Sweet Area using Ambisonics with Simulated Line Arrays

Patrick Heidegger[1], Benedikt Brands[1], Luca Langgartner[1], and Matthias Frank[2]
[1] University of Music and Performing Arts, Graz, Austria
Email: {patrick.heidegger, benedikt.brands, luca.langgartner}@student.kug.ac.at
[2] Institute of Electronic Music and Acoustics, University of Music and Performing Arts, Graz, Austria, Email: frank@iem.at

Introduction

In Ambisonic reproduction systems, the spatial resolution increases with the number of the applied spherical harmonics, i.e. the Ambisonics order [1]. Hence, the area around the center, in which the reproduction can be considered as physically correct, also increases with the Ambisonics order $r < \frac{N}{2}\frac{c}{f}$ [2]. This criterion would result in considerably small listening areas, as for frequencies up to $f = 3\,\text{kHz}$, an order of $N = 5$ is required for a single human head to fit into the area. In contrast, listening experiments have shown that the perceptual sweet area radius is considerably larger than the physical one [3, 4]. The perceptual sweet area is the area around the central listening position, in which the playback can still be perceived in plausible quality and we define it using two criteria: (i) Unambiguous source localization in a dry-audio scenario, using a single source position. The localization of the source has to be localized from the desired direction with an accuracy of roughly $\pm 10°$. (ii) Listener envelopment in a diffuse scenario, defined as the degree of fullness with auditory images surrounding the listener, i.e., the diffuse sound is noticeable from all directions or loudspeakers [5]. Note that current research found that the contribution to envelopment by elevated directions is minor and that directions at ear height are most important [6].

Whereas the perceptual sweet area for localization increases even for orders $N > 3$, there seem to be other limiting factors for envelopment [4]. Results in [7] indicate that the limitation is due to an uneven level distribution at off-center listening positions. Using normal point-source loudspeakers, which possess a pressure decay reciprocal to their distance $p \propto \frac{1}{r}$, envelopment is hard to achieve for listeners outside the center.

Line arrays consist of multiple stacked loudspeakers with narrow vertical radiation patterns, which are harmonized using wave guides to form a single line-like source. This allows for a homogeneous level distribution over large listening areas [8, 9] and therefore these systems are typically used in sound reinforcement. The application of line-array loudspeaker arrangements in Ambisonics was first investigated in [10] on a 3^{rd}-order horizontal system with 8 line arrays. The informal results indicate that in such a configuration, the perceived plausible listening area was large compared to a system with point-source loudspeakers.

This paper experimentally investigates how the application of line-array loudspeakers for Ambisonic playback effects the size of the sweet area for localization and listener envelopment. For the experiment, the line arrays are simulated by position-dependent level compensation of point-source loudspeakers for a single listener. Thereby, we implemented and compared three compensation methods: (a) full compensation of the $p \propto \frac{1}{r}$ decay, adjusts $+6\,\text{dB}$ per doubling of listener-to-loudspeaker distance, (b) semi-compensation that adjusts $+3\,\text{dB}$ per doubling of distance (this method represents a more realistic scenario, since a completely position-independent level distribution is hard to achieve in practice), (c) no compensation, which constitutes standard point-source loudspeakers. The sweet areas were determined using the method proposed in [4], where listeners were asked to walk towards loudspeaker directions and indicate the largest radius within which the playback was still plausible, cf. Figure 1. The radii were evaluated for localization of a single frontal sound object encoded and played back with 1^{st}, 3^{rd}, and 5^{th} order, as well as for listener envelopment of a diffuse signal in 5^{th} order. The experiments took place at the IEM-CUBE, equipped with a 24-channel hemisphere of point-source loudspeakers and a motion tracking system.

The paper is constructed as follows. We first describe the compensation of the position-dependent levels and its implementation, the playback environment, the creation of the stimuli, and the experimental procedure. Subsequently, the results of the experiment are presented and statistically analyzed. Finally, the conclusion discusses the results, summarizes the paper and provides an outlook for future research.

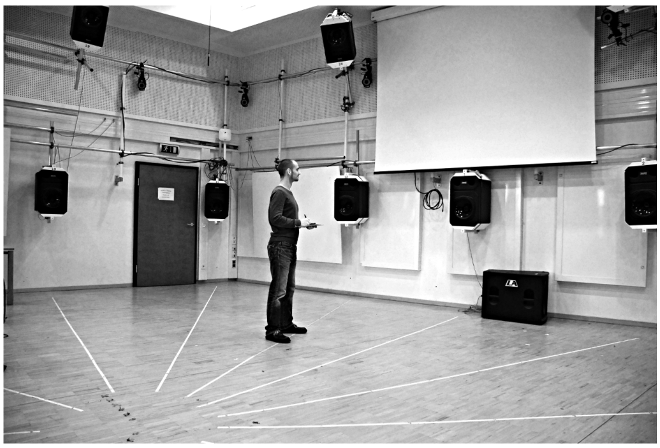

Figure 1: Listener during the listening experiment, using the guidelines on the floor and facing the sound object [4].

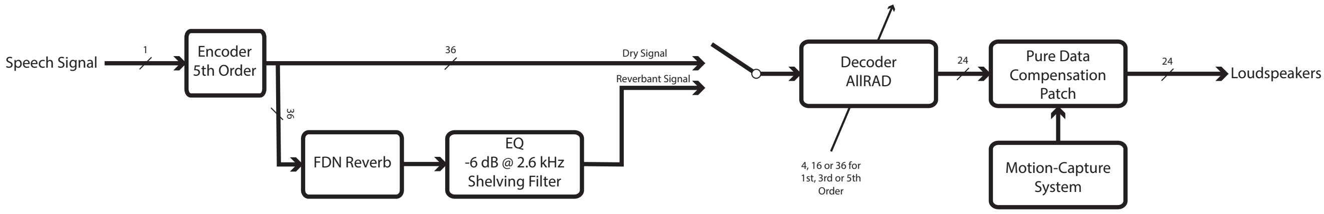

Figure 2: Signal flow of the test signal creation and playback

Experimental Setup

The listening experiment was conducted in the IEM-CUBE (approximately 10 m × 12 m × 5 m) equipped with a hemispherical loudspeaker system and a motion tracking system. The loudspeaker system consists of 24 Tannoy System 1200 coaxial loudspeakers (12 at 0° elevation, 8 at 30° and 4 at 60°) , with a 90° conical dispersion (-6 dB) and thus can be considered point sources with 6 dB decay for every doubling of distance. The motion tracking system, which is linked to the software Motive by OptiTrack, consists of 16 infrared cameras which track the listener's head as a rigid body target.

Level Compensation

To be able to control each loudspeaker's level individually and in real-time, we utilized the graphical programming language Pure Data (Pd). The custom patch we created obtains the position values from the motion tracker and computes the distances from the listener's head to each loudspeaker, as well as to the acoustic center of the system. In order to apply full compensation, which means a +6 dB level increment per doubling of distance, we are then able to compute the linear weights for each loudspeaker as the normalized difference between the listener-to-center-distance and the listener-to-loudspeaker-distance.

$$w_{i,6\mathrm{dB}}(\vec{r}) = \frac{|\vec{r} - \vec{r_i}'|}{|\vec{r_i}'|} \quad (1)$$

For a semi compensation method of approximately +3 dB per doubling of distance, we applied a linear conversion of the full-compensation weights.

$$w_{i,3\mathrm{dB}}(\vec{r}) = \Big(w_{i,6\mathrm{dB}}(\vec{r}) - 1\Big)(\sqrt{2} - 1) + 1 \quad (2)$$

We are aware of that the conversion in Equation 2 does not represent an exact compensation of +3 dB per doubling of distance. The correct function would have been achieved by taking one half of the log-weights:

$$w_{i,3\mathrm{dB}}(\vec{r}) = 10^{\frac{1}{2}log_{10}(w_{i,6\mathrm{dB}}(\vec{r}))} \quad (3)$$

An analysis of Equation 2 and Equation 3 yielded higher gains (less compensation) for loudspeakers that are close to the listener, with our method. For the IEM-CUBE, with a radius $r_0 \approx 5$ m in the horizontal loudspeaker ring, the deviation accounts for less than 1 dB, when the listener is more than 2.5 m far off the loudspeaker. For a deviation of 3 dB, the listener has to move 1.25 m close to a loudspeaker. Therefore, we could assume that our compensation results in slightly smaller values for the obtained radii.

For the real-time compensation, the below described stimuli have been played back and looped directly in the Pd patch. Thereby, each of the 24 loudspeaker channels has been multiplied by the continuously updated weighting coefficients.

Stimuli

The stimuli consisted of a sample of an English male speaker, which stems from the EBU Sound Quality Assessment Material [11, Track No. 50]. Using 5$^{\mathrm{th}}$-order Ambisonics encoding[1], we positioned the speaker at the centered front of the loudspeaker system, at 0° azimuth and 0° elevation. We prepared four sound files, three to investigate localization and one for diffuse envelopment. For the localization task, we decoded[2] the dry Ambisonic signal to our loudspeaker setup in orders 1, 3, and 5. The decoding employed the AllRAD approach [12] with appropriate max-$\mathrm{r_E}$ weighting. For diffuse envelopment, we created a reverberant soundscape by using a feedback delay network [13, 14, 15]. We set the parameters of the plugin[1] to a room size of 25, a reverberation time of 1.6 s and a fade-in of 0.5 s. Additionally, we set the internal high-shelf filter to $f = 8$ kHz, $Q = 0.5$ and a gain of $-43\ \frac{\mathrm{dB}}{\mathrm{s}}$ and the low-shelf filter to $f = 53$ Hz, $Q = 0.5$ and a gain of 6 $\frac{\mathrm{dB}}{\mathrm{s}}$. This filter reduces the reverberation time of frequencies higher than 2 kHz to obtain a natural sounding reverb. Subsequently to the reverberation plugin, we added a second high-shelf filter with at $f = 2.8$ kHz, with $Q = 0.69$, and a gain of -6 dB. We decoded the resulting reverberant signal with a 5$^{\mathrm{th}}$-order Ambisonics decoder, again with AllRAD and appropriate max-$\mathrm{r_E}$ weighting.

Procedure

Given the shape of the room, a nearly symmetrical sweet-area could be assumed and therefore the experiment took place only in one half of the room. Starting from the system's acoustical center, the listeners were asked to move towards a specified loudspeaker on the horizontal plane, facing the front-centered loudspeaker. For orientation, we provided guidelines in each moving direction. While positions were constantly tracked during the experiment, the listeners were asked to indicate the spot at which the questioned criterion was not fulfilled anymore. Figure 1 shows a listener on a guideline during the experiment.

[1]StereoEncoder and FDNReverb, part of the IEM Plug-in Suite. Freely available at https://plugins.iem.at

[2]Matthias Kronlachner's Ambix Decoder, freely available at http://www.matthiaskronlachner.com

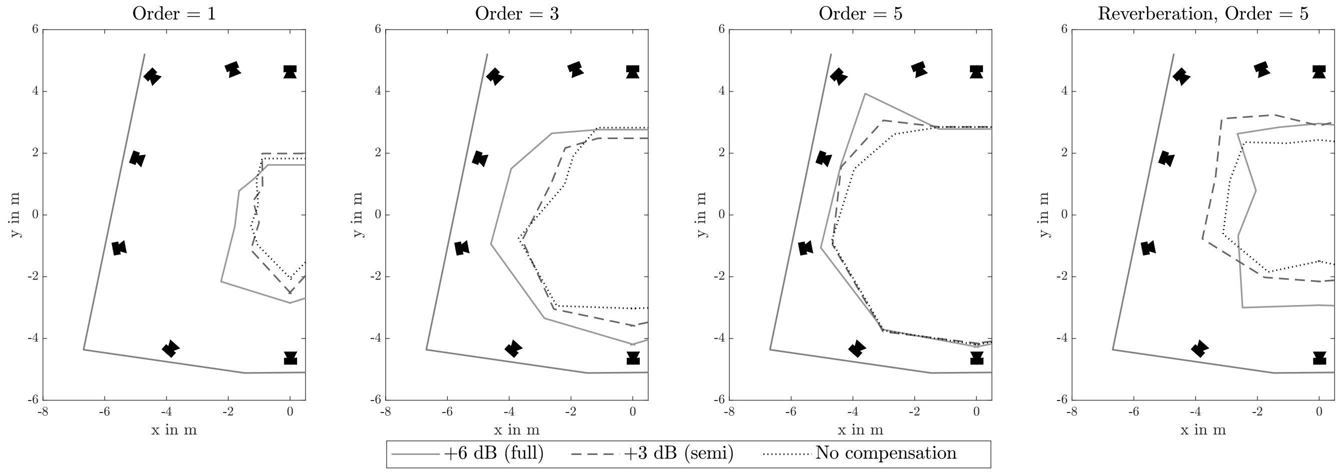

Figure 3: Calculated median sweet area for three different compensation methods, three different Ambisonics orders and reverbant signal.

The criteria were chosen according to [4]: (i) To investigate unambiguous localization, the criterion was the presence of a single sound source, located between the three frontal loudspeakers of the system, i.e. within a localization error of roughly ±10° (ii) as for diffuse envelopment, the criterion was fulfilled when the sound was perceived as "coming from each direction", i.e. each loudspeaker contributed to the sound field equally. The experiment consisted of 75 trials, in four runs. Three runs were to unambiguous localization, for Ambisonic orders 1, 3, and 5 into six directions and for three compensation methods as described above: full compensation, semi compensation and no compensation. We did not include the 7$^{\text{th}}$ direction to the front-centered loudspeaker for these runs, but adopted the values from the direction left beside it. This was because the source position was located exactly in that direction and hence the localization would not have been influenced by compensation. The fourth run was to investigate diffuse envelopment in 5$^{\text{th}}$-order Ambisonics and consisted of seven trials, including the front-centered loudspeaker direction, again for all three compensation methods. For each run, both the order of the compensation methods (and Ambisonics orders) for every single direction and the order of the examined directions were randomized, generating a different test sequence for every participant. This sequence was saved in a text-file, imported in the PD-patch and complemented with the investigated coordinates. The whole experiment took 20 minutes on average for every of the 13 participants (2 female, 11 male, between 23 and 38 years old).

Results

In order to evaluate the results of the listening experiment, the first step consisted in plotting the determined median (and mean) sweet areas. For this purpose, a MATLAB script was written, which reads the results from the output files and orders its lines, since those are still in the random order used during the experiment. After that, the median radii for each direction, compensation method, and Ambisonics order were computed and plotted, as can be seen in Figure 3. Since only one half of the room was considered for the testing, we chose not to mirror our results and thus plot only half of the sweet area. At first glance we can see that especially for localization, the full compensation performs best, resulting in a larger sweet area. However, for diffuse envelopment, the semi compensation seemed to perform best. This was also confirmed by the multivariate analysis.

For a multivariate analysis, a Wilcoxon signed-rank test with Bonferroni correction was conducted. For localization, the results show a significant increase ($p < 0.01$) of the relative sweet area radius for compensation methods: semi vs. no, full vs. no and full vs. semi in the case of Ambisonic orders 3 and 5. Order 1 shows only significance for: full vs. no and full vs. semi compensation. For diffuse envelopment, the semi vs. no compensation yields significantly ($p < 0.01$) greater results, while full vs. no and full vs. semi compensation shows no significant differences ($p = 0.1436$ and $p = 1.3645$, respectively). Therefore, we assume the semi compensation as the best performing method for the reverberant case. Figure 4 shows the median values and their confidence intervals for relative sweet area radii.

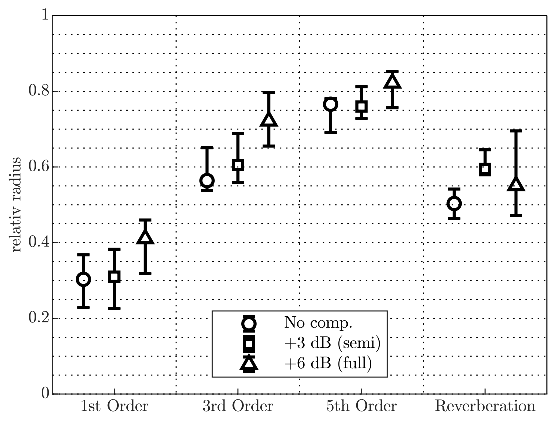

Figure 4: Median values and corresponing 95% confidence intervals for the relative radii for different playback scenarios and compensation methods.

Conclusion

The main goal of the experiment was to investigate how the application of line arrays for Ambisonic playback could affect the sweet area size. For this purpose, the usage of line arrays was simulated by a position-dependent sound level compensation of point-source loudspeakers for a single listener with three different compensation methods: full compensation (+6 dB), semi compensation (+3 dB) and no compensation. The difference between these compensation methods was investigated in 1^{st}-, 3^{rd}-, and 5^{th}-order Ambisonics for localization and using a reverberated stimulus in 5^{th}-order Ambisonics for listener envelopement.

The conducted experiment confirmed that the application of line-array loudspeakers for Ambisonic playback can have a positive impact on the sweet area size both in terms of localization and envelopment. It was determined that for non-reverberant reproduction scenarios, increasing position-dependent sound level compensation results in a significant enlargement of the sweet area size for unambiguous source localization. For reverberant reproduction, where the criterion of diffuse envelopment was investigated, the full compensation has shown no significant advantages compared to the semi compensation. In fact, for some directions the determined median radii were larger when using the semi compensation. This could be traced back to ambiguous individual decision criteria. For full compensation, we could observe an effect that the diffuse soundscape was "evading" the listener, as they were moving close to a loudspeaker. As we had not previously defined if the diffuse envelopment criterion is still satisfied in such cases, there is some ambiguity in the subject's response behavior for these scenarios.

The conducted experiment also confirmed the findings in [2, 4] that the sweet area, in which the playback can still be perceived plausibly, extends with increasing Ambisonic order.

A completely position-independent level distribution can hardly be achieved in practice, thus a compensation level between 3 dB and 6 dB can be considered realistic to allow suitable playback of any type of audio material. It also needs to be taken into consideration that in this experimental scenario, the compensation affected only the direct sound coming from the loudspeakers and that the diffuse sound field coming from the room, which contributed to the perceived sound field, could not be influenced. Since line arrays deliver narrow vertical radiation patterns, there should be fewer reflections in the room and thus less diffuse sound field. This could indicate that the semi compensation could perform better in the localization task using real line arrays.

The next step could be to conduct further experiments, both involving simulated compensation with different scenarios and also using real line arrays or loudspeakers with a wider playback angle. By varying the arrangement of the line-array elements or, if possible, tweaking the playback angle of the single loudspeaker, the desired compensation level can be achieved.

References

[1] F. Zotter and M. Frank, *Ambisonics: A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality.* SpringerOpen, 2019.

[2] F. Zotter, H. Pomberger, and M. Frank, "An alternative ambisonics formulation: Modal source strength matching and the effect of spatial aliasing," *126th Audio Engineering Society Convention 2009*, vol. 1, 01 2009.

[3] M. Frank and F. Zotter, "Spatial impression and directional resolution in the reproduction of reverberation," in *Fortschritte der Akustik, DAGA*, Aachen, March 2016.

[4] ——, "Exploring the perceptual sweet area in ambisonics," in *Audio Engineering Society Convention 142.* Audio Engineering Society, 2017.

[5] M. Morimoto, "Auditory spaciousness and envelopment," in *Proc. 13th ICA*, vol. 2, Belgrade, 1989, pp. 215–218.

[6] M. P. Cousins, F. M. Fazi, S. Bleeck, and F. Melchior, "Subjective diffuseness in layer-based loudspeaker systems with height," in *Audio Engineering Society Convention 139.* Audio Engineering Society, 2015.

[7] M. Blochberger, F. Zotter, and M. Frank, "Sweet area size for the envelopment of a recursive and a non-recursive diffuseness rendering approach," in *International Conference on Spatial Audio, ICSA*, 2019.

[8] M. S. Ureda, "„J"and „Spiral"Line Arrays," *Journal of The Audio Engineering Society*, 2001.

[9] ——, "Line arrays: Theory and applications," *Journal of The Audio Engineering Society*, 2001.

[10] J. Nettingsmeier and D. Dohrmann, "Preliminary studies on large-scale higher-order ambisonic sound reinforcement systems," in *Proc. 3rd Ambisonics Sympopsium*, Lexington, KY, 2011.

[11] EBU. (2008) EBU SQAM CD: Sound Quality Assessment Material recordings for subjective tests. [Online]. Available: https://tech.ebu.ch/publications/sqamcd

[12] F. Zotter and M. Frank, "All-Round Ambisonic Panning and Decoding," *Journal of the Audio Eng. Soc.*, vol. 60, no. 10, pp. 807–820, 2012.

[13] J. Stautner and M. Puckette, "Designing multichannel reverberators," *Computer Music Journal*, vol. 6, no. 1, pp. 52–65, 1982.

[14] J.-M. Jot and A. Chaigne, "Digital delay networks for designing artificial reverberators," in *90th AES Conv., prepr. 3030*, Paris, February 1991.

[15] N. Meyer-Kahlen, S. J. Schlecht, T. Lokki *et al.*, "Fade-in control for feedback delay networks," in *International Conference on Digital Audio Effects*, 2020.