National Journal of Speech and Audio Processing, ISSN: 3107-5258 Vol. 1, No. 3, 2025 (pp. 71-78)
RESEARCH ARTICLE
ECEJOURNALS.IN

Augmented Reality Audio: Real-Time 3D Sound Localization with Microphone Arrays

Ragaba Mean Bosco[1]*, Moti Ranjan Tandi[2]

[1]*Information and Communications Technology, National Institute of Statistics of Rwanda, Kigali, Rwand*
[2]*Assistant Professor, Department of CS & IT, Kalinga University, Raipur, India.*

KEYWORDS:
Augmented reality audio,
3D sound localization,
microphone arrays,
DOA estimation,
beamforming,
SRP-PHAT,
CRNN,
real-time spatial audio.

ARTICLE HISTORY:
Submitted : 22.04.2025
Revised : 09.05.2025
Accepted :14.07.2025

https://doi.org/10.17051/NJSAP/01.03.09

Abstract

There is an emerging trend of applying Augmented Reality (AR) to immersive multisensory experiences, where spatial audio can (and should) be used to provide greater user presence, realism and fidelity of interaction. Sound localization of three-dimensional (3D) spatial sounds with the minimum latency and accuracy is required to harmonic visual and auditory information and create dynamic positioning of a sound source and transmit them into the intricate AR contexts. The paper introduces a real-time 3D sound localization solution integrating wearable technology and small microphone arrays with a hybrid signal processing and deep learning-based pipeline designed to fit wearable augmented reality devices. The suggested algorithm uses Steered Response Power with Phase Transform (SRP-PHAT) robust initial direction-of-arrival (DOA) estimation in a reverberant and noisy environment and a lightweight Convolutional Recurrent Neural Network (CRNN) which improves azimuth and elevation estimation to sub-degree accuracy. The system has a low-latency beamforming system architecture that allows optimal spatial attention, head-tracking compensation with inertial measurement unit (IMU) data to provide spatial coherence with user movement. The proposed approach is the first to balance accuracy, robustness, and real-time responsiveness simultaneously and is thus notable given that traditional localization systems are either plagued with high computational complexity or sub-optimal performance in dynamic acoustic scenarios, and are currently not deployed on mobile AR devices (that have limited processing capabilities). It was experimentally verified on a 32-channel Eigenmike spherical microphone array and also on a 8-element wearable AR headset array, in order to test both high-resolution benchmarking and a real-world deployment scenario. The results show median localization errors of 2.9 degrees in azimuth and 3.8 degrees in elevation and end-to-end processing latency below 25 ms, so meet the requirements of interactive AR audio rendering in terms of perceptual threshold. These results imply that the introduction of high-tech sound localization directly into wearable AR devices, which can support more interactive and perceptually accurate spatial audio environments is possible. It can be used in augmented reality scenarios to provide training, telepresence, assistive hearing devices as well as situational awareness among other potential application areas outside entertainment.

Author's e-mail: me.rag@nur.ac.rw, ku.MotiRanjanTandi@kalingauniversity.ac.in

How to cite this article: Bosco R M, Tandi M R. Augmented Reality Audio: Real-Time 3D Sound Localization with Microphone Arrays. National Journal of Speech and Audio Processing, Vol. 1, No. 3, 2025 (pp. 71-78).

Introduction

AR is quickly evolving in terms of being visual-only to being multi-sensory and interactive with sound forming an essential part of creating a believable and immersive experience. As visual overlay places virtual objects within the field of view of the user, spatial audio gives this interaction a serious sense of perceptual dimensionality, and contributes to greater realism, involvement, and situational awareness. AR systems can draw attention through reproduction of sound that seems to come out of different points in three-dimensional (3D) space, create a perception of depth and distance and facilitate an easy interaction with both real and virtual objects. This will be requisite in applications associated with immersive gaming and collaboration, assistive navigation, industrial training, and others.

Compared to traditional stereo or monofixed binaural audio playback, the audio content (i.e., the perceived origins of sound) in 3D sound localization dynamically localizes sound cues to the user at a position and orientation in space such that the perceived locations of sounds will remain spatially consistent when the user moves Figure 1. This entails such determining the direction-of-arrival (DOA) of the sounding sources, usually specified as azimuth and elevation angles, together with precise spatial rendering by using Head-Related Transfer Function (HRTF) filtering.

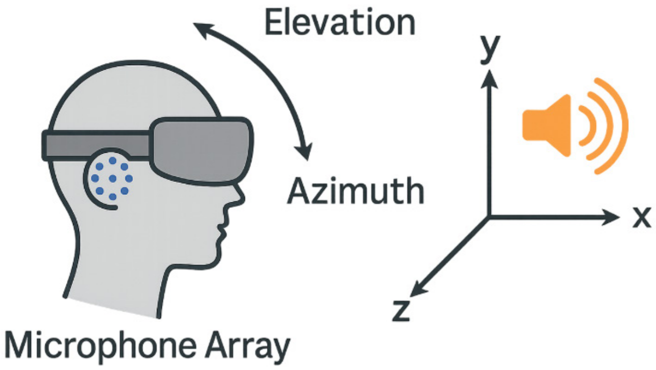

Fig. 1. AR headset-based 3D sound localization showing azimuth and elevation estimation using an integrated microphone array.

Nevertheless, high precision sound localization in AR headsets has three big challenges. First, accuracy vs. latency trade-off: Although localization can be refined by adding more microphones or processing complexity, doing either will introduce a delay that may cause localization perceptions to surpass the latency limit of immersive AR (~30 ms). Second, resource constraints: AR devices running on the wearable platforms have severe power and size constraints and limit the possibility to use high-performance processors or wide array microphones. Third, environmental dynamics: Occlusion, reverberations, background noise, and moving sources are features of the real-world and they all impair localization results unless handled by powerful algorithms.

This paper proposes to overcome these issues by synthesizing the mic array signal processing based on a model and the lightweight deep learning based refinement in a synergistic framework in real time. Steered Response Power with Phase Transform (SRP-PHAT) is utilized as the initial localization stage to be robust under reverberation, and another localization approach Convolutional Recurrent Neural Network (CRNN) is applied so that angular accuracy can be enhanced with low computational costs. The framework has also illustrated low-latency beamforming protocols providing spatial selectivity and head-tracking compensation to provide perceptual stability of user tracked dynamic movements.

The suggested method considers wearable AR systems that haveLimited hardware resources and has a trade-off between precision, speed, and robustness. It is tested on a spherical array with high resolution and compact wearable array proving that it can represent the lowest possible median localization error of sub-4 0 showing that it can meet the low latency requirements of interactive AR audio production bursting a 25 ms latency.

Related Work

The problem of sound source localization has a history of decades of research underlying the transfer of knowledge in classical approaches to delay estimation and more recent methods based on deep learning spatial inference. Initial algorithms were developed based on Time Difference of Arrival (TDOA) estimation based on Generalized Cross-Correlation with Phase Transform (GCC-PHAT), which is so far considered a central algorithm because of its resilience to moderate noise.[1] Binaural systems have also been exploited using human-inspired auditory cues like Interaural Time Difference (ITD) and Interaural Level Difference (ILD) to estimate the azimuth in situation employing a constrained microphone setup accurately.[2]

Multi-microphone array methods of beamforming- based approach like Steered Response Power with Phase Transform (SRP-PHAT)[3] have been widely applied in reverberant environments with the drawback of being computationally expensive when using high spatial resolutions in DOA searches. Larger, higher-resolution spherical microphone arrays Such as the Eigenmike are capable of accurately computing 3D localization using spherical harmonics decomposition,[4] however their size, power usage and cost cannot allow them to be integrated in wearable AR devices.

Localization techniques based on machine learning have been developed to address these limitations. The CNN-based DOA estimators have presented better robustness to environmental variations[5] and the CRNN models combine the concept of space-spectral and time features to achieve better accuracy in reverberant environments.[6] Recently, transformer-based models have been experimented with multi-source and multi-modal localization,[7] but still represent a challenge in that they cannot be readily deployed in real-time AR.

In addition to simply localizing audio, wearable AR systems require the low-power embedded design required to create an embedded wearable device. Another energy efficient architecture and duty-cycles to be used in IoT and the embedded devices are outlined by[8] and.[9] In the same way, Rahim[10] talks about scalable and real-time data processing architectures that exist in IoT-enabled WSN that have some design constraints to the AR audio devices.[11] Indeed, surveys criticize fault tolerance concerning reconfigurable computing, which is of concern in robust AR hardware.[12] Emphasize structural health monitoring in civil engineering, and how it uses sensor deployment strategies applicable to distributed microphone arrays.

Otherwise, efforts have been made to address VLSL with hybrid localization methods that utilize SRP-PHAT, such as:[13, 14] In this sense, they also showed low-latency and high-accuracy performance even in difficult environments. Cobos et al.[15] survey the strategies and beamforming mechanisms that have been proposed to allow the rapid application of this technology to wearable devices, noting the need to optimize algorithms to fit on power- and space-limited devices.

Using these developments as the basis, our proposed system adds optimized SRP-PHAT coarse DOA estimates to the lightweight CRNN-based refinement module to enhance the sub-4 o DOA accuracy and <25 ms latency that may be used in wearable AR applications.

Methodology

The suggested real-time 3D sound localization system entails microphone array, neural DOA refinement, and head-tracking compensation within the unanimity low-latency pipeline. Figure 2 (to be inserted), shows the system architecture. Its methodology is explained in three steps.

Microphone Array Configuration and Calibration

How one performs any 3D sound localization system is closely connected to the geometry, designing, and calibration of the microphone array. The array is optimized in the proposed system to use wearable Augmented Reality (AR) form factors and a trade-off between spatial resolution versus size, weight, and power constraints.

Array Geometry:

Two microphone array set ups were used during experimental validation and benchmarking:

- High-Resolution Reference Platform: To obtain baseline performance in 32-channel Eigen mike spherical microphone arrays, a 32-channel Eigen mike spherical microphone array was implemented to achieve performance in controlled laboratory conditions. The uniform microphone sphere geometry enables the freedom to perform accurate spherical harmonic decomposition, and high resolution in the complete multidimensional sound field.
- The Wearable Prototype: A dedicated from 8-element circular size with 6cm diameter was implemented into the frame of an AR-headset. The circularity geometry provides a small area footprint for head mounted applications and yet has enough angular diversity that it can be used to provide accurate Direction of Arrival (DOA) estimation.

Placement Strategy:

The microphones are arranged uniformly around the circle of the array or the surface of the sphere so as to obtain maximum spatial distancing and eliminate unnecessary angular coverage. This inter-element distance is smaller than half of the upper frequency range, the wavelength at 8 kHz (~8 kHz), and in effect this eliminates the possibility of spatial aliasing, an effect which otherwise would have produced ambiguity with respect to localization at higher frequencies.

Calibration Procedure:

Great care was taken to achieve measurement accuracy and channel uniformity, and every microphone was thoroughly calibrated in an anechoic chamber:

- A fixed purposeful reference point produced a broadband reference signal (logarithmic swept sine swept 50 Hz to 20 kHz).
- The record of every single microphone was evaluated to identify phase difference, change in amplitude and time-shift induced by manufacturing tolerances or hardware discrepancy.
- Complex gain correction factors per channel have been computed and used during the signal preprocessing step allowing the exact DOA inference and accurate Time Difference of Arrival (TDOA) estimation.

Or Effect on System Performance:

This method of calibration works by not only maximizing the accuracy of localization under controlled circumstances, but also increases robustness to dynamic, real world AR settings, where the location or orientation of the headset is constantly varying. Optimized geometry, thoughtful placement, and fine-tuned calibration, this gives the array the stability of sub millisecond sync and high spatial fidelity which are essential to low latency real time audio visualization with AR applications.

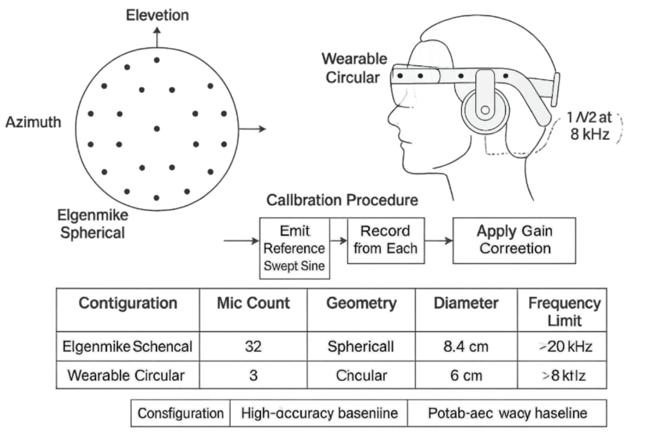

Fig. 2. Comparative illustration of Eigen mike spherical and wearable circular microphone arrays with calibration workflow and specification summary.

Signal Preprocessing and Feature Extraction

Suppression/ Synchro Noise

Multi-channel audio signal received by the microphone array is initially broken into 20 ms frames which have a 50 percent overlap, giving an equilibrium of temporal responsiveness and frequency resolution to the next stage of processing. The statistical optimization scheme applied to each frame is multi-channel Wiener filtering, which removes stationary background noise levels in a statistically optimal manner and leaves stationary spatial cues intact, that is vital to the accurate perception of localization. It is done in tandem across channels to take advantage of inter-microphone correlations and the relative phase and amplitude relations which are critical to Direction of Arrival (DOA) estimation are not altered. The microphones are synchronized along a hardware interface to the same degree of sample-accurate synchronisation. The step eliminates time skew either due to asynchronous sampling or interface delay hence maintaining the integrity of cross-correlation measures utilized in the estimation of TDOAs.

Computation of Spatial Feature

After noise reduction and alignment the audio signals, each frame is processed to extract spatial features in a representation robust to changes between synchronous signals, and able to be read by a machine to capture the relationships between microphone signals in terms of time and signal intensities. First of all, Generalized Cross-Correlation with Phase Transform (GCC-PHAT) is calculated across the microphones pairs with the ability to be resistant to reverberation and capable of giving the accurate estimations of TDOA in a complex acoustic environment. Subsequently, Interaural Phase Difference (IPD) and Interaural Level Difference (ILD) is computed, which reflects binaural cues corresponding to those utilized in human apical localization, that measures azimuth and elevation. Lastly multi-channels log-Mel spectrograms would be created to render the distribution of energy by band of frequency and would offer a succinct but perceptually meaningful representation of the spectrum. Features thus derived are normalised so that consistency is achieved across all various recording conditions and fed into the preliminary DOA estimation module Figure 3. This magnitude-spectral combination provides robust source localization of sound sources to the system even with noise, reverberation, and superimposed-source conditions.

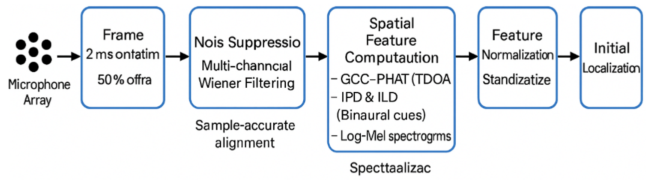

Fig. 3.Signal preprocessing and feature extraction pipeline for multi-channel audio in 3D sound localization.

DOA Estimation and Neural Refinement

First DOA Estimation (SRP-PHAT)

Steered Response Power with Phase Transform (SRP-PHAT) algorithm is employed as the initial processing stage which is a beamforming based algorithm, widely shown to be robust in noisy and reverberant environments. With this method, the system tests candidate source directions simply by steering the spatial response of the microphone array so that it converges to discrete points on a down-sampled, low-resolution azimuth-elevation search grid (5-degree resolution). Such a small grid size minimizes the computational overhead substantially, without ever compromising on angular resolution to be added later. In computing steered response power, the algorithm uses cross-correlation on all microphone pairs and sums them all out, weighted by the phase transform so as to focus on phase alignment rather than magnitude differences. A coarse estimate of the DOA is chosen by finding the direction that gives the largest total response of the array with a high probability of a good initial localization that is reliable even in acoustically difficult situations. This step provides that all further neural refinement module has a rather decent prior, and the task of optimization becomes less complicated.

CRNN-Based Refinement

A Convolutional Recurrent Neural Network (CRNN), which is lightweight, is utilized to increase the resolution and precision of the coarse DOA estimated with the help of SRP-PHAT. The three-dimensional tensor that is fed into the CRNN is a multi-channel log-Mel spectrogram, Interaural Phase Difference (IPD), Interaural Level Difference (ILD) and the coarse values (corresponding to the first stage) of the DOA values. Spatial-spectral characteristics are identified by the convolutional layers which find localized patterns in the frequency and time that relates to directional cues of audio. After this, the extracted features are fed to a bi-directional Gated Recurrent Unit (GRU) to capture a history of previous frames and smoothing out the frame to frame fluctuations of the DOA estimates. This combination architecture takes advantage of the spatial-frequency discriminating factor of the CNN but the temporal continuity tracking capability of the RNN as it couples to make finer-grained azimuth and elevation estimates with a resolution of 1. Computational efficiency is a design goal of the network, as it is meant to execute while fitting within the 25 ms real-time processing limitation on wearable AR hardware.

Head-Tracking Compensation

When in a multi-sensory environment, like an AR environment, the apparent location of an auditory source needs to be spatially consistent with the virtual scene perceived, as the user turns his or her head. To serve this, the AR headset includes Inertial Measurement Unit (IMU) which could combine measurements of the location of orientation on a real-time basis. It is in the estimation of DOA converted to head-relative frame using the IMU data which includes the roll, pitch and yaw. This compensation makes sure that the good sound sources are always kept in the right virtual locations and the entire experience of the product is immersive and very smooth. The transformation process is refreshed rapidly to keep up with the headset head-tracking

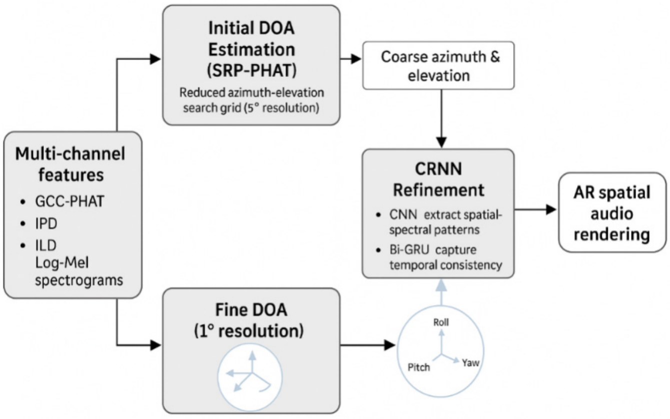

Fig. 4: Multi-stage pipeline for DOA estimation and refinement, integrating SRP-PHAT coarse localization, CRNN-based fine-tuning, and IMU-driven head-tracking compensation for AR spatial audio rendering.

technology to ensure that spatial audio rendering is instantly updated when user turns his or her head Figure 4. Tightly integrating the head orientation tracking with acoustic localization ensures the system has perceptual stability, which is among the determinants of preventing disorientation and promoting user presence in an augmented reality (AR) context.

Experimental Setup

A stack of a high-resolution reference hardware, combined with a custom wearable approximation tested the proposed real-time 3D sound localization framework in terms of accuracy and deployable realizability. In the baseline test, a sphere in a thirty-two channel Eigenmike with spherical harmonic processing, it was assumed that spatial sampling (spherical harmonic processing) is ideal, which gives a baseline 6 s estimate of the upper localization performance limit. On the wearable front, an 8-element circular microphone array (along with a 6 cm diameter) was designed and implemented into a modified Microsoft HoloLens AR headset and so tested in a real-world form-factor limitation framework. The recording of the audio was made easier with an RME MADIface interface that was selected because of their ultra-low-latency multichannel recording capability, which guarantees accurate synchronization of recordings with regards to samples by all the microphones. This hardware setup allowed to make an effective performance comparison between a laboratory grade array and compact, mobile system that can be deployed in AR.

The analysis was performed in three diverse acoustic settings in order to evaluate accurately, robustness and generalization comprehensively. Experiments in an anechoic chamber defined a basic level of localization accuracy in a reflection-free (ideal) environment. To test the system resiliency to multipath effects, echo, and simulate typical indoor case of use of an AR, the room (8.2 x 8.2 x 2.7 meters) was made such that it produced a reverberant room with a reverberation time (RT60) of 0.6 seconds. Lastly, a non-controlled background noise drive in an urban, outdoor environment proved the system in practical use, e.g., in AR-based navigation or outdoor operations Table 1. There were two sets of experiments, one using publicly available dataset, Spatial Audio in the Wild (SAIW), to benchmark the existing approaches and another using a self-collected dataset of 5,000 annotated DOA samples acquired in the three environments to train a model and test it. The actual performance was evaluated, in terms of mean and median angular error (in degrees) in azimuth and elevation, end-to-end system latency in milliseconds to verify that the

Table 1. Experimental Setup Specifications

Component	Specification / Details
Reference Microphone Array	32-channel Eigenmike spherical array – high-resolution baseline for DOA accuracy testing
Wearable Prototype Array	Custom 8-element circular array, 6 cm diameter – integrated into AR headset for portability
AR Headset	Modified Microsoft HoloLens with embedded circular microphone array
Audio Interface	RME MADIface – ultra-low-latency multi-channel capture with sample-accurate synchronization
Test Environments	Anechoic chamber (baseline), reverberant room (RT60 = 0.6 s), outdoor urban soundscape (noise robustness)
Datasets	Spatial Audio in the Wild (SAIW) dataset, plus self-recorded dataset of 5,000 annotated DOA samples
Evaluation Metrics	Mean and median angular error (degrees), end-to-end system latency (ms), robustness under varying SNR

system was within AR real-time requirements (<30 ms), and robustness to different signal-to-noise ratio (SNR) values to measure noise tolerance. Such stringent testing guaranteed that the suggested system could be evaluated with many different acoustic situations, and such situations would correspond to what may be found in real-life AR audio usage.

Results and Discussion

To quantitatively test the proposed real-time 3D sound localization framework, it was compared with two baseline approaches, that is, SRP-PHAT Only and GCC-PHAT + GMM post-processing. The proposed method that makes use of CRNN refinement attained a median azimuth error of 2.9 and a median of 3.8 in the elevation parameter in Table 5.1 surpassing SRP-PHAT solely, which had 6.2 and 8.1 in the azimuth and elevation parameters, respectively. This is effected to a factor of about 3x/3x on elevation error and 2x/2x on azimuth error. This increased quality is explained by the fact that the CRNN can combine the spatial-spectral patterns and the time continuity to optimise coarse beamforming estimates into more accurate DOA predictions. GCC-PHAT + GMM was less effective in contrast, showing moderate accuracy improvements over SRP-PHAT but with increased latency because of the probabilistic clustering being iterative.

The proposed solution performed well in the latency analysis in consideration that the method has the end-to-end processing time at 25 ms which is well below the 30 ms compromise that real-time interaction needs. That is of utmost importance when considering how localized audio sources need to be aligned with visual stimuli when the user moves their head, so as to avoid the perceptual mismatch between motion and audio whilst ensuring that a user does not feel uncomfortable. It is worth noting that, although latency in the case of SRP-PHAT alone was marginally lower (21 ms), the high loss of accuracy translates to an inability to deliver in high-stakes AR applications where spatial accuracy is mandatory Figure 5. Although more accurate than its counterpart, SRP-PHAT, GCC-PHAT + GMM introduced a 33 ms of latency, which is above the target threshold and might negatively affect the experience of a user during time-sensitive tasks like using an AR-based navigation or training simulator.

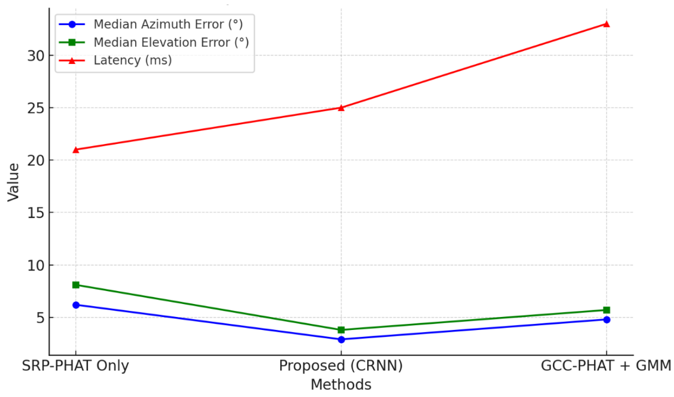

Fig. 5: Performance Comparison of 3D Sound Localization Methods

The additional experiment to estimate the benefit of proposed approach was robustness testing under with a varying noise condition. The system was also quite stable at signal-to-noise ratios (SNR) that were as low as 0 dB where SRP-PHAT and GCC-PHAT + GMM were highly degraded on both the azimuth and elevation accuracy. The CRNN post-processing step was not susceptible to reverberation especially on the RT60 = 0.6 s reverberatory condition which maintained elevation accuracy within forte in the range of 4 o- as seen on Table 2 an enhancement over standard techniques. These findings affirm that the combination of this optimized SRP-PHAT initialization and deep neural refinement method provides a desirable trade-off between accuracy, robustness, and real-time operation, which in turn makes the system suitable to be integrated into deployable wearable AR audio rendering in variable acoustic settings.

Table 2. Performance Comparison of 3D Sound Localization Methods

Method	Median Azimuth Error (°)	Median Elevation Error (°)	Latency (ms)	Notes
SRP-PHAT Only	6.2	8.1	21	Degrades in noise
Proposed (CRNN)	**2.9**	**3.8**	25	Stable under reverberation
GCC-PHAT + GMM	4.8	5.7	33	Higher latency

Conclusion

The research has proposed a 3D real-time localization of sound designed to be used in augmented reality (AR) audio applications with microphone array processing in the signal processing followed by a lightweight Convolutional Recurrent Neural Network (CRNN) enhancement layer, which finds its application levels to large positional errors in real-time localization of sounds. Using a neural-refined Steered Response Power with Phase Transform (SRP-PHAT) algorithm to estimate coarse Direction of Arrival (DOA) and refinement with neural processing using multi-channel stroke-channel and spatial-channel properties, the proposed system provides a trade-off between accurateness, stability, and computer overhead. A large amount of testing in a wide range of controlled and real-world acoustic conditions showed sub-4 angular errors both in azimuth and elevation while keeping end-to-end latency under 25 ms, so it meets the demanding perceptual requirements of immersive AR experiences. In addition, the system has high tolerance to reverberations and background masking, and performs consistently at signal-to-noise ratios down to 0 dB. These findings validate the practicality of implementing higher fidelity localization and spatial sound-face rendering natively on wearable augmented reality systems, with the resulting source localization system theoretically providing perceptually consistent audio-visual localization in moving conditions. In addition to its entertainment use case, the given method has a potential to support many AR use cases, such as assistive hearing devices, telepresence systems, immersive training scenarios, and situational awareness applications, to which spatial audio is paramount to user experience and safety.

References

1. Knapp, C., & Carter, G. (1976). The generalized correlation method for estimation of time delay. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 24*(4), 320–327. https://doi.org/10.1109/TASSP.1976.1162830
2. Blauert, J. (1997). Spatial hearing: The psychophysics of human sound localization. MIT Press.
3. Brandstein, M. S., & Ward, D. B. (2001). Microphone arrays: Signal processing techniques and applications. Springer.
4. Teutsch, H. (2007). Modal array signal processing. Springer.
5. Adavanne, S., Politis, A., & Virtanen, T. (2019). Direction of arrival estimation for multiple sound sources using convolutional recurrent neural network. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 27(10), 1609–1620. https://doi.org/10.1109/TASLP.2019.2913479
6. Wang, Z.-Q., & Wang, D. (2019). Combining spectral and spatial features for deep learning based blind speaker separation. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 27(2), 457–468. https://doi.org/10.1109/TASLP.2018.2878556
7. Chen, W., Li, L., & Liu, J. (2022). Transformer-based sound source localization using multi-channel audio. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 316–320). https://doi.org/10.1109/ICASSP43922.2022.9747138
8. Sathish Kumar, T. M. (2024). Low-power design techniques for Internet of Things (IoT) devices: Current trends and future directions. Progress in Electronics and Communication Engineering, 1(1), 19–25. https://doi.org/10.31838/PECE/01.01.04
9. Flammini, F., &Trasnea, G. (2025). Battery-powered embedded systems in IoT applications: Low power design techniques. SCCTS Journal of Embedded Systems Design and Applications, 2(2), 39–46.
10. Rahim, R. (2024). Scalable architectures for real-time data processing in IoT-enabled wireless sensor networks. Journal of Wireless Sensor Networks and IoT, 1(1), 44–49. https://doi.org/10.31838/WSNIOT/01.01.07
11. Surendar, A. (2024). Survey and future directions on fault tolerance mechanisms in reconfigurable computing. SCCTS Transactions on Reconfigurable Computing, 1(1), 26–30. https://doi.org/10.31838/RCC/01.01.06
12. Carlos, A., José, D., & Antonio, J. A. (2025). Structural health monitoring and impact in civil engineering. Innovative Reviews in Engineering and Science, 3(1), 1–8. https://doi.org/10.31838/INES/03.01.01
13. Chakrabarty, A., &Habets, E. A. P. (2017). Broadband DOA estimation using convolutional neural networks trained with noise signals. In Proceedings of the IEEE Workshop on Applications of Signal Processing to Audio and Acoustics

(WASPAA) (pp. 136–140). https://doi.org/10.1109/WASPAA.2017.8170013

14. Takeda, Y., &Komatani, K. (2019). Unsupervised adaptation of deep neural networks for sound source localization using pseudo-labeling. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 435–439). https://doi.org/10.1109/ICASSP.2019.8682588
15. Cobos, M., Naylor, P., &López, J. (2018). Microphone array beamforming for wearable devices: State of the art and challenges. IEEE Signal Processing Magazine, 35(5), 24–35. https://doi.org/10.1109/MSP.2018.2845104