

Wahrnehmung künstlich erzeugter Schallfelder

Hagen Wierstorf¹, Sascha Spors², Alexander Raake¹

Email: hagen.wierstorf@tu-berlin.de

¹ *Assessment of IP-based Applications, Technische Universität Berlin, 10587 Berlin, Deutschland*

² *Institute of Communications Engineering, Universität Rostock, 18119 Rostock/Warnemünde, Deutschland*

Einleitung

Die Schallfeldsynthese beruht auf dem physikalischen Prinzip, dass ein Schallfeld in einem Zuhörerraum eindeutig durch den Schalldruck an seinem Rand bestimmt ist. In der Praxis wird versucht, den Schalldruck am Rand durch Lautsprecher zu steuern. Zur Berechnung entsprechender Ansteuerungssignale der Lautsprecher gibt es unterschiedliche Verfahren: In diesem Beitrag werden die Wellenfeldsynthese und band-begrenztes Ambisonics höherer Ordnung diskutiert. Die Verwendung von einzelnen Lautsprechern stellt eine räumliche Abtastung des Randes dar und kann zu frequenzabhängigen Fehlern im synthetisiertem Schallfeld führen. Dies wiederum kann Beeinträchtigungen der Lokalisation und der Klangfarbe synthetisierter Quellen zur Folge haben.

Das synthetisierte Schallfeld der Wellenfeldsynthese und Ambisonics unterscheidet sich nur für Frequenzen, für die die räumliche Abtastung durch die Lautsprecher nicht ausreichend ist. In diesem Fall sind im Schallfeld auftretende Fehler für die Wellenfeldsynthese gleichmäßiger verteilt, während es bei Ambisonics einen ungestörten Bereich in der Mitte der Lautsprechergruppe und stärkere Fehler außerhalb dieses Bereiches gibt.

Das folgende Experiment untersucht, wie sich die unterschiedliche Verteilung der Fehler in der Wellenfeldsynthese und Ambisonics auf die Lokalisation einer synthetisierten Punktquelle an unterschiedlichen Positionen im Zuhörerbereich auswirkt.

Experiment

Es wurde an 16 verschiedenen Zuhörerpositionen die Lokalisation einer synthetisierten Punktquelle untersucht. Die Positionen sind in Abbildung 1 an Hand des Mittelpunktes der dort eingezeichneten Pfeile erkennbar. Die Punktquelle wurde jeweils mit Wellenfeldsynthese und Ambisonics für drei verschiedene Lautsprechergruppen untersucht. Die Lautsprechergruppen wiesen alle den selben Durchmesser von 3 m auf, hatten jedoch einen unterschiedlichen Lautsprecherabstand von 17 cm, 34 cm und 67 cm.

Als Stimuli wurden weiße Rauschpulse verwendet. Aufgabe der Versuchspersonen war es in die Richtung zu schauen, aus der sie das Rauschen wahrnehmen und eine Taste zu drücken. Im Anschluss wurde sofort der nächste Stimulus dargeboten. Um die unterschiedlichen Zuhörerpositionen und unterschiedlichen Lautsprechergruppen zusammen untersuchen zu können, wurde das

gesamte Experiment per Kopfhörer und dynamischer Binauralsynthese realisiert. In einer Vorstudie wurde bestätigt, dass diese keinen Einfluss auf die erzielten Ergebnisse hat [1].

An dem Versuch nahmen 12 Versuchspersonen teil. Die Ergebnisse sind in Abbildung 1 zusammengefasst. Dort ist für jede Zuhörerposition jeweils die mittlere wahrgenommenen Richtung aller Versuchspersonen mit Hilfe eines Pfeils eingezeichnet. Die Farbe des Pfeils gibt die absolute Abweichung von der gewollten Richtung der Punktquelle an. Dieses wird im Folgendem als Lokalisationsfehler bezeichnet. Je roter der Pfeil, desto größer die Abweichung. Neben jeder Zeile von Zuhörerpositionen ist die mittlere absolute Abweichung über alle Positionen in der entsprechenden Zeile angegeben. Das Konfidenzintervall liegt bei allen Zuhörerpositionen, Lautsprechergruppen und Schallfeldsyntheseverfahren um die 2°. Es ist zu erkennen, dass der Lokalisationsfehler umso größer wird, je weniger Lautsprecher verwendet werden. Insbesondere bei Ambisonics kommt es außerhalb des Zentrums zu sehr großen Abweichungen für die Lautsprechergruppe mit einem Lautsprecherabstand von 67 cm. Teilweise wurde hier sogar mehr als eine Quelle wahrgenommen, was durch die eingezeichnete Anzahl an Pfeilen angedeutet ist. Bei einem Lautsprecherabstand von 17 cm ist die Lokalisation für die Wellenfeldsynthese hingegen nicht von der einer realen Quelle unterscheidbar.

Modellierung

Um die wahrgenommenen Richtungen mit einem auditorischen Modell zu ermitteln wurde das binaurale Modell nach Dietz et al. [2] verwendet und leicht erweitert. Das Modell startet mit einer monauralen Vorverarbeitung in beiden Kanälen: Bandpass, Gammatonfilterung, Halbwellengleichrichtung, Tiefpass. Anschließend wird die interaurale Zeitdifferenz bestimmt. Diese wird dann in jedem Frequenzkanal in einen Winkel umgerechnet. Hierzu wird eine Zuordnungstabelle verwendet, die mit Hilfe der Kopfübertragungsfunktionen erstellt wurde, welche auch für die Binauralsynthese verwendet wurden. In einem letzten Schritt wird aus den Richtungen der einzelnen Frequenzbänder der Median als Richtung der wahrgenommenen Quelle berechnet. Dieses geschieht mit Hilfe einer gleichmäßigen Gewichtung aller Frequenzbänder und einer einfachen Ausreißerentfernung. Hierbei werden die Richtungen nicht berücksichtigt, die mehr als 30° vom Median abweichen. Das Modell ist als Bestandteil der *Auditory Modelling Toolbox* [3] frei verfügbar.

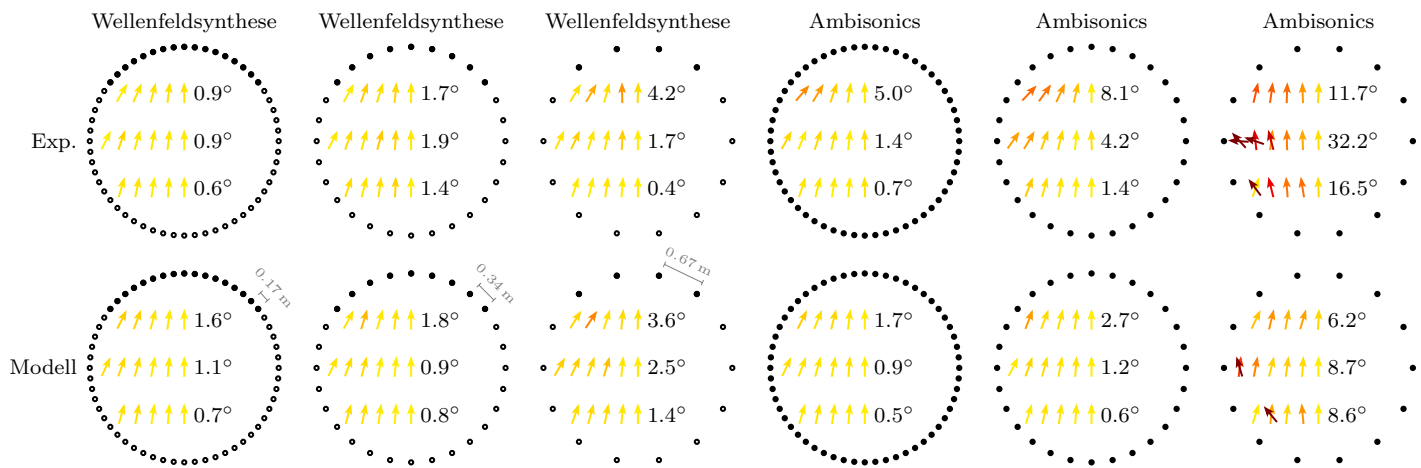


Abbildung 1: Mittelwerte der Lokalisationsergebnisse aus dem Experiment im Vergleich mit den Modellvorhersagen. Die sechs oberen Graphen zeigen die Ergebnisse aus dem Experiment, wobei an jeder Zuhörerposition ein Pfeil die wahrgenommene Richtung anzeigt. Die Farbe des Pfeils kodiert die absolute Abweichung dieser Richtung von der gewollten Richtung der synthetisierten Punktquelle.

Zur Vorhersage der Richtung wurden die gleichen Signale, welche den Versuchspersonen für eine Kopforientierung von 0° dargeboten wurden, in den Eingang des binauralen Modells geleitet. Die Vorhersagen des Modells weisen insbesondere im Fall der Wellenfeldsynthese eine große Übereinstimmung mit den im Experiment beobachteten Richtungen auf, vergleiche Abbildung 1. Für Ambisonics liefert das Modell hingegen eine Unterschätzung des Lokalisationsfehlers. Diese ist umso ausgeprägter je größer der Lautsprecherabstand.

Bei einem Lautsprecherabstand von 67 cm kam es im Experiment zu Fällen, bei denen mehr als eine Quelle von den Versuchspersonen wahrgenommen wurde. Dies wurde an Hand der Verteilung aller Antworten für eine Zuhörerposition ermittelt. Für einige Positionen wiesen diese keine Normalverteilung auf, sondern konnten zwei unterschiedlichen Verteilungen zugeordnet werden. Das gleiche konnte im Modell realisiert werden. Hier wurde die Vorhersage auf insgesamt zehn unterschiedliche Kopforientierungen pro Zuhörerposition ausgeweitet und die Verteilung der Vorhersagen analysiert. An zwei Zuhörerpositionen kam es zu mehr als einer Verteilung, was ebenfalls durch zwei eingezeichnete Pfeile gekennzeichnet ist. Die Zuhörerpositionen mit mehr als einer wahrgenommenen Quelle stimmen nicht komplett überein zwischen Experiment und Modell. Das Modell ist jedoch in der Lage die kritische Lautsprechergruppe, Schallfeldsynthesemethode und grobe Position für diese Fälle zu ermitteln.

Zusammenfassung

Die beiden Schallfeldsyntheseverfahren Wellenfeldsynthese und Ambisonics unterscheiden sich insbesondere in den Fällen räumlicher Unterabtastung durch die Lautsprecher. Während bei Wellenfeldsynthese die Artefakte im Schallfeld gleichmäßig im Zuhörerbereich verteilt sind, bei ungefähr gleichem Abstand zur Position der synthetisierten Quelle, weist Ambisonics eine ungleiche Verteilung auf. Hier kommt es zu einem Bereich

im Zentrum des Zuhörerbereiches, indem keine Fehler auftreten, während diese verstärkt außerhalb des Zentrums auftreten.

In einem Lokalisationsexperiment konnte gezeigt werden, dass sich dies auch in den Ergebnissen der Versuchspersonen niederschlägt. Während die Lokalisationsfehler für die Wellenfeldsynthese gleichmäßig im Zuhörerbereich verteilt sind, kommt es bei Ambisonics zu größeren Fehlern außerhalb des Zentrums der Lautsprechergruppe. Insbesondere bei großen Lautsprecherabständen sind diese Fehler groß und können an einigen Zuhörerpositionen zur Wahrnehmung mehrerer Quellen führen.

Ein binaurales Modell nach Dietz et al. [2] wurde leicht modifiziert und war anschließend in der Lage, die beobachteten Ergebnisse korrekt zu modellieren. Lediglich im Fall von Ambisonics weisen die Modellvorhersagen teilweise größere Abweichungen auf, insbesondere außerhalb des Zentrums.

Danksagung

Diese Studie wurde gefördert durch DFG RA 2044/1-1.

Literatur

- [1] H. Wierstorf, S. Spors, A. Raake - Perception and evaluation of sound fields. 59th Open Seminar on Acoustics (2012), 263-68
- [2] M. Dietz, S. D. Ewert, V. Hohmann - Auditory model based direction estimation of concurrent speakers from binaural signals. *Speech Communication* (2011), 53(5), 592-605
- [3] P. Søndergaard, P. Majdak - The auditory-modeling toolbox. In *The technology of binaural listening* (2013), Hrsg. J. Blauert, 33-56