



IEM Report 38/07

Binaural-Ambisonic 4.Ordnung 3D-Raumsimulationsmodell mit ortsvarianten Quellen und Hörerin bzw. Hörer für PD

Verfasser:

Thomas Musil

Alois Sontacchi

Markus Noisternig

Robert Höldrich

Januar 2007

Zusammenfassung

Dieser Bericht umfasst die folgenden Themen : Herleitung eines reduzierten HRIR-Satzes für 3D-Ambisonic, Vorstellung einer gleichmäßig verteilten Lautsprecher-Anordnung (3D 4.Ordnung), Beschreibung eines Raumsimulationsmodells in 3D für bewegte Quellen und Hörerin bzw. Hörer, Beschreibung des Spiegelquellenmodells, Beschreibung des Nachhallmodells, Bedienungsanleitung des in Pd geschriebenen Simulationsmodells, Beschreibung der dafür notwendigen Bibliotheken `iem_bin_ambi` (Objekte zur Berechnung der reduzierten Dekodermatrix und des reduzierten HRIR-Satzes) und `iem_abs` (Objektsammlung zur Realisierung von Ambisonic-Systemen). Pd ist eine graphische Echtzeit-Multimedia-Programmiersprache von Miller Puckette [1]. Der Bericht setzt die Kenntnis von IEM-Report 15/03 voraus, und die dort vorgestellten Pd-Bibliotheken für ein 3D-Ambisonic-System: `iem_matrix` (Signalobjekte für Matrizenmultiplikation) und `iem_ambi` (Objekte zur Berechnung der Matrizenkoeffizienten bezüglich Kodierung, Rotation und Dekodierung).

Abstract

This report covers the following topics: the derivation of a reduced HRIR set for 3D Ambisonic, the presentation of an evenly spread arrangement of loudspeakers (3D 4.order), a description of a three dimensional room simulation model for moving sources and listener, the mirror source method and the late reverberation, the operation manual of the simulation model. All these Ambisonic related topics are implemented to Pd , a realtime multimedia graphical programming language, developed by Miller Puckette [1]. 3 Pd-libraries are required to realize a 3D Ambisonic spatialisation system: `iem_matrix` (signal matrix multiplication objects), `iem_ambi` (encoding, rotating and decoding) and `ambi_abs` (implementation examples).

INHALTSVERZEICHNIS

1	EINLEITUNG.....	1
2	HERLEITUNG EINES REDUZIERTEN HRIR-SATZES	2
2.1	Berechnung der Dekodiermatrix	2
2.2	Herleitung der beiden reduzierten HRIR-Sätze.....	5
2.3	Herleitung der Verwendung nur eines reduzierten HRIR-Satzes.....	6
2.4	Optimales Lautsprecher-Layout.....	9
2.5	Messergebnisse.....	11
3	AUFBAU UND FUNKTIONSWEISE DES RAUMSIMULATIONSMODELLS ...	13
3.1	Prinzipschaltbild des Spatialisierungsmodells.....	13
3.2	Das Erst- und Zweit-Reflexionsmodul und deren Kodierung	14
3.3	Das Diffusnachhallmodul und dessen Kodierung	17
3.4	Das Dekodiermodul und die HRTF-Filterung	20
4	BEDIENUNGSANLEITUNG DER PD-IMPLEMENTIERUNG	22
4.1	Die grafische Oberfläche und die Bedienung der Parameter.....	22
4.2	Die grafische Simulation des Raums	26
4.3	Der Main-Patch.....	27
4.4	Die Presets	28
4.5	Beschreibung der Steuerparameter des DSP-Moduls	29

5	BESCHREIBUNG DER PD-BIBLIOTHEKEN	33
5.1	Binaural-Ambisonic-Dekoder-Bibliothek „iem_bin_ambi“	33
5.2	Bibliothek „ambi_abs“	35
6	ERKENNTNISSE, AUSSICHTEN	37
7	LITERATURVERZEICHNIS	38
	ANHANG	39

1 Einleitung

Dieser Arbeitsbericht basiert auf den Erkenntnissen der Diplomarbeiten von Johannes Zmöllnig [2] und Markus Noisternig [4], bzw. der Dissertationen von Alois Sontacchi [3] und Jérôme Daniel [5]. Die Herleitung des reduzierten HRIR-Satzes und der reduzierten Dekodermatrix für Binauralanwendungen, bzw. der Hall-Algorithmus [8] wurden etwas ausführlicher erläutert. Bezüglich der Pd-Message-Syntax wird auf [6] verwiesen, bezüglich der Ambisonic-Kodierungen und -Rotationen auf [7]. Die Bibliothek, die hier besprochen wird, besteht aus einer External-Bibliothek (iem_bin_ambi) mit den Endungen *.dll für Windows, *.pd_linux für Linux und einem Sammelordner ambi_abs mit Pd-Subpatches.

Die aktuelle Version kann man beziehen unter:

http://iem.at/~musil/iemlib_extended/. © IEM 2006, Thomas MUSIL [musil@iem.at].

Im weiteren Text wird vorausgesetzt, dass der Begriff Hörer natürlich genauso Hörerinnen mit einschließt.

Die Raumsimulation wird vereinfacht für einen quaderförmigen Raum mit gleichen Wand-, Decken- und Boden-Absorptionseigenschaften berechnet, der Nachhall weist aber keine störenden Flatterechos auf. Diese Raumsimulationsumgebung dient zur Untersuchung des Einflusses verschiedener Parameter (HRIR-Sätze, Zweitreflexionen, HRIR-Längen) auf die räumliche Ortungsgenauigkeit.

2 Herleitung eines reduzierten HRIR-Satzes

2.1 Berechnung der Dekodiermatrix

Die Herleitung des reduzierten HRIR-Satzes für ein 3D-Binaural-Ambisonic-Dekodiersystem wird hier anhand eines Beispiels durchgeführt. Wir wählen ein Ambisonic-System 2.Ordnung mit 13 virtuellen Lautsprecherpositionen, die möglichst gleichmäßig im Raum verteilt werden. Um das Beispiel einfach zu halten, wählen wir einen Lautsprecher am Nordpol, 3 Lautsprecher auf den nördlichen mittleren Breiten, 5 Lautsprecher am Äquator, 3 Lautsprecher auf den südlichen mittleren Breiten und einen am Südpol. Für diese 13 Lautsprecher liegen HRIR-Daten vor (für den Südpol-Lautsprecher muss man sich erst eine Impulsantwort konstruieren, zB. Man nehme die HRIR mit Elevation -40° , Azimut 180° , filtere sie mit einem Tiefpass und senke den Pegel um -6 dB).

$$\Phi = \begin{pmatrix} \begin{array}{c|c|c|c|c|c|c|c|c|c|c|c|c} & LS_4 & & LS_8 & LS_9 & & LS_{12} & & \\ \hline LS_1 & LS_2 & LS_3 & -LS_3 & LS_5 & LS_6 & LS_7 & -LS_7 & -LS_6 & LS_{10} & LS_{11} & -LS_{11} & LS_{13} \\ \hline \theta & 90 & 45 & 45 & 45 & 0 & 0 & 0 & 0 & 0 & -45 & -45 & -45 & -90 \\ \hline \varphi & 0 & 0 & 120 & -120 & 0 & 72 & 144 & -144 & -72 & 180 & 60 & -60 & 0 \end{array} \end{pmatrix} \quad (2.1.1)$$

Bei der Verteilung der Lautsprecher entlang eines Breitengradrings (gleicher Elevation) soll der Azimutwinkelabstand $360^\circ/(\text{Ringlautsprecheranzahl})$ betragen. Den südlichen Ring müssen wir um 60° verdrehen, damit die Pseudoinverse regulär bleibt. Die Anordnung der Lautsprecherpositionen in der Matrix ist willkürlich. Wichtig bei der Wahl der Positionen ist, dass man für jeden Lautsprecherring entweder einzigartige Positionen in der zx-Ebene findet (Azimut = 0° oder 180°), oder Paare mit gegengleichen Azimut wählt. Diese spiegelsymmetrischen Paare (bezüglich der zx-Ebene) werden weiters analog ihres Azimutwinkelvorgezeichens so bezeichnet (zB. LS_3 und $-LS_3$).

$$\varphi_j = -\varphi_k \quad \theta_j = \theta_k$$

Die Kodiervorschriften 2.Ordnung, 3D lauten:

$$Y_{0,0} = 1 = \underline{\cos}(0\varphi)$$

$$Y_{1,+1} = \cos(\theta) \cos(\varphi) = \cos(\theta) \underline{\cos}(\varphi)$$

$$Y_{1,-1} = \cos(\theta) \sin(\varphi) = \cos(\theta) \underline{\sin}(\varphi)$$

$$Y_{1,0} = \sin(\theta) = \sin(\theta) \underline{\cos}(0\varphi)$$

$$Y_{2,+2} = \frac{\sqrt{3}}{2} \cos^2(\theta) \cos(2\varphi) = \frac{\sqrt{3}}{2} \cos^2(\theta) \underline{\cos}(2\varphi)$$

$$Y_{2,-2} = \frac{\sqrt{3}}{2} \cos^2(\theta) \sin(2\varphi) = \frac{\sqrt{3}}{2} \cos^2(\theta) \underline{\sin}(2\varphi)$$

$$Y_{2,+1} = \sqrt{3} \cos(\theta) \sin(\theta) \cos(\varphi) = \sqrt{3} \cos(\theta) \sin(\theta) \underline{\cos}(\varphi)$$

$$Y_{2,-1} = \sqrt{3} \cos(\theta) \sin(\theta) \sin(\varphi) = \sqrt{3} \cos(\theta) \sin(\theta) \underline{\sin}(\varphi)$$

$$Y_{2,0} = \frac{1}{2} (3\sin(\theta)^2 - 1) = \frac{1}{2} (3\sin(\theta)^2 - 1) \underline{\cos}(0\varphi)$$

Dabei wurden Vorschriften, die unabhängig vom Azimutwinkel φ sind mit dem Ausdruck $\cos(0\varphi)$ erweitert. Nun sieht man, dass alle Ausdrücke entweder symmetrisch oder antisymmetrisch bezüglich φ sind ($\cos(k\varphi)$ -Term oder $\sin(k\varphi)$ -Term). Diese Eigenschaft wird für die Zukunft abgekürzt verwendet: c für symmetrische, s für antisymmetrische Verhältnisse.

Die Kodiervorschriften werden zu einem spaltenweise angeordneten Operator zusammengefasst: $Y\left(\frac{\theta}{\varphi}\right)$ (2.1.2)

Nun kodieren wir die 13 Lautsprecher-Positionen in den Ambisonic-Bereich und arrangieren die Ergebnisse spaltenweise in eine Matrix C (Dimension : 9×13).

$$C = Y(\Phi) \quad (2.1.3)$$

$$C = \begin{pmatrix} \begin{array}{c|cccc|cccc|cccc|cccc} & LS_1 & LS_2 & LS_3 & -LS_3 & LS_5 & LS_6 & LS_7 & -LS_7 & -LS_6 & LS_{10} & LS_{11} & -LS_{11} & LS_{13} \\ \hline c & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ c & 0 & 0.707 & -0.354 & -0.354 & 1 & 0.309 & -0.809 & -0.809 & 0.309 & -0.707 & 0.354 & 0.354 & 0 \\ s & 0 & 0 & 0.612 & -0.612 & 0 & 0.951 & 0.588 & -0.588 & -0.951 & 0 & 0.612 & -0.612 & 0 \\ c & 1 & 0.707 & 0.707 & 0.707 & 0 & 0 & 0 & 0 & 0 & -0.707 & -0.707 & -0.707 & -1 \\ c & 0 & 0.433 & -0.217 & -0.217 & 0.866 & -0.701 & 0.268 & 0.268 & -0.701 & 0.433 & -0.217 & -0.217 & 0 \\ s & 0 & 0 & -0.375 & 0.375 & 0 & 0.509 & -0.824 & 0.824 & -0.509 & 0 & 0.375 & -0.375 & 0 \\ c & 0 & 0.866 & -0.433 & -0.433 & 0 & 0 & 0 & 0 & 0 & 0.866 & -0.433 & -0.433 & 0 \\ s & 0 & 0 & 0.75 & -0.75 & 0 & 0 & 0 & 0 & 0 & 0 & -0.75 & 0.75 & 0 \\ c & 1 & 0.25 & 0.25 & 0.25 & -0.5 & -0.5 & -0.5 & -0.5 & -0.5 & 0.25 & 0.25 & 0.25 & 1 \end{array} \end{pmatrix}$$

Die Matrix C besitzt nun folgende Symmetrieeigenschaften: bei spaltenweisem Vergleich zweier symmetrischer Lautsprecher (zB. LS_{11} und $-LS_{11}$) sind die Elemente in einer c-Zeile symmetrisch oder Null, die Elemente in einer s-Zeile sind antisymmetrisch oder Null.

$$c_{ij} = c_{ik} \quad \forall \quad i \in \mathbf{c}\text{-Zeile}, \varphi_j = -\varphi_k \wedge \theta_j = \theta_k$$

$$c_{ij} = -c_{ik} \quad \forall \quad i \in \mathbf{s}\text{-Zeile}, \varphi_j = -\varphi_k \wedge \theta_j = \theta_k$$

Nun bilden wir die Pseudoinverse dieser überbestimmten, rechteckigen Matrix:

$$\tilde{D} = C^T \times (C \times C^T)^{-1} \quad (2.1.4)$$

(Dimension : 13×9).

$$\tilde{D} = \begin{pmatrix} & c & c & s & c & c & s & c & s & c \\ \begin{matrix} LS_1 \\ LS_2 \\ LS_3 \\ -LS_3 \\ LS_5 \\ LS_6 \\ LS_7 \\ -LS_7 \\ -LS_6 \\ LS_{10} \\ LS_{11} \\ -LS_{11} \\ LS_{13} \end{matrix} & \begin{matrix} 0.0509 \\ 0.0732 \\ 0.0732 \\ 0.0732 \\ 0.0894 \\ 0.0894 \\ 0.0894 \\ 0.0894 \\ 0.0894 \\ 0.0732 \\ 0.0732 \\ 0.0732 \\ 0.0301 \end{matrix} & \begin{matrix} 0 \\ 0.1768 \\ -0.0884 \\ -0.0884 \\ 0.25 \\ 0.0773 \\ -0.2023 \\ -0.2023 \\ 0.0773 \\ -0.1768 \\ 0.0884 \\ 0.0884 \\ 0 \end{matrix} & \begin{matrix} 0 \\ 0 \\ 0.1531 \\ -0.1531 \\ 0 \\ 0.2378 \\ 0.1469 \\ -0.1469 \\ -0.2378 \\ 0 \\ 0.1531 \\ -0.1531 \\ 0 \end{matrix} & \begin{matrix} 0.2 \\ 0.1414 \\ 0.1414 \\ 0.1414 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ -0.1414 \\ -0.1414 \\ -0.1414 \\ -0.2 \end{matrix} & \begin{matrix} 0 \\ 0 \\ 0 \\ 0 \\ 0.4619 \\ -0.3737 \\ 0.1427 \\ 0.1427 \\ -0.3737 \\ 0 \\ 0 \\ 0 \\ 0 \end{matrix} & \begin{matrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0.2715 \\ -0.4393 \\ 0.4393 \\ -0.2715 \\ 0 \\ 0 \\ 0 \\ 0 \end{matrix} & \begin{matrix} 0 \\ 0.3849 \\ -0.1925 \\ -0.1925 \\ -0.2309 \\ 0.1868 \\ -0.0714 \\ -0.0714 \\ 0.1868 \\ 0.3849 \\ -0.1925 \\ -0.1925 \\ 0 \end{matrix} & \begin{matrix} 0 \\ 0 \\ 0.3333 \\ -0.3333 \\ 0 \\ 0.1357 \\ -0.2196 \\ 0.2196 \\ -0.1357 \\ 0 \\ -0.3333 \\ 0.3333 \\ 0 \end{matrix} & \begin{matrix} 0.2602 \\ 0.0488 \\ 0.0488 \\ 0.0488 \\ -0.1626 \\ -0.1626 \\ -0.1626 \\ -0.1626 \\ -0.1626 \\ 0.0488 \\ 0.0488 \\ 0.0488 \\ 0.2602 \end{matrix} \end{pmatrix}$$

Die Symmetrie- bzw. Antisymmetrie-Verhältnisse blieben transponiert erhalten. Die Matrix \tilde{D} besitzt die Eigenschaft, dass bei zeilenweisen Vergleich von symmetrischen Lautsprechern (zB. LS_6 und $-LS_6$) die Elemente in einer c-Spalte zueinander symmetrisch, die Elemente in einer s-Spalte zueinander antisymmetrisch sind. Es können noch zusätzlich einige Elemente den Wert 0 erhalten. Diese erfüllen trivialerweise sowohl die Symmetrie- als auch die Antisymmetrie-Bedingung.

Nun wird noch die Gewichtung der Ambisonic-Ordnungen durchgeführt

(z.B. mit $ow_0 = 1$, $ow_1 = 1$, $ow_2 = 0.5$).

$$D = \tilde{D} \times OW \quad (2.1.5)$$

(Dimension von D : 13×9 , Dimension von OW : 9×9).

$$OW = \begin{pmatrix} ow_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & ow_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & ow_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & ow_1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & ow_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & ow_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & ow_2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & ow_2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & ow_2 \end{pmatrix}$$

$$D = \begin{pmatrix} & c & c & s & c & c & s & c & s & c \\ LS_1 & 0.0509 & 0 & 0 & 0.2 & 0 & 0 & 0 & 0 & 0.1301 \\ LS_2 & 0.0732 & 0.1768 & 0 & 0.1414 & 0 & 0 & 0.1925 & 0 & 0.0244 \\ LS_3 & 0.0732 & -0.0884 & 0.1531 & 0.1414 & 0 & 0 & -0.0962 & 0.1667 & 0.0244 \\ -LS_3 & 0.0732 & -0.0884 & -0.1531 & 0.1414 & 0 & 0 & -0.0962 & -0.1667 & 0.0244 \\ LS_5 & 0.0894 & 0.25 & 0 & 0 & 0.2309 & 0 & -0.1155 & 0 & -0.0813 \\ LS_6 & 0.0894 & 0.0773 & 0.2378 & 0 & -0.1868 & 0.1357 & 0.0934 & 0.0679 & -0.0813 \\ LS_7 & 0.0894 & -0.2023 & 0.1469 & 0 & 0.0714 & -0.2196 & -0.0357 & -0.1098 & -0.0813 \\ -LS_7 & 0.0894 & -0.2023 & -0.1469 & 0 & 0.0714 & 0.2196 & -0.0357 & 0.1098 & -0.0813 \\ -LS_6 & 0.0894 & 0.0773 & -0.2378 & 0 & -0.1868 & -0.1357 & 0.0934 & -0.0679 & -0.0813 \\ LS_{10} & 0.0732 & -0.1768 & 0 & -0.1414 & 0 & 0 & 0.1925 & 0 & 0.0244 \\ LS_{11} & 0.0732 & 0.0884 & 0.1531 & -0.1414 & 0 & 0 & -0.0962 & -0.1667 & 0.0244 \\ -LS_{11} & 0.0732 & 0.0884 & -0.1531 & -0.1414 & 0 & 0 & -0.0962 & 0.1667 & 0.0244 \\ LS_{13} & 0.0301 & 0 & 0 & -0.2 & 0 & 0 & 0 & 0 & 0.1301 \end{pmatrix}$$

Auch nach dieser Multiplikation behält Matrix D die zuvor bei Matrix \tilde{D} beschriebenen Symmetrieeigenschaften. Da diese Eigenschaften bisher nur deduktiv erwiesen wurden, werden sie in der praktischen Implementierung auch erst bei Matrix D geprüft.

$$d_{ji} = d_{ki} \quad \forall \quad i \in \text{c-Spalte}, \varphi_j = -\varphi_k \wedge \theta_j = \theta_k$$

$$d_{ji} = -d_{ki} \quad \forall \quad i \in \text{s-Spalte}, \varphi_j = -\varphi_k \wedge \theta_j = \theta_k$$

Die jeweiligen Außenohr-Impulsantworten H_{Li} bzw. H_{Ri} der virtuellen Lautsprecher-Einfallswinkelrichtungen bezüglich linken bzw. rechten Ohrs arrangieren wir als Spaltenvektoren von jeweils n Samples in 2 Matrix-Sätzen:

$$H_L = (H_{L1} \ H_{L2} \ H_{L3} \ H_{L4} \ H_{L5} \ H_{L6} \ H_{L7} \ H_{L8} \ H_{L9} \ H_{L10} \ H_{L11} \ H_{L12} \ H_{L13})$$

$$H_R = (H_{R1} \ H_{R2} \ H_{R3} \ H_{R4} \ H_{R5} \ H_{R6} \ H_{R7} \ H_{R8} \ H_{R9} \ H_{R10} \ H_{R11} \ H_{R12} \ H_{R13}).$$

Die Dimension von H_L bzw. von H_R beträgt: $n \times 13$.

2.2 Herleitung der beiden reduzierten HRIR-Sätze

Nun bilden wir die reduzierten HRIR-Sätze RH_L und RH_R mit Dimension : $n \times 9$.

Dies geschieht durch Multiplikation des jeweiligen HRIR-Satzes mit der Dekodiermatrix D . RH_{Li} und RH_{Ri} stellen die einzelnen Impulsantworten des reduzierten HRIR-Satzes bezüglich der Ambisonic-Kanäle zum linken bzw. zum rechten Außenohr in Form von Spaltenvektoren dar. (2.2.1) (2.2.2)

$$RH_L = H_L \times D = (RH_{L_1} \ RH_{L_2} \ RH_{L_3} \ RH_{L_4} \ RH_{L_5} \ RH_{L_6} \ RH_{L_7} \ RH_{L_8} \ RH_{L_9})$$

$$RH_R = H_R \times D = (RH_{R_1} \ RH_{R_2} \ RH_{R_3} \ RH_{R_4} \ RH_{R_5} \ RH_{R_6} \ RH_{R_7} \ RH_{R_8} \ RH_{R_9})$$

Nun werden die 9 Ambisonic-Kanalsignale mit diesen reduzierten HRIR-Sätzen gefaltet, bzw. es werden mittels Blockfilter-FFT (overlap and add) die komplexen Spektren der Ambisonic-Kanäle mit den jeweiligen Fouriertransformierten der reduzierten HRIR-Sätzen komplex multipliziert und wieder in den Zeitbereich zurück transformiert.

$$S_{Left_Ear} = S_{Ambi} * RH_L \quad (2.2.3)$$

$$S_{Right_Ear} = S_{Ambi} * RH_R \quad (2.2.4)$$

Wobei S_{Ambi} die Dimension von $9 \times m$ hat. Der Matrizenfaltungsoperator „ $*$ “ bedeutet: Faltung der i-ten Zeile der Matrix S_{Ambi} mit der i-ten Spalte der Matrix RH_L (bzw. RH_R) ergibt die i-te Zeile der Teilkomponentensignalmatrix S_{Left_Ear} für das linke Ohr bzw. S_{Right_Ear} für das rechte Ohr (ähnlich wie bei Multiplikation mit einer Diagonalmatrix). Beide Ergebnismatrizen besitzen die Dimension von $9 \times (m + n - 1)$.

$$\begin{pmatrix} S_{Left_Ear_1} \\ S_{Left_Ear_2} \\ S_{Left_Ear_3} \\ S_{Left_Ear_4} \\ S_{Left_Ear_5} \\ S_{Left_Ear_6} \\ S_{Left_Ear_7} \\ S_{Left_Ear_8} \\ S_{Left_Ear_9} \end{pmatrix} = \begin{pmatrix} S_{Ambi_1} \\ S_{Ambi_2} \\ S_{Ambi_3} \\ S_{Ambi_4} \\ S_{Ambi_5} \\ S_{Ambi_6} \\ S_{Ambi_7} \\ S_{Ambi_8} \\ S_{Ambi_9} \end{pmatrix} * (RH_{L_1} \quad RH_{L_2} \quad RH_{L_3} \quad RH_{L_4} \quad RH_{L_5} \quad RH_{L_6} \quad RH_{L_7} \quad RH_{L_8} \quad RH_{L_9})$$

$$\begin{pmatrix} S_{Right_Ear_1} \\ S_{Right_Ear_2} \\ S_{Right_Ear_3} \\ S_{Right_Ear_4} \\ S_{Right_Ear_5} \\ S_{Right_Ear_6} \\ S_{Right_Ear_7} \\ S_{Right_Ear_8} \\ S_{Right_Ear_9} \end{pmatrix} = \begin{pmatrix} S_{Ambi_1} \\ S_{Ambi_2} \\ S_{Ambi_3} \\ S_{Ambi_4} \\ S_{Ambi_5} \\ S_{Ambi_6} \\ S_{Ambi_7} \\ S_{Ambi_8} \\ S_{Ambi_9} \end{pmatrix} * (RH_{R_1} \quad RH_{R_2} \quad RH_{R_3} \quad RH_{R_4} \quad RH_{R_5} \quad RH_{R_6} \quad RH_{R_7} \quad RH_{R_8} \quad RH_{R_9})$$

2.3 Herleitung der Verwendung nur eines reduzierten HRIR-Satzes

Aus den Symmetrieverhältnissen der Lautsprecherpositionen (Glg. 2.1.1) und des menschlichen Gehörs lassen sich die Außenohrimpulsantworten bezüglich des rechten Ohrs durch Impulsantworten des linken Ohrs ersetzen:

$$H_R = (H_{L1} \quad H_{L2} \quad H_{L4} \quad H_{L3} \quad H_{L5} \quad H_{L9} \quad H_{L8} \quad H_{L7} \quad H_{L6} \quad H_{L10} \quad H_{L12} \quad H_{L11} \quad H_{L13})$$

wenn folgendes gilt:

$$\varphi_3 = -\varphi_4 \wedge \theta_3 = \theta_4$$

$$\varphi_6 = -\varphi_9 \wedge \theta_6 = \theta_9$$

$$\varphi_7 = -\varphi_8 \wedge \theta_7 = \theta_8$$

$$\varphi_{11} = -\varphi_{12} \wedge \theta_{11} = \theta_{12}$$

$$RH_L = H_L \times D = (RH_{L_{-1}} \quad RH_{L_{-2}} \quad RH_{L_{-3}} \quad RH_{L_{-4}} \quad RH_{L_{-5}} \quad RH_{L_{-6}} \quad RH_{L_{-7}} \quad RH_{L_{-8}} \quad RH_{L_{-9}})$$

$$RH_R = H_R \times D = (RH_{R_{-1}} \quad RH_{R_{-2}} \quad RH_{R_{-3}} \quad RH_{R_{-4}} \quad RH_{R_{-5}} \quad RH_{R_{-6}} \quad RH_{R_{-7}} \quad RH_{R_{-8}} \quad RH_{R_{-9}})$$

Nach der Matrizenmultiplikation (Glg. 2.2.1, 2.2.2) erhalten wir die Summenterme der Spalten der reduzierten HRIR-Sätze bezüglich des linken und rechten Außenohrs.

$$RH_{L_{-1}} = 0.0569H_{L1} + 0.0732H_{L2} + 0.0732H_{L3} + 0.0732H_{L4} + 0.0894H_{L5} + 0.0894H_{L6} + 0.0894H_{L7} \\ + 0.0894H_{L8} + 0.0894H_{L9} + 0.0732H_{L10} + 0.0732H_{L11} + 0.0732H_{L12} + 0.0569H_{L13}$$

$$RH_{R_{-1}} = 0.0569H_{L1} + 0.0732H_{L2} + 0.0732H_{L4} + 0.0732H_{L3} + 0.0894H_{L5} + 0.0894H_{L9} + 0.0894H_{L8} \\ + 0.0894H_{L7} + 0.0894H_{L6} + 0.0732H_{L10} + 0.0732H_{L12} + 0.0732H_{L11} + 0.0569H_{L13}$$

$$RH_{L_{-2}} = 0H_{L1} + 0.1768H_{L2} - 0.0884H_{L3} - 0.0884H_{L4} + 0.25H_{L5} + 0.0773H_{L6} - \\ - 0.2023H_{L7} + 0.2023H_{L8} + 0.0773H_{L9} - 0.1768H_{L10} + 0.0884H_{L11} + 0.0884H_{L12} + 0H_{L13}$$

$$RH_{R_{-2}} = 0H_{L1} + 0.1768H_{L2} - 0.0884H_{L4} - 0.0884H_{L3} + 0.25H_{L5} + 0.0773H_{L9} - \\ - 0.2023H_{L8} + 0.2023H_{L7} + 0.0773H_{L6} - 0.1768H_{L10} + 0.0884H_{L12} + 0.0884H_{L11} + 0H_{L13}$$

$$RH_{L_{-3}} = 0H_{L1} + 0H_{L2} + 0.1531H_{L3} - 0.1531H_{L4} + 0H_{L5} + 0.2378H_{L6} + 0.1469H_{L7} \\ - 0.1469H_{L8} - 0.2378H_{L9} + 0H_{L10} + 0.1531H_{L11} - 0.1531H_{L12} + 0H_{L13}$$

$$RH_{R_{-3}} = 0H_{L1} + 0H_{L2} + 0.1531H_{L4} - 0.1531H_{L3} + 0H_{L5} + 0.2378H_{L9} + 0.1469H_{L8} \\ - 0.1469H_{L7} - 0.2378H_{L6} + 0H_{L10} + 0.1531H_{L12} - 0.1531H_{L11} + 0H_{L13}$$

.....

$$RH_{L_{-8}} = 0H_{L1} + 0H_{L2} + 0.1667H_{L3} - 0.1667H_{L4} + 0H_{L5} + 0.0679H_{L6} - 0.1098H_{L7} \\ + 0.1098H_{L8} - 0.0679H_{L9} + 0H_{L10} - 0.1667H_{L11} + 0.1667H_{L12} + 0H_{L13}$$

$$RH_{R_{-8}} = 0H_{L1} + 0H_{L2} + 0.1667H_{L4} - 0.1667H_{L3} + 0H_{L5} + 0.0679H_{L9} - 0.1098H_{L8}$$

$$+0.1098H_{L7} - 0.0679H_{L6} + 0H_{L10} - 0.1667H_{L12} + 0.1667H_{L11} + 0H_{L13}$$

$$RH_{L_9} = 0.1301H_{L1} + 0.0244H_{L2} + 0.0244H_{L3} + 0.0244H_{L4} - 0.0813H_{L5} - 0.0813H_{L6} - 0.0813H_{L7} \\ - 0.0813H_{L8} - 0.0813H_{L9} + 0.0244H_{L10} + 0.0244H_{L11} + 0.0244H_{L12} + 0.1301H_{L13}$$

$$RH_{R_9} = 0.1301H_{L1} + 0.0244H_{L2} + 0.0244H_{L4} + 0.0244H_{L3} - 0.0813H_{L5} - 0.0813H_{L9} - 0.0813H_{L8} \\ - 0.0813H_{L7} - 0.0813H_{L6} + 0.0244H_{L10} + 0.0244H_{L12} + 0.0244H_{L11} + 0.1301H_{L13}$$

Hier sehen wir schon, dass $RH_{R_1} = +RH_{L_1}$, $RH_{R_2} = +RH_{L_2}$, $RH_{R_3} = -RH_{L_3}$,

$RH_{R_8} = -RH_{L_8}$ und $RH_{R_9} = +RH_{L_9}$ entspricht.

Es besteht somit Symmetrie bezüglich gleicher c-Spalten, Antisymmetrie bezüglich gleicher s-Spalten in den beiden reduzierten HRIR-Satzmatrizen.

Zum Schluss werden die einzelnen Signalkomponenten der virtuellen Lautsprecher an beiden Ohren aufsummiert und somit das Stereosignal erzeugt. Dies wird mittels der reduzierten Binauraldekodiermatrix B_R erledigt.

$$S_{LR} = B_R \times S_{Left_Ear} \quad (2.3.1)$$

Da wir uns bei der Wahl des HRIR-Satzes auf das linke Ohr entschieden haben, wird die Binauraldekodiermatrix B_R in ihrer ersten Zeile nur Einser-Elemente besitzen. Die zweite Zeile ergibt eine Folge von positiven und negativen Einser-Elementen, die dem Muster der c- und s-Terme der Kodiervorschrift (Glg. 2.1.2) entspricht ($c \hat{=} +1, s \hat{=} -1$).

$$\begin{pmatrix} S_L \\ S_R \end{pmatrix} = \begin{pmatrix} +1 & +1 & +1 & +1 & +1 & +1 & +1 & +1 & +1 \\ +1 & +1 & -1 & +1 & +1 & -1 & +1 & -1 & +1 \end{pmatrix} \times \begin{pmatrix} S_{Left_Ear_1} \\ S_{Left_Ear_2} \\ S_{Left_Ear_3} \\ S_{Left_Ear_4} \\ S_{Left_Ear_5} \\ S_{Left_Ear_6} \\ S_{Left_Ear_7} \\ S_{Left_Ear_8} \\ S_{Left_Ear_9} \end{pmatrix}$$

S_{Left_Ear} hat die Dimension von $9 \times (m+n-1)$, B hat die Dimension von 2×9 und S_{LR} hat die Dimension von $2 \times (m+n-1)$.

Zur Überprüfung der Dimensionen noch einmal:

$$\begin{aligned}
D &= \tilde{D} \times OW & (13 \times 9) &= (13 \times 9) \times (9 \times 9) \\
RH_L &= H_L \times D & (n \times 9) &= (n \times 13) \times (13 \times 9) \\
S_{Left_Ear} &= S_{Ambi} * RH_L & (9 \times [m + n - 1]) &= (9 \times m) * (n \times 9) \\
S_{LR} &= B \times S_{Left_Ear} & (2 \times [m + n - 1]) &= (2 \times 9) \times (9 \times [m + n - 1])
\end{aligned}$$

2.4 Optimales Lautsprecher-Layout

Ein auf einer Kugeloberfläche optimal gleichmäßig verteiltes Lautsprecher-Layout, das für Dekodierungen bis maximal 4. Ordnung genügt, sei hier präsentiert:

32 virtuelle Lautsprecherpositionen in den Mittelpunkt der Fünfecke bzw. Sechsecke des abgestumpften Ikosaeders (Fußball):

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
θ	90	52.7	52.7	52.7	52.7	52.7	26	26	26	26	26	11.6	11.6	11.6	11.6	11.6
φ	0	0	72	144	216	288	36	108	180	252	324	0	72	144	216	288
i	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
θ	-11.6	-11.6	-11.6	-11.6	-11.6	-26	-26	-26	-26	-26	-52.7	-52.7	-52.7	-52.7	-52.7	-90
φ	36	108	180	252	324	0	72	144	216	288	36	108	180	252	324	0

verwendete HRIR-wave-Dateien von Kemar :

1: L90e000a.wav	9: L26e180a.wav	17: L-12e036a.wav	25: L-26e216a.wav
2: L53e000a.wav	10: L26e252a.wav	18: L-12e108a.wav	26: L-26e288a.wav
3: L53e072a.wav	11: L26e324a.wav	19: L-12e180a.wav	27: L-53e036a.wav
4: L53e108a.wav	12: L12e000a.wav	20: L-12e252a.wav	28: L-53e108a.wav
5: L53e216a.wav	13: L12e072a.wav	21: L-12e324a.wav	29: L-53e180a.wav
6: L53e288a.wav	14: L12e144a.wav	22: L-26e000a.wav	30: L-53e252a.wav
7: L26e036a.wav	15: L12e216a.wav	23: L-26e072a.wav	31: L-53e324a.wav
8: L26e108a.wav	16: L12e288a.wav	24: L-26e144a.wav	32: L-90e000a.wav

Die originalen, diffusfeldentzerrten HRIR-Dateien von KEMAR sind Stereo-Wav-Dateien mit 2 mal 128 Samples für linken und rechten Kanal, wobei man den Azimutwinkel leider in mathematisch negativer Richtung zählte. Die Auflösung der Elevation beträgt 10°, beginnend von +90° bis -40°. Die Auflösung des Azimuts

beträgt im Äquatorbereich 5° , und wird Richtung Nord- und Südpol größer. Im Allgemeinen wurde zwischen 4 Dateien linear interpoliert, um eine Datei mit den exakten Koordinaten zu berechnen. Es wurde dabei der Konvertierungs-Patch *convert_kemar.pd* mit dem Datensatz *convert_kemar.txt* verwendet. Der Name der HRIR-Datei (zB. L-26e288a.wav) ist folgendermaßen aufgebaut: *L* steht für links, die nächsten Ziffern, die mit *e* abgeschlossen sind stehen für die Elevation in Grad, die letzten Ziffern, die mit *a* abgeschlossen sind stehen für Azimut in Grad. In der Datensatz-Datei *convert_kemar.txt* stehen folgende Instruktionen:

five_ring hrir_1.wav hrir_2.wav hrir_3.wav hrir_4.wav k_az k_el hrir_5.wav hrir_6.wav

Befehl *five_ring* steht für Berechnung eines Lautsprechers in einem Fünfering;

hrir_1.wav ist die HRIR mit Elevation \leq Zielelevation und Azimut \leq Zielazimut;

hrir_2.wav ist die HRIR mit Elevation \geq Zielelevation und Azimut \leq Zielazimut;

hrir_3.wav ist die HRIR mit Elevation \leq Zielelevation und Azimut \geq Zielazimut;

hrir_4.wav ist die HRIR mit Elevation \geq Zielelevation und Azimut \geq Zielazimut;

(diese Relationen beziehen sich auf die Betragszahlen der Elevation)

k_az ist der Interpolationsfaktor bezüglich der Azimutwinkel (zwischen 0 ... 1);

k_el ist der Interpolationsfaktor bezüglich der Elevationswinkel (zwischen 0 ... 1);

hrir_5.wav ist die HRIR-Zieldatei mit Azimut $0^\circ \dots 180^\circ$ und auf das linke Ohr bezogen;

hrir_6.wav ist die HRIR-Zieldatei mit Azimut $181^\circ \dots 359^\circ$ (linkes Ohr bezogen).

Die Südpol-HRIR berechnet sich aus dem Mittel von den beiden HRIRs auf Elevation -40° und Azimut 0° und 180° , der Mittelwert wird um 6 dB abgeschwächt und bei 2500 Hz tiefpassgefiltert. Die HRIRs auf Elevation -53° werden von den HRIRs auf Elevation -40° abgeleitet.

2.5 Messergebnisse

Die Pegel wurden als Spannung an den Lautsprechern gemessen; 1 V Gleichspannung gilt als 0 dB Referenzpegel des Quellsignals am Eingang der Kodiereinheit. Die Pegelschwankungen wurden errechnet als Differenz zwischen Maximum- und Minimum-Pegel bei zeitlich zufällig variierten Kodierungspositionen über eine hinreichend lange Dauer.

1.Messung:

ambi_weight: $w_0 = 1, w_1 = 1; w_2 = 1; w_3 = 1; w_4 = 1;$

gemittelter Lautsprecherpegel bei w-Kanal-Speisung mit 0 dB: -30.103 dB;

maximale Abweichung der einzelnen Pegel: ± 0.017 dB;

lineare Lautsprecherpegel-Summenschwankung: $\Delta \sum_{i=1}^{32} u_i = 0$ dB;

quadratische Lautsprecherpegel-Summenschwankung: $\Delta \sqrt{\sum_{i=1}^{32} u_i^2} = 0.28$ dB;

Hauptkeule: -2.46 dB;

nächster Nachbarlautsprecher: -18.67 dB $\hat{=}$ -16.21 dB bez. Hauptkeule;

max. Nebenkeule in opposit. Hemisphäre: -16.4 dB $\hat{=}$ -13.94 dB bez. Hauptkeule;

oppositionelle Nebenkeule: -16.4 dB $\hat{=}$ -13.94 dB bez. Hauptkeule;

2.Messung:

ambi_weight: $w_0 = 1.1236; w_1 = 1.1236; w_2 = 1.1236; w_3 = 1.1236; w_4 = 0.5056;$

gemittelter Lautsprecherpegel bei w-Kanal-Speisung mit 0 dB: -29.09 dB;

maximale Abweichung der einzelnen Pegel: ± 0.017 dB;

lineare Lautsprecherpegel-Summenschwankung: $\Delta \sum_{i=1}^{32} u_i = 0$ dB;

quadratische Lautsprecherpegel-Summenschwankung: $\Delta \sqrt{\sum_{i=1}^{32} u_i^2} = 0.255$ dB;

Hauptkeule: -3.4 dB;

nächster Nachbarlautsprecher: -14.98 dB $\hat{=}$ -11.58 dB bez. Hauptkeule;

max. Nebenkeule in opposit. Hemisphäre: -27.6 dB $\hat{=}$ -24.2 dB bez. Hauptkeule;

oppositionelle Nebenkeule: -61.9 dB $\hat{=}$ -58.5 dB bez. Hauptkeule;

3.Messung:

ambi_weight: $w_0 = 1.6667$; $w_1 = 1.5075$; $w_2 = 1.09085$; $w_3 = 0.57582$; $w_4 = 0.15915$;

gemittelter Lautsprecherpegel bei w-Kanal-Speisung mit 0 dB: -25.67 dB;

maximale Abweichung der einzelnen Pegel: ± 0.017 dB;

lineare Lautsprecherpegel-Summenschwankung: $\Delta \sum_{i=1}^{32} u_i = 0$ dB;

quadratische Lautsprecherpegel-Summenschwankung: $\Delta \sqrt{\sum_{i=1}^{32} u_i^2} = 0.088$ dB;

Hauptkeule: -5.76 dB;

nächster Nachbarlautsprecher: -12.37 dB $\hat{=}$ -6.61 dB bez. Hauptkeule;

max. Nebenkeule in opposit. Hemisphäre: -33.8 dB $\hat{=}$ -28.04 dB bez. Hauptkeule;

oppositionelle Nebenkeule: -43.2 dB $\hat{=}$ -37.44 dB bez. Hauptkeule;

4.Messung:

ambi_weight: $w_0 = 1.32767$; $w_1 = 1.33032$; $w_2 = 1.1949$; $w_3 = 0.819172$; $w_4 = 0.327934$;

gemittelter Lautsprecherpegel bei w-Kanal-Speisung mit 0 dB: -27.64 dB;

maximale Abweichung der einzelnen Pegel: ± 0.017 dB;

lineare Lautsprecherpegel-Summenschwankung: $\Delta \sum_{i=1}^{32} u_i = 0$ dB;

quadratische Lautsprecherpegel-Summenschwankung: $\Delta \sqrt{\sum_{i=1}^{32} u_i^2} = 0.174$ dB;

Hauptkeule: -4.44 dB;

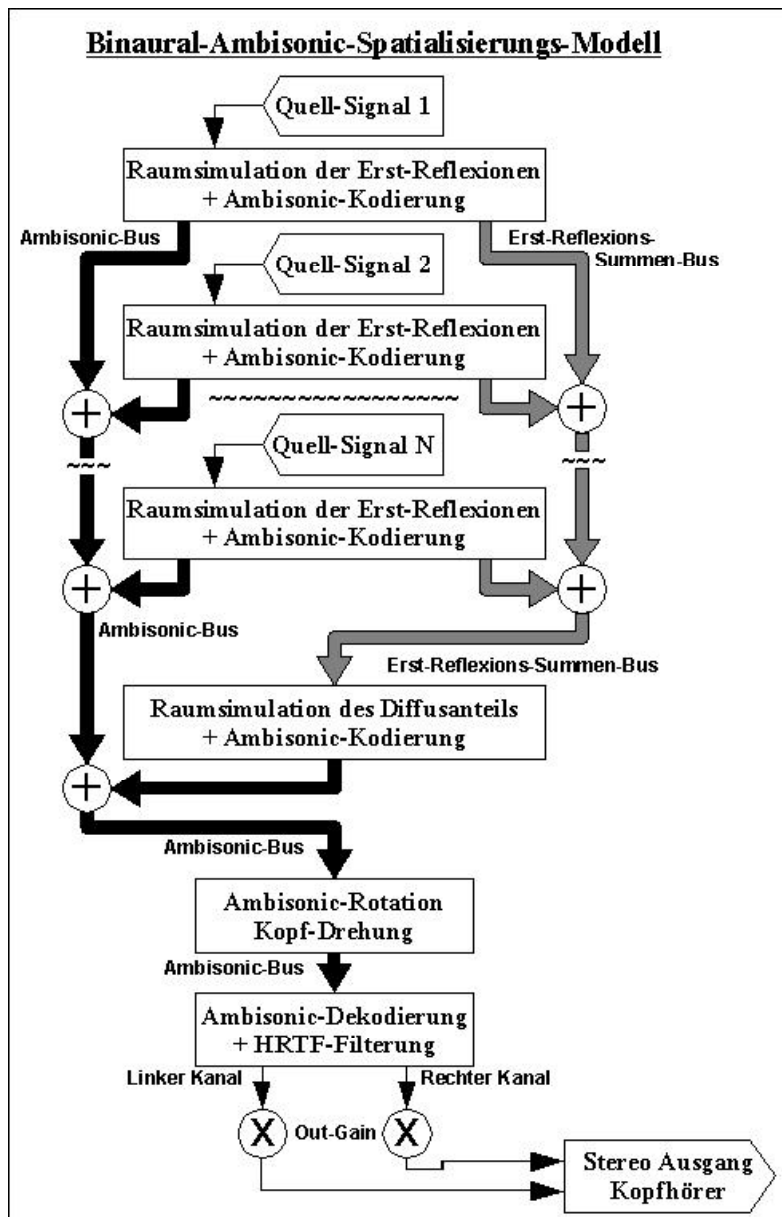
nächster Nachbarlautsprecher: -13.36 dB $\hat{=}$ -8.92 dB bez. Hauptkeule;

max. Nebenkeule in opposit. Hemisphäre: -39.62 dB $\hat{=}$ -35.18 dB bez. Hauptkeule;

oppositionelle Nebenkeule: -39.62 dB $\hat{=}$ -35.18 dB bez. Hauptkeule;

3 Aufbau und Funktionsweise des Raumsimulationsmodells

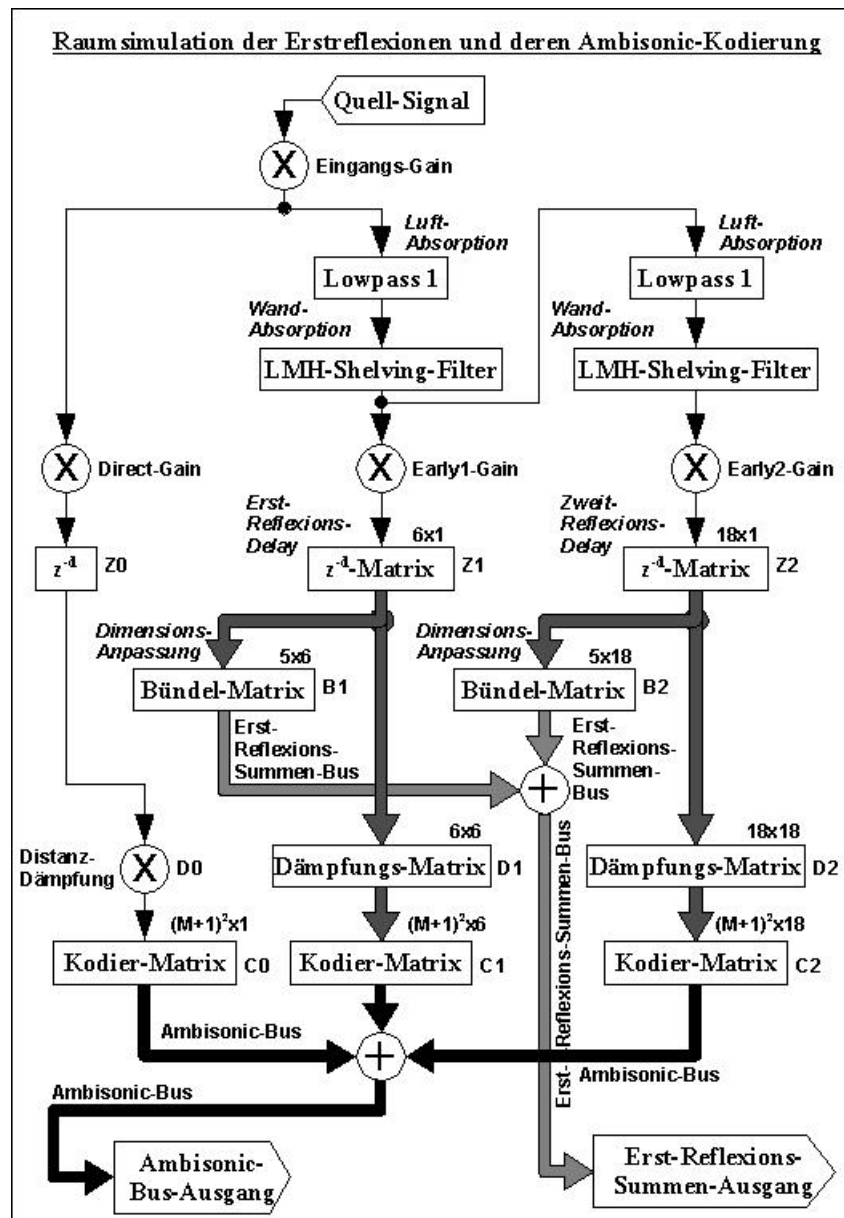
3.1 Prinzipschaltbild des Spatialisierungsmodells



Das Binaural-Ambisonic-Raummodell besteht aus N Signalquellen (Soundfile-Players) und N Erstreflexionsmodulen, die den Direktanteil, die Erst- und die Zweit-Reflexionen berechnen und in den Ambisonic-Bereich kodieren. Die Bussummierungen des Ambisonic-Busses, bzw. des Erstreflexionsbusses befinden sich in Wirklichkeit innerhalb der vorher beschriebenen Module. D.h. man muss die Bussignale immer durchschleifen. Am Ende der Kette von Erstreflexionskodiermodulen münden beide Busse in den Diffusanteilraumsimulator und Diffusanteilkodierer. Heraus kommt nur mehr der Ambisonic-Bus, der noch einmal eine Rotationsmatrix durchwandert und zuletzt im Binaural-Ambisonic-

Dekodierer landet. Ein Stereosignal verlässt dieses Modul, es wird noch die Gesamtlautstärke geregelt, und das Signal geht Richtung Kopfhörer.

3.2 Das Erst- und Zweit-Reflexionsmodul und deren Kodierung



Das Quellsignal wird zuerst einer *Eingangslautstärkenregelung* unterworfen, verzweigt sich dann Richtung *Direktanteilregelung* mit darauf folgendem Delay Z_0 einer weiteren Distanz-Dämpfung D_0 , wobei Z_0 die Schallverzögerung zwischen Quellposition und Hörerposition, D_0 die Dämpfung des Schalldrucks, bedingt durch den Abstand der beiden Positionen, darstellt. Anschließend wird dieses Signal mittels der Kodiermatrix C_0 (in Abhängigkeit der beiden Winkel Elevation und Azimut bezüglich Quelle zu Hörer) in den Ambisonic-Bereich transformiert. Der andere Teil

des Eingangssignals durchwandert ein quasi-starres *Tiefpass-Filter* erster Ordnung, welches vereinfacht die gemittelte Luftabsorption der Erstreflexionen darstellt (man müsste eigentlich nach den Erstreflexions-Delay-Abzapfungen jeweils ein dynamisches Filter setzen, das kostet aber zu viel CPU-Leistung und bringt wenig wahrnehmbaren Effekt, weil die Luftabsorption erst ab 3 – 4 kHz mit –3 dB pro 100 m einsetzt). Weiters passiert dieses Signal ein *Low-Middle-High-Shelving-Filter*, welches die Wand-, Decken- und Boden-Absorption in 3 Frequenzbänder unterteilt darstellt. Die Übergangsfrequenzen sind 450 Hz und 3500 Hz, die Absorptions-Koeffizienten bewegen sich zwischen 0 und 100 %. Dieses Signal verzweigt sich zuerst in die *Early-1-Gain-Regelung*, danach geht es durch eine Diagonal-Delay-Matrix *Z1* der Dimension 1x6 (für 2D: 1x4). Die Erstreflexionen werden einerseits mittels einer Bündelungsmatrix *B1* (Dimension 5x6, pro Spalte nur eine 1, sonst 0) dem Erstreflexionsbus angepasst und zugeführt (3D: 5 Signale, 2D: 4 Signale), andererseits zweigen sie noch Richtung Diagonaldämpfungsmatrix *D1* (3D: 6x6, 2D: 4x4) ab, welche die Abstandsabhängigkeit der Erstreflexionen realisiert. Mittels der Kodiermatrix *C1* (3D: $(M+1)^2 \times 6$, 2D: $(2M+1) \times 4$, wobei M die Ambisonic-Ordnung ist) gelangen die Erstreflexionen nun in den Ambisonic-Bereich.

Der dritte Teil des Eingangs-Signals durchwandert noch ein zweites mal das *Luftabsorptionstiefpassfilter* erster Ordnung, ein zweites mal ein *Low-Middle-High-Shelving-Filter* für die Wandabsorption und die *Early-2-Gain-Regelung*, danach geht es durch eine Diagonal-Delay-Matrix *Z2* der Dimension 1x18 (für 2D: 1x8). Diese 18 verzögerten Zweitreflexionen werden einerseits mittels einer Bündelungsmatrix *B2* (Dimension 5x18, pro Spalte nur eine 1, sonst 0) dem Erstreflexionsbus angepasst und zugeführt, andererseits zweigen sie noch Richtung Diagonaldämpfungsmatrix *D2* (3D: 18x18, 2D: 8x8) ab, welche die Abstandsabhängigkeit der Zweitreflexionen realisiert. Mittels der Kodiermatrix *C2* (3D: $(M+1)^2 \times 18$, 2D: $(2M+1) \times 8$, wobei M die Ambisonic-Ordnung ist) gelangen die Zweitreflexionen nun in den Ambisonic-Bereich.

Damit der Wertebereich für *D0* nach oben hin begrenzt bleibt, wurde folgendes Distanzmodell gewählt:

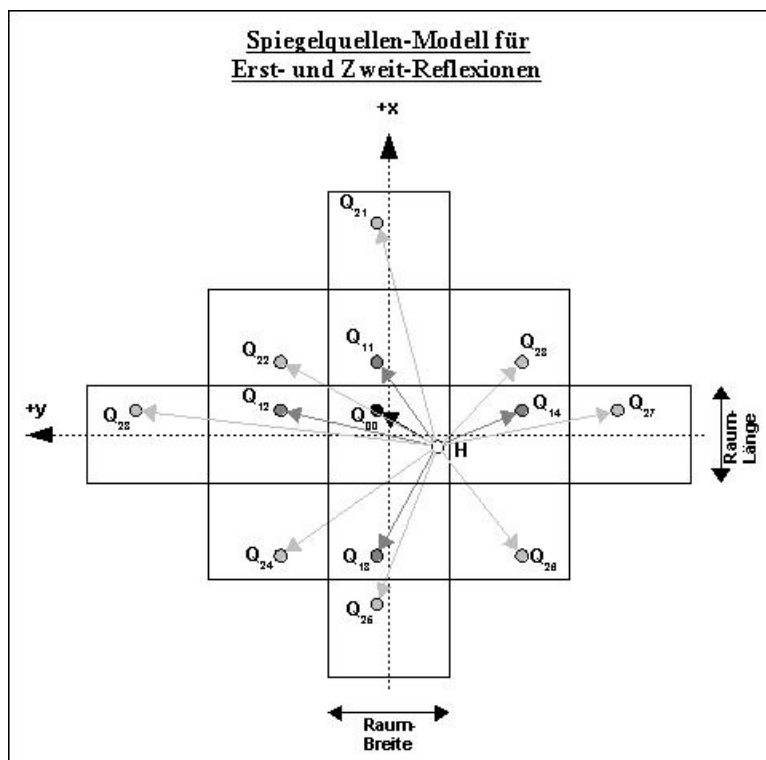
$$D0 = (r_{Ambi} / r_{Distanz})^{\exp_dist_coeff} \quad (3.2.1)$$

wobei $r_{Distanz}$ immer größer gleich r_{Ambi} sein muss. Der Ambisonic-Radius r_{Ambi} ist normalerweise der Abstand der Lautsprecher zum Kopfmittelpunkt bei der Messung

der Außenohrübertragungsfunktionen (bei den KEMAR-HRIR: 1.4 m). Hier stellt er einen Tabubereich um die Hörerin bzw. den Hörer dar, der von Quellobjekten nicht verletzt werden darf, damit die Distanzverstärkung nicht gegen unendlich steigt. Der Exponent „*exp_dist_coeff*“ ist ein experimenteller Faktor im Bereich 0.5 bis 2.0, um die Distanzverstärkung zu übertreiben bzw. zu untertreiben.

Beim Spiegelquellenmodell vernachlässigen wir die Transmissionskoeffizienten, sodass für die Reflexionskoeffizienten gilt: $\rho_{freq} = 100 - \alpha_{freq} - \tau_{freq} = 100 - \alpha_{freq}$ (3.2.2)

Die frequenzabhängigen Reflexionskoeffizienten stellen die Verstärkungsfaktoren des Filters dar. Diese sechs verzögerten Erstreflexionen ergeben sich durch ein Spiegelquellenmodell, das basierend auf einem quaderförmigen Raum, sehr einfach Distanz und Richtung der Spiegelquellen im Bezug auf den Hörer berechnen lässt.

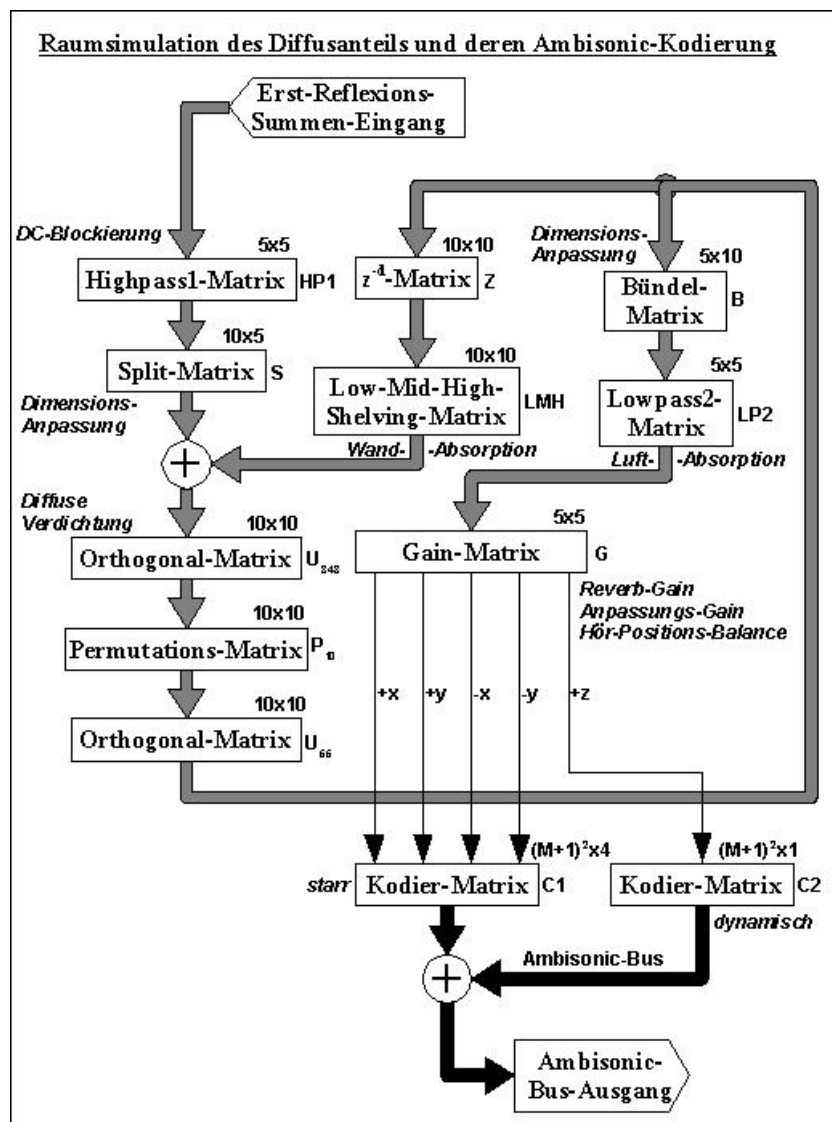


Zu den Delay-Taps sei noch zu erwähnen, es gibt 3 Arten von Erstreflexions-Implementierungen:

- 1.) statische Delays (mit dem Übergabe-Argument „*stat*“ für das Kodiermodul) und statische Kodiermatrizen, diese lassen keine Bewegung von Hörerin bzw. Hörer oder Quellen zu, sonst würden sie Unstetigkeiten im Signalverlauf erzeugen, sie sind billiger und können für eine quasistationäre Abhörsituation verwendet werden, wobei eine Hörerkopfdrehung möglich ist.

- 2.) dynamische Delays (mit dem Übergabe-Argument „dyn“) und Kodiermatrizen, jedes Delay-Tap bedient sich einer 4-Punktinterpolation, es entsteht ein Doppler-Effekt, sobald sich Quelle oder Hörerin bzw. Hörer bewegen.
- 3.) dopplereffektlose dynamische Delays (mit dem Übergabeargument „doppless“) und dynamische Matrizen, die Delay-Taps sind statisch und in zweifacher Ausführung, Veränderungen des Verzögerungswertes werden immer auf dem ausgeblendeten Delay-Tap durchgeführt, anschließend wird leistungserhaltend übergeblendet auf dieses neu eingestellte Delay-Tap. Es ergibt sich dadurch kein Doppler-Effekt, dafür aber bei Signalen mit stark stochastischem Anteil ein Phasing-Effekt.

3.3 Das Diffusnachhallmodul und dessen Kodierung



Die Signale des Erstreflexionsbusses gelangen über eine Diagonalthochpassmatrix *HP1* (3D: 5x5, 2D: 4x4), welche Gleichspannungsanteile blockiert, auf eine

Anpassungsmatrix S (3D: 10x5, 2D: 12x4), welche die Ordnung des Erstreflexionsbusses an die Ordnung des Feedback-Delay-Networks anpasst, bzw. verteilt (pro Zeile eine 1, sonst 0). Nun gelangt das Signalfeld auf ein Summierungsfeld, dessen Ausgangsfeld die erste Orthogonalmatrix U_{343} , die Permutationsmatrix P und die zweite Orthogonalmatrix U_{55} durchwandert und dabei durchmischt wird. U_{343} setzt sich diagonalförmig zusammen aus den 3 Matrizen U_3 , U_4 und U_5 , U_{55} setzt sich aus U_5 und U_5 zusammen.

$$U_{343} = \begin{pmatrix} U_3 & 0 & 0 \\ 0 & U_4 & 0 \\ 0 & 0 & U_5 \end{pmatrix}$$

$$U_{55} = \begin{pmatrix} U_5 & 0 \\ 0 & U_5 \end{pmatrix}$$

Die elementaren Orthogonalmatrizen U_3 , U_4 und U_5 werden nach dem „Householder Schema“ [8] erzeugt.

$$U_K = P_K(E_K) - \frac{2}{K} \bar{u}_K \bar{u}_K^T \quad (3.3.1)$$

K Ordnungszahl

$P_K(E_K)$ permutierte Einheitsmatrix der Dimension $K \times K$

$\bar{u}_K^T = (1 \ 1 \ \dots \ 1 \ 1)$ Zeilenvektor der Dimension $1 \times K$

$$U_3 = \frac{1}{3} \begin{pmatrix} -2 & -2 & 1 \\ 1 & -2 & -2 \\ -2 & 1 & -2 \end{pmatrix} \quad U_4 = \frac{1}{2} \begin{pmatrix} -1 & -1 & -1 & 1 \\ 1 & -1 & -1 & -1 \\ -1 & 1 & -1 & -1 \\ -1 & -1 & 1 & -1 \end{pmatrix} \quad U_5 = \frac{1}{5} \begin{pmatrix} -2 & -2 & -2 & -2 & 3 \\ 3 & -2 & -2 & -2 & -2 \\ -2 & 3 & -2 & -2 & -2 \\ -2 & -2 & 3 & -2 & -2 \\ -2 & -2 & -2 & 3 & -2 \end{pmatrix}$$

Die Eigenschaften einer Orthogonalmatrix sind:

$$U^T = U^{-1} \quad \det(U) = 1$$

Die Eigenschaften der Spalten bzw. der Zeilen einer Orthogonal-Matrix sind:

$$u_{col_i} \cdot u_{col_j} = 0 \quad u_{col_i} \cdot u_{col_i} = 1$$

$$u_{row_i} \cdot u_{row_j} = 0 \quad u_{row_i} \cdot u_{row_i} = 1$$

Die Orthogonalität bleibt auch nach einer Multiplikation mit einer andern orthogonalen Matrix erhalten:

$$U_z = U_x \times U_y$$

Die Permutationsmatrix P sieht folgendermaßen aus:

$$P_{10} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

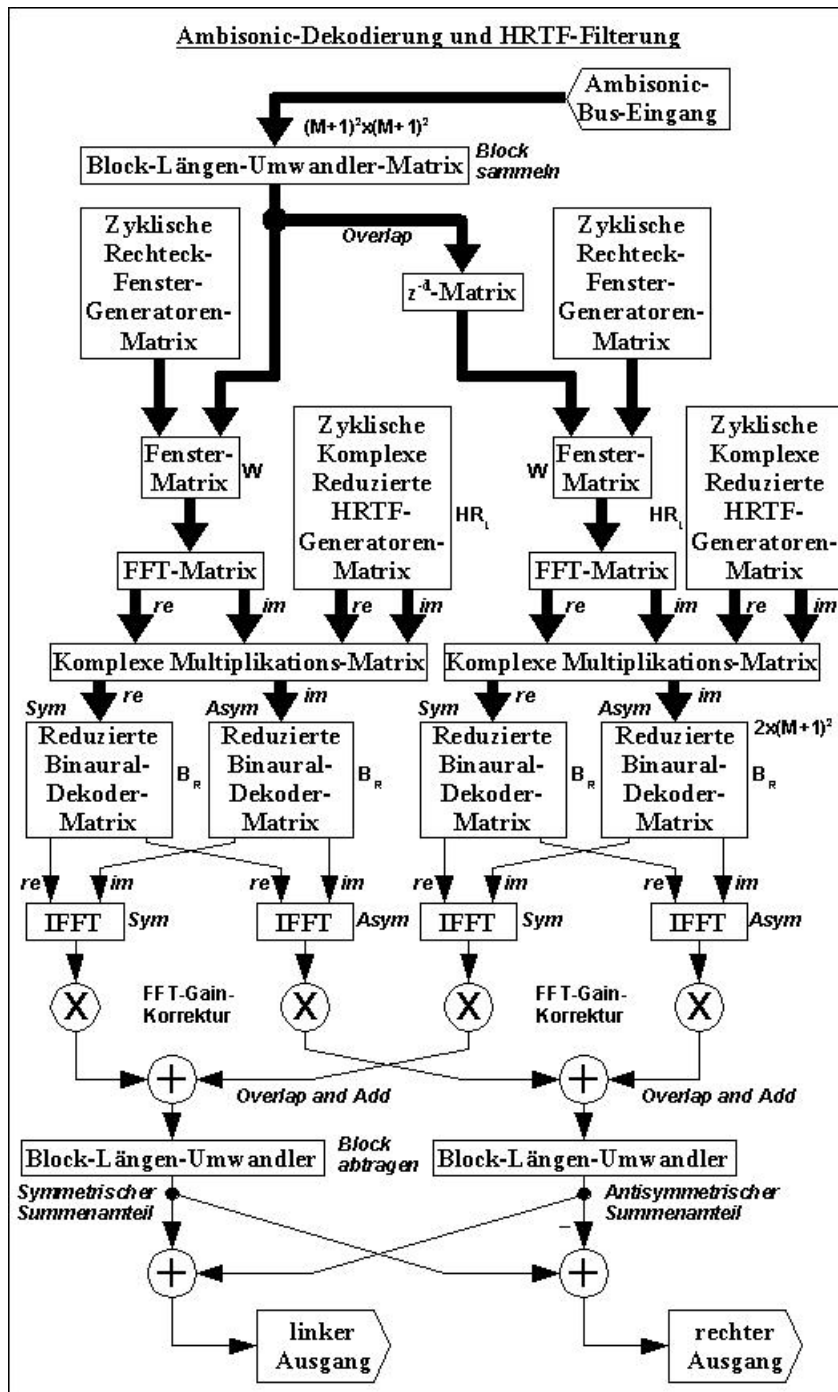
Nach Verlassen der Orthogonalmatrizen geht eine Verzweigung zu einer Anpassungsmatrix B , die die 10 internen Signale auf 5 bündelt (3D: 5x10, 2D: 4x12) und diese auf eine Diagonaltiefpassmatrix 2. Ordnung $LP2$ (2D: 4x4) treffen lässt. Diese stellt vereinfacht die gemittelte Luftabsorption dar. Dieses Fünfer-Feld geht weiter zu einer Diagonalmatrix G , die 3 Funktionen übernimmt:

- 1.) Reverb-Lautstärke (auf alle 5 Elemente gleich wirkend)
- 2.) Anpassungsverstärkung (reziprok zur Zahl der Erst- und Zweit-Reflexionen und Split-Faktor, ca. $0.02 = 1/(24 \cdot 2)$).
- 3.) Hörer-zu-Wand-Abstandseinfluss: die 5 Signale vom Diffusnachhall werden wie bei einer Pyramide von den 4 Grundkantenmitten und von der Spitze in den Ambisonic-Bereich kodiert. Wenn der Hörer sich einer der Wände oder der Decke im virtuellen Quaderraum nähert, wird der Diffusanteil aus dieser Richtung geringer geregelt und kann sogar Richtung 0 gehen, gleichzeitig neigt sich die Kodierungsrichtung der +z-Komponente immer leicht in die entgegen gesetzte verschobene x- oder y-Richtung.

Diese 3 Komponenten pro Diagonal-Element von G werden auf Message-Ebene multipliziert und dann an den Matrixmultiplizierer weiter gegeben. Schließlich trifft das Signalfeld auf die Kodiermatrizen $C1$ und $C2$ (nur mit abgeflachter 3. Ordnung) und summiert zum Ambisonic-Bus.

Der andere Zweig des Feedback-Signalfeldes führt auf eine Diagonal-Delay-Matrix Z mit exponentiell gestuften Verzögerungswerten (Grenzen: halbe kleinste Raumdimension bzw. 1,5-fachen Raumdiagonale geteilt durch Schallgeschwindigkeit). Die anschließende Shelving-Filtermatrix LMH dämpft das Signalfeld in 3 Frequenzbändern unterteilt unterschiedlich. Nun sind wir wieder beim Eingangssummierfeld angelangt.

3.4 Das Dekodiermodul und die HRTF-Filterung

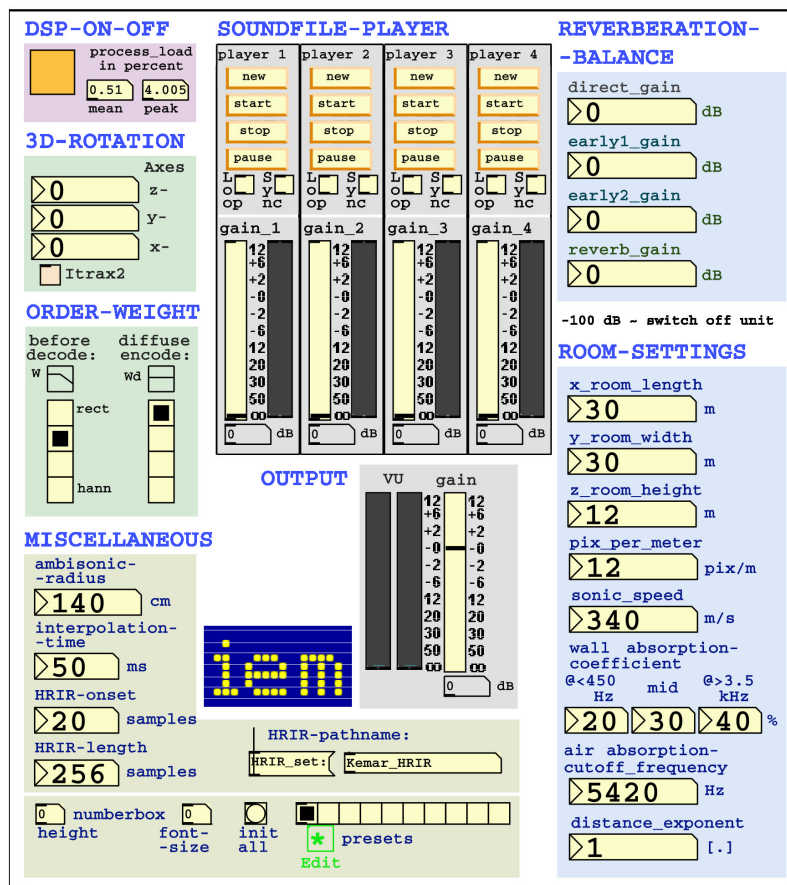


Das gesamte Modul wurde hier schematisch zweifach gezeichnet. In Pd wird diese zweifach Überlappung bzw. die Blocklängenumwandlung von 64 auf 512 und zurück durch die Objekte *block~*, *inlet~* und *outlet~* realisiert. Die Ambisonic-Bus-Signale gelangen vom blocksammelnden Eingang auf eine diagonale Rechteck-Fensterungsmatrix W . Von dort gelangen sie auf eine diagonal angeordnete reale FFT-Matrix, deren Ausgangssignale sich in ihrer Anzahl verdoppeln (Real- und Imaginär-Anteil). Die komplexen, in den Frequenzbereich transformierten Ambisonic-

Bus-Signale werden nun mit den zyklisch ausgelesenen, fourier-transformierten, reduzierten HRIR-Sätzen komplex multipliziert und die komplexen Ergebnisse auf die beiden reduzierten Dekodermatrizen B_R geführt. Deren komplexes Ergebnis wird mittels zweier realer IFFTs wieder in den Zeitbereich transformiert (es benötigt noch einer Gain-Korrektur bezüglich der Blocklänge der FFT; zB. Multiplikation mit Reziprokwert von 256). An den Signalausgängen des FFT-Subpatches wird noch der „overlap and add“-Teil der zyklischen Faltung erledigt. Nun werden noch die symmetrischen und antisymmetrischen Komponenten addiert bzw. subtrahiert und wir erhalten den linken und den rechten Kanal.

4 Bedienungsanleitung der Pd-Implementierung

4.1 Die grafische Oberfläche und die Bedienung der Parameter



DSP-ON-OFF:

Schaltet DSP-Berechnung EIN und AUS, zeigt die DSP-Auslastung des Pd-Prozesses an.

3D-ROTATION:

zeigt die 3 Achsendrehrichtungen des Hörer-Kopfes an in Grad (z-, y- und x-Achse).

Schaltet den INTERTRAX2-Headtracker ein.

ORDER-WEIGHT

before decode:

Ein Satz von 4 Ambisonic-Ordnungs-Gewichten, die kurz vor der Dekodierung multipliziert werden und im Diagramm \bar{w} dargestellt werden; oben rechteckig (nur Einsen), unten hanning-förmig gewichtet.

`diffuse encode:`

Ein Satz von 4 Ambisonic-Ordnungs-Gewichten, die nach der Kodierung des diffusen Nachhalls multipliziert werden und in `Wd` dargestellt werden; oben rechteckig (nur Einsen), unten hanning-förmig gewichtet.

MISCELLANEOUS

`ambisonic-radius [m]:`

Taburadius um den Hörer, bzw.

Bezugsradius für die Berechnung der Entfernungsdämpfung (in m);

Innerhalb dieses Kreises nimmt die Lautstärke einer Quelle nicht mehr zu.

`interpolation-time [ms]:`

Abtastzeit für die Positionen von Quellen und Hörer, bzw.

für die Kopfdrehwinkel des Hörers in ms.

`HRIR-onset [samples]:`

Überspringen der ersten `n` Samples der HRIRs in Samples.

`HRIR-length [samples]:`

Länge der HRIR in Samples, beginnend bei HRIR-Onset,

das 4. Viertel wird mit einer halben Hanning-Funktion ausgeblendet.

`HRIR-Pathname`

`HRIR-set:`

wählt den Satz von HRIRs, indem man innerhalb des Ordners eine beliebige Wav-Datei anklickt (der Ordnername wird dann angezeigt).

`init all:`

sendet alle GUI-Parameter Richtung DSP.

`presets:`

Initialisiert alle GUI-Parameter mit Werten, die im grünen Subpatch „Edit *“ in 10 verschiedenen Message-Boxen stehen

SOUNDFILE-PLAYER

`Player 1`

Befehle für Soundfileplayer 1

`new:` öffnet neues Mono-Soundfile im wav-Format.

`start:` startet Soundfile von Beginn an.

`stop:` stoppt den Abspielvorgang und setzt den Abspielvektor auf Beginn.

`pause`: stoppt den Abspielvorgang und belässt den Abspielvektor auf seiner Position.

`loop`: Wiederhol-Option für Abspielvorgang.

`sync`: Synchronisier-Option zum parallelen Starten und Stoppen aller Player.

`gain [dB]`: Quellsignalverstärkung in technischen dB.

`VU`: Pegelanzeige nach der Quellsignalverstärkung.

OUTPUT

`VU`: Pegelanzeige des linken und rechten Ausgangssignals.

`gain [dB]`: Verstärkung für linkes und rechtes Ausgangssignal in technischen dB.

REVERBERATION-BALANCE

`direct_gain [dB]`:

Verstärkung aller Direktanteile in technischen dB,
beeinflusst nicht die Erst- und Zweit-Reflexionen bzw. den Diffusanteil.

`early1_gain [dB]`:

Verstärkung aller Erstreflexionen in technischen dB,
beeinflusst nicht die Zweitreflexionen bzw. den Diffusanteil.

`early2_gain [dB]`:

Verstärkung aller Zweitreflexionen in technischen dB,
beeinflusst nicht den Diffusanteil.

`reverb_gain [dB]`:

Verstärkung des Diffusanteils in technischen dB,

(Für alle 4 Gains gilt: regelt man den Wert auf -100 dB herunter, werden bei allen Raumsimulationen der Quellen die betroffenen Module von der DSP-Verarbeitungskette herausgenommen.)

ROOM-SETTINGS

`x_room_length [m]`:

Raumabmessung (Länge) in x-Richtung in m.

`y_room_width [m]`:

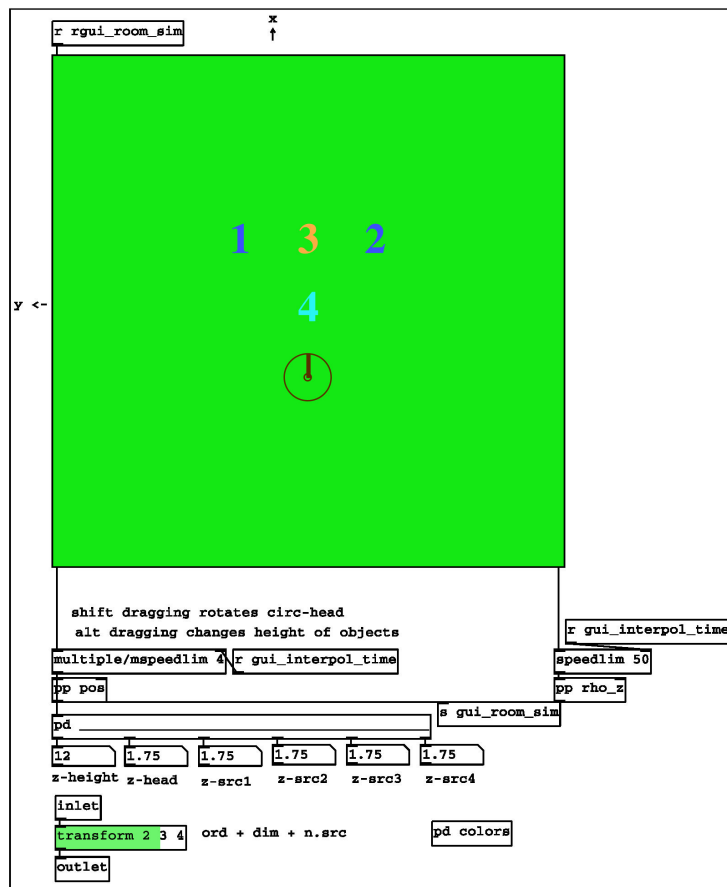
Raumabmessung (Breite) in y-Richtung in m.

`z_room_height [m]`:

Raumabmessung (Höhe) in z-Richtung in m.
`pix_per_meter [pix/m]:`
Darstellungsmaßstab in Pixel/m.
`sonic_speed [m/s]:`
Schallgeschwindigkeit in m/s.
`wall_absorption_coeff < 450 Hz [%]:`
Wandabsorptionskoeffizient für Tiefenfrequenzen in Prozent.
`wall_absorption_coeff mid [%]:`
Wandabsorptionskoeffizient für Mittenfrequenzen in Prozent.
`wall_absorption_coeff > 3.5 kHz [%]:`
Wandabsorptionskoeffizient für Höhenfrequenzen in Prozent.
`air_absorption_cutoff_frequency [Hz]:`
gemittelte Luftabsorptionsgrenzfrequenz in Hz.
`distance_exponent:`
experimenteller Exponent für das Abstandsverhältnis.

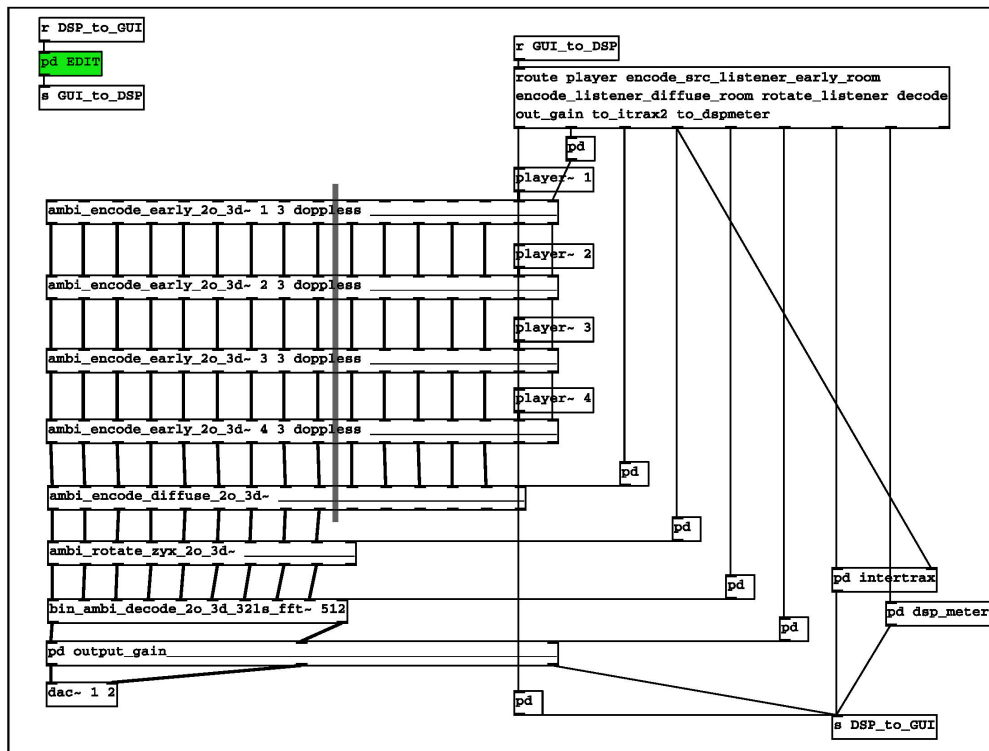
Zur Erweiterung des Patches mit mehr Quellen müsste man hier das Modul „player“ duplizieren und mit der nächsten Quellindexzahl als Übergabeargument versehen.

4.2 Die grafische Simulation des Raums



Durch Anklicken und Bewegen lassen sich die x- und y-Koordinaten der Quellobjekte und der Hörerin bzw. des Hörers verändern (+x nach oben, +y nach links), durch Alt-Taste plus Anklicken lässt sich die z-Koordinate verändern (+z aus dem Bildschirm heraus durch ziehen nach oben). Die Höhe wird signalisiert durch die Größe der Ziffern, bzw. durch die Größe des inneren Kreises. Um die Anzahl der Quellen zu verändern, klickt man in „pd colors“ die horizontale Walleiste an, dann verändert man das Übergabeargument von „multiple/mspeedlim“ auf die gewünschte Zahl und zuletzt das 3. Argument von „transform“.

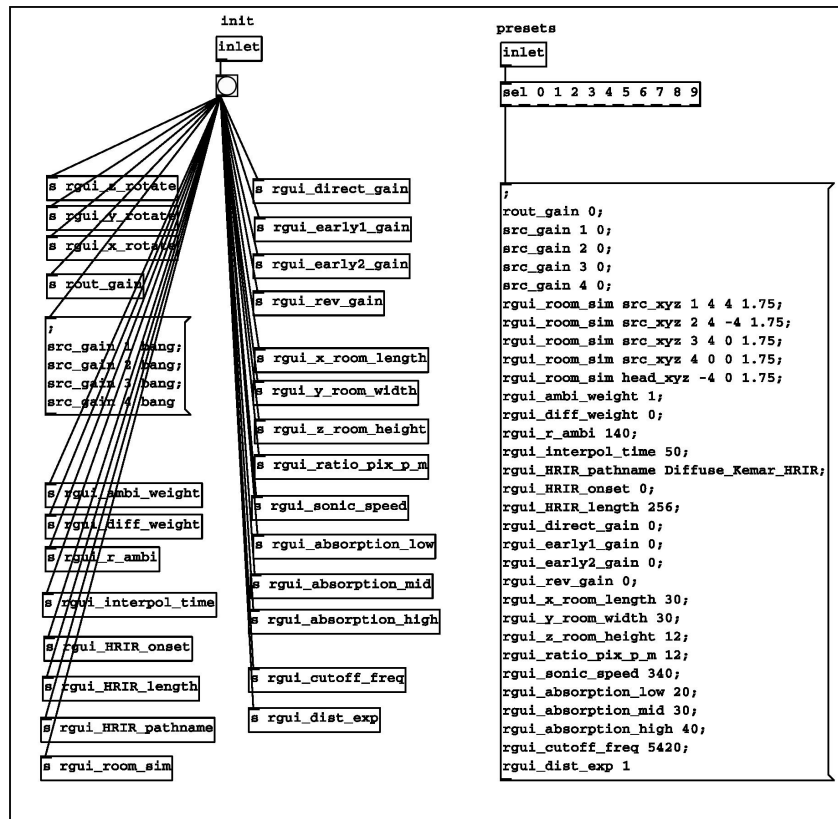
4.3 Der Main-Patch



In „pd EDIT“ befinden sich die beiden in Kapitel 4.1 und 4.2 beschriebenen grafischen Oberflächen, „ambi_encode_early_2o_3d~“ dient zur Kodierung der Direkt- und Erstreflexions-Signale einer Schallquelle „player~“, „ambi_encode_diffuse_2o_3d~“ kodiert die verhallten Erstreflexionssummen, „ambi_rotate_zyx_2o_3d~“ rotiert das Ambisonic-Feld, „bin_ambi_decode_2o_3d_32ls_fft~“ dekodiert das Ambisonic-Feld zu einem Stereosignal, „pd output_gain“ regelt die Lautstärke, „pd intertrax“ verbirgt die Schnittstelle zum Motion-Tracker.

Zur Erweiterung des Patches mit mehr als 4 Quellen müsste man hier die Module „player~“ und „ambi_encode_early_2o_3d~“ duplizieren und jeweils mit der nächsten Indexzahl als Übergabeargument versehen und in die Kanalkette einfügen und verbinden.

4.4 Die Presets



Erklärung der Preset-Parameter:

rout_gain: Ausgangsverstärkung in technischen dB;

src_gain: Quell-Index + Verstärkung der Quelle in technischen dB;

rgui_room_sim src_xyz: Quell-Index + x- +y- + z-Koordinate der Quelle in m;

rgui_room_sim head_xyz: x- +y- + z-Koordinate des Hörers in m;

rgui_ambi_weight: Index für einen Satz von Ambisonic-Ordnungsgewichten
der Dekodierung;

rgui_diff_weight: Index für einen Satz von Ambisonic-Ordnungsgewichten
der diffusen Nachhallkodierung;

rgui_r_ambi: Minimalabstand zwischen Quellen und Hörer in cm;

rgui_interpol_time: Abtastperiode des Motion-Trackers in ms;

rgui_HRIR_pathname: Ordnername des HRIR-Satzes;

rgui_HRIR_onset: Einspringpunkt in HRIR-wav-Datei in Samples;

rgui_HRIR_length: Längenbeschränkung der Impulsantwort in Samples

rgui_direct_gain: Verstärkung des Direktanteils in technischen dB;

rgui_early1_gain: Verstärkung des Erstreflexionsanteils in technischen dB;

`rgui_early2_gain`: Verstärkung des Zweitreflexionanteils in technischen dB;
`rgui_rev_gain`: Verstärkung des Nachhallanteils in technischen dB;
`rgui_x_room_length`: Raumdimension in x-Richtung in m;
`rgui_y_room_width`: Raumdimension in y-Richtung in m;
`rgui_z_room_height`: Raumdimension in z-Richtung in m;
`rgui_ratio_pix_p_m`: Verhältnis Pixel zu Meter bezüglich grünen
Raum-Rechtecks;
`rgui_sonic_speed`: Schallgeschwindigkeit in m/s;
`rgui_absorption_low`: Wandabsorptionskoeffizient kleiner 450 Hz in Prozent;
`rgui_absorption_mid`: Wandabsorptionskoeffizient für mittlere Frequenzen
in Prozent;
`rgui_absorption_high`: Wandabsorptionskoeffizient über 3,5 kHz in Prozent;
`rgui_cutoff_freq`: gemittelte Luftabsorptions-Ersatzgrenzfrequenz in Hz;
`rgui_dist_exp`: experimenteller Exponent für Distanzverhältnis Quelle-Hörer zu
Ambisonic-Minimal-Radius.

4.5 Beschreibung der Steuerparameter des DSP-Moduls

Die Message für jeden Parameter besteht aus einer Reihe von Vornamen, dann folgt der eigentliche Parameterwert bzw. der Parametersatz. Nur nach den Vornamen „source“ bzw. „src“ folgt ein numerischer Index i mit folgender Bedeutung: $i == 0$ für alle Quellkodiermodule, $i > 0$ nur für dieses Quellkodiermodul.

Vom GUI-Modul zum DSP-Modul:

`encode_src_listener_early_room listener position:`
x-, y- und z-Koordinaten der Hörerin / des Hörers in Meter
`encode_src_listener_early_room source 1 gain:`
Verstärkung von Quelle 1 in MIDI-dB
`encode_src_listener_early_room source 0 position:`
x-, y- und z-Koordinaten aller Quellen in Meter
`encode_src_listener_early_room source 1 position:`

x-, y- und z -Koordinaten von Quelle 1 in Meter

`encode_src_listener_early_room source 0 direct_gain:`

Verstärkung des Direktanteils aller Quellen in MIDI-dB

`encode_src_listener_early_room source 0 early1_gain:`

Verstärkung der Erstreflexionsanteile aller Quellen in MIDI-dB

`encode_src_listener_early_room source 0 early2_gain:`

Verstärkung der Zweitreflexionsanteile aller Quellen in MIDI-dB

`encode_src_listener_early_room room size:`

Abmessungen in x-, y- und z-Richtung des virtuellen Raums in Meter

`encode_src_listener_early_room room absorption_coeff low:`

Absorptionskoeffizient für tiefe Frequenzen der Wände [zw. 0 .. 1]

`encode_src_listener_early_room room absorption_coeff mid:`

Absorptionskoeffizient für mittlere Frequenzen der Wände [zw. 0 .. 1]

`encode_src_listener_early_room room absorption_coeff high:`

Absorptionskoeffizient für hohe Frequenzen der Wände [zw. 0 .. 1]

`encode_src_listener_early_room room cutoff_freq:`

gemittelte Luftabsorption, Tiefpassgrenzfrequenz in Hz

`encode_src_listener_early_room room sonic_speed:`

Schallgeschwindigkeit in Meter/Sekunden

`encode_src_listener_early_room distance_exponent:`

experimenteller Exponent für Distanzverhältnis Quelle-Hörer zu Ambisonic-Minimal-Radius [zw. 0.5 .. 2]

`encode_src_listener_early_room ambi_radius:`

Minimalstabschab zur Hörerin / zum Hörer in Meter

`encode_src_listener_early_room interpol_time:`

Abtastperiode des Motion-Trackers in ms

`encode_listener_diffuse_room listener position:`

x-, y- und z-Koordinaten der Hörerin / des Hörers in Meter

`encode_listener_diffuse_room diffuse gain:`

Verstärkung des diffusen Nachhallanteils aller Quellen in MIDI-dB

`encode_listener_diffuse_room diffuse gain ambi_weight:`

**Ambisonic-Ordnungsgewichtungen = (Ordnung+1) Zahlen
(zur Kodierung des diffusen Nachhalls)**

`encode_listener_diffuse_room room size:`

Abmessungen in x-, y- und z-Richtung des virtuellen Raums in Meter

`encode_listener_diffuse_room room absorption_coeff low:`
Absorptionskoeffizient für tiefe Frequenzen der Wände [zw. 0 .. 1]

`encode_listener_diffuse_room room absorption_coeff mid:`
Absorptionskoeffizient für mittlere Frequenzen der Wände [zw. 0 .. 1]

`encode_listener_diffuse_room room absorption_coeff high:`
Absorptionskoeffizient für hohe Frequenzen der Wände [zw. 0 .. 1]

`encode_listener_diffuse_room room cutoff_freq:`
gemittelte Luftabsorption, Tiefpassgrenzfrequenz in Hz

`encode_listener_diffuse_room room sonic_speed:`
Schallgeschwindigkeit in Meter/Sekunden

`encode_listener_diffuse_room interpol_time:`
Abtastperiode des Motion-Trackers in ms

`rotate_listener rho_zyx:`
Rotationswinkel um die z-, y- und die x-Achse in Grad
(in dieser Reihenfolge zu multiplizieren)

`rotate_listener interpol_time:`
Abtastperiode des Motion-Trackers in ms

`decode ambi_weight:`
Ambisonic-Ordnungsgewichtungen = (Ordnung+1) Zahlen
(zur Dekodierung der Ambisonic-Signale)

`decode HRIR_onset:`
dieser Betrag in Samples wird in den
verwendeten HRIRs übersprungen

`decode HRIR_length:`
effektiv verwendete Länge der HRIRs in Samples

`decode HRIR_pathname:`
Pfadname für Ordner, in dem sich die gewünschten HRIRs befinden

`out_gain:`
Gesamtausgangsverstärkung in MIDI-dB

`to_itrax2:`
schaltet Motion-Tracking aus und ein [0 / 1]

`to_dspmeter:`
schaltet Signalverarbeitung aus und ein [0 / 1]

from_dspmeter 1:

liefert den gemittelten DSP-Rechzeitverbrauch in Prozent

from_dspmeter 2:

liefert den Spitzenwert-DSP-Rechzeitverbrauch in Prozent

from_dspmeter 3:

liefert den Ein-/Aus-Status der Signalberechnung [1 / 0]

from_itrax2:

liefert eine Liste mit 3 Rotationswinkel in Grad (ρ_z , ρ_y und ρ_x)

vu src 1:

liefert eine Liste mit Effektivwert und Spitzenwert des verstärkten
Eingangspegels von Quellkodiermodul 1 in technischen dB

vu out_left:

liefert eine Liste mit Effektivwert und Spitzenwert des linken
Ausgangspegels in technischen dB

vu out_right:

liefert eine Liste mit Effektivwert und Spitzenwert des rechten
Ausgangspegels in technischen dB

5 Beschreibung der Pd-Bibliotheken

5.1 Binaural-Ambisonic-Dekoder-Bibliothek „iem_bin_ambi“

bin_ambi reduced decode fft2:

Objekt zur Berechnung eines reduzierten Satzes von HRTF-Feldern für ein Binaural-Ambisonic-System.

Funktionsweise:

Dieses Objekt berechnet zuerst die Koeffizienten der Ambisonic-Lautsprecherkodiermatrix C mittels der Messages „real_ls“ und „pht_ls“. Dann wird die Pseudoinverse gebildet, eventuell werden noch Zeilen der Phantomlautsprecher mittels eines Spiegelgewichts zu den realen Lautsprecher addiert. Nun wird die Existenz der Tables für HRIRs und HRTFs überprüft und erstere geladen. Im nächsten Schritt wird der HRIR-Satz mit der Dekodermatrix multipliziert und das Ergebnis spaltenweise fouriertransformiert und in die HRTF-Felder kopiert.

Im 2-D-Fall müssen die Lautsprecher symmetrisch zur x-Achse (0°-Azimutlinie), im 3-D-Fall müssen die Lautsprecher symmetrisch zur zx-Ebene arrangiert sein.

Übergabe-Parameter:

- 1.Argument: *<float>* Ganzzahl die als Vorname diverser Felder verwendet wird, um den Namen Einzigartigkeit zu verleihen;
- 2.Argument: *<symbol>* Hauptnamen der relevanten HRIR-Felder bezüglich linken Ohrs mit Länge $\geq K$;
- 3.Argument: *<symbol>* Hauptnamen der relevanten realen HRTF-Felder mit Länge $2 \cdot K$;
- 4.Argument: *<symbol>* Hauptnamen der relevanten imaginären HRTF-Felder ($2 \cdot K$);
- 5.Argument: *<symbol>* Name des relevanten HRIR-Fadeout-Feldes mit Länge K ;
- 6.Argument: *<float>* Ambisonic Ordnung M ($1 \leq M \leq 12$ für 2D; $1 \leq M \leq 5$ für 3D);
- 7.Argument: *<float>* Dimension D (2 für 2D; 3 für 3D);
- 8.Argument: *<float>* Anzahl der realen Lautsprecher $N_{L_{re}}$;
- 9.Argument: *<float>* Anzahl der virtuellen Lautsprecher $N_{L_{vi}}$;
- 10.Argument: *<float>* Länge der Blockfilter-FFT $2 \cdot K$;

Eingänge und Ausgänge:

1.Eingang:

- $\{sing_range\} + \langle float \rangle$ Singularitätsgrenze nahe 0 (anfangs 1EE-10);
- oder $\{ambi_weight\} + (M+1)$ mal $\langle float \rangle$ Ordnungsgewichte der Dekodierung
(sind anfangs alle 1);
- oder $\{load_HRIR\} + \langle float \rangle$ Lautsprecherindex n_{L_re} ($1 \leq n_{L_re} \leq N_{L_re}$) +
+ $\langle symbol \rangle$ HRIR-Dateiname des realen Lautsprechers n_{L_re}
gibt am linken Ausgang eine Liste mit 2 $\langle symbol \rangle$ aus
(HRIR-Feldname + HRIR-Dateiname);
- oder $\{real_ls\} + \langle float \rangle$ Lautsprecherindex n_{L_re} ($1 \leq n_{L_re} \leq N_{L_re}$) (3D kodiert) +
+ $\langle float \rangle$ geodätische Elevation θ ($-90^\circ \leq \theta \leq +90^\circ$) +
+ $\langle float \rangle$ Azimut φ ($-180^\circ \leq \varphi \leq +180^\circ$) von realen Lautsprecher n_{L_re} ;
- oder $\{real_ls\} + \langle float \rangle$ Lautsprecherindex n_{L_re} ($1 \leq n_{L_re} \leq N_{L_re}$) (2D kodiert) +
+ $\langle float \rangle$ Azimut φ ($-180^\circ \leq \varphi \leq +180^\circ$) von realen Lautsprecher n_{L_re} ;
- oder $\{pht_ls\} + \langle float \rangle$ Lautsprecherindex n_{L_vi} ($1 \leq n_{L_vi} \leq N_{L_vi}$) (3D kodiert) +
+ $\langle float \rangle$ geodätische Elevation θ ($-90^\circ \leq \theta \leq +90^\circ$) +
+ $\langle float \rangle$ Azimut φ ($-180^\circ \leq \varphi \leq +180^\circ$) von virtuellen Lautsprecher n_{L_vi}
- oder $\{pht_ls\} + \langle float \rangle$ Lautsprecherindex n_{L_vi} ($1 \leq n_{L_vi} \leq N_{L_vi}$) (2D kodiert) +
+ $\langle float \rangle$ Azimut φ ($-180^\circ \leq \varphi \leq +180^\circ$) von virtuellen Lautsprecher n_{L_vi}
- oder $\{calc_pinv\}$: berechnet die Pseudoinverse der kodierten Lautsprechermatrix;
multipliziert die Ordnungsgewichte hinein;
- oder $\{ipht_ireal_muladd\} + \langle float \rangle$ virtueller Lautsprecherindex n_{L_vi} +
+ $\langle float \rangle$ realer Lautsprecherindex n_{L_re} +
+ $\langle float \rangle$ Spiegelgewichtsfaktor, mit dem die virtuelle Zeile multipliziert
und zur realen Zeile addiert wird;
- oder $\{check_HRIR_arrays\} + \langle float \rangle$ realer Lautsprecherindex n_{L_re}
prüft die Existenz der HRIR-Felder und des HRIR-Fadeout-Feldes;
- oder $\{check_HRTF_arrays\} + \langle float \rangle$ Ambisonic-Kanalindex $\leq (M+1)^2$ bzw. $(2M+1)$
prüft die Existenz der HRTF-Felder;
- oder $\{calc_reduced\} + \langle float \rangle$ Ambisonic-Kanalindex $\leq (M+1)^2$ bzw. $(2M+1)$
multipliziert alle Zeilen vom Satz der HRIR-Felder mit jener
Spalte der Dekodermatrix und fouriertransformiert das Ergebnis

zu 2 Spektren, die in den zugehörigen Feldern gespeichert werden;
oder `{calc_sym}`

gibt am rechten Ausgang eine Liste mit 2 *<float>* aus, passend für Objekt `matrix_bundle_stat~` (mit vorangestelltem „element“),
erstes *<float>*: Zeilenindex der reduzierten Dekodermatrix,
zweites *<float>* : Spaltenindex der reduzierten Dekodermatrix.

bin_ambi_reduced_decode_fir2:

Objekt zur Berechnung eines reduzierten Satzes von HRIR-Feldern für ein Binaural-Ambisonic-System. Dieses Objekt besitzt fast die gleichen Eigenschaften wie `bin_ambi_reduced_decode_fir2`, es wird nur nicht fouriertransformiert und die message „check_HRTF_arrays“ wurde durch „check_RED_HRIR_arrays“ ersetzt. Bei den Übergabeargumenten entfällt das 4.Argument, das 3.Argument ist der Hauptname des reduzierten HRIR-Feldes mit der Länge K.

5.2 Bibliothek „ambi_abs“

Für die Kodierung des Direktsignals und der Erstreflexionen gibt es ein Objekt (hier 3-dimensional und 4.Ordnung) „`ambi_encode_early_4o_3d~`“ mit $(2M+1)$ bzw. $(M+1)^2$ Signal-Eingängen und -Ausgängen für Ambisonic, 4 bzw. 5 Signal-Eingängen und -Ausgängen für die Erstreflexionen (abhängig ob 2D oder 3D), einem Signaleingang für das Quellsignal und einem Eingang für die Steuerparameter, die man im Kapitel 4.5 in der Beschreibung der Messages 1 bis 13 hinter dem Vornamen „`encode_src_listener_early_room`“ findet. Die Übergabeargumente sind: Quellindex beginnend bei 1, Dehnfaktor für Interpolationszeit bei Delays und Interpolationsart bei Delays (dyn, stat, doppleless). Diese Objekte gibt es in 2D und in 3D von 2. bis 5. Ordnung.

Die Kodierung des Nachhalls übernimmt ein Objekt (hier 3-dimensional und 4.Ordnung) „`ambi_encode_diffuse_4o_3d~`“ mit $(2M+1)$ bzw. $(M+1)^2$ Signal-Eingängen und -Ausgängen für Ambisonic, 4 bzw. 5 Signal-Eingängen für die Erstreflexionen (abhängig ob 2D oder 3D) und einem Eingang für die Steuerparameter, die man im Kapitel 4.5 in der Beschreibung der Messages 14 bis

23 hinter dem Vornamen „`encode_listener_diffuse_room`“ findet. Diese Objekte gibt es in 2D und in 3D von 2. bis 5. Ordnung.

Die Rotation der Ambisonic-Kanäle realisiert ein Objekt (3-dimensional und 4.Ordnung) „`ambi_rotate_zyx_4o_3d~`“ mit $(2M+1)$ bzw. $(M+1)^2$ Signal-Eingängen und -Ausgängen für Ambisonic und einem Eingang für die Steuerparameter, die man im Kapitel 4.5 in der Beschreibung der Messages 24 bis 25 hinter dem Vornamen „`rotate_listener`“ findet. Diese Objekte gibt es in 2D von 1. bis 10.Ordnung und in 3D von 1. bis 4. Ordnung. In 2D rotiert man nur um die z-Achse, in 3D rotiert man auch gleichzeitig um die z-, y- und die x-Achse.

Zuletzt gibt es ein Objekt für einen 3D-Dekoder 4.Ordnung mit Binaural-Ausgang: „`bin_ambi_decode_4o_3d_32ls_fft~`“ mit $(2M+1)$ bzw. $(M+1)^2$ Signaleingängen und 2 Signalausgängen und einem Eingang für die Steuerparameter, die man im Kapitel 4.5 in der Beschreibung der Messages 26 bis 29 hinter dem Vornamen „`decode`“ findet. Das einzige Übergabeargument ist: die FFT-Länge (doppelte FIR-Länge = maximale HRIR-Länge). Diese Objekte gibt es in 2D von 1. bis 10.Ordnung und in 3D von 1. bis 4. Ordnung.

6 Erkenntnisse, Aussichten

Erste Versuche mit Binaural-Dekodern wiesen Abschwächungen oder Überhöhungen der Stereosummenpegels beim Kodieren von Quellen aus dem Bereich Elevation -40 bis -90 Grad auf, da von KEMAR keine HRIRs zur Verfügung gestellt wurden und deswegen entweder ein regelmäßiges Lautsprecher-Layout mit anschließendem Weglassen der virtuellen Lautsprecher stattfand, oder ein unregelmäßiges Layout von Anfang an gewählt wurde (deshalb der Trick mit der Interpolation / Extrapolation der HRIR-Daten). Der abgestumpfte Ikosaeder wurde schon von Archimedes, Leonardo Da Vinci, Leonard Euler, Buckminster Fuller, Michael Gerzon und Jerome Daniel beschrieben und in verschiedenster Art verwendet. Die Elevationen der Lautsprecherringe wurden optimiert indem man nur mit einem reinen W-Kanalsignal den Dekoder versorgte und die einzelnen Elevationen variierte und auf ein Minimum der Varianz der Lautsprecherpegel hin regelte (siehe Kapitel 2.4).

Eine weitere Erkenntnis war, dass nur isoliert aufgedrehte Erstreflexionen die Richtungsordnung der Quelle unterstützten, isoliert aufgedrehte Zweitreflexionen hingegen die Richtungsordnung eher verwirrender gestalteten, jedoch die Außerkopffortung verbesserten.

Bei den KEMAR-Kurven stellte sich heraus, dass die Ortung bei einer Verkürzung bis auf 128 Samples immer noch gleich gut blieb, unter 128 Samples dann aber die Ortungsqualität rapid abnahm.

Beim Einsatz individueller HRIRs von IRCAM stellte sich eine deutliche Verschlechterung der Ortungsqualität ein.

Bei der Verwendung von Sprache oder Schritten als Signalquelle war der Doppler-Effekt akzeptabel, bei Musiksignalen jedoch war er hochgradig störend. Deshalb wurde neben der statischen und dynamischen Version von Delay-Abzapfungen noch die 3. Version der Überblendtechnik zweier statischer Delays verwendet.

Für die Zukunft könnte man das Raummodell durch richtungsabhängige Quell-Übertragungsfunktionen, oder unterschiedliche Decken-, Wand- und Boden-Absorptionseigenschaften erweitern.

7 Literaturverzeichnis

- [1] Miller Puckette, "*Pd Documentation Rel-0.39*", Dokumentation von Pd im Netz:
http://crca.ucsd.edu/~msp/Pd_documentation/
- [2] J. Zmölning, "*Kap 2 Ambisonic*" in:
 "*Entwurf und Implementierung einer Mehrkanal-Beschallungsanlage*",
 Diplomarbeit am IEM Graz, KUG 2001, pp. 21 - 42
- [3] A. Sontacchi, "*Kap 4.2.1 Implementierung eines 2D-Systems, 3.Ordnung*" in:
 "*Dreidimensionale Schallfeldreproduktion für
 Lautsprecher- und Kopfhöreranwendungen*",
 Dissertation am IEM Graz, KUG 2003, pp. 128 - 133
- [4] P. Majdak, M. Noisternig, "*Kap 4 Ambisonic*" in:
 "*Implementation kopfpositionsbezogener Binauraltechnik*",
 Diplomarbeit am IEM Graz, KUG 2001, pp. 20 - 53
- [5] J. Daniel, "*Représentation de champs acoustiques, application à la transmission
et à la reproduction de scènes sonores complexes dans un
contexte multimédia*",
 Dissertation an der Universität Paris 6, 2000
- [6] T. Musil, "*Kap 2 Begriffsdefinitionen*" in :
 "*IEM-Report 12/03*" Arbeitsbericht am IEM Graz, KUG 2003
- [7] T. Musil, "*Kap 3 Ambisonic Kodiervorschriften*" und
 "*Kap 7 Ambisonic Dekodiervorschriften*" in :
 "*IEM-Report 15/03*" Arbeitsbericht am IEM Graz, KUG 2003
- [8] J.-M. Jot and A.Chaigne, "*Digital delay networks for designing artificial
reverberators*" in "*Audio Engineering Society Convention*", Feb. 1991.

Anhang

Inhalt der beiliegenden CD

Auf der CD befindet sich der Ordner „Binaural_Ambisonic_CD“, darin die Ordner „ambi_abs“ und „ambi_abs_multiple“ für alle Subpatches dieses Projekts, die Ordner „lrcam_HRIR_17“, „lrcam_HRIR_23“, „lrcam_HRIR_37“, „lrcam_HRIR_5“, „Kemar_HRIR“ beinhalten die verschiedenen HRIR-Sätze, in den Ordnern „pd-0.39-2_win“ und „pd-0.39-2_lin“ ist Pd samt Zusatzbibliotheken für Windows bzw. Linux installiert, im Ordner „sounds“ sind Mono-Wav-Dateien (Musik, Schritte, Stimmen) als Quellklangmaterial und 2 Zielformate in Stereo (Tour1.wav und Tour2.wav) vorhanden. Die Startskripte „pd39_asio.bat“ und „pd39_mmio.bat“ für Windows, bzw. „pd39.sh“ für Linux sind noch eventuell anzupassen, falls andere Patches als „bin_ambi_4o_3d_2src_wo_doppler.pd“ verwendet werden sollen („wo“ steht abgekürzt für „without“). Weiters gibt es noch drei Textdokumente, „LICENSE.txt“, „GnuGPL.txt“ und dieses Dokument als „IEM_REPORT_38_06.pdf“. Über das Urheberrecht und Kopierrecht dieses Textes und der Patches auf der CD verfügen

© Thomas Musil 2000 - 2006 und das Institut für Elektronische Musik und Akustik an der Kunstuniversität Graz, Austria.