Orchestrating wall reflections in space by icosahedral loudspeaker: findings from first artistic research exploration

Gerriet K. Sharma, Franz Zotter, and Matthias Frank
Institute of Electronic Music and Acoustics
University of Music and Performing Arts Graz, Austria
sharma@iem.at

ABSTRACT

Can we orchestrate the acoustic reflections in a room within an electroacoustic composition? Doing so appears promising by using the twenty-channel IEM icosahedral loudspeaker (ICO) and its beamforming algorithms. Based on two musical pieces, we present initial results from an investigation about the perceived auditory objects. By means of explorative listening experiments, we bring evidence that the ICO orchestrates wall reflections. Moreover we can roughly explain the responses by a wall reflection model with echo thresholds or by a binaural lateralization model.

1. INTRODUCTION

Holophony [1] can be used to replicate natural sound generators or to excite paths of sound reflection. In particular, the notion of employing sound sources with adjustable acoustic radiation in electroacoustic music was introduced in Paris in the late 1980s by a research group at IRCAM. For this renowned concept study they built "la timée" [2], a cube-shaped loudspeaker with six separately driven channels for the production of freely controllable sound radiation directions. In 2006, a technical in-depth investigation started at the authors' institute. The result was a twenty-channel icosahedral loudspeaker system [3], see Fig. 1. The 20-channel IEM icosahedral loudspeaker (ICO) emits sound whose strength is adjustable for different spatial directions. It is capable of providing a correct and powerful simulation of musical instruments in their lower registers in all their 360° directional transmission range. The device is also suitable for the application of new room acoustic measurements in which controllable directivity is used to obtain a refined spatial characterization.

Currently there exist only few comparable systems in the world: a 120-channel system at CNMAT, Berkeley [4], a 12-channel system at ITA, RWTH-Aachen, Germany, a 12-channel system at the Acoustics Lab, Ben-Gurion University of the Negev, Israel, cubical 6-channel systems at IRCAM, France, hemi-dodecahedral 6-channel systems at the Princeton and Stanford Laptop orchestras [5], and experimental systems in works of Curtis Bahn, Perry Cook,

Figure 1. ICO playing spatial music for a dummy listener for the systematic investigation. Exemplary binaural recordings of stimulus 2 are available at: http://iaem.at/Members/zotter/KK_2.wav

and Dan Trueman [6]. Except for the systems in Berkeley, Princeton, and Stanford, most systems employing spherical beamforming are primarily used for measurement purposes.

The application of the beamforming algorithm developed in [3] allows strongly focused sound beams to be projected onto floors, ceilings, and walls. Thanks to the increased number of loudspeakers, it achieves sound beams that are three times narrower than early prototypes. The beamforming allows to attenuate sounds from the ICO itself while sounds from acoustic reflections can be emphasized. Beams are not only freely adjustable in terms of their radiation angle, also different ones can be blended, or their beam width can be increased. A loose idea behind employing such sound beams in music is to orchestrate reflecting surfaces, maybe yielding useful effects in the perceived impression.

It came as a big surprise that the ICO permits to form three-dimensional auditory objects in space including a useful gradation of depth that was not only noticed by the composer, but also noticed in statements of fellow artists. These three-dimensional auditory objects, although more general, are extensively discussed and pursued in the recent artistic research, e.g. [7].

This contribution is arranged as follows: Section 2 introduces the main questions induced during early performances of the ICO. The subsequent section presents a first step in answering these questions in terms of an exploratory auditory evaluation. The results of the evaluation can partly

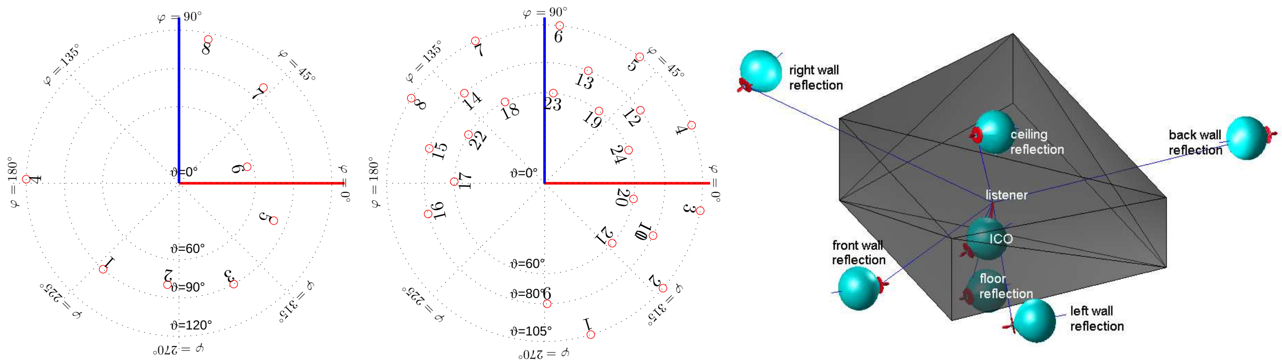

Figure 2. Layouts of the pieces *grrawe* (left) and *firniss* (middle) in terms of beams (see numbers) to which the composition routes and amplitude-pans its signal objects. The image on the right shows the mirror sources representing wall reflections of the ICO, cf. Tab. 1, including an examplary beam pattern pointing at the back-wall reflection.

be explained by simple geometrical considerations. The more elaborated binaural model in section 4 provides more insight into the localization of the auditory objects. Exemplary for a single object, section 5 merges the results of the verbal evaluation, the geometrical considerations, as well as the binaural model. Finally, the last section summarizes our contribution and sketches further steps in understanding the orchestration of space by the ICO.

2. EARLY PERFORMANCES AND INQUIRIES

In the exploratory works [1] grrawe 2010 (10'26") and firniss 2012 (11'23"), the authors explored how to employ the ICO technically and aesthetically for electroacoustic music, using beam/track layouts of Fig. 2. As part of this collaboration, the ICO has been tested since 2009 following the requirements of an artistic work in progress. The results of these "inquiries" were in turn integrated into both the development of the composition and the development of the instrument ICO.

The main questions emerging from early performances and shared experiences were: Can the choreography of electroacoustic auditory objects be reproduced in a composition and in different spaces (e.g. physical, imaginary, social, language)? Or does the choreography mostly develop in the (acoustic) imagination of the composer — hence are they to be seen more as metaphorical and programmatic settings that serve the composer primarily as aids in dealing with the technical equipment? How can one stage the auditory-object choreography and make it tangible for an audience? What spatial conditions are required to do so? How can one describe and verbalize these objects for intra- and interdisciplinary exchange of ideas? In the case of many musical effects that emerge and which are being reinforced by the ICO, the question remains: why and how?

[1] grrawe was presented at the *Forum Alpbach* 2010 and formed part of the concert programme at the *Digital Audio Effects Conference DAFx10*, as well as at the 2011 *next generation Festival* at the Centre for Art and Media Technology ZKM Karlsruhe. Zotter and Sharmas work with the icosahedral loudspeaker was the only Austrian contribution presented in the music category within the framework of the ELIA-Art Schools *NEU/NOW Festivals* 2011. The icosahedral loudspeaker and grrawe were presented to a specialist audience of composers and sound engineers at the *International Conference for Spatial Audio ICSA* 2011. grrawe and firniss were invited to the *International Computermusic Conference ICMC* 2012 and the *Internationale Ferienkurse für Neue Musik Darmstadt* 2014.

sound path	T/ms	$\varphi/°$	$\vartheta/°$
direct (di)	0	42	88
front (fr)	12	18	89
left (le)	15	65	89
back (ba)	28	169	89
right (ri)	30	-67	89
ceiling (ce)	11	47	27
floor (fl)	2	42	124

Table 1. Sound paths arriving at the listening position from azimuth and zenith angles $\varphi/°$ and $\vartheta/°$ with a time delay T relative to the direct sound.

Object	Time in s	Beams
clicks	9-59	Panning 5/6
rays	0-56	1+2+4+7 + Panning 5/6
melody	30-42 & 49-55	Panning 9/10
bass	5-60	Panning 7/8
squeaks	29-33 & 55-59	5+6

Table 2. Appearance times/beam numbers in stimulus 2.

3. EXPLORATORY EVALUATION WITH 7 LISTENERS

Listener-based research is not totally new within this field. Nevertheless, its importance needs to be strongly enforced: Any objective evidence about the qualities of perceived sound objects can only be accessed systematically through listening tests, but still they are only seldom utilized [8]. In this contribution we want to show some initial evidence by a first, small evaluation through listeners. Starting with this kind of research one stops immediately at the first obstacle: How can one ask detailed questions if there is no widely accepted vocabulary for relatively new sonic emergences? However, present terminology still offers only insufficient classification that is relevant in spatial sound-based music. Despite the importance of musicological analysis and its comparability, still researchers have to develop suitable vocabularies on their own, e.g., in [9, 10, 11, 12]. Even more so, the exploration of auditory objects in our study faces the absence of any commonly accepted terminology. To still provide initial results, we used an open questionnaire for the exploratory study below.

Object	How listeners characterized the object's sound	Where listeners perceived sound object
clicks	clicks, waterdrops, pulse	at the ico slightly left, some from behind the ico
rays	sound sphere, drone like, bell like, static sound	proceed out of the ico, invade space from the side
bass	bass, deep, deep drone	beginning in front, towards the end from behind
squeaks	high pitchy squeak, high tv-like whistling, high frequency sound: pitch stable, high feedback like sound	from the right behind me, above my head
melody	phrase, iu, low mid sound: moving pitch going up and down, more smooth than the one before	middle of the wall, from behind the ico, jumps

Table 3. Synopsis of how listeners described sound objects in stimulus 2.

In an exploratory auditory evaluation in October 2012, 7 subjects were presented 8 stimuli, excerpts of the compositions grrawe and firniss, with the ICO in the IEM-CUBE. The length of the stimuli was between 10 and 60 seconds. Each subject was alone in the room when filling a questionnaire with the following three questions for every stimulus:

A. How many distinguishable auditory objects are heard?

B. What characterizes the sounds of these objects?

C. Which object is heard where (including movement)?

Statements and descriptions were transcribed from the paper questionnaires and compared for parallels and significant discrepancies. The terms used by the listeners for characterizing the auditory objects varied, of course, but were surprisingly congruent for some stimuli. For this case study we used the terms that had been similarly used by the listeners. Tab. 2 shows the objects of stimulus 2, exemplarily. The verbalization results for this stimulus are summarized in Tab. 3.
Given the congruence of where listeners perceived sound objects, the question remains whether location and extent can be understood in terms of simple geometric considerations, or are they estimated by psychoacoustic binaural localization models.

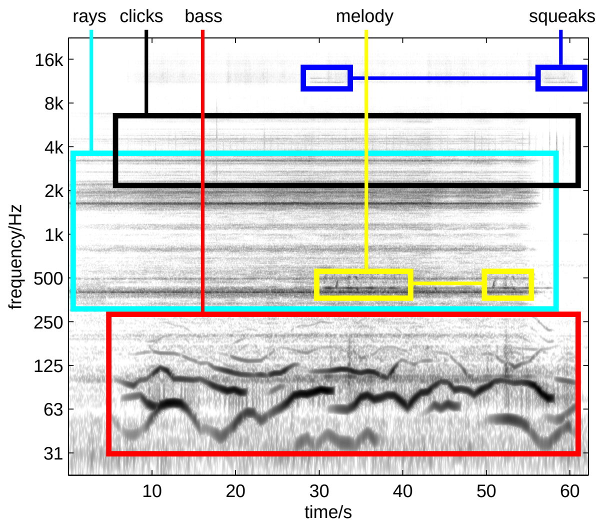

Figure 3. Spectrogram of stimulus 2 based on transformation with logarithmic frequency resolution (short-term constant Q transform, ST-CQT [13]). Sound objects of Tab. 2 largely cover distinct frequency bands (colored boxes).

4. MODELING AUDITORY OBJECTS

Recent work on the sound spatialization technique Ambisonics [14, 15] brought forward new expertise on spatial perception of auditory objects created by multiple active surrounding loudspeakers. This is in contrast to former work on triplet or pairwise stereophony that only uses two or three active loudspeakers per auditory object [16, 17, 18, 19]. The investigations in [20, 21, 22, 23, 24] indicate how to model such phantom sources in terms of direction, width, and coloration.

For auditory events created by the ICO, additional complications, e.g. unequal acoustic delays, need to be considered. Fig. 4 shows that the directivity of the ICO controls the magnitude of each sound propagation path (direct or reflected), of which each one is associated with a direction and time of arrival at the location of the listener. So far, there is only one brief work about the topic [25]. A basic understanding of the perception of reflected sound is given in the work of Hartmann [26, 27, 28, 29] and Morimoto [30, 31, 32], which underlines that sounds, directions, amplitudes, delays, and kind of sounds arriving at the listener determine what is being perceived. What is more, the -0.25dB/ms *echo threshold* given by Rakerd et al. [33] predicts whether sound arriving delayed with regard to preceding sound still affects auditory localization. For instance, the wall reflection *ri* excited by the beam in Fig. 4 lies above this threshold, and accordingly it is expected to strongly determine the perceived auditory object. However, also thresholds discussed in newer works, e.g. by Goupell et al. [34] or Donovan et al. [35], did either only investigate two instances of a sound with varying amplitudes, or more instances of a sound of the same amplitude. For this reason, there has been doubt concerning the applicability of these threshold models during the work on this paper. Only after finishing, our noise burst and speech experiments in [36] could verify the -0.25dB/ms threshold model as a suitable predictor. This has been achieved by combining the echo threshold with the so-called *energy-vector* [23, 37] but couldn't be included here, anymore.

In addition to the experimental exploration here, all stimuli excerpts were recorded with a B&K 4128C dummy head at the listening position, cf. Fig. 1. We fed the recordings into a binaural localization after Lindemann [38, 39] which

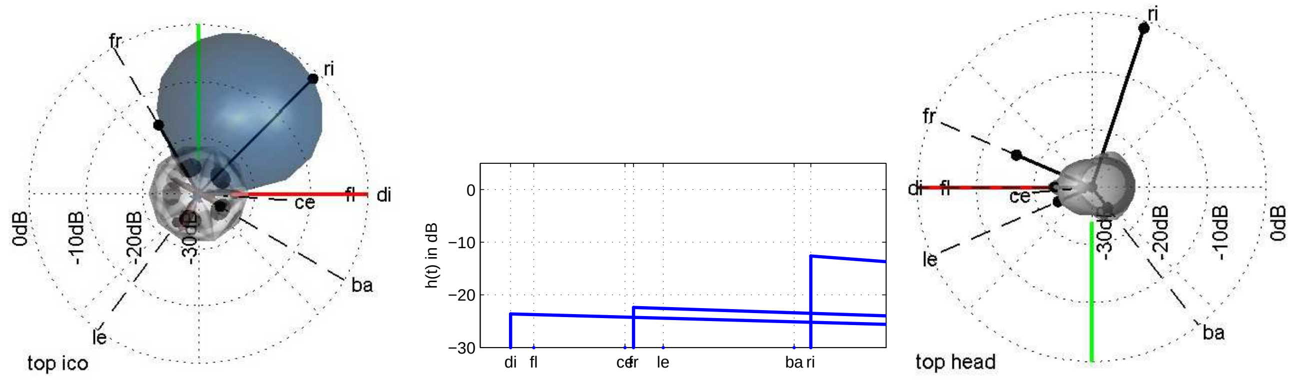

Figure 4. Reflection paths as they are excited by the stimulus example *squeaks* through the ICO are shown in the polar top view (left). Time sequence of arriving reflections at the listener and their *echo thresholds* are shown in the middle. The polar top view on the listener (right) shows from where reflections arrive with which propagation decay.

is part of the Auditory Modeling Toolbox [2]. The model divides the binaural recording into 40 bands with a spacing of 1ERB (equivalent rectangular bandwidth) [40] and center frequencies between 26 Hz and 16.9 kHz. The auditory nerve is modeled by a half-wave rectifier and a low-pass filter at 800 Hz. In each band, the inter-aural level-difference (ILD) is considered by monaural detectors and contra-lateral inhibition that yield a shift of the peaks in the inter-aural cross correlation function. The inter-aural time-difference (ITD) is then computed as the centroid of the inter-aural cross correlation function [41], which delivers one ITD value for each frequency band and each time step of 6 ms.

Exceeding the original model implementation, for each band and time step, our model considers the composite sound pressure level, which is calculated by energetic superposition of both ear signals. The model only keeps ITDs whenever the sound pressure level within the respective bands and time step lies above a certain threshold; note that the model does not use any advanced object identification, except for this manually adjusted level threshold. Moreover, the spectrogram in Fig. 3 shows that the previously named auditory objects mainly cover distinct bandwidths/time segments. To get a single composite ITD value for each object per time instant, the model calculates the level-weighted mean ITD of the frequency bands belonging to the respective object. Finally, to obtain direction estimates from the ITD values, the angle of the binaural impulse response pair (BRIR) with the most similar ITD is selected from a dataset of BRIRs at each frequency; the BRIR dataset is a set of horizontal semicircular measurements taken from the same dummy head that was used to take the binaural stimuli recordings. Fig. 5 shows the resulting directional evolution of each object over time.

5. EXEMPLARY DISCUSSION: *SQUEAKS*

Despite we presently lack evaluation and modeling methods covering dynamical auditory object extent and location description in 3D, it is possible to merge the above-mentioned results, exemplarily. We can gather a conclu-

[2] freely available on `amtoolbox.sourceforge.net/`

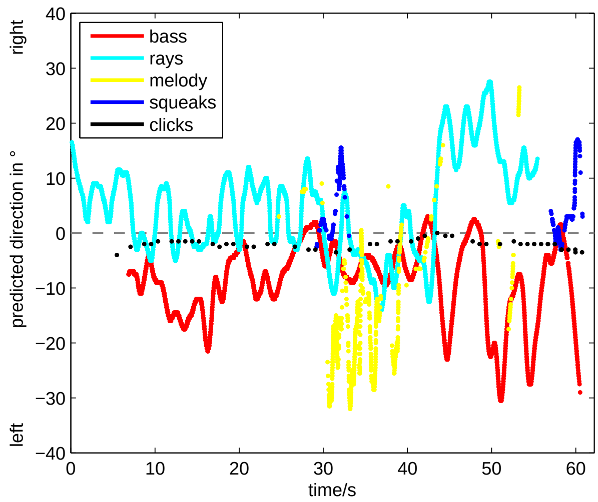

Figure 5. Predicted directions for each object of stimulus 2, calculated by the Lindemann binaural localization model.

sion from the geometrical and echo threshold-based evaluation in Fig. 4, the binaural localization model in Fig. 5, as well as the exploratory verbal evaluation by listeners in Tab. 3.

Most distinctively, the sound object *squeaks* shows a consistent trend:

- Tab. 3: ”*from the right behind me, above my head*”,
- Fig. 4: 15dB stronger right-wall reflection (*ri*) than front-wall reflection, while sound from other directions is negligible; *ri* clearly lies 12dB above the -0.25dB/ms echo threshold,
- Fig. 5: *squeaks* exhibit a lateralization contour that moves from the center to the right.

Notwithstanding, also the other sound objects are largely consistent with the verbally given localization of Tab. 3 and Fig. 5. The lateralization of *clicks* and *bass* matches the verbal description well. However, more accurate models are desirable.

6. CONCLUSION AND OUTLOOK

Our work using the ICO aims at producing stable auditory objects utilizing the reflections of its surrounding space. This contribution exemplarily reveals that individual listeners perceive these objects similarly, although using different verbal descriptions for characterizing the objects and their spatialization. The spatial character of these objects can partly be explained by geometrical, threshold-based models, as well as a binaural lateralization model. However, these models cannot explain all auditory aspects, e.g. three-dimensional localization, extent, and shape.

Steps towards verbalization should be taken in the future to form a consistent and common vocabulary powerful enough to describe the emerging auditory objects. Suitable methods, e.g. repertory grid [42], can help establishing this vocabulary. Additionally, pointing methods [43] can be employed to obtain quantitative localization results. Careful design of experimental environments and stimuli is required for accurate and effective evaluation and should bridge psychoacoustic evaluation and musical composition practice in electroacoustic music. Hence, not only may the stimuli cover typical psychoacoustic test signals, such as pure tones or noise, but also they need to consider simple musical miniatures and complex pieces.

Obviously, psychoacoustic understanding and modeling need to be further developed by new experiments in order to inform the compositional process of new pieces. In turn, these pieces and the new questions they raise inform a further psychoacoustic refinement.

In the future, the interdisciplinary cooperation of the areas of perception, psychoacoustics, cognition, music psychology, and semiotics will enable to better utilize the ICO as a mobile instrument that orchestrates the various spatial environments in which it is played.

Acknowledgments

The work of Gerriet K. Sharma with the ICO and the collaboration for this article were temporarily supported by the University of Music and Performing Arts (KUG) Graz and the Artistic Doctoral School of the KUG. The work of Franz Zotter and Matthias Frank was partly supported by the project ASD, which is funded by Austrian ministries BMVIT, BMWFJ, the Styrian Business Promotion Agency (SFG), and the departments 3 and 14 of the Styrian Government. The Austrian Research Promotion Agency (FFG) conducted the funding under the Competence Centers for Excellent Technologies (COMET, K-Project), a program of the above-mentioned institutions.

7. REFERENCES

[1] M. Jessel, *Acoustique Theorique: Propagation et Holophonie*. Masson, Paris, 1973.

[2] O. Warusfel, P. Derogis, and R. Caussé, "Radiation synthesis with digitally controlled loudspeakers," in *Papers of AES 103rd Conv.*, New York, 1997.

[3] F. Zotter, "Analysis and synthesis of sound-radiation with spherical arrays," Ph.D. dissertation, Univerity of Music and Performing Arts, Graz, 2009.

[4] R. Avizienis, A. Freed, P. Kassakian, and D. Wessel, "A compact 120 independent element spherical loudspeaker array with programmable radiation patterns," in *Papers of AES 120th Conv.*, Paris, 2006.

[5] D. Trueman, P. Cook, S. Smallwood, and G. Wang, "PLOrk: The Princeton Laptop Orchestra, Year 1," in *Proc. of Int. Comp. Mus. Conf., ICMC*, New Orleans, 2006.

[6] P. Cook, G. Essl, G. Tzanetakis, and D. Trueman, "N>>2: Multi-speaker display systems for virtual reality and spatial audio projection," in *Proc. of the Int. Conf. on Auditory Display, ICAD*, Glasgow, 1998.

[7] G. Eckel, M. Rumori, D. Pirrò, and R. González-Arroyo, "A framework for the choreography of sound," in *Proc. of Int. Comp. Mus. Conf., ICMC*, Ljubljana, 2012.

[8] L. Landy, *Understanding The Art of Sound Organization*. MIT Press, 2007.

[9] W. Thies, *Grundlagen der Typologie der Klänge*. Verlag der Musikalienhandlung K.D. Wagner, Hamburg, 1982.

[10] D. Smalley, "Space-form and the acousmatic image," *Organised Sound*, vol. 12, pp. 35–38, 2007.

[11] T. Grill, "Perceptually informed organization of textural sounds," Ph.D. dissertation, University of Music and Performing Arts, Graz, 2012.

[12] A. Lindau, V. Erbes, S. Lepa, H.-J. Maempel, F. Brinkman, and S. Weinzierl, "A spatial audio quality inventory for virtual acoustic environments (SAQI)," in *EAA Joint Symposium on Auralization and Ambisonics, submitted to Acta Acustica united with Acustica*, Berlin, 2014.

[13] C. Schörkhuber, A. Klapuri, N. Holighaus, and M. Dörfler, "A matlab toolbox for efficient perfect reconstruction time-frequency transforms with log-frequency resolution," in *Papers of AES 53rd Int. Conf.*, 2014.

[14] F. Zotter, H. Pomberger, and M. Noisternig, "Energy-Preserving Ambisonic Decoding," *Acta Acustica u. Acustica*, vol. 98, no. 1, pp. 37–47, 2012.

[15] F. Zotter and M. Frank, "All-round Ambisonic panning and decoding," *J. Audio Eng. Soc*, vol. 60, no. 10, pp. 807–820, 2012.

[16] D. M. Leakey, "Some measurements on the effects of interchannel intensity and time differences in two channel sound systems," *J. Acoust. Soc. Am.*, vol. 31, no. 7, pp. 977–986, 1959.

[17] K. Wendt, "Das Richtungshören bei der Überlagerung zweier Schallfelder bei Intensitäts- und Laufzeitstereophonie," Ph.D. dissertation, RWTH Aachen, 1963.

[18] G. Theile, "Über die Lokalisation im überlagerten Schallfeld," Ph.D. dissertation, Technischen Universität Berlin, 1980.

[19] V. Pulkki, "Spatial sound generation and perception by amplitude panning techniques," Ph.D. dissertation, HUT, Finland, 2001.

[20] M. Frank and F. Zotter, "Localization experiments using different 2D Ambisonics decoders," in *Proc. of 25th VDT Int. Convention*, Leipzig, 2008.

[21] M. Frank, "Phantom sources using multiple loudspeakers in the horizontal plane," Ph.D. dissertation, University of Music and Performing Arts, Graz, 2013.

[22] ——, "Source width of frontal phantom sources: Perception, measurement, and modeling," *Archives of Acoustics*, vol. 38, no. 3, pp. 311–319, 2013.

[23] ——, "Localization using different amplitude-panning methods in the frontal horizontal plane," in *Proc. of EAA Joint Symposium on Auralization and Ambisonics*, Berlin, 2014, pp. 42–47.

[24] F. Zotter, M. Frank, M. Kronlachner, and J.-W. Choi, "Efficient phantom source widening and diffuseness in ambisonics," in *Proc. of EAA Joint Symposium on Auralization and Ambisonics*, Berlin, 2014, pp. 69–74.

[25] A. Schmeder, "An exploration of design parameters for human-interactive systems with compact spherical loudspeaker arrays," in *Proc. of 1st Ambisonics Symposium*, Graz, 2009.

[26] W. M. Hartmann, "Localization of sound in rooms," *J. Acoust. Soc. Am.*, vol. 74, pp. 1380–1391, 1983.

[27] B. Rakerd and W. M. Hartmann, "Localization of sound in rooms, II: The effects of a single reflecting surface," *J. Acoust. Soc. Am.*, vol. 78, pp. 524–533, 1985.

[28] ——, "Localization of sound in rooms, III: Onset and duration effects," *J. Acoust. Soc. Am.*, vol. 80, pp. 1695–1706, 1986.

[29] W. M. Hartmann and B. Rakerd, "Localization of sound in rooms IV: The Franssen effect," *J. Acoust. Soc. Am.*, vol. 86, pp. 1366–1373, 1989.

[30] M. Morimoto, "The role of rear loudspeakers in spatial impression," in *Papers of AES 103rd Conv.*, 9 1997.

[31] ——, "The relation between spatial impression and the precedence effect," in *Proc. of Int. Conf. on Auditory Display, ICAD*, 2002.

[32] M. Morimoto, K. Nakagawa, and K. Iida, "The relation between spatial impression and the law of the first wavefront," *Applied Acoustics*, vol. 69, pp. 132–140, 2008.

[33] B. Rakerd, W. M. Hartmann, and J. Hsu, "Echo suppression in the horizontal and median sagittal planes," *J. Acoust. Soc. Am.*, vol. 107, no. 2, pp. 1061–1064, 2000.

[34] M. J. Goupell, G. Yu, and R. Y. Litovsky, "The effect of an additional reflection in a precedence effect experiment," *J. Acoust. Soc. Am.*, vol. 131, no. 4, pp. 2958–2967, 2012.

[35] J. M. Donovan, B. S. Nelson, and T. T. Takahashi, "The contributions of onset and offset echo delays to auditory spatial perception in human listeners," *J. Acoust. Soc. Am.*, vol. 132, no. 6, pp. 3912–3924, 2012.

[36] F. Zotter, M. Frank, A. Fuchs, and D. Rudrich, "Preliminary study on the perception of orientation-changing directional sound sources in rooms," in *Proc. of forum acusticum, Kraków*, 2014.

[37] M. Gerzon, "General metatheory of auditory localisation," in *Papers of AES 92nd Conv.*, Vienna, 1992.

[38] W. Lindemann, "Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals," *J. Acoust. Soc. Am.*, vol. 80, no. 6, pp. 1608–1622, 1986.

[39] ——, "Extension of a binaural cross-correlation model by contralateral inhibition. II. The law of the first wave front," *J. Acoust. Soc. Am.*, vol. 80, no. 6, pp. 1623–1630, 1986.

[40] B. C. J. Moore, R. W. Peters, and B. R. Glasberg, "Auditory filter shapes at low center frequencies," *J. Acoust. Soc. Am.*, vol. 88, no. 1, pp. 132–140, 1990.

[41] L. A. Jeffress, "A place theory of sound localization," *Journal of comparative and physiological psychology*, vol. 41, no. 1, pp. 35–39, 1948.

[42] S. Cunningham, "Applying personal construct psychology in sound design using a repertory grid," in *Proc. of the 5th Audio Mostly Conference*, 2010.

[43] M. Frank, L. Mohr, A. Sontacchi, and F. Zotter, "Flexible and intuitive pointing method for 3-D auditory localization experiments," in *Papers of AES 38th Int. Conf.*, Piteå, 2010.