

Gemessener Schalldruckpegel im Lautsprecherarray für verschiedene Schallfeldsyntheseverfahren

Matthieu Kuntz, Norbert Kolotzek, Bernhard U. Seeber
 Audio-Signalverarbeitung, Technische Universität München
 E-Mail: {matthieu.kuntz, norbert.kolotzek, seeber}@tum.de

Einleitung

Virtuelle akustische Umgebungen und Schallfeldsyntheseverfahren (SFS-Verfahren) ermöglichen die kontrollierte Auralisierung von Schallszenen von einzelnen Quellen als auch von komplexen, Cocktail-Party-ähnlichen Situationen, in denen sowohl einfache Lokalisierungs- und Detektionsexperimente [1,2], wie auch Studien mit Probanden in Bewegung [3] durchgeführt werden können. Darüber hinaus besteht auch die Möglichkeit, audiologische Messungen im Freifeld durchzuführen, was unter anderem Vorteile bei Hörgeräteträgern mit sich bringt. So kann auch die Verbesserung der Hörhilfe direkt am Probanden gemessen werden [4]. Die Verwendung einer Freifeldauralisation umgeht die Nachteile einer HRTF-basierten Wiedergabe, zum Beispiel das in der Praxis häufig nötige Verwenden eines fremden HRTF-Datensatzes, das Verfolgen der Kopfposition und die Verwendung hierfür notwendiger echtzeitfähiger Algorithmen.

Bisherige Arbeiten haben sich mit der numerischen Evaluierung verschiedener Hardware-Setups oder SFS-Verfahren beschäftigt [5,6] oder mit der perzeptiven Evaluierung der Wiedergabeverfahren durch Lokalisierungs- oder Klangfarbenstudien [7-9]. Wenig ist über die praktisch erreichbare Genauigkeit solcher Systeme bekannt, da viele nicht-ideale Parameter (z.B. Lautsprecherrichtcharakteristik, Ungenauigkeiten in deren Positionierung, Reflektionen an Lautsprechern, Halterungen etc.), die maßgebliche Einflüsse haben können, in Simulationen oft nicht mitberücksichtigt werden. Murillo [10] vermaß die Genauigkeit einer Wiedergabe mit Ambisonics für die *basic*, *in-phase* und *max_{RE}* Dekoder in einem 1.4 m × 1.8 m Raster. Seine Messungen fanden in einem sphärischen Lautsprecherarray und 3D Ambisonics 5. Ordnung statt, für 250 Hz, 1 kHz und 2 kHz, wodurch die Ergebnisse nur schwer auf das hier verwendete Setup übertragen werden können. Grandjean [11] führte ebenfalls ähnliche Messungen durch, die laut eigenen Angaben sich aufgrund von Messungenauigkeiten nicht für eine quantitative Studie eignen.

Um das bestehende Setup besser zu evaluieren führten wir zwei *in-situ* Messungen durch: Eine Messung vor dem Hintergrund der ISO 8253-2 [12] für audiometrische Messungen im Freifeld und eine, die die real-erzielbare Genauigkeit der Wiedergabe von vier gängigen SFS-Methoden als Funktion des Empfangsorts untersucht.

Methoden

Die Messungen wurden im reflexionsarmen Raum (RAR) der Professur für Audio-Signalverarbeitung der TU München durchgeführt.

Hardware

Das verwendete Lautsprecherarray ist Teil der *real-time Simulated Open Field Environment* (rtSOFE, [13]) und besteht aus 36 horizontal angeordneten Lautsprechern Dynaudio BM6A (Dynaudio, Skanderborg, Dänemark) in einem Winkelabstand von 10°. Damit kann ein Schallfeld mit 2D Ambisonics bis zur 17. Ordnung wiedergegeben werden. Die Lautsprecher sind auf einem rechteckigen, maßangefertigten Rahmen mit einer Seitenlänge von 4.8 m montiert und auf dessen Mittelpunkt gerichtet. Die Lautsprecher wurden mit 1024-Punkte FIR Filter laufzeit-, pegel- und phasenentzerrt, sodass sie ein virtuelles kreisförmiges Array bilden.

Die Messungen zur ISO 8253-2 Norm wurden mit einem ½“ Klasse 1 Messmikrofon (MM210, Microtech Gefell, Gefell) über einen 48V-zu-IEPE Adapter an einer Soundkarte RME Micstasy (Audio AG; Haimhausen) durchgeführt, welche über MADI mit einer RME HDSPe Soundkarte im Mess-PC verbunden war. Das Messmikrofon wurde an den Positionen in m (0,0), (0.15,0), (-0.15,0), (0,0.15), (0,-0.15), wie in ISO 8253-2 [13] beschrieben platziert.



Abbildung 1: Fahrbares Mikrofonarray mit 23 Halb Zoll-Messmikrofonen in einem Abstand von 8.5 cm. Dahinter ist ein Teil des Lautsprecherarrays zu sehen.

Für die räumlichen Aufnahmen der synthetisierten Schallfelder wurde ein fahrbares Mikrofonarray mit 23 ½“ Mikrofonen in einem Abstand von 8.5 cm innerhalb des Lautsprecherarrays platziert und von -93.5 cm bis zu 93.5 cm in 23 8.5 cm Schritten positioniert, um ein Raster von 23 × 23 Messpunkten abzudecken. Abbildung 1 zeigt das Mikrofonarray im Lautsprecherarray.

Die IEPE-Speisung der Mikrofone verlief über ein PAK mkII Messsystem (Müller-BBM VAS GmbH, Planegg) mit zwei ICM42 Modulen. Das durchgeschleifte, vorverstärkte Signal wurde von 2 RME Micstasys gewandelt und über MADI an eine RME HDSPe Soundkarte im Mess-PC gesendet.

Stimuli

Als Stimulus wurde Gleichmäßig Anregendes Rauschen (GAR, [14]) verwendet, welches mit einer Bandbreite von ± 0.5 Bark um die audiometrischen Frequenzen von 125 Hz, bis 8 kHz bandpassgefiltert wurde. So entstanden 7 Rauschsignale mit einer Bandbreite von 1 Bark, von 125 Hz bis 8 kHz. Der Pegel wurde nach dem Filtern auf 70 dB SPL, gemessen am Arraymittelpunkt, gesetzt und mit einem 30 ms Fenster ein- und ausgeblendet. Die Stimuli wurden einer virtuellen Quelle in 4 m Entfernung zur Mitte des Lautsprecherarrays und einem horizontalen Winkel von 13° zugeordnet.

Schallfeldsyntheseverfahren

Die Messung wurde für vier verschiedene Schallfeldsynthese- und Panningverfahren durchgeführt: Higher-Order Ambisonics mit dem *basic* Dekoder (HOA *basic*) [15], Higher-Order Ambisonics mit dem *max r_E* Dekoder (HOA *max r_E*) [15], Vector Base Amplitude Panning (VBAP, [16]) und Nearest-Loudspeaker-Mapping (NLS). Beim Nearest-Loudspeaker-Mapping wird die Quelle über den Lautsprecher abgespielt, der den geringsten Winkelunterschied zur Quellposition aufweist, hier der Lautsprecher an 10° .

Ergebnisse

Messung nach ISO 8253-2

Die Vorgabe der ISO 8253-2 für ein Freifeld bezieht sich auf den Pegelunterschied zwischen der Mittenposition, auf der sich der Proband befindet, und 4 Positionen an ± 15 cm in *x*- und *y*-Richtung. Dabei wurde für diese Messung der Stimulus von 0° dargeboten.

Abbildung 2 zeigt den Pegelunterschied für die verwendeten audiometrischen Frequenzen. Alle Messpunkte befinden sich innerhalb der Vorgabe der Norm. Zudem lag der Unterschied zwischen der rechten und der linken Position niemals über 3 dB und die Schalldruckpegel an den Positionen vorne und hinten wichen nicht mehr als 1 dB vom *I/R*-Gesetz ab.

Somit ist die rtSOFE Lautsprecherapparatur im RAR der Professur für Audio-Signalverarbeitung der TU München für audiologische Messungen im Freifeld mit Schmalbandsignale geeignet.

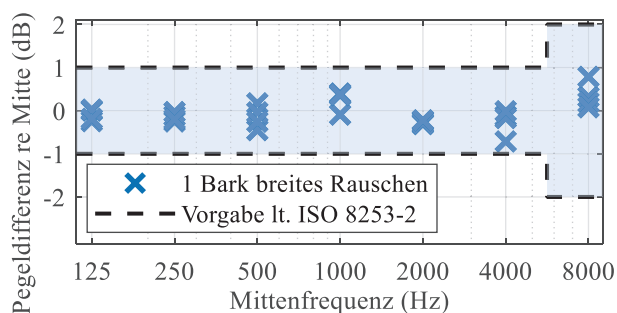


Abbildung 2: Pegelunterschied zwischen den vier exzentrischen Positionen und der Mitte des Lautsprecherarrays (blaue Kreuze), für GAR um die verschiedenen audiometrischen Frequenzen herum. Die Vorgabe der ISO 8253-2 ist durch die gestrichelte Linie gekennzeichnet.

Räumliche Messung des Schallfeldes für verschiedene SFS-Verfahren

Um die verschiedenen SFS-Verfahren evaluieren zu können, werden diese relativ zu einem Referenzschallfeld betrachtet. Hierbei handelt es sich um das berechnete Schallfeld, das von einer Punktquelle an 4 m Entfernung unter 13° abgestrahlt wird und mit der virtuellen Quelle übereinstimmt, die für die SFS-Verfahren verwendet wurde.

Abbildung 3 zeigt den gemessenen, absoluten Pegelfehler relativ zu einer idealen Punktquelle an 13° und 4 m, an den 23×23 Messpunkten im Lautsprecherarray. Die verschiedenen Paneele stellen verschiedene SFS-Verfahren dar, für GAR um 1 kHz und um 2 kHz.

Der Pegelfehler in der Mitte des Lautsprecherarrays liegt bei allen SFS-Verfahren unter 1 dB, mit Ausnahme des VBAP-Verfahrens. Hier beträgt dieser 2.8 dB bei 1 kHz und 1.7 dB bei 2 kHz. Das NLS-Mapping weist bei 1 kHz einen Fehler von 1.3 dB auf.

Gut zu erkennen ist, dass bei beiden HOA Verfahren die Region, in der der Fehler niedrig ist (auch *sweet-spot* genannt) mit steigender Frequenz stark abnimmt und hohe Fehler näher an der Mitte des Lautsprecherarrays gemessen wurden. Ebenfalls zu erkennen sind die Nebenkeulen, die bei HOA *max r_E*, VBAP und bei HOA *basic* an 2 kHz auftreten. Diese zeichnen sich durch eine starke Steigung des absoluten Pegelfehlers aus, der teilweise innerhalb von 20 cm von < 1 dB auf > 5 dB ansteigt. Dies ist auf destruktive Interferenzen zurückzuführen, was beim HOA *max r_E* Dekoder an 2 kHz besonders gut zu sehen ist.

Interessant ist die starke Ähnlichkeit zwischen HOA *max r_E* und VBAP, die vor allem an 2 kHz gut erkennbar ist. Es scheint, als ob der Pegel des synthetisierten Schallfeldes mit VBAP lediglich 1-2 dB höher ist als bei HOA *max r_E*.

Um die verschiedenen SFS-Verfahren auch untereinander zu vergleichen, wurde ein räumlich gefensterter mittlerer Fehler berechnet. Um die Wichtigkeit einer fehlerfreien Synthese in und um die Mitte des Lautsprecherarrays hervorzuheben, wurde eine räumliche Gewichtung der absoluten Pegelfehler durchgeführt. Hierzu wurden die Pegelfehler mit einer 2D Gaussfunktion, normalisiert um den Mittelwert (0,0) mit einer Standardabweichung von 0.5 m, gewichtet. Die danach berechneten gewichteten Pegelfehler sind in Abbildung 4 zusammengefasst. Dort ist zu sehen, dass die Fehler an höheren Frequenzen tendenziell größer sind, mit Ausnahme vom NLS-Verfahren. Hier ist der mittlere absolute Pegelfehler relativ konstant.

Diskussion

Für diese Arbeit wurden zwei Messungen durchgeführt: die Verifizierung des ISO 8253-2 Standards für audiologische Messungen im Freifeld und eine räumliche Messung synthetisierter Schallfelder, die aus der Auralisierung einer Quelle an 4 m und 13° mit vier verschiedenen SFS-Verfahren entstanden.

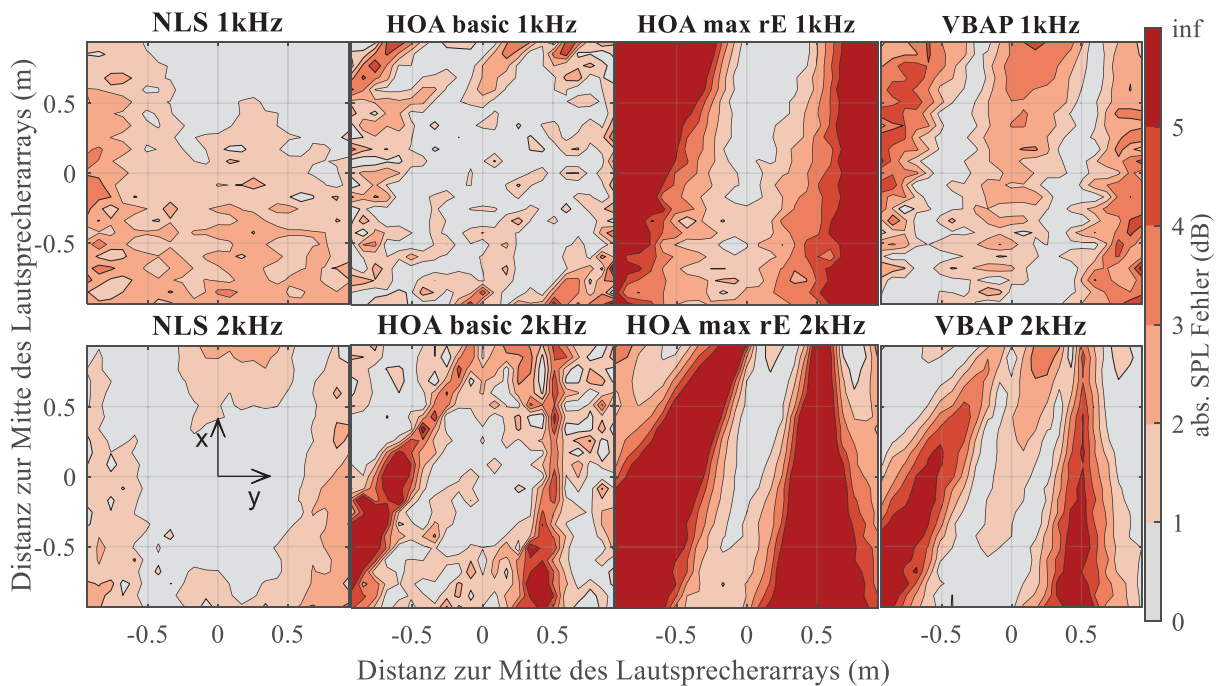


Abbildung 3: Gemessener absoluter Pegelfehler relativ zum theoretischen Schallfeld einer Punktquelle an 4 m Entfernung und 13° zur Mitte des Lautsprecherarrays, farblich dargestellt. Die obere Reihe zeigt die Messung für GAR um 1 kHz, die untere Reihe für GAR und 2 kHz. Die verschiedenen SFS-Verfahren sind in Spalten angeordnet.

Informelle Tests zeigen, dass die Lautsprecherapparatur im RAR der Professur für Audio-Signalverarbeitung der TU München auch für FM Töne der ISO 8253-2 Norm für Freifelder gerecht wird.

Die Pegelfehler in der Mitte des Lautsprecherarrays für das NLS-Verfahren an 1 kHz lassen sich durch die Lautsprecherentzerrung erklären. Die spektrale Auflösung der Filter beträgt 43 Hz, die zum Teil nicht ausreichen um die schmalbandigen Exkursionen des Frequenzgangs der Lautsprecher zu korrigieren. Diese Einflüsse sind bei den anderen SFS-Verfahren auf reduziert, da sich diese bei Verwendung mehrerer Lautsprecher herausmitteln. Der Pegelfehler in der Mitte mit VBAP ist erwartet, da die Normalisierung der Lautsprecher signale bei VBAP von einer inkohärenten Schalladdition ausgeht. Durch die Entzerrung der Lautsprecher im Array addieren sich die Schalldrücke und es wird ein 3 dB Fehler erwartet.

Die Verkleinerung des *sweet-spots* stimmt auch mit den Erwartungen überein. Die Faustregel, die die Ambisonics-Ordnung N mit dem Radius des *sweet-spots* r und der Wellenzahl k verbindet, $N=kr$, lässt einen *sweet-spot* Radius von 93 cm an 1 kHz und 46 cm an 2 kHz vermuten. Der HOA *basic* Dekoder folgt dieser Faustregel, wobei der HOA *max rE* Dekoder einen nur ungefähr halb so großen *sweet-spot* aufweist, der aber auch dem inversen Verhältnis zur Frequenz folgt. Die Beobachtung, dass der gewichtete Pegelfehler mit steigender Frequenz größer wird, ist der Verkleinerung des *sweet-spots* geschuldet, da höhere Fehler schon näher an der Lautsprecherarraymitte auftreten.

Abbildung 4 zufolge ist es am besten, für Stimuli bis 2 kHz das HOA *basic* Verfahren zu verwenden und höhere Frequenzen über NLS-Mapping abzuspielen. Der relativ

konstante mittlere absolute Fehler von VBAP liegt am zu hohen Schalldruckpegel in der die Mitte, sodass ein 0 dB Pegelfehler an den Nebenkeulen erscheinen und den Mittelwert senken. Das Perceptually-Equalized-Panning (PEP) Verfahren korrigiert u.a. diesen Fehler [17].

Unerwartet ist aber der deutlich höhere gewichtete Pegelfehler beim HOA *max rE* Dekoder. Daniel [18], Zotter und Frank [15] und Gerzon [19] beschreiben diesen als Hochfrequenzdekoder und empfehlen dessen Verwendung ab einer bestimmten Grenzfrequenz, abhängig vom der verwendeten Lautsprecherapparatur. Diese Ergebnisse lassen sich durch starke Interferenzen zwischen den kohärenten Lautsprechern und Kammfiltereffekte erklären, die durch die Lautsprecherentzerrung verstärkt werden. Auf Abbildung 3 sind diese deutlich sichtbar.

Jedoch ist die Laufzeitentzerrung essentiell, wenn die physikalisch korrekte Synthese des Schallfelds im Vordergrund steht. Die Laufzeitunterschiede in diesem Lautsprecherarray betragen bis zu 3 ms, was ebenfalls die Lokalisierung einer Schallquelle beeinflussen und potentielle

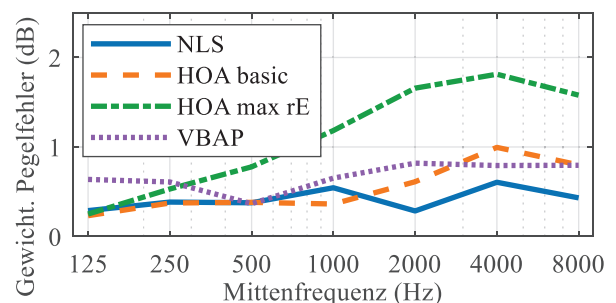


Abbildung 4: Gewichteter Pegelfehler für die verschiedenen Stimuli und die verschiedenen SFS-Verfahren.

Lokalisierungsexperimente erschweren könnte. Aufgrund der hohen Pegelunterschiede zwischen den Lautsprechern ist es allerdings schwer einzuschätzen, wie kritisch diese Laufzeitunterschiede wirklich sind.

Zusammenfassung

In diesem Beitrag wurden zwei Messungen vorgestellt: Die erste bestätigt die Anwendbarkeit des Lautsprecherarrays der SOFE für audiologische Messungen im Freifeld nach ISO 8253-2 [12]. Die zweite Messung evaluierte die Genauigkeit der Reproduktion einer Punktquelle mit verschiedenen SFS-Verfahren. Dabei schnitt der HOA *max r_E* Dekoder schlechter ab als die 3 Vergleichsmethoden. Obwohl er mehrfach für hohe Frequenzen empfohlen wurde, kommt es bei Verwendung von Laufzeitkorrekturen der quadratisch angeordneten Lautsprecher zu Problemen.

Die Messungen weisen darauf hin, dass der HOA *basic* Dekoder für tiefe Frequenzen bis 2 kHz am besten abschneidet. Für höhere Frequenzen sollten Schallquellen nur über das NLS-Verfahren abgespielt werden.

Danksagung

rtSOFE wurde vom BMBF Bernstein Center for Computational Neuroscience, 01GQ1004B, finanziert. Diese Arbeit wurde von der Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 352015383 – SFB 1330 C 5, unterstützt.

Literatur

- [1] Kolotzek, N., Seeber, B.U.: Spatial unmasking of circular moving sound sources in the free field. Proceedings of the 23rd International Congress on Acoustics (2019), Aachen
- [2] Kolotzek, N., Seeber, B.U.: Localizing the end position of a circular moving sound source near masked threshold. Forum Acusticum (2020), Lyon
- [3] Hládek L., Seeber, B.U.: Behaviour and Speech Intelligibility in a changing Multi-talker Environment. Proceedings of the 23rd International Congress on Acoustics (2019), Aachen
- [4] Schulberg, M.: Free field pure tone audiometry as a comparative system for the assessment and fitting of hearing aids. Journal of Biomedical Engineering (1980), 2
- [5] Favrot, S.: A Loudspeaker-Based Room Auralization System for Auditory Research. PhD Thesis, Centre for Applied Hearing Research, Technical University of Denmark (2010)
- [6] Daniel, J.: Spatial Sound Encoding Including Near Field Effect: Introducing Distance Coding Filters and a Viable, New Ambisonic Format. Proceedings of the 23rd AES International Conference (2003), Kopenhagen
- [7] Frank, M., Zotter, F., and Sontacchi, A.: Localization experiments using different 2D Ambisonics decoders. 25th Tonmeistertagung (2008)
- [8] Stitt, P., Bertet, S., Van Walstijn, M.: Off-centre localisation performance of Ambisonics and HOA for large and small loudspeaker array radii. Acta Acustica united with Acustica (2014) 100:937-944
- [9] Yang, L., Bosun, X.: Subjective Evaluation on the Timbre of Horizontal Ambisonics Reproduction. Proceedings of the International Conference on Audio, Language and Image Processing (IEEE)(2014), 11-15
- [10] Murillo, D.M., Fazi, F.M., Shin, M.: Evaluation of ambisonics decoding methods with experimental measurements, Proceedings of the EAA Joint Symposium on Auralization and Ambisonics (2014), 33-40, Berlin
- [11] Grandjean, P.: Captation 3-D de Métriques Directionnelles et Reproduction de Champs Sonores par la Combinaison d'Ambisonies Circulaires et Sphériques. PhD Thesis, Département de génie mécanique, Université de Sherbrooke (2021)
- [12] ISO 8253-2: Acoustics – Audiometric test methods – Part 2: Sound field audiometry with pure-tone and narrow-band test signals. International Organization for Standardization (2009)
- [13] Seeber, B.U., Clapp, S.: Interactive simulation and free-field auralization of acoustic space with the rtSOFE, Journal of the Acoustical Society of America (2017), 141(5):3974
- [14] Fastl, H., Zwicker, E.: Psychoacoustics – Facts and Models. Springer, 2007
- [15] Zotter, F., Frank, M.: Ambisonics – A Practical 3D Audio Theory for Recording, Studio Production, Sound Reinforcement, and Virtual Reality. Springer, 2019
- [16] Pulkki, V.: Virtual Sound Source Positioning Using Vector Base Amplitude Panning. Journal of the Audio Engineering Society (1997), 45(6):456-466
- [17] Seeber, B.U., Hafter, E.: Perceptual equalization in near-speaker panning. Fortschritte der Akustik - DAGA (2007), Berlin
- [18] Daniel, J., Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un context multimédia. PhD Thesis, Université de Paris 6 (2001)
- [19] Gerzon, M., Optimum Reproduction Matrices for Multispeaker Stereo. Journal of the Audio Engineering Society (1992), 40(7-8):571-589
- [20] Frank, M.: How to make Ambisonics sound good. Forum Acusticum (2014), Krakau