AudioPIN: A Spatial Audio-based PIN-like Authentication System

Ryu Naitou
Keio University
Yokohama, Japan
Keio University
Yokohama, Japan
naitou.ryu.1122@keio.jp

Takashi Amesaka
Graduate School of Engineering Science
The University of Osaka
Toyonaka, Japan
Graduate School of Engineering Science
The University of Osaka
Toyonaka, Japan
t.amesaka.es@osaka-u.ac.jp

Tian Min
Lifestyle Computing Laboratory
Keio University
Yokohama, Kanagawa, Japan
Lifestyle Computing Laboratory
Keio University
Yokohama, Kanagawa, Japan
welkinmin@keio.jp

Hiroki Watanabe
Future University Hakodate
Hakodate, Japan
Future University Hakodate
Hakodate, Japan
hwata@fun.ac.jp

Yuta Sugiura
Keio University
Yokohama, Japan
Keio University
Yokohama, Japan
sugiura@keio.jp

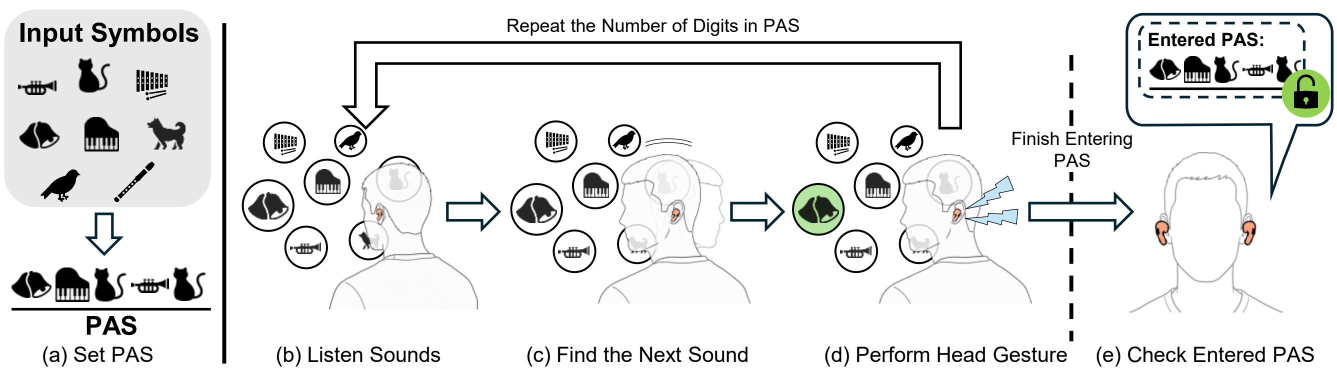

Figure 1: Overview of AudioPIN. Users explore the space-mapped sound sources to locate their designated PAS, then repeatedly aim and input using head tracking.

Abstract

Hearables are expected to evolve into high-performance devices offering functions such as voice assistants and vital sign monitoring. As these functions become integrated into daily life, users access more sensitive information, highlighting the need for secure authentication. While existing methods primarily utilize biometric traits, they often suffer from intra-user variability, environmental susceptibility, and concerns about the sharing of biometric information. To address these issues, we propose AudioPIN, a knowledge-based authentication system for hearables leveraging spatial audio. In AudioPIN, users authenticate by inputting a Personalized Audio Sequence (PAS) by selecting sound sources positioned spatially around them. This study focuses on the auditory interface design at the core of AudioPIN. Specifically, we optimized sound sets, spatial arrangements, and selection areas, and evaluated the system using a prototype. Results confirmed that with an arrangement offering over 10,000 combinations, users completed authentication with 91.0% accuracy in 20.6 seconds.

CCS Concepts

• **Human-centered computing** → **Sound-based input / output**; *Empirical studies in HCI*; • **Security and privacy** → **Authentication**; *Usability in security and privacy.*

Keywords

Earables, Hearables, Audio-based Authentication, Non-visual Interfaces, PIN-like Authentication

ACM Reference Format:

Ryu Naitou, Takashi Amesaka, Tian Min, Hiroki Watanabe, and Yuta Sugiura. 2026. AudioPIN: A Spatial Audio-based PIN-like Authentication System.

CHI EA '26, Barcelona, Spain

ACM ISBN 979-8-4007-2281-3/26/04
https://doi.org/10.1145/3772363.3798635

In *Extended Abstracts of the 2026 CHI Conference on Human Factors in Computing Systems (CHI EA '26), April 13–17, 2026, Barcelona, Spain.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3772363.3798635

1 Introduction

Hearables—earphone-style wearable devices—are increasingly integrating voice assistants and vital-sign sensing [7, 17, 20, 23], allowing them access to sensitive information like personal and health data, which highlights the need for security authentication. Most existing authentication methods for hearables rely on biometric information from built-in sensors, utilizing ear acoustic characteristics [4, 16, 22, 29] and various other traits [8, 10, 15, 18, 30]; however, these methods face challenges when traits change due to injury or illness, or when users avoid storing biometric data for privacy reasons [14].

As an alternative to biometric traits, knowledge-based authentication methods—such as passwords, pattern inputs [1, 26, 34], and image selection [13]—have been proposed and implemented on smartphones and smartwatches. These systems typically rely on touchscreens for user input. However, hearables are generally compact and lack visual display mechanisms, making it difficult to implement such systems without complementary displays.

To address this gap, we propose a novel knowledge-based authentication system, **AudioPIN**, which leverages spatial audio technology and head-tracking features available on hearables. In AudioPIN, users configure a **Personalized Audio Sequence (PAS)** composed of combinations of multiple audio sources. During authentication, they listen to randomly arranged spatial audio sources and identify their locations. The authentication process is completed as the user turns their head toward different spatial audio sources and selects them in the order defined by the PAS. Because AudioPIN does not depend on biometrics or touch-based input, it remains accessible in scenarios where biometric methods are unavailable or inappropriate.

This paper examines the design and evaluation of the auditory interface that underpins AudioPIN. We optimized key design parameters—specifically sound source selection, spatial arrangement, and selection regions—through preliminary investigations to ensure perceptual distinctiveness and usability. Based on this optimized design, we evaluated the system's performance through authentication tasks and usability assessments in a practical setting. We present a knowledge-based authentication system tailored to hearables, which are expected to evolve into personalized multifunctional devices. The contributions of this work are as follows:

- We propose a practical knowledge-based authentication method that can be implemented on hearables without relying on biometric data or touch input.
- We explore the design space of an auditory interface and validate its feasibility through user evaluations in practical scenarios.
- Through the design of an auditory interface tailored for PAS input, we provide new insights into how auditory cues can support multi-layered information presentation, and discuss the broader implications for authentication systems based on auditory and motor interactions.

2 Related Work: Interface Using Spatial Audio Technology

Spatial audio has been widely explored as a channel for conveying information in interactive systems. Prior works have shown its effectiveness for guiding user actions, for example, Xia et al. [31] demonstrated how auditory rather than visual instructions can support limb movements, while Blum et al. [9] designed a navigation system for visually impaired users that relied on spatial cues. Beyond guidance, spatial audio has also been used to enhance situational awareness: Zhao et al. [33] found that audio cues in smart glasses were more intuitive and preferred, and Yang et al. [32] enabled everyday objects to "speak" pseudo-voice announcements, allowing users to perceive their spatial relations.

Building on this foundation, we leverage spatial audio in a different way. Rather than providing guidance or feedback, we investigate how spatially arranged sound sources can themselves become input channels. By addressing this unexplored challenge, we introduce a knowledge-based authentication interface for hearables.

3 Method

Figure 1 illustrates an overview of AudioPIN. During the initial setup, the user creates a PAS by selecting a preferred sequence of sound sources from a candidate set (e.g., xylophone, bell, cat). This set was constructed based on a preliminary survey ($N = 36$) involving user preference ranking and spectral cosine similarity analysis to ensure perceptual distinctiveness. The system stores this sequence as the authentication reference. During authentication, the system spatially arranges the sound source sets using spatial audio technology and plays them sequentially at fixed intervals. The user aligns their head direction with the location of each sound source corresponding to their PAS and confirms the selection using head-related gestures such as tooth clicks or head shakes. This process is repeated for the number of digits in the PAS, completing authentication once the full sequence has been entered.

3.1 Personalized Audio Sequence

The total number of possible combinations (PAS space) is b^d, where b denotes the number of input symbols and d is the number of input digits. A larger PAS space provides stronger security, but also lengthens authentication time, which may hinder usability. Increasing the number of input symbols b enlarges the spatial arrangement and risks reducing localization accuracy, while decreasing the number of input symbols b requires longer sequences d to compensate for equivalent security. To address these trade-offs, we selected input symbol counts ranging from 4 to 7. This range was chosen to ensure reliable spatial localization by avoiding overcrowding, while maintaining a practical sequence length.

To further mitigate observation attacks, the spatial arrangement of sound sources is randomized for each authentication session. This ensures that even when the same PAS is used, the user's head movement patterns vary, making it difficult for unauthorized observers to infer the sequence from motion alone.

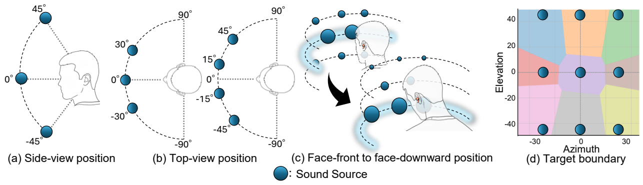

Figure 2: (a) Definition of sound source placement in elevation direction; (b) Definition of sound source placement in the azimuth direction, showing an odd number of sources on the left and an even number on the right; (c) Visual representation of volume changes associated with elevation direction. The sphere size represents the volume; (d) 3×3 grid sound source presentation positions and target boundaries. Different colors are used to visually distinguish the selection regions for each sound source.

3.2 Sound Design

3.2.1 Placement of Sound Sources. In this study, sound source positions are defined as the intersections of latitude and longitude lines on a spherical surface centered around the user's head. Figure 2(a) and Figure 2(b) illustrate an overview of this spatial arrangement. The point where the user's natural forward-facing direction intersects with the spherical surface is defined as the origin. The angular distance between sound sources on the same longitude is represented as the elevation angle, while the distance along the same latitude is represented as the azimuth angle. For the elevation direction, three positions are defined: 0°, 45°, and −45°. When using one row, sources are placed at 0°; when using two rows, they are placed at 0° and −45°. For the azimuth direction, the spacing between adjacent sound sources was set to 30° based on prior studies on spatial audio discrimination [2, 21]. Furthermore, preliminary localization tasks indicated that placing four or more sources in a single horizontal row significantly increased search time. Therefore, we limited the arrangement to a maximum of three sources per row and selected four multi-row arrangements defined by their source-row counts: 4-2, 5-2, 6-2, and 7-3.

3.2.2 Sound Presentation Interval. When multiple sounds are presented simultaneously, sound localization imposes a high cognitive load. Preliminary experiments with four participants showed that localization accuracy dropped below 50% when four or more sources were presented in a horizontal arrangement. Therefore, this study adopts a sequential presentation method, where sound sources are played one after another. For the designed sound set, the maximum decay time—defined as the time for the RMS amplitude (0.01 s moving average window) to decrease to 10% of its maximum value (approx. 0.54 s). Based on this, the minimum interval between sound presentations was set to 0.6 s to avoid acoustic interference between sources. The sound sources are presented in a row-major scanning order: the system traverses each row in turn, and within each row, progresses from left to right.

3.2.3 Sound Localization Assistance via Volume Changes. Users need to efficiently identify the correct sound source. However, with spatial audio reproduction using standard head-related transfer function (HRTF), achieving perfectly accurate sound image localization is difficult. Moreover, human sound localization ability also involves uncertainty, especially in accurately perceiving vertical (elevation) directions of sound sources [25].

To facilitate vertical sound localization, AudioPIN introduces an auditory aid based on volume modulation corresponding to the user's head orientation in the vertical direction. As shown in Figure 2(c), the volume of a sound source is continuously modulated based on the vertical angular difference from the user's head orientation, maximizing when aligned.

3.2.4 Sound Selection Design. Our method requires a dedicated spatial sound selection interface for hearables, consisting of two steps: (1) aiming at the intended sound and (2) issuing the selection command.

Regarding (1) the aiming process, the user's perceived direction of a sound source does not always align with their actual head orientation. To address this, we conducted a pilot study to record users' head orientations when targeting perceived sound sources. Based on the collected data, we defined optimized selection boundaries using a nonlinear SVM (RBF kernel, $C = 1$, $\gamma = 10^{-4}$), ensuring that the selection regions cover 98.4% of the input distribution. Accordingly, we investigated the aiming regions indicated by users in a 3×3 sound source arrangement. Figure 2(d) shows the target boundary of the sound source. In cases where two sources are placed in a row, the central region is evenly divided between the left and right regions.

For (2) command input, possible options include head-related gestures such as tooth clicks [6, 27], facial expressions [3, 11, 28], and head shaking [19]. Although our system can support these inputs, the focus of this study is on how sound presentation and interface design shape user behavior. To avoid confounding effects from gesture recognition, we employed a remote control for input during experiments, ensuring a stable input and enabling a clean evaluation of the auditory design.

4 Evaluation: System Performance

In this experiment, Apple's AirPods Pro 2 were used as the hearables, and the experimental application was implemented using Unity. The audio processing library for spatial audio reproduction employed Resonance Audio [24], utilizing standard HRTF created from the SADIEⅡ database [5]. Head tracking information was acquired via HeadphoneMotion [12], using head posture data from the AirPods Pro's 9-axis sensor.

We implemented a prototype system incorporating the sound source sets, spatial arrangements, and selection regions optimized through the preliminary investigations described in Section 3. We targeted a PAS space ≥ 10,000 (comparable to a 4-digit numeric PIN). The required number of digits was determined by $N^x \geq 10{,}000$, where N is the number of input symbols and x the number of digits (e.g., with $N = 4$, $x = 7$ since $4^7 = 16{,}384$). To examine the effect of digit length, we also evaluated 9-symbol conditions with PAS lengths of 4 to 7 digits. This experiment was approved by the ethics committee of the authors' affiliated institution.

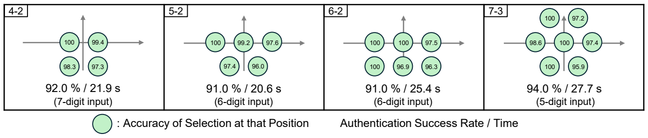

Figure 3: Authentication success rate and authentication time for each sound source arrangement. The accuracy rate for selecting the correct position is indicated within the sound source.

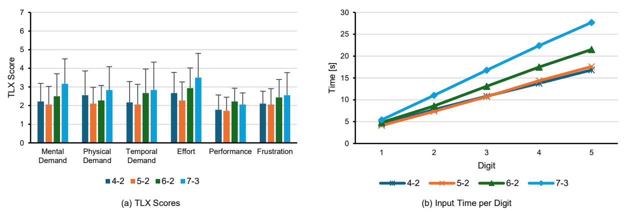

Figure 4: (a) NASA-TLX results (error bars represent the standard deviation), (b) input time required for each digit of PAS at each sound source arrangement.

4.1 Procedure

Participants were first introduced to the system and familiarized with spatial audio and selection regions. In the main experiment, they viewed illustrations of randomly assigned PAS and repeatedly performed localization, aiming, and entering PAS digits one-by-one using a remote button press. Completion time for each PAS entry was recorded. Participants were instructed to continue to the end even if a mistake occurred. Each spatial arrangement was tested five times, followed by the NASA-TLX (using a 7-point Likert scale). Twenty participants (19 males, 1 female; mean age = 23.2, SD = 2.09), eighteen of whom had prior spatial audio experience, took part in the experiment.

4.2 Results

Figure 3 shows the authentication success rate and input completion time, as well as the selection accuracy for each position. Completion time was shortest for 5-2 (20.6 s) and longest for 7-3 (27.7 s). Similarly, accuracy was highest for 7-3 (94.0%) and lowest for 5-2 and 6-2 (91.0%).

The NASA-TLX scores for each arrangement are shown in Figure 4(a). Arrangement 5-2 achieved the best scores (indicating the lowest workload), while arrangement 4-2 performed comparably. Based on these results, arrangement 5-2 was judged highly suitable for balancing usability and time efficiency while ensuring a PAS space larger than 10,000. However, since arrangement 7-3 achieved the highest success rate, it remains a viable option where accuracy is prioritized over speed.

5 Discussion

Overall, there was scope for improvement in the success rate. Since this study employed standard HRTFs, the reduced quality of spatial audio reproduction likely affected performance. As a system improvement, the use of individualized HRTFs should be considered. In addition, personalizing the interface—such as adjusting selection regions based on individual accuracy trends—may enable a more comfortable and accurate system.

In terms of security and efficiency, we examined the relationship between authentication time and the PAS space. Figure 4(b) shows the input completion times up to the required number of digits. The findings suggest that input up to three digits can be completed within approximately 12 s when using 4-6 sound sources. In practical use, this time may be shorter if the PAS is memorized. In such cases, the PAS space ranges from 64 to 216. Although security is reduced, practical deployment may still be feasible by considering constraints on the number of authentication attempts, the shoulder-surfing resistance of our method, trade-offs with the sensitivity of the protected information, and overall usability.

While this study focused on user evaluations of the auditory interface design, realizing a fully practical system requires further investigation into both integration and optimization. First, task performance and usability when integrating head gestures—which are assumed for practical deployment—were not investigated. Future studies should evaluate the system under such conditions and in real-world environments. Second, even within the auditory interface design, there are factors yet to be addressed, such as presentation intervals and error-handling mechanisms. Through comprehensive investigations into these external implementation factors and internal design parameters, the optimal design of knowledge-based authentication systems should be clarified.

6 Conclusion

In this paper, we proposed AudioPIN, a knowledge-based authentication system for hearable devices that leverages spatial audio technology and head tracking. We focused on designing the auditory interface at the core of AudioPIN, optimizing sound source sets, spatial arrangements, and selection regions through preliminary investigations. Through experiments, we confirmed that with a sound source arrangement offering a PAS space greater than 10,000, authentication could be completed with an average accuracy of 91.0% and within 20.6 seconds for arrangement 5-2.

Acknowledgments

This research was partially supported by JSPS KAKENHI (Grant Numbers JP25K03168 and JP23KJ1884); and KGRI Challenge Grant.

References

[1] Yomna Aly, Cosmin Munteanu, Stefania Raimondo, Alan Yusheng Wu, and Molly Wei. 2016. Spin-lock gesture authentication for mobile devices. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct (MobileHCI '16)*. Association for Computing Machinery, New York, NY, USA, 775–782. doi:10.1145/2957265.2961863

[2] M. Alzaher, P. Serrano, J. Tardieu, P. Barone, M. Marx, and P. Nieto. 2021. Contribution of a method of assessing minimum audible angle in headphones. *European Annals of Otorhinolaryngology, Head and Neck Diseases* 138, 5 (2021), 333–336. doi:10.1016/j.anorl.2020.12.006

[3] Takashi Amesaka, Hiroki Watanabe, and Masanori Sugimoto. 2019. Facial expression recognition using ear canal transfer function. In *Proceedings of the 2019 ACM International Symposium on Wearable Computers (ISWC '19)*. Association for Computing Machinery, New York, NY, USA, 1–9. doi:10.1145/3341163.3347747

[4] Takashi Amesaka, Hiroki Watanabe, Masanori Sugimoto, Yuta Sugiura, and Buntarou Shizuki. 2023. User Authentication Method for Hearables Using Sound

Leakage Signals. In *Proceedings of the 2023 ACM International Symposium on Wearable Computers (ISWC '23)*. Association for Computing Machinery, New York, NY, USA, 119–123. doi:10.1145/3594738.3611376

[5] Cal Armstrong, Lewis Thresh, Damian Murphy, and Gavin Kearney. 2018. A Perceptual Evaluation of Individual and Non-Individual HRTFs: A Case Study of the SADIE II Database. *Applied Sciences* 8, 11 (2018). doi:10.3390/app8112029

[6] Daniel Ashbrook, Carlos Tejada, Dhwanit Mehta, Anthony Jiminez, Goudam Muralitharam, Sangeeta Gajendra, and Ross Tallents. 2016. Bitey: an exploration of tooth click gestures for hands-free user interface control. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Florence, Italy) *(MobileHCI '16)*. Association for Computing Machinery, New York, NY, USA, 158–169. doi:10.1145/2935334.2935389

[7] Ananta Narayanan Balaji, Andrea Ferlini, Fahim Kawsar, and Alessandro Montanari. 2023. Stereo-BP: Non-Invasive Blood Pressure Sensing with Earables. In *Proceedings of the 24th International Workshop on Mobile Computing Systems and Applications (HotMobile '23)*. Association for Computing Machinery, New York, NY, USA, 96–102. doi:10.1145/3572864.3580341

[8] Hongliang Bi, Yuanyuan Sun, Jiajia Liu, and Lihao Cao. 2022. SmartEar: Rhythm-Based Tap Authentication Using Earphone in Information-Centric Wireless Sensor Network. *IEEE Internet of Things Journal* 9, 2 (2022), 885–896. doi:10.1109/JIOT.2021.3063479

[9] Jeffrey R. Blum, Mathieu Bouchard, and Jeremy R. Cooperstock. 2012. What's around Me? Spatialized Audio Augmented Reality for Blind Users with a Smartphone. In *Mobile and Ubiquitous Systems: Computing, Networking, and Services*, Alessandro Puiatti and Tao Gu (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 49–62.

[10] Yetong Cao, Chao Cai, Fan Li, Zhe Chen, and Jun Luo. 2024. Enabling Passive User Authentication via Heart Sounds on In-Ear Microphones. *IEEE Transactions on Dependable and Secure Computing* (2024), 1–15. doi:10.1109/TDSC.2024.3429574

[11] Garvit Chugh, Suchetana Chakraborty, and Sandip Chakraborty. 2025. Unlocking Eye Gestures with Earable Inertial Sensing for Accessible HCI. In *2025 17th International Conference on COMmunication Systems and NETworks (COMSNETS)*. 828–832. doi:10.1109/COMSNETS63942.2025.10885707

[12] Anastasia Devana. 2020. Headphone Motion. https://github.com/anastasiadevana/HeadphoneMotion Accessed: 2025-8-23.

[13] Paul Dunphy, Andreas P. Heiner, and N. Asokan. 2010. A closer look at recognition-based graphical passwords on mobile devices. In *Proceedings of the Sixth Symposium on Usable Privacy and Security (SOUPS '10)*. Association for Computing Machinery, New York, NY, USA, Article 3, 12 pages. doi:10.1145/1837110.1837114

[14] Mohamad El-Abed, Romain Giot, Baptiste Hemery, and Christophe Rosenberger. 2010. A study of users' acceptance and satisfaction of biometric systems. In *44th Annual 2010 IEEE International Carnahan Conference on Security Technology*. 170–178. doi:10.1109/CCST.2010.5678678

[15] Yang Gao, Yincheng Jin, Jagmohan Chauhan, Seokmin Choi, Jiyang Li, and Zhanpeng Jin. 2021. Voice In Ear: Spoofing-Resistant and Passphrase-Independent Body Sound Authentication. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 1, Article 12 (March 2021), 25 pages. doi:10.1145/3448113

[16] Yang Gao, Wei Wang, Vir V. Phoha, Wei Sun, and Zhanpeng Jin. 2019. EarEcho: Using Ear Canal Echo for Wearable Authentication. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 3, Article 81 (Sept. 2019), 24 pages. doi:10.1145/3351239

[17] Ghena Hammour, Metin Yarici, Wilhelm von Rosenberg, and Danilo P. Mandic. 2019. Hearables: Feasibility and Validation of In-Ear Electrocardiogram. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 5777–5780. doi:10.1109/EMBC.2019.8857547

[18] Feiyu Han, Panlong Yang, Yuanhao Feng, Haohua Du, and Xiang-Yang Li. 2024. Exploring Earable-based Passive User Authentication via Interpretable In-ear Breathing Biometrics. *IEEE Transactions on Mobile Computing* (2024), 1–17. doi:10.1109/TMC.2024.3453412

[19] Matias Laporte, Preety Baglat, Shkurta Gashi, Martin Gjoreski, Silvia Santini, and Marc Langheinrich. 2021. Detecting Verbal and Non-Verbal Gestures Using Earables. In *Adjunct Proceedings of the 2021 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2021 ACM International Symposium on Wearable Computers* (Virtual, USA) *(UbiComp/ISWC '21 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 165–170. doi:10.1145/3460418.3479322

[20] Keiji Matsumoto, Yuksel Temiz, Hamidreza Taghavi, Elrick L. Cornelius, Hiroyuki Mori, and Bruno Michel. 2019. An earbud-type wearable (A hearable) with vital parameter sensors for early detection and prevention of heat-stroke. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 7049–7055. doi:10.1109/EMBC.2019.8856821

[21] A. W. Mills. 1958. On the Minimum Audible Angle. *The Journal of the Acoustical Society of America* 30, 4 (04 1958), 237–246. arXiv:https://pubs.aip.org/asa/jasa/article-pdf/30/4/237/18737801/237_1_online.pdf doi:10.1121/1.1909553

[22] Yurina Mizuho, Yohei Kawasaki, Takashi Amesaka, and Yuta Sugiura. 2024. EarAuthCam: Personal Identification and Authentication Method Using Ear Images Acquired with a Camera-Equipped Hearable Device. In *Proceedings of the Augmented Humans International Conference 2024 (AHs '24)*. Association for Computing Machinery, New York, NY, USA, 119–130. doi:10.1145/3652920.3653059

[23] Kyle D. Olson, Parker O'Brien, Andy S. Lin, David A. Fabry, Steve Hanke, and Mark J. Schroeder. 2023. A Continuously Worn Dual Temperature Sensor System for Accurate Monitoring of Core Body Temperature from the Ear Canal. *Sensors* 23, 17 (2023). doi:10.3390/s23177323

[24] Resonance Audio Team. 2022. Resonance Audio documentation. https://resonance-audio.github.io/resonance-audio/ Accessed: 2025-8-23.

[25] Riedel, Stefan, Frank, Matthias, and Zotter, Franz. 2025. Effect of HRTFs and head motion on auditory-visual localization in real and virtual studio environments. *Acta Acust.* 9 (2025), 21. doi:10.1051/aacus/2025005

[26] Hai Tao and Carlisle Adams. 2008. Pass-Go: A Proposal to Improve the Usability of Graphical Passwords. *International Journal of Network Security* 7, 2 (Sep 2008), 273–292. doi:10.6633/IJNS.200809.7(2).18

[27] Tomás Vega Gálvez, Shardul Sapkota, Alexandru Dancu, and Pattie Maes. 2019. Byte.it: Discreet Teeth Gestures for Mobile Device Interaction. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI EA '19)*. Association for Computing Machinery, New York, NY, USA, 1–6. doi:10.1145/3290607.3312925

[28] Dhruv Verma, Sejal Bhalla, Dhruv Sahnan, Jainendra Shukla, and Aman Parnami. 2021. ExpressEar: Sensing Fine-Grained Facial Expressions with Earables. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 3, Article 129 (Sept. 2021), 28 pages. doi:10.1145/3478085

[29] Zi Wang, Sheng Tan, Linghan Zhang, Yili Ren, Zhi Wang, and Jie Yang. 2021. An ear canal deformation based continuous user authentication using earables. In *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking (MobiCom '21)*. Association for Computing Machinery, New York, NY, USA, 819–821. doi:10.1145/3447993.3482858

[30] Zi Wang, Yilin Wang, and Jie Yang. 2024. EarSlide: a Secure Ear Wearables Biometric Authentication Based on Acoustic Fingerprint. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 8, 1, Article 24 (March 2024), 29 pages. doi:10.1145/3643515

[31] Chengshuo Xia, Tian Min, and Yuta Sugiura. 2024. AudioMove: Applying the Spatial Audio to Multi-Directional Limb Exercise Guidance. *Proc. ACM Hum.-Comput. Interact.* 8, MHCI, Article 244 (Sept. 2024), 26 pages. doi:10.1145/3676489

[32] Jing Yang, Yves Frank, and Gábor Sörös. 2019. Hearing Is Believing: Synthesizing Spatial Audio from Everyday Objects to Users. In *Proceedings of the 10th Augmented Human International Conference 2019 (AH2019)*. Association for Computing Machinery, New York, NY, USA, Article 28, 9 pages. doi:10.1145/3311823.3311872

[33] Yuhang Zhao, Elizabeth Kupferstein, Hathaitorn Rojnirun, Leah Findlater, and Shiri Azenkot. 2020. The Effectiveness of Visual and Audio Wayfinding Guidance on Smartglasses for People with Low Vision. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3313831.3376516

[34] Yue Zhao, Zhongtian Qiu, Yiqing Yang, Weiwei Li, and Mingming Fan. 2017. An empirical study of touch-based authentication methods on smartwatches. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers (ISWC '17)*. Association for Computing Machinery, New York, NY, USA, 122–125. doi:10.1145/3123021.3123049