

NEUE ANSÄTZE DER SCHALLFELDREPRODUKTION

DISSERTATIONSEXPOSE

von Alois Sontacchi

INHALTSVERZEICHNIS

Einleitung	1
Holophonie	2
Ambisonic	3
Verknüpfung der beiden Theorien.....	8
Realisierung mittels adaptiver Filter.....	10
Realisierung mittels Ambisonic und Raumimpulsantwort.....	14
Ausblick.....	18
Literaturverzeichnis	19

Ansätze der Schallfeldreproduktion

EINLEITUNG

"...Gegeben sei eine primäre Quelle S_p , die ein Feld F_p in einem Volumen V ausstrahlt; weiters eine Oberfläche Z_0 , die S_p umgibt und V in zwei Bereiche teilt: V_1 innerhalb von Z_0 und V_2 außerhalb von Z_0 (Abb. 1a). Es existiert nun ein Satz von sekundären Quellen S_H , die über Z_0 verteilt sind und ein Feld F_H ausstrahlen, welches Null innerhalb V_1 und gleich F_p innerhalb V_2 ist (Abb. 1b). Dieser Satz von sekundären 'Huygens-' Quellen S_H kann die primäre Quelle zur Beschreibung des ausgestrahlten Feldes in V_2 ersetzen..."[1].

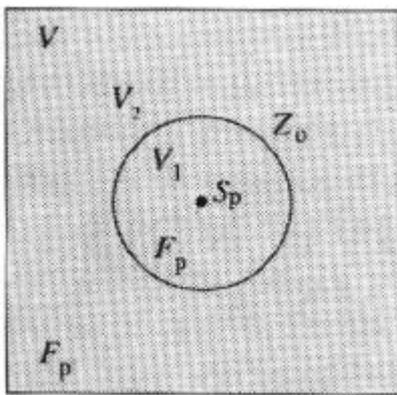


Abb. 1a

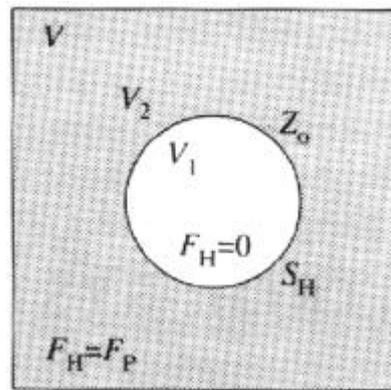


Abb. 1b

Die Wellenfront einer sich ausbreitenden Welle kann durch räumlich verteilte sekundäre Quellen nachgebildet werden. Daraus folgt, dass die verteilten sekundären Quellen die primäre Quelle ersetzen können. Eine Vielzahl von Untersuchungen, die zu Verbesserungen der Abbildung von Stereoapplikationen durchgeführt wurden, führten zu Systemen, die sich in zwei Gruppen einordnen lassen. Die erste Gruppe von Systemen versucht das originale Schallfeld bzw. Schalldruckfeld in einem 2 bzw. 3dimensionalen Bereich um einen Hörer zu reproduzieren. Diese Systemgruppe wird mit dem Begriff "Holographie - Holophonie" bezeichnet. Die zweite Gruppe von Systemen strebt die originalgetreue Wiedergabe des Schalldrucks an den Ohren des Hörers an. Diese Systeme werden mit dem Begriff "Transaural" bezeichnet [2].

Im Folgenden werden die "holophonen Systeme" behandelt, da sie im Vergleich zur anderen Systemgruppe eine größere Hörumgebung schaffen. Dadurch vergrößert sich die Bewegungsfreiheit eines einzelnen Hörers, und es bietet sich die Möglichkeit einer Wiedergabe für mehrere Hörer.

HOLOPHONIE

Der Ausgangspunkt für das Holophonieverfahren stellt die oben genannte Theorie von Huygens dar. Das Kirchhoff-Helmholtz Integral (siehe Gleichung 1) beschreibt diesen Zusammenhang in mathematischer Form [3].

$$p(\vec{r}) = \iint_{\partial\Omega} \left[\vec{\nabla} p_0 \cdot \vec{n} - \frac{\vec{R}}{R} \cdot \vec{n} (1 + jkR) \frac{p_0}{R} \right] \frac{e^{-jkR}}{4\pi R} dS_0 \quad \forall \vec{r} \in \Omega \quad (1)$$

Diese Gleichung besagt, dass das Schallfeld, beschrieben durch den Schalldruck p , innerhalb eines definierten Volumen Ω durch das Integral über dieses Volumen entlang der begrenzenden Oberfläche $\partial\Omega$, wobei p_0 dem Druck entlang $\partial\Omega$ entspricht, beschrieben werden kann. In Gleichung (1) ist k die Wellenkennzahl, und \vec{n} ist der normierte nach außenzeigende Flächennormalvektor. Der Vektor \vec{R} beschreibt den Ausbreitungspfad zwischen der sekundären Quelle, positioniert an der Stelle \vec{r}_0 und dem Hörpunkt (Aufpunkt) \vec{r} (Abb. 2).

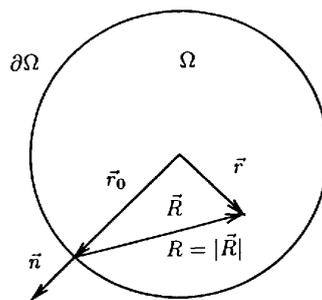


Abb. 2

In Gleichung (1) wird die Verteilung der sekundären Quellen durch Monopol- und Dipolquellen beschrieben. Die Amplituden dieser Quellen entsprechen dem Druckgradienten $\vec{\nabla} p_0$ sowie dem Druck p_0 des Schallfeldes entlang der begrenzenden Oberfläche $\partial\Omega$. Um ein bestimmtes Schallfeld zu reproduzieren, erfolgt zuerst die Aufnahme von Druckgradient und Druck entlang der begrenzenden Oberfläche mittels Schnelle- und Druckmikrofonen. Die Wiedergabe erfolgt durch Lautsprecher, die Monopol- bzw. Dipolquellen entsprechen, wobei die Wiedergabe in einem schalltoten Raum erfolgt.

Zusammenfassung

Zu den Vorteilen des Holophonieverfahrens zählt die relativ einfache Aufnahme mittels Schnelle- und Druckmikrofonen. Für eine exakte Abbildung in Abhängigkeit von der Größe des abzubildenden Volumens ist jedoch ein entsprechend hoher Aufwand notwendig, um ein räumlich-spektrales Aliasing zu vermeiden. Den Nachteil des Verfahrens bildet die praktische Realisierung einer idealen Monopol- bzw. Dipolquelle.

AMBISONIC

Ambisonic bietet die Möglichkeit ein Schallfeld über 360° bzw. dreidimensional durch symmetrisch entlang eines Kreises bzw. einer Kugeloberfläche angeordnete Lautsprecher zu reproduzieren. Im Folgenden wird der 2 dimensionale Fall betrachtet, da der 3 dimensionale Fall direkt daraus folgt (keine neuen Erkenntnisse) und die mathematische Beschreibung an Komplexität stark zunimmt.

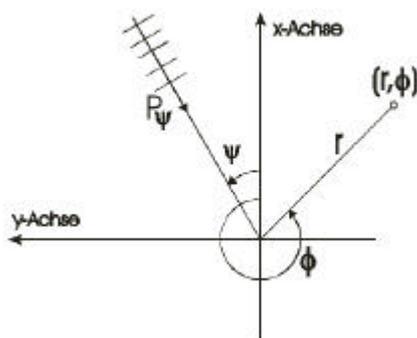
Die Idee hinter dem Ambisonicmodell basiert auf dem Vergleich einer aufgenommenen Referenzwelle mit der reproduzierten Welle des Systems [4]. Dabei werden folgende Annahmen vereinbart:

1. Annahme: Die aufgenommene Welle (Referenzwelle) ist eine ebene Welle. Diese Annahme ist gültig, solange die Schallquelle weit genug vom Hörer entfernt ist.

2. Annahme: Die reproduzierte Welle eines Lautsprechers ist ebenfalls eine ebene Welle. Auch diese Annahme ist gültig, solange die Lautsprecher weit genug vom Zuhörer entfernt sind. Die sich ergebende Gesamtwelle des Ambisonicsystems erfolgt durch Superposition der abgestrahlten Teilwellen der einzelnen Lautsprecher. Sie ist wiederum eine ebene Welle.

Die beiden oben genannten Annahmen beschränken sich auf die Analyse des direkten Schallanteils, d.h. es wird die Freifeld Situation betrachtet.

Beschreibung und Darstellung des Ambisonic Systems



Der Hörer befindet sich an der Stelle (r, ϕ) . Die ebene Welle trifft aus der Richtung \mathbf{y} ein (siehe *Annahme 1*). Der Druck dieser Welle wird durch P_y beschrieben. Es ergibt sich im Punkt (r, ϕ) folgende Welle¹:

$$S_{\text{Referenz}}(r, \mathbf{f}) = P_y \cdot e^{jk \cdot r} = P_y \cdot e^{jk \cdot r \cdot \cos(\mathbf{f} - \mathbf{y})}$$

$k = \frac{2\pi}{\lambda}$ ist die Wellenzahl und λ ist die Wellenlänge.

¹ \vec{k} beschreibt eine Welle der Wellenzahl k , die sich in eine bestimmte Richtung (hier \mathbf{y}) bewegt. $S_{\text{Referenz}}(r, \mathbf{f})$ beschreibt hier die Referenz (aufgenommene) Welle im Punkt (r, ϕ) .

Eine ebene Welle kann in eine Reihe durch Kosinus- und Besselfunktionen entwickelt werden.

$$\begin{aligned}
 S_{\text{Referenz}}(r, \mathbf{f}) &= P_y \cdot e^{jkr \cos(\mathbf{f}-\mathbf{y})} = P_y J_0(kr) + 2P_y \sum_{m=1}^{\infty} i^m J_m(kr) \cos[m(\mathbf{f}-\mathbf{y})] \\
 &= P_y \left(J_0(kr) + 2 \sum_{m=1}^{\infty} i^m J_m(kr) [\cos(m\mathbf{f}) \cos(m\mathbf{y}) + \sin(m\mathbf{f}) \sin(m\mathbf{y})] \right) \quad (2)
 \end{aligned}$$

Die Reihe in Gl. (2) besitzt unendlich viele Reihenelemente, d.h. man benötigt für eine exakte Reproduktion unendlich viele Übertragungskanäle und ebenso viele Lautsprecher. In der Praxis kann diese Reihe daher nur bis zu einem bestimmten Grad approximiert werden. Die Anzahl der verwendeten M Reihenelemente stellt die Ordnung des Systems dar.

Bei einem 2-dimensionalen, symmetrischen Ambisonic System sind N Lautsprecher entlang eines Kreises, äquidistant angeordnet. Jeder Lautsprecher strahlt eine ebene Welle ab (siehe *Annahme 2*). Die abgestrahlte Welle des n -ten Lautsprechers an der Position (\mathbf{f}_n) bewirkt im Punkt (r, ϕ) eine Welle, die wie folgt beschrieben werden kann:

$$\begin{aligned}
 S_n(r, \mathbf{f}) &= P_n \cdot e^{jk_n \cdot r} = P_n \cdot e^{jkr \cos(\mathbf{f}-\mathbf{f}_n)} = \\
 &= P_n \left(J_0(kr) + 2 \sum_{m=1}^{\infty} i^m J_m(kr) [\cos(m\mathbf{f}_n) \cos(m\mathbf{f}) + \sin(m\mathbf{f}_n) \sin(m\mathbf{f})] \right) \quad (3)
 \end{aligned}$$

Da sich die abgestrahlten Wellen der einzelnen Lautsprecher überlagern, folgt für die Gesamtwellen (System reproduzierte Welle, ebene Welle) im Punkt (r, ϕ):

$$\begin{aligned}
 S_{\text{Ambisonic}}(r, \mathbf{f}) &= \sum_{n=1}^N S_n(r, \mathbf{f}) = \sum_{n=1}^N P_n \cdot e^{jk_n \cdot r} = \sum_{n=1}^N P_n \cdot e^{jkr \cos(\mathbf{f}-\mathbf{f}_n)} = \\
 &= \sum_{n=1}^N P_n \cdot J_0(kr) + 2 \sum_{m=1}^{\infty} i^m J_m(kr) \left[\sum_{n=1}^N P_n \cdot [\cos(m\mathbf{f}_n) \cos(m\mathbf{f}) + \sin(m\mathbf{f}_n) \sin(m\mathbf{f})] \right] \quad (4)
 \end{aligned}$$

Durch einen Koeffizientenvergleich bzgl. m in Gl. (2) und Gl. (4) ergeben sich die sogenannten "*matching conditions*" (siehe Gl.(5)). Diese "*matching conditions*" müssen erfüllt werden, damit die reproduzierte (Ambisonic-) Welle mit der originalen (Referenz-) Welle übereinstimmt. Jede Ordnung m entspricht einer der m -ten Kugelharmonischen, die zur Übereinstimmung gebracht werden muss.

$$\begin{aligned}
 P_y &= \sum_{n=1}^N P_n & W \\
 P_y \cdot \cos(m\mathbf{y}) &= \sum_{n=1}^N P_n \cdot \cos(m\mathbf{f}_n) & m=1 \rightarrow X \quad m=2 \rightarrow U \\
 P_y \cdot \sin(m\mathbf{y}) &= \sum_{n=1}^N P_n \cdot \sin(m\mathbf{f}_n) & m=1 \rightarrow Y \quad m=2 \rightarrow V
 \end{aligned} \quad (5)$$

Aus den "matching conditions" ergeben sich bei Übereinstimmung die jeweiligen dazugehörigen Ambisonicsignale W, X, Y, U, V, usw..

Unter Verwendung der 0.Ordnung wird im Mittelpunkt (0,0) der Anordnung der Druck exakt reproduziert (W-Signal, omnidirektional - ungerichtet), da alle Besselfunktionen, deren Ordnung von 0 verschieden ist, für das Argument $kr=0$ keinen weiteren Anteil liefern. Bei der Verwendung der 1.Ordnung wird die Schallschnelle $v(r, \mathbf{f}) \equiv \text{grad}(p(r, \mathbf{f}))$ (mit den richtungsabhängigen Signalen X und Y) im Mittelpunkt (0,0) der Anordnung exakt reproduziert (d.h. nur die Ableitung der Besselfunktion 1.Ordnung liefert einen Anteil). Je höher die Ordnung des Systems ist, desto genauer wird die Referenzwelle auch in einer Umgebung² um den Koordinatenursprung durch die systemproduzierte Welle rekonstruiert.

Die Kodierung der einzelnen Kanäle erfolgt unabhängig von der Lautsprecheranordnung (siehe "matching conditions" Gl.(5)). Bei der Dekodierung wird jeder Lautsprecher in Abhängigkeit von seiner Position (ϕ_n) durch eine Summe der entsprechend gewichteten Ambisonicsignale versorgt. Somit ergibt sich für den i-ten Lautsprecher an der Position (ϕ_i):

$$LS_i = G_1 \cdot \sum_{n=1}^N P_n + G_2 \cdot \left[\sum_{m=1}^{\infty} \left(\left(\sum_{n=1}^N P_n \cdot \cos m \mathbf{f}_n \right) \cos(m \mathbf{f}_i) + \left(\sum_{n=1}^N P_n \cdot \sin m \mathbf{f}_n \right) \cdot \sin(m \mathbf{f}_i) \right) \right] \quad (6a)$$

$$LS_i = G_1 \cdot W + G_2 \cdot (X \cos \mathbf{f}_i + Y \sin \mathbf{f}_i + U \cos(2\mathbf{f}_i) + V \sin(2\mathbf{f}_i) + \text{usw.}) \quad (6b)$$

Die Wahl der Gewichte G_1 und G_2 ist abhängig von der Anzahl der Lautsprecher. Zur Unterstützung der psychoakustischen Lokalisierungsmechanismen (siehe unten, *Lokalisierbarkeit bei Ambisonic Systemen*) können G_1 und G_2 ebenfalls eine Frequenzabhängigkeit aufweisen. In der Literatur werden unterschiedliche Faktoren in Abhängigkeit vom Anwendungsgebiet angegeben [5]. Die benötigte Anzahl der Übertragungskanäle für ein 2 dimensionales System ist $(2M+1)$ und für ein 3 dimensionales $(M+1)^2$.

Das resultierende Schallfeld für ein Ambisonic System M.Ordnung ergibt sich aus der Superposition der Teilwellen die von den N Lautsprecher abgestrahlt werden unter Verwendung von Gl. (6a) zu:

$$\begin{aligned} S_{\text{Ambisonic}}(r, \mathbf{f}) &= \sum_{i=1}^N LS_i \cdot e^{j k \mathbf{f}_i \cdot \mathbf{r}} = \sum_{i=1}^N LS_i \cdot e^{j k r \cdot \cos(\mathbf{f} - \mathbf{f}_i)} = \\ &= \sum_{i=1}^N \left\{ G_1 \cdot \sum_{n=1}^N P_n + G_2 \cdot \left[\sum_{m=1}^M \left(\left(\sum_{n=1}^N P_n \cdot \cos m \mathbf{f}_n \right) \cos(m \mathbf{f}_i) + \left(\sum_{n=1}^N P_n \cdot \sin m \mathbf{f}_n \right) \cdot \sin(m \mathbf{f}_i) \right) \right] \right\} \cdot e^{j k r \cdot \cos(\mathbf{f} - \mathbf{f}_i)} = \\ &= \sum_{i=1}^N \left\{ G_1 \cdot P_y + G_2 \cdot \left[\sum_{m=1}^M (P_y \cdot \cos m y \cdot \cos m \mathbf{f}_i + P_y \cdot \sin m y \cdot \sin m \mathbf{f}_i) \right] \right\} \cdot e^{j k r \cdot \cos(\mathbf{f} - \mathbf{f}_i)} \end{aligned}$$

² Mit zunehmender Entfernung vom Koordinatenursprung (kr) wird der Anteil der höheren Besselfunktionen immer relevanter.

$$S_{\text{Ambisonic}}(r, \mathbf{f}) = P_y \cdot \sum_{i=1}^N \left[G_1 + G_2 \cdot \sum_{m=1}^M \cos[m(\mathbf{y} - \mathbf{f}_i)] \right] \cdot e^{jkr \cdot \cos(\mathbf{f} - \mathbf{f}_i)} \quad (7)$$

Die Kodierungsvorschrift für Ambisonic 1.Ordnung, 2 dimensionaler Fall (B-Format) und dem Quellensignal S aus der Richtung \mathbf{y} lautet:

$$W = 0.707 \cdot S$$

$$X = S \cdot \cos \mathbf{y}$$

$$Y = S \cdot \sin \mathbf{y}$$

Für das B-Format ergeben sich Lautsprechersignale, die abhängig von der Position \mathbf{f}_i der N Lautsprecher sind (bei der Wahl von $G_1=1/N$ und $G_2=2/N$, d.h. $G_2/G_1=2$ siehe [5], Gerzon.):

$$LS_i = \frac{1}{N} (W + 2X \cos \mathbf{f}_i + 2Y \sin \mathbf{f}_i)$$

Untersuchung der Abbildung von Phantomquellen³ und deren Lokalisierbarkeit bei Ambisonic Systemen.

Bei der Lokalisation kann das Ohr als ein System aufgefasst werden, welches auf eine Funktion aus einer bestimmten Richtung reagiert. Diese Funktion kann durch die Amplitude einer Schallwelle (lineare Größe, Theorie 1.Ordnung), durch die Energie einer Schallwelle (quadratische Größe, Theorie 2.Ordnung), oder durch eine komplexere Größe dargestellt werden [7]. Nach M. Gerzon kann die Theorie 1.Ordnung durch das "Velocity Model" beschrieben werden. Dabei wird der Schnellevektor einer reproduzierten Schallquelle berechnet (siehe Berechnung des Schnellektors S.7). Diese Form der Lokalisationsbestimmung wird auch "Makita Localization" genannt und ist nach psychoakustischen Untersuchungen für Frequenzen unter 700Hz von großer Bedeutung. Die Theorie 2.Ordnung kann durch das "Energy Model" beschrieben werden. Für eine reproduzierte Schallquelle wird der Energievektor berechnet (siehe Berechnung des Energievektors S.7) und die damit bestimmte Richtung wird als "Energy Vector Localization" bezeichnet. Die Bedeutung dieser Größen liegt in einem Frequenzbereich von 500 bis 5000Hz.

In einem natürlichen Schallfeld stimmt die Richtung der Schallschnelle und die des Energieflusses einer Quelle überein. Bei einer Schallfeldreproduktion erfolgt die Berechnung des Schnelle- und des Energievektors durch die Kenntnis der Lautsprecherpositionen \mathbf{f}_i und der Verstärkungsfaktoren g_i der N Lautsprecher.

Es folgt die Berechnung für ein Ambisonic System 1.Ordnung in 2 Dimensionen (Alle Lautsprecher befinden sich gleich weit vom Hörer entfernt, und es wird ein Koordinatensystem

³ Unter einer Phantomquelle versteht sich die Abbildung einer "beliebig positionierten Quelle" (d.h. eine Quelle außerhalb des Lautsprecherarrays, die eine ebene Welle abstrahlt), deren Schallfeld sich durch Überlagerung der von den Lautsprechern abgestrahlten Feldern ergibt.

eingeführt, dessen Ursprung mit der Position des Hörers übereinstimmt. Der Winkel \mathbf{f}_i ist jener Winkel, den die x -Achse des gewählten Koordinatensystems mit der Lage des i -ten Lautsprechers einschließt.).

Theorie 1.Ordnung: Berechnung für den Schnellektor.

$$P = \sum_{i=1}^N g_i \quad (8a)$$

$$P_x = \sum_{i=1}^N g_i \cdot \cos(\mathbf{f}_i) \quad (8b)$$

$$P_y = \sum_{i=1}^N g_i \cdot \sin(\mathbf{f}_i) \quad (8c)$$

$$|\vec{r}_V| = \sqrt{\left(\frac{P_x}{P}\right)^2 + \left(\frac{P_y}{P}\right)^2} \quad \text{und} \quad \mathbf{q}_V = \arctan\left(\frac{P_y}{P_x}\right) \quad (8d)$$

Theorie 2.Ordnung: Berechnung für den Energievektor.

$$E = \sum_{i=1}^N g_i^2 \quad (9a)$$

$$E_x = \sum_{i=1}^N g_i^2 \cdot \cos(\mathbf{q}_i) \quad (9b)$$

$$E_y = \sum_{i=1}^N g_i^2 \cdot \sin(\mathbf{q}_i) \quad (9c)$$

$$|\vec{r}_E| = \sqrt{\left(\frac{E_x}{E}\right)^2 + \left(\frac{E_y}{E}\right)^2} \quad \text{und} \quad \mathbf{q}_E = \arctan\left(\frac{E_y}{E_x}\right) \quad (9d)$$

Im Idealfall, wie bereits oben erwähnt wurde, müsste der Schnellektor \vec{r}_V und der Energievektor \vec{r}_E einer Phantomschallquelle, unabhängig von ihrer Lage, die gleiche Richtung besitzen und der Betrag dieser beiden Vektoren gleich eins sein [6-9]. Da dieser Idealfall im allgemeinen nicht realisierbar ist, ergeben sich folgende Bedingungen:

$$\mathbf{q}_V \equiv \mathbf{q}_E \quad (10)$$

$$\max |\vec{r}_V| \quad (11)$$

$$\min \left(1 - |\vec{r}_E|\right) \quad (12)$$

Zusammenfassung

Der Vorteil von Ambisonic ist die relativ einfache Kodierung und Dekodierung. Ein realisiertes B-Format Schallfeld kann wie ein grafisches Objekt beliebig mit Rotationsmatrizen gedreht (rotate)⁴, gekippt (tilt)⁵ oder geneigt (tumble)⁶ werden. Da es sich um ein koinzidentes Aufnahmeverfahren handelt, die Mikrofone aber eine physikalische Ausdehnung besitzen und komplexe Mikrofoncharakteristika (d.h. $\cos(2\mathbf{y})$, $\sin(2\mathbf{y})$, etc.) nur begrenzt existieren, ist dieses Verfahren für Ordnungen größer 1 nur begrenzt realisierbar.

VERKNÜPFUNG DER BEIDEN THEORIEN

Holophonie ist eine allgemeine Methode zur Schallfeldreproduktion, während Ambisonic einen Spezialfall darstellt. Dieser Spezialfall entsteht durch folgende drei Annahmen und kann mathematisch nachgewiesen werden [10].

- Das originale Schallfeld ist eine ebene Welle, siehe Gl. (2) (keine Einschränkung der Allgemeinheit, weil jede Welle als Superposition von ebenen Wellen dargestellt werden kann).
- Die sekundären Quellen sind entlang einer kreisförmigen Linie angeordnet.
- Die sekundären Quellen sind vom "Abhörgebiet" weit genug entfernt, sodass ihre abgestrahlten akustischen Felder wiederum durch ebene Wellen beschrieben werden können.

Der Übergang vom 3 dimensional Volumen, beschrieben durch das Kirchhoff - Helmholtz'sche Integral (siehe Gl. (1)), zum 2 dimensional Fall (Wellenzustände in einer Ebene), bei dem die sekundären Quellen entlang des Kreisumfangs verteilt sind, erfolgt durch die "*Stationary Phase Approximation*". Es wird dabei angenommen, dass sich ein 3 dimensionales Gebilde, z.B. ein unendlich ausgedehnter Zylinder (entlang der z-Achse) aus lauter endlichen Linienquellen zusammensetzt (siehe Gl. (13)).

⁴ Drehung um die z-Achse

⁵ Drehung um die x-Achse

⁶ Drehung um die y-Achse

$$p(\vec{r}) = \int_0^{2p} \int_{-\infty}^{+\infty} \left[\vec{\nabla} p_0 \cdot \vec{n} - \frac{\vec{R}}{R} \cdot \vec{n} (1 + jkR) \frac{p_0}{R} \right] \frac{e^{-jkR}}{4R\mathbf{p}} r_0 \cdot dz_0 \cdot d\mathbf{f}_0 \quad \forall \vec{r} \in \Omega \quad (13)$$

Will man nun die Feldzustände innerhalb einer bestimmten Ebene z.B. an der Stelle $z=z_0$ beschreiben, so kann man das unendliche Integral unter Voraussetzung der "Stationary Phase Approximation"⁷ wie folgt berechnen:

$$I = \int_{-\infty}^{+\infty} \left[\vec{\nabla} p_0 \cdot \vec{n} - \frac{\vec{R}}{R} \cdot \vec{n} (1 + jkR) \frac{p_0}{R} \right] \frac{e^{-jkR}}{4R\mathbf{p}} dz_0 \cong \frac{1}{2\sqrt{jk2\mathbf{p}}} \cdot \left[\vec{\nabla} p_0 \cdot \vec{n} - \frac{\vec{R}}{R} \cdot \vec{n} (1 + jkR) \frac{p_0}{R} \right] \frac{e^{-jkR}}{\sqrt{R}} \Bigg|_{z_0=0}$$

Wird für das originale Schallfeld p_0 eine horizontale ebene Welle (Amplitudenbetrag a , Wellenzahl k und Einfallswinkel θ) angenommen, die durch $p_0(r_0, \mathbf{f}_0) = a \cdot e^{jk r_0 \cos(\mathbf{f}_0 - \mathbf{q})}$ beschrieben wird, so ergibt sich für den Druck innerhalb des Kreises mit dem Radius r_0 :

$$p(\vec{r}) = p(r, \mathbf{f}) = \frac{r_0}{2\sqrt{jk2\mathbf{p}}} \int_0^{2p} \left[\frac{\partial p_0}{\partial r_0} - \cos \mathbf{a} (1 + jkR) \frac{p_0}{R} \right] \frac{e^{-jkR}}{\sqrt{R}} d\mathbf{f}_0 \quad \text{für } r < r_0 \quad (14)$$

$$\text{mit } R = \sqrt{r^2 + r_0^2 - 2rr_0 \cos(\mathbf{f} - \mathbf{f}_0)} \quad \text{und} \quad \cos \mathbf{a} = \frac{r \cos(\mathbf{f} - \mathbf{f}_0) - r_0}{R}$$

\vec{R} beschreibt den Ausbreitungspfad zwischen der Sekundärquelle an der Stelle \vec{r}_0 und dem Aufpunkt \vec{r} . Die Gl. (13) beschreibt sowohl die Ausbreitung von Monopol- ($\frac{e^{-jkR}}{4R\mathbf{p}}$) als auch von

Dipolquellen ($\frac{\vec{R}}{R} \cdot \vec{n} (1 + jkR) \cdot \frac{e^{-jkR}}{4R^2\mathbf{p}}$).

Unter Verwendung der dritten Annahme, dass sich die sekundären Quellen weit genug vom Hörer (Aufpunkt) entfernt befinden d.h. $|\vec{r}_0| > I_{\max}$, gilt der Übergang $\lim_{r_0 \rightarrow \infty}$ und somit ergeben sich folgende Näherungen:

$$\lim_{r_0 \rightarrow \infty} \frac{1 + jkR}{R} = jk$$

$$\lim_{r_0 \rightarrow \infty} \cos \mathbf{a} = -1$$

$$\lim_{r_0 \rightarrow \infty} \frac{e^{-jkR}}{\sqrt{R}} = \frac{e^{-jk r_0}}{\sqrt{r_0}} \cdot e^{jk r \cos(\mathbf{f} - \mathbf{f}_0)}$$

⁷ Das Ergebnis für das approximierte Integral ist nicht von den globalen Eigenschaften des Integranden, sondern von der unmittelbaren Umgebung des stationären Punktes abhängig.

Unter Verwendung dieser Näherungen ergibt sich für den Druck aus Gl. (14) folgender Zusammenhang:

$$p(\vec{r}) = p(r, \mathbf{f}) = \frac{\sqrt{r_0}}{2\sqrt{jk2\mathbf{p}}} \cdot e^{-jk r_0} \int_0^{2\mathbf{p}} \left[\frac{\partial p_0}{\partial r_0} + jk p_0 \right] e^{jkr \cos(\mathbf{f}-\mathbf{f}_0)} d\mathbf{f}_0 \quad \text{für } r < r_0 \quad (15)$$

In Gl. (10) ist die ebene Welle durch eine Reihenentwicklung von Besselfunktionen dargestellt. Für die Annahme $\lim_{r_0 \rightarrow \infty}$ kann die Besselfunktion durch ihre asymptotische Näherung ersetzt werden.

$$\lim_{r_0 \rightarrow \infty} J_n(kr_0) = \sqrt{\frac{2}{kr_0\mathbf{p}}} \cdot \cos \left[kr_0 - \left(n + \frac{1}{2}\right) \cdot \frac{\mathbf{p}}{2} \right]$$

und weiter gilt:

$$\lim_{r_0 \rightarrow \infty} \left[\frac{\partial p_0}{\partial r_0} + jk p_0 \right] = a \cdot \sqrt{\frac{2jk}{r_0\mathbf{p}}} \cdot e^{jkr_0} \left[1 + 2 \sum_{n=1}^{+\infty} \cos[n(\mathbf{f}_0 - \mathbf{q})] \right]$$

Somit ergibt sich für N sekundäre Quellen, die entlang einer Kreislinie angeordnet sind

$$p(\vec{r}) = p(r, \mathbf{f}) = \frac{1}{N} \sum_{i=0}^{N-1} \left[1 + 2 \sum_{n=1}^M \cos[n(\mathbf{f}_i - \mathbf{q})] \right] \cdot e^{jkr \cos(\mathbf{f}-\mathbf{f}_i)} \quad (16)$$

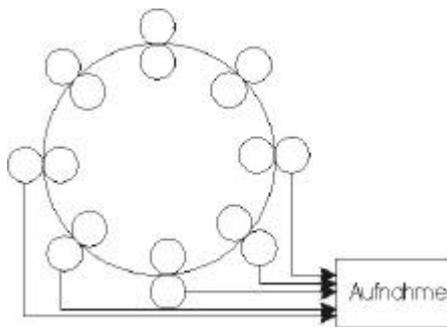
mit: N Anzahl der sek. Quellen, \mathbf{f}_i = Position der sek. Quellen, \mathbf{q} = Einfallswinkel der primären Welle und M Approximationsgrad der asymptotisch genäherten Besselreihe.

Zusammenfassung

Da eine Aufnahme des Schallfeldes mittels Holophonie und eine Schallfeldreproduktion mittels Ambisonic theoretisch relativ einfach zu realisieren sind, liegt eine Verknüpfung dieser beiden Methoden sehr nahe. Die Verknüpfung dieser beiden Ansätze ist schon durch die ähnliche äquivalente mathematische Struktur der beiden Systeme (vergleiche Gl. (7) mit (16)) gegeben. Mit Hilfe einer Holophonie-Anordnung werden die N-Mikrofonsignale aufgenommen und über eine Transformationsmatrix (siehe [10]) die entsprechenden Ambisonicsignale berechnet. Dadurch kann man das Problem der Mikrofoncharakteristika für höhere Ambisonic Ordnungen umgehen. Entsprechend der Lautsprecherkonfiguration werden über eine sogenannte Dekodiermatrix (siehe Gl.(6a)) die Lautsprechersignale berechnet. Die Überlagerung der von den einzelnen Lautsprecher abgestrahlten Schallwellen führt zu einer Reproduktion des aufgenommenen Schallfeldes, wobei bei der Wiedergabe eine Freifeldbedingung (schalltoter Raum) vorausgesetzt wird.

REALISIERUNG MITTELS ADAPTIVER FILTER

Wie oben bereits erwähnt wurde, ist eine exakte Wiedergabe d.h. eine Schallfeldreproduktion an einem beliebigen Ort, mit Hilfe von Ambisonic bzw. Holophonie nur dann realisierbar, wenn man die akustischen Eigenschaften des Wiedergaberaumes vernachlässigen kann. Diese



Voraussetzung ist nur dann erfüllt, wenn die Anordnung in einem nicht geschlossenen bzw. schalltoten Raum realisiert wird. In der Praxis handelt es sich jedoch zumeist um Anwendungen, bei denen der Wiedergaberaum nicht vernachlässigt werden kann. Darum wird im folgenden ein neuer Ansatz zur Schallfeldreproduktion vorgestellt, der mit Hilfe von adaptiven Filtern diese Aufgabe erfüllt.

Abb.3a Aufnahmeschema

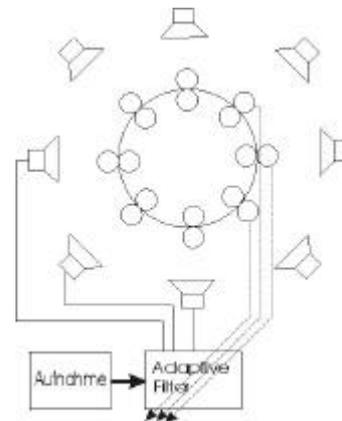


Abb. 3b Wiedergabeschema

Zu Beginn steht eine Aufnahme mit einem Mikrofonarray, bestehend aus M Druck- und Schnellemikrofonen. Die aufgezeichneten Mikrofonensignale werden über eine Wiedergabematrix W auf ein definiertes Lautsprecherlayout mit L Lautsprechern aufgeteilt und von diesen abgestrahlt. Gleichzeitig wird das Schallfeld mit $S (=M)$ Sensoren abgetastet. Die Positionierung der Sensoren erfolgt entsprechend gleich dem verwendeten Mikrofonarray bei der Aufnahme. Mittels adaptiver Filter, die durch die entstehenden Fehlersignale an den räumlichen Abtastpunkten des Sensorarrays gesteuert werden, wird die Wiedergabematrix W modifiziert. Dadurch ergibt sich nach Konvergenz der Filter die äquivalente Situation, wie bei der Aufnahme, an den kontrollierten Abtastpunkten (Sensoren).

In Abbildung 4 ist das Strukturbild für den adaptiven Prozess dargestellt [11],[12].

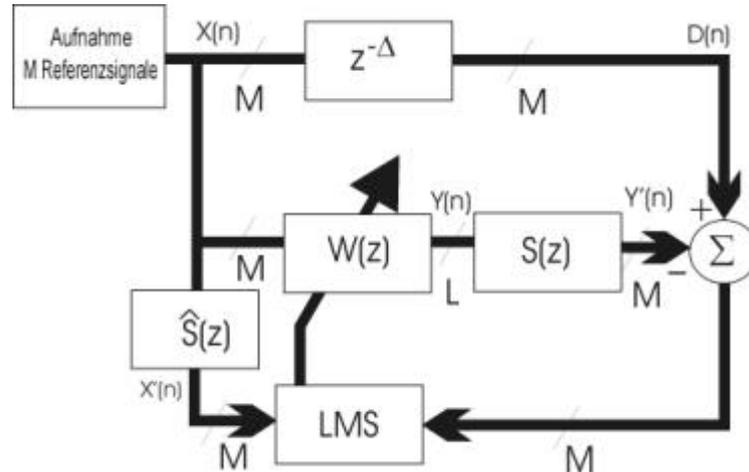


Abb. 4 Blocksachbild eines adaptiven Mehrkanal "Feedforward" Systems zur inversen Modellierung.

Die Ausgangssignale der adaptiven Filter W (d.h. die L Lautsprechersignale) ergeben sich zu:

$$y_l(n) = \sum_{m=1}^M \mathbf{w}_{l,m}^T(n) \cdot \mathbf{x}_m(n) \quad l = 1, 2, \dots, L \quad (17)$$

Dabei stellt $\mathbf{w}_{l,m}(n) \equiv [w_{l,m,0}(n) \ w_{l,m,1}(n) \ \dots \ w_{l,m,N-1}(n)]^T$ die Gewichtsvektoren und $\mathbf{x}_m(n) \equiv [x_m(n) \ x_m(n-1) \ \dots \ x_m(n-N+1)]^T$ mit $m=1, 2, \dots, M$ die Eingangsvektoren (die M Mikrofonsignale) der adaptiven Filter dar. Bei der Realisierung mittels adaptiver Filter werden $L \times M$ FIR Filter der Länge N benötigt. Findet die Wiedergabe in einem Raum mit langer Nachhallzeit statt, so werden die entsprechenden Raumübertragungsfunktionen (Raumimpulsantworten) nur durch sehr lange FIR Filter approximiert werden können⁸. Es existieren $L \times M$ verschiedene sekundäre Übertragungspfade $S_{l,s}(z)$, da die Anzahl S der Sensoren gleich der Anzahl M der Aufnahmemikrofone ist. Für den FXLMS Algorithmus ("filtered referencesignal least mean square") müssen die M Eingangssignale $\mathbf{x}_m(n)$ durch die modellierten sekundären Übertragungspfade $\hat{S}_{l,s}(z)$ gefiltert werden⁹. Die Bestimmung der Übertragungspfade $\hat{S}_{l,s}(z)$ erfolgt durch messen der Raumimpulsantworten vom jeweils l -ten Lautsprecher zum s -ten Sensormikrofon.

Die gefilterten Referenzsignale berechnen sich wie folgt:

$$\mathbf{x}'_{l,m,s}(n) \equiv \hat{S}_{l,s}(z) * \mathbf{x}_m(n) \quad (18)$$

Die Kostenfunktion für den FXLMS Algorithmus lautet:

$$\mathbf{x}(n) = \sum_{m=1}^M e_m^2(n) = \mathbf{e}^T(n) \cdot \mathbf{e}(n) \quad (19)$$

⁸ Eine Abhilfe für die praktische Realisierbarkeit erreicht man durch eine entsprechende Raumadaption oder durch den Entwurf der inversen Modellierung mittels IIR Filter.

⁹ Diese Vorgangsweise gewährleistet die zeitliche Synchronizität zwischen Fehler- und Referenzsignal.

Somit ergibt sich wie folgt, für den FXLMS Algorithmus eines Mehrkanalreferenz- und Mehrkanalausgangsystems die rekursive Formel für die adaptiven Gewichte.

$$\mathbf{w}_{l,m}(n+1) = \mathbf{w}_{l,m}(n) + \mathbf{m} \sum_{s=1}^S \mathbf{x}'_{l,m,s}(n) \cdot e_s(n) \quad (20)$$

Der Faktor \mathbf{m} in Gl. (20) stellt die Geschwindigkeit der Adaption ein. Dieser kann frei aus dem Intervall $[0 < \mathbf{m} < \mathbf{m}_{\max}]$ gewählt werden. Eine schnelle Adaption bei einem stationären Eingangssignal kann durch die Wahl eines großen Wertes für \mathbf{m} erreicht werden, jedoch reagiert dieses System sehr stark auf kleine Änderungen des Eingangssignals.

$$\mathbf{m}_{\max} = \frac{1}{P_{x'}(N + \Delta)} \quad (21)$$

$P_{x'} = \frac{1}{M} \sum_{m=1}^M E[\mathbf{x}_m^2(n)]$ ist die mittlere Leistung der gefilterten Mikrofon-signale, N ist die Ordnung der verwendeten adaptiven Filter und Δ ist die Anzahl der Samples, die der maximalen Verzögerung im sekundären Pfad entspricht.

Eine Modifizierung der Kostenfunktion hinsichtlich der Regularisierung der Gewichte bewirkt eine Minimierung der abgegebenen Leistung der Lautsprecher.

$$\mathbf{x}(n) = \mathbf{e}^T(n) \cdot \mathbf{e}(n) + \mathbf{g} \cdot \sum_{l=1}^L \sum_{m=1}^M \mathbf{w}_{l,m}^T(n) \cdot \mathbf{w}_{l,m}(n) \quad (22)$$

Dadurch ergibt sich der sogenannte "Leaky FXLMS" Algorithmus. Die rekursive Formel für die Gewichte ergibt sich zu:

$$\mathbf{w}_{l,m}(n+1) = \mathbf{n} \cdot \mathbf{w}_{l,m}(n) + \mathbf{m} \sum_{s=1}^S \mathbf{x}'_{l,m,s}(n) \cdot e_s(n) \quad (23)$$

Der Faktor \mathbf{m} berechnet sich wie in Gl. (21). Der Gewichtungsfaktor \mathbf{n} ("leakage factor") für die gesteuerte Leistung liegt im Intervall $[0 < \mathbf{n} < 1]$ und wird bestimmt durch $\mathbf{n} = 1 - \mathbf{m} \cdot \mathbf{g}$. Durch die Wahl von \mathbf{m} und γ ergibt sich der Faktor \mathbf{n} . Der "leakage factor" hat eine stabilisierende Wirkung auf den adaptiven Algorithmus. Der Preis dafür ist eine komplexere Rekursionsformel der Filtergewichte. Die konvergierende Lösung ist jedoch mit einem systematischen Fehler behaftet. Der Faktor γ ist ein Kompromiss zwischen dem systematischen Fehler der konvergierten Gewichte und der möglichen Leistungskontrolle.

Zusammenfassung

Die aufgezeichneten Mikrofon-signale werden über eine Wiedergabematrix auf beliebig positionierte Lautsprecher aufgeteilt. Die Wiedergabematrix wird durch adaptive Filter modifiziert. Dabei werden die inversen Übertragungsfunktionen der sekundären Pfade, d.h. die Eigenschaften des Wiedergaberaumes, durch die adaptiven Filter nachgebildet. Die

aufgenommenen Mikrofonensignale werden mit diesen invertierten Übertragungsfunktionen entsprechend vorgefiltert. Die Steuerung mittels adaptiver Filter ermöglicht es, zumindest theoretisch für jede beliebige Lautsprecheranordnung, an den Kontrollmikrofonen (Sensoren) den exakt gleichen Druckverlauf wie bei der Aufnahme des originalen Schallfeldes zu erhalten. Durch den Einsatz der adaptiven Filter kann man die akustischen Eigenschaften des Wiedergaberaumes, zumindest an den kontrollierten Stellen, ausblenden. Der Vorteil dieser Methode liegt zum einen in der beliebigen Anordnung der Lautsprecher und zum anderen in der Unabhängigkeit des Wiedergaberaumes, da man die Sekundärpfade der Lautsprecher - Mikrofonanordnung durch eine Messung bestimmen kann.

REALISIERUNG MITTELS AMBISONIC UND RAUMIMPULSANTWORT

Im Folgenden wird eine neue Vorgangsweise zur Schallfeldreproduktion in einem geschlossenen Raum vorgestellt. Die Aufnahme liegt bereits in einem Ambisonic Format vor. Bei der Wiedergabe im freien Schallfeld treten bei der Schallfeldreproduktion keine zusätzlichen Fehler¹⁰ auf. Bei der Wiedergabe in einem geschlossenen Raum treten jedoch im Punkt¹¹ (0,0) die Reflexionen des Wiedergaberaumes auf. Diese sind unerwünscht und können, wie im Folgenden gezeigt wird, kompensiert werden.

Mit einem Ambisonic System kann für jede "beliebige Schallquelle" (siehe *Ambisonic S.3*) ein natürliches Schallfeld rekonstruiert werden. Jede Reflexion kann als eine "Schallquelle" interpretiert und somit synthetisiert werden. Wenn aber jede Reflexion synthetisiert werden kann, dann ist es auch möglich, durch entsprechende Wahl der Ambisonic Signale, Komponenten zu generieren, die diese Reflexionen kompensieren.

Die neue Vorgangsweise zur Kompensation von Raumreflexionen wird im folgenden Abschnitt anhand eines Ambisonic Systems 1. Ordnung für den 2 dimensionalen Fall im Punkt (0,0) vorgestellt. Die Wiedergabe erfolgt mit einem kreisförmigen Lautsprecherarray bestehend aus L Lautsprechern. Durch die begrenzenden Flächen des Wiedergaberaumes kommt es zu Reflexionen. Diese Reflexionen bringen das Originalsignal mit einer entsprechenden Verzögerung, Abschwächung und Klangfarbenveränderung aus einer anderen Richtung wieder.

Ambisonic 1.Ordnung:

¹⁰ Wenn man von der Modellannahme ausgeht, dass die abgestrahlten Wellen eines Lautsprechers in einer entsprechenden Entfernung als ebene Welle betrachtet werden können und die Ordnung des Systems hoch genug gewählt ist.

¹¹ Beliebiger Punkt innerhalb des geschlossenen Raumes, der dem Mittelpunkt der Anordnung bzw. der idealen Position des Zuhörers entspricht.

Für eine Freifeld Wiedergabe berechnen sich die Lautsprecher Signale L_i aus den Ambisonic Signalen (W_{org} , X_{org} und Y_{org}) wie folgt:

$$\begin{pmatrix} L_1 \\ \vdots \\ L_L \end{pmatrix} = \begin{pmatrix} \text{Decoder} \end{pmatrix} \cdot \begin{pmatrix} W_{org} \\ X_{org} \\ Y_{org} \end{pmatrix} \quad (24)$$

Der Decoder gewichtet die originalen Drucksignale W_{org} , X_{org} und Y_{org} in Abhängigkeit von der Lautsprecherposition und bildet die entsprechenden Summen der Lautsprecherpfade (siehe Gl.(6a,b)).

Im folgenden werden die Reflexionen hinsichtlich Amplitude a_k , Verzögerung t_k und Richtung \mathbf{q}_k modelliert. Die Raumimpulsantwort des i -ten Lautsprechers h_i aus der Richtung \mathbf{f}_i kann wie folgt formuliert werden:

$$h_i(t) = \mathbf{d}_{\mathbf{f}_i}(t - t_0) + \sum_{k=1}^{+\infty} a_k \cdot \mathbf{d}_{\mathbf{q}_{i,k}}(t - t_i) \quad (25)$$

Die Raumimpulsantworten können mit einem "Soundfield microphone"¹², mit der Methode von A. Farina [5], [13] oder mit dem *Holophonie-Ambisonic System* (siehe S.8) gemessen werden und liegen somit im Ambisonic Format (1.Ordnung) vor. Somit können die Ambisonic Signale der Impulsantwort für den i -ten Lautsprecher wie folgt dargestellt werden:

$$W_{h_i}(t) = \mathbf{d}(t - t_0) + \sum_{k=1}^{+\infty} a_k \cdot \mathbf{d}(t - t_{i,k}) \quad (26a)$$

$$X_{h_i}(t) = \mathbf{d}(t - t_0) \cdot \cos(\mathbf{f}_i) + \sum_{k=1}^{+\infty} a_k \cdot \mathbf{d}(t - t_{i,k}) \cdot \cos(\mathbf{q}_{i,k}) \quad (26b)$$

$$Y_{h_i}(t) = \mathbf{d}(t - t_0) \cdot \sin(\mathbf{f}_i) + \sum_{k=1}^{+\infty} a_k \cdot \mathbf{d}(t - t_{i,k}) \cdot \sin(\mathbf{q}_{i,k}) \quad (26c)$$

Die ersten Terme der Gl. (26a bis 26c) beschreiben die direkte Schallkomponente und werden für die folgenden Betrachtungen entfernt. Für den Anteil der Reflexionen unter der Berücksichtigung, dass diese zur Eliminierung der Raumreflexionen verwendet werden (negatives Vorzeichen) ergeben sich folgende FIR Filter ($t \rightarrow n$):

$$W_{h_{i,r}}(n) = - \sum_{k=1}^{+\infty} a_k \cdot \mathbf{d}(n - m_{i,k}) \quad (27a)$$

$$X_{h_{i,r}}(n) = - \sum_{k=1}^{+\infty} a_k \cdot \mathbf{d}(n - m_{i,k}) \cdot \cos(\mathbf{q}_{i,k}) \quad (27b)$$

¹² Dieses Mikrofon wurde von SoundField in Zusammenarbeit mit Gerzon entwickelt.

$$Y_{h_{i,r}}(n) = -\sum_{k=1}^{+\infty} a_k \cdot \mathbf{d}(n - m_{i,k}) \cdot \sin(\mathbf{q}_{i,k}) \quad (27c)$$

Im Folgenden wird die Vorgangsweise der Kompensierung der räumlichen Filterung anhand eines Lautsprechersignals gezeigt. Das Lautsprechersignal des i ten Lautsprechers für das freie Schallfeld lautet (siehe Gl. (24)):

$$L_i(n) = a_i \cdot W(n)_{org} + b_i \cdot X(n)_{org} + c_i \cdot Y(n)_{org} \quad (28)$$

Die Koeffizienten a_i , b_i und c_i in Gl. (28) sind die Elemente der i . Zeile der Ambisonic Decodermatrix. Die Kompensationssignale für die Reflexionen werden durch Faltung¹³ der Lautsprechersignale mit den dazugehörigen Ambisonic Raumimpulsantworten berechnet.

$$W_{L_{i,r,comp}}(n) = L_i(n) \otimes W_{h_{i,r}}(n) \quad (29a)$$

$$X_{L_{i,r,comp}}(n) = L_i(n) \otimes X_{h_{i,r}}(n) \quad (29b)$$

$$Y_{L_{i,r,comp}}(n) = L_i(n) \otimes Y_{h_{i,r}}(n) \quad (29c)$$

Somit ergeben sich die Lautsprechersignale L^* aus den Ambisonic Signalen wie folgt unter Berücksichtigung der Reflexionen von L_i ...

$$\begin{pmatrix} L_1^* \\ \vdots \\ L_L^* \end{pmatrix} = \begin{pmatrix} Decoder \\ \vdots \\ Decoder \end{pmatrix} \cdot \begin{pmatrix} W_{org} \\ X_{org} \\ Y_{org} \end{pmatrix} + \begin{pmatrix} Decoder \\ \vdots \\ Decoder \end{pmatrix} \cdot \begin{pmatrix} W_{L_{i,r,comp}} \\ X_{L_{i,r,comp}} \\ Y_{L_{i,r,comp}} \end{pmatrix} \quad (30a)$$

$$\begin{aligned} &= \begin{pmatrix} Decoder \\ \vdots \\ Decoder \end{pmatrix} \cdot \begin{pmatrix} W_{org} \\ X_{org} \\ Y_{org} \end{pmatrix} + \begin{pmatrix} Decoder \\ \vdots \\ Decoder \end{pmatrix} \cdot \begin{pmatrix} W_{h_{i,r}}(n) \otimes (a_i W_{org} + b_i X_{org} + c_i Y_{org}) \\ X_{h_{i,r}}(n) \otimes (a_i W_{org} + b_i X_{org} + c_i Y_{org}) \\ Y_{h_{i,r}}(n) \otimes (a_i W_{org} + b_i X_{org} + c_i Y_{org}) \end{pmatrix} \\ &= \begin{pmatrix} Decoder \\ \vdots \\ Decoder \end{pmatrix} \cdot \left[\mathbf{d}(n) \cdot I \cdot \begin{pmatrix} W_{org} \\ X_{org} \\ Y_{org} \end{pmatrix} + \begin{pmatrix} W_{h_{i,r}}(n) \\ X_{h_{i,r}}(n) \\ Y_{h_{i,r}}(n) \end{pmatrix} \otimes \begin{pmatrix} a_i W_{org} + b_i X_{org} + c_i Y_{org} \\ a_i W_{org} + b_i X_{org} + c_i Y_{org} \\ a_i W_{org} + b_i X_{org} + c_i Y_{org} \end{pmatrix} \right] \\ &= \begin{pmatrix} Decoder \\ \vdots \\ Decoder \end{pmatrix} \cdot \left[\begin{pmatrix} \mathbf{d}(n) & 0 & 0 \\ 0 & \mathbf{d}(n) & 0 \\ 0 & 0 & \mathbf{d}(n) \end{pmatrix} + \begin{pmatrix} a_i W_{h_{i,r}}(n) + b_i W_{h_{i,r}}(n) + c_i W_{h_{i,r}}(n) \\ a_i X_{h_{i,r}}(n) + b_i X_{h_{i,r}}(n) + c_i X_{h_{i,r}}(n) \\ a_i Y_{h_{i,r}}(n) + b_i Y_{h_{i,r}}(n) + c_i Y_{h_{i,r}}(n) \end{pmatrix} \right] \otimes \begin{pmatrix} W_{org} \\ X_{org} \\ Y_{org} \end{pmatrix} \end{aligned}$$

¹³ Für die Faltungsoperation wird in den folgenden Gleichungen das Symbol \otimes verwendet.

...und man erhält bei der Berücksichtigung der Reflexionen, die durch den Lautsprecher L_i verursacht wurden, folgendes Ergebnis:

$$\begin{pmatrix} L_1^* \\ \vdots \\ L_L^* \end{pmatrix} = \begin{pmatrix} \text{Decoder} \end{pmatrix} \begin{pmatrix} \mathbf{d}(n) + a_i W_{h_{i,r}}(n) & b_i W_{h_{i,r}}(n) & c_i W_{h_{i,r}}(n) \\ a_i X_{h_{i,r}}(n) & \mathbf{d}(n) + b_i X_{h_{i,r}}(n) & c_i X_{h_{i,r}}(n) \\ a_i Y_{h_{i,r}}(n) & b_i Y_{h_{i,r}}(n) & \mathbf{d}(n) + c_i Y_{h_{i,r}}(n) \end{pmatrix} \otimes \begin{pmatrix} W_{org} \\ X_{org} \\ Y_{org} \end{pmatrix} \quad (30b)$$

Für 2 dimensionales Ambisonic 1.Ordnung erhält man bei der Berücksichtigung der Reflexionen die durch die L Lautsprecher hervorgerufen wurden die kompensierten Lautsprechersignale L_{comp} :

$$\begin{pmatrix} L_{1,comp} \\ \vdots \\ L_{L,comp} \end{pmatrix} = \begin{pmatrix} \text{Decoder} \end{pmatrix} \begin{pmatrix} \mathbf{d}(n) + \sum_{i=1}^L a_i W_{h_{i,r}} & \sum_{i=1}^L b_i W_{h_{i,r}}(n) & \sum_{i=1}^L c_i W_{h_{i,r}}(n) \\ \sum_{i=1}^L a_i X_{h_{i,r}}(n) & \mathbf{d}(n) + \sum_{i=1}^L b_i X_{h_{i,r}}(n) & \sum_{i=1}^L c_i X_{h_{i,r}}(n) \\ \sum_{i=1}^L a_i Y_{h_{i,r}}(n) & \sum_{i=1}^L b_i Y_{h_{i,r}}(n) & \mathbf{d}(n) + \sum_{i=1}^L c_i Y_{h_{i,r}}(n) \end{pmatrix} \otimes \begin{pmatrix} W_{org} \\ X_{org} \\ Y_{org} \end{pmatrix} \quad (31)$$

Durch diese Vorgangsweise benötigt man, abhängig von der Anzahl N der Übertragungskanäle (realisierte Ordnung M von Ambisonic), N^2 (2 dimensional $(2M+1)^2$ bzw. 3 dimensional $(M+1)^4$) Filter zur Eliminierung der Raumreflexionen des Wiedergaberaumes. Die Anzahl der Filter ist jedoch unabhängig von der Anzahl der Lautsprecher.

Bei der Realisierung mittels adaptiver Filter werden $L \cdot M$ ($=L \cdot S$) Filter benötigt, d.h. die Anzahl der benötigten Filter ist abhängig von der Anzahl der Aufnahme- (Sensor-) Mikrofone und der verwendeten Lautsprecher.

Zusammenfassung

Es wurde eine neuer Ansatz zur Schallfeldreproduktion in einem geschlossenen Raum mit einem Ambisonic System und den gemessenen Raumimpulsantworten vorgestellt. Der neue Vorschlag zur Kompensation von Raumreflexionen mit einem Ambisonic System beruht darauf, dass jede Reflexion als eine Schallquelle interpretiert werden kann. Da jede Reflexion synthetisiert werden kann, ist es auch möglich, durch entsprechende Wahl der Ambisonic Signale, Komponenten zu generieren, die diese Reflexionen in Summe eliminieren. Die Anzahl der dafür benötigten Filter ist unabhängig von der Anzahl der Lautsprecher. Somit ist dieser Ansatz im Vergleich mit der "adaptiven Filterrealisierung" recheneffizienter und stellt dahingehend eine Verbesserung dar.

AUSBLICK

Simulation in MATLAB:

Adaptive Filterrealisierung vs. Realisierung mit Ambisonic & Raumimpulsantwort.

- *Vergleich im freien Schallfeld*

Inwieweit stimmen die reproduzierten Schallfelder der beiden Realisierungen mit dem aufgenommenen Schallfeld, für einen bestimmten Bereich, überein?

Untersuchung der Einflüsse von Parameteränderungen wie Anzahl und Anordnung der Sensoren und Lautsprecher, Entfernung der Lautsprecher, etc..

- *Vergleich in einem einfachen, geschlossenen Raum*

(wie oben)

Realisierung eines Systems

Aufnahmen (Ambisonic aus Holophonie)

Hörvergleiche

Literaturverzeichnis:

- [1] Ch. Huygens, 1690.
- [2] M. Poletti, "The Design of Encoding Functions for Sterophonic and Polyphonic Sound Systems", J. Audio Eng. Soc., Vol. 44, No. 11, (1996 November).
- [3] E.N.G. Verheijen, "Sound Reproduction by Wave Field Synthesis." Ph.D. Thesis, Delft University of Technology, Netherlands, 1998.
- [4] J.S. Bamford, "An Analysis of Ambisonic Sound Systems of First and Second Order", Thesis presented to the University of Waterloo, Waterloo, Ontario, Canada, (1995).
- [5] A. Farina, E. Ugonotti, "Software Implementation of B-Format Encoding and Decoding", Preprint of the 104th AES Convention, Amsterdam, (1998, May).
- [6] M.A. Gerzon, "Surround Sound Psychoacoustics", Wireless World, Vol. 80, pp. 483-486, (1974, December).
- [7] M.A. Gerzon, "General Metatheory of Auditory Localisation", presented at the 92nd AES Convention, Vienna, (1992, March).
- [8] M.A. Gerzon, "The Optimum Choice of Surround Sound Encoding Specification", Preprint 1199 of the 56th AES Convention, Paris, (1977, March).
- [9] M.A. Gerzon, "Optimal Reproduction Matrices for Multispeaker Stereo", Preprint 3180 of the 91st AES Convention, New York, (1991, October).
- [10] R. Nicol, M. Emerit, "3D-Sound Reproduction over an extensive Listening Area. A Hybrid Method derived from Holophony and Ambisonic.", The Proceedings of the AES 16th International Conference on Spatial Sound Reproduction, pp.436-453, Rovaniemi, Finland, (1999, April).
- [11] S. M. Kuo, D. R. Morgan, "Active Noise Control", Proceedings of the IEEE, Vol. 87, No. 6, (1999, June).
- [12] S. D. Stearns, "Fundamentals of Adaptive Signal Processing", '*Advanced Topics in Signal Processing*', chapter 5, p.246-288, edited by J.S: Lim, A. V. Oppenheim, ISBN 0-13-013129-6 025 (1988).
- [13] A. Farina, L. Tronchin, "3D Impulse Response measurements on S. Maria dl Fiore Church, Florence, Italy", Proceedings of ICA98, International Conference on Acoustics, Seattle, (1998, June).