

Schallfeldsynthese: Stand der Technik und Perspektiven

Sascha Spors

Quality and Usability Lab, Deutsche Telekom Laboratories, Technische Universität Berlin, Germany.

Email: Sascha.Spors@telekom.de

Einleitung

Seit Anbeginn der Audioaufnahme und -wiedergabe strebt man neben der Übertragung des Musikmaterials auch die Übertragung der Raumakustik an. Die ersten Systeme zur Aufnahme und Wiedergabe, wie zum Beispiel der Phonograph, hatten nur einen Lautsprecher zur Wiedergabe. Damit lässt sich die Raumakustik nur sehr begrenzt abbilden. Man erkannte relativ früh, dass die räumlichen Aspekte einer akustischen Szene wesentlich besser durch mehrere Übertragungskanäle und Lautsprecher abgebildet werden können. Dies führte zur Entwicklung der Stereophonie, bei der zwei getrennte Kanäle genutzt werden. Die Stereophonie ermöglicht die Erzeugung gewisser räumlicher Aspekte zwischen den Lautsprechern. Sie stellt sicherlich das derzeit am weitesten verbreitete Wiedergabeverfahren dar. Um die Wiedergabe der Räumlichkeit weiter zu verbessern, wurden eine Reihe von verschiedenen Verfahren mit mehr als zwei Kanälen entwickelt. Diese werden im Allgemeinen als Surround-Verfahren bezeichnet. Angefangen von dem weit verbreiteten 5.1-Standard [1] mit 5 Lautsprechern und einem Subwoofer bis hin zu 22.2 [2] mit 22 Lautsprechern und zwei Subwoofern. Allerdings haben die Surround-Verfahren meist den entscheidenden Nachteil, dass der korrekte räumliche Eindruck nur innerhalb einer relativ kleinen Hörposition entsteht. Dieser Bereich wird meist als *'sweet spot'* bezeichnet. Aufgrund dieser und anderer Limitationen wurde unter anderem die Schallfeldsynthese entwickelt. Diese strebt die physikalische Synthese eines Schallfeldes durch ein Ensemble von Lautsprechern über einen ausgedehnten Zuhörerbereich an. Dieser Artikel gibt einen Überblick über die physikalischen Grundlagen der Schallfeldsynthese am Beispiel zweier konkreter Verfahren. Deren praktische Umsetzung und die sich daraus ergebenden Eigenschaften werden diskutiert. Des Weiteren werden aktuelle Trends in der Forschung vorgestellt und die Perspektiven der Schallfeldsynthese diskutiert.

Grundlagen der Schallfeldsynthese

In den letzten Jahrzehnten wurden diverse Verfahren der Schallfeldsynthese entwickelt, welche sich in zwei Klassen einordnen lassen. Bei der ersten Klasse werden eine Reihe von Kontrollpunkten in dem gewünschten Zuhörerbereich platziert. Die Ansteuerungssignale für die Lautsprecher werden nun so berechnet, dass sich das gewünschte Schallfeld an den Kontrollpunkten einstellt. Dazu wird das Problem meist als Matrixgleichung im Zeitbereich, Frequenzbereich oder einem räumlich trans-

formierten Bereich formuliert, welches dann durch numerische Methoden auf Basis der Minimierung des mittleren quadratischen Fehlers gelöst wird [3, 4, 5]. Nachdem es sich um ein inverses Problem handelt, ist dies in der Praxis oft schlecht konditioniert [6]. Dies gilt speziell für hohe Frequenzen. Der Nachteil dieser Ansätze ist die numerische Komplexität und Abhängigkeiten der Eigenschaften des synthetisierten Schallfeldes von der Anordnung der Kontrollpunkte und Lautsprecher.

Die zweite Klasse von Verfahren basiert auf einer analytischen Lösung des zugrundeliegenden physikalischen Problems. Bekannte Vertreter dieser Klasse sind die Wellenfeldsynthese ('wave field synthesis', WFS) [7], das Nahfeld-kompensierte Ambisonics höherer Ordnung ('near-field compensated higher-order Ambisonics', NFC-HOA) [8] und die Methode der spektralen Division ('spectral division method', SDM) [9]. Die Verfahren basieren auf der Annahme einer räumlich kontinuierlichen Verteilung von Sekundärquellen, welche in der Praxis durch eine räumlich diskrete Verteilung von Lautsprechern realisiert wird. Wesentliche Unterscheidungsmerkmale der verschiedenen Ansätze sind das Vorgehen zur Lösung des physikalischen Problems und die Geometrie der Lautsprecheranordnung. Die folgenden Abschnitte führen in die physikalischen Grundlagen der WFS und von NFC-HOA ein.

Wellenfeldsynthese

Die ursprüngliche Herleitung der WFS erfolgte auf Basis der Rayleigh Integrale für eine lineare Verteilung von Sekundärquellen [7]. Später wurde die theoretische Grundlage auf die Verwendung von gekrümmten Verteilungen erweitert [10]. Im Folgenden wird eine verallgemeinerte Theorie der WFS auf Basis des Kirchhoff-Helmholtz Integrals vorgestellt [11].

Das Kirchhoff-Helmholtz Integral [12] stellt die Lösung der homogenen Wellengleichung für inhomogene Randbedingungen dar. Eine kontinuierliche Verteilung von Sekundärquellen um den Zuhörerbereich V stellt in diesem Kontext eine inhomogene Randbedingung dar. Ziel der Schallfeldsynthese ist es, ein gewünschtes Schallfeld $S(\mathbf{x}, \omega)$ so gut wie möglich innerhalb von V zu synthetisieren. Der Schalldruck $P(\mathbf{x}, \omega)$ für $\mathbf{x} \in V$ ist eindeutig gegeben als

$$P(\mathbf{x}, \omega) = \oint_{\partial V} G(\mathbf{x} - \mathbf{x}_0, \omega) \frac{\partial}{\partial \mathbf{n}} P(\mathbf{x}_0, \omega) - S(\mathbf{x}_0, \omega) \frac{\partial}{\partial \mathbf{n}} G(\mathbf{x} - \mathbf{x}_0, \omega) dS_0, \quad (1)$$

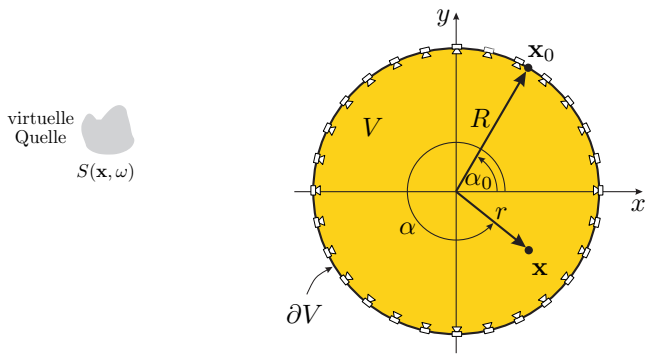


Abbildung 1: Schallfeldsynthese mit einer zirkulären Verteilung von Sekundärquellen.

wobei $G(\mathbf{x} - \mathbf{x}_0, \omega)$ die Greensche Funktion bezeichnet, $\partial/\partial\mathbf{n}$ den Gradienten in Richtung des Normalenvektors \mathbf{n} evaluiert am Punkt \mathbf{x}_0 und dS_0 ein geeignetes Oberflächen- bzw. Linienelement für die Integration. Die Greensche Funktion $G(\mathbf{x} - \mathbf{x}_0, \omega)$ ist die Lösung der inhomogenen Wellengleichung für einen raum-zeitlichen Dirac Impuls an der Position \mathbf{x}_0 . Sie kann daher als das Schallfeld eines akustischen Monopols interpretiert werden, der gerichtete Gradient der Greenschen Funktion als das Schallfeld eines Dipols. Strebt man nun die Synthese des Schallfeldes $S(\mathbf{x}, \omega)$ für $\mathbf{x} \in V$ an, so besagt Gl. (1), dass dies durch eine kontinuierliche Verteilung von Monopolen und Dipolen auf dem Rand ∂V von V realisiert werden kann, wobei die Monopole mit dem gerichteten Gradienten $\partial/\partial\mathbf{n}(S(\mathbf{x}_0, \omega))$ und die Dipole mit dem Schalldruck $S(\mathbf{x}_0, \omega)$ des gewünschten Schallfeldes auf dem Rand ∂V gewichtet werden müssen. Diese Gewichtung wird im Folgenden als Ansteuerungsfunktion bezeichnet. Die vorgestellte Interpretation von Gl. (1) für die Realisierung der Schallfeldsynthese stellt eine implizite Lösung des Syntheseproblems dar.

Zur Illustration des Prinzips dient die Synthese einer monofrequenten ebenen Welle in einem kreisförmigen Bereich V mit Radius R . Abbildung 1 zeigt die zugrundeliegende Geometrie. Zur besseren Vergleichbarkeit der verschiedenen Ansätze wird diese für die meisten Beispiele in diesem Artikel verwendet. Abbildung 2 zeigt das synthetisierte Schallfeld für eine kontinuierliche Verteilung von Monopolen und Dipolen. Das gewünschte Schallfeld wird innerhalb von V ohne Artefakte synthetisiert, außerhalb ist kein Schallfeld vorhanden. Für die praktische Realisierung der Schallfeldsynthese ist jedoch eine Synthese durch Monopole alleine erstrebenswert, da die Eigenschaften dieser gut durch Lautsprecher genähert werden können. Bei der WFS erfolgt dies durch eine Näherung des Kirchhoff-Helmholtz Integrals (1). Durch eine Auswahl der aktiven Sekundärquellen, die vom gewünschten Schallfeld abhängt, und eine Beschränkung auf konvexe Ränder ∂V können die Dipolquellen für die Synthese vernachlässigt werden. Die Approximation des gewünschten Schallfeldes ist gegeben durch [11]

$$S(\mathbf{x}, \omega) \approx \oint_{\partial V} \underbrace{2a(\mathbf{x}_0) \frac{\partial}{\partial \mathbf{n}} S(\mathbf{x}_0, \omega)}_{D_{\text{WFS}}(\mathbf{x}_0, \omega)} G(\mathbf{x} - \mathbf{x}_0, \omega) dS_0, \quad (2)$$

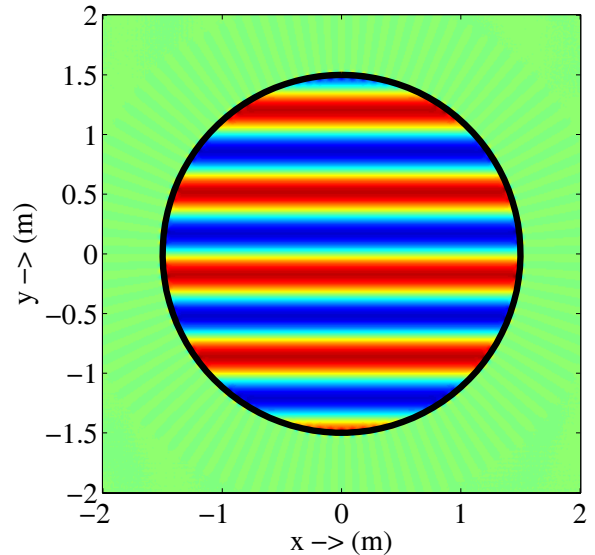


Abbildung 2: Zweidimensionale Synthese einer monofrequenten ebenen Welle ($f_{\text{pw}} = 500$ Hz, $\alpha_{\text{pw}} = 270^\circ$) auf Basis der Kirchhoff-Helmholtz Integrals durch eine kreisförmige Verteilung von Monopolen und Dipolen.

für $\mathbf{x} \in V$. Die Fensterfunktion $a(\mathbf{x}_0)$ wählt die aktiven Sekundärquellen für ein gewünschtes Schallfeld $S(\mathbf{x}, \omega)$ aus und $D_{\text{WFS}}(\mathbf{x}_0, \omega)$ stellt die Ansteuerungsfunktion der Sekundärquellen bei der WFS dar. Die Fensterfunktion kann für bestimmte Schallfelder, wie zum Beispiel eine Punktquelle oder ebene Welle, geschlossen angegeben werden [13]. Abbildung 3a zeigt die Synthese einer ebenen Welle durch die Wellenfeldsynthese. Im Vergleich mit Abbildung 2 sind Artefakte im synthetisierten Schallfeld sichtbar, welche eine Konsequenz der verwendeten Näherungen sind. Es kann gezeigt werden dass die Genauigkeit der Näherung (2) mit steigender Frequenz und/oder größerem Abstand zu den Sekundärquellen zunimmt [14]. Beim aktuellen Stand der Forschung scheinen die Artefakte aber keine perzeptive Relevanz zu besitzen. Weiterhin verschwindet das Schallfeld außerhalb von V nicht, so wie das bei der Synthese mittels Monopolen und Dipolen der Fall ist (siehe Abb. 2).

Nahfeld-kompensiertes Ambisonics höherer Ordnung

Im Gegensatz zur WFS strebt man bei NFC-HOA eine explizite Lösung des physikalischen Problems an, um die Ansteuerungsfunktion für die Sekundärquellen zu gewinnen. In aktuellen Formulierungen geht man dabei direkt von einer kontinuierlichen Verteilung von Monopolen aus [15, 16, 17]. Das synthetisierte Schallfeld ist in diesem Falle gegeben als

$$P(\mathbf{x}, \omega) = \oint_{\partial V} D_{\text{HOA}}(\mathbf{x}_0, \omega) G(\mathbf{x} - \mathbf{x}_0, \omega) dS_0. \quad (3)$$

Die Integralgleichung (3) kann als Fredholm Operator mit einem Fredholm-Index von Null interpretiert werden [18, 19, 20]. Dieser ist somit kompakt und es existiert eine Reihendarstellung von (3) in den orthogonalen Basisfunktionen des Kerns $G(\mathbf{x} - \mathbf{x}_0, \omega)$. Wird das

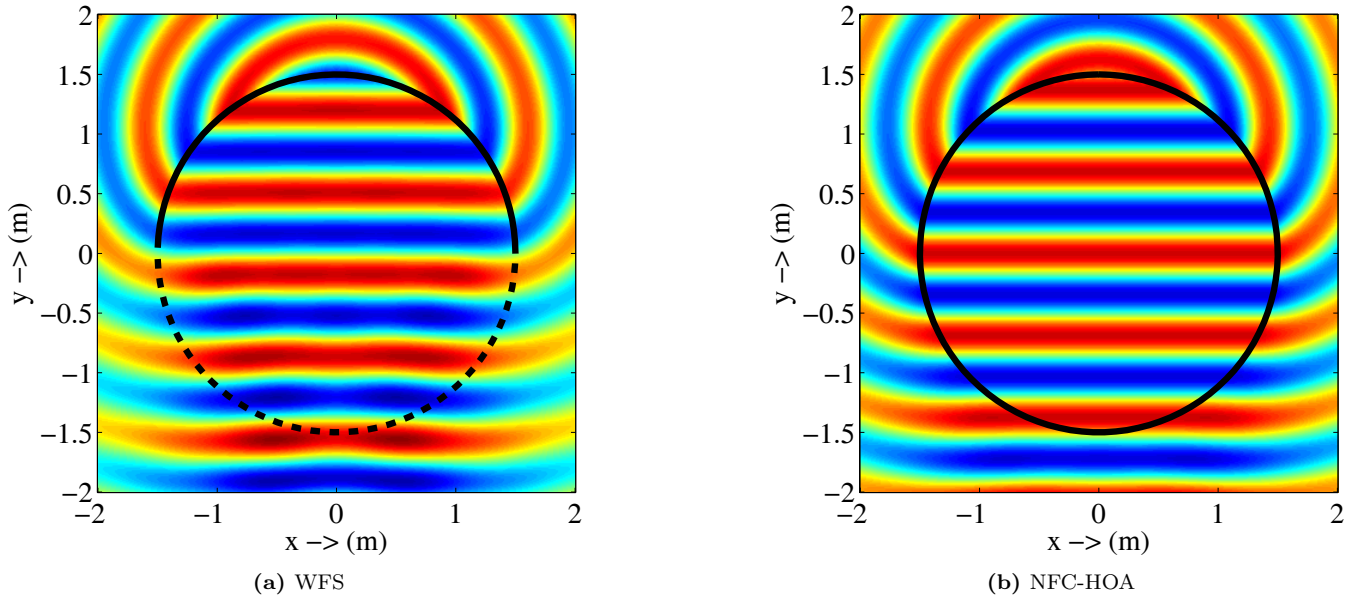


Abbildung 3: Zweidimensionale Synthese einer monofrequenten ebene Welle ($f_{pw} = 500$ Hz, $\alpha_{pw} = 270^\circ$) durch eine kreisförmige Verteilung von Monopolen mittels WFS und NFC-HOA. Die inaktiven Bereiche der Verteilung von Sekundärquellen sind gestrichelt dargestellt.

gewünschte Schallfeld auch bezüglich dieser Basisfunktionen ausgedrückt, entsteht aus (3) eine algebraische Gleichung, die durch Koeffizientenvergleich gelöst werden kann. Im Allgemeinen hängen die benötigten Basisfunktionen von der Geometrie des Randes ∂V ab. Geschlossene Lösungen existieren nur für einfache Geometrien wie Kreise, Kugeloberflächen, Linien oder Flächen.

Wie schon bei der WFS wird das Prinzip der expliziten Lösung am Beispiel eines kreisförmigen Randes ∂V aufgezeigt (siehe Abb. 1). Für diesen Fall wird (3) zu

$$P(\mathbf{x}, \omega) = \int_0^{2\pi} D_{\text{HOA}}(\alpha_0, R, \omega) G(\mathbf{x} - \mathbf{x}_0, \omega) R d\alpha_0, \quad (4)$$

wobei $\mathbf{x} = r[\cos \alpha \sin \alpha]^T$ und $\mathbf{x}_0 = R[\cos \alpha_0 \sin \alpha_0]^T$ durch Polarkoordinaten dargestellt werden. Eine geeignete Basis ist durch die Exponentialfunktionen gegeben, aufgrund der Periodizität des Winkels können die Schallfelder als Fourierreihe dargestellt werden:

$$P(\mathbf{x}, \omega) = \sum_{\nu=-\infty}^{\infty} \hat{P}(\nu, r, \omega) e^{j\nu\alpha}. \quad (5)$$

Führt man dies für alle Größen in (4) durch und nutzt die Orthogonalität der Exponentialfunktionen aus, so erhält man eine algebraische Gleichung welche durch Koeffizientenvergleich gelöst werden kann. Die Ansteuerungsfunktion ist dann gegeben als [15]

$$D_{\text{HOA}}(\alpha, R, \omega) = \sum_{\nu=-\infty}^{\infty} \frac{\hat{S}(\nu, r, \omega)}{\hat{G}(\nu, r, \omega)} e^{j\nu\alpha}, \quad (6)$$

falls $\hat{G}(\nu, r, \omega)$ keine Nullstellen aufweist. Für akustische Monopole und andere einfache Schallfelder können die Fourierkoeffizienten $\hat{G}(\nu, r, \omega)$ und $\hat{S}(\nu, r, \omega)$ geschlossen angegeben werden. In der Praxis ist es nicht möglich

die unendliche Reihe (6) auszuwerten, daher wird diese meist nach einer endlichen Anzahl von Gliedern abgebrochen. Es wird später gezeigt, dass dies im Zusammenhang mit der Abtastung sinnvoll ist. Abbildung 3b zeigt die Synthese einer monofrequenten ebene Welle durch NFC-HOA. Im Vergleich zur WFS in Abbildung 3a fällt die exakte Synthese des gewünschten Schallfeldes auf.

Neben der expliziten Lösung der Synthesegleichung (3) für Sekundärquellenverteilungen auf Kreisen oder Kugeloberflächen, wurden auch Verteilungen auf Linien und Flächen betrachtet [9, 21]. Die Synthesegleichung wird in diesem Falle mit einer räumlichen Fouriertransformation in eine skalare Gleichung umgewandelt, die dann im räumlichen Frequenzbereich gelöst werden kann. Dieses Verfahren wurde 'Spectral Division Method (SDM)' getauft.

2.5-dimensionale Synthese

In der praktischen Umsetzung der Schallfeldsynthese durch WFS oder NFC-HOA ist es meist nicht möglich den Zuhörerbereich mit Sekundärquellen zu umhüllen. In der Praxis wird meist nur die Synthese in einer Ebene realisiert [22]. Idealerweise befinden sich die Ohren des Zuhörers dann in dieser Ebene. Ein solches Szenario wird als 2.5-dimensionale Synthese bezeichnet.

Selbst wenn die 2.5-dimensionale Synthese explizit bei der Herleitung der Ansteuerungsfunktionen berücksichtigt wird, können unerwünschte Artefakte im synthetisierten Schallfeld nicht verhindert werden. Die auftretenden Artefakte sind ein inkorrekt Amplitudenverlauf und geringfügige Phasenfehler im synthetisierten Schallfeld [23, 11, 15].

Abbildung 4 zeigt das synthetisierte Schallfeld für das Szenario aus Abbildung 3, jedoch für die 2.5-dimensionale Synthese. Die auftretenden Amplitudenfehler sind klar zu

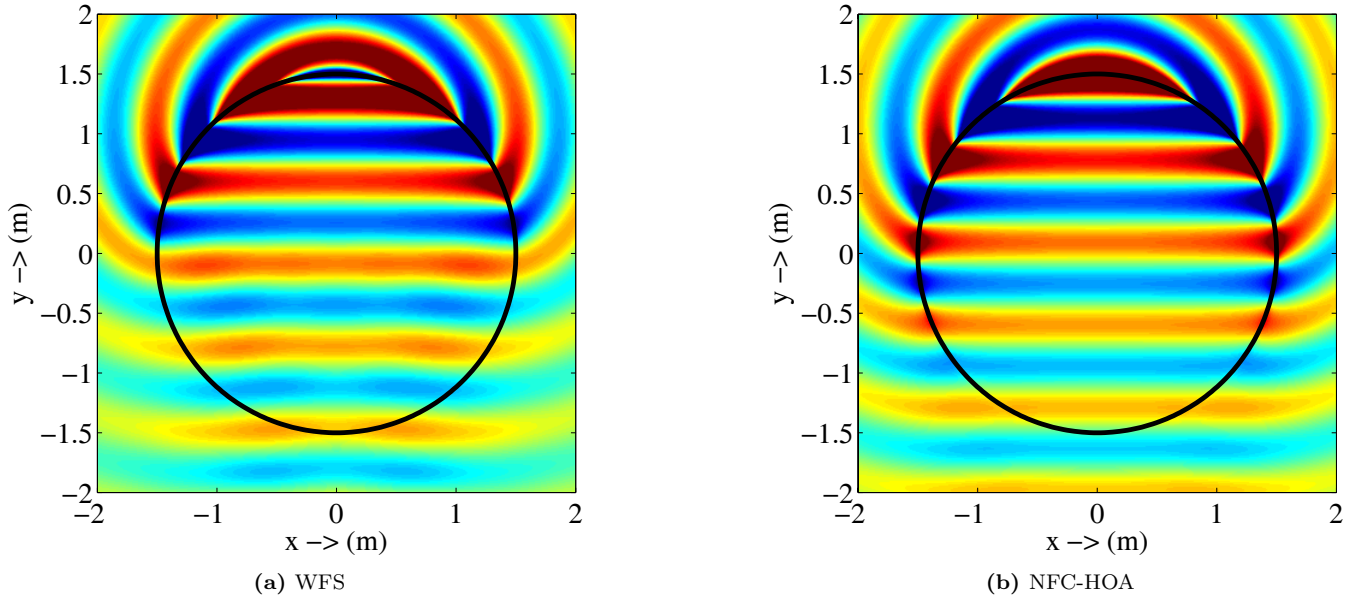


Abbildung 4: 2.5-dimensionale Synthese einer monofrequenten ebene Welle ($f_{pw} = 500$ Hz, $\alpha_{pw} = 270^\circ$).

erkennen.

Die Artefakte können durch Betrachtung der Greenschen Funktion $G(\mathbf{x} - \mathbf{x}_0, \omega)$ erklärt werden. Die explizite Form der Greenschen Funktion hängt von der Dimensionalität des betrachteten Problems ab. Die Synthese in einer Ebene stellt in diesem Kontext ein zweidimensionales Problem dar. Die zweidimensionale Greensche Funktion kann als das Feld einer Linienquelle interpretiert werden [12]. Für die Synthese in einer Ebene müssten daher Linienquellen als Sekundärquellen zum Einsatz kommen. Tatsächlich war dies bei Abbildung 3 der Fall, weshalb hier keine Artefakte auftreten. In der Praxis werden Lautsprecher genutzt, die das Feld einer Punktquelle annähern. Linienquellen haben gegenüber Punktquellen, unter anderem, einen geringeren Abfall der Amplitude über den Abstand von der Quelle.

Räumliche Abtastung

In den vorherigen Abschnitten wurden die physikalischen Grundlagen der WFS und vom NFC-HOA vorgestellt. Beide Verfahren sind mit einer kontinuierlichen Verteilung von Sekundärquellen in der Lage, ein gewünschtes Schallfeld ohne oder mit nur geringen Artefakten zu synthetisieren. Daher sind die Ergebnisse beider Verfahren vergleichbar. In der Praxis wird die kontinuierliche Verteilung der Sekundärquellen durch eine endliche Anzahl von Lautsprechern realisiert. Dies stellt eine räumliche Abtastung der Sekundärquellenverteilung dar. Es wurde gezeigt das WFS und NFC-HOA wesentliche Unterschiede in Bezug auf die räumliche Abtastung aufweisen [24, 25]. Die Untersuchung der Effekte der räumlichen Abtastung ist derzeit nur für bestimmte Geometrien der Sekundärquellenverteilung durchgeführt worden. So zum Beispiel für lineare Verteilungen bei der WFS und SDM [26, 27, 28, 9], für zirkuläre Verteilungen bei der WFS und NFC-HOA [27, 20] und für sphärische Verteilungen bei NFC-HOA [20].

Im Folgenden wird die Abtastung für die betrachtete zirkuläre Verteilung von Sekundärquellen vorgestellt.

Die Abtastung der kontinuierlichen Verteilung von Sekundärquellen kann durch die räumliche Abtastung der Ansteuerungsfunktion $D(\alpha, R, \omega)$ modelliert werden. Abbildungen 1 und 5 illustrieren dies für eine äquiangulare Abtastung. Die Abtastung führt zu Wiederholungen im transformierten Bereich

$$\dot{D}_S(\nu, R, \omega) = \sum_{\eta=-\infty}^{\infty} \dot{D}(\nu + \eta N, R, \omega), \quad (7)$$

wobei $\dot{D}_S(\nu, R, \omega)$ die Fourierkoeffizienten der abgetasteten Ansteuerungsfunktion bezeichnet und N die Anzahl der Abtastpunkte auf dem Kreis. Die synthetisierte Schallfeld $P_S(\mathbf{x}, \omega)$ für eine diskrete Verteilung von Sekundärquellen ist gegeben als (siehe Abbildung 5)

$$P_S(\mathbf{x}, \omega) = \sum_{\nu=-\infty}^{\infty} \dot{D}_S(\nu, R, \omega) \dot{G}(\nu, r, \omega) e^{j\nu\alpha}. \quad (8)$$

Die Eigenschaften des synthetisierten Schallfeldes $P_S(\mathbf{x}, \omega)$ hängen daher von den Eigenschaften der raum-zeitlichen Spektren der (abgetasteten) Ansteuerungsfunktion $\dot{D}(\nu, R, \omega)$ und der Sekundärquellen $\dot{G}(\nu, r, \omega)$ ab. Es konnte gezeigt werden [24], dass das Spektrum der Ansteuerungsfunktion der WFS $\dot{D}_{WFS}(\nu, R, \omega)$ nicht bandbegrenzt ist. In der abgetasteten Ansteuerungsfunktion tritt daher Aliasing aufgrund der überlappenden spektralen Wiederholungen auf. Wie bereits vorher im Abschnitt zu NFC-HOA diskutiert, wird die Fourierreihe der Ansteuerungsfunktion von NFC-HOA $\dot{D}_{HOA}(\nu, R, \omega)$ typischerweise bei $\pm N/2$ Gliedern abgebrochen. Dies führt zu einer räumlichen Bandbegrenzung, welche räumliches Aliasing bei der Abtastung verhindert.

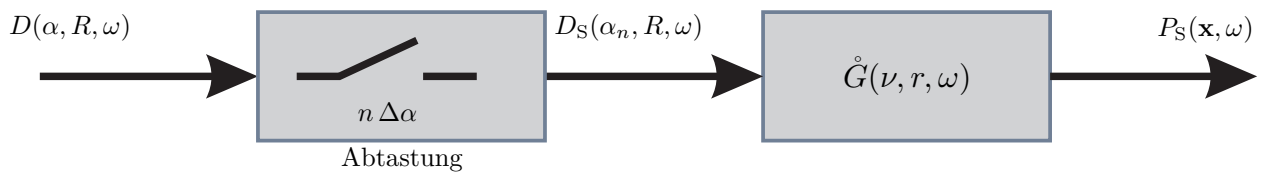


Abbildung 5: Systemtheoretisches Modell der räumlichen Abtastung einer kreisförmigen kontinuierlichen Sekundärquellenverteilung.

Die spektralen Wiederholungen der Ansteuerungsfunktion werden mit dem raum-zeitlichen Spektrum der Sekundärquelle $\dot{G}(\nu, r, \omega)$ gewichtet (siehe Abb. 5). Damit diese Wiederholungen nicht im synthetisierten Schallfeld auftreten, müsste das Spektrum der Sekundärquelle räumlich bandbegrenzt sein, was nicht der Fall ist.

Die auftretenden Artefakte werden im Folgenden am Beispiel einer zirkularen Verteilung von 56 Punktquellen aufgezeigt. Die Geometrie entspricht den vorhergegangenen Beispielen (siehe Abb. 1). Abbildungen 6a und 6b zeigen das synthetisierte Schallfeld für eine monofrequente ebene Welle mit einer Frequenz von $f_{pw} = 500$ Hz, Abbildungen 6c und 6d bei $f_{pw} = 3$ kHz. Bei tiefen Frequenzen zeigt die räumliche Abtastung bei WFS und NFC-HOA keinen Einfluss auf das synthetisierte Schallfeld. Erhöht man allerdings die Frequenz, so sind deutliche Unterschiede zwischen beiden Verfahren sichtbar. Während hier bei NFC-HOA im Zentrum eine Zone ohne sichtbare Artefakte auftritt, sind diese bei der WFS über einen großen Bereich verteilt. Die Zone ohne sichtbare Artefakte bei NFC-HOA wird mit zunehmender Frequenz immer kleiner. Um den perzeptiven Einfluß der Artefakte der räumlichen Abtastung bewerten zu können, ist es sinnvoll die Synthese einer breitbandigen ebenen Welle zu betrachten. Das synthetisierte Schallfeld für einen festen Zeitpunkt t_0 ist in den Abbildungen 6e und 6f dargestellt. Auch hier sind wieder deutliche Unterschiede zu beobachten. Bei NFC-HOA bildet sich in einem kleinen Bereich um das Zentrum eine Zone ohne sichtbare Artefakte aus. Außerhalb dieses Bereichs sind starke Artefakte zu beobachten. Neben einer unerwünschten Krümmung der Wellenfront, tritt hier weiterhin eine zeitliche Aufspaltung der Wellenfront auf. Bei der WFS wird die erste Wellenfront gut synthetisiert, danach treten Artefakte auf. Die Abbildungen 6e und 6f können als die raum-zeitliche Impulsantwort der jeweiligen Systeme für die Synthese einer ebenen Welle interpretiert werden.

Vergleich von WFS und NFC-HOA

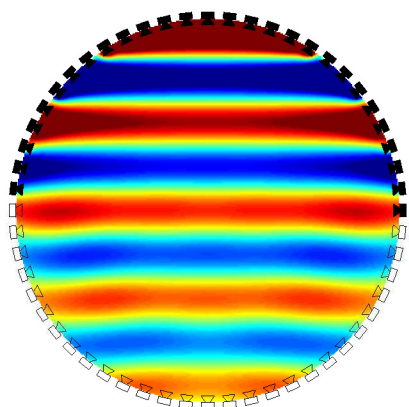
Praktische Umsetzungen der Schallfeldsynthese basieren derzeit meist auf der WFS oder dem NFC-HOA. Daher ist ein Vergleich der Eigenschaften dieser beiden Verfahren von speziellem Interesse. Neben den objektiven Eigenschaften in Bezug auf das synthetisierte Schallfeld sind die perzeptiven Eigenschaften für ein Audiowiedergabesystem besonders wichtig.

Während WFS auf einer impliziten Lösung der approxi-

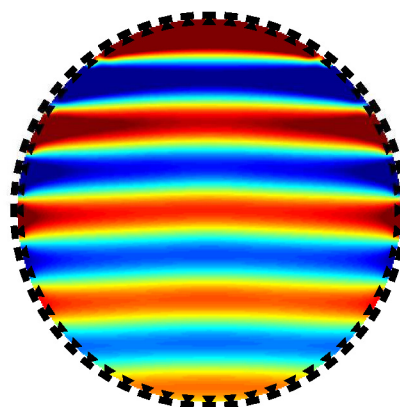
mierten Synthesegleichung beruht, wird bei NFC-HOA eine explizite Lösung verwendet. In den vorhergegangenen Abschnitten wurde gezeigt, dass sich die grundlegenden Eigenschaften des synthetisierten Schallfeldes bei der WFS und dem NFC-HOA, unter der Annahme einer kontinuierlichen Verteilung von Sekundärquellen, nicht wesentlich unterscheiden. Beide Verfahren ermöglichen es, ein nahezu beliebiges Schallfeld innerhalb der Sekundärquellenverteilung zu synthetisieren. Für die Anwendung in der Audiowiedergabe gibt es derzeit keine Anzeichen, dass der Umstand das die WFS-Lösung eine Approximation darstellt zu Beeinträchtigungen führt. Beide Verfahren zeigen auch ein nahezu gleiches Verhalten bei der 2.5-dimensionalen Synthese.

Wesentliche Unterschiede zwischen beiden Verfahren bestehen allerdings in Bezug auf die räumliche Abtastung der Verteilung von Sekundärquellen. Für die Synthese monofrequenter ebener Wellen sind die Artefakte der räumlichen Abtastung bei der WFS im gesamten Zuhörerbereich verteilt. Bei NFC-HOA stellt sich im Zentrum ein kreisförmiger, nahezu artefaktfreier Bereich ein. Dieser Bereich wird mit steigender Frequenz kleiner. Eine Analyse der breitbandigen raum-zeitlichen Impulsantworten beider Systeme (siehe Abb. 6e und 6f) liefert im Grundsatz die gleichen Ergebnisse. Jedoch sieht man bei NFC-HOA bei hohen Frequenzen eine deutliche Krümmung der Wellenfronten außerhalb des Zentrums, wohingegen bei der WFS die ersten Wellenfront über den gesamten Zuhörerbereich nahezu korrekt synthetisiert wird. Eine tiefere Analyse der Eigenschaften von WFS und NFC-HOA im Zeitbereich wurde in [25] veröffentlicht.

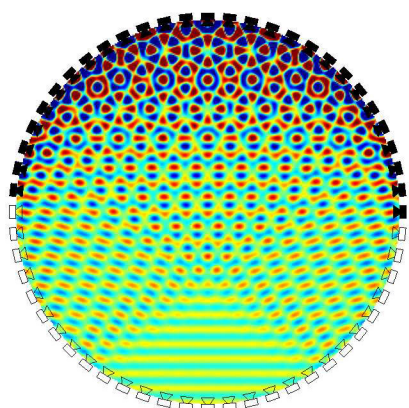
Die gezeigten Beispiele entsprechen in ihren Parametern aktuellen Realisierungen. Es kann daher gefolgert werden, dass derzeit bei voller Audiobandbreite von 20 kHz bei vertretbarem Aufwand keine physikalisch korrekte Synthese eines Schallfeldes in einem ausgedehnten Bereich möglich ist. Die Eigenschaften von WFS und NFC-HOA hängen daher wesentlich von der Wahrnehmung synthetischer Schallfelder ab. Für beide Verfahren wurden in der Vergangenheit eine Reihe von Untersuchungen durchgeführt, um die perzeptiven Eigenschaften zu evaluieren. Für die WFS wurden im Wesentlichen Untersuchungen zur Lokalisation virtueller Punktquellen [26, 29, 30, 31, 32] und der Klangfarbe durchgeführt [33]. Meist nur für eine oder wenige Zuhörerpositionen. Es konnte gezeigt werden das die Lokalisation sehr genau ist, aber Veränderungen in der Klangfarbe der virtuellen Quelle auftreten. Für NFC-HOA existieren nur wenige Studien. In [34] wurde die Lokalisation für einen Zuhörer im Zen-



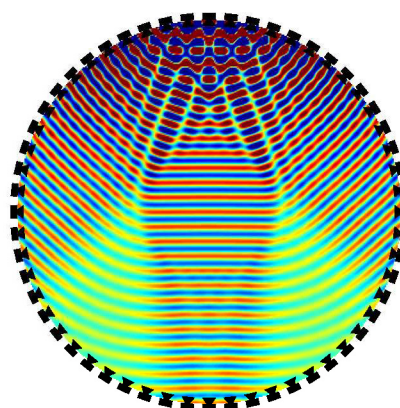
(a) Monochromatische ebene Welle synthetisiert durch WFS bei $f_{pw} = 500$ Hz



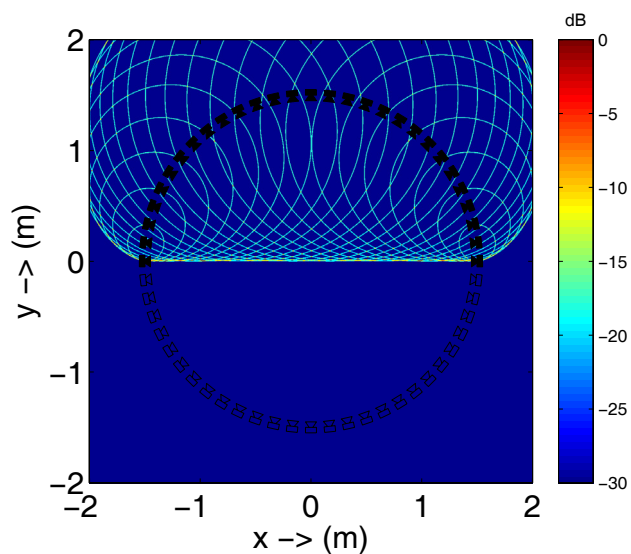
(b) Monochromatische ebene Welle synthetisiert durch NFC-HOA bei $f_{pw} = 500$ Hz



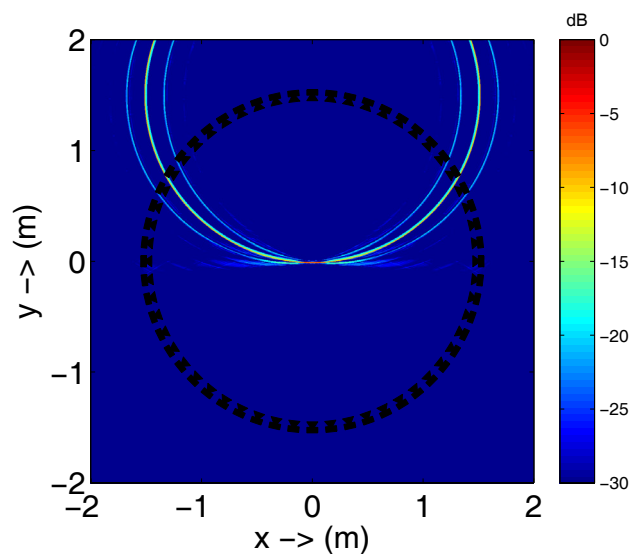
(c) Monochromatische ebene Welle synthetisiert durch WFS bei $f_{pw} = 3$ kHz



(d) Monochromatische ebene Welle synthetisiert durch NFC-HOA bei $f_{pw} = 3$ kHz



(e) Breitbandige ebene Welle mit WFS



(f) Breitbandige ebene Welle mit NFC-HOA

Abbildung 6: 2.5-dimensionale Synthese einer monochromatischen/breitbandigen ebene Welle durch ortsdiskrete Punktquellen ($N = 56$, $\alpha_{pw} = 270^\circ$). Die aktiven Sekundärquellen sind durch ausgefüllte Lautsprechersymbole gekennzeichnet.

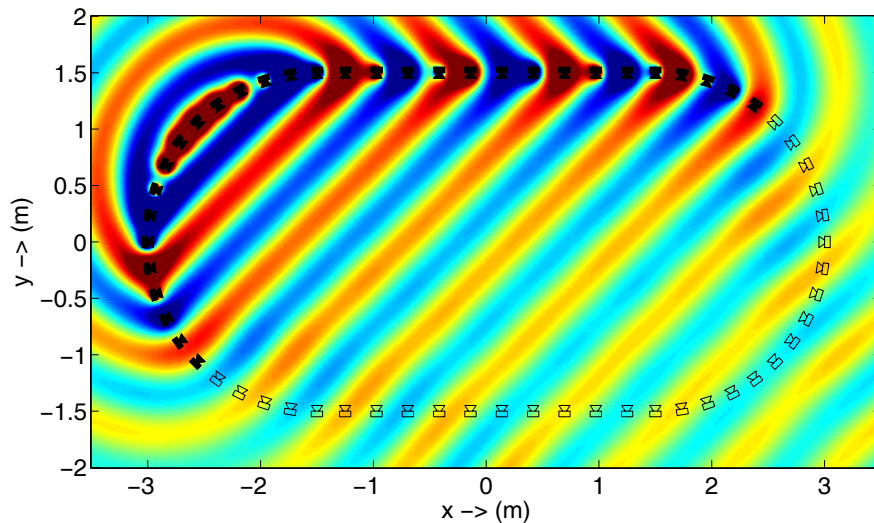


Abbildung 7: 2.5-dimensionale Synthese einer monofrequenten ebenen Welle mit einem ovalen Lautsprecherarray ($f_{pw} = 500$ Hz, $N = 62$, $\alpha_{pw} = 315^\circ$). Die aktiven Sekundärquellen sind ausgefüllt.

trum in Abhängigkeit von der Anzahl der verwendeten Lautsprecher evaluiert.

Gegenüber NFC-HOA bietet die WFS in der praktischen Anwendung einige Vorteile. So erlaubt sie die Nutzung von beliebig geformten konvexen Verteilungen von Sekundärquellen, die Ansteuerungsfunktion hängt nur von räumlich lokalen Größen ab und die Berechnung der Ansteuerungsfunktionen kann sehr effizient realisiert werden. Der erste Punkt ist in Abbildung 7 illustriert, wo die Synthese einer ebenen Welle mit einem ovalen Lautsprecherarray gezeigt ist. Derzeit ist für eine solche Geometrie keine NFC-HOA Lösung verfügbar.

Beschreibung und Aufnahme von virtuellen Szenen

Bisher wurde davon ausgegangen, dass das Schallfeld $S(\mathbf{x}, \omega)$ der zu synthetisierenden akustischen Szene bekannt ist. Es können zwei grundsätzlich verschiedene Ansätze zur Beschreibung der virtuellen Szene unterschieden werden: (1) die objektorientierte oder (2) die datenbasierte Beschreibung [35]. Im ersten Fall wird die Szene durch einfache akustische Objekte und deren Parameter beschrieben. Als Objekte kommen hier häufig virtuelle Punktquellen und/oder ebene Wellen zum Einsatz. Im zweiten Fall wird die raum-zeitliche Struktur einer Szene durch Verfahren der Schallfeldanalyse aufgenommen. Dies geschieht in der Praxis meist durch zirkuläre [36] oder sphärische Mikrofonarrays. Falls, was in der Praxis meist der Fall ist, die Geometrie des Lautsprecherarrays von der Geometrie des Mikrofonarrays abweicht, werden Verfahren der Schallfeldextrapolation genutzt, um das Feld vom Mikrofonarray zu den Lautsprecherpositionen zu extrapolieren.

Die objektorientierte Beschreibung virtueller Szenen bietet eine Reihe von Vorteilen. So ist die Beschreibung im Prinzip unabhängig vom verwendeten Verfahren der Schallfeldsynthese und der Geometrie des Lautsprecherarrays, die Szene kann leicht (auch nachträglich) be-

arbeitet werden und die Berechnung der Ansteuerungssignale kann meist sehr effizient für einfache Quellenmodelle durchgeführt werden. Dies wird dann als modellbasiertes Rendering bezeichnet [35].

Der SoundScape Renderer (SSR) [37] ist ein flexibles Werkzeug für modellbasiertes Rendering von objektorientierten Szenen durch verschiedene Syntheseverfahren. Neben der WFS und dem HOA, wurden auch 'Vectorbase Amplitude Panning (VBAP)' [38] und Verfahren zur binauralen Synthese implementiert. Die Beschreibung der virtuellen Szene erfolgt auf Basis des 'Audio Scene Description Formats (ASDF)' [35]. Der SSR ist unter <http://www.tu-berlin.de/?id=ssr> als quelloffene Software erhältlich.

Aktuelle Arbeiten

In den vorherigen Abschnitten wurden die grundlegenden Funktionsweisen und Eigenschaften von WFS und NFC-HOA vorgestellt. Für beide Verfahren wurden eine Vielzahl von Erweiterungen entwickelt. Einige werden im Folgenden kurz vorgestellt.

Praktisch realisierbare Lautsprecher haben meist nicht die bei der Herleitung der Verfahren angenommene Charakteristik einer Punktquelle. Es wurden eine Reihe von Verfahren entwickelt um die Synthese für Lautsprecher mit bekanntem Abstrahlverhalten zu optimieren [39, 40, 41]. Als Verallgemeinerung kann man die Kompensation des Einflusses des Wiedergaberaumes betrachten. Auch hier wurde eine Reihe von Verfahren vorgeschlagen [42, 43].

Für das modellbasierte Rendering sind virtuelle Quellen mit komplexer Richtcharakteristik entwickelt worden [44, 45, 46], sowie ausgedehnte virtuelle Quellen [47]. Weiterhin wurde die Synthese von virtuellen Quellen die sich mit hoher Geschwindigkeit bewegen optimiert [48, 49]. Ein weiterer Fokus von Forschungsarbeiten ist die Optimierung der Artefakte der räumlichen Diskretisierung der Sekundärquellen [50, 51]. Da bei vielen Anwendungen oft eine Synthese in einem kleinen Bereich ausrei-

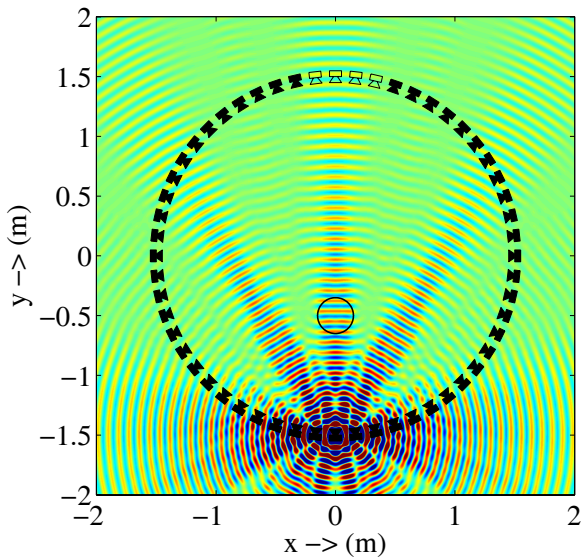


Abbildung 8: Synthese eines Schallfeldes durch lokale WFS innerhalb des durch den Kreis gekennzeichneten lokalen Zuhörerbereichs ($N = 56$, $f_{pw} = 4$ kHz, $\alpha_{pw} = 90^\circ$).

chend ist, wurden eine Reihe von Verfahren zur lokalen Schallfeldsynthese vorgeschlagen [52, 53, 54]. Diese bieten eine erhöhte Genauigkeit der Synthese in einem lokalen Zuhörerbereich auf Kosten stärkerer Artefakte außerhalb dieses Bereichs. Abbildung 8 zeigt ein Beispiel des in [53] vorgestellten Verfahrens. In dem durch die Kreislinie gekennzeichneten lokalen Zuhörerbereich wird eine monochromatische ebene Welle mit einer Frequenz von $f_{pw} = 4$ kHz ohne sichtbare Artefakte synthetisiert. In Abbildung 6c waren bei der Nutzung der konventionellen WFS bei $f_{pw} = 3$ kHz schon erhebliche Artefakte zu beobachten.

Die vorgestellten Verfahren der Schallfeldsynthese können auch nutzbringend für die effiziente Extrapolation von gemessenen Außenohrübertragungsfunktionen genutzt werden [55, 56, 57].

Zusammenfassung und Perspektiven

Der vorliegende Artikel gibt einen Überblick über zwei der bekanntesten Verfahren der Schallfeldsynthese, WFS und NFC-HOA. Die physikalischen Grundlagen beider Verfahren sowie die räumliche Diskretisierung der Sekundärquellenverteilung sind bei aktuellem Stand der Forschung bereits gut verstanden. Die Artefakte der räumlichen Diskretisierung wurden als der wesentliche Unterschied zwischen den mit WFS und NFC-HOA synthetisierten Schallfeldern identifiziert. Eine Grundannahme der Schallfeldsynthese ist, dass eine im physikalischen Sinne perfekte Synthese eines Schallfeldes einen vom Original nicht unterscheidbaren Höreindruck erzeugt. Allerdings ist eine artefaktfreie Synthese in einem ausgedehnten Bereich über die gesamte Audiobandbreite derzeit technisch mit vertretbarem Aufwand nicht umsetzbar. Die Eigenschaften aktueller Implementierungen beruhen daher wesentlich auf der menschlichen Wahrnehmung synthetischer Schallfelder, ohne dass diese derzeit

explizit durch die Verfahren berücksichtigt werden. Es liegt die Vermutung nahe, dass beide Verfahren durch das bessere Verständnis und die Ausnutzung der relevanten Wahrnehmungsmechanismen verbessert werden können. Synthetische Schallfelder weisen einige Eigenschaften auf, die so bei natürlichen Schallfeldern nicht auftreten. Daher sind bekannte Ergebnisse aus der Psychoakustik zum Teil nur eingeschränkt auf synthetische Schallfelder übertragbar.

Erste psychoakustisch motivierten Übertragungs- und Wiedergabeverfahren wurden bereits entwickelt. So zum Beispiel 'Directional Audio Coding (DirAc)' [58] und 'Binaural Cue Coding (BCC)' [59]. Allerdings können auch diese noch nicht die hochgesteckten Ziele der Schallfeldsynthese, die Wiedergabe in einem ausgedehnten Bereich, vollständig erfüllen.

Die Vortragsfolien zu diesem Artikel und weiteres Zusatzmaterial sind unter <http://audio.qu.tu-berlin.de/?p=599> erhältlich.

Danksagungen

Die in diesem Artikel vorgestellten Beiträge zur Schallfeldsynthese wären ohne die Mithilfe von vielen Kollegen nicht möglich gewesen. Ich möchte mich ganz besonders bei meinen Kollegen aus der Arbeitsgruppe 'Audio Technology' an den Deutschen Telekom Laboratories für die sehr freundschaftliche und produktive Zusammenarbeit der letzten Jahre bedanken. Besonderer Dank gilt dabei Alexander Raake, Jens Ahrens, Matthias Geier, Hagen Wierstorf und Karim Helwani. Ein Teil der Arbeiten entstand in Zusammenarbeit mit externen Partnern. Auch diesen möchte ich für die offene fachliche Diskussion und Zusammenarbeit danken. Besonderer Dank gilt hier Alexander Lindau und Stefan Weinzierl vom Fachbereich Audiokommunikation der Technischen Universität Berlin, sowie Benjamin Bernschütz und Christoph Pörschmann an der Fachhochschule Köln.

Literatur

- [1] F. Rumsey. *Spatial Audio*. Focal Press, 2001.
- [2] K. Hamasaki, K. Hiyama, and R. Okumura. The 22.2 multichannel sound system and its application. In *118th AES Convention*, Barcelona, Spain, May 2005. Audio Engineering Society (AES).
- [3] O. Kirkeby and P.A. Nelson. Reproduction of plane wave sound fields. *Journal of the Acoustic Society of America*, 94(5):2992–3000, Nov. 1993.
- [4] O. Kirkeby and P.A. Nelson. Digital filter design for inversion problems in sound reproduction. *Journal of the Audio Engineering Society*, 47:583–595, July/August 1999.
- [5] J. Hannemann and K.D. Donohue. Virtual sound source rendering using a multipole-expansion and method-of-moments approach. *Journal of the Audio Engineering Society (JAES)*, 56(6):473–481, June 2008.
- [6] F.M. Fazi and P.A. Nelson. The ill-conditioning problem in sound field reconstruction. In *123rd AES Convention*, New York, USA, 2007. Audio Engineering Society (AES).
- [7] A.J. Berkhout, D. de Vries, and P. Vogel. Acoustic control by wave field synthesis. *Journal of the Acoustic Society of America*, 93(5):2764–2778, May 1993.

- [8] J. Daniel. *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia*. PhD thesis, Université Paris 6, 2000.
- [9] J. Ahrens and S. Spors. Sound field reproduction using planar and linear arrays of loudspeakers. *IEEE Transactions on Audio, Speech and Language Processing*, 18(8):2038 – 2050, November 2010. doi:10.1109/TASL.2010.2041106.
- [10] E.W. Start. Application of curved arrays in wave field synthesis. In *110th AES Convention*, Copenhagen, Denmark, May 1996. Audio Engineering Society (AES).
- [11] S. Spors, R. Rabenstein, and J. Ahrens. The theory of wave field synthesis revisited. In *124th AES Convention*. Audio Engineering Society (AES), May 2008.
- [12] E.G. Williams. *Fourier Acoustics: Sound Radiation and Nearfield Acoustical Holography*. Academic Press, 1999.
- [13] S. Spors. Extension of an analytic secondary source selection criterion for wave field synthesis. In *123th AES Convention*, New York, USA, October 2007. Audio Engineering Society (AES).
- [14] J. Ahrens and S. Spors. On the secondary source type mismatch in wave field synthesis employing circular distributions of loudspeakers. In *127th AES Convention*, New York, USA, October 2009. Audio Engineering Society (AES).
- [15] J. Ahrens and S. Spors. An analytical approach to sound field reproduction using circular and spherical loudspeaker distributions. *Acta Acustica united with Acustica*, 94(6):988–999, December 2008.
- [16] F.M. Fazi, P.A. Nelson, J.E.N. Christensen, and J. Seo. Surround system based on three dimensional sound field reconstruction. In *125th AES Convention*, San Francisco, USA, 2008. Audio Engineering Society (AES).
- [17] Y.J. Wu and T. Abhayapala. Soundfield reproduction using theoretical continuous loudspeaker. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Las Vegas, USA, 2008.
- [18] J. Giroire. Integral equation methods for the helmholtz equation. *Integral Equations and Operator Theory*, 5(1):506–517, 1982.
- [19] S. Spors and J. Ahrens. Towards a theory for arbitrarily shaped sound field reproduction systems. *Journal of the Acoustical Society of America*, 123(5):3930, May 2008.
- [20] J. Ahrens. *The Single-Layer Potential Approach Applied on Sound Field Synthesis and Its Extension to Non-Enclosing Distributions of Secondary Sources*. PhD thesis, Technische Universität Berlin, 2010.
- [21] J. Ahrens and S. Spors. Applying the ambisonics approach on planar and linear arrays of loudspeakers. In *Ambisonics Symposium*, Paris, France, May 2010.
- [22] D. de Vries. *Wave Field Synthesis*. Audio Engineering Society, 2009.
- [23] J.-J. Sonke, D. de Vries, and J. Labeeuw. Variable acoustics by wave field synthesis: A closer look at amplitude effects. In *104th AES Convention*, Amsterdam, Netherlands, May 1998. Audio Engineering Society (AES).
- [24] S. Spors and J. Ahrens. A comparison of wave field synthesis and higher-order ambisonics with respect to physical properties and spatial sampling. In *125th AES Convention*. Audio Engineering Society (AES), October 2008.
- [25] J. Ahrens, H. Wierstorf, and S. Spors. Comparison of higher order ambisonics and wave field synthesis with respect to spatial discretization artifacts in time domain. In *AES 40th International Conference on Spatial Audio*, Tokyo, Japan, October 2010. Audio Engineering Society (AES).
- [26] E.N.G. Verheijen. *Sound Reproduction by Wave Field Synthesis*. PhD thesis, Delft University of Technology, 1997.
- [27] S. Spors and R. Rabenstein. Spatial aliasing artifacts produced by linear and circular loudspeaker arrays used for wave field synthesis. In *120th AES Convention*, Paris, France, May 2006. Audio Engineering Society (AES).
- [28] S. Spors and J. Ahrens. Spatial aliasing artifacts of wave field synthesis for the reproduction of virtual point sources. In *126th AES Convention*. Audio Engineering Society (AES), May 2009.
- [29] E.W. Start. *Direct Sound Enhancement by Wave Field Synthesis*. PhD thesis, Delft University of Technology, 1997.
- [30] W. de Bruijn. *Application of Wave Field Synthesis in Videoconferencing*. PhD thesis, Delft University of Technology, 2004.
- [31] P. Vogel. *Application of Wave Field Synthesis in Room Acoustics*. PhD thesis, Delft University of Technology, 1993.
- [32] E. Corteel, M. Nogues, and O. Warusfel. Monitoring distance effect with wave field synthesis. In *6th International Conference on Digital Audio Effects (DAFX)*, London, UK, 2003.
- [33] H. Wittek. *Perceptual differences between wavefield synthesis and stereophony*. PhD thesis, University of Surrey, 2007.
- [34] S.S. Bertet. *Formats Audio 3D Hiérarchiques: Caractérisation Objective et Perceptive des Systèmes Ambisonics d'Ordres Supérieurs*. PhD thesis, Institut National des Sciences Appliquées de Lyon, 2009.
- [35] M. Geier, J. Ahrens, and S. Spors. Object-based audio reproduction and the audio scene description format. *Organized Sound*, 15(3):219–227, October 2010. doi:10.1017/S1355771810000324.
- [36] E. Hulsebos, D. de Vries, and E. Bourdillat. Improved microphone array configurations for auralization of sound fields by Wave Field Synthesis. *Journal of the Audio Engineering Society (AES)*, 50(10), Oct. 2002.
- [37] M. Geier, J. Ahrens, and S. Spors. The soundscape renderer: A unified spatial audio reproduction framework for arbitrary rendering methods. In *124th AES Convention*, Amsterdam, The Netherlands, May 2008. Audio Engineering Society (AES).
- [38] V. Pulkki. Virtual sound source positioning using vector base amplitude panning. *Journal of the AES*, 45(6):456–466, June 1997.
- [39] E. Corteel, U. Horbach, and R.S. Pellegrini. Multichannel inverse filtering of multiexciter distributed mode loudspeakers for wave field synthesis. In *112th AES Convention*, Munich, Germany, May 2002. Audio Engineering Society (AES).
- [40] J. Ahrens and S. Spors. Sound field reproduction employing non-omnidirectional loudspeakers. In *126th AES Convention*, Munich, Germany, May 2009. Audio Engineering Society (AES).

- [41] J. Ahrens and S. Spors. An analytical approach to 2.5D sound field reproduction employing circular distributions of non-omnidirectional loudspeakers. In *European Signal Processing Conference (EUSIPCO)*, Glasgow, UK, August 2009.
- [42] E. Corteel. *Caractérisation et Extensions de la Wave Field Synthesis en conditions réelles d'écoute*. PhD thesis, Université de Paris VI, 2006.
- [43] S. Spors. *Active Listening Room Compensation for Spatial Sound Reproduction Systems*. PhD thesis, University of Erlangen-Nuremberg, 2006.
- [44] E. Corteel. Synthesis of directional sources using wave field synthesis, possibilities, and limitations. *EURASIP Journal on Advances in Signal Processing*, 2007, 2007. Article ID 90509.
- [45] J. Ahrens and S. Spors. Rendering of virtual sound sources with arbitrary directivity in higher order ambisonics. In *123th AES Convention*, New York, USA, October 2007. Audio Engineering Society (AES).
- [46] J. Ahrens and S. Spors. Implementation of directional sources in wave field synthesis. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, USA, October 2007.
- [47] M.A.J. Baalman. *On Wave Field Synthesis and electro-acoustic music, with particular focus on the reproduction of arbitrarily shaped sound sources*. PhD thesis, Technische Universität Berlin, 2008.
- [48] A. Franck, A. Gräfe, T. Korn, and M. Strauss. Reproduction of moving sound sources by wave field synthesis: An analysis of artifacts. In *32nd International AES Conference*, Hillerod, Denmark, September 2007. Audio Engineering Society (AES).
- [49] J. Ahrens and S. Spors. Reproduction of moving virtual sound sources with special attention to the doppler effect. In *124th AES Convention*, Amsterdam, The Netherlands, May 2008. Audio Engineering Society (AES).
- [50] E. Corteel, R. Pellegrini, and C. Kuhn-Rahloff. Wave field synthesis with increased aliasing frequency. In *124th AES Convention*, Amsterdam, The Netherlands, May 2008. Audio Engineering Society (AES).
- [51] H. Wittek. Perceptual enhancement of wavefield synthesis by stereophonic means. *Journal of the Audio Engineering Society*, 55(9), September 2007.
- [52] J. Ahrens and S. Spors. An analytical approach to sound field reproduction with a movable sweet spot using circular distributions of loudspeakers. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Taipei, Taiwan, April 2009.
- [53] S. Spors and J. Ahrens. Local sound reproduction by virtual secondary sources. In *AES 40th International Conference on Spatial Audio*, pages 1–8, Tokyo, Japan, October 2010. Audio Engineering Society (AES).
- [54] J. Ahrens and S. Spors. An analytical approach to local sound field synthesis using linear arrays of loudspeakers. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2011. Submitted.
- [55] S. Spors and J. Ahrens. Interpolation and range extrapolation of head-related transfer functions using virtual local sound field synthesis. In *130th AES Convention*. Audio Engineering Society (AES), May 2011.
- [56] S. Spors and J. Ahrens. Generation of far-field head-related transfer functions using sound field synthesis. In *German Annual Conference on Acoustics (DAGA)*, March 2011.
- [57] S. Spors and J. Ahrens. Efficient range extrapolation of head-related impulse responses by wave field synthesis techniques. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2011.
- [58] Ville Pulkki. Spatial sound reproduction with directional audio coding. *Journal of the Audio Engineering Society (JAES)*, 55(6):503–516, June 2007.
- [59] F. Baumgarte and C. Faller. Binaural cue coding – part I: Psychoacoustic fundamentals and design principles. *IEEE Transactions on Speech and Audio Processing*, 11(6):509–519, November 2003.