

institut für elektronische musik und akustik



Toningenieur – Projekt:

Distanzkodierung über Kopfhörer

Distance coding via headphones

Bearbeitung: Michael Neffe

Betreuung: A. Sontacchi, R.Höldrich

KURZFASSUNG

A couple of Head Related Transfer Functions (HRTF) exists with an angle-resolution of 5 degrees azimuth and 10 degrees elevation. By performing measurements for one loudspeaker at a distance of 1.4 m to a mock-up head the distance-perception of reproduction through one of these HRTF-pairs is given. By combination of multiple Transfer Functions the distance of a virtual (primary) sound source can be varied.

The calculation of the signals of the secondary sources for the desired position of the primary source is done by the Kirchhoff-Helmholtz-Integral (KHI). Adequate filtering procedures and spatial distribution of the primary signal are followed by coding to ambisonic domain which gives possibilities to vary time as well as the position of the virtual source in azimuth and elevation. In Ambisonic domain the resulting sonic field is decoded to the virtual loudspeakers. From these virtual loudspeaker signals and after filtering with appropriate HRTFs the output (headphone) signals are obtained.

Primary goal of this project is the positioning of the virtual sources in various distances with focus on the difference of given distance and perceived distance. The influence of the distance-perception is to be investigated depending on different filter parameters.

ABSTRACT

Ein Satz von Außenohrübertragungsfunktionen (Head Related Transfer Functions, HRTF) liegt mit einer Winkelauflösung von 5° in der Horizontalebene und 10° in der Vertikalebene vor. Diese wurden für einen Lautsprecher in jeweils 1,4m Entfernung zu einem Kunstkopf vermessen. Dadurch ist für eine Wiedergabe bei Verwendung eines zusammengehörenden HRTF - Paares der Entfernungseindruck bereits vorgegeben. Durch Kombination mehrerer Übertragungsfunktionen kann die Entfernung einer virtuellen Quelle variiert werden.

Mittels der Kirchhoff Helmholtz Integral (KHI) werden für die gewünschte Position der Primärquelle (virtuelle Quelle) die Signale der Sekundärquellen bestimmt. Nach entsprechender Filterung und räumlicher Verteilung des Primärsignals kann eine Kodierung in den Ambisonic Bereich erfolgen. Dadurch besteht die Möglichkeit die virtuelle Quelle frei in Elevation und Azimuth zu positionieren und zeitvariant zu bewegen. Im Ambisonic Bereich wird das resultierende Schallfeld auf virtuelle Wiedergabelautsprecher dekodiert. Aus diesen virtuellen Lautsprechersignalen wird durch Filterung mit den entsprechenden HRTFs die Kopfhörer-(Ausgangs-)signale gewonnen.

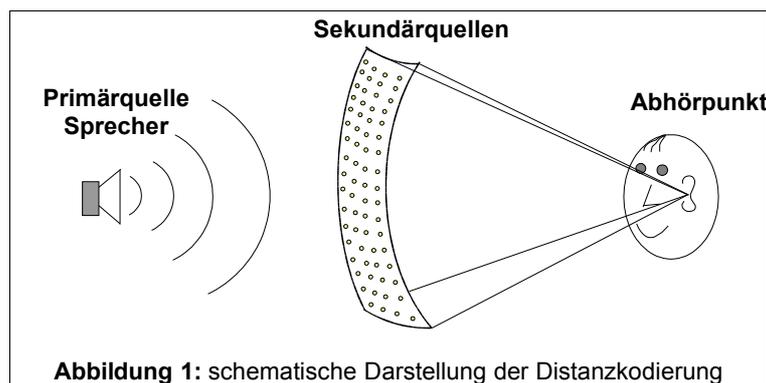
Ziel dieses Projektes ist die Positionierung von virtuellen Quellen in unterschiedlichen Distanzen, wobei der Unterschied von vorgegebener Distanz und wahrgenommener Distanz von Interesse ist. Der Einfluss auf die Distanzwahrnehmung soll durch unterschiedliche Parameterwahl der verschiedenen Filter eruiert werden.

INHALT:

1 .Einleitung.....	3
2 .Kirchhoff – Helmholtz – Integral (KHI).....	5
2.1 .Hygens-Fresnel` sches Prinzip.....	5
2.3 .Ableitung des KHI.....	5
2.6 .Rayleigh-Integral I.....	5
2.7 .Fokussierung.....	6
2.8 .Zusammenfassung.....	7
3 .Realisation.....	7
3.1 .Diskretisierung.....	7
3.2 .Apertur.....	8
3.3 .Richtcharakteristik.....	9
3.4 .Grenzfrequenzen.....	10
3.4.1 .Obere Grenzfrequenz.....	10
3.4.2 .Untere Grenzfrequenz.....	10
3.5 .Aperturanpassung.....	10
3.6 .Nachbearbeitung bzw. Equalizing.....	10
3.7 .Anzahl der Lautsprecher für gute Fokussierung.....	10
4 .Informeller Hörversuch.....	11
5 .Zusammenfassung.....	14
6 .Referenzen.....	14

1 .Einleitung

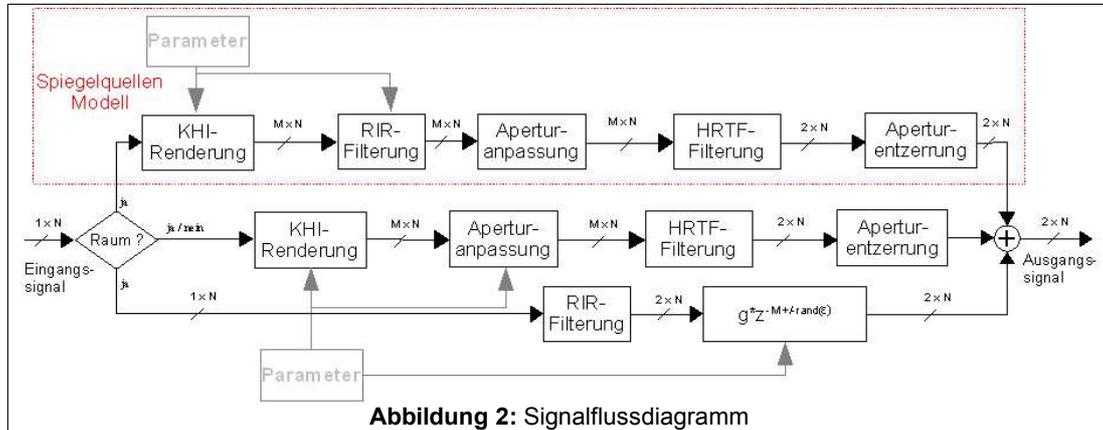
Die Kodierung von Distanz gewinnt in letzter Zeit in Zusammenhang mit der räumlichen Abbildung von Quellen zunehmend an Bedeutung. Für 3-D TV-Anlagen wie für "virtual reality" Systeme spielt der auditive Stimulus ergänzend zum visuellen eine entscheidende Rolle um eine realistischere Umgebung nachbilden zu können. Koinzidenz, nicht nur zwischen visuellen und auditiven Ereignissen in Richtung, sondern zusätzlich in der Tiefenstaffelung von auditiven Ereignissen, lässt Bild und Ton erst realistisch wirken. Bei den meisten Wiedergabesystemen ist eine Kodierung von Distanz nicht oder nur marginal möglich. Stereolautsprecher zum Beispiel können Quellen nicht näher als die Distanz der Lautsprecher abbilden. Weiters darf der Einfluss des Wiedergaberaums nicht außer acht gelassen werden. Dies führt schließlich zu Binauralsystemen die davon entkoppelt sind, jedoch bisher auf die Wiedergabe von gleichweit entfernten Schallquellen beschränkt waren. Erweiterte Systeme haben durch Zufügen von frühen Reflexionen Entfernungsbildung imitiert. Hier wird ein System der Entfernungskodierung vorgeschlagen, das mit Hilfe des Kirchhoff-Helmholtz-Intergrals (KHI) eine Primärquelle in variabler Entfernung abbilden kann. Dafür werden eine große Anzahl von Sekundärquellen, die sich in konstanter Entfernung zum Abhörpunkt (Hörer) befinden, verwendet. Die Sekundärquellsignale leiten sich mittels des KHI aus dem Signal und der Position der Primärquelle ab.



Ausgangspunkt der Überlegungen dieser Arbeit ist das Vorhandensein von Außenohrübertragungsfunktionen (Head Related Transfer Functions, HRTF). Diese wurden für einen Lautsprecher in einer konstanten Entfernung zu einem Kunstkopf vermessen und liegen in einer bestimmten Winkelauflösung in Horizontalebene und Vertikalebene vor. Dadurch ist für die Wiedergabe bei Verwendung eines zusammengehörenden HRTF - Paares der Entfernungseindruck bereits vorgegeben. Die Externalisierung der Quelle wird durch die Qualität des HRTF-Satzes entscheidend mitbestimmt. Durch Kombination mehrerer Übertragungsfunktionen kann die Entfernung einer virtuellen Quelle variiert werden. Zur Reduktion des Berechnungsaufwandes erfolgt die Berechnung der Sekundärquellen nur in einem vordefinierten Bereich (siehe Abb. 1), wobei diese Apertur für eine konstante Richtcharakteristik optimiert ist.

Nach der räumlichen Verteilung des Primärsignals erfolgt als erste Möglichkeit, die sich wiederum in zwei Teile unterteilt, einerseits die direkte Filterung mit den entsprechenden HRTFs um so das gewünschte Kopfhörer-Ausgangssignal zu erhalten, was der Freifeldsituation entspricht. Andererseits kann es zu einem vollständigen Raummodell erweitert werden. Das bedeutet zusätzlich die Berechnung von frühen direktionalen Reflexionen und eines funktionalen Diffusanteils. Diese müssen räumlich und

zeitlich mit dem Direktsignal passend zusammengefügt werden, wie aus Abbildung 2 ersichtlich ist. Als zweite Möglichkeit kann die räumliche Verteilung der Primärquelle in den Ambisonic-Bereich transformiert werden. Zusätzlich können auch hier frühe Reflexionen und ein Diffusanteil berechnet und danach in den Ambisonic-Bereich transformiert werden. Schallfeldrotationen können dann unabhängig von der HRTF-Verteilung durchgeführt werden [1]. Diese Möglichkeit wird jedoch erst in weiteren Veröffentlichungen vorgestellt.



Der verbleibende Teil dieser Arbeit ist folgendermaßen gegliedert: Zu Beginn wird ausgehend vom Hygens-Fresnel'schen Prinzips die allgemeine Form des KHI angegeben und danach die für dieses Projekt interessierenden Zusammenhänge abgeleitet. Am Ende dieses Abschnitts wird gezeigt, dass die Distanzkodierung in drei Entfernungsabschnitte in Bezug zur HRTF-Entfernung unterteilt ist. Danach wird auf praktische Problemfälle, wie der der diskreten Quellverteilung, der endlichen Apertur und Fensterung im Frequenz-Raum-Bereich eingegangen. Im letzten Abschnitt werden die in diesem Zusammenhang gemachten Ergebnisse eines informellen Hörversuchs präsentiert und Einflüsse in Bezug auf Parameterwahl diskutiert.

2 .Kirchhoff – Helmholtz – Integral (KHI)

2.1 .Hygens-Fresnel'sches Prinzip

Eine Primärquelle Ψ in einem homogenen Medium sendet eine Welle aus, die sich mit der Schallgeschwindigkeit c ausbreitet. Nach der Zeit t hat sich am Ort $c \cdot t$ eine Wellenfront ausgebildet. Nach dem Hygen'schen Prinzip kann jeder Punkt auf dieser $c \cdot t$ entfernten Wellenfront nun wiederum als Ausgangspunkt einer Elementarwelle gesehen werden. Die Summe aller Sekundärquellen bildet zum Zeitpunkt $t + \Delta t$ eine Wellenfront Σ die mit der der Primärquelle, zum gleichen Zeitpunkt übereinstimmt.

2.3 .Ableitung des KHI

Das KHI ist die mathematische Formulierung des Hygens – Fresnel'schen Prinzips. Dieses Prinzip besagt, dass die Wellenfront einer sich ausbreitenden Welle durch räumlich verteilte Sekundärquellen nachgebildet werden kann. Daraus folgt, dass die verteilten Sekundärquellen eine bzw. mehrere Primärquellen ersetzen können [2,3].

Die mathematische Formulierung mit dem KHI (Glg.1) besagt, dass man eine beliebige

Primärquellenverteilung in einem quellfreien Volumen nur durch die Kenntnis von Druck (Dipolquelle) und Normalkomponente der Schnelle (Monopolquelle) auf seiner umgebenden Oberfläche vollständig rekonstruieren kann.

$$p_a = \frac{1}{4\pi} \int_S (p \nabla G - G \nabla p) \mathbf{n} dS \quad (\text{Glg. 1})$$

Wobei G die Green'sche Funktion ist und die Lösung der inhomogenen Wellengleichung im inneren des quellfreien Volumens unter Freifeldbedingungen darstellt.

$$G = \frac{e^{-jkr}}{r} + F \quad (\text{Glg. 2})$$

Mit Hilfe der Euler'schen Bewegungsgleichung erhält man schließlich das KHI für ein homogenes Medium

$$\text{grad}(p) = -\rho \frac{dr}{dt} \rightarrow \frac{dp}{dn} = -j\omega \rho_0 v_n \quad (\text{Glg. 3})$$

$$p_a = \frac{1}{4\pi} \int_S \left[\left(p \frac{1+jkr}{r^2} * \cos(\varphi) * e^{-jkr} \right) + \left(j\omega \rho_0 v_n \frac{e^{-jkr}}{r} \right) \right] dS \quad (\text{Glg. 4})$$

2.6 .Rayleigh-Integral I

Um nun das Schallfeld nur durch eine Verteilung von Monopol- bzw. Dipolquellen auf der das Volumen umschließenden Oberfläche darstellen zu können, werden nun Vereinfachungen bezüglich der Oberflächengeometrie getätigt und eine spezielle Wahl für F in der Green'sche Funktion vorgenommen um so das Rayleigh Integral I zu erhalten [2]. In dieser Formel wird das Schallfeld ausschließlich durch die Normalkomponente der Schallschnelle bestimmt.

$$p_A = \frac{1}{2\pi} \int_S j\omega \rho_0 v_n \frac{e^{-jkr}}{r} dS \quad (\text{Glg. 5})$$

Zur Ableitung für die hier verwendete Anwendung sei eine Punktquelle $S(\omega)$ im Raum positioniert. Nun ist der Schalldruck im Abstand r von der Punktquelle durch $S(\omega)$ und $G(\varphi, \vartheta, \omega)$, der Richtcharakteristik, gegeben

$$P(r, \omega) = S(\omega) * G(\varphi, \vartheta, \omega) \frac{e^{-jkr}}{r} \quad (\text{Glg. 6})$$

Durch Positionierung einer sekundären Quellverteilung mit Monopolcharakteristik auf einer Kugeloberfläche kann das Schallfeld der Primärquelle innerhalb der Kugeloberfläche unter Anwendung des Rayleigh Integral I vollständig rekonstruiert werden.

$$P_{\text{synth}}(r, \omega) = \frac{1}{2\pi} \int_S j\omega \rho_0 \frac{e^{-jkr_0}}{r_0} dS \quad (\text{Glg. 7})$$

Durch Anwendung der Euler'schen Bewegungsgleichung erhält man

$$j\omega \rho_0 v_n = -\frac{dP}{dn} = -S(\omega) \frac{d}{dn} \left[G(\varphi, \vartheta, \omega) \frac{e^{-jkr}}{r} \right] \quad (\text{Glg. 8})$$

Nachdem die einhüllende Oberfläche eine Kugeloberfläche ist wird $\frac{d}{dn}$ durch $-\frac{d}{dr}$ ersetzt.

Mit $G(\varphi, \vartheta, \omega) \equiv 1$ für eine Primärquelle mit Monopolcharakteristik erhält man schließlich

$$-\frac{dP}{dn} = -S(\omega) * \left[-jk \frac{e^{-jkr}}{r} - \frac{e^{-jkr}}{r^2} \right] \quad (\text{Glg.9})$$

$$P_{\text{synth}}(r, \omega) = \frac{1}{2\pi} \int_S S(\omega) \frac{1+jkr}{r} \frac{e^{-jkr}}{r} \frac{e^{-jkr_0}}{r_0} dS \quad (\text{Glg.10})$$

mit $k \cdot r \gg 1$

$$P_{\text{synth}}(r, \omega) = \frac{jk}{2\pi} \int_S S(\omega) \frac{e^{-jkr}}{r} \frac{e^{-jkr_0}}{r_0} dS \quad (\text{Glg. 11})$$

Daraus resultiert das Antriebssignal der Sekundärquellen zu:

$$Q_m(r, \omega) = \frac{jk}{2\pi} S(\omega) \frac{e^{-jkr}}{r} \quad (\text{Glg. 12})$$

2.7 .Fokussierung

In der Akustik gilt das Prinzip, dass Quelle und Senke vertauscht werden können ohne das sich die akustischen Eigenschaften des Ausbreitungspfadens ändern. Dadurch scheint es logisch das man Schallwellen die von einer Verteilung von Sekundärquellen ausgehen auf einen gewissen Punkt fokussieren kann [3]. Im Bereich zwischen Primär- und Sekundärquellen breiten sich die jeweiligen Schallfelder in entgegengesetzter Richtung aus. Im Bereich zwischen Primärquelle und Abhörpunkt ist die Ausbreitungsrichtung des Schallfeldes der Primär- und Sekundärquelle gleich, wodurch die Rekonstruktion des Primärschallfeldes durch eine Verteilung von Sekundärquellen möglich wird. Wird für die Primärquelle eine Monopolcharakteristik angenommen, so ist das Schallfeld der Primärquelle wie folgt gegeben:

$$P(r, \omega) = S(\omega) \frac{e^{-jk\rho}}{\rho} \quad (\text{Glg.13})$$

ρ ...Abstand Sekundärquelle Abhörpunkt (Aufpunkt)

Das synthetisierte Schallfeld einer Verteilung von Sekundärquellen mit Monopolcharakteristik ist dann gegeben durch

$$P_{\text{synth}}(r, \omega) = \int_S Q_m^{\text{foc}}(r, \omega) \frac{e^{-jkr_0}}{r_0} dS, \quad (\text{Glg. 14})$$

wobei das Antriebssignal der Sekundärquellen folgenderweise gegeben ist:

$$Q_m^{\text{foc}}(r, \omega) = \frac{k}{j2\pi} S(\omega) \frac{r_0}{r_0 - r} e^{+jkr} \quad (\text{Glg. 15})$$

2.8 .Zusammenfassung

In diesen Abschnitt wurden die Vorschriften zur Berechnung der Sekundärquellen für drei unterschiedliche Fälle angegeben. Zum ersten für Primärquellen außerhalb der Kugeloberfläche, zweitens für Primärquellen innerhalb der Kugeloberfläche und drittens für den Fall das sich die Primärquelle auf der Kugeloberfläche befindet. Hier wird die Berechnung der Sekundärquellen auf einer Kugeloberfläche aus der Primärquelle in Beziehung zum Radius der Kugel durchgeführt:

für $r > r_0$

$$Q_m(r, \omega) = \frac{jk}{2\pi} S(\omega) \frac{e^{-jkr}}{r} \quad (\text{Glg. 16})$$

für $r < r_0$

$$Q_m^{foc}(r, \omega) = \frac{k}{j2\pi} S(\omega) \frac{r_0}{r_0 - r} e^{+jkr} \quad (\text{Glg. 17})$$

für $r = r_0$

$$Q_m(\omega) = S(\omega) \quad (\text{Glg. 18})$$

3 .Realisation

3.1 .Diskretisierung

Mit dem KHI erfolgt die Generierung der Sekundärsignale aus der am gewünschten Ort positionierten Primärquelle (virtuelle Quelle), wobei das KHI die kontinuierliche Verteilung der Sekundärquellen auf der Kugeloberfläche beschreibt. Aus praktikablen Gründen ist dies jedoch nicht möglich und es wird auf eine diskrete Sekundärquellenverteilung übergegangen. Das Integral kann durch eine Summe ersetzt werden.

$$\int F(r) dr \rightarrow \sum_n F_n \Delta r \quad (\text{Glg. 19})$$

mit $F_n = F(r = r_n)$ und Δr dem Samplingabstand

Auch hier gilt gleich einer zeitlichen Signalabtastung das Shannon`sche Abtasttheorem

$$f_{max} = \frac{c}{2 \Delta r} \rightarrow \lambda \geq 2 \Delta r, \quad (\text{Glg. 20})$$

wobei die Winkelabhängigkeit des Abtasttheorems noch nicht berücksichtigt wurde. Die Auswirkung auf die Distanzkodierung durch die Diskretisierung wurde weiters untersucht. Dabei wurde die Ergiebigkeit eingeführt die eine quasikontinuierliche Berechnung der durch eine Fläche tretenden Schalldruck ermöglicht. Die Ergebnisse der mit der hier implementierten Näherung zeigten eine gute Übereinstimmung.

In diesem Projekt war für die direkte Verarbeitung die maximale Auflösung, das heisst der minimale Abstand der einzelnen virtuellen Sekundärquellen auf der Kugeloberfläche, durch die Auflösung der gemessenen HRTF`s beschränkt. Der hier verwendete CIPIC-HRTF – Satz [5] liegt in einer Auflösung von -90° - $+90^\circ$ in der horizontalen Ebene und von -90° - $+270^\circ$ in der Mediane vor.

Genauere Definition der Auflösung:

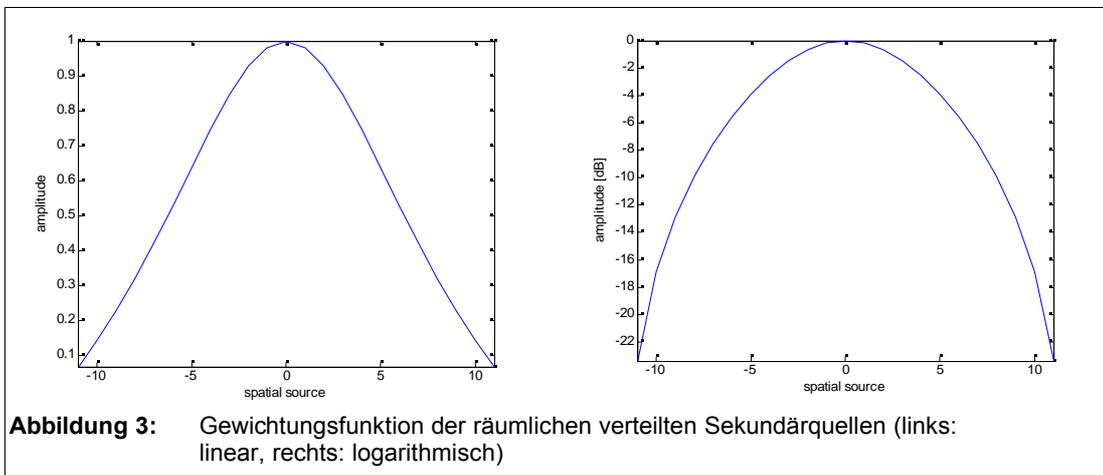
$$azimut = [-80 \quad -65 \quad -55 \quad -45 : 5 : 45 \quad 55 \quad 65 \quad 80]$$

$$elevation = -45 + 5.625 * 0 : 49$$

Für die Weiterverarbeitung in der Ambisonic Domäne kann grundsätzlich jede beliebige Auflösung (d.h. beliebig kleine Abstände zwischen den einzelnen Sekundärquellen) gewählt werden, jedoch ist man bei der Dekodierung immer noch auf die HRTF-Auflösung beschränkt.

3.2 .Apertur

Als Apertur wird hier jene Fläche auf der Kugeloberfläche bezeichnet die zur diskreten Verteilung der Sekundärquellen herangezogen wird. Nachdem aus der Position der Primärquelle maximal die Hälfte der gesamten Kugeloberfläche einzusehen ist (d.h. der Beitrag der Sekundärwellen die auf der Hälfte der Kugeloberfläche liegen die der Primärquelle zugewandt ist am größten ist), ist diese auch als maximale Apertur definiert. Das andere Extremum ist eine einzige Sekundärquelle. Die Verteilung kann je nach Definition von Azimut- und Elevationswinkel der Apertur von punktförmigen, kreislinienförmigen bis zu einer flächigen Verteilung der Sekundärquellen reichen. Diese endliche Apertur entspricht einer räumlichen Fensterung mit einem Rechteck. Damit die Artefakte durch den Rand nicht zu stark auftreten [3] wird eine Gewichtungsfunktion über die Apertur gelegt wobei die Sekundärquellen die weiter außen liegen stärker abgeschwächt (Abb. 3) werden als jene die in der Mitte der Apertur liegen.



3.3 .Richtcharakteristik

Um nun die Richtcharakteristik der Apertur berechnen zu können wird die Fouriertransformation in den Ortsfrequenzbereich herangezogen. Der Einfachheit halber wird nur die Fouriertransformierte entlang einer Linie ("virtuelle" Lautsprecherzeile) gebracht, der Übergang auf Flächen bzw. 3 dimensionale Formen stellt keine Schwierigkeit dar, wird aber aus Gründen der Übersichtlichkeit nicht gebracht. Die Transformierte der Strahlerschnelle $v(y)$ ergibt sich dann entlang einer Lautsprecherzeile im Fernfeld zu [6]

$$p_{fern}(R, \varphi) = \frac{\omega \rho}{2} \sqrt{\frac{2}{\pi k_0 R}} e^{\frac{j\pi}{4}} e^{-jk_0 R} \int_{-\infty}^{\infty} v(y_q) e^{k_0 y_q \sin \varphi} dy_q \quad (\text{Glg. 21})$$

mit: ω ...Kreisfrequenz $2\pi \cdot f$
 ρ ...Dichte des Mediums
 k_0 ...freie Wellenzahl
 R ...Abstand von Mittelpunkt der Lautsprecherzeile zum Aufpunkt
 φ ...Winkel von Normalen-Richtung der Lautsprecherzeile

Beispiel einer kohas bewegten Kolbenmembran deren FT $V(k)$ von der Spaltfunktion gegeben ist durch

$$V(k) = v_0 l \frac{\sin(\frac{kl}{2})}{\frac{kl}{2}} = v_0 l \frac{\sin(\frac{\pi l}{\lambda})}{\frac{\pi l}{\lambda}} \quad (\text{Glg. 22})$$

Bei tiefen Frequenzen $l/\lambda_0 \ll 1$ wird nur ein sehr kleiner, praktisch konstanter Ausschnitt für die Richtwirkung sichtbar, die Abstrahlung erfolgt in alle Richtungen gleichmäßig. Bei hohen Frequenzen $l/\lambda_0 \gg 1$ hingegen werden viele Halbwellen des Spektrums für die Richtwirkung überdeckt. Dadurch ergibt sich eine starke Richtwirkung mit Hauptkeule und Nebenkeulen in Abstrahlrichtung.

Um nun eine konstante Richtwirkung zu erzielen wird die breite des räumlichen Fensters (in allen Raumrichtungen $W(\varphi, \vartheta)$) an die Frequenz angepasst. Dies geschieht durch Fensterungstechniken bei der ortsdiskreten Fensterung wie bei der Fensterung der zeitdiskreten Fouriertransformation. Dadurch können die Nebenkeulen gedämpft und die Hauptkeulenbreite beeinflusst werden.

3.4 .Grenzfrequenzen

3.4.1 .Obere Grenzfrequenz

Die untere Grenzfrequenz ist durch die Länge bzw. Längsausbreitung der Apertur gegeben. Sie gibt die Frequenz an ab der Auslöschungen stattfinden können (also richtungsabhängige Abstrahlung erfolgen kann).

$$L > \frac{\lambda_0}{2} \rightarrow f_{ug} > \frac{c}{2L} \quad (\text{Glg. 23})$$

3.4.2 .Untere Grenzfrequenz

Um räumliches Aliasing zu vermeiden ergibt sich die obere Grenzfrequenz, unter Einhaltung des Shannon Abtasttheorem's, zu:

$$\Delta x \leq \frac{\lambda_0}{2 \sin(\varphi)} \rightarrow f_{og} = \frac{c}{2 \Delta x \sin(\varphi_{max})} \quad (\text{Glg. 24})$$

3.5 .Aperturanpassung

Um nun die Auswirkung der schmaler werdenden Hauptkeule mit steigender Frequenz entgegenzuwirken und somit die Überhöhte Leistung in Hauptabstrahlrichtung für höhere Frequenzen zu bedämpfen wird die Apertur frequenzabhängig gemacht. Laut Theorie stellt sich eine konstante Richtcharakteristik, bei ortskontinuierlicher Quellverteilung ein wenn in jedem räumlichen Fenster die gleiche Anzahl an Wellenlängen platzfindet. Das heißt für hohe Frequenzen sind weniger Quellen in Betrieb als für tiefe Frequenzen.

3.6 .Nachbearbeitung bzw. Equalizing

Damit der Frequenzgang des Summenschallpegels der Apertur den gleichen Verlauf aufweist wie das Primärsignal, ist eine Filterung des Gesamtausgangssignals notwendig. Es kommt daher zu einer inversen Filterung mit den Gewichtungskoeffizienten die für die Raumfrequenzfilterung notwendig waren um eine konstante Richtwirkung zu erzielen.

3.7 .Anzahl der Lautsprecher für gute Fokussierung

In [7] wurde eine "äquivalente Monopol Distanz" definiert. Für ein Lautsprecherarray das auf einen bestimmten Punkt fokussiert, beschreibt sie die äquivalente Distanz die durch eine Monopolquelle wahrgenommen wird. Diese Berechnung basiert auf dem Verhältnis von Direkt- zu Diffusschall bzw. spät eintreffenden Diffusschall.

$$r_{eq} = \sqrt{\frac{W_a}{W}} * \left| \sum_{i=1}^N \frac{1}{r_i} e^{-jk(r_i - d_i)} \right|^{-1} \quad (\text{Glg. 25})$$

Mit W_a der Gesamtleistung die vom Lautsprecherarray abgestrahlt wird, W der Schalleistung der Monopolquelle, N der Anzahl der Arraylautsprecher, r_i der Distanz Fokus und i -tem LS und d_i der Distanz des Beobachters zum i -tem LS. Der Wert von r_{eq} gibt die Anzahl der benötigten Lautsprecher an damit eine aus dem Lautsprecherarray herauspringende Quelle wahrgenommen werden kann.

Ist nun eine Arrayanordnung gegeben bei der die einzelnen LS zueinander einen geringeren Abstand als $\lambda/2$ besitzen und unter der Bedingung einer konstanten Arraygröße, ist r_{eq} gleich dem Wert von $\lambda/2$. Daraus kann nun die Anzahl der benötigten Lautsprecher bzw. der Arraygröße für eine konstante Entfernungswahrnehmung abgeschätzt werden.

$$\frac{A}{\left(\frac{\lambda}{2}\right)^2} > \left(\frac{d}{r_{eq}}\right)^2 \quad (\text{Glg. 26})$$

A...Lautsprecherarrayfläche bzw. Aperturfläche

d...mittlere Abstand Beobachter (Abhörpunkt) – Lautsprecherarray

Inwieweit dieser Zusammenhang eine Verbesserung der Lokalisation und Entfernung bringt muss noch in folgenden Hörversuchen untersucht werden.

4 .Informeller Hörversuch

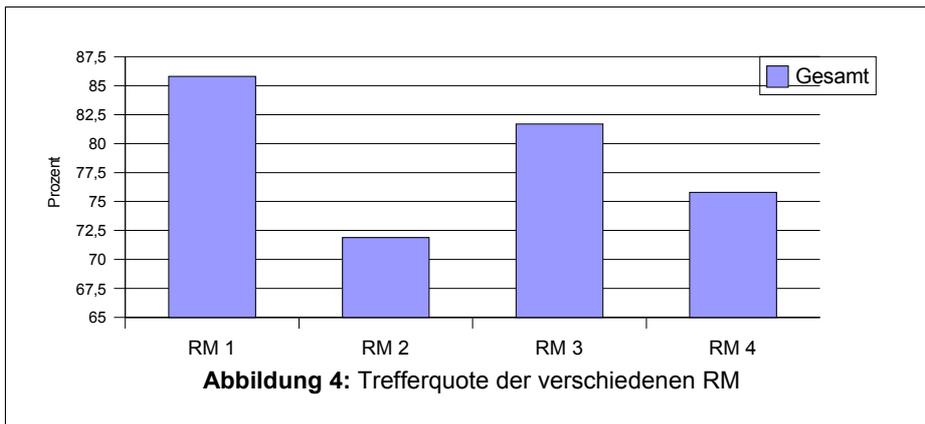
Um diese Methode zur Distanzkodierung auch quantitativ bewerten zu können, wurde ein Hörversuch durchgeführt. An diesem Hörversuch nahmen 18 Probanden teil. Davon waren 16 männlich und 2 weiblich. 14 Personen studieren Toningenieur und sind mit Hörversuchen vertraut, die restlichen 4 Personen nicht. Der Hörversuch dauerte circa 40 Minuten wobei 10 Minuten für eine Einführung durch den Versuchsleiter verwendet wurden. Die Probanden wurden dabei mündlich, graphisch und durch drei exemplarische Hörproben auf den Hörversuch vorbereitet. Danach wurden die Probanden in 30 Minuten mit 152 Hörbeispielen plus 8 zusätzlichen Fragen konfrontiert, was ein hohes Maß an Konzentrationfähigkeit von den Probanden abverlangte. Der Hörversuch wurde in zwei Aufgabenstellungen unterteilt. Die Erste beschäftigte sich mit der Distanzwahrnehmung ganz allgemein und unterteilte sich in 4 Raummodelle zu je 32 Hörbeispielen. Wobei die 4 Raummodelle unterschiedliche physikalische Räume darstellen, wie aus Tabelle 1 ersichtlich ist. Die zweite Frage blendete das Kriterium Pegel für die Distanzwahrnehmung aus. Es wurden zwei unterschiedliche Distanzen kodiert, wobei beide jeweils um ± 3 und ± 6 dB im Pegel absichtlich verändert wurden, daraus ergaben sich 24 Hörbeispiele. Ein Hörbeispiel gestaltete sich in Form eines A-B Vergleichs. A und B waren kurze Sprachsignale von einem männlichen oder weiblichen Sprecher, die Sprecher wurden zufällig zugeteilt. Zwischen den Hörproben A und B wurde eine Pause von 500 ms gewählt. Die Aufgabe der Probanden war nun B bezüglich A zu bewerten, wobei sie auf einer diskreten Skala entscheiden mussten. Die Skala reichte von "B ist viel näher als A" bis "B ist viel weiter entfernt als A". Zwischen den einzelnen Hörbeispielen hatten die Probanden 5 Sekunden um sich zu entscheiden. Das Sprachsignal wurde in 4 Distanzen (0.35m, 0.7m, 1.4m, 5.6 m) kodiert, wobei jede Kombinationsmöglichkeit zweimal überprüft wurde.

	Raummodell 1 (RM 1)	Raummodell 2 (RM 2)	Raummodell 3 (RM 3)	Raummodell 4 (RM 4)
Hallradius r_H [m]	2,8	1	∞	∞
Raummaße	7,1 x 5,2 x 3,8	11 x 8,1 x 5,5	-	-
HRTF-Satz	CIPIC	CIPIC	KEMAR	KEMAR

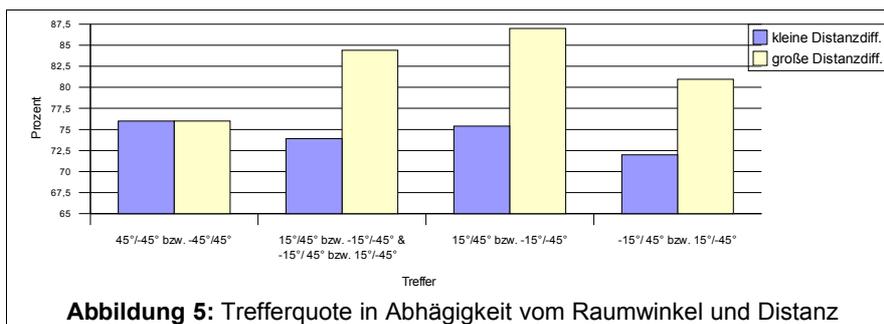
Tabelle 1: Difinition der Raummodelle

Weiters wurde angenommen, dass methathetische Faktoren wie weiblicher oder männlicher Sprecher bzw. die Richtung in Azimut aus der die Sprachsignale kommen, keinen Einfluss auf den prothetischen Faktor der Distanz haben. Das bedeutet, A und B konnten beliebig aus einer der vier verwendeten Richtungen (-45° , -15° , 15° und 45°) kommen [8]. Für die zweite Aufgabenstellung wurden die methathetischen Faktoren nicht als unabhängig betrachtet, um diesen Einfluss ausschließen zu können und es den Probanden zu ermöglichen sich ganz auf die Entfernungswahrnehmung zu konzentrieren. Jene Probanden, die keine reliablen oder validen Aussagen beim Hörversuch lieferten, wurden aus der Wertung gestrichen. Bei der ersten Aufgabenstellung waren dies 3 Probanden.

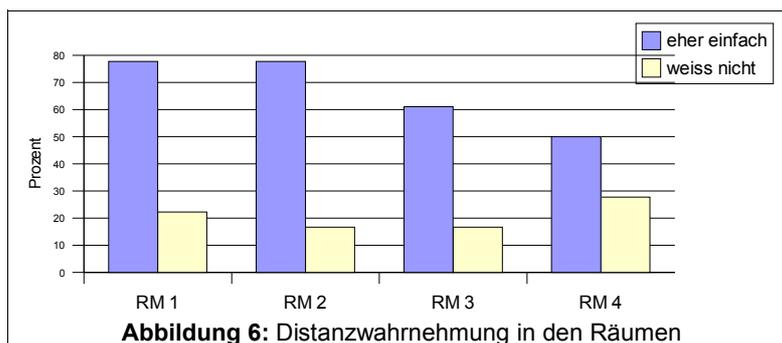
Sehr interessante Ergebnisse ergab die Auswertung vom RM 1 und RM 2 für die Aufgabenstellung 1. Der Einfluss des Hallradius beziehungsweise der Raumgröße ist deutlich erkennbar (siehe Abb. 3). Für das RM 2 lag der Hallradius genau zwischen den beiden mittleren kodierten Distanzen. Es lagen also zwei



kodierte Distanzen außerhalb des Hallradius, für das RM 1 war es nur Eine. Aus diesem Ergebnis ist ersichtlich dass die Distanzwahrnehmung für Quellen die sich im Diffusschallfeld befinden ungleich schwieriger wird. Die Abweichung der Trefferquote zwischen RM 3 und RM 4 die sich ja nur durch einen anderen Satz von HRTFs unterschieden, kann nur den genannten Einfluss beziehungsweise durch Ermüdungserscheinungen der Probanden erklärt werden.

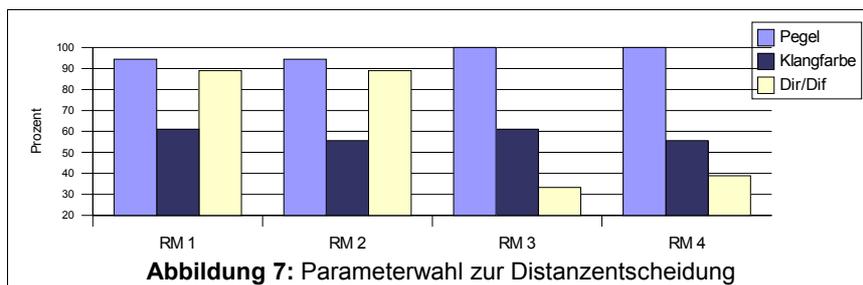


In Abbildung 4 ist eine Aufschlüsselung der richtigen Antworten (Trefferquote) abhängig vom Raumwinkel zwischen den Vergleichspaaren dargestellt. Die Anzahl der Trefferquote ist für Vergleichspaare die in Distanz nahe beieinander und weit auseinander liegen aufgeschlüsselt. Für Vergleichspaare die in Distanz nahe beieinander und weit auseinander liegen aufgeschlüsselt. Für Vergleichspaare aus der Richtung 45° oder -45° für A und -45° oder 45° für B spielt es keine Rolle ob die Distanzdifferenz zwischen dem Vergleichspaar groß oder klein ist. Jedoch kann dies nicht für ±15° Richtungen kombiniert mit den ±45° bestätigt werden. Der Grund für diesen Sachverhalt ist in der unterschiedlich wahrgenommenen Externität abhängig von der Richtung der HRTFs zu suchen, was Untersuchungen von [9] bestätigen.

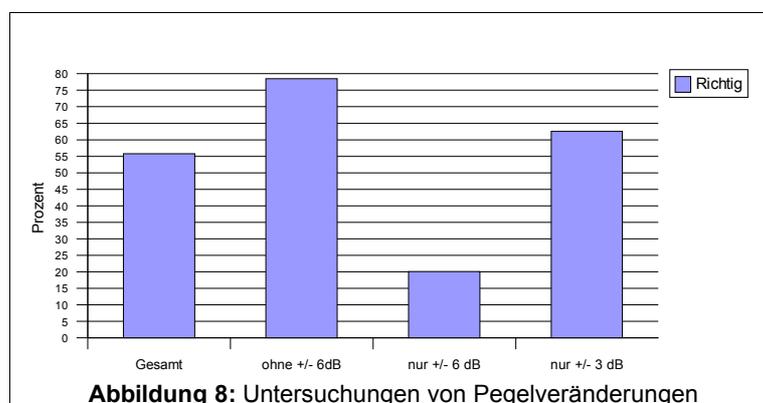


Am Ende jedes Raummodells folgten je zwei Fragen: Zum Einen mussten die Probanden angeben ob sie Entfernungskodierung leicht oder schwer wahrgenommen haben (siehe Abb. 6). Zum Anderen nach welchen Kriterien sie die Distanzunterscheidung vorgenommen hatten. Zur Auswahl standen Pegel, Klangfarbe und Direkt- zu Diffusschall, wobei dies auf einer Skala von „trifft zu“ bis „trifft nicht zu“ zu kennzeichnen war. Wie in Abbildung 5 zu erkennen ist, war für die Raummodelle RM 1 und RM 2, bei denen eine vollständige Raumsimulationen erfolgte, die Entfernungskodierung erheblich leichter als für RM 3 und RM 4.

Die Auswertung der Kriterien nach der die Probanden ihrer Entscheidungen getroffen haben, ergab folgendes (siehe Abb. 7): Für das Raummodell RM 1 und RM 2 ist zur Entfernungskodierung zusätzlich zum Pegel der Raumeindruck, d.h. das Verhältnis von Direkt- zu Diffusschall, wesentlich.



In der Aufgabenstellung zwei wurden, wie oben erwähnt, die Pegel der kodierten Signale um ± 3 dB und ± 6 dB verändert. Dabei sollte der Einfluss des Pegels auf die Distanzwahrnehmung bestimmt und die Frage geklärt werden, um bis zu wieviel dB ein Signal verändert werden kann um dennoch in der richtigen Distanz wahrgenommen zu werden. Abbildung 8 illustriert den Anteil der richtigen Antworten in Prozent. Der erste Balken zeigt die Gesamttrefferquote. Diese gibt zu erkennen, dass die Entfernungskodierung bei vollständiger Ausblendung des Kriteriums Pegel nicht zu einer signifikanten Aussage führt. Der zweite Balken liefert das Ergebnis ohne die Vergleichspaare die um ± 6 dB verändert wurden. Der Dritte repräsentiert die Erkennungsrate jener Vergleichspaare die im Pegel mit ± 6 dB verändert und der Vierte nur jene die um ± 3 dB verändert wurden.



Die Schlussfolgerung aus diesem Ergebnis ist, dass der Pegel nachwievor einen entscheidenden Einfluss auf die Distanzwahrnehmung hat, jedoch bei guter räumlicher Kodierung um ± 3 dB verändert werden kann, wobei sich der Entfernungseindruck nicht entscheidend ändert.

5 .Zusammenfassung

Ausgehend vom KHI wurde hier ein leistungsfähiges Tool zur binauralen Distanzkodierung vorgestellt. Die Auswertungen des informellen Hörversuchs lieferte positive Ergebnisse und ist somit ein gute Möglichkeit zur Distanzkodierung. Darüber hinaus konnten Übereinstimmungen mit anderen Untersuchungen festgestellt werden. Neben den üblichen Kriterien zur Entfernungswahrnehmung, wie der Pegelabnahme mit Entfernungszunahme, dem Verhältnis von Direkt- zu Diffusanteil und Klangfarbenänderung wurden die Primärquelle und einige frühe Reflexionen nicht nur zeitlich und örtlich sondern auch entfernungsabhängig physikalisch modelliert. Als weitere Schritte sind, ein Performancevergleich mit anderen distanzkodierenden Verfahren und die Einbindung in ein bestehendes recheneffizientes Ambisonic-Verfahren geplant.

6 .Referenzen

- [1] A. Sontacchi, R. Höldrich: „Distance Coding in 3-D Sound Fields“, AES 21st Conference, St. Petersburg, Russia, 2002 June 1–3.
- [2] E. Start, “Direct sound enhancement by wave field synthesis”, 1997.
- [3] Verheijen, E.: „Sound Reproduction by Wave Field Synthesis“, Dissertation, Delft University of Technology, The Netherlands, 1997.
- [4] Start, E. W.: „Application of Curved Arrays in Wave Field Synthesis”, Preprint Nr. 4143, presented at the 100th AES Convention, Copenhagen, Denmark, 1996.
- [5] V. R. Algazi, R. O. Duda, D. M. Thompson and C. Avendano, "The CIPIC HRTF Database," Proc. 2001 IEEE Workshop on Applications of Signal Processing to Audio and Electroacoustics, pp. 99-102, Mohonk Mountain House, New Paltz, NY, Oct. 21-24, 2001.
http://interface.cipic.ucdavis.edu/CIL_html/CIL_HRTF_database.htm
- [6] M. Möser: „Analyse und Synthese akustischer Spektren”, Springer,1988.
- [7] S. Komiyama, Y. Nakayama, K. Ono and S. Koizumi, “A loudspeaker-array to control sound image distance”, Acoust. Sci. & Tech. 24, 5 (2003) S. 242-249
- [8] Stevens
- [9] P. Majdak, M. Noisternig: „Implementation kopfpositionsbezogener Binauraltechnik“, Institut für Elektronische Musik und Akustik, Kunstuniversität Graz, 2002.