

Binaurale Modellierung der Wahrnehmung mit Wellenfeldsynthese generierter Schallfelder

Hagen Wierstorf, Sascha Spors, Alexander Raake

Deutsche Telekom Laboratories, Technische Universität Berlin, Email: hagen.wierstorf@telekom.de

Einleitung

Bei der Wellenfeldsynthese (WFS) handelt es sich um ein vielkanaliges Schallfeldsyntheseverfahren, das in der Lage ist das Schallfeld äußerer virtueller Quellen in einem ausgedehnten Zuhörerbereich zu synthetisieren [1]. Unter Ausnutzung der zeit-inversen akustischen Fokussierung ist zudem die Wiedergabe von sogenannten fokussierten Quellen im Zuhörerbereich möglich. Da zur Synthese des Schallfelds eine räumlich kontinuierliche Verteilung von Lautsprechern nötig wäre, kommt es bei den in der Praxis verwendeten Lautsprecherarrays zu einer räumlichen Unterabtastung. Dieses führt zu Aliasartefakten oberhalb der räumlichen Aliasfrequenz f_{al} , welche für heutige Lautsprecheranordnungen zumeist unterhalb von 2000 Hz liegt. Diese machen sich durch zusätzliche gefilterte Wiederholungen der Wellenfronten bemerkbar. Dabei trägt jeder einzelne Lautsprecher eine Wiederholung bei. Für synthetisierte Schallfelder von virtuellen Quellen außerhalb des Zuhörerbereiches treten diese Wiederholungen erst nach der ersten, gewünschten Wellenfront auf. Dies führt dazu, dass die Artefakte in der Wahrnehmung des Schallfelds in den Hintergrund treten, da diese auf Grund des Präzedenzeffektes [2] durch die erste Wellenfront dominiert werden. Für fokussierte Quellen erreichen die ungewollten Wiederholungen jedoch zuerst den Zuhörer, wie in Abbildung 1 zu erkennen ist. Wird davon ausgegangen, dass auch im Fall der fokussierten Quellen der Präzedenzeffekt die Wahrnehmung bestimmt, sollten folgende zwei Hypothesen zutreffen: (1) für genügend kurze Lautsprecheranordnungen sind keine Artefakte hörbar; (2) die Lokalisierung der fokussierten Quelle wird durch den Lautsprecher bestimmt, dessen Wellenfront als erstes beim Zuhörer eintrifft. Bei kürzeren Lautsprecherarrays kommt es neben einem offensichtlich kleinerem Zuhörerbereich auch zu frequenzabhängigen Beugungseffekten im Schallfeld, da das Lautsprecherarray wie ein Spalt fungiert, durch den das Schallfeld hindurchtritt. Die Beugung führt zu Nebenmaxima und -minima im Schallfeld, was wiederum die interaurale Pegeldifferenz (ILD) beeinflusst.

Um die Hypothesen zu untersuchen wurde ein Versuch mit verschiedenen Lautsprecheranordnungen und Zuhörerpositionen für die Synthese einer fokussierten Quelle durchgeführt. Zudem wurde ein binaurales Modell nach Lindemann [3] verwendet um den Einfluss des Präzedenzeffektes und der Beugung auf die Lokalisation der fokussierten Quelle genauer zu untersuchen. Das Modell wertet die interaurale Laufzeitdifferenz (ITD) mit Hilfe einer Kreuzkorrelation aus und integriert die ILD in die Korrelation, indem durch kontralaterale Inhibiti-

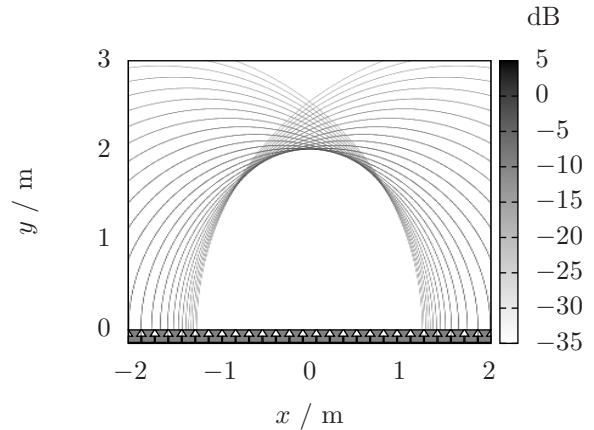


Abbildung 1: Simulation des Schallfelds einer fokussierten Quelle, die einen breitbandigen Impuls aussendet, positioniert bei $\mathbf{x}_s = (0, 1)$ m zum Zeitpunkt $t = 2,9$ ms nach dem Durchqueren des Fokuspunktes. Die weiteren Parameter der Simulation waren $L = 4$ m, $\Delta x = 0,15$ m, $\mathbf{x}_{ref} = (0, 2)$ m.

on das Maximum der Kreuzkorrelation entsprechend der ILD verschoben wird. Das Modell bezieht jedoch nicht den Präzedenzeffekt ein. Damit sollte mit Hilfe des Modells unterscheidbar sein, wann zur Erklärung der Lokalisation der Präzedenzeffekt benötigt wird und wann nicht.

Methode

Stimuli: Es wurden keine realen Lautsprecherarrays verwendet, sondern diese über dynamische binaurale Resynthese realisiert. Für Details hierzu und den verwendeten Ansteuerungsfunktionen für die WFS sei auf [4] verwiesen. Die verwendeten Lautsprecherarrays hatten eine Länge von 10 m, 4 m, 1,8 m, 0,75 m und 0,3 m. Die fokussierte Quelle war am Punkt $\mathbf{x}_s = (0, 1)$ m positioniert (siehe Abb. 1). Sechs verschiedene Zuhörerpositionen wurden ausgewählt, von denen je drei auf einem Radius von 1 m bzw. 4 m um die fokussierte Quelle herum angeordnet waren. Dabei haben sie einen Winkel von 0° , 30° und 60° zur y -Achse eingenommen. Für den Lautsprecherarray mit einer Länge von $L = 10$ m wurden nur die Positionen mit einem Radius von $R = 4$ m betrachtet für einen Lautsprecherarray mit $L = 4$ m nur Positionen mit $R = 1$ m. Die anfängliche Kopforientierung der Zuhörer war immer in Richtung der fokussierten Quelle, so dass sie von vorne wahrgenommen werden sollte. Zusätzlich wurde als Referenz ein einzelner Lautsprecher an der Position der fokussierten Quelle verwendet. Als Audiomaterial kamen weibliche Sprache und Kastagnet-

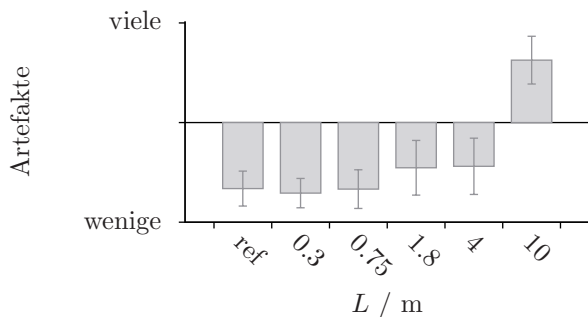


Abbildung 2: Mittelwert und Varianz für die unterschiedlichen Lautsprecherarrays für das Attributpaar *wenig Artefakte* vs. *viele Artefakte*.

ten zum Einsatz.¹

Darbietung: Aufgabe der Versuchspersonen war es, die dargebotenen Bedingungen für die zwei Attributpaare *viele Artefakte* vs. *wenig Artefakte* und *links* vs. *rechts* zu bewerten. Hierzu wurde den Personen ein Bildschirm mit dem Attributpaar im oberen Bereich eingeblendet. Darunter befanden sich dann neun verschiedene Bedingungen (jeweils alle für einen Winkel möglichen), zwischen denen die Versuchsperson beim Anhören instantan hin und her schalten konnte. Jede Bedingung war mit einem Schieberegler ausgestattet, den die Versuchsperson entsprechend ihrer Wahrnehmung und dem Attributpaar einstellen sollte.

Durchführung: An dem Versuch nahmen 6 normalhörende Versuchspersonen, die für ihre Teilnahme keine Aufwandsentschädigung erhielten. Die Darbietung erfolgte mit Hilfe dynamischer binauraler Resynthese über einen Kopfhörer AKG K601 und mit Hilfe eines Headtrackers Polhemus Fastrack.

Ergebnisse

In Abbildung 2 sind die Ergebnisse als Mittelwert und Varianz für das Attributpaar *wenig Artefakte* vs. *viele Artefakte* über alle sechs Versuchspersonen, beide verschiedenen Audiomaterialien und alle Zuhörerpositionen eingezeichnet. Es ist für die Artefakte zu erkennen, dass die beiden Lautsprecherarrays mit einer Länge von 0,3 m oder 0,75 m sich nicht signifikant von der Referenz unterscheiden. Erst ab einer Länge von 1,8 m sind zusätzliche Artefakte hörbar. Ebenso ist zu erkennen, dass diese umso stärker werden, je längere Arrays verwendet werden.

In Abbildung 3 sind die Ergebnisse für die Lokalisation der fokussierten Quelle dargestellt, die für ein fehlerfreies Schallfeld immer in der Mitte liegen sollte. Es ist zu erkennen, dass für kürzere Arrays eine Lateralisierung auf die linke Seite stattfindet. Dies ist besonders bei dem kleineren Radius $R = 1$ m zu beobachten. Ebenfalls sind die Ergebnisse des binauralen Modells dargestellt. Das Modell wurde mit den Standardparametern angewendet und lediglich mit einem Faktor skaliert, um die Ergebnis-

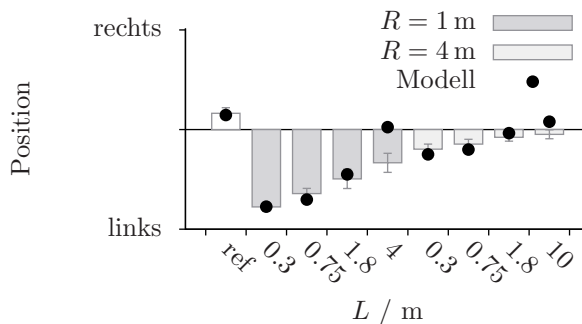


Abbildung 3: Mittelwert und Varianz für die unterschiedlichen Konditionen für das Attributpaar *links* vs. *rechts*. Zusätzlich sind die Ergebnisse des binauralen Modells eingetragen.

se in den selben Graphen eintragen zu können. Es ist zu erkennen, dass das Modell für die Referenzkondition und die drei kurzen Arrays in der Lage ist die beobachtete Lokalisation richtig vorherzusagen. Lediglich für die beiden langen Arrays ist eine Abweichung zu beobachten.

Diskussion

In der Einleitung wurden zwei Hypothesen aufgestellt, die sich aus der Verwendung des Präzedenzeffektes zur Erklärung der Wahrnehmung fokussierter Quellen ergeben. Eine davon besagte, dass für kürzere Lautsprecherarrays keine wahrnehmbaren Artefakte auftreten, da der Präzedenzeffekt die Wahrnehmung einzelner Wiederholungen unterdrückt. Dies ist offenbar der Fall (vgl. Abb. 2). Die Lokalisierung der fokussierten Quellen wird hingegen für kurze Lautsprecherarrays vornehmlich durch die Beugungsartefakte bestimmt. Dies ist daran zu erkennen, dass kürzere Arrays eine Lokalisation weiter nach links bewirken und das binaurale Modell, welches nicht in der Lage ist den Präzedenzeffekt zu berücksichtigen, dies richtig vorherzusagen kann. Lediglich für die beiden langen Arrays mit $L = 4$ m und $L = 10$ m wird die Lokalisation wahrscheinlich durch den Präzedenzeffekt bestimmt, da bei dieser Arraylänge Beugungseffekte keine große Rolle mehr spielen und das binaurale Modell auch eine mittige Lokalisation erwarten würde.

Literatur

- [1] A.J. Berkhout et al.: Acoustic control by Wave Field Synthesis. *JASA* 93 (1993), 2764–2778.
- [2] H. Wallach et al.: The precedence effect in sound localization. *AJP* 57 (1949), 315–336.
- [3] W. Lindemann: Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals. *JASA* 80 (1986), 1608–1622.
- [4] H. Wierstorf et al.: Reducing artifacts of focused sources in Wave Field Synthesis. *Proc. 2010 129th AES Conv.*

¹Audiobeispiele für die einzelnen Bedingungen finden sich unter <http://audio.qu.tu-berlin.de/?p=574>