

Komfortable Messung und Bereitstellung individueller kopfbezogener Impulsantworten als OpenDAFF

Michael Weinert¹, Gerald Enzner¹, Jan-Mark Batke², Peter Jax², Christiane Antweiler³

¹ Institut für Kommunikationsakustik, Ruhr-Universität Bochum, Email: {Michael.Weinert, Gerald.Enzner}@rub.de

² Technicolor Research & Innovation, Hannover, Email: {Jan-Mark.Batke, Peter.Jax}@technicolor.com

³ Institut für Nachrichtengeräte und Datenverarbeitung, RWTH Aachen University, Email: Antweiler@ind.rwth-aachen.de

Einleitung

Eine präzise Binauraltechnik beruht auch auf der Nutzung individueller, hoch aufgelöster, kopfbezogener Impulsantworten (HRIRs: Head-Related Impulse Responses). Eine komfortable Realisierung erfordert jedoch eine schnelle akustische HRIR-Messtechnik. In diesem Beitrag wird ein Verfahren vorgestellt, welches mit einer Messdauer von lediglich ein bis zwei Minuten einen dreidimensionalen HRIR-Datensatz mit hoher Qualität ermitteln kann. Durch kontinuierliche Rotation des Probanden innerhalb der Messdauer und durch Verwendung adaptiver Filter zur anschließenden Berechnung der HRIRs aus den gemessenen Ohrsignalen wird eine praktisch beliebig feine azimutale Auflösung erzielt. Mehrkanalige adaptive Filter können zur Bestimmung von HRIR-Datensätzen mit mehreren Elevationsebenen und azimutal quasi-kontinuierlicher Auflösung genutzt werden. Durch geeignete Wahl des Messsignals kann die vorgestellte Prozedur sogar akustisch besonders angenehm für die Probanden gestaltet werden. Für eine effiziente Nutzbarkeit, z.B. in Echtzeitsystemen, können die Datensätze nach azimutaler Abtastung als OpenDAFF (Open Directional Audio File Format) zur Verfügung gestellt werden. Dabei wird der gesamte Datensatz innerhalb einer einzigen Datei gespeichert, deren unkomprimierte Größe bei acht Elevationsebenen und einer azimutalen Abtastung im 0.1 Grad Intervall ca. 35 MB beträgt. Die komfortable Messtechnik, die hohe räumliche Auflösung der resultierenden individuellen HRIR-Datensätze sowie die freie Verfügbarkeit von Leserroutinen für OpenDAFF bieten neue technische Möglichkeiten zur Rekonstruktion und experimentellen Untersuchung von binauralen Schallfeldern.

Dynamisches Messverfahren

Mit Hilfe des in Abb. 1 schematisch dargestellten Messaufbaus werden zunächst die Daten erfasst, mit welchen der adaptive NLMS-Algorithmus dann die Berechnung der HRIRs für den Azimut Φ_k und die Elevation Θ_ν leistet [1]:

$$\hat{\mathbf{h}}_{\Theta_\nu}^{l,r}(\Phi_{k+1}) = \hat{\mathbf{h}}_{\Theta_\nu}^{l,r}(\Phi_k) + \mu_0 \cdot \frac{e^{l,r}(k)\mathbf{x}_{\Theta_\nu}(k)}{\sum_{\nu=1}^M \|\mathbf{x}_{\Theta_\nu}(k)\|_2^2} \quad \forall \Theta_\nu \quad (1)$$

$$e^{l,r}(k) = y^{l,r}(k) - \sum_{\nu=1}^M \mathbf{x}_{\Theta_\nu}^T(k) \hat{\mathbf{h}}_{\Theta_\nu}^{l,r}(\Phi_k) \quad (2)$$

Dabei bezeichnen $\mathbf{x}_{\Theta_\nu}(k)$ die Vektoren der zeitdiskreten Messsignale $x_\nu(k)$, $\hat{\mathbf{h}}_{\Theta_\nu}^{l,r}(\Phi_k)$ die HRIR-Koeffizientenvektoren, $y^{l,r}(k)$ die an den blockierten Eingängen der Ohrkanäle gemessenen Signale und $e^{l,r}(k)$ die jeweils gemeinsamen Fehler-signale für P -kanalige Messungen ($\nu = 1, 2, \dots, P$):

$$\mathbf{x}_{\Theta_\nu}(k) = [x_{\Theta_\nu}(k), \dots, x_{\Theta_\nu}(k - N + 1)]^T \quad (3)$$

$$\hat{\mathbf{h}}_{\Theta_\nu}^{l,r}(\Phi_k) = [h_{\Theta_\nu,0}^{l,r}(\Phi_k), \dots, h_{\Theta_\nu,N-1}^{l,r}(\Phi_k)]^T. \quad (4)$$

Die einzelnen $\mathbf{x}_{\Theta_\nu}(k)$ werden mittels Referenzaufnahmen in der Position der Kopfmittte ohne den Probanden gewonnen.

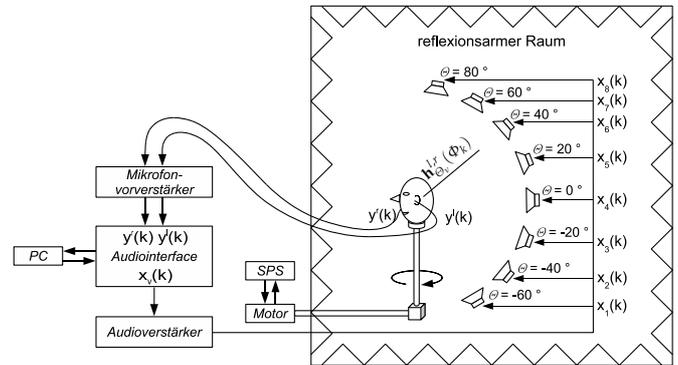


Abbildung 1: Messaufbau zur HRIR-Erfassung, $P = 8$

Abb. 2 zeigt die auf die Filterlänge $N = 308$ normierte Fehlersignaldämpfung $(\sigma_e^2/\sigma_y^2)/N$ in Abhängigkeit der Umdrehungsdauer für 1, 2, 4 und 8-kanalige Messungen.

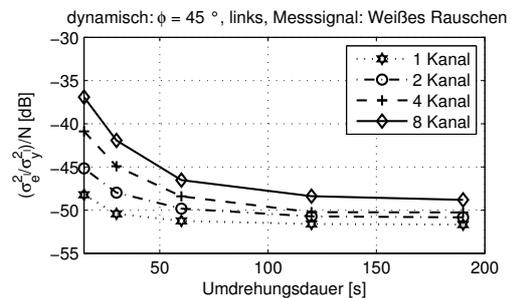


Abbildung 2: Fehlersignaldämpfung vs. Umdrehungsdauer

Nachfolgend wird ermittelt, welche akustischen Einflussgrößen für die Sättigung der Fehlersignaldämpfung in Abb. 2 bei ca. -52 dB verantwortlich sind.

Validierung des Verfahrens

Durch Reduktion der Messtechnik auf 1-kanalige, statische Kunstkopfmessungen ($\Phi = 45^\circ$, $\Theta = 0^\circ$) können die das Messergebnis limitierenden Einflussgrößen ermittelt werden. Als zusätzliches Messsignal und diagnostisches Mittel kommen perfekte periodische Sweeps zum Einsatz, die speziell für die Verwendung mit adaptiven Filtern entworfen wurden [2, 3]. Die Periodendauer entspricht dabei der adaptiven Filterlänge.

Zunächst wird in Abb. 3 die Fehlersignaldämpfung in Abhängigkeit vom SNR der Messung, welches durch die Wiedergabelautstärke bestimmt wird, dargestellt. Während sich die Fehlersignaldämpfung für perfekte Sweeps (b) linear mit dem SNR bzw. Schalldruckpegel (SPL) verhält, ist für weißes Rauschen (a) wiederum eine Sättigung bei ca. -52 dB zu verzeichnen. Zur Interpretation dieses Ergebnisses wird auf die bekannte Eigenschaft von kurzzeit-periodischen Mess-

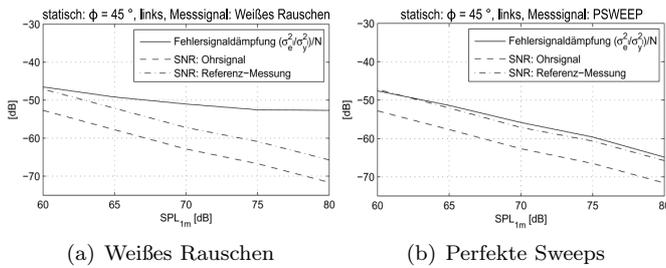


Abbildung 3: Fehlersignaldämpfung vs. SPL

sequenzen verwiesen, dass akustische Reflexionen den eigentlichen HRIRs messtechnisch überlagert werden [4]. Damit können mit perfekten Sweeps trotz kurzer Filterlänge hohe Fehlersignaldämpfungen erreicht werden. Bei Verwendung von weißem Rauschen hingegen wirken sich untermodellierete Reflexionen effektiv wie eine unsystematische Messstörung aus. Abb. 3 legt insgesamt also nahe, dass akustische Reflexionen für die in Abb. 2 beobachtete Sättigung verantwortlich sind. Mit dieser Hypothese müsste das Spektrum $|E(j\omega)|$ des NLMS-Fehlersignals bei Verwendung von weißem Rauschen einen ausgeprägteren Tiefpasscharakter aufweisen, da ein reflexionsarmer Messraum gerade bei tiefen Frequenzen seine Reflexionsarmut verliert. Außerdem können weitere Reflexionen durch die Messanordnung kaum vermieden werden. Abb. 4 verifiziert den vermuteten Tiefpasscharakter des Fehlersignals, der bei weißem Rauschen um etwa 10 dB stärker ausgeprägt ist als bei perfekten Sweeps. Dieser Unterschied befindet sich im Einklang mit demjenigen der Fehlersignaldämpfungen in Abb. 3 bei SPL = 80 dB.

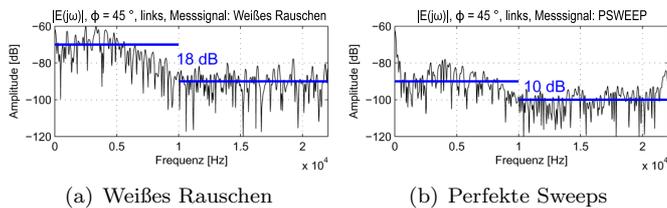


Abbildung 4: Betragsspektren des Fehlersignals, SPL=80dB

Trotz der beinahe unvermeidlichen Störung des Messergebnisses durch akustische Reflexionen kann weißes Rauschen dennoch gut als Messsignal im beschriebenen Verfahren zum Einsatz kommen. Insbesondere wurde von Probanden im Rahmen der Messprozedur über eine besonders angenehme auditive Wahrnehmung des weißen Rauschens berichtet.

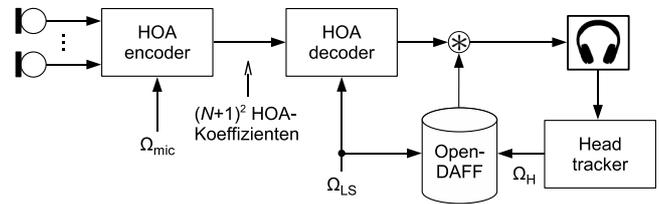
Durch Verwendung von Referenzaufnahmen, in denen systematische Nichtlinearitäten der HRIR-Messung gleichermaßen enthalten sind, treten Nichtlinearitäten in der HRIR-Berechnung als Ursache für die -52dB-Sättigung in den Hintergrund.

RUB-HRIRs als OpenDAFF

Bereitgestellt werden die HRIRs im OpenDAFF-Format [5], sodass die jeweiligen Datensätze in kompakter Weise auch für die Anwendung in Echtzeitsystemen zur Verfügung stehen. Die HRIRs werden hierbei äquidistant in Azimut und Elevation gespeichert. Die zu wählende Rasterung der HRIR-Daten in Elevationsrichtung ergibt sich aus der Lautsprecheranordnung in Abb. 1. In azimuthaler Richtung kann durch geeignete Unterabtastung der quasi-kontinuierlichen HRIR-Messung eine nahezu beliebige feine Rasterung erreicht werden. Weil OpenDAFF als open-source-Projekt angelegt ist, besteht eine gute Schnittstelle zu den Anwendungen der Binauraltechnik.

Anwendungsbeispiel: Wiedergabe von Higher-Order-Ambisonics Signalen

Beispielhaft wurde die 3D-Kopfhörerwiedergabe natürlicher Schallfelder im Ambisonics-Format implementiert. Das Blockdiagramm des Gesamtsystems ist in Abb. 5 dargestellt.

Abbildung 5: Schallfelderfassung und binaurale Reproduktion mittels der virtuellen Lautsprecherpositionen Ω_{LS}

Für die Ambisonics-Aufnahmen wurde ein Mikrofonarray (*EigenMike, mh acoustics*) verwendet. Die 32 Kapselsignale werden zuerst unter Kenntnis der Mikrofonpositionen Ω_{mic} in Koeffizienten einer Higher-Order-Ambisonics (HOA) Darstellung der Ordnung $N = 4$ überführt. Für jeden zeitlichen Abtastpunkt entstehen so $(N+1)^2 = 25$ Koeffizienten, welche die räumlichen Eigenschaften des Schallfeldes am Aufnahmeort beschreiben und grundsätzlich unabhängig von Mikrofon- und Lautsprecheranordnung sind.

Die Schallfeldwiedergabe wird durch eine Decodierung der HOA-Koeffizienten unter Annahme virtueller Lautsprecherpositionen Ω_{LS} realisiert, gefolgt von der Aufprägung entsprechender kopfbezogener Impulsantworten. Hierbei werden die HRIRs der OpenDAFF Datenbasis für Positionen entnommen, die sich aus der Verknüpfung der Lautsprecherpositionen Ω_{LS} mit der Kopforientierung Ω_H ergeben. Letztere wird über einen Headtracker (*Polhemus Fastrak*) in das Echtzeitsystem eingespeist. Die Ohr-Signale werden über ein entzerrtes Kopfhörersystem (*STAX SR Lambda*) auralisiert.

Erste informelle Hörtests lieferten vielversprechende Ergebnisse. Die auf den HRIRs basierende Schallfeldwiedergabe ergibt ein detailliertes und natürlich empfundenes räumliches Klangbild. Die Lokalisation einzelner Schallobjekte entspricht den Erwartungen an eine HOA-Darstellung der Ordnung 4.

Im Rahmen der Implementierung des Systems erwies sich die Verfügbarkeit von wechselbaren und individuellen HRIR-Datensätzen in Form von OpenDAFF-Dateien als sehr praktisch. Die hohe Auflösung der HRIRs ermöglicht eine flexible Auswahl virtueller Lautsprecherpositionen sowie eine unkomplizierte Anpassung an die augenblickliche Kopforientierung.

Literatur

- [1] G. Enzner: 3D-Continuous-Azimuth Acquisition of Head-Related Impulse Responses Using Multi-Channel Adaptive Filtering, WASPAA, 2009
- [2] A. Telle, C. Antweiler, P. Vary: Der perfekte Sweep - Ein neues Anregungssignal zur adaptiven Systemidentifikation zeitvarianter akustischer Systeme, DAGA, 2010
- [3] C. Antweiler: Perfect-Sweep NLMS for Time-Variant Acoustic System Identification, ICASSP, 2012
- [4] C. Antweiler, G. Enzner: Perfect Sequence LMS for Rapid Acquisition of Continuous-Azimuth Head Related Impulse Responses, WASPAA, 2009
- [5] F. Wefers: OpenDAFF - A Free, Open-Source Software Package for Directional Audio Data, DAGA, 2010