

Exposé

Messtechnische Akquise akustischer Szenen mit (pseudo-)binauralen Aufnahmeverfahren

Julian Domann BEARBEITUNG
Matrikelnummer: 394104

Prof. Dr. Stefan Weinzierl BEGUTACHTUNG
David Ackermann

Inhaltsverzeichnis

1	Einleitung und Stand der Technik	1
2	Methoden	2
2.1	Arbeitspaket 1 - Modellerstellung	2
2.2	Arbeitspaket 2 - Sweepbasierte Impulsantworten	3
2.3	Arbeitspaket 3 - Lautsprecherorchester Aufnahmen	3
3	Form der Datenveröffentlichung	4
4	Arbeits- und Zeitplan	5
	Literaturverzeichnis	7

1 Einleitung und Stand der Technik

Das vorliegende Exposé beschreibt die Vorgehensweise zur Erstellung einer wissenschaftliche Datenpublikation für eine Masterarbeit im Studiengang Audiokommunikation und -technologie an der Technischen Universität Berlin. Der Datensatz soll in einem Anschlussprojekt die perzeptive Evaluation von dynamischen (pseudo-)binauralen Übertragungsverfahren ermöglichen. Inhalt des Datensatzes werden Übertragungsfunktionen akustischer Szenen sein, die mit verschiedenen binauralen und pseudo-binauralen Mikrofonverfahren erfasst werden. Genauer werden diese Szenen für jedes Aufnahmeverfahren einen Satz an mehrkanaligen und binauralen Raumimpulsantworten, sowie mehrkanaligen und binauralen Aufnahmen eines Lautsprecherorchesters bereitstellen.

Eine der Öffentlichkeit zugängliche Datenbank, ist mit den hier aufgeführten Inhalten nicht bekannt. Veröffentlichte Aufnahmen eines Lautsprecherorchesters dienen bisher nur dem Zweck, die Auralisierung von verschiedenen Konzertsälen bei Aufnahme mit dem gleichen Mikrofon und Ausgabe über Lautsprecher miteinander zu vergleichen. So haben Lokki et al. (2011) Aufnahmen von klassischen Sinfonien unterschiedlicher Epochen aus einem 32-kanaligen Lautsprecherorchesters (Aufbau nach Pätynen et al. (2008)) in verschiedenen Konzertsälen erstellt, bei denen ein 6-kanaliges Mikrofonarray verwendet wurde. Bei Lokki et al. (2012) wurde dies auf weitere Konzertsäle erweitert, jedoch wurden hier nur noch die RIRs der einzelnen Lautsprecher ermittelt. Eine Auswertung von verschiedenen (pseudo-)binauralen Verfahren mit den gleichen Quell- und Empfangspositionen bei binauralisierter Wiedergabe ist hiermit nicht möglich.

In der messtechnischen Akquise kommt der FABIAN (Fast and Automatic Binaural Impulse response Acquisition) Head-and-Torso Simulator (HATS) der TU Berlin (Lindau und Weinzierl, 2007), welcher die dynamische Erfassung von binauralen Raumimpulsantworten (BRIR, binaural room impulse response) mit Hilfe eines servomotorisches Halsgelenks ermöglicht, zum Einsatz. Daneben wird das 16-kanalige MTB-Mikrofon der TU Berlin (Fiedler et al., 2017), welches in Kooperation mit der Firma Sennheiser entstanden ist, verwendet. Das mit dem MTB Mikrofonarray aufgenommene mehrkanalige Signal kann später durch geeignete Auswahl eines sich gegenüberliegendem Mikrofonpaares dem Hörer wiedergegeben werden. Eine dynamische Nachführung der aktuellen Kopfposition ist über ein Head-tracking System und einem Algorithmus zur Interpolation zwischen den Mikrofonen ebenso möglich. Als First-Order Ambisonics Mikrofon wird das AMBEO VR Mic der Firma Sennheiser (Sennheiser electronic GmbH & Co. KG, 2021) eingesetzt. Hierbei wird durch eine bestimmte Anordnung von 4 Mikrofonkapseln mit Nierencharakteristik das räumliche Schallfeld im A-Format aufgenommen. Nach anschließender Encodierung ins B-Format (nicht Gegenstand dieses Datensatzes) erlaubt dieses Verfahren ebenso ein dynamisches Nachführen der Kopfposition über ein Head-Tracker. Zur Messung von Higher-Order-Ambisonics (HOA) Signalen dient zum einen das 19-kanalige Mikrofonarray 3. Ordnung ZM-1 des Herstellers Zylia (Zylia Sp.

z. o. o., 2021), welches auf einer starren Kugeloberfläche 19 omnidirektionale Mikrofonkapseln eingelassen hat, zum anderen das 86-kanalige Mikrofonarray HØSMA der TH Köln (Dziwis et al., 2019). Hiermit können HOA-Signale bis zur 7. Ordnung encodiert werden.

Die aufbereiteten und dokumentierten Messdaten sollen im Anschluss im Repositorium für Forschungsdaten und Publikationen der TU Berlin *DepositOnce* veröffentlicht werden.

Im nächsten Abschnitt werden im Methodenteil die einzelnen Arbeitspakete vorgestellt, die den Inhalt des Datensatzes liefern, Danach folgt eine Beschreibung in welcher Form die Daten veröffentlicht werden und zuletzt wird ein Zeitplan zur Bearbeitung dieses Projektes aufgestellt.

2 Methoden

Das folgende Kapitel beschreibt die Aufgaben der einzelnen Arbeitspakete zur Erstellung des Datensatzes.

2.1 Arbeitspaket 1 - Modellerstellung

Der erste Teil zur Erstellung des Datensatzes betrifft die simulationsfähige Modellierung der Räume, in denen die Aufnahmen durchgeführt werden. Als raumakustische Umgebung dient zum einen ein typischer Raum zur Darbietung von klassischer Musik. Hierbei handelt es sich um den große Saal der Berliner Philharmonie ($21\,000\text{ m}^3$, $EDT = 1,9\text{ s}$). Zum anderen werden Messungen in einem Raum des Berlin Open Labs (2500 m^3 , $T_{20} = 1\text{ s}$) durchgeführt.

Zunächst werden dreidimensionale Modelle der Räume mit der 3D-Design-Software SketchUP des Entwicklers Trimble Inc. erstellt. Hierfür ist eine Erfassung der Raumgeometrien, der Materialeigenschaften der Grenzflächen und sich im Raum befindenden Objekten (Absorptions- und Streuungsgrade) und der Quell- und Empfängerpositionen nötig. Außerdem ist eine detaillierte Fotodokumentation zu erstellen. Auf Grundlage der realen raumakustischen Eigenschaften der Räume werden die Modelle dann mit der Raumakustik-Simulationssoftware RAVEN (Schröder und Vorländer, 2011) gefittet, sodass monaurale, sowie binaurale Impulsantworten simuliert werden können. Diese ermöglichen eine spätere Integration von etablierten Gestaltungsmitteln wie Stützmikrofonen oder künstlichen Nachhall, welche nicht ohne Weiteres dem binauralen bzw. binauralisierten Audiomaterial beigemischt werden können, ohne die Plausibilität bei dynamischer binauralisierter Wiedergabe zu verringern.

Die Bereitstellung der Modelldaten soll dem Benutzer des Datensatzes eine exakte Bestimmung der Positionen der verwendeten Sender und Empfänger ermöglichen und durch symbolhafte Darstellung der akustischen Szenen während der Messung den Messaufbau dokumentieren. Dies ist aufgrund des Anspruches auf Reproduzierbarkeit der Messungen und eventueller Erweiterung des Datensatzes ein wichtiger Bestandteil.

2.2 Arbeitspaket 2 - Sweepbasierte Impulsantworten

Im zweiten Arbeitspaket geht es um die Erzeugung von mehrkanaligen und binauralen Raumimpulsantworten. Diese werden mit einem logarithmischen Sweep als Stimulus (Müller und Massarani, 2001) und anschließender Entfaltung des vom jeweiligen Mikrofonarrays aufgenommenen Signals gewonnen.

Als Schallquellen dienen 15 Lautsprecher vom Typ QSC K8, Genelec 8030 und Meyer UPL-1 mit Bündelungsgraden, die sich dem von Orchesterinstrumenten ähneln. Dies unterscheidet sich zu den Bedingungen unter denen typischerweise Messungen zu raumakustischen Parametern durchgeführt werden und bei denen normalerweise omnidirektionale Lautsprecher zum Einsatz kommen. Dies geschieht hier, da die IRs direkt auf den Einsatz von echten Instrumenten Bezug nehmen sollen. Die genaue Zuweisung der Lautsprechermodelle zu den Instrumenten(-gruppen) erfolgt im weiteren Verlauf des Projekts und wird unter dem Aspekt der Übereinstimmung der Abstrahlcharakteristik ermittelt. In einem früheren Projekt der Forschungsgruppe SEACEN (Simulation und Evaluation akustischer Umgebungen) wurde die Richtcharakteristik einiger Lautsprecher bereits ausführlich vermessen und dokumentiert (Aspöck et al., 2020). Die Positionierung der Lautsprecher erfolgt analog zu den Messungen des Lautsprecherorchesters, welche in Abschnitt 2.3 ausführlicher beschrieben werden.

Als Empfangspositionen der Mikrofone werden zur Evaluation der Übertragungsqualität in Bezug auf den Einfluss des Verhältnisses von Direkt- zu Diffusschall eine bühnennahe (ungefähr im Abstand des doppelten Hallradius) und eine bühnenferne Position ausgewählt und sind bei allen verwendeten Verfahren gleich. Zur Generierung der BRIRs mit verschiedenen Kopf-über-Torso Ausrichtungen (HATO, Head-above-Torso Orientation) wird der FABIAN HATS verwendet und die Auslenkung nach jeder Messung um 2° geändert (zwischen den Grenzen -82° und $+82^\circ$). Die Mikrofonarrays messen die mehrkanaligen RIRs. Um das in 2.1 beschriebene Fitten des Modells vorzunehmen wird zusätzlich noch eine monaurale RIR durch Verwendung eines ungerichteten Mikrofons aufgezeichnet.

Die Erfassung der bis zu 86-kanaligen Messsignale erfolgt mit der professionellen Studio- und digitalen Netzwerktechnik der Berliner Philharmonie. Zur weiteren Signalverarbeitung wird AkTools verwendet (Brinkmann und Weinzierl, 2017). Dies stellt Matlab-Methoden zur Verfügung mit denen Audiodaten aufgenommen, bearbeitet und grafisch dargestellt werden können. Außerdem können mit AkTools ebenfalls die Stimuli (log. Sinus-Sweeps) erzeugt und ausgespielt werden. Die ermittelten mehrkanaligen RIRs, BRIRs und monauralen RIRs werden zuletzt im SOFA-Format (AES Standards Comitte, 2015), sowie als wav-Datei abgespeichert (48 kHz, 32 bit). Während dieses Arbeitspaketes ist bei allen Schritten auf eine ausführliche Dokumentation zu achten.

2.3 Arbeitspaket 3 - Lautsprecherorchester Aufnahmen

Im dritten Arbeitspaket geht es um die Erzeugung von mehrkanaligen und binauralen Aufnahmen eines Lautsprecherorchesters. Diese werden durch Simulation eines Orchesters, d.h. durch Ausspielung von nachhall- und übersprechfreien Audioaufnahmen über Lautsprecher, gewonnen. Als Audiomaterial wird eine, im reflexionsarmen Raum der TU Berlin produzierte, Aufnahme des Konzerthausorchesters Berlin (Claude Debussys Nr. 6 Golliwogg's Cakewalk aus der Suite Children's Corner) verwendet, welche

die Einzelspuren der Instrumente beinhaltet und somit eine separate Ausspielung und Messung der Instrumente ermöglicht.

Die Positionierung von 15 Lautsprechern folgt den gängigen Vorgaben der europäischen Sitzordnung innerhalb eines Orchesters (Meyer, 2008). Die Lautsprecher sollen jeweils die ausgespielten Streich-, Blas- oder Schlaginstrumente repräsentieren. Es werden die selben Lautsprecher mit der selben Abstrahlcharakteristik wie in AP2 verwendet.

Entsprechend zu Abschnitt 2.2 werden zur binauralen Aufnahme des Lautsprecherorchesters der FABI-AN HATS und für die mehrkanaligen Aufnahmen die Mikrofonarrays benutzt. Die Positionierung der Mikrofone, die Signalverarbeitung und Datenspeicherung erfolgt ebenso analog zu Abschnitt 2.2 und ist bei allen Aufnahmen gleich. Abweichend hierzu wird aufgrund der Tatsache, dass für einen HATS, die binaurale Aufzeichnung von längerem Audiomaterial in hoch aufgelösten HATOs einen für dieses Projekt zu hohen Zeitaufwand bedeutet, nur für eine frontal ausgerichtete Kopfposition aufgezeichnet. Wie im vorherigen AP muss auch hier auf eine ausführliche Dokumentation geachtet werden.

3 Form der Datenveröffentlichung

In diesem Kapitel wird die Form und Struktur der Datenbank zur Veröffentlichung auf DepositOnce vorgestellt (s. Abb. 3.1). Da es bereits einige Veröffentlichungen zum Thema der virtuellen Akustik vom Fachgebiet Audiokommunikation der TU Berlin gibt (Ackermann et al., 2020; Aspöck et al., 2020; Brinkmann et al., 2017), wird zur vereinfachten Integration in übergeordnete Projekte und besseren Gegenüberstellung untereinander, die grobe Formatierung übernommen.

Die Datenbank enthält 3 Hauptkategorien in denen zwischen der Beschreibung der akustischen Szenen (01_Scene_descriptions), der Beschreibung der Quellen und Empfänger (02_Source_and_Receiver_descriptions), sowie zusätzlichen Daten (03_Additional Data) unterschieden wird.

01_Scene_descriptions enthält die Aufteilung in zwei Räume, in denen die Messungen durchgeführt werden. Die nächste Ebene enthält die Unterscheidung zwischen der Raumgeometrie, mit 3D-modellierten Darstellungen des Raumes und den verwendeten Positionen der Messinstrumente, aufgenommenen Bildern zur Dokumentation und der Unterteilung der einzelnen Mikrofonverfahren (MP1-MP5). Die einzelnen Verfahren beinhalten jeweils die Impulsantworten und die Aufnahmen des Lautsprecherorchesters. Jedes Aufnahmeverfahren enthält eine Zusammenfassung der Messergebnisse als Datei im SOFA-Format, sowie für jedes Einzelmikrofon die zugehörigen Audiodaten im wav-Format.

02_Source_and_Receiver_descriptions enthält Daten zur Abstrahlcharakteristik und Konstruktion der verwendeten Mikrofone und Lautsprecher.

03_Additional Data stellt zusätzliche Informationen und Daten zur Verfügung. Es wird zwischen Audiodaten, in denen die Stimuli zur IR-Generierung, das nachhallfreie Audiomaterial und verschiedenen Auralisationen bereitgestellt werden und simulierten raumakustischen Parametern der Räume aufgeteilt.

Zuletzt wird eine ausführliche Dokumentation und Beschreibung verfasst und ebenfalls dem Datensatz hinzugefügt.

4 Arbeits- und Zeitplan

Tabelle 4.1 zeigt den geplanten Zeitaufwand zur Abarbeitung der einzelnen Arbeitspakete.

Zeitraum	Aufgaben
März	Erstellung des Exposés, Erarbeitung der Grundlagen
April	Konzeptionelle Vorbereitung der Messungen, Durchführung von Testmessungen
Mai	Messtechnische Akquise der akustischen Szenen, Erfassung der Raumeometrien und Materialeigenschaften
Juni	Erstellung der Modelle, Auswertung und Aufbereitung der Messergebnisse
Juli	Aufbau des Datensatzes, Erstellung der Dokumentation
August	Schriftliche Ausfertigung der Thesis

Tabelle 4.1: Arbeits- und Zeitplan zur Anfertigung der Masterarbeit

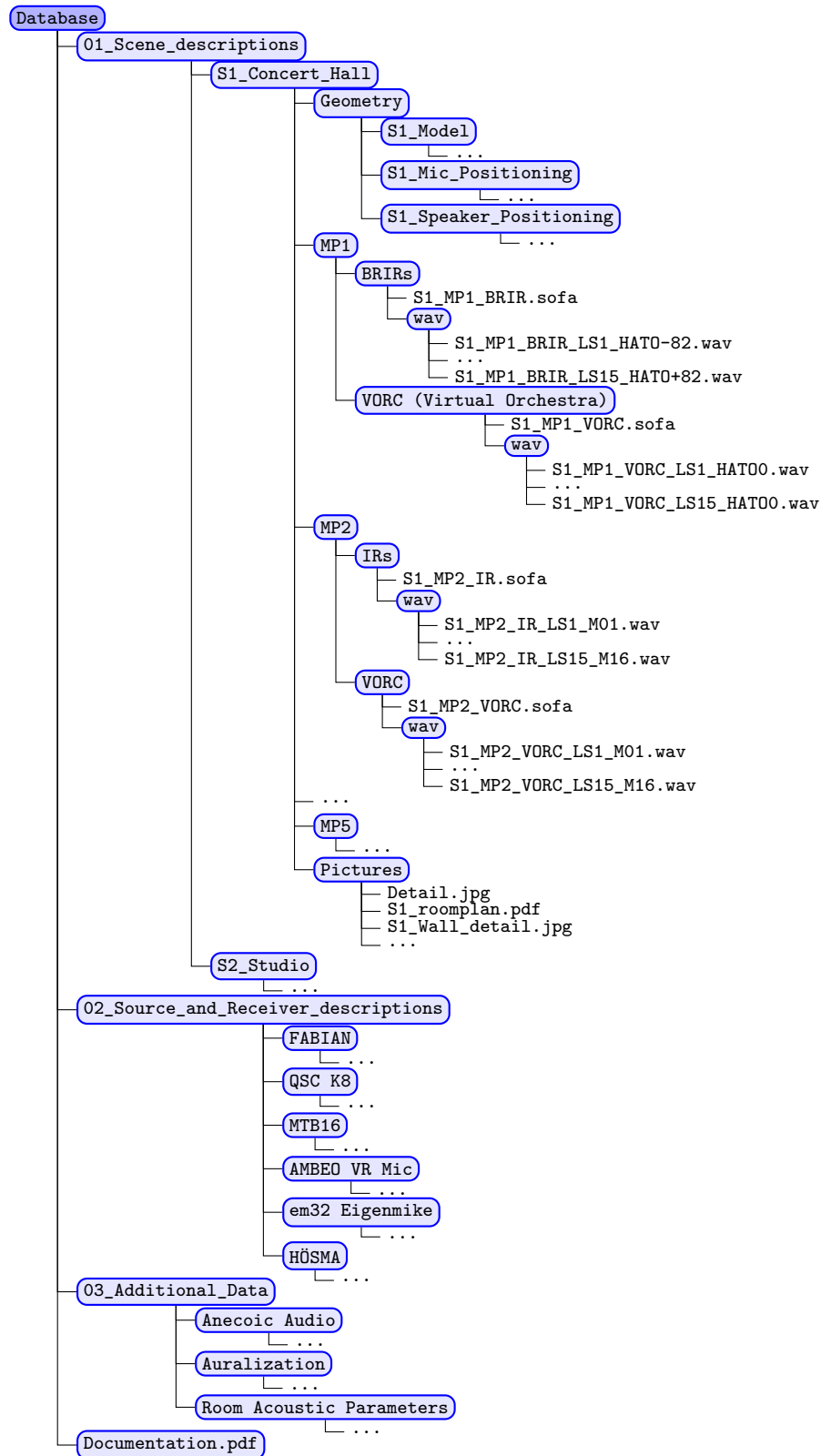


Abbildung 3.1: Ordner- und Datenstruktur der Datenbank

Literaturverzeichnis

- Ackermann, D., Ilse, M., Grigoriev, D., Lepa, S., Pelzer, S., Vorländer, M., und Weinzierl, S. (2020). A Ground Truth on Room Acoustical Analysis and Perception (GRAP). DOI: 10.14279/depositonce-7003.4.
- AES Standards Comitte (2015). *AES69-2015: AES standard for file exchange - Spatial acoustic data file format*. Audio Engineering Society, Inc.
- Aspöck, L., Vorländer, M., Brinkmann, F., Ackermann, D., und Weinzierl, S. (2020). Benchmark for Room Acoustical Simulation (BRAS). DOI: 10.14279/depositonce-6726.3.
- Brinkmann, F., Lindau, A., Weinzierl, S., Geissler, G., van der Par, S., Aspöck, L., Obdam, R., und Vorländer, M. (2017). The FABIAN head-related transfer function data base. DOI: 10.14279/depositonce-5718.5.
- Brinkmann, F. und Weinzierl, S. (2017). AKtools – An Open Software Toolbox for Signal Acquisition, Processing, and Inspection in Acoustics. In: *142nd AES Convention*, e-Brief 309.
- Dziwis, D. T., Lübeck, T., Arend, J. M., und Pörschmann, C. (2019). Development of a 7th Order Spherical Microphone Array for Spatial Audio Recording. In: *Fortschritte der Akustik – DAGA Rostock*, S. 883–885.
- Fiedler, F., Ackermann, D., Brinkmann, F., Schneider, M., und Weinzierl, S. (2017). Entwicklung und Evaluation eines Mikrofonarrays für die Aufnahme von räumlichen Schallfeldern nach dem Motion-Tracked Binaural (MTB) Verfahren. In: *Fortschritte der Akustik – DAGA Kiel*, S. 1115–1117.
- Lindau, A. und Weinzierl, S. (2007). FABIAN - Schnelle Erfassung binauraler Raumimpulsantworten in mehreren Freiheitsgraden. In: *Fortschritte der Akustik – DAGA Stuttgart*, S. 633–634.
- Lokki, T., Pätynen, J., Kuusinen, A., und Tervo, S. (2011). Concert hall acoustics assessment with individually elicited attributes. In: *The Journal of the Acoustical Society of America*, 130, S. 835–849. DOI: 10.1121/1.3607422.
- Lokki, T., Pätynen, J., Kuusinen, A., und Tervo, S. (2012). Disentangling preference ratings of concert hall acoustics using subjective sensory profiles. In: *The Journal of the Acoustical Society of America*, 132(5), S. 3148–3161. DOI: 10.1121/1.4756826.
- Meyer, J. (2008). Musikalische Akustik. In: *Handbuch der Audiotechnik*, S. 123–180. Springer, Berlin, Heidelberg.

- Müller, S. und Massarani, P. (2001). Transfer-Function Measurement with sweeps. Director's Cut including previously unreleased material and some corrections. In: *J. Audio Eng. Soc.*, 49(6), S. 443–471.
- Pätynen, J., Tervo, S., und Lokki, T. (2008). A loudspeaker orchestra for concert hall studies. In: *The Seventh International Conference On Auditorium Acoustics*, S. 45–52, Institut of Acoustics, Oslo, Norway.
- Schröder, D. und Vorländer, M. (2011). RAVEN: A Real-Time Framework for the Auralization of Interactive Virtual Environments. In: *Proc. of Forum Acusticum*, S. 1541–1546, Aalborg, Dänemark.
- Sennheiser electronic GmbH & Co. KG (2021). Produkthomepage: AMBEO VR Mic. <https://de-de.sennheiser.com/mikrofon-3d-audio-ambeo-vr-mic>. Aufgerufen am: 04.03.2021.
- Zylia Sp. z o. o. (2021). Produkthomepage: ZM-1. <https://www.zylia.co/zylia-zm-1-microphone.html>. Aufgerufen am: 30.03.2021.