International Journal of Communication Networks and Information Security
2024, 16(4)
ISSN: 2073-607X,2076-0930
https://ijcnis.org/
Research Article

Dynamic Sweet Spot Audio System: Enhancing Personalized Audio Signal Through ILD and ITD Techniques

Indra gandhi.K[1] ,Kannan.G[2] , Iniyavan R[3], Jawahar.P.K[4] , Padma Usha.M[5]

[1](https://orcid.org/0000-0002-7299-8565) ,[2](https://orcid.org/0000-0001-9060-1317)
[3](https://orcid.org/0009-0007-5359-0971) [4](https://orcid.org/0000-0003-3559-178X)
[5](https://orcid.org/0000-0002-0174-4250)

[1,2,3,4,5] Department of Electronics and Communication Engineering, B.S.Abdur Rahman Crescent Institute of Scienceand Technology, Chennai, India.

*__Corresponding Author:__ kannan@crescent.education

Citation: Indra gandhi.K, "Dynamic Sweet Spot Audio System: Enhancing Personalized Audio Signal Through ILD and ITD Techniques," *International Journal of Communication Networks and Information Security (IJCNIS)*, vol. 16, no 4, 2024. pp.1673-1684.

ARTICLE INFO

Received: 1 Sep 2024
Accepted: 18 Oct 2024

ABSTRACT

The proposed model aims to create a audio system that can adjust the optimal listening area, known as the "sweet spot," to the listener's specific location. The sweet spot is the ideal position where the sound is most balanced and clear. Traditional audio systems have a fixed sweet spot where the audio is heard as intended by the mixer. This new system, however, allows for the adjustment of the sweet spot so listeners can have a personalized listening experience with high-quality sound tailored to their preferences. Instead of moving the speakers, this model directs the audio signals from the speakers towards the listener's position. This approach is innovative, as no existing product on the market offers this capability. The system utilizes ILD (interaural level difference) and ITD (interaural time difference) methodologies to tailor the audio signals to the listener's specific needs.

Keywords: Sweet spot , Spatial adjustment, Audio system, Inter-aural time difference, Inter-aural level difference

INTRODUCTION

In traditional audio systems, optimal sound reproduction is confined to a limited area known as the "sweet spot"—the ideal location where listeners experience the stereo audio mix as intended by sound engineers. Stereo sound is produced through distinct left (L) and right (R) channels, creating a spatial audio field that enhances the listening experience. However, this fixed sweet spot presents significant limitations, as it restricts listener movement and fails to accommodate multiple listeners. Outside this optimal zone, audio quality is often compromised [1,2].To address these challenges, key audio localization techniques such as Interaural Time Difference (ITD), Interaural Level Difference (ILD), and Head-Related Transfer Function (HRTF) are utilized. ITD refers to the time it takes for sound to reach each ear, helping the brain to determine the direction of the sound source. ILD refers to the difference in sound intensity between the ears, allowing for further precision in locating sound sources. HRTF goes a step further, personalizing sound perception by factoring in the way sound is filtered by the shape of the listener's ears, head, and torso [3,4].

This paper presents a novel approach that dynamically adjusts the sweet spot in real-time, matching it to the listener's position within a stereo system. Our system leverages ITD and ILD techniques to adjust the sweet spot based on the listener's position, while HRTF ensures personalized audio perception by accounting for individual anatomical variations. The proposed system considers crucial factors such as high-quality sound sources, room acoustics, and the spatial relationship between speakers and listeners, delivering a personalized and immersive audio experience.

The innovation of our approach lies in its ability to provide high-quality, personalized sound without requiring physical adjustments to speaker placement. By factoring in room acoustics and environmental conditions, our system significantly enhances the listening experience, particularly in challenging acoustic settings [5,6].

This research fills a critical gap in current audio systems by providing a flexible, software-based solution that maintains high audio quality throughout the adjustment process. With an accessible interface, even for non-technical users, the application dynamically fine-tunes the sweet spot, offering high-quality sound tailored to individual preferences and environments [7].

LITERATUREREVIEW

Several researchers have explored methods to adjust the sweet spot in audio systems, each with their own strengths and limitations. A comprehensive review of these approaches highlights the need for our proposed solution.

Merchel et al. [5] introduced a stereophonic playback system that adjusts the sweet spot to the listener's position using an optical face tracker. Their PC-based system, implemented in C++, allows real-time adjustment of loudspeaker signals by tracking the listener's x-y location. While successful in achieving the intended signal shift, this approach faces significant drawbacks. The reliance on a camera module raises privacy concerns and limits its effectiveness in low-light conditions, hampering its commercial viability [6].

Lee et al. [7] presented an alternative approach, designing a stereophonic playback system that cancels cross-talk signals at any listener position using acoustic signals like voice or hand claps. Their system employs a direction of arrival (DOA) algorithm to estimate acoustic source directions. However, this method only identifies the general direction of the sound source, not the precise x-y position, limiting the accuracy of sweet spot localization.

Theoretical foundations for spatial audio processing were explored by Lipshitz et al. [8] and Braasch et al. [9]. They investigated modeling techniques such as Interaural Time Difference (ITD) and Head-Related Transfer Functions (HRTF) for various angles [15,16]. While their work provides valuable theoretical concepts and algorithms, it stops short of presenting complete practical models for sweet spot adjustment.

In a related field, Comanducci et al. [10] developed techniques for generating conditional chord progressions based on harmonic complexity. Using CVAE and RVAE architectures, they demonstrated a correlation between complexity values and perceived complexity. While not directly addressing sweet spot adjustment, their work showcases advancements in conditional audio processing that could inform more sophisticated approaches to spatial audio optimization.

This review of existing literature reveals several key limitations in current sweet spot adjustment techniques. Many systems rely on specialized hardware like cameras, which can compromise privacy and performance, especially in low-light conditions. Acoustic-based methods often lack precise localization, identifying only general sound directions rather than exact positions. There's also a significant gap between theoretical models and practical, user-friendly implementations, with many approaches remaining in the conceptual stage. Furthermore, there's a limited focus on maintaining audio quality throughout the adjustment process. Our proposed system aims to address these limitations by integrating ITD and ILD techniques into a software-based solution that doesn't rely on specialized hardware. By focusing on high-quality audio processing and user accessibility, we seek to bridge the gap between theoretical advancements and practical applications in dynamic sweet spot adjustment, offering a more versatile and user-friendly approach to personalized audio experiences.

METHODOLOGY

Shift the Sweet Spot

The sweet spot can be shifted using various methods. In our proposed system, we utilize inter-aural time difference (ITD), which refers to the inter-channel time delay, and inter-aural level difference (ILD), which represents the inter-channel amplitude. These two methods enhance sound localization across the entire off-center listening area. The objective is to create a personalized listening position for the user, allowing them to sit in any desired position without concerns about the sweet spot, as it can now be shifted to match their location.

The method comprises two sources responsible for generating the left and right channels in a stereophonic system. The left and right channels are modeled, and binaural localization cues such as inter-aural time differences (ITDs) and inter-aural level differences (ILDs) are estimated and analyzed. The adjustment of the speaker signals is crucial for shifting the sweet spot. Without the adjustment of the loudspeaker signals, there can be no shift in the sweet spot. The sweet spot represents the focal point of the audio signals emanating from the channels. Its size is typically influenced by the speaker's size and amplitude. Shifting such a focal point may introduce noise and distort the audio signals. Conventional speakers are designed to emit audio signals primarily at the focal point. Altering this focal point requires careful attention, as any errors in the computation of the delay or level difference will result in a noisy and distorted output audio[17].

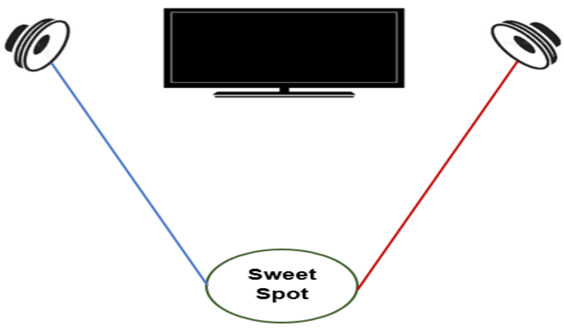

Fig 1. Original Sweet Spot

Figure 1: This illustration depicts the original sweet spot configuration in a traditional stereo audio system. The diagram indicates the ideal listening position (marked as "Initial Position") where the sound is balanced. The shaded area represents the limited zone of optimal audio perception, emphasizing the constraints of fixed speaker placements. To ensure an excellent and immersive audio experience, the sweet spot must be adjusted if the user wants to listen to music from various positions.Figure 2: This figure illustrates the transition to a dynamic sweet spot adjustment within the stereo system. The diagram shows the adjusted listening position (marked as "Adjusted Position"), where the sweet spot has been effectively moved. Arrows indicate the direction of audio signal adjustments, highlighting the improved audio experience in the new position.

Adjusting several audio characteristics, including the listener's position, player settings, volume, and more, is necessary to change the sweet spot. These settings enable users to see changes in real-time and make accurate adjustments to find their ideal sweet spot. Users can visualize the listener's position in a representation of a room, for example, and input their preferences accordingly.By changing the gain and time delay parameters, the sweet spot can be effectively moved. Signal processors (DSPs) or specialized equipment designed for regulating delay and gain can be employed to adjust these parameters. Gain refers to the volume or amplitude of an audio signal, while time delay denotes the length of time it takes for a signal to reach a listener's ears.The sweet spot can be shifted by modifying the gain and time delay of each channel. This method of shifting the sweet spot using stereo channels can greatly improve the listening experience, especially in challenging acoustic conditions. To create a more engaging and entertaining sound stage, adjustments must be made to each channel's gain, phase, time delay, and frequency response. However, it is important to apply these techniques judiciously to avoid creating an unbalanced or unnatural sound stage.

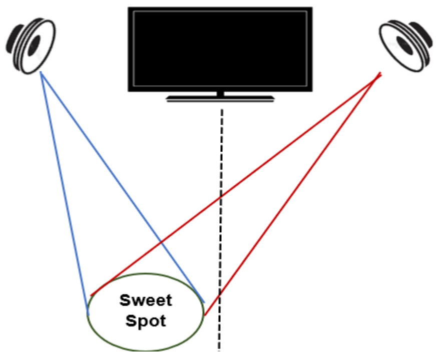

Fig 2. Shifted Sweet Spot

Adjustment of speaker signal

The perceptual shift in sound localization primarily results from the varying delays in speaker signals, which depend on the distances between the speakers and the listener. To ensure simultaneous signal reception at the listener's ears, delays must be adjusted for the specific listening position. Additionally, the amplitudes of the speaker signals are modified to minimize level differences at the listener's location. Accurate calculation of delay and level differences is crucial, as any adjustments can shift the sound away from the intended location. Proper delay calculation enables the creation of a small, position-dependent region within the sweet spot.

The delay is calculated to ensure signals from both speakers reach the center of the listener's head simultaneously. While HRTF (Head-Related Transfer Functions) achieves a similar goal by customizing sound based on individual ear characteristics, our model directs audio signals to any user in a specific position rather than tailoring them to a single person. This allows anyone in the target position to experience the optimal sweet spot. HRTF requires unique modeling for each individual, while our approach delivers consistent quality to any listener in the designated area.We adjust stereo signals in WAV files to shift the sweet spot. WAV (Waveform Audio File) is chosen for its superior audio quality and uncompressed format, which ensures accurate sound reproduction. Although WAV files are larger than MP3 files, they are preferred for their high-quality audio in signal processing [18,19].

ARCHITECTURE OF THE PROPOSED SYSTEM

The block diagram provides an overview of the proposed audio system's operation. Accurate estimation of the listener's position and calibration of sound signals are vital for achieving the desired output. This algorithm focuses on enhancing the sweet spot by dynamically shifting it to a location tailored specifically for the user.Once the listener's position is determined, the system calculates the interaural time difference (ITD) and interaural level difference (ILD) for that particular position and applies these values to the audio signals. Figure 3 illustrates the architecture of the proposed system, showcasing key components such as the audio source, signal processing unit, and the listener's position. The diagram emphasizes how the system adapts to changes in listener location using ITD and ILD techniques.

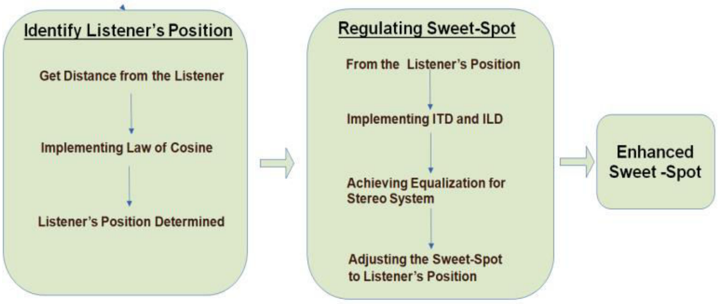

Fig.3 Architecture of Proposed System

ITD and ILD values differ for each listener position within the environment. For every specific location, the corresponding ITD and ILD are calculated and integrated into the audio signals. Precision in these calculations is essential, as any errors may result in inaccurate localization of the sweet spot. Consequently, the final output audio is directed toward the user's position, ensuring that only this specific location receives optimal sound quality.

PROPOSED MODEL

In our model, we utilize interaural time difference (ITD) and interaural level difference (ILD) as methods to improve localization across the entire off-center listening area. ITD refers to the time delay between the arrival of sound at each ear, while ILD refers to the difference in sound level or amplitude between the two ears. By manipulating these two parameters, we can enhance the localization of sound sources, ensuring a more accurate perception of their position, even when the listener is not at the center or sweet spot of the stereo system.

Inter-aural Time Difference (ITD)

ITD (Interaural Time Difference) refers to the difference in time it takes for a sound to reach each ear. It is one of the key cues used by the auditory system to determine the spatial location of sound sources. ITDs arise from the separation of the two ears in space, leading to differences in the path length that sound must travel to reach each ear. By analyzing these time differences, the brain can localize the source of a sound and perceive its direction in space.

$$\mathrm{ITD} = 630 \cos\theta \ \mathrm{microseconds} \qquad (1)$$

Equation 1 represents general ITD at greater than 1m distance (i.e., in terms of speed of sound in air).

$$\mathrm{ITD} = \frac{d}{c}\cos\theta \qquad (2)$$

Equation 2 represents the derived ITD for any distance at any angle upfront. Where,ITD – Inter-aural Time Difference in seconds, d- Distance between the ears (meters), c - Speed of sound in air. θ - Angle of the sound source relative to the listenerand it is obtained from the Law of Cosine.ITD (Interaural Time Difference) measures the time difference for a sound to reach one ear compared to the other. A positive ITD means the sound source is closer to the ear it reaches first, while a negative ITD indicates it's closer to the ear it reaches second. The auditory system uses ITD values to determine sound direction and location.

Accurate ITD reproduction in a stereo system enhances spatial awareness and directionality, creating a more immersive sound experience.

Law of cosines

When the lengths of the other two sides and the included angle of a triangle are known, a mathematical theorem called the law of cosines, commonly referred to as the cosine formula, is used to determine the length of the unknown side. When the side lengths of a triangle are known, it may also be used to calculate the angles of the triangle.
The law of cosines is expressed as follows:

$$c^2 = a^2 + b^2 - 2ab\cos C \qquad (3)$$

Applying the law of cosines to calculate the interaural time difference (ITD) based on an angle is shown in Equation 3. The ITD can be determined using this equation, enabling additional examination and interpretation of sound localization cues.

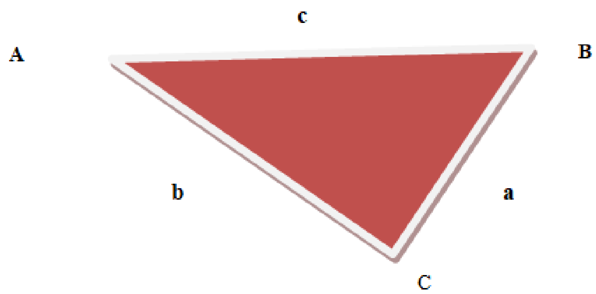

Fig 4. Law of Cosines (Triangle is formed from the User Entered Values and it specifies Listener's position)

Based on the input values given by the user, the position of the listener has been determined in the suggested model using the law of cosines. As seen in Figure 4, this mathematical rule is especially helpful for resolving oblique triangles, which are triangles without a right angle. The rule of cosines is a more broad formula that may be used to compute the sides and angles of triangles with any angle configuration, in contrast to the Pythagorean theorem, which only applies to right triangles.

Inter-aural Level Difference (ILD)

ILD (Interaural Level Difference) refers to the variation in sound intensity between the two ears due to attenuation as sound travels through the skull. For example, a sound coming from the right will be louder in the right ear and quieter in the left ear, and vice versa for sounds from the left. ILD is crucial for binaural hearing as it provides the brain with spatial cues to locate the sound source. The difference in loudness between the ears helps the brain determine the direction of the sound. Equation 4 outlines the method for calculating the gain factor.

$$\frac{x}{x_0} = 10^{\frac{Lx}{20}} \qquad (4)$$

where Lx is Inter-aural Level Difference (ILD) in dB, x0 is the Actual distance of speaker to the Listener's Position and x isdistance between two speakers.The ITD and ILD work together to give the brain the spatial signals required for sound localization. We may create a more immersive and realistic listening experience by faithfully recreating these cues in a stereo system. Binaural audio signals can be produced using the ITD and ILD and moved to the desired place. The audio signals can be precisely positioned for a personalized listener's experience when the ITD and ILD values are precise and devoid of noise or distortion.

RESULTS AND DISCUSSIONS

The paper presents quantitative results, including changes in RMS amplitude and dB levels, which are critical for evaluating the effectiveness of our audio system. However, it's important to interpret these metrics in the context of perceived audio quality and listener satisfaction.

RMS Amplitude Analysis

RMS (Root Mean Square) amplitude provides a measure of the average audio level over time, serving as a more accurate representation of perceived loudness than peak amplitude. While peak amplitude indicates the highest point in an audio signal, it does not reflect the overall listening experience, as it may not account for variations that can lead to listener fatigue or discomfort. By analyzing the RMS values of the original and customized audio tracks, we observe that our system maintains consistent audio levels. This is crucial for preventing sudden volume changes that can disturb the listener. Figures 6, 7, and 8 demonstrate how the RMS values are adjusted smoothly, resulting in a listening experience that is both immersive and comfortable.

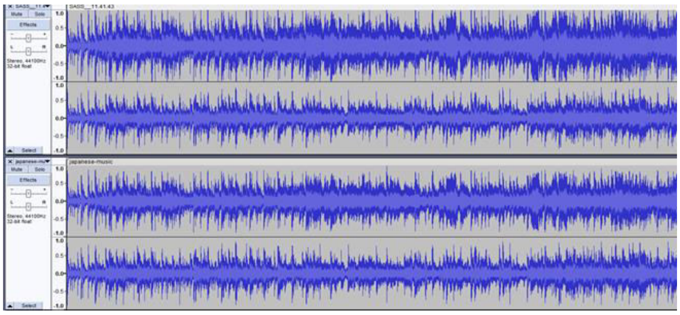

Fig. 6. Audio Processed for Experimental Verification

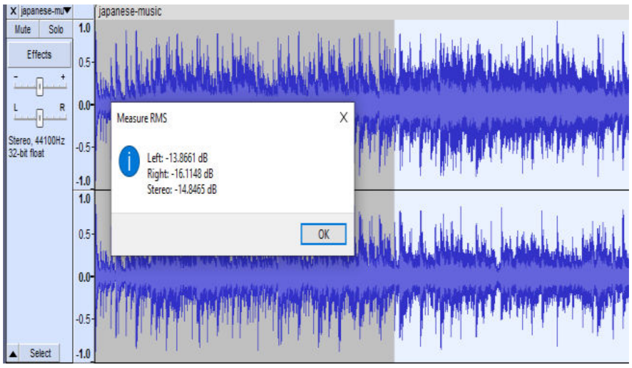

Fig. 7. RMS Value of Original Audio Track

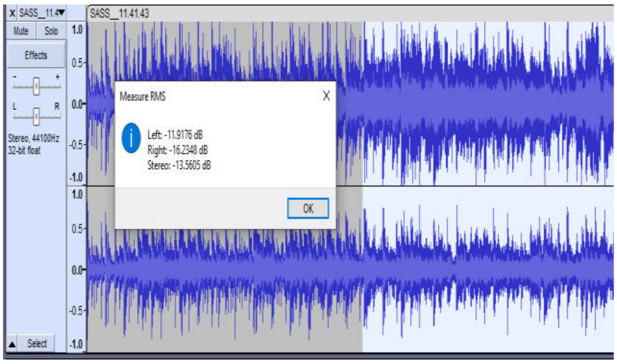

Fig. 8. RMS Value of Customized Audio Track

The customized audio track, which reflects the adjustments made to accommodate the listener's position, shows improved RMS amplitude consistency. This indicates that the system effectively preserves audio quality and listener satisfaction, particularly in environments where traditional fixed sweet spots would result in audio discrepancies.

dB Level Comparison

Table 1 compares the sound levels of the original and customized audio signals for each channel (left, right, and stereo) after applying ITD and ILD techniques. The changes in dB levels highlight the adjustments made to suit the listener's position. The ability to tailor these levels indicates that our system not only enhances the stereo experience but also optimizes spatial audio cues that are essential for sound localization.

Table 1. Comparison of Audio Signals

Audio Track	Amplitude Level of Left Channel (in dB)	Amplitude Level Of Right Channel (in dB)	Amplitude Level of Stereo Channel (in dB)
Original	-10.1613	-10.2285	-10.1948
Customized	-11.9176	-16.2348	-13.5605

By ensuring that the audio track is adjusted with specific gain and delay values, the system creates a balanced sound profile tailored to individual preferences. This personalization is key to improving perceived audio quality, as listeners are more likely to enjoy a sound experience that caters to their specific acoustic environment.

Listener Satisfaction Validation

To further validate the effectiveness of the system, we propose incorporating subjective listener evaluations in future studies. By correlating the quantitative metrics with user feedback on perceived audio quality and satisfaction, we can provide a comprehensive assessment of our system's impact. This dual approach—combining objective measurements with subjective experiences—will strengthen the analysis and demonstrate the real-world applicability of our model.

System Development Overview

The proposed model is developed using Python as the programming language for application development. The graphical user interface (GUI) is created using the "Tkinter" library, which provides the required functionality for window configuration and layout is shown in figure 9. An "Open" button is implemented as a file dialogue box, allowing users to provide input by selecting a file. Once the file is loaded, the user is required to enter the values necessary to shift the sweet spot. The GUI displays an error if the submit button is clicked without entering the required input values. To ensure proper functionality, it is mandatory to load the audio file and enter the values before clicking the submit button.

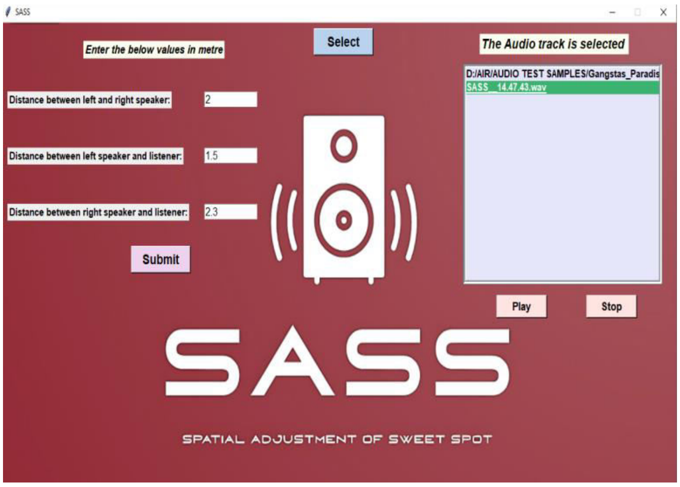

Fig.9. GUI window

Audio Quality Verification

To verify audio quality and accuracy, we utilize Audacity to examine waveforms, spectrograms, clipping, and audio statistics. The steps for analyzing these statistics are outlined in Figure 10. This comprehensive analysis confirms that our system maintains audio fidelity while dynamically adapting to the listener's position.

i. The audio file is opened in Audacity.

ii. The "Plot Spectrum" option is selected from the "Analyze" menu to view the frequency distribution of the audio.

iii. Additionally, the "RMS Amplitude" option is chosen from the "Measure" submenu under the "Analyze" menu to measure the average level of the audio.

iv. The spectrum plot displays the distribution of frequencies in the audio, allowing the spectral characteristics to be assessed.

v. The RMS amplitude measurement provides the average level of the audio signal, giving an indication of its overall loudness.

vi. The audio statistics of the original track can be compared with those of the customized track to evaluate the effectiveness of the sweet spot shifting.

Fig. 10. Steps involved to analyze audio statistics and quality

User Experience and Subjective Testing

While the evaluation of the system is based primarily on objective metrics, such as RMS amplitude and dB level adjustments, incorporating subjective testing would add significant depth to the assessment. By including listener surveys or gathering qualitative feedback from users, we could gain valuable insights into how the adjustments in ITD (Interaural Time Difference) and ILD (Interaural Level

Difference) are perceived in terms of spatial audio quality and overall listening experience. Subjective testing would allow us to better understand listener satisfaction, particularly with regard to how natural the audio adjustments feel and how well the system adapts to varying positions. Such feedback could highlight aspects of the system that objective metrics alone may not capture, providing a more comprehensive evaluation of its performance in real-world scenarios.

PRACTICAL IMPLEMENTATION CHALLENGES

While the proposed system offers a novel approach to dynamically adjusting the sweet spot, there are several practical challenges that must be considered for real-world deployment. Addressing these challenges will provide a more comprehensive understanding of the system's applicability and potential limitations.

Room Acoustics:

One of the most significant factors influencing the effectiveness of ILD and ITD adjustments is the acoustic properties of the room. Reflections, reverberations, and standing waves can interfere with the precise timing and level differences required for accurate sound localization. In environments with poor acoustics, such as rooms with hard surfaces that reflect sound excessively, the system's ability to accurately adjust the sweet spot may be compromised. Effective room treatments, such as acoustic panels or diffusers, may be necessary to optimize sound quality.

Speaker Placement:

The placement of speakers plays a crucial role in delivering balanced sound to the listener. Inconsistent or suboptimal speaker positioning can lead to inaccurate ITD and ILD calculations, reducing the system's effectiveness. For the system to work as intended, it is essential to ensure that the speakers are positioned symmetrically and that the listener is within an appropriate range. Asymmetrical speaker placement or placing them too far apart can cause distortion in the perceived sound field, making it difficult to achieve the desired sweet spot.

Listener Mobility:

One of the core benefits of the system is its ability to adjust the sweet spot dynamically as the listener moves. However, rapid or unpredictable listener movements could pose a challenge. The system must be capable of real-time recalibration to maintain accurate sound localization, especially in environments where the listener moves frequently. If the recalibration process is too slow or lacks precision, the listener may experience audio artifacts or shifts in sound quality.

Environmental Noise:

Noise from external sources, such as HVAC systems, traffic, or background conversations, can interfere with the listener's ability to perceive spatial cues, particularly the subtle differences in ITD and ILD. In noisy environments, the system's effectiveness may be diminished, as the listener may struggle to detect the fine adjustments made to the audio signals. To mitigate this, noise-cancellation technologies or sound-isolation measures could be incorporated to improve the listener's experience.

Computational and Hardware Constraints:

The system's reliance on precise real-time calculations of ITD and ILD means that computational performance is critical. In real-world applications, the system may face challenges due to hardware limitations or delays in processing the required audio adjustments. This is particularly relevant for consumer-grade devices, where processing power may be limited. Future work should explore optimization strategies to ensure the system can operate efficiently on a range of devices without sacrificing performance.

Calibration and User Setup:

In practical scenarios, users may face difficulties during the setup process, particularly when calibrating the system for their specific room and preferences. Providing clear guidelines or automated calibration tools would help users optimize the system without requiring technical expertise. Additionally, recalibration may be necessary if the room layout changes or if new audio equipment is introduced.

CONCLUSION

The proposed model describes the development of an application that can shift the sweet spot of audio signals to the listener's position in a stereo system. The sweet spot for the desired position is achieved by considering all the environmental and room acoustic factors. The proposed application is not a conventional audio player; it is fully capable software that can fine-tune the sweet spot and maximize the listener's listening experience. The developed application is tailored to provide high-quality sound as per the specific needs and preferences of the user. It optimizes the listening experience without compromising the speaker position.In future extensions, we can develop a 5.1-channel system where the five independent channels can be fine-tuned for the respective listener's position

FUTURE WORK

The proposed system effectively adjusts the sweet spot, future research can address its current limitations. A key area is expanding the system to accommodate multiple listeners without compromising audio quality. Additionally, the system is not yet optimized for rapid listener movements, requiring improvements in real-time recalibration and tracking. Extending the system to 5.1 surround sound also presents challenges, particularly in multi-channel setups. Future studies should explore its performance in varied acoustic environments, including those with significant reverberation or noise interference.

ETHICAL DECLARATION

Conflict of interest: No declaration required. **Financing:** No reporting required. **Peer review:** Double anonymous peer review.

REFERENCES

[1] Simon Galvez, Marcos F., Dylan Menzies, Filippo Maria Fazi, Teofilo de Campos, and Adrian Hilton. "A Listener Position Adaptive Stereo System for Object-Based Reproduction." In Audio Engineering Society Convention 138. Audio Engineering Society, 2015.

[2] Owrutsky, Zoe L., Victor Benichoux, and Daniel J. Tollin. "Binaural hearing by the mammalian auditory brainstem: joint coding of interaural level and time differences by the lateral superior olive." Binaural Hearing: With 93 Illustrations (2021): 113-144.

[3] Gutierrez-Parera, Pablo, Jose J. Lopez, Javier M. Mora-Merchan, and Diego F. Larios. "Interaural time difference individualization in HRTF by scaling through anthropometric parameters." Eurasip Journal On Audio, Speech, And Music Processing 2022, no. 1 (2022): 9.

[4] S. Merchel and S. Groth, "Adaptive adjustment of the sweet spot to the listener's position in a stereophonic play back system," in Proceedings of NAG/DAGA, Rotterdam, vol. 5, pp. 255-258, 2009.

[5] Merchel, Sebastian, and Stephan Groth. "Adaptive Adjustment of the "Sweet Spot" to the Listener's Position in a Stereophonic Play Back System–Part 1." In Proceedings of ICA, pp. 1-2. 2009.

[6] Tan, Johann Kay Ann. "Binaural recording methods with analysis on inter-aural time, level, and phase differences." arXiv preprint arXiv:2312.16884 (2023).

[7] G.-T. Lee, S.-M. Choi, B.-Y. Ko, and Y.-H. Park, "HRTF measurement for accurate sound localization cues," KAIST Korea, vol. 6, no. 1, pp. 1-8, 2022.

[8] S. P. Lipshitz, "Stereo Microphone Techniques: Are the Purists Wrong?" Journal of the Audio Engineering Society, vol. 34, no. 9, pp. 718-731, 1986.

[9] J. Braasch, "Modelling of binaural hearing," in Communication Acoustics (J. Blauert ed.), Springer Verlag, vol. 75, pp. 75-108, 2005.

[10] Comanducci, Luca, Davide Gioiosa, Massimiliano Zanoni, Fabio Antonacci, and Augusto Sarti. "Variational Autoencoders for chord sequence generation conditioned on Western harmonic music complexity." EURASIP Journal on Audio, Speech, and Music Processing 2023, no. 1 (2023): 24.

[11] C. Armstrong, L. Thresh, D. Murphy, G. Kearney, A perceptual evaluation of individual and nonindividual HRTFs: A case study of the SADIE II database, Appl. Sci.-Basel 8 (11) (2018) 1–21.

[12] K.-S. Lee and S.-P. Lee, "A Real-time Audio System for Adjusting the Sweet Spot to the Listener's Position," in IEEE Xplore, vol. 18, no. 2, pp. 218-226, 2010.

[13] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," in IEEE Trans. Acoustic., Speech, Signal Processing, vol. 24, no. 4, pp. 320-327, 1976.

[14] K. Iida, "Head-Related Transfer Function and Acoustic Virtual Reality," Springer, Singapore, vol. 1, pp. 1-204, 2019.

[15] M. Zhang, W. Zhang, R. A. Kennedy, and T. D. Abhayapala, "HRTF measurement on KEMAR manikin," in Proceedings of ACOUSTICS, Adelaide, vol. 31, pp. 1-6, 2009.

[16] M. Pollow, K.-V. Nguyen, O. Warusfel, and T. Carpentier, "Calculation of Head-Related Transfer Functions for Arbitrary Field Points Using Spherical Harmonics Decomposition," Applied Sciences, vol. 2, no. 3, pp. 551-566, 2012.

[17] R. Sridhar and E. Y. Choueiri, "A Formula for Low-Frequency Interaural Level Difference," 3D Audio and Applied Acoustics Lab, Princeton University, vol. 1, pp. 1-8, 2020.

[18] B. Bauer, "Broadening the Area of Stereophonic Perception," Journal of the Audio Engineering Society, vol. 8, no. 2, pp. 72-75, 1960.

[19] D. Pressnitzer, C. Suied, S. Shamma, Auditory scene analysis: the sweet music of ambiguity. Front. Hum. Neurosci. (5), 158 (2011)