Comparing Ambisonics Multichannel Reproduction Across Venues

First Author's Name, Initials, and Last name

Second Author's Name, Initials, and Last Name

The study explores how environmental factors and acoustic characteristics affect the reproduction of multichannel systems, offering insights valuable for creating spatial audio artistic works. Tests were conducted using the AUTHOR multichannel system, consisting of 12 speakers, and a 16-speaker system arranged in a rectangular prism at the AUTHOR. Pilot tests refined the algorithm and identified challenges, such as difficulties in horizontal localization and confusion in height perception. Listening tests results showed consistent performance in horizontal localization but challenges in vertical localization, likely due to the natural limitations of human sound perception. Results also showed a minor influence of reverberation time on system performance, highlighting the importance of controlling early reflections for accurate multichannel reproduction. Future developments of the system tested will explore integrating emerging technologies like virtual reality (VR) and mixed reality (MR) to expand spatial audio's applications into art and educational projects.

Keywords and Phrases: spatial audio, listening tests, sound installations and virtual reality

1 INTRODUCTION

The recent increasing demand for immersive and realistic auditory experiences has driven the advancement of spatial audio technologies in recent decades [1]. These technologies enable the recreation of three-dimensional sound environments that immerse the listener in an enveloping and authentic acoustic experience. Among the various available audio options, Ambisonics stands out as a versatile and effective method in capturing, processing, and reproducing spatial audio [1]. The Ambisonics method, a technique for synthesizing and reproducing sound fields, employs an array of microphones to capture sound from all directions around an acoustic source. This enables a precise recreation of the direction and location of sound sources in three-dimensional space [2]. As a playback tool provides a powerful alternative to closed and for-profit audio systems like Dolby Atmos, which lack flexible options for creative and educational applications beyond conventional cinema and entertainment formats. Its flexibility and its compatibility with various loudspeaker configurations, making it suitable for application in a wide range of artistic environments and audio platforms. It’s use in artistic contexts by musicians and sound artists is often constrained by room acoustics and a lack of knowledge by artists of the impact that the early and late reflections in a room can have in the accuracy of the reproduced sound. This often leads to poor presentation of sound installations and concerts in venues at festivals and concerts. This study seeks to offer pertinent insights into how environmental factors and acoustic characteristics influence the perceived reproduction of field recordings in an historic museum site considering technical and architectural limitations.

2 SPATIALISATION ALGORITHM DESIGN

The research presented in this article was conducted as part of the AUTHOR research project [3]. This interdisciplinary endeavor explores the potential of wildlife field recordings as a catalyst for developing site-specific projects that bridge art, science and environmental education. Over the past few years, the project has undertaken various activities, including multichannel concerts, acoustic monitoring, development of machine learning tools, creation of bird identification applications, educational projects, and more recently, virtual reality tools. Immersive audio has been a cornerstone of the project, drawing inspiration from the ISO/DIS 12913 standard for soundscape quantitative and qualitative assessment methods [4].

Building upon a pilot study by AUTHOR, which focused on designing and assessing binaural spatialization techniques using Ambisonics methodology, a spatialization algorithm tailored for multichannel loudspeaker reproduction was developed and implemented [5]. This algorithm, created using the Max/MSP modular programming software [6] in conjunction with the SPAT 5 spatial audio library [7], allows for the encoding, decoding, and reproduction of spatial audio using Ambisonics tools.

3 COMPARED LISTENING EXPERIENCES

Two multichannel loudspeaker configurations were compared in two different acoustic environments. Initially, the system was tested in the main acoustic studio at the AUTHOR, employing a 12-channel loudspeaker system arranged semi-spherically around the listener [5]. In contrast, at the AUTHOR [6], a 16-speaker setup arranged in the shape of a rectangular prism was utilized. This latter configuration was specifically designed to replicate a setup for an upcoming audiovisual art installation to be carried out later this year by the authors at the museum's site.

Figure 2: 12-channel loudspeaker setup in acoustically controlled environment used for the first part of the test. [AUTHOR].

The loudspeaker setup in the laboratory consisted of two vertical levels: one ring at ear level and another ring above the listener's head, with eight and four speakers, respectively, arranged in a circle at the same distance from the listener. This setup was tailored for conducting listening tests for a single person within a confined listening area.

Figure 3: 16-channel loudspeaker setup at the MUSEUM [AUTHOR].

Subsequently, this configuration was adjusted to fit the size and shape of the museum's installation space. Modifications were made to the laboratory loudspeaker setup to better cater to the requirements of the sound installation. These changes included expanding the listening area and incorporating delays to compensate for time differences between speakers and listeners.

4 LISTENING TESTS

To optimize the algorithm and assess the performance of the multichannel system, listening tests were conducted at different stages of the project. Initially, pilot tests were carried out to refine the questionnaire, algorithm design, and other aspects based on feedback from the participants. It's worth noting that although not all participants were experts in the field, they had at least basic knowledge about audio tools and their evaluation. These pilot tests provided valuable insights, such as the challenge in identifying sounds over the horizontal plane, necessary improvements for test accessibility, and the selection of appropriate sound materials. The definitive auditory tests focused on pinpointing sound sources, a critical aspect for spatial creative sound design and evaluating the capabilities of the proposed multichannel system. Given the project's context, bird songs were chosen as source materials due to their availability in soundscapes recordings from wetlands near AUTHOR. A 5-second bird call sound sample was utilized for the tests.

The listening tests were designed with two sections in mind, each comprising a series of source localization tasks for the listener. The main difference between the two test sections lay in the elevation of the virtual sound sources. In the first section, listeners were tasked with identifying sources on a horizontal plane at ear level. In contrast, the second section added the vertical component, requiring participants to indicate both the direction and elevation of each emitted sound.

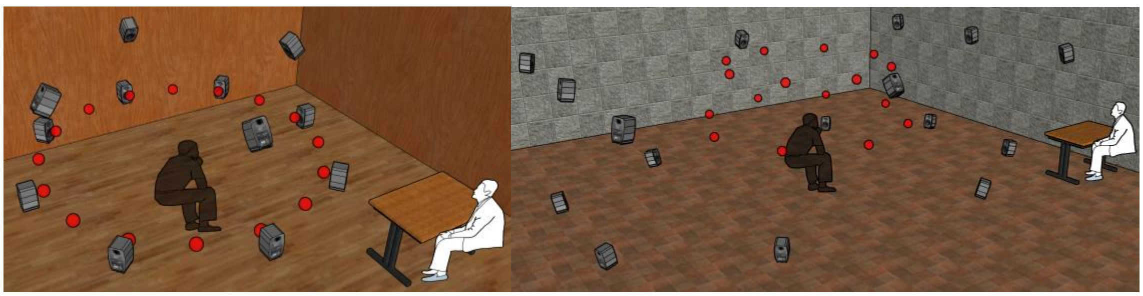

Figure 5: Diagrams showing the setups for both listening tests [AUTHOR].

Figure 5 illustrates the architecture of the spatial reproduction systems and the available options for the various tests. Figure 5a shows the laboratory setup, depicting the positions of the tests on the horizontal plane. Figure 5b showcases the museum's setup, including the positions of the tests with heights. It's worth noting that the positions of both tests were equidistant from the central point of both systems, maintaining consistency in test parameters across different venues.

5 RESULTS

The sample for this experiment comprised 29 students, with 23 participating in both the laboratory and museum tests. This resulted in 528 collected results per test in the laboratory and 552 in the museum, totaling 88 and 92 data points, respectively, per position studied in each test at both locations. These data were processed and organized into various types of graphs to visualize the results more effectively.

The first graph, a violin plot, succinctly presents information about the density of responses within a range of values, allowing for easy comparison between the two situations, which is ideal for representing this comparison. Additionally, a box-and-whisker plot was overlaid to provide insights into data dispersion and the presence of outliers. Polar patterns were created to visually represent the system's performance. Initially, data were categorized into correct or incorrect responses. However, after detailed analysis, a third category called "minor error" was added to provide more detailed and realistic information about the algorithm's performance. This categorization aimed to distinguish almost correct responses from much larger errors, which the initial option grouped together. Two polar plots were proposed to better understand the systems' performance.

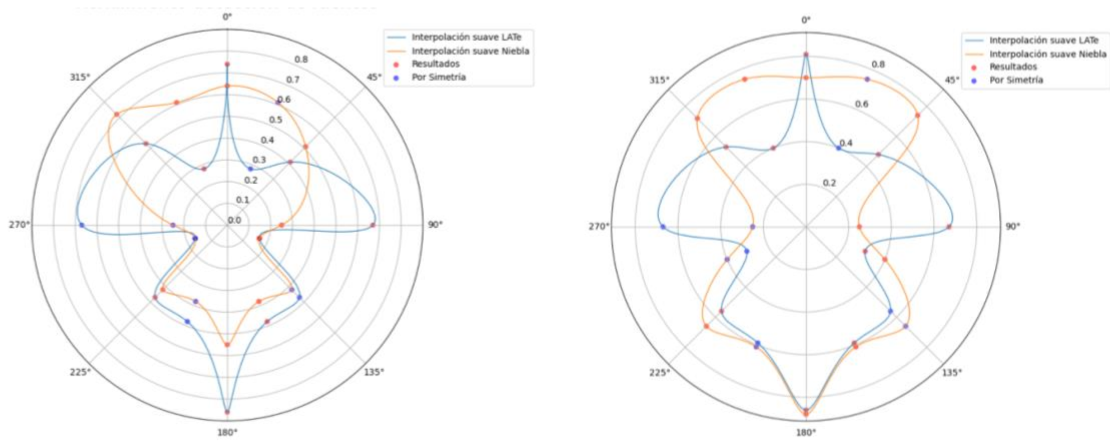

Figure 6: Localisation tests results in the horizontal plane for both venues showing correct responses (left) and correct results within a ±22.5 margin (right) [AUTHOR].

Figure 6 illustrates the comparison of responses between both venues for correct answers. At the rear of the listener, the generated spectrum is quite similar in both cases, with some differences in magnitude. However, significant differences between venues are observed at the front, particularly in the performance improvement at ±45 degrees.

The polar pattern in Figure 6 not only displays correct responses but also errors within a ±22.5 degrees difference from the specific correct position. Thus, a significant improvement is observed compared to the previous graph, mostly maintaining the original pattern. It's noticeable that the sides behind the listener show lower performance in both systems, consistent with the results of previous work.

6 DISCUSSION

To optimize the algorithm and assess the performance of the multichannel system, various listening tests were conducted. Below are the most relevant reflections from the project, presented in sections that encompass the key characteristics of the experiment. This includes comparing both venues to visualize the main differences in the system's performance and their potential causes.

6.1 Influence of speaker system architecture

The difference in system design and loudspeaker configuration between the two locations is evident, influenced by the physical conditions of each space and the intended purpose of the audio reproduction. Consequently, the museum's loudspeaker system is larger and has more loudspeakers than the laboratory's version. While the difference in the number of speakers did not significantly affect the results because the systems were calibrated to reproduce sound at the same level in their respective listening areas, the positioning of the loudspeakers could have influenced participants' perception.

The noticeable differences both at the sides and the front of the system, may have arisen due to differences in speaker density in the two spaces. The frontal area of the museum's system included more speakers, which could have generated a more detailed sound image and thus elicited better responses. In contrast, at the sides of the system, where the speakers were more spaced out, the effect might have been the opposite.

Another characteristic associated with speaker positioning is the listening area, which is primarily affected by the area covered by the speaker array. The greater the distance between the listener and the speakers, the larger the listening area. This aspect was considered when proposing a different system for the exhibition venue tests, aiming to achieve as large a listening area as possible.

6.2 Source localisation

In the horizontal plane tests, there's generally a positive trend towards the correct response in most positions. However, some outliers were observed, with several responses significantly deviating from the correct position.

Many of these deviations are linked to the original position, likely due to confusion caused by symmetrical points in the system.
Additionally, there's evidence of front-back confusion, where the listener's perception of the source's location is distorted when the source is placed symmetrically behind or in front of its original position. This occurs not only at 0° and 180° but also at angles like 157.5° or 112.5°, along with their symmetrical positions in the system.
In the second test, which introduced height as a variable, the results were less consistent, making detailed analysis challenging. This is partly due to the limited response options (head height and above the head) and the difficulty in accurately identifying sources at different heights, influenced by the physiology of the human auditory system. The inclusion of a third height option (ground level) was considered but ultimately discarded due to confusion among pilot test participants. This confusion might be attributed to the naturalness of hearing birds at elevated heights and the subtle differences in sound perception at different elevations.

6.3 Impact of room acoustic characteristics

In this analysis, we focused on comparing specific acoustic parameters critical to multichannel reproduction: early and late reflections within the studied venues. Both locations exhibited considerable early reflections, but there was a notable difference in late reflections, with a 0.4-second discrepancy in reverberation time.
Based on the experiment results and insights from expert Natasha Barrett [8], it was concluded that late reflections had little impact on the Ambisonics system's performance. Despite the disparity in reverberation time between the two venues, the outcomes on-site were largely positive.
In contrast, early reflections are crucial. Reflections occurring within the first 10 milliseconds can significantly confuse listeners [9]. This is likely because these reflections provide "erronous" information, altering the expected auditory perception when reproducing sound sources in the system [10].
The study's findings suggest that venues can indeed influence multichannel reproduction significantly. However, if certain acoustic characteristics can be controlled, multichannel reproduction within the same venue can yield positive results.

7 CONCLUSIONS AND FUTURE WORK

Based on tests conducted in both an acoustic lab and a historic museum, it can be concluded that the spatial performance of a multichannel system using Ambisonics remains relatively stable when transitioning from a controlled to an uncontrolled acoustic environment. However, consistent height localization is challenging in both environments.
Localization tests in both venues demonstrated an acceptable and consistent performance of the spatialization algorithm. However, deficiencies were identified, such as confusion zones and the front-back confusion phenomenon. These issues may stem from inherent limitations of the human auditory system, the architecture of the loudspeaker configuration, involuntary head movements, and the listener's specific position in the room.
Results from tests conducted at various heights were inconclusive. This difficulty may be attributed to human sound perception and our limited ability to accurately locate sources in height, which is often compensated for by involuntary movements.
Regarding the acoustic characteristics of the venues, it is notable that despite the considerable differences in reverberation time between the two locations, there was only a minor influence on the perceived performance of the system. This highlights the importance of controlling early reflections for accurate multichannel reproduction and an enhanced immersive experience.
A new study will explore innovative approaches in spatial audio and multichannel reproduction, focusing on the impact of environments on sound perception. Integrating VR and MR with spatial audio, alongside gamification and visual elements, can effectively enhance immersive auditory experiences and user engagement.
The goal will be to incorporate Ambisonics audio into VR, using Unreal Engine and Max for sound localization tests. Challenges like increasing audio channel precision and platform compatibility will be addressed with hybrid workflows.

Figure 7: Lab tests with virtual reality engine and multichannel loudspeaker system [AUTHOR].

ACKNOWLEDGMENTS

This article acknowledges the support provided by FUNDING INSTITUTION.

REFERENCES

[1] Frank, M. & Zotter, F. (2017) Exploring the perceptual sweet area in Ambisonics. AES (*Audio Engineering Society*). http://www.aes.org/e-lib/browse.cfm?elib=18604

[2] Agrawal, S., M., & Bech, S., S. (2022). Immersion in audiovisual experiences. In Sonic interactions in virtual environments (pp. 319-351). Springer Nature

[3] Wiggins, B. (2021). Ambisonics comes of age. Audio Engineering Society, 69(9), 696–700.

[4] International Standards Organisation, ISO/DIS 12913-1, Acoustics - Soundscape - Part 1: Definitions and conceptual framework, Ginebra: ISO, 2014.

[5] AUTHOR.

[6] *What is Max? | Cycling '74*. (s. f.). https://cycling74.com/products/max

[7] *Spat | IRCAM Forum*. (s. f.). https://forum.ircam.fr/projects/detail/spat/

[8] *Natasha Barrett's website*. (s. f.). https://www.natashabarrett.org/

[9] Strong, J. T. Braasch, J. & Xiang, N. (2010) A Binaural Model with Head Motion That Resolves Front-Back Confusions for Analysis of Room Impulse Responses. AES (*Audio Engineering Society*). http://www.aes.org/e-lib/browse.cfm?elib=15701

[10] ITU (2015). Subjective assessment methods for audio systems.

[11] Rees-Jones, J., & Murphy, D. (2017). A comparison of player performance in a gamified localisation task between spatial loudspeaker systems. *Proceedings of the 20th International Conference on Digital Audio Effects*.

[12] Huisman, T., Ahrens, A., & MacDonald, E. (2021). Ambisonics sound source localization with varying amount of visual information in virtual reality. *Frontiers in Virtual Reality*.