Influence of directivity pattern order on perceived distance

Florian Wendt, Matthias Frank, Franz Zotter, Robert Höldrich

University of Music and Performing Arts Graz, Austria

Institute of Electronic Music and Acoustics, Email: {wendt, frank, zotter, hoeldrich}@iem.at

Introduction

The metaphoric idea behind employing the icosahedral loudspeaker array (ICO) in music is to "orchestrate" reflecting surfaces, resulting in the perception of auditory objects at distinct locations in the room [1].
An early study [2] about the perceived direction of auditory objects created by directional sound sources such as the ICO confirms the influence of orientation of directivity on localization deviating from the direct path.
In extension to this preliminary study, this contribution focuses on the perceived distance of auditory objects created by the ICO. The controllability of the perceived distance is examined, which is essential for the application of the ICO as an instrument.

We briefly discuss the method to control the perceived auditory distance using a variable-order directional source and subsequently design a listening experiment based on an auralized room. After this, we discuss results, which are modeled in the last section.

Controlling the auditory distance

In electro-acoustic music, the notion of adjustable-directivity loudspeakers was introduced in the late 1980s by researchers at IRCAM. For the renowned concept study "la timée" [3], a cube housing six separately controlled loudspeakers was built to achieve freely controllable directivity. In 2006, researchers at IEM reconsidered the theory and built a larger and more powerful 20-sided, 20-channel playback device yielding controllable 3rd-order directivity patterns [4] (see Fig. 1).

Recently Laitinen [5] presented a method to control the perceived distance of an auditory object by changing the directivity pattern of a cubical loudspeaker array.
Directivity control was used to modify the amount of reverberant energy. The direct-to-reverberant energy ratio

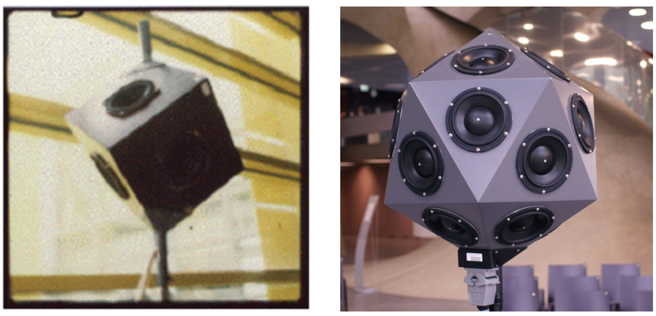

Figure 1: Adjustable-directivity loudspeakers "la timée" (left) and ICO (right).

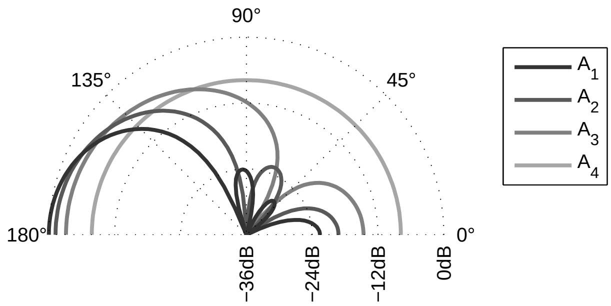

Figure 2: Directivity patterns $A_{1...4}$ normalized to constant energy.

(D/R-ratio) is known to be a prominent cue for distance perception (see [6] for a thorough review).
The D/R-ratios in numbers does not only depend on the directivity pattern but is essentially shaped by the room response. Still, directivity-pattern designs can be defined that accomplish room-independent control of the auditory distance. Following the idea in [5], the controllability of auditory distance is examined in closer detail here, after allowing a more thorough control of directivity.

Considered directivity patterns are based on frequency-independent max-$\mathbf{r}_\mathrm{E}$ beampattern designs, which exhibit a relatively narrow main lobe and permits sufficiently well suppressed side lobes [7]. Seven different directivity patterns were tested, denoted as $A_{1...7}$. Table 1 lists these patterns $A_1 \ldots 7$ in particular, and Figure 2 shows the patterns $A_{1...4}$ (normalized to constant energy).

Table 1: Properties of tested directivity patterns.

A_1	3rd order max-$\mathbf{r}_\mathrm{E}$ beam towards listener
A_2	2nd order max-$\mathbf{r}_\mathrm{E}$ beam towards listener
A_3	1st order max-$\mathbf{r}_\mathrm{E}$ beam towards listener
A_4	omnidirectional beampattern
A_5	1st order max-$\mathbf{r}_\mathrm{E}$ beam away from listener
A_6	2nd order max-$\mathbf{r}_\mathrm{E}$ beam away from listener
A_7	3rd order max-$\mathbf{r}_\mathrm{E}$ beam away from listener

Experimental Setup

We investigated the influence of beampatterns on the perceived distance in a listening experiment. The experiment was done at IEM in Graz and is part of a more comprehensive study submitted to DAFx 2016 [8].

Direct sound and early reflections were auralized using the image source method [9], whereas diffuse sound was simulated with the software toolbox MCRoomSim [10].

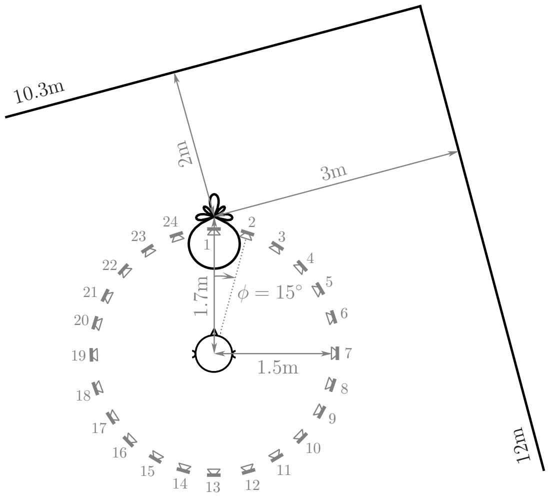

Figure 3: Setup of the auralized room and sound source using the 24-channel loudspeaker ring.

The auralized room is shoebox shaped ($10.3\,\mathrm{m} \times 12\,\mathrm{m} \times 4.8\,\mathrm{m}$) with a frequency-independent absorption coefficient of 0.3 and a mean reverberation time of 700 ms. Playback employed a ring of 24 equally-distributed Genelec 8020 loudspeakers with a radius of $r = 1.5\,\mathrm{m}$, placed in the anechoic laboratory of the IEM. Figure 3 shows the setup and positioning of the auralized room.

Each listener's task was to indicate the perceived distance on a graphical user interface displaying a continuous slider for each sample in a multi-stimulus set to permit comparative rating along the ordinal scale *very close* (vc), *close* (c), *moderate* (m), *distant* (d), and *very distant* (vd). The subjects were allowed to repeat each sample at will, and sound files were played back in a loop.

During the listening session, the listener was requested to face loudspeaker 1 ($\phi = 0°$), which corresponds to the direction of the auralized sound source.

Fifteen listeners participated in the test. All of them were experienced listeners with normal hearing.

Conditions

To allow comparison, sounds comply with sounds from earlier experiments [2, 5]. We used female speech (S_1, sample taken from CD *B&O* 101, 1992) and Gaussian white noise (S_2) with signal spectrum and envelope shaped to the speech signal [11, 12]. Both sounds were equalized in level.

The listening test was carried out as a multi-stimulus test in a MUSHRA-like procedure [13]. Each sample represents a directivity pattern and sound.

Both sounds $S_{1/2}$ were tested with directivities $A_{1\ldots7}$ in individual sets yielding responses $x^{\mathrm{I}}_{1\ldots7}$ for each subject and sound. To compare results from different sounds, $A_{1\ldots7}(S_1)$ and $A_{1\ldots7}(S_2)$, an additional multi-stimulus set included the selected conditions $A_{1,4,7}(S_{1,2})$ for both sounds, yielding the responses $x^{\mathrm{II}}_{1,4,7}$ per subject and sound. The responses $x^{\mathrm{I}}_{2,3}$ and $x^{\mathrm{I}}_{4,5}$ were re-mapped for each listener and sound by linear scaling and shifting to match $x^{\mathrm{I}}_{1,4}$ with $x^{\mathrm{II}}_{1,4}$, and $x^{\mathrm{I}}_{4,7}$ with $x^{\mathrm{II}}_{4,7}$, respectively:

$$x_i = \begin{cases} x_i^{\mathrm{II}} & \text{for } i \in \{1,4,7\}, \\ \dfrac{x_4^{\mathrm{II}} - x_1^{\mathrm{II}}}{x_4^{\mathrm{I}} - x_1^{\mathrm{I}}}(x_i^{\mathrm{I}} - x_1^{\mathrm{I}}) + x_1^{\mathrm{II}} & \text{for } i \in \{2,3\}, \\ \dfrac{x_7^{\mathrm{II}} - x_4^{\mathrm{II}}}{x_7^{\mathrm{I}} - x_4^{\mathrm{I}}}(x_i^{\mathrm{I}} - x_4^{\mathrm{I}}) + x_4^{\mathrm{II}} & \text{for } i \in \{5,6\}, \end{cases} \tag{1}$$

i.e., a complete response set $x_{1\ldots7}$ per listener and sound.

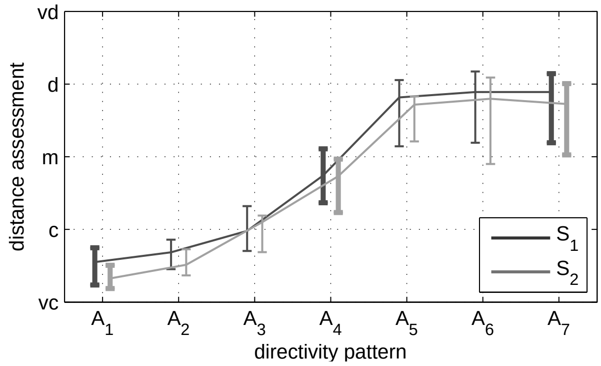

Figure 4: Median and 95% confidence intervals for tested sounds $S_{1,2}$ and directivity patterns $A_{1\ldots7}$.

Results

Figure 4 shows median values and corresponding 95% confidence intervals of $x_{1\ldots7}$ using Eq. (1).

For conditions $A_{1\ldots5}$, the auditory distance increases monotonically for both sounds. An analysis of variance (ANOVA) of neighboring values reveals conditions $A_{1\ldots5}$ to be significantly different ($p < 0.09$). By contrast, conditions $A_{5\ldots7}$ do not yield a significant change ($p \geq 0.74$). The choice of the sound S_1 or S_2 does not yield a significant difference ($p = 0.52$), so that the considerations below only uses pooled responses for both sounds $S_{1,2}$.

Modeling the auditory distance

The question is whether the responses can be explained by characteristic metrics used to characterize the spatial sound field in psychoacoustics. Linear regression based on the following simulated metrics were tested:

- Direct-To-Reverberant Energy Ratio,
- Lateral Energy Fraction,
- Inter-aural Cross Correlation Coefficient.

The D/R-ratio is widely accepted for prediction of auditory source distance and hence the most obvious predictor. By contrast, the Lateral Energy Fraction (LF) and Inter-aural Cross Correlation Coefficient (IACC) are both used to describe either listener envelopment or apparent source width [14, 15, 16]. While the D/R and Lateral Energy ratios are positively correlated with features relevant for auditory distance, the IACC is negatively correlated. Regression uses 1−IACC.

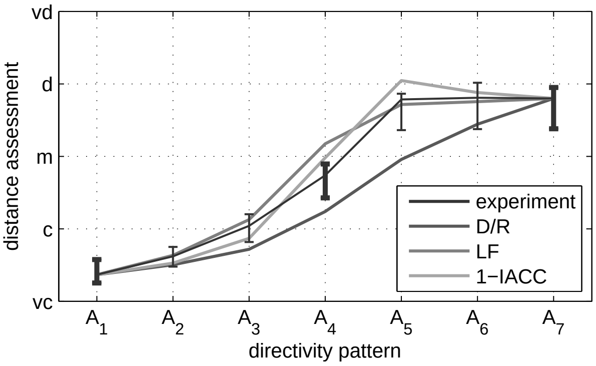

Figure 5: Comparison of median and 95% confidence intervals of assessed distance pooled over sounds $S_{1,2}$ with predictors based on D/R-ratio, LF, and 1−IACC.

Figure 5 compares the median and 95% confidence intervals with the linear regression models for the conditions $A_{1/7}$.

All models yield curves that are highly correlated with the experimental data. Interestingly, spatial measures quantyifing the apparent source width turn out almost perfect. This is underlined by their correlations of $R^2 = 0.97$ for the LF and $R^2 = 0.99$ for 1−IACC, whereas the D/R-ratio reaches $R^2 = 0.93$.

Conclusion

We investigated the influence of frequency-independent max$-\boldsymbol{r}_\mathrm{E}$ directivities on the perceived auditory distance. We could show that for a variable-directivity source pointing its beam towards and away from the listener is able to evoke a series of pronounced and graduated distance impressions.

The mapping of the directivities $A_{1\dots7}$ to the perceived distance curve is sigmoid-shaped and thus we could confirm a signal-independent range of controllability.

Finally, we showed that each the D/R-ratio, the LF, and 1−IACC are suitable in linear regression models yielding curves highly correlated to the responses.

Acknowledgments

Our research was partly funded by the Austrian Science Fund (FWF) project nr. AR 328-G21, Orchestrating Space by Icosahedral Loudspeaker (OSIL). We thank all subjects taking part in the experiment, which was supported by the project Acoustic Sensing and Design (ASD), granted by the Austrian ministries BMVIT, BMWFJ, the Styrian Business Promotion Agency (SFG), and departments 3, 14 of the Styrian Government. For these institutions hosting the Competence Centers for Excellent Technologies (COMET, K-Project), the Austrian Research Promotion Agency (FFG) processes and monitors funding.

References

[1] G. K. Sharma, F. Zotter, and M. Frank, “Orchestrating wall reflections in space by icosahedral loudspeaker: findings from first artistic research exploration,” in *Proc. ICMC/SMC, Athens*, pp. 830–835, 2014.

[2] F. Zotter and M. Frank, “Investigation of auditory objects caused by directional sound sources in rooms,” *Acta Physica Polonica A*, vol. 128, no. 1, pp. A5–A10, 2015.

[3] O. Warusfel, P. Derogis, and R. Causse, “Radiation Synthesis with Digitally Controlled Loudspeakers,” in *Audio Engineering Society Convention 103*, sep 1997.

[4] F. Zotter, *Analysis and synthesis of sound-radiation with spherical arrays.* PhD thesis, Univ. Musik u. darstellende Kunst Graz, 2009.

[5] M.-V. Laitinen, A. Politis, I. Huhtakallio, and V. Pulkki, “Controlling the perceived distance of an auditory object by manipulation of loudspeaker directivity,” *The Journal of the Acoustical Society of America*, vol. 137, no. 6, pp. EL462—-EL468, 2015.

[6] P. Zahorik, D. S. Brungart, and A. W. Bronkhorst, “Auditory distance perception in humans: A summary of past and present research,” *Acta Acustica united with Acustica*, vol. 91, no. 3, pp. 409–420, 2005.

[7] J. Daniel, *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia.* PhD thesis, Université Paris 6, 2001.

[8] F. Wendt, M. Frank, F. Zotter, and R. Höldrich, “Directivity patterns controlling the auditory source distance,” submitted to *DAFx*, 2016.

[9] J. B. Allen and D. A. Berkley, “Image Method for Efficiently Simulating Small-room Acoustics,” 1979.

[10] A. Wabnitz, N. Epain, C. T. Jin, and A. Van Schaik, “Room acoustics simulation for multichannel microphone arrays,” *International Symposium on Room Acoustics*, no. August, pp. CDROM: 1–6, 2010.

[11] D. Byrne, “An international comparison of long-term average speech spectra,” *The Journal of the Acoustical Society of America*, vol. 96, no. 4, p. 2108, 1994.

[12] A. Kohlrausch, R. Kortekaas, M. van der Heijden, S. van de Par, A. J. Oxenham, and D. Püschel, “Detection of Tones in Low-noise Noise: Further Evidence for the Role of Envelope Fluctuations,” *Acta Acustica united with Acustica*, vol. 83, pp. 659–669, 1997.

[13] ITU, “ITU-R BS.1534-1: Method for the subjective assessment of intermediate sound quality (mushra),” 2006.

[14] J. Blauert, *Spatial hearing - the psychophysics of human sound source localization.* The MIT Press, 1983.

[15] A. H. Marshall, “A note on the importance of room cross-section in concert halls,” *Journal of Sound and Vibration*, vol. 5(1), pp. 100–112, 1967.

[16] M. Barron and A. H. Marshall, “Spatial impression due to early lateral reflections in concert halls: The derivation of a physical measure,” *Journal of Sound and Vibration*, vol. 77, no. 2, pp. 211–232, 1981.