JOURNAL OF MULTIMEDIA AND IT
VOL. X NO. 01, PP. 28-36
P-ISSN: 2548-9534 | E-ISSN: 2548-9550

Exploration of Spatial Design and Interactive AI in Development of 3D Audio Games for Visually Impaired Players: A Systematic Literature Review

Kurniawan Kaka Dwi Sukardi[a] , Agus Juhana[b]

[a,b]Universitas Pendidikan Indonesia, Jawa Barat

ARTICLE INFORMATION

Article History:
Received by Editorial Board: December 02
Final Revision: July 12, 26
Published Online: July 13, 26

KEYWORDS

Audio Games, Visually Impaired, AI Voice Assistant, Systematic Literature Review, PRISMA Framework

CORRESPONDENCE

Kurniawan Kaka Dwi Sukardi
Departement of Multimedia Education
Universitas Pendidikan Indonesia
kurniawankaka21@upi.edu

ABSTRACT

The paradox of digital accessibility in the global gaming industry has systematically marginalized more than 2.2 billion visually impaired people, creating an immersion gap between visual and auditory gaming experiences. This study investigates the integration of 3D audio games and AI based voice assistants to create immersive and inclusive gaming experiences through a Systematic Literature Review (SLR) using the PRISMA framework and snowballing technique. 975 studies identified, 56 articles met the inclusion criteria after a rigorous selection process. The synthesis results reveal key findings, namely that immersive 3D audio design requires spatial audio precision, adaptive sonification, and haptic feedback to help players build mental maps of the virtual environment. Additionally, the need for narrative design and voice-based interaction emphasizes the importance of exploration flexibility, audio information balance, and adaptive narrative systems capable of adjusting responses based on situational context. Furthermore, the integration of artificial intelligence as an interactive companion plays a crucial role in providing contextual narratives and spatial navigation guidance through two-way voice interaction. The research conclusion affirms that the synergy between spatial audio technology, adaptive narrative design, and artificial intelligence creates a new paradigm in inclusive game design, where sound is no longer merely a substitute for visuals but the primary medium that unifies perception and interaction in an equitable gaming experience.

DOI: https://doi.org/10.46961/jommit.v10i1.1874

1. INTRODUCTION

Digital accessibility in the global gaming industry proves to be a paradox. Although the gaming industry is growing with increasing magnitude, becoming a huge cultural as well as economic force, there seems to be a systematic disregard for a large segment of the population, particularly individuals with disabilities, particularly those with visual impairment. Statistics from the World Health Organization (WHO) suggest that between 2019 and 2023, at least 2.2 billion people in the world live with some form of visual impairment [1]. Such a large segment of the population is left with no access to this recreational activity. This gap is not only about entertainment, but also about social participation, community building, and even cognitive opportunities [2]. However, the dominance of visual elements in conventional game design creates an almost insurmountable "visual wall." Most games are generally designed with high visuals without considering audio accessibility and navigation, thus marginalizing the visually impaired from the exciting experience of playing games as entertainment [3], [4]. Visually impaired individuals have damaged and limited eyesight, making it difficult for them to process visual information. However, research shows that visually impaired individuals have sharper audio abilities, so special methods are needed to create games for people with disabilities, especially the visually impaired [4], [5].

The challenges faced by visually impaired people in playing games are not only a lack of access, but also the very uneven quality of the gaming experience. The majority of games that are claimed to be "accessible" to the visually impaired, such as A Blind Legend or AudioWizards, are often limited to puzzle or runner genres with linear and repetitive mechanics. Despite being laudable for making efforts to address this challenge, they do not really portray the reality associated with modern gaming by virtue of being rich in open world exploration, narrative complexities, dynamic elements, and social interactions among others. Consequently, this generates what may refer to as the "immersion gap" between the gaming experience for the visually impaired and visual experience for normally sighted individuals [4]. The technical challenges are multidimensional. Firstly spatial navigation in 3D environments is one of the severe difficulties, without visual cues, spatial disorientation can be easily encountered, and poor environmental sound design and absence of audio landmarks are some of the major causes of disorienting conditions [4], [6], [7], [8]. Second, it is very hard to understand the context and story, players in visual games know what's going on with the story, the characters feelings, and their mission goals by using cutscenes, facial expressions, and text in the environment [8]. Third, it's very hard to interact with complicated game objects and mechanics, without a well audio feedback system, it's impossible to tell the difference between usable objects, enemies, and decorative elements [7].

An overarching approach to overcome the fundamental limitations of existing accessible games, a synergistic integration between an engaging 3D audio based gaming environment and a contextual An AI-based voice assistant could be the answer. This solution is designed to create a complete "perception ecosystem," in which individual components are interdependent in some fill the cognitive and sensory gaps left by the absence of visuals. First, 3D or binaural audio acts as a "spatial canvas" for players, it simulates how human ears pick up (interaural time and interaural level differences) using the concept of auditory space to manipulate digital sounds to be as similar as possible to the real world by changing effects and filters in certain circumstances [9]. Development of the Via Echo game prototype, players can accurately localize the direction of enemy sounds, estimate enemy distance based on echoes, or perceive the depth of a room based on reverberation [9], [10]. Several studies have 3D audio representation has been proven to greatly navigation performance and increases the user's sense of presence in virtual environments for users with visual impairments [7], [10], [11]. However, if just using 3D is not enough to convey complex narratives, this AI assistant is not just a passive text reader, but acts as an intelligent narrative guide. In the EnVisionVR game prototype, which integrates Virtual Reality (VR) with AI, it has been proven successful in making AI an assistant that helps understand the game and create adaptive narratives. It can proactively describe the location of objects, provide contextual navigation instructions, and identify the missions that must be completed to finish the game [12]. This integration transforms the experience from simply "hearing a world" to "interacting with a living and responsive world."

In order to have a solid theoretical and empirical basis for designing research for game development, a Systematic Literature Review (SLR) was conducted to analyze any preceding literature associated with game audio, 3D audio, AI Voice Assistant technology, or Voice interaction for games. Through the SLR, it was possible to discover how to create audio-based games that are suitable for people with visual impairments. These findings will be the subject of further research. Based on these objectives, this study formulates three main research questions:
RQ1: How to design an immersive 3D audio-based game that can be accessed and played by visually impaired people without relying on visual elements?
RQ2: To what extent can the integration of artificial intelligence (AI) as an interactive companion help visually impaired players understand the context, navigation, and storyline of the game?
RQ3: What are the narrative design and voice-based interaction mechanism requirements that can deliver an exciting, inclusive, and equal gaming experience with visual-based games?
Through the answers to these RQs, this SLR will provide a foundation for further research that is appropriate, relevant, and innovative in the development of 3D audio-based games with AI voice assistants for the visually impaired.

2. METHOD

To compile this research, the Systematic Literature Review (SLR) method was used, implemented with the PRISMA framework and supplemented with Backward and Forward Snowballing techniques. SLR is a research method used to identify, evaluate, and interpret all findings available on a particular research topic in a systematic, explicit, and replicable manner[13] . Unlike traditional narrative literature reviews, which tend to be subjective and selective, SLR follows a rigorous process through predetermined stages to minimize bias and produce a comprehensive synthesis of evidence. SLR is designed to provide a comprehensive survey of a field of research by answering specific research questions[14] . The main benefit of SLR is its ability to provide a strong foundation for new research by identifying gaps in existing literature, consolidating scattered empirical findings, and providing a solid theoretical background. The function of the PRISMA framework is to guide the process of identifying, screening, eligibility, and inclusion of studies using a structured flowchart [15].

The SLR procedure in this study adopts the guidelines proposed by Kitchenham & Charters (2007) [13] and follows the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework [15] to ensure transparency and complete reporting. This process is carried out through the following key stages:

2.1. Problem Formulation

This initial stage forms the foundation of the entire SLR process. Based on the background described above, the core problem formulated is "How can design and technology principles be used to develop 3D audio-based games integrated with AI Voice Assistants that are effective and immersive for the visually impaired?" This clear problem formulation guides all subsequent stages, from formulating research questions to inclusion and exclusion criteria.

2.2. Keyword Formulation

The problem formulation and RQ were used to formulate the main keywords and their derivatives. The latter were then combined into a comprehensive search string using Boolean operators AND and OR. The following search strategy was used: *("game-based audio" OR "spatial sound" OR "accessibility game") AND ("artificial intelligence" OR "AI" OR "voice assistant") AND ("visual impairment" OR "blindness").*

2.3. Data Mining

The A primary search of the literature was done using Google. Scholar database accessed through Publish or Perish software on September 22, 2025. The search focused on articles this was published within the last 10 years (2015-2025) to ensure the relevance of the findings to contemporary technology. This process yielded a large number of potential papers; 975 were and were subsequently filtered in the following step.

2.4. Data Screening and PRISMA Framework

The article selection process was described using the PRISMA diagram to ensure a systematic and transparent flow. The identification stage from various databases yielded 975 studies, with no additional sources, after selecting titles with the words "systematic literature review," writings in the form of books, conferences, chapters, and languages other than English were excluded, leaving 852 studies. During the screening stage, 814 articles were excluded due to irrelevance based on the title, abstract, and missing files, leaving 38 studies to proceed to the inclusion stage for analysis. The selection process also involved analysis based on the inclusion and exclusion tables.

Table 1
Inclusion and Exclusion

Inclusion	Exclusion
English language	Non-English
Journal articles	Non-journal articles
Publication range 2015-2025	Published before 2015
Discussing games and AI	Discussing the hardware/tools used

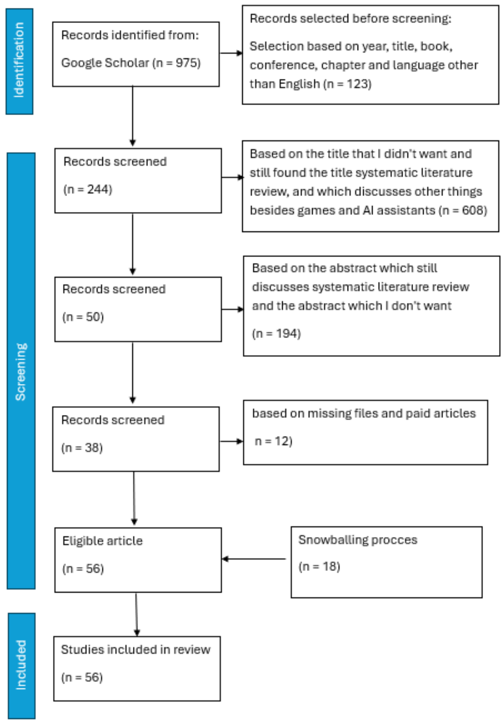

Figure 1. PRISMA Framework

2.5. Snowballing Process

To ensure that no relevant research is overlooked, this study applies the Snowballing technique. Snowballing is a method of searching for additional literature by examining backward snowballing reference lists and forward snowballing articles that cite core studies obtained from the initial database search [16]. This technique is very effective for finding seminal articles and recent research that may not be indexed by the keywords used [17]. From this snowballing process, 18 additional articles that met inclusion criteria. These articles provided deeper insights into 3D audio design, AI in accessibility, and psychology of visually impaired users in interaction with games. These 18 articles were combined with the 38 articles from the primary search, bringing the total number of studies analyzed in this SLR to 56 articles.

3. RESULTS AND DISCUSSION

3.1. Designing Immersive 3D Audio-Based Games for the Visually Impaired

The design of immersive 3D audio based games enables visually impaired games focuses on crafting a gaming experience devoid of depend more on visual elements but on spatial perception and interaction through sound. From the 25 studies, it has been observed that this type of research needs a multidisciplinary approach related to spatial audio, spatial cognition, auditory design, AI integration, and multimodal support like haptics and body movement.

One of the most important elements in 3D audio based game design involves the process of spatial perception construction using audio. It has been found in numerous studies that blind gamers have the ability to form spatial representations of space based solely on the location, range, and strength of spatialized audio cues [6], [18], [19], [20]. Systems such as Echo House make use of the phenomenon of sound reflection (echos) in order to help the user understand the size of the space through audibility [20]. These mechanisms take advantage of the blind brain's ability to make use of cross-modal plasticity, which refers to the brain's adaptation that translates visuals to complex sounds [18]. For this reason, the role of audio cues characterized by accurate localization in three dimensional space serves not only as a guiding tool based on direction but can in fact replace the visual element. The precision of the spatial location of the source of the audio cue, the predictability of the direction from which the cue comes, and the perceived difference between the main cue and the ambient noise all contribute towards achieving the goal of navigating through the game space [10], [19].

In addition to spatial aspects, the role of acoustic design and sonification is critical in achieving immersion. Every parameter of the acoustic signal like volume, pitch, tempo, timbre, or duration could potentially be utilized to communicate information about objects, distance, or game state. It has been found that the above setting of parameters influences the perception of auditory space, where pitch expresses vertical direction, and changes in intensity and reverberation simulate perceptions of distance [9]. Changes in tone and tempo are much more useful than plain and repetitive sounds in helping gamers distinguish directions [11], but achieving realism poses some issues in terms of balancing realness with readability, very realistic sounds can lead to confusion in the absence of visualization [9]. therefore, there is a need for an adaptive sonification system that can vary the style and level of sounds depending on the scenario in every game, as per the gamers choice.

Artificial Intelligence (AI) integration offers new avenues for advancing audio accessibility and immersion. AI can function as an assistant or narrator, providing real-time answers, descriptions, and feedback pertinent to the user [21], [22], [23]. One such example would be the VRSight System [22] , in which computer vision technology, natural language processing, and spatial audio come together to generate descriptive speech based on the player character's location in the world and the direction they are facing. Under this paradigm, players can detect spatial information both auditorily and visually, such as the positions of objects, motion paths, and the overall affective ambiance of the game space. Similar technology was developed in the VRBubble [23], where the AI regulates both the strength and spatial orientation of the sound source based on the inter-avatar distance in order to support social awareness through audition in multi player scenarios.

Besides research in the area of sound and artificial intelligence, there has been significant attention paid to the importance of multimodal interactions that involve both auditory feedback and body movements. These interactions have shown the ability to help players in understanding surroundings and react accordingly at a faster pace. In Virtual Showdown, both verbal and vibration cues help players determine the direction of the ball and predict where it is likely to go when there are no visual cues [24]. It is suggested that vibration feedback when combined with spatial audio can increase the precision of navigation in virtual environments [25]. There are other studies which introduce the Natural Audio Interface (NAI), which combines motion control with context-dependent sounds in order to support more natural interaction. Not only does such technology replace the function of vision but expands the spatial perception through the usage of several other human senses, thus meeting the requirements of visually impaired people [26].

All of these studies above highlight the challenges pertinent to carrying out inclusive design and achieving universal accessibility, assertion in accessibility by design is that the issue of accessibility needs to be considered in the early stages of game development [24], [27]. Common factors include auditory overload or too much information resulting in confusion, inability to distinguish vertical directions [7], [28], and storylines that are too long or irrelevant in context [21], [29]. In attempt to counter these challenges, many researchers such as in [30], [31] have ventured to rectify these challenges through the utilization of unconventional audio and manipulation of sound through time shifting, style shifting, and transparency shifting to manage the dynamic nature of sound in three-dimensional space environments. The process makes it easier to enjoy an accurate and clear listening experience. Other factors that can play important roles include volume, speech pacing, and the nature of the sound itself due to the different sensitivities of listeners [19].

3.2. The Need for Narrative Design and Sound-Based Interaction Mechanisms in Inclusive Games

The design of interactive mechanisms based on narratives and sound for immersive gaming that is inclusive of visually impaired persons would require the substitution of visual elements with meaningful sound experiences. The primary need is creating a flexible, dynamic experience that allows users to build their own sense emotions through sound, movement, and adaptive system responses. Essentially, the design would focus on finding ways for users to create their own affective experiences through sound, gesture, and system response [32].

The interaction design point of view, the most significant factors have been identified as movement, expression, interaction, and gesture. Movement and locomotion require spatial audio, but adaptive text-to-speech synthesis is capable of emotional expression. Foley sound helps to represent physical interactions, while gestures can also benefit from sound [8]. Nevertheless, all these sounds need to be perceptually balanced to avoid overwhelming the user's senses, which is helped by haptic and AI driven microexpression [8]. Another successful approach is to use voice input and bodily movement to create efficient core gaming activities for a VR game to find characters through "Marco Polo" communication [33]. This confirms that voice can be a central gameplay element, not just an accessibility feature.

Interactivity also extends to narrative. Stories put together using spatial audio or binaural rendering may give players the opportunity to decide what happens to the story simply by adjusting their own head position [34]. Moreover, research about the ways partially sighted players adapt existing games also confirms the development of their own distinctive sonic landmarks within said games, thereby supporting the need to be receptive to player adaptation of sound based games themselves [35]. In terms of technology, the repetitive use of the sonification of variables through pitch and volume to express surface properties helps enhance the formation of a sound language that facilitates navigation as well as the progression of narratives [36]. Experimentation occurs through the use of sound perspective shift, allowing users to shift between the first-person perspective and third person perspective, thus incorporating sound into the gaming aspect, much like cinematography [37].

Overall, the basic requirements for inclusive storytelling and audio interaction, as identified, are flexible auditory exploration, well balanced sonic information, combined multimodal inputs, as well as adaptive narrative systems. An exciting and equitable gaming experience for visually impaired players is not just the product of advanced audio technology, but of a design that empowers players to construct meaning, interact naturally, and feel in complete control of the game world. Through concepts like embodied sense-making, where sound and haptic feedback make players "feel" interactions [38], sound is transformed from a mere visual substitute into the primary medium that unifies narrative, emotion, and action into an equitable and inclusive experience for all.

3.3. Integration of Artificial Intelligence (AI) as an Interactive Companion in Inclusive Gaming

The integration of artificial intelligence for voice-based games is highly promising for offering an adaptive, interactive, and equal gaming experience for visually impaired individuals. A total of 23 articles are presented to describe the integration of games and artificial intelligence. This AI acts as an interactive companion for players to not only deliver information but also to facilitate understanding of the context of the story being depicted, directions being followed, and information regarding the gaming environment presented through natural voice conversation. Using NLP and context awareness, this AI helps to create a highly human like gaming experience for users based on their sensory requirements.

The role of AI in developing contextual narratives is important in helping them understand the world without explicitly showing or telling them visuals. Various studies demonstrate that scene comprehension and natural language processing technologies, which can be used to generate adaptive and personalized descriptions of the environment. For example, AI can provide dynamic narratives about the atmosphere of a space and interaction of objects so that players understand the context through auditory means experience rather than sight [39]. In other studies, AI is used to detect objects and translate the recognition results into real-time narratives relevant to the game situation, enabling players to know the position and role of elements surrounding them [40]. AI can also be used as the explanation of particular events in the story development, allowing players to create a better spatial understanding of the virtual world they are exploring [41]. This AI based narrative approach also uses binaural rendering capabilities of adjusting the sound direction to the player position, making the listening experience more immersive [42]. Moreover, other research confirms that AI is able to build the narrative context especially effective when combined with multimodal models that connect voice, text and visual space recognition, such as those developed in vision language model based systems to describe scenes in real time [12], [43], [44].

Other than its use as a narrative component, AI is also vital for aiding navigation and orientation for visually impaired gamers. AI based spatial audio mapping techniques have also been demonstrated to facilitate gamers perception of objects direction, distance, and shape within this virtual setting using differences in audio frequency and intensity [45]. In various studies, AI is also utilized to create audio virtual maps which change depending on the gamers movements and provide audio information for any changes in position [41]. This strategy also further enhances directional perception by utilizing AI as an adaptive navigation agent for interactive gamers responses [46]. Some systems also make use of machine learning and reinforcement learning in order to estimate optimal paths and give indications by means of sound patterns or haptic vibrations whenever players are about to hit an obstacle [42]. In the broader context of virtual reality, the combination of AI with spatialized sound rendering enables visually impaired users to construct three-dimensional mental maps based on the direction of sound sources and environmental feedback [47]. Application of this concept can be seen in research on multimodal AI systems able to merge voice input, head direction, and spatial position in order to help users intuitively understand the spatial context [12], [48]. These findings demonstrate that AI is not only a provider of navigation guidance; it further acts as a system for the comprehension of space that changes for every player according to individual style.

AI also serves as a communicative companion that is able to interact with the player in a natural way, using voice. Many studies created AI as an interactive companion capable of holding conversations, giving instructions, or answering players' emotions based on what is going on in the game [49]. This approach is adopted in speech based interaction systems whereby AI spontaneously answers questions posed by players or comments made by the latter as part of their efforts to restore the continuity of the narrative flow. In one more experiment, AI companions have been used with the intention of helping players solve puzzles and provide emotional support during gameplay [50]. More interactively, AI can even change its language style, intonation, and speed of utterance to feel more natural to the visually impaired user [51]. Several works have also presented AudioGPT and SpeechGPT technologies that allow AI to combine speech recognition and language generation in order to create a real-time two-way dialogue experience [52], [53]. Companion AI is also used in further studies in enriching branching narratives where players can decide upon the direction of the story through conversations handled by an AI system understanding the context and goals of the player [54], [55]. Through such an approach, AI would not be only a tool but an active interactive companion influencing the shaping of the player's emotional experience.

AI integrated with sensors and adaptive emotion systems is another focus area of personalized gaming experience development. Many studies show that artificial intelligence can identify physiological signals such as heart rate, tone, or respiratory rate in order to evaluate the player emotion as well as self regulate their gaming experience. [56]. With such a method as MFF, artificial intelligence is able to fuse information from user behavior and voice in determining the degree of stress or confusion that the user experiences, a suitable response can be given via non player characters through controlling storytelling speed or through the repetitive addition of directional information. [54]. Other studies involved adaptive artificial intelligence, showing its ability to adapt the scenario in a game based on player changes in their emotional states, thus increasing levels of player personalization as well as empathy [55]. Such studies show that AI not only assists as a cognitive tool, as commonly seen, but as a social actor that can respond to human emotional processes.

However, the use of AI technology as an interactive companion in applications designed for the visually impaired faces some technical and conceptual issues. Many studies highlight that the computing requirements of real-time AI applications often cause delays in the processing necessary for voice output, thus disrupting the continuity of interaction [40], [41]. Possibility of directional perception errors and audio overload is likely given that AI may generate too much information at any one moment in time [42], [46]. These models may also cause hallucination bias, as the generated descriptions do not relate to the real context [12], [56]. Another disadvantage is that there would be no feedback correction system in place on the users side for rectifying the errors in the descriptions produced by the AI [54], [55]. Current researchers stress the importance of finding a balance between the complexity of the information and the simplicity of sound design to prevent user cognitive fatigue [44]. Thus, efficiency, accuracy, and adaptability with regard to the auditory requirements of visually impaired users must be considered in the design of future AI models.

4. DISCUSSION

In conclusion, the results of this systematic literature review offer empirical evidence and a framework that supports the need for a holistic perspective in addressing the “visual wall” that visually impaired persons face in the gaming industry. In fact, the compilation of the 56 studies analyzed not only sufficiently responded to the three research questions (RQs) but effectively supported the original hypothesis that the complementary blend of immersive 3D audio with the use of contextual AI assistants is a basic solution [9], [12].

This synthesis also has enormous implications for game developers. Firstly, it is essential to ensure the principle of accessible by design, where 3D audio and AI are integrated from day one in development rather than towards the end stages of development as demonstrated in studies [24], [27]. Secondly, focus must be on audio design consistency to ensure that designs are coherent [36] and individual demands for sound customization to fulfill human perception requirements [19]. Thirdly, it would be appreciable to focus on AI systems developed to act as pro active and context sensitive interactive gaming partners. Fourthly, visual impaired usability testing would be an increasing source of learning for further development of game designs and understanding human perceptions according to visual impaired testing conducted iteratively [35]. Conversely, for this SLR study, there can be several drawbacks such as possibly unincluded data status from different databases and the increasing dynamics related to this gaming research area, in upcoming research study areas, focusing on optimizing AI gaming technologies to present less latency issues and hallucination biases [12], [40], [56], sound management strategies for overcoming sound overflow phenomena related to gamer auditory perceptions [28], [44], establishment of standardization for designs, and further application of these concepts to uncharted areas of game genres.

This systematic review has successfully established a thorough foundation, showing that the path to truly inclusive gaming for the visually impaired lies in building a complete "perception ecosystem". Within this ecosystem, where the sound is rendered in engaging 3D auditory space and contextual AI voice assistants fill it with meaning interaction. Supported by flexible and multimodal narrative design and interaction, this synergistic integration has the has the potential to turn the paradox of digital accessibility into a reality of thrilling, inclusive, and entertaining gaming all experiences must be available to everyone, without exclusion.

5. CONCLUSION

From the analysis of the three primary focuses of research, it can be said that it could be concluded that the design of 3D audio-based games for the rehabilitation of a visually impaired requires an approach that is multidisciplinary in nature. Combines spatial audio technology with adaptive narrative design and the integration of artificial intelligence as an interactive companion. From the perspective of immersive 3D audio design, sound has proven to play a major role in building spatial perception orientation and environmental context. Through application, of precise spatial audio and adaptive sonication, visually the impaired players can create mental representations, such as mental mapping without the use of a display. The realism sound aspect represents a central element that must be achieved in sound interaction design in order for the experience to remain enjoyable. From the results, sound immersion necessitates a flexible form of storytelling through gaming that encourages individual discovery, sound is not only a replacement for visuals but also acts as a means to express, communicate, convey emotion by adding body movements, haptic input, and speech to enhance natural interaction between players and the environment. Interactive sonic narratives form the essential component required for developing a fair gaming experience for visually impaired individuals, also AI is significantly involved in such technology, as it serves as the interactive companion that is able to comprehend the scenario, explain the environment, as well as provide path guidance. However, the use of AI is impeded by technology limitations such as lag, loud background noise, as well as the possibility of discriminatory expressions.

REFERENCES

[1] D. Wudie, H. Nigussie, and T. Gelana, “Prevalence of eye diseases, refractive errors and visual impairments in the Debre Markos referral hospital,” *Sci. Rep.*, vol. 15, no. 1, pp. 1–12, 2025, doi: 10.1038/s41598-025-04262-3.

[2] S. Mambela, “TINJAUAN UMUM MASALAH

PSIKOLOGIS DAN MASALAH SOSIAL INDIVIDU PENYANDANG TUNANETRA," *J. Buana Pendidik.*, no. 25, pp. 65–73, 2018.

[3] L. V. Neto, P. H. F. Fontoura Junior, R. A. Bordini, J. L. Otsuka, and D. M. Beder, "Design and implementation of an educational game considering issues for visually impaired people inclusion," *Smart Learn. Environ.*, vol. 7, no. 1, 2020, doi: 10.1186/s40561-019-0103-4.

[4] E. Agrimi, C. Battaglini, D. Bottari, G. Gnecco, and B. Leporini, "Game accessibility for visually impaired people: a review," *Soft Comput.*, vol. 28, no. 17–18, pp. 10475–10489, 2024, doi: 10.1007/s00500-024-09827-4.

[5] R. R. Handoyo, "PENGEMBANGAN KOMUNIKASI ANAK TUNANETRA DALAM PERMAINAN KOOPERATIF TRADITIONAL," *J. Pendidik. Khusus*, vol. 12, no. 2, pp. 30–45, 2016.

[6] R. Andrade, M. J. Rogerson, J. Waycott, S. Baker, and F. Vetere, "Playing Blind," pp. 1–14, 2019, doi: 10.1145/3290605.3300346.

[7] C. S. C. Rodrigues, V. Nazareth, R. O. Azevedo, P. Barbosa, and C. Werner, "Unseen: Advancing Digital Accessibility with Binaural Audio Technology in an Immersive Gaming Prototype," *J. Interact. Syst.*, vol. 16, no. 1, pp. 98–108, 2025, doi: 10.5753/jis.2025.4439.

[8] B. Biggs, S. Murgaski, P. Coppin, and B. N. Walker, "Creating Non-Visual Non-Verbal Social Interactions in Virtual Reality," *Virtual Worlds*, vol. 4, no. 2, p. 25, 2025, doi: 10.3390/virtualworlds4020025.

[9] T. Farkaš, "Understanding Auditory Space in Digital Games for Visually Impaired People," *Acta Ludologica*, vol. 7, no. 1, pp. 136–150, 2024, doi: 10.34135/actaludologica.2024-7-1.136-150.

[10] G. Vontzalidis, S. Mystakidis, A. Christopoulos, and K. Moustakas, "Spatial Audio Cues in an Immersive Virtual Reality STEM Escape Room Game: A Comparative Study," *Commun. Comput. Inf. Sci.*, vol. 2271 CCIS, pp. 317–328, 2025, doi: 10.1007/978-3-031-80475-5_23.

[11] S. Poudratchi, Y. A. Sekhavat, M. R. Azadehfar, and S. Roohi, "BatSight: A Navigation Game to Map Environmental Information into Audio Cues," *Int. J. Serious Games*, vol. 11, no. 1, pp. 101–119, 2024, doi: 10.17083/ijsg.v11i1.718.

[12] J. Chen, R. P. G. Esparza, V. Garaj, P. O. Kristensson, and J. Dudley, "EnVisionVR: A Scene Interpretation Tool for Visual Accessibility in Virtual Reality," pp. 1–14, 2025, [Online]. Available: http://arxiv.org/abs/2502.03564

[13] B. Kitchenham, "Kitchenham , B .: Guidelines for performing Systematic Literature Reviews in software engineering . EBSE Technical Report EBSE-2007-01 Guidelines for performing Systematic Literature Reviews in Software Engineering," *Icse*, no. January 2007, pp. 1–57, 2007.

[14] C. Cronin, "Doing your literature review: traditional and systematic techniques," *Eval. Res. Educ.*, vol. 24, no. 3, pp. 219–221, 2011, doi: 10.1080/09500790.2011.581509.

[15] M. J. Page *et al.*, "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *Rev. Panam. Salud Publica/Pan Am. J. Public Heal.*, vol. 46, pp. 1–11, 2022, doi: 10.26633/RPSP.2022.112.

[16] C. Wohlin, "Guidelines for snowballing in systematic literature studies and a replication in software engineering," *ACM Int. Conf. Proceeding Ser.*, 2014, doi: 10.1145/2601248.2601268.

[17] T. Greenhalgh and R. Peacock, "Effectiveness and efficiency of search methods in systematic reviews of complex evidence: Audit of primary sources," *Br. Med. J.*, vol. 331, no. 7524, pp. 1064–1065, 2005, doi: 10.1136/bmj.38636.593461.68.

[18] O. Balan, A. Moldoveanu, and F. Moldoveanu, "Spatial Auditory Representation in the Case of the Visually Impaired People," *J. Inf. Syst. Oper. Manag.*, no. May, pp. 1–11, 2015, [Online]. Available: https://manchester.idm.oclc.org/login?url=https://www.proquest.com/scholarly-journals/spatial-auditory-representation-case-visually/docview/1771399885/se-2?accountid=12253%0Ahttps://media.proquest.com/media/hms/OBJ/kKt0S?_a=ChgyMDI1MDMxMjExMTE1NjIyOTo4OTI

[19] V. Nair *et al.*, *Uncovering Visually Impaired Gamers' Preferences for Spatial Awareness Tools Within Video Games*, vol. 1, no. 1. Association for Computing Machinery, 2022. doi: 10.1145/3517428.3544802.

[20] R. Andrade, S. Baker, J. Waycott, and F. Vetere, "Echo-house: Exploring a virtual environment by using echolocation," *ACM Int. Conf. Proceeding Ser.*, no. August, pp. 278–289, 2018, doi: 10.1145/3292147.3292163.

[21] Y. Yensathit, P. Ariya, K. Intawong, and K. Puritat, "Accessible virtual reality for library orientation: A user-centered approach for visually impaired learners in inclusive education," *Soc. Sci. Humanit. Open*, vol. 11, no. May, p. 101630, 2025, doi: 10.1016/j.ssaho.2025.101630.

[22] D. Killough *et al.*, *VRSight: An AI-Driven Scene Description System to Improve Virtual Reality Accessibility for Blind People*, vol. 1, no. 1. arXiv, 2025. doi: 10.1145/3746059.3747641.

[23] T. F. Ji, B. Cochran, and Y. Zhao, *VRBubble: Enhancing Peripheral Awareness of Avatars for People with Visual Impairments in Social Virtual Reality*, vol. 1, no. 1. Association for Computing Machinery, 2022. doi: 10.1145/3517428.3544821.

[24] R. Wedoff, L. Ball, A. Wang, Y. X. Khoo, L. Lieberman, and K. Rector, "Virtual Showdown: An accessible virtual reality game with scaffolds for youth with visual impairments," *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 1–15, 2019, doi: 10.1145/3290605.3300371.

[25] O. BĂlan, A. Moldoveanu, and F. Moldoveanu,

"Binaural sound analysis and spatial localization for the visually impaired people," *Proc. Int. Conf. Interfaces Hum. Comput. Interact. 2015, IHCI 2015, Game Entertain. Technol. 2015, GET 2015 Comput. Graph. Vis. Comput. Vis. Image Process. 2015, CGVCVIP 2015 - P*, no. July, pp. 331–335, 2015.

[26] N. Interfaces and J. Li, "applied sciences Beyond Sight : Enhancing Augmented Reality Interactivity with," 2024.

[27] M. Mott *et al.*, "Accessible by design: An opportunity for virtual reality," *Adjun. Proc. 2019 IEEE Int. Symp. Mix. Augment. Reality, ISMAR-Adjunct 2019*, pp. 451–454, 2019, doi: 10.1109/ISMAR-Adjunct.2019.00122.

[28] J. Mendes, M. Piçarra, I. Gonçalves, A. Rodrigues, and J. Guerreiro, "Exploring Aiming Techniques for Blind People in Virtual Reality," *IEEE Trans. Vis. Comput. Graph.*, vol. 31, no. 5, pp. 3267–3274, 2025, doi: 10.1109/TVCG.2025.3549847.

[29] J. A. C. Castaneda *et al.*, "Designing inclusive tech playful educative solutions for visually impaired learners in STEM education," *Smart Learn. Environ.*, vol. 12, no. 1, 2025, doi: 10.1186/s40561-024-00358-x.

[30] M. Krzyzaniak, D. Frohlich, and P. J. B. Jackson, "Six types of audio that DEFY reality!: A taxonomy of audio augmented reality with examples," *ACM Int. Conf. Proceeding Ser.*, pp. 160–167, 2019, doi: 10.1145/3356590.3356615.

[31] R. C. Chang, C. S. Hung, B. Y. Chen, D. Jain, and A. Guo, "SoundShif: Exploring Sound Manipulations for Accessible Mixed-Reality Awareness," *Proc. 2024 ACM Des. Interact. Syst. Conf. DIS 2024*, pp. 116–132, 2024, doi: 10.1145/3643834.3661556.

[32] H. Balasubramanian *et al.*, "Enable Blind Users' Experience in 3D Virtual Environments: The Scene Weaver Prototype," *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 2–5, 2023, doi: 10.1145/3544549.3583909.

[33] Lucas R. Lebrao, Robert E. McKenna, and Justin J. Morrow, "Creating an Audio Game Platform for the Visually Impaired," 2015, [Online]. Available: https://web.wpi.edu/Pubs/E-project/Available/E-project-050715-223127/unrestricted/IQP_JM_RM_LL.pdf

[34] M. Geronazzo *et al.*, "Creating an Audio Story with Interactive Binaural Rendering in Virtual Reality," *Wirel. Commun. Mob. Comput.*, vol. 2019, 2019, doi: 10.1155/2019/1463204.

[35] D. Gonçalves, M. Piçarra, P. Pais, J. Guerreiro, and A. Rodrigues, "'My Zelda Cane': Strategies Used by Blind Players to Play Visual-Centric Digital Games," *Conf. Hum. Factors Comput. Syst. - Proc.*, 2023, doi: 10.1145/3544548.3580702.

[36] J. Schito and S. I. Fabrikant, "Exploring maps by sounds: using parameter mapping sonification to make digital elevation models audible," *Int. J. Geogr. Inf. Sci.*, vol. 32, no. 5, pp. 874–906, 2018, doi: 10.1080/13658816.2017.1420192.

[37] I. de Villiers Bosman, O. O. Buruk, R. Tordiffe, K. Jørgensen, T. Huang, and J. Hamari, "SoundTrek: A Virtual Reality Audio Perspective-Switching Game," *CHI-PLAY Companion 2024 - Companion Proc. Annu. Symp. Comput. Interact. Play*, pp. 352–358, 2024, doi: 10.1145/3665463.3678824.

[38] R. van der Veen, J. Peeters, and A. Trotto, "Charged Utopia VR," pp. 292–298, 2018, doi: 10.1145/3173225.3173286.

[39] J. Vilares, C. Gómez-Rodríguez, L. Fernández-Núñez, D. Penas, and J. Viteri, "Bringing Roguelikes to Visually-Impaired Players by Using {NLP}," *Work. Games Nat. Lang. Process.*, no. May, pp. 59–67, 2020, [Online]. Available: https://aclanthology.org/2020.gamnlp-1.9/

[40] T. Qiu *et al.*, *GamerAstra: Enhancing Video Game Accessibility for Blind and Low-Vision Players through a Multi-Agent AI Framework*, vol. 1, no. 1. arXiv, 2025. [Online]. Available: http://arxiv.org/abs/2506.22937

[41] X. Cao, K. P. Ju, C. Li, and D. Jain, *SceneGenA11y: How can Runtime Generative tools improve the Accessibility of a Virtual 3D Scene?*, vol. 1, no. 1. Association for Computing Machinery, 2025. doi: 10.1145/3706599.3720265.

[42] X. Cao, K. P. Ju, C. Li, V. Potluri, and D. Jain, *RAVEN: Realtime Accessibility in Virtual ENvironments for Blind and Low-Vision People*, vol. 1, no. 1. arXiv, 2025. [Online]. Available: http://arxiv.org/abs/2510.06573

[43] F. De La Torre, C. M. Fang, H. Huang, A. Banburski-Fahey, J. A. Fernandez, and J. Lanier, *LLMR: Real-time Prompting of Interactive Worlds using Large Language Models*. 2024. doi: 10.1145/3613904.3642579.

[44] R. Gonzalez, J. Collins, C. Bennett, and S. Azenkot, "Investigating Use Cases of AI-Powered Scene Description Applications for Blind and Low Vision People," *Conf. Hum. Factors Comput. Syst. - Proc.*, 2024, doi: 10.1145/3613904.3642211.

[45] R. D. Gaina and M. Stephenson, "'did you hear that?' Learning to play video games from audio cues," *IEEE Conf. Comput. Intell. Games, CIG*, vol. 2019-Augus, 2019, doi: 10.1109/CIG.2019.8848088.

[46] C. M. Tran, N. G. Bach, P. X. Tan, E. Kamioka, and M. Kanamaru, "Enabling Self-Practice of Digital Audio–Tactile Maps for Visually Impaired People by Large Language Models," *Electron.*, vol. 13, no. 12, pp. 1–15, 2024, doi: 10.3390/electronics13122395.

[47] M. Chen *et al.*, "SAVVY: Spatial Awareness via Audio-Visual LLMs through Seeing and Hearing," 2025, [Online]. Available: http://arxiv.org/abs/2506.05414

[48] P. D. G. Vyawahare, "Image to Audio Conversion for Blind People Using Neural Network," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 11, no. 12, pp. 1949–1957, 2023, doi: 10.22214/ijraset.2023.57712.

[49] C. Zhong, J. Yu, Y. Cao, S. Wu, W. Wu, and K. Zhang, "SoundScape: A Human-AI Co-Creation System

Making Your Memories Heard," 2024, [Online]. Available: http://arxiv.org/abs/2410.08136

[50] A. El Saddik, J. Ahmad, M. Khan, S. Abouzahir, and W. Gueaieb, "Unleashing Creativity in the Metaverse: Generative AI and Multimodal Content," *ACM Trans. Multimed. Comput. Commun. Appl.*, vol. 21, no. 7, pp. 1–43, 2025, doi: 10.1145/3713075.

[51] R. Huang *et al.*, "AudioGPT: Understanding and Generating Speech, Music, Sound, and Talking Head," *Proc. AAAI Conf. Artif. Intell.*, vol. 38, no. 21, pp. 23802–23804, 2024, doi: 10.1609/aaai.v38i21.30570.

[52] D. Zhang *et al.*, "SpeechGPT: Empowering Large Language Models with Intrinsic Cross-Modal Conversational Abilities," *Find. Assoc. Comput. Linguist. EMNLP 2023*, pp. 15757–15773, 2023, doi: 10.18653/v1/2023.findings-emnlp.1055.

[53] A. Kuzdeuov, O. Mukayev, S. Nurgaliyev, A. Kunbolsyn, and H. A. Varol, "ChatGPT for Visually Impaired and Blind," *6th Int. Conf. Artif. Intell. Inf. Commun. ICAIIC 2024*, no. March, pp. 722–727, 2024, doi: 10.1109/ICAIIC60209.2024.10463430.

[54] S. A. Selvamani and N. D. Ganapathy, "A Multi-Agent AI Framework for Immersive Audiobook Production through Spatial Audio and Neural Narration," pp. 1–14, 2025, [Online]. Available: http://arxiv.org/abs/2505.04885

[55] S. Xu, X. Jin, W. Zhang, H. Qu, and Y. Yan, *Branch Explorer: Leveraging Branching Narratives to Support Interactive 360° Video Viewing for Blind and Low Vision Users*, vol. 1, no. 1. arXiv, 2025. doi: 10.1145/3746059.3747791.

[56] A. Raza *et al.*, "Optimized virtual reality design through user immersion level detection with novel feature fusion and explainable artificial intelligence," *PeerJ Comput. Sci.*, vol. 10, pp. 1–25, 2024, doi: 10.7717/PEERJ-CS.2150.