International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

R-OTT in StereoPilot -An Efficacious Wearable Target Location System for Blind

Athira P[a], Nidhin K V[b]

[a]MTech Scholar Dept.Of ECE NSS College of Engg Palakkad Kerala India
[b]Assistant Professor Dept.Of ECE NSS College of Engg Palakkad Kerala India

ABSTRACT :

StereoPilot is a device designed to support spatial cognition in the BVI (blind and visually impaired). By wearing a head-mounted RGB-D camera, you can measure and process spatial 3Denvironmental information. This is possible thanks to the use of Spatial Sound Processing (SAR) technology, which is dif- ferent from the sound in the instinctive region in humans. StereoPilot offers a new wearable solution for object localization that quickly and intuitively delivers environmental information to individuals in the BVI virtual world. Return 1 to 2 (R-OTT) is a standard used in MPEG surround to combine two channels into one, originally configured as a closed loop. This closed-loop module can reduce the quantization error spatial parameters, thus improving the quality of the synthesized signal. This study demonstrates the effective- ness of SteroPilotin combination with R-OTT, thus being able to reduce signal degradation introduced during the encoding process. The proposed system achieves a different score than experimental audio encodings

Keywords: Spatial Audio, Blind assistance, Sensory feedback

Introduction :

Vision loss significantly hampers object recognition and spatial cognition for individuals with limited sight. The inability to clearly perceive visual cues and spatial relation- ships makes daily tasks challenging. Recognizing objects, navigating spaces, and understanding the layout of the environment become formidable obstacles. Assistive technologies, such as tactile aids or auditory cues, play a crucial role in compensating for these challenges, helping individuals with limited vision enhance their object recognition and spatial awareness to improve their overall independence and quality of life. The Blind or visually impaired people have problems in daily life, such as choosing the same color clothes or finding their favorite products from the market, and they are more likely to fall. The development of machine vision has led to the BVI gaining environmental insights. There are many methods and programs available to assist activities in the BVI. 3D audio technology converts negative sounds into simulated natural sound waves radiating from a point in 3D space. It allows the ear and nerve to be used to allow the brain to place the sound in 3D space after hearing the sound through two headphones.

Spatial audio rendering :

Spatial audio rendering is a process in audio technology that aims to cre- ate a three-dimensional sound experience for listeners, replicating the way humans perceive sound in the real world. This technique goes beyond tra- ditional stereo or surround sound by incorporating additional dimensions such as height, depth, and distance, enhancing the sense of immersion. The rendering process involves the use of sophisticated algo- rithms and signal processing techniques to manipulate audio signals. It considers factors like the listener's position, the location of sound sources, and the acoustic char- acteristics of the environment. By simulating the complex interplay of sound waves, spatial audio rendering can create the illusion of sounds coming from specific directions and distances.

StereoPilot :

The concept of StereoPilot is shown in Figure1. This is a visual light that can identify and locate objects in the environ- ment. Stereo headphones designed for the visually impaired provide information on object recognition and assistance from spatial sound cognition. Object uses a head-mounted RGB- D camera and allows tracking algorithms to generate the 3D position of an object (people) under the control of the camera. Ensuring that the information is consistent with the user's location cognition. .Transfer control from the camera control system to the user coordinate system using an esti- mate of the user's head pose. There is a feedback mechanism to achieve the target position. It can see the sound of the target object in spatial sound or virtual world with virtual sound. Wearing a head-mounted RGB-D camera introduces a novel approach to capturing and processing 3D spatial informa- tion for navigation cues. RGB-D cameras combine traditional color (RGB) imaging with depth (D) sensing capabilities, typically utilizing technologies such as structured light or time-of-flight to measure the distance to objects in the environment. This integration of color and depth information allows for a more comprehensive understanding of the surroundings. As the wearer moves through the environment, the RGB-D camera captures the visual scene in real-time, creating a depth map that represents the distance of objects from the camera. This depth information is crucial for understanding the spatial layout, as it enables the differentiation between objects and surfaces at vary- ing distances. The processed 3D spatial information serves as a foundation for generating navigation cues. Algorithms analyze the depth map, identifying obstacles, pathways, and key environmental features. By interpreting this in- formation, the system can provide navigation cues to the user, guiding them through the environment. These cues may include indicators for obstacles to avoid, directional arrows for navigation, or even augmented reality overlays that enhance the user's perception of their surroundings. This technology finds applications in various fields, such as augmented reality navigation, as- sistance for individuals with visual impairments, or immersive experiences in virtual reality. By leveraging the capabilities of head-mounted RGB- D cameras, the system creates a dynamic and responsive spatial understanding, enhancing navigation by transforming the captured 3D spatial information into actionable cues for the user.

R-OTT :

The R-OTT module outputs two parameters: Channel Level Difference (CLD) and ICC. CLD is the ratio of the power of the first input channel to the power of the second input channel.ICC explains how the two strategies are related. Both CLD and ICC are calculated in a limited way to obtain good coding results.The R-OTT module extracts CLD and ICC as spatial parameters and then down-mixes both input channels into a single down-mix signal. To process more than two audio channels, a number of R-OTT mod- ules are applied in a tree scheme. To process more than two audio channels, a number of R-OTT modules are applied in a tree scheme. The Reverse One-to-Two (R-OTT) module serves a crucial role in the context of MPEG Surround technology, particularly in the process of down-mixing two audio channels into a single channel. This module operates within the framework

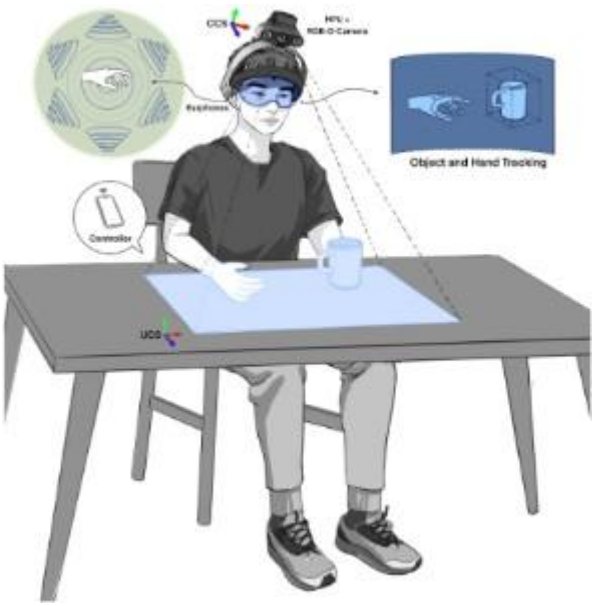

Figure 1: Design concept of stereopilot

of a closed-loop system, indicating a self-contained mechanism where the output of the system influences its further behavior. In the MPEG Surround context, down-mixing is the process of combining multiple audio channels into a single channel, often done to ensure compatibility with systems that support fewer channels. The R-OTT module reverses this down-mixing oper- ation, taking two audio channels and condensing them into a singular output channel. The closed-loop configuration suggests that the module continually adjusts its output based on feedback or input signals, optimizing the down- mixing process dynamically. This adaptability can be crucial for achieving high-quality audio output under varying conditions or in response to changes in the input signal. The Reverse One-to-Two (R-OTT) module in MPEG Sur- round is a closed-loop system designed to efficiently down- mix two audio channels into a single channel. Its closed-loop nature implies a dynamic and adaptive process, potentially enhancing the overall quality and adaptability of the down- mixing operation within the MPEG Surround framework.

Integrating R-OTT into the stereo pilot :

It utilizes advanced spatial parameter quantization tech- niques to en- hance its effectiveness. The device employs a closed-loop system, ensuring real-time adaptation and feed- back, contributing to a more responsive and accurate user experience. Additionally, StereoPilot integrates the Reverse One-to-Two Module (ROTT), a crucial component that plays a pivotal role in reducing quantization errors associated with spatial parameters. The ROTT not only refines the precision of the device but also contributes significantly to the overall quality of synthesized audio signals. By seamlessly incor- porating the ROTT into the StereoPilot framework, we aim to synergize cutting-edge spatial parameter processing with the device's assistive features. This inte- gration is poised to elevate the user experience for individuals with visual impair- ments, offering not only enhanced spatial awareness but also deliver- ing synthesized audio signals of superior quality. The combination of Stere- oPilot and ROTP stands at the forefront of assistive technology, providing an innovative solution for improved navigation and audio perception for the visually impaired. This sophisticated device relies on cutting-edge spatial pa- rameter quantization techniques, aiming to provide an unparal- leled as- sistance experience. The incorporation of the Reverse One-to-Two Module (ROTT) further enhances the device's capabilities. The ROTT serves as a pivotal component, actively working to mitigate quantization errors associ- ated with spatial parameters. This not only refines the precision of the device but also significantly contributes to an elevated quality of the synthesized audio signals generated by StereoPilot.The Reverse One-to-Two Module op- erates as a key player in the closed-loop system, facilitating a dynamic feed- back mecha- nism that continuously refines the spatial parameters. By intel- ligently analyzing and adjusting these parameters in real time, ROTT con- tributes to a more accurate representation of the user's environment. This dynamic adaptation is instrumental in reducing errors and enhancing the overall effectiveness of StereoPilot in aiding navigation and environmental awareness for individuals with visual impairments.Furthermore, the in- cor- poration of ROTT into StereoPilot doesn't merely stop at error reduction; it actively drives improvements in the quality of synthesized audio signals. The module's ability to fine-tune spatial parameters not only enhances ac- curacy but also results in a more immersive and lifelike auditory experience for the user. This holistic approach to spatial awareness and audio synthesis positions StereoPilot as a comprehensive solution for individuals with visual impairments, offering not just assis- tance but an enriched perceptual envi- ronment.The synergy be- tween StereoPilot and ROTT marks a significant stride forward in assistive technology. By seamlessly integrating the spatial refinement capabilities of ROTT with the already advanced features of Stere- oPilot, we create a harmonious fusion of pre- cision and adaptability. Users will experience a newfound level of confidence and independence as they nav- igate the world with enhanced spatial awareness and crystal-clear audio cues. This innovative combination not only addresses the unique challenges faced by individuals with visual impairments but also sets a new standard for the potential of assistive devices in providing a more inclusive and enriched daily experience. In the context of audio processing, especially in technologies like MPEG Surround, a closed-loop module is designed to provide a dynamic and adaptive system. The closed-loop configuration involves a continuous feedback mechanism, al- lowing the system to adjust its parameters based on the output or other relevant signals. This adaptability is particularly bene- ficial for mitigating quantization errors associated with spatial parameters, ultimately enhancing the quality of synthesized audio signals. Quantization errors occur when analog signals are converted into digital format, resulting in a loss of precision due to the finite representation of values. In the case of spatial parameters in audio processing, these errors can affect the accuracy of representing the spatial characteristics of the sound source.

Conclusion :

This paper devises a wearable system for locating targets, aiming to assist individuals with visual impairments in intu- itively understanding and en- gaging with their surroundings. Initially, we presented a method for perceiv- ing environmental information using SAR (Synthetic Aperture Radar).And incor- perating R-OTT module makes the device more effective.The Reverse One to One module in spatial audio coding appears to be a valuable con- tribution to the field. The implementa- tion of this module demonstrates an innovative approach to spatial audio processing, likely enhancing the im- mersive and realistic experience for listeners. The method, as described, seems to offer advantages in terms of perceptual accuracy or computational efficiency. However, further empirical testing and comparison with existing methods would be necessary to fully assess its performance and potential applications. Overall, the module holds promise for advancing spatial audio coding techniques, contributing to the evolving landscape of audio technol ogy. The closed-loop Reverse One-to-Two (R-OTT) module, cou- pled with an algorithm for selecting sub-optimal signals and parameters, represents a significant advancement in audio processing technology. The closed-loop ar- chitecture of the R-OTT module introduces a dynamic and adaptive system, allowing continuous adjustments based on feedback. This feature is particu- larly valuable for minimizing quantization errors associated with spatial pa- rameters, thereby elevating the quality of the synthesized audio signals. The algorithm for selecting sub-optimal signals and parameters complements the closed-loop nature of the R-OTT module by providing an intelligent and effi- cient method for signal processing. By identifying and utilizing sub-optimal signals, the algorithm enhances the adaptability of the module, ensuring op- timal performance under diverse conditions. [1] [2] [3] [4] [5] [6] [7] [8] [9] [10]

REFERENCES:

1. M. Cobos, J. Ahrens, K. Kowalczyk, A. Politis, Data-based spatial audio processing (2022).
2. Elfitri, X. Shi, A. Kondoz, Analysis by synthesis spatial audio coding, IET Signal Processing 8 (1) (2014) 30–38.

3. J. Herre, J. Hilpert, A. Kuntz, J. Plogsties, Mpeg-h 3d audio—the new standard for coding of immersive spatial audio, IEEE Journal of selected topics in signal processing 9 (5) (2015) 770–779.
4. X. Hu, A. Song, H. Zeng, D. Chen, Intuitive environmental percep- tion assistance for blind amputees using spatial audio rendering, IEEE Transactions on Medical Robotics and Bionics 4 (1) (2022) 274–284.
5. X. Hu, A. Song, Z. Wei, H. Zeng, Stereopilot: A wearable target location system for blind and visually impaired using spatial audio rendering, IEEE Transactions on Neural Systems and Rehabilitation Engineering 30 (2022) 1621–1630.
6. K. Nowak, L. Tankelevitch, J. Tang, S. Rintel, Hear we are: Spatial audio benefits perceptions of turn-taking and social presence in video meetings, in: Proceedings of the 2nd Annual Meeting of the Symposium on Human-Computer Interaction for Work, 2023, pp. 1–10.
7. L. Picinali, B. F. Katz, M. Geronazzo, P. Majdak, A. Reyes-Lecuona,
8. Vinciarelli, The sonicom project: Artificial intelligence-driven im- mersive audio, from personalization to modeling [applications corner], IEEE Signal Processing Magazine 39 (6) (2022) 85–88.
9. Rajguru, G. Brianza, G. Memoli, Sound localization in web-based 3d environments, Scientific Reports 12 (1) (2022) 12107.
10. Rinaldi, F. Franchi, A. Marotta, F. Graziosi, C. Centofanti, On the exploitation of 5g multi-access edge computing for spatial audio in cul- tural heritage applications, IEEE Access 9 (2021) 155197–155206.

    Wincott, J. Martin, I. Richards, Spatial awareness: State of the art and future needs of spatial audio journalism (2020).