

Perzeptive Evaluation transauraler Binauralsynthese unter Berücksichtigung des Wiedergaberaums

Gunar Schlenstedt¹, Fabian Brinkmann¹, Sönke Pelzer², Stefan Weinzierl¹

¹ Technische Universität Berlin, Fachgebiet Audiokommunikation, Email: fabian.brinkmann@tu-berlin.de

² Audioborn, Köln, Email: pelzer@audioborn.com

Einleitung

Die Schalldruckverläufe an den Ohren einer Person enthalten sämtliche räumliche Cues eines auditiven Ereignisses und können mittels raumakustischer Modelle und Binauralsynthese simuliert und reproduziert werden. Wird eine solche Simulation nicht mit Kopfhörern, sondern unter Verwendung von zwei oder mehreren Lautsprechern realisiert (transaurale Binauralsynthese), so hängt die Güte der Simulation maßgeblich von der durch digitale Filter realisierten Übersprehdämpfung der Lautsprecher zum jeweils abgewandten Ohr ab (Crosstalk cancellation, CTC). Obwohl diese Filter in der Regel nur den Direktschall, nicht aber Reflexionen des Wiedergaberaums kompensieren, wird die transaurale Binauralsynthese häufig auch in nachhallbehafteten Umgebungen angewandt, beispielsweise um eine gleichzeitige Präsentation visueller und akustischer Reize zu ermöglichen. Demzufolge werden die binauralen Signale durch Reflexionen des Wiedergaberaums verändert und somit womöglich auch das resultierende auditive Ereignis beeinträchtigt. Diese Studie beschäftigt sich daher mit den perzeptiven Auswirkungen nachhallbehafteter Wiedergaberäume auf die Qualität der transauralen Binauralsynthese im Vergleich zu einer binauralen Referenz. Die Ergebnisse zeigen einen mit steigender Nachhallzeit des Wiedergaberaums einhergehenden Qualitätsverlust der transauralen Synthese. Im Extremfall ist dieser größer, als die durch die Verwendung nicht-individueller CTC-Filter verursachte Beeinträchtigung und die Qualität ist vergleichbar mit herkömmlicher Stereo-Wiedergabe ohne Verwendung von CTC-Filtern.

Transaurale Binauralsynthese

Bei herkömmlicher Stereo-Wiedergabe erreichen Signale des linken und rechten Lautsprechers beide Ohren nahezu gleichermaßen (s. Abb. 1, rechts). Ziel der transauralen Binauralsynthese ist es, das in diesem Fall ungewollte Übersprechen der Lautsprecher zum jeweils abgewandtem Ohr H_{LR} , bzw. H_{RL} durch digitale CTC-Filter

$$\mathbf{C} = (\mathbf{H}^* \mathbf{H} + \beta \mathbf{I})^{-1} \mathbf{H}^* \quad (1)$$

zu unterdrücken [1, 2]. Gleichung (1) liefert die Filter \mathbf{C} durch regularisierte Invertierung der Lautsprecherübertragungsfunktionen \mathbf{H} , gegeben durch

$$\mathbf{C} = \begin{bmatrix} C_{LL} & C_{RL} \\ C_{LR} & C_{RR} \end{bmatrix}, \text{ und } \mathbf{H} = \begin{bmatrix} H_{LL} & H_{RL} \\ H_{LR} & H_{RR} \end{bmatrix}. \quad (2)$$

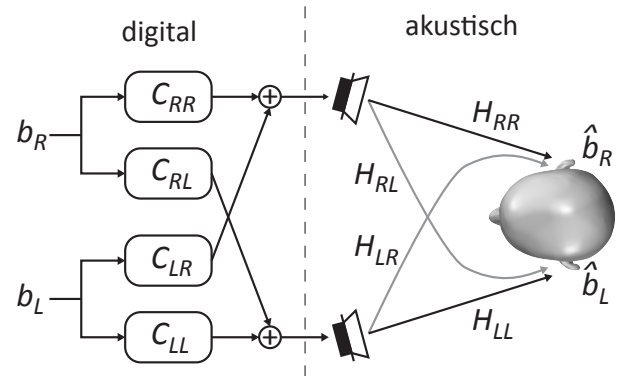


Abbildung 1: Blockschaltbild der transauralen Binauralsynthese. Links: Die auf das binaurale Eingangssignal $b_{L,R}$ angewandten CTC-Filter C_{LL} , C_{LR} , C_{RR} und C_{RL} . Rechts: Die akustische Übertragung des linken Lautsprechers zum linken (H_{LL}) und rechten Ohr (H_{LR}), sowie vom rechten Lautsprecher zum rechten (H_{RR}) und linken Ohr (H_{RL}) und die daraus resultierenden Ohrsignale \hat{b}_L und \hat{b}_R .

Dabei sind \mathbf{I} die Einheitsmatrix und $(*)$ die hermitesche Transponierung. Der Regularisierungsparameter β limitiert die Filterverstärkung für den Fall $\beta \neq 0$. Andernfalls liefert (1) die Moore-Penrose Pseudoinverse von \mathbf{H} . Für bessere Lesbarkeit wurde die Frequenzabhängigkeit in (1) und (2) ausgespart. Moore *et al.* [3] konnten im Hörversuch zeigen, dass mit einem System nach Abb. 1 in einem nachhallfreien Wiedergaberaum eine virtuelle Schallquelle erzeugt werden kann, die sich von einer entsprechenden realen Quelle nahezu nicht unterscheiden lässt.

In nachhallbehafteten Umgebungen, setzen sich die in Abb. 1 angedeuteten Übertragungsfunktionen \mathbf{H} neben dem Direktschall auch aus zahlreichen Reflexionen zusammen. Werden etwaige Reflexionen nicht kompensiert, fügen sie dem wiederzugebendem Signal $b_{L,R}$ Energie hinzu. Dies verschlechtert die Übersprehdämpfung am abgewandten Ohr [4], beeinträchtigt das Ausklingverhalten [5], die Lautheit und bedingt durch Kammfiltereffekte auch die wahrgenommene Klangfarbe [6, 7].

Sæbø [8], Politis *et al.* [9] und Kosmidis *et al.* [10] haben den Einfluss einzelner früher Reflexionen auf die Lokalisierbarkeit einer Quelle unter Verwendung transauraler Synthese untersucht. Während Kosmidis *et al.* keinen Einfluss nachweisen konnten, berichten Politis *et al.* und Sæbø von einem Einbruch der Lokalisationsleistung, bzw. von einer Verschiebung der Wahrgenommenen Quellposition in Richtung der Reflexionen ohne diese jedoch inferenzstatistisch zu belegen. Bei Berücksichtigung der

Reflexionen in der CTC-Filterberechnung, verbesserte sich die Lokalisierbarkeit bei Politis *et al.*, verschlechterte sich jedoch bei Sæebø, wofür dieser die Empfindlichkeit der Filter auf Positionsänderungen der Versuchspersonen anführte.

Den Einfluss realer Wiedergaberäume auf die Lokalisationsleistung transauraler Binauralsynthese wurde zahlreich untersucht [11, 12, 13, 14]. Für Räume mit geringen Nachhallzeiten um 0.4 s wurden in allen Studien gute Ergebnisse erzielt, unabhängig davon, ob der Direktschall oder die gesamte binaurale Raumimpulsantwort für die Berechnung der CTC-Filter herangezogen wurde. Für eine Nachhallzeit von 1 s verschlechterte sich die Lokalisierbarkeit von Quellen jedoch besonders in der Medianebene [12].

Der widersprüchliche Einfluss der Reflexionen auf die Lokalisierbarkeit könnte auf den Präzedenzeffekt zurückgeführt werden. Demzufolge werden 5 - 40 ms nach dem Direktschall eintreffende Reflexionen nicht als eigenständiges auditives Objekt wahrgenommen, sondern fusionieren mit dem durch den Direktschall hervorgerufenen Hörereignis. Eine Veränderung der wahrgenommenen Position einer Schallquelle in Richtung einer Reflexion ist jedoch nur für Verzögerungen unter 5 ms oder distinkte Reflexionen, deren Pegel deutlich über dem des Direktschalls liegen, zu erwarten [15, S. 271 ff.], [16, S. 266 ff.].

Die oben genannten Studien beschränkten sich fast ausschließlich auf die Lokalisationsleistung bei transauraler Synthese unter Verwendung nachhallfreier binaurale Signale $b_{L,R}$. In dieser Studie wurde daher der perzeptive Einfluss nachhallbehalteter Wiedergabeumgebungen auf die transaurale Synthese mittels eines qualitativen Fragenkataloges unter Verwendung ebenfalls nachhallbehalteter binauraler Signale untersucht.

Synthese der Wiedergabeumgebungen

Um später einen direkten Vergleich verschiedener Wiedergabeumgebungen zu ermöglichen, wurden drei geometrisch identische Räume mittels raumakustischer Simulation erzeugt [17]. In Anlehnung an existierende audiovisuelle Simulationsumgebungen [18] wurde eine Grundfläche von 5.33 x 5.38 m bei einer Höhe von 4 m gewählt und je ein Lautsprecher in der Wandmitte auf einer Höhe von 3.5 m, platziert. Der leicht unregelmäßige Grundriss wurde gewählt, um durch die Simulation entstehende, unnatürliche Bündelungseffekte zu vermeiden. Für die Wiedergabeumgebung *Freifeld* wurden alle Absorptionskoeffizienten auf eins gesetzt, während für einen *bedämpften* und *unbedämpften* Raum perforierte Wandflächen ($\alpha_{\text{mean}} = 0.52$), bzw. Acrylglas ($\alpha_{\text{mean}} = 0.08$) verwendet wurden. Die Decke wurde in allen Fällen absorbierend ($\alpha_{\text{mean}} = 0.95$) gestaltet, woraus Nachhallzeiten von 0 s, 0.2 s und 0.9 s resultierten.

Die in **H** enthaltenen Übertragungsfunktionen wurden für zwei im Zentrum des Raums stehende Empfänger (FABIAN [19], Neumann KU100 [20]) und horizontale Kopforientierungen im Bereich von $\pm 44^\circ$ mit einer Auflösung von 2° simuliert. Dabei entspricht die 0° -

Orientierung einer in die Raumecke blickenden Person. Die ausschließlich den Direktschall kompensierenden CTC-Filter wurden mittels regularisierter Invertierung ($\beta = 0,001$) unter Verwendung der ITA-Toolbox [21] für die mit FABIAN simulierten Daten berechnet. Entsprechend der Forschungsfrage blieben die Reflexionen der Wiedergabeumgebung bei der Filterberechnung unberücksichtigt.

Perzeptive Evaluation

Die perzeptive Evaluation umfasste acht Bedingungsvariationen, die sich aus einer vollständigen Variation der beiden Faktoren *Wiedergabesystem* (A: Freifeld mit CTC; B: bedämpfter Raum mit CTC; C: unbedämpfter Raum mit CTC; D: unbedämpfter Raum ohne CTC) und *Individualisierung* (a: angepasst; u: unangepasst) ergaben. Dabei wurde die transaurale Wiedergabe ohne CTC-Filter als Ankerbedingung eingeführt. Im Fall der angepassten Binauralsynthese basierten alle Signale der Wiedergabekette ($b_{L,R}$, **H** und **C**) auf FABIAN, während bei der nicht angepassten Synthese die Raumsimulation (**H**) mit KU 100 verwendet wurde.

Als binaurales Eingangssignal $b_{L,R}$ wurde gepulstes rosa Rauschen (1 s Rauschen, 0.75 s Stille, 0.02 s Rampen), gefaltet mit binauralen Raumimpulsantworten (BRIR) des unbesetzten Renaissance-Theaters Berlin, verwendet (T=0.7 s; QSC K8 Lautsprecher auf Bühnenmitte; FABIAN in 10 m Entfernung auf den Lautsprecher ausgerichtet). Es wurde so gewählt, dass dessen Nachhallzeit zwischen denen von Wiedergabesystem B und C liegt.

Unter den Annahmen mittlerer bis großer Effekte ($f=0,33$) und einer späteren varianzanalytischen Auswertung ($\alpha = 5\%$, Teststärke $1 - \beta = 0.8$, Korrelation zwischen Messwiederholungen 0.15, Sphärizitätskorrektur $\epsilon = 0.7$) wurde eine Stichprobengröße von 19 Personen festgelegt [22] die aus 3 Frauen und 16 Männer mit einem durchschnittlichen Alter von 31 Jahren bestand. Der Versuch wurde in einer ruhigen Laborumgebung durchgeführt, wobei die aus den acht Bedingungsvariationen resultierenden Ohrsignale $\hat{b}_{L,R}$ über geeignet entzerrte Kopfhörer dargeboten wurden. Eine Beschreibung der für die Auralisation verwendeten Softwareumgebung findet sich bei Brinkmann *et al.* [23].

Die Aufgabe der Versuchspersonen war die transaurale Binauralsynthese anhand 16 ausgewählter Qualitäten des Spatial Audio Quality Inventories (SAQI [24]) gegen eine Referenz zu bewerten. Als Referenz wurden die direkt über Kopfhörer wiedergegebenen Eingangssignale $b_{L,R}$ verwendet. Aufgrund der Kürze der Darstellung beschränken wir uns in der Präsentation der Ergebnisse auf eine Analyse der Mittelwerte und Konfidenzintervalle von neun der 16 getesteten Qualitäten.

Abb. 2 (links) zeigt, dass deutliche perzeptive Unterschiede zur binauralen Referenz bestehen. Lediglich für eine angepasste CTC im Freifeld (Aa) sind sie nicht signifikant, da das Konfidenzintervall die Null umschließt. Die Unterschiede zwischen Freifeld (Aa), bedämpfter (Ba) und unbedämpfter Wiedergabeumgebung (Ca) sind signifikant, und steigen entsprechend der Nachhallzeit des Wiedergaberaums. Dazukommend verschlechtert ei-

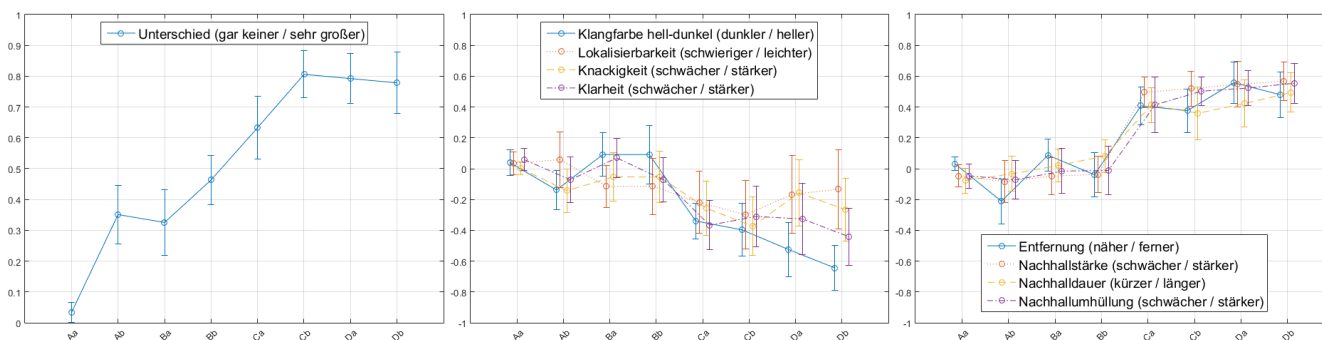


Abbildung 2: Über 19 Versuchspersonen gemittelte Bewertungen und 95% Konfidenzintervallen. Qualitäten und Skalenpole können der Legende entnommen werden, wobei der erstgenannte Skalenpol einem Rating von 0, bzw. -1 entspricht.

ne unangepasste CTC in diesen Fällen die Bewertungen. Bemerkenswert ist, dass ein Wiedergabesystem in unbedämpfter Umgebung ohne CTC (Da, Db) keine signifikanten Unterschiede zu dem mit CTC (Ca, Cb) aufwies und somit ab einem gewissen Nachhallanteil des Wiedergaberaums durch eine CTC keine bedeutende Qualitätssteigerung mehr erreichbar ist.

Weitere Qualitäten mit ähnlichen Mittelwertverläufen sind in Abb. 2 (mitte, rechts) abgebildet. Die klare Abstufung von Bb zu Ca für sämtliche Qualitäten lässt darauf schließen, dass Unterschiede zwischen bedämpften und unbedämpftem Wiedergaberaum deutlich wahrnehmbar und zudem bedeutender sind, als die Unterschiede zwischen angepasster und unangepasster CTC. Dabei nehmen Lokalisierbarkeit, Klarheit und Knackigkeit mit steigender Halligkeit ab, während die wahrgenommene Entfernung, sowie die Nachhallstärke, -dauer und -umhüllung zunehmen, sowie die Klangfarbe als dunkler empfunden wird. Für die Qualitäten Klangfarbe, Entfernung, Nachhallstärke, -dauer und -umhüllung sind diese Unterschiede aufgrund nicht überlappender Konfidenzintervalle als signifikant zu bewerten.

Der Faktor Individualisierung zeigte nur für Wiedergabe im Freifeld (Aa, Ab) signifikant Unterschiede (Unterschied, Entfernung), weswegen davon auszugehen ist, dass die Individualisierung weniger Einfluss auf die Qualität hat, als der Wiedergaberaum.

Diskussion und Ausblick

Die obigen Ergebnisse belegen einen deutlichen Einfluss des Wiedergaberaums auf die Qualität der transauralen Binauralsynthese. Dabei hat sich gezeigt, dass die Nachhallzeit des Wiedergabesystems die des virtuellen Raums – gegeben durch die binauralen Eingangssignale $b_{L,R}$ – nicht überschreiten sollte. Die Überlagerung von Wiedergabe- und virtuellem Raum entspricht dabei einer Faltung der entsprechenden BRIRs und hat eine verlängerte und energiereichere Impulsantwort zur Folge. Dies erklärt die Veränderung der wahrgenommenen Nachhallstärke, -dauer und -umhüllung. Die deutlich wahrgenommenen Klangverfärbungen, veränderte Lokalisierung und Verringerung der Klarheit und Knackigkeit beim unbedämpften System können auf den Einfluss der frühen Reflexionen des Wiedergaberaums und de-

ren Fusionieren mit dem Direktsignal zurückgeführt werden. Möglicherweise wird die Knackigkeit darüber hinaus auch durch das mit der Überlagerung von Wiedergabe- und virtuellem Raum einhergehende zeitliche Verwischen des Direktschallanteils in den BRIRs beeinträchtigt. Die durch die Individualisierung hervorgerufenen Unterschiede konnten als weniger bedeutsam für die Qualität der transauralen Synthese eingestuft werden, als die durch den Wiedergaberaum bedingten. Trotz deutlich geringerem Übersprechen und geringerer Klangverfärbung des angepassten Systems konnten kaum signifikante Unterschiede zum unangepassten festgestellt werden. Dies führt zu dem Schluss, dass die Channel Separation als einzelner Indikator zur Einschätzung der Qualität eines CTC-Wiedergabesystems nicht ausreichend ist.

Insofern sollten wiedergabeseitig raumakustische Parameter wie z.B. die Nachhallzeit RT60, frühe Abklingzeit EDT, Stärkemaß G, Deutlichkeitsmaß C_{50} , sowie frühe Reflexionen (z.B. C_{20}) berücksichtigt werden. Des Weiteren könnten psychoakustische, binaurale Maße, wie beispielsweise die Nachhallumhüllung (LEV) und die wahrgenommene Quellgröße (ASW), für die Qualitätsbewertung der gesamten Wiedergabesituation herangezogen werden.

Im Hinblick auf die Raumakustik von virtuellen Umgebungen muss also weiter nach Möglichkeiten zur Verringerung der Reflexionen an den Raumbegrenzungsflächen geforscht werden. Vor dem Hintergrund der Anforderungen an die Wände, die auch der Videoprojektion dienen, ergeben sich aber konstruktive Einschränkungen. Als Alternativen zu Acrylglasswänden wären Leinwände oder spezielle Projektionsfolien zu nennen, welche jedoch weitaus empfindlicher sind und im Regelfall eine deutlich kürzere Lebensdauer aufweisen.

Ein weiterer Weg frühe Reflexionen zu kompensieren, wäre deren Berücksichtigung in den CTC-Filtern. Dies setzt jedoch entweder eine Kombination aus Positionstracking der Hörer und einem Raummodell oder eine Echtzeitmessung der Übertragungsfunktionen am Ohr der Hörer voraus, um die entsprechenden Reflexionen zu erfassen. Dabei könnten sich jedoch schon kleine Ungenauigkeiten negativ auf das Resultat auswirken.

Die hier beschriebene explorative Studie lieferte erste Erkenntnisse zum Einfluss des Wiedergaberaums auf die Qualität der transauralen Binauralsynthese. Eine geplante varianzanalytische Auswertung kann darüber hinaus systematische Einflüsse der einzelnen Faktoren aufdecken. Sicherlich wäre in diesem Sinne auch interessant, wie groß beispielsweise die Nachhallzeit des Wiedergaberaums sein darf, ohne die in dieser Studie untersuchten Qualitäten zu beeinträchtigen. Das würde allerdings eine größere Stichprobe zur Aufdeckung kleiner Effekte voraussetzen. Im Zuge weiterer Studien könnte zudem überprüft werden, wie stark der Einfluss eines visuellen Stimulus, sowie sich frei bewegender Personen auf die wahrgenommene Qualität ist.

Literatur

- [1] O. Kirkeby and P. A. Nelson, "Digital filter design for inversion problems in sound reproduction," *J. Audio Eng. Soc.*, vol. 47, no. 7/8, pp. 583–595, July/August 1999.
- [2] P. Majdak, B. Masiero, and J. Fels, "Sound localization in individualized and non-individualized crosstalk cancellation systems," *J. Acoust. Soc. Am.*, vol. 133, no. 4, pp. 2055–2068, April 2013.
- [3] A. H. Moore, A. I. Tew, and R. Nicol, "An initial validation of individualised crosstalk cancellation filters for binaural perceptual experiments," *J. Audio Eng. Soc. (Engineering Report)*, vol. 58, no. 1/2, pp. 36–45, January/February 2010.
- [4] D. B. Ward, "On the performance of acoustic crosstalk cancellation in a reverberant environment," *J. Acoust. Soc. Am. (Letter to the Editor)*, vol. 110, no. 2, pp. 1195–1197, August 2001.
- [5] M. Kohnen, S. Pelzer, and M. Vorländer, "3-d audio reproduction in non-ideal environments," in *Fortschritte der Akustik – DAGA 2015*, Nürnberg, Germany, March 2015, pp. 1625–1628.
- [6] M. Barron, "The subjective effects of first reflections in concert halls - the need for lateral reflections," *J. Sound Vib.*, vol. 15, no. 4, pp. 475–494, 1971.
- [7] E. Georganti, N. Dillier, and J. N. Mourjopoulos, "Measuring perception of coloration due to early reflections in binaural room impulse responses," in *Forum Acusticum*. Kraków, Poland: Europ. Acoust. Assoc., September 2014.
- [8] A. Sæbo, "Effect of early reflections in binaural systems with loudspeaker reproduction," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, New York, October 17–20 1999.
- [9] A. Politis, T. Papadopoulos, and P. A. Nelson, "Subjective evaluation of the performance of virtual acoustic imaging systems in suboptimal conditions of implementation," in *4th National Conference of HELINA (Acoustics 2008)*, Xanthi, Greece, October 2008, pp. 483–493.
- [10] D. Kosmidis, Y. Lacoutre-Parodi, and E. A. P. Habets, "The influence of low order reflections on the interaural time differences in crosstalk cancellation systems," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 2873–2877.
- [11] N. Sakamoto, T. Gotoh, T. Kogure, M. Shimbo, and A. H. Clegg, "Controlling sound-image localization in stereophonic reproduction: Part II," *J. Audio Eng. Soc.*, vol. 30, no. 10, pp. 719–722, October 1982.
- [12] J. Köring and A. Schmitz, "Simplifying cancellation of cross-talk for playback of head-related recordings in a two-speaker system," *Acustica*, vol. 79, no. 3, pp. 221–232, November 1993.
- [13] P. A. Nelson, F. Orduna-Bustamante, and D. Engler, "Experiments on a system for the synthesis of virtual acoustic sources," *J. Audio Eng. Soc.*, vol. 44, no. 11, pp. 990–1007, November 1996.
- [14] F. Pausch, L. Aspöck, and J. Fels, "Evaluation of binaural sound reproduction systems with focus on perceptual quality," in *Fortschritte der Akustik – DAGA 2015*, Nürnberg, Germany, March 2015, pp. 1456–1459.
- [15] J. Blauert, *Spatial Hearing. The psychophysics of human sound localization*, revised ed. Cambridge, Massachusetts: MIT Press, 1997.
- [16] B. C. J. Moore, *An introduction to the psychology of hearing*, 6th ed. Leiden, Boston: Brill, 2013.
- [17] S. Pelzer, L. Aspöck, D. Schröder, and M. Vorländer, "Integrating real-time room acoustics simulation into a cad modeling software to enhance the architectural design process," *Building Acoustics*, vol. 4, no. 2, pp. 113–138, April 2014.
- [18] F. Wefers, S. Pelzer, R. Bomhardt, M. Müller-Trapet, and M. Vorländer, "Audiotechnik des aixCAVE Virtual Reality-Systems," in *Fortschritte der Akustik – DAGA 2015*, Nürnberg, Germany, March 2015, pp. 467–470.
- [19] F. Brinkmann, A. Lindau, S. Weinzierl, G. Geissler, and S. van de Par, "A high resolution head-related transfer function database including different orientations of head above the torso," in *AIA-DAGA 2013, International Conference on Acoustics*, Merano, Italy, March 2013, pp. 596–599.
- [20] B. Bernschütz, "A spherical far field HRIR/HRTF compilation of the Neumann KU 100," in *AIA-DAGA 2013, International Conference on Acoustics*, Merano, Italy, March 2013, pp. 592–595.
- [21] P. Dietrich, M. Guski, M. Pollow, M. Müller-Trapet, B. Masiero, R. Scharer, and M. Vorländer, "ITA-Toolbox – An open source Matlab toolbox for acousticians," in *Fortschritte der Akustik – DAGA 2012*, Darmstadt, Germany, March 2012, pp. 151–152.
- [22] F. Faul, E. Erdfelder, A.-G. Lang, and A. Buchner, "G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences," *Behavior Research Methods*, vol. 39, no. 2, pp. 175–191, 2007.
- [23] F. Brinkmann, A. Lindau, M. Vrhovnik, and S. Weinzierl, "Assessing the authenticity of individual dynamic binaural synthesis," in *Proc. of the EAA Joint Symposium on Auralization and Ambisonics*, Berlin, Germany, April 2014, pp. 62–68.
- [24] A. Lindau, V. Erbes, S. Lepa, H.-J. Maempel, F. Brinkmann, and S. Weinzierl, "A Spatial Audio Quality Inventory (SAQI)," *Acta Acust. united Ac.*, vol. 100, no. 5, pp. 984–994, 2014.