Exposée zur Dissertation:

Envelopment and Distance in Spatial Audio

Stefan Riedel

Graz, August 2019

Abstract

This work investigates the perception of envelopment and distance in spatial audio systems. While previous research in the field of room acoustics has proposed predictors and simple objective measures for listener envelopment, recent studies revealed contradictory results. The listening experiments in this work shall lead to psychoacoustic criteria that explain listener envelopment in real and virtual immersive listening environments.

The experiments are conducted with two types of loudspeaker arrays. Firstly, traditional surrounding loudspeaker arrays to investigate different directions of arriving sound and the effects of signal decorrelation. Secondly, experiments with compact spherical loudspeaker arrays which explicitly excite wall reflections to create listener envelopment.

Kurzfassung

Diese Arbeit untersucht die Wahrnehmung von Einhüllung und Distanz in räumlichen Wiedergabesystemen. Vorhergehende Untersuchungen im Feld der Raumakustik schlagen objektive Maße zur Beurteilung von Einhüllung vor, jedoch zeigen neuere Studien widersprüchliche Ergebnisse. Die Hörversuche in dieser Arbeit sollen zu psychoakustischen Kriterien führen, die Einhüllung in realen und virtuellen immersiven Hörumgebungen erklären.

Hörversuche sollen mit zwei Typen von Lautsprecheranordnungen durchgeführt werden. Einerseits soll mit traditionellen umgebenden Lautsprecheranordnungen der Einfluss verschiedener Schalleinfallsrichtungen und der Signaldekorrelation untersucht werden. Andererseits sollen Experimente mit kompakten Kugellautsprecheranordnungen, welche gezielt Wandreflektionen anregen um Einhüllung zu erzeugen, durchgeführt werden.

Contents

1 Introduction

Listener envelopment is one of the key achievements of new spatial audio systems, therefore they are often referred to as 'immersive audio' or '3D Audio' technologies. Compared to stereophony [4], where auditory events are perceived on a line connecting the two frontal loudspeakers, surrounding loudspeaker arrays allow to create an experience that makes the listener feel inside the audio scene, a spatial attribute that has been termed as *envelopment* [18]. In fact, the term 'listener envelopment' (LEV) was introduced by concert hall acoustics research and refers to the quality of a concert hall, when listeners feel surrounded by the provided reverberation [3, 16].

Berg and Rumsey introduced the term 'room envelopment' for an enveloping reverberation of a perceivable room and 'source envelopment' for an auditory event or sound that is enveloping by its extent around the listener [1].

Technically, both forms are created from a *single source signal* that becomes an enveloping auditory event through audible spatial reverberation or widening/decorrelation filters that have minimal coloration as side effect [15, 21, 20].

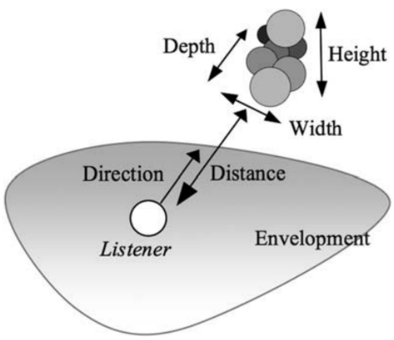

Figure 1.1: Illustration of the attributes of spatial imagery [14].

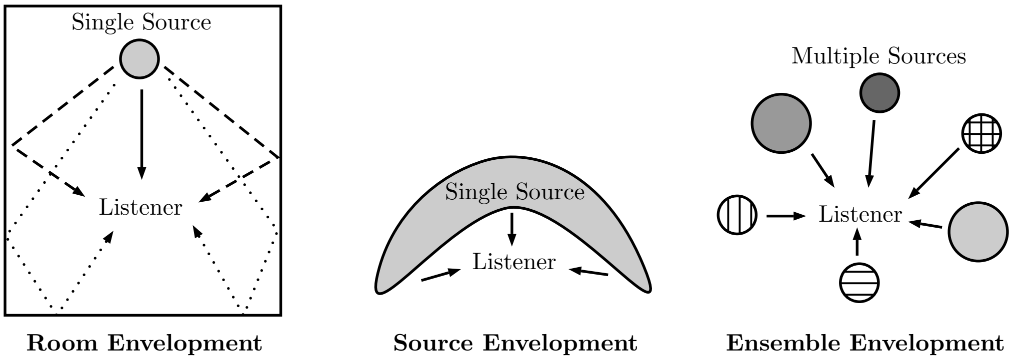

Figure 1.2: Illustration of the types of Listener Envelopment (LEV).

The author proposes a term for the form of envelopment that is created by an enveloping scene composed of *multiple source signals*. This shall be termed as 'ensemble envelopment', in accordance to the term 'ensemble width' by Rumsey et. al. (see Fig. 1.2).

This differentiation on a technical level must be investigated with listening experiments to find the fundamental psychoacoustic criteria for listener envelopment. Intuitively, one could formulate: 'Different or *decorrelated* signals presented from surrounding directions (frontal, lateral, behind, above) lead to the perception of envelopment'. This assumes that all sounds are perceived at a distance outside the listeners head (externalized), which is typically guaranteed with loudspeaker arrays, but is delicate in auralizations over headphones. The relationship between the perception of distance/depth and envelopment in spatial sound reproduction needs further investigation.

The following section discusses listening experiments and objective measures for listener envelopment (LEV) and depth perception proposed in the literature. After the literature review, own listening experiments are proposed that investigate all parameters like direction and spatial resolution of arriving sound, signal properties and decorrelation algorithms.

2 Literature Review

Early works in the context of concert hall acoustics investigated the correlation between perceived listener envelopment and objective measures such as the interaural cross-correlation (IACC) function, computed from recorded binaural impulse responses of a space [11]. The assumption is that a low interaural cross-correlation, especially for late arriving sound after 80 ms, relates to a diffuse and enveloping sound field. In 1995, Bradley and Soulodre conducted listening experiments in anechoic conditions and drew the conclusion that 'listener envelopment depends on having strong lateral reflections arriving at the listener 80 ms or more after the direct sound'. They found that their measure 'late relative lateral sound level' $\mathrm{LG}_{80}^{\infty}$ has an even higher correlation with perceived listener envelopment:

$$\mathrm{LG}_{80}^{\infty} = 10 \log \left[\int_{80\,\mathrm{ms}}^{\infty} p_L(t)\, dt \Bigg/ \int_{0\,\mathrm{ms}}^{\infty} p_A(t)\, dt \right] \tag{2.1}$$

, with $p_L(t)$ being the instantaneous pressure response of the lateral sound obtained from a figure-of-eight microphone and $p_A(t)$ is the response of the same source at a distance of 10 m in free-field (anechoic) conditions [5].

Furuya et. al. investigated the importance of other arrival directions of late sound energy and they conclude that 'lateral sound level gives the highest correlation with LEV, while late sound arriving from overhead and behind the listener also correlates very strongly' [8], see Fig. 2.1.

Meanwhile D. Griesinger developed a theory that explains envelopment through rapid fluctuations of the ITD (Interaural Time Difference) and IID (Interaural Intensity Difference). A well designed concert hall produces diffuse reflections from side walls and modulates the ITD rapidly, which is best perceived 80 - 200 ms after the direct sound, thus referred to as background spatial impression

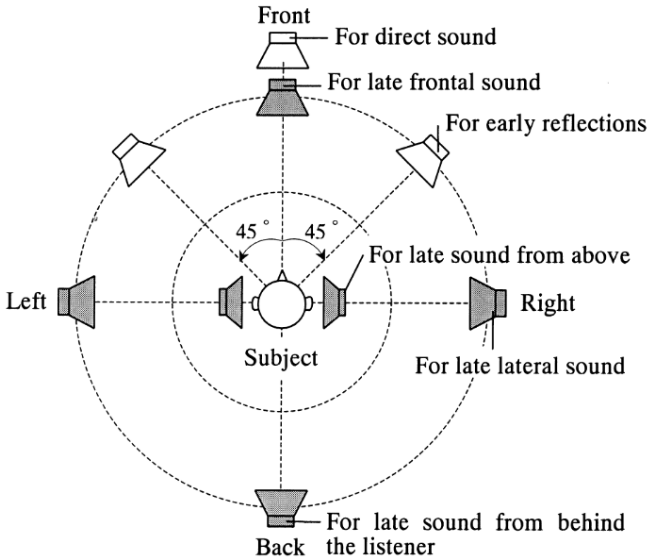

Figure 2.1: Illustration of the experimental setup to investigate the influence of late sound arrival direction on LEV (Furuya et. al. [8]).

(BSI) [10]. With early spatial impression (ESI) he explains the perception of source width determined by early reflections during the onset of a sound event. A continous spatial impression (CSI) is present during ongoing musical notes or intransparent speech, which causes less modulations in ITDs/IIDs and generally causes a lower perception of envelopment. He subsequently developed a method to detect ITD fluctuations and compute a so called Diffuse-Field Transfer Function as another objective measure to characterize the listener envelopment in rooms [9].

More recent work by Dick et. al. [7] used spherical microphone arrays to record spatial impulse responses of a concert hall. They investigated the established hypothesis that mostly the late spatial distribution is relevant for listener envelopment. Through experiments that presented hybrid IRs (containing portions of an enveloping IR and an unenveloping IR with crossover times ranging from 40 to 140 ms) they concluded that the early part of the spatial IR contributes significantly to the perceived envelopment, which is somewhat contradictory to earlier results.

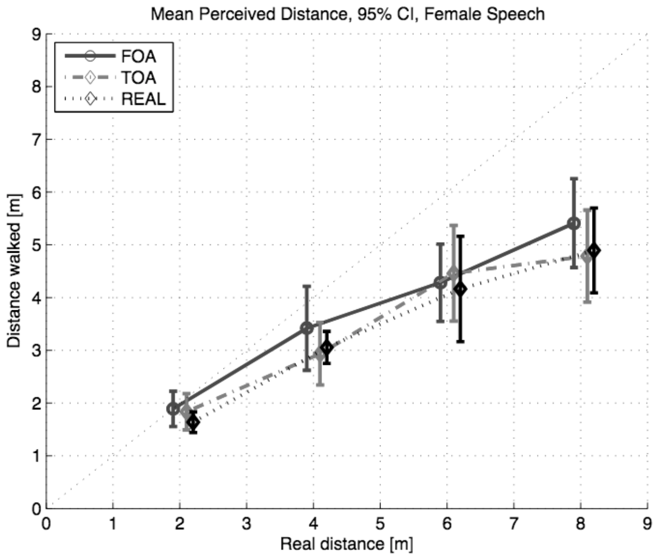

Figure 2.2: Perceived distance vs. simulated/reference distance. *FOA* and *TOA* denote ambisonic loudspeaker array auralization of measured room impulse responses of different distances (first and third order). [REAL] denotes a reference loudspeaker distance estimation. Note that distances are underestimated even for the reference condition. (Kearney et. al. [12]).

Regarding the aspect of distance perception the literature suggests that sound intensity and direct-to-reverberant ratio are the predominant cues for distance estimation, although distance is typically underestimated in real/reference and auralized conditions. This has been investigated both in loudspeaker array experiments [12] (Fig. 2.2) and headphone auralization experiments [13] (Fig. 2.3). Both studies used the method of direct blind walking as answering protocal (participants walked the estimated distance).

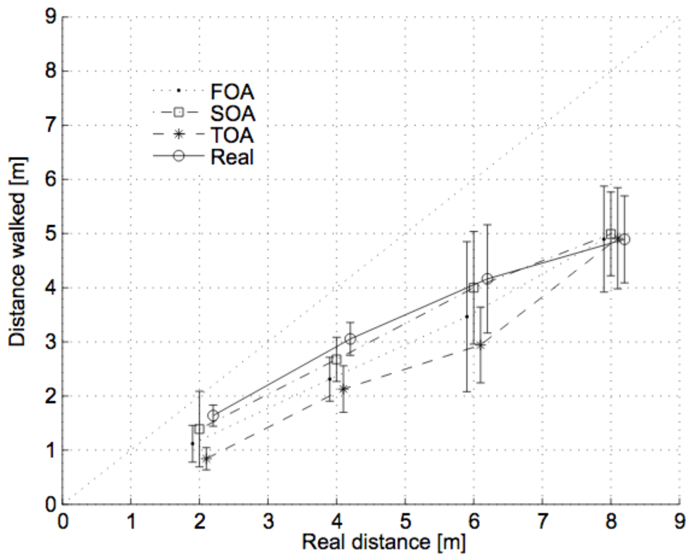

Figure 2.3: Perceived distance vs. simulated/reference distance. *FOA* to *TOA* denote ambisonic order in headphone auralizations of spatial room impulse responses of different distances. *Real* denotes reference loudspeaker distance estimation (Kearney et. al. [13]).

3 Objectives and Listening Experiments

Surrounding Loudspeaker Arrays

Listening experiments with surrounding loudspeaker arrays could investigate the effect of the following variables on perceived envelopment:

- Arrival Directions: Can we confirm or extend previous findings?
- Signal Decorrelation: Decorrelation algorithms for Room/Source Envelopment compared to Ensemble Envelopment (multiple source signals).
- Ambisonic Order: What spatial resolution is necesarry to transmit these informations to the ears (measuring ITD/ILD fluctuations, IACC)?
- Off-Center Listening: How large is the sweet spot of envelopment under typical conditions?

Compact Spherical Loudspeaker Arrays

Compact spherical loudspeaker arrays are beamformers that excite wall reflections in order to achieve a sense of envelopment and spaciousness [19]. Recent work at the Insitute of Electronic Music and Acousitcs (IEM) resulted in low-cost mixed-order arrays with increased horizontal resolution (4th order instead of 3rd order with the IEM icosahedral array) [17]. A listening experiment shall investigate the conditions under which this increased resolution allows for the perception of auditory events behind the listener. Informal listening experiments indicated that the array, placed in front of the listener, can produce auditory objects with more than 90° lateralization, which could subsequently allow for a higher degree of listener envelopment.

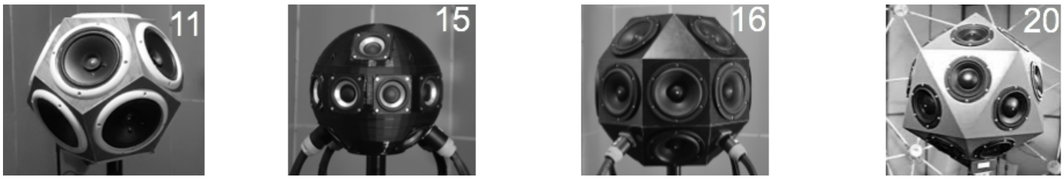

Figure 3.1: Compact spherical loudspeaker arrays at the IEM and their number of transducers. The 3|9|3 array ($3 + 9 + 3 = 15$ loudspeakers) is a promising prototype of a mixed-order array that needs further evaluation through listening experiments.

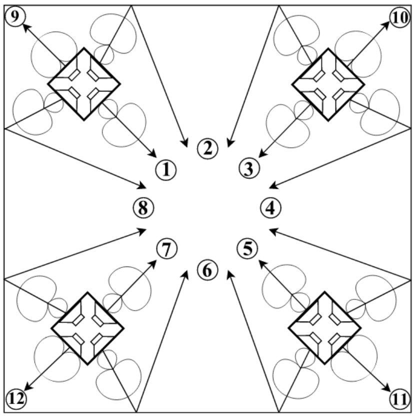

Figure 3.2: Principle of surround-with-depth using first-order cubical loudspeakers as published by Deppisch et. al. [6]. With multiple higher order beamformers (higher spatial resolution) a more enveloping sound scene might be achieved.

Surrounding Compact Spherical Arrays

While the idea of employing multiple compact arrays for sound field synthesis is not new [2, 6], the case of multiple higher order beamformers is more or less uninvestigated in practice. With e.g. three compact mixed-order arrays (4th order horizontal beamforming), a new platform for performances and research will be established. The higher resolution compared to first-order arrays might lead to a more enveloping experience, especially considering larger performance spaces and listening areas. Subsequently, experiments with one of the arrays being tilted (to employ the high resolution to excite ceiling reflections) could be conducted to investigate if the perception of envelopment can be further enhanced.

Bibliography

[1] J. Berg and F. Rumsey. "Systematic evaluation of perceived spatial quality." In: *AES 24th International Conference on Multichannel, Audio26-28 June 2003, Banff, Alberta, Canada*. 2003.

[2] T Betlehem, C Anderson, and MA Poletti. "A directional loudspeaker array for surround sound in reverberant rooms." In: *Proc. Int. Conf. Acoust.(ICA'10)*. 2010, pp. 1–6.

[3] J. Blauert and W. Lindemann. "Auditory spaciousness: Some further psychoacoustic analyses." In: *The Journal of the Acoustical Society of America* 80.2 (1986), pp. 533–542.

[4] A. D. Blumlein. "Improvements in and relating to Sound-transmission." In: *Sound-recording and Sound-reproducing Systems, UK Patent* 394 (1931), p. 325.

[5] J. S. Bradley and G. A. Soulodre. "Objective measures of listener envelopment." In: *The Journal of the Acoustical Society of America* 98.5 (1995), pp. 2590–2597.

[6] T. Deppisch et al. "Surround with depth on first-order beam-controlling loudspeakers." In: *Audio Engineering Society Convention 144*. Audio Engineering Society. 2018.

[7] D. A. Dick and M. C. Vigeant. "An investigation of listener envelopment utilizing a spherical microphone array and third-order ambisonics reproduction." In: *The Journal of the Acoustical Society of America* 145.4 (2019), pp. 2795–2809.

[8] H. Furuya et al. "Arrival direction of late sound and listener envelopment." In: *Applied Acoustics* 62.2 (2001), pp. 125–136.

[9] D. Griesinger. "General overview of spatial impression, envelopment, localization, and externalization." In: *Audio Engineering Society Conference: 15th International Conference: Audio, Acoustics & Small Spaces.* Audio Engineering Society. 1998.

[10] D. Griesinger. "The psychoacoustics of apparent source width, spaciousness and envelopment in performance spaces." In: *Acta Acustica united with Acustica* 83.4 (1997), pp. 721–731.

[11] T. Hidaka, T. Okano, and L. Beranek. "Interaural cross correlation (IACC) as a measure of spaciousness and envelopment in concert halls." In: *The Journal of the Acoustical Society of America* 92.4 (1992), pp. 2469–2469.

[12] G. Kearney, M. Gorzel, and F. Boland. "On loudspeaker rendering of auditorydistance in higher order Ambisonics." In: *Proc. of the Acoustics 2012 Nantes Conference.* 2012.

[13] G. Kearney et al. "Distance perception in interactive virtual acoustic environments using first and higher order ambisonic sound fields." In: *Acta Acustica united with Acustica* 98.1 (2012), pp. 61–71.

[14] G. S. Kendall. "Spatial perception and cognition in multichannel audio for electroacoustic music." In: *Organised Sound* 15.3 (2010), pp. 228–238.

[15] G. S. Kendall. "The decorrelation of audio signals and its impact on spatial imagery." In: *Computer Music Journal* 19.4 (1995), pp. 71–87.

[16] M. Morimoto. "Auditory spaciousness and envelopment." In: *Proc. 13th ICA Belgrade* 2 (1989), pp. 215–218.

[17] S. Riedel, F. Zotter, and R. Hoeldrich. "Design and Control of Mixed-Order Spherical Loudspeaker Arrays." In: *International Computer Music Conference.* ICMA. 2019.

[18] S. Spors et al. "Spatial sound with loudspeakers and its perception: A review of the current state." In: *Proceedings of the IEEE* 101.9 (2013), pp. 1920–1938.

[19] F. Zotter et al. "A beamformer to play with wall reflections: The icosahedral loudspeaker." In: *Computer Music Journal* 41.3 (2017), pp. 50–68.

[20] F. Zotter et al. "Efficient phantom source widening and diffuseness in ambisonics." In: *Proc. of the EAA Joint Symposium on Auralization and Ambisonics* (2014).

[21] F. Zotter et al. "Phantom source widening with deterministic frequency dependent time delays." In: *International conference on digital audio effects (DAFx). Paris, France*. 2011.

Envelopment and Distance in Spatial Audio

Stefan Riedel

Abstract

This work investigates the perception of envelopment and distance in spatial audio systems. While previous research in the field of room acoustics has proposed predictors and simple objective measures for listener envelopment, recent studies revealed contradictory results. The listening experiments in this work shall lead to psychoacoustic criteria that explain listener envelopment in real and virtual immersive listening environments.

The experiments are conducted with two types of loudspeaker arrays. Firstly, traditional surrounding loudspeaker arrays to investigate different directions of arriving sound and the effects of signal decorrelation. Secondly, experiments with compact spherical loudspeaker arrays which explicitly excite wall reflections to create listener envelopment.

Kurzfassung

Diese Arbeit untersucht die Wahrnehmung von Einhüllung und Distanz in räumlichen Wiedergabesystemen. Vorhergehende Untersuchungen im Feld der Raumakustik schlagen objektive Maße zur Beurteilung von Einhüllung vor, jedoch zeigen neuere Studien widersprüchliche Ergebnisse. Die Hörversuche in dieser Arbeit sollen zu psychoakustischen Kriterien führen, die Einhüllung in realen und virtuellen immersiven Hörumgebungen erklären.

Hörversuche sollen mit zwei Typen von Lautsprecheranordnungen durchgeführt werden. Einerseits soll mit traditionellen umgebenden Lautsprecheranordnungen der Einfluss verschiedener Schalleinfallsrichtungen und der Signaldekorrelation untersucht werden. Andererseits sollen Experimente mit kompakten Kugellautsprecheranordnungen, welche gezielt Wandreflektionen anregen um Einhüllung zu erzeugen, durchgeführt werden.