



TECHNISCHE UNIVERSITÄT ILMENAU

Fakultät für Elektrotechnik und Informationstechnik

Masterarbeit

Individuelle Auswahl nicht individueller Außenohrübertragungsfunktionen

| | |
|------------------------------|---|
| vorgelegt von: | Hui Fei |
| eingereicht am: | 13.04.2012 |
| geboren am: | 17.07.1985 in Zhejiang |
| Studiengang: | Elektrotechnik und Informationstechnik |
| Studienrichtung: | Informations- und Kommunikationstechnik |
| Anfertigung im Fachgebiet: | Audiovisuelle Technik Institut für Medientechnik |
| Verantwortlicher Professor: | Prof. Dr.-Ing. Hans-Peter Schade |
| Wissenschaftlicher Betreuer: | Dipl.-Ing. Stephan Werner Dipl.-Ing. Florian Klein |

Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die mich bei dieser Arbeit immer wieder motiviert und unterstützt haben.

Mein herzlicher Dank gilt in erster Linie meinen Betreuern Herrn Dipl.-Ing. Stephan Werner und Herrn Dipl.-Ing. Florian Klein für ihre ausgezeichnete und engagierte Betreuung dieser Arbeit. Ohne ihre fachliche und menschliche Unterstützung wäre diese Arbeit nicht möglich gewesen.

Außerdem bedanke ich mich bei Herrn Prof. Dr.-Ing. Hans-Peter Schade, der mir es ermöglicht hat, diese Arbeit am Fachgebiet Audiovisuelle Technik machen zu können.

Ebenfalls möchte ich noch meinen Freunden Denis Hoffmann und Steffi Hoffmann danken, dass sie mir bei der sprachlichen Überarbeitung geholfen haben. Dank gilt auch meinen Freunden Yongbin Zhou und Sven Sammer, dass sie mir für den Aufbau des Auswahl-Tools und der Struktur der Arbeit viel geholfen haben.

Danken möchte ich auch meinen Eltern, die mich in Zeiten des Zweifels während meines gesamten Studiums und beim Erstellen dieser Arbeit immer wieder motivierten. Sie gaben mir den nötigen Halt, mein Masterstudium erfolgreich zu absolvieren.

Kurzfassung

Ein virtuelles Hörsystem (Virtual Auditory Display) dient der Erzeugung des räumlichen Schalls. Durch die Faltung der Außenohrimpulsantwort (AOIA) bzw. Außenohrübertragungsfunktion (AOÜF) mit den Audiosignale der Schallquellen kann die Richtungsdarstellung realisiert werden. Die AOÜFs, die im VAD verwendet werden, müssen für jeden Benutzer individualisiert werden. Fehler wie "Vorne-Hinten-Vertauschungen", "Im-Kopf-Lokalisierung" usw. würden häufig eintreten, wenn die AOÜFs nicht individualisiert würde. Obwohl die aktuelle Methode, die direkte Messung der AOÜFs, die die besten individualisieren AOÜFs liefert, erfordert sie für die Zuhörer eine besondere Messeinrichtung und eine lange Zeit für das Messen, die nicht für eine Client-Anwendung geeignet ist. In dieser Arbeit, wird ein zur Zeit neuer Trend der Individualisierung der AOÜFs - individuelle Auswahl nicht individueller AOÜFs diskutiert. Zwei Auswahlverfahren werden durch Matlab aufgebaut und programmiert. Außerdem werden beide Auswahlverfahren durch ein Lokalisierungstest ausgewertet. Die Ergebnisse zeigen die Durchführbarkeit der Methoden und die Optimierungsmöglichkeit für bessere Ergebnisse in der Zukunft.

Abstract

A virtual auditory display (VAD) is a system for rendering spatialized sound to a listener, which can be achieved by convolving head-related transfer functions (HRTFs; head-related impulse response im zeitbereich) to a sound source. The HRTFs used in the VAD must be individualized to the listener. The localization errors like front-back confusion, in-head-localization and so on would happen frequently, if the HRTFs were not individualized. Although the current method, a direct measurement of HRTFs for the listeners, can afford the most accurate individual HRTFs, it requires a special measuring apparatus and a long time for measuring, which is not suitable for a Client application. In this thesis, a currently new trend of individualization of HRTFs - individual selection of non-individual HRTFs will be discussed. Two selection methods will be built and programmed by Matlab. Besides, both of them will be evaluated by a localization test. The results shows the feasibility and the direction of the optimization for a future prospect.

Inhaltsverzeichnis

| | | |
|----------|---|-----------|
| 1 | Einleitung | 1 |
| 1.1 | Hintergrund | 1 |
| 1.2 | Motivation | 2 |
| 1.3 | Aufgabenstellung | 3 |
| 1.4 | Aufbau der Arbeit | 3 |
| 2 | Theoretische Grundlagen | 5 |
| 2.1 | Binaurale Technologie | 5 |
| 2.1.1 | Binaural Encoding | 6 |
| 2.1.2 | Außenohrübertragungsfunktion | 8 |
| 2.1.3 | Virtuell Auditorisches Display | 9 |
| 2.1.4 | Binaural Decoding für den Kopfhörer | 10 |
| 2.2 | Methodik der Individualisierung der AOÜFs | 10 |
| 2.3 | Einführung der individuellen Auswahl | 12 |
| 2.3.1 | Methode einer subjektiven Auswahl der nicht individuellen AOÜFs | 12 |
| 2.3.2 | DOMISO | 14 |
| 2.4 | Bewertung eines virtuellen auditorischen Systems | 15 |
| 3 | Aufbau des Auswahl-Tools | 18 |
| 3.1 | Konstruktion der Datenbank | 18 |
| 3.1.1 | Vergleich der vorhandenen Datenbanken | 18 |
| 3.1.2 | Verknüpfung der Datenbanken | 20 |
| 3.2 | Wiedergabe des Testschalls | 21 |
| 3.3 | Programm und Schnittstelle | 24 |
| 3.3.1 | Zweistufige individuelle Auswahl durch Seeber | 24 |
| 3.3.2 | DOMISO | 26 |
| 3.3.3 | Fußball-Stil-Turnier | 27 |
| 3.4 | Vergleich und Bestimmung der Verfahren | 28 |

| | |
|--|-----------|
| 4 Hörtest und Auswertung | 31 |
| 4.1 Rahmen und Vorbereitung | 31 |
| 4.2 Ablauf | 31 |
| 4.3 Ergebnisse und Bewertung | 34 |
| 5 Ausblick | 43 |
| 6 Fazit | 44 |
| A Figuren | 46 |
| Literaturverzeichnis | 51 |
| Abbildungsverzeichnis | 55 |
| Tabellenverzeichnis | 57 |
| Thesen zur Masterarbeit | 58 |
| Erklärung | 59 |

1 Einleitung

1.1 Hintergrund

Die Multimedia-Technologien werden immer schneller entwickelt. Die Interaktionstechnologie, das Heimkino, der 3D-Film änderten unser tägliches Leben (Abbildung 1.1). Trotz der schnellen Entwicklung werden immer noch höhere Anforderungen an Multimedia-Technologien in der vergangenen Dekade gestellt.



Abbildung 1.1: Heimkino [Home]

Audio-Technologie, als ein wichtiger Teil der Virtual-Reality-Technologie und der Multimedia-Technologie, beeinflusst auf die Authentizität der Wahrnehmung der Anhörung. Virtual Auditory Display, ein neues System für die Erzeugung der virtuellen Schallsignale, wurde in den letzten Jahren immer wichtiger. Der Begriff "Auditory Display" bedeutet die Verwendung der Schallsignale, um die Information von einem Computer für den Benutzer zu kommunizieren. Im Gegensatz zu den vorhandenen

Technologien wie Stereophonie, welche kostspielige Anlagen benötigt, kann diese Technik durch einen Computer realisiert werden, was bequemer und billiger ist. Es ist eine wunderbare Nachricht für den Endverbraucher, wie zum Beispiel Computerspieler. Grund hierfür ist eine große benötigte Benutzerfreundlichkeit. Seit der ersten Erforschung dieser Techniken in International Community for Auditory Display (ICAD) zieht sie immer mehr Aufmerksamkeit auf sich.

1.2 Motivation

Der Mensch hat 2 Ohren kann aber einen Schall in einem dreidimensionalen Raum lokalisieren. Seit langer Zeit wird das Hörsystem analysiert, um die räumliche Illusion der Schallquelle (i.e. Virtual Auditory Display, VAD) zu realisieren [Nico10]. Vorhanden wird die binaurale Technologie als die beste Lösung der virtuellen Realität zur Verfügung gestellt. Die binaurale Technologie begreift eine Menge der Werkzeuge für die Synthese und Wiedergabe der binauralen Signale an die Ohren des Zuhörers in sich. Im Gegensatz zu der Technologie der Klangräumlichkeit wie die Stereophonie, die Mehrkanal-Tonsysteme, die Wellenfeldsynthese oder Ambisonics, deren Zweck Vermittlung der Illusion der räumlichen lokalisierten Schallquelle zum Zuhörer ist, fokussiert die binaurale Technologie auf die wiedergegebenen Signale am Trommelfell, um dem Hörsystem eine originalgetreue Kopie aller auditiver Informationen zu bieten, die potentiell für Wahrnehmung und Lokalisierung eines Schalls verwendet werden können.

Außenohrübertragungsfunktion (AOÜF) ist der Rohstoff der binauralen Technologie, der die Beugung der akustischen Welle durch den Körper beschreibt. Sie beinhaltet die Richtungsinformation der Schallquelle. Mit Hilfe von AOÜFs können die virtuellen Schallquellen genau wiedergegeben werden. Jedoch gibt es ein Problem: Die AOÜFs sind streng individualisiert. Eine Methode zur Bestimmung der individuellen AOÜF ist die Messung der Übertragungsfunktionen mittels Sondenmikrofonen am Eingang des Gehörgangs oder in Trommelfellnähe. [Nico10] Diese Methode verspricht die genaueste Bestimmung der persönlichen AOÜFs eines Probanden. Die Messung von AOÜFs ist allerdings keine geeignete Methode für anwendungsorientierte Systeme. Die Durchführung der Messung ist zeitlich und technisch aufwendig. Eine alternative Methode stellt die individuelle Auswahl von nicht individuellen AOÜFs dar. Diese Methode ist zeitsparend und billig, was den kommerziellen Anwendungen am wichtigsten ist.

1.3 Aufgabenstellung

Diese Arbeit legt den Fokus auf eine schnelle Auswahl der AOÜFs. Der Proband wählt dabei anhand von geeigneten Parametern ein Set von AOÜFs aus einer Datenbank aus. Das Auswahlverfahren kann dabei durch einen Hörversuch realisiert werden. In dem Versuch können relevante Qualitätsmerkmale, wie z.B. Externalisierung der Schallereignisse (Im-Kopf-Lokalisierungsfehler), Lokalisationsgenauigkeit, Vorne-Hinten-Vertauschungen und Klangfarbe abgefragt werden. Ein schneller Hörversuch kann durch einen Turnier-Ansatz umgesetzt werden. Der Proband soll die dargebotenen Schallsignale je Qualitätsmerkmal auf einer Qualitätsskala bewerten. Die Ausprägungen der Skala können von gut-schlecht bis hin zu einigen wenigen Zwischenstufen reichen oder mit der Priorität der Kriterien realisieren. Eine Evaluierung der Auswahlmethodik kann durch einen subjektiven Vergleich mit den besten, den schlechten Ausgaben und ggf. individuell gemessenen AOÜFs durchgeführt werden.

Aufgabenstellung dieser Masterarbeit ist die Entwicklung und Evaluierung eines Systems zur individuellen Auswahl von AOÜFs aus einer Vielzahl von AOÜFs. Dabei sind folgende Punkte zu beachten:

- Festlegung von Anforderungen an das Auswahlverfahren,
- Auswahl eines geeigneten Verfahrens zur individuellen Auswahl von AOÜFs,
- Entwicklung des Verfahrens zur individuellen Auswahl von AOÜFs,
- Bestimmung der Gewichtung nach der Relevanz der Qualitätsmerkmale,
- Entwicklung einer subjektiven Evaluationsmethodik,
- Evaluation des entwickelten Verfahrens durch einen Hörversuch kleineren Ausmaßes.

1.4 Aufbau der Arbeit

Im folgenden Kapitel 2 werden die theoretischen Grundlagen dieser Arbeit diskutiert. Der erste Abschnitt gibt einen Einblick über die binaurale Technologie. Das binaurale Encoding, die Außenohrübertragungsfunktion, Virtuell Auditorisches Display und das binaurale Decoding für den Kopfhörer werden prinzipiell detailliert. Die Methodik der Individualisierung der AOÜFs wird danach eingeführt. Die Methodik der individuellen Auswahl, welche Schwerpunkt dieser Arbeit ist, wird daraufhin erklärt. Zwei Verweise

werden auch vorgestellt. Zum Schluss wird die Grundlage der Methodik der Bewertung diskutiert.

Im Kapitel 3 wird der Aufbau des Auswahl-Tools detailliert beschrieben. Zuerst wird die Gewinnung der Datenbank diskutiert. Verschiedene öffentliche Datenbanken werden verglichen, ausgewählt und dann verknüpft. Danach muss der Testschall wiedergegeben werden. Schließlich werden das Programm und die Benutzerschnittstellen dargestellt und erläutert. Zwei vorhandene und ein entwickeltes Verfahren werden zur Diskussion gestellt. Am Ende dieses Kapitels wird ein endgültiger Plan für den Hörtest festgelegt.

Im Kapitel 4 wird der Hörtest erläutert. 4.1 beschreibt den Rahmen und die Vorbereitung, während 4.2 den Ablauf des Hörtestes erklärt. Alles, in diesem Test benötigte, wird in diesen beiden Abschnitten diskutiert. Die Ergebnisse werden im folgenden Abschnitt gezeigt und bewertet.

Kapitel 5 zeigt den Ausblick der Verbesserung, während Kapitel 6 das Fazit verfasst.

2 Theoretische Grundlagen

In diesem Kapitel werden die dieser Arbeit gehörenden Grundlagen eingeführt. In dem ersten Abschnitt wird die binaurale Technologie diskutiert. Danach wird die Methodik der Individualisierung der AOÜFs vorgestellt. Die individuelle Auswahl bzw. zwei vorhandene Verweise werden schließlich erläutert. Die Methodik der Bewertung wird zum Schluss zur Diskussion gestellt.

2.1 Binaurale Technologie

Die binaurale Technologie ist eine Lösung für die Klangräumlichkeit, die dem Hören im realen Leben ähnlichst ist. Diese Technologie ahmt die menschliche Lokalisierung des Schalls im 3D-Raum nach [Nico10], siehe Abbildung 2.1. In diesem Abschnitt wird die binaurale Technologie eingeführt. Das binaurale Encoding, die AOÜF, das virtuell Auditorische Display und das binaurale Decoding für den Kopfhörer werden hier zur Diskussion gestellt.

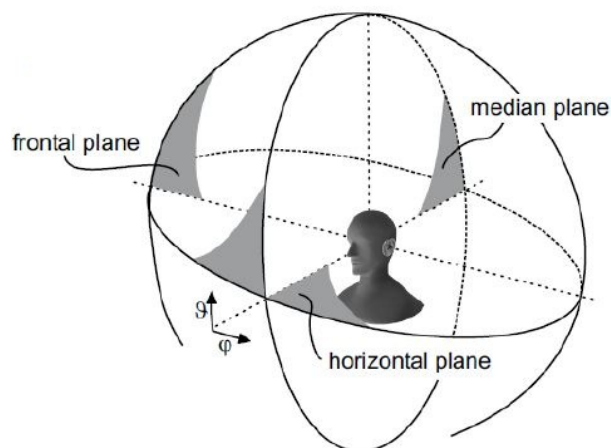


Abbildung 2.1: Klangräumlichkeit [Bran11]

2.1.1 Binaural Encoding

Das binaurale Encoding einer Schallquelle besteht entweder in der Aufnahme oder der Synthese des Schalldrucks am Eingang der Gehörgänge des Zuhörers [Møll]. Die Signale an den beiden Gehörgängen sind eine relevante Darstellung der Schallquelle, denn sie sind die minimalen Informationen, die das Hörsystem braucht, um eine Schallquelle zu lokalisieren. Das binaurale Encoding basiert auf der Synthese der Lokalisationsmerkmale, nämlich die ITD (Interaural Time Difference), die ILD (Interaural Level Difference) und die SC (Spectral Cues). [Nico10] Die ITD und ILD beschreiben die Unterschiede der empfangenden Schallsignale zwischen den beiden Ohren. Aber durch ITD und ILD kann eine Schallquelle nicht genau lokalisiert werden. Die Abbildung 2.2 zeigt den sogenannten “Kegel der Verwirrung” (cones of confusion). Die Übertragungsfunktionen auf dem gleichen Kreis haben die gleichen ITDs und ILDs. Die Folge davon ist die Verwirrung der Richtungen der Schallquelle. Deswegen wird SC mit den spektralen Merkmalen eingeführt, um die genaue räumliche Information zu beschreiben und dadurch die AOÜFs von der Verwirrung zu unterscheiden.

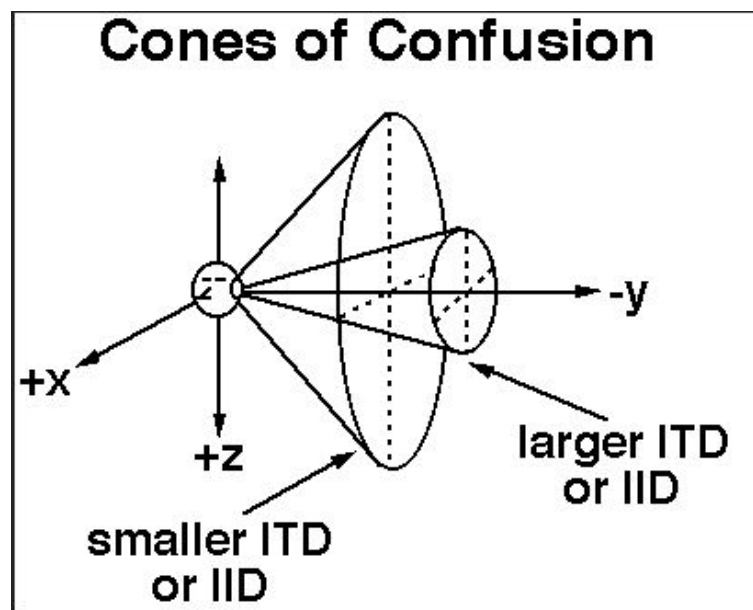


Abbildung 2.2: Cones of Confusion mit ITD und ILD [Nasa]

Viele Varianten sind vorhanden, um die binauralen Signale (normalerweise am Eingang des Gehörgangs oder Trommelfell) zu gewinnen. Eine typische Methode davon ist die Aufnahme einer natürlichen Schallquelle durch einen Kunstkopf, siehe Abbildung 2.3. Statt des Kunstkopfes kann sicherlich diese Methode auch am Menschen durchgeführt werden, siehe Abbildung 2.4.



Abbildung 2.3: Kemar Kunstkopf [LORE]



Abbildung 2.4: Binaural Recording am Menschen [Chri]

Die binaurale Synthese ist eine Variante, die häufig verwendet wird. Durch die Faltung einer Schallquelle mit einem Paar binauralen Filtern, die die Übertragungsfunktion des Schallwegs zwischen der Schallquelle und die Gehörgang des Zuhörers beschreiben, können die binauralen Signale wiedergegeben werden. Diese Übertragungsfunktion heißt Außenohrübertragungsfunktion (AOÜF) im Frequenzbereich oder Außenohrimpulsantwort (AOIA) im Zeitbereich. Durch den Aufbau einer Datenbank der AOÜFs kann die binaurale Synthese realisiert werden, siehe Abbildung 2.5. Die Parameter r, θ, ϕ sind der Radius, der Azimut und die Elevation. \tilde{h}_L und \tilde{h}_R sind die extrahierten

AOÜFs für das linke und rechte Ohr.

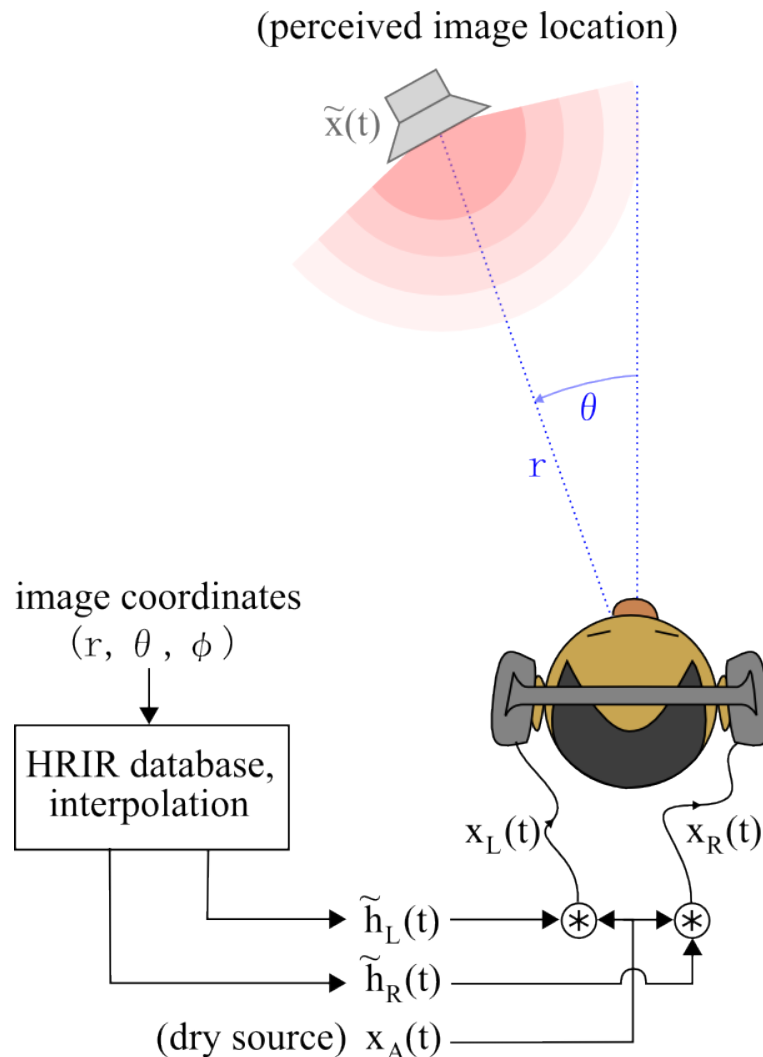


Abbildung 2.5: Binaurale Synthese durch ein Paar AOÜFs (\tilde{h}_L und \tilde{h}_R) aus einer Datenbank [Wiki]

2.1.2 Außenohrübertragungsfunktion

Außenohrübertragungsfunktion (AOÜF) charakterisiert die spektralen Änderungen in Betrag und Phase, die das Schallsignal auf dem Weg zum Trommelfell durch den Einfluss hauptsächlich der Ohrmuschel und des Kopfes, aber auch durch den Gehörgang und den Körper erfährt [Seeb03]. AOÜFs können die gesamten akustischen Wellen beinhalten, nicht nur die Ausbreitung im Freifeld, sondern auch die Beugung durch den Kopf des Zuhörers, die Reflexion des Körpers und die Reflexion und Beugung durch die Ohrmuscheln.

Für AOÜFs werden hier 3 Typen definiert [Bran11]. Dabei bedeutet H die Übertragungsfunktion. α, ϕ, f sind der Azimut, die Elevation und die Frequenz.

- Monaural: “reference” ist dasselbe Ohr mit definierten Bezugsschallquelle und Bezugsrichtungen.

$$HRTF_{monaural_left|right}(\alpha, \phi, f) = \frac{H_{arbitrary}(\alpha, \phi, f)}{H_{reference}(\alpha_{ref}, \phi_{ref}, f_{ref})} \quad (2.1)$$

- Interaural: Verhältnis zwischen linken und rechten Ohr.

$$HRTF_{interaural}(\alpha, \phi, f) = \frac{H_{left}(\alpha, \phi, f)}{H_{right}(\alpha, \phi, f)} \quad (2.2)$$

- Freifeld: “reference” ist die Messung des Mittelpunktes des Kopfes durch das Mikrofon.

$$HRTF_{freefield_left|right}(\alpha, \phi, f) = \frac{H_{left|right}(\alpha, \phi, f)}{H_{reference}(\alpha, \phi, f)} \quad (2.3)$$

In dieser Arbeit werden die Freifeld-AOÜFs verwendet. Sie ergeben sich aus dem Verhältnis der Fouriertransformation der Schalldrücke am Trommelfell und dem Schalldruck am Ursprung des Koordinatensystems (Abbildung 2.1) ohne Anwesenheit des Kopfes.

2.1.3 Virtuell Auditorisches Display

Mit Hilfe der AOÜFs kann das virtuell auditorische Display (VAD), realisiert werden. Ein VAD ist ein System für die Erzeugung der räumlichen Illusion des Schalls [Iway06]. Akio Honda definiert ein VAD als ein System für die Erzeugung und Übertragung der Schallsignale, die die räumlichen Charakteristiken beinhalten [Akio]. Es wird verwendet, um den Endbenutzern die Informationen der beiden Sprachsignale und nicht Sprachsignale zu präsentieren. Viele Forschungen in den verschiedenen Bereichen zeigen, dass das VAD eine Verbesserung der Benutzerfreundlichkeit leisten kann [Jack98] [Simo02] [Dere06] [Bruc06]. Sogar das Militär-Standard der USA, MIL-STD-1472F, weist darauf hin, dass das VAD unter einigen bestimmten Bedingungen bevorzugt benutzt werden sollte [MIL-99].

2.1.4 Binaural Decoding für den Kopfhörer

Das binaurale Decoding ist für die Wiedergabe der lokalisierten Signale. Die Wahrnehmung einer wiedergegebenen Schallquelle soll der realen Schallquelle so ähnlich wie möglich sein. Normalerweise sollen die binauralen Signale am Gehörgang oder Trommelfell, wo sie aufgenommen werden, wiedergegeben werden. Aber es ist nicht möglich, mit den normalen elektroakustischen Anlagen (Kopfhörer) das in der Praxis zu realisieren, zumindest ist es ohne Änderung der Wellenausbreitung unmöglich. Denn diese Anlagen sind nicht akustisch transparent [Nico10]. Trotzdem steht der Kopfhörer für die Wiedergabe der binauralen Signale an erster Stelle. Der Grund dafür ist, dass der Kopfhörer über 2 getrennte Kanäle, die dem Trommelfell so nah wie möglich sind, für die beiden Ohren verfügt. Vor der Benutzung eines Kopfhörers für das binaurale Decoding soll der Kopfhörer kompensiert werden, damit die Fehlanpassung und die Frequenzantwort minimiert werden [Møll].

Kopfhörerentzerrung

Werden Schalle über Kopfhörer wiedergegeben, so ist das Schallsignal im Gehörgang nicht exakt dem elektrischen Kopfhörersignal proportional, da es durch akustische Resonanzen und Nichtlinearitäten im Frequenzgang des Kopfhörers verändert wird. Um ein definiertes Schallsignal im Gehörgang einstellen zu können, sind diese Veränderungen in der Kopfhörerübertragungsfunktion zu charakterisieren [Seeb03].

Durch eine Korrektur der AÖUF mit der inversen Kopfhörerübertragungsfunktion kann die bei der Kopfhörerwiedergabe entstehende Veränderung des Schallsignals reduziert werden. Diese Methode heißt Kopfhörerentzerrung. Die Gleichung 2.3 wird umgeschrieben:

$$HRTF_{freefield_eq_left|right}(\alpha, \phi, f) = \frac{HRTF_{freefield_left|right}(\alpha, \phi, f)}{H_{left|right}^{kh}(f)} \quad (2.4)$$

Dabei ist $H_{left|right}^{kh}(f)$ die Kopfhörerübertragungsfunktionen der linken und rechten Ohren.

2.2 Methodik der Individualisierung der AOÜFs

Die Methodik der Individualisierung der AOÜFs lenkt die Aufmerksamkeit der Forscher auf sich. Viele Experimente wurden durchgeführt, um die Lösungen zu beweisen. Zusammengefasst sind die folgenden Methoden [Song07].

1. Die direkte Messung der AOÜFs [Pern03]: Durch dieses Verfahren können die genauesten AOÜFs gewonnen werden. Aber das ist auch zeitlich und technisch aufwendig. Es ist praktisch nicht möglich, die AOÜFs jedes Benutzers zu messen [Dmit03]. Daher ist diese Methode nicht für den kommerziellen Einsatz praktikabel.
2. Teilweise Individualisierung durch Mittelung der AOÜFs oder unter Verwendung typischer AOÜFs: die durchschnittlichen AOÜFs sind teilweise individualisierte AOÜFs. Diese AOÜFs können durch Mittelung der gemessenen AOÜFs der Probanden, des typischen Kunstkopfs oder durch theoretische Berechnung, die auf geometrischen Modellen wie Schneemann-Modell [Rich93] basiert, erhalten werden.
3. Individuelle Auswahl nicht individueller AOÜFs: Diese Lösung ist wahrscheinlich die einfachste. Die Probanden werden gefragt, aus einer Datenbank die ihnen am besten wahrgenommenen AOÜFs auszuwählen. Die Auswahl kann auf einen Hörtest basieren, damit die Probanden die Qualität der Räumlichkeit der AOÜFs einschätzen können [Iway06] [Bern03]. Eine andere Variante ist die Auswahl der ähnlichsten AOÜFs durch Vergleich der Körperparameter des Zuhörers mit den Parametern der AOÜFs in der Datenbank [Dmit03].
4. Individualisierung der AOÜFs durch Skalierung/Gruppierung nicht individueller AOÜFs: Zum Vergleich mit der direkten Messung ist die Gruppierung nicht individueller AOÜFs ein Kompromiss, um die relativ individualisierten AOÜFs zu erhalten. Experimente, die durch Wightman durchgeführt wurden, zeigen, dass mehrere hilfswise AOÜFs nötig sind, um ein verwendbares auditorisches System aufzubauen [Fred93].
5. Individualisierung durch theoretische Berechnung: Die AOÜFs werden durch theoretische Berechnung der individualisierten Körperparameter des Zuhörers (die Simulation der AOÜFs anhand von gescannten 3D Kopf- und Ohrmodelle) bestimmt [Ralp02] [Ralp99]. Diese Methode klingt einfach und machbar. Aber zum Vergleich der experimentellen Methoden ist sie nicht so genau. Denn die Berechnung basiert normalerweise auf vereinfachten geometrischen Modellen. Außerdem ist die Berechnung mit komplexen Körperparametern nicht leicht.
6. Indirekte Individualisierung durch physikalische Merkmale: Durch Wenzel wurde herausgefunden, dass die individualisierten Unterschiede der Beurteilung einer

Schallquelle in der Elevation am größten sind [Eliz88]. Weiter wurde auch gezeigt, dass die Genauigkeit der Lokalisierung eines Zuhörers durch die Charakteristiken der Außenohren vorhergesagt werden kann. Deswegen glauben die Forscher, dass die individualisierten AOÜFs durch Vergleich der Charakteristiken der Außenohren bzw. der anthropologisch Parameter gewonnen werden können [Dmit03].

7. Individualisierung durch Tuning: Active Sensory Tuning (AST) ist eine allgemeine Technologie für Optimierung der subjektive Kriterien durch die Suche der großen multidimensionalen Parameterräume. Runkle baute ein Model und stellte den Einsatz von AST für die Anpassung der generischen AOÜFs für Individuen [Paul00]. Eine andere Methode durch Andreas Silzle zeigt, dass Tuning und Auswahl der AOÜFs durch Tuning-Experten aus irgendwelcher Richtung individuell wiedergegeben werden können [Silz02].

2.3 Einführung der individuellen Auswahl

Diese Arbeit stützt sich auf die individuelle Auswahl nicht individueller AOÜFs. Wie oben genannt, basiert diese Methode auf die Auswahl des Benutzers. Die Auswahl wird normalerweise durch einen Hörtest realisiert. Diese Methode wird in dieser Arbeit verwendet, weil die Aufgaben des Benutzers einfach sind. Das heißt, dass diese Methode benutzerfreundlich ist. Das ist einem anwendungsorientierten System am wichtigsten. Laut der vorher durchgeführten Experimente werden hier zwei Verfahren als Verweise vorgestellt. Die beiden Verweise sind schon bewiesen worden, dass eine individuelle Auswahl der nicht individuellen AOÜFs funktioniert.

2.3.1 Methode einer subjektiven Auswahl der nicht individuellen AOÜFs

Ein Referenzexperiment mit 2 Stufen für die subjektive Auswahl der nicht individuellen AOÜFs wurde erst durch Bernhard Seeber und Hugo Fastl im Jahr 2003 an der TU München durchgeführt [Bern03].

- Vorbereitung: In diesem Experiment werden 5 Pulse des weißen Rauschens (20 Hz-20 kHz, 30 ms Dauer, 70 ms Pause, 5 ms gaußsche Steigung, 60 dB SPL) als Testsignal verwendet. Durch Faltung der Pulse mit den AOÜFs werden die virtuellen Richtungen in -40 , -20 , 0 , 20 , 40° in der Horizontalebene generiert. Der AUDIS-Katalog wird hier benutzt, um 12 verschiedene Sätze der AOÜFs zu

bieten. Diese Datenbank wird auch interpoliert, weil die Winkelauflösung dieser Datenbank 15° betragen. Weil eine vordefinierte Ordnung der Darstellung der AOÜFs mit der Gewöhnung des Zuhörers wahrscheinlich zu einem schlechten Ergebnis führt, wird hier der zufällige Zugriff der AOÜFs verwendet.

Die Sätze der AOÜFs mit den entsprechenden Nummern werden durch den Zuhörer eingeschätzt. Die schlechten Sätze sollen verworfen werden, während die ähnlichen Sätze wiederholt und verglichen werden können.

- Vorauswahl: 5 aus den 12 Sätze AOÜFs mit der besten räumlichen Wahrnehmung werden zuerst ausgewählt.
- Endauswahl: In der zweiten Stufe dieses Auswahlverfahrens wird ein Satz AOÜFs aus den fünf vorausgewählten AOÜFs ausgewählt. Für die Auswahl werden die folgenden Kriterien verwendet:
 - Der Schall wird von -40° links bis 40° rechts wahrgenommen, aber kann niemals außerhalb sein.
 - Der Schall bewegt sich horizontal in gleichmäßigen Schritten.
 - Der Schall hat immer eine konstante Elevation.
 - Der Schall wird in dem frontalen Bereich wahrgenommen (keine Vorne-Hinten-Vertauschung).
 - Der Schall wird mit einer konstanten Entfernung wahrgenommen.
 - Die Entfernung kann so weit wie möglich sein.
- Auswertung: Für die Auswertung dieser Methode werden 3 Stufen durchlaufen.
 1. Die 12 Sätze AOÜFs außer den durch die Vorauswahl ausgewählten 5 Sätzen werden von dem Besten zum Schlechtesten geordnet.
 2. Während der Endauswahl werden die 5 vorausgewählten Sätze auch geordnet.
 3. Ein zusätzliches Lokalisierungsexperiment wird ausgeführt. Die 5 durch die Vorauswahl ausgewählten AOÜFs werden ausgewertet. Die Zuhörer lokalisieren den virtuellen Schall von $\pm 50^\circ$, $\pm 40^\circ$, $\pm 20^\circ$, und 0° in der frontalen Horizontalebene mit Hilfe eines Laser-pointers. 10 Teilnehmer zählen 10 Versuche pro Richtung und AOÜF (insgesamt 7 Richtungen*5 AOÜFs). Das Lokalisierungsexperiment liefert die objektiven Kriterien für die Lokalisationsfähigkeit der ausgewählten AOÜFs.

Durch die 3 Experimente können die AOÜFs ausgewertet werden.

2.3.2 DOMISO

Die zweite Variante der individuellen Auswahl der AOÜFs wird durch den Prozess namens Determination method of Optimum Impulse-response by Sound Orientation (DOMISO) realisiert [Iway06]. Die Abbildung 2.6 zeigt das Schema von DOMISO. Diese Prozedur läuft wie folgend:

- Zufällige Auswahl von 32 Sätze AOÜFs aus der Datenbank
- Eine Laufbahn inklusive 13 virtuellen Klangbilder (Ein-Sekunden-langes Rosa Rauschen, 16 Bits, $f_s = 48$ kHz) für je 30 Grad gegen den Uhrzeigersinn von der Vorderseite eines Zuhörers in der horizontalen Ebene wird vorbereitet. Die ausgewählten 32 Sätze AOÜFs werden auch eingesetzt.
- Die Laufbahn der virtuellen Klangbilder wird den Probanden durch eine Illustration vor dem Hörtest gezeigt.
- Die Turnierspiele für 32 Klangbilder, die durch die 32 Sätze AOÜFs wiedergegeben werden, werden geplant.
- In jedem Spiel wählen die Probanden ein besser wahrgenommenes Klangbild aus. Er geht mit den nachfolgenden Spielen weiter.
- Zum Schluss gewinnt ein Satz AOÜFs das ganze Turnier. Das ist der angepasste Satz der AOÜFs.

In dieser Prozedur wird ein schweizerisches-Stil-Turnier durchgeführt. Die Spiele werden zwischen den gleichzeitigen Gewinnern geplant. Eine Niederlage von 2 guten Sätze AOÜFs in der frühen Phase des Turniers wird vermieden.

Sogar die perfektste Datenbank der AOÜFs kann nicht hundertprozentig zu den AOÜFs aller Menschen passen, deswegen soll die Durchführbarkeit von DOMISO ausgewertet werden. Um die Genauigkeit der Darstellung der Schalllokalisation zu bewerten, wird ein kleiner Hörtest mit 5-7 Probanden durchgeführt.

Die individualisierten Sätze AOÜFs werden zuerst durch DOMISO bestimmt. Die AOÜFs der Probanden in einigen Richtungen werden durch die direkte Messung mit Sondenmikrofonen gemessen. Die Genauigkeit der Darstellung der Schalllokalisation wird unter 3 Bedingungen mit 3 Sätzen AOÜFs eingeschätzt (die eigenen AOÜFs, die durch DOMISO individualisierten AOÜFs und ein Satz AOÜFs, das im Turnier völlig unmöglich zu gewinnen ist). Die Ergebnisse werden analysiert und verglichen.

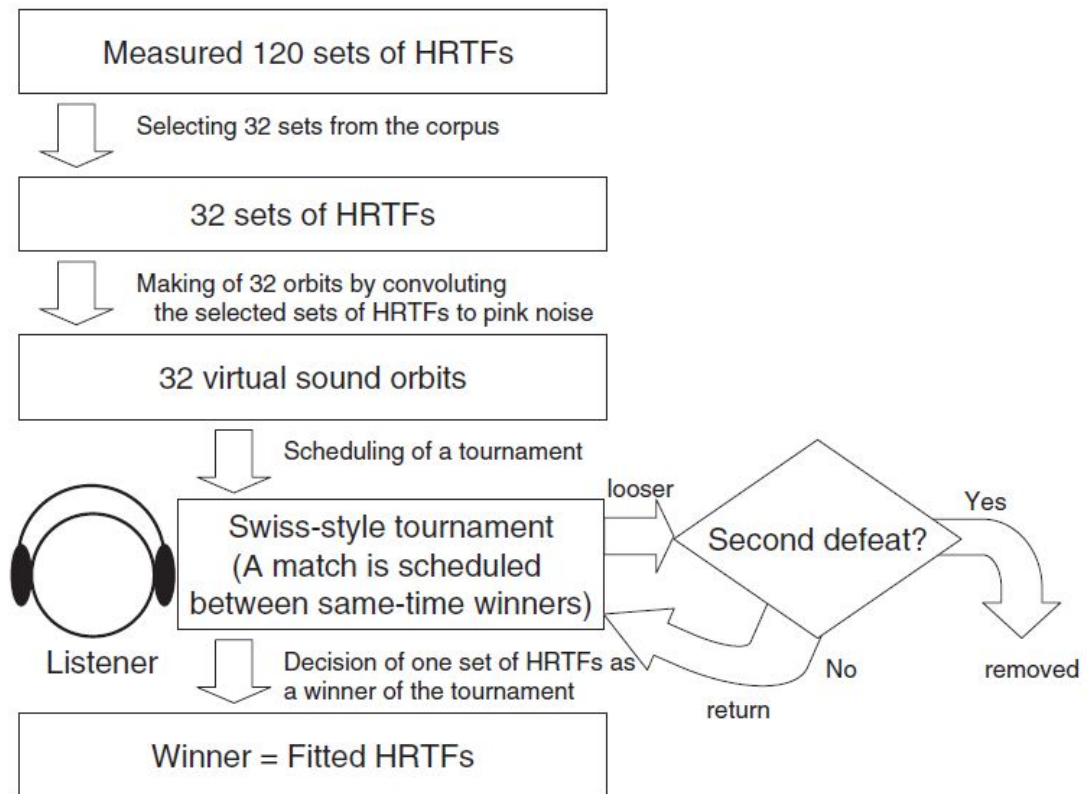


Abbildung 2.6: Schema von DOMISO [Iway06]

2.4 Bewertung eines virtuellen auditorischen Systems

Nach dem Aufbau des Auswahl-Tools sollen die Methoden bewertet werden, ob sie eine gute Illusion der virtuellen Realität schaffen. Obwohl die oben genannten Verfahren die eigenen Bewertungsverfahren erklären, muss ein bestimmter Plan zur Bewertung in dieser Arbeit festgelegt werden. Normalerweise kann eine Bewertung aus zwei Ansichten geprüft werden [Nico10].

- Objektive Bewertung: die gemessene eigene AOÜF definiert das Ziel der ausgewählten AOÜF. Durch Berechnung und Vergleich der Ähnlichkeit kann die ausgewählte AOÜF eingeschätzt werden.
- Subjektive Bewertung: In einem Hörtest werden die Probanden gefragt, ob sie die Räumlichkeit der Wiedergabe gut wahrnehmen können. Wegen der mehrdimensionalen Dimension eines Raums wird die Genauigkeit eines binauralen Filters mehr oder weniger beeinflusst. Dafür müssen mehr Aufgaben der Bearbeitung der Ergebnisse gemacht werden.

Objektive Bewertung

Die Rohstoffe der objektiven Bewertung sind die gemessene eigene AOÜF und die ausgewählte AOÜF. Zum Vergleich der Ähnlichkeit der beiden AOÜFs sind viele Kriterien vorhanden [Nico06]. Die meisten davon basieren auf der Berechnung der spektralen Fehler (linear oder logarithmisch) der Magnitude der beiden AOÜFs. Für jede Richtung wird ein durchschnittlicher Fehler erhalten. Eine Gewichtung der Frequenz-Bins kann dazu angewendet werden. Die Tabelle 2.1 zeigt einige Beispiele [Nico10]. Davon bedeutet $H(i)$ die eigene AOÜF und $\hat{H}(i)$ die ausgewählte AOÜF.

| Kriterium | Definition |
|-------------------------|--|
| MSE (Mean Square Error) | $C_{MSE} = \frac{1}{N} \sum_{i=1}^N [H(i) - \hat{H}(i)]^2$ |
| CB (Critical Band) | $C_{CB} = \frac{1}{N} \sum_{i=1}^N \{\alpha(i)[H(i) - \hat{H}(i)]\}^2$ $\alpha(i)$ ist die Gewichtungsfaktor von Frequenz-Bins. |
| Fahn | $C_F = \frac{\sum_{i=1}^N [H(i) - \hat{H}(i)]^2}{\sum_{i=1}^N [H(i)]^2}$ |
| Avendano | $C_A = 10 \log_{10} \left\{ \frac{\sum_{i=1}^N [H(i) - \hat{H}(i)]^2}{\sum_{i=1}^N [H(i)]^2} + 1 \right\}$ |
| Durant | $C_D = \frac{1}{N} \sum_{i=1}^N \{20 \log_{10} [\frac{\hat{H}(i)}{H(i)} - \bar{d}]\}^2$ $\bar{d} = \frac{1}{N} \sum_{i=1}^N 20 \log_{10} [\frac{\hat{H}_2(i)}{\hat{H}_1(i)}]$ |

Tabelle 2.1: Beispiele der objektiven Kriterien

Subjektive Bewertung

Im Gegensatz zur objektiven Bewertung basiert die subjektive Bewertung auf der indirekten Messung. Tatsächlich ergibt sie keine Beurteilung einer AOÜF an sich, sondern die Bewertung des Aussehens durch Verhalten des Zuhörers. Ein Hörtest muss dafür durchgeführt werden. Die Qualität der Räumlichkeit wird dann gemäß der Lokalisierungsfehler beurteilt. Vier Typen der Fehler werden hier definiert. Die Vorne-Hinten-Vertauschung und Oben-Unten-Vertauschung gehören zum Kegel der Verwirrung (cones of confusion).

- Im-Kopf-Lokalisierungsfehler: Die Wiedergaben sind sehr diffus. Schlechte Räumlichkeit wird wahrgenommen, sodass der Zuhörer gar nicht lokalisieren kann.
- Vorne-Hinten-Vertauschung: Die Richtung der Wiedergabe wird wegen der unpassenden SCs (Spectral Cues) zwischen vorne und hinten vertauscht.
- Oben-Unten-Vertauschung: Die Richtung der Wiedergabe wird wegen der unpassenden SCs zwischen oben und unten vertauscht.

- Generell Lokalisierungsfehler: Die simulierte Schallquelle wird in einer falschen Richtung wahrgenommen.

Dafür sind die folgenden Kriterien einer Bewertung sehr wichtig [Bern03]:

- Externalisation: Minimierung der Im-Kopf-Lokalisierungsfehler;
- Minimierung der Anzahl der Vorne-Hinten- und Oben-Unten-Vertauschung;
- Anpassung der simulierten und wahrgenommenen Richtung.

Vergleich und Festlegung der Bewertungsverfahren

Die objektive und subjektive Bewertung eines Auswahl-Tools sind ganz unterschiedlich. Der größte Vorteil der subjektiven Bewertung ist die intuitive Ansicht der Qualität der AOÜFs. Deswegen wird diese Methode breit verwendet. Trotzdem sind die Mängel der subjektiven Bewertung auch offensichtlich. Die Einschätzung der Lokalisierung hängt von dem Zuhörer ab. Menschen sind bei dieser Ausgabe ganz unterschiedlich. Ein gut trainierter Proband kann bei der Lokalisierung die Aufgabe viel besser bewältigen. Ein anderer Mangel ist das Bias der Lokalisierung. Es passiert sogar bei der realen Situation kaum, dass eine Schallquelle hundertprozentig lokalisiert werden kann. Wie soll ein Fehler in einem Hörtest gezählt werden? Wie groß kann ein Bias sein? Über diese Fragen müssen sich vor einem Hörtest Gedanken gemacht werden.

Die objektive Bewertung basiert auf die Berechnung der Ähnlichkeit. Das ist mathematisch überzeugend. Die Berechnung ist einfach. Aber der fatale Mangel ist, dass das menschliche Hörsystem so kompliziert ist, dass es nicht einfach mathematisch modelliert und berechnet werden kann. Die AOÜF beinhaltet nicht nur die Lokalisierungsinformation, sondern auch andere Informationen wie Klangfarbe, Externalisierung, Natürlichkeit und so weiter. Alle Informationen können zu den spektralen Unterschieden der AOÜFs führen. Ein virtuell auditorisches System ist anwendungsorientiert. Ein anwendungsorientiertes System soll durch Anwendung bewertet.

Aus diesen Gründen wird ein solches System in dieser Arbeit durch einen subjektiven Hörtest geprüft.

3 Aufbau des Auswahl-Tools

In diesem Kapitel wird der Aufbau des Auswahl-Tools erklärt. Dafür werden zwei Kernprobleme diskutiert: Die Konstruktion der Datenbank und der Aufbau des Programms jedes Verfahrens. Im folgenden ersten Abschnitt wird die Lösung des ersten Problems gegeben. Danach wird die Wiedergabe des Testschalls eingeführt. Schließlich wird das Programm erläutert. 3 Verfahren mit der Benutzerschnittstelle werden dargestellt. Zum Schluss werden sie zur Diskussion gestellt, um einen endgültigen Plan für den Hörtest zu erstellen. Diese Arbeit wird mit Hilfe von Matlab programmiert und simuliert.

3.1 Konstruktion der Datenbank

Dieser Abschnitt stellt die Konstruktion der Datenbank vor.

Zum Erstellen des Entwurfs der Datenbank von AOÜFs, die in diese Arbeit verwendet wird, gibt es zwei Varianten. Die erste Methode ist die direkte Messung der AOÜFs mittels Sondenmikrofonen wie oben genannt, siehe Abbildung 3.1. Bei dieser Methode können die Rahmen der Datenbank und die Vielfältigkeit der AOÜFs unter Kontrolle gebracht werden. Eine gute Datenbank sollte nicht zu wenige Individuen beinhalten (normalerweise mehr als 40 Individuen) [Nico10]. Allerdings ist die direkte Messung dieser Arbeit nicht geeignet, weil diese Methode zeitlich und technisch aufwendig ist. Die andere Variante ist die Nutzung einer existierenden Datenbank, was einfach und effizient ist. Diese Arbeit basiert auf der individuellen Auswahl, deswegen werden die existierenden Datenbanken verwendet. Die Auswahl einer Datenbank ist trotzdem noch sehr wichtig. Aus diesem Grund werden hier einige Kandidaten vorgestellt und verglichen [Nico10] [Song07].

3.1.1 Vergleich der vorhandenen Datenbanken

Vorhanden sind vier Datenbanken zum Vergleich. Die erste heißt CIPIC-AOÜF-Datenbank, die durch Prof. V. Ralph Algazi beim CIPIC-Interface-Labor an der Uni-



Abbildung 3.1: Messung der AOÜFs [Hear]

versity of California entwickelt wurde [CIPIC] [Ralp99]. Die zweite ist LISTEN-AOÜF-Datenbank, die durch das IRCAM (Institut de Recherche et Coordination Acoustique/Musique) aufgebaut wurde [Insta]. Eine andere Variante, die ARI-Datenbank, stellt das Institut für Schallforschung von Wien zur Verfügung [Instb]. Die letzte Datenbank wurde durch das Akustische Information-Systeme-Labor an der Universität Tohoku gemessen [Acos].

Tabelle 3.1 zeigt den Vergleich der existierenden Datenbanken. Die Anzahl der Individuen ist einer Datenbank sehr wichtig, weil sie auf die Vielfältigkeit der AOÜFs beeinflusst [Bern03]. In dieser Tabelle wird gezeigt, dass Pr. Suzuki nur 3 Individuen hat. Deswegen wird sie nicht als eine Variante überlegt. Der Winkelbereich und die Winkelauflösung einer Datenbank sind auch sehr wichtig. Diese Tabelle zeigt, dass die Winkelbereiche von CIPIC, LISTEN und ARI sehr ähnlich sind, während die Winkelauflösungen sich unterscheiden. LISTEN verfügt über eine Winkelauflösung mit 15° in dem Azimut und der Elevation. ARI hat die beste Winkelauflösung sowohl im Azimut (2.5° inner $\pm 45^\circ$, 5° außerhalb) als auch in der Elevation (5°). CIPIC hat eine gute Winkelauflösung mit 5.625° in der Elevation aber eine unbeständige und grobe Auflösung im Azimut. Die unbeständige Auflösung hat die schwierige Aufgabe der Lokalisierung zur Folge. Außerdem lässt sich CIPIC schwierig zusammen mit den anderen wegen den fehlenden Richtungen verwenden. Aus dieser Sicht betrachtet, sind die LISTEN und ARI die besten Varianten. Weil eine gute Auswahl von HRTFs normalerweise über 100 Sätze braucht [Iway06], werden die beiden Datenbanken verknüpft.

| Parameter \ Datenbank | CIPIC | LISTEN | ARI | Pr.Suzuki |
|----------------------------|---|---|--|---|
| Anzahl der Individuen | 45 | 51 | 67 | 3 |
| Winkelbereich ¹ | Azimut: -80°~80°, Elevation: -45°~ 230.625° | Azimut: 0°~360°, Elevation: -45°~90° | Azimut: 0°~360°, Elevation: -30°~80° | Azimut: 0°~360°, Elevation: 0°~90° |
| Winkelauflösung | Azimut: ² , Elevation: 5.625°, | 15° | Azimut: 2.5°(inner ±45°), 5° (außerhalb), Elevation: 5° | Azimut: 5°~30°, Elevation: 10° |

(1)Die Front wird als 0°in Elevation und Azimut gezählt. Gegen den Uhrzeigersinn wird als minus Grad gezählt. (2)Der Azimut der CIPIC ist -80, -65, -55, -45:5:45, 55, 65, 80°.

Tabelle 3.1: Vergleich der Datenbanken

3.1.2 Verknüpfung der Datenbanken

Für die Verknüpfung der Datenbanken sollen nicht nur die Struktur der Datenbanken sondern auch die akustischen Merkmale der AOÜFs moduliert werden, damit ein Vergleich mit den AOÜFs aus den unterschiedlichen Datenbanken angestellt werden kann. Die erste Stufe ist die Verknüpfung der Mat-Files. Die Dateien der verschiedenen Datenbanken werden in einem Dateiordner nummeriert und sortiert. Die Dateien von LISTEN werden mit 1 bis 51 nominiert, während die von ARI mit 52 bis 118 genannt werden.

In der zweiten Stufe werden die Strukturen der beiden Datenbanken wegen des Unterschiedes verändert. Die Grundidee dafür ist: Sie bleiben, was sie sind. Sowohl die AOÜFs als auch die Parameter werden in 2 Structs für die beiden Ohren unter die Dateien sortiert und gespeichert. Sie werden mit "l_hrir" und "r_hrir" genannt. Die "position" und "content_m" Matrizen werden dann für die Tabellen der Richtungen und die AOÜFs unter den beiden Structs erzeugt. Die anderen Parameter, z.B. die Länge der AOÜFs bzw. die Abtastfrequenzen werden auch in den originalen Matrizen gespeichert.

Ein anderes Problem ist der Unterschied der Winkelauflösungen der beiden Datenbanken. Tabelle 3.1 zeigt, dass die Auflösung von ARI feiner als die von LISTEN ist. Allerdings müssen die Winkelauflösungen der Laufbahn der Testsignale im Azimut

gleich sein.

Abbildung 3.2 zeigt eine Laufbahn der Testsignale in der Horizontalebene. Die Vorderseite wird als 0° gezählt. Die Signale laufen gegen Uhrzeigersinn. D.h., dass die linke Seite 90° ist.

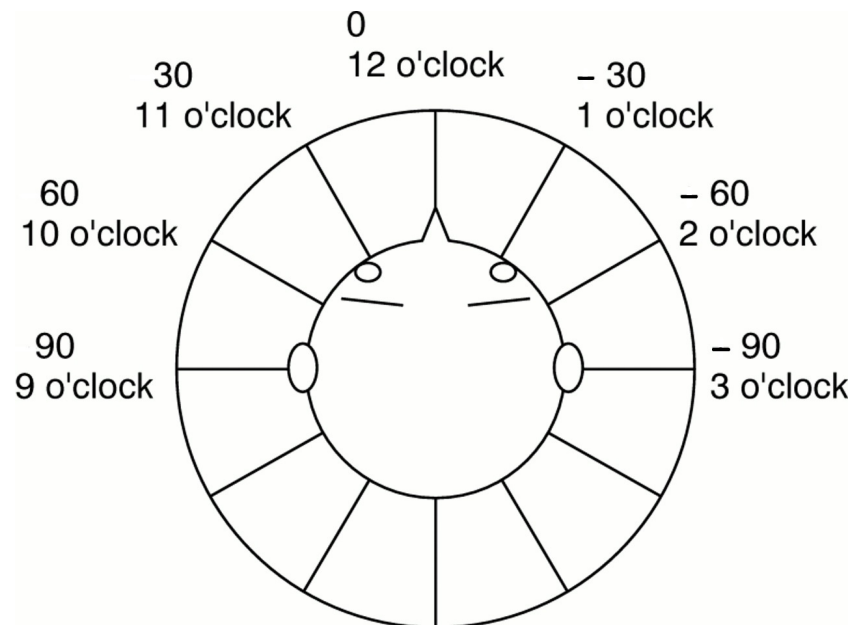


Abbildung 3.2: Horizontalebene der Laufbahn [Impa]

Im ersten Entwurf dieser Arbeit wurde eine Laufbahn mit 10° Schritten gedacht. Das bedeutet, dass die Winkelauflösungen der beiden Datenbanken im Azimut feiner als 10° sein müssen. Aber die Auflösung von LISTEN ist 15° . Eine Lösung dafür ist die Interpolation. Die AOÜFs dazwischen wurden mit den beiden nächsten AOÜFs linear interpoliert [Taka99]. Eine Folge dafür ist die Unvergleichbarkeit der Datenbanken, weil die ARI keine Interpolation braucht. Außerdem wirkt die Interpolation unangenehm, da sie unnatürlich ist.

Aus diesen Gründen werden eine Laufbahn mit 15° Schritten in dieser Arbeit in Anspruch genommen, damit keine Ansprüche an der Interpolation der Datenbanken gestellt werden.

Alle Programme für die Verknüpfung der Datenbanken können in der beiliegenden CD gefunden werden.

3.2 Wiedergabe des Testschalls

In diesem Abschnitt wird die Wiedergabe der Testsignale erklärt.

Das Testsignal wird als Rosa Rauschen empfunden, bei dem ein durchschnittlicher Mensch alle Frequenzbereiche des hörbaren Schallspektrums etwa gleich laut empfindet [Wiki12]. Das Signal wird mit den AOÜFs in den vorhergehenden Richtungen gefiltert. Die Ausgaben werden als Testschall verwendet. Dieser Arbeit soll die Durchführbarkeit der individuellen Auswahl unterstreichen. Bei Verwendung von nicht-individuellen HRTFs kommt es sehr häufig zu Fehlern in der Elevation. Trotzdem wird eine Laufbahn in der Horizontalebene dafür verwendet, weil die horizontalen Wiedergaben einfach lokalisiert und bewertet werden können. Für jeden Satz von AOÜFs werden insgesamt 24 Richtungen mit 15° Schritte wiedergegeben, siehe Abbildung 3.3. Das Signal in jeder Richtung ist ein 0.1-Sekunden-langes Rosa Rauschen. Die ganze Laufbahn dauert 2.4 Sekunden.

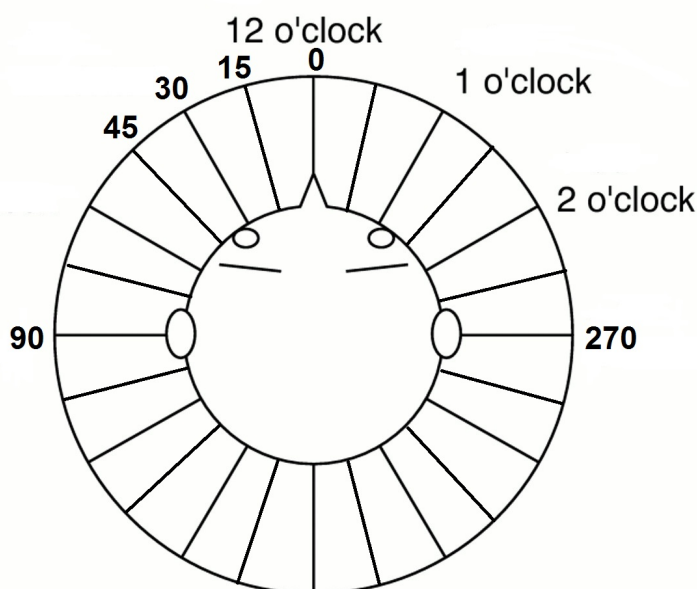


Abbildung 3.3: Winkelauflösung [Impa]

Wegen der unterschiedlichen spektralen Eigenschaften der AOÜFs von verschiedenen Datenbanken (kleine Magnitude der AOÜFs von ARI bei der Niederfrequenz unter 200 Hz) müssen die ausgegebenen Testsignale von LISTEN mit einem bestimmten Hochpassfilter (Grenzfrequenz 200 Hz) moduliert werden, damit die Klangfarben der Ausgaben aus den unterschiedlichen Datenbanken ähnlich klingen.

Wegen Verwendung des Kopfhörers müssen die wiedergegebenen Testsignale auch mit dem Kopfhörerentzerrungsfilter entzerrt werden. Der Kopfhörerentzerrungsfilter ist eine inverse Kopfhörerübertragungsfunktion. Durch Faltung des Testsignals mit der inversen Impulsantwort im Zeitbereich oder Multiplikation des Testsignals mit der

inversen Übertragungsfunktion können die Kopfhörerwiedergaben korrigiert werden. Die Abbildung 3.4 und 3.5 zeigen die Magnituden der beiden Kanäle des Kopfhörerentzerrungsfilters.

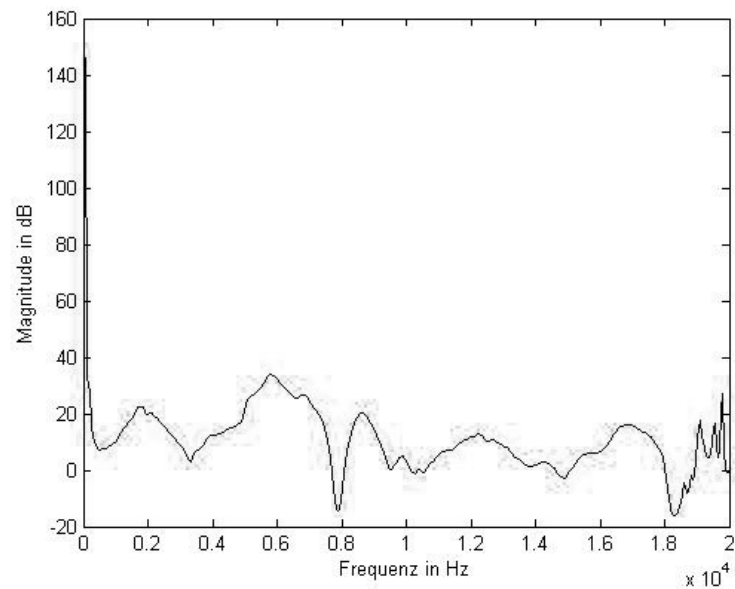


Abbildung 3.4: inverse Kopfhörerübertragungsfunktion des linken Kanals

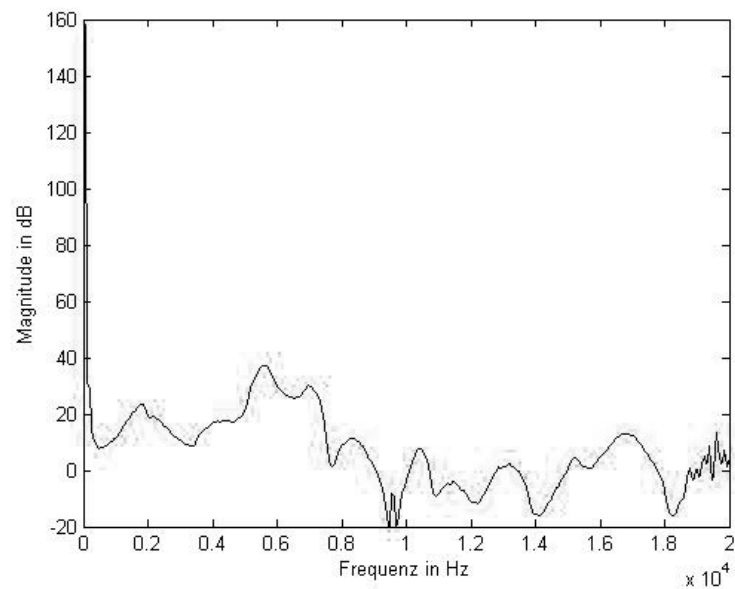


Abbildung 3.5: inverse Kopfhörerübertragungsfunktion des rechten Kanals

Zwei Anzeigemodi werden zur Auswahl für den Probanden gestellt. Der Standard ist eine kontinuierliche Laufbahn. Das heißt, dass die Signale in allen Richtungen zusammen berechnet und wiedergegeben werden. Ein Hann-Fenster mit 128 Punkten wird für die Ein- und Ausblendung (fade in/fade out) der benachbarten Signale verwendet, damit die Signale aus verschiedenen Richtungen natürlich klingen. Der andere Modus ist ein Anzeige-Stopp-Modus. Im Gegensatz zum Standard werden die Signale in 24 Richtungen nacheinander berechnet und wiedergegeben. Die Signale jeder Laufbahn sind nicht kontinuierlich, sondern intermittierend. Zwischen der Wiedergaben von 2 Richtungen wird eine Pause (kürzer als 0.1 Sekunde) gestellt. Der Grund für die Erstellung der beiden Modi ist, dass die Fähigkeit des Hörens bei jedem Individuum unterschiedlich ist. Es kann passieren, dass für einen Probanden die kontinuierliche Laufbahn in dem Hörtest ungewohnt ist. Während einer Auswahlphase muss einer der beiden Modi ausgewählt und nicht geändert werden.

Alle Programme für die Wiedergabe der Testsignale können in der beiliegenden CD gefunden werden.

3.3 Programm und Schnittstelle

In diesem Abschnitt werden die Programme und Benutzerschnittstellen für 3 Verfahren dargestellt. Nach dem Vergleich werden zum Schluss 2 geeigneten Verfahren davon bestimmt.

Alle Programme für das Auswahl-Tool können in der beiliegenden CD gefunden werden.

3.3.1 Zweistufige individuelle Auswahl durch Seeber

Die Grundidee dieser Methode wird vom durch Seeber und Fastl durchgeführten Experiment [Bern03] hergeleitet. Diese Auswahlprozedur besteht aus 2 Teilen: Einer Vorauswahl und einer Endauswahl.

Die Abbildung 3.6 zeigt die Benutzerschnittstelle der Vorauswahl. 32 von 118 Individuen werden aus der Datenbank für die Vorauswahl zufällig ausgewählt. Sie werden von 1 bis 32 nummeriert. Die "Go"-Taste hat die Funktion des Programmstartes. Nach dem Klicken wird die Laufbahn der derzeitigen AOÜFs wiedergegeben und der nächste Satz von AOÜFs für das nächste Klicken danach eingesetzt. Die Nummer der derzeitigen AOÜFs wird mit "Current No." dargestellt. Das Popup-Menü hat die ähnliche Funktion wie die "Go"-Taste nach dem Klicken der eingegebenen Nummer. Nach der

Wiedergabe einer Laufbahn sollen die Probanden sich entscheiden, ob sie diesen Satz von AOÜFs löschen wollen. Das Kriterium dazu ist einfach: Ob die Laufbahn auf einem guten Kreis in der Horizontalebene bleibt. Das heißt, dass sich die Laufbahn kreisförmig auf der Horizontalebene bewegen soll. Die “Y”-Taste bedeutet die Löschung dieses jetzigen Satzes, während die “N”-Taste Beibehaltung heißt.

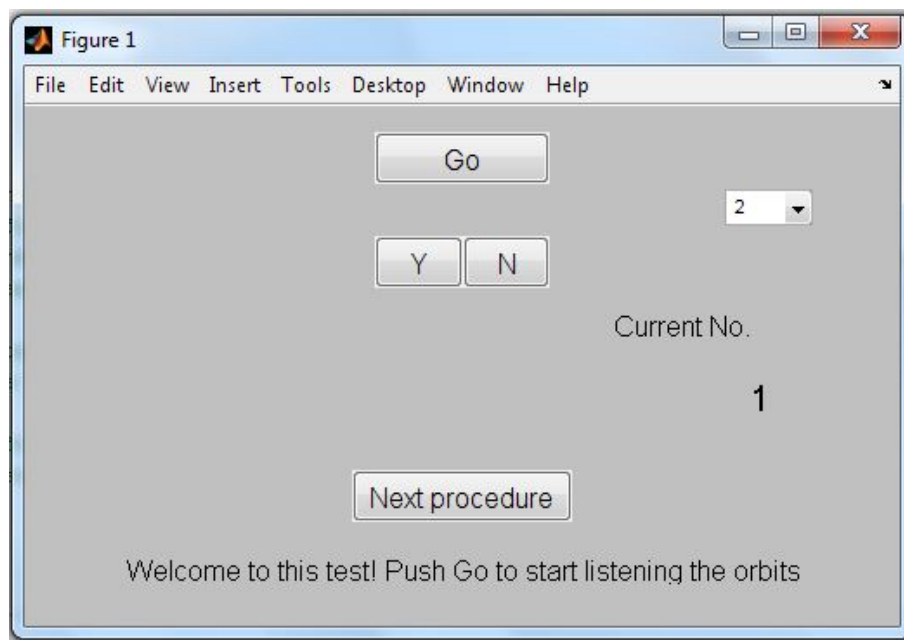


Abbildung 3.6: Schnittstelle der Vorauswahl der subjektiven Auswahl

Am Anfang dieser Prozedur können die Probanden sich natürlich nicht mit dem Hören weniger Sätze von AOÜFs entscheiden. Deswegen sollen sie die anfänglichen Sätze mit der “N”-Taste beibehalten. Nach dem Hören einiger Sätze können sie dann mit dem Popup-Menü die Sätze wiederholen und löschen. Zum Schluss werden 12 Sätze bleiben, siehe Abbildung A.1 im Anhang. Diese Schnittstelle wird den Probanden einen Tipp zeigen. Nachdem die Probanden auf die “Next procedure”-Taste klicken, wird die Schnittstelle der Endauswahl aufgerufen. Die Abbildung A.2 zeigt das Flussdiagramm der Endauswahl.

Die Abbildung 3.7 zeigt die Schnittstelle der Endauswahl. Diese Schnittstelle hat die gleiche “Go”-Taste, das Popup-Menü und das “Current No.”-Textfeld. Die “0”-Taste ist der Bezug des rosa Rauschens. Während der Endauswahl werden die 12 übrigen Sätze von AOÜFs mit 3 Kriterien eingeschätzt. Die Kriterien sind “Horizontal”, “Timbre” und “Smooth”, die im Mittel der Schnittstelle dargestellt werden. Die Noten dafür werden von 1 bis 10 im Textfeld danach eingegeben. Davon bedeutet 10 am besten. Mit dem Klicken der “OK”-Taste werden die eingegebenen Noten für den jetzigen Satz

berechnet. Der Rang für alle Sätze wird im “Rank”-Textfeld gezeigt. Am Ende dieser Prozedur wird der Gewinner dieser Methode im “Winner”-Textfeld dargestellt. Die Abbildung A.2 im Anhang zeigt das Flussdiagramm der Endauswahl.



Abbildung 3.7: Schnittstelle der Endauswahl der subjektiven Auswahl

3.3.2 DOMISO

Das DOMISO-Verfahren in dieser Arbeit wird vom durch Iwaya durchgeführten Test [Iway06] moduliert. Für eine schnelle Anwendung werden 16 aus 118 Sätzen von AOÜFs zufällig aus der Datenbank ausgewählt. Die Auswahlprozedur ist wie die in der originalen DOMISO-Prozedur. Die 16 Sätze werden für das schweizerische-Stil-Turnier arrangiert. Die Probanden wählen den ihnen besser passenden Satz in jedem Turnier aus. Nach dem Turnier aller Sätze wird eine neue Ordnung für die nächste Runde geordnet. Die Sätze, die zweimal ein Turnier verloren haben, werden gelöscht. Die Sätze, die gleich oft gewonnen haben, werden gesammelt und für die neuen Turniere arrangiert. Zum Schluss bleibt ein Satz, der als Gewinner gezählt wird.

Die Benutzerschnittstelle wird in der Abbildung 3.8 gezeigt. Das prinzipielle Kriterium dieser Methode ist, dass die Laufbahn auf einem Kreis in der Horizontalebene bleiben muss. Es wird mit der blauen Farbe einmal auf Englisch und einmal auf Deutsch im Textfeld dargestellt. Die “Repeat 1”- und “Repeat 2”-Tasten sind für die Darstellung der Laufbahn der 2 Sätze in einem Turnier. Die “1”- und “2”-Tasten sind zuerst

nicht sichtbar. Nach dem Hören der 2 Sätze werden die Tasten dann dargestellt. Die Probanden klicken auf eine Taste davon für den ihnen besseren Satz. Danach werden die nächsten 2 Sätze für das nächste Turnier eingesetzt. Die "Next Round"-Taste ist für das Arrangement der nächsten Runde. Die "Break"-Taste hat die Funktion des Abbruchs des Hörtestes falls sich ein Proband z.B. während der Auswahl unwohl fühlt. Am Ende dieses Verfahrens wird der Gewinner im "Final Winner"-Textfeld gezeigt. Die Abbildung A.3 zeigt das Flussdiagramm von DOMISO.

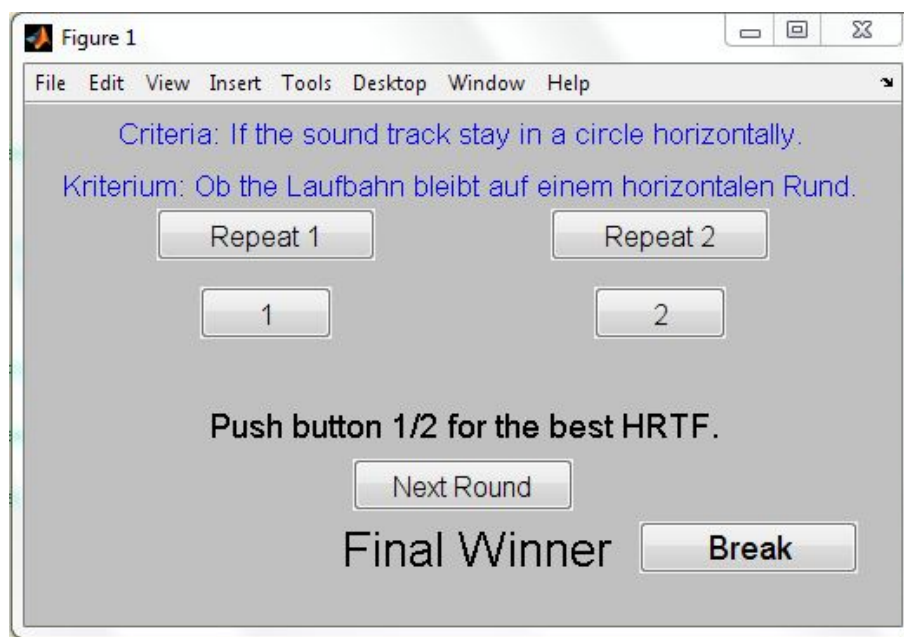


Abbildung 3.8: Schnittstelle von DOMISO

3.3.3 Fußball-Stil-Turnier

Hierbei wird eine ganz neue Auswahlmethode zur individuellen Auswahl in dieser Arbeit entwickelt und vorgestellt. Diese Methode wird mit der Name "Fußball-Stil-Turnier" genannt.

Diese Methode basiert auch auf Turnieren wie DOMISO. Aber das Organisationsmodell ist ein anderes. Die Grundidee ist von einem Fußballturnier wie z.B. UEFA Champions League oder UEFA European Football Championship. Im Unterschied zum Fußballspiel existiert hier in einem Spiel kein Remis. Außerdem wird zur Zeitersparnis nur ein Spiel zwischen 2 Teilnehmer anstatt 2 Spiele in einem Aufeinandertreffen der UEFA Champions League durchgeführt. Als Teilnehmer werden 16 Sätze aus der Datenbank zufällig ausgewählt.

Am Anfang werden die Vorrunden durchgeführt. Die 16 Sätze von AOÜFs werden in 4 Gruppen zufällig getrennt. In jeder Gruppe werden 6 Spiele zwischen 4 Teilnehmer ausgeführt. Nachdem werden alle 16 Teilnehmer geordnet. Die besten Sätze in jeder Gruppe werden für die K.o.-Spiele ausgewählt. Wenn die besten Sätze in einer Gruppe mehr als einmal existieren, wird einer davon zufällig ausgewählt. Nach den K.o.-Spiele bleibt dann der beste Satz übrig. Dieser ist der Gewinner dieses Verfahrens.

Die Abbildung 3.9 zeigt die Benutzerschnittstelle dieser Methode. Diese Schnittstelle ist DOMISO ähnlich. Die Unterschiede sind, dass die Nummern der Teilnehmer in der Datenbank und die derzeitige Gruppe im Mittel der Schnittstelle dargestellt werden. Außerdem wird der jetzige Progress neben der "Next Stage"-Taste gezeigt. Das Flussdiagramm ist auch DOMISO ähnlich, deswegen wird es hier nicht gezeigt.

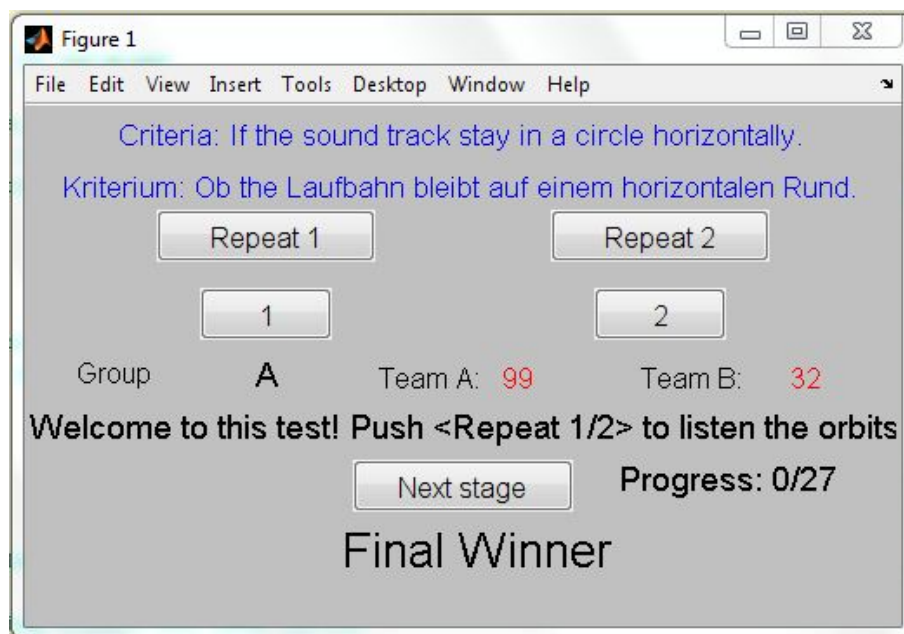


Abbildung 3.9: Schnittstelle des Fußball-Stil-Turniers

3.4 Vergleich und Bestimmung der Verfahren

In diesem Abschnitt werden die 3 Methoden verglichen und analysiert, um einen angemessenen Plan für den Hörtest zu bestimmen. Die 3 Methoden können in 2 Gruppen getrennt werden. Die zweistufige individuelle Auswahl durch Seeber basiert auf der Vielfältigkeit der Kriterien. Mehrere Kriterien werden verwendet, um die Qualität der AOÜFs einzuschätzen. Obwohl alle Methoden anhand von Kriterien entscheiden, wird diese Methode kriterienorientiert genannt. Im Gegensatz dazu basieren DOMISO

und Fußball-Stil-Turnier auf den Turnieren. Die Unterschiede dazwischen werden in der Tabelle 3.2 erläutert. [Iway06] [Bern03] [Sato08]

| Verfahren | Kriterienorientiert | Turnierorientiert |
|-----------|---|---|
| Vorzüge | Minimierung der Varianz der Lokalisierung und der Anzahl der Im-Kopf-Lokalisierung; | Einfache Aufgaben für die Zuhörer, keine spezifischen Kriterien brauchen (Auswahl für die großen Datenbank (120 Sätze) dauert 10-20 min); |
| | Optimum der AOÜFs normalerweise möglich; | Zufällige Zugriff in allen Turnierspiele, i.e. die Gewöhnung des Gehirns unmöglich; |
| | Einfacher objektive Auswertung durch das Experiment; | |
| Mängel | Zeitaufwand mit zu vielen Kriterien für die Auswahl aus einer großen Datenbank (Normalerweise über 100 Sätze von AOÜFs benötigt); | Keine Minimierung der Varianz der Fehler zwischen den ausgewählten AOÜFs und den eigenen AOÜFs, i.e. Kein Optimum. Für große Datenbanken existieren normalerweise eine Gruppe von AOÜFs, die alle durch die Prozedur als "das beste Satz" ausgewählt werden können; |
| | Während der Endauswahl führt die schnelle Anpassung des Gehirns zu den wiederholenden AOÜFs wahrscheinlich zu den schlechten Ergebnisse anstatt eines Optimums der AOÜFs; | Objektive Auswertung schwierig, weil das Ranking keine genaue Beziehung zur Minimierung der Varianz hat. Sogar der ausgewählte "beste" Satz ist sehr wahrscheinlich dem Zuhörer nicht der wirklich beste Satz in der Datenbank; |
| | Relativ schwere subjektive Auswertung, i.e. schwere Einschätzungsaufgaben für die Zuhörer; | Auswahlverfahren nur in der Horizontalebene, keine Auswertung in der Elevation. |
| | Auswahlverfahren nur in der frontalen Horizontalebene, keine Auswertung in der Elevation und Rückseite. | |

Tabelle 3.2: Vergleich der Methoden

In dieser Tabelle wird gezeigt, dass die beiden Typen Vor- und Nachteile haben, insbesondere der Mangel der Auswertung in der Elevation. Tatsächlich laut eines folgenden Experimentes für DOMISO hat eine Laufbahn der Wiedergabe auf 30° in der Elevation ein besseres Ergebnis der Auswahl zur Folge [Sato08]. Deswegen wird eine Folgerung aufgedeckt: die verschiedenen Methoden sollen für die verschiedenen Anwen-

dungen verwendet werden, z.B. für die Telefonkonferenz oder Computerspiele, in der normalerweise keine Bewegung in der Elevation existiert und keine Anforderung an der strengen Genauigkeit der individualisierten AOÜFs gestellt wird, wird ein turnierorientiertes Verfahren in der Horizontalebene verwendet; für die militärische Ausbildung, wird die kriterienorientierte Methode oder eine direkte Messung durchgeführt, weil großer Bedarf an der strengen Genauigkeit der Individualisierung der AOÜFs besteht. Diese Idee kann eine denkbare Erweiterung dieser Arbeit sein.

In dieser Arbeit ist die Geschwindigkeit eines anwendungsorientierten Systems sehr wichtig. Das kriterienorientierte Verfahren hat die guten Ergebnisse zur Minimierung der Fehlervarianz der Lokalisierung. Trotzdem ist dieses Verfahren zeitaufwendig. Mit der gleichen Anzahl der Sätze (16 Sätze) von AOÜFs dauert es über 30 Minuten ein Ergebnis auszugeben, während die turnierorientierte Methoden knapp 10 Minuten brauchen. Außerdem sind zu viele Kriterien den Benutzern sehr unangenehm. Die Kriterien müssen den Probanden erst erklärt werden. Danach ist ein Training dafür nötig, was sehr zeitaufwendig ist. Aus diesen Gründen werden DOMISO und Fußball-Stil-Turnier für den Hörtest ausgewählt.

4 Hörtest und Auswertung

In diesem Kapitel wird der zu dieser Arbeit gehörende Hörtest erklärt. Der folgende Abschnitt beschreibt den Rahmen und die Vorbereitung des Hörtestes. Dann wird der Ablauf vorgestellt. Die Ergebnisse werden am Ende dieses Kapitels gezeigt und ausgewertet. Die dazu gehörenden Testprogramme und die Rohdaten werden in der beiliegenden CD gespeichert.

4.1 Rahmen und Vorbereitung

Der ganze Hörtest wurde im Hörlabor durchgeführt. 7 Probanden, 2 weiblich und 5 männlich, nahmen daran teil. Alle Probanden sind Studierende oder Mitarbeiter an der TU Ilmenau. Die Altersstufe der Probanden geht von 20 bis 31 Jahren. Jeder Proband hatte einen einzigen Termin, der 2 Stunden dauerte. Dafür wurde ein Computer (Dell Laptop Studio 1535) zur Verfügung gestellt. Die Illusion der Schallquellen wurde durch einen Kopfhörer (STAX SRS-202) und den dazugehörenden Verstärker (SRM-252) wiedergegeben.

Desweiteren wurden 2 Programme für das Training und den Test des Lokalisierungstestes vorbereitet. Die dazu verwendeten Signale sind rosa Rauschen und ein telefonisches Sprachsignal für den Tipp: "We are sorry. You have dialed a number which cannot be reached from your calling area." [Phon]. Die beiden Signale sind breitbandig. Die Rohdaten des Hörtestes werden durch den Computer des Hörlabors in Form von Excel-Datei gespeichert.

4.2 Ablauf

Der Ablauf des Hörtestes besteht aus 3 Teilen: die Auswahlphase, die Trainingphase und die Testphase. Nach der Einführung wird der Test beginnen. Während des ganzen Hörtestes nehmen die Probanden auf einem Stuhl Platz. Dabei sollen sie eine angenehme Sitzposition wählen und diese halten. Während des gesamten Hörtestes wird der Versuchsleiter bei den Probanden im Raum sein und den Testverlauf kontrollieren.

Auswahlphase

Vor der Auswahlphase werden die Probanden an das Programm von DOMISO und Fußball-Stil-Turnier und die GUI gewöhnt. Die Gewöhnung dauert ca. 5 Minuten, damit die Probanden mit den Aufgaben vertraut sind. Danach werden sie mit der Auswahlphase anfangen.

Während dieser Phase werden den Probanden immer 2 Teilnehmer in jedem Turnier zur Auswahl gestellt. Das Kriterium dafür ist, dass die Laufbahn der Schallquelle auf einem Kreis in der Horizontalebene bleibt. Sie sollen den ihnen besser passenden Teilnehmer jedes Turniers auswählen. Die Anzahl der Turniere von DOMISO beträgt ca. 30, während 27 Turniere in Fußball-Stil-Turnier durchgeführt werden. Diese Phase dauert 15-20 Minuten (beide Auswahlverfahren ohne Training). Am Ende dieser Phase werden 2 Ausgaben als die besten Sätze der beiden Verfahren ausgewählt.

Trainingphase

In der Trainingphase für den Hörtest werden den Probanden insgesamt 14 Signale aus verschiedenen Richtungen vorgespielt. Jedes Signal wird zweimal vorgespielt. Die durch den "KEMAR"-Kunstkopf mit den großen Ohren gemessene AOÜFs [CIPI] werden in 6 gegebenen repräsentativen Richtungen zufällig wiedergegeben. Die Richtungen werden mit 12, 11, 7, 6, 4, 3 Uhren gezeigt, was den 0° , 30° , 150° , 180° , 240° und 270° im Azimut äquivalent sind. Keine Auswertung wird in der Elevation eingeschätzt, siehe Abbildung 4.1. Der Im-Kopf-Fehler, den die Probanden gar nicht lokalisieren können, wird mit 0 Uhr genannt. Der Testschall wird mit 2 Signalen wiedergegeben: einmal rosa Rauschen und einmal das Sprachsignal. Außerdem werden immer die Verweise der originalen Signale vor der Änderung der Signale vorgespielt.

Der Ablauf der Trainingphase ist folgenderweise:

- Vorspiel des originalen rosa Rauschens.
- Darstellung des rosa Rauschens in 6 Richtungen mit einem zufälligen Zugang.
- Vorspiel des originalen Sprachsignals.
- Darstellung des Sprachsignals in 6 Richtungen mit einem zufälligen Zugang.

Die Abbildung A.4 im Anhang zeigt die Zustandsdiagramm der Trainingphase. Nach der Phase wird der Lokalisierungstest durchgeführt.

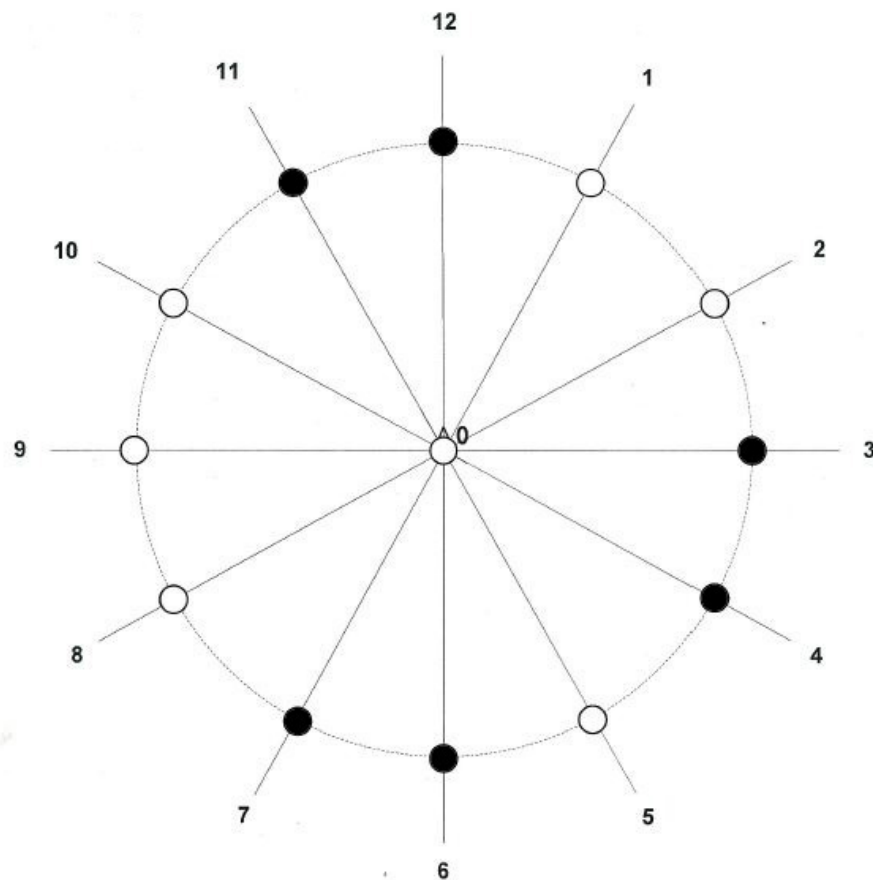


Abbildung 4.1: Zufälliger Zugang der 6 Richtungen, 12 Uhr bedeutet die Vorderseite des Zuhörers. Die schwarzen Kreise sind die gegebenen Richtungen.

Testphase

In der Testphase werden den Probanden insgesamt 56 Signale aus verschiedenen Richtungen vorgespielt. Jedes Signal wird zweimal wiedergegeben. Die beiden durch die Auswahlphase ausgewählten Sätze von AOÜFs werden in den gleichen gegebenen Richtungen zufällig wiedergegeben. Außerdem werden 2 Sätze von AOÜFs, die die schlechten Ergebnisse in den beiden Verfahren haben, zum Vergleich im Test eingesetzt. Die ganze Phase läuft wie die Trainingsphase. Im Unterschied dazu sind die Sätze von AOÜFs. Anstatt der KEMAR-AOÜFs werden der beste Satz von DOMISO, einer der schlechten Sätze von DOMISO, der beste Satz von Fußball-Stil-Turnier und einer der schlechten Sätze von Fußball-Stil-Turnier vorgespielt. Für die Klarheit werden sie als Domiso_b (der beste Satz von DOMISO), Football_b (der beste Satz von Fußball-Stil-Turnier), Domiso_w (einer der schlechten Sätze von DOMISO) und Football_w (einer der schlechten Sätze von Fußball-Stil-Turnier) abgekürzt. Außerdem werden immer die

Verweise der originalen Signale vor der Richtungsdarstellungen vorgespielt.

Der Ablauf der Testphase ist wie folgt:

- Vorspiel des originalen rosa Rauschens.
- Darstellung von Domiso_b mit dem rosa Rauschen aus 6 Richtungen mit einem zufälligen Zugang.
- Vorspiel des originalen Sprachsignals.
- Darstellung von Domiso_b mit dem Sprachsignal aus 6 Richtungen.
- Vorspiel des originalen rosa Rauschens.
- Darstellung von Football_b mit dem rosa Rauschen aus 6 Richtungen.
- Vorspiel des originalen Sprachsignals.
- Darstellung von Football_b mit dem Sprachsignal aus 6 Richtungen.
- Vorspiel des originalen rosa Rauschens.
- Darstellung von Domiso_w mit dem rosa Rauschen aus 6 Richtungen.
- Vorspiel des originalen Sprachsignals.
- Darstellung von Domiso_w mit dem Sprachsignal aus 6 Richtungen.
- Vorspiel des originalen rosa Rauschens.
- Darstellung von Football_w mit dem rosa Rauschen aus 6 Richtungen.
- Vorspiel des originalen Sprachsignals.
- Darstellung von Football_w mit dem Sprachsignal aus 6 Richtungen.

Die Abbildung A.5 zeigt das Zustandsdiagramm der Testphase.

4.3 Ergebnisse und Bewertung

Im diesem Abschnitt werden die Ergebnisse des Hörtestes dargestellt und ausgewertet. Die 7 Probanden wurden mit der gleichen Umgebung (Raum, Geräte, Auswahlverfahren, Testsignale und Lautstärke der Wiedergabe) die Lokalisierung getestet. Um die Wirkung der Lautstärke der Wiedergabe einzuschätzen, wurde ein Proband davon mit

den gleichen Ausgaben der AOÜFs aber unterschiedlicher Lautstärke der Schallsignale zum Vergleich gestellt.

Ein Proband, der die Vorderseite mit der simulierten Umgebung gar nicht lokalisieren konnte, wurde herausgegriffen und am Ende dieses Kapitels getrennt analysiert.

Die folgenden Abbildungen 4.2- 4.9 sind die zusammengefassten Ergebnisse der übrigen Probanden des Hörtestes. Alle Daten der 6 Probanden werden gemäß der Ausgaben und Testsignale dargestellt. Z.B. Domiso_b_noise 4.2 bedeutet die Ergebnisse aller Probanden mit den besten Ausgaben von DOMISO mit rosa Rauschen (Domiso_b_speech fürs Sprachsignal). In jeder Zeile werden 2 Abbildungen für die beste Ausgabe und einer der schlechten Ausgaben des gleichen Verfahrens mit den gleichen Testsignalen zum Vergleich dargestellt.

Die horizontale Achse zeigt die simulierten Richtungen (Uhren, siehe Abbildung 4.1), während die vertikale Achse die wahrgenommenen Richtungen aufweist. Die 0 Uhr bedeutet die Im-Kopf-Lokalisierung. Die Kreise in diesen Abbildungen zeigt zu der Größe die Häufigkeit der wahrgenommenen Richtungen proportional. Die Im-Kopf-Lokalisierungsfehler, die Vorne-Hinten-Vertauschung bzw. das Entsprechen der simulierten und wahrgenommenen Richtungen können in dieser Abbildung beobachtet werden. Die durch Spearman-Correlation ausgerechneten Abweichungen (ρ) von der idealen Lokalisationskurve werden in die Bildunterschrift geschrieben.

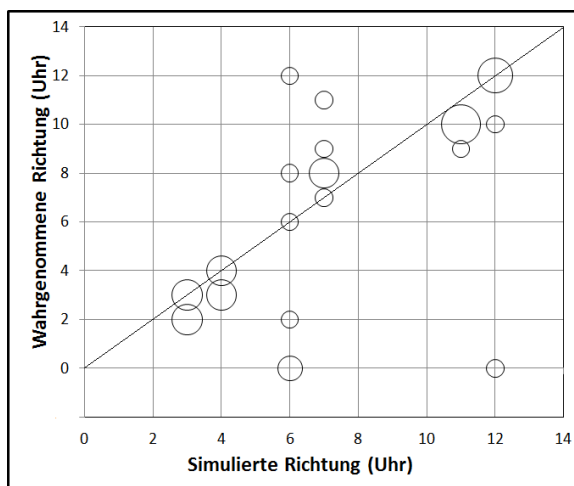


Abbildung 4.2: Domiso_b_noise (36 Datensätze), $\rho = 0.9879$

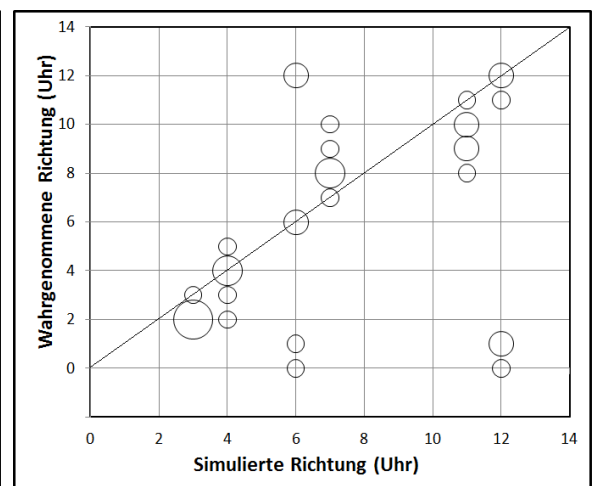


Abbildung 4.3: Domiso_w_noise (36 Datensätze), $\rho = 0.9812$

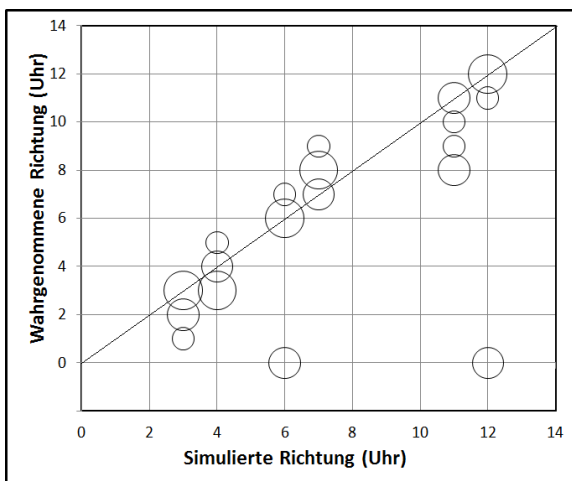


Abbildung 4.4: Domiso_b_speech (36 Datensätze), $\rho = 0.9946$

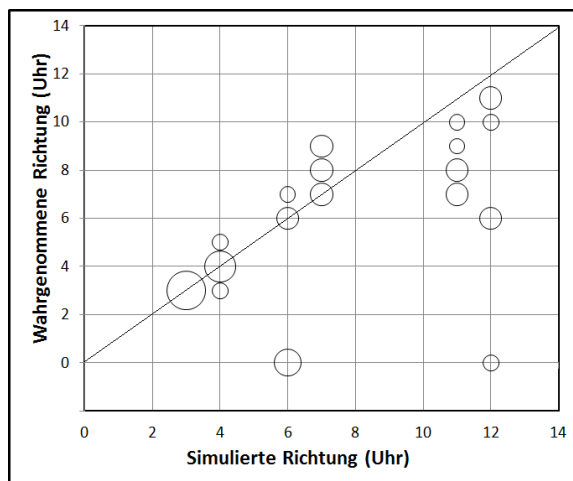


Abbildung 4.5: Domiso_w_speech (36 Datensätze), $\rho = 0.9905$

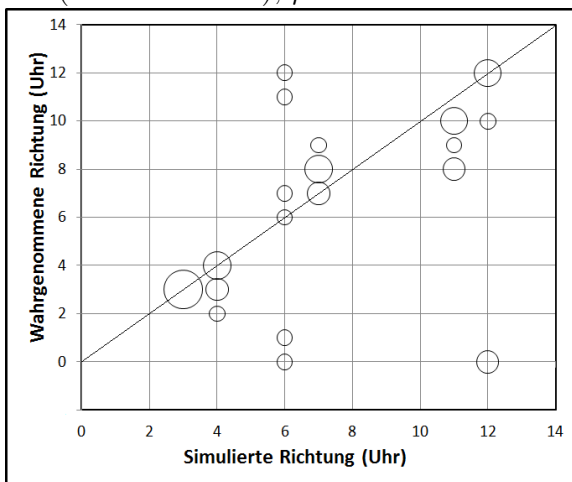


Abbildung 4.6: Football_b_noise (36 Datensätze), $\rho = 0.9834$

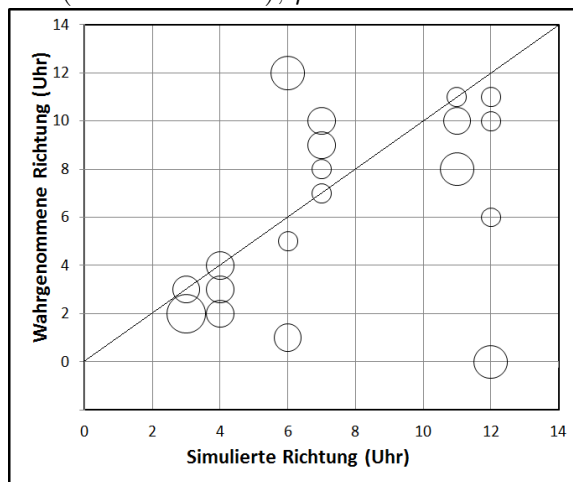


Abbildung 4.7: Football_w_noise (36 Datensätze), $\rho = 0.9653$

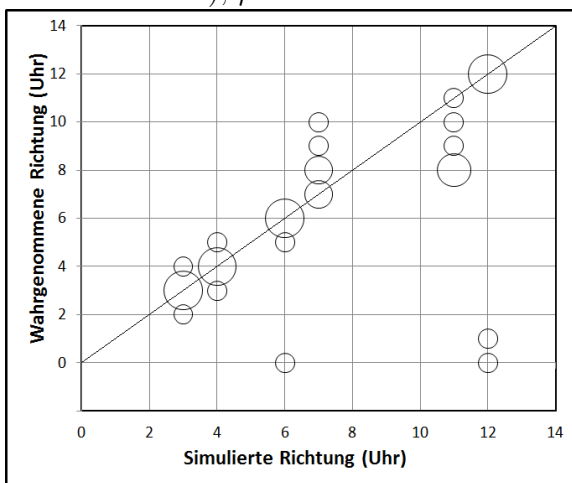


Abbildung 4.8: Football_b_speech (36 Datensätze), $\rho = 0.9932$

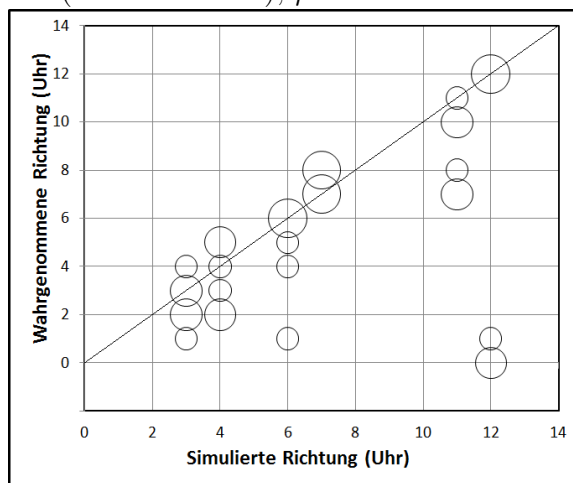


Abbildung 4.9: Football_w_speech (36 Datensätze), $\rho = 0.9878$

Bei einem guten Ergebnis sollte die Diagonale der beiden Achsen (die idealen Lokalisationskurve) konvergent sein. Aus den Abbildungen können wir generell einschätzen, dass die Ergebnisse der besten Ausgaben der Auswahlphase zu einer konvergenteren Linie als die Ergebnisse der schlechten Ausgaben führen. Durch die Abweichung von Spearman-Correlation wird es auch unterstrichen. Außerdem scheint es, dass das Sprachsignal im Lokalisierungstest besser als das Rauschen funktioniert. Mehr Schlussfolgerungen sollen mit Quantifizierung der Ergebnisse zur Diskussion gestellt werden.

3 Typen der Fehler traten in diesem Test auf. In dieser Arbeit werden die Richtungen, die bei den angrenzenden Richtungen wahrgenommen wurden (z.B. Der Proband hat ein mit 11 Uhr simuliertes Signal 10 Uhr wahrgenommen), nicht als Fehler gezählt. Die 3 Fehler sind:

- Im-Kopf-Lokalisierungsfehler: Die Wiedergaben sind sehr diffus. Schlechte Räumlichkeit wird wahrgenommen, sodass der Zuhörer gar nicht lokalisieren kann (hier wird als 0 Uhr definiert).
- Vorne-Hinten-Vertauschung: Die Richtung der Wiedergabe wird wegen der unpassenden SCs (Spectral Cues) ein Spiegelbild der Frontalebene (siehe Abbildung 2.1) zwischen vorne und hinten vertauscht (z.B ein 11 Uhr wird 6, 7 oder 8 Uhr wahrgenommen).
- Generell Lokalisierungsfehler: Die simulierte Schallquelle wird in einer falschen Richtung wahrgenommen (z.B ein 11 Uhr wird 9 Uhr wahrgenommen). Alle Fehler außer der Im-Kopf-Lokalisierungsfehler und der Vorne-Hinten-Vertauschung werden als generelle Lokalisierungsfehler gezählt.

Die folgende Tabelle 4.1 zählt die insgesamt Fehlerrate (Im-Kopf, Vorne-Hinten-Vertauschung und generelle Lokalisierungsfehler der falschen Richtungen) der vier Ausgaben (Domiso_b, Football_b, Domiso_w, Football_w) je Proband in Prozent auf. Dabei werden die unterschiedlichen Testsignale auch getrennt. Z.B, Domiso_b_noise bedeutet die beste Ausgabe von DOMISO mit rosa Rauschen. Die insgesamten Fehlerraten aller Probanden der besten Ausgaben und der schlechten Ausgaben betragen 27.08% und 38.89%.

Aus diese Tabelle können wir sehen, dass die individuellen Ergebnisse der Probanden eine große Varianz (0%-66.67% Fehlerrate) haben. Ein Grund dafür sind die zu wenigen Testdaten wegen der kleineren Ausmaße (7 Probanden, 6 Richtungen, 2 Wiederholungen). Aufgrund der Zeitdauer des Testes wurden in diesem Hörtest immer nur die 6 Richtungen für jeden Satz und jedes Testsignal einmal getestet. Das bringt die Varianz.

| AOÜF \ Proband | 1 | 2 | 3 | 4 | 5 | 6 | Durchschnitt |
|-----------------------|--------|--------|--------|--------|--------|--------|--------------|
| Domiso_b_noise | 50.00% | 33.33% | 16.67% | 0.00% | 16.67% | 50.00% | 27.78% |
| Domiso_w_noise | 33.33% | 33.33% | 50.00% | 16.67% | 33.33% | 33.33% | 33.33% |
| Domiso_b_speech | 16.67% | 16.67% | 50.00% | 0.00% | 16.67% | 50.00% | 25.00% |
| Domiso_w_speech | 50.00% | 33.33% | 33.33% | 33.33% | 33.33% | 50.00% | 38.89% |
| Football_b_noise | 50.00% | 16.67% | 33.33% | 50.00% | 16.67% | 33.33% | 33.33% |
| Football_w_noise | 16.67% | 50.00% | 66.67% | 50.00% | 66.67% | 66.67% | 52.78% |
| Football_b_speech | 16.67% | 16.67% | 33.33% | 16.67% | 33.33% | 16.67% | 22.22% |
| Football_w_speech | 33.33% | 0.00% | 50.00% | 33.33% | 16.67% | 50.00% | 27.78% |
| Durchschnitt_best | 33.33% | 20.83% | 33.33% | 16.67% | 20.83% | 38.50% | 27.08% |
| Durchschnitt_schlecht | 33.33% | 29.17% | 50.00% | 33.33% | 38.50% | 50.00% | 38.89% |

Tabelle 4.1: Darstellung der individuellen Fehlerrate

Ein anderer Grund ist wahrscheinlich der zufällige Zugang der Datenbank. Die beiden Auswahlverfahren haben einen zufälligen Zugang mit 16 aus 118 ausgewählten Sätzen der AOÜFs. Laut der Literatur [Nico10] [Iway06] [Sato08] sind die räumlichen Unterschiede zwischen den AOÜFs einem Benutzer graduell. Das heißt, dass die Sätze der AOÜFs in einer Datenbank sich gruppieren. Der zufällige Zugang bringt eine unbestimmte Gruppe der AOÜFs, die an der Auswahlphase teilnehmen sollen. Die beste Situation ist, dass die ausgewählten Sätze in der Gruppe große Unterschiede zwischen den Besten und den Schlechtesten haben, wie z.B. Proband 4, der ein exzellentes Ergebnis mit der besten Ausgabe von DOMISO (0% Fehlerrate mit beiden Testsignalen) hat. Das lässt vermuten, dass für ihn diese Ausgabe so gut wie die eigenen AOÜFs ist. Proband 2 und Proband 5 haben auch die Ausgaben, die ihren eigenen AOÜFs ähnlich sind. Aber es kann auch passieren, dass die Sätze in einer Gruppe den Probanden die ähnlichen Räumlichkeiten haben. Die Folgen sind die ähnlichen Fehlerraten der besten und der schlechten Sätze (z.B. Proband 1), eine große Varianz der Fehlerraten und schwierige Aufgaben der Auswahl für die Probanden. Tatsächlich tritt diese Situation sehr häufig auf. Die möglichen Lösungen werden im Ausblick diskutiert.

Trotz der Varianz scheint es, dass die meisten der individualisierten besten Ausgaben der beiden Auswahlverfahren zu den besseren Ergebnissen als die schlechten Ausgaben führen. Nur ein Proband (Proband 1) hat die gleichen Fehlerraten (33.33%). Das weist die Folgerung der vorher durchgeführten Experimente wie subjektive Auswahl durch Seeber und DOMISO nach, dass die Methode, individuell Auswahl nicht individueller AOÜFs, funktioniert.

Durch die durchschnittlichen Daten können wir konvergentere Ergebnisse bekommen. Die Ergebnisse der besten Sätze betragen ca. 27%, während die schlechten Sätze

die größeren Fehlerrate zwischen 30% und 50% haben. Der Vergleich der durchschnittlichen Daten beweist auch die obengenannte Folgerung.

Die Tabelle 4.2 und die Abbildung 4.10, 4.11, 4.12 zeigen die getrennten Fehlerraten der Im-Kopf-Lokalisierungsfehler, der Vorne-Hinten-Vertauschung und der generellen Lokalisierungsfehler. Die schwarzen Balken sind die Ergebnisse der besten Ausgaben, während die weißen Balken die Ergebnisse der schlechten Sätze sind. Daraus wird auch gezeigt, dass die besten Ausgaben zu den besseren Ergebnissen führen. Die Anzahl der generellen Lokalisierungsfehler der besten und schlechten Sätze sind gleich. Der Grund dafür ist wahrscheinlich die Diffusion wegen des intransparenten Kopfhörers. Allerdings erscheinen die Ergebnisse des Im-Kopf-Fehlers und der Vorne-Hinten-Vertauschung unterschiedlich. Die Anzahl des Im-Kopf-Fehlers der besten Sätze sind etwas kleiner als die Anzahl der schlechten, während die Anzahl der Vorne-Hinten-Vertauschung der schlechten Sätze den Anzahl der besten doppelt sind.

| AOÜF \ Fehler | Im-Kopf | Vorne-Hinten | Generell |
|-----------------------|---------|--------------|----------|
| Domiso_b | 9.72% | 6.94% | 9.72% |
| Domiso_w | 11.11% | 15.28% | 9.72% |
| Football_b | 6.94% | 13.89% | 6.94% |
| Football_w | 9.72% | 23.61% | 6.94% |
| Durchschnitt_best | 8.33% | 10.42% | 8.33% |
| Durchschnitt_schlecht | 10.42% | 19.44% | 8.33% |

Tabelle 4.2: Darstellung der Fehlerrate mit den getrennten Typen

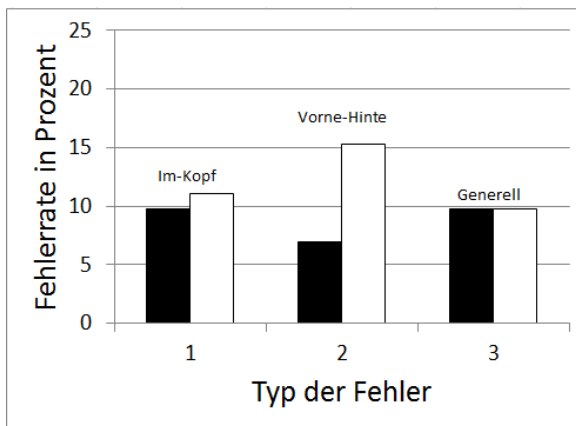


Abbildung 4.10: Vergleich der Fehlerraten verschiedener Typen aller Probanden von DOMISO (schwarze Balken: beste AOÜFs; weiße Balken: schlechte AOÜFs)

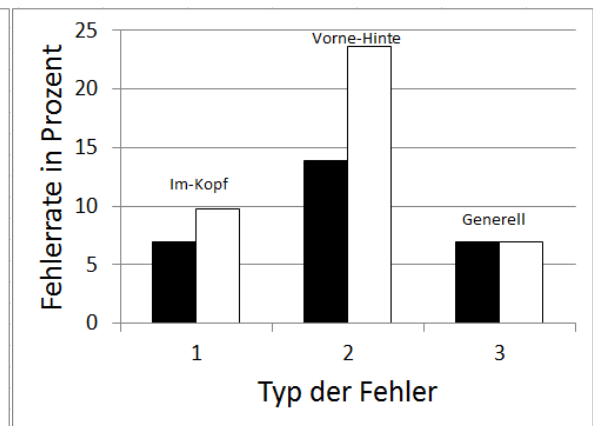


Abbildung 4.11: Vergleich der Fehlerraten verschiedener Typen aller Probanden von FST

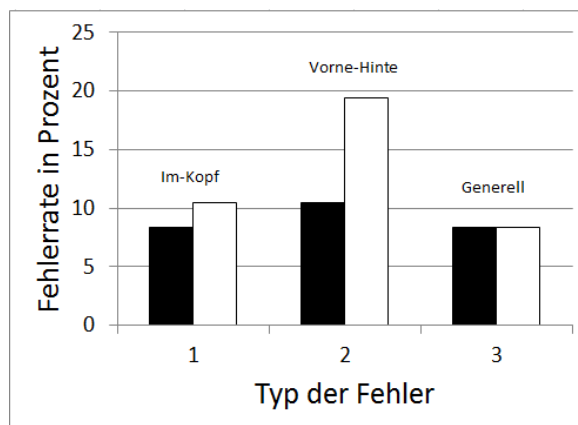


Abbildung 4.12: Vergleich der Fehlerraten verschiedener Typen der beiden Verfahren

Zum Vergleich der 2 Auswahlverfahren wird die folgende Tabelle 4.3 verfasst. Die durchschnittlichen Fehlerraten der beiden besten Ausgaben scheint ähnlich (ca. 27%). Die Fehlerrate bei der besten von DOMISO sieht geringfügig kleiner als die bei der besten von Fußball-Stil-Turnier aus. Allerdings unterscheiden sie sich mit den schlechten Sätzen. Die Differenz der Fehlerraten zwischen den besten und den schlechten sind 9.72% für DOMISO und 12.5% für Fußball-Stil-Turnier.

| AOÜF \ Proband | 1 | 2 | 3 | 4 | 5 | 6 | Durchschnitt |
|----------------|--------|--------|--------|--------|--------|--------|--------------|
| Domiso_best | 33.33% | 25.00% | 33.33% | 0.00% | 16.67% | 50.00% | 26.39% |
| Football_best | 33.33% | 16.67% | 33.33% | 33.33% | 25.00% | 25.00% | 27.78% |
| Domiso_worse | 41.67% | 33.33% | 41.67% | 25.00% | 33.33% | 41.67% | 36.11% |
| Football_worse | 25.00% | 25.00% | 58.33% | 41.67% | 41.67% | 58.33% | 40.28% |

Tabelle 4.3: Darstellung der Fehlerrate verschiedener Verfahren

Eine interessante Folgerung wurde auch hier gefunden. Ein Proband (Proband 4) wurde zum Vergleich der Wirkung der verschiedenen Umgebung mit verschiedener Lautstärke (einmal normal und einmal laut) bei der Wiedergabe getestet. Ansonsten wurden die Bedingungen bzw. Sätze der AOÜFs nicht geändert.

Tabelle 4.4 zeigt den Vergleich der verschiedenen Umgebung. Normal und loud bedeuten die Wiedergabe mit der normalen Umgebung und der lauterer Umgebung. Z.B heißt Domiso_b_normal der beste Satz von DOMISO mit der normalen Umgebung.

| Fehler \ Bedingung | Im-Kopf | Vorne-Hinten | Generell | Insgesamt |
|-----------------------|---------|--------------|----------|-----------|
| Domiso_b_normal | 16.67% | 8.33% | 8.33% | 33.33% |
| Domiso_b_loud | 0.00% | 33.33% | 16.67% | 50.00% |
| Football_b_normal | 16.67% | 16.67% | 0.00% | 33.33% |
| Football_b_loud | 8.33% | 33.33% | 8.33% | 50.00% |
| Durchschnitt_b_normal | 16.67% | 12.50% | 4.16% | 33.33% |
| Durchschnitt_b_loud | 4.17% | 33.33% | 12.50% | 50.00% |
| Domiso_w_normal | 16.67% | 25.00% | 0.00% | 41.67% |
| Domiso_w_loud | 8.33% | 25.00% | 16.67% | 50.00% |
| Football_w_normal | 16.67% | 33.33% | 8.33% | 58.33% |
| Football_w_loud | 8.33% | 25.00% | 8.33% | 41.67% |
| Durchschnitt_w_normal | 16.67% | 29.17% | 4.16% | 50.00% |
| Durchschnitt_w_loud | 8.33% | 25.00% | 12.50% | 45.83% |

Tabelle 4.4: Vergleich der Fehlerrate verschiedener Umgebung für Proband 4

In der Tabelle wird festgestellt, dass die Anzahl des Im-Kopf-Lokalisierungsfehlers bei der lauterer Situation sich reduziert. Aber die Anzahl der Vorne-Hinten-Vertauschung und des generellen Fehlers steigt. Das bedeutet, dass die lautere Wiedergabe zu einer besseren Externalisierung führt, dennoch auch zu einer Verwirrung der Lokalisierung. Das kann auch durch den Vergleich der Anzahl der durchschnittlichen Fehler unterstrichen werden. Außerdem lässt das vermuten, dass die ausgewählten besten Sätze

AOÜFs wie die schlechten Sätze bei der lauterer Umgebung funktionieren (50% Fehlerrate beim besten Satz und 45.83% beim schlechten in der lauterer Umgebung).

Der Grund weshalb Proband 7 die Vorderseite mit der simulierten Umgebung gar nicht lokalisieren konnte, ist wahrscheinlich die strenge Besonderheit der Körperparameter oder des Hörsystems des Probanden. Auch die allgemeinen AOÜFs von dem Kemar-Kunstkopf funktionierten bei ihm gar nicht.

5 Ausblick

Diese Arbeit unterstreicht die durch Seeber und Iwaya entdeckte Folgerung. Aber die Mängel des Auswahl-Tools sind auch offensichtlich. Wie die Probleme zeigen, ist die Varianz der Fehlerrate groß. Die Gründe wurden erläutert.

Es scheint, dass der Zugang des auszuwählenden AOÜFs unbedingt wichtig ist. Eine einfache Lösung für die Verbesserung des Zugangs ist die Erhöhung der Anzahl der auszuwählenden AOÜFs. 16 zu 118 ist einem stabilen Zugang zu wenig. Wenn sich die Anzahl erhöht (z.B. 32 oder 64 aus 118), steigt natürlich die Wahrscheinlichkeit, dass der Proband den passenden Satz der AOÜFs auswählen kann. Aber diese Lösung ist zeitaufwendig und benutzerunfreundlich. Mit 32 Sätzen dauert die Auswahlprozedur 20-30 Minuten, während die Auswahlprozedur von 64 Sätzen über eine Stunde dauert.

Eine effizientere Variante wird durch die sogenannte Clustering-Methode [Shoj94] inspiriert. Wie schon gesagt, sind die Unterschiede der AOÜFs in einer Datenbank graduell. Deswegen können die AOÜFs durch die Korrelation gruppiert werden. Dazu werden die Spatial Frequency Response Surfaces (SFRS) anstatt AOÜFs verwendet [Nico10]. Die SFRSs beschreiben die gleichen Informationen wie AOÜFs. Aber eine SFRS zeigt allen Richtungsinformation einer Frequenz-Bin, während eine AOÜF alle Information im Frequenzbereich einer Richtung beinhaltet. Mit Kreuzkorrelation kann die Ähnlichkeit von 2 SFRSs offenbar gesehen werden. Durch Vergleich der Ähnlichkeit kann die AOÜFs in einer Datenbank geclustert werden. Für jedes Cluster wird ein Prototyp ausgewählt. Durch diese Methode kann der zufällige Zugang der Datenbank aus den Prototypen ausgewählt werden oder kann die Datenbank reduziert werden. Das Problem dazu ist die Komplexität der Berechnung. Darum sollten sich die Forscher mehr bemühen.

6 Fazit

Mit der vorliegenden Arbeit wurde ein Auswahl-Tool durch individuelle Auswahl nicht individueller AOÜFs konzipiert und experimentell nachgewiesen.

Dabei wurden zuerst die theoretischen Grundlagen dieser Arbeit erläutert. Die durch Seeber und Iwaya realisierten Auswahlmethoden motivierten diese Arbeit. Danach wurde das Auswahl-Tool aufgestellt. 3 Entwürfe mit Schnittstellen und Flussdiagramm wurden zur Diskussion gestellt. Einer davon wurde durch diese Arbeit entwickelt (Fußball-Stil-Turnier). Zum Schluss wurde ein endgültiger Plan mit 2 Auswahlverfahren für den Hörtest festgelegt.

Der Hörtest war für den Nachweis der individuellen Auswahl nicht individueller AOÜFs. Der Rahmen, die Vorbereitung und der Ablauf wurden im Kapitel 5 erklärt. Die Ergebnisse wurden detailliert diskutiert. Trotz des kleineren Ausmaßes wurde die Durchführbarkeit der individuellen Auswahl unterstreicht. Der Mangel war auch offensichtlich. Für die Verbesserung der weiteren Forschung wurde eine mögliche Methode am Ende dieser Arbeit vorgestellt.

Die Forschung der AOÜF und des auditorischen Systems wurde seit langer Zeit entwickelt. Aber für die virtuelle Realität wurden immer größere Anforderungen an die höhere Technologie gestellt. Für eine zeitsparende, effiziente und benutzerfreundliche Anwendung, setzt die individuelle Auswahl nicht individueller AOÜFs in den kommenden Jahren einen Trend.

Anhang

A Figuren

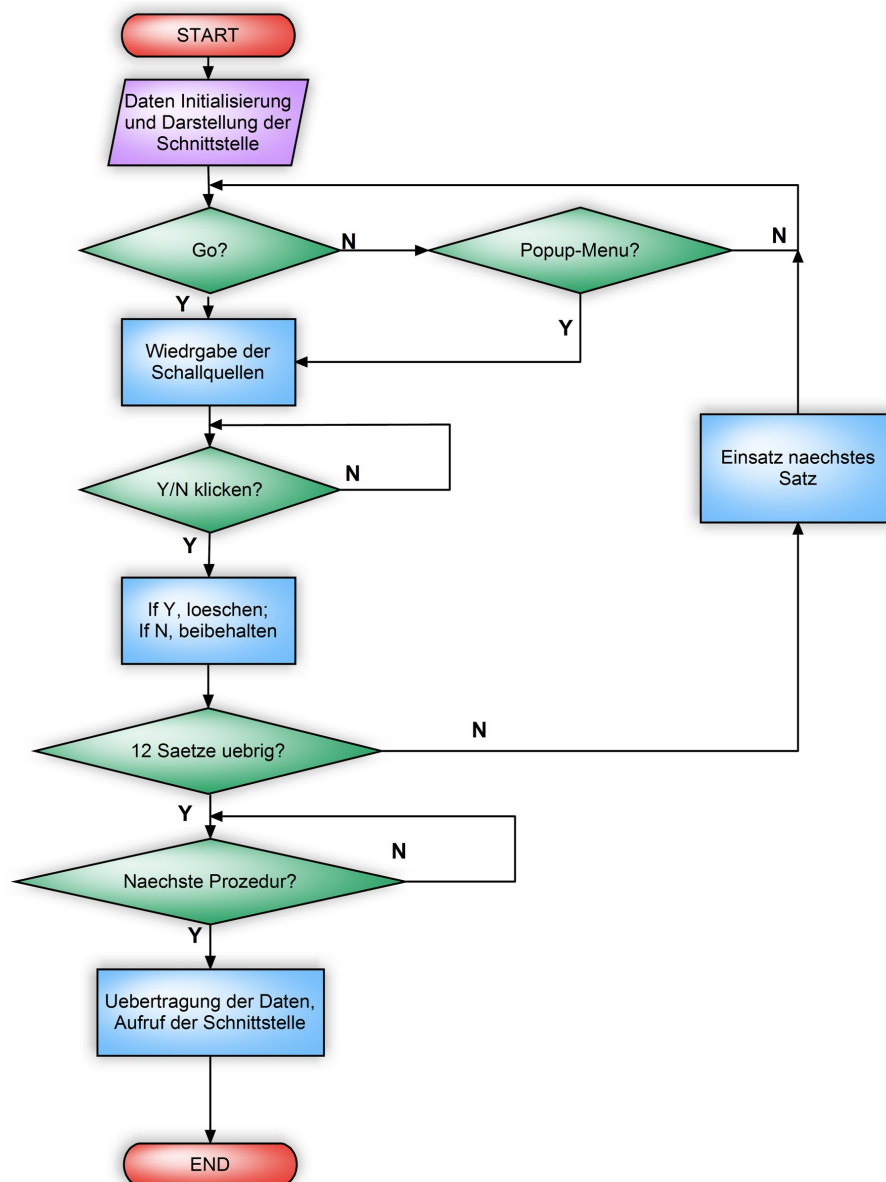


Abbildung A.1: Flussdiagramm der Vorauswahl der subjektiven Auswahl

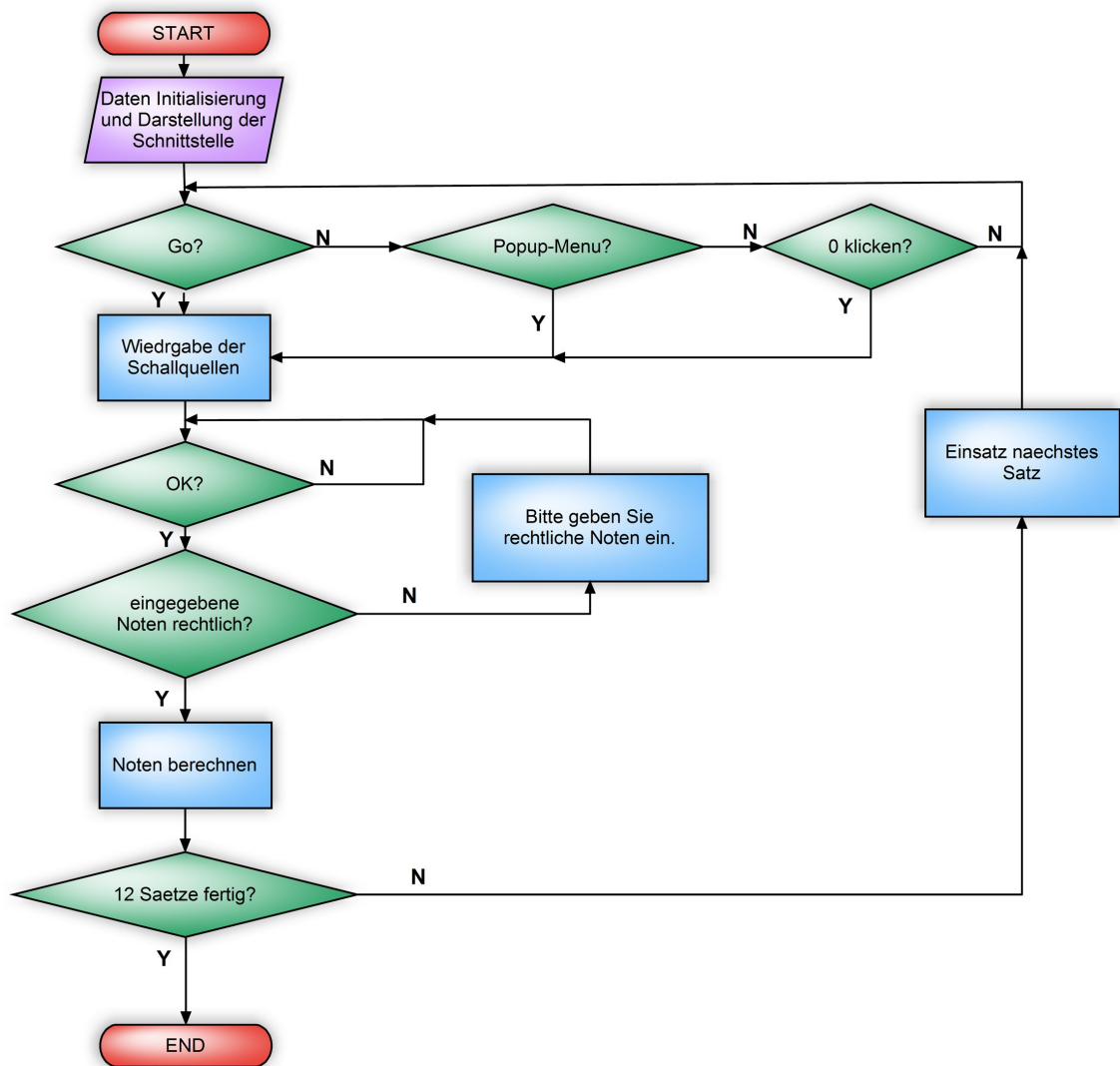


Abbildung A.2: Flussdiagramm der Endauswahl der subjektiven Auswahl

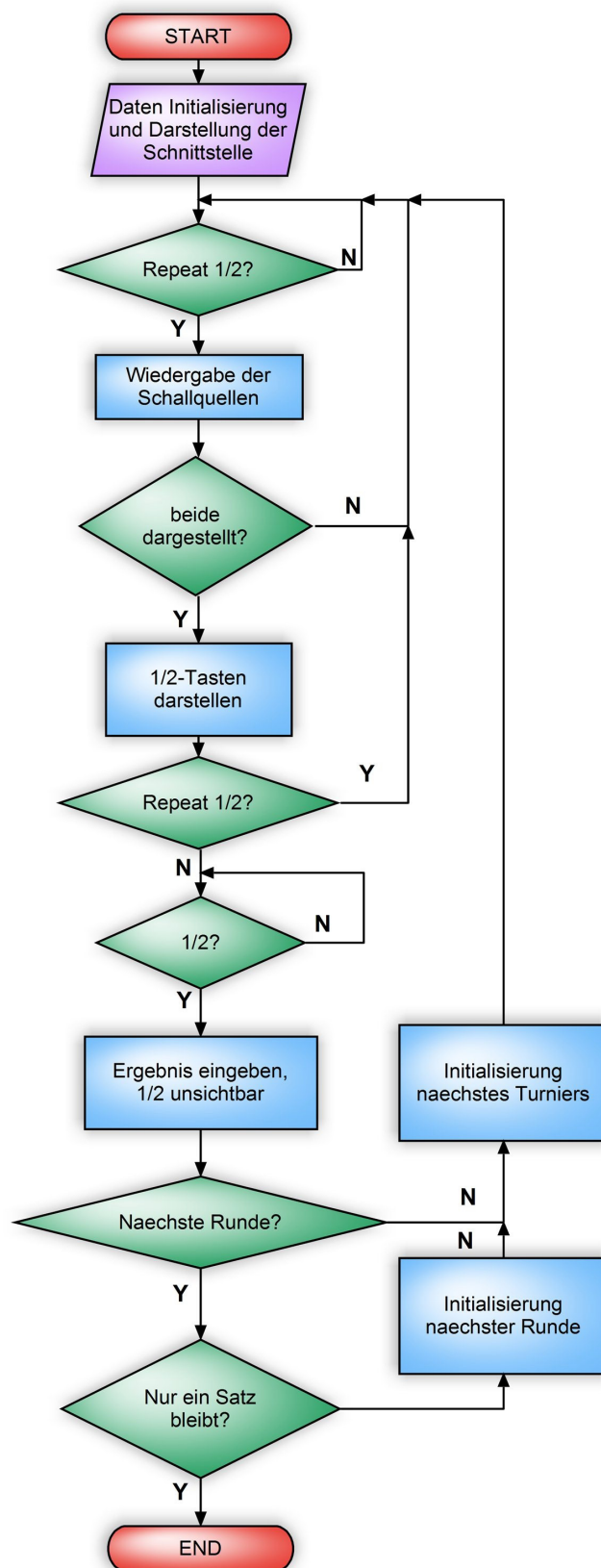


Abbildung A.3: Flussdiagramm von DOMISO

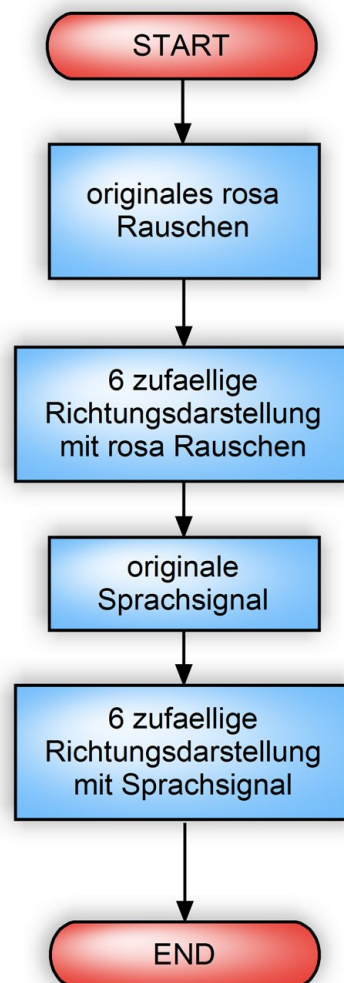


Abbildung A.4: Zustandsdiagramm der Trainingphase

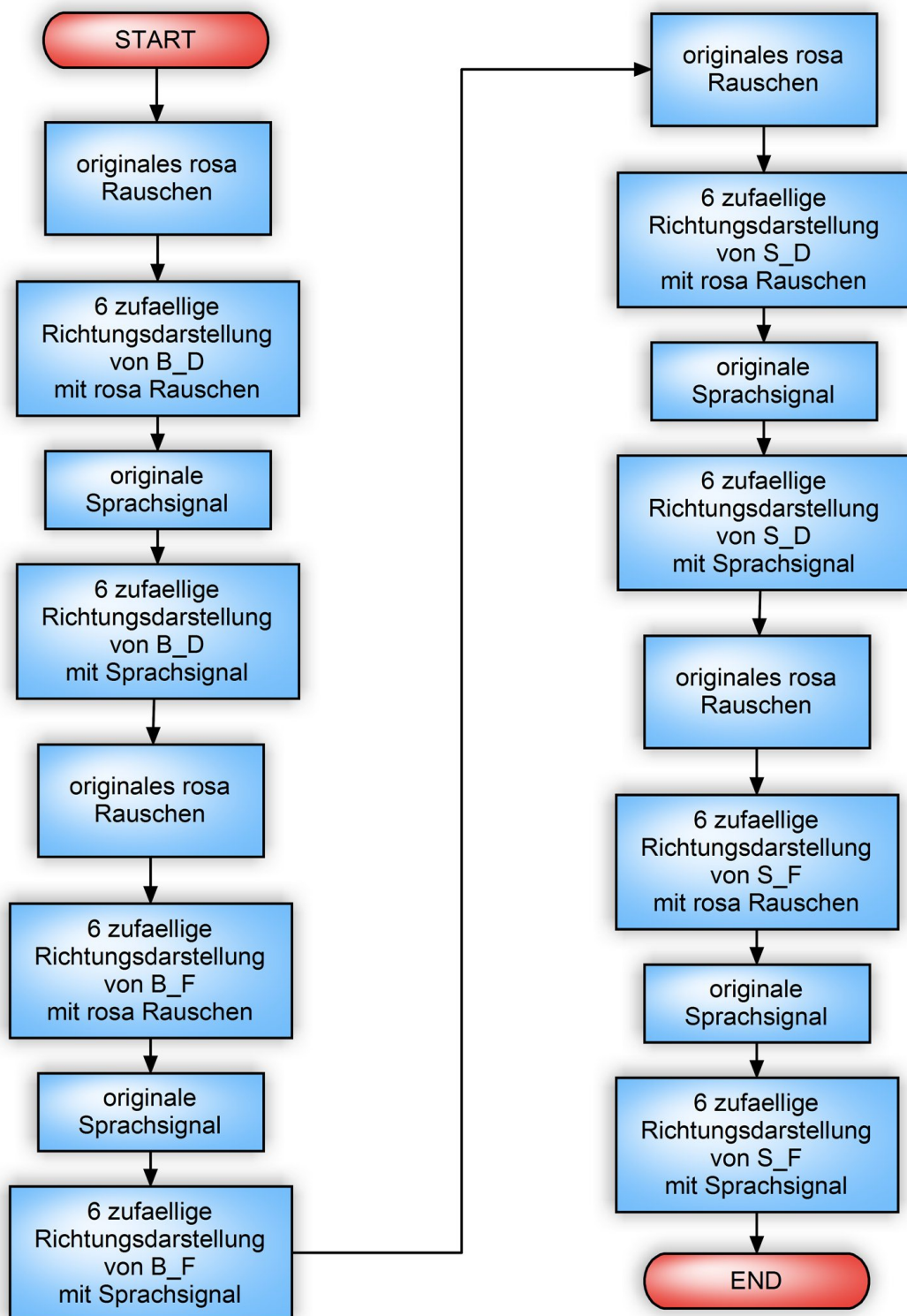


Abbildung A.5: Zustandsdiagramm der Testphase

Literaturverzeichnis

- [Acos] Acoustic Information Systems Laboratory, Tohoku University, <http://www.ais.riec.tohoku.ac.jp/lab/db-hrtf/index.html>. *Database of HRTFs*. last update on May 26, 2001.
- [Akio] Akio Honda, Hiroshi Shibata, Jiro Gyoba, Kouji Saitou, Yukio Iwaya, Yôiti Suzuki. *Transfer Effects on Sound Localization Performances from Playing a Virtual Three-Dimensional Auditory Game*. *Applied Acoustics*, Volume 68, Issue 8, August 2007, Pages 885-896.
- [Bern03] Bernhard Seeber, Hugo Fastl. *Subjective selection of non-individual Head-Related Transfer Functions*. Proc. ICAD, Boston, MA, USA, July 6-9. Script of 2003.
- [Bran11] Prof. Dr.-Ing. Karlheinz Brandenburg. *Advanced Psychoacoustics, Lecture 3*. Institut für Medientechnik, TU Ilmenau, SS 2011.
- [Bruc06] Bruce N. Walker, Jeffrey Lindsay. *Navigation Performance with a Virtual Auditory Display: Effects of Beacon Sound, Capture Radius, and Practice*. *Human Factors* 48(2), 265-278. 2006.
- [Chri] Christopher Baker, <http://christopherbaker.net/projects/27-may-2003/>. *Binaural Recording*. last update on May 27, 2003.
- [CIPI] CIPIC Interface Laboratory, University of California, <http://interface.cipic.ucdavis.edu/sound/hrtf.html>. *CIPIC HRTF Database, published by Prof. V. Ralph Algazi*. last updated February 25, 2011.
- [Dere06] Derek Brock , Brian Mcclimens , Anthony Hornof , Tim Halverson. *Cognitive Models of the Effect of Audio Cueing on Attentional Shifts in a Complex Multimodal Dual-Display Dual-Task*. Meeting of the Cognitive Science Society, Vancouver. 2006.

- [Dmit03] Dmitry N. Zotkin, Ramani Duraiswami, Larry S. Davis. *HRTF Personalization Using Anthropometric Measurements*. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2003 pp. 157-160, 2003.
- [Eliz88] Elizabeth Wenzel, Frederic Wightman, Doris Kistler, Scott Foster. *Acoustic Origins of Individual Differences in Sound Localization Behavior*. Journal of the Acoustic Society of America 84 (Suppl.1S79), 1988.
- [Fred93] Frederic Wightman, Doris Kistler. *Multidimensional Scaling Analysis of Head-Related Transfer Functions*. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 17-20 October, 1993.
- [Hear] Hearing Review, http://www.hearingreview.com/issues/articles/2011-04_02.asp. *HRTF Measurement*. last update on April 2011.
- [Home] Home Theater Systems, <http://www.athometheatersystems.com/>. *High technique home theater*. last check on April 10, 2012.
- [Impa] Impact Commentaries, <http://jnnp.bmj.com/content/70/1/43.full>. *Horizontal plane of the dummy head*. last check on April 04, 2012.
- [Insta] Institut de Recherche et Coordination Acoustique/Musique, <http://recherche.ircam.fr/equipes/salles/listen/>. *LISTEN HRTF Database*. last update on May 25, 2003.
- [Instb] Institut für Schallforschung, Wien, <http://www.kfs.oeaw.ac.at/content/view/608/606/>. *ARI HRTF Database*. last update on February 2012.
- [Iway06] Yukio Iwaya. *Individualization of head-related transfer functions with tournament-style listening test: Listening with other's ears*. Acoust. Sci. & Tech. June 27. Script of 2006.
- [Jack98] Jack Loomis, Reginald Golledge, Roberta Klatzky. *Navigation System for the Blind: Auditory Display Modes and Guidance*. Teleoperators and Virtual Environments 7(2), 193-203. 1998.
- [LORE] LORENGEI, <http://www.lorengei.com/castellano/Binaural-Sound.html>. *Kemar Dummy Head*. last check on April 04, 2012.
- [MIL-99] MIL-STD-1472F. *Human Engineering Design Criteria for Military Systems Equipment and Facilities*, pp. 42-53. 1999.

- [Møll] Henrik Møller. *Fundamentals of binaural technology*. Applied Acoustics, Volume 36, Issues 3-4, 1992, Pages 171-218.
- [Nasa] Nasa, http://humansystems.arc.nasa.gov/groups/ACD/projects/dynamic_info.php. *Cones of Confusion*. last update on December 14, 2005.
- [Nico06] Alexis Bondu, Sylvain Busson, Vincent Lemaire, Rozenn Nicol. *Looking for a Relevant Similarity Criterion for HRTF Clustering: A Comparative Study*. In Proc. of 120th AES Convention, 2006.
- [Nico10] Rozenn Nicol. *Binaural Technology*. AES Monograph. Audio Engineering Society Inc. 2010.
- [Paul00] Paul Runkle, Anastasia Yendiki, Gregory H. Wakefield. *Active Sensory Tuning for Immersive Spatialized Audio*. Proceeding of International Conference on Auditory Display, Atlanta, GA, 2000.
- [Pern03] Jean-Marie Pernaux. *Spatialisation du son par les techniques binaurales: Application aux services de telecommunications*. PhD thesis, Institut National Polytechnique de Grenoble, 2003.
- [Phon] Phone Recording Archive, <http://www.payphone-directory.org/sounds.html>. *Sprachsignal*. last check on April 04, 2012.
- [Ralp99] Ralph Algazi, Richard O. Duda, Dennis M. Thompson. *THE CIPIC HRTF DATABASE*. CIPIC, U.C. Davis, Davis, CA 95616, USA, 1999.
- [Ralp02] Ralph Algazi, Richard O. Duda, Dennis M. Thompson. *Use of Head-and-Torso Methods for Improved Spatial sound Synthesis*. Proceeding of AES 113th Convention, Los Angeles, CA, 2002.
- [Rich93] Richard O. Duda. *Modeling Head Related Transfer Functions*. Proceeding of Twenty-Seventh Annual Asilomar Conference on Signals, Systems and Computers, Asilomar CA, USA 1993, vol. 2, pp. 996-1000, 1993.
- [Sato08] Satoshi Yairi, Yukio Iwaya and Yoiti Suzuki. *Individualization Feature Of Head-Related Transfer Functions Based On Subjective Evaluation*. The 14th International Conference on Auditory Display, Paris, France, June 24-27. Script of 2008.

- [Seeb03] Bernhard Seeber. *Untersuchung der auditiven Lokalisation mit einer Lichtzeigermethode*. Doktor-Ingenieurs Dissertation, 2003.
- [Shoj94] Shoji Shimada, Nobuo Hayashi, Shinji Hayashi. *A clustering methods for sound localization transfer functions*. J. Audio. Eng. Soc. 42: 577-584, 1994.
- [Silz02] Andreas Silzle. *Selection and Tuning of HRTFs*. AES 112th Convention, Munich, Germany, May 10-13, pp 1-14, 2002.
- [Simo02] Simon Holland, David R. Morse, Henrik Gedenryd. *Audio GPS: Spatial Audio Navigation with a Minimal Attention Interface*. Personal and Ubiquitous Computing 6(4), 253-259. 2002.
- [Song07] Song Xu, Zhizhong Li, und Gavriel Salvendy. *Individualization of Head-Related Transfer Function for Three-Dimensional Virtual Auditory Display*. A Review, ICVR'07 Proceedings of the 2nd international conference on Virtual reality. 2007.
- [Taka99] Takanori Nishino, Shoji Kajita, Kazuya Takeda and Fumitada Itakura. *INTERPOLATING HEAD RELATED TRANSFER FUNCTIONS IN THE MEDIAN PLANE*. Proc. 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, New York, October 17-20. 1999.
- [Wiki] Wikipedia, the free encyclopedia, http://en.wikipedia.org/wiki/File:Hrir_binaural_synthesis.png. *Binaurale Synthese*. last update on February 13, 2005.
- [Wiki12] Wikipedia, the free encyclopedia, <http://de.wikipedia.org/wiki/1/f-Rauschen>. *Pink Noise*, Feb. 2012. last modified on February 01, 2012.

Abbildungsverzeichnis

| | | |
|-----|--|----|
| 1.1 | Heimkino [Home] | 1 |
| 2.1 | Klangräumlichkeit [Bran11] | 5 |
| 2.2 | Cones of Confusion mit ITD und ILD [Nasa] | 6 |
| 2.3 | Kemar Kunstkopf [LORE] | 7 |
| 2.4 | Binaural Recording am Menschen [Chri] | 7 |
| 2.5 | Binaurale Synthese durch ein Paar AOÜFs (\tilde{h}_L und \tilde{h}_R) aus einer Datenbank [Wiki] | 8 |
| 2.6 | Schema von DOMISO [Iway06] | 15 |
| 3.1 | Messung der AOÜFs [Hear] | 19 |
| 3.2 | Horizontalebene der Laufbahn [Impa] | 21 |
| 3.3 | Winkelauflösung [Impa] | 22 |
| 3.4 | inverse Kopfhörerübertragungsfunktion des linken Kanals | 23 |
| 3.5 | inverse Kopfhörerübertragungsfunktion des rechten Kanals | 23 |
| 3.6 | Schnittstelle der Vorauswahl der subjektiven Auswahl | 25 |
| 3.7 | Schnittstelle der Endauswahl der subjektiven Auswahl | 26 |
| 3.8 | Schnittstelle von DOMISO | 27 |
| 3.9 | Schnittstelle des Fußball-Stil-Turniers | 28 |
| 4.1 | Zufälliger Zugang der 6 Richtungen, 12 Uhr bedeutet die Vorderseite des Zuhörers. Die schwarzen Kreise sind die gegebenen Richtungen. . . | 33 |
| 4.2 | Domiso_b_noise (36 Datensätze), $\rho = 0.9879$ | 35 |
| 4.3 | Domiso_w_noise (36 Datensätze), $\rho = 0.9812$ | 35 |
| 4.4 | Domiso_b_speech (36 Datensätze), $\rho = 0.9946$ | 36 |
| 4.5 | Domiso_w_speech (36 Datensätze), $\rho = 0.9905$ | 36 |
| 4.6 | Football_b_noise (36 Datensätze), $\rho = 0.9834$ | 36 |
| 4.7 | Football_w_noise (36 Datensätze), $\rho = 0.9653$ | 36 |

| | | |
|------|---|----|
| 4.8 | Football_b_speech (36 Datensätze), $\rho = 0.9932$ | 36 |
| 4.9 | Football_w_speech (36 Datensätze), $\rho = 0.9878$ | 36 |
| 4.10 | Vergleich der Fehlerraten verschiedener Typen aller Probanden von DO- MISO (schwarze Balken: beste AOÜFs; weiße Balken: schlechte AOÜFs) | 40 |
| 4.11 | Vergleich der Fehlerraten verschiedener Typen aller Probanden von FST | 40 |
| 4.12 | Vergleich der Fehlerraten verschiedener Typen der beiden Verfahren . . | 40 |
| A.1 | Flussdiagramm der Vorauswahl der subjektiven Auswahl | 46 |
| A.2 | Flussdiagramm der Endauswahl der subjektiven Auswahl | 47 |
| A.3 | Flussdiagramm von DOMISO | 48 |
| A.4 | Zustandsdiagramm der Trainingphase | 49 |
| A.5 | Zustandsdiagramm der Testphase | 50 |

Tabellenverzeichnis

| | | |
|-----|---|----|
| 2.1 | Beispiele der objektiven Kriterien | 16 |
| 3.1 | Vergleich der Datenbanken | 20 |
| 3.2 | Vergleich der Methoden | 29 |
| 4.1 | Darstellung der individuellen Fehlerrate | 38 |
| 4.2 | Darstellung der Fehlerrate mit den getrennten Typen | 39 |
| 4.3 | Darstellung der Fehlerrate verschiedener Verfahren | 41 |
| 4.4 | Vergleich der Fehlerrate verschiedener Umgebung für Proband 4 | 41 |

Thesen zur Masterarbeit

1. Für ein Virtual Auditory Display ist die individuelle Auswahl nicht individueller Außenohrübertragungsfunktionen geeignet.
2. Für eine Datenbank der AOÜFs stehen die Parameter Winkelbereich und Winkelauflösung auf dem ersten Platz.
3. Für eine effiziente Auswahl sollte ein turnierorientiertes Verfahren verwendet werden.
4. Kopfhörerentzerrung ist bei einer intransparenten Wiedergabe durch den Kopfhörer unbedingt wichtig. Aber trotz der Entzerrung hat der Kopfhörer Einfluss auf die spektralen Merkmale der Wiedergabe.
5. Die subjektive Bewertung durch einen Hörtest, ist für ein anwendungsorientiertes System tauglich.
6. Ein Hörversuch kleineren Ausmaßes ist zum Vermeiden der Varianz der Datensätze unmöglich. Für die genauen Ergebnisse eines Lokalisierungstestes soll ein Hörversuch großen Ausmaßes durchgeführt werden.
7. Zum Vergleich der Signale mit normaler Lautstärke, führen die lauten Signale zu einer besseren Externalisierung, aber auch zu einer Verwirrung der Lokalisierung.
8. Zur Verbesserung der Effizienz der individuellen Auswahl nicht individueller AOÜFs kann die Anzahl der AOÜFs in einer Datenbank durch Clustering-Methode reduziert werden.

Ilmenau, den 11. 04. 2012

Hui Fei

Erklärung

Die vorliegende Arbeit habe ich selbstständig ohne Benutzung anderer als der angegebenen Quellen angefertigt. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten Quellen entnommen wurden, sind als solche kenntlich gemacht. Die Arbeit ist in gleicher oder ähnlicher Form oder auszugsweise im Rahmen einer oder anderer Prüfungen noch nicht vorgelegt worden.

Ilmenau, den 11.04.2012

Hui Fei