

Masterthesis im Studienfach
Master of Computer Science

**Design und Implementierung einer modularen
Bibliothek zur Einbindung von 3-D Sound in
interaktive Echtzeit – Anwendungen**

Dipl. Medieninform. (FH) Severin S. Todt
28. August 2003

Inhaltsverzeichnis

INHALTSVERZEICHNIS	I
ABBILDUNGSVERZEICHNIS	IV
LITERATURVERZEICHNIS.....	V
ABSTRAKT	1
1. EINLEITUNG	1
2. GRUNDLAGEN	2
2.1. Entstehung und Wahrnehmung von Sound	2
2.1.1. Physikalische Beschreibung von Schallereignissen	2
2.1.1.1. Leistung einer Soundquelle	2
2.1.1.2. Intensität einer Soundquelle	2
2.1.1.3. Schalldruck.....	3
2.1.1.4. Schalldruckpegel	3
2.1.1.5. Lautstärkepegel	3
2.1.2. Schallwahrnehmung	3
2.1.2.1. Head Related Transfer Function	3
2.1.2.2. Wahrnehmung der Position von Soundquellen.....	4
2.1.3. Schallausbreitung	4
2.1.3.1. Raumwahrnehmung.....	4
2.1.3.2. Doppler Effekt.....	5
2.2. Grenzen des menschlichen Hörens	5
2.2.1. Genauigkeit bei der Ortung von Soundquellen.....	5
2.2.2. Frequenzabhängige Ortung von Soundquellen	6
2.3. Spatial Manipulation	6
2.4. Techniken für 3-D Auditory Displays.....	6
2.4.1. 3-D Auditory Displays auf Basis stereophoner Techniken	6
2.4.1.1. Einkanal Lautsprecher Setup.....	7
2.4.1.2. Zweikanal Lautsprecher Setup	7
2.4.1.3. Vierkanal Lautsprecher Setup	8
2.4.1.4. Mehrkanal Lautsprecher Setup.....	8
2.4.1.5. Kubisches Achtkanal Lautsprecher Setup.....	8
2.4.2. 3-D Auditory Displays auf Basis binauraler Techniken.....	9
2.4.2.1. Kopfhörer Nutzung zur Umsetzung binauraler Techniken.....	9
2.4.2.2. Lautsprecher Nutzung zur Umsetzung binauraler Techniken.....	9
3. ANFORDERUNG AN INTERAKTIVE ECHTZEIT 3-D SOUND SYSTEME.....	10
3.1. Anforderungen gemäß I3DL2.....	10
3.1.1. I3DL2 Minimalanforderungen	10
3.1.2. Kritik an den I3DL2 Minimalanforderungen	10
3.2. Entwurf eines Anforderungs-katalogs für interaktive 3-D Sound Systeme	11
3.2.1. Datenaspekte des Anforderungskataloges.....	11

3.2.1.1.	Virtuelle Soundquelle	11
3.2.1.2.	Virtueller Hörer	12
3.2.1.3.	Virtuelles Environment.....	12
3.2.1.4.	Physikalischer Hörer.....	13
3.2.1.5.	Physikalisches Environment.....	13
3.2.2.	Verarbeitungsaspekte des Anforderungskataloges	13
3.2.3.	Wiedergabeaspekte des Anforderungskataloges.....	13
3.2.4.	Abstraktionsgrade des Anforderungskataloges	13
3.2.4.1.	Abstraktionslevel 1	14
3.2.4.2.	Abstraktionslevel 2	14
3.2.4.3.	Abstraktionslevel 3	14
3.2.4.4.	Abstraktionslevel 4	14
3.2.5.	Anforderungskatalog für 3-D Sound Systeme.....	14
4.	BEURTEILUNG EXISTIERENDER INTERAKTIVER 3-D SOUND SYSTEME	15
4.1.	Beurteilung von Special – Purpose 3-D Sound Systemen.....	15
4.2.	Beurteilung von Multi – Purpose 3-D Sound Systemen.....	16
5.	DESIGN EINES INTERAKTIVEN 3-D SOUND SYSTEMS	18
5.1.	Zielsetzung	19
5.1.1.	Anforderungskatalog konformer Entwurf.....	19
5.1.2.	Betriebssystem - Unabhängigkeit.....	19
5.1.3.	Modulares Design.....	19
5.1.4.	Special – Purpose Entwurf mit möglichem Multi – Purpose Einsatz	19
5.2.	Erweiterung existierender 3-D Sound Systeme zur Zielerfüllung.....	19
5.3.	Design Aspekte beim Entwurf eines 3-D Sound Systems	20
5.3.1.	Analyse 3-D Sound System	20
5.3.1.1.	Analyse der Wiedergabe Aspekte.....	20
5.3.1.2.	Analyse der physikalischen Aspekte.....	21
5.3.1.3.	Analyse der virtuellen Aspekte.....	21
5.3.1.4.	3-D Sound System im Detail.....	21
5.4.	Entwurf eines 3-D Sound Systems	21
5.4.1.	Entwurf der Datenstrukturen	21
5.4.1.1.	Datenstruktur Physikalisches Environment.....	22
5.4.1.2.	Datenstruktur Physikalischer Hörer	22
5.4.1.3.	Datenstruktur Sound Device.....	22
5.4.1.4.	Datenstruktur Virtuelle Soundquelle.....	22
5.4.1.5.	Zentrale Datenstruktur 3-D Sound System.....	23
5.4.2.	Entwurf der Verarbeitungsprozesse	23
5.4.2.1.	Identifizierung der Verarbeitungsprozesse	23
5.4.2.2.	3-D Sound Verarbeitungs – Pipeline.....	24
5.4.3.	Entwurf der Schichtstruktur	25
6.	IMPLEMENTIERUNG EINES INTERAKTIVEN 3-D SOUND SYSTEMS	26
6.1.	Methoden und Techniken der Implementierung.....	26
6.2.	Modularisierung.....	26
6.2.1.	Audio Source Modul	27
6.2.1.1.	Datenstrukturen und Funktionen	27
6.2.1.2.	Aktualisierung von Audio Source Daten.....	27

6.2.1.3.	Synchronisation des Audio Source Moduls.....	28
6.2.1.4.	Ein- und Ausgabedaten	28
6.2.2.	Audio Buffer Modul	29
6.2.2.1.	Datenstrukturen und Funktionen	29
6.2.2.2.	Anpassung der Intensität über die Distanz	29
6.2.2.3.	Synchronisation des Audio Buffer Moduls	30
6.2.2.4.	Ein- und Ausgabedaten	30
6.2.3.	Audio Pipe Modul	30
6.2.3.1.	Datenstrukturen und Funktionen	30
6.2.3.2.	Berechnung des 3-D Klangbildes	31
6.2.3.3.	Klangbild Berechnung	31
6.2.3.4.	Synchronisation des Audio Pipe Moduls	32
6.2.3.5.	Ein- und Ausgabedaten	33
6.2.4.	Audio Device Modul	33
6.2.4.1.	Datenstrukturen und Funktionen	33
6.2.4.2.	Multidevice Implementierung	33
6.2.4.3.	Berechnung von Intensitätsanpassungen für wechselnde Hörerpositionen	34
6.2.4.4.	Ein- und Ausgabedaten	34
6.2.5.	Audio Processing Modul.....	34
6.2.5.1.	Datenstrukturen und Funktionen	34
6.2.6.	Audio Byteorder Modul.....	35
6.2.7.	Audio Error Modul.....	35
6.3.	Integration in Multimedia Entwicklungsumgebungen.....	35
6.4.	Implementierung der Schichtstruktur.....	36
6.5.	Erweiterungsmöglichkeiten	36
6.5.1.	Erweiterungsmöglichkeiten Audio Device Modul	36
6.5.2.	Erweiterungsmöglichkeiten Audio Source Modul.....	36
6.5.3.	Erweiterungsmöglichkeiten Audio Buffer Modul	36
6.5.4.	Erweiterungsmöglichkeiten Audio Integration Modul.....	37
7.	ZUSAMMENFASSUNG UND AUSBLICK	37
	EIDESSTATTLICHE ERKLÄRUNG	38

Abbildungsverzeichnis

Abbildung 1 : Hörschwelle und Kurven gleicher Lautstärke	3
Abbildung 2 : ITD und IID.....	4
Abbildung 3 : Environment Reverberation [IAS_99].....	5
Abbildung 4 : Zweikanal Lautsprecher Setup	7
Abbildung 5 : Stereophone Fläche	8
Abbildung 6 : Kubisches Achtkanal Lautsprecher Setup	9
Abbildung 7 : Ausbreitungsverhalten virtueller Soundquellen	12
Abbildung 8 : Übersicht Anforderungskatalog.....	15
Abbildung 9 : Beurteilung existierender interaktiver 3-D Sound Systeme.....	18
Abbildung 10 : Top – Level Sicht 3-D Sound System	20
Abbildung 11 : Detaillierte Sicht 3-D Sound System.....	20
Abbildung 12 : Wiedergabe – Modul im Detail.....	21
Abbildung 13 : Physikalische Aspekte im Detail.....	21
Abbildung 14 : Virtuelle Aspekte im Detail	21
Abbildung 15 : 3-D Sound System im Überblick	22
Abbildung 16 : Entwurf Datenstruktur	23
Abbildung 17 : Entwurf der Verarbeitungsprozesse	24
Abbildung 18 : Verarbeitungspipeline	25
Abbildung 19 : Verarbeitungspipeline mit externem konstanten Einfluß.....	25
Abbildung 20 : Schichtstruktur	25
Abbildung 21 : Modulübersicht.....	26
Abbildung 22 : Auszug der Datenstruktur des Audio Source Modul	27
Abbildung 23 : Auszug der Funktionen des Audio Source Modul.....	27
Abbildung 24 : Auszug der Datenstruktur des Audio Buffer Modul.....	29
Abbildung 25 : Auszug der Funktionen des Audio Buffer Modul	29
Abbildung 26 : Zyklische Koppelspeicher.....	30
Abbildung 27 : Auszug der Datenstruktur des Audio Pipe Modul	31
Abbildung 28 : Auszug der Funktionen des Audio Pipe Modul	31
Abbildung 29 : Klangbild Berechnung.....	31
Abbildung 30 : Ambisonics.....	32
Abbildung 31 : Auszug der Datenstruktur des Audio Device Modul	33
Abbildung 32 : Auszug der Funktionen des Audio Device Modul	33
Abbildung 33 : Wiedergabeprozess Multidevice.....	34
Abbildung 34 : Auszug der Datenstruktur des Audio Device Modul	34
Abbildung 35 : Auszug der Funktionen des Audio Processing Modul	35
Abbildung 36 : Auszug der Funktionen des Audio Processing Modul	35
Abbildung 37 : Module der Schichtenstruktur	36

Literaturverzeichnis

- BEG_91** Begault, Durand R. : Preferred Sound Intensity Increase For Sensation of Half Distance, National Research Council, NASA Ames Research Center, Kalifornien, USA, 1991
- BEG_94** Begault, Durand R. : 3-D sound For Virtual Reality And Multimedia, Moffet Field, NASA Ames Research Center, Academic Press Professional, Kalifornien, USA, 1994, Reprint 2000
- BEG_99** Begault, Durand R. : Auditory And Non-Auditory Factors That Potentially Influence Virtual Acoustic Imagery, Moffet Field, Human Factors and Technology Division - NASA Ames Research Center, Kalifornien, USA, 1999
- DIC_97** Dickreiter, Michael : Handbuch der Tonstudioteknik, München, Saur, K.G, 1997
- FAU_00** Faust, Simon : 3-D Akustik mit Lautsprechern, Aachen, RWTH Aachen Virtual Reality and 3D Visualization, 2000
- FUN_99** Funkhouser, Thomas u.a. : A Beam Tracing Approach to Acoustic Modeling for Interactive Virtual Environments, Princeton, 1999
- GAR_98** Gardner, William G. : 3-D Audio Using Loudspeakers, Boston, Massachusetts Institute of Technology, Kluwer Academic Publishers, Norwell, Massachusetts, 1998
- GAR_WWW_1** Gardner, Bill, Martin Keith : HRTF Measurements of a KEMAR Dummy-Head Microphone, <http://sound.media.mit.edu/KEMAR.html>, 04.08.2003
- GRO_90** Groening, Matt : The Call of the Simpsons, Twentieth Century Fox, 1990
- HAS_03** Hass, Jeffrey, An Acoustics Primer, <http://www.indiana.edu/~emusic/acoustics/amplitude.htm>, Center for Electronic and Computer Music, School of Music Indiana University, Bloomington, Indiana, 2003
- IAS_99** IASIG : IASIG Interactive 3D Audio Guidelines (Level 2), 1999
- JOT_96** Jot, Jean-Marc : Synthesising Three-Dimensional Sound Scenes in Audio or Multimedia Production and Interactive Human – Computer Interfaces, Ircam - Centre Georges-Pompidou, 5th International Conference, Interface to Real & Virtual Worlds, Montpellier, 1996
- KEN_WWW_1** Kendall, Gary : A 3-D sound Primer, <http://www.northwestern.edu/musicschool/classes/3D/pages/3DsoundPrimer.html>, 11.02.2002
- KUH_89** Kuhn, Wilfried : Physik Band II, Braunschweig, Westermann Schulbuchverlag, 1989
- PUL_01** Pulkki, Ville : Spatial Sound Generation and Perception by Amplitude Panning Techniques, Helsinki, Laboratory of Acoustics and Audio Signal Processing - Helsinki University of Technology, Espoo 2001, Report 62, 2001
- STO_95** Stolz, Dieter : Computergestützte Audio- und Videotechnik, Berlin, Springer Verlag Berlin Heidelberg, 1995
- TOD_02_1** Todt, Severin S. : Integration of Multichannel Sound in Virtual Environments, Fachhochschule Wedel, Februar 2002
- TOD_02_2** Todt, Severin S. : Spatialized Sound Server Implementation, Electronic Visualization Laboratory, University of Illinois at Chicago, Summer 2002
- TOD_03** Todt, Severin S. : Untersuchung und Bewertung von Soundbibliotheken auf ihre Einsetzbarkeit in Echtzeit 3D Grafikanwendungen, Juni 2003

TON_WWW_93 Tonnesen, Cindy; Steinmetz, Joe : 3-D sound Synthesis, Washington, The Encyclopedia of Virtual Environments, <http://www.hitl.washington.edu/scivw/EVE/I.B.1.3DSoundSynthesis.html>, 1993

Abstrakt

Der Immersionsgrad interaktiver Multimedia – Systeme steigt mit denen in die Anwendung integrierten Sinne des Anwendungsnutzer. Während sich die visuelle Informationsvermittlung über grafische Displaysysteme stetig und mit einem exponentiellen Wachstum entwickelt, bleibt die Integration weiterer Rückkopplungskanäle, speziell die Integration von interaktivem 3-D Sound, zumeist unbeachtet.

Diese Arbeit führt in die Thematik der 3-D Soundwahrnehmung und –verarbeitung ein und definiert Anforderungen, die an interaktive 3-D Sound Systeme zu stellen sind, um eine störungsfreie Integration in interaktive Echtzeit – Anwendung und damit eine Erhöhung des Immersionsgrades zu gewährleisten. Der entworfene Anforderungskatalog dient im weiteren Verlauf der Arbeit als Grundlage zur Analyse vorangegangener Ansätze zur Integration von 3-D Sound. Aus den Analyse – Ergebnissen, in Verbindung mit dem Anforderungskatalog, erwächst der Entwurf eines modularen 3-D Sound Systems. Der Entwurf zeigt einen möglichen Weg zur Umsetzung eines Systems, das in der Lage ist, Sound unter Berücksichtigung von räumlicher Position und Ausrichtung in Echtzeit zu verarbeiten, um ein Klangbild zu erzeugen, das möglichen Nutzern des Systems erlaubt, Rückschlüsse auf das räumliche Verhalten von Soundobjekten zu ziehen. Der mit dem Entwurf beschriebene Weg zur Umsetzung wird mit einer beispielhaften Implementierung eines 3-D Sound Systems im Laufe der Arbeit besprochen. Die mit der Implementierung umgesetzte Kopplung der Soundverarbeitung an die Entwicklungsumgebung OpenGL® zeigt mögliche Wege zur Synchronisation der visuellen und akustischen Rückkopplung und zur Anbindung von Soundeigenschaften an Grafikobjekte. Als Ergebnis der Arbeit entsteht ein interaktives System zur Einbindung von 3-D Sound in interaktive Echtzeit – Grafik – Anwendungen.

1. Einleitung

Die Einbindung von Sound in interaktive Multimedia – Anwendungen hat schon Ende der 80er Jahre öffentliches Interesse erfahren können. Die Soundverarbeitung und –manipulation, speziell die Manipulation von Soundeigenschaften aufgrund der räumlichen Anordnung einer Soundquelle in Echtzeit, bleibt derzeit allzu oft bei der Erstellung multimedialer Anwendungen unberücksichtigt.

Inbesondere ist der Bereich der Virtual Reality Anwendung den Sound betreffend ein nur wenig berührtes Gebiet. Virtual Reality Anwendungen dienen in den meisten Fällen der Visualisierung von Informationen. Nur ein kleiner Teil in diesem Bereich befasst sich überhaupt mit der Thematik Sound. In der Simulation von hochwertigen oder sicherheitssensitiven Gerätschaften, bei denen die Akustik zur Bedienung eine entscheidende Rolle spielt, oder aber im noch sehr kleinen Bereich der unterhaltenden Virtual Reality, wird sich mit der Thematik Sound ernsthaft auseinandergesetzt. Selbst wenn es sich hier zumeist um interaktive Echtzeit-Anwendungen handelt, werden den Sound betreffend andere Maßstäbe angelegt als bei der Visualisierung.

Wird die Grafik zumeist in Echtzeit und interaktiv berechnet, genügen sich Entwickler von Virtual Reality Anwendungen sowie deren Anwender mit Sound, der zwar interaktiv beim Auftreten eines gewissen Ereignisses abgespielt wird, jedoch auf im Voraus berechneten Sounddaten basiert.

Außerhalb des Bereiches Virtual Reality, im Entertainment - Bereich von Multimedia – Anwendungen, insbesondere im Gebiet der Spieleentwicklung, findet die Verwendung von Sound weitaus größere Beachtung. Hier wird seit Ende der 90er Jahre verstärkt Sound in Multimedia – Anwendungen integriert, der nicht nur durch Ereignisse getriggert, weiter noch in Echtzeit berechnet wird. Dies wird unter Anderem durch die vermehrt verfügbaren Bibliotheken

zur Einbindung von Sound in interaktive Echtzeit – Anwendungen begünstigt. Diese erleichtern zum Teil die Einbindung von Sound in spezielle Multimedia – Entwicklungsumgebungen, oder stellen ein Kollektiv von Datenkonstrukten und Funktionen zur Implementierung von Soundverarbeitung in Echtzeit zur Verfügung. Nur ein Bruchteil dieser ermöglicht jedoch die interaktive Verarbeitung und Berechnung von 3-D Sound.

3-D Sound ist Sound, dessen Daten unter Berücksichtigung von räumlichen Attributen, wie Position und Orientierung, berechnet und/oder manipuliert werden. Er wird in einer Art und Weise wiedergegeben, die den Hörer einen den räumlichen Attributen entsprechenden Raumeindruck erfahren lässt.

Kaum eine der verfügbaren Bibliotheken geht auf die speziellen Belange von Virtual Reality Systemen und deren Anwendungsumfeld ein, um den Immersionsgrad positiv zu beeinflussen.

Virtual Reality Systeme sind Systeme, die unter Nutzung aller angeschlossener Subsysteme, die des Feedbacks oder der Mensch – Computer Interaktion dienen, Echtzeit – Interaktivität ermöglichen und der Vermittlung von Daten dienen.

Immersiongrad ist das Maß eines Virtual Reality Systems, das darüber eine Aussage trifft, in wie weit sich der Anwendungsnutzer als Bestandteil einer Anwendung fühlt.

Ein Weg, den Immersionsgrad von Virtual Reality Systemen zu erhöhen, führt über die Einbindung einer Vielzahl verschiedener menschlicher Sinnesorgane für das Feedback. Die Einbindung von 3-D Sound wird den Grad der Immersion erhöhen helfen.

Zur Erhöhung des Immersionsgrades ist weiter die Berücksichtigung des Verhältnisses zwischen

Anwendungsnutzer und Virtual Reality System, wie die Position und Orientierung des Benutzers relativ zum System, von essentieller Wichtigkeit.

Ziel dieser Arbeit ist der Entwurf und die Implementierung einer Bibliothek, die eine interaktive 3-D Sound – Verarbeitung und Wiedergabe ermöglicht. Die Bibliothek soll das Anwendungsumfeld des genutzten Virtual Reality Systems, wie das Verhältnis des Nutzers relativ zum System bei der Verarbeitung berücksichtigen und die Einbindung von 3-D Sound in interaktive Echtzeit – Grafikanwendungen vereinfachen. Sie soll dann in Virtual Reality Systemen wie auch im Umfeld von Multimedia – Anwendungen zum Einsatz kommen und so den Einsatz von Echtzeit 3-D Sound im Multimedia- und Virtual Reality Bereich weiter etablieren helfen.

In Kapitel 2 soll zunächst mit den Grundlagen der Soundwahrnehmung des Menschen und den speziellen Aspekten von 3-D Sound ein Überblick über die vielseitigen Aspekte der 3-D Sound – Verarbeitung gegeben werden.

In Kapitel 3 soll auf die Anforderungen an Systeme zur Echtzeit - Verarbeitung von 3-D Sound eingegangen werden. Es wird ein Anforderungskatalog erstellt, nach dem existierende Sound Systeme und - Bibliotheken zur 3-D Sound – Verarbeitung bewertet werden können und der als Richtlinie beim Design und der Implementierung neuer 3-D Sound Systeme dienen kann.

In Kapitel 4 werden existierende Systeme auf Grundlage des in Kapitel 3 erstellten Bewertungskataloges auf ihre Leistungsfähigkeit und Einsetzbarkeit in Virtual Reality Systemen hin untersucht und beurteilt.

In Kapitel 5 wird eine neue Bibliothek zur Verarbeitung von 3-D Sound in Echtzeit unter Berücksichtigung der im Bewertungskatalog aus Kapitel 3 festgehaltenen Kriterien entworfen.

In Kapitel 6 werden die zur Implementierung des Designs aus Kapitel 5 gewählten Strategien erläutert. Die Kernaspekte der entwickelten Bibliothek werden in ihrer Funktion kurz erläutert und die zur Umsetzung genutzten Algorithmen beschrieben.

Das Kapitel 7 fasst die erzielten Ergebnisse kurz zusammen und gibt einen Ausblick auf zukünftige Entwicklungen.

Die Ausarbeitung wie auch die Implementierung der 3-D Sound Bibliothek stellt keine vollständige Beschreibung aller Aspekte der 3-D Sound – Verarbeitung dar, noch sind alle Aspekte der 3-D Sound – Verarbeitung in der Bibliothek vollständig implementiert. Vielmehr stellt die Arbeit die Konzepte und eine mögliche Umsetzung dieser in der Programmiersprache „C“ dar. Nicht zur Diskussion stehen hier die Einsatzgebiete und die beim Einsatz von 3-D Sound entstehenden Vorteile. Diese sind aus der im Literaturverzeichnis angegebenen Literatur erkenntlich.

2. Grundlagen

Dieses Kapitel führt in die Grundlagen der Akustik und der Soundwahrnehmung des Menschen ein und zeigt die wichtigsten Aspekte der 3-D Sound – Verarbeitung auf. Für die 3-D Sound – Verarbeitung sind im speziellen die Techniken zur Erzeugung eines dreidimensionalen Klangbildes beim Hörer sowie die zur Verarbeitung und Erzeugung eines real wirkenden Eindrucks notwendigen Attribute von besonderer Wichtigkeit.

2.1. Entstehung und Wahrnehmung von Sound

In der Akustik sind zwei Arten von Ereignissen zu unterscheiden. Das Ereignis der Soundentstehung und –ausbreitung und das Ereignis der Wahrnehmung des Schalls. Eine Schallwelle, die von einer Soundquelle ausgeht, sich im Raum ausbreitet und auf das Ohr eines Hörers trifft, ist eindeutig physikalisch definierbar. Dieses Schallereignis trifft als Reiz auf das Gehör, wandert durch das komplizierte mechanische System des Gehörs und wird schließlich als nervöse Information zum Gehirn geleitet. Erst jetzt wird dem Hörer der Reiz als Hörereignis, als Empfindung bewusst.

2.1.1. Physikalische Beschreibung von Schallereignissen

Schallereignisse lassen sich durch die charakterlichen Größen Leistung, Intensität und Schalldruck einer Soundquelle beschreiben.

Der Schalldruck steht in direktem Zusammenhang zur akustischen Energie oder Intensität einer Schallquelle. Beide, Schalldruck und Intensität, sind direkt abhängig von der Leistung der Soundquelle.

2.1.1.1. Leistung einer Soundquelle

Stellt man sich die von einer Soundquelle ausgehenden Schallwellen als sich kugelförmig im Raum ausbreitende Energie vor, beschreibt die Leistung die Summe der kinetischen Energie auf der Oberfläche dieser Kugel. Leistung wird in Watt angegeben. 1 Watt entspricht 1 Newton/Sekunde. [HAS_03]

2.1.1.2. Intensität einer Soundquelle

Die Betrachtung der Leistung einer Soundquelle in Verbindung mit dem Abstand von der Soundquelle bei der Messung formulieren den Zusammenhang der Intensität. Intensität wird angegeben in Watt/Sekunde oder Watt/m².

Mit der Ausbreitung der Oberfläche der Kugel wird die Energie, die von einer Soundquelle ausgeht, über eine sich exponential vergrößernde Fläche verteilt. Die Energie in einem Quadratmeter der Kugeloberfläche nimmt mit dem inversen Quadrat der Entfernung ab.

$$\text{Intensitätsminderung} = \frac{1}{\text{Entfernung}^2}$$

z (Oberfläche einer Kugel : $4\pi r^2$, Intensität auf der Kugeloberfläche $s/4\pi r^2$ mit s = Intensität der Soundquelle)
Dieser funktionale Zusammenhang wird bezeichnet als Inverse Square Law.

Inverse Square Law beschreibt das physikalische Modell bei der Ausbreitung von Wellen, wonach die Intensität mit dem Quadrat der Entfernung abnimmt, bzw. sich die Intensität mit der Verdopplung der Distanz um 6 dB vermindert. [BEG_91]

Die akustische Energie einer Soundquelle wird somit an einer doppelt soweit von ihr entfernten Position über eine viermal so große Fläche verteilt. Die Intensität beträgt dann nur $\frac{1}{4}$ der ursprünglichen Energie. [HAS_03]

2.1.1.3. Schalldruck

Der Schalldruck ist das objektive Maß zur Beschreibung des durch Schallwellen verursachten Grad der Luftdruckveränderung. Schalldruck wird gemessen als das Maß an Kraft, die auf eine Fläche wirkt. Der Schalldruck wird als wirkende Kraft in Newton/m^2 oder als Druck auf die Fläche in Pascal angegeben. ($1 \text{ N/m}^2 = 1 \text{ Pa}$)

Der Schalldruck wird auch als Amplitude bezeichnet. Der Schalldruckbereich, begrenzt durch Hörschwelle und Schmerzgrenze, liegt zwischen $2 \cdot 10^{-5} \text{ Pa}$ und 20 Pa . Dies entspricht einer Spanne vom Faktor 1000000.

2.1.1.4. Schalldruckpegel

Zur verbesserten mathematischen Erfassung des Schalldrucks wurde die logarithmische Verhältnisgröße Dezibel (dB) eingeführt, der Schalldruckpegel. Um die Größen des Schalldruckpegels von anderen Größen abzugrenzen, die ebenfalls in dB angegeben werden, ist der Schalldruckpegel in dB SPL anzugeben. (SPL – Sound Pressure Level)

So entspricht ein Schalldruckpegel von 0 dB einem Schalldruck von $2 \cdot 10^{-5} \text{ Pa}$ und ein Pegel von 120 dB einem Schalldruck von 20 Pa . Eine Schallpegeldifferenz von 1 dB ist durch das menschliche Gehör gerade noch wahrnehmbar. Eine Schalldruckverdopplung wirkt sich als Erhöhung des Schalldruckpegels um 6 dB aus. Der logarithmische Zusammenhang wird berechnet unter Verwendung des Schalldrucks für die Hörschwelle.

$$\text{Schalldruckpegel} = 20 \log \left(\frac{p}{2 \cdot 10^{-5}} \right) \text{ dB SPL}$$

2.1.1.5. Lautstärkepegel

Als Lautstärke wird die Hörempfindung bezeichnet, die durch den Schalldruck hervorgerufen wird. Beschrieben wird die Lautstärke über den Lautstärkepegel. Der Lautstärkepegel wird in phon angegeben. Ein Sinuston mit einem konstanten Schalldruckpegel von 20 dB SPL, der den gesamten Frequenzbereich, also z.B. von 125 – 16000 Hz durchläuft, wird mit steigender Frequenz zunächst lauter und ab 4000 Hz wieder leiser.

Um diese Feststellung für verschiedene Schalldruckpegel genau zu erfassen, hat man die „Kurven gleicher Lautstärke“ ermittelt. (Siehe Abbildung 1) Sie geben in Abhängigkeit von der Frequenz den Schalldruckpegel an, der die jeweils gleichen Lautstärkeempfindungen hervorruft. Man ordnet jeder der Kurven einen bestimmten Lautstärkepegel zu, der in phon angegeben wird. Für 1000 Hz stimmen Schalldruckpegel in dB SPL und Lautstärkepegel zahlenmäßig überein.

Einer Verdopplung der empfundenen Lautstärke in dem praktisch wichtigen Lautstärkebereich über 30 phon entspricht eine Lautstärkepegel – Zunahme von 10 phon. [DIC_97]

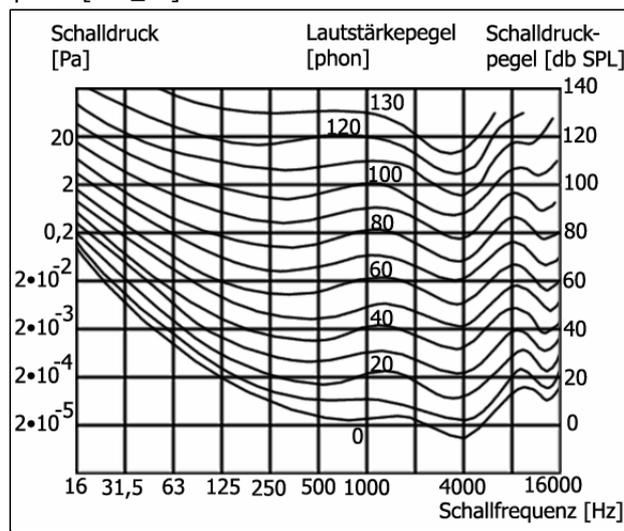


Abbildung 1 : Hörschwelle und Kurven gleicher Lautstärke

2.1.2. Schallwahrnehmung

Die bei einem Hörer eintreffende physikalisch beschreibbare Schallwelle wird vom individuellen Hörer unterschiedlich wahrgenommen. Mit dem Auftreffen der Schallwelle am Körper des Hörers erfährt die Schallwelle eine letzte Manipulation, bis sie zum Trommelfell vordringt und durch das mechanische System des Hörers wahrgenommen wird.

Neben den unterschiedlichen physikalisch beschreibbaren Schallwellen, die das Trommelfell erreichen, ist die Leistungsfähigkeit des Hörapparates wie auch der subjektive aus der Wahrnehmung gewonnene Eindruck von Hörer zu Hörer verschieden.

Die sich auf die Wahrnehmung auswirkenden Faktoren finden hier eine kurze Erläuterung.

2.1.2.1. Head Related Transfer Function

Mit dem Eintreffen der Schallwelle am Körper eines Hörers, erfährt die Schallwelle, abhängig von der Kopfform, dem Volumen des Torso sowie der Position der Schultern relativ zu den Ohren als auch die Beschaffenheit des Ohres selbst, eine letzte spektrale Manipulation des Schallspektrums. Diese Manipulation ist für jeden Hörer individuell als funktionaler Zusammenhang beschreibbar.

Als Head Related Transfer Function (HRTF) wird der funktionale Zusammenhang bezeichnet, der die individuelle Soundwahrnehmung eines Menschen beschreibt. Als spektraler Filter wirkt die HRTF auf einfallende Schallwellen, bevor sie das Trommelfell erreichen. Primär durch das Außenohr verursacht, kann die HRTF als frequenzabhängiger Filter bezeichnet werden, der im einfallenden Soundspektrum, für Amplitudenmanipulation und Zeitverzögerungen im μsec Bereich führt. Dabei führt vor allem die asymmetrische und komplexe Form des äußeren Ohres, abhängig vom Einfallswinkel der Schallwellen, zu einem einzigartigen Satz von Zeitverzögerungen, Resonanzen und Reflektionen, die als eine einzigartige HRTF für jede Soundquellen – Position zusammengefasst werden können. [BEG_94]

2.1.2.2. Wahrnehmung der Position von Soundquellen

Das durch einen Hörer erfahrene Hörereignis lässt den Hörer Rückschlüsse auf die Position der Soundquelle schließen, von der aus sich die Schallwelle ausgebreitet hat. Der Rückschluss basiert auf dem unterschiedlichen Zeitpunkt des Eintreffens der Schallwellen an den Ohren sowie der an den Ohren unterschiedlichen Intensität.

Die Beschreibung der Position einer Soundquelle erfolgt relativ zum Hörer. Ausgehend vom Referenzpunkt mit dem Ursprung im Zentrum des Kopfes, direkt zwischen den Ohren, ungefähr auf Augenhöhe, erfolgt die Angabe der Position mit Hilfe von kartesischen Koordinaten, oder unter Angabe von Azimuth und Elevation in Verbindung mit einer Entfernung. [KEN_WWW_1]

Azimuth beschreibt die Position einer Soundquelle in Grad auf der horizontalen Ebene. Azimuth wird in Grad im Bereich von 0 – 360 Grad entgegen des Uhrzeigersinns, oder aber im Bereich 0 – 180 Grad und dem Zusatz „links“, bzw. „rechts“ angegeben (z.B. links 75 Grad). Eine Soundquelle direkt voraus hat einen Azimuth Wert von 0 Grad.

Elevation beschreibt die Position einer Soundquelle in Grad auf der vertikalen Ebene (auch als Median Ebene bezeichnet). Elevation wird in Grad im Bereich 0 – 90 Grad mit dem Zusatz „oben“, bzw. „unten“ angegeben (z.B. oben 75 Grad). Eine Soundquelle direkt voraus hat einen Elevation Wert von 0 Grad.

Bei der Lokalisierung von Soundquellen werden lediglich die zuerst eintreffenden zu einer Soundquelle gehörenden Wellenfronten genutzt. Dieser als Precedence Effect bezeichnete Effekt ermöglicht dem Menschen auch in stark hallenden Umgebungen, eine Soundquelle in der Position zu orten. Nachfolgende Sound Information trägt kaum weiter zur Ortung bei. [PUL_01]

Precedence Effect bezeichnet den Effekt, dass die Ortung der Position von Schallquellen hauptsächlich auf Basis der zuerst am Trommelfell eintreffenden Schallwellen erfolgt.

Eine Lokalisierung wird dabei erst möglich durch die an den beiden Ohren des Hörers unterschiedliche Wahrnehmung derselben Schallwelle. Aufgrund der Geometrie des Kopfes unterscheiden sich die Wege, die eine Schallwelle von der Soundquelle zu jedem einzelnen Ohr zurückzulegen hat. Die unterschiedliche Länge der Wege zu den Ohren verursacht ein zeitlich versetztes Auftreffen des Schalls an den Ohren mit unterschiedlichen Intensitäten. Dieser Zeit- und Intensitätsunterschied wird als Interaural Time Difference und Interaural Intensity Difference bezeichnet. (Siehe Abbildung 2)

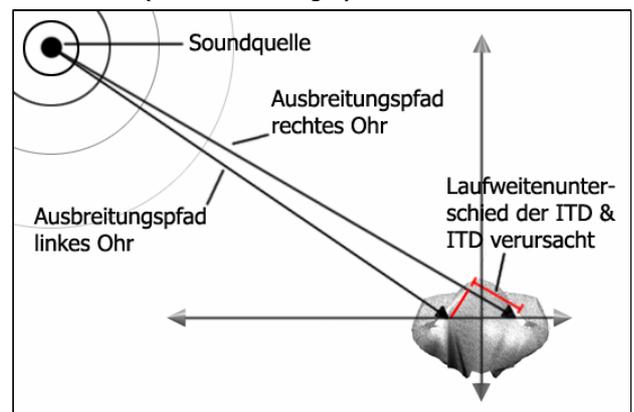


Abbildung 2 : ITD und IID

Interaural Time Difference (ITD) bezeichnet die zeitliche Verzögerung zwischen dem Auftreffen der ersten Schallwellenfront an einem Ohr und dem Auftreffen derselben Wellenfront am anderen Ohr.

Interaural Intensity Difference (IID) bezeichnet den Intensitätsunterschied in der Amplitude der an beiden Ohren auftreffenden Schallwellen.

ITD und IID sind beim menschlichen Gehör die primären Faktoren zur Lokalisierung von Schallquellen. [KEN_WWW_1]

2.1.3. Schallausbreitung

Ausgehend von einer im Raum positionierten Schallquelle breiten sich die Schallwellen im Raum aus. Abhängig von den Eigenschaften des Raumes – besser bezeichnet als Environment –, den im Raum befindlichen Geometrien und den Eigenschaften der Soundquelle kommt es während der Ausbreitung zur Verfälschung des initialen Soundspektrums.

2.1.3.1. Raumwahrnehmung

Abhängig vom Medium, in dem sich der Schall ausbreitet, erfährt dieser über die zurückgelegte Distanz

eine Abschwächung der Intensität. Die Intensitätsminderung ist abhängig von der Dichte des Mediums wie auch von der Distanz der Soundquelle. Neben der Intensitätsminderung erfährt die Schallwelle eine Manipulation durch die sich im Raum befindlichen Objekte.

Durch den Einfluß der sich im Environment befindlichen Objekte wird es neben den sich auf dem direkten Pfad von der Soundquelle zum Hörer ausbreitenden Schallwellen weiter auch Schallwellen geben, die nach einigen wenigen Reflektionen oder nach einer Vielzahl Reflektionen auf den Hörer treffen.

Diese unterschiedlich eintreffenden Schallereignisse werden als direkter Schall, frühe Reflektionen und späte Reflektionen oder auch Reverberation bezeichnet. [IAS 99] (Siehe Abbildung 3)

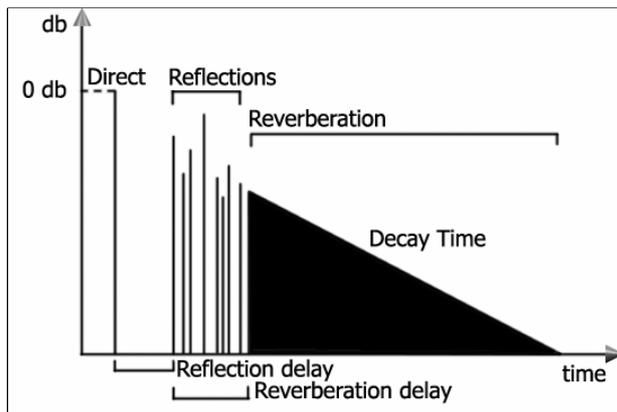


Abbildung 3 : Environment Reverberation [IAS_99]

Direkter Schall bezeichnet die Soundwellen, die auf direkten Weg von der Soundquelle ausgehend auf den Hörer treffen.

Frühe Reflektion bezeichnet die Schallwellen, die nach einer vom Environment abhängigen als Reflection Delay bezeichneten Verzögerung auf den Hörer treffen, nachdem sie an verschiedenen Objekten im Raum reflektiert wurden. Frühe Reflektionen können als einzelne Schallereignisse identifiziert werden. (Siehe Abbildung 3)

Reverberation / Späte Reflektion bezeichnet die Schallereignisse, die nach dem Reflection Delay nach einer weiteren als Reverb Delay bezeichneten Verzögerung beim Hörer eintreffen. Beim Eintreffen von Reverberation sind keine einzelnen Schallereignisse durch den Hörer zu identifizieren, sie verschmelzen zu einem Schalleindruck, der über eine vom Environment abhängige Zeitspanne, bezeichnet als Decay Time, anhält und über diese in der Intensität bis zu einer nicht mehr wahrzunehmenden Lautstärke abfällt.

2.1.3.2. Doppler Effekt

Einer der wichtigsten bei der Soundausbreitung sich bewegender Soundquellen ergebener Effekt ist als Doppler Effekt bekannt. Der Doppler Effekt (nach

Christian Doppler, 1842) ermöglicht dem Hörer auf Basis des Hörereignisses auf die relative Geschwindigkeit zwischen Soundquelle und ihm selbst zu schließen.

Bewegen sich Schallquelle und Hörer relativ zueinander, so registriert der Hörer eine Frequenzveränderung. Wir erleben diese Frequenzveränderung, wenn wir mit dem Wagen an einer läutenden Glocke vorbeifahren, oder ein Motorrad auf der Landstraße schnell an uns vorbeifährt.

Bei sich bewegenden Soundquellen drängen sich die Schallwellen bei einer Bewegung auf den Hörer zusammen, d.h. die Wellenlänge wird kleiner. Entfernt sich die Soundquelle, wächst die Wellenlänge.

Der Hörer nimmt bei einer entgegenkommenden Soundquelle also einen höheren Ton, bei einer sich entfernende Soundquelle einen tieferen Ton wahr.

Falls dem Hörer ein bekannter Sound als Referenz zur Verfügung steht, wird dieser über die Frequenzverschiebung abschätzen können, wie sich die Soundquelle relativ zu ihm bewegt. [KUJ_89]

2.2. Grenzen des menschlichen Hörens

Das menschliche Gehör stößt in verschiedenen Kategorien schnell an seine Grenzen. Diese Grenzen sollten dem Entwickler eines 3-D Auditory Displays bekannt sein, um ein leistungsfähiges System entwickeln zu können.

2.2.1. Genauigkeit bei der Ortung von Soundquellen

Die Ortung von Soundquellen erfolgt beim menschlichen Gehör mit unterschiedlicher Genauigkeit.

Wird die Größe Azimuth genau in Blickrichtung mit 2 Grad Genauigkeit aufgelöst, nimmt die Auflösung hier zu den Seiten mit 10 Grad Genauigkeit ab und verbessert sich im rückwärtigen Bereich wieder auf eine 6 Grad genaue Auflösung.

Die Elevation unterliegt einer im Vergleich zu Azimuth größeren Ungenauigkeit. Können Soundquellen direkt voraus noch mit 9 Grad Genauigkeit geortet werden, verschlechtert sich die Genauigkeit auf bis zu 22 Grad direkt über oder unter dem Hörer. [KEN_WWW_1]

Die größte Ungenauigkeit bei der Ortung von Soundquellen ist mit der Abschätzung der Entfernung einer Soundquelle verbunden. Die Fähigkeit, Entfernungen von Soundquellen einzuschätzen, variiert jedoch individuell. Nur für den Fall, dass es sich um eine bekannte Soundquelle handelt, ist es dem Hörer möglich, aufgrund der Intensität bzw. des Intensitätsabfalls die Entfernung zu bestimmen. Bei Soundquellen unbekanntem Ursprungs fehlt dem Hörer zur Entfernungsbestimmung jegliche Referenz. Aus diesem Grund wird, sobald vorhanden, der zur Soundquelle gehörende visuelle Eindruck zur Entfernungsbestimmung verwendet.

Stehen dem Menschen sowohl visuelle als auch akustische Eindrücke zur Positionierung eines Objektes zur Verfügung, zeigt sich eine Dominanz der visuellen

Ortungsfähigkeiten des Menschen. Ein Sound, der von einem visuell erfassten Objekt zu kommen scheint, kann um bis zu 11 Grad von der visuell determinierten Position abweichen, bevor der Sound durch die menschliche Wahrnehmung als separates Objekt identifiziert wird. Sollen beide Eindrücke als eine Einheit wahrgenommen werden, ist nicht nur die lokale Nähe von Bedeutung, auch die zeitliche Nähe beim Auftreten des Eindrucks ist zu berücksichtigen. Liegen weniger als 40 ms zwischen dem Auftreten des visuellen Reizes und der akustischen Wahrnehmung, werden beide Eindrücke als zusammenhängende Einheit wahrgenommen. Erst bei vergrößertem Zeitversatz von über 40 ms werden beide Eindrücke als separate Wahrnehmungen wahrgenommen. [BEG_99]

2.2.2. Frequenzabhängige Ortung von Soundquellen

Das menschliche Gehör funktioniert optimal in einem Frequenzbereich von 500 Hz bis zu 6000 Hz. Wohingegen steigende Frequenzen stetig schlechter wahrgenommen werden, beginnt der Hörer mit tiefer werdenden Frequenzen diese schlechter zu hören, gleichzeitig beginnt er bei entsprechender Intensität, diese über das Bauchgefühl als Schwingung des eigenen Körpers wahrzunehmen. Ab einer Frequenz von 20 Hz sind Schwingungen für den Menschen ausschließlich fühl- und nicht mehr hörbar.

Die Ortung niederfrequenter Soundanteile ergibt sich ebenfalls zu der von hohen Frequenzen verschieden. Soundquellen mit Informationsanteil einer individuell verschiedenen Frequenz im Infraschallbereich (unter 1000 Hz) können nicht geortet werden und erscheinen als im Referenzpunkt des Hörer Koordinaten Systems positionierte Soundquellen.

2.3. Spatial Manipulation

Von besonderem Interesse bei der Betrachtung der psychoakustischen Wahrnehmung sind die Aspekte, die bei der computergestützten Verarbeitung durch ein 3-D Sound System und die Wiedergabe durch ein 3-D Auditory Display beeinflussbar sind.

3-D Sound System beschreibt ein System, das zum Einen die Manipulation der 3-D Sound Daten implementiert und zum Anderen über die Nutzung eines 3-D Auditory Display eine entsprechende Wiedergabe ermöglicht.

3-D Auditory Display bezeichnet ein System, das in der Lage ist, 3-D Sound in der Art und Weise wiederzugeben, wie es durch die psychoakustischen Attribute einer Soundquelle bestimmt ist. 3-D Auditory Displays dienen zumeist als Feedbackmöglichkeit in interaktiven Systemen.

Der Betreiber eines 3-D Sound Systems wird vorhersagen wollen, welches Hörereignis beim Hörer durch das 3-D Auditory Display erzeugt wird. Diese Form

des Feedbacks wird als Spatial Manipulation bezeichnet und bedingt die Berücksichtigung der wesentlichen psychoakustischen Attribute bei der Verarbeitung.

Spatial Manipulation bezeichnet die Soundverarbeitung in der Art und Weise, dass der Hörer ein Klangbild wahrnimmt, wie es durch die, in Verarbeitung verwendeten und durch den Betreiber des Virtual Acoustic Displays bestimmten, Attribute charakterisiert ist. [BEG_94]

2.4. Techniken für 3-D Auditory Displays

Zur Wiedergabe von dreidimensionalen Klangbildern werden 3-D Auditory Display Systeme verwendet. Diese sind in der Lage, dem Hörer den gewünschten akustischen Raumeindruck zu vermitteln. 3-D Auditory Display Systeme unterscheiden sich vor allem durch die von ihnen genutzte Hardware, das Setup dieser und die zur Wiedergabe verwendete Methodik. Demnach kann unterschieden werden zwischen 3-D Auditory Displays, die ein Lautsprecher Setup verwenden, und solchen, die auf die Verwendung von Kopfhörern zurückgreifen, um ein dreidimensionales Klangbild zu erzeugen. Die Leistungsfähigkeit von Lautsprecher Setups wird durch die Anzahl der verwendeten Lautsprecher maßgeblich bestimmt. Nur eine geringe Anzahl Lautsprecher Setups ist überhaupt in der Lage, einen Raumeindruck zu erzeugen. Bei der Verwendung von Lautsprechern hängt es von der gewählten Methode ab, ob ein Raumeindruck bei der Wiedergabe erzeugt werden kann. Bei den vorgestellten Methoden kann zwischen binauralen Techniken und stereophonen Techniken zur räumlichen Positionierung einer Schallquelle unterschieden werden. [TOD_02_2]

2.4.1. 3-D Auditory Displays auf Basis stereophoner Techniken

Primäre Methode bei der Verwendung von Lautsprechern zur räumlichen Positionierung sind stereophone Techniken. Stereophone Techniken versuchen, ein Schallereignis zu simulieren. Das bei der Wiedergabe entstehende Schallereignis soll dem gleichen, als wäre an einer gewünschten Stelle tatsächlich eine Soundquelle positioniert.

Über die verwendete Lautsprecher – Installation lassen sich die ITD und IID, die bei einer real existierenden Soundquelle entstehen würden, nachbilden.

Die Basis stereophon arbeitender Systeme bildet ein im Mono – Format gemixtes Soundsamples, das mit angepassten Intensitäten über die angeschlossenen Lautsprecher wiedergegeben wird und so die Richtungscharakteristik vermittelt. [JOT_96]

Einige der hier vorgestellten Setups sind in der Form auch für Kopfhörernutzung geeignet. Sie sind für Lautsprecher Setups erläutert und ihre Besonderheiten bei einer möglichen Kopfhörernutzung kurz beschrieben.

Auch wenn nur wenige Lautsprecher basierte Setups in der Lage sind, ein dreidimensionales Klangbild zu erzeugen, sollen hier doch alle vorgestellt werden, um ihre individuellen Stärken und Schwächen kurz erläutern zu können.

Lautsprecher Setups zeichnen sich vor allem dadurch aus, dass für jedes Setup eine beste Position für den Hörer, der so genannten Hotspot, existiert, in dem sich der Hörer befinden sollte, um den zu erzielenden Effekt erfahren zu können. Um den vollen Effektgenuss garantieren zu können, haben Betreiber eines 3-D Auditory Displays auf Lautsprecherbasis lediglich die Möglichkeit, die Bewegungsfreiheit des Hörers auf den Hotspot des Setups einzuschränken, oder aber Aufwand zu betreiben, um die verwendete Methodik der sich ändernden Hörerposition anzupassen. Diese Tatsache schlägt für alle Lautsprecher Setups als großer Nachteil zu Buche.

Vorteilhaft bei allen Lautsprecher Setups wirkt sich aus, dass diese den Hörer nicht durch das Tragen von Kopfgeschirr in Form eines Kopfhörers behindern, oder aber akustisch von seiner realen Umwelt abkapseln. Dem Hörer bleiben so alle Möglichkeiten der realen Interaktion mit seiner Umgebung, während er die Effekte des 3-D Auditory Displays erfährt.

2.4.1.1. Einkanal Lautsprecher Setup

Bei der Nutzung eines einzelnen Lautsprechers zur Soundwiedergabe, scheint der Sound seinen Ursprung immer in der Richtung zu haben, in der auch der Lautsprecher liegt. Ein einkanal Lautsprecher Setup ist somit in der Lage, ein eindimensionales Klangbild wiederzugeben. Nahezu keines der räumlichen Attribute der wiederzugebenden Soundquelle kann hier in der Wiedergabe Berücksichtigung finden. Lediglich über die Regelung der Intensität bei der Wiedergabe der Soundinformationen einer Soundquelle kann eine Simulation der Distanz erfolgen. Die eindimensionale Soundwiedergabe kann somit zu einer zweidimensionalen Soundwiedergabe ausgebaut werden. Eine Positionierung vor oder hinter dem Lautsprecher wird mit der Entfernungssimulation möglich.

Die Richtung, aus der der Sound zu kommen scheint, bleibt dabei auf die Richtung, in der sich der Lautsprecher befindet, begrenzt.

Für den Hörer gibt es bei diesem Setup keinen eigentlichen Hotspot. Der Hörer kann sich auf einem Kreis um den Lautsprecher herum positionieren. Ein einmal festgelegter Hotspot sollte jedoch zur Anpassung der Intensität bei der Entfernungssimulation als Referenzpunkt gelten.

Bewegt sich der Hörer während der laufenden Wiedergabe weiter an den Lautsprecher heran, ist dem mit einer Intensitätsminderung entgegenzuwirken, damit sich für den Hörer dann stets ein einheitliches Intensitätsbild ergibt. Das Entgegengesetzte gilt für den umgekehrten Fall.

Ein einkanal Setup lässt sich auch mit Hilfe eines Kopfhörers betreiben. Dabei wird über beide Kanäle des Kopfhörers das gleiche Klangbild wiedergegeben.

2.4.1.2. Zweikanal Lautsprecher Setup

Bei der Verwendung von zwei Lautsprechern, ergibt sich ein zweidimensionales Setup, das es ermöglicht, ein Klangbild zu erzeugen, bei dem ein Sound aus einer Richtung zu kommen scheint, die auf einer Achse zwischen den beiden verwendeten Lautsprechern liegt. Unter Hinzunahme der Simulation der Distanz über die Regelung der Intensität lässt sich die Position einer Soundquelle bei der Wiedergabe simulieren. Dies jedoch nur in der horizontalen Ebene und nur für Positionen die sich in der durch die zwischen dem Hörer und den Lautsprechern verlaufenden Achsen begrenzten Fläche befinden. (Siehe Abbildung 4)

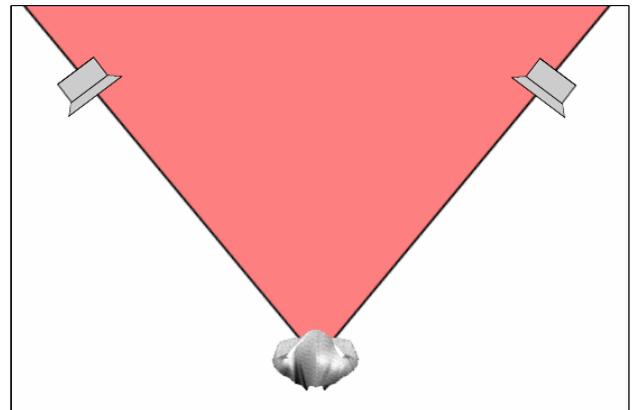


Abbildung 4 : Zweikanal Lautsprecher Setup

Die Erzeugung des Klangbildes erfolgt durch Anpassung der Wiedergabeintensitäten beider Lautsprecher. Die Intensität eines Kanals ergibt sich aus dem Verhältnis des Winkels zwischen der Blickrichtung des Hörers und der zu simulierenden Soundposition zum Winkel zwischen der Blickrichtung des Hörers und der Lautsprecherposition.

Die Gesamtintensität des Systems von 100% wird auf beide Lautsprecher verteilt, um die Positionierung der Soundquelle zwischen diesen zu realisieren. Zur Simulation der Entfernung wird die Gesamtintensität des Systems von 100% ausgehend reduziert.

Um den erzeugten Stereo Effekt voll erfahren zu können, ist der Hörer gezwungen, sich in einer kleinen Fläche zwischen den Lautsprechern, Stereophone Fläche genannt, aufzuhalten. [DIC_97] (Siehe Abbildung 5)

Wie auch bei der einkanal Wiedergabe, sollte sich der Hörer in einem initialen Hotspot positionieren und diesen nicht verlassen. Tut er dies doch, muss dem durch eine gleichmäßige Intensitätsanpassung begegnet werden, falls es sich lediglich um eine Bewegung auf den Mittelpunkt der Achse zwischen den beiden Lautsprechern zu handelt. Bewegt sich der Hörer seitlich zu einem der beiden Lautsprecher, ist dies durch eine unsymmetrische Intensitätsanpassung auszugleichen.

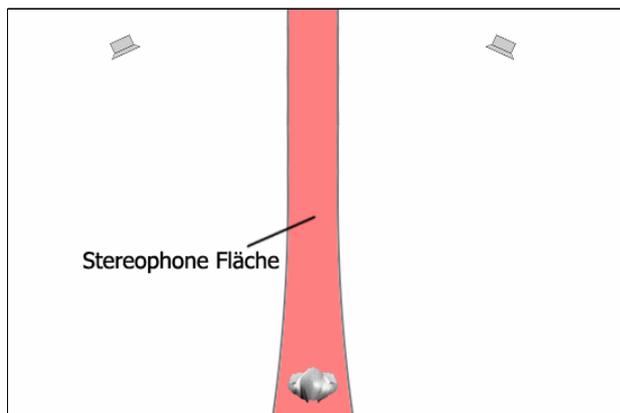


Abbildung 5 : Stereophone Fläche

Nach der Bewegung des Hörers wird dem dann im Vergleich zur initialen Position jetzt näheren Lautsprecher eine geringere Intensität bei der Wiedergabe ermöglicht. Der jetzt weiter entfernte Lautsprecher erhält einen größeren Intensitätsspielraum.

Auch diese Technik ist unter Nutzung von Kopfhörern realisierbar. Die unveränderliche Hörerposition entspricht dann einer Position unmittelbar zwischen den zwei Lautsprechern.

Wird statt der stereophonen Technik eine Technik verwendet, die den Einfluß der HRTF nutzt, können anstatt der mit dieser Technik möglichen zweidimensionalen Klangbilder dreidimensionale Klangbilder erzeugt werden. Diese als Binaural Spatialization bezeichnete Technik wird im späteren Verlauf näher erläutert. (Siehe Abschnitt 2.4.2.)

2.4.1.3. Vierkanal Lautsprecher Setup

Die Funktionsweise eines vierkanal Lautsprecher Setups orientiert sich an der Funktionsweise eines zweikanal Setups. Anstatt der verwendeten zwei Lautsprecher kommen vier Lautsprecher zum Einsatz, die den Hörer möglichst in Form eines Quadrates umgeben. Es ist mit dieser Technik möglich, ein Klangbild zu erzeugen, das einen Eindruck zu vermitteln vermag, dass ein Sound aus dem rückwärtigen Bereich des Hörers zu kommen scheint.

Auch bei diesem Setup sollte der Hörer in einem initialen Hotspot positioniert bleiben, der möglichst im Zentrum des Setups liegen sollte. Anpassungen an mögliche Hörerbewegungen erfolgen analog zur zweikanal Behandlung.

2.4.1.4. Mehrkanal Lautsprecher Setup

Als mehrkanal Lautsprecher Setups werden solche Setups bezeichnet, die mehr als vier Lautsprecher zur Erzeugung eines Klangbildes verwenden. Mehrkanal Lautsprecher Setups sind vor allem bekannt durch Surround Sound Systeme, wie sie im Unterhaltungsbereich in z.B. Kino, Video, DVD oder inzwischen auch in der Konsolen- und Computerspiele

Industrie Verwendung finden. Diese System, wie Dolby Surround®, Dolby ProLogic®, Dolby Surround Digital®, Dolby Digital Surround EX® oder Sony's® SDDS beschreiben Vorschriften für Lautsprecher Setups, damit Hörer in definierten Positionen, z.B. die Loge im Kino, ein optimales Klangbild erfahren können. Alle Systeme dieser Art arbeiten dabei im Prinzip auf stereophonen Techniken. Diese im Entertainment Bereich genutzten Systeme arbeiten alle auf einem „N.1“ Lautsprecher Setup, bei dem N Lautsprecher genutzt werden, um eine räumliche Positionierung des Sounds unter Nutzung der stereophonen Techniken zu bewirken. Ein zusätzlicher Lautsprecher („1“) wird dabei lediglich zur Wiedergabe nicht ortbarer, niederfrequenter Soundeindrücke genutzt. Dieser kann somit quasi frei im Raum positioniert werden. [FAU_00]

Ein 6.1 Setup arbeitet dabei zum Beispiel mit 3 Lautsprechern im vorderen Bereich, von denen einer zentral, die anderen beiden seitlich angebracht sind sowie drei in gleicher Weise rückwärtig angebrachten Lautsprechern. Der frequenzbegrenzte Lautsprecher für die niedrigen Frequenzen wird zumeist zentral im vorderen Bereich positioniert.

Alle bis hier vorgestellten Setups beschreiben Installationen, bei denen sich die verwendeten Lautsprecher in einer gemeinsamen Ebene befinden. Sie alle sind damit außer Stande, einen wirklich räumlichen Eindruck zu vermitteln. Soundquellen, die mit Hilfe dieser Systeme positioniert werden, werden alle ihren Ursprung in einem Punkt haben, der sich ebenfalls in der Ebene der Lautsprecher befindet.

Wirklich dreidimensionale Raumeindrücke werden erst durch Setups erreicht, bei denen die Lautsprecher nicht in einer gemeinsamen Ebene positioniert, denn im Raum verteilt sind.

2.4.1.5. Kubisches Achtkanal Lautsprecher Setup

Eine Möglichkeit eines räumlichen Lautsprecher Setups besteht in der Verwendung von acht Lautsprechern, die in Form eines Würfels positioniert sind. Diese Form des Setups hat sich als besonders praktisch bei Virtual Reality Systemen gezeigt, die ein CAVE® System integrieren. Für die meisten Fälle bestehen CAVE® Systeme aus quadratischen Projektionsflächen, die die Seiten eines Würfels bilden. (Siehe Abbildung 6)

Auch für den Fall, dass nicht alle sechs Seiten der CAVE® Installation vorhanden sind, wird sich der Anwendungsnutzer / Hörer im Zentrum der CAVE®, umgeben von den acht Lautsprechern, befinden. Über die Verwendung stereophoner Techniken lassen sich mit diesem Lautsprecher Setup Klangbilder mit für den Hörer dreidimensionalem Eindruck erzeugen.

Der Hörerbewegung innerhalb der CAVE® ist mit Intensitätsanpassungen über alle acht Lautsprecher zu begegnen. Die Intensitätsanpassung zur

Entfernungssimulation erfolgt analog zu den oben besprochenen Setups.

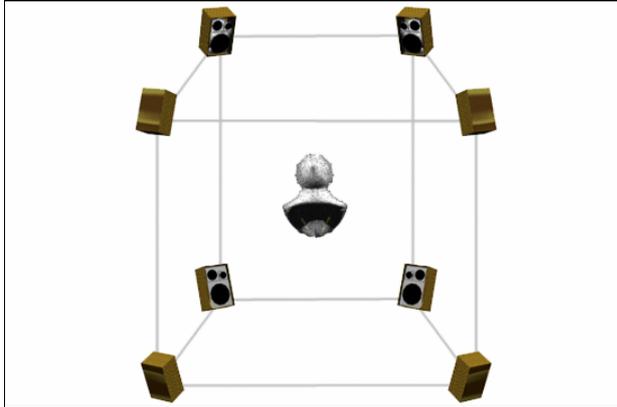


Abbildung 6 : Kubisches Achtkanal Lautsprecher Setup

2.4.2. 3-D Auditory Displays auf Basis binauraler Techniken

Binaurale Techniken versuchen nicht das Schallereignis zum simulieren. Bei binauralen Techniken soll vielmehr das Hörereignis simuliert werden. Zur Erreichung dieses Ziels werden für jedes Ohr und die entsprechende Position der Soundquelle die entsprechende HRTF zur Filterung des Soundspektrums der Soundquelle genutzt. Da HRTF jedoch stets individuelle Funktionen sind, die mit jeder Person variieren, ist es notwendig, auf Standard Funktionen zurückzugreifen, um nicht für jeden Hörer die HRTF neu bestimmen zu müssen. Zu diesem Zweck können gemittelte HRTF verschiedener Personen zur Verwendung kommen, oder aber auch solche, die anhand eines Kunstkopfes entstanden sind. Bei der Verwendung gemittelter oder synthetischer HRTF kommt es zu Abstrichen in der Positionierungsgenauigkeit, doch können diese in Anbetracht der geringen Genauigkeit des menschlichen Gehörs vernachlässigt werden. Allgemeine Akzeptanz haben synthetische Werte auf Basis eines KEMAR® Kunstkopfes gefunden. Die notwendigen HRTF sowie zusätzliche Informationen, die Kopfhörer oder Lautsprecherimplementierung betreffend, sind frei verfügbar. [GAR_WWW_1]

2.4.2.1. Kopfhörer Nutzung zur Umsetzung binauraler Techniken

Die präziseste Methode zur Umsetzung binauraler Techniken ist die, bei der Kopfhörer zur Wiedergabe Verwendung finden. Abgeschirmt von äußeren Einflüssen, kann die Wiedergabe der mit der HRTF gefilterten Soundinformation unmittelbar vor dem Trommelfell erfolgen. Lediglich der in ihrer Auswirkung kaum wahrzunehmender Einfluss des Innenohrs nimmt erneut Einfluß auf die Soundwahrnehmung. So ist es möglich, dem Hörer ein exaktes Klangbild zu vermitteln.

Problematisch bei der Verwendung von Kopfhörern in Kombination mit HRTF ist die Orientierung des Hörers, die bei der Verarbeitung initial als direkt geradeaus

angenommen wird. Dreht der Hörer während der Wiedergabe seinen Kopf, so ist dies in der Verarbeitung zu berücksichtigen. Die Kopfdrehung geht aufgrund des mit der Kopfdrehung gedrehten Hörer – Koordinatensystems mit einer Änderung der relativen Position der Soundquelle zum Hörer einher. Bei Änderung der Orientierung hat die Filterung dann mit der HRTF der neuen relativen Position zu erfolgen. Um eine volle Funktionsfähigkeit zu gewährleisten, ist somit der Hörer einem Tracking zu unterziehen, damit die Soundwiedergabe angepasst an etwaige Änderungen der Hörerorientierung erfolgen kann.

Zur Simulation der Entfernung kann auch hier eine Anpassung der Intensität bei der Wiedergabe erfolgen.

2.4.2.2. Lautsprecher Nutzung zur Umsetzung binauraler Techniken

Die Methode, wie sie bei der Verwendung von Kopfhörern zur Anwendung kommt, wird prinzipiell in der Form auch bei der Nutzung von zwei Lautsprechern verwendet. Vorausgesetzt wird dabei ein Hörer, der sich im Hotspot des Stereo Lautsprecher Setups befindet. Gefiltert mit der zu simulierenden Soundquellenposition gehörenden HRTF wird die resultierende Soundinformation für das linke Ohr über den linken Lautsprecher, die Soundinformation für das rechte über den rechten Lautsprecher wiedergegeben.

Da jedoch die wiedergegebene Soundinformation nicht direkt am Trommelfell des Hörers angebracht werden kann, wird die wiedergegebene Soundinformation während der Ausbreitung im Raum durch das Environment beeinflusst, in dem die Wiedergabe erfolgt sowie durch eine weitere HRTF gefiltert, die sich ergibt, wenn der Hörer den wiedergegebenen Sound wahrnimmt. Da es nicht vorhersagbar ist, in welchem Environment die Wiedergabe erfolgt, lässt sich der Fehler durch die Beeinflussung des Environment nicht wieder korrigieren. Die erneute Filterung durch die HRTF des Hörers jedoch lässt sich korrigieren, sofern bekannt ist, in welcher Position sich die Lautsprecher befinden. Mit einer inversen Filterung der HRTF für die entsprechende Lautsprecherposition wird der Einfluss der erneuten HRTF Filterung ausgeglichen.

Als weiterer großer Nachteil bei dieser Art der Wiedergabe wird der Hörer Soundinformation, die über den linken Lautsprecher für das linke Ohr wiedergegeben wird, auch mit dem rechten Ohr wahrnehmen. Durch die gleichzeitige Wiedergabe des inversen Spektrums der für das rechte Ohr aufbereiteten Soundinformation über den linken Lautsprecher und ebenso im umgekehrten Fall, lässt sich dieser Eindruck ebenfalls beheben. Dieses als Crosstalk Cancellation bezeichnete Verfahren setzt jedoch die genaue Kenntnis der relativen Position der Lautsprecher vom Hörer voraus, um die passende Intensität des Crosstalk Cancellation Anteils bestimmen zu können. [GAR_98] Großer Aufwand wird also auch hier für das Tracking des Users und die Übernahme der geänderten Positions- und Orientierungsparameter betrieben werden müssen.

3. Anforderung an interaktive Echtzeit 3-D Sound Systeme

Grundlage eines erfolgreichen Designs und einer erfolgreichen Implementierung eines Echtzeit 3-D Sound Systems sollte ein Anforderungskatalog sein. Der Anforderungskatalog sollte zum Einen die minimalen Voraussetzungen definieren, die zur interaktiven Erzeugung von 3-D Soundeindrücken zwingend erforderlich sind. Zum Anderen sollte der Katalog Anforderungen definieren, die für eine realitätsgetreue interaktive 3-D Sound – Verarbeitung von Nöten sind.

Der Anforderungskatalog kann so nicht nur als Grundlage für eine Neuimplementierung dienen, sondern weiter helfen, bereits existierende Echtzeit 3-D Sound Implementierungen zu bewerten.

Unter verschiedenen Gesichtspunkten betrachtet ergeben sich Teilaspekte des gesamten Anforderungskataloges. Der Datenaspekt definiert Mindestanforderungen an das im System zu implementierende Datenmodell, während der Verarbeitungsaspekt zu berücksichtigende Verarbeitungsschritte und deren empfohlene Umsetzung festhält. Unter dem Aspekt der Wiedergabe werden mögliche Wiedergabetechniken betrachtet.

Als Basis des Anforderungskataloges kann ein Katalog an Mindestanforderungen dienen, der bereits im Vorfeld als „Interactive 3-D Audio Minimum Requirements (Level 2)“ (I3DL2) entwickelt wurde. [IAS_99] Dieser wird hier kurz vorgestellt, bevor auf erweiterte Daten- und Verarbeitungsaspekte eingegangen und so ein Anforderungskatalog neu erstellt wird.

3.1. Anforderungen gemäß I3DL2

Als Expertenkommission, bestehend aus Vertretern der Audio Hardware- und Softwareentwicklung, hat sich die „3-D Working Group“ (3DWG) innerhalb der „Interactive Audio Special Interest Group“ (IASIG) gefunden mit dem Ziel, Mindestanforderungen zu definieren, an denen interaktive 3-D Sound Anwendungen gemessen werden können und die Grundlage neuer Implementierungen sein können. [IAS_99] Die Minimalanforderungen seien hier vorerst vorgestellt und im Anschluss diskutiert.

3.1.1. I3DL2 Minimalanforderungen

Die I3DL2 Mindestanforderungen definieren für Soundquellen, welche Attribute zu berücksichtigen sind und welche Effekte für die Soundquellen unter welchen definierten Randbedingungen zu berechnen sind. [IAS_99]

Als Randbedingung ist eine Wiedergabe bei einer Frequenz von mindestens 22050 Hz bei einer 16 Bit Quantisierung zu garantieren. Bei der Wiedergabe dürfen keine akustischen Artefakte zu hören sein, noch darf die Latenz bei der Berücksichtigung sich ändernder Parameter zu groß sein. Es sind mindestens 16, besser 32 virtuelle Soundquellen gleichzeitig zu verarbeiten.

Für die virtuellen Soundobjekte sowie einen virtuellen Hörer ist die Position innerhalb eines Kartesischen Koordinatensystems im Datenmodell wie auch später in der Verarbeitung zu berücksichtigen. Die Orientierung des virtuellen Hörers ist ebenso zu implementieren wie auch die Geschwindigkeit. Über die für die Soundquellen zu implementierende Geschwindigkeit der Soundquelle lässt sich die relative Geschwindigkeit zwischen Soundquelle und Hörer bestimmen. In Verbindung mit der zu berücksichtigenden Orientierung der Soundquelle ist ein Ausbreitungsmodell zu definieren, das die Richtungscharakteristik der Soundquelle beschreibt.

Die Verarbeitung von Effekten verlangt nach der Implementierung eines Entfernungsmodells in Form von Intensitätsanpassung und der Implementierung von Reverberation auf Basis der Entfernung zwischen Hörer und Soundquelle. Weiter ist in der Verarbeitung ein Reverberation Modell zu implementieren mit frühen und späten Reflektionen, wie es im Abschnitt 2.1.3.1. beschrieben ist. Die Umsetzung dieses Modells beinhaltet die Berechnung von Raumeffekten, die sich aus der Geometrie des Environments ergeben, sowie eines Tiefpassfilters, der zum Einen auf den direkten Schall und zum Anderen auf den indirekten, reflektierten Schall wirkt, und somit die Dämpfung durch das Ausbreitungsmedium simuliert.

Um dem Hörer weiter die Möglichkeit zu geben, aus dem erfahrenen Klangbild auf die räumliche Beziehung zwischen virtuellem Hörer und virtuellen Soundquellen zu schließen, ist die Berechnung und Einbindung des Doppler Effektes in die Soundverarbeitung notwendig.

3.1.2. Kritik an den I3DL2 Minimalanforderungen

Die durch die 3DWG definierten Minimalanforderungen beschreiben zum Teil sehr detailliert, welche Aspekte zu berücksichtigen sind, dann jedoch auch wieder andere Aspekte lückenhaft oder unpräzise. Gerade speziell in Bezug auf die Integration in Virtual Reality Systeme werden viele Aspekte nicht berücksichtigt.

Ein 3-D Sound System muß in der Lage sein, ein 3-D Klangbild zu erzeugen, das die Positionscharakteristik vermittelt, nicht jedoch den Anspruch an Realitätsnähe erfüllt. Solch ein System wäre auch bereits als interaktives 3-D Sound System zu bezeichnen, wengleich es auf einer abstrakteren Ebene arbeitet und so nach den I3DL2 Minimalanforderungen nicht als interaktives 3-D Sound System gilt. In der Praxis reicht oft ein geringerer Detaillierungsgrad aus, um die notwendigen Informationen zu vermitteln, oder die Informationsvermittlung zu fördern. Wie in der Grafik ist Realitätsnähe nicht immer erwünscht, so dass es sich bei den durch die 3DWG definierten Anforderungen nicht um Minimalanforderungen handeln kann, sondern um Anforderungen, die für 3-D Sound Systeme gelten, die bereits einen gewissen Grad an Realitätsnähe implementieren.

Die Aufteilung der Anforderungen in verschiedene Abstraktionsstufen käme einer qualifizierten Einordnung und Beurteilung zugute.

Zusammen mit einer Ergänzung und präzisierten Formulierung von Parametern kann so ein Katalog entstehen, der eine vollständige Beurteilung zulässt. Dies erfordert eine weitere kritische Betrachtung der bisherigen Parameter.

Bei der Definition der Randbedingungen wird für die Wiedergabe und Verarbeitung bei einer Sampling Rate von 22050 Hz und einer Quantisierung von 16 Bit bestanden, doch wird nur sehr schwammig formuliert, dass es zu keinen Artefakte und nur geringe Latenz bei der Verarbeitung kommen darf.

Artefakte in der Verarbeitung, die durch zu lang laufende Verarbeitungsprozesse, ungünstig organisierte Speicherverwaltung oder Wiedergabemethoden verursacht werden, liegen im Einflussbereich des Entwicklers eines 3-D Sound Systems und müssen vermieden werden. Artefakte jedoch, die rein aus der Rechnerhardware resultieren, wie zum Beispiel ein Knacken beim Anlaufen der Festplatte, liegen außerhalb des Einflussbereiches. Hier muss eine genaue Differenzierung erfolgen. Die Ursache der Artefakte muss eine Rolle spielen.

Eine „geringe“ Latenz bei der Verarbeitung kann bei den Randbedingungen als Wert keine Gültigkeit haben. Die zulässige Latenz lässt sich aufgrund des maximalen zeitlichen Versatzes zwischen zusammengehörenden visuellen und akustischen Eindrücken (Siehe Abschnitt 2.2.1.) als feststehende Größe definieren. Zwischen der grafischen und der akustischen Anpassung an geänderte Verhältnisse darf maximal eine Zeitspanne von 40 ms verstreichen.

Betrachtet man allein die Reaktion auf aufkommenden User Input, so ist die Reaktionszeit in diesem Fall dann doch als möglichst gering festzuhalten. Die Reaktion auf User Input kann jedoch den Verarbeitungsprozessen des Virtual Reality Systems zugesprochen werden, so dass die nach User Input entstehende Latenz nicht durch das 3-D Sound System beeinflussbar ist. Der Einfluss des 3-D Sound Systems besteht einzig in einer möglichst schnellen und für alle beteiligten Elemente synchronen Übernahme geänderter Verarbeitungsparameter. Synchroner Übernahme meint hier, dass in einem Aktualisierungsschritt alle zu berechnenden Soundquellen auf den gleichen Parametern arbeiten und nicht ein Teil, noch unter Berücksichtigung der alten Parameter, berechnet wird, wohingegen ein anderer schon mit den neuen arbeitet.

Die Position und Orientierung der virtuellen Soundquellen wie auch die des virtuellen Hörers sollen gemäß den I3DL2 Minimalanforderungen in der Verarbeitung Berücksichtigung finden.

Die Arbeitsweise verschiedener 3-D Auditory Displays, und hier gerade derjenigen, deren Einsatz speziell im Bereich der Virtual Reality Systeme als sehr Erfolg versprechend gilt, verlangt neben der Berücksichtigung der rein virtuellen Attribute des virtuellen Environments, des virtuellen Hörers und der virtuellen Soundquellen, nach der Berücksichtigung der physikalischen

Verhältnisse, die bei der Wiedergabe beim Anwendungsnutzer herrschen.

In den I3DL2 Minimalanforderungen werden zu den zur Wiedergabe über ein Auditory Display notwendigen Parameter keine Aussagen getroffen. Die definierten Anforderungen reichen hier nicht weit genug, um eine Bewertung der Wiedergabe und Wiedergabetechniken zu durchzuführen.

Wie bereits im Abschnitt 2.4.1. erläutert, sind die Position der Lautsprecher relativ zum Anwendungsnutzer, also dem physikalischen Hörer, sowie seine Orientierung bei der Berechnung des finalen Klangbildes erforderlich. Für eine erfolgreiche Integration des 3-D Sound Systems in Virtual Reality Systeme und die Bewertung solcher Systeme müssen diese Aspekte in den Anforderungskatalog aufgenommen werden.

Neben den aufzunehmenden Parametern zur Beschreibung des physikalischen Environments sind weiter Datenkonstrukte notwendig, die das virtuelle Environment beschreiben können. Aspekte, wie die Beschaffenheit des virtuellen Raumes, die Geometrien des Raumes und das Verhalten des Ausbreitungsmediums innerhalb dieses virtuellen Raumes, sollten hier erfasst werden.

Auch dieser Aspekt ist in den I3DL2 Minimalanforderungen nicht erwähnt.

3.2. Entwurf eines Anforderungskatalogs für interaktive 3-D Sound Systeme

Eine Identifizierung der von einem 3-D Sound System zu berücksichtigenden Parameter und die Gruppierung unter gemeinsamen Aspekten ermöglicht eine Bewertung von Systemen auf verschiedenen Ebenen. Inspiriert durch die I3DL2 Minimalanforderungen können Anforderungen für die Datenaspekte, die Verarbeitungsaspekte sowie die Aspekte der Wiedergabe definiert werden. Eine detaillierte Erläuterung der einzelnen Aspekte erfolgt in den nachfolgenden Abschnitten.

Der Anforderungskatalog selbst kann dann auf Basis der dargelegten Parameter in tabellarischer Form präsentiert werden.

3.2.1. Datenaspekte des Anforderungskataloges

Der Datenaspekt umfasst die notwendigen Datenstrukturen zur Repräsentation der zu simulierenden virtuellen Soundquellen, des virtuellen Hörers sowie des virtuellen Environments. Des Weiteren müssen Strukturen für das physikalische Environment und den physikalischen Hörer im Datenmodell aufgenommen sein.

3.2.1.1. Virtuelle Soundquelle

Eine virtuelle Soundquelle bezeichnet eine durch das 3-D Sound System zu simulierende Soundquelle.

Für die Soundquelle ist deren Position, angegeben in Koordinaten eines kartesischen Koordinatensystems, für die Positions- und Richtungscharakteristik notwendig. Das Ausbreitungsverhalten der Soundquelle kann ungerichtet oder gerichtet sein. Handelt es sich um eine gerichtete Soundquelle, ist die Orientierung ebenso notwendig wie die Definition des Ausbreitungsverhaltens der Soundquelle. Zur Definition des Ausbreitungsverhaltens können zwei Winkel dienen, mit Hilfe derer ein den Orientierungsvektor umgebender Zylinder definierbar ist. Zusätzlich kann die Konzentration der Schallausbreitung definiert sein, das Maß also, mit dem der Schall von der zentralen Achse des beschriebenen Zylinders zur Oberfläche hin abnimmt. (Siehe Abbildung 7)

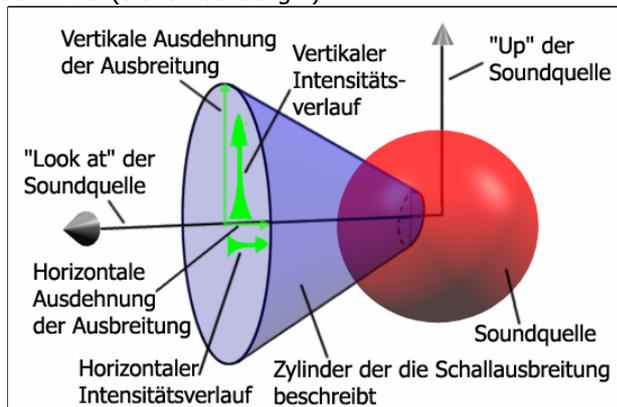


Abbildung 7 : Ausbreitungsverhalten virtueller Soundquellen

Zur Beschreibung der Geschwindigkeit einer Soundquelle kann eine skalare Größe nicht genügen, da sich eine Soundquelle nicht zwingend entlang ihrer Ausbreitungsrichtung bewegt. Eine vektorielle Größe beschreibt dann die Bewegung einer Soundquelle. Aus der Lage des Vektors ergibt sich die Bewegungsrichtung und die Länge definiert die Geschwindigkeit. (Siehe Abbildung 9)

Neben der Richtung, in die sich der Schall, ausgehend von der Position der Schallquelle, ausbreitet, ist festzulegen, wie groß die Intensität des Schalls, also die Lautstärke der Soundquelle, an ihrer Position ist.

3.2.1.2. Virtueller Hörer

Eine einzigartige Datenstruktur nimmt alle notwendigen Attribute auf, die helfen, einen virtuellen Hörer im virtuellen Environment zu beschreiben.

Die Position bildet mit der Orientierung die Hauptmerkmale des virtuellen Hörers. Beide beziehen sich auf den Kopf des virtuellen Hörers und sind, wie die Position der virtuellen Soundquelle, in kartesischen Koordinaten angegeben.

Bewegt sich ein Hörer nicht in Richtung der angegebenen Orientierung, kann auch beim virtuellen Hörer eine skalare Größe zur Beschreibung der Bewegung nicht ausreichen. Auch hier ist die Bewegung als vektorielle Größe unabhängig von der Orientierung anzugeben.

3.2.1.3. Virtuelles Environment

Das zu berechnende Klangbild basiert neben den Parametern der virtuellen Soundquellen und denen des virtuellen Hörers auf dem virtuellen Environment, in dem das Schallereignis stattfindet.

Das Environment ist primär geprägt durch das Ausbreitungsmedium, die Geometrien der sich im Environment befindlichen Objekte sowie der Materialien dieser.

Das Ausbreitungsmedium nimmt durch seiner Dichte Einfluß auf die Ausbreitungsgeschwindigkeit und die Dämpfung des Schalls. Das Medium kann so über die Angabe der Dichte beschrieben werden, oder aber über die Angabe eines Dämpfungsfaktors für die Intensitätsanpassung über die Entfernung und die bei der Ausbreitung im Raum verursachte Tiefpaß Filterung.

Die Beschreibung der Geometrien im Environment ermöglicht die exakte Berechnung von frühen und späten Reflektionen sowie von Beugung, Brechung und weiterer Effekte. Sie sorgt für ein realistisches räumliches Klangbild. Die Geometrien können in der Datenstruktur des Environments aufgenommen werden, oder aber aus der interaktiven Anwendung abgeleitet werden, die das 3-D Sound System nutzt. Ein Ableiten der Geometrien bedingt jedoch mit dem virtuellen Environment eine Referenz auf die Geometriebeschreibungen der interaktiven Grafikanwendung in die Datenstruktur aufzunehmen. Werden die für die visuelle Darstellung modellierten Objekte aus der interaktiven Anwendung als Berechnungsgrundlage für Reflektionen und weitere Effekte genutzt, ist mit dem höchsten Maß an Realitätsnähe zu rechnen, sofern die zur Berechnung verwendeten Verfahren ein realitätsnahes Verhalten umsetzen.

Da die Verfahren zur exakten Berechnung der eben erwähnten Effekte auf Basis von gegebenen Geometrien sehr rechenaufwendig sind und keine Echtzeitverarbeitung erlauben [FUN_99], kann auch ein Ansatz zur Beschreibung des Environments gewählt werden, bei dem nicht das Environment selbst, denn die Charakteristik des Environments als Parameter herangezogen wird. Die Charakteristik eines Environments zeichnet sich durch die zeitliche Verzögerung zwischen dem direkten Schall und der ersten frühen Reflektion (Reflection Delay), die Verzögerung nach der ersten Reflektion bis zum Auftreten der ersten späten Reflektion (Reverberation Delay) sowie die Zeit bis zum Abklingen der späten Reflektionen aus(Decay Time). (Siehe Abschnitt 2.1.3.1.) Durch die Angabe der Parameter zur Beschreibung der zeitlichen Aspekte und der Parameter, die die Intensität der frühen- und späten Reflektionen beschreiben, lassen sich so Methoden implementieren, die durch ihre starke Vereinfachung echtzeitfähig bleiben. Für die einzelnen Parameter können dann in der laufenden Anwendung Presets von Werten genutzt werden, die sich aus Messungen in Beispiel Environments ergeben haben.

Je nach gewählter Methode zur Beschreibung der Geometrien im virtuellen Environment kann so die Datenstruktur in ihrem Umfang variieren.

3.2.1.4. Physikalischer Hörer

Die Abbildung des physikalischen Hörers ist insbesondere für eine effektvolle Wiedergabe unter Nutzung eines 3-D Auditory Displays notwendig.

Die Position und die Orientierung des Hörers im physikalischen Environment sind hier die wesentlichen Aspekte, die zu berücksichtigen sind, um eine Anpassung der Wiedergabe an Bewegungen des Anwendungsnutzers zu ermöglichen.

3.2.1.5. Physikalisches Environment

Die Struktur des physikalischen Environment zeichnet sich vor allem dadurch aus, dass es die Position der Lautsprecher zu beschreiben vermag sowie die Art des verwendeten 3-D Auditory Displays. Position und Art des Displays bilden die Grundlage für die Berechnung des finalen wiedergegebenen Raumklangs.

3.2.2. Verarbeitungsaspekte des Anforderungskataloges

Die Verarbeitungsaspekte umfassen die in der Echtzeit 3-D Sound – Verarbeitung zu berücksichtigenden Effekte und die Anzahl der in Echtzeit verarbeitbaren virtuellen Soundquellen.

Effekte können unterschieden werden zwischen Soundeffekten und Raumeffekten. Soundeffekte basieren auf Parametern, die primär von einer virtuellen Soundquelle abhängig sind. Raumeffekte hingegen basieren aus Parametern, die durch das virtuelle Environment gegeben sind.

Der wichtigste Effekt im Zusammenhang mit der Ortung von Soundquellen ist die Berechnung der Beeinflussung des Soundspektrums über die Distanz der virtuellen Soundquelle zum virtuellen Hörer. Ein weiterer für die räumliche Wahrnehmung wichtiger Effekt ist der Doppler Effekt, der einen Schluss auf die relative Geschwindigkeit einer virtuellen Soundquelle zulässt. Zusätzlich sollte noch die Berücksichtigung der Ausbreitungscharakteristik, also die Implementierung eines Ausbreitungsmodells, erfolgen.

Zu den Raumeffekten zählen Effekte wie Beugung, Brechung, Absorption, frühe- und späte Reflektionen. Die Methode und Exaktheit, nach der diese berechnet werden, ob auf Basis exakter Geometrien oder aber von Environment Charakteristika, bestimmt zum Einen die Performance eines 3-D Sound Systems, zum Anderen die Realitätsnähe dessen.

3.2.3. Wiedergabeaspekte des Anforderungskataloges

Welche Verfahren zur Wiedergabe des 3-D Klangbildes genutzt werden und in wie weit der physikalische User

dabei Beachtung findet, lässt sich unter den Wiedergabeaspekten zusammenfassen.

Von besonderem Interesse bei den Wiedergabetechniken sind dabei solche, die besonders im Bereich von Virtual Reality Systemen ihre Einsatzgebiete haben und solche, die sich als Quasi Standard im Entertainment Bereich bereits etabliert haben. Die Unterstützung von stereo, vierkanal, 5.1 oder kubischen achtkanal Lautsprecher Setups wird die Integrierbarkeit in Virtual Reality Systeme und Multimedia - Anwendungen verbessern.

Stereo Wiedergabe hat sich schon seit nunmehr Jahrzehnten im Unterhaltungsbereich als Standard durchgesetzt. Es kann davon ausgegangen werden, dass jede verfügbare Soundhardware zumindest Stereo Wiedergabe unterstützt. Des Weiteren wird mit der Verfügbarkeit von stereofähiger Hardware auch die Umsetzung binauraler Techniken auf HRTF Basis ermöglicht. Vierkanal Systeme sind leicht in CAVE® basierten Systemen zu installieren und werden teilweise bereits systemseitig durch Standard Soundkomponenten unterstützt.[TOD_02_2]

Als gängiges Format im computerbasierten Entertainmentbereich wird für 5.1 Setups eine große Vielfalt an verfügbarer Hardware existieren, bzw. der Anwendungsnutzer bereits über ein solches Setup verfügen.

Für eine optimale Integrierbarkeit in CAVE® basierte Virtual Reality Systeme sollten kubische achtkanal Lautsprecher Setups als mögliches 3-D Auditory Display unterstützt werden.

Von den hier erwähnten Setups ermöglichen nur die binauralen Wiedergabetechniken auf HRTF Basis sowie das kubische achtkanal Lautsprecher Setup die Wiedergabe dreidimensionaler Klangbilder. Sie müssen zum Pflichtprogramm einer erfolgreichen Umsetzung eines 3-D Sound Systems zählen.

Um dem physikalischen Hörer neben der Position im Hotspot des Setups wechselnde Positionen zu ermöglichen, ist die Berücksichtigung von geänderten User Parametern auf Basis eines User Trackings zwingend notwendig. (Siehe Abschnitt 2.2.)

3.2.4. Abstraktionsgrade des Anforderungskataloges

Der I3DL2 Minimalkatalog kennzeichnet nicht die minimalen Anforderungen an ein interaktives 3-D Sound System, es definiert vielmehr Anforderungen an ein System, das bereits in einem gewissen Rahmen Realitätsnähe zu implementieren versucht. (Siehe Abschnitt 3.1.2.) 3-D Sound Systeme, die auf einer abstrakteren Ebene Information mittels auditiver Techniken vermitteln, also keine Realitätsnähe versuchen zu implementieren, können durch einen solchen Anforderungskatalog nicht ausreichend beurteilt werden. Es ist daher wichtig, dass nicht nur die ergänzenden Parameter und Detaillierungen der einzelnen Aspekte in einen Anforderungskatalog aufgenommen werden, denn mehr eine Unterteilung des Kataloges in verschiedene Grade der Abstraktion erfolgt, so dass dann eine Beurteilung verschiedener

Implementierungen gleichen Abstraktionsgrades erfolgen kann.

Neben dem Minimalsystem, das lediglich die Primärinformation einer Soundquelle, also die Position, vermittelt und dem Maximalsystem, das eine maximale Realitätsnähe erreicht, gibt es mindestens eine Abstufung für Systeme, die ein Mittelmaß implementieren.

Systeme die sich auf die Vermittlung der Position beschränken haben hier den höchsten Abstraktionsgrad, stehen also in einer Rangliste der Abstraktion an oberster, erster Stelle. Je mehr Details in die Implementierung des 3-D Sound Systems aufgenommen werden, desto geringer wird von der Realität abstrahiert, so dass sich für solche Systeme höhere Abstraktionsgrade ergeben. Dies hat zur Folge, dass man später möglichen höheren Detaillierungsgraden durch die Einführung weiterer höherer Abstraktionsgrade gerecht werden kann.

Die Abstraktionsgrade können auch als Abstraktionslevel bezeichnet werden. Die einzelnen Level bauen aufeinander auf, d.h. Level mit steigender Ordnungszahl implementieren alle Aspekte des vorangegangenen Abstraktionslevels.

3.2.4.1. Abstraktionslevel 1

Der Abstraktionslevel 1 beschreibt den Level des Anforderungskataloges mit dem höchsten Abstraktionsgrad. Lediglich die zur Erzeugung eines dreidimensionalen Klangbildes notwendigen Informationen werden verarbeitet.

Ein entsprechendes Klangbild lässt sich allein unter Berücksichtigung der Basisparameter der virtuellen Soundquelle, der Parameter des physikalischen Hörers und der Parameter des physikalischen Environments erzeugen und wiedergeben.

Als Basisparameter der virtuellen Soundquelle reicht schon die Position der Quelle aus, um diese räumlich zu positionieren. Die Aspekte des physikalischen Hörers und des physikalischen Environments mit dem entsprechend berücksichtigten Tracking des Users hingegen müssen komplett abgebildet sein, um eine präzise Wiedergabe sicherzustellen.

Die Wiedergabe muss dabei jedoch über ein 3-D Auditory Display erfolgen, das über die Möglichkeit zur Wiedergabe eines 3-D Eindrucks verfügt.

Eine Mindestanzahl von 16 Soundquellen ist dabei in Echtzeit bei einer Sampling Rate von 22050 Hz und einer Quantisierung von 16 Bit zu verarbeiten. Die Verarbeitung darf dabei keine durch zu lange Verarbeitungsprozesse oder verzögerte Speicherzugriffe bedingten Artefakte erzeugen. Die Latenz bei der Verarbeitung muss so gering sein, dass sich eine Änderung von Attributen innerhalb von maximal 40 ms im Klangbild widerspiegelt. Die Übernahme von Parametern hat synchron zu erfolgen.

3.2.4.2. Abstraktionslevel 2

Abstraktionslevel 2 implementiert eine detaillierte Simulation des Verhaltens der Soundquelle. Es werden sämtliche Aspekte der virtuellen Soundquelle bei der Verarbeitung berücksichtigt.

Zur Umsetzung und Berechnung des Ausbreitungsmodells ist die Implementierung weiterer Aspekte des virtuellen Hörers notwendig. Erst durch die Orientierung des virtuellen Hörers und der virtuellen Soundquelle sowie durch die vollständige Beschreibung des Ausbreitungsmodells der Soundquelle kann eine exakte Berechnung der Schallausbreitung erfolgen.

Bei der Verarbeitung von Soundeffekten wird neben dem Ausbreitungsmodell zumindest der Doppler Effekt berechnet, um dem User die relative Bewegung der Soundquelle zu vermitteln.

3.2.4.3. Abstraktionslevel 3

Mit dem Abstraktionslevel 3 werden zusätzlich Informationen über das virtuelle Environment in der Verarbeitung berücksichtigt. Die Berücksichtigung erfolgt im Level 3 jedoch auf Basis der Charakteristik des virtuellen Environments. Es werden frühe und späte Reflektionen berechnet. Die Berechnung von Beugung und Brechung erfolgt hier aufgrund mangelnder Informationen über die Position und Beschaffenheit der sich im virtuellen Environment befindlichen Objekte nicht.

Nicht das virtuelle Environment, denn das akustische Verhalten dessen wird in den Parametern erfasst und in der Verarbeitung berücksichtigt. (Siehe Abschnitt 3.2.1.3.)

3.2.4.4. Abstraktionslevel 4

Eine detailliertere Berechnung der durch das Environment verursachten Effekte wird mit der Berücksichtigung der Geometrien im Raum möglich. Sowohl die Position als auch die Modellierung der Objekte werden zusammen mit der Beschaffenheit der Materialien bei der Effektberechnung berücksichtigt, so dass eine exakte Berechnung der frühen und späten Reflektionen sowie Beugung und Brechung erfolgt.

3.2.5. Anforderungskatalog für 3-D Sound Systeme

Zusammenfassend lässt sich der Anforderungskatalog tabellarisch darstellen. Die Kennzeichnung der einzelnen Parameter ermöglicht die schnelle Identifizierung der jeweiligen Anforderungen für die einzelnen Abstraktionslevel. (Siehe Abbildung 8)

	Abstraktionslevel 1	Abstraktionslevel 2	Abstraktionslevel 3	Abstraktionslevel 4
Randbedingungen Verarbeitung				
22050 HZ Sampling Rate	✓	✓	✓	✓
16 Bit Quantisierung	✓	✓	✓	✓
max. 40 ms Latenz	✓	✓	✓	✓
min. 16 virtuelle Soundquellen	✓	✓	✓	✓
Keine Artefakte	✓	✓	✓	✓
Datenaspekte				
Virtuelle Soundquelle				
Position*	✓	✓	✓	✓
Orientierung*		✓	✓	✓
Ausbreitungsverhalten		✓	✓	✓
Geschwindigkeit**		✓	✓	✓
Lautstärke	✓	✓	✓	✓
Virtueller Hörer				
Position (in kartesischen Koordinaten)	✓	✓	✓	✓
Orientierung (in kartesischen Koordinaten)		✓	✓	✓
Geschwindigkeit (als vektorielle Größe)		✓	✓	✓
Virtuelles Environment				
Environment Charakteristik (Angabe von Reflektions-, Reverbintensitäten und -verzögerungen sowie von Reverb Decay)			✓	
Detailliertes Environment				✓
Physikalischer Hörer				
Position (in kartesischen Koordinaten)	✓	✓	✓	✓
Orientierung (in kartesischen Koordinaten)	✓	✓	✓	✓
Physikalisches Environment				
Lautsprecher Position (in kartesischen Koordinaten)	✓	✓	✓	✓
Verarbeitungsaspekte				
Soundeffekte				
Abschwächung über die Distanz	✓	✓	✓	✓
Doppler Effekt		✓	✓	✓
Berechnung der Ausbreitungscharakteristik		✓	✓	✓
Raumeffekte				
Beugung				✓
Brechung				✓
Absorption				✓
Frühe Reflektionen / Hall			✓	✓
Späte Reflektionen / Reverberation			✓	✓
Wiedergabeaspekte				
(eine der Techniken ist zu implementieren)				
Stereo Wiedergabe				
Binaural HRTF Kopfhörer Wiedergabe	✓	✓	✓	✓
Binaural HRTF Lautsprecher Wiedergabe	✓	✓	✓	✓
Vierkanal Wiedergabe				
5.1 Wiedergabe				
Kubische Achtkanal Wiedergabe	✓	✓	✓	✓
User Tracking berücksichtigt	✓	✓	✓	✓

Abbildung 8 : Übersicht Anforderungskatalog

4. Beurteilung existierender interaktiver 3-D Sound Systeme

Einem Neuentwurf eines interaktiven 3-D Sound Systems sollte eine Recherche vorausgehen, die existierende Systeme auf ihre Einsetzbarkeit hin prüft. Mit Hilfe des erarbeiteten Anforderungskataloges und auf Basis vorangegangener Untersuchungen [TOD_03] wird in diesem Abschnitt die Beurteilung verschiedener 3-D Sound System Ansätze.

Bei der Untersuchung von Sound Systemen unterscheidet man zwischen solchen, die speziell für Multimedia – Entwicklungsumgebungen konzipiert wurden und sich besonders effizient einbinden lassen und solchen, die, losgelöst von existierenden Umgebungen zur Erstellung multimedialer Inhalte, einen

Satz an Funktionen und Datenkonstrukten zur interaktiven 3-D Sound – Verarbeitung zur Verfügung stellen.

Solche Systeme können als „Special – Purpose“ und „Multi – Purpose“ unterschieden und bezeichnet werden.

4.1. Beurteilung von Special – Purpose 3-D Sound Systemen

Als Special Purpose 3-D Sound Systeme werden solche Systeme bezeichnet, die speziell vor dem Hintergrund entwickelt wurden, eine effektive Lösung zur Einbindung von interaktivem 3-D Sound für eine bestimmte Entwicklungsumgebung bereitzustellen. Als Entwicklungsumgebung ist nicht eine integrierte Entwicklungsumgebung im Sinne eines Microsoft Visual Studio® zu verstehen, denn viel mehr ein Satz von Bibliotheken, deren gemeinsames Ziel die Erstellung multimedialer Inhalte ist. Mit Virtual Reality als einem besonderen Aspekt von Multimedia Anwendungen werden solche Umgebungen auch zur Erstellung von Virtual Reality Anwendungen genutzt.

Special – Purpose Systeme identifizieren sich primär dadurch, dass ihre Syntax stark an die der Umgebung angelehnt ist, für die sie entwickelt wurden. Neben der ähnlichen Syntax werden häufig gemeinsame Datentypen verwendet. In manchen Fällen ergibt sich eine besonders enge Bindung, wenn das 3-D Sound System nach dem Aufruf von Methoden der Entwicklungsumgebung, oder aber nach Referenzen auf Objekte der Entwicklungsumgebung verlangt.

Abhängig von der Stärke der Abhängigkeit des Sound Systems von der Umgebung, sind solche Systeme auch in anderen Umfeldern einsetzbar. Existiert lediglich eine geringe Kopplung über ähnliche Syntax oder aber gemeinsame Datenstrukturen, ist in der Regel eine „Fremdverwendung“ möglich. Besteht weiter jedoch eine Verbindung über notwendige Funktionsaufrufe oder aber Objekte, die mit Hilfe der Entwicklungsumgebung zu erstellen sind, ist zumeist eine anderweitige Verwendung uneffektiv zu bewerten.

Microsoft® stellt mit der DirectSound® Bibliothek als Bestandteil von DirectX® das wohl meistgenutzte interaktive Special – Purpose 3-D Sound System zur Verfügung.

DirectSound® fügt sich optimal in die Strukturen und Abläufe von DirectX® basierten Anwendungen ein. DirectX® zählt aufgrund der engen Verzahnung mit dem Betriebssystem des Herstellers zu der Multimedia – Bibliothek mit dem höchsten Verbreitungsgrad. Zur alleinstehenden Nutzung der DirectSound® Komponenten ist die Einbindung weiterer Komponenten des DirectX® Paketes erforderlich.

DirectSound® erfüllt mit Einschränkungen die Anforderungen des Abstraktionslevel 3 des Anforderungskataloges. Mit Ausnahme der Berücksichtigung der physikalischen Hörerattribute sind alle für den Level 3 notwendigen Parameter in der Verarbeitung berücksichtigt. (Siehe Abbildung 9) Primär für den Heimbereich ausgelegt, wird auf Seiten von

DirectX® nicht von einem möglichen User Tracking ausgegangen. [TOD_03]

Aus Entwicklersicht bietet sich mit DirectSound® ein umfangreiches und leistungsfähiges System, das auch ohne ausgeprägtes Akustik – Hintergrundwissen die einfache Möglichkeit zur Einbindung von 3-D Sound in interaktive Echtzeitanwendungen einräumt.

Steigender Verbreitungsgrad bei der Entwicklung von Multimedia – Anwendungen zeichnet ebenso die Programmiersprache Java® aus. Java® bietet neben der mit ihr implementierbaren Multimedialität weiter einen Satz von Paketen zur Umsetzung von interaktiven Echtzeit 3-D Anwendungen. Mit dem Java 3D® Paket wird ein Interface zur Erstellung von interaktiven Anwendungen zur Verfügung gestellt. Mit der Einführung von Java 3D® stieg der Verbreitungsbereich von Java im Bereich der Virtual Reality Anwendungen sprunghaft an. Es ist mit Java 3D® ein Paket gegeben, das sowohl im Heim- als auch im professionellen Virtual Reality Bereich einen hohen Verbreitungsgrad vorweist.

Die „3D Sound“ Komponente von Java3D® stellt Funktionalitäten und Strukturen zur Verfügung, die eine Einbindung von 3-D Sound in Java 3D® basierte Anwendungen ermöglicht.

In der Szenengraph basierten Struktur der grafischen Java3D® Anwendung, fügen sich die 3-D Sound Knoten ohne Probleme in den Szenengraph ein.

Eine externe Verwendung der 3-D Sound Pakete ist indes nicht möglich. [TOD_03]

Die Leistungsfähigkeit von Java 3D® ist jedoch begrenzt. Es werden die Parameter des virtuellen Hörers wie auch die Position, Orientierung und das Ausbreitungsmodell der virtuellen Soundquellen berücksichtigt. Die Aspekte des virtuellen- und physikalischen Environments wie auch des physikalischen Hörers bleiben jedoch unberücksichtigt. Java 3D® erfüllt keinen der Abstraktionslevel des Anforderungskataloges. (Siehe Abbildung 9)

Wird das endgültige Klangbild unter Berücksichtigung der notwendigen Parameter berechnet, ist eine Wiedergabe lediglich als zweikanal Wiedergabe möglich. Es besteht keine Möglichkeit, ein 3-D Auditory Display zu betreiben. Mit Java 3D® bietet sich aus Entwicklersicht eine Komponente zur 3-D Sound Einbindung, die sich elegant eingliedert, jedoch die mit den zu manipulierenden Parametern implizierte Erwartung, aufgrund der mangelhaften Wiedergabe, nicht zu erfüllen vermag.

Die zwei hier vorgestellten Special – Purpose 3-D Sound Systeme unterscheiden sich stark in ihrer Leistungsfähigkeit. Ihnen gemein ist jedoch die schwere oder unmöglichen Nutzbarkeit losgekoppelt von ihren Entwicklungsumgebungen. Weiter zeichnen sich beide der hier vorgestellten Implementierungen durch eine mangelnde Berücksichtigung der physikalischen Hörer Parameter aus, da sie für den Desktop PC Bereich und nicht für den Einsatz in Virtual Reality Systemen konzipiert sind.

Sie stellen innerhalb der Entwicklungsumgebung, für die sie konzipiert wurden, die Ideallösung dar.

Eine gesteigerte Leistungsfähigkeit gekoppelt mit einer flexibleren Nutzbarkeit, auch außerhalb der

Entwicklungsumgebung, wird den Verbreitungsgrad von Special – Purpose 3-D Sound Systemen weiter erhöhen können.

4.2. Beurteilung von Multi – Purpose 3-D Sound Systemen

Multi – Purpose 3-D Sound Systeme stellen eine Sammlung von Funktionen, Datenstrukturen und internen Prozessen zur Verfügung, mit Hilfe derer Entwickler in die Lage versetzt werden, 3-D Sound in beliebige Anwendungen zu integrieren.

Multi – Purpose Systeme weisen in der Regel keine Verbindung zu ausgewählten Entwicklungsumgebungen im oben beschriebenen Sinn auf. Sie bieten damit große Flexibilität bei der Wahl des Basis Multimedia Systems, in das das 3-D Sound zu integrieren ist, können aus diesem Grund jedoch nicht auf spezielle Aspekte bestimmter Entwicklungsumgebungen eingehen und diese zur Performancesteigerung nutzen. Weiter verlangen Multi – Purpose Systeme zumeist nach einer redundanten Datenhaltung, d.h. Positionen von grafischen Objekten können nicht direkt im Sound System verwendet werden. Die fehlende Verbindung zwischen 3-D Sound System und der Anwendung, in die Sound eingebunden wird, bedingt eine durch den Entwickler zu steuernde Strategie zur Synchronisation der grafischen und akustischen Abläufe.

Es existiert eine Vielzahl von Multi – Purpose Sound Systemen. Eine geringe Anzahl dieser jedoch berücksichtigt die Implementierung von 3-D Sound Aspekten, so dass hier nur eine kleine Auswahl kurz vorgestellt und beurteilt wird.

OpenAL® bezeichnet die Open Audio Library. Die Bibliothek wurde als Multi – Purpose 3-D Sound System entwickelt, das dem Entwickler einen einfachen Zugang zur Einbindung von 3-D Sound in beliebige Anwendungen verschaffen soll.

Neben der allgemeinen Verwendbarkeit wurde bei der Syntax eine Namenskonvention eingeführt, die der von OpenGL® sehr ähnelt. So soll die Einbindung von 3-D Sound speziell in OpenGL® basierte Anwendungen vereinfacht werden.

Da weder gemeinsame Datenaspekte noch gemeinsame Routinen zwischen OpenGL® und OpenAL® existieren, ist die Unterstützung bei der Einbindung jedoch stark limitiert. So müssen für die Grafikdarstellung der einzelnen Objekte bereits durchgeführte Transformationen separat ein zweites Mal für die virtuellen Soundquellen von OpenAL® durchgeführt werden. Es resultiert daraus eine redundante Datenhaltung und ein programmiertechnischer Mehraufwand. OpenAL® zeigt damit bei der Einbindung in OpenGL® basierte Anwendungen kein in besonderem Maße unterstütztes Verhalten, so dass es auch für OpenGL® als Multi – Purpose 3-D Sound System gelten muß.

Die Leistungsfähigkeit von OpenAL® ist nur schwer beurteilbar. Die erzielbaren Resultate sind stark von dem gewählten Betriebssystem abhängig. So implementiert

OpenAL® unter Windows® Betriebssystemen mit Einschränkungen einen Abstraktionslevel 3 des Anforderungskataloges, unter Linux Betriebssystemen wird jedoch keiner der Abstraktionslevel implementiert. (Siehe Abbildung 11)

Der Grund für die wechselnde Leistungsfähigkeit liegt in der, je nach Betriebssystem abweichenden, Implementierung. Unter Windows® Betriebssystemen wird die Funktionalität direkt auf Aufrufe von DirectSound® abgebildet. Intern werden dann DirectSound® Datenstrukturen verwendet und die Berechnung und Wiedergabe von 3-D Sound erfolgt über DirectSound®. Aufgrund der nicht vorhandenen Verfügbarkeit von DirectSound® unter Linux Betriebssystemen wird hier eine speziell für OpenAL® entwickelte 3-D Sound Engine genutzt. Diese bleibt in der Leistungsfähigkeit gegenüber DirectSound® zurück. Die Parameter der virtuellen Soundquelle werden fast komplett berücksichtigt und auch Soundeffekte, wie der Doppler Effekt, oder aber Raumeffekte werden berechnet. Die Wiedergabe jedoch kann nur unter Nutzung eines zweikanal Stereo Setup erfolgen. [TOD_03]

OpenAL® bietet somit für Linux Betriebssystem ein in den Leistungen eingeschränktes 3-D Sound System. Für Windows® Betriebssysteme kann es aufgrund der nicht vollständigen Eingliederung in die Strukturen und Abläufe von OpenGL® keine Konkurrenz zu DirectSound® darstellen, so dass hier vom Entwickler eher die direkte Nutzung von DirectSound® der indirekten Nutzung dessen durch OpenAL® vorgezogen werden dürfte.

In vorangegangenen Arbeiten des Autors wurden weitere Multi – Purpose 3-D Sound Systeme entwickelt. [TOD_02_1][TOD_02_2]

Als Portaudio 3D Extension entstand im Rahmen einer Diplomarbeit ein 3-D Sound System, das ein kubisches Achtkanal Lautsprecher Setup zur Wiedergabe nutzt, doch bei Berücksichtigung der primären Parameter der virtuellen Soundquelle und des virtuellen Hörers nicht in der Lage ist, ausreichend virtuelle Soundquellen in Echtzeit zu berechnen. Während das physikalische Environment bei der Verarbeitung Berücksichtigung findet, bleibt der physikalische Hörer unberücksichtigt.

Die Portaudio 3D Extension erfüllt somit keinen der Abstraktionslevel des Anforderungskataloges. (Siehe Abbildung 9)

Die Portaudio 3D Extension bleibt damit, trotz des implementierten 3-D Auditory Display, als wenig einsetzbares Produkt hinter den Anforderungen an 3-D Sound Systeme zurück.

Auf Basis eines existierenden Sound Systems wurde mit dem Spatialized Snerd Server ein verteiltes 3-D Sound System entwickelt, das mit einem Client- und einem Servermodul die Soundverarbeitung implementiert. Während die Clientkomponenten in die Anwendung integriert werden können, kann die Soundverarbeitung auf einem getrennt existierenden System durchgeführt werden. Die zur Verarbeitung der Basisanwendung verfügbare Performance wird damit durch die Einbindung von Sound kaum geschmälert. Die

Auslagerung des 3-D Sound Systems ermöglicht weiter den Einsatz von auf Soundverarbeitung spezialisierter Hardware. Bei der Verarbeitung werden die primären Parameter der virtuellen Soundquelle sowie des virtuellen Hörers berücksichtigt. Die Wiedergabe erfolgt angepasst an die aktuelle Position des physikalischen Hörers.

Aufgrund der initialen Implementierung für IRIX® Betriebssysteme unter Nutzung von SGI Indigo® Rechnerhardware kann eine Wiedergabe lediglich unter Nutzung eines vier- oder zweikanal Lautsprecher Setup erfolgen. Eine erfolgte Portierung auf Linux Betriebssysteme, die zurzeit lediglich eine zweikanal Wiedergabe unterstützt, legt jedoch den Grundstein für eine weite Verbreitung des 3-D Sound Systems.

Es kann aufgrund der fehlenden Implementierung eines 3-D Auditory Displays keine Bewertung gemäß eines der Abstraktionslevel des Anforderungskataloges erfolgen. (Siehe Abbildung 9)

Dem Anwendungsentwickler steht mit dem Spatialized Snerd Sound Server ein System zur Verfügung, das eine Verarbeitung dem Abstraktionslevel 1 entsprechend durchführt, jedoch keine entsprechende Wiedergabe implementieren kann. Unter Abstrichen ist dieses System somit als performante Lösung zur Einbindung von 3-D Sound nutzbar.

Wie bei den Special – Purpose Systemen unterscheiden sich auch die Multi – Purpose Systeme stark in ihrer Leistungsfähigkeit. Solange die Funktionalität nicht direkt auf Special – Purpose Systeme abgebildet ist, liegen alle Multi – Purpose Systeme hinter der Leistungsfähigkeit der Special – Purpose Implementierungen zurück. Dieses Phänomen lässt sich jedoch dadurch erklären, dass mit der Verfügbarkeit von 3-D Sound das Komplettpaket der Entwicklungsumgebung an Nutzen und damit an Popularität gewinnen soll, so dass hier ein größeres Augenmerk darauf gelegt wird, dass eine 3-D Sound Implementierung nicht hinter dem Leistungsstandard der Umgebung zurücksteht.

Wenngleich die variable Einbindbarkeit der Multi – Purpose Implementierungen die Einbindung von 3-D Sound in Anwendungen einer beliebigen Basisumgebung vereinfacht, bleibt der zu erwartende Leistungsgewinn bei einer möglichen Einbindung einer Special – Purpose Lösung zu dem dadurch entstehenden Mehraufwand für den Entwickler in Relation zu setzen.

In den meisten Fällen wird sich ein Mehraufwand zur Einbindung eines Special – Purpose Systems als vorteilhaft für den gewonnenen 3-D Sound Eindruck erweisen.

	Abstraktionslevel 1	Abstraktionslevel 2	Abstraktionslevel 3	Abstraktionslevel 4	DirectSound®	Java 3D®	OpenAL® (auf DirectSound Basis)	OpenAL®	Portaudio 3D Extension	Spatialized Snerd Sound Server
Randbedingungen Verarbeitung										
22050 HZ Sampling Rate	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
16 Bit Quantisierung	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
max. 40 ms Latenz	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
min. 16 virtuelle Soundquellen	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Keine Artefakte	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Datenaspekte										
Virtuelle Soundquelle										
Position*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Orientierung*	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Ausbreitungsverhalten	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Geschwindigkeit**	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Lautstärke	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Virtueller Hörer										
Position (in kartesischen Koordinaten)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Orientierung (in kartesischen Koordinaten)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Geschwindigkeit (als vektorielle Größe)	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Virtuelles Environment										
Environment Charakteristik (Angabe von Reflektions-, Reverbintensitäten und -verzögerungen sowie von Reverb Decay)			✓		✓		✓			
Detailliertes Environment				✓						
Physikalischer Hörer										
Position (in kartesischen Koordinaten)	✓	✓	✓	✓						✓
Orientierung (in kartesischen Koordinaten)	✓	✓	✓	✓						
Physikalisches Environment										
Lautsprecher Position (in kartesischen Koordinaten)	✓	✓	✓	✓	✓		✓		✓	✓
Verarbeitungsaspekte										
Soundeffekte										
Abschwächung über die Distanz	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Doppler Effekt		✓	✓	✓	✓	✓	✓	✓	✓	✓
Berechnung der Ausbreitungscharakteristik		✓	✓	✓	✓	✓	✓	✓	✓	✓
Raumeffekte										
Beugung				✓						
Brechung				✓						
Absorption				✓			✓	✓		
Frühe Reflektionen / Hall			✓	✓	✓		✓	✓		
Späte Reflektionen / Reverberation			✓	✓	✓		✓	✓		
Wiedergabeaspekte										
(eine der Techniken ist zu implementieren)										
Stereo Wiedergabe						✓		✓		✓
Binaural HRTF Kopfhörer Wiedergabe	✓	✓	✓	✓	✓		✓			
Binaural HRTF Lautsprecher Wiedergabe	✓	✓	✓	✓	✓					
Vierkanal Wiedergabe					✓					✓
5.1 Wiedergabe					✓					
Kubische Achtkanal Wiedergabe	✓	✓	✓	✓					✓	
User Tracking berücksichtigt	✓	✓	✓	✓						

Abbildung 9 : Beurteilung existierender interaktiver 3-D Sound Systeme

5. Design eines interaktiven 3-D Sound Systems

Einschränkungen in der Leistungsfähigkeit existierender 3-D Sound Systeme und deren mitunter aufwendige Einbindung in existierende interaktive Echtzeit Anwendungen begründen den Entwurf eines neuen 3-D Sound Systems.

Ein neu zu entwerfendes System sollte sich an den Stärken bereits existierender Systeme orientieren, aber auch deren Schwächen beim Entwurf berücksichtigen.

Vor der tatsächlichen Entwurfsphase erfolgt eine präzise Formulierung der Ziele, anhand derer die zu

lösenden Problematiken abgeleitet und entsprechende Lösungs- und Implementierungsstrategien entwickelt werden.

Beginnend mit der Definition der zu erreichenden Ziele, wird anhand dieser eine Diskussion geführt, inwieweit die Erweiterung eines vorhandenen 3-D Sound Systems zur Erreichung der Ziele dienen kann.

Aufbauend auf einer Analyse von 3-D Sound Systemen kann ein Neuentwurf unter Berücksichtigung der gesetzten Ziele erfolgen.

Der so in diesem Kapitel entstehende Entwurf eines 3-D Sound Systems dient dann im nächsten Abschnitt der Implementierung des 3-D Sound Systems.

5.1. Zielsetzung

Ziel eines Neuentwurfes ist die Erstellung eines echtzeitfähigen interaktiven 3-D Sound Systems, das die Einbindung von 3-D Sound in interaktive Echtzeit Anwendungen ermöglicht bzw. vereinfacht. Die Einbindung soll speziell für solche interaktive Anwendungen erleichtert werden, wie sie zumeist in Virtual Reality Anwendungen implementiert sind, sie soll aber auch für abweichende Umgebungen möglich sein.

Mit dem Gesamtziel verbunden ist eine Vielzahl von Teilzielen, die in diesem Abschnitt ihre Erläuterung finden.

5.1.1. Anforderungskatalog konformer Entwurf

Das 3-D Sound System muss die im Anforderungskatalog formulierten Anforderungen erfüllen. Ein stabiler und effektiver Entwurf, der den Abstraktionslevel 1 erfüllt, soll Grundlage für die erste Implementierung sein und als Basis für weiterführende Entwürfe und Implementierungen der nachfolgenden Abstraktionslevel dienen.

Das im Rahmen dieser Arbeit zu erstellende Design geht über den Abstraktionslevel 1 nicht hinaus und legt die Schwerpunkte auf die Aspekte Stabilität und Performance.

5.1.2. Betriebssystem - Unabhängigkeit

Um sich als weit verbreitetes 3-D Sound System etablieren zu können und in Anwendungen auf verschiedenen Rechner - Plattformen eingesetzt werden zu können, ist ein Layout zu wählen, das eine einfache Portierung des Systems auf wechselnde Plattformen zulässt.

Für die initiale Implementierung ist eine Plattform zu wählen, die sowohl im Heim – Multimedia Bereich als auch im professionellen Virtual Reality- und Multimedia Bereich einen hohen Verbreitungsgrad hat, um mit dem ersten Release bereits ein weites Anwendungsfeld zu finden.

5.1.3. Modulares Design

Nur ein modulares Design kann eine gute Wartbarkeit garantieren. Ein modulares Design begünstigt weiter die Kapselung einzelner Problemaspekte.

Auftretendes Fehlverhalten im System kann bei einem modularen Entwurf schnell auf einzelne Module beschränkt werden. Im Problemfall ergibt sich so eine schnelle Wartung und Problembeseitigung.

Die einzeln in den Modulen gekapselten Aspekte können getrennt voneinander auf ihre Performance geprüft und auf diese hin optimiert werden. Für den

Fall, dass zukünftig Aspekte der Soundverarbeitung durch Hardwarekomponenten zur Verfügung gestellt werden, können die entsprechenden Module in diesem Fall durch hardwarebasierte Verarbeitung implementiert werden.

Mit der Umsetzung eines modularen Entwurfes wird eine Anpassbarkeit des 3-D Sound Systems an wechselnde Voraussetzungen, wie z.B. variierende Hardwarekomponenten, gewährleistet, so dass eine Vielzahl verschiedener Plattformen und Hardwarekomponenten unterstützt werden kann.

Neben der Anpassbarkeit kann mit einem modularen Design durch die Erweiterung um zusätzliche Module schnell auf neue Anforderungen eingegangen werden. Eine hohe Aktualität des Sound Systems bleibt damit gewahrt.

5.1.4. Special – Purpose Entwurf mit möglichem Multi – Purpose Einsatz

Um eine optimale Integrierbarkeit in interaktive Anwendungen sicherzustellen, sollte das Design des Systems auf eine enge Verschmelzung zwischen dem 3-D Sound System und der Entwicklungsumgebung ausgelegt sein, um redundante Datenhaltung und unnötige Funktionsaufrufe möglichst zu vermeiden und weiter eine Synchronisation zwischen Basisanwendung und 3-D Sound System zu ermöglichen.

Um auf der anderen Seite jedoch einen hohen Verbreitungsgrad zu erzielen, sollte das System – Layout einen Multi – Purpose Ansatz verfolgen, der die Integration von 3-D Sound in beliebige Anwendungen ermöglicht, ohne spezielle Datenaspekte oder Routinenaufrufe eines Special – Purpose Ansatzes beachten zu müssen.

Hier ist nicht ein Kompromiss aus beiden Ansätzen zu wählen, denn vielmehr mit Hilfe modularer Strukturen für die jeweiligen Anforderungen gekapselte Module zu entwickeln.

5.2. Erweiterung existierender 3-D Sound Systeme zur Zielerfüllung

Die Erweiterung existierender Ziele kann in vielen Situationen die Entwicklungszeit verkürzen helfen.

Keine der hier untersuchten bereits implementierten 3-D Sound Systeme kann jedoch als Grundlage eines erweiterten 3-D Sound Systems dienen. Individuell verschiedene Aspekte lassen eine Erweiterung der einzelnen Systeme nicht zu.

DirectSound® als Special – Purpose Implementierung ist durch die Abhängigkeit von Microsoft® Windows® Betriebssystemen zu stark eingeschränkt. Im Bereich von Virtual Reality Anwendungen spielen Windows® Betriebssysteme eine nur geringe Relevanz. Eine Portierung des Systems auf alternative Betriebssysteme ist aufgrund der „Closed Source“ Politik des Herstellers nicht möglich. Weiter bedarf es zur Einbindung von DirectSound® in nicht DirectX® basierte Anwendungen ein großes Maß an Aufwand, der als Zielvorgabe jedoch zu minimieren ist.

Mit der Beschränkung auf Anwendungen, die in Java® implementiert sind, bietet Java3D® lediglich ein semi – performantes System. Andere Programmiersprachen wie C, C++ lassen hier eine hardwarenahe Implementierung und dadurch bessere Performance zu. Die zu erzielende höhere Performance ist einer der Hauptgründe für den Einsatz abweichender Programmiersprachen im Bereich Virtual Reality. Eine Implementierung von Java3D® auf Basis einer abweichenden Programmiersprache käme einem Neudesign eines 3-D Sound Systems gleich.

Die wechselnden Implementierungsstrategien auf variierenden Betriebssystemen verlangen bei einer Erweiterung von OpenAL® nach Anpassungen an verschiedenen Stellen, unter wechselnden Voraussetzungen und mit abweichenden Implementierungsstrategien. Ein einheitliches, klares Design der zu berücksichtigenden Erweiterungen ist aus diesen Gründen nicht umsetzbar.

Exotische Implementierungen, wie die oben vorgestellten Portaudio 3D Extension und Spatialized Snerd Sound Server, implementieren jeweils nur kleine Bereiche eines 3-D Sound Systems und erfüllen nur einen geringen Teil der Zielvorgaben, so dass diese nicht als Basis, denn mehr als Vorlage für einzelne Aspekte dienen können.

Die Untersuchten 3-D Sound Systeme können als Vorlage und Orientierung für das Design und die Implementierung eines neuen Systems dienen, nicht jedoch als Basis einer Erweiterung.

5.3. Design Aspekte beim Entwurf eines 3-D Sound Systems

Zur Identifizierung der beim Entwurf zu berücksichtigenden Design Aspekte ist eine Betrachtung der Abläufe und Komponenten bei der Verarbeitung von 3-D Sound in Echtzeit notwendig.

Anhand einer Top – Down Analyse lassen sich, ausgehend von einer Gesamtsicht, einzelne Aspekte identifizieren und getrennt voneinander entwerfen. Über die detaillierte Sicht auf das System können sowohl die beteiligten Verarbeitungsschritte als auch die Kommunikations- und Datenverbindungen der einzelnen Aspekte abgeleitet werden, die die Basis für den Entwurf bilden.

5.3.1. Analyse 3-D Sound System

Ein 3-D Sound System verarbeitet externe Sounddaten unter Berücksichtigung externer Steuerparameter, um ein 3-D Klangbild wiederzugeben. (Siehe Abbildung 10) Die Wiedergabe der Sounddaten des finalen Klangbildes erfolgt durch ein geeignetes Sound Device. Die Anwendung, die die Soundverarbeitung beeinflusst sowie der Ursprung der Sounddaten sind nicht Teil des Systems.

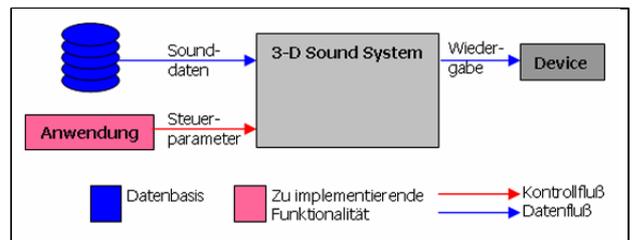


Abbildung 10 : Top – Level Sicht 3-D Sound System

Aus dem Anforderungskatalog lässt sich ableiten, dass eine Soundverarbeitung unter Berücksichtigung der in der Wiedergabeumgebung herrschenden physikalischen Gegebenheiten erfolgen soll. Diese während der Laufzeit fortlaufend über ein Tracking System aktualisierten Daten finden im 3-D Sound System ihre Repräsentation als physikalische Aspekte.

Als virtuelle Aspekte finden dann die zur Verarbeitung individueller Soundquellen notwendigen Komponenten ihre Repräsentation im System.

Abhängig von der zur Verfügung stehenden Hardware und den virtuellen und physikalischen Aspekten erfolgt die finale Aufbereitung der Daten zur Wiedergabe. Die bei der Aufbereitung zu berücksichtigenden Aspekte sind als Wiedergabeaspekte im 3-D Sound System repräsentiert. (Siehe Abbildung 11)

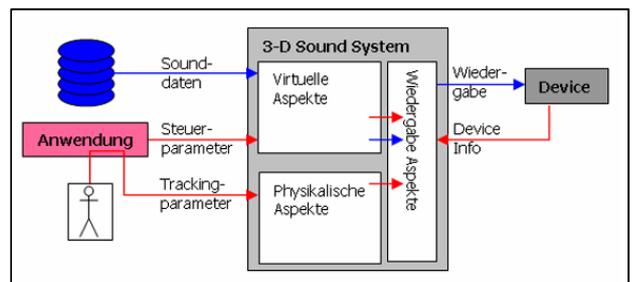


Abbildung 11 : Detaillierte Sicht 3-D Sound System

Die Detailanalyse für die einzelnen Aspekte erfolgt getrennt voneinander in den folgenden Abschnitten.

5.3.1.1. Analyse der Wiedergabe Aspekte

Die Wiedergabe der Sounddaten des finalen Klangbildes erfolgt über ein ausgesuchtes Device.

Die Berechnung des finalen Klangbildes erfolgt zum Einen abhängig von den unterstützten Formaten des Devices und zum Anderen abhängig von der zur Wiedergabe des Raumklangs gewählten Methode.

Während einzelne Effekte einer Soundquelle losgelöst von der zur Raumklang Wiedergabe gewählten Technik berechnet werden können, kann die Positionierung einer Soundquelle im Raum hingegen nur im Rahmen der finalen Aufbereitung der Sounddaten zur Wiedergabe erfolgen.

Diese als Wiedergabe – Modul kapselbare Einheit eines 3-D Sound Systems arbeitet auf vorverarbeiteten Sounddaten individueller Soundquellen unter Berücksichtigung der Positionsparameter dieser Quellen, den physikalischen Aspekten der Wiedergabe Umgebung sowie den Kapazitäten des genutzten Sound Devices.

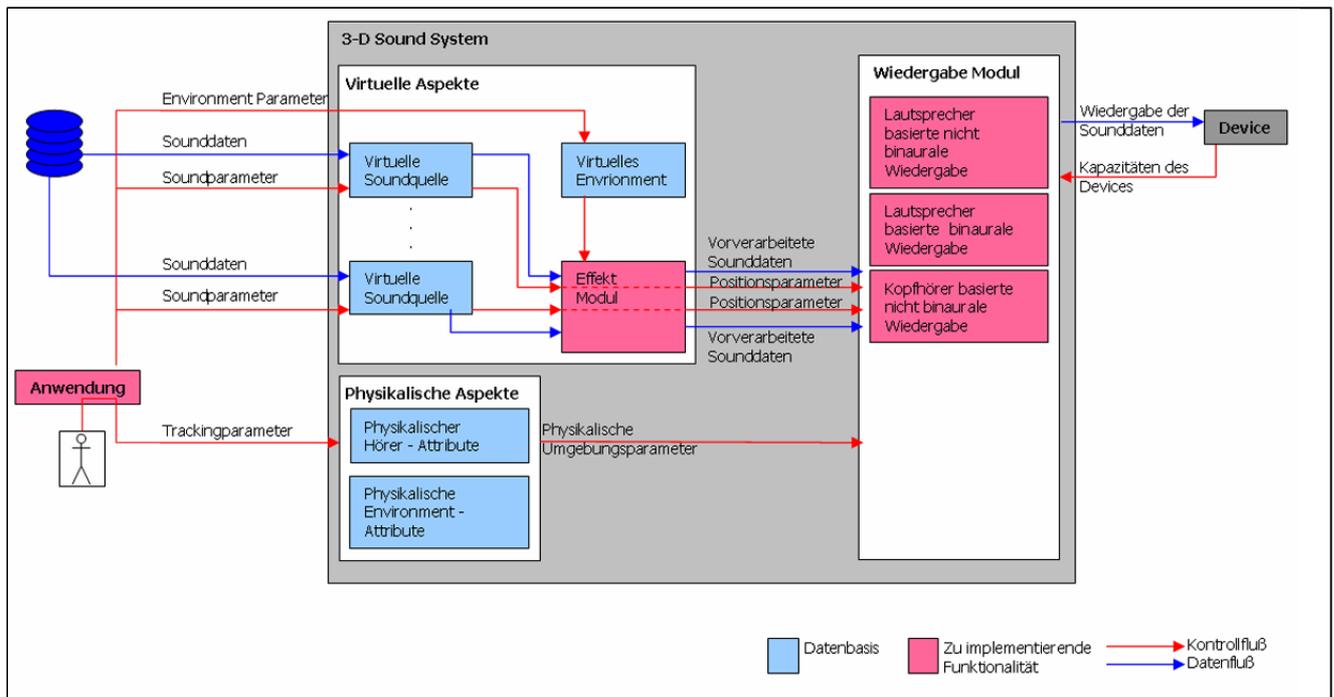


Abbildung 15 : 3-D Sound System im Überblick

Bei Entwurf eines Level 1 konformen Systems sind die Aspekte des virtuellen Environments nicht zu berücksichtigen.

Zusätzlich werden im Verlauf der Anwendung Informationen über nutzbare Sound Devices von Nöten sein, um eine Entscheidung darüber treffen zu können, ob und wie eine gewählte Wiedergabetechnik zum Einsatz kommen kann.

Eine zentrale Datenstruktur wird als Zugang zum 3-D Sound System einen einzigen Zugangspunkt darstellen, über den alle beteiligten Datenstrukturen verfügbar sind.

5.4.1.1. Datenstruktur Physikalisches Environment

Die Auswahl einer Methode zur Berechnung des 3-D Klangbildes, bedingt ein Datenfeld, das einen entsprechenden Eintrag einer Methode aufnehmen kann.

Abhängig von der gewählten Methode ist eine unterschiedliche Anzahl Lautsprecher notwendig. Neben der Anzahl der verwendeten Lautsprecher sind Datenstrukturen einzusetzen, die für jeden Lautsprecher individuell die Position und für das gesamte Lautsprecher Setup die Position des Hot Spot aufnehmen können. Die Angabe der Positionen erfolgt in Form kartesischer Koordinaten.

Werden binaurale Techniken zur Wiedergabe des 3-D Klangbildes verwendet, ist ein Satz von HRTF Daten notwendig. Das physikalische Environment muß, da es das Wiedergabe Setup beschreibt, auch eine Referenz auf eine HRTF Datenbasis in die Datenstruktur aufnehmen. (Siehe Abbildung 16)

5.4.1.2. Datenstruktur Physikalischer Hörer

Zur Anpassung der Wiedergabe an Hörer Positionen die vom Hot Spot abweichen, ist die Position des Hörers wie auch seine Orientierung in der Datenstruktur zur Repräsentation des physikalischen Hörers aufzunehmen. (Siehe Abbildung 16)

5.4.1.3. Datenstruktur Sound Device

Neben den Leistungsmerkmalen eines Sound Devices, wie die verfügbaren Quantisierungen und die nutzbaren Sample Raten sind Informationen in Datenstrukturen aufzunehmen, die eine Beurteilung erlauben, ob eine Methode zur 3-D Klangbild Wiedergabe Verwendung finden kann.

Es müssen weiter die zum Zugriff auf die Hardware notwendigen Strukturen aufgenommen werden die als Referenz auf das Device Verwendung finden. (Siehe Abbildung 16)

5.4.1.4. Datenstruktur Virtuelle Soundquelle

Konform zum Level 1 des Anforderungskataloges sind für die virtuellen Soundquellen lediglich deren Position in kartesischen Koordinaten, die Geschwindigkeit und die initiale Intensität in der Datenstruktur zu berücksichtigen.

Die Datenstruktur einer virtuellen Soundquelle wird eine Referenz auf eine beliebige Datenquelle, die die Sounddaten der Quelle bereitstellt, aufnehmen müssen. (Siehe Abbildung 16)

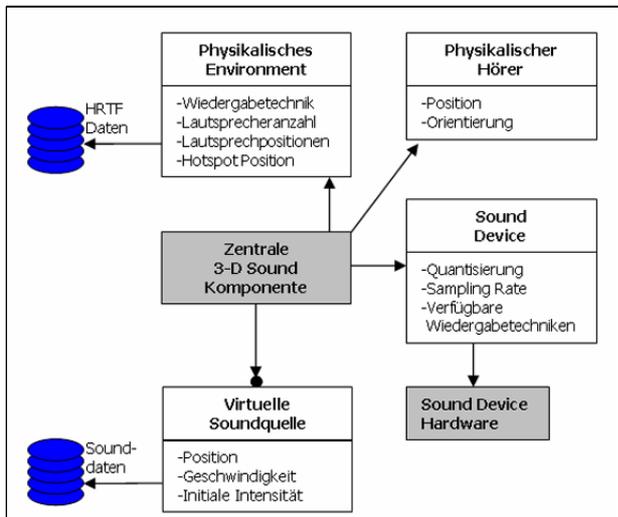


Abbildung 16 : Entwurf Datenstruktur

5.4.1.5. Zentrale Datenstruktur 3-D Sound System

Um einen zentralen Zugang auf die Datenstrukturen eines 3-D Sound Systems zur Verfügung zu stellen, sind Referenzen auf die einzelnen Datenaspekte in einer zentralen Datenstruktur aufzunehmen.

Je eine Referenz ist dabei notwendig für das physikalische Environment, den physikalischen Hörer und das für die Wiedergabe gewählte Device.

Zum Zugriff auf die virtuellen Soundquellen ist ein Feld mit Referenzen auf die Soundquellen in der zentralen Datenstruktur zu berücksichtigen. (Siehe Abbildung 16)

5.4.2. Entwurf der Verarbeitungsprozesse

Auf Basis der entworfenen Datenstrukturen und des Anforderungskataloges lassen sich die auf Ihnen zur Umsetzung der 3-D Sound Verarbeitung arbeitenden Verarbeitungsprozesse identifizieren, ihre Position im 3-D Sound System festlegen und ihre Aufgaben im Detail beschreiben.

Die Prozessbeschreibung mit der für individuelle Prozesse identifizierten Position im 3-D Sound System ermöglicht eine Gliederung der Prozesse in einer sequentiellen Reihenfolge, in Form einer Pipeline.

5.4.2.1. Identifizierung der Verarbeitungsprozesse

Die Datenstrukturen beschreiben die zu verarbeitenden Daten. Die Betrachtung der Datenstruktur führt über die Überlegung, wie die Daten zu verarbeiten sind, zur Identifizierung der notwendigen Prozesse.

Die Daten des physikalischen Environment müssen ausgelesen und ausgewertet werden, um in der Soundverarbeitung Berücksichtigung finden zu können. Die Auswertung muss zu jedem Zeitpunkt erfolgen, an dem sich einzelne Werte des physikalischen Environments ändern.

Das Auswerten der Daten erfolgt abhängig von der zur 3-D Klangbild – Wiedergabe gewählten Technik.

Basiert die Wiedergabe auf stereophonen Techniken, sind initial und bei einer Änderung der zum physikalischen Hörer relativen Lautsprecherpositionen die Ausrichtung der Lautsprecher neu zu bestimmen. Auf Basis der berechneten Ausrichtung der Lautsprecher ist die Ermittlung der Intensitätsanpassung für jeden beteiligten Kanal möglich.

Bei Nutzung binauraler Techniken ist zusätzlich zu der Bestimmung der Ausrichtung der Lautsprecher bzw. Kopfhörer eine Änderung der relativen Position von virtuellen Soundquellen zu berücksichtigen. Abhängig von der Position der virtuellen Soundquelle ist aus der HRTF Datenbasis der für den Kanal entsprechende HRTF Filter zu bestimmen und aus der Datenbasis zu lesen.

Der Verarbeitungsprozess lässt sich als Kanalfilter – Berechnung bezeichnen. (Siehe Abbildung 17)

Neben dem physikalischen Environment hat weiter auch der physikalische Hörer Einfluss auf die Berechnung des wiederzugebenden Klangbildes. Initial für eine Hörerposition im Hotspot eines Lautsprecher – Setups berechnet, muss die Wiedergabe an abweichende Positionen angepasst werden. Mit jedem Positionswechsel des Hörers ist eine Anpassung vorzunehmen. Die Anpassung erfolgt in Form einer Intensitätsanpassung wie in Abschnitt 2.2.1. beschrieben, oder einer Anpassung der Crosstalk Cancellation und der inversen HRTF Filterung für den Fall der Nutzung binauraler Techniken, wie in Abschnitt 2.2.2. beschrieben.

Eine Änderung der ursprünglich in der Verarbeitung vorausgesetzten Orientierung des physikalischen Hörers ist bei der Nutzung binauraler Techniken zu berücksichtigen. Mit der Änderung der Orientierung ist ein entsprechender HRTF Filter für beide Kanäle zu wählen und dieser durch ein Auslesen der HRTF Daten in die Datenstruktur zu übernehmen.

Dieser Anpassungsprozess kann als Wiedergabe – Anpassung bezeichnet werden. (Siehe Abbildung 17)

Um die Aktualität der Attribute der physikalischen Elemente zu gewährleisten, sind Prozesse zum Setzen der Attributwerte notwendig. Diese Prozesse sind zusammen mit der Verarbeitung von User Tracking – Daten nicht als Teile eines 3-D Sound Systems zu sehen, denn vielmehr extern umzusetzen.

Neben dem physikalischen Environment und dem physikalischen Hörer, als die 3-D Klangbild – Berechnung beeinflussenden Faktoren, bestimmen die in der Datenstruktur des Sound Devices aufgenommenen Randbedingungen des Hardware – Devices die entgültige Klangbild – Berechnung. Die Berechnung muss in einem durch das Device unterstützten Format erfolgen. Die Bestimmung der Device Attribute ist zu diesem Zweck zwingend notwendig. (Siehe Abbildung 17)

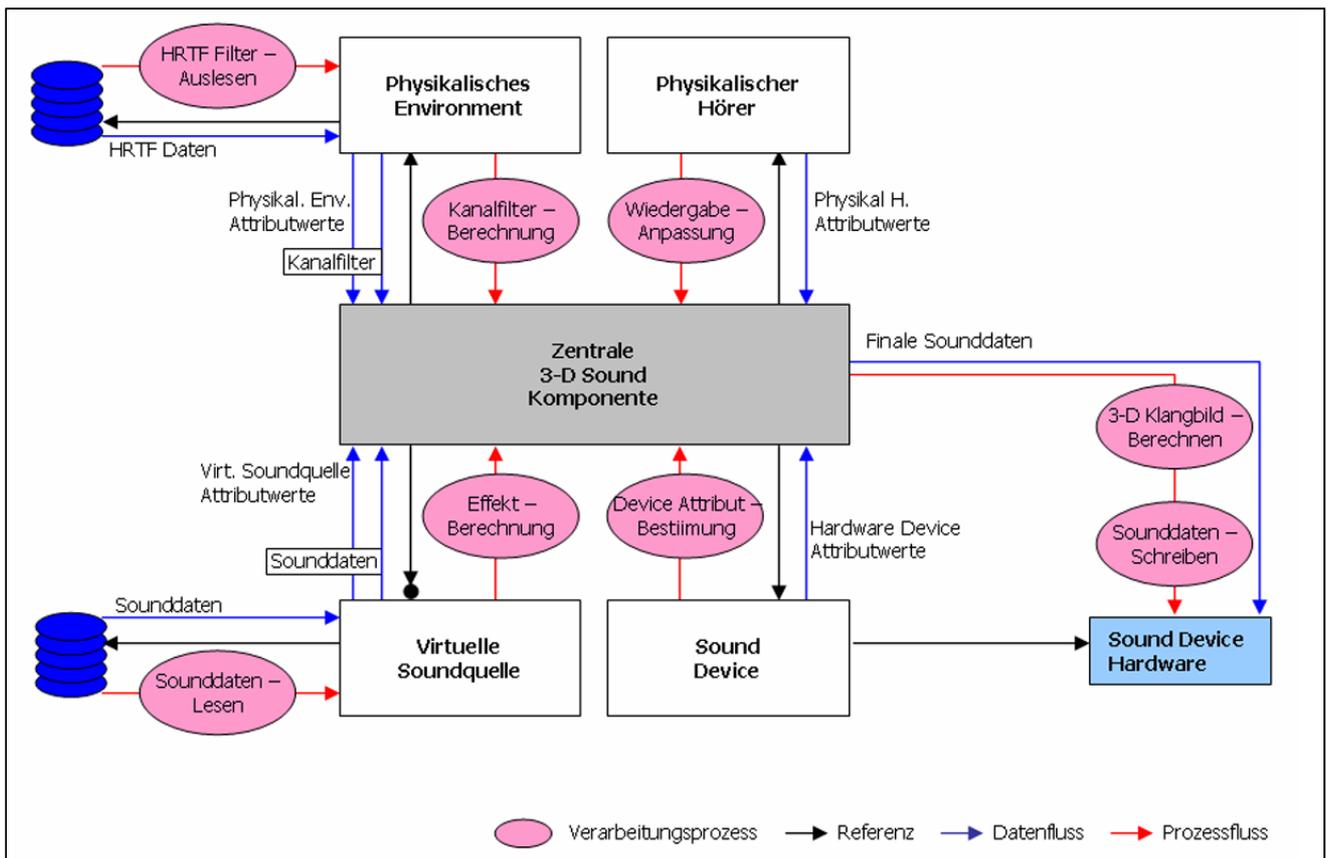


Abbildung 17 : Entwurf der Verarbeitungprozesse

Bilden die Attribute der physikalischen Elemente mit den Randbedingungen des Devices die beeinflussenden Parameter der Berechnung des 3-D Klangbildes, erfolgt diese auf Basis von Sounddaten der beteiligten virtuellen Soundquellen.

Durch die zentrale Komponente des 3-D Sound Systems koordiniert und getriggert, werden die Sounddaten aktualisiert. Unter Nutzung der mit der virtuellen Soundquelle gehaltenen Referenz auf die zu nutzenden Sounddaten, werden diese stückweise ausgelesen. Das Auslesen der Daten kann als Sounddaten – Lesen bezeichnet und in einem separaten Verarbeitungsprozess beschrieben werden.

Das Aktualisieren von Sounddaten geschieht mit jedem ausgelösten Trigger der zentralen Komponente und für jede der eingebundenen virtuellen Soundquellen. (Siehe Abbildung 17)

Die aktualisierten Sounddaten einer virtuellen Soundquelle dienen im weiteren Verlauf als Basis zur Berechnung von Effekten. Bei der Effektberechnung werden für den Level 1 des Anforderungskataloges ausschließlich Soundeffekte auf Basis der Position und der Geschwindigkeit berechnet. Die dazu notwendigen Attributwerte stehen in der Datenstruktur der zugehörigen virtuellen Soundquelle zur Verfügung und sind mit dem Verarbeitungsprozess der Effektberechnung auszulesen. (Siehe Abbildung 17)

Die aus der Effektberechnung resultierenden gefilterten Sounddaten werden im Verarbeitungsprozess der 3-D Klangbild – Berechnung unter Berücksichtigung der

räumlichen Attribute der virtuellen Soundquelle wie auch der Werte des physikalischen Environment und – Hörer Attribute weiterverarbeitet.

In der 3-D Klangbild Berechnung erfolgt die Berechnung der finalen Sounddaten zur Wiedergabe. In diesem Prozess werden die Sounddaten unter Nutzung des berechneten Kanalfilters zur Wiedergabe über die individuellen Kanäle aufbereitet, so dass bei der Wiedergabe ein Raumklang entsteht. Die kanalgefilterten Daten werden dann mit denen in der Wiedergabeanpassung berechneten Werte gefiltert und zur Wiedergabe verwendet. (Siehe Abbildung 17)

Die Wiedergabe erfolgt durch das verwendete Sound Device durch eine digital – analog Wandlung. Die digitalen Daten, das Ergebnis der 3-D Klangbild – Berechnung, werden im letzten Verarbeitungsprozess unter Nutzung der in der Datenstruktur des Sound Devices gehaltenen Referenz auf den Datenkanal des Devices geschrieben. (Siehe Abbildung 17)

5.4.2.2. 3-D Sound Verarbeitungs – Pipeline

Die Verarbeitungsprozesse können aufgrund ihrer zeitlichen und inhaltlichen Zusammenhänge in einer sequentiellen Verarbeitungspipeline angeordnet werden.

Die Verarbeitungspipeline definiert die Reihenfolge soundverarbeitender Prozesse zur Erzeugung und Wiedergabe von 3-D Klangbildern. Die einzelnen Prozesse arbeiten in der Pipeline auf Basis der als

Resultat aus dem vorangegangenen Prozess hervorgehenden Daten. Der in der Verarbeitung an erster Stelle stehende Prozess arbeitet im Unterschied zu den nachfolgenden Prozessen auf Basis kontinuierlich aktualisiert von außen zugelieferter, variierender Sounddaten. (Siehe Abbildung 18)

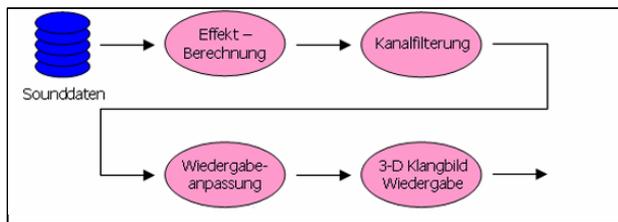


Abbildung 18 : Verarbeitungspipeline

Neben denen aus vorangegangenen Prozessen hervorgehenden Daten arbeiten einzelne Verarbeitungsschritte teilweise auf zusätzlichen externen Daten, die keine Sounddaten beschreiben, dennoch aber zur Verarbeitung unabdingbar sind. (Siehe Abbildung 19)

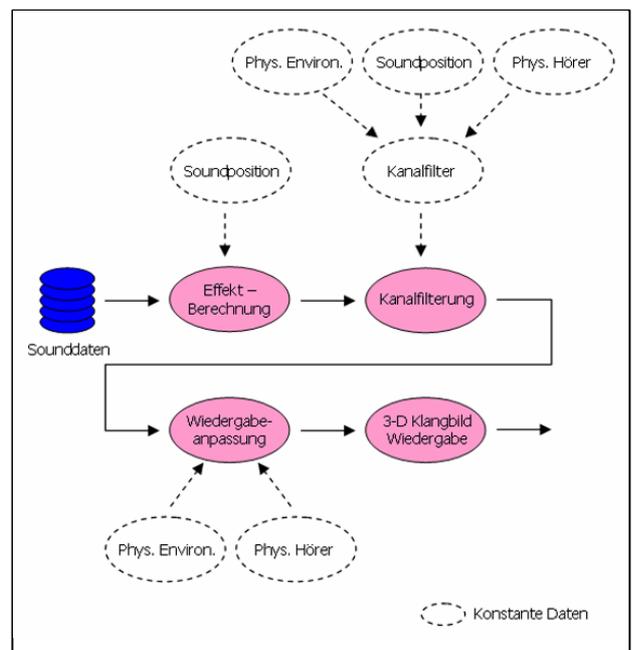


Abbildung 19 : Verarbeitungspipeline mit externem konstantem Einfluß

Bei den zusätzlichen externen Daten handelt es sich im Gegensatz zu den Sounddaten nicht um einen Datenstrom, denn mehr um Daten mit einer Aktualität, die über einen endlichen Zeitraum hinweg fortbesteht. In nicht definierten zeitlichen Abständen kommt es durch äußere Impulse zu einer Aktualisierung der externen Daten. Die externen Daten können für einen individuellen Pipelinedurchlauf als konstant angenommen werden.

Zu den externen Daten zählen sowohl die das physikalische Environment beschreibenden Attribute als auch die physikalischen Attribute des Hörers sowie die Attribute, die die räumliche Anordnung einer virtuellen Soundquelle beschreiben.

Die sich durch weitere Verarbeitungsschritte aus den externen Daten ergebenden Größen können in diesem Zusammenhang ebenso als konstant angenommen werden.

5.4.3. Entwurf der Schichtstruktur

Soll dem entwickelten 3-D Sound System ein flexibles Einsatzgebiet zuteil werden, ist eine Schichtstruktur bei der Implementierung umzusetzen, die es ermöglicht, das 3-D Sound System auf verschiedenen Ebenen zu nutzen.

Es muss Entwicklern möglich sein, die zur direkten Einbindung in eine Entwicklungsumgebung gebotenen Funktionalitäten und Datenstrukturen zu nutzen, um einen möglichst schnellen und einfachen Zugang zur Materie 3-D Sound zu bekommen.

Entwicklern, die Multimedia Anwendung in einer Entwicklungsumgebung erstellen, für die keine direkte Einbindungsstrategie umgesetzt wurde, muss die Möglichkeit eingeräumt werden, das 3-D Sound System direkt zu nutzen. Eine direkte Nutzung kommt ohne die zur vereinfachten Einbindung in Entwicklungsumgebungen entwickelten Funktionen und Datenkonstrukte aus.

Neben der umgesetzten 3-D Sound Funktionalität ist für eine Auswahl an Multimedia Anwendung ein allgemeines Soundfeedback ohne Richtungscharakteristik ausreichend. Entwicklern solcher Anwendungen muss ein Zugang zu einer Basisschicht im System geboten werden, die neben einem Sound System keine darüber hinausgehenden Funktionen umsetzt. Ein einfaches Sound System bietet Funktionalitäten zur kontinuierlichen Wiedergabe von Sounddaten über ein Standard Sound Device, ohne eine räumliche Positionierung oder Soundeffekte zu berechnen.

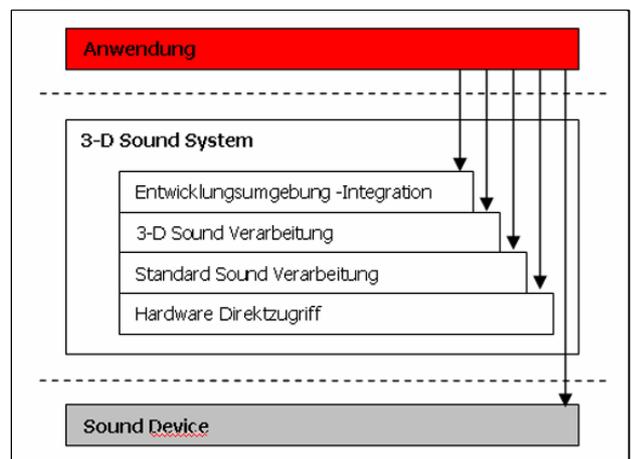


Abbildung 20 : Schichtstruktur

Anwendungen, die aus Performancegründen eigene Strategien zur Sounddaten – Verarbeitung implementieren, nutzen unter Umständen lediglich den im 3-D Sound System implementierten Zugang zum Sound Device. Für solche Anwendungstypen muss eine Schnittstelle existieren, die ein direktes Schreiben von

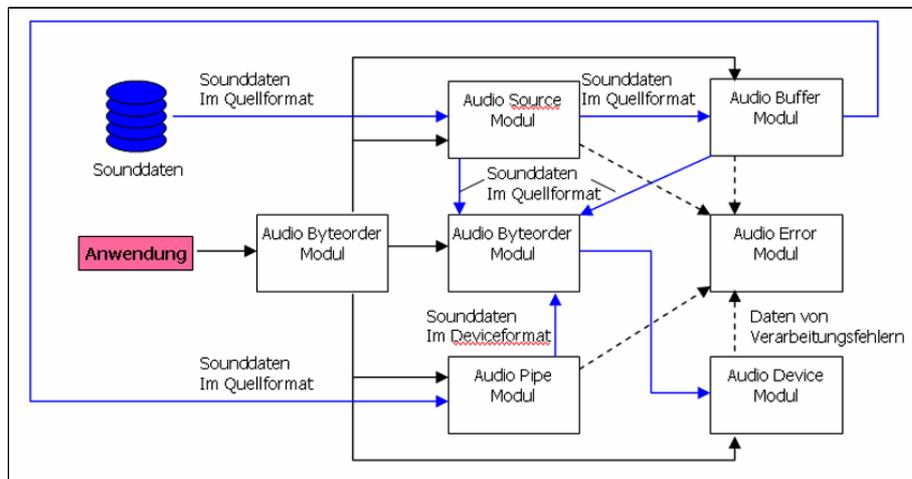


Abbildung 21 : Modulübersicht

Sounddaten auf den Datenkanal der Sound Hardware ermöglicht.

Aus den wechselnden Ansprüchen ergibt sich eine breite Schnittstelle zum 3-D Sound System, die einen Einstieg auf verschiedenen Schichten ermöglicht. (Siehe Abbildung 20)

6. Implementierung eines interaktiven 3-D Sound Systems

Die vorangegangene Analyse stellt mit dem Anforderungskatalog und dem Entwurf die Voraussetzungen für die Implementierung eines 3-D Sound Systems bereit. Anhand des Entwurfes der Datenstrukturen und der auf ihnen arbeitenden Verarbeitungsprozesse, können modulare Pakete aus ihnen geschnürt werden, die die Zielvorgaben implementieren und der entworfenen Schichtstruktur genügen.

Zur Implementierung eines stabilen 3-D Sound System wird entsprechend dem Entwurf mit der Open Sound Library (OpenSL) ein System umgesetzt, das dem Level 1 des Anforderungskataloges entspricht. Die Umsetzung erfolgt mit Hilfe von Techniken, durch die eine größtmögliche Betriebssystemunabhängigkeit erreicht werden kann.

Die Sourcecode Dateien, wie auch die Sourcecode Dokumentation sind als Anlage auf dem beigefügten Datenträger verfügbar.

Die Wahl der Technik und der Werkzeuge zur Implementierung bilden die Basis der Umsetzung.

Mit den gewählten Techniken und Methoden ist die Umsetzung der gewünschten Pakete möglich. Dazu werden die Inhalte und Zusammenhänge der Pakete als Modul im Detail beschrieben, bevor die Implementierung für die 3-D Sound Verarbeitung besonders einflussreiche Komponenten beschrieben wird.

6.1. Methoden und Techniken der Implementierung

Mit der Wahl der zu verwendenden Techniken wird der Grundstein zur Erreichung einzelner Teilziele gelegt.

Um eine performante Verarbeitung zu ermöglichen, wird zur Entwicklung mit „C“ eine hardwarenahe Programmiersprache verwendet, die einen schnellen Zugriff auf Betriebssystemroutinen und im System installierte Hardware ermöglicht sowie über einen hohen Verbreitungsgrad verfügt.

Eine als ANSI konform entwickelte Anwendung stellt sicher, dass bei der Umsetzung nur standardisierte Spracherweiterungen der Programmiersprache zur Verwendung kommen.

Als POSIX konform entwickelte Bibliothek nutzt diese nur Funktionen die für alle POSIX konformen Betriebssysteme definiert sind. Damit wird eine Lauffähigkeit oder zumindest einfache Portierung auf wechselnde Betriebssysteme ermöglicht.

Kommen ANSI und POSIC konforme Entwicklung bei der Umsetzung des 3-D Sound Systems zur Anwendung, sind diese Maßstäbe auch an eingebundene Bibliotheken zu legen.

Das Maß an Betriebssystem spezifischen Abhängigkeiten kann so auf den kleinen Teil des Hardware Zugriffs begrenzt bleiben. Die für den Zugriff auf die Hardware zu nutzende Implementierung des Treibers ist nur in Ausnahmefällen als Betriebssystem unabhängig zu erachten. Dieser Abhängigkeit ist mit einer Anpassung an wechselnde Betriebssystem zu begegnen.

6.2. Modularisierung

Die entworfenen Datenaspekte und Verarbeitungsprozesse ermöglichen eine Kapselung von inhaltlich verbundenen Themengebieten in Modulen.

Als Module von OpenSL können ein Audio Source Modul, ein Audio Buffer Modul, ein Audio Pipe Modul, ein Audio Device Modul und ein Audio Byteorder Modul, als auch ein Modul zur Fehlerbehandlung, das Audio Error Modul identifiziert werden. (Siehe Abbildung 21)

Das Audio Source Modul implementiert das Auslesen von Sounddaten aus der Datenbasis. Die durch das Audio Source Modul bereitgestellten Daten finden Verwendung im Audio Buffer Modul, in dem die Weiterverarbeitung der Daten in Form einer Effektverarbeitung erfolgt. Die Effekt bearbeiteten Daten einer Vielzahl von Audio Buffer werden im Audio Pipe Modul in ein Format gemischt, wie es die gewählte 3-D Wiedergabe – Methode verlangt. Es werden die Kanalfilter passend zu den Attributen der virtuellen Soundquellen angewendet und es findet die Anpassung der Wiedergabe auf Basis der aktuellen Hörerposition statt. Zur endgültigen Wiedergabe der gemischten und gefilterten Daten die aus dem Audio Pipe Modul resultieren wird das Audio Device Modul und für den Fall, dass die Hardware eine abweichende Byte Order verlangt, ein weiteres Modul Audio Byteorder zur Anpassung der Byteorder genutzt. Zur Ablaufsteuerung und als zentrale Komponente dient ein Audio Processing Modul.

Auftretende Fehler werden im Modul Audio Error protokolliert und können zur Laufzeit abgefragt werden. Auf Basis der Inhalte und Zusammenhänge der identifizierten Module kann eine Detailbeschreibung der einzelnen Module erfolgen, die die Implementierung im Detail erläutert.

6.2.1. Audio Source Modul

Das Audio Source Modul dient als Datenfabrik von OpenSL. Es liest Daten aus einer Datenbasis definierten Typs oder generiert Daten, für den Fall, dass es sich um einen generischen Datentyp handelt. Als generische Typen stehen Sinustöne bestimmter Frequenz wie auch weißes Rauschen zur Verfügung.

Die zum Lesen von Daten aus einer Audio Datei notwendigen Vorbereitungsschritte, wie das initiale Öffnen und abschließende Schließen der Datei ist mit jeder Instanz des Audio Source Moduls genau einmal auszuführen. Das anschließende Lesen von Daten erfolgt ohne nennenswerte Verzögerung. Lediglich die Repositionierung der Datei – Leseposition führt zu einer geringen Latenz.

6.2.1.1. Datenstrukturen und Funktionen

Das Audio Source Module ist implementiert in den Dateien `slAudioSource.h` und `slAudioSource.c`. Das Modul nutzt OpenSL Funktionen und Strukturen der Dateien `slAudioError.h`, `slAudioError.c` zur Fehlerbehandlung sowie die Datenstrukturen aus `slAudioTypes.h`.

Die vom Audio Source Modul benötigten Daten sind in einem Struct organisiert. Sie umfassen neben dem Typ der Datenquelle eine Referenz auf diese sowie Hilfsstrukturen, die das Lesen der Daten, bzw. ein Erzeugen von Daten ermöglichen (Siehe Abbildung 22).

Type	Name	Info
SLaudioSourceType	sourceType	Typ der Soundquelle (WAV File, AIFF File, Tone, Noise)
void*	sourceDataHandle	Referenz auf die Datenbasis
off_t	sourceSize	Aktuelle Leseposition, Aktueller Parameter bei der Generierung
SLsampleFormat	sourceSampleFormat	Sample Format des Audio Files (8bit (un)signed, 16bit (un)signed, 32bit signed)
SLsampleRate	sourceFrequency	Sampling Rate des Audio Files (8000, 11025, 16000, 22050, 32000, 44100, 48000, 96000)
SLchannelSetup	sourceChannelSetup	Channel Setup des Audio Files (Mono, Stereo)
int	sourceUpdateInProgress	Synchronisation des Updatevorgangs
char*	sourceFilename	Filename des Audio Files

Abbildung 22 : Auszug der Datenstruktur des Audio Source Modul

Das Modul stellt Funktionen zum Erzeugen und Zerstören von Audio Source Konstrukten zur Verfügung sowie eine Update Funktion, die für eine angegebene Audio Source Instanz aktuelle Daten liefert. (Siehe Abbildung 23)

Return Type	Name	Parameter
SLaudioSource *	slSourceCreateAudioSource	(SLdevice *device, SLaudioSourceType type, void *data)
void	slSourceDestroyAudioSource	(SLaudioSource *source)
void *	slSourceUpdateAudioSource	(SLaudioBuffer *buffer, int finalize)

Abbildung 23 : Auszug der Funktionen des Audio Source Modul

Von besonderem Interesse im Audio Source Modul ist vor allem die Update Funktion, die aktuelle Sounddaten liefert.

6.2.1.2. Aktualisierung von Audio Source Daten

Abhängig vom Typ der Audio Source, möglich sind statische Typen auf Basis von Audio Dateien und generische Typen als Ton oder Rauschen, erfolgt die Aktualisierung von Daten. Sie wird aufgerufen mit einer Instanz des Audio Buffer Moduls als Parameter.

Bei Datei basierten Typen wird über die Referenz auf die Datei, die schon mit der Erzeugung einer Audio Source Modul Instanz geöffnet wurde, auf die Sounddaten zugegriffen. Das Auslesen der Datei geschieht unter Nutzung von Funktionen der Bibliothek „libsndfile“, die Dateityp- und Betriebssystem unabhängige Routinen zum Öffnen und Lesen von Dateien implementiert.

Für generische Typen werden auf Basis des bei der Erzeugung übergebenen Parameters Daten generiert. Für den Noise Typ werden unter Nutzung der Random Funktion Soundsamples generiert. Eine Sinusschwingung einer angegebenen Frequenz ist Grundlage der Generierung bei Ton – Quellen.

Die Anzahl wie auch das Format zu generierender oder aus einer Datei zu lesender Soundsamples bestimmt sich über das Audio Device, auf das über die als Parameter übergebene Instanz des Audio Buffer Moduls zugegriffen werden kann. Für das Device ist die Größe des Device – Buffers angegeben, der beim Schreiben von Daten auf das Device verwendet wird. Die Device – Buffergröße gibt an, wie viele Samples pro Schreibvorgang vom Speicher auf den Datenkanal des Devices geschrieben werden. Die Größe des Buffers, der im letzten Schritt der Verarbeitungspipeline, der Wiedergabe, Verwendung findet bestimmt die Größe

der in jedem Schritt der Pipeline durch die einzelnen Prozesse zu verarbeitenden Datenmengen. Das Öffnen eines Devices zur Wiedergabe von Sounddaten, geschieht unter Festlegung auf eine Sampling Rate sowie ein Sample Format. Nach dem Öffnen eines Device ist eine Änderung der Sampling Rate oder des Sample Formates nicht mehr möglich.

Bei der Aktualisierung der generischen Audio Source Typen werden die zu erzeugenden Daten im entsprechenden Format mit der festgelegten Sampling Rate generiert. Liegen die Daten der Audio Datei nicht im Format und der Sampling Rate vor, wie es durch das Device festgelegt ist, erfolgt mit dem Auslesen der Daten aus der Datei ein Resampling.

Beim Resampling erfolgt eine Anpassung des Formates durch eine Normalisierung der Daten über den durch das Audio Device Format festgelegten Wertebereich. Bei der Anpassung der Sampling Rate ergibt sich aus dem Quotienten aus Audio Datei Sampling Rate und Audio Device Sampling Rate ein Resamplingfaktor.

$$\text{Resamplingfaktor} = \frac{\text{Sampling Rate AudioDatei}}{\text{Sampling Rate AudioDevice}}$$

Ergibt sich ein Resamplingfaktor < 1 , so liegt die Audio Datei mit einer kleineren Sampling Rate vor, als die Audio Device Sampling Rate. Der umgekehrte Fall liegt vor für einen Faktor > 1 .

Ist die Audio Datei Sampling Rate kleiner als die des Devices, so ist durch eine Vervielfältigung einzelner Sample Werte eine Anpassung der Sampling Rate durchzuführen. Dabei wird der Kehrwert des Resamplingfaktors gebildet. Der ganzzahlige Anteil dessen ergibt den Faktor der Sample – Vervielfältigung. Ergibt sich zum Beispiel ein Resamplingfaktor von 0,5, bedeutet dies eine kleinere Sampling Rate der Audio Datei.

Der Kehrwert von 2 bestimmt, dass einzeln gelesene Samples der Audio Datei 2 mal in den Ausgabestream zu schreiben sind.

Auftretende gebrochene Anteile eines Resamplingfaktors werden hierbei mit jedem aus der Audio Datei gelesenen Sample summiert, bis die Summe der gebrochenen Anteile 1 überschreitet. Ist dies der Fall, ist das gelesene Sample einmal mehr zu schreiben, als es der Kehrwert vorgibt.

Bei einem Kehrwert von 2,2 werden pro gelesenen Sample 2 Samples geschrieben. Beim fünften Lesevorgang jedoch, werden 3 Samples geschrieben.

Bei einer im Vergleich zur Sampling Rate des Audio Devices größeren Sampling Rate des Audio Files, ist die Anpassung durch Reduzierung der Werte durchzuführen. Der Resamplingfaktor gibt dabei an, um welchen Faktor die Reduzierung zu erfolgen hat.

Aus Performancegründen wird in diesem Fall eine stark vereinfachte Methode zur Wertereduktion verwendet.

Bei dieser Methode gibt der Resamplingfaktor eine Schrittweite an. Die Schrittweite beschreibt den Abstand zwischen 2 Samplewerten aus der Quelldatei, die in die Ausgabedaten geschrieben werden.

So ergibt ein Resamplingfaktor von 2, dass nur jedes 2. Sample der Datei in die Ausgabedaten geschrieben wird. Bei gebrochenen Resamplingfaktoren wird nach dem Verfahren des Nearest Neighbour das Sample verwendet, dass sich am dichtesten an der errechneten Schrittweite befindet. Bei einer Schrittweite von 2,7 würde somit das 3. Sample und nicht das 2. in die Ausgabedaten geschrieben. Andere Verfahren, die auf Basis von Mittelwerten von einer Menge von Samples arbeiten, können aufgrund längerer Laufzeiten hier nicht zum Einsatz kommen.

Der mit dem vereinfachten Verfahren auftretende Informationsverlust, ist bei der Soundverarbeitung als geringer einzuschätzen als der durch die Vereinfachung erreichte Performancegewinn.

Nach dem Theorem von Shannon ist für die höchste Signalfrequenz eine zweimalige Abtastung erforderlich, damit eine Schwingung erhalten bleibt und keine Störfrequenzen entstehen. [STO_95] Somit ist auch bei einem Resamplingfaktor von 12, der sich ergibt aus der größten unterstützten Sampling Rate von 96000 Hz und der kleinsten unterstützten Sampling Rate von 8000 Hz, die Verarbeitung aller durch das menschliche Gehör optimal wahrnehmbaren Signalfrequenzen möglich. Mit einer bei einer Sampling Rate von 8000 Hz maximal kodierbaren Signalfrequenz von 4000 Hz wird eine Signalfrequenz erreicht, die sich am oberen Rand des optimalen Hörbereichs befindet.

Die durch das Audio Source Modul implementierte Funktionalität findet im Audio Buffer Modul Verwendung. Aus dem Audio Buffer Modul wird in regelmäßigen zeitlichen Abständen die Update Funktion des Audio Source Moduls aufgerufen.

Es ist möglich, dass mehrere verschiedene Instanzen des Audio Buffer Moduls die gleiche Instanz des Audio Source Moduls verwenden.

6.2.1.3. Synchronisation des Audio Source Moduls

Um Zugriffsverletzungen bei der Aktualisierung zu vermeiden, ist die Update Funktion innerhalb des Audio Source Moduls über die Verwendung eines binären Semaphors, das als „sourceUpdateInProgress“ bezeichnetes Flag Teil der Datenstruktur ist, abgesichert. Somit ist sichergestellt, dass ein gleichzeitiger Aufruf der Update Funktion durch mehrere verschiedene Instanzen des Audio Buffer Moduls der zu Inkonsistenten Datenständen oder asynchronen Datenzugriffen führen kann, nicht möglich ist.

6.2.1.4. Ein- und Ausgabedaten

Unter Verwendung der implementierten Funktionen des Audio Source Moduls werden so Sounddaten der Eingabe, die in einem beliebigen Quellformat vorliegen, gelesen oder generiert und als Sounddaten für die Ausgabe produziert .

6.2.2. Audio Buffer Modul

Das Audio Buffer Modul existiert losgelöst vom Audio Source Modul, da Situationen denkbar sind, in denen mehrere virtuelle Soundquellen existieren, die auf der gleichen Basis von Sounddaten basieren. Die Entkopplung beider Module ermöglicht die Vermeidung redundanter Datenhaltung und eine performante Verarbeitung durch mehrfach von verschiedenen Audio Buffer Instanzen referenzierte Audio Source Instanzen.

6.2.2.1. Datenstrukturen und Funktionen

Das Audio Buffer Modul ist implementiert in den Dateien slAudioBuffer.h und slAudioBuffer.c.

Es nutzt OpenSL Funktionen und Strukturen der Dateien slAudioError.h, slAudioError.c zur Fehlerbehandlung, slAudioSource.c zur Aktualisierung von Sounddaten sowie die Datenstrukturen aus slAudioTypes.h

Das die im Audio Buffer Modul benötigte Daten umfassende Datenkonstrukt beinhaltet den Typ des Buffers - ob es sich um einen positionierten oder nicht positionierten Buffer handelt -, eine Referenz auf eine Audio Source, einen zyklischen Buffer zur Kopplung der Ausgabedaten mit nachfolgenden Modulen, Kontrollstrukturen für die Wiedergabe sowie Angaben zur räumlichen Anordnung des Audio Buffers. (Siehe Abbildung 24).

Type	Name	Info
SLAudioSource*	bufferSource	Referenz auf eine Instanz des Audio Source Moduls
SLdevice*	bufferDevice	Referenz auf eine Instanz des Audio Device Moduls, des genutzten Audio Devices
void*	bufferData[SL_INTERNAL_BUFFER_FACTOR]	Zyklischer Kopplungspeicher
short	bufferActDataPoint	Aktuelle Position im zyklischen Speicher
SLAudioBufferState	bufferState	Aktueller Status des Buffer (PLAY, STOP, PAUSE)
SLAudioBufferType	bufferType	Typ des Buffer (POSITIONED, UNPOSITIONED)
SLAudioBufferSoundEffect	bufferSoundEffect	Bitmaske der aktiven Effekte (ALL, NONE, DOPPLER)
int	bufferLoopFlag	Fortlaufende Wiederholte Wiedergabe der Audio Source ?
float	bufferGainInit	Initiale Intensität
int	bufferUpdateInProgress	Synchronisation des Updatevorgangs
float	bufferPosition[3]	Position des Buffers rel. zum virt. Hörer
float	bufferVelocity	Geschwindigkeit des Buffers
float	bufferAttenuationFactor	Faktor der Intensitätsabschwächung
float	bufferReferenceDistance	Referenz Entfernung für Berechnung der Intensitätsabschwächung über die Distanz
float	bufferFalloffDistance	Distanz bei der Buffer nicht mehr zu hören

Abbildung 24 : Auszug der Datenstruktur des Audio Buffer Modul

Das Modul stellt Funktionen zum Erzeugen und Zerstören von Audio Buffer Konstrukten zur Verfügung sowie eine Update Funktion, die für eine angegebene Audio Buffer Instanz aktuelle Daten liefert. Die Aktualisierungsroutine des Audio Buffer Moduls ruft intern die zur Effektverarbeitung notwendigen Funktionen auf.

Return Type	Name	Parameter
SLAudioBuffer *	slBufferCreateAudioBuffer	(SLAudioSource *source, SLdevice *device)
void	slBufferDestroyAudioBuffer	(SLAudioBuffer *buffer)
void *	slBufferUpdateAudioBuffer	(void *data, int finalize)

Abbildung 25 : Auszug der Funktionen des Audio Buffer Modul

Aufgrund des geringen Leistungsumfanges – es ist für den Level 1 des Anforderungskataloges ausschließlich die Intensitätsanpassung über die Entfernung zu

berechnen – bleibt die Effektverarbeitung Bestandteil des Audio Buffer Moduls und wird nicht als isoliertes Effekt Modul implementiert. (Siehe Abbildung 25)

Zu betrachten ist im Audio Buffer Modul vor allem die Effektverarbeitung innerhalb der Update Funktion wie auch die Kopplung zu nachfolgenden Modulen über zyklische Speicher.

6.2.2.2. Anpassung der Intensität über die Distanz

Die Effektverarbeitung erfolgt im Rahmen der Update Funktion des Buffers. Nach einem Aufruf der Update Funktion der durch den Audio Buffer referenzierten Audio Source, der einen aktualisierten Satz an Sounddaten aus der Sound – Datenbasis liefert, erfolgt der Aufruf der zu berechnenden Effektroutinen. Diese arbeiten auf Basis der vom Audio Source Modul gelieferten Daten und der Raumattribute der Instanz des Audio Buffer Moduls.

Bei der Intensitätsanpassung finden die über die Position des Audio Buffers, relativ zum virtuellen Hörer, ermittelbare Entfernung sowie weitere teilweise für die Berechnung notwendige und für wechselnde Methoden der Entfernungsanpassung verschiedener Parameter Berücksichtigung.

Es existieren verschiedene Ansätze zur Anpassung der Intensität von entfernten Soundquellen. Sie unterscheiden sich durch die erzielbare Realitätsnähe und die Performance bei der Berechnung.

Untersuchungen haben bewiesen, dass eine Verdopplung der Entfernung einer ungerichteten Soundquelle, gemessen an einer Referenzentfernung, mit einer Schalldruckhalberung oder einer Schalldruckpegelverminderung 6 dB zur Einschätzung der Entfernung einer Soundquelle, dem Hörer die genauesten Entfernungsabschätzung ermöglicht. Dieses Verhalten ist beschrieben durch das Inverse Square Law. [BEG_94]

Der Schalldruckpegel eines Schalldrucks lässt sich berechnen über das Verhältnis des gemessenen Schalldrucks zum Schalldruck der Hörschwelle. (Siehe Abschnitt 2.1.1.4.)

$$\text{Schalldruckpegel} = 20 \log \left(\frac{p}{2 \cdot 10^{-5}} \right) \text{ db SPL}$$

Setzt man den an einer Messstelle gemessenen Schalldruck zu dem in einer Referenzentfernung herrschenden Schalldruck, bei der Berechnung des Schalldruckpegels in Bezug, lässt sich der Schalldruckpegelunterschied berechnen.

$$\Delta \text{Schalldruckpegel} = 20 \log \left(\frac{p}{p_{\text{referenz}}} \right) \text{ db SPL}$$

Das Verhältnis aus dem Schalldruck, gemessen in einer Referenzdistanz zu dem in einer Messdistanz gemessenen Schalldruck verhält sich wie das umgekehrte Verhältnis der Entfernungen zueinander. Der Schalldruckpegel lässt sich über die Entfernungsänderung beschreiben mit :

$$\Delta \text{Schalldruckpegel} = 20 \log \left(\frac{\text{Distanz}_{\text{referenz}}}{\text{Distanz}} \right) \text{ db SPL}$$

Da Schalldruck als Beschreibung der subjektiven Empfindung wenig geeignet ist, lässt sich die Entwicklung des Schalldrucks auch über die vom Schalldruck abhängige Lautstärke beschreiben. Da jedoch eine Verdopplung der Lautstärke einer Zunahme des Lautstärkepegels um 10 phon entspricht, ergibt sich die Lautstärkeänderung zu:

$$\Delta \text{Lautstärkepegel} = 30 \log \left(\frac{\text{Distanz}_{\text{referenz}}}{\text{Distanz}} \right) \text{ db SPL}$$

Neben denen auf Basis des Schalldrucks und der Lautstärke arbeitenden Ansätze können stark vereinfachte Methoden zum Einsatz kommen, die ein weniger realistisches Entferungsverhalten simulieren, dabei jedoch dem Anwender ein größeren Maß an Einflussmöglichkeiten einräumen.

Eine lineare Abschwächung der Amplitude über die Distanz ermöglicht dem Entwickler neben der Angabe einer Referenzdistanz die Angabe einer Falloff Distanz. Bis zu einer durch die Referenzdistanz festgelegten Entfernung wird keine Anpassung der Amplitude vorgenommen. Bei Entfernungen größer als die Referenzdistanz, kleiner als die Falloff Distanz, erfolgt eine lineare Anpassung der Amplitude bis zu einer Amplitude von 0 beim Erreichen der Falloff Distanz.

Die Angabe eines Abschwächungsfaktors im Rahmen einer weiteren Methode erlaubt eine Abschwächung der Amplitude um den Abschwächungsfaktors mit der Zunahme der Entfernung.

Für jede der verfügbaren Methoden ist die Angabe von minimalen und maximalen Werten für die Amplitude möglich. Bei Angabe von Werten im Bereich 0...1, erfolgt ein Clipping der aus der Berechnung resultierenden Amplitudenwerte gegen den angegebenen Min-Max Bereich. Das Clipping dient als weiterer Einflussfaktor des Entwicklers.

6.2.2.3. Synchronisation des Audio Buffer Moduls

Um Zugriffsverletzungen bei der Aktualisierung zu vermeiden, ist die Update Funktion innerhalb des Audio Buffer Moduls über die Verwendung eines binären Semaphors, das als „bufferUpdateInProgress“ bezeichnetes Flag Teil der Datenstruktur ist, abgesichert.

6.2.2.4. Ein- und Ausgabedaten

Das Audio Buffer Modul nutzt die durch ein Audio Source Modul bereitgestellten Daten. Diese als Rohdaten bezeichneten Daten liegen im Quellformat vor und werden im selbigen verarbeitet und weitergegeben.

Das Audio Buffer Modul implementiert einen Teil der Verarbeitungspipeline der 3-D Sound Verarbeitung. Es sind Situationen denkbar, in denen nachfolgende Verarbeitungsschritte länger oder kürzer laufen als die ihnen vorgeschalteten. Damit solche Situationen keine

Inkonsistenten Speicherbereiche bei denen zur Kopplung zweier Verarbeitungsmodulen genutzten Koppelspeicher verursachen, sind die Koppelspeicher aller Verarbeitungsmodulen der Verarbeitungspipeline durch zyklische Speicher gekoppelt.

Die Verwendung zyklischer Speicher erlaubt das Ausgleichen kurzzeitig auftretender Laufzeitunterschiede der Verarbeitungsmodulen. Laufen zwei Module, wie das Audio Buffer- und das Audio Pipe Modul, nacheinander, schreibt das vorlaufende Modul, hier das Audio Buffer Modul, an einer definierten Schreibposition in den zyklischen Buffer, während das nachlaufende Modul an einer Lese-Position aus diesem liest. Die Schreibposition befindet sich im optimalen Fall im zyklischen Speicher direkt hinter der Lese-Position. (Siehe Abbildung 26)

Bei einem gleichzeitigen Start der Verarbeitungsmodulen müssen an der Eingabe der Module bereits Daten zur Verarbeitung anliegen. Der zyklische Speicher stellt sicher, dass, wenn dieser zuvor gefüllt wurde, Daten anliegen und nicht aus dem Speicherbereich gelesen wird, in den die aktuellen Ergebnisdaten des Vorläufers geschrieben werden. Arbeiten alle in der Vorverarbeitungspipeline berücksichtigten Module synchron, kann der zyklische Speicher mit 2 Elementen angelegt werden, so dass ein Vertauschen der Lese- und Schreibposition zur sicheren Kopplung der Module ausreicht.

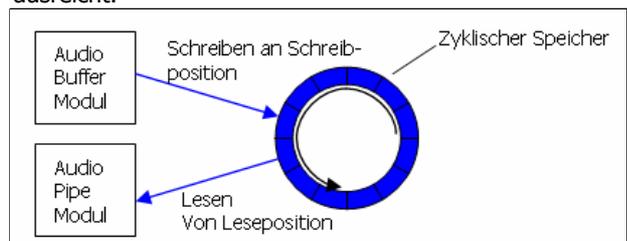


Abbildung 26 : Zyklische Koppelspeicher

6.2.3. Audio Pipe Modul

Das Audio Pipe Modul implementiert die räumliche Positionierung der Soundquellen. Die von einem für eine Instanz des Pipemoduls registrierten Audio Buffer berechneten Sounddaten, werden unter Berücksichtigung der Positionsparameter zur Berechnung des Klangbildes verwendet, das zur Ausgabe gelangt.

6.2.3.1. Datenstrukturen und Funktionen

Das Audio Pipe Modul ist implementiert in den Dateien sIAudioPipe.h und sIAudioPipe.c.

Es nutzt OpenGL Funktionen und Strukturen der Dateien sIAudioError.h, sIAudioError.c zur Fehlerbehandlung, sIAudioBuffer.c zur Aktualisierung von Sounddaten sowie die Datenstrukturen aus sIAudioTypes.h .

Die Datenstruktur des Audio Pipe Moduls umfasst vor allem Referenzen auf die registrierten Buffer als Datenlieferanten sowie das Audio Device, für welches das Klangbild zu berechnen ist. (Siehe Abbildung 27)

Type	Name	Info
SLdevice*	bufferDevice	Referenz auf eine Instanz des Audio Device Moduls, des genutzten Audio Devices
int	pipeBufferCount	Anzahl für Pipe registrierte Buffer
SaudioBuffer **	SaudioBuffer ** pipeBuffers	Liste von Referenzen auf die registrierten Buffer
void *	pipeData ISL_INTERNAL_BUFFER_FACTOR1	Zyklischer Kopplungspeicher
int	pipeUpdateInProgress	Synchronisation des Updateorgans

Abbildung 27 : Auszug der Datenstruktur des Audio Pipe Modul

Die Kernfunktionalität des Audio Pipe Moduls liegt in der Berechnung der über das Device wiederzugebenden Klangbilder. Neben den Funktionen zum Erstellen und Zerstören einer Instanz einer Audio Pipe, implementiert das Audio Pipe Modul Funktionen zum Registrieren und Deregistrieren von Audio Buffer Instanzen, die mit der Registrierung in der Verarbeitung Berücksichtigung finden. Die Berechnung des Klangbildes ist wiederum in der Update Funktion der Audio Pipe implementiert. (Siehe Abbildung 28)

Return Type	Name	Parameter
SaudioPipe *	slPipeCreateAudioPipe	(SLdevice *device, int maxBufferCount)
void	slPipeDestroyAudioPipe	(SaudioPipe *pipe)
void	slPipeRegisterBufferInPipe	(SaudioPipe *pipe, SaudioBuffer *buffer)
void	slPipeDeregisterBufferInPipe	(SaudioPipe *pipe, SaudioBuffer *buffer)
void *	slPipeUpdateAudioPipe	(void *data, int finalize)

Abbildung 28 : Auszug der Funktionen des Audio Pipe Modul

6.2.3.2. Berechnung des 3-D Klangbildes

Die in der Update Funktion gekapselte Berechnung des für die Wiedergabe bestimmten Klangbildes ist abhängig von den durch das Device bereitgestellten Leistungsmerkmalen und der zur Wiedergabe des 3-D Klangbildes gewählten Techniken unterschiedlich zu implementieren.

Um dem Level 1 des Anforderungskataloges zu genügen, ist zumindest ein 3-D Auditory Display zu unterstützen. Neben der stereo und vierkanal Wiedergabe unterstützt die Implementierung der Klangbild – Berechnung die Berechnung von 3-D Klangbildern für die Nutzung eines achtkanaligen Lautsprecher Setups.

6.2.3.3. Klangbild Berechnung

Die Berechnung des für die Wiedergabe bestimmten Klangbildes erfolgt in zwei Schritten. Basierend auf den Positions Attributen der für eine Audio Pipe registrierten Audio Buffer, wird zuerst 3-D Klangbild berechnet, das in einem zweiten Schritt der Position des Hörers im physikalischen Environment angepasst wird.

Die zur Zeit unterstützten stereophonen Techniken erfordern die Berechnung individueller Daten für jeden der im genutzten Setup vorhandenen Kanäle und für jeden der registrierten Audio Buffer. Die Ergebnisdaten basieren alle auf den Eingabe Daten, die aus dem vorgelagerten Audio Buffer Modul erzeugt wurden. Mit der Anpassung der Amplitude der einzelnen Kanäle wird die Vermittlung der Richtungscharakteristik ermöglicht. In Verbindung mit der bereits im Audio Buffer Modul berücksichtigten Entfernung einer virtuellen Soundquelle ist dann eine Positionierung bei der Wiedergabe möglich.

Die pro zu berücksichtigendem Kanal berechneten Amplituden werden auf jedes Sample des Eingabe –

Datenstroms angewandt. Die aus einem Sample resultierenden Samples verschiedener Amplituden werden als interleaved Stream in einen Audio Buffer spezifischen Zwischenstream geschrieben. Alle aus einem gleichen Sample der Eingabedaten resultierenden Samples bilden dabei einen Frame. Die Größe des Frames ergibt sich aus der Anzahl verwendeter Kanäle. Bei einer Wiedergabe über ein vierkanal Audio Display ergibt sich eine Frame Größe von 4. Mehrere Frames in Folge ergeben den Datenstrom.

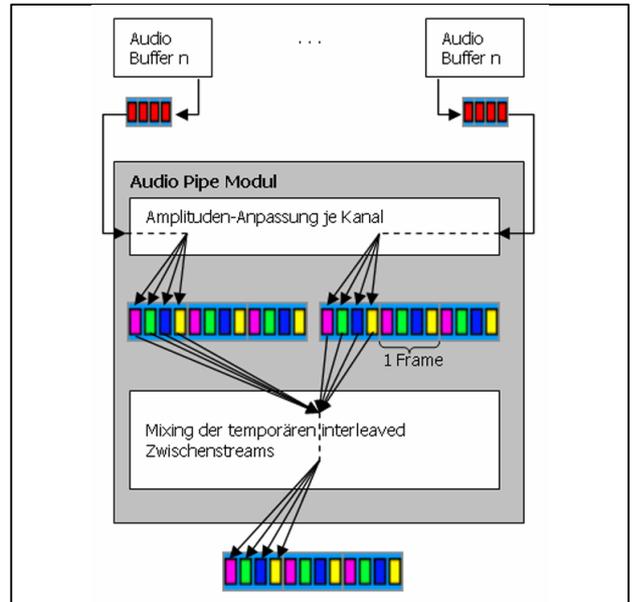


Abbildung 29 : Klangbild Berechnung

Die aus jedem für die Audio Pipe registrierten Audio Buffer resultierenden Datenströme werden für die Wiedergabe über das genutzte Device zu einem einzigen Datenstrom gemixt. Sie stehen dann als einzigartiger Datenstrom als Ergebnis zur Verfügung. (Siehe Abbildung 29)

Um das Schallereignis erfolgreich simulieren zu können, sind die Amplituden der Lautsprecher entsprechend der Richtung der entsprechenden virtuellen Soundquelle und der Richtung der Lautsprecher relativ zum physikalischen Hörer zu berechnen. Da bei stereophonen Techniken das Schallereignis simuliert wird, ist die Hörerorientierung nicht zu berücksichtigen. Mit Ambisonic bietet sich eine stereophone Technik, mit Hilfe derer harmonische Klangbilder für gleichmäßig verteilte Lautsprecher Setups mit einer geraden Lautsprecheranzahl (>2) berechnet werden können. Bei der Verwendung der Technik Ambisonic, wird ein Audio Signal auf alle für das Setup notwendigen Kanäle verteilt.

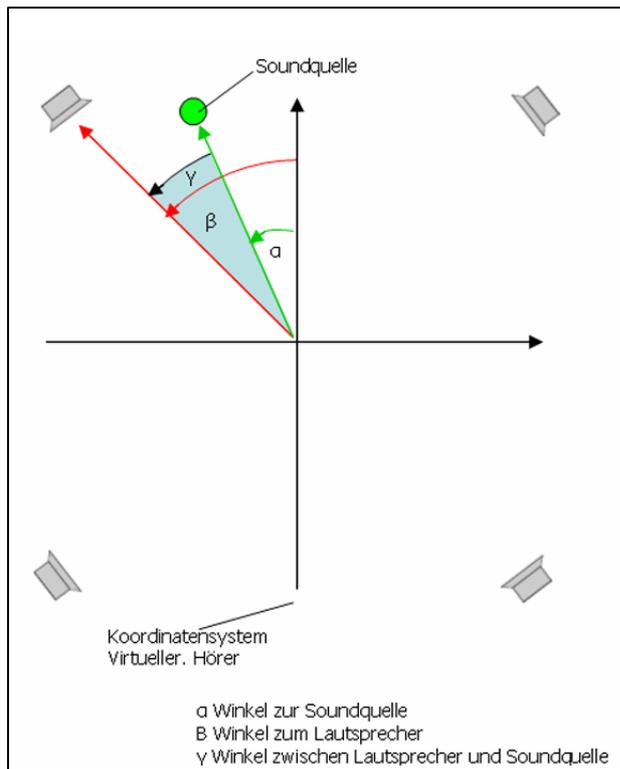


Abbildung 30 : Ambisonics

Die Intensität, mit der einer der Lautsprecher angesteuert wird, ergibt sich aus dem relativen Winkel zwischen der vom physikalischen Hörer ausgehenden Ausrichtung des Lautsprechers und der vom virtuellen Hörer ausgehenden Ausrichtung der virtuellen Soundquelle (Siehe Abbildung 30) [PUL_01]:

$$g_i = \frac{1}{N} (1 + 2 \cos \alpha_i)$$

g : Intensität des Lautsprechers (gain)

i : Lautsprecherindex

N : Anzahl Lautsprecher im Setup

α : relativer Winkel zwischen Soundsource und Lautsprecher

Die Verwendung der Cosinus – Funktion zur Berechnung bedingt eine höhere Intensität für Lautsprecher, die einen kleineren Winkel zur virtuellen Soundquelle haben. Mit steigendem Winkel fällt diese bis zum Winkel von 90 Grad ab.

Für Winkel, die den Wert von 90 Grad überschreiten, ergibt sich so ein negativer Wert. Um weiter auch die Lautsprecher mit einem Winkel > 90 Grad in die Wiedergabe einzubinden, erfolgt eine Normierung auf den Intensitätsbereich $0 \dots 2$. Lediglich der der Soundquelle direkt gegenüberliegende Lautsprecher wird auch bei einer Normierung nicht eingebunden.

$$g_i = 1 + \cos \alpha_i$$

Mit der Einführung der Normierung verringert sich der Einfluß des relativen Winkels auf 50% des errechneten Intensitätswertes. Der zur Normierung verwendete Summand macht 50% des absoluten Maximum der Cosinus – Funktion aus.

Eine Faktorisierung des Cosinus – Anteils in der Berechnung um einen Faktor 2 verdoppelt den Einfluß des Winkels. Er macht bei einer Verdopplung rund 67% aus.

$$g_i = 1 + 2 \cos \alpha_i$$

Mit der Faktorisierung wird eine harmonische Mischung aus guter Richtungscharakteristik und einem räumlichen Gefühl bei der Wiedergabe gewährleistet. [PUL_01]

Die Faktorisierung bewirkt eine Normalisierung des möglichen Intensitätsbereiches auf einen Bereich zwischen $-1 \dots 3$. Intensitäten als Anteil einer maximal verfügbaren Intensität bei der Wiedergabe bewegen sich im positiven Bereich bis 1. Um die Einhaltung dieses Bereichs zu garantieren, ist die errechnete Intensität ins Verhältnis zur Lautsprecheranzahl zu setzen. Eine erneute Normalisierung durch eine Division mit dem Divisor 3 hätte eine starre Normalisierung auf den Bereich $-1/3 \dots 1$ erwirkt und würde so zu einer mit der Anzahl der verwendeten Lautsprecher steigenden Gesamtintensität geführt. Die Verwendung der Lautsprecheranzahl reduziert diesen Effekt.

Da Ambisonic aufgrund der zuletzt notwendigen Division mit einem Divisor > 3 ausschließlich für Lautsprecher Setups mit einer Lautsprecheranzahl > 2 definiert ist, erfordert ein zweikanal Setup hier eine Sonderbehandlung.

Aufgrund der Tatsache, dass keine rückwärtigen Lautsprecher existieren, sind keine Normalisierungen von Nöten.

Spiegelt man zur Berechnung der Intensitäten die hinter dem virtuellen Hörer liegenden virtuellen Soundquellen an der x-Achse des Koordinaten Systems des virtuellen Hörers, können nicht nur die vor dem virtuellen Hörer liegenden Soundquellen berücksichtigt werden, denn auch die hinter ihm liegenden.

Mit den berechneten Intensitäten der einzelnen Kanäle lässt sich das Zielklangbild für die Ausgangssituation, mit dem physikalischen Hörer im Hotspot des Setups positioniert, erzeugen. Die Anpassung an vom Hotspot abweichende Hörerpositionen erfolgt unter Berücksichtigung der aktuellen Position des physikalischen Hörers. Das physikalische Environment mit dem physikalischen Hörer ist Bestandteil des Setup und somit im Audio Device Modul verfügbar in dem auch die Berechnung der Intensitätsanpassung erfolgt. Mit einer im Device Modul registrierten Änderung der Position des physikalischen Hörers, werden die pro Kanal notwendigen Intensitätsanpassungen berechnet. Diese werden im letzten Schritt zusammen mit denen aus der Klangbild Berechnung resultierenden Intensitäten auf die Sounddaten angewendet.

6.2.3.4. Synchronisation des Audio Pipe Moduls

Um Zugriffsverletzungen bei der Aktualisierung zu vermeiden, ist die Update Funktion ist auch innerhalb des Audio Pipe Moduls über die Verwendung eines

binären Semaphors, das als „pipeUpdateInProgress“ bezeichnetes Flag Teil der Datenstruktur ist, abgesichert.

6.2.3.5. Ein- und Ausgabedaten

Das Audio Pipe Modul arbeitet auf Basis der durch das vorgelagerte Audio Buffer Modul erzeugten Daten. Die Eingabedaten des Pipe Moduls liegen im Quellformat vor und werden in das für die am Device festgelegte Methode passende Format gewandelt. (Siehe Abbildung 29).

6.2.4. Audio Device Modul

Das Audio Device Modul ermöglicht den Zugriff auf die im Rechner verfügbare Audio Hardware. Passend zur gewählten Wiedergabetechnik erfolgt die Wiedergabe der im Pipe Modul erzeugten Daten.

6.2.4.1. Datenstrukturen und Funktionen

Das Audio Device Modul ist implementiert in den Dateien `slAudioDevice.h` und `slAudioDevice.c`. Es nutzt OpenSL Funktionen und Strukturen der Dateien `slAudioError.h`, `slAudioError.c` zur Fehlerbehandlung, `slAudioProcessing.c` zur Aktualisierung von Sounddaten sowie die Datenstrukturen aus `slAudioTypes.h`

In der Datenstruktur des Audio Device Moduls werden vor allem Attribute zur Beschreibung der Fähigkeiten eines Devices und zum Zugriff auf ein Audio Device gehalten. Die vom Device unterstützten Sampleraten, die unterstützten Sample Formate gehören ebenso zur Datenstruktur wie die aktuell genutzte Rate und das genutzte Format.

Die Auswahl möglicher Wiedergabetechniken wird durch die synchrone Nutzung mehrerer Devices erweitert. Zu diesem Zweck ist die Möglichkeit zur Bildung von Subdevices möglich. Die im Datenfeld „subDevices“ referenzierten Devices beschreiben die synchron zu nutzenden Audio Devices.

Mit jedem Device ist eine Callbackfunktion zu registrieren, die die Datenaktualisierung implementiert. Eine Referenz auf die Callbackfunktion ist Bestandteil der Datenstruktur.

Der letztendliche Zugriff auf das Device erfolgt über die in der Datenstruktur berücksichtigte Referenz auf das physikalische Audio Device. (Siehe Abbildung 31)

Mit dem Audio Device, das die Wiedergabemöglichkeiten beschreibt, sind die physikalischen Aspekte, physikalischen Environment und physikalischer Hörer, als Teil des Device Moduls implementiert. Device, physikalisches Environment und physikalischer Hörer beschreiben zusammen das verwendete Wiedergabe – Setup.

Zum Zugriff auf die im Rechnersystem verfügbaren Audio Devices ist es notwendig, Informationen über alle verfügbaren Devices abfragen und anzeigen zu können. Auf Basis der Informationen kann das passende Device gewählt und geöffnet werden.

Type	Name	Info
char *	name	Interne Bezeichnung des Device
SLDeviceID	deviceId	Interne ID des Device
void *	deviceHandle	Referenz auf den Datenkanal des Device
int	totalSubDevices	Anzahl SubDevices
SL_DEVICE_T **	subDevices	Liste mit Referenzen auf Sub Devices
slCallbackFunction	audioProcessingCallbackFunction	Callbackfunktion zur Datenaktualisierung
SLbyteOrder	deviceByteOrder	ByteOrder des Audio Devices
int	numberOfBuffers	Anzahl Buffer für gebufferte Ausgabe
int	totalOutputChannels	Anzahl verfügbarer Ausgabekanäle
int	usedOutputChannels	Anzahl genutzte Ausgabekanäle
SLsampleFormat	availableSampleFormats	Verfügbare Sample Formate
SLsampleFormat	currentSampleFormat	Aktuelles Sample Format
SLsampleRate	availableSampleRates	Verfügbare Sample Raten
SLsampleRate	currentSampleRate	Genutzte Sample Rate
SLchannelSetup	availableChannelSetups	Verfügbare Wiedergabe Setups
SLchannelSetup	currentChannelSetup	Genutztes Wiedergabe setup
SLspatializationTechnique	spatializationTechnique	Zur Wiedergabe genutzte Raumklangtechnik
float	listenerPosition [3]	Position des phys. Hörers im Setup
float	listenerLookAt [3]	Orientierung des phys. Hörers
float	listenerUp [3]	Orientierung des phys. Hörers
float *	speakerPosition	Position der Speaker rel. Zum phys. Hörer
float *	speakerActualSpeakerDistance	Entfernung der Speaker vom phys. Hörer

Abbildung 31 : Auszug der Datenstruktur des Audio Device Modul

Die zur Abfrage, zum Öffnen und zum Schreiben der Sounddaten notwendigen Funktionen sind Bestandteil des Audio Device Moduls (Siehe Abbildung 32)

Ist die Leistungsfähigkeit eines Devices nicht zur Umsetzung einer gewünschten Wiedergabemöglichkeit ausreichend, besteht die Möglichkeit, mehrere Audio Devices zu kombinieren und die Kombination als Multidevice in der Verarbeitung zu nutzen.

Return Type	Name	Parameter
void	slDeviceInitDevices	(void)
void	slDeviceFreeDevices	(void)
void	slDeviceFreeDevice	(SLDevice *device)
int	slDeviceCountDevices	(void)
SLdevice *	slDeviceEnumDevices	(void)
SLdevice *	slDeviceGetDefaultDevice	(void)
SLdevice *	slDeviceGetDeviceByID	(int deviceId)
char *	slDeviceGetDeviceInfoString	(SLDevice *device)
void	slDeviceOpenDevice	(SLdevice *deviceToOpen)
void	slDeviceCloseDevice	(SLdevice *deviceToClose)
int	slDeviceWriteSamples	(SLdevice *device, void *samples)

Abbildung 32 : Auszug der Funktionen des Audio Device Modul

6.2.4.2. Multidevice Implementierung

Bei der Bildung eines Multidevices werden mehrere einzelne im Rechnersystem verfügbaren Audio Devices synchronisiert und zur Ausgabe verwendet. So ist es möglich, durch eine Kombination in der Leistung beschränkter Audio Devices, ein Device zu schaffen, das in der Lage ist, auch eine acht- oder mehrkanalige Wiedergabe zu implementieren.

Eine Instanz des Audio Device Moduls dient dabei als logisches Multidevice. Das Multidevice referenziert eine Vielzahl von Subdevices, repräsentiert selbst aber keine im Rechnersystem verfügbare Hardware. Aus den referenzierten Audio Devices dienen die kleinsten gemeinsamen Werte zur Bestimmung der verfügbaren Sampleraten, der Sample Formate und schließlich der verfügbaren Wiedergabetechniken.

Die physikalischen Aspekte des physikalischen Environment und des physikalischen Hörers werden nur einmalig im logischen Multidevice gehalten.

Das Multidevice dient bei vorgelagerten Modulen der Verarbeitungspipeline als Referenz. Die in den vorgelagerten Modulen, speziell im Audio Pipe Modul berechneten Daten, entsprechen dem gewählten, vom logischen Multidevice unterstützten Wiedergabeformat,

nicht jedoch den Möglichkeiten der einzelnen Subdevices.

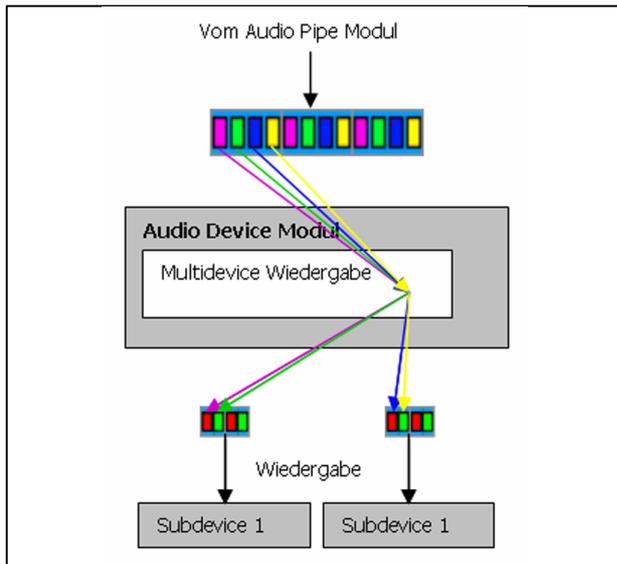


Abbildung 33 : Wiedergabeprozess Multidevice

Eine speziell zur Wiedergabe über ein Multidevice entwickelte Funktion, implementiert die Wiedergabe des im Pipe Modul implementierten Formats über die Subdevices. Der für das Wiedergabeformat im Audio Pipe Modul erstellte interleaved Stream der Sounddaten, den das Audio Device als Eingabedaten erhält, wird dabei Frame - weise bearbeitet. Einzelne Samples der Frames der Eingabedaten werden zu neuen Frames zusammengestellt, die den Wiedergabemöglichkeiten des individuellen Subdevices entsprechen. So entsteht pro Subdevice ein neuer interleaved Stream, der durch die Subdevices synchron wiedergegeben wird. (Siehe Abbildung 33)

6.2.4.3. Berechnung von Intensitätsanpassungen für wechselnde Hörerpositionen

Bei vom Hotspot abweichender Position des physikalischen Hörers innerhalb des durch die Position der Lautsprecher begrenzten Setups, ist die Intensität der Wiedergabe anzupassen.

Mit einer Verkürzung der Distanz des Hörers zu einem Lautsprecher muss die Intensität dieses Lautsprechers reduziert werden. Eine Vergrößerung der Distanz zu einem Lautsprecher bedingt eine Vergrößerung der Intensität für diesen Lautsprecher. Nur durch eine Anpassung der Intensitäten ist zu erreichen, dass der Hörer in wechselnden Positionen einen identischen Richtungseindruck der simulierten virtuellen Soundquelle erhält.

Die Anpassung erfolgt auf Basis des Inverse Square Law, da dieses die menschliche Entfernungswahrnehmung am genauesten beschreibt.

Um eine Anpassung zu ermöglichen, ist die Angabe des Hotspot für das aktuelle Setup notwendig. Entspricht die Position des physikalischen Hörers der des Hotspots, ist keine Anpassung nötig. Bewegt sich der physikalische

Hörer auf die Position des Lautsprechers, der die größte Distanz vom Hotspot hat, so befindet er sich in der Position, die die maximale Anpassung verursacht. Für diese Position ist vor Beginn der Soundverarbeitung die maximale Intensitätsverstärkung zu berechnen. Da eine Intensitätsüberschreitung von mehr als 100% zu Artefakten bei der Wiedergabe führt und durch das Audio Device nicht abgebildet werden kann, ist eine Basiskorrektur zu berechnen, die den verfügbaren Intensitätsrahmen begrenzt. Ist für die Extremposition eine Verstärkung von 15% notwendig, so darf auch bei einer Hörerposition im Hotspot des Setups die Wiedergabe lediglich mit einer Intensität von 85% erfolgen um Raum zur Intensitätsanpassung nach oben zur Verfügung zu haben. Nur so kann der Intensitätsbereich voll zur Anpassung der Hörerposition Verwendung finden.

Im Verlauf der Soundverarbeitung ist das Korrekturmaß für jeden Lautsprecher neu zu bestimmen, sobald sich die Position des physikalischen Hörers verändert.

6.2.4.4. Ein- und Ausgabedaten

Das Audio Device Modul arbeitet auf dem durch das Audio Pipe Modul generierten interleaved Audio Stream. Die Eingabedaten liegen bereits im Wiedergabeformat vor und können so direkt wiedergegeben werden. Lediglich bei der Wiedergabe über ein Multidevice ist eine erneute Aufspaltung des interleaved Streams notwendig, um diesen in ein Format zu transformieren, das durch die Subdevices zu verarbeiten ist.

6.2.5. Audio Processing Modul

Das Audio Processing Modul implementiert die Steuerung der Soundverarbeitung. Über das Audio Processing Modul wird die fortlaufende Wiedergabe über ein gewähltes Device ermöglicht.

6.2.5.1. Datenstrukturen und Funktionen

Das Audio Processing Modul verfügt lediglich über einige wenige Datenstrukturen. Von besonderer Wichtigkeit ist hier die in der Datenstruktur gehaltene Referenz auf eine Callbackfunktion, die mit der Initialisierung des Audio Processing Moduls übergeben wird.

Type	Name	Info
pthread_t	siSystemMainAudioProcessingThread	Referenz auf die zur Verarbeitung aufzurufende Callback Funktion

Abbildung 34 : Auszug der Datenstruktur des Audio Device Modul

Die referenzierte Callbackfunktion ist durch den Anwendungsentwickler zu erstellen. Sie hat zur Aufgabe, Daten zur Wiedergabe zu produzieren. Diese Funktion kann eigene Funktionalitäten beinhalten, oder aber auf den Update Funktionen des Audio Source-, Audio Buffer- oder Audio Pipe Moduls basieren. Durch die variierenden Implementierungsstrategien ist ein Zugriff auf die verschiedenen Schichten der Schichtstruktur möglich.

Zur Steuerung der Audioverarbeitung stehen Funktionen zur Initialisierung – dem Registrieren des verwendeten Audio Devices und der zu nutzenden Callback Funktion – zum Starten und Beenden der Soundverarbeitung zur Verfügung.

Die Steuerung der Audioverarbeitung ist so mit einem geringen Befehlssatz möglich. (Siehe Abbildung 32)

Return Type	Name	Parameter
void	slProcessingInitializeAudioProcessing	(SLdevice *device, slCallbackFunction deviceCallback)
void	slProcessingStartAudioProcessing	(SLdevice *device)
void	slProcessingStopAudioProcessing	(SLdevice *device)
void	slProcessingTerminateAudioProcessing	(SLdevice *device)

Abbildung 35 : Auszug der Funktionen des Audio Processing Modul

6.2.6. Audio Byteorder Modul

Neben der im Betriebssystem bei der Datenverarbeitung genutzten Byteorder variiert die Byteorder auch zwischen variierenden Audio Hardware Devices.

Arbeiten Betriebssystem und Audio Device auf Basis der gleichen Byteorder, ergeben sich keine Probleme. Ist dies jedoch nicht der Fall, müssen die in der Source, im Buffer oder in der Pipe generierten Daten in das entsprechende andere Byteorder Format transformiert werden.

Die Transformation der Byteorder erfolgt im Audio Byteorder Modul. Die Entscheidung, ob eine Transformation zu erfolgen hat, geschieht dabei unter Berücksichtigung der System Byteorder des Betriebssystems. Diese zur Compilezeit festgestellte Byteorder wird zur Laufzeit gegen die Byteorder des Devices geprüft. Es existiert keine Information, die zur Laufzeit direkt vom Device abgefragt werden könnte. Die vom Entwickler bei der Initialisierung eines Devices gemachten Angaben müssen hier als Informationsquelle ausreichen.

Da die interne Soundverarbeitung in der Byteorder des Betriebssystems erfolgt, ist der Transformationsschritt unmittelbar vor dem Schreiben der binären Audiodaten auf den Datenkanal des Audio Devices durchzuführen. Die Verwendung des 3-D Sound Systems auf verschiedenen Ebenen durch die Nutzung von Funktionen der verschiedenen Schichten, macht bei den Update Funktionen der Daten generierenden Module die Angabe eines Parameters erforderlich, ob die generierten Daten direkt in das Byteorder Format des Audio Devices zu bringen sind, oder im original Format zu belassen sind, wenn sie in der Verarbeitungspipeline weiter verarbeitet werden.

6.2.7. Audio Error Modul

Das Audio Error Modul ermöglicht eine umfangreiche Fehlerbehandlung. Auftretende Fehler werden auf einem Error Stack gespeichert und können zur Laufzeit abgefragt werden. Neben den Fehlercodes stehen Informationen zu den Fehlern zur Verfügung, die zur Laufzeit Verwendung finden können.

Die Tiefe des Error Stacks lässt sich zur Compilezeit festlegen.

6.3. Integration in Multimedia Entwicklungsumgebungen

Die modulare Umsetzung des im vorangegangenen Kapitel entworfenen Designs, ermöglicht es, das 3-D Sound System ohne tiefgehende Probleme durch ein weiteres Modul zu ergänzen, das eine Einbindung in eine ausgewählte Multimedia Entwicklungsumgebung ermöglicht.

Beispielhaft erfolgte die Entwicklung eines Moduls zur Integration von OpenSL in OpenGL®.

Bei der Integration wurden die speziellen Abläufe innerhalb von OpenGL® zunächst analysiert und dann in die Umsetzung des Moduls zur Integration integriert. Das Integrations Modul umfasst nunmehr einen Satz von Funktionen, die es ermöglichen, 3-D Sound in einer Art und Weise in OpenGL® basierte Anwendungen zu integrieren, wie es auch OpenGL® intern für Licht notwendig ist.

Mit dem „Sound Utilization for GL“ Modul (SLUGL) ist es möglich, bis zu 16 virtuelle Soundquellen separat zu aktivieren, zu positionieren und zu manipulieren. Neben der expliziten Angabe der Position einer Soundquelle, ist es zusätzlich möglich, die innerhalb eines OpenGL® Szenengraphen aufgebauten Transformationsketten für die integrierten virtuellen Soundquellen zu nutzen.

Zur Synchronisation der Soundverarbeitung mit der grafischen Verarbeitung wurde ein Mechanismus umgesetzt, bei dem Transformationen auf den virtuellen Soundquellen solange unberücksichtigt bleiben, bis sie über einen „slFlush“ Befehl als neue Parameter in die Verarbeitungspipeline übernommen werden. Zum Einen wird somit die Synchronisation mit der grafischen Darstellung in OpenGL® ermöglicht, zum Anderen ermöglicht diese Methodik die parallele Übernahme sequentiell durchgeführter separater Transformationen.

Der geringe Umfang des Befehlssatzes erleichtert zusätzlich eine schnelle Integration von 3-D Sound in interaktive Echtzeit Grafik – Anwendungen. (Siehe Abbildung 36)

Return Type	Name	Parameter
void	slSoundf	(SLenum sound, SLenum pname, SLfloat param)
void	slSoundi	(SLenum sound, SLenum pname, SLint param)
void	slSoundfv	(SLenum sound, SLenum pname, const SLfloat *params)
void	slSoundiv	(SLenum sound, SLenum pname, const SLint *params)
void	slGetSoundfv	(SLenum sound, SLenum pname, SLfloat *params)
void	slGetSoundiv	(SLenum sound, SLenum pname, SLint *params)
void	slSoundModelf	(SLenum pname, SLfloat param)
void	slSoundModeli	(SLenum pname, SLint param)
void	slSoundModelfv	(SLenum pname, const SLfloat *params)
void	slSoundModeliv	(SLenum pname, const SLint *params)
void	slFlush	(void)
void	slInt	()
void	slInitSpatializationMode	(SLenum mode)
void	slCreateSoundDevice	(void)
void	slEnable	(SLenum pname)
void	slDisable	(SLenum pname)
void	slSound	(SLenum sound)

Abbildung 36 : Auszug der Funktionen des Audio Processing Modul

Das SLUGL Modul ist implementiert in den Quellcodedateien slugl.c und slugl.h und nutzt direkt die Funktionen des Audio Processing Moduls und des Audio Device Moduls. Die detaillierte Funktionsbeschreibung findet sich in der bereitgestellten Dokumentation zu OpenGL.

6.4. Implementierung der Schichtstruktur

Die implementierten Module bilden die entworfene Schichtstruktur des 3-D Sound Systems in einer verfeinerten Granularität ab. Jedes der implementierten Module implementiert eine der Systemschichten oder dient allen Modulen als Service – Modul, wenn es für diese einen Dienst erbringt.

Durch die implementierte Schichtenstruktur entstehen so einseitige, in der Schichtstruktur nach unten gerichtete Abhängigkeiten. Module, die höhere Schichten implementieren, sind ausschließlich von unter ihnen liegenden Schichten abhängig. (Siehe Abbildung 36)

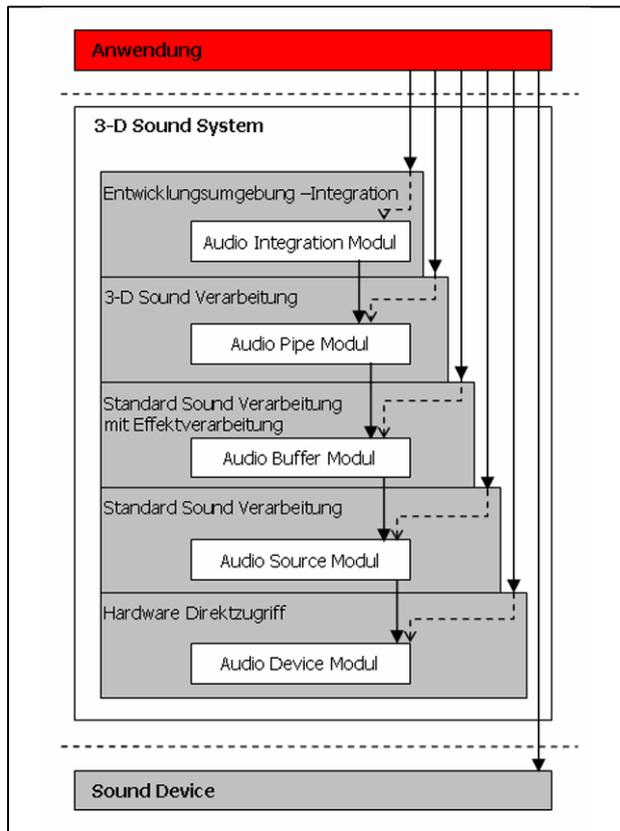


Abbildung 37 : Module der Schichtenstruktur

6.5. Erweiterungsmöglichkeiten

Die OpenSL 3-D Sound System Implementierung implementiert ein Level 1 des Anforderungskataloges konformes 3-D Sound System, bietet aufgrund des modularen Designs ausreichend Raum und Schnittstellen für beliebige Erweiterungen.

6.5.1. Erweiterungsmöglichkeiten Audio Device Modul

Die zur Zeit implementierten Routinen zur Initialisierung der Audio Devices und der zur Wiedergabe genutzten Setups verlangen nach einem hohen Arbeitsaufwand. Einzelne Lautsprecherpositionen müssen separat

angegeben und die zur Berechnung des 3-D Klangbildes zu verwendende Technik explizit ausgewählt werden. Dies geschieht derzeit direkt im Quellcode der entsprechenden Anwendung. Eine anwenderfreundliche Unterstützung zur Wahl des verfügbaren Setups oder zur Eingabe der Positionen der zu berücksichtigenden Lautsprecher ist nicht verfügbar.

Ein GUI gestütztes Tool zur Einstellung der bei der 3-D Soundverarbeitung relevanten Attributwerte sollte das Audio Device Modul in naher Zukunft ergänzen.

6.5.2. Erweiterungsmöglichkeiten Audio Source Modul

Die zur Zeit unterstützten Formate des Audio Source Moduls beschränken sich derzeit auf eine Handvoll ausgesuchter Formate. Lediglich WAV oder AIFF formatige Dateien sowie Ton- und Rausch – Quellen als generische Sound Source Typen werden unterstützt.

Die Erweiterung um weitere Dateiformate wird die Akzeptanz von OpenSL verbessern helfen. Großer Nutzen wird durch die Integration von komprimierten Datenformaten gezogen werden können. Der Gewinn im Speicherbedarf wird die Verbreitung im Heim – Bereich verbessern können.

Um verteilte Anwendungen besser unterstützen zu können, sollten Audio Streams als Sound Source Berücksichtigung finden. Ein über Netzwerke eintreffender Audio Strom als Sound Source, etwa einer Live – Übertragung, birgt ein großes Potential an möglichen Anwendungsgebieten. Durch die Einbindung von Datenströmen wären dann Videokonferenz – Systeme mit dreidimensional positionierbaren Gesprächspartnern möglich. In Verbindung mit CAVE® Display Systemen können sich durch die Kopplung mehrerer Virtual Reality Systeme neue Kommunikationsmuster entwickeln.

6.5.3. Erweiterungsmöglichkeiten Audio Buffer Modul

Das Audio Buffer Modul berücksichtigt derzeit bei der Bearbeitung ausschließlich die Position einer virtuellen Soundquelle sowie die initiale Intensität dieser. Zusätzlich sollte in naher Zukunft die Berücksichtigung eines Ausbreitungsmodells für virtuelle Soundquellen sowie die Orientierung der Soundquellen Berücksichtigung finden.

Bei der Implementierung weiterer Soundeffekte traten schnell Einschränkungen der 3-D Sound System Implementierung zutage. Aufgrund der Level 1 des Anforderungskataloges basierten Implementierung fehlen die zur Umsetzung von Soundeffekten notwendigen Strukturen. Zur Umsetzung von Soundeffekten wäre die Implementierung dieser aufgrund wechselnder Charakteristik der Effekte bei einer Beibehaltung des Designs auf das Audio Source und das Audio Buffer Modul zu verteilen. Die Einbindung eines Audio Effekt Moduls ist zwingend notwendig, um die Effektverarbeitung zu zentralisieren.

Mit dem Audio Effekt Modul können sowohl die Sound Effekte wie auch die Raum Effekte zur Umsetzung gelangen.

interaktiver Grafik – Anwendungen wird zukünftig ein zu bearbeitender Bereich sein.

6.5.4. Erweiterungsmöglichkeiten Audio Integration Modul

Die exemplarische Integration von OpenSL in OpenGL® deckt eine der am stärksten verbreiteten Spracherweiterung zur Erstellung interaktiver Grafik – Anwendungen ab.

Um im Bereich der interaktiven Anwendung schnell einen hohen Verbreitungsgrad zu erreichen, sind in naher Zukunft weitere Integrationsmöglichkeiten zu schaffen. Basierend auf einer Analyse, welche Entwicklungsumgebungen verstärkt zum Erzeugen interaktiver Multimedia – Anwendungen genutzt werden, sollte die Integration für die einflussreichsten Entwicklungsumgebungen vorangetrieben werden.

7. Zusammenfassung und Ausblick

In dieser Arbeit wurde zuerst ein Überblick über die Aspekte der 3-D Sound Wahrnehmung und –verarbeitung gegeben. Es wurde ein Anforderungskatalog für interaktive 3-D Sound Systeme entwickelt, auf Basis dessen bereits existierende 3-D Sound Systeme analysiert und bewertet wurden. Im weiteren Verlauf der Arbeit wurde mit den Ergebnissen der Analyse und dem exemplarischen Entwurf eines 3-D Sound Systems, das die interaktive Verarbeitung von 3-D Sound in Echtzeit ermöglicht, ein möglicher Weg zur Integration von 3-D Sound in interaktive Echtzeit – Grafik – Anwendungen aufgezeigt.

Die anschließende beispielhafte Implementierung des im Entwurf beschriebenen System, bestätigt die Funktionsfähigkeit der Grundlagen des System – Designs. Das umgesetzte 3-D Sound System ermöglicht die Integration von 3-D Sound in interaktive Echtzeit – Anwendungen bei gleichzeitiger Kopplung von Soundquellen an grafische Objekte und einer synchronen Wiedergabe der aufbereiteten grafischen und akustischen Daten.

Ein speziell zur Integration von 3-D Sound entwickeltes Modul beschleunigt und vereinfacht die Integration von Sound in den Anwendungsablauf. Mit dem Ergebnis dieser Arbeit wird die Möglichkeit zur Erstellung von Anwendungen höheren Immersionsgrades geboten.

Zu diesem Zeitpunkt ist mit dem 3-D Sound System keine realitätsnahe Soundverarbeitung verfügbar. Die modulare Umsetzung des Systems erlaubt in zukünftigen Arbeiten jedoch eine effektive Erweiterung einzelner Module, um so schrittweise von der jetzigen abstrakten Umsetzung eine realitätsnahe Soundverarbeitung zu implementieren.

Unterstützt das System derzeit explizit nur die Einbindung in OpenGL® basierte Anwendungen, ist eine von der OpenGL® - Kopplung losgelöste Verwendung jederzeit möglich. Auch die explizite Kopplung an weitere Umgebungen zur Erstellung

Eidesstattliche Erklärung

Hiermit erkläre ich an Eides Statt, dass ich diese Arbeit selbständig und ohne Benutzung anderer als den angegebenen Hilfsmitteln angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht.

Diese Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungskommission vorgelegt und auch nicht veröffentlicht.

Ellerau, den 28.08.2003

Diplom Medieninformatiker (FH) Severin S. Todt