Assisting the Multi-directional Limb Motion Exercise with Spatial Audio and Interactive Feedback

Tian Min
welkinmin@keio.jp
Keio University
Yokohama, Japan

Chengshuo Xia
chengshuoxia@xidian.edu.cn
Guangzhou Institute of Technology, Xidian University
Guangzhou, China
Keio University
Yokohama, Japan

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

ABSTRACT

Guiding users with limb motion exercises can assist them in their physical recovery and muscle training. However, traditional visual tracking methods often require multiple camera angles to help the user understand their movements and require the user to be within the range of the screen. This is not effective for more ubiquitous situations. Therefore, we propose a method that uses spatial audio to guide the user with a multiple-directional limb motion exercise. Depending on the user's perception of spatial audio's direction, we designed feedback to help the user adjust the movement height and complete the move to the designated position. We attached a smartphone to the body limb to capture the user's motion data and generate feedback. The experiment showed that with our designed methods, the user could conduct the multiple-directional limb exercise to the designated position. The spatial audio-based limb's motion exercise system could create a natural, pervasive, and non-visual exercise training system in daily life.

CCS CONCEPTS

• **Human-centered computing → Empirical studies in HCI**; **Interactive systems and tools**.

KEYWORDS

Auditory feedback, motion guidance, exercise, spatial audio

ACM Reference Format:
Tian Min, Chengshuo Xia, and Yuta Sugiura. 2023. Assisting the Multi-directional Limb Motion Exercise with Spatial Audio and Interactive Feedback. In *Companion Proceedings of the Conference on Interactive Surfaces and Spaces (ISS Companion '23), November 05–08, 2023, Pittsburgh, PA, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3626485.3626542

1 INTRODUCTION

Motion exercise has played an essential role in various fields, including restoring physical function or keeping daily health and fitness. The essential part of a supportive training system is guidance toward the designated posture. Among the ways of guiding, visual information, as the most direct and effective modality for receiving information input, is fully utilized in existing exercise systems. Despite its many advantages, visual information in certain applications is still subject to certain limitations. These may include the imposition of a substantial cognitive load on the user, which is required to process and analyze the visual information on the screen. Additionally, such approaches may require specialized environments and equipment to be effective, limiting their overall practicality and accessibility. Notably, visual information-based approaches are also inaccessible to individuals with visual impairments, underscoring the need for alternative modalities or innovative solutions to address these limitations.

To cope with the issue, systems based on other modalities, such as auditory and tactile, have been extensively investigated. Auditory feedback is commonly used as an enhancement together with other modalities, using the pitch variation to remind the user to focus in Mindless Attractor [2], or combined with a digital human motion in VoLearn [13]. There remains considerable potential for leveraging auditory cues as an independent modality for motion guidance and exercise training.

Designing an exercise training system independent of visual information is challenging due to the complexity of conveying 3D movement information using a single modality. To address this issue, in this demo, we propose using spatial audio to provide multi-directional motion guidance. Our system utilizes interactive auditory feedback based on a smartphone and earphones. With the ability to rely solely on auditory feedback information, users would move away from the need for traditional visual feedback and combine relevant motion with their personalized audio. Our system provides a seamless and effective solution for limb-based exercise training at the user end, with the potential for widespread adoption in various healthcare and fitness settings.

2 RELATED WORK

Guiding the user in an exercise motion has attracted intensive attention in HCI field. Most auditory-based exercise training systems utilize sound to support the post-feedback. Such post feedback typically provides the user with an evaluation of their motion and helps the user to improve their movement, e.g., the score in [10], and the emoji in [12]. The verbal information could be a straightforward and effective way to convey the assessment. COPD Trainer [11] introduced a mobile application for chronic pulmonary obstructive

ISS Companion '23, November 05–08, 2023, Pittsburgh, PA, USA

ACM ISBN 979-8-4007-0425-3/23/11…$15.00
https://doi.org/10.1145/3626485.3626542

disease patients' rehabilitation motion. And the auditory feedback was designed to present the error information to the user. Similarly, VoLearn [13] utilized verbal information to offer the improvement strategy regarding motion speed. Casamassima et al. [5] used the audio bio-feedback message to Parkinson's patients to analyze their gait pattern and help improve gait rehabilitation.

For spatial audio, it has been successfully applied to many entertainment systems to increase the immersion and reality of users [6], such as in video games [9], VR games [4]. With pure earphones used for spatial audio, many products have emerged with spatial audio, such as AirPods [1], Microsoft Soundscape [7], and Bose [3], among others. Those products could follow the concept of acoustic augmented reality (AAR), which coordinates the user's movement with spatial audio. Ear-AR [14] introduced an indoor AAR design that integrates head motion with spatial and indoor localization. Spatial audio has been widely used due to its directional characteristic, while its exploration in exercise systems is still limited.

3 DEMO: ASSISTING USERS WITH MULTI-DIRECTIONAL LIMB MOTION EXERCISE WITH SPATIAL AUDIO

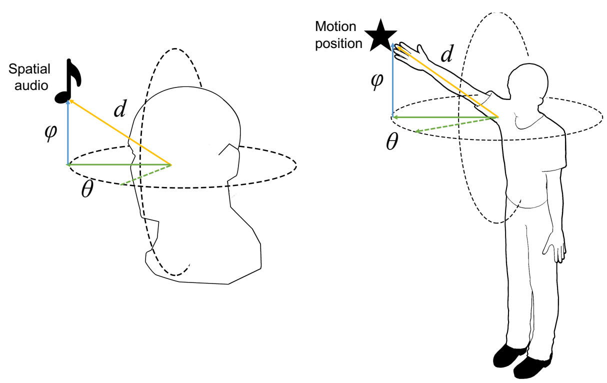

Figure 1: Audio source located in the human head's spherical coordinate, which can also be used to describe the movement of the limbs.

3.1 Background of Spatial Audio

The sense of audio spatiality comes from the temporal and volume difference in how the sound waves interact with human ears: interaural time difference (ITD) and interaural level differences (ILD), helping us determine the azimuth angle of the audio source. The elevation angle of an audio source is difficult to localize due to the symmetric alignment of ears, which produces identical ITDs and ILDs. However, this limitation can be partially overcome by spectral effects [8], which can be simulated using head-related transfer functions (HRTFs)[1].

Spatial audio as the source would provide the basis of directional information to users via the spatial perception ability of each individual. Generally, the position of the audio source can be represented by a spherical coordinate system with the head as the origin: azimuth (θ), elevation (ϕ), and distance (d). Similarly, we could also characterize a body limb motion with the same spherical coordinate system. The user moves the specific body limb to the desired height and direction. The target position could also be represented by (θ, ϕ, d), where d is the length of the body limb. (Figure 1).

3.2 Guiding Exercises with Spatial Audio and Interactive Feedback

In the proposed demo, we develop a virtual environment from Unity3D, where a *virtual smartphone* synchronizes the gyroscope sensor data of the real smartphone attached to the user's limb. The *virtual smartphone* generates a ray (α) that originates from its bottom and projects in the direction of the limb's movement in the physical world.

Due to the limited ability of users to perceive spatial audio, it isn't easy to control the user's motion to an exact position. Thus, as in Figure 2 (a), we applied a sphere object as the ultimate target motion position (i.e., *target area.*) To assist the user could reach the *target area*, a bigger sphere object (i.e., the *effective area*) was also created as a guide. The size of *effective area* can be downsized via the user's motion. We aimed to form an 'approximate' process in the user-end so that the user could gradually reach the ultimate tiny target area.

To realize this, as in Figure 2 (b), we detected the collision between the ray and the *effective area* in a virtual environment to determine whether the user has reached the area. When the user gets into the *effective area*, the size of the *effective area* would be downsized, and an audio cue 1 could be triggered. The new size of the *effective area* was calculated by the minimum angel δ_{min} between the user's body limb direction and the target's direction during the movement, namely, $s = \sin(\delta_{min}) \cdot d$, where d is the distance between the body and target. If the user's movement over the *effective area*, the system is able to generate an audio cue 2 is produced. And if the user successfully reaches the defined *target area*, an audio cue 3 is played.

Following the designed method, the motion from the user-end could be customized by adjusting the size of both the effective area and the target area via the θ and ϕ. It is, therefore, able to control the difficulty of the motion exercise to reach the target position. We believed that this would provide a boarder application range of motion exercises for different individuals, like rehabilitation, anaerobic exercises, joint relaxation, dance motion, etc.

4 EVALUATION

After obtaining consent, 14 participants (5 females and 9 males) who averagely aged 25.6 (SD = 2.6) are invited to investigate the viability of spatial audio guidance. All Participants are right-handed with normal hearing abilities, and each of them are compensated with $15. We took the upper arm limb motion as the object given its widespread application in rehabilitation and weight-training settings.

The experimental task involves touching 6 targets (Figure 3 (b)), appearing in a predetermined randomized order, positioned in front

[1]In this paper, Unity and Resonance Audio (Google Inc.) are used as the audio rendering engine.

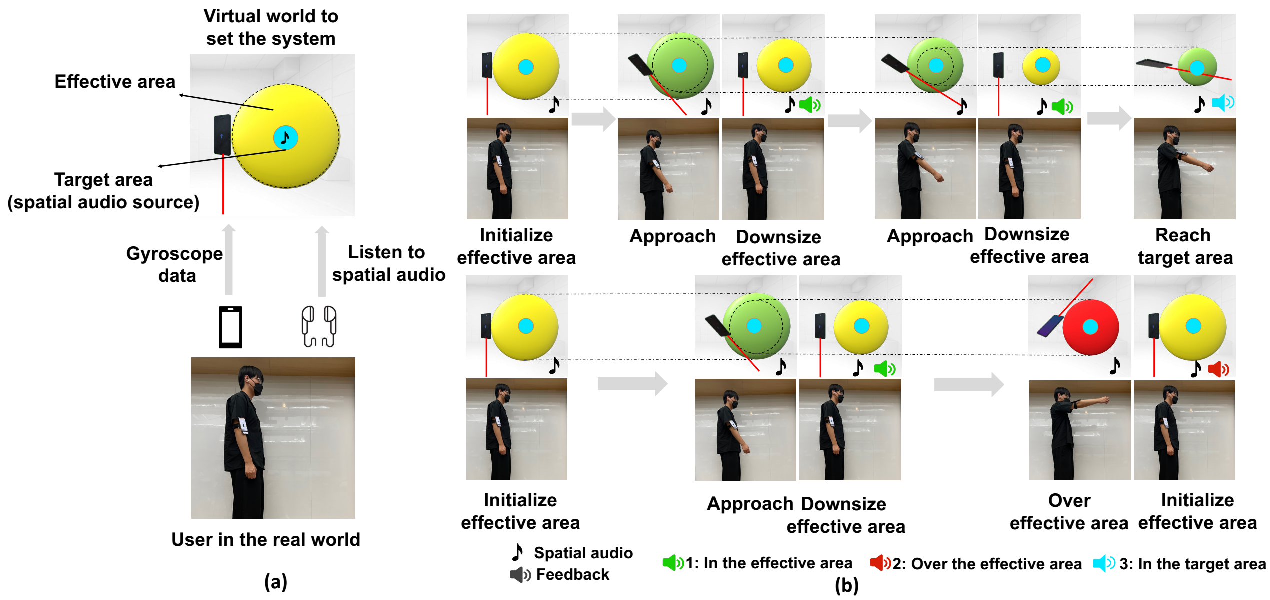

Figure 2: The principle of proposed method using spatial audio and interactive feedback to guide a multi-directional limb motion. (a) An illustration of real and virtual system configuration. (b) The process of leading the user motion exercise with feedback.

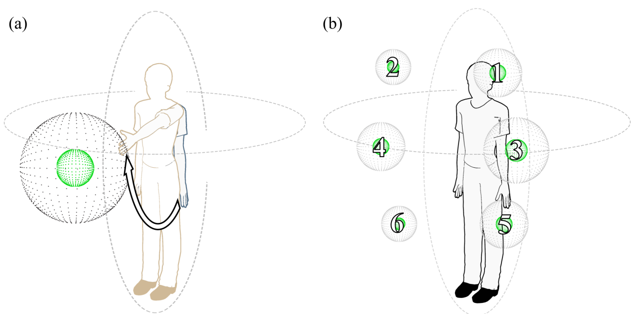

Figure 3: (a) Cyclic multi-directional motion; (b) 6 positions of targets.

of the user. The target locations were specifically designed to align with the range of motion of the shoulder joint. The user interacts with the system by executing motions of raising and lowering the hand (Figure 3 (a)). Upon a successful trial, it automatically advances to the next target. During the experiment, a smartphone is attached to the user's right upper arm to capture data. Before the commencement, a preliminary trial was conducted with each participant to ensure familiarity with the system. A music clip is used as the spatial audio source [2].

Attempts where the angle error between the user's arm and the target was less than 20 degrees (i.e., the *target area* range), were considered successful. And the effective area was ± 42 degrees for θ and ϕ. The results indicate that, on average, **5.26** attempts (variance = 25.94) were required to reach the correct target position; the average error when the user successfully touched the target was **9.76** degrees (variance = 25.98), irrespective of the target location. Thus, through our developed method, the users could reach the designated motion position successfully.

[2]We used the audio from this video: https://www.youtube.com/watch?v=Opp9nqiN5m0.

5 DISCUSSION

In our test, only the left upper arm was selected as the subject, and all the defined target positions were based on the range of motion of the shoulder joint. According to our method, all body's main limbs can be attached by IMU sensor and conduct the motion via the leading of spatial audio Figure 4. However, in addition to each joint should have its related range of motion, the user's perception of moving the corresponding limb to the intended position will be different. There is a deviation in the direction of the user's perception and the direction of the moving limb. The effect of this bias on different limbs is yet to be studied. This will make it necessary to balance the bias effect and the margin of the corresponding target area to enable the user to better exercise training.

In practical applications, we claimed two main advantages of the proposed system, i.e., customized target setting and only the auditory modality employed. Therefore, the system is able to suitable for many exercise-based scenarios, such as rehabilitation, muscle training, and more natural motion exercise. For example, the patient could conduct the motion with different target positions according to the individual's situation. The user could train the muscle in multi-angle with weights. Besides, users can read/watch other content and simultaneously utilize the associated audio for fitness, as in Figure 5.

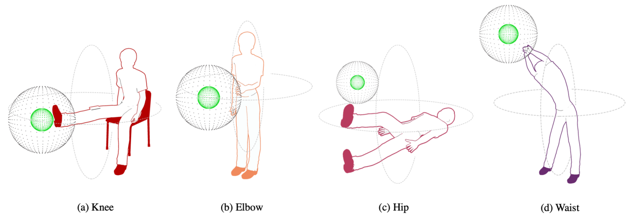

Figure 4: Various body limbs can be trained by spatial audio in a multi-direction.

Figure 5: Potential application scenarios of proposed system.

6 CONCLUSION

In this demo, we presented using spatial audio to assist the users in conducting the limb motion exercise, particularly in a multi-directional limb motion. The method maintains the advantages as it is a non-visual system, and only the audio modality is required, which has high ubiquity and is close to the users' exercise habits. We designed the feedback to help reach the designated height and used spatial audio to provide the directional information. We believed that a spatial-audio-based limb motion exercise method could have more natural, ubiquitous and broader application scenarios.

ACKNOWLEDGMENTS

This work was partially funded by JST PRESTO (Grant Number JPMJPR2134).

REFERENCES

[1] Apple. 2023. Airpods. https://www.apple.com/airpods/.
[2] Riku Arakawa and Hiromu Yakura. 2021. Mindless Attractor: A False-Positive Resistant Intervention for Drawing Attention Using Auditory Perturbation. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, USA, 1–15.
[3] Bose. 2023. Bose. https://www.bose.com/en_us/landing_pages/developer-portal.html.
[4] James Broderick, Jim Duggan, and Sam Redfern. 2018. The importance of spatial audio in modern games and virtual environments. In *2018 IEEE Games, Entertainment, Media Conference (GEM)*. IEEE, 1–9.
[5] Filippo Casamassima, Alberto Ferrari, Bojan Milosevic, Laura Rocchi, and Elisabetta Farella. 2013. Wearable audio-feedback system for gait rehabilitation in subjects with Parkinson's disease. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*. ACM, New York, NY, USA, 275–278.
[6] Catarina Mendonça and Victoria Korshunova. 2020. Surround sound spreads visual attention and increases cognitive effort in immersive media reproductions. In *Proceedings of the 15th International Audio Mostly Conference*. 16–21.
[7] Microsoft. 2023. Microsoft Soundscape. https://www.microsoft.com/en-us/research/product/soundscape/.
[8] John Middlebrooks and David Green. 1991. Sound Localization by Human Listeners. *Annual review of psychology* 42 (02 1991), 135–59. https://doi.org/10.1146/annurev.ps.42.020191.001031
[9] Konstantin Semionov and Iain McGregor. 2020. Effect of various spatial auditory cues on the perception of threat in a first-person shooter video game. In *Proceedings of the 15th International Audio Mostly Conference*. 22–29.
[10] Maurício Sousa, João Vieira, Daniel Medeiros, Artur Arsenio, and Joaquim Jorge. 2016. SleeveAR: Augmented reality for rehabilitation using realtime feedback. In *Proceedings of the 21st international conference on intelligent user interfaces*. ACM, New York, NY, USA, 175–185.
[11] Gabriele Spina, Guannan Huang, Anouk Vaes, Martijn Spruit, and Oliver Amft. 2013. COPDTrainer: a smartphone-based motion rehabilitation training system with real-time acoustic feedback. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing*. ACM, New York, NY, USA, 597–606.
[12] Milka Trajkova and Francesco Cafaro. 2018. Takes Tutu to ballet: designing visual and verbal feedback for augmented mirrors. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 1 (2018), 1–30.
[13] Chengshuo Xia, Xinrui Fang, Riku Arakawa, and Yuta Sugiura. 2022. VoLearn: A Cross-Modal Operable Motion-Learning System Combined with Virtual Avatar and Auditory Feedback. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 6, 2 (2022), 1–26.
[14] Zhijian Yang, Yu-Lin Wei, Sheng Shen, and Romit Roy Choudhury. 2020. Ear-ar: indoor acoustic augmented reality on earphones. In *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*. 1–14.

Received 2023-08-15; accepted 2023-09-15