinguish the difference between the earcons. Therefore, the following problems are raised as the future work: Our method must obtain a stable localization sensation of the sound and improve the design index of the earcons in order to make it easy to distinguish under noise.

Author Contributions: Writing—original draft, Y.M. and R.I.; Software, Y.M. and R.I.; Investigation, Y.M., R.I. and H.E.B.S.; Methodology, H.E.B.S. and K.Z.; Writing—review & editing, N.W. and K.M.; Supervision, M.K.; Project administration, K.Z.; All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Tsukuba Collaboration-Boost Project, University of Tsukuba and Tsukuba Society 5.0 Social Implementation Trial Support Project, Tsukuba City.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bourne, R.R.; Flaxman, S.R.; Braithwaite, T.; Cicinelli, M.V.; Das, A.; Jonas, J.B.; Keeffe, J.; Kempen, J.H.; Leasher, J.; Limburg, H.; et al. Magnitude, temporal trends, and projections of the global prevalence of blindness and distance and near vision impairment: A systematic review and meta-analysis. *Lancet Glob. Health* **2017**, *5*, e888–e897. [CrossRef]
2. World Health Organization; World Bank. *World Report on Disability 2011*; World Health Organization: Geneva, Switzerland, 2011.
3. National Federation of the Blind. *The Braille Literacy Crisis in America*; National Federation of the Blind: Baltimore, MD, USA, 2009.
4. Jacobson, R.D. Navigating maps with little or no sight: An audio-tactile approach. In *Content Visualization and Intermedia Representations (CVIR'98)*; University of California: Santa Barbara, CA, USA, 1998.
5. Miele, J.A.; Landau, S.; Gilden, D. Talking TMAP: Automated generation of audio-tactile maps using Smith-Kettlewell's TMAP software. *Br. J. Vis. Impair.* **2006**, *24*, 93–100. [CrossRef]
6. Brock, A.M.; Truillet, P.; Oriola, B.; Picard, D.; Jouffrais, C. Interactivity improves usability of geographic maps for visually impaired people. *Hum. Interact.* **2015**, *30*, 156–194. [CrossRef]
7. Yelamarthi, K.; Haas, D.; Nielsen, D.; Mothersell, S. RFID and GPS integrated navigation system for the visually impaired. In Proceedings of the 53rd IEEE International Midwest Symposium on Circuits and Systems, Seattle, WA, USA, 1–4 August 2010; pp. 1149–1152.
8. Na, J. The blind interactive guide system using RFID-based indoor positioning system. In *International Conference on Computers for Handicapped Persons*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 1298–1305.
9. Murad, M.; Rehman, A.; Shah, A.A.; Ullah, S.; Fahad, M.; Yahya, K.M. RFAIDE—An RFID based navigation and object recognition assistant for visually impaired people. In Proceedings of the 2011 7th International Conference on Emerging Technologies, Islamabad, Pakistan, 5–6 September 2011; pp. 1–4.
10. Tsirmpas, C.; Rompas, A.; Fokou, O.; Koutsouris, D. An indoor navigation system for visually impaired and elderly people based on Radio Frequency Identification (RFID). *Inf. Sci.* **2015**, *320*, 288–305. [CrossRef]
11. Öktem, R.; Aydin, E. An RFID based indoor tracking method for navigating visually impaired people. *Turk. J. Electr. Eng. Comput. Sci.* **2010**, *18*, 185–198.
12. Moreira, A.J.; Valadas, R.T.; de Oliveira Duarte, A. *Reducing the Effects of Artificial Light Interference in Wireless Infrared Transmission Systems*; IET: London, UK, 1996.
13. Gionata, C.; Francesco, F.; Alessandro, F.; Sabrina, I.; Andrea, M. An inertial and QR code landmarks-based navigation system for impaired wheelchair users. In *Ambient Assisted Living*; Springer: Berlin/Heidelberg, Germany, 2014; pp. 205–214.
14. Chang, Y.J.; Tsai, S.K.; Chang, Y.S.; Wang, T.Y. A novel wayfinding system based on geo-coded QR codes for individuals with cognitive impairments. In Proceedings of the 9th International ACM SIGACCESS Conference on Computers and Accessibility, Tempe, AZ, USA, 15–17 October 2007; pp. 231–232.
15. Kim, J.E.; Bessho, M.; Kobayashi, S.; Koshizuka, N.; Sakamura, K. Navigating visually impaired travelers in a large train station using smartphone and bluetooth low energy. In Proceedings of the 31st Annual ACM Symposium on Applied Computing, Pisa, Italy, 6–8 April 2016; pp. 604–611.
16. Sato, D.; Oh, U.; Naito, K.; Takagi, H.; Kitani, K.; Asakawa, C. NavCog3: An evaluation of a smartphone-based blind indoor navigation assistant with semantic features in a large-scale environment. In Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility, Baltimore, MA, USA, 29 October–1 November 2017; pp. 270–279.
17. Nakajima, M.; Haruyama, S. New indoor navigation system for visually impaired people using visible light communication. *EURASIP J. Wirel. Commun. Netw.* **2013**, *2013*, 37. [CrossRef]
18. International Organization for Standardization. Accessible Design-Auditory Guiding Signals in Public Facilities. (ISO Standard No.19029:2016). 2016. Available online: <https://www.iso.org/standard/63762.html> (accessed on 20 July 2020).

19. Miyachi, T.; Balvig, J.J.; Kisada, W.; Hayakawa, K.; Suzuki, T. A quiet navigation for safe crosswalk by ultrasonic beams. In *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 1049–1057.
20. Miyachi, T.; Balvig, J.J.; Kuroda, I.; Suzuki, T. Structuring spatial knowledge and fail-safe expression for directive voice navigation. In *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 664–672.
21. Aoki, S.; Toba, M.; Tsujita, N. Sound localization of stereo reproduction with parametric loudspeakers. *Appl. Acoust.* **2012**, *73*, 1289–1295. [[CrossRef](#)]
22. Gaver, W.W. Auditory icons: Using sound in computer interfaces. *Hum.-Comput. Interact.* **1986**, *2*, 167–177. [[CrossRef](#)]
23. Gaver, W.W. Synthesizing auditory icons. In Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems, Amsterdam, The Netherlands, 24–29 April 1993; pp. 228–235.
24. Gaver, W.W. What in the world do we hear?: An ecological approach to auditory event perception. *Ecol. Psychol.* **1993**, *5*, 1–29. [[CrossRef](#)]
25. Gaver, W.W. The SonicFinder: An interface that uses auditory icons. *Hum. Interact.* **1989**, *4*, 67–94. [[CrossRef](#)]
26. Blattner, M.M.; Sumikawa, D.A.; Greenberg, R.M. Earcons and icons: Their structure and common design principles. *Hum. Interact.* **1989**, *4*, 11–44. [[CrossRef](#)]
27. Brewster, S.; Raty, V.P.; Kortekangas, A. Earcons as a method of providing navigational cues in a menu hierarchy. In *People and Computers XI*; Springer: Berlin/Heidelberg, Germany, 1996; pp. 169–183.
28. Karshmer, A.I.; Brawner, P.; Reiswig, G. An experimental sound-based hierarchical menu navigation system for visually handicapped use of graphical user interfaces. In Proceedings of the First Annual ACM Conference on Assistive Technologies, Marina Del Rey, CA, USA, 31 October–3 November 1994, pp. 123–128.
29. Leplâtre, G.; Brewster, S.A. Designing non-speech sounds to support navigation in mobile phone menus. In Proceedings of the International Conference on Auditory Display, Atlanta, GA, USA, 2–5 April 2000, pp. 190–199.
30. Bustoni, I.A.; Hidayatulloh, I.; Azhari, S.; Augoestin, N.G. Multidimensional Earcon Interaction Design for The Blind: A Proposal and Evaluation. In Proceedings of the 2018 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), Yogyakarta, Indonesia, 21–22 November 2018; pp. 384–388.
31. Willert, V.; Eggert, J.; Adamy, J.; Stahl, R.; Korner, E. A probabilistic model for binaural sound localization. *IEEE Trans. Syst. Man, Cybern. Part B (Cybern.)* **2006**, *36*, 982–994. [[CrossRef](#)]
32. Perrott, D.R.; Musicant, A. Minimum auditory movement angle: Binaural localization of moving sound sources. *J. Acoust. Soc. Am.* **1977**, *62*, 1463–1466. [[CrossRef](#)] [[PubMed](#)]
33. Garzonis, S.; Jones, S.; Jay, T.; O'Neill, E. Auditory icon and earcon mobile service notifications: Intuitiveness, learnability, memorability and preference. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Boston, MA, USA, 4–9 April 2009; pp. 1513–1522.
34. Brewster, S.A. Using nonspeech sounds to provide navigation cues. *ACM Trans. Comput. -Hum. Interact. (TOCHI)* **1998**, *5*, 224–259. [[CrossRef](#)]
35. Brewster, S.A.; Wright, P.C.; Edwards, A.D. Experimentally derived guidelines for the creation of earcons. *Adjun. Proc. HCI* **1995**, *95*, 155–159.
36. Brewster, S.A.; Wright, P.C.; Edwards, A.D. An evaluation of earcons for use in auditory human-computer interfaces. In Proceedings of the INTERACT'93 and CHI'93 Conference on Human Factors in Computing Systems, Amsterdam, The Netherlands, 24–29 April 1993; pp. 222–227.
37. Roberts, L. Consonance judgements of musical chords by musicians and untrained listeners. *Acta Acust. United Acust.* **1986**, *62*, 163–171.
38. Ochiai, Y.; Hoshi, T.; Suzuki, I. Holographic whisper: Rendering audible sound spots in three-dimensional space by focusing ultrasonic waves. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 6–11 May 2017; pp. 4314–4325.

