Hearing and Feeling: Immersion in Audio

Roman Kiyan, Stephan Preihs, Jürgen Peissig

Institut für Kommunikationstechnik, Leibniz Universität Hannover

{roman.kiyan, stephan.preihs, peissig}@ikt.uni-hannover.de

Introduction

Spatial audio and virtual acoustics are concerned with recreating or synthesizing sound fields that produce sensations similar to those experienced in natural spatial hearing. This is achieved by reproducing relevant physical properties of a sound field and and therefore certain perceptual cues. Whilst analyzing these objective metrics, current research is also concerned with how sensory cues evoke higher-level reactions in a cognitive and emotional sense. In this context, the notion of immersion is understood as a psychological construct – the definition of immersion thus lies in the domain of psychology, as does the development and validation of experimental paradigms used to assess it. The role of acoustics and audio research, on the other hand, is to examine how measurable sound field properties are related to immersive experience beyond basic perceptual attributes.

Paradigms in spatial audio evaluation

In the evaluation of audio technologies, one major goal is to assess the merits of technical systems in terms of their effects on human perception. Typically, perceptual evaluation methods used in audio evaluation attempt to exclude personal preference or emotion from their assessments [1]. As a consequence, common inventories of perceptual attributes – such as the Spatial Audio Quality Inventory (SAQI) [2] – focus on the "perceptual characterization of a simulation's technical shortcomings" [2]. The overall experience provided by a spatial audio system, however, does include aspects beyond technical properties. This disparity may be formalized by distinguishing between between the concepts of basic audio quality (BAQ) and overall listening experience (OLE) [3], with immersion being related to the latter.

It is important to note that there is no clear terminological consensus on *immersion*, *presence* and related attributes as different terms are used in describing the sensation of being spatially included in a scene and being absorbed in an activity in a cognitive sense [4, 5]. A distinction between these aspects can be made through the distinction between *spatial* and *emotional* immersion [6]. The Immersive Musical Experience Inventory (IMEI) [7], an inventory for the assessment of immersion in music listening, is based on a definition of immersion synthesized from that of Witmer and Singer [8] and that of Georgiou and Kyza [9], combining aspects of spatial envelopment as well as cognitive absorption.

Current research on immersion in audio

Several studies investigating perceptual differences between spatial audio reproduction formats in terms of immersion and similar concepts have been conducted. The results of such studies differ for BAQ and OLE [10], suggesting that the incorporation of OLE-like aspects represents a non-redundant addition to the evaluation of a spatial audio system. Specifically, OLE (unlike BAQ) was found to tend towards a saturation limit with an increasing number of loudspeakers [10]. The phenomenon of more loudspeakers delivering diminishing returns with respect to immersion is observed across several studies [10, 11, 12]. The exact number of loudspeakers leading to this saturation effect varies between studies, however. This is to be expected given the different experimental conditions. However, being able to analyze this phenomenon independently from a particular experimental setup would be desirable.

Another common observation is the content-dependency of immersion and similar attributes [11, 12, 13] as well as the dependency on production and downmixing techniques employed in the creation of mixes for different loudspeaker configurations [14, 15]. While this dependency should ideally be controlled, when evaluating immersion and OLE it is essential that ecological validity is ensured by using stimuli representing examples of real-world audio content. Hence, ways to account for the content-dependency need to be developed.

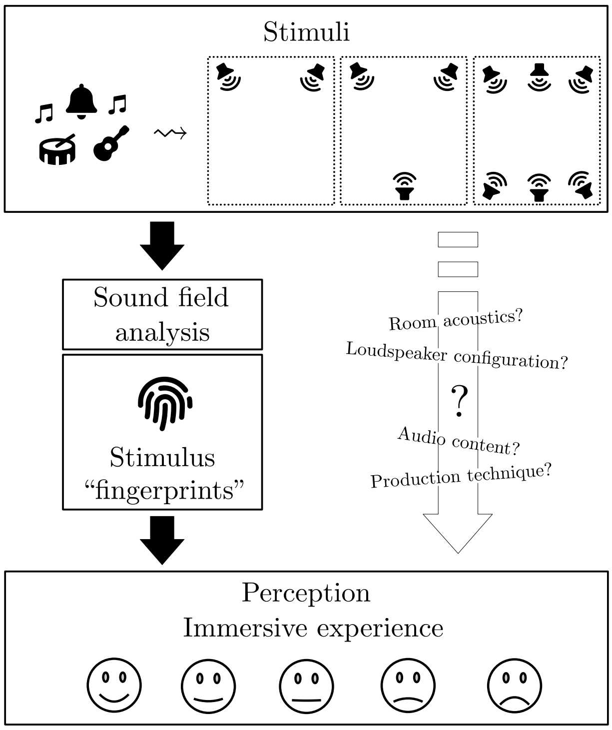

Figure 1: Illustration of an immersive experience modeling approach through sound field analysis.

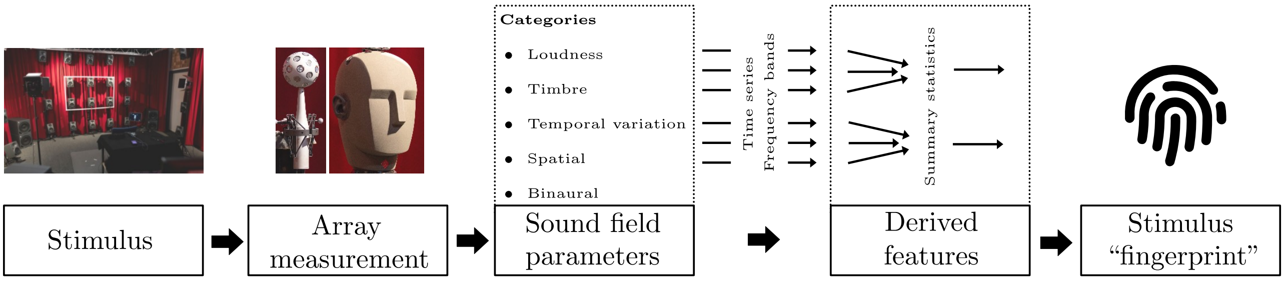

Figure 2: Sound field analysis approach.

Sound field analysis framework

When studying immersion in audio rendered by different reproduction techniques, a potential modeling approach would preferably be de-coupled from the specific experimental configuration and the content used to assess immersion. Based on concepts and techniques originating from the field of soundscape fingerprinting [16, 17], the immersive experience induced by a stimulus may be modeled through a more general acoustic *fingerprint* of each stimulus rather than by comparing between specific experimental conditions. Such a fingerprint can be derived by evaluating acoustic properties of the sound field produced by a particular stimulus in a given experimental setup. As illustrated by Figure 1, this can help to alleviate the uncertainty in investigating immersive experience introduced by the experimental conditions.

The methodology used by the authors in a recent study on immersive musical experience in multichannel music reproduction [18, 19] is depicted in Figure 2. Based on Higher Order Ambisonics (HOA) re-recordings of the musical stimuli using a spherical microphone array, various sound field parameters were computed which can be categorized into loudness, timbral, temporal, spatial and binaural (computed from the HOA recordings by binaural rendering). In the framework used for sound field analysis [20, 21], the sound field parameters are computed as time series in frequency bands, from which simplified features forming a stimulus *fingerprint* may be derived by computing summary statistics. However, even these simplified features are numerous, requiring further feature selection to be applied in order to arrive at a reduced set of meaningful features.

Modeling immersive experience based on sound field features

In the recent study by the authors [18] within the scope of the research project *Richard Wagner 3.0* [22], immersion was modeled based on sound field features using a linear mixed-effects approach. Immersion ratings were obtained from a listening experiment with 57 naive subjects in the Immersive Media Lab (IML) listening room [23] at the Institute of Communications Technology. A feature selection procedure resulted in inter-aural cross-correlation (IACC) and diffuseness emerging as particularly relevant features. Such features being commonly associated with attributes such as listener envelopment or spaciousness [24, 25, 26] is in line with the spatial presence component of the assumed definition of immersion. Notably, the phenomenon of immersion reaching a point of saturation as well as – to a certain extent – the content-dependency of immersion ratings were shown to be predictable from the sound field features of the stimuli. However, a strong inter-personal dependency of immersion ratings and an interaction between personal and content effects remain, highlighting the need for further investigation of the emotional and cognitive aspects of immersive musical experience.

Summary and outlook

In audio research, notions of listening experience beyond the common perceptual attributes are becoming increasingly relevant. One such notion is that of *immersion* which can be defined to be formed by a sense of spatial presence combined with cognitive absorption in the listening process. In the evaluation of immersion evoked by varying technical parameters of an audio system, using a sound field analysis approach can be argued to be advantageous over comparing immersion ratings for different stimulus conditions, as this can help to explain differences in immersion ratings as well as implicitly accounting for different experimental conditions. This may help to more generally model perceived immersion based on the actual sound field experienced by a listener.

In this context, it is crucial to validate that the sound field features used in immersion modeling are representative of what a listener is hearing. This concerns both the computation of the sound field features from acoustic measurement results as well as the acoustic measurements themselves – with the characterization of the area around a reproduction system's sweet spot yielding a consistent perceptual experience being of particular interest.

References

[1] F. Rumsey. "Spatial quality evaluation for reproduced sound: Terminology, meaning, and a scene-based paradigm". In: *Journal of the Audio Engineering Society* 50.9 (2002), pp. 651–666. URL: `https://www.aes.org/e-lib/browse.cfm?elib=11067`.

[2] A. Lindau, V. Erbes, S. Lepa, H.-J. Maempel, F. Brinkman, and S. Weinzierl. "A Spatial Audio Quality Inventory (SAQI)". In: *Acta Acustica united with Acustica* 100.5 (2014), pp. 984–994. DOI: `10.3813/aaa.918778`.

[3] M. Schoeffler and J. Herre. "The relationship between basic audio quality and overall listening ex-

perience". In: *The Journal of the Acoustical Society of America* 140.3 (2016), pp. 2101–2112. DOI: 10.1121/1.4963078.

[4] C. Jennett, A. L. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton. "Measuring and defining the experience of immersion in games". In: *International Journal of Human-Computer Studies* 66.9 (2008), pp. 641–661. DOI: 10.1016/j.ijhcs.2008.04.004.

[5] N. C. Nilsson, R. Nordahl, and S. Serafin. "Immersion Revisited: A review of existing definitions of immersion and their relation to different theories of presence". In: *Human Technology* 12.2 (2016), pp. 108–134. DOI: 10.17011/ht/urn.201611174652.

[6] C. Zhang, A. Perkis, and S. Arndt. "Spatial immersion versus emotional immersion, which is more immersive?" In: *2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX)*. IEEE, 2017. DOI: 10.1109/qomex.2017.7965655.

[7] Y. Wycisk, K. Sander, R. Kopiez, F. Platz, S. Preihs, and J. Peissig. "Wrapped into sound: Development of the Immersive Music Experience Inventory (IMEI)". In: *Frontiers in Psychology* 13 (2022). DOI: 10.3389/fpsyg.2022.951161.

[8] B. G. Witmer and M. J. Singer. "Measuring Presence in Virtual Environments: A Presence Questionnaire". In: *Presence: Teleoperators and Virtual Environments* 7.3 (1998), pp. 225–240. DOI: 10.1162/105474698565686.

[9] Y. Georgiou and E. A. Kyza. "The development and validation of the ARI questionnaire: An instrument for measuring immersion in location-based augmented reality settings". In: *International Journal of Human-Computer Studies* 98 (2017), pp. 24–37. DOI: 10.1016/j.ijhcs.2016.09.014.

[10] M. Schoeffler, A. Silzle, and J. Herre. "Evaluation of Spatial/3D Audio: Basic Audio Quality Versus Quality of Experience". In: *IEEE Journal of Selected Topics in Signal Processing* 11.1 (2017), pp. 75–88. DOI: 10.1109/jstsp.2016.2639325.

[11] S. Agrawal, S. Bech, K. D. Moor, and S. Forchhammer. "Influence of Changes in Audio Spatialization on Immersion in Audiovisual Experiences". In: *Journal of the Audio Engineering Society* 70.10 (2022), pp. 810–823. DOI: 10.17743/jaes.2022.0034.

[12] J. Francombe, T. Brookes, R. Mason, and J. Woodcock. "Evaluation of Spatial Audio Reproduction Methods (Part 2): Analysis of Listener Preference". In: *Journal of the Audio Engineering Society* 65.3 (2017), pp. 212–225. DOI: 10.17743/jaes.2016.0071.

[13] E. Hahn. "Musical Emotions Evoked by 3D Audio". In: *Audio Engineering Society Conference: 2018 AES International Conference on Spatial Reproduction-Aesthetics and Science*. 2018. URL: https://www.aes.org/e-lib/browse.cfm?elib=19640.

[14] A. Silzle, S. George, E. A. Habets, and T. Bachmann. "Investigation on the quality of 3D sound reproduction". In: *Proceedings of the International Conference on Spatial Audio*. 2011, pp. 334–341.

[15] C. Eaton and H. Lee. "Subjective evaluations of three-dimensional, surround and stereo loudspeaker reproductions using classical music recordings". In: *Acoustical Science and Technology* 43.2 (2022), pp. 149–161. DOI: 10.1250/ast.43.149.

[16] J. Bergner and J. Peissig. "On the identification and assessment of underlying acoustic dimensions of soundscapes". In: *Acta Acustica* 6 (2022). DOI: 10.1051/aacus/2022042.

[17] J. Bergner, D. Schössow, S. Preihs, and J. Peissig. "Identification of discriminative acoustic dimensions in stereo, surround and 3D music reproduction". In: *Journal of the Audio Engineering Society* 71.7/8 (2023), pp. 420–430. DOI: 10.17743/jaes.2022.0071.

[18] R. Kiyan, J. Bergner, S. Preihs, Y. Wycisk, D. Schössow, K. Sander, J. Peissig, and R. Kopiez. "Towards predicting immersion in surround sound music reproduction from sound field features". In: *Acta Acustica* 7.45 (2023). DOI: 10.1051/aacus/2023040.

[19] R. Kiyan, J. Bergner, S. Preihs, Y. Wycisk, D. Schössow, K. Sander, J. Peissig, and R. Kopiez. *Immersive music experience in surround sound music reproduction*. [Data set]. DOI: 10.25835/3vx9ls5h.

[20] J. Bergner. *Soundscape Analysis - Indicators*. [Code repository]. URL: https://gitlab.com/janywhere/sosca-indicators/.

[21] R. Kiyan. *immersionmodeling*. [Code repository]. URL: https://gitlab.uni-hannover.de/roman.kiyan.jr/immersionmodeling.

[22] *Richard Wagner 3.0*. URL: https://richard-wagner3-0.de/.

[23] R. Hupke, M. Nophut, S. Li, R. Schlieper, S. Preihs, and J. Peissig. "The immersive media laboratory: Installation of a novel multichannel audio laboratory for immersive media applications". In: *Audio Engineering Society Convention 144*. 2018. URL: https://www.aes.org/e-lib/browse.cfm?elib=19522.

[24] S. Klockgether and S. van de Par. "A Model for the Prediction of Room Acoustical Perception Based on the Just Noticeable Differences of Spatial Perception". In: *Acta Acustica united with Acustica* 100.5 (Sept. 2014), pp. 964–971. DOI: 10.3813/aaa.918776.

[25] M. Morimoto, K. Iida, and Y. Furue. "Relation between auditory source width in various sound fields and degree of interaural cross-correlation". In: *Applied Acoustics* 38.2-4 (1993), pp. 291–301. DOI: 10.1016/0003-682x(93)90057-d.

[26] 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. DOI: 10.1121/1.404472.