

## Zur binauralen Simulation verteilter Schallquellen

Alexander Lindau, Martin Klemmer, Stefan Weinzierl

TU Berlin, Fachgebiet Audiokommunikation

alexander.lindau@tu-berlin.de, martin.klemmer@gmail.com, stefan.weinzierl@tu-berlin.de

### Fragestellung

Die binaurale Simulation räumlich verteilter Klangquellen (Orchester, Kammermusikensembles, Chöre) erfordert im Idealfall eine große Anzahl binauraler Datensätze für jede Position von Schallquelle und Hörer. Im Hinblick auf den Aufwand sowohl für die Messung als auch für die Auralisation solcher komplexer Quellen wurde untersucht, für welche räumliche Konfiguration eine Simulation mit für jede Einzelquelle separat aufgenommenen Datensätzen binauraler Raumimpulsantworten tatsächlich erforderlich ist. Dazu wurde in einem Hörversuch für verschiedene Abstände zwischen Quelle und Hörer und verschiedene Quellsignale ermittelt, ab welchem Öffnungswinkel 2 bzw. 4 Quellen als räumlich getrennt wahrgenommen werden.

### Stand der Forschung

Zur Lokalisation multipler, inkohärenter Schallquellen im direkt/diffusen Schallfeld [1, Kap. 3] existieren nur wenige systematische Untersuchungen. Sie zeigen generell eine starke Abhängigkeit der Lokalisationsleistung vom Signaltyp sowie von der Entfernung und damit vom Diffusfeldanteil des Quellsignals. Weiterhin ist eine mit zunehmender Entfernung ansteigende wahrgenommene Quellausdehnung beobachtbar [2][3]. Neuere Studien auf der Grundlage von binaural simulierten Szenarien zeigen, dass etwa ein Orchester - auralisiert mit Einzelquellen - gegenüber einer einzelnen Monopolquelle oder einer Verteilung von kohärenten Einzelquellen als realistischer bewertet wird [4][5]. Beide Untersuchungen nutzten numerische raumakustische Simulationen und statische, d.h. nicht auf Kopfbewegungen des Hörers reagierende Auralisationen. In den durchgeführten Hörversuchen wurde zwar eine Rangfolge für die Plausibilität der Quelldarstellungen ermittelt, nicht jedoch eine konkrete räumliche Distanz, oberhalb derer die Auralisation räumlich verteilter Quellen gegenüber der Auralisation einer Einzelquelle zuverlässig unterschieden werden kann.

### Untersuchungsmethode

Um zu bestimmen, ab welchem Öffnungswinkel Hörer in der Lage sind, Klangquellen als räumlich verteilt wahrzunehmen, wurden mit dem Messsystem FABIAN [6] in einem konzertsaalähnlichen Raum (Audimax der TU Berlin,  $V = 8600\text{m}^3$ ,  $T = 2,1\text{ s}$ ,  $r_H = 3,6\text{m}$ ) für zwei Hörpositionen Datensätze mit binauralen Raumimpulsantworten (BRIRs) für verschiedene Öffnungswinkel zwischen mehreren Einzelquellen gemessen. Als Distanz zwischen Quelle und Empfänger wurden ein Abstand in der Größenordnung des Hallradius (4 m) und dessen vierfache Wert (16 m) gewählt. Auf der Grundlage dieser binaural und dynamisch (mit head tracking) synthetisierten Datensätze wurden in einem 3AFC-Versuch Schwellwerte für eine eben wahrnehmbare Separation

der Quellen ermittelt. Hierbei wurde der Winkel für die Separation der Quellen adaptiv auf den Schwellwert nachgeführt. Als Stimuli wurden 2- bzw. 4-kanalige Rauschpulse (individuell oktavbreit gefiltertes Rauschen von 250 ms Länge und 750 ms Pause, mehrfach wiederholt und je Kanal zeitlich versetzt) bzw. ein ‚natürlicher‘ Stimulus präsentiert (nachhallfreie Aufnahmen eines Streichquartetts im reflexionsarmen Raum der TU Berlin, nachhallfreie Holzbläser- und Sprecheraufnahmen von ca. 5 s Dauer). Die BRIR-Datensätze wurden mit einer Winkelauflösung von  $1^\circ$  in der Horizontalebene vermessen und gestatteten den VPn während des Hörversuchs horizontale Kopfbewegungen im Bereich von  $\pm 80^\circ$ . Die für die Messung der binauralen Datensätze verwendete Variation der Winkel und der gesamte Öffnungswinkelbereich wurde auf der Grundlage eines Vorversuchs mit 5 VP und zwei Schallquellen ermittelt.

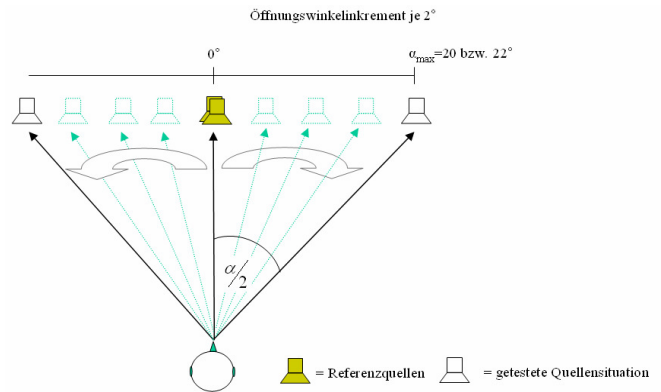


Abbildung 1: Im Hörversuch auralisierte Situation 1 (2 Quellen) - die VP hört je Trial zweimal die Referenzprobe und einmal die ‚heiße Probe‘ (Quellen mit variablem Öffnungswinkel  $\alpha$ ) und detektierte jeweils die ‚heiße‘ Probe

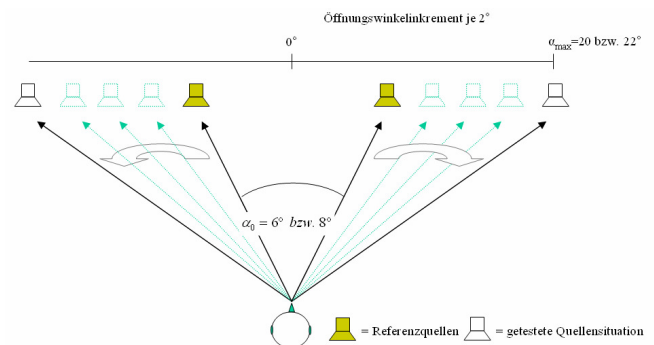


Abbildung 2: Im Hörversuch auralisierte Situation 2 (4 Quellen) - 2 Quellen bilden die Referenzsituation, der Öffnungswinkel der äußeren Quellen wird sukzessive adaptiert

Die Öffnungswinkelschrittweite wurde daraufhin auf  $2^\circ$ , der Öffnungswinkelbereich sitzplatzabhängig auf  $20^\circ$  bzw.  $22^\circ$  festgelegt (Abbildung 1). Der Hörversuch wurde mit 22, überwiegend musikalisch vorgebildeten VP (90% männl.)

als univariater, 3-faktorieller, zweifach gestufter und vollständiger Test durchgeführt. Da die Auswirkungen aller Faktorstufenkombinationen an derselben Stichprobe gemessen wurden, sind diese intraindividuellen Beobachtungen nicht unabhängig (Test mit Messwiederholungen). Für jede VP wurde ein Schwellwert für den eben erkennbaren Öffnungswinkel der Einzelquellen in Abhängigkeit von der Entfernung (4m, 16m), dem Stimulus (Rauschpulskomplex, natürliches Klangbeispiel) und der Quellenanzahl (2/4) gemessen. Zu Beginn der Versuche wurden individuelle Trainingssessions anhand deutlich unterscheidbarer Stimuli durchgeführt.

## Ergebnisse

Die aufgrund des Designs mit Messwiederholungen vorliegende individuelle Reliabilität war für eine explorative Fragestellung zufrieden stellend hoch (Kronbachs  $\alpha = 0,779$ ). Die Verteilung der individuellen Schwellwerte des für die Separation von Einzelquellen erforderlichen Öffnungswinkels über die eingesetzten VPn war normal (Kolmogoroff-Smirnoff Test). Die Streuung war relativ gering  $\sigma = 2,84^\circ$ , der Mittelwert lag bei  $10,42^\circ$  (95%-KI d. MW:  $[9,2^\circ \ 11,6^\circ]$ ). Die Verteilung aller erhobenen Antworten (also inkl. intraindividuelle Varianz) war jedoch nicht normalverteilt. Ebenso wenig war die Varianzhomogenität der Antworten gegeben (Levene Test). Zur Untersuchung der Haupteffekte sind daher nur nichtparametrische Tests (Wilcoxon Test, abhängige Stichproben) indiziert. Die Ergebnisse der statistischen Untersuchung lassen sich wie folgt zusammenfassen:

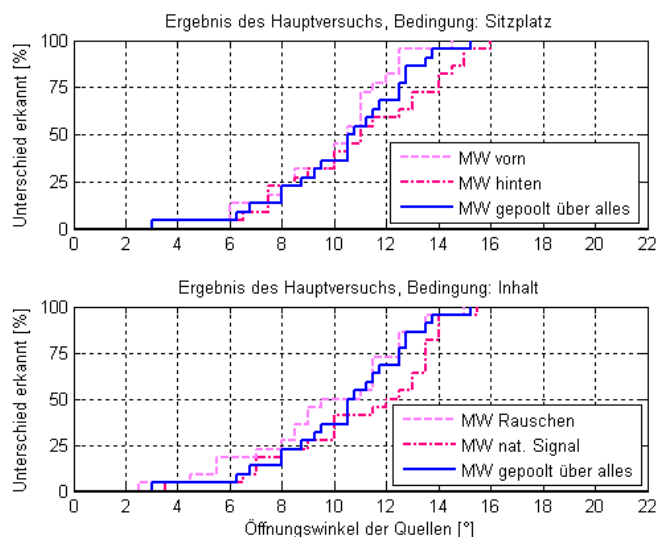
1. Die Anzahl der Quellen wirkt sich (innerhalb des untersuchten Bereichs mit 2 bzw. 4 Quellen) nicht auf den zur Unterscheidung notwendigen Winkel aus.
2. Mit zunehmender Entfernung zwischen Quelle und Empfänger bzw. zunehmendem Diffusanteil nimmt auch der eben wahrnehmbare Winkelunterschied zu ( $p_{\text{Wilcoxon}} = .101$ ,  $\Delta = 1^\circ$  für die untersuchten Abstände von 4 bzw. 16 m).
3. Rauschpuls sind gegenüber den verwendeten natürlichen Klangbeispielen bereits bei signifikant geringem Winkel als räumlich getrennt wahrnehmbar ( $p_{\text{Wilcoxon}} = .044$ ,  $\Delta = 1,4^\circ$ ).

Der erste Punkt zeigt, dass der für das Verfolgen von 4 Quellsignalen erforderliche, höhere kognitive Aufwand nicht auf Kosten der Lokalisationsleistung geht. Somit kann eine Schwelle für die Separation von Einzelquellen zunächst unabhängig von der Anzahl der vorhandenen Einzelschallquellen angegeben werden. Abbildung 3 zeigt die kumulative Darstellung der individuellen Schwellwertverteilung und entspricht damit einer Schätzung der tatsächlichen psychometrischen Funktion, zusätzlich sind die laut Punkt 2 und 3 eingetretenen Unterschiede dargestellt.

## Diskussion

Aus den Ergebnissen lassen sich Kriterien für den Einsatz von separaten binauralen Auralisationen für komplexe Klangquellen ableiten. So kann man erwarten, dass mindes-

tens 50% der Hörer den Unterschied zwischen monofonen und räumlich verteilten Quellen nicht wahrnehmen, wenn die Quellen vom Hörer aus unter einem Öffnungswinkel von  $\alpha \leq 10^\circ$  sichtbar sind. Um für 95% der Hörer eine von der Realität nicht unterscheidbare Simulation zu gewährleisten, kann man nur Quellen mit einem Öffnungswinkel von  $\alpha \leq 5^\circ$  zu einem Quellort zusammenfassen.



**Abbildung 3:** Kumulierte Verteilungsfunktion der im Hauptversuch ermittelten mittleren individuellen Schwellwerte für den zur auditiven Separation multipler Klangquellen notwendigen Öffnungswinkel im Vergleich zu den Bedingungsvariationen Abstand und Audioinhalt

Die genannten Werte gelten für gut lokalisierbare Signale (wie gepulstes Rauschen) und für ein etwa ausgeglichenes Direkt/Diffus-Verhältnis. Für Musik und Sprache sowie für höhere Anteile des Diffusalls liegt der Schwellwert um etwa  $1-2^\circ$  höher. Dies bedeutet, dass etwa für eine Streichquartettbesetzung mit einem Abstand zwischen den Musikern von etwa 1.5m an einem Hörerplatz in 4 m Entfernung mit 4 individuellen Quellen simuliert werden muss, während an einem Hörerplatz in 16 m Entfernung eine Darstellung mit zwei oder drei Quellen ausreichend sein kann.

## Literatur

- [1] Blauert, J.: Spatial Hearing. The Psychophysics of Human Sound Localization. 2. Aufl., Cambridge, MA.: MIT Press, 1997.
- [2] Damaske, P.: Subjektive Untersuchungen von Schallfeldern. *Acustica*, 19 (1967/68), 199-213.
- [3] Wagener, B.: Räumliche Verteilung der Hörrichtungen in synthetischen Schallfeldern. *Acustica* 25 (1971), 203-219.
- [4] Vigeant, M. C.; Wang, L. M.; Rindel, J. H.: Investigations of Multi-Channel Auralization Technique for Solo Instruments and Orchestra, ICA Proc. Madrid, 2007.
- [5] Witew, I.B.; Paprotny, J.; Behler, G.: Auralization of orchestras in concert halls using numerous uncorrelated sources, Proc. of the Institute of Acoustics of the RWTH Aachen 28 (2006), No. 2, 293-296.
- [6] Lindau, A.; Weinzierl, S.: Fabian - Schnelle Erfassung binauraler Raumimpulsantworten in mehreren Freiheitsgraden. Fortschritte der Akustik, DAGA Stuttgart 2007, 633-634.
- [7] Lindau, A.; Hohn, T.; Weinzierl, S.: Binaural resynthesis for comparative studies of acoustical environments, 122nd AES Convention, Vienna 2007, preprint no. 7032.