Virtual Stage: an Immersive Musical Game for People with Visual Impairment

Patricia A. de Oliveira, Erich P. Lotto, Ana Grasielle D. Correa, Luis G. G. Taboada, Laisa C. P. Costa, Roseli D. Lopes
Laboratory of Integrated Systems - LSI - Department of Electronic Systems Engineering
Polytechnic School, University of São Paulo
São Paulo, SP, Brazil
Email: {paaraujo, erich, anagrasi, ltaboada, laisa, roseli}@lsi.usp.br

***Abstract*—Musical games help in motor and cognitive development and provide pleasure because of their playfulness and challenge. However, visually impaired people are unable to interact with conventional music games, because audiovisual aspects are inaccessible to this public. The objective of this study was to investigate if three-dimensional sounds could guide blind users during interaction with a tablet music game. As a proof of concept, a musical game for Android was developed with binaural audio techniques. Three fully blind users evaluated the game. The results show that the developed game was able to immerse the users in a virtual environment.**

***Keywords*-Visual Impairment; Audiogame; Three-dimensional Sounds; Virtual Reality**

I. Introduction

In the last decade, the rapid integration of mobile devices into society, particularly tablets and smartphones, enabled the growth of innovation opportunities in the entertainment industry. Such devices are no longer simple tools of communication and access to information, but also entertainment platforms, providing various types of games and multimedia content. Many current devices have high processing power and feature various types of sensors such as accelerometers, gyroscopes, barometers, compasses, touches and gestures on the screen. Thus, new opportunities for interaction in applications developed for mobile devices attract the interest of people of all age groups. For this reason, games for tablets are becoming increasingly popular due to portability and the high degree of interaction provided.

Over the years, it is possible to observe significant changes in games: graphics, stories, soundtrack and increasingly complex gameplay [1]. This set of possibilities provides the necessary incentive to a satisfying user's gaming experience. The complexity of the current games requires a greater reflective and cognitive effort, motor responses and faster reasoning to enable a more meaningful level of interaction between player and game [2]. However, the feedback traditionally used in games is difficult or impossible to be perceived by people with visual impairment. For this reason, [1] listed the main problems of accessibility in video games for the blind: (1) difficulty in using the interaction devices; (2) difficulty in receiving feedback (3); difficulty in identifying the set of answers.

Strategies to replace the visual interface components, allowing visually impaired players to interact with video games, can be employed to minimize this problem. Many studies have reported the use of audio to provide feedback to blind players [3] [4]. This game mode, called audio game, is a game access tool in digital media; therefore, with information transmitted by sounds, these players can experience and understand the universe in which the game scenario is inserted.

Audio in games is an important asset, because they reproduce sounds almost identical to those naturally heard, improving the individual immersion sensation. 3D audio techniques allow perceiving the loudness, distance and depth of sound, allowing locating an object in space just by listening to the sound. Only with a sound, for example, the user can identify the exact position of the object in space.

The fundamental premise of 3D audio is to play sound effects identical to those naturally perceived by the human ear through headphones or speakers. It is possible to determine the point of origin of sounds. The audio reproduction technique used in this study is similar to the binaural audio technique, with which the user can locate sound sources using stereo headphones. The technique enables an immersive listening experience, emulating the way the sound should be heard in real life.

Our aim is to explore the use of an audio technique, inspired by the binaural audio technology, in a music game for tablets. In this paper, we investigated whether a blind person is able to immerse, interact and engage with a virtual environment using only audio. The interaction design followed accessibility guidelines, using gestures on the touch screen and the tablet gyroscope as inputs. This work designed sensory feedback targeting the visually impaired, so that the game experience can be inclusive and satisfying.

II. Concepts and Fundamentals

This section introduces the concepts that guided the development of this work: the types of visual impairment, audio games and binaural audio guidelines for accessible games development.

A. *Types of Visual Impairment*

Visual impairment is defined as vision loss, total or partial, congenital or acquired. The level of visual acuity may vary

2159-6662/16 $31.00 © 2016 IEEE
DOI 10.1109/SBGames.2015.26

depending on the type of disability [5]:

- Blindness: when there is total loss of vision or little ability to see, causing the person to need assistive technology to read and write.
- Low vision: characterized by impaired visual function of the eyes, even after treatment or correction. People with low vision can read texts with enlarged letters.

Visual impairment is the greatest impact on the world's population. Currently, there are estimated to be 180 million visually impaired people worldwide, of which 45 million are blind and 135 million have low vision [6]. Most cases of blindness occur in underdeveloped or developing countries. In Brazil, where this research was developed, more than 6.5 million people have some visual impairment. Of this total, more than 582,000 people have total blindness and just over 6 million have low vision [6].

B. Guidelines for Accessible Games Development

Accessibility refers to equal opportunities in any fields of society: physical spaces, care for people with disabilities, interpersonal communication, use of tools, rules and regulations and Internet access [7]. Accessibility in games can be defined as the ability to play a game even with a given impairment [8]. An important factor in accessible games is their interaction mechanism: the interface (visual and audio elements) and input and output devices.

Game interactions consist of game events, user action commands and the game responses to these actions. User commands use input devices such as keyboard, mouse, joystick, microphone or movements. The game events and responses are provided by sensory stimulations: vision, hearing, touch and smell. In tablets, the input mechanisms can be gestures on the touch screen, device movimentation, text input through a virtual keyboard, virtual buttons or audio input. Responses can be visual, audible or tactile (vibration).

The game feedback is very important to the user with visual impairment. Design decisions must be made to assist impaired players with three challenges: difficulty in using the interaction devices, difficulty in receiving feedback and difficulty in identifying the set of answers [1].

Through the analysis of some accessible games, we identified as the main feedbacks used: a) Text-to-Speech (TTS) b) 3D audio; c) audio metaphor; d) tactile feedback. In addition, different options are recommended to change/enhance the contrast between background and text/font size adjustment.

The Text-to-Speech (a) read the interface elements, either using predetermined recording or by software generating speech from text. 3D audio (b) is a spatial audio reproduction technique that can be used to aid in the identification the position of sound sources, locating game elements, serving as orientation for the user in the scenario (eg. obstacles or items, enemies or allies, near and distant objects, among others). Audio metaphor (c) is an intuitive audio feedback that helps to perceive game events (eg. open or close a door, walk, drop, etc.), as well as orienting the user in the scenario (eg. forest with owl sounds, river sound, sound of wind in the trees). Tactile feedback (d) provides response as to interface elements (eg. vibrate when the user touches or swipes the screen) as user actions (eg. vibrate when bumping into an obstacle).

Sounds indicating the result of an interaction are important. Accessible games must standardize sounds to indicate the result of actions, such as sounds of success or failure, sounds to indicate opening doors, collections of items, enemies defeated, items used, etc. The sound outputs may still have variations to aggregate more information (intensity change, volume, depth, repetition and speed). All these points should be planned before the game development.

C. Audio Games

The Merrian-Webster dictionary [7] defines video as the moving images that are seen in a recording or broadcast and video game as an electronic game in which players control images on a television or computer screen. The audio outputs are usually present enriching the playing experience, however visual outputs are the major interface.

Audio outputs as music and sound effects are mainly responsible for increasing players' immersion in the story and scenes of the game. The sound effects are also known to be capable of generating emotional feelings more than any other event, for example, fast sounds are associated with action activities and slow sounds with passivity and stillness. Moreover, unlike the visual information, aural information cannot be easily ignored; therefore, it is an important source of information [9].

Despite its relevance, audio is still used only as a complement in video games. Nowadays, it is not possible to play a game without the video on whereas it is frequently possible to play it without the audio. Thus, as alternative to video games, audio games have emerged as games for visually impaired people.

Audio games are those that do not depend on visual outputs because they use audio to represent all game events and feedbacks.

There are initiatives to provide popular games in accessible format using sound features and textual information. For example, the website audiogames.net offers about 195 audio games. However, many games, in this website, cannot be downloaded, some games are no longer available and some have functional problems to be executed. In Brazil, the initiatives emerge from the same idea, as the website audiogames.com.br provides 34 other games based in the audiogame.net or developed exclusively for it.

D. Binaural Audio

Binaural audio is a type of 3D audio[10]. The binaural audio characteristic is to help to identify and to distinguish

Figure 1. Mannequin of a human head used for binaural recording [12]

the number and location of sound sources in a given environment. If recorded correctly, the binaural audio faithfully reproduces any type of environment; the listener can practically perfectly hear the sound as recorded [11]. It is, therefore, a feature that provides notions of distance and location in space, both horizontally and vertically.

[11] explains that binaural is a method of recording multidimensional sound fields using microphones arranged one in each ear canal of an artificial head, making it possible to "hear" as the artificial head does. This recording enables an imaginary translocation in space due to the way the sound is formed in a person's head.

To create a binaural recording, a mannequin of a human head (Figure 1) is used. High sensitivity microphones are accurately arranged in the inner ears, to seamlessly mimicking human beings ears, with the same angles and distances in between them. This balance makes it possible to recreate the human hearing system for the recording. This method allows the listener to spatially identify the location source of the sound.

In this kind of recording, the ear captures an audio source position in the environment: the sound direction and distance to the ears. As there is a head between the two ears, it is possible to discern whence the sound came through the slight difference between the capture of one ear to another.

Binaural audio systems have been used in a wide range of applications with different requirements and rely on factors such as the realism desired, the necessary precision, the cost and the ease of use. Several researches have confirmed the usefulness of binaural audio as an interface for a person with visual impairment. For instance, many studies conclude that the user performance in activities as the navigation by a Global Positioning System (GPS) increases by using binaural audio as a user interface [12] [13] [14] [4]. Besides, several studies have suggested using 3D audio as user interface for the visually impaired using web browsers [15]. Moreover, several prototypes of sensory substitution systems and augmented reality aimed at the visually impaired proposed 3D audio interface [12] [14].

III. 3D Audio Technique

This work implemented a technique based on binaural audio similar to the one used in a few games such as Top Speed and Defender [1], where the sound gets louder on one of the audio channels, indicating that there is an interaction on that side, in order to create a virtual environment that enables the user to identify the spatial location of a sound source. The sound source position defines the panning of the sound, i.e., the volume of each of the auditory channels is modified with the position of the player avatar in the virtual environment. The sound volume in each auditory channels is controlled by using a stereo headset.

Each sound source has a correction factor applied to the volume in each of the stereo channels. This correction factor is defined in the following way: $Sb = (A - B)/180$. Where Sb is the sound balance, A is the angle where the listener is looking and B is the angle between A and the instrument. That returns a value between -1 and 1, indicating the sound balance (-1 = all left and 1 = all right) that is then divided to the distance between the player and the instrument, calculating the volume according to the position of the sound source as regards the player's avatar position. Both distance and the angular distance are considered.

Figure 2 and Figure 3 illustrates the technique used: five sound sources are represented by S1, S2, S3, S4 and S5; each sound source receives a pair of numbers that is the correction factor in percentage to each of the audio channels (left / right). In the example presented in Figure 2, the player's avatar is positioned in front of Source 1; his/her right ear will receive: 50% of Source 1 audio volume, 75% of that of Source 2, 0% of that of Source 3, 50% of that of Source 4 and 84% of that of Source 5. While his/her left ear will receive: 50% of that of Source 1 audio volume, 25% of that of Source 2, 100% of that of Source 3, 50% of that of Source 4 and 16% of that of Source 5.

If the player's avatar changes its position (rotating to the left) in the virtual environment, the audio sources will change their respective positions. Figure 3 presents a new configuration, where the correction factors are changed. The player's right ear will receive: 75% of Source 1 audio volume, 100% of that of Source 2, 25% of that of Source 3, 25% of that of Source 4 and 58% of that of Source 5. While his/her left ear will receive: 25% of that of Source 1, 0% of that of Source 2, 75% of that of Source 3, 75% of that of Source 4 and 42% of that of Source 5.

IV. Virtual Stage Game Development

The game named "Virtual Stage", was developed in Java programming language and was projected to run on devices with Android operating system.

[1]Both games are available for download at http://www.audiogames.com.br

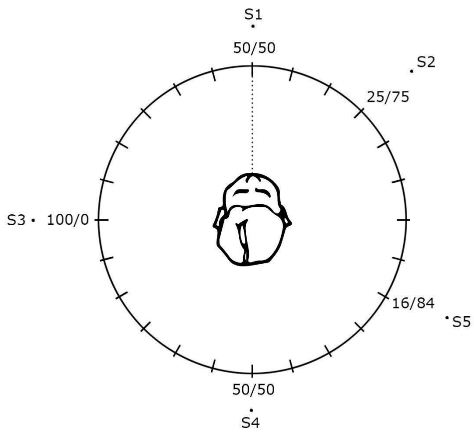

Figure 2. Audio spatialization technique scheme positioned in front

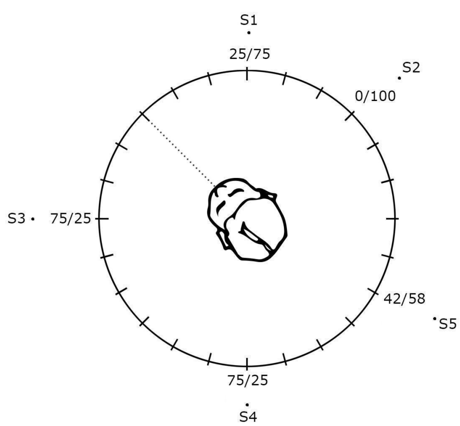

Figure 3. Audio spatialization technique scheme new configuration

"Virtual Stage" consist in virtual musical band, which can be formed by various instruments and vocals. To create a scenario for the game each musical instrument, and the vocal composing the music, must be in a separated track. Once the game starts, the audio tracks are played simultaneously and synchronously, mixed back so as to form the music.

Each instrument track should be stored in a specific file. Each of these files will be played with the volume correction factor dynamically defined according to the movement of the avatar in the virtual scenario. The correction factor is different for the left and for the right channel, according to which direction the sound is coming from.

The player has an avatar at the virtual stage and is challenged to find an instrument on the stage, walking towards it. Both the position of the instruments and the challenge instrument selection are randomly defined when the game starts. Figure 4 illustrates the game operation example in accordance with the technique applied.

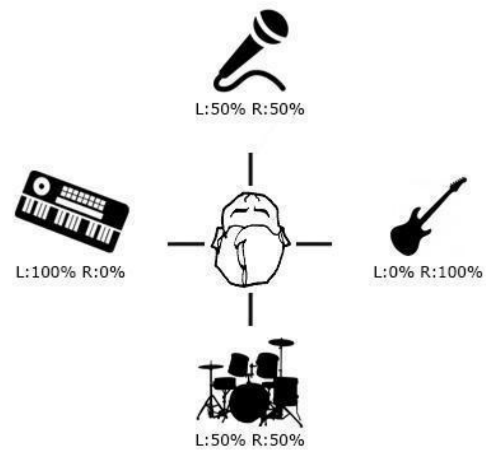

Figure 4. Example user position with four sound sources

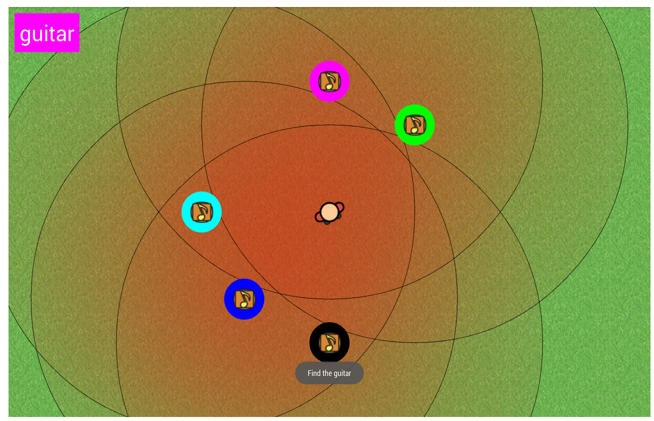

Figure 5. Game screenshot

A. Validation scene

The structure of the game code has been developed so that the developer can set different scenarios. For this paper, it was created a scenario with four musical instruments, plus vocal, positioned at fixed points in the scenario (sound sources).

Figure 5 presents the scenarios used for validation in which the avatar is in the center of the screen and the challenge is to locate the guitar on the stage.

The song used was a rock song. The original song consists of an OGG file size 34.9 MB. This file provides a single slot for each musical instrument composing the music. These tracks were extracted using the Audacity program, resulting in five audio files in MP3 format, representing the vocal and the instruments: guitar, drums, bass and keyboard.

In order to locate the requested instrument, the user must walk towards the instrument following the sound of this specific instrument. There are two ways to command the avatar movements: moving the tablet or using gestures.

The idea of moving the avatar by moving the tablet is that the player should hold the tablet in a fixed position close to his/her body and perform rotational movements with his/her body in the real environment. The angular velocity of the movement performed is detected through the device's gyroscope. With this measure, the instruments are re-positioned considering the player's avatar.

The avatar walks ahead when the player touches the screen with a finger. While the finger is touching the screen, the avatar moves forward.

The position (angle and distance) between the avatar and the instruments are calculated and correction factors are applied to the volume of each instrument in order to mix the right and left audio channels. Through these interaction mechanisms, the game provides the information necessary to move the avatar in the scene.

V. Game's Validation

A questionnaire was developed with subjective questions to assess the players' satisfaction when playing the game. Two questions were prepared with answers varying according to the Likert scale and one open question for the general review of the participant. The questions were:

I. Did you enjoy the experience you have just performed?
[Enjoyed Very Much/Enjoyed/Indifferent/ I didn't Like it/ I Hate it]
II. To find the instrument was:
[Very Easy / Easy / Indifferent / Hard / Very Hard]

Observing success or failure of the challenge to meet the requested instrument, information about the game usability was collected. The notes during the challenge were:

I. The user found the instrument:
[In the first attempt/ In the second attempt/ In the third attempt/ In the fourth attempt/ In the fifth attempt/ Didn't find it]
II. In which way and direction the user has made his/her first move:
[to the target / Next to the target / Opposite the target]
III. Do you have any suggestions or criticism about the experience you have just performed?

A. Validation by Blind Users

The game was designed for users with any degree of visual impairment (blindness or low vision) because it is believed that the immersion through spatialized sound is capable of assisting the users in playing the game without full vision information.

However, blind users were invited for this validation phase, as preliminary tests of this first version of the game. The purpose of this phase is to know the degree of immersion caused by spatialized sound with blind users.

The game was used by three blind users. A 27 year-old male and two females of 30 and 50 years-old respectively.

As the difficulty of using the device can influence the ease of the use of the application, the participants were asked how often they used touchscreen devices in order to know the level of familiarity with this type of device. We obtained from all participants, the answer "I use it more than once a day" meaning that the frequency use of touchscreen devices was high.

The answers of the satisfaction questionnaire about interaction are presented in Table I.

Table I
User satisfaction evaluation

Participant	Easiness	Experience	Commentaries
1	Very easy	Enjoy it	On a scale from 0 to 10, the user would grade 5, since he/she had difficulty in moving in the real environment.
2	Easy	Enjoy it	Enjoyed it, but the movement was confusing.
3	Easy	Enjoy it a lot	Enjoyed it a lot, but it was hard to interact with the device.

Table II
User actions, as noted by the spectator

Participant	Indicator	User action
1	Progress	Found the instrument on the first attempt
	Localization	His first move was towards the target
2	Progress	Found the instrument in the second attempt
	Localization	His first move was on the opposite direction from the target
3	Progress	Found the instrument in the fourth attempt
	Localization	His first move was next to the target

Table II indicates each participant's action while playing, as a observer took notes according to the questionnaire presented.

VI. Results and Discussions

Considering the test results with blind participants, we observed important aspects regarding the user interaction with the Virtual Stage game. The game interaction requires the player to rotate the entire body together with the tablet. This way of interaction, coupled with aural feedback from the sound sources has caused movement difficulties in the real environment, as pointed out by participant 1. The reason for this is that the user must turn round his/her own axis to hear the sounds from the sources and the 3D audio technique developed makes the main sense used by the blind person, hearing, be isolated from the real world. It causes an orientation loss, generating fear to move in the environment without the notion of the real space around.

Another important observation is that, during the interaction with the game, the participants had to move a lot to find the sound sources. Thus, participant 2 lost his/her sense of direction in the real world. That is, at the end of the game, the participant had no reference of his/her current direction, which led to the claimed “confusion”.

Both results, although described as discomfort, are interesting because they show that the total immersion of a blind user is possible in a virtual environment since their senses were fully involved in the surroundings, causing them to loose track of the real world. Tests with virtual reality games demonstrate that blocking the sight and hearing are key factors in the process of immersion in games [White et al., 2008]. It happens with users experiencing head mounted displays equipped with headphones and motion sensors to sync the in game camera with a head movement. This system plays very well with first-person games, such as Call of Duty, Medal of Honor and Battlefield. The headphones even more increase the sense of immersion, since the audio used in these games is very captivating.

As for the “difficulty in interacting with the device”, described by participant 3, it can be remedied with a better explanation about the input mechanism, because the user have a high frequency of touchscreen device use, the ability of interaction differs for each user. This difficulty and the reason for the discomfort of the participants were observed to be related to the fact that they are not used in the prototype, especially with the complete isolation of their hearing.

All the users managed to address the task successfully, but the game does not distinguish when the sound source is in front or behind the user. For this reason, one of the users was able to identify the target angular position, but walked away from it.

VII. Conclusion

The results from these preliminary experiments are promising. We will improve the prototype considering the findings in this work and will proceed new experiments with visually impaired users. The main modifications will focus on:

1) Input mechanisms: usage of a device with gyroscope attached to the head to improve the user’s motion capture, working with his/her head’s own axis, making this interaction as natural as possible. Besides, the movement of walking in the virtual environment will be reconsidered.
2) Output mechanisms: Besides the spacialized audio and the talkback functionality, it is necessary to use other audio feedbacks, such as the sound of footsteps to indicate the movement in the environment, for example.
3) Universal Design: Following the principles of universal design, an immersive environment will be created not only through audio, but also using visual immersion, with a virtual cave or head mounted displays, to provide the gaming experience to any user.

This is the first step towards achieving the main goal of helping people with visual disabilities regarding social inclusion, evaluating whether an approach focusing on participation through communication and entertainment of visually impaired with other user groups with different characteristics is possible in order to make these types of approaches not mutually exclusive.

Being able to play a game is a matter of quality of life that, especially to a younger audience, either with or without a disability, may affect their psychological welfare.

Regarding the musical work with the visually impaired, mainly with the blind, it does not differ much from those with sighted individuals. It encompass auditive perception and music making, contextualized inside a musicalization process. People participate in a wide and enriching musical experience with the use of vocal, instrumental and movement expressions. Digital musical games may favor the development of analysis and perception of the constructive elements of music. All these opportunities must be inclusive.

References

[1] B. Yuan, E. Folmer, and F. C. Harris Jr, “Game accessibility: a survey,” *Universal Access in the Information Society*, vol. 10, no. 1, pp. 81–100, 2011.

[2] C. S. Bruggers, R. A. Altizer, R. R. Kessler, C. B. Caldwell, K. Coppersmith, L. Warner, B. Davies, W. Paterson, J. Wilcken, T. A. DAmbrosio *et al.*, “Patient-empowerment interactive technologies,” *Science translational medicine*, vol. 4, no. 152, pp. 152ps16–152ps16, 2012.

[3] E. E. Cunha, “Jogavox: Ferramenta e estratégias para construção de jogos educacionais para deficientes visuais,” 2007.

[4] A. Rodríguez, J. J. Yebes, P. F. Alcantarilla, L. M. Bergasa, J. Almazán, and A. Cela, “Assisting the visually impaired: obstacle detection and warning system by acoustic feedback,” *Sensors*, vol. 12, no. 12, pp. 17 476–17 496, 2012.

[5] W. H. Organization *et al.*, “International classification of impairments, disabilities, and handicaps: a manual of classification relating to the consequences of disease, published in accordance with resolution wha29. 35 of the twenty-ninth world health assembly, may 1976,” 1980.

[6] ——, “Strategy for cancer prevention and control in the eastern mediterranean region 2009-2013,” 2010.

[7] Merriam-Webster, *Merriam-Webster’s Collegiate Dictionary: Eleventh Edition*, ser. Logos Bible Software. Merriam-Webster, Incorporated, 2004. [Online]. Available: https://books.google.es/books?id=TAnheeIPcAEC

[8] J. Torrente, A. del Blanco, P. Moreno-Ger, I. Martínez-Ortiz, and B. Fernández-Manjón, "Implementing accessibility in educational videogames with <e-adventure>," in *Proceedings of the First ACM International Workshop on Multimedia Technologies for Distance Learning*, ser. MTDL '09. New York, NY, USA: ACM, 2009, pp. 57–66. [Online]. Available: http://doi.acm.org/10.1145/1631111.1631122

[9] J. R. Parker and J. Heerema, "Audio interaction in computer mediated games," *International Journal of Computer Games Technology*, vol. 2008, p. 1, 2008.

[10] F. Rumsey, "Spatial audio: Binaural challenges," *Journal of the Audio Engineering Society*, vol. 62, no. 11, pp. 798–802, 2014.

[11] H. Møller, "Fundamentals of binaural technology," *Applied acoustics*, vol. 36, no. 3, pp. 171–218, 1992.

[12] E. Carrasco, E. Loyo, O. Otaegui, C. Fösleitner, J. Spiller, D. Patti, R. Olmedo, and M. Dubielzig, "Autonomous navigation based on binaural guidance for people with visual impairment," *Assistive Technology: From Research to Practice. IOS Press*, 2013.

[13] S. Kammoun, M. J.-M. Macé, B. Oriola, and C. Jouffrais, *Towards a geographic information system facilitating navigation of visually impaired users*. Springer, 2012.

[14] B. F. Katz, S. Kammoun, G. Parseihian, O. Gutierrez, A. Brilhault, M. Auvray, P. Truillet, M. Denis, S. Thorpe, and C. Jouffrais, "Navig: augmented reality guidance system for the visually impaired," *Virtual Reality*, vol. 16, no. 4, pp. 253–269, 2012.

[15] G. R. White, G. Fitzpatrick, and G. McAllister, "Toward accessible 3d virtual environments for the blind and visually impaired," in *Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts*. ACM, 2008, pp. 134–141.