

Simulation und psychoakustische Bewertung von Übertragungsfehlern bei der Wellenfeldsynthese

Florian Lindner, Florian Völk, Hugo Fastl

AG Technische Akustik, MMK, TU München, 80333 München, Deutschland, E-Mail: florian.voelk@mytum.de

Einleitung

Wellenfeldsynthese (WFS) ist ein Audio-Wiedergabeverfahren, mit dessen Hilfe das Schallfeld einer Referenzszene in einem begrenzten Volumen, dem Hörraum, möglichst fehlerfrei synthetisiert werden soll. Theoretisch ist die korrekte Synthese des von einer Quelle (Primärquelle) erzeugten Wellenfeldes unter bestimmten Bedingungen im Hörraum zum Beispiel mit einer unendlich großen Anzahl von Sekundärquellen, basierend auf dem Kirchhoff-Helmholtz-Integral (KHI), möglich. Als Primärquellen werden häufig Quellen von Kugelwellen oder ebenen Wellenfronten verwendet. Für Implementierungen der Wellenfeldsynthese müssen jedoch die theoretischen Anforderungen auf ein realisierbares Maß angepasst werden, was unweigerlich zu Abweichungen zwischen synthetischem und primärem Schallfeld führt.

In diesem Beitrag soll zur Klärung der Frage beigetragen werden, unter welchen Bedingungen diese Abweichungen hörbar sind. Zunächst werden dafür die von ausgesuchten Wellenfeldsynthesensystemen erzeugten Schalldruck-Zeitfunktionen an einem bestimmten Ort im Hörraum mittels einer Rechnersimulation synthetisiert. Die so erhaltenen Signale werden anschließend in psychoakustischen Experimenten nach dem "three-alternative forced-choice"-Verfahren bei diotischer Kopfhörerwiedergabe mit dem entsprechenden Originalsignal verglichen. Auf diese Weise ist es möglich, die Hörbarkeit prinzipieller Fehler im Wellenfeldsyntheseverfahren und seiner Herleitung unabhängig von einer speziellen Implementierung zu untersuchen. Darüber hinaus erlaubt es dieses Vorgehen, die Ursache für bestimmte Fehler zielgerichtet im Hörversuch einzugrenzen, und so die physikalischen Auslöser zu bestimmen.

Grundlagen und mögliche Übertragungsfehler

Gemäß dem KHI sind kontinuierlich an jedem Punkt einer geschlossenen (dreidimensionalen) Volumenoberfläche zwei Sekundärquellen, jeweils eine mit dem entsprechenden driving signal angesteuerte Dipol- und Monopolquelle, zu platzieren, um ein Schallfeld innerhalb des Volumens korrekt synthetisieren zu können, und gleichzeitig kein Schallfeld außerhalb des Volumens zu erzeugen (vgl. für die theoretische Herleitung hier und im Folgenden [2]). Um zu realisierbaren Situationen zu gelangen, geht man zumeist von diskreten (zweidimensionalen) Kreis- oder Linienanordnungen aus Monopolquellen aus. Die driving signals müssen entsprechend angepasst werden, was Näherungen in deren auf dem KHI basierender Herleitung erfordert. So ist durch die Eliminierung der Dipolquellen das Feld schon im dreidimensionalen Fall außerhalb des Volumens nicht mehr gleich null und auch im Volumen entstehen im allgemeinen Fall Fehler. Neben einem geschlossenen Volumen können auch zwei Halbräume mit ebener Begrenzung betrachtet werden. In diesem speziellen Fall ist die korrekte Synthese innerhalb eines Halbraumes mit einer kontinuierlichen, un-

endlich ausgedehnten Monopol-Sekundärquellenverteilung möglich. Zusätzliche Fehler sind für finite und diskrete Anordnungen zu erwarten. Die Begrenzung von unendlich ausgedehnten Sekundärquellenanordnungen führt unweigerlich zu Fehlern im Wellenfeld, die üblicherweise durch örtliche Fensterung an den Rändern der Verteilung (Truncation) reduziert werden. Die Diskretisierung der Sekundärquellenanordnung entspricht einer örtlichen Abtastung wodurch sich eine obere Grenzfrequenz, die sogenannte Aliasfrequenz f_a ergibt, ab der keine fehlerfreie Synthese mehr möglich ist.

Für die korrekte Reduktion auf die Schallfeldreproduktion in einer Ebene wäre eine Anordnung aus senkrecht zur betrachteten Ebene angeordneten Linienquellen erforderlich, die in praktischen Anwendungen typischerweise als in der Ebene befindliche Punktquellen angenähert werden. Durch diese Näherung ergeben sich weitere Fehler im resultierenden Wellenfeld, die durch auf eine Position (Optimierungspunkt) optimierte Korrekturen minimiert werden können. Für die Simulation primärer Kugelquellen entsteht in diesem Fall ein zusätzlicher Amplitudenfehler, der bei Verwendung einer Linienquelle als Primärquelle nicht auftritt. Bezeichnet f die Frequenz und r den Abstand von einer Linienquelle, so wird das Spektrum des abstrahlenden Schalls bei der Abstrahlung im Vergleich zur Kugelquelle mit $f^{-1/2}$ gedämpft, die Amplitude fällt mit $r^{-1/2}$ (vgl. [3]) statt wie bei der Kugelquelle mit r^{-1} . Im Rahmen dieser Arbeit werden diese Abweichungen korrigiert, was für den Amplitudenfehler wiederum nur für eine bestimmte Position möglich ist.

Simulationsverfahren

Um im Hörversuch die Unterscheidbarkeit von Original und Synthese untersuchen zu können, werden zunächst die Übertragungsfunktionen jeweils bestimmter Primärquellen und deren Synthese zum Hörort basierend auf [2] berechnet und im Zeitbereich als FIR-Filter implementiert. Betrachtet werden kontinuierliche und diskrete kugel- und kreisförmige Sekundärquellenverteilungen mit einem Radius von 1.3m, sowie ebene (3.9m x 3.3m) und lineare (3.9m) Anordnungen. Als Primärquellen werden auf den Hörort aus Arrayrichtung einfallende ebene Wellenfronten und punktförmige Kugelquellen außerhalb des Volumens in 2m Abstand zum Hörort betrachtet. Tabelle 1 zeigt alle betrachteten Konfigurationen. Da die Diskretisierung mit 8.5cm Sekundärquellenabstand vorgenommen wird, ergibt sich $f_a \approx 2\text{kHz}$. Der Hörort wird am Optimierungspunkt des jeweiligen Syntheseverfahrens gewählt, der für die kreis- bzw. kugelförmigen Anordnungen im Mittelpunkt sowie für die linearen und ebenen auf der Symmetrieachse festgelegt wird. Verwendet wird für alle Korrekturen der gleiche Optimierungspunkt. Für lineare und ebene Anordnungen wird entlang jeder Dimension kosinusförmige Truncation über 20% der entsprechenden Abmessung vorgenommen. Auftretende Integrationen werden numerisch mit 32 Bit Wortbreite berechnet.

Tabelle 1: Im Hörversuch betrachtete Wellenfeldsynthesekonfigurationen. Synthetisiert wird der von Monopol-Sekundärquellen erzeugte Schalldruck am Optimierungsort.

Konfiguration	Sekundärquellenverteilung	Primäre Wellenfront
1	3D kugelförmig, kontinuierlich	eben
2		kugelförmig
3	3D kugelförmig, diskret	eben
4		kugelförmig
5	3D eben, kontinuierlich	eben
6		kugelförmig
7	2.5D kreisförmig, kontinuierlich	eben
8		kugelförmig
9		kugelförmig, korr.
10	2.5D linear, kontinuierlich	eben

Hörversuchsmethode und Vorgehen

Die Hörversuche wurden in einer abgedunkelten und schallisolierten Kabine durchgeführt, die Schallwiedergabe erfolgte diotisch über Kopfhörer (Beyer DT 48) mit vorgeschaltetem Freifeldentzerrer nach [1]. Als Versuchsmethode kam ein „three-alternative forced-choice“ (3-AFC) Verfahren zum Einsatz: Die Probanden hörten drei Schalle, davon zwei identische, und mussten in jedem Fall den sich von den beiden anderen unterscheidenden Schall bestimmen. Die Wahrnehmbarkeitsschwelle wird als 67%-Punkt der psychometrischen Funktion festgelegt. Testschalle waren ein gleichmäßig anregender Rauschpuls (GAR, vgl. [1], Pulsdauer 1s) sowie ein Ausschnitt eines in reflexionsarmer Umgebung aufgezeichneten Sprachsignals (weibliche Sprecherin, 1s Dauer), jeweils breitbandig und für alle diskreten Anordnungen zusätzlich oberhalb von f_a bandbegrenzt. Vor der Darbietung wurden die Schalle mit dem die betrachtete Synthese oder Originalsituation repräsentierenden FIR-Filter gefaltet und mit 44.1kHz Abtastrate und 24 Bit Wortbreite mit einem freifeldäquivalenten Pegel um 63dB SPL in zufälliger Reihenfolge ausgegeben. Acht Probanden beurteilten jeden Stimulus dreimal. Der Versuch wurde in vier Sitzungen mit einer mittleren Dauer von etwa 17 Minuten unterteilt.

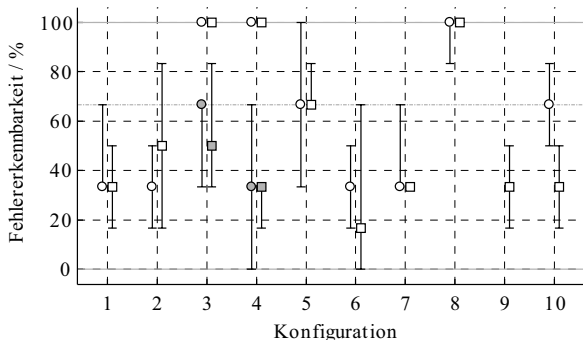


Abbildung 1: Mediane und Interquartilbereiche der relativen Häufigkeit der Hörbarkeit von Übertragungsfehlern der betrachteten Wellenfeldsynthese (vgl. Tabelle 1). Kreise symbolisieren Sprache, Quadrate einen Rauschpuls als Stimulus. Offene Symbole repräsentieren breitbandige, gefüllte unterhalb der Aliasfrequenz bandbegrenzte Wiedergabe.

Ergebnisse und Diskussion

Die Simulation des KHI zeigte erwartungsgemäß keine hörbaren Unterschiede zur originalen Primärquelle, weshalb diese Situation hier nicht dargestellt ist. Wird die Sekundärquellenanordnung jedoch diskretisiert, ist es perzeptiv möglich, die Synthese aufgrund der auftretenden Artefakte vom Original zu unterscheiden, falls der Testschall Spektralanteile oberhalb von f_a enthält. Unter dieser Voraussetzung ist die Diskretisierung für alle betrachteten Situationen hörbar. Abbildung 1 zeigt die Ergebnisse des Hörversuchs, dargestellt als Mediane und Interquartilbereiche der intra-individuellen Mittelwerte. Kreise symbolisieren den verwendeten Sprachschall, Quadrate den Rauschpuls. Offene Symbole kennzeichnen breitbandige Wiedergabe, gefüllte Bandbegrenzung unterhalb der Aliasfrequenz (bei diskreten Anordnungen). Wird ein kugelförmiges Array aus Monopolquellen betrachtet, entstehen Fehler bei tiefen Frequenzen, deren Erkennbarkeit im Median für beide Schalle und beide Primärquellenarten unterhalb der Hörbarkeitsschwelle liegt, vgl. Abbildung 1, Konfiguration 1 (ebene Welle) und 2 (Kugelwelle). Vergleichbare Werte ergeben sich für die entsprechenden diskreten Konfigurationen (3 und 4) bei Verwendung unterhalb der Aliasfrequenz bandbegrenzter Schalle (gefüllte Symbole). Breitbandige Schalle unterscheiden sich hingegen aufgrund der Aliasartefakte hörbar vom Original. Verwendet man das betrachtete endliche ebene Array, so entstehen Fehler bei tiefen Frequenzen, die bei Synthese ebener Wellen hörbar sein können (vgl. 5 und 6). Bei weiterer Reduzierung der Wellenfeldsynthese auf kreisförmige Arrays mit Monopol-Sekundärquellen entstehen im Allgemeinen Amplitudenfehler, die im betrachteten Fall für den Hörort korrigiert werden. Die Konfigurationen 7 und 8 zeigen die Ergebnisse für primäre ebene Wellenfronten (7) und eine (nicht korrigierte) primäre Kugelquelle (8). Ebene Wellen können im betrachteten Fall nicht vom Original unterschieden werden, während die bei Verwendung der Kugelquelle auftretenden Fehler deutlich hörbar sind. Wird dagegen die hier vorgeschlagene Korrektur angewandt, fällt die Hörbarkeit der Übertragungsfehler unter die Wahrnehmbarkeitsschwelle (die Verifikation der Korrektur wurde nur mehr mit Rauschen als Stimulus durchgeführt, das sich bei allen vorhergehenden Experimenten als kritischster Stimulus erwies, vgl. Konf. 9). Finite lineare Arrays verhalten sich vergleichbar zu kreisförmigen. Allerdings treten hier ähnlich wie bei finiten ebenen Arrays Fehler bei tiefen Frequenzen auf, die hörbar sein können (vgl. Konf. 10). Hier nicht im Detail vorgestellte Experimente zeigen, dass die Synthese bei einer Verschiebung des Hörortes weg vom Optimierungspunkt hier im Allgemeinen keine hörbaren Fehler erzeugt, falls die Verschiebung senkrecht zur Ausbreitungsrichtung der Primärwelle erfolgt und nicht mehr als 0.5m beträgt.

Diese Arbeit wurde von der Deutschen Forschungsgemeinschaft (DFG) im Rahmen des Projekts FA 140/4 gefördert.

Literatur

- [1] Fastl H., E. Zwicker: *Psychoacoustics. Facts and Models*. 3. Auflage, (Springer, Berlin, Heidelberg, 2007)
- [2] S. Spors, R. Rabenstein, J. Ahrens: The Theory of Wave Field Synthesis Revisited, 124th AES Convention (2008)
- [3] M. Zollner, E. Zwicker: *Elektroakustik*, 3. Auflage (Springer, Berlin 1993)