

Diplomarbeit

Die Beschreibung und Kategorisierung von Audiosignalen

Ausgeführt zum Zweck der Erlangung des akademischen Grades
Dipl. Ing. (FH) für Telekommunikation und Medien
am Fachhochschul-Diplomstudiengang
Telekommunikation und Medien St. Pölten
Vertiefungsrichtung Medientechnik

von:

Wolfgang Csacsinovits
tm0210038021

Begutachter:
Dipl. Ing. Hannes Raffaseder

Zweitbegutachter:
Dipl. Ing. Franz Zotlöterer

St.Pölten, Sommersemester 2006

Ehrenwörtliche Erklärung

Ich versichere, dass

- ich diese Diplomarbeit selbstständig verfasst habe, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt habe und mich auch sonst keiner unerlaubten Hilfe bedient habe.
- ich dieses Diplomarbeitsthema bisher weder im Inland noch im Ausland einem Begutachter/einer Begutachterin zur Beurteilung oder in irgendeiner Form als Prüfungsarbeit vorgelegt habe.

Diese Arbeit stimmt mit der vom Begutachter beurteilten Arbeit überein.

.....
Ort, Datum

.....
Unterschrift

Zusammenfassung

In dieser Arbeit wird das Forschungsprojekt „AllThatSounds“ vorgestellt, bei der es um die Entwicklung eines effizienten Algorithmus zur leichten, assoziativ - semantischen Suche von Sounddateien geht. Daher werden Möglichkeiten zur Beschreibung, Benennung und Kategorisierung von Audiodateien aufgezeigt.

Es wird das Thema Metadaten in Bezug auf Audiofiles durchleuchtet und mögliche Entstehungsprozesse von Audio – Metadaten werden anhand von bereits existierenden Sound – Such – Systemen analysiert. Es wird generell auf den Einsatz von Sprache für die Beschreibung von Klängen eingegangen, wobei abschließend quantitative Usertests mehr über die verwendeten Möglichkeiten zur Kategorisierung von kurzen Audio - Samples verraten.

Abstract

This diploma thesis presents the research project „AllThatSounds“, which deals with the development of an efficient algorithm for an associative search of sound - files. Therefore it shows possibilities for the description, the naming and the categorization of audio – files.

The theme metadata will be arranged and possible accrument processes of audio – metadata will be analysed, based on existing sound – search – systems. Generally this work also shows the exertion of speech to describe sounds, whereat finally quantitative usertests try to show more about the possibilities to categorize short audio – samples.

Inhaltsverzeichnis

Ehrenwörtliche Erklärung.....	2
Zusammenfassung.....	3
Abstract.....	3
Inhaltsverzeichnis.....	4
1 Einleitung	
1.1 Forschungsprojekt „AllThatSounds“.....	6
1.2 Ziele der Arbeit.....	11
2 Metadaten im Audibereich	
2.1 Arten von Metadaten.....	12
2.2 Formate.....	15
2.3 Entstehungsarten von Metadaten.....	23
3 Klangbeschreibung,- und Kategorisierung	
3.1 Beschreibung von Klangobjekten.....	35
3.2 Evaluierung der bestehenden Soundlibraries.....	42
3.3 Ansatz zur Kategorisierung von Klängen.....	46
I. Schallquelle.....	47
II. Art des Sounds bzw. mediale Anwendung.....	49
III. Time.....	51
IV. Tone.....	53
V. Klangcharakter.....	54
VI. Intervall.....	56
VII. Dichte.....	58
VIII. Überlagerung.....	58
IX. Lautstärkeänderung & Dynamik.....	59
X. Perspektive.....	60
XI. weitere Parameter.....	61

4	Testreihen	
4.1	Ziele der Usertests und Ansatz der Methoden.....	65
4.2	Gefahren und Probleme der Testreihen.....	67
4.3	Auswahl der Testsounds und Methoden der Durchführung.....	68
4.4	Testfragen, Auswertungsmethode und Ergebnisse.....	70
5	Aussblick und Schlussbemerkungen.....	91
ANHANG		
	Anhang A: Literatur,- und Quellennachweis.....	93
	Anhang B: Abbildungs,- und Tabellenverzeichnis.....	95
	Anhang C: Danksagung.....	96
	Anhang D: Glossar – Fachausdrücke.....	97

1 Einleitung

1.1 Forschungsprojekt „AllThatSounds“

Forschung zur assoziativ, semantischen Erfassung von Audiofiles

„Nachfrage und Angebot an digital gespeicherten Soundfiles haben innerhalb weniger Jahre sprunghaft zugenommen. Anzahl, stilistische Vielfalt, Qualität digital verfügbarer Sounddateien haben ein unüberschaubares Ausmaß erreicht. Effektive Ansätze für die Suche nach Audiodaten spielen dabei eine entscheidende Rolle.

Bei einem Großteil der vorhandenen Tools zur Suche nach Audiomaterial müssen bereits vor der Suche der Titel, Interpret oder ähnliche wichtige Eckdaten bekannt sein, damit man eine realistische Chance hat, den gewünschten Song oder Sound zu finden. Vieles bleibt daher unentdeckt, worauf früher vielleicht der persönliche Berater im Plattengeschäft hingewiesen hätte.

AllThatSounds soll das Finden von, für den einzelnen User relevanten Audiodaten vereinfachen. Spezielle Features unterstützen vor allem Musik, Medien- oder Werbeproduzenten bei der Suche nach Musik, Drumloops und Soundeffekten, die am besten zum Image, zur Stimmung, zum Inhalt des Produkts passen.

Eine Suche nach semantischen und an menschlichen Hörgewohnheiten angelehnten Kriterien ist bisher kaum möglich. Anfragen nach Musik zum Entspannen, Sounds für gewisse emotionale Untermahlungen in Filmen oder für Produktpräsentationen scheitern. Persönlicher Geschmack, kulturelle und soziale Herkunft etc. sind wichtige Aspekte, die bisher kaum Beachtung finden.

In der ersten Projektphase wurde qualitative Recherche und Evaluierung von neuen Erkenntnissen der Grundlagenforschung und aktuellen Produkten

1.1. Einleitung – Forschungsprojekt „AllThatSounds“

durchgeführt. Auf einen intensiven Austausch mit internationalen Forschungseinrichtungen und VertreterInnen aus der Medienbranche wird dabei Wert gelegt. Im Anschluss daran wird nun an der Weiterentwicklung, Optimierung und Anpassung an die speziellen Anforderungen von AllThatSounds gearbeitet.“¹

Das Projekt hat eine Laufzeit von insgesamt 2,5 Jahren und startete im Oktober 2005 mit der ersten Projektphase.

Die 4 Teilbereiche in AllThatSounds

Damit AllThatSounds die bestmöglichen Ergebnisse liefern kann, werden mehrere unterschiedliche Ansätze bei der Suche nach Sounds kombiniert. AllThatSounds baut daher auf 4 Säulen auf, wobei in dieser Arbeit vor allem auf die ersten beiden Punkte, die deskriptive Analyse, sowie die semantische Analyse eingegangen wird:

Deskriptive Analyse

„Diese Teilaufgabe beschäftigt sich mit der Beschreibung der Audiodaten aus der Sicht des Urhebers bzw. Produzenten. Es sollen hier u.a. folgende Fragen geklärt werden: Wofür bzw. für welchen Zweck wurde komponiert bzw. produziert? In welchem sozialen, kulturellen, geographischen Kontext wurde komponiert? Wo, wie, warum etc. wurden die Sounds aufgezeichnet? Sind die Sounds metronomische, loop-ähnliche Sounds oder eher flächige Sounds ohne Takt? Haben die Sounds tonalen oder eher geräuschehaltigen Charakter?“² usw.

Zentrales Ziel dieser Teilaufgabe ist die Entwicklung eines einfach handhabbaren, effizienten, möglichst umfassenden und erweiterbaren Modells für die deskriptive Analyse.

¹Projektseite „AllThatSounds“ (2005): <http://www.allthatsounds.net/projekt.html>

²Raffaseder, H. (2005): Antragsformular für Kooperationsvorhaben „AllThatSounds“, S.34

1.1. Einleitung – Forschungsprojekt „AllThatSounds“

Je mehr solcher Fragen bereits vom Urheber beantwortet werden, desto besser wird die Kategorisierung der Audiodaten in diesem Bereich funktionieren. Nun besteht aber ein Hauptproblem bestehender Ansätze (z.B. ID3-Tags bei mp3-Dateien) darin, dass genau diese Informationen oft nicht oder nur unzureichend vorhanden sind.

Es müssen im System AllThatSounds daher Anreize geschaffen werden, die Urheber dazu veranlassen, ihre Audiodaten mit den gewünschten Zusatzinformationen in Form von Metadaten zu versehen. In dieser Diplomarbeit werden daher Ideen zur Kategorisierung und Beschreibung von Klängen geliefert, die als Anstoß für die Erstellung einer Eingabemaske für Metadaten dienen, bei der der Urheber oder auch erst der Soundbenutzer auf einfache Weise den betreffenden Sound so exakt wie möglich beschreiben kann.

Semantische Analyse

„Dem Bereich des semantischen Gehalts von akustischen Ereignissen kommt gerade für den Bereich der Medienproduktion eine herausragende Bedeutung zu. Beispielsweise werden „afrikanische Rhythmen“, „furchterregende Geräusche“ oder „traurige Musik“ benötigt, um im Medienprodukt den geographischen Kontext zu beschreiben, um Spannung zu erzeugen oder um die Stimmung eines Protagonisten oder des Umfeldes zu verdeutlichen.

Wodurch erhalten Audiodaten aber die Bedeutung „afrikanisch“, „furchterregend“ oder „traurig“ etc.? Wie können unterschiedliche Bedeutungen kategorisiert werden? Wie ist es möglich, die Semantik von Audiodaten zumindest teil-automatisiert zu erfassen?

Die aktuelle Forschung gibt auf derartige Fragen nur unbefriedigende Antworten. Das soziale und kulturelle Umfeld scheint eine wichtige Rolle zu

1.1.Einleitung – Forschungsprojekt „AllThatSounds“

spielen. Wechselwirkungen zwischen visueller und akustischer Wahrnehmung dürften ebenfalls große Bedeutung haben.

In dieser Teilaufgabe werden zunächst die wenigen bisher bestehenden Ansätze in den Bereichen „akustische Semantik“ und „Akustisch-musikalische Wirkungsforschung“ für das Forschungsvorhaben AllThatSounds und die FH St. Pölten verfügbar gemacht, analysiert und evaluiert. Darauf aufbauend werden konkrete Maßnahmen für die Weiterentwicklung erarbeitet.

Wichtig erscheint auch, für die Bedeutung des semantischen und symbolischen Gehalts von akustischen Ereignissen im Zuge der Gestaltung der Tonspur eines Medienprodukts das allgemeine Problembewusstsein zu vertiefen.“³

Mehrere Versuchsreihen, bzw. qualitative Interviews mit Medienproduzenten sollen neue Aufschlüsse über die Semantik von akustischen Ereignissen geben. Wesentliches Ziel ist die Definition eines erweiterbaren Datenmodells, das eine möglichst umfassende Kategorisierung von Audiodaten nach semantischen Kriterien erlaubt.

Weiters wurde gleich in der Startphase des Projekts damit begonnen, eine Clipdatenbank aufzubauen. Diese besteht aus kurzen Ausschnitten diverser Filme von unterschiedlichsten Genres, bei denen Sounds entweder sehr gut, interessant oder anderweitig bedeutend eingesetzt wurden. Die Datenbank soll im Verlauf des Projekts noch mit vielen Clips erweitert werden und mit diversen Suchfunktionen versehen werden. Pro Clip gibt es jeweils Zusatzinformationen über die Symbolik, die Ereignisse und die vorkommenden Objekte und Materialien in den Filmszenen. Darauf aufbauend sollen mit Hilfe von Studenten, die beispielsweise ihre Bakkelaufsätze – Arbeit darüber verfassen, Bild-Ton-Beziehungen genauestens untersucht werden.

³Raffaseder, H. (2005): Antragsformular für Kooperationsvorhaben „AllThatSounds“, S.35

1.1.Einleitung – Forschungsprojekt „AllThatSounds“

Folgende Fragen können danach beantwortet werden: Wie sehr ist der Ton für die auslösende Stimmung in der Szene verantwortlich? Wirkt die Szene auch mit konträrer Musik gleichermaßen? Wie sehen die Wechselwirkungen zwischen Bild und Ton aus?

Hörer-spezifische Analyse

„Welche Anwender eine Audiodatei unter welchen Voraussetzungen für welche Produktionen bzw. in welchem Kontext einsetzen bzw. hören, stellt im Allgemeinen eine ganz wichtige Charakterisierung dar: „Musik die vom User X bzw. der Usergruppe Y bevorzugt gehört wird.“ Die Teilaufgabe 2 „Hörer-spezifische Analyse“ trägt diesem Umstand Rechnung. Durch Vergleich von Hörer-Profilen kann das System AllThatSounds dann entsprechende Vorschläge auf bestimmte Suchanfragen machen: „Hörer A hat zur Entspannung Musik X gerne gehört und User B gerne Song Y. Dein Profil liegt zwischen A und B, wähle daher Musik Z.“ Ähnliche Ansätze werden in verschiedenen Branchen bereits erfolgreich eingesetzt (vgl. z.B. Amazon). In dieser Teilaufgabe werden bestehende Systeme analysiert, evaluiert und für die Anforderungen von AllThatSounds adaptiert.“⁴

Musikalisch-technische Analyse

„Zahlreiche internationale Forschungsinstitute erarbeiten zurzeit wichtige Grundlagen, um aus den Ergebnissen einer technischen Signalanalyse auf den musikalischen Kontent einer Audiodatei rückschließen zu können. Beispiele hierfür sind automatische Beat- und Tempodetection, Extraktion von Tonhöhen bzw. Melodien, Erkennung von Instrumenten bzw. Klangfarben etc. In dieser Teilaufgabe werden Algorithmen und Ergebnisse aus der Grundlagenforschung getestet, evaluiert und an die Anforderungen von AllThatSounds angepasst.“⁵

⁴Raffaseder, H. (2005): Antragsformular für Kooperationsvorhaben „AllThatSounds“, S.34

⁵Raffaseder, H. (2005): Antragsformular für Kooperationsvorhaben „AllThatSounds“, S.35

1.2 Ziele der Arbeit

Ziel dieser Arbeit ist es einen neuen Ansatz zur Erfassung, zur Beschreibung und letztendlich Ideen zur Erleichterung der Suche nach Audiofiles zu entwickeln. Dabei soll dem Problem der mangelnden assoziativen Möglichkeit zur Suche nach Sounddateien auf mehreren Ebenen entgegengewirkt werden. Diese verschiedenen Ebenen beinhalten in dieser Arbeit die sinnvolle Vergabe von Metadaten und die bestmögliche semantische Beschreibung von Sounds. Diese unterschiedlichen Herangehensweisen an das Thema „Suche nach Sounds“ sollen gewisse Probleme wie die Subjektivität des emotionalen Gehaltes von Soundfiles, oder die Schwierigkeit, abstrakte Sounds zu kategorisieren und zu beschreiben so weit wie möglich umgehen.

- Im Bereich der Metadaten soll durchleuchtet werden, in welchen Formaten solche Metadaten im Bezug auf Audiodateien bereits vorkommen. Anhand von bestehenden Software – Beispielen werden unterschiedliche Entstehensarten von Audio – Metadaten beschrieben.
- Des Weiteren wird auf die Frage eingegangen, welches Vokubular grundsätzlich zur Beschreibung von Klängen zur Verfügung steht. Anhand der für AllThatSounds angeschafften Sound – Libraries werden die darauf befindlichen Ansätze zur Benennung von Klängen evaluiert, und eine für die Suche nach Sounds für Medienproduktionen nützliche Kategorisierung von Klängen erstellt. Wie kann man Audiosignale zur leichten Wiederauffindung am besten einteilen und mit Eigenschaften versehen?
- Testreihen mit Versuchs – Hörern sollen letztlich Aufschluß darüber geben, wie wir Sounds generell beschreiben. Welche sprachlichen Mittel werden am häufigsten verwendet und mit welchen emotionalen Qualitäten werden unterschiedliche Testsounds verknüpft? Werden dieselben Sounds von den Probanden stets gleich oder ähnlich beschrieben?

2 Metadaten im Audibereich

2.1 Arten von Metadaten

Metadaten sind allgemein Daten über die Daten. Sie enthalten Informationen über die eigentlichen Daten, um sie entweder besser zu verstehen, leichter zu finden oder die Herkunft eindeutig festzulegen etc. Effizienter Zugriff und eine einfache Verwaltung der Daten können mögliche Gründe für die Anlegung von Metadaten sein.

Grundsätzlich gibt es mehrere Möglichkeiten um Metadaten voneinander zu unterscheiden. Einige davon wären:

1) Art der Metadaten:

Inhaltsabhängige und inhaltunabhängige Metadaten

2) Sinn,- und Nutzungsarten:

Informationsabfrage, Datenaustausch, Datafitness (Dokumentation von erstellten Daten für das Verständnis Dritter)

3) Verfahren zur Generierung:

Manuelle Eingabe, Vergleich mit Referenzwerten, Messung bzw. automatische Berechnung, Ableitung aus dem Inhalt

4) Natur der Metadaten:

Laien,- oder Experten Metadaten

5) Status:

Statische, dynamische, langfristige oder kurzfristige Metadaten

6) Level:

Sammlungs- Metadaten bezogen auf ganze Informationssammlungen oder Element – Metadaten, bezogen auf individuelle Objekte⁶

Im Audiobereich macht der Einsatz von Metadaten so wie auch in anderen Bereichen, wie z.B: bei Webseiten, für die Beschreibung der Daten ohne Zweifel sehr viel Sinn.

Die am häufigsten verwendeten Audio – Metadaten sind großteils lediglich identifizierende Metadaten, also weitgehend inhaltsunabhängig. Das bedeutet, dass sie hauptsächlich eingesetzt werden, um sie eindeutig von anderen Audiodaten abzugrenzen. Bei Musikstücken werden meistens Interpret, Titel, eventuell Album, Entstehungsjahr und Genre ergänzt, wie es bei den bekannten ID3-Tags beim MP3 – Audioformat der Fall ist.

Im Gegensatz dazu beinhalten beispielsweise Metadaten bei Webseiten meist auch viel mehr inhaltsbeschreibende, also semantische Informationen über die Seite.

Bei Musikstücken sind diese identifizierenden Metadaten vielleicht in den meisten Fällen noch ausreichend, vor allem dann, wenn man gezielt nach einem einzelnen Musikstück sucht, und Titel bzw. Interpret bekannt ist. Bei Sounds, dh. einzelne Drumloops, kurze Instrumentalparts, Geräuschen und Ähnlichem, die nur als einzelner Baustein für ein Musikstück oder für die emotionale Untermahlung eines Medienprodukts dienen sollten, reichen diese Informationen nicht aus.

⁶vgl. Cisneros, A. (2005): Metadaten

http://www.ibr.cs.tu-bs.de/courses/ss05/skm/ausarbeitungen/Cisneros_Metadaten.pdf

Ein Beispiel:

Ein Sounddesigner sucht für die Vertonung einer spannenden Autoverfolgungsjagd die passende Untermahlung. Er weiß intuitiv, dass er ein eher Rhythmus – betontes, schnelleres Musikstück mit einer einprägsamen Basslinie für die Vertonung braucht. Da Titel und Interpret des Stücks also noch unbekannt sind, würden ihm die angesprochenen Zusatzinformationen der Audiodaten nicht helfen.

Hier wäre die Suche nach semantischen Kriterien absolut notwendig.

2.2 Formate

Diverse Audioformate bieten die Möglichkeit, Metadaten im Audiofile selbst mitzuspeichern. Zusatzinformationen, die in Textformat in Audiodateien abgespeichert sind, können verschiedene Formate aufweisen. Diese Formate, die so genannten Tagging – Formate haben ihren Ursprung in der Erfindung des populären Audio – Kompressionsverfahrens „MPEG Audio Layer 3“ – kurz „MP3“.

Das 1985 vom Fraunhofer-Institut entwickelte Audio – Format beinhaltet vorerst keine zusätzlichen Daten, abgesehen von relevanten Informationen über den Dekodierungsprozess und wenige spärliche Copyright - Infos. Aufgrund der steigenden Popularität des MP3 – Algorithmus als Standard für digitale Musik und Austausch im Webbereich, wurde die Notwendigkeit, Textinformationen zu inkludieren erkannt und der ID3 – Tag eingeführt.

Der daraus resultierende Vorteil eines solchen Tagging - Formates ist jener, dass man sämtliche Infos über das Soundfile nicht im Datei,- oder Ordnernamen bzw. gar in einem externen Textfile unterbringen muss.

ID3v1:

Die ursprüngliche Version des ID3 Tags, die speziell für das MP3 Format von Eric Kemp entwickelt wurde hat eine fixe Länge von 128 bytes. Diese werden für folgende Metadaten genutzt:

1. Songtitel (30 characters)
2. Artist (30 characters)
3. Album (30 characters)
4. Jahr (4 characters)
5. Kommentar (30 characters)
6. Genre (1 byte)

Die Inhalte werden hintereinander als Clear Text an die Datei angehängt. Durch die fix vorgegebene Länge und Reihenfolge der einzelnen Felder ist es nicht notwendig, die Feldnamen extra abzuspeichern.

Das Feld „Genre“ speichert lediglich eine Zahl zwischen 0 und 79. Aus einer vordefinierten Liste kann also zwischen 80 Genres ausgewählt werden.

Am Ende des ID3 Tags bleiben noch 3 bytes, um den Tag als solchen zu markieren. Dh. diese 3 bytes beinhalten stets „TAG“ für die Identifikation des ID3 Tags.

Ein großer Vorteil des ID3v1 Tags ist nach wie vor die große Kompatibilität zu fast jeder Software bzw. Hardware. Die Nachteile des etwas veralteten Tagging - Formates sind die starr auf 30 Zeichen beschränkten Felder und das Fehlen der Möglichkeit zum Anlegen weiterer Felder. Weiters gibt es keine Unicode – Unterstützung für internationale Zeichensätze.

Die nächste kleine Weiterentwicklung ist der ID3v1.1 Tag, der durch Setzen eines Null-Bytes am Ende eines Feldeintrags das nächste Feld automatisch um diese auf 30 Zeichen fehlende Anzahl an Zeichen vergrößert. Hat also der Songtitel lediglich 10 Zeichen, so hat man beim Artist- bzw. Album Feld bereits 20 Zeichen mehr zur Verfügung. Bei ID3v1.1 werden die 128 bytes also besser genutzt und man ist bei Titel, Artist und Genre nicht zwingend auf die jeweils 30 Zeichen beschränkt. Weiters wurde das Feld „Album – Track“ hinzugefügt.

ID3v2:

Der Version 2 des ID3 Tags wurde ein grundlegend neues Design verpasst. Das Format ist weitaus flexibler und erweiterbarer als die beiden Vorgänger ID3v1 und ID3v1.1. Längenbeschränkungen der Felder wurden aufgehoben und mehr vordefinierte Felder, auch Frames genannt, mit ausführlichen Beschreibungen ihres jeweiligen Einsatzzwecks wurden hinzugefügt.

Zusätzlich zu den Cleartext – Feldern aus ID3v1 können auch eigene weitere Felder angelegt werden, so wie auch binäre Objekte wie Bilder im Tag abgespeichert werden.

Im Gegensatz zu ID3v1 wurde hier der Dateianfang als Speicherort gewählt, um die Tags bei Streaming Anwendungen als erstes zu senden und damit dem Empfänger anzeigen zu können.

Die Unicode-Unterstützung wird zwar nicht vom Standard aus sichergestellt, allerdings können Zeichen in der systemeigenen Codepage abgespeichert werden.

„Die bereits mehrfach überarbeitete Version des ID3v2 Tags ist mit den meisten aktuellen Softwareproduktion zur Bearbeitung von Audio – Metadata – Tags oder MP3 bezogenen Programmen kompatibel, jedoch sind Inkompatibilitäten nicht ausgeschlossen.“⁷

⁷Nilsson, M. (2001): ID3 made easy
<http://id3lib.sourceforge.net/id3/id3v1.html>

APE:

Der Tagging – Standard APE wurde für das Audioformat MPC, ein auf MP2 Algorithmen basierender Audio Codec von Franz Klemm entwickelt. Die überarbeitete Version APEv2 wird heute jedoch auch von einigen anderen Audioformaten unterstützt wie beispielsweise WavPack, APE (Monkey's Audio) oder OptimFROG.

Immer mehr Anwendungen wie zB: Foobar, MP3tag oder Helium 2005 unterstützen aufgrund der höheren Flexibilität als ID3 APEv2 – Tags auch in Verbindung mit MP3 - Files.

APEv2-Tags erlauben dem Benutzer das Anlegen völlig frei definierbarer Felder, die Feldnamen werden dabei als Cleartext in die Datei mitgespeichert. Ein Feld kann auch mehrfach angelegt werden, was sinnvoll ist um einer Audiodatei z.B. mehrere Genres zuzuteilen. Für viele Zwecke sind Felder vordefiniert; dies geschieht jedoch weniger streng als es bei ID3v2 der Fall ist.

Unicode wird ebenfalls unterstützt, wobei alle Felder im UTF8-Format abgespeichert werden, was einen Vorteil gegenüber ID3v2 darstellt.

Vorbis Comment:

Vorbis Comment wurde für das Audioformat Ogg Vorbis entwickelt und ist auch offizieller Bestandteil des verlustfreien Audioformats FLAC. Es unterscheidet sich im Aufbau geringfügig von APEv2, hat jedoch weitgehend die gleichen Vorteile.

Broadcast Wave Format:

Das ursprünglich von Microsoft entwickelte WAV-Format stellt quasi den Standard für digitale Audioverarbeitung dar.

Broadcast – Wave Format (BWF), welches von der EBU (European Broadcasting Union) mit Partnern in der Audioindustrie weiterentwickelt wurde, wurde für den besseren Austausch von Audiodaten zwischen mehreren Computerplattformen konzipiert.

Zusätzlich bietet das Containerformat BWF die Möglichkeit, Metadaten abzuspeichern, sowie auch einen plattformunabhängigen Timecode einzubetten.

Die Tags, die dabei hinzugefügt werden können, sehen jedoch auch lediglich identifizierende Metadaten vor und sind nicht für semantische Eigenschaften vorgesehen:

Diese Tags wären:

Description, Originator, OriginatorReference, OriginationDate, OriginationTime, TimeReference, Version (des BWF), UMID (einmalige Nummer zur Identifizierung), Reserved, Coding History.⁸

⁸vgl. AudioHQ (2005): Tags und Tagging – Formate
<http://www.audiohq.de/index.php?showtopic=22>

MPEG7:

MPEG7 stellt einen neuen Multimedia Content Description Standard dar und wurde von der Moving Picture Experts Group entwickelt. Er stellt einige Deskriptoren bereit, die Eigenschaften von Audio- bzw. Videomaterial beschreiben sollen. Eine spezielle Sprache, die „Description Definition Language“ ermöglicht das Hinzufügen weiterer Deskriptoren. „Description Schemes“ beschreiben letztlich die Struktur und die Semantik der Deskriptoren. Die vom MPEG7 – Standard vordefinierten Audio Deskriptoren basieren auf elementaren Algorithmen der Audioanalyse und sind Low-Level Deskriptoren. Jedoch wurden bereits mehrere Ansätze entwickelt, um aus diesen Low-Level - Informationen High-Level - Deskriptoren abzuleiten. Diese bestehen folglich aus textlichen Beschreibungen bzw. quantisierten Werten, die aus der Low-Level – Extraktion hervorgehen.

Folgende Low-Level - Deskriptoren sind vom Standard vorgesehen:

Descriptor Gruppen	Basic	Basic Spectral	Signal Parameter	Timbral Temporal	Timbral Spectral	Spectral Basis
Descriptor	Audio Waveform	Audio Spectrum Envelope	Audio Fundamental Frequency	LogAttack Time	Harmonic Spectral Centroid	Audio Spectrum Basis
	AudioPower	Audio Spectrum Centroid	Audio Harminicity	Temporal Centroid	Harmonic Spectral Deviation	Audio Spectrum Projection
		Audio Spectrum Spread			Harmonic Spectral Spread	
		Audio Spectrum Flatness			Harmonic Spectral Variation	
					Spectral Centroid	

Abb 1: MPEG7 Standard mit Deskriptoren

Hier ein kleiner Auszug aus einem MPEG7 – xml – File, welches aus der Extraktion eines kurzen Glockensounds entstanden ist:

```
<DescriptionUnit xsi:type="AudioSpectrumFlatnessType" loEdge="250" hiEdge="16000">
  <SeriesOfVector hopSize="PT10N1000F" vectorSize="21" totalNumOfSamples="64">
    <Raw mpeg7:dim="64 21">
      0.5666 0.7668 0.6673 0.5891 0.6397 0.6393 0.8119 0.604 0.7694 0.2487 0.196 0.05455 0.5123 0.7109 0.8002
      0.8629 0.8603 0.948 0.6375 0.8786 0.9598
      0.9263 0.7833 0.5216 0.6397 0.725 0.6345 0.8029 0.6452 0.8104 0.1987 0.03965 0.007968 0.5243 0.6611 0.4888
      0.8761 0.871 0.9615 0.5603 0.9113 0.9289
      0.4822 0.3374 0.5636 0.7288 0.778 0.7553 0.6856 0.522 0.6086 0.5604 0.005113 0.001168 0.8399 0.9151 0.5495
      0.8907 0.9701 0.9354 0.7237 0.9339 0.9441
      0.4711 0.7685 0.9193 0.8652 0.6584 0.7054 0.8087 0.7121 0.8415 0.562 0.003417 0.001166 0.8326 0.9165 0.7481
      0.8595 0.8771 0.8161 0.6529 0.666 0.6814
      0.7239 0.7945 0.8821 0.9323 0.9685 0.9478 0.9431 0.9187 0.9122 0.8379 0.004413 0.001283 0.8171 0.8409 0.8125
      0.862 0.5611 0.4313 0.5384 0.5555 0.6087...
    </Raw>
  </SeriesOfVector>
</DescriptionUnit>
```

Durch die Audioanalyse und schließlich der Extraktion dieser Deskriptoren entsteht ein Feld aus Vektoren, Float – Values und Werte – Matrizen, die als solches lediglich eine unlesbare Anreihung von Analysewerten sind. Durch Vergleiche der Vektoren mit anderen aus weiteren Soundsamples entstandenen Vektoren, können jedoch zB: audioteknisch – ähnliche Sounds auf einer mehrdimensionalen Map projiziert und gesammelt werden.

Radoslav Ivanov beschreibt weiters den Einsatz von 5 Audiofeatures, welche als High – Level Audio Description Tools gelten und aus den Low-Level Deskriptoren bereits Aussage über musikalische Qualitäten der Audiofiles machen. Diese Qualitäten sind Audio Signature, Musical Instrument Timbre, Melody Description, General Sound Recognition und Spoken Content.

Nachdem die technische Analyse der Audiofiles in dieser Projektphase noch nicht sehr weit fortgeschritten und auch kein Hauptthema dieser Diplomarbeit

2.2. Metadaten im Audiobereich – Formate

ist, soll hier nun auch nicht näher darauf eingegangen werden. In dieser Arbeit soll jedoch noch ein großes Augenmerk auf Metadaten gelegt werden, die weniger durch mathematische Algorithmen entstehen, sondern eher durch sinnvolle manuelle Vergabe von Metadaten.

2.3 Entstehungsarten von Metadaten

Es gibt einige verschiedene Möglichkeiten, wie man zu sinnvollen Metadaten von Audiofiles kommt. Anhand von Beispielen soll dieses verdeutlicht werden:

Dezentrale Vergabe von Metadaten durch eine Community:

Bei einem web – basiertem Tool zur Suche nach Audiofiles kann man darauf bauen, dass man bei entsprechender Anzahl an Usern zu einer Menge an Metadaten kommt, ohne dafür eigene Metadaten - Administratoren zu benötigen. Ein Beispiel für ein Web – Tool, das darauf beruht, dass jeder registrierte User Metadaten zu Songs vergeben kann, ist z.B. das Audioscrobber – Tool:

Das Audioscrobber - Plug In auf last.fm ist ein Tool, das den persönlichen Musikgeschmack eines Users erfasst, um danach z.B. Vorschläge für weitere Downloads zu erstellen, und Kontakt zu Usern mit ähnlichen Geschmack (Neighbours) zu ermöglichen.

2.3. Metadaten im Audiobereich – Entstehungsarten von Metadaten



Abb 2: last.fm

Es handelt sich um ein PlugIn, welches für Windows Media Player, Winamp, iTunes, XMMS etc. funktioniert. Nach der Installation und Konfiguration des PlugIns schickt es beim Abspielen jedes Titels alle ID3-Tags an den Audioscrobbler - Server. So wird das musikalische Profil des Users immer wieder neu aktualisiert. Jeder User hat weiters eine eigene Page auf last.fm, wo seine gehörten Titel aufgelistet sind, und er seine persönliche Hörerstatistik abrufen kann. Danach macht das System Vorschläge für weitere "benachbarte" Musikstücke. Es werden also die "Hörergewohnheiten" am Server gespeichert und beim nächsten Einstieg auf die persönliche Seite wieder abgerufen.

Wenn zur Zeit keine Internet - Verbindung zum Audioscrobbler - Server besteht, werden die gespielten Tracks in einem lokalen Cache gespeichert und bei der nächsten Verbindung aktualisiert.

Nachdem man ca. 300 Tracks gehört hat, analysiert das Tool die Hörergewohnheiten und findet "neighbours", dh. andere registrierte Personen, die einen ähnlichen Musikgeschmack haben.

2.3. Metadaten im Audibereich – Entstehungsarten von Metadaten

Das Tool ist weiters mit üblichen Community - Features wie "friends hinzufügen", "Gruppen erstellen", "persönliche Charts erstellen" etc. ausgestattet.

Es gibt auch die Möglichkeit, tag-basiert nach Bands, Titeln, Usern etc. zu suchen.

Weiters gibt es noch den Last.fm Radioplayer, welcher Playlisten nach dem Userprofil und Geschmack ermittelt und abspielt.

Hinter dem Audioscrobbler – Tool steckt die große Musikdatenbank v. MusicBrainz (musicbrainz.org). Metadaten zu Songs und Artists werden hier von den Usern händisch eingetragen.

Das ist ein Community – Portal mit derzeit ca. 220 000 Usern, 250 000 eingetragenen Artists und über 4,8 mio. Songs. Nach Registrierung auf der Webseite kann man sich die Software „Tagger“ runterladen. Mit diesem Tool kann man selbst Artists und Songs hinzufügen und mit Metadaten versehen. Songs und Artists können auch mit ähnlichen Songs und Artists verknüpft werden. Auf diese Weise werden diese ähnlichen Artists nach Anzahl der Verknüpfungen unterschiedlich stark verbunden. Für die Eingabe von Metadaten gibt es gewisse Style Guides, um die Einheitlichkeit der Daten zu gewähren.

Ein dezentrales System zur Vergabe von Metadaten, wie es bei Musicbrainz.org der Fall ist, hat gewisse Schwächen, aber auch Vorteile:

Eher unbekannte oder neue Songs werden meist nur von Usern einmalig beschrieben und haben daher eine sehr subjektive Charakterisierung. Je bekannter der Song ist, desto mehr User legen ihre Metadaten dazu ab und man kann quasi dann auf objektivere Durchschnittswerte zurückgreifen.

2.3. Metadaten im Audiobereich – Entstehungsarten von Metadaten

Bei Systemen dieser Art muss man auch mit einem gewissen Graubereich rechnen, bei dem Metadaten von einzelnen Usern unzureichend oder auch unpassend vergeben wurden.

Zusätzlich ist es nur schwer möglich, eine Einheitlichkeit der Metadaten zu gewährleisten, wenn eine riesige Anzahl an Personen alle auf ihre eigene Art und Weise diese Daten vergibt.

Ein großer Vorteil bietet natürlich die stetig mit den Usern wachsende Anzahl an Sounds und Metadaten, ohne jegliche zentrale Administration dieser Daten.

Manuelle Vergabe von Metadaten durch Administratoren:

Eine andere Möglichkeit bietet die Erschaffung von Audio – Metadaten durch eingeschulte Administratoren. Damit verbunden ist natürlich ein Verwaltungsaufwand, da jeder Song bzw. Sound durch ein Team einzeln kategorisiert werden muss. Vor allem ist es auch ein permanenter Verwaltungsaufwand, sofern man stets aktuelle Musik in der Datenbank haben will. Andererseits sind diese Metadaten in einheitlicher Form und man kann ihnen bei guter Administration eher vertrauen als bei einem dezentralen System.

Ein Beispiel für diese Art der Metadaten – Vergabe ist das Tool „Pandora“ auf pandora.com:

2.3. Metadaten im Audiobereich – Entstehungsarten von Metadaten



Abb 3: pandora.com

Hierbei handelt es sich um ein personalisiertes Webradio von „Music Genome Project“. Das Webradio hat ein Flash Frontend, wo man einen Titel oder Interpret eingeben kann, und das Tool stellt daraus ein 10 Std. Radioprogramm zusammen, welches aus ähnlichen Stücken wie das Eingeebene besteht. D.h. aus einem ausgewählten Song oder Artist wird automatisch die Geschmacksrichtung des Users berechnet und ähnliche Songs gespielt.

Dahinter steckt eine Datenbank aus ca. 300 000 Songs. Diese Songs werden Track für Track von Teammitgliedern mit Metadaten versehen.

Mit ca. 400 verschiedenen Attributen werden die Songs versehen, wobei für die Stimme alleine ca. 30 Attribute vorgesehen sind. Ansonsten beinhalten diese Attribute musikalische Qualitäten wie etwa Melodie, Harmonie, Rhythmus, stimmliche Performance und Texte etc.

Diese Attribute werden bei einer Anfrage mit den Eigenschaften anderer Songs verglichen und ein Algorithmus bestimmt die Songs, die dem zuerst ausgewähltem Song am ähnlichsten sind.

2.3. Metadaten im Audiobereich – Entstehungsarten von Metadaten

Voll – automatisierte Vergabe von Metadaten

Bei Musicsurfer, ein Sound – Search Tool der Music Technology Group der Universität Pompeu Fabra in Barcelona handelt es sich im Prinzip um ein vollautomatisiertes Tool zur Suche nach Musik:

The screenshot shows the 'UPF - MusicSurfer Demo' interface. It features three main panels: Genres, Artists, and Albums. The 'Artists' panel has 'Rapoon' selected. Below these panels is a table of tracks with columns for #, Title, Album, Artist, Time, and Genre. To the right of the table, there are buttons for 'Similar Songs' and 'Similarity Map' for each track.

#	Title	Album	Artist	Time	Genre
1	Refraction	Cidar	Rapoon	8:38	Ambient
2	Borderlands	Cidar	Rapoon	4:08	Ambient
3	Black Feel	Cidar	Rapoon	2:37	Ambient
4	Estuary	Cidar	Rapoon	4:39	Ambient
5	Red Water	Cidar	Rapoon	6:04	Ambient
6	Noord	Cidar	Rapoon	6:09	Ambient
7	Goost	Cidar	Rapoon	5:46	Ambient
8	Cimbala	Cidar	Rapoon	9:43	Ambient
9	Lake	Cidar	Rapoon	6:50	Ambient
1	Sign	Easterly 6 or 7	Rapoon	7:08	Ambient
2	Our trespasses	Easterly 6 or 7	Rapoon	4:45	Ambient
3	Blow (lethal blow)	Easterly 6 or 7	Rapoon	5:30	Ambient
4	All criminals	Easterly 6 or 7	Rapoon	4:28	Ambient
5	Waters reaching	Easterly 6 or 7	Rapoon	8:49	Ambient
6	Falling more Slowly	Easterly 6 or 7	Rapoon	16:41	Ambient
7	6 or 7	Easterly 6 or 7	Rapoon	6:38	Ambient
8	Variable 2	Easterly 6 or 7	Rapoon	7:52	Ambient
9	Rattling Sabers	Easterly 6 or 7	Rapoon	13:13	Ambient
1	Sanctum	Fallen Gods	Rapoon	9:09	Ambient

Abb 4: Musicsurfer

Es extrahiert aus einem vorgegebenen Pool an Songs per automatischer Audio - Analyse Beschreibungen zur Musik bezüglich Harmonien, Rhythmus, Instrumentierung etc. Weitere Vergabe von Metadaten von Menschenhand ist nicht notwendig.

Mit Hilfe eines Similarity – Algorithmus und der Similarity – Map werden ähnliche Musikstücke automatisch vorgeschlagen und nach dem Maß der Ähnlichkeit angereicht.

2.3. Metadaten im Audibereich – Entstehungsarten von Metadaten

Die Vorgangsweise als User ist folgende:

Man kann sich mit Hilfe der Attribute Genre, Artist und Album durch ein Pool an Musikstücken durchbrowsen und sich den Einstellungen entsprechend die einzelnen Musikstücke in einer Liste anzeigen lassen.

Nun hat man die Möglichkeit, sich für jeden Song eine Liste mit ähnlich – klingenden Songs anzeigen zu lassen. Dieses Ähnlichkeitsmaß wird aus rein audiotecnischen Analysen und der automatischen Extraktion von Metadaten ermittelt.

Ausserdem gibt es eine Similarity – Map, die die verwandten Songs übersichtlich auf einer grafischen Oberfläche anzeigt. Der ursprünglich ausgewählte Song steht in der Mitte und die ähnlich – klingenden Musikstücke werden nach auswählbaren Kriterien von der Mitte nach außen hin angeordnet. Jeder „Kreis“ steht für einen bestimmten Song, wobei beim Rollover mit der Maus über den „Kreis“ die genauen Daten zum Song angezeigt werden:

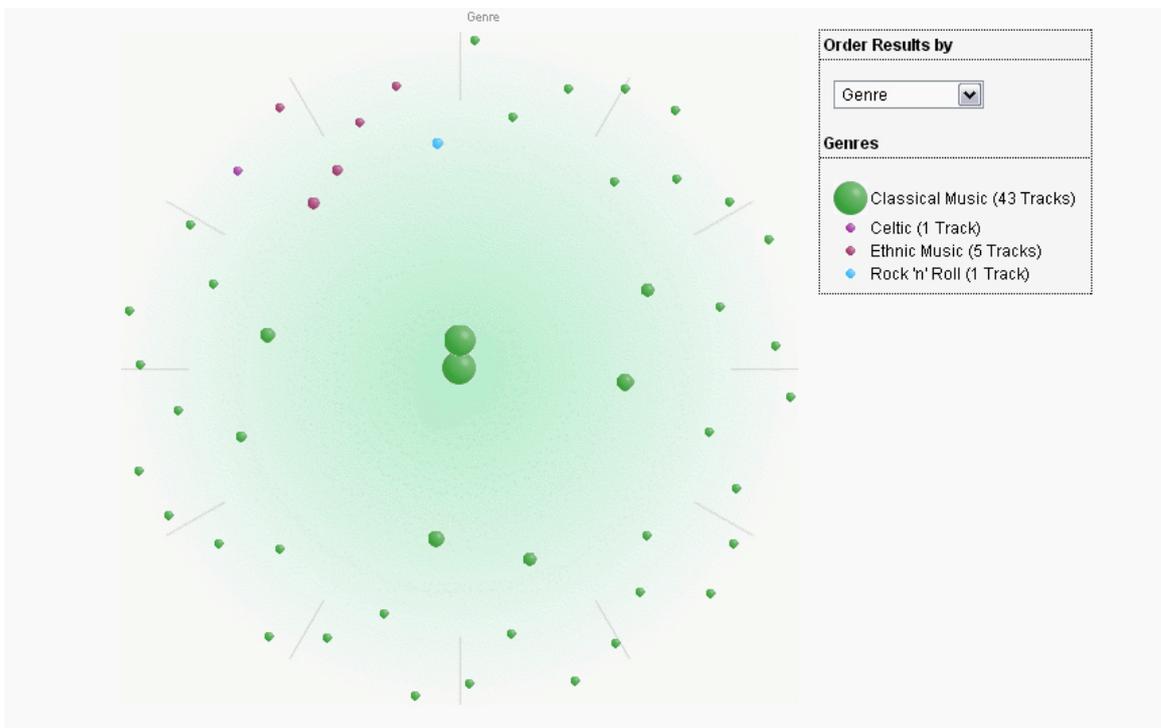


Abb 5: Musicsurfer – Similarity Map

2.3. Metadaten im Audibereich – Entstehungsarten von Metadaten

Ein weiteres Feature von Musicsurfer ist „Query By Example“:

Man kann hierbei eine Audiodatei in 3 verschiedenen Formaten (mp3, ogg oder wav) uploaden. Diese Datei wird anschließend server – seitig analysiert und mit Metadaten versehen. Nun hat man erneut die Möglichkeit, zu dem hochgeladenen Sound ähnliche Sounds aus dem vorgegebenen Pool anzeigen zu lassen.

Halb - automatisierte Vergabe von Metadaten

Die ultimative Parallele zu Suchmaschinen wie Google, aber auf die Suche nach Sounds bezogen, ist z.B. das Webtool auf findsounds.com:

FindSounds
Search the Web for Sounds

Search for [Help](#)
[Need Examples?](#)

File Formats	Number of Channels	Minimum Resolution	Minimum Sample Rate	Maximum File Size
<input checked="" type="checkbox"/> AIFF	<input checked="" type="checkbox"/> mono	8-bit	8000 Hz	2 MB
<input checked="" type="checkbox"/> AU	<input checked="" type="checkbox"/> stereo			
<input checked="" type="checkbox"/> WAVE				

Sounds 1-10 of 200 labelled "gun"

- 
  <http://www.pic00.com/sound/pao/RICO6.WAV>
gun
7k, mono, 8-bit, 11025 Hz, 0.7 seconds ([show page](#) | [e-mail this sound](#))
- 
  <http://sep800.mine.nu/files/sounds/loadclip.wav>
gun
3k, mono, 8-bit, 8000 Hz, 0.4 seconds ([show page](#) | [e-mail this sound](#))

Abb 6: Findsounds

2.3. Metadaten im Audiobereich – Entstehungsarten von Metadaten

Dieses Tool ist von der Weise wie hierbei Metadaten entstehen eine Art Mischung aus dem manuellen System von pandora.com und dem automatisierten Musicsurfer – Tool.

Findsounds konzentriert sich ganz auf kürzere Samples und Effekte. Dabei kann man im gesamten Web nach Audiofiles in den Formaten AIFF, AI und WAVE und mit unterschiedlichen Abtastraten, Filegrößen und Auflösungen (8-bit, 16-bit) etc. suchen.

Der User kann über ein Textfeld mit einem oder mehreren Keywords nach den Beschreibungen der Sounds suchen. Nach Auflistung der Sounds kann er sich für jeden Sound wieder ähnliche Klänge auflisten lassen. Diese ähnlichen Klänge werden von einem automatisierten Algorithmus ermittelt, der rein audiotekhnische Ähnlichkeiten heranzieht, unabhängig von Filenamen oder Beschreibungen rund um den Sound – Link.

Die Suche nach den Sounds ist in diesem Fall also einerseits textbasiert, d.h. ein Eingabefeld nach einem Schlüsselwort genügt, um eine Reihe von Sounds zu erhalten. Andererseits gibt es auch die interessante Möglichkeit, die Klangfarbe und andere audiotekhnische Eigenschaften an sich als Suchparameter heranzuziehen.

Hinter diesem Tool steckt der Findsounds – „Spider“, der permanent das gesamte Web nach neuen Sounddateien absucht und Links zu neuen Sounds in einer Datenbank ablegt. Jeder Sound wird vom „Spider“ heruntergeladen und analysiert. Da die Dateinamen der Sounds, sowie auch der Text rund um den Link für zuverlässige Metadaten nicht genügen würden, werden die Sounds einzeln abgehört und von Teammitgliedern mit Metadaten versehen. Sounds, die mit unserer Sprache nur schwer zu beschreiben sind, werden ohne wörtliche Beschreibung belassen, aber für die Ähnlichkeiten – Suche ebenfalls herangezogen.

Die Sounds werden also auch audiotekhnisch analysiert und mit klanglich - ähnlichen Sounds verknüpft. Oft mag die rein textbasierte Suche genügen,

2.3. Metadaten im Audibereich – Entstehungsarten von Metadaten

sofern die benötigte Schallquelle schon vorher bekannt ist. Aber zum Entdecken neuer Sounds mit gewissen Klangqualitäten kann die Ähnlichkeiten – Suche für Sounddesigner bestimmt sehr sinnvoll sein.

Es kann z.B. das Geräusch eines Elefanten sehr ähnlich klingen wie ein bestimmtes Saxophon – Sample. In einer bestimmten Szene kann es etwa ein tolles Stilmittel sein, wenn man unter dem Bild eines Elefanten eben nicht den eigentlichen Schrei des Elefanten legt, sondern ein Saxophon – Sample, das sich allmählich in ein ganzes Instrumentalstück verwandelt.

Personalisierte Metadaten - Vergabe durch den Sounddesigner

Für den Sounddesigner, der sich mit der Zeit eine große Bibliothek an Sounds zusammengestellt hat, ist es ungemein wichtig, dass er stets schnell den richtigen Sound finden kann. Dafür ist es sinnvoll, wenn der Sounddesigner seine eigenen Metadaten zu seinen Sounds auf der Festplatte hinzufügen kann.

Tut er sich einmal die Arbeit an, alle Sounds zu beschreiben, zu kategorisieren und in sinnvollen Gruppen einzuteilen, dann kann er sich später bei der Produktion immens viel Zeit ersparen.

Eine Software, die genau diese Funktion erfüllt ist das Soundminer Audio File Management System der kanadischen Firma Sound Ideas:

2.3. Metadaten im Audiobereich – Entstehungsarten von Metadaten



Abb 7: Software Soundminer

Dieses weitgehend plattformunabhängige Datenbank - System wurde speziell für die Erleichterung der Suche nach den eigenen Soundfiles konzipiert. Es unterstützt gängige Audioformate wie WAVE, AIFF, MP3, AAC (mpeg4) und übernimmt dabei auch bereits vorhandene Metadaten aus den bestehenden Files wie z.B. die ID3 - Tags der MP3 - Files und die Broadcast Wave Metadaten.

Bis zu 50 Metadaten - Felder können pro File hinzugefügt werden. Als Datenbankquelle kann sowohl eine lokale Festplatte, als auch ein Netzlaufwerk benutzt werden. Neben den Datenbank - Funktionen wie Import, Export, Datenverschmelzung und die Suche nach Duplikaten gibt es auch einen internen Wave - Editor sowie einen Waveform - Überblick und diverse Stapelverarbeitungsmöglichkeiten.

Auch hier gibt es über den internen Sound - Digger die Möglichkeit, nach ähnlich - klingenden Audiofiles zu suchen.

2.3. Metadaten im Audiobereich – Entstehungsarten von Metadaten

Die Oberfläche und die Datenbank – Eigenschaften sind sehr individuell modifizierbar. Ein VST – Rack für die Audibearbeitung ist ebenfalls inkludiert, wobei nach Bearbeitung stets das Originalfile unberührt bleibt und die erledigten Schritte auch nach langer Zeit noch zurückzuverfolgen sind. Die Audibearbeitung erfolgt also nicht – destruktiv.

3 Klangbeschreibung,- und Kategorisierung

Bis jetzt wurde in dieser Arbeit hauptsächlich darauf eingegangen, in welcher Form bzw. in welchen Formaten und mit welchen Methoden Metadaten von Audiodateien entstehen können. Nun soll auch näher darauf eingegangen werden, wie und mit welchen sprachlichen Mitteln diese Audiodateien beschrieben und eingeteilt werden können.

3.1 Beschreibung von Klangobjekten

Barbara Flückinger setzte sich in ihrem Buch „Sounddesign, die virtuelle Klangwelt des Films“, erschienen im Schüren Presseverlag, intensiv mit der Beschreibung von Klangereignissen, insbesondere auch mit Sounddesign im Film auseinander.

Ihrer Meinung nach ist die ungenügende sprachliche Formulierung sowohl Folge als auch Grund unserer mangelnden Wahrnehmungsschulung. Diese beiden Dinge würden sich gegenseitig bedingen, d.h. umso differenzierter die sprachliche Beschreibung von Klängen ist, desto feiner wird auch die Wahrnehmung von ihnen.

Um diese Wahrnehmung zu schulen und eine Analyse möglich zu machen, ist es notwendig, die subjektiven Empfindungen beim Hören in deklaratives Wissen – also sprachlich zugängliche Bewusstseinsinhalte – umzuwandeln und sie damit allgemein zugänglich zu machen. Die Versprachlichung von Klangwahrnehmungen dient also einerseits der eigenen Erkenntnis und andererseits der Kommunikation darüber.

3.1. Klangbeschreibung,- und Kategorisierung – Beschreibung von Klangobjekten

Barbara Flückiger hat ein modulares Beschreibungssystem entwickelt, um eine solche Versprachlichung zu ermöglichen. Es orientiert sich an dem informationstheoretischen Modell der Differenzierung und Auflösung, um die klangbezüglichen Empfindungen möglichst einfach zu beschreiben. Bei diesem Modell werden mehrere Fragen an die Klangobjekte gestellt, aus deren Antworten sich die einzelnen Module bilden, die wiederum ihrerseits der Klangbeschreibung dienen. Bei der Beantwortung der Fragen erweist es sich als sinnvoll, auf konventionelle umgangssprachliche Begriffe zurückzugreifen, da es uns – wie bereits erwähnt – an einem spezifischen Wortschatz mangelt.

„Die fünf Module der Klangobjektbeschreibung setzen sich aus folgenden Fragen zusammen:

Was klingt? Was bewegt sich? Welches Material klingt? Wie klingt es? Wo klingt es?

1. Was klingt?

Anthropologisch betrachtet war es schon immer von großer Bedeutung, anhand von Geräuschen auf ihre Quelle schließen zu können, da man nur dadurch zwischen „Beute“ und „Feind“ unterscheiden konnte. Heutzutage spielen solche automatisierten Formen der Verhaltensteuerung, beispielsweise im Straßenverkehr, eine wichtige Rolle.

Dieses erste Modul der Klangbeschreibung dient allein der Ermittlung des Geräuschverursachers. Allerdings ist es nur eine Durchgangsstation, denn für eine analytische Zuordnung des Klangs reicht dieses Modul nicht aus. Die Genauigkeit, mit der man eine Klangquelle sprachlich erfassen kann, hängt sowohl vom spezifischen Erfahrungshintergrund und dem allgemeinen Weltwissen des Rezipienten, als auch größtenteils von der gleichzeitigen Visualisierung der Klangquelle ab. Allerdings bleibt die Quelle in manchen Fällen trotz allem unidentifizierbar.

3.1. Klangbeschreibung,- und Kategorisierung – Beschreibung von Klangobjekten

2. Was bewegt sich?

Das zweite Modul bezieht sich auf die Prozesskomponente einer Klangquelle, denn Klang ist immer untrennbar an Bewegung gebunden. Wenn der Geräuschverursacher, bei Flückiger auch als Agens bezeichnet, ein Lebewesen ist, ist die Prozessbeschreibung meist sehr einfach, da es hauptsächlich Tätigkeiten oder sprachliche sowie nichtsprachliche stimmliche Äußerungen sind.

Dennoch entstehen, vor allem bei menschlichen Tätigkeiten, fast unbegrenzte Variationen der Klangerzeugung. Allerdings ist das handelnde Agens nicht immer gleichzeitig auch das klingende Agens. Wenn jemand z.B. eine Tür öffnet, ist es die Tür (die passiv bewegt wird), die knarrt – die also das klingende Agens ist – und nicht der Mensch (handelndes Agens), der sie bewegt.

Die Prozesse sind jedoch um einiges schwieriger zu entschlüsseln, wenn sie in der unbelebten Welt auftreten, da sie dort häufig im Inneren ablaufen: Der Lüfter summt etwa im Innern des Computers. Hauptsächlich entstehen Geräusche durch die Interaktion zweier Gegenstände. Dabei unterscheidet man, ob beide Objekte, also das treibende und das bewegte, oder nur eines von beiden in Schwingung versetzt wird. So kann es sein, dass man bei Schritten sowohl das Schuhmaterial (z.B. Leder) als auch den Untergrund (etwa den Holzfußboden) hört. Dadurch erfährt man nicht nur etwas über das Agens oder den ablaufenden Prozess, sondern auch maßgeblich etwas über die Materialien der Objekte, was uns zum dritten Modul bringt:

3.1. Klangbeschreibung,- und Kategorisierung – Beschreibung von Klangobjekten

3. Welches Material klingt?

Das Geräusch eines Klangs verrät, mehr noch als über die Art der Quellen, etwas über deren materielle Beschaffenheit. Material wie Stein, Holz, Wasser oder Papier hinterlassen so eindeutige (Klang-) Spuren, dass diese klingenden Materialien besonders leicht zu beschreiben sind. Materialien erzeugen häufig so markante Klänge, dass man teilweise selbst den Aggregatzustand eines Materials bestimmen kann. Wasser ist beispielsweise als Dampf, als Regen oder auch als Eis hörbar.

4. Wie klingt es?

Im Gegensatz zur materiellen Beschaffenheit lassen sich die klanglichen Eigenschaften eher schwer sprachlich ausdrücken. Deshalb eignen sich hauptsächlich lautmalerische Ausdrücke, wie 'blubbernd' oder 'zischend', um klangliche Aspekte sprachlich greifbar zu machen. Bei ihnen besteht eine Verwandtschaft zwischen Form und Inhalt, weil der Ausdruck selbst eine bildhafte Empfindung erzeugt.

Die bereits anfänglich erwähnte hohe Subjektivität bei der Beschreibung von Klängen ist besonders bei diesem Modul sehr problematisch, da es keine sachlichen Kriterien gibt, um z.B. Wörter wie 'brausend' oder 'tosend' voneinander zu unterscheiden. Noch komplizierter wird eine eindeutige und allgemeingültige Beschreibung dadurch, dass es selbst bei der Verwendung eines Wortes durch ein und dieselbe Person zu Schwankungen kommen kann.

Ein weiteres Problem ist, dass manche Wörter – wie 'scheppernd' oder 'ächzend' – sowohl Prozesse als auch eine gewisse Klanglichkeit beschreiben. Weiterhin kann z.B. der Frequenzbereich so auffällig sein, dass er als 'bassig', 'mittig' oder 'hoch' in die Klangbeschreibung einfließt. Genauso kann durch die Beschreibung als 'laut' oder 'leise' die Lautstärke in eine Objektbeschreibung integriert werden. Als 'hektisch', 'regelmäßig' oder 'beschleunigt' werden

3.1. Klangbeschreibung,- und Kategorisierung – Beschreibung von Klangobjekten

Klänge bezeichnet, die durch einen auffälligen Rhythmus geprägt sind. Auch eine Zuordnung zu technischen Geräten, wie Fernseher, Radio oder Telefon, kann erfolgen, wenn sie klangentscheidend ist.

5. Wo klingt es?

Mit dem letzten Modul wird die räumliche Situierung von Klangobjekten beschrieben und damit ihr Ort sowie die akustischen Spuren verschiedener Räume. Dabei drückt sich der Ort der Klangerzeugung als räumliche Distanz aus, wobei das Objekt etwa als 'nah' oder 'fern' bzw. als 'im Vordergrund' oder 'im Hintergrund' beschrieben wird. Allerdings ist die Entfernung nur bedingt eine Funktion der Lautstärke, denn mit zunehmender Distanz wird der Klang mittiger.

Des Weiteren verändert sich dadurch auch das Verhältnis von direktem und diffusem Schall: die diffusen Schallanteile nehmen zu und werden als Hall hörbar. Die von einem Klangobjekt erzeugten Schallwellen überwinden nicht nur Distanzen, sondern dringen z.B. auch durch Wände und machen damit diese Hindernisse ebenfalls hör- und beschreibbar. Im Stereoklangbild ist es möglich den Klang auf der Horizontalen von links nach rechts wandern zu lassen. All diese fünf Module dienen einzig und allein einem Ziel, der Identifikation einer Quelle.

Identifikation einer Quelle

Die Genauigkeit, mit der wir einen Klang bestimmen können, hängt einerseits von der sinnlichen Erfahrung des Klangs ab, als auch von der Relevanz, die diesem Klang im Leben des Einzelnen zukommt. Ein Hühnerzüchter könnte so vielleicht 20 verschiedene Hähne an ihrem Krähen unterscheiden, was anderen nie gelingen würde.

3.1. Klangbeschreibung,- und Kategorisierung – Beschreibung von Klangobjekten

Unsere auditive Wahrnehmung ist dort differenzierter, wo ein reichhaltiger Erfahrungshintergrund zur Verfügung steht. Daraus ergibt sich, dass unsere genaueste Wahrnehmung die menschliche Stimme betrifft. Daher erkennen wir z.B. unsere Mutter am Telefon – also ohne gleichzeitige Visualisierung – sofort, obwohl sie vielleicht erkältet ist oder nebenbei einen Apfel isst.

Um ein Objekt auch in schwierigen Situationen so genau zu erkennen, muss ein Bündel von verschiedenen Merkmalen in unserem Gehirn gespeichert sein. Ein Grundproblem bei der Identifikation einer Quelle ist, dass wir sie durch alle momentanen Störungen hindurch wahrnehmen und erkennen müssen. Das Vorhandensein spezifischer, physikalischer Reizeigenschaften, die in einer regelhaften, typischen Gruppierung auftreten, sind die Minimalvoraussetzung für das Erkennen einer Klangquelle.

Beispielsweise besteht eine quietschende Tür aus einer Verknüpfung von akustischen Merkmalen – wie Frequenzverteilung, zeitliche Ausdehnung oder Lautstärke –, die sie eindeutig beschreiben. Diese Merkmale unterscheiden sich in obligatorische (unerlässliche) und fakultative (zusätzliche). Dabei wären die obligatorischen in unserem Beispiel die Merkmale, welche die Tür eindeutig als Tür identifizieren.

Die fakultativen Merkmale hingegen würden hier die Tür z.B. als eine Holztür (materielle Eigenschaft) in einer Kirche (räumliche Eigenschaft) definieren. Ein weiteres Stilmittel der Filmemacher ist das unidentifizierte Klangobjekt (UKO).

Das unidentifizierte Klangobjekt (UKO)

Als Hauptcharakteristik eines UKOs gilt die unterbrochene Verbindung zum Klangobjekt, die dadurch geschaffen wird, dass die Klangquelle weder im Bild sichtbar ist noch aus dem Kontext zu erschließen wäre. Das hat die Erhaltung der Mehrdeutigkeit zum Ziel.

3.1. Klangbeschreibung,- und Kategorisierung – Beschreibung von Klangobjekten

In den klassischen Hollywood-Filmen wurde dieses Stilmittel nur sehr selten verwendet, da es aufgrund der damals noch nicht ausgereiften Technik leicht als Fehler interpretiert werden konnte. Ohne zusätzliches Bild war es damals sozusagen fast unmöglich, beispielsweise Klatschen von Regen zu unterscheiden.

Die Regel 'see a dog – hear a dog' war die damalige vorherrschende Auffassung der Filmemacher. Außer in *Modern Times* von Charlie Chaplin und *Citizen Kane* von O. Welles fand man UKOs in der klassischen Periode noch nicht. Erst mit dem Fortschritt der Technik und der damit verbundenen Verfeinerung der Auflösung kam es zu einer genaueren klanglichen Definition. In Folge dessen fanden UKOs immer öfter – vor allem in Science-Fiction-Filmen – Verwendung.

Jedoch kann es immer noch passieren, dass diese unidentifizierbaren Klangobjekte unidentifizierbar bleiben und damit als eine Störung eingestuft werden. Solche unerwünschten Störungen können den technischen Apparat hinter dem Gesamtwerk sichtbar machen, und dadurch könnte der Film als Artefakt entlarvt und die gesamte Illusion zerstört werden. Diese Störungen minimieren sich allerdings mit der ständig zunehmenden Verfeinerung der Auflösung. So ist zu erklären, dass UKOs in zeitgenössischen Science-Fiction-, Action- und Katastrophen-Filmen bis zu 20% aller wahrnehmbaren Tonelemente ausmachen. Dort werden UKOs gezielt zur Frustration der Zuschauer eingesetzt, indem sie in ihnen ein Gefühl von Ohnmacht oder auch Angst erzeugen. Denn am bedrohlichsten werden immer die Geräusche eingestuft, die sich auch nach der gewohnten Identifikationszeit nicht einer bestimmten Quelle zuordnen lassen.“^{9,10}

⁹Schaffartzik, S. (2004): Beschreibung von Klangobjekten
<http://server4.medienkomm.uni-halle.de/filmsound/kap1-2.1.htm>

¹⁰vgl. Flückinger, B. (2001): Sounddesign. Die virtuelle Klangwelt des Films, Schüren Presseverlag

3.2 Evaluierung der bestehenden Soundlibraries

Die für das Projekt „AllThatSounds“ angekauften Sample - CDs sind in unterschiedlichster Art und Weise eingeteilt und strukturiert. Die Benennung der einzelnen Samples basiert auf unterschiedlichsten Methoden.

Insgesamt wurden 14 CDs angeschafft. Bei den Namen der Sounddateien bzw. der Unterordner, in denen sie sich befinden, wurde bei den meisten CDs versucht, bereits wichtige Informationen über den Sound preiszugeben.

Dabei gibt es die unterschiedlichsten Ansätze zur Einteilung und Benennung der Klänge.

Allgemein findet man in Soundlibraries folgende verschiedene Ansätze zur Beschreibung bzw. Kategorisierung von Sounds:

Einteilung / Beschreibung nach:

- 1) Emotion, Stimmung
(z.B. funny, melancholic moods...)
- 2) auslösendes Objekt bzw. Material
(z.B. train, water – noise...)
- 3) Ort der Klängenstehung, Umgebung
(z.B. in the wood, space...)
- 4) Umgangssprachliche Beschreibung
(z.B. woosh, klirr...)

- 5) Anwendungsgebiet bzw. Art des Klangs
(z.B. button sounds, film atmo theme...)
- 6) Freie Assoziationen
(z.B. rhythmsphera, thrill me...)
- 7) Zeitliche Beschreibung, Tempo
(z.B. loop 240bpm)
- 8) tonale Beschreibung, Harmonien
(z.B. Bass 1G, Piano f#major)
- 9) Instrument bzw. Spielweise des Instruments
(z.B. Bass Drum1, pick guitar)
- 10) fortlaufende Nummer

Viele Sounds auf den Sample CDs sind vor allem durch Soundnamen charakterisiert, die aus freien Assoziationen entstanden sind und rein aus der Phantasie des Herausgebers entspringen. Viele dieser Dateinamen sagen leider nur sehr wenig über den tatsächlichen Klang aus.

Über Command – Lines der Software Putty wurden alle auf einen Server gestellten Ordner,- und Dateinamen der Sample CDs zuerst aufgelistet, danach einzelne Wörter gezählt.

Die Verteilung der Stichwörter sieht folgendermaßen aus:

99 mal: Loops
77 mal: Atmos
72 mal: Hits
72 mal: Explode
51 mal: Noizes
38 mal: Bonus
29 mal: Lazars
20 mal: Noise
19 mal: Snare
17 mal: Game
14 mal: ROBOT, DeformedDrumloop
12 mal: ROCKET
11 mal: PercussiveNoise
10 mal: HiHat
9 mal: Stinger, MECHAN, Drum, CathedralOneshot, Bass
8 mal: PillarDrone, LoNoise, film,
7 mal: theme, TempleDrone, STARSHIP, elements
6 mal: SHOT, layers, kit, FluxDrone, events, Door, Cracker, atmospheric,
Ambient, ALIEN, Pads
5 mal: WAVER, TombDrone, THUND, Shot, samples, movements,
MetalGlide, Looming, L.A., HARPPi, funny, Drones, DeadValley,
CRISIS, CavernSteel, BeatMutilation, atmo
4 mal: XMAS, Woodnotes, TWIND, Thunderbolt, SSPACE, SFPass, Rocket,
RAZOR, RANDOM, PHAN, OUTER, moves, Mix, MajorMist, light,
hydrau, HARPD, elevator, DRAC, Doorslide, DEEP, DANG, creatures, CI,
ASIA, ALARM
3 mal: WINDFX, WATER, VOXTRN, VOCALW, up, uneasy, UBOOTF,

3.2.Klangbeschreibung,- und Kategorisierung – Evaluierung der bestehenden Soundlibraries

Trolley, SYNATM, streets, Steelpad, StarSparkles, Space, SFSHOT, SDREAM, romance, Robotic, rhythmsphaera, PH, Perc.Loop, paranoid, NSWEEP, NOISEQ, Nimbus, NCRAK, mystic, mystery, MudHammer, movement
...

Dann gibt es auch noch unzählige Begriffe, die jeweils nur ein,- oder zweimal in den Datei,- bzw Ordnernamen der Sound CDs vorkommen.

Diese Stichwörter bieten sozusagen die einzige Möglichkeit, um einen einzelnen Sound von einem anderen ohne Anhören zu unterscheiden. Bei genauerer Betrachtung sieht man, dass man nur bei wenigen der Wörter eine vage Vorstellung bekommt, wie sich ein solcher Sound anhören könnte.

Natürlich muss man auch ganz klar zwischen den Sounds unterscheiden. D.h. es hängt sehr stark von der Art des Sounds ab, wie leicht und gut man ihn beschreiben kann.

Generell kann man sagen, je eindeutiger die Schallquelle ist, desto einfacher ist es, den Sound in Worte zu fassen. Jeder kann sich unter einem Sound, der z.B. „startendes Flugzeug“ oder „Trommelwirbel“ benannt wurde, leicht etwas vorstellen.

Schwieriger wird es eben bei abstrakteren Sounds, die eventuell elektronisch erzeugt wurden oder wo die Schallquelle nicht mehr herauszuhören ist.

Vermutlich ist auch für jeden Sound eine andere Art der Beschreibung sinnvoll. Für das „startende Flugzeug“ wäre eine technischere Beschreibung, wie etwa „Sound mit rauschhaltiger Hüllkurve und steigender Grundfrequenz von 200 Hz auf 4 kHz“ unbrauchbar. Schließlich würde man in diesem Fall ja ganz exakt nach einer klar definierten Schallquelle suchen.

Für andere Sounds könnte eine ähnliche technische oder musikalische Beschreibung jedoch sehr brauchbar oder auch die einzige Möglichkeit zur Kategorisierung sein.

3.3 Ansatz zur Kategorisierung von Klängen

In den folgenden Seiten soll ein Ansatz beschrieben werden, wie Sounds grundsätzlich eingeteilt werden können. Es handelt sich dabei nicht um eine allumfassende Einteilung in Eigenschaften, sondern eine für den praktischen Einsatz in Medienproduktionen ausgerichtete Kategorisierung.

Diese Aufgliederungsmöglichkeiten der Sounds nach verschiedenen Kriterien könnten später auf mehreren Ebenen in die Architektur des ausprogrammierten Tools einfließen.

Grundsätzlich kann man sagen, dass eine gesamt - baumartige Struktur bei der Einteilung von Sounds eher wenig Sinn macht. Zu viele, vielleicht passende Sounds würden so dieser Hierarchie zum Opfer fallen. Müsste man z.B. vorerst zwischen elektronischen und natürlichen Klängen auswählen, so würden bereits bei der ersten Auswahl zu viele Sounds wegfallen.

D.h.: Die Suche muss von mehreren Seiten begonnen werden können.

Man müsste sich einzelne Eigenschaften von Sounds zusammenstöpseln können und mit einem Ausschließungsverfahren in die Nähe des richtigen Sounds geführt werden. Bei jeder neuen Eigenschaft die man wählt, fallen entweder wieder einige Sounds weg, oder sie werden weiter nach unten gereiht, bis man letztlich durch Anhören einiger Sounds zum Gewünschten kommt.

Zusätzlich wäre eine Möglichkeit zur Gewichtung der einzelnen Eigenschaften denkbar. D.h. man kann z.B. dem Spektrum eine geringere Bedeutung als der Time zuweisen.

Hier sind alle Eigenschaften aufgezählt und kommentiert:

1. Schallquelle:

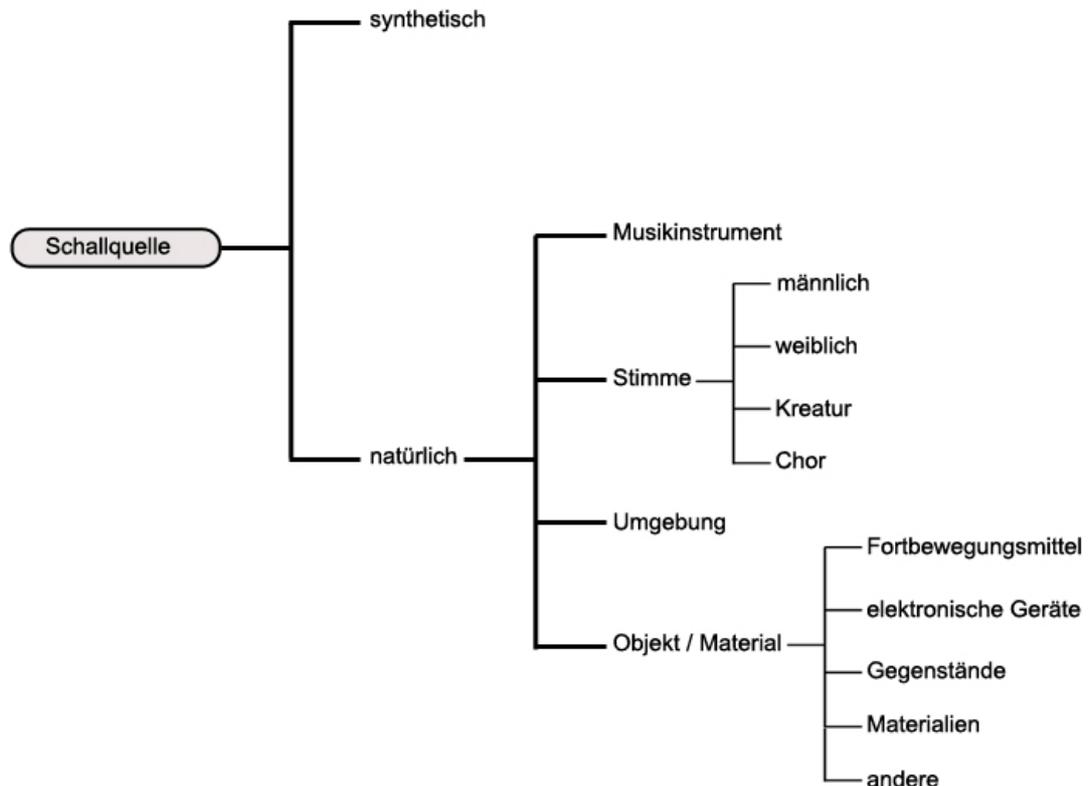


Abb 8: Kategorie „Schallquelle“

Die Schallquelle kann in erster Linie natürlicher oder synthetischer Natur sein. Damit ist der Entstehungsprozess des Schalls gemeint und weniger die Beschaffenheit des auslösenden Objekts. Dh. in „synthetisch“ würden alle Sounds reinfallen, die eindeutig von einem Synthesizer oder ähnlichen Soundmodulen stammen – also künstlich erzeugt wurden.

Alle real in der Welt vorkommenden und per Mikrophon aufgenommenen Objekte würden unter die Rubrik „natürlich“ fallen.

Die weitere Einteilung dieser natürlichen Schallquellen könnte auch vorerst z.B. nur in unterschiedliche Materialien eingeteilt werden. Für die Suche nach

3.3.Klangbeschreibung,- und Kategorisierung – Ansatz zur Kategorisierung von Klängen

einem Sound, bei dem man die Schallquelle schon vorher exakt weiß, sind jedoch andere Einteilungen geeigneter.

Die nächsten sinnvollen Auswahlmöglichkeiten wären: Musikinstrument, Stimme, Objekt/Material oder der Ort/Umgebung der Schallquelle.

Bei den Musikinstrumenten wäre eine Auflistung der Instrumente, und weiters gegebenenfalls in Spielweisen der Instrumente sinnvoll.

Unter Stimme macht eine Einteilung zwischen Mann, Frau, Chor und Kreatur Sinn.

Die Einteilung in Objekte bzw. Materialien in diesen fünf Kategorien ist zwar bei genauer Betrachtung der AllThatSounds – Soundfiles brauchbar, jedoch würde man bei der Suche nach dem Sound eines bestimmten Objekts eher eine Texteingabe vorziehen, als sich durch diesen Strukturbaum zu browsen. D.h. es müsste ein Textfeld geben, in dem man z.B. den String Flugzeug eingibt und alle entsprechenden Sounds werden aufgelistet.

Bei Umgebung ist hier der Ort der Schallerzeugung gemeint, weniger die Objekte, die diesen Schall auslösen. Sucht man beispielsweise nach einen Umgebungssound wie „Straßenlärm“, wäre es sehr unpraktisch sich einzelne Sounds von Autos und sprechenden Menschen erst übereinander legen zu müssen, damit man im Endeffekt die richtige Atmosphäre schaffen kann. Also würde man bei der Auswahl einer bestimmten Umgebung bereits die Zusammensetzung aus dort typischerweise auftretenden Schallereignissen bekommen.

2. Art des Sounds bzw. mediale Anwendung:

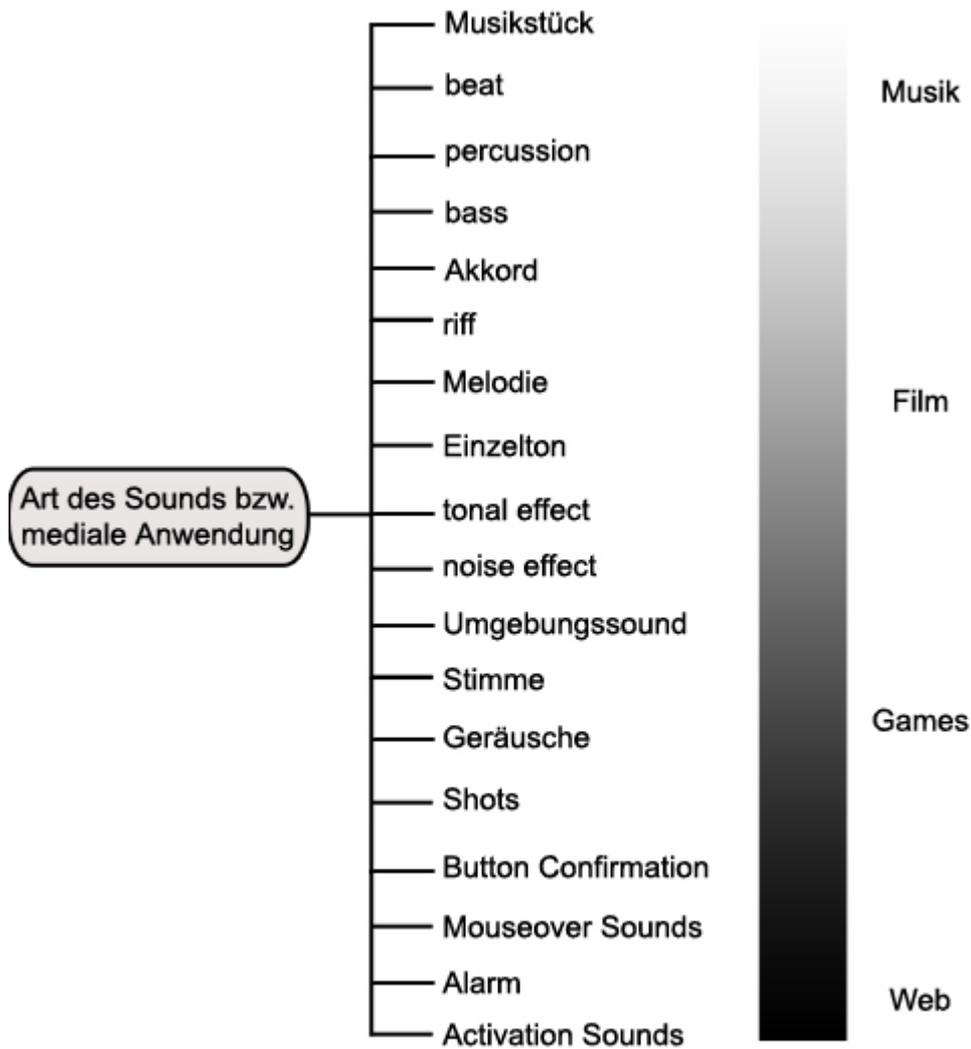


Abb 9: Kategorie „Art des Sounds bzw. mediale Anwendung“

Diese Aufgliederung nach der „Art des Sound bzw. mediale Anwendung“ ist zwar die am wenigsten Exakteste in dieser ganzen Struktur, jedoch sehr praktisch wenn man einen Sound ganz schnell klassifizieren will. Ausdrücke wie beat, riff oder Effekt kommen unter Musikern und Sounddesignern sehr oft

3.3.Klangbeschreibung,- und Kategorisierung – Ansatz zur Kategorisierung von Klängen

vor und dienen zur raschen Beschreibung und Positionierung des Sounds im gesamten Klangbild. Die Abtrennung der einzelnen Kategorien voneinander ist natürlich subjektiv und nicht immer eindeutig festzulegen, da hier von keinen eindeutigen technischen oder musikalischen Parametern die Rede ist.

Praktischerweise müssten deshalb Sounds, die wirklich im Grenzbereich zu mehreren Begriffen stehen, auch mehreren Soundarten zugeordnet sein.

Diese Unterteilung soll mit „Button Confirmation“ oder „Mouseover Sounds“ vor allem auch typischen Sounds von Computerspielen und solche im Webbereich Rechnung tragen.

Eine gröbere, fließende Unterteilung in die Bereiche Musik, Film, Games und Web wäre denkbar.

3. Time:

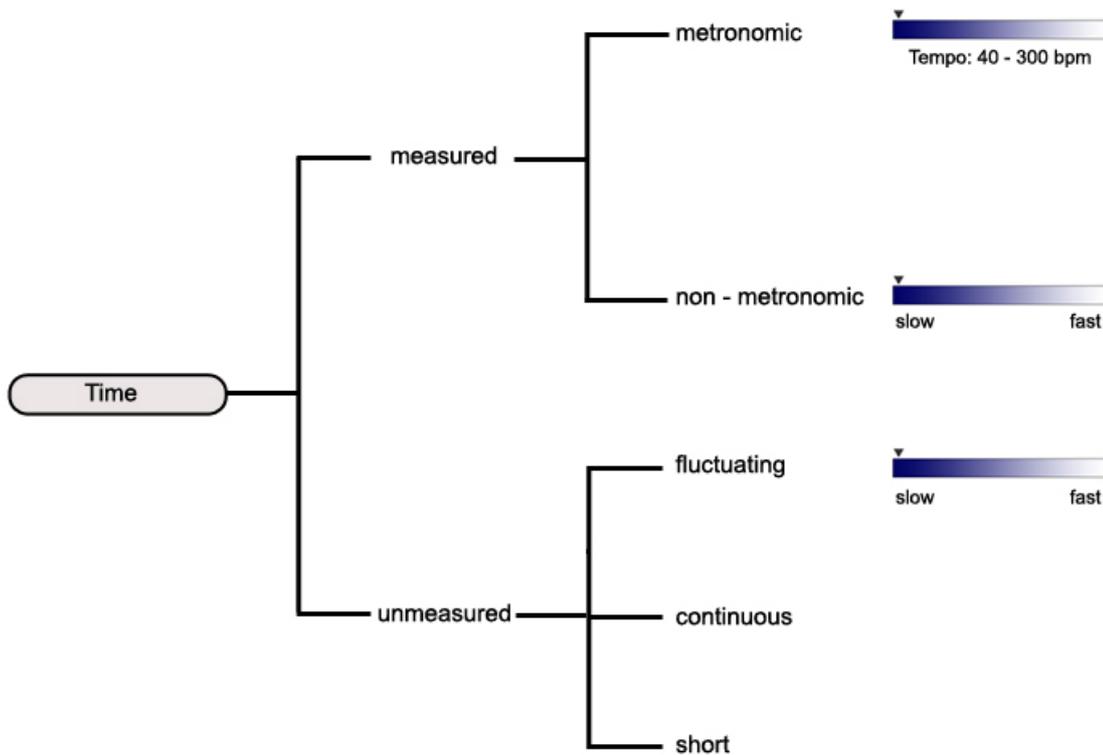


Abb 10: Kategorie „Time“¹¹

Der zeitliche Faktor bei einem Audiosignal ist sehr entscheidend für den Charakter des Sounds. Gängige Begriffe wie Loop, Flächensound, percussiver Sound u.ä. beschreiben im Prinzip schon diesen zeitlichen Faktor.

Unter „measured“ fallen also alle eindeutig rhythmischen und zeitlich messbaren Sounds. Weiters würden unter „metronomic“ sämtliche gleichmäßigen Beats mit kontinuierlichem Tempo und regelmäßige Rhythmen fallen, also im Prinzip auch fast jedes fertige Musikstück in unserem Kulturkreis.

„Measured, non-metronomic“ bedeutet, dass es zwar rhythmische Akzente gibt, jedoch keinen regelmäßigen Takt oder gleich bleibendes Tempo. Sprache wäre das Paradebeispiel für diese Art von Sounds.

¹¹vgl. Van Leeuwen, T. (1999): Speech, Music, Sound, S.51, Palgrave Macmillan

3.3.Klangbeschreibung,- und Kategorisierung – Ansatz zur Kategorisierung von Klängen

„Unmeasured – fluctuating“ beinhaltet alle Sounds ohne erkennbaren Rhythmus, aber pulsierende, wiederkehrende Akzente.

Mit „unmeasured – continuous“ sind länger anhaltende, flächige Sounds ohne hörbaren Rhythmus oder wiederkehrenden Impulsen gemeint.

Schließlich sind mit „unmeasured – short“ kurze, percussive Sounds mit kurzer Decay- und Releasetime gemeint, die sich nicht regelmäßig wiederholen.

Mit dieser zeitlichen Differenzierung kann schon eine Menge über das Audiosignal ausgesagt werden. Sie ist stark mit der ADSR – Hüllkurve, also dem zeitlichen Verlauf einer Schallquelle verbunden. Eine exakte Abgrenzung der einzelnen Kategorien zueinander ist weiters auch kein größeres Problem und kann im besten Fall auch völlig automatisiert über technische Analysen erfolgen.

Man könnte dann noch eine weitere Eigenschaft hinzufügen, die hier in diese zeitliche Kategorisierung passt, und zwar „loop“. Loops charakterisieren sich dadurch, dass sie so geschnitten sind, dass man sie problemlos hintereinander setzen kann, wobei der rhythmische Fluss erhalten bleibt.

4. Tone:

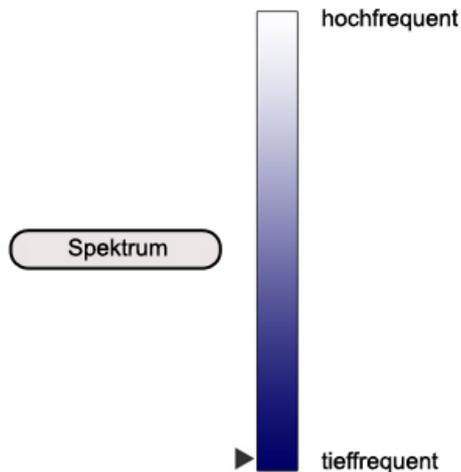


Abb 11: Kategorie „Tone“

Mittels audiotechnischer Analyse und dem MPEG7 Standard kann man auch bereits mit verschiedenen Algorithmen diverse Parameter zum Spektrum ermitteln. Diese Parameter wären Audio Spectrum Envelope, Audio Spectrum Centroid, Audio Spectrum Spread und Audio Spectrum Flatness. Mit Hilfe dieser Parameter kann man dominante Grundfrequenzen oder die Breite des Spektrums ermitteln.

Das Spektrum eines Audiosignals kann also problemlos automatisiert ermittelt werden. Dennoch kann es hilfreich sein, wenn der Sounddesigner auch die Möglichkeit hat, manuell Eigenschaften über das Spektrum zu vergeben. Denn oft können aus psychoakustischen Gründen Abweichungen zwischen dem Ergebnis der Spektralanalyse und dem tatsächlich empfundenen Spektrum bestehen.

5. Klangcharakter:

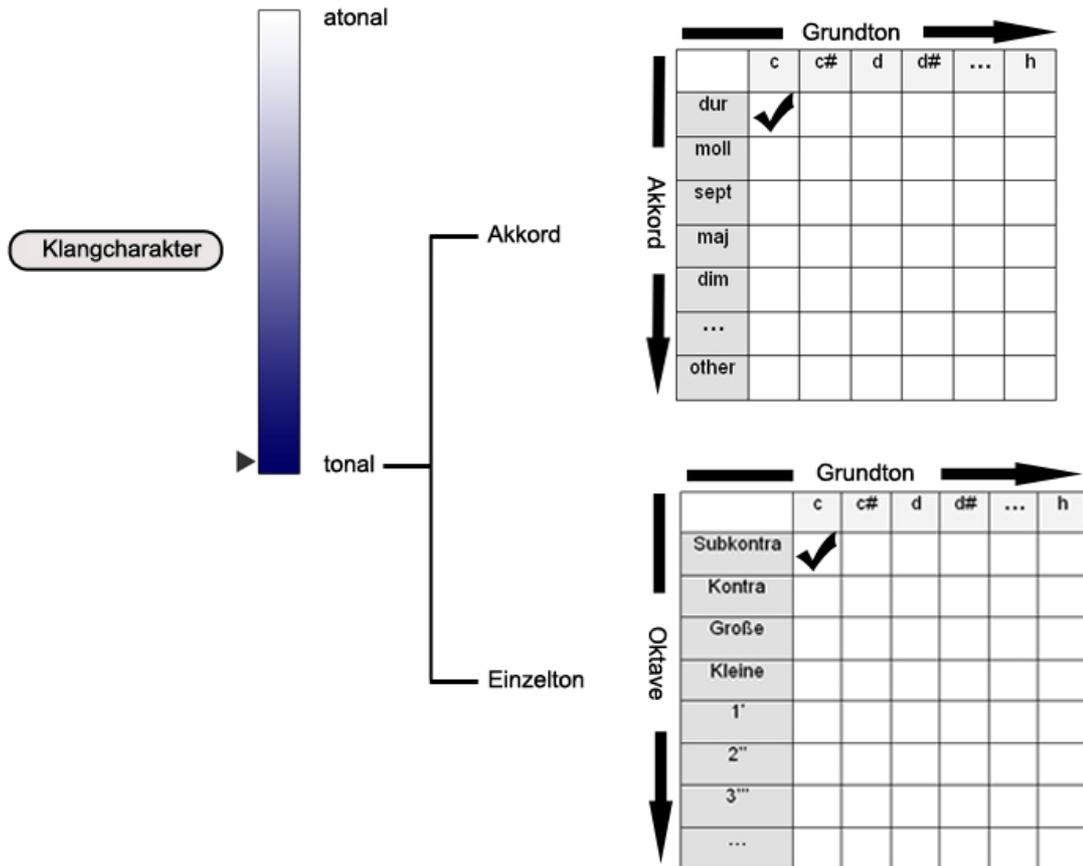


Abb 12: Kategorie „Klangcharakter“

Der Klangcharakter eines Sounds kann geräuschartig, also atonal oder tonal sein. Als Geräusch bezeichnet man ein nichtperiodisches Schallereignis mit unregelmäßigem Zeitverlauf, dem keine Tonhöhe zugeordnet werden kann. Weitere tonale Unterteilungen sind also hinfällig. Jedoch kann ein Rauschen sowohl eher hochfrequent oder tieffrequent sein, was bei der vorherigen Unterteilung „Spektrum“ ausgewählt werden kann.

Alle tonalen Klänge, also Klänge im eigentlichen Sinn, werden hier nach musikalischen Aspekten weiter unterteilt.

3.3.Klangbeschreibung,- und Kategorisierung – Ansatz zur Kategorisierung von Klängen

Als Akkorde bezeichnet man das gleichzeitige Erklingen von mehreren Tönen, die sich harmonisch in gewisser Weise deuten lassen. Meistens werden Akkorde als das Zusammenklingen von mindestens drei Tönen definiert. Das Zusammenklingen zweier Töne kann jedoch auch bereits akkordische Funktionen erfüllen. In der Musik können einzelne unabhängige Stimmen erklingen, die vom Hörer als Bestandteil einer harmonischen Struktur erlebt werden, also auch akkordisch sind.

In dieser Unterscheidung dient die Gegenüberstellung Einzelton / Akkord mehr dem Auswahlverfahren an sich. D.h. man kann sich auch einen Akkord zusammenstellen, indem man den Einzelton wählt und dann eine Mehrfachauswahl trifft, die wiederum einen Akkord ergibt. Die Auswahl „tonal à Einzelton à c, e und g“ würde als C – Dur ergeben. Dasselbe Ergebnis hätte man mit „tonal à Akkord à c – Dur“.

Auf diese Art wäre es sowohl möglich, Akkorde exakt mit Tönen auswählen zu können, als auch schnell einen gängigen Akkord in einer Datenbank von Sounds zu finden.

6. Intervall:

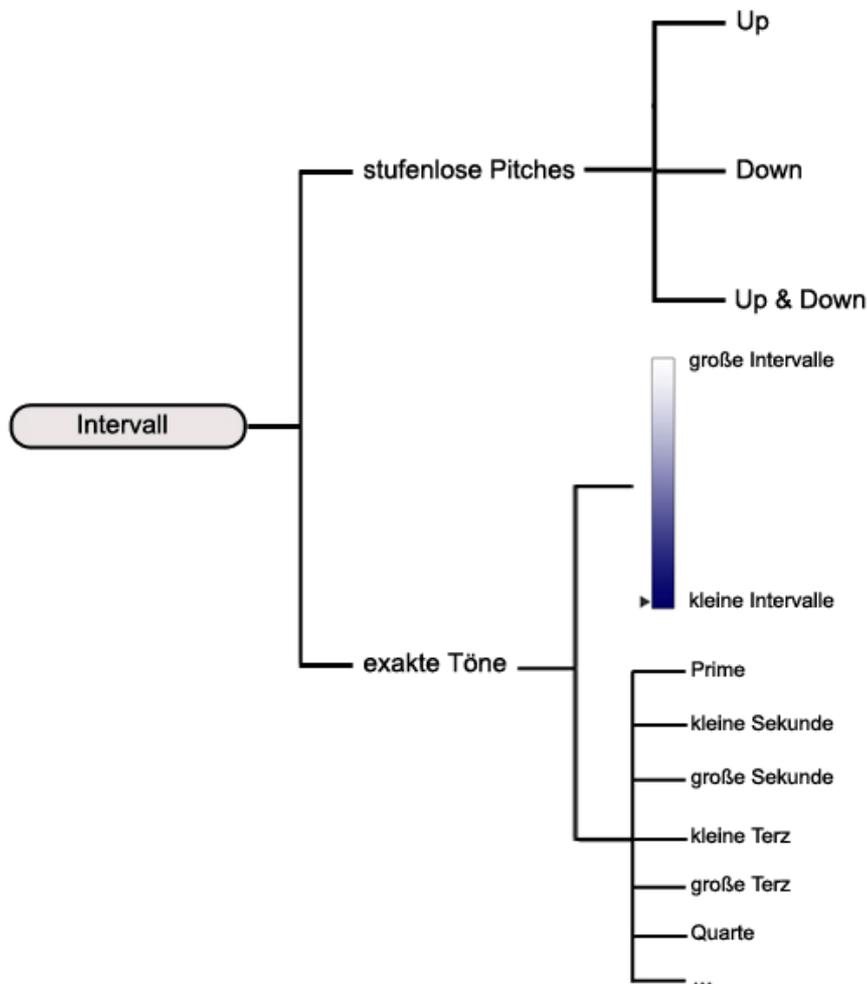


Abb 13: Kategorie „Intervall“

Das Intervall bezeichnet allgemein den Tonhöhenabstand zweier hintereinander- oder gleichzeitig klingender Töne. Nachdem das gleichzeitige Erklängen mehrerer Töne bereits im vorherigen Punkt abgehandelt wurde, bezieht sich dieser Punkt ausschließlich auf hintereinander – erklingende Töne, wenn man von exakten Tönen ausgeht.

3.3.Klangbeschreibung,- und Kategorisierung – Ansatz zur Kategorisierung von Klängen

Stufenlose Pitches kann mit dem musikalischen Begriff „glissando“, also die kontinuierliche und gleitende Veränderung der Tonhöhe innerhalb eines größeren Intervalls, gleichgesetzt werden.

Die typische Sirene zum Alarmieren der Feuerwehr wäre also ein stufenloser Pitch, welcher „Up & Down“ geht, und zwar genau 3 mal.

7. Dichte:

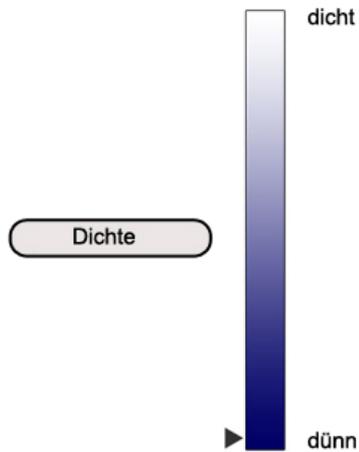


Abb 14: Kategorie „Dichte“

8. Überlagerung:

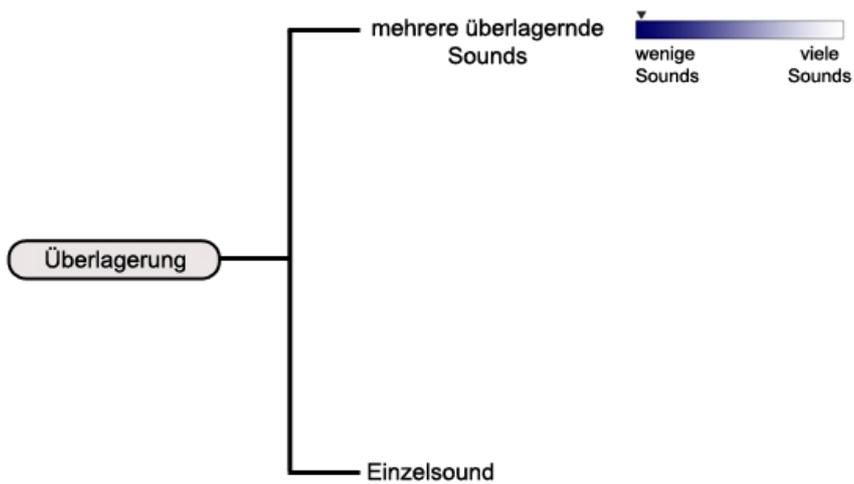


Abb 15: Kategorie „Überlagerung“

9. Lautstärkeänderung & Dynamik:

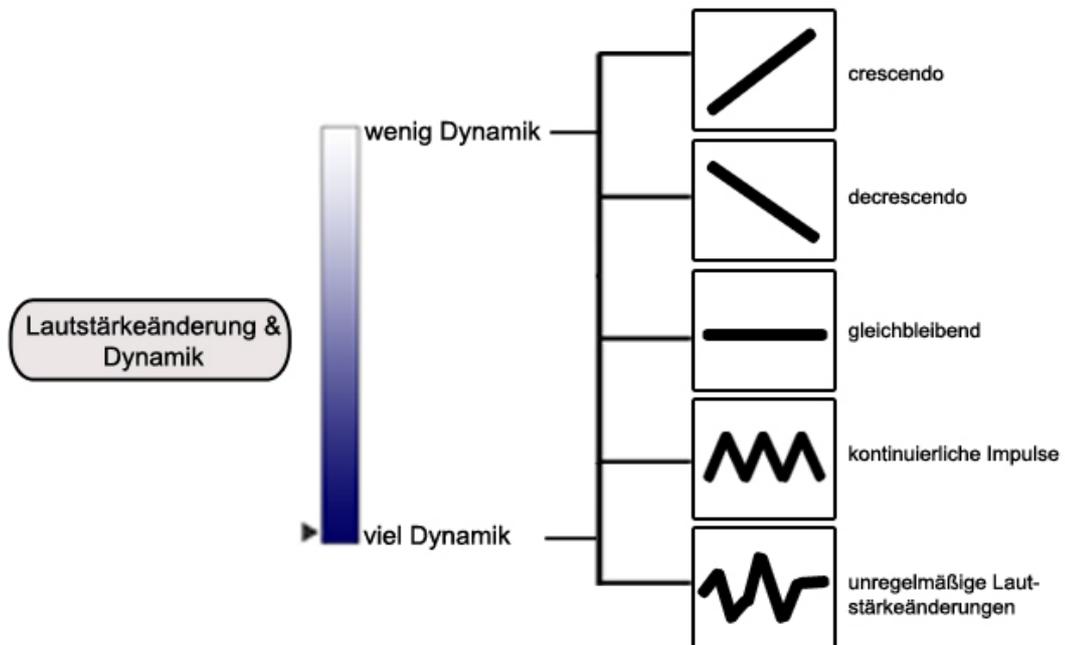


Abb 16: Kategorie „Lautstärkeänderung & Dynamik“

Die Dynamik bezeichnet den Unterschied zwischen der leisesten und der lautesten Stelle eines Klanges oder Musikstückes. Abgesehen von der Höhe der Dynamik können Lautstärkeschwankungen auch eine unterschiedliche Form aufweisen. Das oft zum Spannungsaufbau eingesetzte „Crescendo“, also eine stetige Steigerung der Lautstärke steht z.B. dem „Decrescendo“ gegenüber.

10. Perspektive:

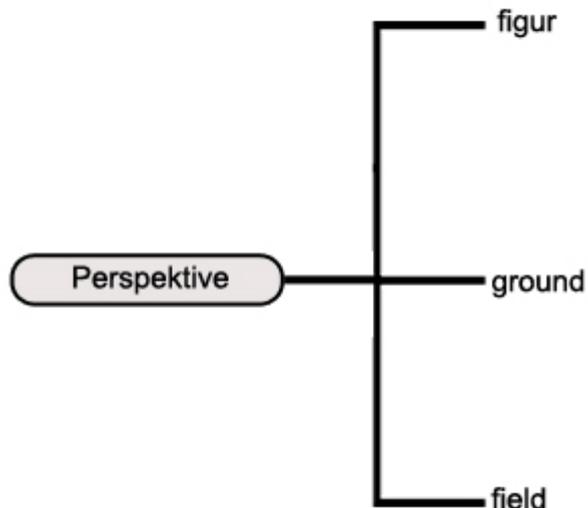


Abb 17: Kategorie „Perspektive“¹²

Diese Unterscheidung hängt in gewisser Hinsicht auch mit dem Punkt Nr. 2 „Art des Sounds und mediale Anwendung“ zusammen, da sich die Differenzierung auch auf die Anwendung in der Medienproduktion bezieht:

figur:

Diese Bezeichnung betrifft Sounds, die eher für den Vordergrund verwendet werden, also für etwas, das real in einer Szene oder allgemein in Situationen von Medienproduktionen vorkommt. D.h. Gewehrschüsse bei einem Computerspiel oder brüllende Aliens in einem Film würden mit figur – sounds vertont werden.

¹²vgl.Van Leeuwen, T. (1999): Speech, Music, Sound, S.16ff, Palgrave Macmillan

3.3.Klangbeschreibung,- und Kategorisierung – Ansatz zur Kategorisierung von Klängen

ground:

Das quasi Gegenteil dazu sind Hintergrundsounds, die eher zur emotionalen Untermahlung einer Szene dienen, jedoch mit den realen Gegebenheiten der Situation nicht unmittelbar zu tun haben. Dazu würden demnach auch Flückinger´s UKO – Sounds zählen.

field:

Hierzu zählen Sounds, die zwar im Hintergrund sind, jedoch auch tatsächlich in der Szene real vorkommen. Dazu zählt z.B. Straßenlärm, Waldgeräusche etc.

11. Weitere Parameter:

Es gibt noch einige weitere Parameter, mit denen man klangliche Ereignisse beschreiben kann. Für die Charakterisierung von kürzeren Samples, welche man eben typischerweise auf Sound – Libraries findet, sind jedoch einige Parameter schwer anwendbar.

Eine häufige Einteilung, die auf ganze Musikstücke zutrifft, ist klarerweise die Einteilung in Genre. Da findet man unzählige verschiedene Ansätze, die in ihrer Detailliertheit äußerst unterschiedlich sind. Manche unterscheiden vorerst mal in U-Musik und E-Musik. Andere teilen nach kulturellen oder geographischen Aspekten. Manchmal findet man auch beispielsweise unter „Rockmusik“ noch etliche weitere Unterscheidungen und Subgruppen, bei denen Abgrenzungen zu anderen Subgenres vermutlich schon sehr schwierig und subjektiv sind.

Für kürzere Soundsamples ist eine Einteilung in Genre oft schwierig, da man es erst beim Gesamtwerk und nicht beim einzelnen Baustein bestimmen kann.

Allerdings können einzelne Hinweise auf den kulturellen oder geografischen Hintergrund, sowie auch dem musikalischen Genre bei bestimmten Samples

3.3.Klangbeschreibung,- und Kategorisierung – Ansatz zur Kategorisierung von Klängen

schon sehr brauchbar sein. Etwa macht es Sinn, ein kurzes Sample einer typischen spanischen Flamenco – Gitarre oder ein typisches Heavy – Metal Riff auch als solche zu kennzeichnen.

Andere Parameter wären z.B. psychoakustische Parameter wie Lautheit, Schärfe, Tonhaltigkeit, Tonhöhe (Geräusche), Impulshaftigkeit etc. Sie sind auch Beschreibungsmittel für klangliche Ereignisse und zeigen Zusammenhänge der menschlichen Empfindung von Schall und den physikalischen Größen, die mit einem Schallereignis in Verbindung stehen. Psychoakustische Ergebnisse werden großteils mittels größerer Hörertests erzielt. Individuelle Eindrücke werden summiert, und statistische Durchschnittswerte herangezogen.

Aus diesem Grund sind solche Parameter für die schnelle Kategorisierung einer,- oder mehrerer Einzelpersonen eher ungeeignet.

Eine weitere Möglichkeit zur Beschreibung von Klängen bietet die umgangssprachliche Beschreibung. Mit Namen wie „woosh“ oder „klirr“ können bestimmte Sounds sehr treffend beschrieben werden.

Es kann auch vorkommen, dass eine Einzelperson, also in diesem Fall der Sounddesigner, aus persönlichen Gründen eine ganz bestimmte freie Assoziation zu einem Klang hat, welche andere Personen aber nicht nachvollziehen könnten.

Deshalb sollte es immer zusätzlich die Möglichkeit geben, weitere freie Metadaten zu Sounds hinzuzufügen.

Diese Aufgliederung kann also z.B. als Vorlage für eine Eingabemaske für Metadaten dienen. Das heißt, dass der Produzent oder erst der Anwender des betreffenden Sounds eine Vorlage bekommt, mit der er mit wenigen Klicks einen Sound sehr detailliert mit Metadaten versehen kann, um ihn in Folge leicht wieder zu finden und passend einsetzen zu können.

3.3.Klangbeschreibung,- und Kategorisierung – Ansatz zur Kategorisierung von Klängen

Andererseits kann diese Einteilung bereits als Vorlage für den Suchprozess bei einem fertigen Sound – Suchtool gelten. Man könnte sich vorstellen, dass dieses Suchtool nach einem ausschließenden Verfahren vorgeht. Das heißt, dass aus einem großen Pool an Sounds durch Anwählen bestimmter Eigenschaften die Anzahl an Sounds immer mehr eingegrenzt wird und letztlich aus einer kleinen Liste an Sounds durch Durchhören binnen kurzer Zeit der gewünschte Sound gefunden werden kann. Bei der Suche nach Sounds ist es generell vorteilhaft, wenn nach vielen Kriterien gesucht werden kann.

Diese Idee implementiert, dass diese Auswahlmöglichkeiten der Eigenschaften vorerst alle deaktiviert sind und einzeln aktiviert werden können. Bei jeder Auswahl würde die Liste der resultierenden Sounds entsprechend verkleinert werden.

Allerdings ist der richtige Suchvorgang wie gesagt auch stark abhängig vom gesuchten Sound:

Um bei dem Beispiel des Flugzeugs zu bleiben:

Würde man einfach nach dem Geräusch eines startenden Flugzeugs suchen, so wäre eine Suche nach der auslösenden Schallquelle das einzig Sinnvolle. Man definiert also das „Flugzeug“ als Schallquelle (am besten über ein Textfeld) und hat dann vielleicht noch die Auswahlmöglichkeit eines startenden, landenden oder gleitenden Flugzeugs und eventuell mehrere Varianten von Samples.

Richtig interessant wird es dann, wenn beispielsweise die Schallquelle nicht vorausbestimmt oder einfach unklar ist:

Ein Filmsounddesigner sucht z.B. für eine spannende Verfolgungsjagd in einem Science – Fiction - Thriller die passende Vertonung. Nun weiß er intuitiv, dass er einen schnellen rhythmischen Grundsound braucht, und als weitere Schicht

3.3.Klangbeschreibung,- und Kategorisierung – Ansatz zur Kategorisierung von Klängen

darüber „irgendetwas“ helleres mit Synthesizersound, passend zum restlichen Sounddesign im Film.

Nun würde er sich die einzelnen Bausteine folgendermaßen zusammenstellen können:

Für den rhythmischen Grundsound wählt er als „Art des Sounds“ „beat“ aus und für die „Time“ – Auswahl „measured“ und das Tempo um die 140 bpm. Aus der restlichen kleineren Liste an resultierenden Sounds wählt er dann einen passenden Sound aus oder modifiziert die Suche nach weiteren, z.B. tonalen Eigenschaften.

Für die nächste darüber liegende „Schicht“ wählt er zunächst eine synthetische Schallquelle aus. Weiters wählt er für einen flächigen Sound als „Time“ „unmeasured – continuous“ aus und in Folge für die „Schärfe/Helligkeit“ einen graduell eher „höhenlastigen“ Sound. Wieder kann er im Anschluss daran aus den übrig gebliebenen Sounds nach dem am meisten geeigneten Sound auswählen...

Diese Suchbeispiele sollen nur als Anschauung dafür gelten, dass die Suche nach vielen unterschiedlichen Kriterien möglich sein muss, damit sie effizient sein kann. Weiters sollte der Sinn eines ausschließenden Verfahrens bei der Soundsuche untermauert werden. Die genannte Einteilung in Eigenschaften könnte in ihrer Detailliertheit noch beliebig geändert werden, und soll wie gesagt als Vorlage dienen. Jedoch ist eine allzu detaillierte Einteilung in Eigenschaften möglicherweise eher kontraproduktiv und erlaubt keine intuitive Bedienung des Tools.

Wichtig wäre auch, dass die einzelnen Eigenschaften intelligent verknüpft sind, und zwar so, dass sie sich bei Widersprüchlichkeit gegenseitig ausschließen.

Der Begriff „Beat“ z.B. impliziert einen regelmäßigen (wenn vielleicht auch ungeraden) Takt. Als „Art des Sounds“ kann ein „Beat“ also von der „Time“ her nicht „unmeasured“ sein. Diese 2 Eigenschaften müssten sich also gegenseitig ausschließen etc.

4 Testreihen

4.1 Ziele der Usertests und Ansatz der Methoden

In den folgenden Seiten wird beschrieben, welche Ziele diese Usertests verfolgen und welche Methoden bez. Fragen aus diversen Gründen ausgewählt wurden. Es wurde zuerst ein Online – Pretest programmiert, um die angepeilte Testmethode und die Testfragen zu prüfen. Danach wurde der Haupttest im Audiostudio der Fachhochschule St.Pölten durchgeführt.

Grundsätzlich geht es darum, aus einigen Testsounds, und den Beurteilungen der Testuser so viel Information wie möglich zu gewinnen. Nachdem das Durchhören von Sounds eben eine gewisse Zeit beansprucht, muss darauf geachtet werden, dass man mit einigen präzisen Fragen über eine geringere Anzahl von Sounds auskommen muss.

Einerseits soll mittels freier Assoziation der Testpersonen auch über das Thema Metadaten und allgemein Soundbeschreibungen Aufschluss gewonnen werden.

Andererseits soll die emotionale Qualität von Klängen, und die Möglichkeit, diese in Worte zu fassen, untersucht werden.

Das Sound und Musik eng mit Emotionen verknüpft sind, ist unbestreitbar:

„Akustische Ereignisse, seien es Klänge, Geräusche oder gar Musik, eignen sich am besten, um Emotionen mitzuteilen! Diese Behauptung stimmt doch mit unserer Erfahrung überein: Man denke etwa an Tanzmusik, New-Age-Meditationsmusik, Trinklieder, Schlachtgesänge in den Fußballstadien, Freudenlieder etc. Oder wie leblos wären die meisten Filme, würden Musik und Geräusche nicht die zum Bild passende Stimmung vermitteln?

Es werden aber nicht unbedingt Pauken und Trompeten oder „butterweiche“ Geigen benötigt, die vielen Filmen ihre emotionale „Tiefe“ verleihen sollen. Oft wird – auch oder gerade in der alltäglichen zwischenmenschlichen

4.1. Testreihen – Ziele der Usertests und Ansatz der Methoden

Kommunikation – durch kleine akustische Gesten eine Vielzahl von Gefühlsinhalten transportiert. In einem Seufzer eines Mitmenschen kann Erschöpfung oder Depression, Mitgefühl oder überhebliche Ablehnung usw. mitschwingen.“¹³

Kernfragen beim Test sind: Werden gewisse Sounds von allen Testusern in ähnlicher oder gleicher Weise beschrieben? Welche Wörter werden allgemein bei Soundbeschreibungen am häufigsten verwendet? Werden gewissen Sounds stets dieselben emotionalen Qualitäten zugesprochen? Besitzen gewisse Sounds allein stehend (ohne Kombination mit visuellen oder anderen Reizen) überhaupt schon emotionalen Charakter? Kann man später sogar Zusammenhänge zu technischen Eigenschaften sehen (z.B: die großteils als „traurig“ eingestuften Sounds haben sehr ähnliche Spektren oder Melodieverläufe etc.)?

Weitere wichtige Themen in diesem Zusammenhang sind Untersuchungen über Bild – Ton Beziehungen. Da dieses Thema aber für sich ein sehr großer eigener Punkt ist und den Rahmen dieser Diplomarbeit gesprengt hätte, wurden Tests über dieses Thema von einem Kollegen durchgeführt.

¹³ Raffaseder, H. (2002): Audiodesign, Hanser Fachbuchverlag, S.15

4.2 Gefahren und Probleme der Testreihen

Da die Ressourcen bei der Testdurchführung begrenzt waren, sind gewisse Gefahren zu bedenken gewesen, um die Wissenschaftlichkeit der Tests zu gewähren. Da es sich um eine quantitative Testmethode handelt, müssten so viele Testsounds wie möglich von so vielen Testhörern wie möglich abgehandelt werden. Ein weiteres Problem ist die Subjektivität in diesem Kontext: Ein für den ersten Testhörer sehr „trauriger“ Sound kann im extremsten Fall für den anderen Testhörer „fröhlich“ klingen.

Ein anderes Problem bei derartigen Tests ist die Schwierigkeit, manche zusätzliche Faktoren auszuschalten. Z.B: kann bereits die Reihenfolge, in der der Testhörer die Sounds hört einen Einfluss auf das Ergebnis haben. Nach einem extrem unangenehmen Sound empfindet man einen nur etwas mehr entspannten Sound vielleicht gleich als übermäßig positiv. Zusätzlich kann schon die Tagesverfassung des Probanden das Testergebnis beeinflussen.

Man muss also auch die Möglichkeit in Betracht ziehen, dass die Tests nur wenige Aussagen über die semantischen Eigenschaften von Sounds treffen. Wenn dieselben Sounds dermaßen unterschiedlich bewertet und beschrieben werden, dass man nichts daraus herleiten kann, dann würde es zumindest eine andere Aussage machen: Und zwar dass der semantische bzw. emotionale Gehalt eines Klanges möglicherweise einfach zu subjektiv ist, um Verallgemeinerungen zu treffen.

Aus diesen Gründen ist dieser Test eher als tendenzielle Zuordnung der Sounds zu Beschreibungen, Emotionen und passenden Filmgenres zu sehen.

4.3 Auswahl der Testsounds und Methoden der Durchführung

Um die Testmethode und die Testfragen zu überprüfen wurde wie bereits erwähnt ein Pre – Test entwickelt. Dieser ist ein Online – Test, der vorerst nur an ausgesuchten Bekannten per Mail ausgesendet wurde. Dieser Pre – Test floss in die allgemeine Auswertung nicht mit ein, sondern hatte wie gesagt nur die Funktion, die Testmethode zu überprüfen und gegebenenfalls weiterzuentwickeln.

Nach den ersten Ergebnissen aus dem Pre - Test zeichnete sich bereits eine relativ hohe Einstimmigkeit der Ergebnisse bei gewissen Testsounds ab.

Für den Endtest im Audiostudio sollten nun insgesamt 20 Testsounds zum Einsatz kommen. Die Testuser waren aus praktischen Gründen zum Großteil Studenten, die während des Audiotechnik – Unterrichts einzeln ins Audiostudio gebeten wurden, um jeweils 5 Testsounds zu bewerten und zu beschreiben.

So erreichte man eine angemessene Zahl an Testergebnissen.

Technisch gesehen ist der Test als eine Form einer PHP – Webapplikation programmiert.

Alle Testergebnisse werden pro Sound in eine Datenbank geschrieben, wobei die jeweilige Sound ID zur Unterscheidung der Sounds in eine Spalte eingetragen wird. Jeder Testuser erhält auch beim Aufruf der Startseite eine User ID zugewiesen. So kann jedes Ergebnis im Nachhinein eindeutig sowohl dem Sound als auch dem User zugeordnet werden.

Die 20 Testsounds sind unterschiedlichster Art und bestehen aus diesen 6 Kategorien:

(Die 20 Testsounds befinden sich auf der beigelegten Audio – CD, um die Beschreibungen der Probanden besser nachvollziehen zu können).

4.4. Testreihen – Testfragen, Auswertungsmethode und Ergebnisse

1) Beats, Percussion:

Für diese Art der Testsounds ist vor allem die Frage interessant, ob rein perkussive Sounds bzw. Drumloops ohne Töne oder Akkorde allein stehend bereits eindeutigen emotionalen Charakter für die Testpersonen aufweisen.

2) Natur- Atmos, Umgebungsgeräusche:

Hierbei handelt sich um Umgebungsgeräusche, die in einer realen Umwelt zu hören sind wie Wind/Sturm, Funkgespräche aus der Ferne, Gebrüll eines wilden Tieres,...

3) Flächensounds:

Typische atmosphärische, flächige Sounds die als emotionale Untermahlung in diversen Filmen sehr geeignet sind.

4) Melodiefragmente:

Hier geht es um kurze Melodiefragmente mit unterschiedlicher Klangfarben.

5) Geräusche, Sound FX:

Diese Sounds betreffen elektronische Effekte und Geräusche, bei denen die Schallquelle entweder schwer erkennbar oder elektronischer Natur ist.

6) Akkorde:

Diese Kategorie beinhaltet einzelne Akkorde bzw. Akkordfolgen. Interessant hierbei ist z.B. die Überprüfung der über die Zeit „gelernten Tatsache“, dass Dur – Akkorde als fröhlich und Moll – Akkorde als traurig empfunden werden. Es wurde deshalb aus demselben Sample (einem Chor-Arrangement) einmal der Dur – Akkord und ein weiteres Mal der Mol – Akkord in der gleichen Tonart herausgeschnitten und zu den Testsounds hinzugefügt.

4.4 Testfragen, Auswertungsmethode und Ergebnisse

Insgesamt nahmen 95 Personen bei den Usertests teil. Der Großteil davon sind Studenten, aber auch einige andere Personen unterschiedlichsten Alters. 72,6% der Personen waren männlich, 27,4% waren weiblich. Der Altersschnitt der Probanden liegt bei 23,2 Jahren. 52,2% davon spielen oder spielten ein Musikinstrument. 61,1% haben bereits Erfahrungen mit Sound – Design oder Musikproduktion.

Nach einigen kleineren Änderungen nach dem Pre – Test beinhaltete der Endtest nun folgende 5 Fragen zu jedem Testsound:

1.) Beschreibe diesen Sound kurz:

Hierbei sind jegliche freie Assoziationen und Beschreibungen erwünscht.



(Pflichtfeld)

Ein entscheidendes Kriterium beim Test war es, vorerst eine unvoreingenommene, möglichst spontane Beschreibung des Sounds vom Probanden zu bekommen. Deshalb ist vorerst auch nur die erste Frage des Tests sichtbar, damit die Auswahleigenschaften bei den weiteren Fragen den Testuser nicht bereits beeinflussen.

Da dieses Feld ein Pflichtfeld ist, ist der Proband sozusagen „gezwungen“ seinen ersten Eindruck zu diesem Klang zu vermitteln. Hierbei ist interessant, ob hier eher emotionale Adjektive oder eher mehr oder weniger weit – hergeholte Assoziationen zum Einsatz kommen, und mit welchen Worten die Sounds generell beschrieben werden.

4.4. Testreihen – Testfragen, Auswertungsmethode und Ergebnisse

Diese Assoziationen der Testpersonen könnten später bei der Einteilung und Beschlagwortung der Sounds sehr hilfreich sein.

Ergebnisse:

Insgesamt wurden bei dieser Frage 1431 Wörter eingegeben. Davon wurden 839 verschiedene Begriffe eingegeben. 541 Begriffe davon wurden nur einmal genannt. 204 Begriffe kamen 2 mal vor, 37 Begriffe 3 mal und 17 Begriffe 4 mal.

Einige Begriffe wurden 5 mal und öfter genannt. Hier sind sie im Einzelnen aufgelistet:

16 mal: Spannung, entspannend

15 mal: ruhig

14 mal: bedrohlich

12 mal: Wind

10 mal: Chor, Weite

9 mal: Windows, metallisch, Kirche, fremd, dunkel, beruhigend

8 mal: Gefahr, angenehm

7 mal: Bedrohung

6 mal: Sirene, Monster, Meer, mächtig, indisch, entspannt, asiatisch

5 mal: Wald, verwirrend, sphärisch, orientalisch, monoton, leicht, langsam, Krieg, Geister, aufregend

Hier sieht man bereits, dass emotionale Begriffe wie „Spannung“, „bedrohlich“ oder „beruhigend“ oft genannt wurden.

Nachdem dieses Feld ja ein völlig freies Feld war, ist es interessant zu sehen, ob häufiger emotionale Begriffe, objektbezogene Begriffe oder ähnliches genannt wurden.

4.4. Testreihen – Testfragen, Auswertungsmethode und Ergebnisse

Aus diesem Grund wurden Begriffe nach der Art der Beschreibung zusammengefasst und die Häufigkeit der Beschreibungsarten prozentuell ausgewertet:

Beschreibung nach:	emotionale Begriffe	Objekt, Material	Ort der Klangentstehung	freie Assoziationen	andere Eigenschaften
Häufigkeit	39 %	20 %	17 %	5 %	19 %
Beispiele	Spannung ruhig Bedrohung aufregend geheimnisvoll	Chor metallisch Sirene Geister Streicher	Kirche indisch Urwald Weltraum Wald	Windows Horror Urlaub Distanz Computerspiel	tief dumpf langsam leicht flächig

Abb 18: Tabelle „Häufigkeiten der Beschreibungsarten“

4.4. Testreihen – Testfragen, Auswertungsmethode und Ergebnisse

Generell muss man natürlich sagen, dass diese Verteilung stark von den ausgewählten Sounds abhängt.

Betrachtet man die Auswertung der Sounds in Bezug auf Frage Nr.1 im Einzelnen, so kann man folgendes feststellen:

Bei Sounds, in denen die Schallquelle elektronischer Natur oder einfach schwer herauszuhören ist, wurde mehr auf emotionale Begriffe zurückgegriffen.

Sobald die Schallquelle erkennbar war, wurde meist entweder die Schallquelle selbst zuerst genannt (Harfe, Chor, Streicher, Sirene etc.) oder der Ort mit denen solche Sounds häufig verbunden werden (Urwald, Hafen, China etc.).

Erst danach wurden auch emotionale Begriffe verwendet.

Manche Sounds rufen bei den Probanden stets dieselben Assoziationen hervor.

Es kam auch vor, dass Objekte bzw. Materialien vermehrt genannt wurden, ohne dass sie wirklich im Soundmaterial zu hören sind:

Hochinteressant ist da z.B., dass bei Sound Nr.12 häufig Assoziationen zu Wasser im weitesten Sinne genannt wurden, ohne dass bei diesem Sound auch nur im Entferntesten Wasser zu hören ist.

Begriffe wie Meer, Wassertropfen, Regen, Wasser, Farbe blau etc. wurden durchgehend bei diesem Sound als Assoziation genannt.

Mit der Gegenüberstellung „weit, distanziert“ bzw. „eng, nahe“ konnten viele Probanden in Frage Nr. 3 bei den meisten Sounds eher wenig anfangen, was man aus den Befragungen und den Ergebnissen in Erfahrung bringen konnte. Bei Sound Nr. 20 ist es jedoch auffällig, dass häufige Begriffe wie Weite, große Kuppel, Raum, Freiheit auf etwas Großes und Weites hindeuten. Es ist so, als ob für die meisten Sounds diese Begriffe einfach nicht anwendbar sind, jedoch für diesen speziellen Sound plötzlich sehr passend seien.

4.4. Testreihen – Testfragen, Auswertungsmethode und Ergebnisse

Als speziellen „Test im Test“ wurde wie bereits erwähnt ein Chor – Sound jeweils einmal als Dur – Akkord, (Sound Nr.16) und einmal als Moll – Akkord (Sound Nr.13) in derselben Tonart unter den 20 Testsounds gemischt.

Es sollte ja das musikalische „Klischee“ überprüft werden, dass Dur - Klänge eher fröhlich und Moll – Klänge eher traurig empfunden werden.

Tatsächlich wurden mit dem Dur – Klang eher Begriffe wie frei, leicht, himmlisch, festlich etc. verbunden. Bei der nachfolgenden Frage Nr. 2 (emotionale Begriffe) dominieren Begriffe wie Freude, friedlich, positiv, feierlich etc. Bei den Grundemotionen (Frage Nr. 4) wurde 12 mal Lösung, Gelassenheit und 3 mal Freude, aber auch 4 mal Trauer genannt. 3 mal wurde keine Angabe zur Grundemotion gemacht.

Auffällig bei diesem Sound ist die häufige Nennung der Begriffe Himmel und himmlisch.

Beim Moll – Klang wurden zwar auch häufig dieselben Begriffe wie Kirche und Chor genannt, diese sind jedoch auf die gleiche Klangfarbe der beiden Sounds zurückzuführen. Die emotionalen Begriffe gehen weit auseinander (unheimlich, bestimmend, energievoll, Trauer etc.). Bei der Auswahl der Grundemotionen wurde beim Moll – Klang 7 mal Trauer genannt, 13 mal Anspannung/Furcht und 3 mal keine Angabe gemacht.

Man kann also zusammenfassend sagen, dass der Dur – Klang tatsächlich um vieles „positiver“ beschrieben wurde.

Die Begriffe „Dunkelheit“ und „dunkel“ wurden insgesamt 10 mal genannt. 5 mal davon wurde alleine der Sound Nr. 17 mit diesen Begriffen von unterschiedlichen Probanden beschrieben. Alle anderen Nennungen dieser Begriffe teilen sich auf andere Sounds auf, aber immer nur einmal pro Sound.

Aus diesen Beispielen kann man die Tendenz herauslesen, dass gewisse Begriffe, obwohl sie generell seltener genannt werden, dann vermehrt bei speziellen Sounds auftauchen.

2.) Welche Stimmung, welche Emotion würden Sie mit diesen Sound verbinden?

(optional)

Bei dieser Frage nach der Emotion wurde ein optionales Textfeld gewählt, da erwartet wurde, dass bei manchen Sounds eine emotionale Verbindung so dermaßen schwierig ist (zB: bei perkussiven Sounds), dass man hierbei auch die Möglichkeit haben müsste, das Feld frei zu lassen. Zudem wurden von vielen Testpersonen wie gesagt bereits im Feld der Frage 1 rein emotionale Eigenschaften genannt.

Ergebnisse:

Hier wurden insgesamt 736 Begriffe, davon 379 unterschiedliche Begriffe eingegeben. 279 Begriffe davon wurden nur einmal genannt. 92 Begriffe kamen 2 mal vor, 66 Begriffe 3 mal. Hier sind alle Begriffe, die mindestens 4 mal genannt wurden, aufgelistet:

43 mal: Angst

31 mal: Spannung

25 mal: Furcht

20 mal: Entspannung

16 mal: Ruhe

14 mal: Gelassenheit

11 mal: Anspannung

9 mal: Freude

4.4. Testreihen – Testfragen, Auswertungsmethode und Ergebnisse

8 mal: Trauer, Gefahr

7 mal: Ungewissheit, ruhig

6 mal: Unsicherheit, Unruhe, Stimmung, spannend, Bedrohung, Aufregung

5 mal: Unbehagen, neutral, mystisch, entspannt, entspannend, Beklemmung

4 mal: Stress, Neugierde, Neugier, Hektik, geheimnisvoll, düster

Fasst man nun Begriffe wie Spannung und spannend, sowie auch ähnlich – semantische Begriffe wie z.B.: Furcht und Angst gewissermaßen zu Überbegriffen und Wortfamilien zusammen, so bekommt man folgende Verteilung:

(Hier sind bereits die Begriffe der ersten und zweiten Frage zusammengefasst. Diese Aufzählung betrifft also alle in Textfeldern eingegebenen Begriffe).

146 mal (6,7 %): Begriffe der Wortfamilie „Spannung“

Spannung, Entspannung, spannend, entspannend, entspannt, Anspannung, angespannt, Entspanntheit, spannungserzeugend, spannungsaufbauend, Spannungsaufbau, entspannen, Entspannungsmusik, Angespanntheit

86 mal (3,9 %): Begriffe der Wortfamilien „Angst“ und „Furcht“

Angst, Furcht, ängstlich, beängstigend, angsterregend, angsteinflößend, furchteinflößend, Ehrfurcht, ehrfürchtig

63 mal (2,9 %): Begriffe der Wortfamilie „Ruhe“

Ruhe, ruhig, beruhigend, Unruhe, beunruhigend, unruhig, Ruhephase

31 mal (1,4 %): Begriffe der Wortfamilie „Bedrohung“

Bedrohung, bedrohlich, bedrohend

24 mal (1,1 %): Begriffe der Wortfamilie „Gelassenheit“

4.4. Testreihen – Testfragen, Auswertungsmethode und Ergebnisse

Gelassenheit, gelassen, Losgelassenheit, losgelassen, Ausgelassenheit, ausgelassen

21 mal (< 1 %): Begriffe der Wortfamilie „Gefahr“

Gefahr, gefährlich

17 mal (< 1 %): Begriffe der Wortfamilie „Aufregung“

Aufregung, aufregend

16 mal (< 1 %): Begriffe der Wortfamilie „Mystik“

Mystik, mysteriös, mystisch

15 mal (< 1 %): Begriffe der Wortfamilie „Weite“

Weite, weit

14 mal (< 1 %): Begriffe der Wortfamilie „Fremdheit“

fremd, Fremdheit, verfremdend

11 mal (< 1 %): Begriffe der Wortfamilie „Freude“

Freude, freudig

11 mal (< 1 %): Begriffe der Wortfamilie „Hektik“

Hektik, hektisch

11 mal (< 1 %): Begriffe der Wortfamilie „Sicherheit“

Unsicherheit, unsicher, verunsichert, sicher

11 mal (< 1 %): Begriffe der Wortfamilie „Verwirrung“

Verwirrung, Verwirrtheit, verwirrt, verwirrend

4.4. Testreihen – Testfragen, Auswertungsmethode und Ergebnisse

10 mal (< 1 %): Begriffe der Wortfamilie „Dunkelheit“

Dunkelheit, dunkel

10 mal (< 1 %): Begriffe der Wortfamilie „Trauer“

Trauer, traurig

9 mal (< 1 %): Begriffe der Wortfamilie „Leichtigkeit“

Leichtigkeit, leicht, Erleichterung

Weitere Wortfamilien, die weniger als 9 mal vorkommen:

Helligkeit, Neugierde, Stimmung, Macht, Düsterteit, Beklemmung, Lösung...

Begriffe wie „indisch“, „orientalisch“, „Weltraum“ etc. wurden hier nicht berücksichtigt, da sie zu spezifisch auf die Sound – Auswahl zurückzuführen sind und über generelle Beschreibungsmöglichkeiten von Sounds wenig aussagen.

Der genaue Prozentsatz der einzelnen Wortfamilien ist im Grunde nicht entscheidend. Es geht viel mehr darum, zentrale Begriffe zur Soundbeschreibung herauszufiltern. So kommt man zum Beispiel an Begriffe wie „Spannung“ etc. kaum vorbei.

Begriffe, die in die Wortfamilie „Spannung“ fallen, wurden ausnahmslos bei jedem der 20 Sounds mindestens einmal, im Schnitt jedoch beachtliche 7,3 mal pro Sound genannt.

Ganz wichtig ist z.B. auch der Begriff „Spannungsaufbau“. Sounds, die zunehmend mehr Spannung erzeugen, werden in Medienproduktionen sehr häufig gebraucht, vor allem in der Filmvertonung, aber auch in der Musik. (Man stelle sich zum Beispiel eine Mordszene vor, in der man den Mörder mit einem Messer in der Hand dem Opfer langsam näher kommen sieht). Diese „Spannungsaufbauer“ (Bsp: Sound Nr.1) gehen oft mit einer stetigen Steigerung

4.4. Testreihen – Testfragen, Auswertungsmethode und Ergebnisse

der Amplitude (Crescendo) einher und lassen sich daher vermutlich auch per technische Analysen bis zu einem gewissen Grad isolieren.

3.) Wie empfindest Du diesen Sound?

(tendenziell jeweils von 1-5, 3 = neutrale Mitte)

	1-5						nicht sicher bzw. keines von beiden
sanft, entspannend	<input type="radio"/>	aufdringlich, unangenehm	<input checked="" type="radio"/>				
hell	<input type="radio"/>	dunkel	<input checked="" type="radio"/>				
fremd	<input type="radio"/>	vertraut	<input checked="" type="radio"/>				
groß, mächtig, dicht	<input type="radio"/>	klein, dünn	<input checked="" type="radio"/>				
hektisch, schnell	<input type="radio"/>	ruhig, langsam	<input checked="" type="radio"/>				
weit, distanziert	<input type="radio"/>	eng, nahe	<input checked="" type="radio"/>				
antreibend	<input type="radio"/>	schleppend	<input checked="" type="radio"/>				
einfach	<input type="radio"/>	komplex	<input checked="" type="radio"/>				

Die dritte Testfrage beinhaltet jeweils Gegenüberstellungen von zwei konträren Eigenschaften.

Sie bezieht sich auf Beschreibungen von Klängen, die zwar relativ häufig verwendet werden, aber eigentlich nur sehr schwer definiert und erklärt werden können.

Was macht wirklich einen hellen Sound aus? Wann ist ein Sound groß und mächtig und wann wird er als hektisch empfunden?

Ähnliche Fragestellungen sollen mit dieser Testfrage überprüft werden. Würden beispielsweise gewisse Sounds von nahezu allen Probanden als sanft und entspannend empfunden werden, so könnte man anschließend versuchen, Zusammenhänge zwischen den Sounds zu finden und optimalerweise eine Art gemeinsames Muster für z.B. angenehme, entspannende Sounds finden.

4.4. Testreihen – Testfragen, Auswertungsmethode und Ergebnisse

Ergebnisse:

Aus den Ergebnissen und den Befragungen der Probanden über den Test an sich kann man sagen, dass es sehr auf die vorgegebenen Wörter in den Gegenüberstellungen ankam: Mit manchen Bezeichnungen taten sich die Probanden schwer, bei anderen Bezeichnungen war es den Testpersonen klar, was gemeint war.

Berechnet man nun von allem angekreuzten Kästchen Sound für Sound den Durchschnitt, so erhält man folgende Ergebnisse:

(1 = Maximum der ersten Eigenschaft; 5 = Maximum der zweiten Eigenschaft; 3 = Mitte, also neutral)

4.4. Testreihen – Testfragen, Auswertungsmethode und Ergebnisse

Sound	(1) sanft – unangenehm (5) ∅	(1) hell – dunkel (5) ∅	(1) fremd – vertraut (5) ∅	(1) groß– klein (5) ∅	(1) hektisch – ruhig (5) ∅	(1) weit – nahe (5) ∅	(1) antreibend– schleppend (5) ∅	(1) einfach – komplex (5) ∅
1	3,8	3,1	2,1	2,4	2,7	3,1	2,6	3,4
2	1,2	2,3	4,3	2,9	4,7	3,4	3,4	2,8
3	3,6	3,1	2,2	2,0	2,2	3,0	2,4	3,5
4	4,0	1,3	3,6	4,2	2,4	3,6	2,4	1,8
5	3,9	4,0	1,5	2,2	3,4	2,2	3,1	3,8
6	2,0	2,4	2,6	3,1	4,2	2,2	3,4	2,4
7	1,5	1,9	2,6	3,6	4,4	3,2	3,9	2,8
8	3,7	4,3	2,2	1,4	3,9	3,0	3,0	3,3
9	1,4	1,9	4,0	3,3	4,0	3,6	2,8	2,7
10	4,6	4,3	1,7	1,6	3,0	3,8	2,7	3,8
11	2,5	2,8	3,5	2,8	3,2	3,1	2,8	2,5
12	1,1	1,7	3,7	3,4	4,6	2,9	3,6	2,6
13	3,8	2,7	2,4	1,8	3,6	2,4	3,2	2,8
14	3,9	2,6	2,0	3,6	1,6	3,4	1,7	2,6
15	4,6	2,6	4,0	2,0	3,0	1,7	2,0	1,5
16	2,2	1,8	3,3	2,1	4,1	2,7	3,4	2,7
17	3,3	4,1	2,3	1,5	4,3	2,5	3,5	3,2
18	3,8	3,5	2,1	2,5	4,0	2,0	3,9	2,9
19	3,5	3,1	2,4	3,8	2,7	1,9	3,1	3,2
20	2,0	2,2	2,8	2,3	3,8	2,6	3,0	3,2

Abb 19: Tabelle „Durchschnittswerte der Gegenüberstellungen (Frage 3)“

sanft, entspannend – aufdringlich, unangenehm:

Die erste Gegenüberstellung brachte die eindeutigsten Ergebnisse. Dh. es wurden oft die beiden Extreme ausgewählt. Im Schnitt wich man um 1,05 „Kästchen“ von der Mitte ab.

Versucht man bei dieser Gegenüberstellung Zusammenhänge zu Anderen zu finden, so gibt es die am ehesten bei „fremd vs. vertraut“.

4.4. Testreihen – Testfragen, Auswertungsmethode und Ergebnisse

Klänge, die als überdurchschnittlich „sanft, entspannend“ gewertet wurden, bekamen meist auch die Eigenschaft „vertraut“. Umgekehrt wurden „aufdringliche“ Klänge meist auch als eher „fremd“ eingestuft.

Nach diesem Ergebnis wäre Sound Nr.12 der „entspannenste“ Klang und Sound Nr.10 und Nr.15 wären ex aequo die „unangenehmsten“ Klänge.

Sound Nr.17 scheint mit einem Schnitt von 3,3 der neutralste Sound zu sein.

hell – dunkel:

Auch bei dieser Gegenüberstellung gibt es für viele Sounds sehr eindeutige Ergebnisse. Die Probanden konnten mit diesen beiden Begriffen, bezogen auf Sounds, relativ viel anfangen.

Der psychoakustische Begriff „Helligkeit“ ist mit dem Begriff „Brillanz“ gleichzusetzen und bedeutet eine Betonung im oberen Frequenzbereich zwischen 2000 Hz und 6000 Hz. Anders ausgedrückt: ein Klang wirkt umso heller, je obertonreicher er ist.

Tatsächlich kann man sagen, dass Sounds, die im Bassbereich am schwächsten und eher mittel,- und hochtönig sind, am hellsten empfunden wurden (Bsp: Sounds Nr. 4, 7 und 16). Andererseits wurden Sounds, die im Bassbereich hörbar einen großen Anteil haben oftmals als dunkel eingestuft.

Vielen der Probanden ist die genaue Bedeutung des psychoakustischen Begriffs „Helligkeit“ sicherlich nicht bewusst. Aber die meisten scheinen intuitiv eine Ahnung von der Bedeutung dieses Begriffs in diesem Kontext zu haben.

„Dunkelheit“ und „Helligkeit“ scheinen also durchaus berechnete Begriffe für die intuitive Beschreibung von Klängen zu sein.

fremd – vertraut:

Klänge, die als „vertraut“ eingestuft wurden, waren Sounds wie ein Streichorchester, eine Sirene und eine Harfe etc. Also Sounds, bei denen die

4.4. Testreihen – Testfragen, Auswertungsmethode und Ergebnisse

Schallquelle eindeutig ist und die man aus dem alltäglichen Leben oder einfach aus gängiger Musik kennt.

Synthetische Sounds oder Science Fiction – artige Sounds bekamen eher die Eigenschaft „fremd“ zugesprochen.

groß, mächtig – klein, dünn:

Hier gibt es eindeutig einen Zusammenhang zwischen den Begriffen „groß, mächtig“ und dem Begriff „dunkel“. „Dunkle“ Sounds sind meist auch als „groß“ eingestuft worden.

Hochtönige, helle Sounds bekamen oftmals auch eher die Eigenschaften „klein, dünn“ zugesprochen.

Sounds, die hier eher im mittleren Frequenzbereich liegen, liegen oftmals auch bei der Gegenüberstellung „hell vs. dunkel“ eher in der Mitte.

hektisch, schnell – ruhig, langsam:

Die meisten der 20 Sounds liegen im Schnitt über der Mitte, sind also als ruhig eingestuft worden. Die wenigen Sounds, die eindeutig als hektisch eingestuft wurden (Sounds Nr. 14, 3, 4), haben sonst keine auffälligen Gemeinsamkeiten.

Es gibt jedoch die Tendenz, dass Klänge mit eindeutigen Rhythmus oder zumindest wiederkehrenden Impulsen als eher hektisch empfunden wurden.

Einige sehr „ruhige“ Klänge wurden auch als „sanft, angenehm“ empfunden.

Man kann im Gegensatz aber nicht sagen, dass „hektische“ Klänge zum Großteil als „unangenehm“ empfunden wurden.

weit, distanziert – eng, nahe:

Mit diesen Begriffen hatten die Probanden nach Befragungen teilweise große Probleme, sie mit Klängen in Verbindung zu bringen. Der Durchschnitt liegt hier auch Großteils ziemlich in der Mitte. Es gibt nur ganz wenige Sounds, die hier deutlich von der Mitte abweichen.

4.4. Testreihen – Testfragen, Auswertungsmethode und Ergebnisse

Als eindeutig „weit, distanziert“ wurden beispielsweise nur die Sirene (Sounds Nr.15) und das Geräusch einer Stimme aus einem Lautsprecher bzw. Funkgerätes (Sound Nr. 19) empfunden.

Als sehr „eng“ bzw. „nahe“ wurde kein Sound eindeutig bewertet. 3,8 ist die höchste Bewertung, welche also nur um 0,8 von der Mitte abweicht. Bei anderen Eigenschaften wurden viel eindeutigere Ergebnisse erzielt.

antreibend – schleppend:

Diese beiden Extreme weichen nur gering von der Gegenüberstellung „hektisch, schnell“ und „ruhig, langsam“ ab. Dh. dass „antreibend“ tendenziell mit „hektisch, schnell“ gleichgesetzt wurde.

Jedoch weicht dieser Vergleich noch viel weniger von der Mitte ab, als es bei den Begriffen „hektisch, schnell“ und „ruhig, langsam“ der Fall war.

Lediglich der rhythmische Sound Nr. 14 wurde als sehr antreibend empfunden.

einfach – komplex:

Die „einfachsten“ Sounds sind die Sirene (Sound Nr.15) und der „Windows – artige“ Sound Nr.4.

Generell ist auch hier die Abweichung von der Mitte weit geringer als bei anderen Gegenüberstellungen. Es gibt z.B. kaum Sounds, die sehr eindeutig als komplex empfunden wurden.

Es sind auch keine erkennbaren Zusammenhänge zu den anderen Gegenüberstellungen zu finden.

Nach sporadischen Befragungen der Probanden zu den Begriffen wurde dieses Begriffspaar auch als sehr schwierig genannt.

4.) Welche Grundemotion passt zu diesem Sound am ehesten?

- Trauer
- Lust, Liebe
- Hass, Abneigung
- Anspannung, Furcht
- Lösung, Gelassenheit
- Freude
- bin mir nicht sicher bzw.
keine der angegebenen
Emotionen treffen zu

Hierbei geht es um die Zuordnung des Klangs zu einer Emotion. Es wurden diverse Ansätze aus der Psychologie herangezogen, die die Einteilung in Grundemotionen wagten. In der kategorialen Klassifikation von Emotionen etwa werden alle Emotionen auf Basisemotionen zurückgeführt, aus denen sich alle anderen komplexeren Emotionen zusammensetzen. Mit Lothar Schmidt-Atzert und der Weiterentwicklung von Plutchik extrahierte man schließlich acht Basisemotionen, die jeweils noch in ihrer Intensität verschieden stark ausgeprägt sein konnten.

Diese, auf Untersuchungen mit Testpersonen basierte Theorie, steht jedoch auch anderen Theorien in der Psychologie gegenüber, wie z.B. der Emotionstheorie von Izard, der die Emotionen vorerst grob in angenehme und unangenehme Emotionen einteilte, oder der dimensionalen Klassifikation von Wundt.¹⁴

Aus diesen teilweise konträren Theorien wurden für diesen Usertest schließlich 6 Emotionsgruppen herauskristallisiert, die den Sounds zugeordnet werden sollten.

Auch hier hat man als Testperson die Möglichkeit, die Frage offen zu lassen, wenn man sich bezüglich eines Sounds unsicher ist. Es wurde auch erwartet, dass es bei gewissen Sounds einfach sehr schwer ist, sich für eine derartige Emotion zu entscheiden.

¹⁴ vgl. Trimmerl, M. (2002): Emotionstheorie von Robert Plutchik
http://homepage.univie.ac.at/michael.trimmel/motivation_ss2002/DieEmotionstheorievonRobertPlutchik.pdf

4.4. Testreihen – Testfragen, Auswertungsmethode und Ergebnisse

Ergebnisse:

(Die Zahlen bedeuten die Anzahl der Nennungen für die jeweilige Emotion).

Sound	Trauer	Lust, Liebe	Hass, Abneigung	Anspannung, Furcht	Lösung, Gelassenh.	Freude	keine Angabe
1	-	-	2	21	-	-	-
2	4	3	-	-	15	-	1
3	-	-	4	8	2	-	9
4	-	-	3	1	6	6	7
5	-	-	2	18	1	-	2
6	7	2	-	7	6	-	2
7	-	4	-	-	16	1	3
8	1	-	1	18	-	-	3
9	-	5	-	-	12	4	2
10	-	-	1	21	-	-	-
11	-	1	1	2	13	2	4
12	1	7	-	-	14	-	1
13	7	-	-	13	-	-	3
14	-	-	2	12	1	1	7
15	-	-	3	20	-	-	-
16	4	-	-	-	12	3	3
17	-	-	1	18	1	-	3
18	-	-	4	17	-	-	2
19	-	-	1	6	2	-	14
20	1	-	-	2	14	-	6

Abb 20: Tabelle „Anzahl der Nennungen für Emotionsgruppen (Frage 4)“

Summiert man die Nennungen, so erhält man folgende Ergebnisse:

Trauer: 25 mal

4.4. Testreihen – Testfragen, Auswertungsmethode und Ergebnisse

Lust, Liebe: 22 mal

Hass, Abneigung: 25 mal

Anspannung, Furcht: 184 mal

Lösung, Gelassenheit: 115 mal

Freude: 17 mal

keine Angabe: 72 mal

Grundsätzlich sieht man, dass eindeutig auf die beiden Emotionsgruppen Anspannung, Furcht bzw. Lösung, Gelassenheit am häufigsten zugegriffen wurde. Dies hängt klarerweise wie auch bei allen anderen Testfragen auch speziell von den 20 ausgewählten Sounds ab. Bei diesem Test gab es jedenfalls eindeutig die Tendenz, dass diese beiden Emotionen eher auf die Beschreibung von Sounds angewendet wurden. Man muss auch bedenken, dass diese 20 Sounds so derart unterschiedlich sind, und trotzdem wurden diese beiden Begriffspaare mit großem Abstand am häufigsten angekreuzt.

Bei den meisten Sounds war das Ergebnis relativ eindeutig, d.h. bei fast allen Sounds hat eine Emotionsgruppe klar die meisten Nennungen.

Es gibt jedoch 4 Sounds, bei denen sich die Testpersonen deutlicher uneinig waren. Das sind Sounds Nr. 4, Nr. 6, Nr.11 und Nr.14.

Am meisten verschiedene Nennungen gab es bei Sound Nr. 11. Dieser „Hip – Hop Beat“ – artige Klang rief gleich 5 verschiedene Nennungen hervor. Auch konträre Nennungen wie sowohl Freude als auch Furcht wurden genannt. Interessanterweise gibt es auch bei Sound Nr. 14, ein weiterer perkussiver Sound, fünf verschiedenen Nennungen. Perkussive Sounds ohne hörbaren harmonischen Elementen scheinen also eher Uneinigkeit bei den Testpersonen ausgelöst zu haben.

Am häufigsten „verweigert“ wurde die Angabe bei Sound Nr.19 (Funkgespräch aus der Ferne). Hier wurde gleich 14 mal keine Angabe zur Emotion gemacht.

4.4. Testreihen – Testfragen, Auswertungsmethode und Ergebnisse

Die eindeutigsten Ergebnisse wurden bei den Sounds Nr. 1 und Nr. 10 erzielt. Beide wurden fast geschlossen mit „Anspannung, Furcht“ in Verbindung gebracht.

Die Emotionsgruppe „Anspannung, Furcht“ hatte generell auch mit Abstand die eindeutigsten Nennungen. D.h: Wurden diese Emotionen bei einem Klang am häufigsten genannt, so gibt es meist auch nur ganz wenige andere Nennungen beim selben Klang. Anders ausgedrückt könnte man sagen, dass „Anspannung, Furcht“ am leichtesten herausgehört wurden.

5.) In welcher Filmkategorie würde dieser Sound typischerweise eher vorkommen?

		1-5					
Science Fiction	unpassend	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr passend
Fantasy	unpassend	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr passend
Liebesfilm, Erotik	unpassend	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr passend
Komödie	unpassend	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr passend
Actionfilm	unpassend	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr passend
Krimi	unpassend	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr passend
Horrorfilm, Thriller	unpassend	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr passend
Trickfilm, Kinderfilm	unpassend	<input type="radio"/>	<input type="radio"/>	<input checked="" type="radio"/>	<input type="radio"/>	<input type="radio"/>	sehr passend

Die Zuordnung zu einem Filmgenre dient dem groben Vorschlag, welcher medialen Anwendung der Sound dienen könnte. Es ist klar, dass grundsätzlich jeder Sound in jeder Art von Film vorkommen kann, da es auch sehr szenenabhängig ist. Ein extrem horror- mäßiger Sound kann beispielsweise natürlich auch in einer satirischen Komödie als Stilmittel Verwendung finden. Deshalb lautet die Fragestellung auch, in welcher Filmkategorie der Sound „typischerweise eher“ vorkommt.

Aber wenn man als Medienproduzent nach einem wirklich typischen Sound für einen Horrorfilm sucht (zB: Spannungsaufbau in einer heiklen Szene), dann

4.4. Testreihen – Testfragen, Auswertungsmethode und Ergebnisse

kann eine solche Einteilung bestimmt sehr sinnvoll sein. Vermutlich wäre es klug, nur eindeutig klassifizierbare Sounds dann wirklich einem Filmgenre zuzuteilen, und den Rest in einer Grauzone zu lassen.

Ergebnisse: (Durchschnittswerte der Sounds im Einzelnen)

Sound	Sci – Fi Ø	Fantasy Ø	Liebesfilm, Erotik Ø	Komödie Ø	Actionfilm Ø	Krimi Ø	Horrorfilm, Thriller Ø	Trickfilm, Kinderfilm Ø
1	3,6	3,4	1,4	1,5	3,2	4,0	4,9	1,2
2	3,1	3,5	4,1	2,5	1,9	1,8	1,6	3,4
3	3,9	3,6	1,6	1,6	3,3	2,6	3,7	1,4
4	1,7	3,3	1,5	2,3	1,6	1,6	2,1	3,7
5	3,8	3,9	1,3	1,4	3,4	3,1	4,5	1,2
6	2,8	3,6	2,8	1,3	2,5	2,8	3,5	2,0
7	2,3	3,0	3,8	2,2	1,8	1,8	1,6	2,0
8	4,6	4,0	1,3	1,3	4,0	3,2	4,2	1,3
9	2,6	4,0	4,2	2,5	1,6	2,0	1,4	3,5
10	4,3	4,5	1,0	1,4	3,3	2,2	4,7	1,4
11	3,2	2,4	2,5	1,8	3,4	2,7	2,3	1,3
12	2,6	3,4	4,0	1,7	1,3	1,6	1,4	2,5
13	4,1	4,0	1,6	1,3	2,6	3,6	3,9	1,3
14	2,3	3,1	1,4	1,7	2,5	2,4	3,1	1,8
15	4,0	2,5	1,5	1,6	4,3	3,0	3,6	1,4
16	3,9	4,0	2,2	2,4	1,8	1,9	2,6	2,6
17	4,4	3,9	1,2	1,4	3,2	3,3	4,4	1,6
18	3,4	4,0	1,2	1,3	2,7	3,6	4,8	1,8
19	2,2	2,9	1,2	1,6	3,9	2,6	3,4	1,5
20	3,7	4,4	2,6	1,8	2,3	2,9	3,4	2,6

Abb 21: Tabelle „Durchschnittswerte der Filmgenre – Auswahl (Frage 5)“

4.4. Testreihen – Testfragen, Auswertungsmethode und Ergebnisse

Um eine brauchbare Auswertung dieser Frage zu erhalten, macht es wohl Sinn, Spitzenwerte aus den Ergebnissen herauszulesen.

Wenn ein Klang bei einem Genre gegenüber Anderen einen hohen Durchschnittswert hat, so könnte man den Sound als typisch für dieses Filmgenre einstufen.

So scheinen folgende Sounds passend zu sein für folgende Filmkategorien:

Science Fiction: Nr. 8, 10, 17

Fantasy: Nr. 10, 20

Liebesfilm, Erotik: Nr. 9

Komödie: kein eindeutiger Sound

Actionfilm: Nr. 15

Krimi: Nr. 1

Horrorfilm, Thriller: Nr. 1, 5, 8, 10, 17, 18

Trickfilm, Kinderfilm: Nr. 4, 9

Die eindeutigsten Zuordnungen gab es hier zu den Genres Horrorfilm, Fantasy und Science - Fiction. Die Sounds, die am häufigsten mit Horrorfilm, Thriller in Verbindung gebracht wurden, hatten logischerweise auch bei den Grundemotionen „Anspannung, Furcht“ die meisten Nennungen.

Eine Eingabemaske, bei der der Medienproduzent seine eigenen Audiofiles zu Filmgenres zuordnen kann wäre denkbar, wenn man sich die Eindeutigkeit mancher Nennungen ansieht.

5 Ausblick und Schlussbemerkungen

Zu den Usertests kann man abschließend sagen, dass sie größtenteils sehr gut verlaufen sind. Dadurch dass immer nur zwei Personen auf einmal den Test durchführten, bekam man auch einen guten Einblick darüber, mit welchen Fragen bzw. Sounds es die Testpersonen schwieriger oder leichter hatten. Vor allem wollten die Probanden den Test auch gewissenhaft durchführen, da auch immer eine Aufsichtsperson neben saß. Bei einem Online – „Massentest“ wäre das wahrscheinlich weniger der Fall gewesen.

Dadurch dass hauptsächlich Studenten am Werk waren, die im selben Zweig studieren und mit dieser Materie zu tun haben, war das Interesse an einem solchen Test und generell an dem Projekt „AllThatSounds“ auch von Haus aus gegeben.

Es wird interessant zu beobachten, wie sich das Projekt noch weiter entwickelt. Wie wird das Endprodukt aussehen? Es könnte z.B. eine fixe Soundlibrary mit gut – strukturierten Einteilungen und effizientem Suchsystem werden, ähnlich der Ring Music Library. Oder es könnte auch ein Tool werden, mit dem man seine eigenen bereits vorhandenen Sounds strukturieren und beschreiben kann, ähnlich dem Soundminer – System. Oder auch eine Mischung aus beiden Systemen: Es werden bereits fixe Samples mitgeliefert; zusätzlich kann man andere Samples ins System hineinbringen, um sie in Folge leichter zu finden.

Auch eine web – basierte Methode wäre natürlich denkbar. Dabei könnte man mehr auf semantische Kriterien oder den Einsatz in Medienproduktionen eingehen, als es zum Beispiel bei Findsounds.com der Fall ist.

Die Frage ist auch, wie viel Information sich bei diesem Sound – Such – System automatisieren lässt. Was kann man durch technische Analysen erreichen und reicht es aus, um Verknüpfungen zum inhaltlichen und emotionalen Charakter des Klanges zu finden? Eine voll – automatisierte Analyse aller Sounds, die danach perfekt eingeteilt und somit bestens auffindbar sind wird es wohl nicht

5.Ausblick und Schlussbemerkungen

geben. Dafür ist die Kluft zwischen technischen Eigenschaften eines Klages und dem was wir wirklich hören und damit verbinden wohl zu groß. Aber je mehr aus technischen Analysen gewonnen werden kann, umso besser ist es und umso weniger muss manuell geschehen.

Man hofft, dass die bestmögliche Mischung aus unterschiedlichen Ansätzen bald sämtlichen Sounddesignern zugänglich ist und das ewige Suchen nach dem richtigen Sound demnächst ein Ende hat.

ANHANG A:

Literatur und Quellennachweis:

AudioHQ (2005): Tags und Tagging – Formate

<http://www.audiohq.de/index.php?showtopic=22>

Cisneros, A. (2005): Metadaten

http://www.ibr.cs.tu-bs.de/courses/ss05/skm/ausarbeitungen/Cisneros_Metadaten.pdf

Flückinger, B. (2001): Sounddesign. Die virtuelle Klangwelt des Films, Schüren Presseverlag

Nilsson, M. (2001): ID3 made easy

<http://id3lib.sourceforge.net/id3/id3v1.html>

Raffaseder, H. (2002): Audiodesign, Hanser Fachbuchverlag;

(2005): Antragsformular für Kooperationsvorhaben „AllThatSounds“

Schaffartzik, S. (2004): Beschreibung von Klangobjekten

<http://server4.medienkomm.uni-halle.de/filmsound/kap1-2.1.htm>

Trimmel, M. (2002): Emotionstheorie von Robert Plutchik

http://homepage.univie.ac.at/michael.trimmel/motivation_ss2002/DieEmotionstheorievonRobertPlutchik.pdf

Van Leeuwen, T. (1999): Speech, Music, Sound, Palgrave Macmillan

Weiterführende Literatur:

Bammer, D. (2001): Wirkung von Musik auf den Menschen
<http://www.hausarbeiten.de/faecher/hausarbeit/mus/16453.html>

Cano, P. (2004): Sound Effects Taxonomy Management In Production Environments
<http://www.iaa.upf.es/mtg/publications/AES25-pcano.pdf>

Ivanov, R. (2004): Metadata Management Standard – Part 4
<http://www.ifs.univie.ac.at/ec2002/24>

Krautschik, C. (2005): Musikpsychologie
http://www.hisf.no/sts/Musikkterapi/hovudfag/semh01_krautschick.html

Wake, S. (1998): Sound Retrieval with intuitive verbal expressions
<http://www.icad.org/websiteV2.0/Conferences/ICAD98/papers/WAKE/WAKE.PDF>

ANHANG B:

Abbildungs,- und Tabellenverzeichnis:

- Abb 1: MPEG7 – Standard mit Deskriptoren
- Abb 2: last.fm
- Abb 3: pandora.com (<http://www.pandora.com>)
- Abb 4: Musicsurfer (<http://musicsurfer.iaa.upf.edu/>)
- Abb 5: Musicsurfer – Similarity Map (<http://musicsurfer.iaa.upf.edu/>)
- Abb 6: findsounds.com (<http://www.findsounds.com>)
- Abb 7: Soundminer (<http://www.soundminer.com/>)
- Abb 8: Kategorie „Schallquelle“
- Abb 9: Kategorie „Art des Sounds bzw. mediale Anwendung“
- Abb 10: Kategorie „Time“
- Abb 11: Kategorie „Tone“
- Abb 12: Kategorie „Klangcharakter“
- Abb 13: Kategorie „Intervall“
- Abb 14: Kategorie „Dichte“
- Abb 15: Kategorie „Überlagerung“
- Abb 16: Kategorie „Lautstärkeänderung & Dynamik“
- Abb 17: Kategorie „Perspektive“
- Abb 18: Tabelle „Häufigkeiten d. Beschreibungsarten“
- Abb 19: Tabelle „Durchschnittswerte d. Gegenüberstellungen (Frage 3)“
- Abb 20: Tabelle „Anzahl der Nennungen für Emotionsgruppen (Frage 4)“
- Abb 21: Tabelle „Durchschnittswerte d. Filmgenre - Auswahl (Frage 5)“

ANHANG C:

Danksagung:

An dieser Stelle sei all jenen Personen gedankt, welche maßgeblich an der Entstehung dieser Arbeit beteiligt waren.

Auf der einen Seite Herrn Hannes Raffaseder, Lehrbeauftragter an der Fachhochschule für Telekommunikation und Medien in St. Pölten, Projektleiter des Forschungsprojekts „AllThatSounds“ und Hauptbegutachter dieser Arbeit; auf der anderen Seite auch allen anderen Mitarbeitern dieses interessanten Projekts.

Des Weiteren möchte ich mich bei meinen Studienkollegen bedanken, die mir mit Ratschlägen bei der Erstellung dieser Arbeit geholfen haben.

Des Weiteren sei meiner Freundin Kerstin, meinen Bandkollegen sowie auch allen anderen Freunden und nahestehenden Personen gedankt, die mir die notwendige Ablenkung zu dieser Arbeit gebracht haben.

Mein Dank geht auch an meine Familie, die mich während der Verfassung dieser Arbeit unterstützt hat und mir die Möglichkeit geboten hat, ein Studium zu absolvieren. Danke!

ANHANG D:

Glossar - Fachausdrücke:

AAC: Audio – Kompressionsverfahren, welches von der MPEG (Moving Picture Experts Group) als Weiterentwicklung des MPEG-2 Standards spezifiziert wurde

Agens: Jenes Objekt, welches ein Geräusch verursacht (in Barbara Flückinger's Buch „Sounddesign – Die virtuelle Klangwelt des Films“)

AIFF: (Audio Interchange File Format) ist ein von Apple für den Macintosh entwickeltes, verlustfreies Datenformat zum Speichern von Audiodaten.

EBU: (European Broadcasting Union) ist die 1950 gegründete Europäische Rundfunkunion. 74 Fernseh,- und Rundfunkanstalten in 54 Ländern sind derzeit involviert.

FLAC: (Free Lossless Audio Codec) ist ein frei verfügbarer, verlustfreier Audio – Codec von der Xiph.Org Foundation.

Hüllkurve: Die Hüllkurve zeigt in grafischer Form den Verlauf des Pegels eines Tons oder Geräuschs an und besteht aus den Parametern Attack, Decay, Sustain und Release.

Putty: ist ein von Simon Tatham entwickeltes freies SSH- und Telnet-Client-Programm für und dient dazu, eine sichere und bei Bedarf auch verschlüsselte Verbindung von einem Rechner zu einem anderen Rechner, meist einem Server, herzustellen.

Synthesizer: ein elektronisches oder elektronisch gesteuertes Gerät zur Synthese von Schwingungen bzw. zur künstlichen Klangerzeugung. Am häufigsten wird als

Synthesizer ein Tasteninstrument bezeichnet, das auf elektronischem Wege Klänge erzeugen kann.

Tag: Schlagwörter, die in Metadaten – Felder vorkommen, um Dateien für die leichtere Auffindung zu beschreiben.

UKO: Unidentifiziertes Klangobjekt nach B. Flückinger = Klangobjekt, das weder im Bild sichtbar, noch aus dem Kontext zu erschließen wäre.

UTF-8: (Unicode Transformation Format) ist eine Methode, Unicode-Zeichen auf Folgen von Bytes abzubilden. Zeichencodes mit einem Wert von weniger als 128 werden in UTF-8 als ASCII-Code dargestellt.

VST: (Virtual Studio Technology) ist eine von der Firma Steinberg entwickelte Schnittstelle zwischen einem VST – Host (z.B. Programme wie Cubase) und virtuellen Instrumenten und Effekten, die sich als Plug-In in der Software betreiben lassen.

XML: (Extensible Markup Language) ist ein Standard zur Erstellung maschinen- und menschenlesbarer Dokumente in Form einer Baumstruktur, der vom World Wide Web Consortium (W3C) definiert wird.