

Studie zur Anwendung von DirAC höherer Ordnung auf Eigenmikesignale

Patrick Ziegler
Toningenieurprojekt

Betreuung: Dr. Franz Zotter
Graz, 10. März 2016



institut für elektronische musik und akustik



Zusammenfassung

Die Tonaufnahme mit ambisonischen Mikrofonen höherer Ordnung erlaubt eine relativ genaue Richtungslokalisierung von Schallquellen. Daher bieten diese Mikrofone interessante Möglichkeiten in der Audiosignalverarbeitung, wobei mehrfache Richtungsbestimmung eine davon ist. In dieser Arbeit wird ein Verfahren zur Mehrfachrichtungsdetektion mit dem Eigenmike[®] behandelt. Die 32 Kanäle des Eigenmike[®] erlauben im Gegensatz zu Ambisonics 1. Ordnung, wo eine Richtung pro Zeit-/Frequenzeinheit zugeordnet werden kann, eine Zerlegung in Ambisonics 4. Ordnung, was eine Mehrfachrichtungslokalisierung ermöglicht. Basierend auf der Zerlegung des Tonsignals in mehrere Richtungssektoren wird die Lokalisation in Anlehnung an ein bekanntes Verfahren für Ambisonics 1. Ordnung, namens Directional Audio Coding, durchgeführt. Hierbei kann durch die Berechnung eines Intensitätsvektors in jedem Sektor direkt eine Quellenrichtung bestimmt werden und in Summe aller Sektoren ergibt sich eine Mehrfachrichtungslokalisierung. Die Arbeit beschreibt verschiedene Tests des Algorithmus und seine Eigenschaften bei Anwendung auf Raumimpulsantworten und aufgenommene Ambisonic-Signale.

Abstract

The sound recording with ambisonic microphones higher order allows a relatively precise directional localization of sound sources. Therefore, these microphones offer interesting possibilities in audio signal processing, where multiple directional determination is one of them. In this work a method for multiple directional detection is shown with the Eigenmike[®]. The 32 channels of the Eigenmike[®] allow to detect multiple directions by a decomposition in Ambisonics 4th order. Whereas Ambisonics 1st order only one direction per time- / frequency bin can be assigned. Based on the decomposition of the audio signal into several sectors, the localization is performed according to a known method for Ambisonics 1st order, called Directional Audio Coding. Here, by calculating an intensity vector in each sector a source direction can be determined directly and the sum of all sectors results in a multiple directional localization. The paper describes various tests of the algorithm and its properties when it is applied to room impulse responses and recorded Ambisonics signals.

Inhaltsverzeichnis

| | | |
|----------|---|-----------|
| 1 | Einleitung | 4 |
| 2 | Testset aus bekannten Quellrichtungen | 5 |
| 3 | Berechnung der Ambisonicssignale | 7 |
| 3.1 | Radialfilter | 7 |
| 3.2 | Interpolierte Darstellung | 9 |
| 4 | Richtungsanalyse | 11 |
| 4.1 | Richtungsanalyse 1. Ordnung | 11 |
| 4.2 | Richtungsanalyse höherer Ordnung | 13 |
| 4.2.1 | Aufteilung in Winkelbereiche (Sektoren) | 13 |
| 4.2.2 | Bestimmung der Gewichtungsfaktoren | 14 |
| 4.2.3 | Gewichtung der Kugelflächenfunktionen | 16 |
| 4.3 | Sektorendefinition | 18 |
| 5 | Auswertung des Algorithmus | 21 |
| 5.1 | Beispiel: Direktschall | 21 |
| 5.2 | Beispiel: Signal mit Reflexionskomponente | 21 |
| 5.3 | Beispiel: Diffuses Schallfeld | 24 |
| 6 | Ambisonics Kodierung | 26 |
| 7 | Schlussfolgerungen | 28 |
| 7.1 | Objektive Betrachtung | 28 |
| 7.2 | Subjektive Betrachtung | 29 |

1 Einleitung

Viele aktuelle Anwendungen in der Audiosignalverarbeitung erfordern beispielsweise zur Wellenfeldanalyse- oder Synthese immer detailliertere Informationen über das Schallfeld im Raum. Dabei spielen exakte Richtungslokalisationen eine große Rolle. Da in einem komplexen Schallfeld aber auch frühe Reflexionen entscheidenden Einfluss haben können, reicht eine einfache Richtungsbestimmung des Direktschalls in einigen Fällen nicht aus. In dieser Arbeit wird deshalb das Hauptaugenmerk auf eine Mehrfachrichtungsdetektion mithilfe ambisonischer Aufnahmen höherer Ordnung gelegt.



Abbildung 1 – MH-Acoustics Eigenmike[®] EM32 [EM3]

Als Grundlage für die Aufnahme dienen Impulsantworten des IEM-CUBE [ZRS02], einem Aufführungsraum mit 24 über einer Halbkugeloberfläche angeordneten Lautsprechern. Zur Aufnahme der Signale wurde das MH-Acoustics Eigenmike[®] EM32 [EM3] mit 32 kugelförmig angeordneten Mikrofonen, wie in Abbildung 1 zu sehen, verwendet. Die gemessenen Raumimpulsantworten dienen als Basisdaten zur Berechnung des sphärischen Wellenspektrums. Nach einer Anpassung der Signale durch räumliche Filterung kann anschließend die Richtungsdetektion analog zum Verfahren Directional Audio Coding [PP11] erfolgen. In den folgenden Abschnitten wird dieser Algorithmus vorgestellt. Im Anschluss daran wird das Verfahren bewertet und die Qualität und Genauigkeit der Mehrfachrichtungsdetektion analysiert.

2 Testset aus bekannten Quellrichtungen

Die verwendeten Daten zur Richtungsdetektion sind Aufnahmen aus dem IEM-CUBE. In der vorliegenden Messanordnung wurden nacheinander exponentielle Sinus-Sweeps mit einer Dauer von einer Sekunde an allen 24 Lautsprechern ausgespielt. Die Anordnung der Lautsprecher ist in Abbildung 2 grafisch dargestellt.

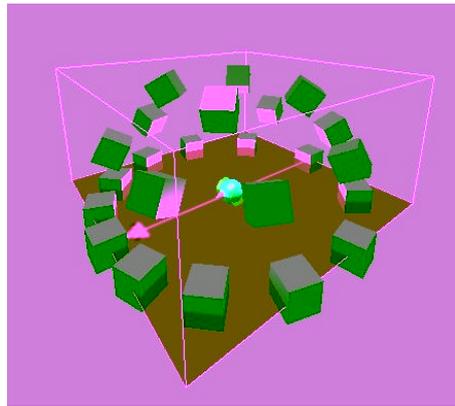


Abbildung 2 – Lautsprecheranordnung im IEM-CUBE [ZRS02]

Für jeden der 24 ausgespielten Sinus-Sweeps wurden die Antwortsignale mit dem 32-kanaligen Eigenmike[®] aufgenommen, sodass man $24 \times 32 = 768$ Antwortsignale erhält. Zur Berechnung der Raumimpulsantworten wird diese Signalmatrix mit einer schnellen Fouriertransformation in den Frequenzbereich transformiert. Gleichzeitig unterzieht man auch den originalen Sweep einer FFT. Mithilfe der einfachen Beziehung in Gleichung 1 können die Raumimpulsantworten des IEM-CUBE für jeden Lautsprecher berechnet werden, wobei die Abtastfrequenz $f_s = 44100\text{Hz}$ beträgt

$$\mathbf{h}(n) = \text{iFFT} \left\{ \frac{\text{FFT}\{\mathbf{y}(n)\}}{\text{FFT}\{x(n)\}} \right\}. \quad (1)$$

In obiger Gleichung repräsentiert $x(n)$ den Sinus-Sweep, $\mathbf{y}(n)$ die Signalmatrix der 24×32 Sweepantworten des Eigen-mike[®] und $\mathbf{h}(n)$ die Raumimpulsantwortmatrix zum Abtastzeitpunkt n . Durch die Fouriertransformation mit einer hohen Anzahl an Punkten erhält man in den ersten Samples viele unbrauchbare Werte, die weitere Berechnungen verfälschen können und entsprechend vernachlässigt werden. Auch spätere Samples sind für die Richtungsbestimmung in dieser Arbeit nicht weiter relevant, da lediglich frühe Reflexionen zum Richtungshören beitragen. Es werden deshalb etwa 50ms der gemessenen Impulsantworten zur weiteren Berechnung herangezogen. In Abschnitt 6 wird ein Verfahren zur Anwendung des Algorithmus auf kontinuierliche Tonaufnahmen vorgestellt, wobei die Aufnahmen in Ambisonicssignale kodiert werden.

Um ein einzelnes Schallereignis (Direktschall eines Lautsprechers) für die Richtungsbestimmung herauszulösen, werden sämtliche Impulsantworten mit einer Fensterfunktion zu gewichtet. An den letzten Samples einer Impulsantwort dient ein $\cos^2(n)$ -Fenster zum weichen Ausblenden des Signals. Zum Einblenden wird ein $\sin^2(n)$ -Fenster verwendet [Har78]. Da allerdings aufgrund der geometrischen Anordnungen die Impulse zu unterschiedlichen Zeitpunkten an den einzelnen Mikrofonen auftauchen, legt man das Maximum des $\sin^2(n)$ -Fensters an den Abtastzeitpunkt, an dem ein bestimmter Schwellwert das erste mal überschritten wird. Dieser Wert wird durch das quadratische Mittel
$$h_{thres} = \sqrt{1/N \cdot \sum_{n=0}^N \mathbf{h}^2(n)}$$
 aller Impulsantworten $\mathbf{h}(n)$ bestimmt, wie in Abbildung 3 durch die horizontal gestrichelte rote Linie verdeutlicht wird. Dadurch stellt man sicher, dass auch der Direktschall vollständig in die Berechnungen mit einfließt.

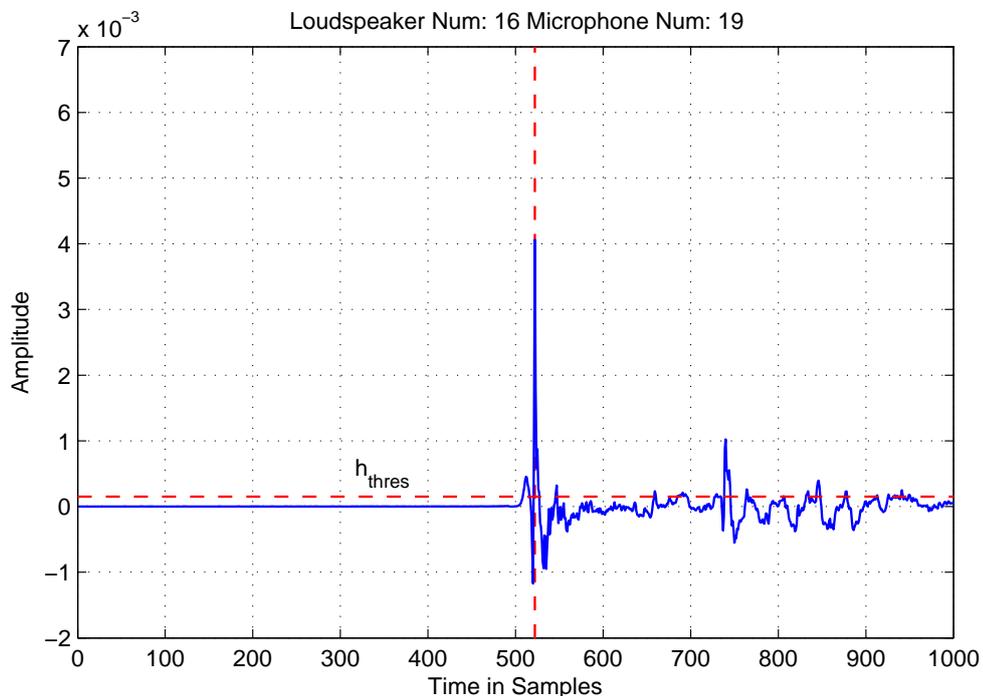


Abbildung 3 – Impulsantwort im Mikrofonkanal 19, Lautsprecher 16 mit h_{thres}

Betrachtet man das Beispiel der Impulsantwort in Abbildung 3 genauer, erkennt man gut den teilweise starken Einfluss der frühen Reflexionen. Während für manche Mikrofone der Direktschall die höchste Amplitude aufweist, sind bei anderen Mikrofonen die frühen Reflexionen sogar stärker ausgeprägt als der Direktschall. Das bedeutet, dass das Mikrofon für gewisse Frequenzen eine sehr hohe Abschattung aufweist, sofern der Direktschall zu diesem Zeitpunkt auf der gegenüberliegenden Seite des Mikrofons ankommt. Somit wirkt die Reflexion von einer nahegelegenen Grenzfläche direkt auf das Mikrofon und der Ausschlag ist höher. Zur Richtungsanalyse ist diese Darstellung allerdings zu unscharf bei tiefen Frequenzen. Erst durch eine Zerlegung in Kugelflächenfunktionen und die resultierende Berechnung eines Ambisonicssignals höherer Ordnung wird eine Richtungsbestimmung ermöglicht.

3 Berechnung der Ambisonicssignale

Der erste Schritt zur Ambisonics Enkodierung der Eigenmikesignale ist die Zerlegung in Kugelflächenfunktionen $Y_l^m(\varphi, \vartheta)$, wobei φ und ϑ jeweils ein Winkelpaar eines der 32 Mikrofone darstellen. Der Index l beschreibt die Ordnung von 0 bis L und m jeweils den Grad von $-l$ bis $+l$. Aus den 32 Kanälen des Eigenmike[®] kann man ein Ambisonicssignal der maximalen Ordnung $L = 4$ berechnen. Aus der Ungleichung $32 \geq (L + 1)^2 = 25$ ist damit die maximale Anzahl an Ambisonicskanälen festgelegt. Der Vektor \mathbf{y}_L^T mit der Länge $(L + 1)^2$ beschreibt alle Kugelflächenfunktionen an einem Winkelpaar (φ, ϑ) bis zur Ordnung L folgendermaßen:

$$\mathbf{y}_L^T(\varphi, \vartheta) = [Y_0^0(\varphi, \vartheta) Y_1^{-1}(\varphi, \vartheta) Y_1^0(\varphi, \vartheta) Y_1^1(\varphi, \vartheta) \dots Y_L^L(\varphi, \vartheta)]. \quad (2)$$

Die Gleichung 3 für die Berechnung des sphärischen Wellenspektrums $\boldsymbol{\psi}(n)$ zerlegt die 32 gemessenen Schalldrucksignale $\mathbf{p}(n) = [\mathbf{p}_1(n), \dots, \mathbf{p}_{32}(n)]$ des Eigenmike[®] durch Einpassung der abgetasteten Kugelflächenfunktionen

$$\boldsymbol{\psi}(n) = \begin{pmatrix} \mathbf{y}_L^T(\varphi_1, \vartheta_1) \\ \mathbf{y}_L^T(\varphi_2, \vartheta_2) \\ \vdots \\ \mathbf{y}_L^T(\varphi_{32}, \vartheta_{32}) \end{pmatrix}^\dagger \mathbf{p}(n) = \mathbf{Y}_L^\dagger \mathbf{p}(n) \quad (3)$$

Unter der Annahme, dass sich \mathbf{Y}_L^\dagger bilden lässt (Pseudoinversion), kann der an 32 Abtastpunkten gemessene Schalldruck beliebig fein interpoliert werden. Als Testsignale lassen sich für $\mathbf{p}(n)$ die Impulsantworten $\mathbf{h}(n)$ aus dem vorherigen Abschnitt einsetzen.

3.1 Radialfilter

Um die ambisonische Surroundwiedergabe zu ermöglichen, sind so genannte Radialfilter erforderlich [ZL15]. Niederfrequente Signale bilden sich aufgrund der Abmessungen des Mikrofons auf dessen Oberfläche beinahe mit konstantem Schalldruck ab, weshalb eine Verstärkung für höhere Ordnungen der Kugelflächenfunktionen notwendig ist. Die Komponenten höherer Ordnung werden dabei so angehoben, dass schwach vorhandene Unterschiede bei tiefen Frequenzen verstärkt werden.

Die Frequenzgänge der verwendeten Radialfilter aus [BPF11] sind in Abbildung 4 illustriert. Um die Ambisonicssignale für die Richtungsanalyse zu berechnen, wird das sphärische Wellenspektrum mit den Filterfunktionen der Filterbank gefaltet. Das Ambisonicssignal $\boldsymbol{\phi}(n)$ ergibt sich, kompakt geschrieben, zu

$$\boldsymbol{\phi}(n) = \boldsymbol{\psi}(n) * \mathbf{h}_{\text{rad}}(n). \quad (4)$$

wobei $\mathbf{h}_{\text{rad}}(n)$ einen Vektor von $(L+1)^2$ Radialfiltern für jede Ordnung der Kugelflächenfunktionen beschreibt

$$\mathbf{h}_{rad}(n) = [h_{rad0}(n) \ h_{rad1}(n) \ h_{rad1}(n) \ h_{rad1}(n) \ \dots \ h_{rad4}(n)]^T. \quad (5)$$

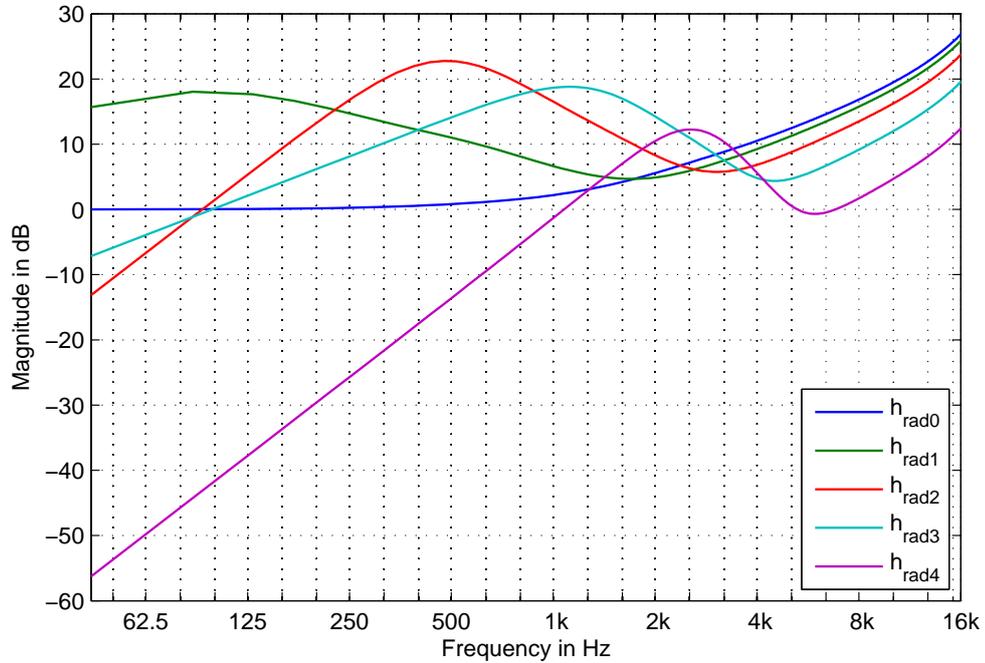


Abbildung 4 – Amplitudenfrequenzgang der Radialfilter $h_{rad0} \dots h_{rad4}$

Zur Unterdrückung von räumlichem Aliasing, welches durch die zu geringe Mikrofonanzahl zur räumlichen Abtastung hoher Frequenzen entsteht, wird zusätzlich zu den Radialfiltern ein Bandpassfilter mit den Grenzfrequenzen $50\text{Hz} - 3000\text{Hz}$ eingesetzt. Die Faltungsoperation aus Gleichung 4 erweitert sich zu

$$\phi(n) = \psi(n) * \mathbf{h}_{rad}(n) * h_{BP}(n). \quad (6)$$

Der strukturelle Aufbau dieser Filterbank ist in Abbildung 5 zu sehen. Pro Ordnung werden die Signale mit den jeweiligen Radialfiltern gefiltert. Der Bandpass wird für das komplette sphärische Spektrum gleich verwendet. Das gefilterte sphärische Wellenspektrum (Ambisonicssignale) wird im nächsten Unterpunkt zur räumlichen Darstellung der Impulsantworten verwendet.

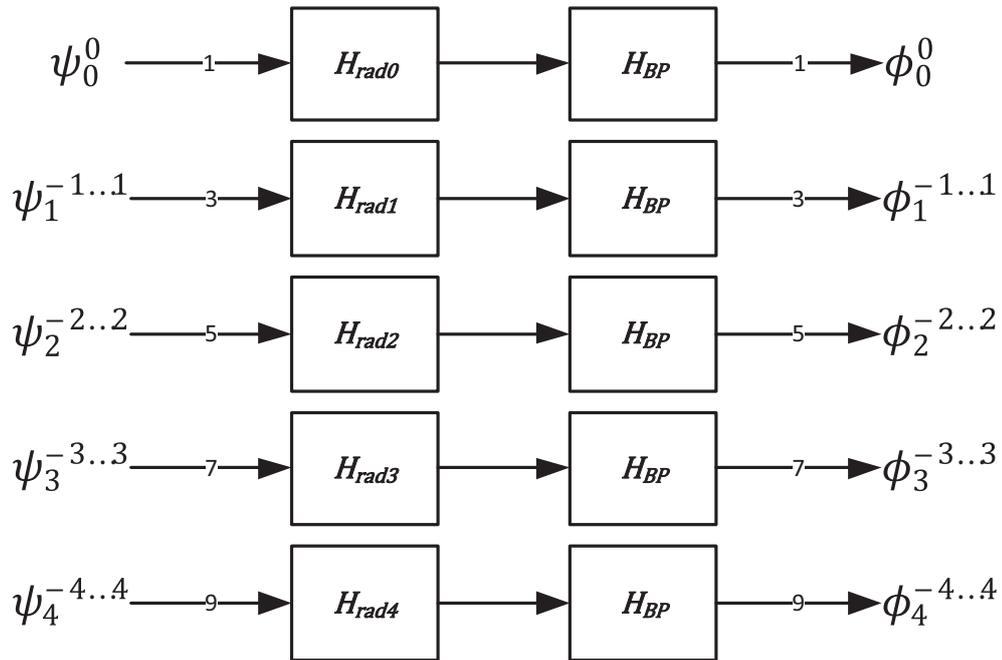


Abbildung 5 – Struktureller Aufbau der Radialfilterbank mit Bandpassfilterung

3.2 Interpolierte Darstellung

Eine interpolierte Darstellung mit Kugelflächenfunktionen mithilfe einer gleichwinklig in (φ, ϑ) abgetasteten Kugeloberfläche [Zot14] ist in Abbildung 6 skizziert.

Die Reflexionen eines Lautsprechers beziehungsweise die Abschattung des Eigenmike[®] sind in der Abbildung im Gegensatz zur Darstellung als Impulsantwort nun einer definierten Richtung zugeordnet und werden durch den Anteil des omnidirektionalen W -Signals in der Amplitude gewichtet. Die Schalldruckverteilung ergibt sich nach [Zot14] formal aus den mit $\phi(n)$ überlagerten, hochabgetasteten Kugeloberflächen, kurz, zu

$$p(n, \varphi, \vartheta) = \mathbf{y}_L(\varphi, \vartheta) \cdot \phi(n). \quad (7)$$

Für eine verbesserte Darstellung wird die Schalldruckverteilung in ein normiertes, logarithmisches Maß umgerechnet und eine maximale Dynamik von 30dB zugelassen.

$$P(n, \varphi, \vartheta) = 20 \cdot \log |p(n, \varphi, \vartheta)| \text{ dB} \quad (8)$$

$$R(n, \varphi, \vartheta) = 1 + \frac{P(n, \varphi, \vartheta) - P_{max}(n)}{30} \text{ dB} \quad (9)$$

wobei $P_{max}(n)$ dem maximalen Pegel der logarithmierten Schalldruckverteilung $P(n, \varphi, \vartheta)$ über alle Winkel zum Abtastzeitpunkt n entspricht.

Zur Darstellung der Ergebnisse als Mollweide Projektion werden die winkelbezogenen Signale wie folgt umgerechnet:

$$\begin{aligned} w &= R(n, \varphi, \vartheta) \\ x &= \varphi \cdot \sin^{0.4} \vartheta. \\ y &= \vartheta \end{aligned} \quad (10)$$

In Gleichung 10 stellt y den vertikalen Verlauf von 0° bis 90° dar. Abhängig von ϑ wird x durch $\sin^{0.4}$ verzerrt und beschreibt die horizontale Achse von -180° bis 180° . Bestimmt durch $R(\varphi, \vartheta)$ für einen Abtastzeitpunkt n steht w für den Farbwert (Pegel von -30dB bis 0dB .) abhängig von ϑ und φ .

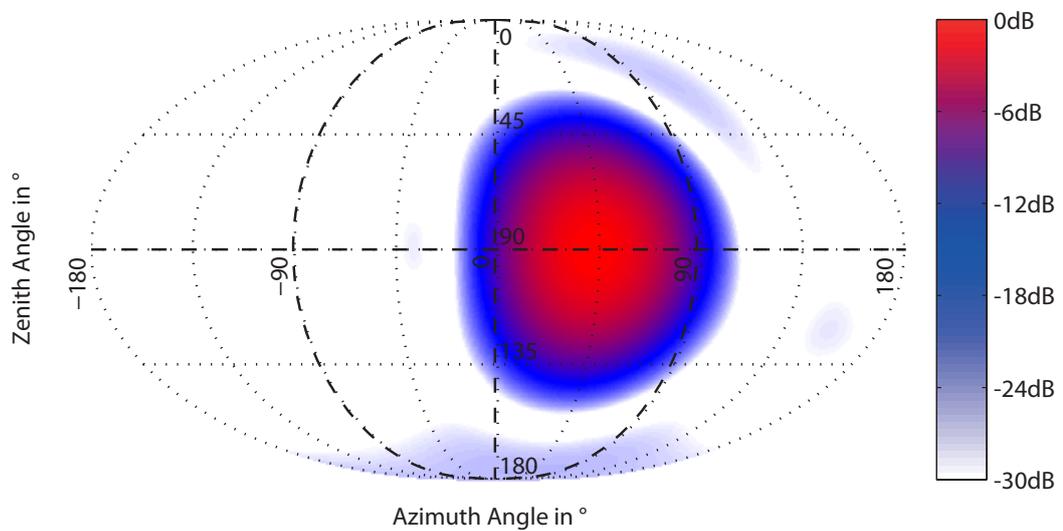


Abbildung 6 – Mollweide Projektion der interpolierten Ambisonicssignale am Beispiel des Direktschalls des Lautsprechers 4 im IEM-CUBE

4 Richtungsanalyse

4.1 Richtungsanalyse 1. Ordnung

Die Richtungsbestimmung 1. Ordnung verwendet den Algorithmus des bereits bekannten Verfahrens Directional Audio Coding [Puk07]. Hier dienen die Ambisonicssignale 1. Ordnung $X(n), Y(n), Z(n)$ als Daten zur Berechnung des Intensitätsvektors und somit zur Richtungsbestimmung. Das $W(n)$ -Signal dient zur Bestimmung des Richtungszeichens $\pm 180^\circ$ und kann zusätzlich als Indikator für die Signalstärke verwendet werden (analog zu Abschnitt 3.2). Die Signale 0. und 1. Ordnung können direkt aus $\phi(n)$ extrahiert werden. Durch die einfache Berechnung des Intensitätsvektors in Gleichung 11 können mit $\mathbf{X}(n) = [X(n) Y(n) Z(n)]^T$ die Winkelpaare der Quellrichtungen direkt aus dem Richtungskontinuum bestimmt werden.

$$\mathbf{u}(n) = -\frac{\mathbf{I}_a(n)}{\|\mathbf{I}_a(n)\|} = -\frac{\mathbf{X}(n) \cdot \text{sign } W(n)}{\|\mathbf{X}(n)\|} = \begin{bmatrix} \cos \vartheta \cos \varphi \\ \sin \vartheta \cos \varphi \\ \sin \varphi \end{bmatrix}. \quad (11)$$

Daraus ergeben sich die Winkel der Quellrichtungen zu:

$$\boldsymbol{\theta}_{DOA}(n) = \begin{bmatrix} \arccos(u_X(n)) \\ \arctan2(u_Y(n), u_Z(n)) \end{bmatrix} \quad (12)$$

Die Indizes beschreiben jeweils den Richtungsanteil in X, Y bzw. Z -Richtung. Die Funktion $\arctan2$ wird „Arkustangens mit zwei Argumenten“ genannt und liefert den vollständigen Polarwinkel ϑ im Wertebereich $[-\pi; +\pi]$.

Als Beispiel wird der Direktschall (z.B. bei Abtastzeitpunkt $n = 974$) des 13. Lautsprechers gezeigt. Das Richtungssignal wird mit den Gleichungen 11 und 12 berechnet. In Abbildung 7 erkennt man die detektierte Richtung anhand des blau eingezeichneten Kreuzes. Da es sich beim Direktschall um eine klar zugeordnete Richtung ohne nennenswerte Reflexionen handelt (keine Nebenkeulen erkennbar), funktioniert die Richtungsbestimmung 1. Ordnung sehr gut. Der Intensitätsvektor zeigt exakt in die Richtung des stärksten Signalanteils.

Ist jedoch zusätzlich zum Direktschall eine relativ hohe Reflexion im Signal enthalten, funktioniert die Richtungsdetektion 1. Ordnung nicht mehr exakt. Der Intensitätsvektor verschiebt sich durch die Reflexion aus der Richtung des Maximums einer Keule heraus, wie beispielsweise in Abbildung 8 zu sehen. Da diese Reflexionen aber häufig auftreten können und maßgeblichen Einfluss auf die Wiedergabe haben, ist es wichtig auch hier eine zuverlässige Richtungsbestimmung verwenden zu können. Im nächsten Unterpunkt wird deshalb eine Variante zur Richtungsanalyse höherer Ordnung vorgestellt, die die Verschiebung des Intensitätsvektors durch eine Mehrfachdetektion ersetzt.

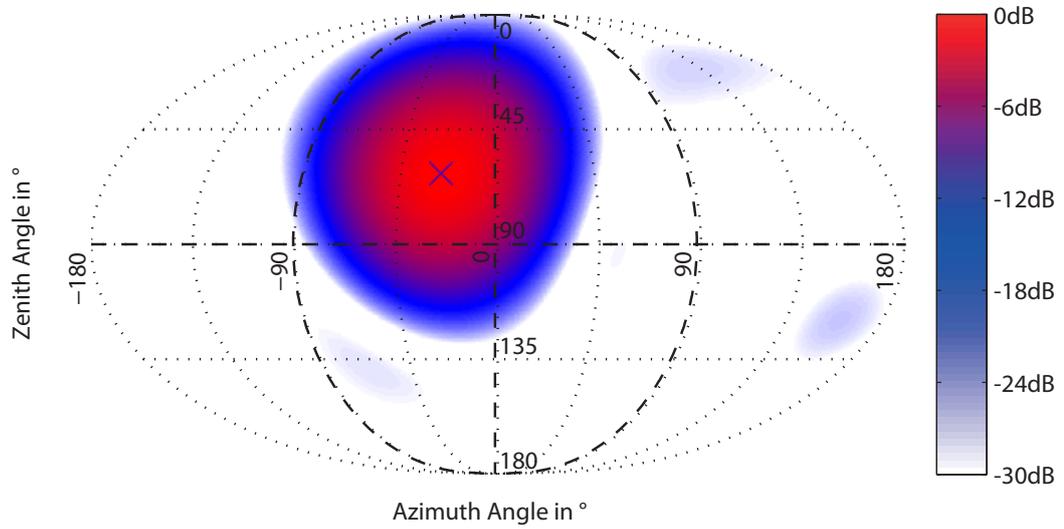


Abbildung 7 – Richtungsdetektion 1. Ordnung am Direktschall des 13. Lautsprechers

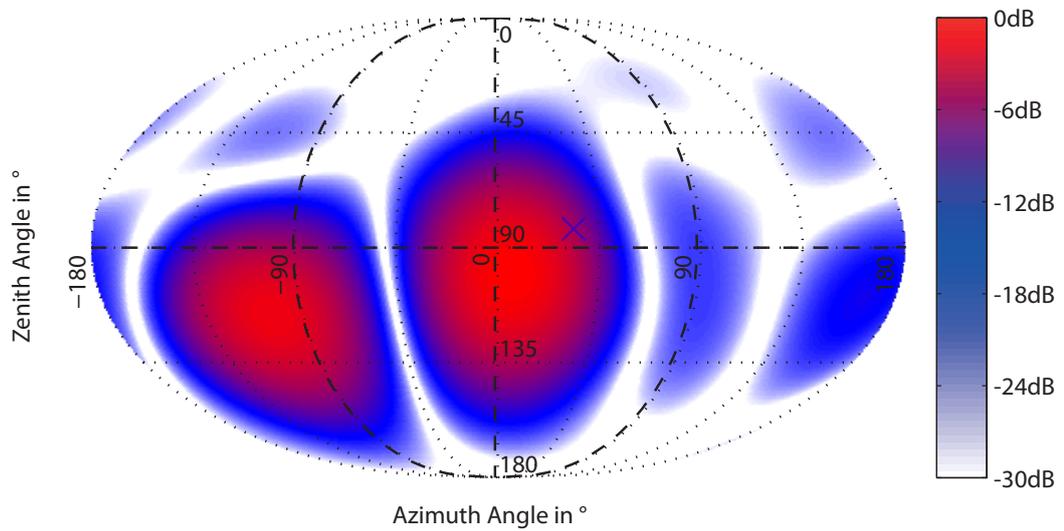


Abbildung 8 – Fehler bei der Richtungsdetektion 1. Ordnung beim Direktschall des 1. Lautsprechers mit zusätzlicher Bodenreflexion

4.2 Richtungsanalyse höherer Ordnung

Im Unterschied zum vorhergehenden Verfahren werden bei der Richtungsanalyse höherer Ordnung die Ambisonicssignale in Sektoren aufgeteilt. Durch eine derartige Aufteilung können die einzelnen Winkelbereiche gezielt und isoliert betrachtet werden. In Abschnitt 4.3 werden verschiedene Ansätze zur Richtmusterzerlegung genauer behandelt. In den Sektoren kann anschließend eine Richtungsanalyse 1. Ordnung erfolgen, wodurch in Summe mehrfache Richtungsbestimmungen für S Sektoren ermöglicht werden. Für die Herleitung der Sektorenaufteilung wird hier von einer allgemeinen Transformationsvorschrift ausgegangen, die auch mehrere Varianten von Gewichtsfunktionen erlaubt.

4.2.1 Aufteilung in Winkelbereiche (Sektoren)

Nach [KZ14] sind allgemein die räumlich transformierten Ambisonicssignale $\tilde{\phi}(n)$ für den Abtastzeitpunkt n wie folgt definiert:

$$\tilde{\phi}(n) = \mathbf{T} \cdot \phi(n). \quad (13)$$

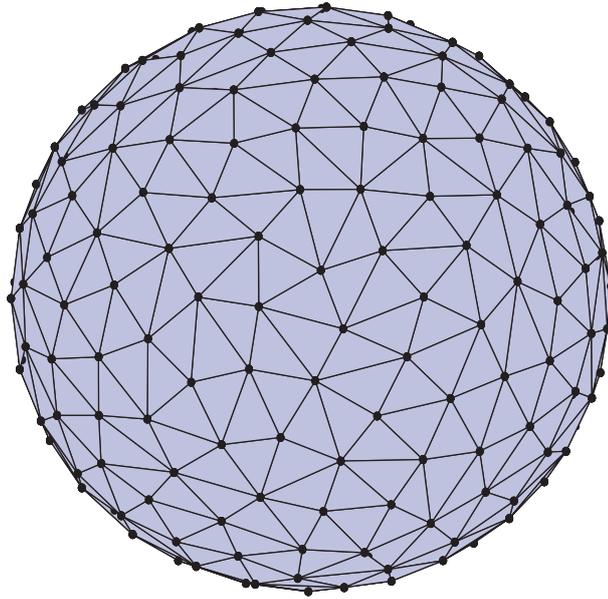
Ziel ist es nun die Transformationsmatrix \mathbf{T} so zu entwerfen, dass vorher festgelegte Sektoren isoliert betrachtet werden können und zueinander amplitudenkomplementär sind. Eine Amplitudenkomplementarität (konstante Amplitude in der Summe der W -Signale aller Sektoren) ist notwendig, um Überhöhungen oder Absenkungen der Signalamplitude bedingt durch Rechenoperationen zu vermeiden. Nach [KZ14] gilt unter der Verwendung von t -Designs [SH96] zur Abtastung der Kugeloberfläche für die Transformationsmatrix der Ausdruck

$$\mathbf{T} = \text{diag}\{\mathbf{b}\} \mathbf{Y}_L^T(\Theta_p) \text{diag}\{\mathbf{g}\} \mathbf{Y}_L(\mathcal{T}^{-1}\{\Theta_p\}), \quad (14)$$

wobei $\mathbf{b} = [\frac{2n+1}{L}]$ als Vektor von Ausgleichskonstanten, \mathbf{g} die Gewichtsfaktoren und $\Theta_p = [\theta_1 \dots \theta_P]$ die Richtungen der P Abtastpunkte des t -Designs sind. $\mathcal{T}^{-1}\{\Theta_p\}$ repräsentiert eine allgemeine, bijektive Abbildungsvorschrift im Winkel, die in diesem Fall nicht benötigt wird. Die Gleichung vereinfacht sich zu

$$\mathbf{T} = \text{diag}\{\mathbf{b}\} \mathbf{Y}_L^T(\Theta_p) \text{diag}\{\mathbf{g}\} \mathbf{Y}_L(\Theta_p). \quad (15)$$

Für eine Ambisonicsordnung L werden zur Berechnung der Transformationsmatrix \mathbf{T} , abgesehen von \mathbf{g} , zwei Kugelflächenfunktionen der Ordnung L multipliziert. Damit die Berechnung der Transformationsmatrix korrekt ist (Transformationsintegral gleich diskreter Summe über t -Design Abtastpunkte), muss ein t -Design von $t \geq 2N$ zur Abtastung der Kugeloberfläche verwendet werden. Hier wird ein 21-Design mit $P = 240$ Punkten [tde] angewandt, welches in Abbildung 9 zu sehen ist. Eine höhere Anzahl an Abtastpunkten hat keine wesentliche Verbesserung gezeigt. Umgekehrt erwiesen sich kleinere t -Designs als ungenügend, da die Kugelflächenfunktionen bei einer niedrigen Anzahl an Abtastpunkten eine Unterabtastung einzelner Sektoren hervorrufen (räumliches Aliasing).

Abbildung 9 – Sphärisches 21-Design mit $P = 240$ Punkten

4.2.2 Bestimmung der Gewichtungsfaktoren

Wie in Gleichung 15 zu erkennen ist, beeinflussen die Gewichtungsfaktoren g die Form der räumlichen Fensterfunktion und somit die Transformation für einen Sektor. Die einfachste Methode zur Bestimmung der Gewichtungsfaktoren ist, für t -Design Richtungen Θ_p innerhalb eines Sektors binär auf 1 und außerhalb eines Sektors auf 0 zu setzen. Bei dieser Methode werden allerdings alle Sektoren sehr scharf voneinander getrennt, was aus zweierlei Sichtweisen Probleme ergeben kann:

Die endliche Anzahl an Abtastpunkten kann demzufolge Fehler bei der Richtungsanalyse bewirken. Zudem erlaubt eine Zerlegung in endlich viele Kugelflächenfunktionen keine unendlich scharf ortsbegrenzten Muster. Abbildung 10 zeigt die Transformationsmatrix \mathbf{T} nullter Ordnung für einen Oktanten-Sektor (0° bis 90° in der Elevation und 0° bis -90° im Azimutwinkel) der Kugeloberfläche mit einer binären Gewichtung. Man erkennt, dass auch in Bereichen außerhalb des Sektors relativ starke Nebenkeulen auftreten, was auf die begrenzte Ordnung der Kugelflächenfunktionen zurückzuführen ist. Dies beeinträchtigt später die Genauigkeit der Richtungsbestimmung.

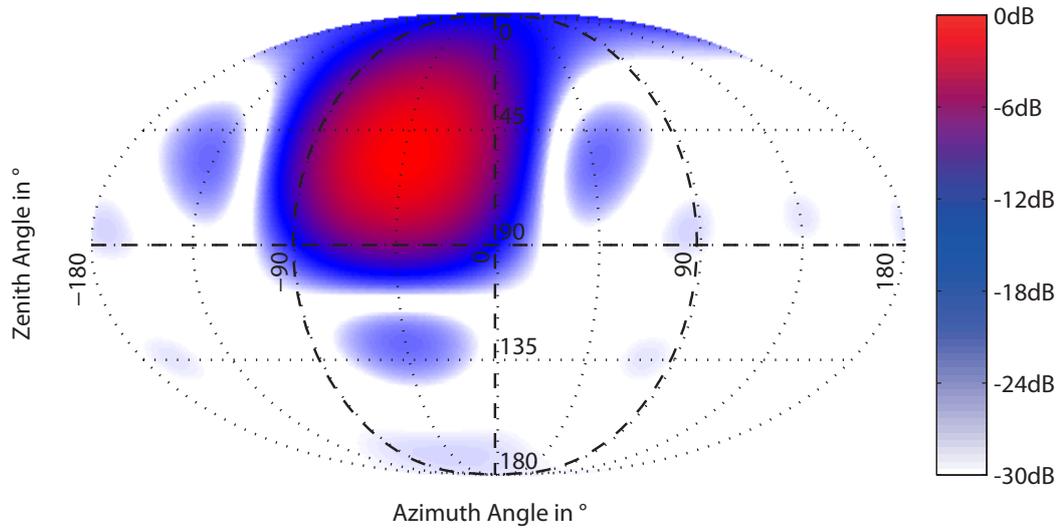


Abbildung 10 – Transformationsmatrix \mathbf{T} nullter Ordnung mit binärer Gewichtung für den Oktanten $-90^\circ \leq \vartheta \leq 0^\circ, 0 \leq \varphi \leq 90^\circ$

Eine andere Möglichkeit zur Bestimmung der Gewichtungsfaktoren ist die Verwendung von sphärischen Richtmustern in die Raumrichtung Θ für eine Gewichtungsmatrix. Hierbei wird für einen Sektor eine um ein Zentrum (in diesem Fall der Koordinatenursprung) rotationssymmetrische Rechteckfunktion (Kappenfunktion) verwendet und die Gewichtungsmatrix \mathbf{g} selbst in Kugelflächenfunktionen zerlegt

$$\mathbf{g} = \mathbf{Y}_L(\Theta_p) \mathbf{y}_L(\theta), \quad (16)$$

wobei $\mathbf{y}_L(\theta)$ die Transformierte der sphärischen Richtmuster darstellt. Um anschließend Nebenkeulen zu unterdrücken, werden in-phase- oder max-rE-Gewichte (siehe Abschnitt 4.2.3) \mathbf{a} als Gewichtung für die Kugelflächenfunktionen eingesetzt. Die Gewichtungsmatrix \mathbf{g} für einen Sektor in Richtung θ ergibt sich folglich zu

$$\mathbf{g} = \mathbf{Y}_L(\Theta_p) \text{diag}(\mathbf{a}) \mathbf{y}_L(\theta). \quad (17)$$

Wie in Abschnitt 3 erwähnt kann mit dem Eigenmike[®] ein Ambisonicssignal 4. Ordnung berechnet werden. Um die Analyse einer Richtungszerlegung 1. Ordnung zu erreichen, ist eine Multiplikation des Zerlegungsmusters 1. Ordnung mit Gewichtungsmustern 3. Ordnung notwendig. Das Gewichtungsmuster reicht damit bis zur 4. Ordnung. Dadurch ergibt sich folgende Gleichung für eine Gewichtungsmatrix mit $L = 3$:

$$\mathbf{g} = \mathbf{Y}_3(\Theta_p) \text{diag}(\mathbf{a}) \mathbf{y}_3(\theta). \quad (18)$$

Nun ergibt sich \mathbf{T} aus einer Multiplikation von Kugelflächenfunktionen 1., 3. und 4. Ordnung. Damit muss für die Berechnung der Transformationsmatrix zumindest ein $t \geq 8$ -Design zur Abtastung verwendet werden. Die Gewichte \mathbf{a} werden für jede Ordnung und deren Grad der Kugelflächenfunktionen analog zur Struktur der Radialfilter in Abschnitt 3.1 angewandt. Für den Vektor \mathbf{a} ergibt sich demnach mit der Ambisonics Ordnung L folgender Ausdruck:

$$\mathbf{a} = [a_0 \ a_1 \ a_1 \ a_1 \ a_2 \ a_2 \ a_2 \ a_2 \ a_2 \ \dots \ a_L]^T \quad (19)$$

Die Ordnung der Kugelflächenfunktionen für die Gewichtung stellt einen deutlichen Einfluss auf die Richtungsverteilung dar. Vor allem für eine max-rE Gewichtung kann die Wahl einer zu hohen Ordnung $L > 3$ die Detektion von Richtungen unkontrolliert hin und her springen lassen und somit eine ungewollte Klangverfärbung hervorrufen. In Abschnitt 5 wird dies an Beispielen näher erläutert.

4.2.3 Gewichtung der Kugelflächenfunktionen

Wie erwähnt sollen zum flacheren Verlauf der Fensterfunktionen Nebenkeulen unterdrückt werden. Jérôme Daniel beschreibt in seiner Dissertation [Dan01] verschiedene Arten zur Gewichtung von Kugelflächenfunktionen, die zur Unterdrückung von störenden Nebenkeulen verwendet werden können. In diesem Abschnitt sollen kurz die Funktionen zur Gewichtung „in-phase“ und „max-rE“ vorgestellt werden.

max-rE Bei der max-rE Gewichtung wird der energetische Richtungsschwerpunkt r_E maximiert. Sie bietet die richtungsabhängig stärkste Bündelung der Signale, wobei zur Unterdrückung der Nebenkeulen Legendre-Polynome $P_l(r_E)$ die Gewichte beschreiben. Gleichung 20 beschreibt mit einer Approximation für die größte Nullstelle des $L + 1$ Legendre-Polynoms [ZF12] L Gewichtungsfaktoren für $l = 0 \dots L$. In Abbildung 11 erkennt man eine gute Unterdrückung der Nebenkeulen, ohne die Hauptkeule nennenswert zu verändern. Die Abgrenzung einzelner Sektoren ist daher sehr gut, allerdings können die Nebenkeulen den Einfluss anderer Sektoren nicht vollkommen ausschließen

$$a_l = P_l(r_E) = P_l\left(\cos\left(\frac{137.9^\circ}{L + 1.51}\right)\right). \quad (20)$$

in-phase Ziel der in-phase Gewichtung ist es Nebenkeulen abhängig von ihrer Richtung relativ zu einer virtuellen Quelle (Hauptkeule) vollständig zu unterdrücken. Das Richtungsgewicht fällt von der Zielrichtung weg streng monoton bis zur Mehrfachnullstelle in der Gegenrichtung. Diametral entgegengesetzt zur virtuellen Quelle ist die Gewichtung somit $g(180^\circ) = 0$. Die Berechnung der L Gewichte erfolgt nach Gleichung 21

$$a_l = \frac{L!(L + 1)!}{(L + l + 1)!(L - l)!}. \quad (21)$$

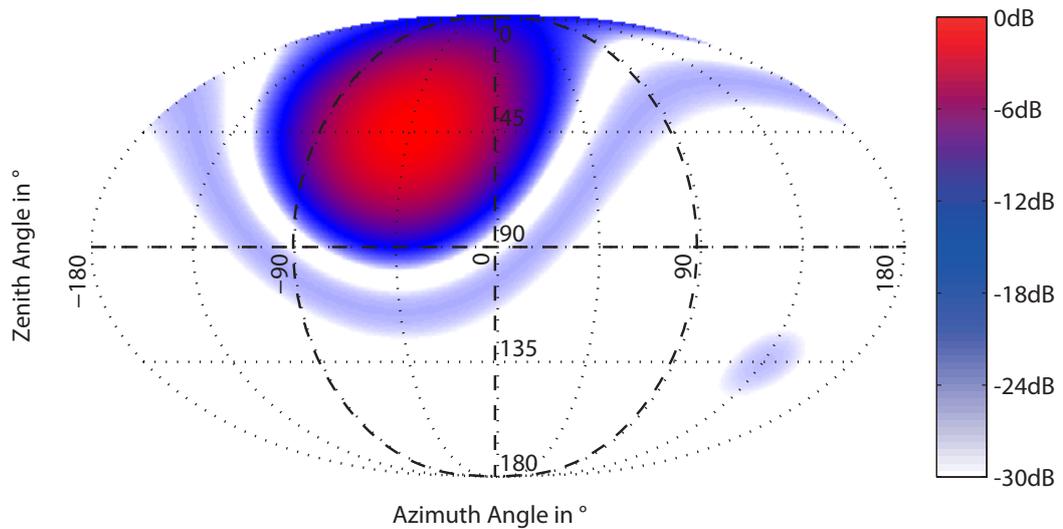


Abbildung 11 – Transformationsmatrix \mathbf{T} nullter Ordnung mit max-rE Gewichtung für den Oktanten $-90^\circ \leq \vartheta \leq 0^\circ, 0 \leq \varphi \leq 90^\circ$

Wie in Abbildung 12 zu sehen, ergibt sich ein weicherer, jedoch wesentlich breiterer Verlauf der Hauptkeule. Diese überlappt stark mit den angrenzenden Sektoren, wodurch die Richtungsdetektion im Sektor beeinflusst wird. Bei der Analyse zeitlich diskreter Impulse mag das zwar ungünstig erscheinen, aber erst bei der Richtungsschärfung von klingenden Ambisonicssignalen wird sich zeigen, welche Form der Abgrenzung günstigere Eigenschaften besitzt.

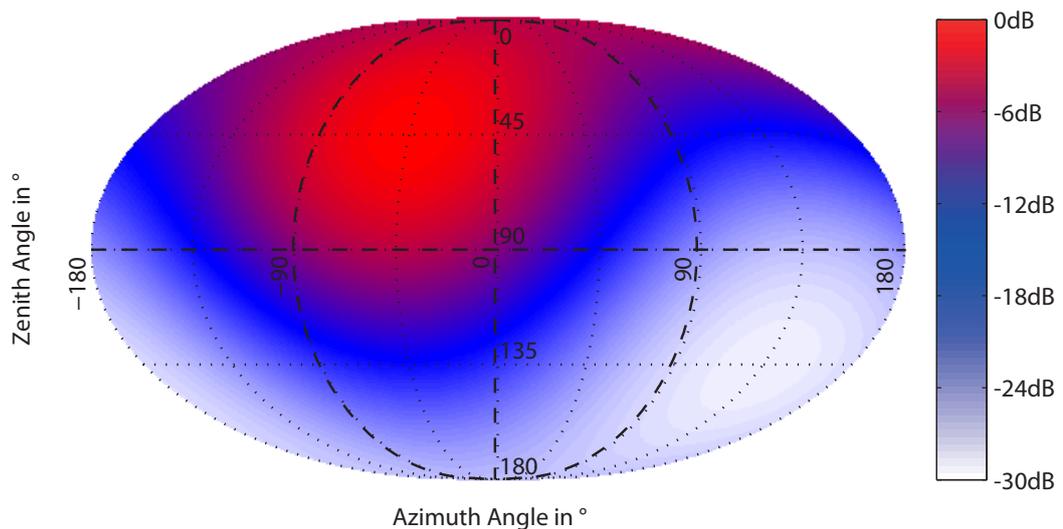


Abbildung 12 – Transformationsmatrix \mathbf{T} nullter Ordnung mit in-phase Gewichtung für den Oktanten $-90^\circ \leq \vartheta \leq 0^\circ, 0 \leq \varphi \leq 90^\circ$

4.3 Sektorendefinition

Setzt man die Gewichtungsfaktoren aus Gleichung 18 in Gleichung 15 ein und berücksichtigt jeweils die Ordnung der Kugelflächenfunktionen gemäß einer maximalen Ambisonics-Ordnung von $L = 4$, erhält man den Ausdruck in Gleichung 22

$$\mathbf{T}_i = \text{diag}\{\mathbf{b}_1\} \mathbf{Y}_1^T(\Theta_p) \text{diag}\{\mathbf{Y}_3(\Theta_p) \cdot \text{diag}(\mathbf{a}) \cdot \mathbf{y}_3(\theta_i)\} \mathbf{Y}_4(\Theta_p). \quad (22)$$

für $i = 1 \dots S$ Sektoren. Um eine Amplitudenkomplementarität zu erreichen, muss die Summe aller Sektorensignale konstant und winkelunabhängig sein. Da die Zerlegung der Sektoren durch die Gewichtungsfunktionen bestimmt wird, müssen sich diese zu 1 aufsummieren. Somit gilt für S Gewichtungsfunktionen \mathbf{g}

$$\frac{1}{\alpha} \sum_{i=1}^S \mathbf{g}_i \stackrel{!}{=} \mathbf{1} \neq \mathbf{f}(\Theta_p).$$

wobei α als Konstante zur Normierung dient. Unter der Verwendung von t -Designs ergibt sich für die Gewichtungsfunktionen, definiert durch Gleichung 18, mit der Verwendung sphärischer Richtmuster 3. Ordnung die Forderung:

$$\sum_{l=0}^3 a_l \sum_{i=1}^S \sum_{m=-l}^l Y_l^m(\Theta_p) Y_l^m(\theta_i) = \alpha. \quad (23)$$

Nach dem Additionstheorem der Kugelflächenfunktionen kann das aufsummierte Produkt zweier Kugelflächenfunktionen bis zur 3. Ordnung auch durch sphärische Legendre-Polynome bis zur 3. Ordnung ausgedrückt werden

$$P_l(\underbrace{\theta_i^T \Theta_p}_{:=\mu_i}) = 4\pi \sum_{m=-l}^l Y_l^m(\Theta_p) Y_l^m(\theta_i). \quad (24)$$

Wird Gleichung 23 mit 4π multipliziert und damit umformuliert, erhält man:

$$\sum_{l=0}^3 a_l \sum_{i=1}^S P_l(\mu_i) = 4\pi\alpha \quad (25)$$

Die Summe $\frac{1}{S} \sum_{i=1}^S P_l(\mu_i)$ schätzt den bekannten Gleichanteil der Legendre-Polynome $\int_{-1}^1 P_l(\mu) d\mu = 2\delta_{l,0}$. Wird zur Bestimmung der Richtungen für die Sektoren ein t -Design mit $t \geq 3$ verwendet, so entspricht die Summe für $l \leq 3$ exakt dem Integral:

$$\frac{1}{S} \sum_{i=1}^S P_l(\theta_i^T \Theta_p) = \int_{-1}^1 P_l(\mu) d\mu = 2\delta_{l,0}. \quad (26)$$

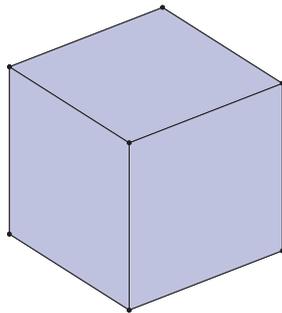
Ist dies der Fall, vereinfacht sich nach einer Division von Gleichung 25 durch S die Zweifachsumme zu

$$2a_0 = \frac{4\pi}{S}\alpha. \quad (27)$$

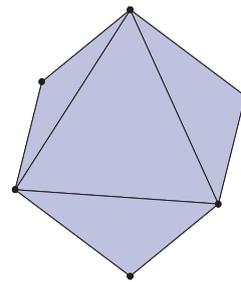
Gleichung 27 ist offensichtlich konstant, womit die Amplitudenkomplementarität unter der Verwendung von t -Designs zur Sektorendefinition gewährleistet ist. Es ergibt sich dann je nach Gewichtung a (in-phase oder max-rE) für die Normierung α :

$$\alpha = \frac{S}{2\pi}a_0. \quad (28)$$

In Abbildung 13a ist ein Würfel-3-Design mit acht Punkten zur Zerlegung in Oktanten abgebildet, die bereits aus Abschnitt 4.2.2 bekannt sind. Abbildung 13b zeigt ein Oktaeder-3-Design mit sechs Punkten. In [PVP15] sind diverse weitere Ansätze zur Zerlegung einer Einheitskugel in Sektoren unter Verwendung von t -Designs aufgeführt.



(a) Würfel-3-Design mit 8 Punkten



(b) Oktaeder-3-Design mit 6 Punkten

Abbildung 13 – Auswahl an t -Designs zur Sektorenzerlegung

Ein Oktaeder-3-Design weist der südlichen Hemisphäre nur einen, der horizontalen Ebene dagegen vier Sektor zu. Ein weiterer steht der nördlichen Hemisphäre zur Verfügung. Bei einem Würfel-3-Design jedoch werden der südlichen Hemisphäre vier Sektoren zugeschrieben. In Abbildung 14 ist die Sektorenzerlegung durch ein Würfel-3-Design (Richtungsvektoren blau) und ein Oktaeder-3-Design (Richtungsvektoren rot) illustriert. Betrachtet man den IEM-CUBE als Beispiel, ist eine Sektorenzerlegung durch ein Oktaeder-3-Design aufgrund der halbkugelförmigen Lautsprecheranordnung geeigneter. Im Vergleich zum Oktaeder-3-Design findet durch das Würfel-3-Design aufgrund der höheren Anzahl an Sektoren zusätzlich eine schärfere Trennung einzelner Bereiche statt, was unter Umständen zu einer schlechteren Richtungsauflösung führen kann. Aus diesen Gründen soll im nächsten Abschnitt ein Oktaeder-3-Design zur Sektorendefinition verwendet und analysiert werden.

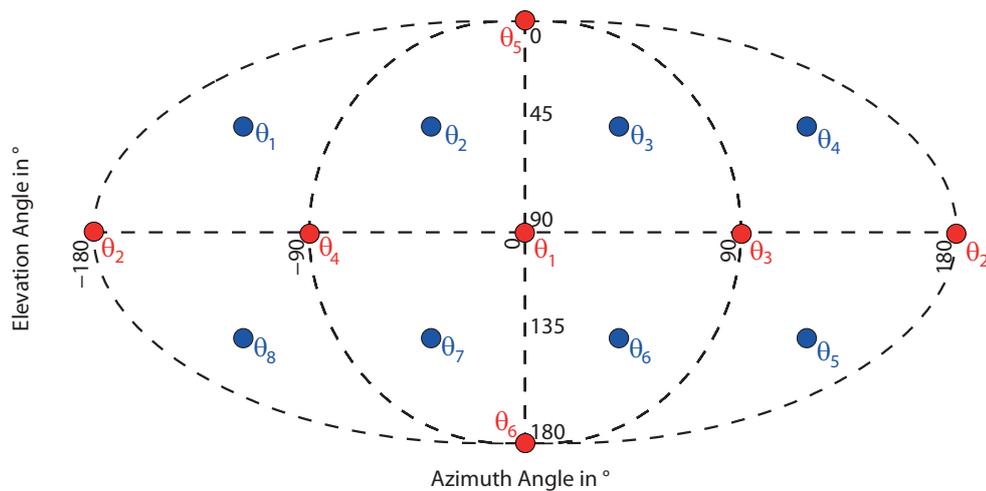


Abbildung 14 – Winkelpaare θ_i in Richtung des Normalenvektors eines Kugeloktanten

Die räumlich transformierten Ambisonicssignale $\tilde{\phi}_i(n)$ aus Gleichung 22 können nach einer Sektorenzerlegung für $i = 1 \dots S$ Sektoren zur Richtungsanalyse 1. Ordnung mithilfe der Gleichungen 11 und 12 verwendet werden. Für das obige Beispiel in Abbildung 8 ergibt sich unter der Verwendung einer Zerlegung durch ein Oktaeder-3-Design mit max-rE Gewichtung bei Mehrfachrichtungsdetektion die Veränderung in Abbildung 15. Nun können die Richtungen der beiden hohen Signalkomponenten separat detektiert werden. Im nächsten Abschnitt sollen Vor- und Nachteile der verschiedenen Gewichtungen und der Mehrfachrichtungsbestimmung anhand von konkreten Beispielen behandelt werden.

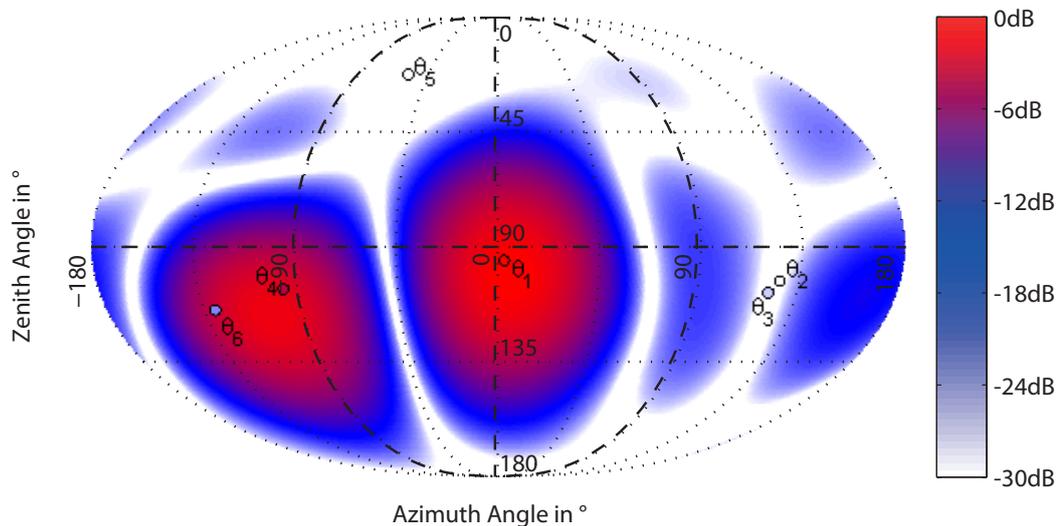


Abbildung 15 – Mehrfachrichtungsdetektion in den Sektoren S_1 bis S_6 beim Direktschall des 1. Lautsprechers mit zusätzlicher Bodenreflexion

5 Auswertung des Algorithmus

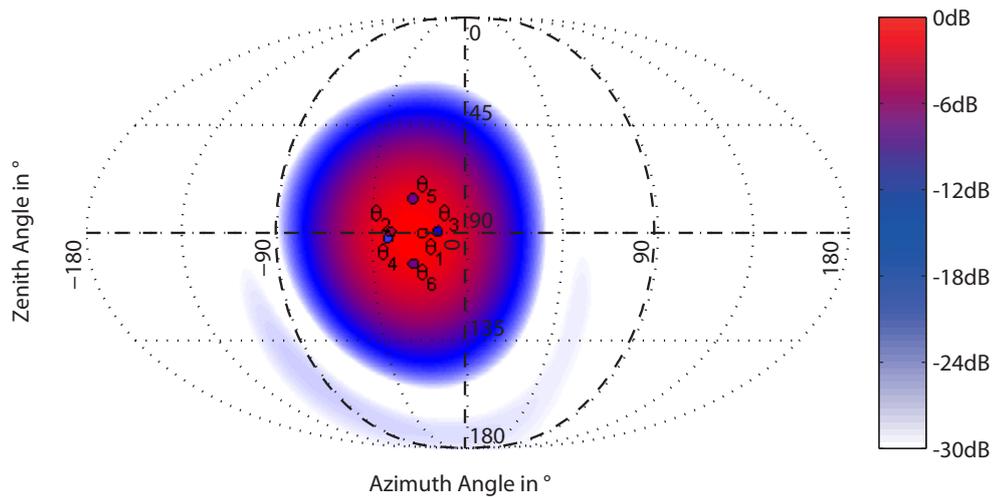
Um die objektive Qualität der Mehrfachdetektion zu verifizieren, werden in diesem Abschnitt mehrere Beispiele und verschiedene Ausführungen (Fensterfunktionen) der Transformationsmatrix \mathbf{T} betrachtet. Es soll dabei jeweils der Unterschied zur Richtungsbestimmung 1. Ordnung und den verschiedenen Gewichtungen max-rE und in-phase untersucht werden. Das Transformationsverfahren mit binärer Gewichtung wird in dieser Arbeit nicht berücksichtigt. Es wurde empirisch festgestellt, dass diese Art der sphärischen Fensterfunktionen keine brauchbaren Werte zur Richtungsbestimmung liefert.

5.1 Beispiel: Direktschall

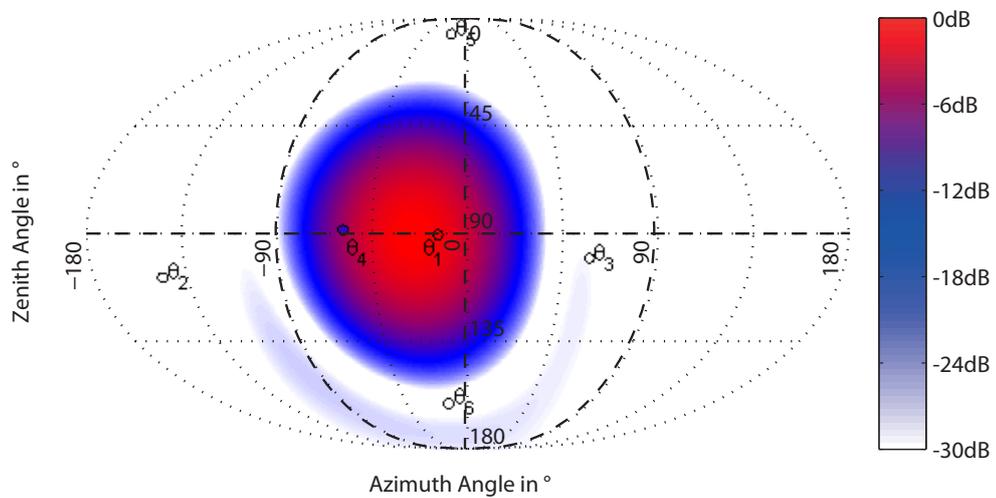
Abbildung 16 zeigt das Signal des 2. Lautsprechers zum Abtastzeitpunkt $n = 1027$. Hier erscheint der Direktschall am Mikrofon und befindet sich nicht komplett innerhalb eines Sektors. Der Algorithmus detektiert die Richtung der höchsten Intensität in der Hauptkeule ungenau. Unter diesen Umständen zeigt sich eine geringfügige Schwachstelle der Mehrfachdetektion. Bei einer Richtungsanalyse 1. Ordnung zeigt der Intensitätsvektor in diesem Fall direkt auf das Zentrum der Hauptkeule. Für eine in-phase Gewichtung in Abbildung 16a werden aufgrund der wesentlich breiteren Überlappung der Sektoren alle Richtungen in der Gegend um die Direktschallkeule detektiert. Nur die Amplituden der W -Signale unterscheiden sich. Für Sektor S_1 besitzt die Richtung θ_1 die größte Amplitude, da in diesen Sektor der größte Anteil des Signals fällt. Zu den Grenzen der Sektoren hin fällt die max-rE Gewichtung in Abbildung 16b teilweise sehr steil ab, was in einigen Fällen sogar Nullstellen innerhalb eines Sektors hervorruft. In diesem Beispiel befindet sich der Direktschall fast ausschließlich in den Sektoren S_1 und S_4 , weshalb aufgrund des steilen Abfalls der max-rE Gewichtung in die anderen Sektoren nahezu kein Signalanteil fällt (weiße Detektionspunkte). Es ist also weitestgehend die Detektion in Sektor S_1 für die Richtungszerlegung ausschlaggebend, was aufgrund des vorliegenden Beispiels mit einem Direktschall sinnvoll erscheint.

5.2 Beispiel: Signal mit Reflexionskomponente

Zur Untersuchung des Unterschieds zwischen den Gewichtungen bei einem Signal mit zusätzlich hoher Reflexionskomponente wird an dieser Stelle das bereits bekannte Beispiel aus Abbildung 8 betrachtet. Hier liegt der Hauptsignalanteil wieder zwischen mehreren Sektoren. Zusätzlich ist allerdings eine sehr hohe Reflexionskomponente im Signal enthalten. Hier offenbart sich der Vorteil der Mehrfachrichtungsdetektion: Sowohl die in-phase in 17a als auch die max-rE 17b Gewichtung können beide Signalkomponenten sehr gut detektieren. Bei der in-phase Gewichtung fallen sämtliche Detektionen wieder in die Bereiche der größten Amplituden und max-rE detektiert wie im obigen Beispiel nur in den Sektoren mit nennenswertem Signalanteil auf sinnvolle Weise. Eine Richtungsbestimmung 1. Ordnung würde wegen der deutlichen Reflexion eine Fehldetektion hervorrufen, wie bereits in Abschnitt 4.1 gezeigt wurde.

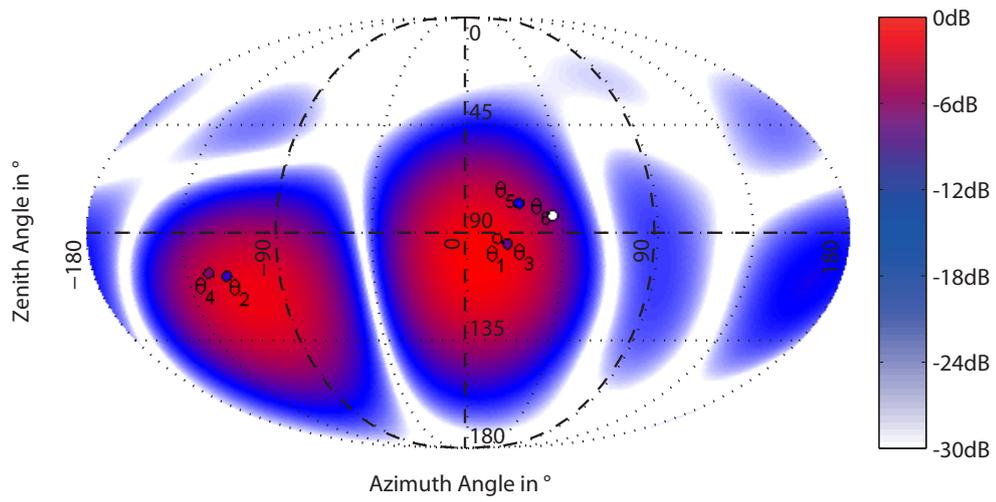


a) in-phase Gewichtung

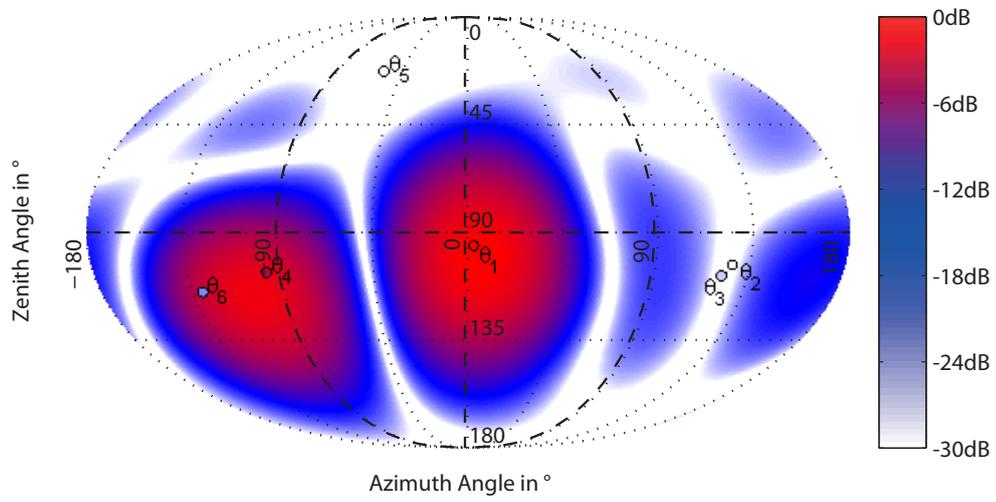


b) max-rE Gewichtung

Abbildung 16 – Mehrfachdetektion am Direktschall des 2. Lautsprechers



a) in-phase Gewichtung



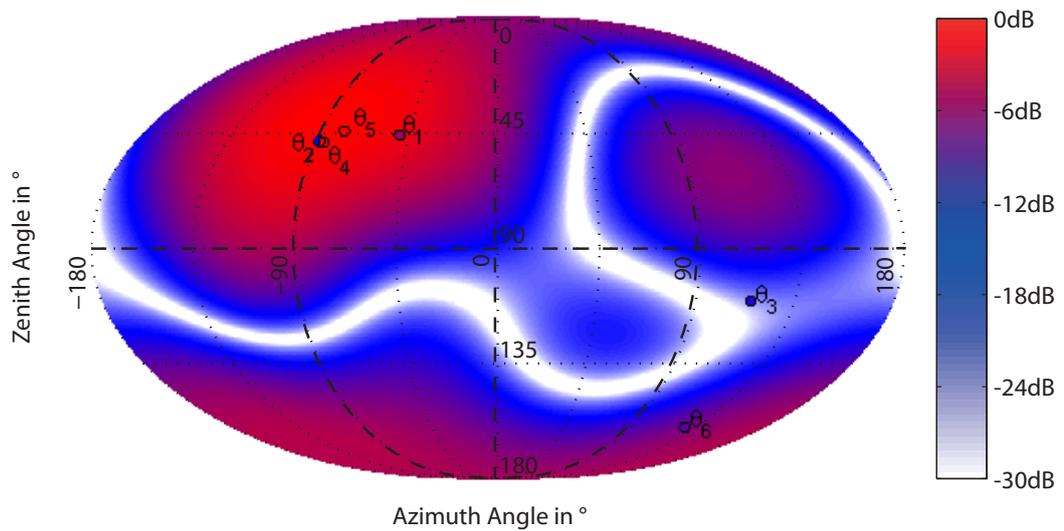
b) max-rE Gewichtung

Abbildung 17 – Mehrfachdetektion beim Direktschall des 1. Lautsprechers mit zusätzlicher Bodenreflexion

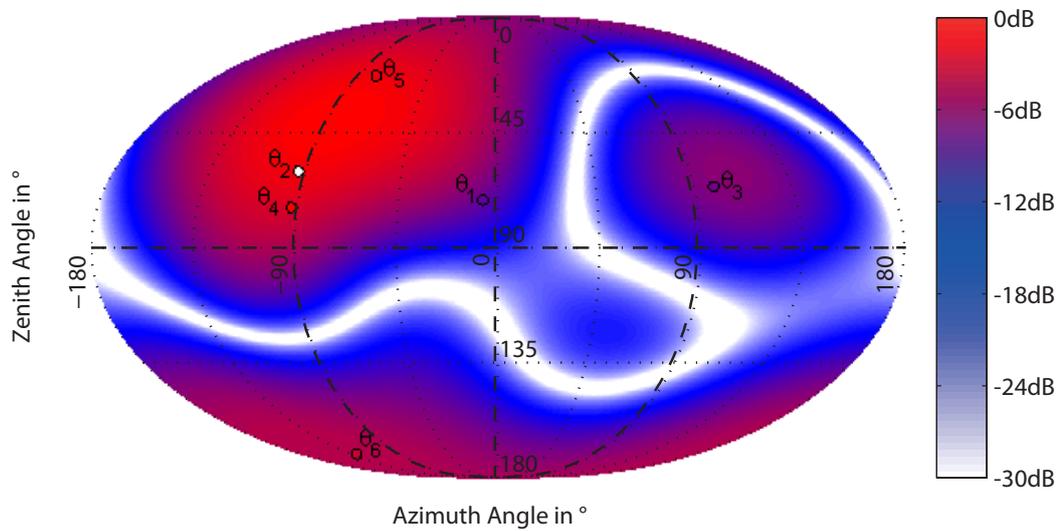
5.3 Beispiel: Diffuses Schallfeld

Deutlich wird der Vorteil von scharf abgetrennten Mehrfachrichtungsdetektionen bei zusätzlichen Reflexionen, die kleinere Amplitudenwerte aufweisen. In Abbildung 18 ist eine ausgeprägte Schallkomponente bei $\varphi = -90^\circ$ und $\vartheta = 45^\circ$, eine relativ schwache Reflexionskomponente bei etwa $\varphi = 100^\circ$ und $\vartheta = 60^\circ$ und eine breitere Reflexionskomponente in der südlichen Hemisphäre illustriert. In diesem wesentlich komplexeren Beispiel eines diffusen Schallfeldes kann eine Richtungsdetektion 1. Ordnung keinen brauchbaren Wert mehr liefern. Analysiert man dieses Signal mit der max-rE Gewichtung in Abbildung 18b, erkennt man, dass alle nennenswerten Signalkomponenten sinnvoll detektiert werden können. Bis auf eine leichte Verschiebung der Hauptkomponente findet eine gute Richtungsdetektion statt.

Betrachtet man dieses komplexere Beispiel mit einer in-phase Gewichtung in 18a, erkennt man lediglich zwei Richtungsbereiche, in denen eine Detektion erfolgt. Nur bei der Direkt-schallkomponente kommt eine sinnvolle Richtungsbestimmung zustande. Die breite Reflexionskomponente in der südlichen Hemisphäre wird nicht detektiert. Aufgrund der oben genannten breiten Sektorenüberlappung können viele Signalanteile mit kleineren Amplituden nicht mehr richtig detektiert werden. Die Gewichtung mit max-rE zeigt also bei Reflexionen mit relativ geringer Amplitude ein deutlich besseres Ergebnis.



a) in-phase Gewichtung



b) max-rE Gewichtung

Abbildung 18 – Mehrfachdetektion am Beispiel eines diffusen Schallfeldes

6 Ambisonics Kodierung

Zur praktischen Anwendung des Algorithmus auf übliche Tonaufnahmen sind Berechnungen notwendig, die Aufnahmen entweder Blockweise als Echtzeitverarbeitung oder offline als komplette Tonaufnahme verarbeiten können. Im Folgenden wird ein Kodierer vorgestellt, der aus der Kurzzeit-Fouriertransformierten einer Ambisonics-Aufnahme höherer Ordnung eine Mehrfachdetektion für alle Sektoren anwendet und das kodierte Signal dann über ein Overlapp-Add Verfahren zusammenführt. Das daraus entwickelte Signal X_{ambiX} kann im Anschluss über entsprechende Plugins in geeigneten Umgebungen abgespielt werden. In Abbildung 19 ist der strukturelle Aufbau des Ambisonics Encoders für 6 Sektoren dargestellt.

Die transformierten Ambisonicssignale 0. und 1. Ordnung für einen Sektor i dienen dem Encoder als Eingangsdaten und werden zunächst durch ein Wurzel-von-Hann Fenster der Länge $N = 1024$ in Blöcke unterteilt. Diese Blocklänge ging in einer parallel am IEM durchgeführten Seminararbeit zur Richtungsanalyse 1. Ordnung als perzeptiv sinnvoll hervor. Durch Transformation in den Frequenzbereich mittels FFT wird jedem Frequenz-Bin f durch den Algorithmus eine Richtung θ_{DOA} zugeordnet. Mit diesen Richtungen werden zur Re-Enkodierung Kugelflächenfunktionen 5. Ordnung ausgewertet und mit dem W -Signal multipliziert. Somit erhält man für den Sektor i ein kodierte Signal $\xi_i(f)$ im Frequenzbereich, mit dem die Signale der anderen Sektoren überlagert werden. Eine inverse Fouriertransformation und erneute Multiplikation mit dem Wurzel-von-Hann-Fenster ermöglicht die Rücktransformation in den Zeitbereich und eine entsprechende Gewichtung. Im letzten Schritt werden durch ein Overlapp-and-Add Verfahren nachstehende bzw. vorangehende Blöcke jeweils zur Hälfte mit dem aktuellen überlagert. Das Ergebnis ist ein 36-kanaliges Audiosignal mit der Länge der Originalaudiodatei und kann direkt abgespielt werden.

Es hat sich bei der Überprüfung von Eigenschaften des Algorithmus gezeigt, dass auch die 36-kanaligen Zeitsignale der einzelnen Sektoren aufschlussreiches Tonmaterial liefern.

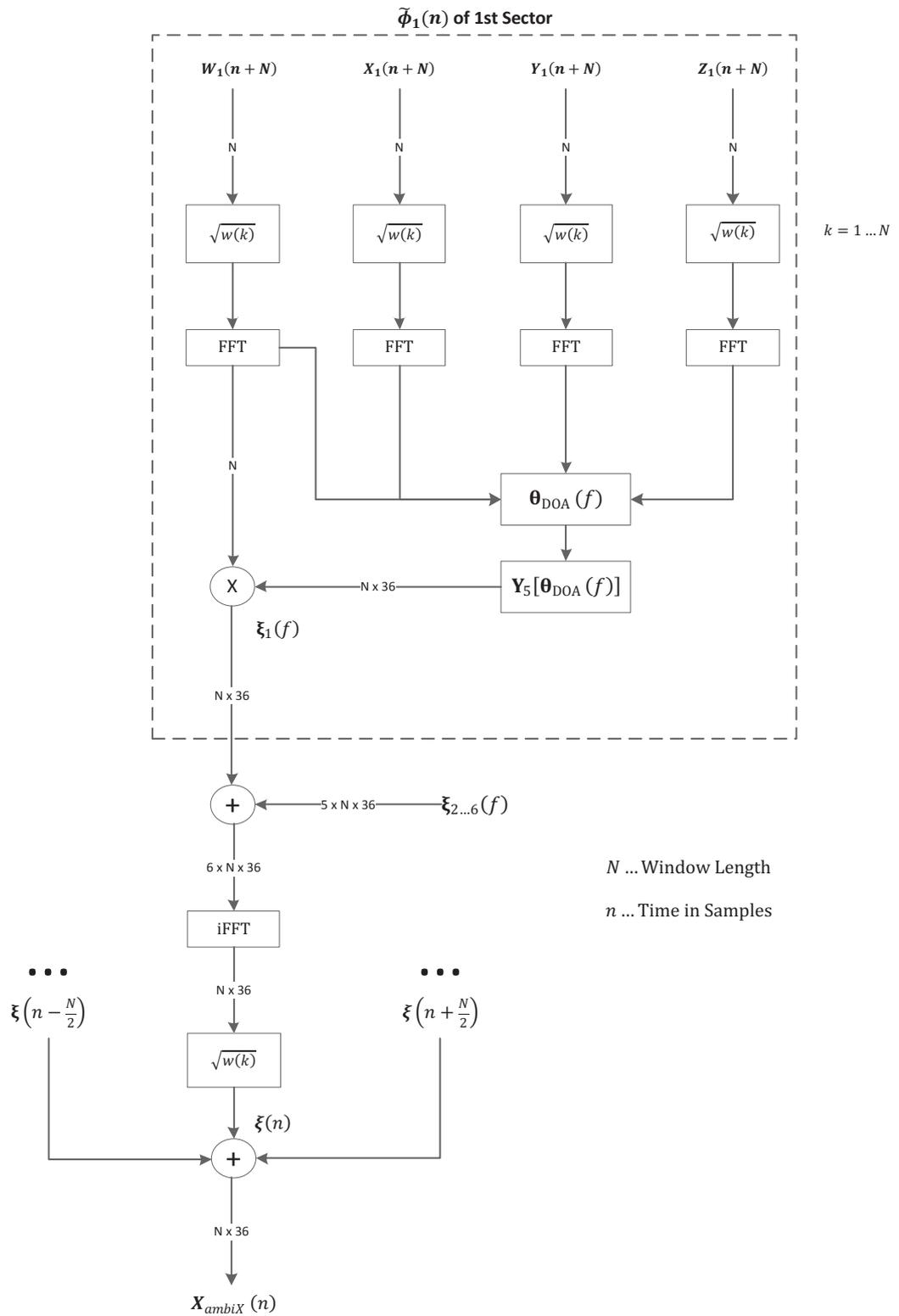


Abbildung 19 – Struktureller Aufbau des Ambisonics-Kodierers

7 Schlussfolgerungen

7.1 Objektive Betrachtung

In den vorhergehenden Abschnitten wurde gezeigt, dass eine Mehrfachrichtungsdetektion von Tonaufnahmen eine sinnvolle Ergänzung zu bereits bekannten Verfahren 1. Ordnung darstellt. Vor allem durch die detailliertere Beschreibung des Schallfeldes können auch frühe und starke Reflexionen detektiert und die Ambisonics-Wiedergabe realistischer gestaltet werden. Die Probleme von Richtungsdetektionen 1. Ordnung, wie die Verschiebung des Intensitätsvektors bei mehreren stark ausgeprägten Keulen können durch die Mehrfachdetektion ebenfalls eliminiert werden.

Für bestimmte Situationen mit zeitlich isolierten Einzelereignissen, wie Direktschall ohne Reflexionen, ist allerdings eine Detektion 1. Ordnung sinnvoller. Da die Sektoren bei Mehrfachrichtungsbestimmungen durch die Transformation statisch festgelegt sind, werden die Ergebnisse solcher Einzelereignisse zwischen einzelnen Sektoren ungenauer. Sowohl die in-phase als auch die max-rE Gewichtung zeigen Vor- und Nachteile. Bei vielen Reflexionen mit kleinen Amplituden liefert die max-rE Gewichtung die wesentlich besseren Ergebnisse. Bei sehr starken Signalanteilen, vor allem an den Grenzen von Sektoren, detektiert der Algorithmus mit in-phase Gewichtung die genauere Richtung. In den gängigsten Anwendungen wird dennoch die max-rE Gewichtung sinnvollere Werte liefern.

Sonderfälle, in denen eine Lokalisation mit in-phase oder max-rE 1. Ordnung bessere Ergebnisse liefert, tauchen nur sehr selten auf, wobei auch hier die Abweichungen der Mehrfachdetektionen für das menschliche Gehör überaus gering einzustufen sind. Zusammenfassend lässt sich sagen, dass aus den genannten Gründen für Eigenmike-Aufnahmen die Vorteile der Mehrfachdetektion zur ambisonischen Wiedergabe von Tonaufnahmen gegenüber des Directional Audio Coding 1. Ordnung überwiegen.

7.2 Subjektive Betrachtung

Bei der Anwendung auf Tonaufnahmen kann perzeptiv eine Verbesserung der Richtungsauflösung wahrgenommen werden. In mehreren Durchläufen wurden die Ergebnisse des Algorithmus aus Abbildung 19 mit drei Personen im IEM-Lehrstudio und im IEM-CUBE gemeinsam einem Hörvergleich unterzogen (Frank, Zotter, Ziegler). Dazu wurden bekannte Aufnahmen (Essen, Maroni, Asteroid, Acapella) in Summe aller Sektoren und in Einzelsektoren abgehört.

Konzentriert man sich auf einen speziellen Ton einige Sekunden nach Beginn der Asteroid-Aufnahme (beispielsweise einen Beckenschlag), erscheint die Quellenrichtung beim Direktschall deutlich schärfer. Diffuse Anteile werden dann auch geschärft auf der Kugeloberfläche wiedergegeben, allerdings ohne merklichen Unterschied. Die deutliche Verbesserung zu einer Originalaufnahme ist in diesen besonderen Fällen nur bedingt möglich. Vergleicht man eine bestimmte Aufnahme mit der entsprechend kodierten Aufnahme, kann nur sehr schwer ein Unterschied festgestellt werden. Die Wahl der Sektoren, die Ordnung der Sektorengewichtung und die Wahl der Gewichtungsfunktion $\max\text{-rE}$ oder in-phase spielen aber eine tragende Rolle. Hier ergibt sich ein deutlich hörbarer Unterschied und es können viele Optimierungsaufgaben zu einem verbesserten Ergebnis führen. Im Hörvergleich wurde beobachtet, dass die Ordnung der Kugelflächenfunktionen zur Sektorengewichtung nicht höher als $L = 3$ sein sollte. Höhere Ordnungen verursachen durch die zu scharfe Abtrennung von Sektoren unerwünschte Nebengeräusche (Musical Noise). Durch eine $\max\text{-rE}$ Gewichtung konnte für sämtliche Hörbeispiele, in den einzelnen Sektoren, eine wesentlich schärfere und detailliertere Richtungszuordnung wahrgenommen werden. Des Weiteren konnte für den IEM-CUBE ein Oktaeder-3-Design, welches der nördlichen Hemisphäre (halbkugelförmige Lautsprecheranordnung) mehr Sektoren zuweist als ein Würfel-3-Design, eine bessere Richtungsauflösung gehört werden. Eine entsprechende Überlappung der Sektoren (z.B. durch Gewichtung mit $L = 3$ und $\max\text{-rE}$ Gewichtung) scheint für sämtliche Lautsprecheranordnungen und Gewichtungsfunktion sinnvoll zu sein.

Verbessertes Richtungshören konnte im IEM-CUBE und mit den verfügbaren Eigenmike-Aufnahmen in Sonderfällen festgestellt werden. Erscheinen die Signale räumlich diffus zu werden, ist der Algorithmus in der Lage eine stärkere Richtungszuweisung herzustellen. Wird berücksichtigt, dass sich bei sonstigen Diffusfeldern und Nachhall die Klangqualität kaum verändert, könnte der Algorithmus Aufnahmen mit dem Eigenmike® einen subtilen Feinschliff in der Abbildungsschärfe verleihen. Zur technischen Richtungsdetektion ist der vorgeführte Ansatz zweifelsohne eine hilfreiche Technik.

Literatur

- [BPF11] R. Baumgartner, H. Pomberger, and M. Frank, "Practical implementation of radial filters for ambisonic recordings," 2011.
- [Dan01] J. Daniel, "Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia," 2001.
- [EM3] "Mh-acoustics eigenmike em32," <http://www.mhacoustics.com/>.
- [Har78] F. J. Harris, "On the use of windows for harmonic analysis with the discrete fourier transform," 1978.
- [KZ14] M. Kronlachner and F. Zotter, "Spatial transformations for the enhancement of ambisonic recordings," 2014.
- [PP11] A. Politis and V. Pulkki, "Broadband analysis and synthesis for directional audio coding using a-format input signals," 2011.
- [Puk07] V. Pukki, "Spatial sound reproduction with directional audio coding," 2007.
- [PVP15] A. Politis, J. Vilkamo, and V. Pulkki, "Sector-based parametric sound field reproduction in the spherical harmonic domain," 2015.
- [SH96] N. J. A. Sloane and R. H. Hardin, "McClarens improved snub cube and other new spherical designs in three dimensions," 1996.
- [tde] "Spherical t-designs," <http://neilsloane.com/sphdesigns/>.
- [ZF12] F. Zotter and M. Frank, "All-round ambisonic panning and decoding," 2012.
- [ZL15] F. Zotter and S. Loesler, "Comprehensive radial filter design for practical higher-order ambisonic recording," 2015.
- [Zot14] F. Zotter, *Akustische Holografie und Holografie*, 2014.
- [ZRS02] J. M. Zmoelnig, W. Ritsch, and A. Sontacchi, "Der iem cube," 2002.