A Case Study on the Effects of Auditory Cues in Influencing Spatial Memory and Representation for a Person who is Blind

Minh Nguyen
minh.t.nguyen-4@student.uts.edu.au
School of Computer Science,
University of Technology Sydney
Sydney, Australia

Howe Zhu
howe.zhu@uts.edu.au
School of Computer Science,
University of Technology Sydney
Sydney, Australia

Huiyuan Sun
huiyuan.sun@sydney.edu.au
School of Electrical and Information Engineering, The University of Sydney
Sydney, Australia

Vincent Nguyen
vincent.nguyen@uts.edu.au
Graduate School of Health, University of Technology Sydney
Sydney, Australia

Avinash Singh
avinash.singh@uts.edu.au
School of Computer Science,
University of Technology Sydney
Sydney, Australia

Craig Jin
craig.jin@sydney.edu.au
School of Electrical and Information Engineering, The University of Sydney
Sydney, Australia

Chin-Teng Lin
chin-teng.lin@uts.edu.au
School of Computer Science,
University of Technology Sydney
Sydney, Australia

ABSTRACT

Navigation is a coordinated and goal-oriented movement through the environment, in which vision is an integral part of acquiring spatial information. People with blindness and vision impairment rely on alternative senses to navigate in daily life. Auditory cues are critical in building a spatial representation and effective navigation. Our work investigates the neurophysiological response to explore how auditory cues can shape a person's spatial representation and memory of an environment. In our experiment, a person who is blind was presented with verbal and non-verbal cues, either at close-reaching distance or globally. Electroencephalography, walking trajectories and tactile drawing data were collected to evaluate spatial representation and memory formed by different auditory cues and the role of theta oscillations in such relationships. Our results indicated a consistent increase in theta power at turning and starting points, a higher cognitive load with verbal strategies, and enhanced path recollection with global strategies. These findings emphasise the role of theta as a key indicator of cognitive workload and spatial memory processing. The preliminary results contribute to a broader understanding of the underlying processing mechanism of auditory information processing to support spatial navigation.

BCIMM '24, October 28-November 1, 2024, Melbourne, VIC, Australia

ACM ISBN 979-8-4007-1189-3/24/10
https://doi.org/10.1145/3688862.3689115

CCS CONCEPTS

• **Human-centered computing** → *User studies*; • **Social and professional topics** → **People with disabilities**.

KEYWORDS

hearing; perception; audio; spatial representation; navigation; cognitive neuroscience; assistive technology

ACM Reference Format:
Minh Nguyen, Howe Zhu, Huiyuan Sun, Vincent Nguyen, Avinash Singh, Craig Jin, and Chin-Teng Lin. 2024. A Case Study on the Effects of Auditory Cues in Influencing Spatial Memory and Representation for a Person who is Blind. In *Proceedings of the 1st International Workshop on Brain-Computer Interfaces (BCI) for Multimedia Understanding (BCIMM '24), October 28-November 1, 2024, Melbourne, VIC, Australia.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3688862.3689115

1 INTRODUCTION

Navigating an environment is a fundamental aspect of daily life. Visual perception is critical in sensing and building a spatial representation of the surrounding environment. For individuals with blindness and vision impairment, this task involves a heavy dependence on alternative sensory inputs and complex cognitive processes that differ significantly from those used by the sighted [1]. These strategies consist of sensory perception, spatial memory, and decision-making. Sensory perception involves interpreting environmental cues, spatial memory allows individuals to store and recall information about their surroundings, and decision-making processes enable them to make a turn on the path, avoid obstacles, and correct errors. For blind individuals, the reliance on auditory and tactile information necessitates a heightened level of cognitive processing to integrate these inputs effectively for navigation [1].

Hearing is critical in building a spatial representation of the surrounding environment through decoding semantics and spatial information. The human brain utilises binaural hearing to localise sounds both passively and actively [9, 19, 23]. Through training, certain individuals can utilise sound cues to construct a mental map of their surroundings to aid in navigating new environments [12–14]. Many assistive and navigational aids utilise auditory signals and associated feedback to aid users with visual impairment in comprehending their surrounding environment [21, 22]. In our work, we utilise Electroencephalography (EEG) sensors as a non-invasive method to measure the electrophysiological activity in the brain to study the temporal dynamics of cognitive functions associated with listening to sound cues [6, 8, 11]. In this regard, the theta band (4-8 Hz) activity is known to be involved in spatial memory [10, 15] and the integration of sensory information [10], facilitating the transformation of auditory cues into spatial representations [2].

In this paper, we aim to explore the effects of different auditory cues in spatial representation and memory and the role of theta band oscillations in the interpretation of the underlying cognitive processes of a person who is blind. Our hypotheses are: 1) verbal instructions require higher cognitive load due to the interpretation of linguistic information; 2) the presentation of global compared to local spatial and navigational information using auditory cues would facilitate more accurate spatial representations. We examined the behaviour and neurological response of one participant who is blind when using different auditory cues to perceive and navigate an unknown route. We found that the type of auditory cue (verbal or non-verbal) and the extent of the spatial information (local or global) can actively alter the participant's spatial representation. Our results indicate that the theta band can be a promising biomarker for evaluating the user's ability to build spatial representations. The study provides a glimpse into the interplay between hearing and spatial cognition and motivates further study with a larger cohort of participants. A deeper understanding of these issues can contribute to developing more effective navigational aids that enhance the independence of people with blindness.

2 METHODOLOGY

2.1 Participant

This study involves a male participant who has lived with blindness with no light perception for nine years (age = 62). The participant's main mobility method is a long cane. The participant reported no history of hearing loss. We acquired informed consent from the participant through an accessible form for a screen reader. Upon arrival, the research team introduced the person to the experiment area, protocol, and devices (Fig. 1). The research team verbally repeated the risks and consent statement to ensure the participant fully understood the experiment's requirements. The study was approved by the University of Technology Sydney's human research ethics committee (ETH22–6883).

2.2 Experiment Apparatus

The *Razer Anzu smart glasses* (Fig. 1B) was used to deliver high-fidelity spatialised audio with a low Bluetooth latency of 60ms [16]. Reflective markers are attached to the glasses to track the user's head position using the *Optitrack (Flex13) motion capture (mocap) cameras*. The binaural spatial sound was generated based on the relative distance between the user and the virtual sound sources [17, 18]. We recorded the EEG data with the *Unicorn Hybrid Black EEG headset* (Fig. 1D), a wireless hybrid (wet and dry electrodes) EEG headset with eight channels (Fz, Cz, C3, C4, Pz, PO7, PO8, and Oz). Saline fluid was applied to each electrode and sampled at 250 Hz. A *remote controller* (Fig. 1C) was used to prompt the auditory stimuli.

2.3 Experiment paradigm

The participant received instructions on how to navigate according to each experimental condition. Once the participant understood the process, he was briefed on the trial phases. At the beginning of each trial, the participant was guided to the starting position, handed a remote and then navigated a specific path at his own pace (Fig. 1E). Upon completing the walking phase, the operator guided him to a table to perform the task of drawing the perceived path. After completing the drawing phase, participants were guided back to the central point to start a new trial.

We presented the user with different auditory stimuli for each experiment condition. For the non-verbal cues, we used spatialised earcons - abstract sounds that have been processed to create the perception of those sounds originating from specific locations in three-dimensional space. Regarding the spatial extent of the cues, the auditory stimuli were delivered within a local context (i.e., acoustic sources were positioned one or two meters from the user) or in a more global context (i.e., acoustic sources were positioned along the entire path). Hence, we have four conditions in comparison.

Local earcon (LE): the two virtual loudspeaker positions nearest the participant's location were used to deliver spatialised earcons. The earcons were spatially to the left and right of the user. Sounds were presented on the listener's audiovisual horizon, and the sound sources were always positioned to orient the participant along the correct path. The guiding cue was a repeated sound burst imitating the noise of a white cane tapping.

Local speech (LS): non-spatialised verbal instructions were used to relay navigational instructions. Each cue lasted between 2.5 and 4 seconds, followed by a break of 4 to 5.5 seconds, providing sufficient time for the participant to follow them. For instance, "The path goes straight one meter" guides the user forward.

Global earcon (GE): spatialised sound bursts with a higher-pitch were presented at the participant's position and continuing along the path towards the destination. The participant was expected to keep the cues centred and follow them. At the corners of the path, a spatialised sound stimulus imitating a "flowing wind breeze" was played to indicate a turn. If the stimulus moves from left to right, it signals a right turn, and vice versa.

Global speech (GS): provides longer non-spatialised verbal instructions that describe the shape of the route. For instance, the participant would hear, "The path goes straight 3 meters, then turns right 90 degrees and straight 1 meter." The global speech also offers reminders if the participant loses track of his direction. For example, "Turn left to be on track," or "Turn around to be on track."

A brief burst of high-pitched tone acting as a warning sound would play continuously when the participant turned incorrectly. Upon reaching the end of the path, a "You have arrived" message indicated that the participant had reached the destination.

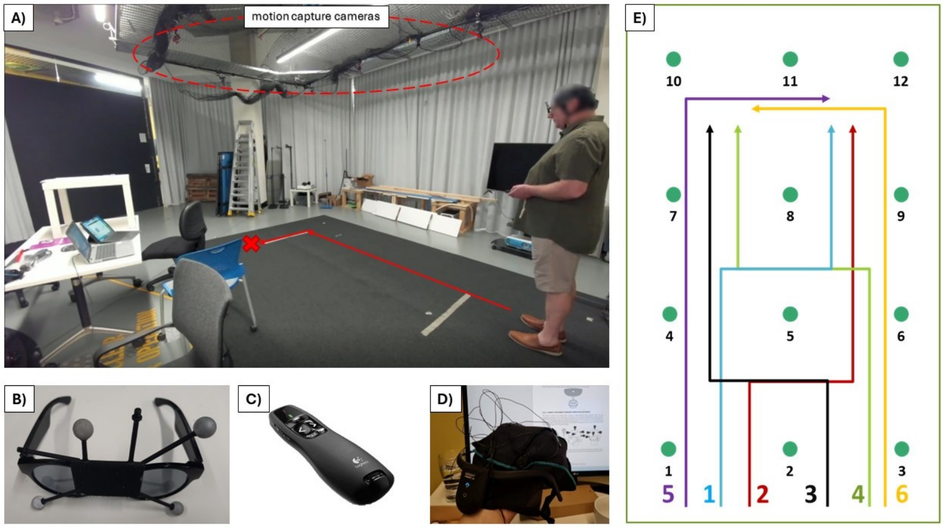

Figure 1: Experiment setup. A) The motion capture cameras track the participant's position. The red arrows show one of the paths of the participant, and the red "X" marks the destination. B) Smart glasses with reflective markers. C) A remote controller. D) The Unicorn Hybrid Black EEG headset. E) A layout of six different routes and the location of twelve virtual speakers.

2.4 Data analysis

We compare the differences in spatial representation resulting from the different auditory conditions. The coordinates (x,y and z from mocap) of the executed paths are plotted onto a 2D graph compared with the pre-defined routes (ground truth). In addition, the EEG theta power of the frontal lobe signal (Fz) per second was calculated and colour-coded onto the paths to examine the cognitive processes during navigation. The frontal theta signal reflects the higher cognitive control and spatial memory required to update the mental map and reorientation. Also, theta power in the frontal regions is associated with language processing encoding verbal instructions.

First, raw EEG data was band-pass filtered between 2 Hz and 60 Hz using a fourth-order Butterworth filter to retain cortical activity while removing slow drifts and high-frequency noise. Line noise at 50 Hz and its harmonics were attenuated using the CleanLine function from the EEGLAB toolbox [7]. Portions of the data with flat signal and excessively noisy channels were identified and removed. Removed channels were then re-added through spherical spline interpolation. Next, we used Artifact Subspace Reconstruction to remove high-amplitude artifacts (e.g., motion artifacts from movement) [4]. The data was re-referenced to the common average to remove any residual non-cortical activity. EEG data was epoched based on each trial and every condition. Finally, power spectral density (PSD) based on Welch's method was used to compute the power of each EEG band on each channel, with a Hamming window of one second and no overlapping.

Regarding spatial memory, the drawings were digitised from paper onto the computer and plotted against the walked trajectory and the ground truth. EEG captures the underlying neural processes, walking trajectories provide behavioural evidence, and tactile drawings externalise the participant's mental maps. Together, these modalities allow for a multidimensional understanding of spatial cognition, ensuring a more accurate and holistic evaluation.

3 RESULTS

To evaluate the effects of different auditory stimulus conditions on spatial representation and memory, Figure 2 compares the user trajectories and memory recall drawings of a 2-turn route (route 2) and a 1-turn route (route 6) in the four conditions (see Supplementary for the results of other routes). The colour bars show the range of theta power changes over time within a condition, with blue indicating lower power and red indicating higher power. The first observation from the results is an increased theta power at the start of each route and right after or before a path corner. In addition, higher theta power can also be seen in verbal strategies (LS and GS) compared to earcon strategies (LE and GE). Based on the scale of the colour bar, it seems that GS elicited theta power three times higher than GE. Another noteworthy point is that global strategies have more accurate drawings (five out of six for GE and six out of six for GS) than local strategies (three out of six for LE and two out of six for LS) (see Supplementary for more results), even though triggered warnings are spotted more frequent in global strategies. It can also be seen that frontal theta power in global strategies is at least ten times higher than in local strategies. Consequently, the results show that different auditory stimulus conditions provided the user with different spatial comprehension and mental mapping, and theta power in the frontal lobe corroborates such circumstances.

4 DISCUSSION

4.1 Heightened Cognitive Activity at the Start and Turning Points of the Path

Firstly, we observed a consistent increase in theta power at the beginning of navigation (when the participant is introduced to the path) and each path's turning points (corners) across all conditions. This increase in theta power may suggest heightened cognitive activity associated with spatial orientation and decision-making

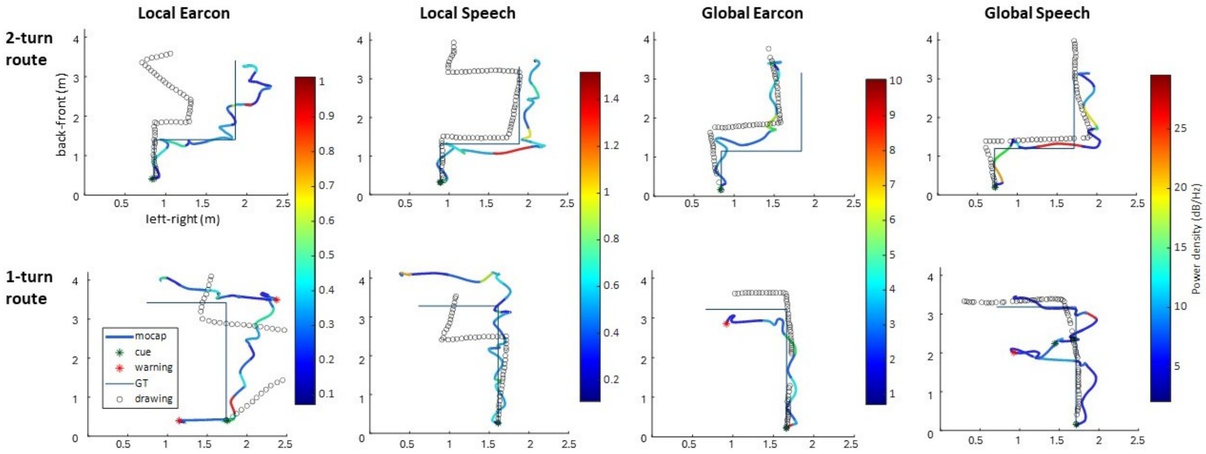

Figure 2: Trajectory and drawing of a 2-turn route (Route 2) and 1-turn route (Route 6) in the four conditions. Ground truth path (thin straight lines), mocap-based trajectory with theta power colour-coded (multi-colour lines) and drawing (black dotted lines) are plotted within a route. Green and red asterisks denote triggered guidance and warning, respectively.

at these critical points. The movement initiation and the requirement to change direction during turns will likely demand more mental resources, as the participant needs to process spatial information and update his mental maps accordingly. This finding aligns with previous research indicating that theta oscillations are closely linked to navigational decision points and the encoding of spatial information [2, 3, 5]. The increased theta power at these points may imply that the brain engages in more intense spatial processing and memory retrieval. This suggests that any navigational aid should ensure clarity and accuracy of cues at these critical moments to support the cognitive load experienced by users.

4.2 Earcons versus Speech in Navigation

Spatialised earcons can provide users guidance with a lower cognitive load by enabling them to interact with their surroundings more naturally and intuitively [20]. Our results reveal that theta power in the frontal lobe is higher when the participant uses verbal strategies (LS and GS) than non-verbal ones (LE and GE). The frontal lobe is known for its role in complex cognitive functions, including working memory, attention, and executive control. The increased theta activity in this region during verbal strategies suggests that these tasks impose greater cognitive demands on participants. Verbal instructions require linguistic processing, likely due to the increased cognitive load observed. This result underscores the importance of considering the type of auditory cue used in navigation aids, as verbal instructions may require more cognitive resources than non-verbal cues. Additionally, this finding suggests that verbal strategies may provide detailed guidance and contribute to cognitive fatigue over prolonged use.

4.3 Mental Mapping and Path Recollection

We hypothesised that global strategies likely facilitate a more holistic understanding of the travelled routes, allowing the users to form more accurate mental representations of the paths and retain more spatial information (e.g., the number of straight lines and corners). Our results indicate that the global strategies drawings are more accurate than those in local strategies. The improved recall performance in the global strategies suggests that these approaches provide more comprehensive spatial information, enabling better mental mapping and spatial memory. These complex cognitive processes require a higher cognitive load, indicated by higher theta power in the frontal lobe, associated with better mental mapping and spatial memory. It can also be inferred that the increased theta power at the navigation start and turning points in global strategies account not only for initiating movement and orienting but also for mental mapping and landmarks recalling (e.g., the first turn is right and the second is left). This finding implies that navigation aids incorporating global strategies might be more effective when users need to remember complex routes or navigate large areas.

5 FUTURE WORKS AND CONCLUSION

Our findings provide a promising early indication of the underlying variations in auditory-based spatial representation and motivate further research to explore the interplay between auditory cues and cognitive processing. More behavioural and physiological measures on a larger population can help better understand the relationship between these cognitive processes and guiding strategies. Furthermore, investigating activity in the occipital lobe can provide insights into the cross-modal plasticity of people with blindness and vision impairment.

In conclusion, our study highlights the critical role of theta power in understanding the cognitive processes underlying navigation with auditory stimuli. The consistent increase in theta power at critical navigational points, the higher cognitive load with verbal strategies, and the enhanced recall performance with global strategies all underscore the importance of selecting appropriate auditory stimuli and context for navigation aids. These findings contribute to the broader understanding of how different types of auditory information can support navigational performance, providing valuable insights for designing more effective assistive technologies.

ACKNOWLEDGMENTS

This work was supported in part by Australian Cooperative Research Centres Projects (CRC-P) Round 11 CRCPXI000007, the Australian Research Council (ARC) under discovery grants DP210101093 and DP220100803, Australian National Health and Medical Research Council (NHMRC) Ideas Grant APP2021183, and the UTS Human-Centric AI Centre funding sponsored by GrapheneX (2023-2031).

REFERENCES

[1] Amandine Afonso-Jaco and Brian FG Katz. 2022. Spatial knowledge via auditory information for blind individuals: spatial cognition studies and the use of audio-VR. *Sensors* 22, 13 (2022), 4794.

[2] Oana Bălan, M-Iu Dascălu, Alin Moldoveanu, Simona Caraiman, Florina Ungureanu, Robert Lupu, and Andrei Stan. 2018. Brain activation during virtual and real-world mobility tasks with single and multimodal sensory input provided by an assistive navigational device for visually impaired people. In *2018 25th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS)*. IEEE, 1–6.

[3] Walter F Bischof and Pierre Boulanger. 2003. Spatial navigation in virtual reality environments: an EEG analysis. *Cyberpsychology & behavior* 6, 5 (2003), 487–495.

[4] Chi-Yuan Chang, Sheng-Hsiou Hsu, Luca Pion-Tonachini, and Tzyy-Ping Jung. 2018. Evaluation of artifact subspace reconstruction for automatic EEG artifact removal. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 1242–1245.

[5] Miriam Clemente, Alejandro Rodríguez, Beatriz Rey, and Mariano Alcañiz. 2013. Measuring presence during the navigation in a virtual environment using EEG. *Annual Review of Cybertherapy and Telemedicine 2013* (2013), 136–140.

[6] Alexandre Delaux, Jean-Baptiste de Saint Aubert, Stephen Ramanoël, Marcia Bécu, Lukas Gehrke, Marius Klug, Ricardo Chavarriaga, José-Alain Sahel, Klaus Gramann, and Angelo Arleo. 2021. Mobile brain/body imaging of landmark-based navigation with high-density EEG. *European Journal of Neuroscience* 54, 12 (2021), 8256–8282.

[7] Arnaud Delorme and Scott Makeig. 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods* 134, 1 (2004), 9–21.

[8] Tien-Thong Nguyen Do, Chin-Teng Lin, and Klaus Gramann. 2021. Human brain dynamics in active spatial navigation. *Scientific Reports* 11, 1 (2021), 13036.

[9] Craig Jin, JA Bell, Lil Deverell, Felicity Gates, Ibai Gorodo, S Shayikh, Chin-Teng Lin, Marx Melencio, V Nguyen, A Singh, et al. 2023. Acoustic touch: An auditory sensing paradigm to support close reaching for people who are blind. In *24th International Congress on Acoustics*.

[10] C-T Lin, T-C Chiu, and Klaus Gramann. 2015. EEG correlates of spatial orientation in the human retrosplenial complex. *NeuroImage* 120 (2015), 123–132.

[11] Makoto Miyakoshi, Lukas Gehrke, Klaus Gramann, Scott Makeig, and John Iversen. 2021. The AudioMaze: An EEG and motion capture study of human spatial navigation in sparse augmented reality. *European Journal of Neuroscience* 54, 12 (2021), 8283–8307.

[12] Minh Nguyen, Howe Zhu, Huiyuan Sun, Vincent Nguyen, Lil Deverell, Avinash Singh, Craig Jin, and Chin-Teng Lin. 2023. An evaluation of the presentation of acoustic cues for shorelining techniques. In *2023 Immersive and 3D Audio: from Architecture to Automotive (I3DA)*. IEEE, 1–9.

[13] Minh Nguyen, Howe Zhu, Huiyuan Sun, Vincent Nguyen, Craig Jin, and Chin-Teng Lin. 2023. An evaluation of various spatial audio rendering and presentation techniques to enhance active navigation with sensory augmentation. *The Journal of the Acoustical Society of America* 154, 4_supplement (2023), A196–A196.

[14] Phoebe Peng, Huiyuan Sun, Alexandre Marcireau, Minh Nguyen, Howe Zhu, Chin-Teng Lin, and Craig Jin. 2023. Auditory sensory augmentation to support table tennis games for people with vision loss. *The Journal of the Acoustical Society of America* 154, 4_supplement (2023), A197–A197.

[15] Markus Plank, Hermann J Müller, Julie Onton, Scott Makeig, and Klaus Gramann. 2010. Human EEG correlates of spatial navigation within egocentric and allocentric reference frames. In *Spatial Cognition VII: International Conference, Spatial Cognition 2010, Mt. Hood/Portland, OR, USA, August 15-19, 2010. Proceedings 7*. Springer, 191–206.

[16] Razer. 2023. *RAZER ANZU - SMART GLASSES*. https://www.razer.com/au-en/mobile-wearables/razer-anzu-smart-glasses

[17] Huiyuan Sun, Minh Nguyen, Howe Zhu, Vincent Nguyen, Chin-Teng Lin, and Craig Jin. 2023. A binaural room impulse response dataset and Shorelining psychophysical task for the evaluation of auditory sensory augmentation. *The Journal of the Acoustical Society of America* 154, 4_supplement (2023), A195–A195.

[18] Huiyuan Sun, Howe Y Zhu, Minh TD Nguyen, Vincent Nguyen, Chin-Teng Lin, and Craig T Jin. 2024. From RIR to BRIR: A Sparse Recovery Beamforming Approach for Virtual Binaural Sound Rendering. In *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 1231–1235.

[19] Lore Thaler, Stephen R Arnott, and Melvyn A Goodale. 2011. Neural correlates of natural human echolocation in early and late blind echolocation experts. *PLoS one* 6, 5 (2011), e20162.

[20] Jan-Niklas Voigt-Antons, Zhirou Sun, Maurizio Vergari, Navid Ashrafi, Francesco Vona, and Tanja Kojic. 2024. Impact of spatial auditory navigation on user experience during augmented outdoor navigation tasks. *arXiv preprint arXiv:2404.16473* (2024).

[21] Howe Yuan Zhu and Chin-Teng Lin. 2023. Virtual/augmented/mixed reality technologies for enabling metaverse. *Metaverse Communication and Computing Networks: Applications, Technologies, and Approaches* (2023), 125–155.

[22] Howe Yuan Zhu, Nguyen Quang Hieu, Dinh Thai Hoang, Diep N Nguyen, and Chin-Teng Lin. 2024. A human-centric metaverse enabled by brain-computer interface: A survey. *IEEE Communications Surveys & Tutorials* (2024).

[23] Howe Yuan Zhu, Shayikh Nadim Hossain, Craig Jin, Avinash K Singh, Minh Tran Duc Nguyen, Lil Deverell, Vincent Nguyen, Felicity S Gates, Ibai Gorordo Fernandez, Marx Vergel Melencio, et al. 2023. An investigation into the effectiveness of using acoustic touch to assist people who are blind. *Plos one* 18, 10 (2023), e0290431.