

„Signalkorrelation und Raumeindruck bei Stereo- und
5.1 Surround-Aufnahmen“

Dissertation

Universität für Musik und darstellende Kunst, Graz

IEM – Institut für Elektronische Musik und Akustik

Edwin Pfanzagl-Cardone

September 2010

Abstract

Ausgehend von bereits durch den Dissertanten durchgeführten Messungen an 2-Kanal Stereo-Mikrofonssystemen, wird die Problemstellung auf die empirische Untersuchung von unterschiedlichen 5.1 Surround-Mikrofonierungstechniken erweitert. Ziel der Arbeit ist es, auf zwei voneinander unabhängigen Wegen vier verschiedene 5.1 Surround-Mikrofonensysteme (OCT, DECCA, KFM, AB-PC) dahingehend zu bewerten, wie gut sie in der Lage sind, den im Aufnahmeraum auftretenden Klangeindruck in einer Surround-Abhörsituation zu reproduzieren.

1) Im Rahmen von A/B Hörvergleichen wurden die oben genannten 5.1 Mikrofonierungstechniken von Probanden hinsichtlich verschiedener Aspekte (z.B. Räumlichkeit, Lokalisation, Klangfarbe, etc.) sowie der Hörerpräferenz bewertet. Die Daten wurden mittels Varianzanalyse, Clusteranalyse, Korrelationsanalyse und Faktorenanalyse ausgewertet. Statistisch signifikante Unterschiede ergaben sich vor allem für das KFM-System, welches am schlechtesten beurteilt wurde.

2) Parallel dazu wurde auf rein messtechnischem Weg ein Vergleich der frequenzabhängigen interauralen Kreuz-Korrelation (FIACC) sowie der frequenzabhängigen Kohärenz von Kunstkopfaufnahmen aus dem Aufnahmeraum mit jener der Kunstkopfaufnahme aus dem Wiedergaberaum durchgeführt. Um für die Surround-Aufnahmen noch detailliertere Information zu erhalten, wurde auch die frequenzabhängige Kreuz-Korrelation (FCC) und frequenzabhängige Kohärenz von benachbarten Kanälen ermittelt.

Zwei Mikrofonensysteme fielen hierbei besonders auf:

A) das KFM-System wies bei der Korrelationsanalyse in breiten Frequenzabschnitten hohe Gegenphasigkeit sowohl bei der Kunstkopfaufnahme, als auch bei manchen Kanal-Paarungen des Surround-Signals auf;

B) mit dem AB-PC-System konnte die Korrelationsfunktion der Kunstkopfaufnahme aus dem Konzertsaal bei 5.1 Wiedergabe am ‚Sweet spot‘ am genauesten nachgebildet werden, was dafür spricht, dass Aufnahmen mit dieser Mikrofontchnik das Originalschallfeld des Konzertsaals am besten approximieren.

Theoretische Überlegungen zur Bedeutung der Wichtigkeit der Signaldekorrelation vor allem bei tiefen Frequenzen für den in einer Aufnahme enthaltenen Raumeindruck wurden anhand der Ergebnisse des Hörvergleichs und der messtechnischen Auswertung des Datenmaterials verifiziert. Der Vergleich der Messergebnisse mit den Resultaten der Hörerbeurteilung zeigte insofern gute Übereinstimmung, als z.B. messtechnisch auffällige Anomalien des KFM-Systems (Gegenphasigkeit) mit einer im Hörvergleich festgestellten relevanten ‚Klangfärbung‘ und mangelnden ‚Natürlichkeit‘ in Zusammenhang gebracht werden können. Weiters decken sich die Ergebnisse sowohl des Hörvergleichs, als auch der messtechnischen Auswertung gut mit einer Reihe anderer Veröffentlichungen, was für ihre Plausibilität spricht.

Basierend auf den Ergebnissen der vorliegenden Untersuchung wurde der Vorschlag für die Einführung eines ‚Binauralen Qualitäts-Index für reproduzierte Musik‘ (BQI_{rep}) gemacht, der aus dem in Bereich der Konzertsaal-Akustik bekannten ‚Binaural Quality Index‘ abgeleitet wurde.

Im Rahmen der vorliegenden Arbeit wurde auch die Qualität einer vom Dissertanten speziell für 5.1 Surround entwickelten Mikrofonierungsmethode (AB-PC: ‚AB-Polycardioid Centerfill‘) bewertet, welche auf dem Ansatz beruht, das in einem Aufnahmerraum entstehende Schallfeld möglichst ‚dekorreliert‘ zu erfassen. Aufgrund der vorliegenden Ergebnisse darf diese auf gleicher qualitativer Höhe mit bereits arrivierten Systemen wie OCT und DECCA eingestuft werden.

Außerdem wurde eine weitere Mikrofontechnik vorgeschlagen, welche für die Aufnahme kleiner Schallquellen in stereo und surround optimiert ist, und ebenfalls einen ‚dekorrelierten‘ Ansatz verfolgt: das BPT-System (‚Blumlein-Pfanzagl-Triple‘).

Das Ergebnis der vorliegenden Untersuchung hat im Prinzip große Relevanz für den Klassik-Bereich der Audiobranche, liefert es doch einen eindeutigen Fingerzeig, wie natürlicher klingende Aufnahmen zu erzielen wären.

Abstract

Based on results from a previous analysis of correlation coefficients of stereo-recordings obtained from the most common stereo-microphone techniques an empirical study was undertaken concerning a number of popular microphone techniques that are currently used for 5.1 surround.

It is the aim of this research project to compare four different 5.1 microphone techniques (OCT, DECCA, KFM, AB-PC) in respect to their ability to reproduce the sound field present in the recording space as faithfully as possible in the listening room:

1) Through A/B comparison tests a number of listeners had to grade several parameters of the surround-recordings (e.g. spatial impression, localization, sound-coloration, etc.), as well as listener preference. The data were subject to analysis of variance, cluster analysis, correlation analysis and principal component analysis. Statistically significant differences were obtained mainly for the KFM-system, which was rated worst.

2) Parallel to this, a comparison of the frequency-dependent inter-aural cross-correlation (FIACC) as well as frequency dependent coherence of artificial head recordings from the original venue were compared with the artificial head recordings made in the ‘sweet spot’ of the listening room. In order to obtain more detailed information for the surround-recordings the frequency dependent cross-correlation (FCC) and frequency dependent coherence of neighboring channels were analyzed as well.

Two systems stood out among the others:

A) the KFM-microphone exhibited broadband out-of-phase signal components in the artificial head recordings, as well as for some of the surround-channel pairings;
B) the AB-PC-system was able to replicate the correlation-function of the original artificial head recording most accurately, which is also an indicator that this microphone technique is best in recreating the original sound field of the recording venue.

Theoretical considerations concerning the importance of signal de-correlation in relation to spatial impression were verified through analysis of results of the listening test as well as the acoustic measurements. The comparison of the measurements with the results of the subjective listening test coincided well, as the KFM’s out-of-phase

anomalies found a counterpart in relevant ‘sound-coloration’ and lack of ‘naturalness’, as evaluated by the test listeners.

Also, both the results of the subjective listening tests, as well as the measurements go well with the findings of a number of publications, which is an indication for their plausibility.

Based on the results of this research, a ‘Binaural Quality Index for reproduced music’ (BQI_{rep}) has been proposed, which has been derived from the ‘Binaural Quality Index’, which is well known in the field of concert hall acoustics.

As part of the thesis the performance of a new 5.1 surround microphone technique (AB-PC: ‘AB-Polycardioid Centerfill’), developed by the researcher, has been evaluated. This microphone-technique, which is based on a ‘de-correlated’ recording approach, proved to be on a par with well established systems such as OCT and DECCA.

In addition, the BPT-system (‘Blumlein-Pfanzagl-Triple’) has been proposed, mainly aimed at the recording of small sound-sources in stereo or surround, which also follows a ‘de-correlated’ recording approach.

The results of this research project should be of high interest to the classical segment of the recording industry, as they are able to provide a clear indication on how to achieve more naturally sounding recordings.

| | |
|---|-------|
| Inhaltsverzeichnis | I-VI |
| Abstract | 1-4 |
| Vorwort | 5-9 |
| 1 Einleitung | 10-15 |
| 1.1 Angaben zur Problemstellung und kurze Einführung in das Thema | 10 |
| 1.2 Überblick über den Aufbau der Arbeit | 13 |
| 1.3 Ziel der Arbeit | 15 |
| 2 Räumliches Hören | 16-51 |
| 2.1 Mechanismen der Lokalisation | 17 |
| HRTF-Phasencharakteristik | 25 |
| Lokalisation mit HRTFs | 28 |
| 2.2 Mechanismen der Distanzwahrnehmung | 29 |
| 2.2.1 Schallintensität | 29 |
| 2.2.2 Diffusschallanteil | 30 |
| 2.2.3 Frequenzgang | 30 |
| 2.2.4 Binaurale Differenzen | 31 |
| 2.3 Raumeindruck | 31 |
| 2.4 Physikalische Größen im Zusammenhang mit dem Raumeindruck | 35 |
| 2.5 Einflüsse der Eigenschaften von Lautsprecher und Akustik des Abhörspaces auf die Hörerpräferenz | 41 |
| 2.6 Psychoakustische Effekte bei Signalwiedergabe über Lautsprecher betreffend Lokalisation und Räumlichkeit | 44 |
| 2.6.1 Frequenzabhängige Lokalisationsverzerrungen in der horizontalen Ebene | 47 |
| 2.6.2 Frequenzabhängige Lokalisationsverzerrungen in der vertikalen Ebene | 48 |
| 2.6.3 Effekte betreffend die Wiedergabe des Räumlichkeits- eindrucks | 48 |
| 3 Korrelation und Kohärenz | 52-68 |
| 3.1 Signalkorrelation bei Stereo-Mikrofonsystemen | 53 |

| | |
|--|--------|
| 3.2 Optimale Signal-Kohärenz bei Surround-Mikrofonsystemen | 61 |
| 3.3 Optimaler interauraler Korrelationsgrad | 64 |
| 3.4 Gedanken zur Wechselwirkung zwischen Lautsprecher-Richtwirkung und Hörer-Envelopment | 65 |
| 4 Stereo-Mikrofontechniken | 69-75 |
| 4.1 Koinzidente Techniken | 69 |
| 4.1.1 XY mit Nieren und Hypernieren | 69 |
| 4.1.2 XY mit Achten (Blumlein-Paar) | 70 |
| 4.1.3 MS-Stereo | 71 |
| 4.1.4 Koinzidente Omnidirektionale Mikrofone | 71 |
| 4.2 Äquivalenz-Stereophonie | 71 |
| 4.2.1 ORTF | 72 |
| 4.2.2 NOS | 72 |
| 4.2.3 Faulkner ‚Phased Array‘ | 72 |
| 4.2.4 OSS – Jecklin Disk | 73 |
| 4.3 AB-Systeme (Laufzeit Stereophonie) | 73 |
| AB-Polycardioid Centerfill (AB-PC) | 75 |
| 5 Surround-Mikrofontechniken | 76-110 |
| 5.1 Koinzidente und ‚quasi-koinzidente‘ Arrays | 76 |
| 5.1.1 Soundfield-Mikrofon | 76 |
| 5.1.2 DMS (Double-MS Technik) | 77 |
| 5.1.3 Double-MS + Kunstkopf | 78 |
| 5.1.4 BPT (Blumlein-Pfanzagl-Triple) | 79 |
| 5.1.5 Martin-Triple | 83 |
| 5.1.6 XY-Nierenanordnung (Back-to-Back) | 85 |
| 5.1.7 MILAB 360 (360° Surround-Mikrofon) | 85 |
| 5.1.8 Compact 4-3-4 Phasor Array | 85 |
| 5.2 Spaced Arrays (2-dimensionale Arrays) | 87 |
| 5.2.1 Circular Microphone Array | 87 |
| 5.2.2 Trinnov-Array | 88 |
| 5.2.3 ABC 5 (Omni Array) | 89 |
| 5.2.4 INA 5 (Ideale Nieren Anordnung) | 90 |
| 5.2.5 OCT-Surround (Optimal Cardioid Triangle – Surround) | 91 |

| | | |
|--------|---|---------|
| 5.2.6 | OCT V.2 | 92 |
| 5.2.7 | Klepko-Array | 92 |
| 5.2.8 | Atmo-Kreuz | 93 |
| 5.2.9 | Hamasaki-Square | 94 |
| 5.2.10 | Multiple-AB (,Mikrofonvorhang') | 94 |
| | Tree-Anordnungen | 96 |
| 5.2.11 | Fukada-Tree | 96 |
| 5.2.12 | Div. Tree-Anordnungen (z.B. Polyhymnia-Pentagon, etc.) | 97,98 |
| 5.2.13 | AB-PC (AB-Polycardioid Centerfill) | 99 |
| 5.3 | Trennkörper und ,3D'-Techniken | 100 |
| 5.3.1 | KFM 360 (Kugelflächenmikrofon) | 101 |
| 5.3.2 | Diverse Trennkörper-Surroundmikrofonsysteme | 101 |
| 5.3.3 | Holophone ,H2-Pro' und ,PortaMic 5.1' | 102 |
| 5.3.4 | Sony Kugelanordnung | 103 |
| 5.3.5 | Multichannel Microphone Array | 103 |
| 5.3.6 | Eigenmike | 103 |
| 5.3.7 | OSIS-System (Optimal Sound Image Space) | 104 |
| 5.3.8 | Pan-Ambiophonic 2D/3D System | 105 |
| 5.4 | Allgemeines Betrachtungen über Mikrofon-Arrays | 107 |
| 5.5 | Zusammenfassung | 108 |
| 6 | Kunstkopfaufnahmen | 111-118 |
| 6.1 | Verwendung eines Kunstkopfs als ,Bezugsnormal' | 111 |
| 6.2 | Studien betreffend die Qualität verschiedener Kunstköpfe | 112 |
| 6.3 | Zusammenfassung und Schlussfolgerung zu den Kunstkopf-Studien | 117 |
| 7 | Hörvergleich | 119-136 |
| 7.1 | Grundlegendes zur Problematik von Hörvergleichen | 119 |
| 7.2 | Anforderungen an die Wiedergabeanordnung | 120 |
| 7.3 | Einflussfaktoren auf die Hörerbewertung | 122 |
| 7.4 | Auswahl der zu bewertenden Klangattribute | 124 |
| 7.5 | Art des Hörvergleichs | 131 |
| 7.5.1 | Bewertungs-Skala | 132 |

| | |
|---|---------|
| 7.5.2 ‚Kalibrierung‘ der Testhörer | 132 |
| 7.5.3 Vorbemerkungen zur Anwendbarkeit der Varianzanalyse | 132 |
| 8 Auswertung des Hörvergleichs ‚ORCH 5.1‘ (Lautsprecher) | 137-192 |
| 8.1 Durchführung der Varianzanalyse | 137 |
| 8.1.1 Signifikante Unterschiede bei den Mittelwerten der 4 Mikrofontechniken | 140 |
| 8.1.2 Reliabilität und Validität der Hörerbewertungen | 144 |
| 8.1.3 ‚Spezialfall 6‘ | 145 |
| 8.1.4 ‚Spezialfall 9‘ | 147 |
| 8.2 Präferenz-Abschätzung durch Mittelwertberechnung (gestrichen) | 150 |
| 8.3 Datenauswertung mittels Hauptfaktorenanalyse (Principal Component Analysis – PCA) | 152 |
| 8.3.1 Anzahl der zu berücksichtigenden Faktoren | 157 |
| 8.3.2 Interpretation der Faktoren | 157 |
| 8.3.3 Grenzen der Faktorenanalyse | 161 |
| 8.4 Clusteranalyse | 163 |
| Grenzen der Clusteranalyse | 167 |
| 8.5 Korrelationsanalyse | 167 |
| 8.6 Analytische Betrachtung der Signale der Rear-Kanäle LS, RS | 171 |
| 8.7 Zusammenfassung und Vergleich mit den Ergebnissen anderer Untersuchungen | 178 |
| 8.7.1 Varianzanalyse | 178 |
| 8.7.2 Faktorenanalyse | 178 |
| 8.7.3 Clusteranalyse | 185 |
| 8.7.4 Korrelationsanalyse | 189 |
| 8.7.5 Schlussfolgerung | 190 |
| 8.8 Qualitative Reihung der Surround-Mikrofonarrays – ein Vergleich mit anderen Untersuchungen | 191 |
| 9 Hörvergleich Orchesteraufnahme-Kunstkopf ORCH-KK (Kopfhörer) | 193-209 |
| 9.1 Programm-Material | 193 |
| 9.2 Kopfhörer-Wiedergabe | 194 |

| | |
|--|---------|
| 9.3 Ermittlung der Transfer-Funktion ‚Kunstkopf + Kopfhörer‘ | 195 |
| 9.4 Fragebogen und Ablauf des Hörversuchs | 201 |
| 9.5 Analyse: Hörvergleich Orchesteraufnahme unter Kopfhörerwiedergabe | 205 |
| | |
| 10 Untersuchung der Korrelation und Kohärenz bei Surround-Signalen | 210-257 |
| 10.1 Auswahl der Surround-Mikrofonsysteme und des Aufnahmewinkels | 210 |
| 10.2 Mikrofon- und Kunstkopfpositionen | 214 |
| 10.3 Programm-Material | 216 |
| Orchesteraufnahme, Duo Aufnahme | 216 |
| 10.4 Messung der frequenzabhängigen Kreuz-Korrelation (FCC) und frequenzabhängigen interauralen Kreuz-Korrelation (FIACC) | 218 |
| 10.5 Messung der frequenzabhängigen Kohärenz | 226 |
| 10.6 Paarweise Korrelation und Kohärenz | 229 |
| 10.7 Frequenzabhängige Kreuz-Korrelation (FCC) bei Stereo- Hauptmikrofonsignalen | 242 |
| Zusammenfassung | 251 |
| 10.8 Binaural Quality Index für reproduzierte Musik (BQI_{rep}) | 251 |
| 10.9 Zusammenfassung | 256 |
| 11 Zusammenfassung, Diskussion, Schlußfolgerung und Ausblick | 258-265 |
| 11.1 Einleitung | 258 |
| 11.2 Zusammenfassung und Diskussion | 259 |
| 11.3 Schlußfolgerung | 264 |
| 11.4 Ausblick | 265 |
| | |
| Literaturverzeichnis | 266-288 |
| Diskographie, Videoquellennachweis | 289 |

ANHANG

| | |
|--|---------|
| Abb. A1: Graphik Schallabschattung am Kopf (aus [Steinberg and Snow, 1934]) | 287 |
| Abb. A2: MW+Konfidenzintervall-Darstellung Gruppe IEM | 291 |
| Abb. A3: MW+Konfidenzintervall-Darstellung Gruppe SF | 291 |
| Tabelle AT1: Vergleich der p-Werte für 15 Aspekte für die Hörergruppen ‚SF‘ und ‚IEM‘ | 292 |
| Abb. A4: Matrixplot der Hörerbewertung (Aspekte 1-8) | 293 |
| Abb. A5: Matrixplot der Hörerbewertung (Aspekte 9-15, zzgl. Präferenz) | 294 |
| Ergebnisse des Surround-Hörvergleich des ORF (2001) | 295-299 |
| Dokumentation zur raumakustischen Messung des IEM Produktionsstudios | 300-306 |
| Tab. AT2 und AT3: frequenzabhängiger Hallradius Regieraum IEM und SF | 307 |
| Abb. A24: Dokumentationsphotos Regieraum IEM (Hörtest mit Proband) | 308 |
| Abb. A25: Zuspelplatz (Pro-Tools) im IEM-Regieraum | 308 |
| Abb. A26: Grundriss IEM-Regieraum | 309 |
| Abb. A27: Nachhallzeit IEM-Regieraum gemäß ITU-R BS.1116-1 | 310 |
| Abb. A28: Fragebogen Hörvergleich ‚ORCH 5.1‘ (Lautsprecher) | 311 |
| Abb. A29: Fragebogen Hörvergleich ‚ORCH-KK‘ (Kopfhörer) | 312 |
| Abb. A30: Fragebogen Hörvergleich ‚DUO 5.1‘ (Lautsprecher) | 313 |
| Referenz Email 1 zu OCT-Array (Cornelis von der Gragt) | 314 |
| Referenz Email 2 zu OCT-Array (Jan Korte) | 315 |
| Dokumentation zum Hörvergleich ‚DUO 5.1‘ (Lautsprecher) | 316-321 |

Vorwort

Gegen Ende meiner Tätigkeit als Lehrbeauftragter am Institut für Elektroakustik der Universität für Musik und darstellende Kunst, Wien (1994-1999) konfrontierte mich eines Tages ein Student damit, er hätte in Unterlagen, die im Rahmen der Tonmeisterausbildung an der Hochschule der Künste in Berlin verwendet würden gelesen, dass nur ‚Klein-AB‘ (mit einem Kapselabstand von ca. 50cm) den ‚richtigen‘ Klangeindruck (vor allem in Hinblick auf Lokalisation) liefern würde, und sich mit wachsendem Kapselabstand eigentlich ein immer kleiner werdender (korrekt abgebildeter) Aufnahmewinkel ergäbe. Dies schien dem ‚gesunden Hausverstand‘ zu widersprechen, denn wie sollte dann Groß-AB (mit einem Kapselabstand von bis zu einigen Metern) klanglich überhaupt ‚funktionieren‘ ?

Ich erklärte mich damals gerne bereit, die Unterlagen – welche aus dem Skriptum von Eberhard Sengpiel, Tonmeister der ehemaligen Plattenfirma Teldec und Lehrbeauftragter an der HdK Berlin, stammten – näher anzusehen, um fachlich Stellung beziehen zu können.

Und tatsächlich: in diesen wurde – mithilfe einer (gezwungenermaßen) starken Vereinfachung der akustischen Grundannahmen – recht plausibel und u.a. auf Basis trigonometrischer Überlegungen die auftretenden Signallaufzeit-Differenzen zwischen dem linken und rechten Ohr eines Hörers berechnet und aufgrund dessen Aussagen über die Korrektheit im Lokalisationseindruck verschiedener Mikrofonssysteme getroffen.

Die Ergebnisse dieser Überlegungen standen jedoch – wie bereits oben angedeutet – in krassem Gegensatz zur Arbeitspraxis vieler Tonmeister, die sehr erfolgreich ‚Groß-AB‘ Systeme – oftmals auch in Verbindung mit anderen Hauptmikrofonverfahren - bei ihren Aufnahmen einsetzen. (siehe hierzu [Betz, 2000], [La Grou, 1994]).

Erfolgreich deswegen, weil es u.a. so prominente historische Beispiele gibt, wie die ersten Stereo-Aufnahmen der RCA in den frühen 1950-er Jahren, die damals unter dem ‚Red Seal‘-Label veröffentlicht wurden und alle (auch) eine Groß-AB Kapselanordnung - zumindest als sog. ‚Outriggers‘ - verwenden (siehe [Valin, 1994]). Diese Aufnahmen werden von Musikliebhabern weltweit geschätzt und deshalb Mitte der 1990-er Jahre erneut als CDs herausgebracht; um das Jahr 2003 wurden dann sogar die originalen 3-

Kanal Masterbänder im audiophilen ‚Multichannel SA-CD‘ Format veröffentlicht.

Kurzum – hier gab es eine Diskrepanz zwischen Theorie und Praxis, die einer Überprüfung bedurfte: wenn einzig ‚Klein-AB‘ die ‚richtige‘ Abbildung lieferte, dann mußte ‚Groß-AB‘ ja gezwungenermaßen ‚falsch‘ sein – und umgekehrt.

Aufgrund meiner Fixanstellung als Tonmeister für das Salzburger Festspielhaus befand ich mich ab Herbst 2000 in der glücklichen Lage, nicht nur einen – nach objektiven Kriterien beurteilt (siehe [Hidaka et al., 1995]) - sehr gut klingenden Konzertsaal praktisch ständig zur Verfügung zu haben, sondern auch die erforderlichen (authentischen) Schallquellen in Form verschiedener Orchester, welche es mir ermöglichten, unter gleichbleibenden Bedingungen empirische Untersuchungen anstellen zu können: ich wollte die oben aufgeworfene Frage ganz einfach für mich persönlich beantworten.

Testaufnahmen mit verschiedenen Stereo-Hauptmikrofontechniken erfolgten im Laufe der Jahre 2001 und 2002, deren Auswertung in der Veröffentlichung erster Ergebnisse im Rahmen eines Papers für die 112. AES-Convention in München ihren Niederschlag fand. [Pfanzagl-Cardone, 2002]

Verbunden mit dem Studium einschlägiger Literatur ergab sich bald ein etwas differenzierteres Bild, welches nun nicht mehr die einfachen Schlussfolgerungen von Sengpiel zuließ: seine Überlegungen sind im Grunde natürlich richtig, allerdings nur dann haltbar, wenn die Grundargumentation auf stark vereinfachten Sachverhalten aufbauen darf, die so in der Praxis eigentlich nicht haltbar sind, da die Psychoakustik des menschlichen Hörens doch viel zu komplex ist, um sich durch einfache trigonometrische Berechnungen approximieren zu lassen. (siehe hierzu [Benjamin and Brown, 2007]) Diese Widersprüche (bzw. die mitunter auftretende ‚Uneindeutigkeit‘ bei der Lokalisationsgenauigkeit von Klein-AB Systemen) wurden interessanter Weise zur gleichen Zeit von namhaften Akademikern – allerdings mit anderen Schlussfolgerungen - thematisiert (siehe hierzu [Wittek und Theile, 2002]).

Im Rahmen meines AES-Papers für die 112.Convention war auch ein auf Groß-AB basierendes Hauptmikrofonsystem („AB-Polycardioid Centerfill“; abgekürzt: AB-PC), geeignet für Stereo- wie auch Surroundaufnahmen, vorgeschlagen worden, dessen klangliche Eigenschaften nun denen in der Branche bereits etablierter Surround-Mikrofontechniken gegenübergestellt werden sollte. Aus diesem Grunde führte ich im Frühling des Jahres 2003 eine Testaufnahme unter Verwendung von insgesamt 5 Surround-Mikrofonsystemen (OCT, DECCA, KFM, ABPC, Soundfield) im Großen Festspielhaus Salzburg durch.

Durch die Zulassung zum Doktoratsstudium an der KUG mit Wintersemester 2003 ergab sich auch die Möglichkeit, im Laufe der folgenden Jahre diese Aufnahmen nicht nur einer messtechnisch genaueren Auswertung, sondern auch einer umfangreichen Hörerbewertung durch geeignete Probanden zuzuführen. Somit konnten den subjektiven Eindrücken von Testhörern die Ergebnisse objektiver Messungen gegenübergestellt werden, die genauere Rückschlüsse über den entstandenen Höreindruck zuließen.

Die Ergebnisse der Untersuchung liegen nun im Rahmen dieser Dissertation vor und wären – wie auch im Abstract bereits thematisiert – von großer Relevanz für die zu erzielende Qualität bei 2-Kanal Stereo, wie auch Surroundaufnahmen (5.1, 7.1, etc.), da – neben der ‚Natürlichkeit‘ - der in einer Aufnahme enthaltene Räumlichkeitseindruck einen großen Einfluss auf die Hörerpräferenz hat. (siehe hierzu [Berg and Rumsey, 2001], [Gabrielsson and Lindström, 1985], sowie [Toole, 1985])

Im Rahmen der Dissertation werden zwei neue Mikrofonierungstechniken vorgeschlagen und untersucht, die für die ‚dekorrelierte‘ Aufnahme großer Schallquellen (AB-PC), wie auch kleiner Schallquellen („Blumlein-Pfanzagl-Triple“ – BPT) geeignet, und sowohl für 2-Kanal-Stereo Anwendungen, wie auch Surround gedacht sind. Aufgrund der vorliegenden Hörerbeurteilung dürfen beide Techniken als mit bereits anerkannten Systemen (wie z.B. OCT) qualitativ vergleichbar eingestuft werden.

Ein aktueller Zwischenstand der Analyseergebnisse konnte im Rahmen eines Papers auf der 124. AES-Convention in Amsterdam präsentiert werden. [Pfanzagl-Cardone und Höldrich, 2008]

Die in der vorliegenden Dissertation zusammengefassten Ergebnisse sind das Resultat einer kritischen Auseinandersetzung mit dieser Materie, die sich – bereits geraume Zeit vor Beginn des Doktoratsstudiums - bis in das Jahr 1998 zurückerstreckt und somit einen Zeitraum von über 10 Jahren umfasst. Als Legitimation der langen Dauer bis zur Fertigstellung darf eventuell auch die Tatsache ins Treffen geführt werden, dass die Arbeit an dieser Dissertation zeitgleich zu meiner beruflichen Vollanstellung erfolgte.

An dieser Stelle möchte ich mich sehr herzlich bei folgenden Personen und Institutionen bedanken, die mich beim Entstehen dieser Arbeit unterstützt haben:

IEM – Institut für Elektronische Musik und Akustik:

- für die Zur-Verfügung- Stellung von Material und Räumlichkeiten,
- o. Prof. Dr. tech. Robert Höldrich – für die Betreuung dieser Dissertation,
Erstbegutachter
- Prof. Dr. Gerhard Eckel – Zweitbegutachter
- Dr. Alois Sontacchi – organisatorische Unterstützung (Hörtests), sowie raumakustische Vermessung des Produktionsstudios (Regieraum)
- DI Markus Noisternig – Unterstützung bei der Orchesteraufnahme, sowie div. Gedankenaustausch und Beratung

Folgenden Firmen/Personen gebührt ebenfalls Dank:

- Prof. DI(FH) Karlheinz Müller (Fa. Müller-BBM, München), für die leihweise Überlassung des im Rahmen der Dissertation verwendeten Neumann KU81i Kunstkopfes
- DI Jörg Wuttke (Fa. Schoeps), für das zur Verfügung gestellte Kugelflächenmikrofon (KFM), DSP-4 Prozessor und die OCT-Surround Spezialhalterung.
- der Fa. Soundfield, England für das zur Verfügung gestellte MKV-Mikrofon + Prozessor

- der Leitung der Salzburger Festspiele, welche die Durchführung der Testaufnahmen genehmigt hat, dem beteiligten Orchester (Warschauer Kammerphilharmonie) und Musikern des Duos Klavier-Cello (Francesca Cardone, Stefan Grove), sowie

Tontechnikern, die bei den Aufnahmen unterstützend mitgewirkt haben: Günther Harner, Ilja Bayerl,

- Mag. Markus Waldner, für wiederholte Unterstützung bei der Literatursuche, und

- den Studierenden am IEM, sowie div. Professionisten aus der Audio-Branche, welche als Testpersonen für die Hörvergleiche zur Verfügung gestanden haben.

1 Einleitung

1.1 Angaben zur Problemstellung und kurze Einführung in das Thema

Der ursprüngliche Anstoß für Untersuchungen zu diesem Themenkreis begann mit einer simplen Fragestellung: wenn die Mikrofontechnik ‚Klein-AB‘ (zwei Mikrofone mit kugelförmiger Richtcharakteristik [Druckempfänger] mit Kapselabstand unter einem Meter) - im Gegensatz zu ‚Groß-AB‘ - aufgrund der Gesetze der Psychoakustik eine korrekte Lokalisation der Schallquelle ermöglicht, und Groß-AB hingegen das inhärente Problem des sogenannten ‚Mittenlochs‘ aufweist, warum kommt es dann so häufig vor, dass Tonmeister oft ‚Groß-AB‘ basierte Techniken (sogenannte ‚Outriggers‘) in Ergänzung zu den von ihnen favorisierten Hauptmikrofonsystemen einsetzen ... ? (siehe hierzu [La Grou, 1994], pg.32, [Valin, 1994], pg. 123/124, [Betz, 2000])

Das bereits im Vorwort angesprochene Paper von [Wittek und Theile, 2002] brachte allerdings zu Tage, dass die Lokalisation von Klangquellen einer Musikaufnahme unter Anwendung von ‚Klein-AB‘ Mikrofontechniken - basierend auf psychoakustischen Gesetzmäßigkeiten – mitunter stark unterschiedlich ausfallen kann, da hierbei Abweichungen von teilweise über 100 % gemessen wurden (vgl. hierzu die Angaben der Aufnahmewinkel für ‚Klein-AB‘ [omnis 50cm] bei Sengpiel und Wittek)

| Setup | Hugonnet | Sengpiel | Williams | Wittek |
|--|----------|----------|----------|--------|
| AB omnis 50cm | 130° | 180° | 100° | 74° |
| AB omnis 100cm | - | 62° | - | 36° |
| ORTF cardioids 110°, 17cm | 90° | 96° | 100° | 102° |

Tab. 1.1: Aufnahmewinkel von 3 verschiedenen Mikrofonsystemen, wie in der Literatur angegeben (nach [Wittek und Theile, 2002])

Einige Tonmeister - darunter z.B. auch Eberhard Sengpiel von der ehemaligen Teldec (siehe [Sengpiel, 1992]) - sind davon überzeugt, dass bei Anwendung der AB-Technik

von zwei Mikrofonen mit Kugelcharakteristik nur ‚Klein-AB‘ eine korrekte Lokalisation ermöglicht, ignorieren dabei jedoch völlig, dass hierbei aufgrund physikalischer Gesetzmäßigkeiten das Stereo-Signal zu tiefen Frequenzen hin immer stärker monophon wird, da der Korrelationsgrad eines AB-Paares zu den tiefen Frequenzen hin ansteigt. (siehe dazu Abb.1.1)

Dieser Sachverhalt wurde bereits in der zweiten Hälfte der 1950-er Jahre erstmals bei [Cook et al., 1957] theoretisch wie messtechnisch untersucht; die führenden Plattenfirmen der damaligen Zeit hatten sich bereits einige Jahre zuvor - zumeist in der ersten Hälfte der 1950-er Jahre - praktisch mit der neuen Problematik der Stereo-Tonaufnahmetechnik auseinandergesetzt (siehe hierzu z.B. [Valin, 1994]).

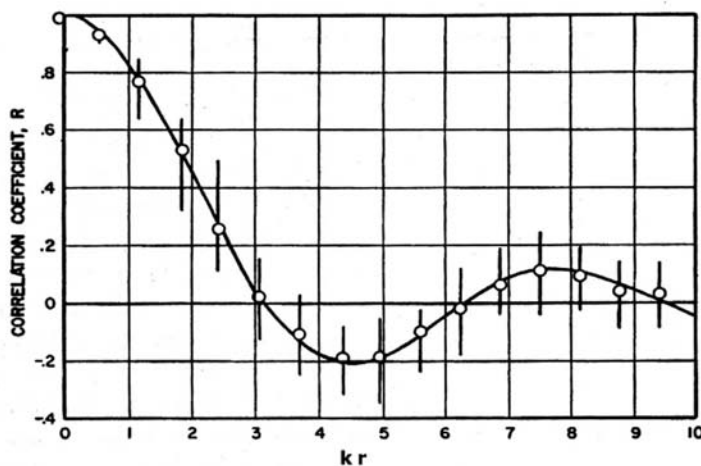


Abb. 1.1:
Erste Berechnung (Kurve) und Messung (o) der Korrelation der Signale zweier Druckempfänger im diffusen Schallfeld (mit $k r = 2 \pi r / \lambda$)

(aus [Cook et al., 1957])

Dass die Stärke der Signal-Korrelation in Abhängigkeit von Frequenz und Kapselabstand größte Relevanz für das klangliche Endresultat hat, war den Audio-Engineers also bereits von mehr als 50 Jahren durchaus bewusst. Auch nach der Einführung des damals noch jungen Stereo-Tonformats beschäftigten sich praktisch arbeitende Tonmeister weiterhin mit dieser Problematik: so gibt es beispielsweise von Carl Ceoen zwei Veröffentlichungen, in denen u. a. auch der Korrelationsgrad der Signale von Stereo-Mikrofonverfahren untersucht wird (siehe [Ceoen, 1972] und [Ceoen, 1974]).

Bei Blauert finden sich in mehreren Kapiteln seines Standardwerks ‚Räumliches Hören‘ [Blauert, 1974] Erörterungen zur Problematik von kohärent und inkohärent abstrahlenden Schallquellen, und er widmete sich auch in speziellen Veröffentlichungen ähnlich gelagerten Themen (siehe z.B. [Blauert, 1976]).

Konsequenter Weise beschäftigen sich mit dieser Problematik - im Zusammenhang mit Surround-Mikrofonierungstechniken – auch in jüngerer Zeit verschiedene Fachleute (siehe hierzu u.a. [Theile, 2000], [Martin, 2005]).

An dieser Stelle soll angemerkt werden, dass heutzutage viele Musikaufnahmen mit ‚Industrie-Standardklangbild‘, welche oftmals eine Vielzahl von Stützmikrofonen verwenden, dem Hörer in erster Linie nur im hohen Frequenzbereich eindeutige Lokalisations-Reize liefern. Wie in einer bereits vor Beginn der Dissertation veröffentlichten Studie [Pfanzagl-Cardone, 2002] gezeigt werden konnte, sind die Signale vieler der häufig verwendeten Stereo-Hauptmikrofone im Bereich tiefer Frequenzen stark korreliert, was in diesem Frequenzbereich auch keine eindeutige Lokalisation mehr ermöglicht.

Warum sich viele Stereo-Hauptmikrofontechniken, die über hohe Korrelation im tiefen Frequenzbereich verfügen, sich so großer Beliebtheit (zumindest in Tonmeister-Kreisen) erfreuen, liegt möglicherweise daran, dass zu Zeiten der Vinylschallplatte gegenphasige Signalinformation im Bassbereich aus im Bereich des Platten-Umschnitts gelegenen technischen Gründen vermieden werden mussten.

(Anm.: Gegenphasige Information bei tiefen Frequenzen mit hohen Pegeln führt entweder zum Abreißen des ‚Schneidspans‘ oder später, beim Abspielen der gepressten Patte auf einer billigeren Heimanlage, zum Herausspringen der Plattennadel aus der Rille)

Ein weiterer Grund mag in der - zumindest bis in die 1980-er Jahren geltenden - Forderung des Rundfunks nach möglichst hoher ‚Mono-Kompatibilität‘ des auszustrahlenden Stereosignals bestehen. Unter anderem ist auch deswegen die XY-Mikrofontechnik mit zwei Nierenmikrofonen und Öffnungswinkel von 90° weit verbreitet. Bedingt durch die Einführung der CD gelten die oben genannten technischen Einschränkungen - betreffend mechanische Tonträger - nicht mehr.

Das Gehör des Testhörers greift die im hohen Frequenzbereich angebotenen Lokalisationsreize dankbar auf und ist somit in der Regel - mangels der Möglichkeit eines AB-Vergleichs mit anderen Aufnahmetechniken, die auch im Bassbereich eine höhere Signaldekoration aufweisen - zufrieden mit dem angebotenen Signal. Ein selektives Abhören von einzelnen Frequenzbändern eines Stereo- Hauptmikrofonsignals (siehe [Griesinger, 1997], pg.3), kann hier akustisch sehr interessante Resultate bringen

(eine Möglichkeit, über die allerdings ein durchschnittlicher Hörer zuhause normalerweise nicht verfügt, erfordert diese doch Filter, die nur im Bereich der professionellen Studioteknik oder im Messlabor zu finden sind.).

1.2 Überblick über den Aufbau der Arbeit

Die Arbeit gliedert sich im Wesentlichen in folgenden Teile:

Im Rahmen der Einleitung wird eine kurze Einführung in den Themenkomplex Korrelation von Mikrofonsignalen gegeben. Neben Angaben zur Problemstellung im Allgemeinen und Hinweisen zur inhaltlichen Abgrenzung, werden die Hauptziele der Arbeit definiert.

Der Hauptteil besteht aus Grundlagen-Kapiteln zu den Themenbereichen Psychoakustik des räumlichen Hörens, Korrelation und Kohärenz, sowie Stereo- und Surround-Mikrofonssysteme.

Im Rahmen der Arbeit wurden mehrere Hörvergleiche durchgeführt: neben Gedanken zur Problematik von Hörvergleichen werden die verschiedenen klanglichen Aspekte, deren Bewertung Gegenstand der Untersuchung war, erläutert und der Fragebogen vorgestellt. Die Methodik des Hörvergleichs wird erklärt; die Daten der Hörerbefragung nach statistischen Methoden ausgewertet (Varianzanalyse, Clusteranalyse, Korrelationsanalyse), sowie einer Faktorenanalyse unterworfen.

Hierbei werden die Ergebnisse der Hörerbefragung verschiedener Test-Aufnahmen untersucht:

- Surround-Aufnahme eines Orchester (Lautsprecherwiedergabe),
- im Regieraum bei Wiedergabe verschiedener Surround-Techniken neu entstandene Kunstkopfaufnahmen werden mit der Original-Kunstkopfaufnahme des Orchesters aus dem Konzertsaal verglichen (Kopfhörerwiedergabe)
- Surround-Aufnahme eines Instrumental-Duos (Cello-Klavier)

Im zweiten, messtechnisch orientierten Teil werden die frequenzabhängige Korrelation und Kohärenz verschiedener Surround-Mikrofontechniken sowie deren binauraler Entsprechungen (Kunstkopfaufnahme) untersucht. Darüber hinaus wird für eine zweite, davon unabhängige Gruppe von Stereo-Aufnahmen die frequenzabhängige Korrelation sowie ebenso von deren binauraler Entsprechung (Kunstkopf-Simulation) ermittelt und diskutiert.

Abschließend werden die Ergebnisse der statistischen Auswertung der Hörerbefragungen der messtechnischen Analyse gegenübergestellt und diskutiert.

1.3 Ziel der Arbeit

Ziel der Arbeit ist es, auf zwei voneinander unabhängigen Wegen verschiedene 5.1 Surround-Mikrofonsysteme dahingehend zu bewerten, wie gut sie im Stande sind, den im Konzertsaal auftretenden Klangeindruck in einer ‚Surround‘-Abhörsituation (im Studio oder Wohnzimmer) zu reproduzieren. Das System, welches dazu am besten in der Lage ist, sollte dementsprechend als das am natürlichsten klingende aus der Studie hervorgehen.

Vermutlich ist die objektivste Methode zur Ermittlung, welche Mikrofontchnik ein akustisches Ereignis am natürlichsten erfasst, folgende:

1. Das Klangereignis (z.B. Orchesterkonzert) wird mit dem/den zu untersuchenden Mikrofonsystem/en aufgenommen.
2. Zur gleichen Zeit wird mittels Kunstkopf an einer möglichst guten Hörposition im Konzertsaal eine zweite Aufnahme gemacht.
3. Während die erste Aufnahme auf einer Stereo- oder Surround-Anlage abgespielt wird, fertigt man mittels einem an der optimalen Abhörposition des ‚sweet spot‘ aufgestellten Kunstkopf eine dritte Aufnahme an (‚Re-Recording‘).
4. Anschließend erfolgt die messtechnische Auswertung (Ermittlung der frequenzabhängigen Kreuz-Korrelation) der verschiedenen Mikrofonsysteme und ein Vergleich mit der Korrelation der Ohrensinnale der im Konzertsaal gemachten Original-Kunstkopfaufnahme, um feststellen zu können, wie sehr die bei der Aufnahme des Re-Recordings aufgetretene Signal-Korrelation jener der Original-Aufnahme ähnelt.
5. Parallel dazu werden im Rahmen von A/B Hörvergleichen die 5.1 Surround-Aufnahmen der ausgewählten Mikrofonierungstechniken von Testhörern hinsichtlich verschiedener Parameter (z.B. Räumlichkeit, Lokalisation, Klangfarbe, etc.) bewertet. In diesem Zusammenhang kann außerdem die Frage untersucht werden, ob – aufgrund der Beeinflussung durch Hörgewohnheiten (Stichwort: ‚Industrie-Standardklangbild‘) – die natürlichste Klangreproduktion auch als die ‚am schönsten‘ klingende (Stichwort: Hörer-Präferenz) beurteilt wird.
6. Abschließend wird ein Vergleich der Messergebnisse mit den Resultaten der Hörerbeurteilung vorgenommen.

2 Räumliches Hören

Nachfolgend sollen die Grundlagen des räumlichen Hörens kurz in ihren Grundzügen dargestellt werden. Dabei erfolgt eine Beschränkung auf die physikalische Akustik und Psychoakustik, da diese beiden Bereiche hinsichtlich der Art der vorliegenden Untersuchung relevant sind. Auf neurophysiologische Erkenntnisse oder Theorien (siehe z.B. Theile's ‚Assoziationsmodell‘ zu den Mechanismen der Summenlokalisierung [Theile, 1978]) soll hier nicht näher eingegangen werden, da diese nicht zum Gegenstand der Untersuchung zählen.

Ebenso soll auf eine Darstellung der historischen Entwicklung des Wissensstandes zum menschlichen Hören weitgehend verzichtet werden, sondern in erster Linie eine Kurzfassung des aktuellen Wissenstandes gegeben werden.

Große Teile des Kapitels stammen aus folgenden Quellen: Übersetzung eines Auszugs eines Papers von [Merimaa und Pulkki, 2005], mit teilweisen Ergänzungen aus [Rumsey, 2001] und [Toole, 2008], sowie der Übernahme von Textteilen aus [Fellner und Höldrich 1998a, 1998b].

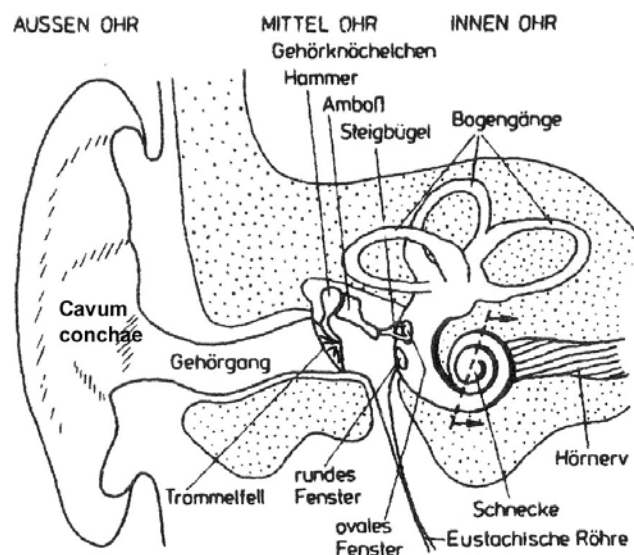


Abb. 2.1 Anatomie des menschlichen Ohres (nach [Zollner und Zwicker, 2003])

Der Schall wird zunächst durch die Pinna (das Außenohr), sowie nahe Körperteile (Schultern und Kopf) beeinflusst. Danach folgen die Effekte des Gehörgangs (Meatus),

der zum Mittelohr führt. Dieses besteht aus dem Trommelfell und den Gehörknöchelchen. Die mechanische Energie wird durch Bewegung am ovalen Fenster in Flüssigkeitsdruck im Innenohr (Cochlea) umgewandelt. Dieser Druck verursacht frequenzabhängige Erregungsmuster auf der Basilarmembran, durch die Haarzellen angeregt werden, welche wiederum elektrische Aktionspotentiale in den Neuronen des Auditiven Systems auslösen. Diese werden auf höheren Niveaus mit den Informationen des zweiten Ohrs kombiniert und verarbeitet.

Die Wahrnehmung einer Schallquelle im Raum ist durch folgende akustische Parameter bestimmt:

Lokalisation in der horizontalen Ebene (Azimut) und vertikalen Ebene (Elevation), Distanzwahrnehmung, Räumlichkeitseindruck (erzeugter Diffusfalls bzw. Impulsantwort des Raumes, wenn durch Nutzsignal angeregt).

Im Folgenden sind die hinter diesen Begriffen stehenden psychoakustischen Mechanismen, die dem menschlichen Hören zugrunde liegen kurz erklärt.

2.1 Mechanismen der Lokalisation

Die Lokalisation von Klangereignissen basiert beim Menschen primär auf folgenden fünf frequenzabhängigen Schallinformationen:

- 1) Interaurale Zeitunterschiede (engl. Interaural Time Differences - ITD), und
- 2) Interaurale Pegelunterschiede (engl. Interaural Level Differences - ILD) stellen die zwei wichtigsten Informationen dar, anhand denen das menschliche Gehör entscheidet, in welchem sog. 'cone of confusion' eine Klangquelle angesiedelt ist.

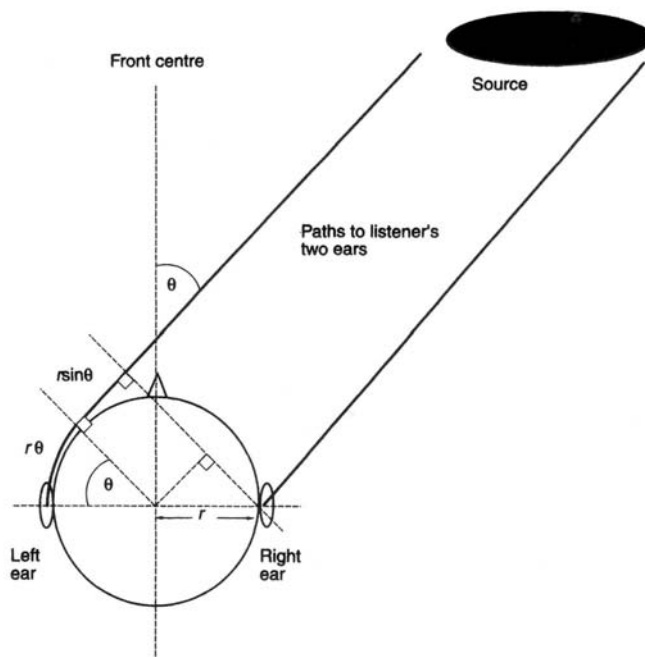


Abb. 2.2 Die interaurale Zeitdifferenz (ITD) für einen Hörer hängt vom Winkel des Schalleinfalls ab, da dieser den zusätzlichen Weg bedingt, den der Schall zum abgewandten Ohr zurücklegen muss. In diesem Modell lässt sich die ITD berechnen wie folgt: $\Delta t = r(\theta + \sin \theta)/c$ (mit der Schallgeschwindigkeit $c = 340\text{m/s}$ und θ in radians) (nach [Rumsey, 2001])

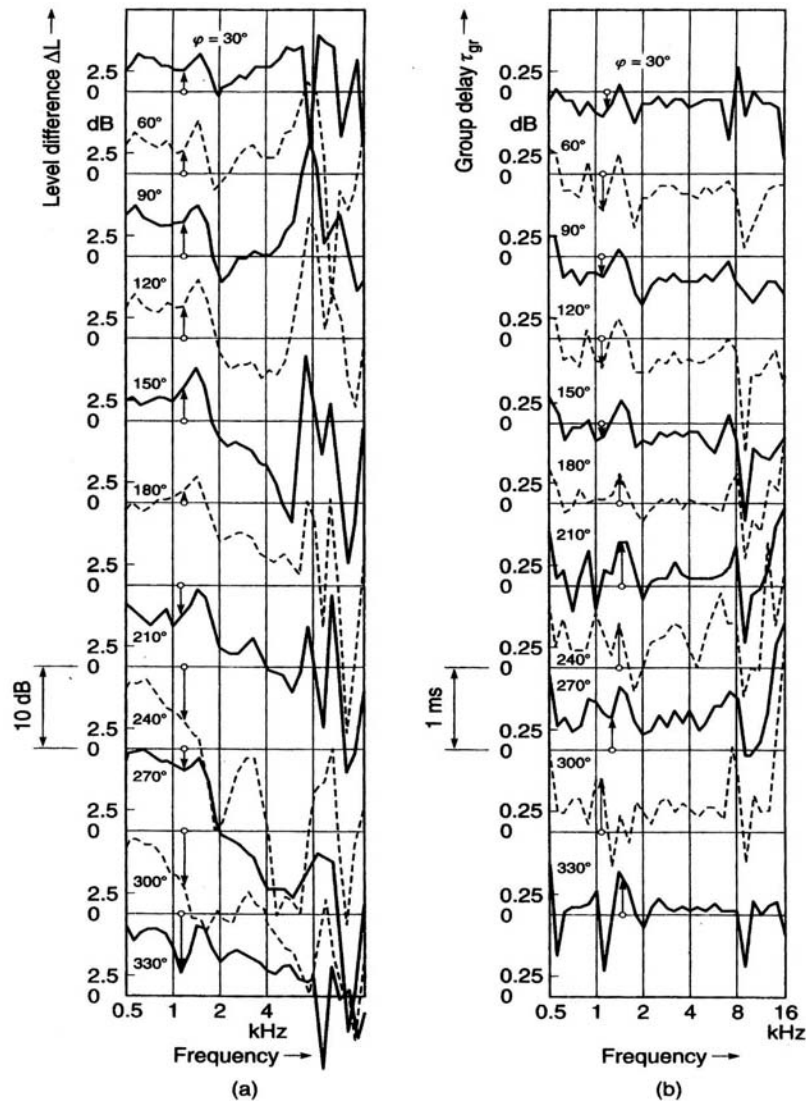


Abb.2.3: Monaurale Übertragungsfunktion des linken Ohres für mehrere Richtungen der Horizontalebene, (vorne: $\varphi=0^\circ$), reflexionsarmer Raum. Lautsprecherentfernung 2m, 25 Personen, komplex gemittelt. (a) Pegeldifferenz, (b) Zeitdifferenz (nach [Blauert, 1997])

Als einer der ersten hatte Lord Rayleigh im Rahmen seiner Experimente zur menschlichen Hörwahrnehmung gegen Ende des 19. Jahrhunderts erkannt, dass bei Schallwellen mit kleiner Wellenlänge im Vergleich zu den Abmessungen des Kopfes, durch diesen eine abschattende Wirkung für das der Schallquelle abgewandte Ohr auftritt, was bewirkt, dass an den Ohren ein Pegelunterschied (ILD) auftritt. Darüber hinaus entstehen aufgrund der unterschiedlich langen Schallwege von der Schallquelle

zu den beiden Ohren (sofern sich die Schallquelle nicht genau in der Median-Ebene des Hörers befindet) unterschiedliche Laufzeiten (ITD) an den Ohren.

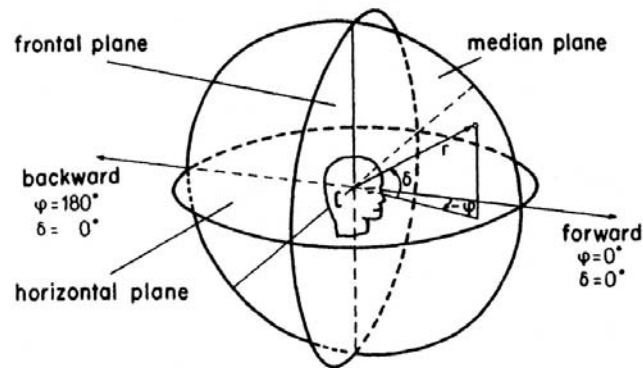


Abb. 2.4: Ein kopfbezogenes Koordinatensystem, wie es bei Hör-Experimenten verwendet wird; Definition von Median-, Frontal-, und Horizontalebene; r ist die Distanz zur Schallquelle, φ der Azimut, δ die Elevation (aus [Blauert, 1997])

Die zwischen den Ohren auftretende maximale Laufzeitdifferenz (bei seitlichem Schalleinfall bei 90°) liegt für natürliche Schallquellen in der Größenordnung von $650\mu\text{s}$ ($=0.65\text{ms}$), wobei ein mittlerer Ohrabstand von 21cm zugrunde gelegt wird (vgl. [Rumsey, 2001], S. 22 sowie [Blauert, 1997], S. 143).

Weitere Forschungen [Yanagawa et al., 1976] und [Suzuki and Tohyama, 1981], [Tohyama and Suzuki, 1989] belegen, dass die akustische wirksame Distanz zwischen den Ohren tatsächlich höher ist und bei ca. 33cm liegt (siehe hierzu auch [Blauert, 1997], S. 146-149). Dem würde bei einer Schallgeschwindigkeit von 340m/s eine ITD von 0.96ms entsprechen.

Lord Rayleigh konnte zeigen, dass die Laufzeitdifferenzen (ITD) besonders bei tiefen Frequenzen zum Tragen kommen, bei denen auch noch keine ILDs in relevanter Größe auftreten („Duplex-Theorie“, 1907). Seine Schlussfolgerung war, dass die Lokalisation bei tiefen Frequenzen durch die ITDs und bei hohen Frequenzen durch ILDs bestimmt wird. Der Übergangsbereich zwischen diesen beiden hörphysiologischen Mechanismen liegt bei ca. 1.5kHz .

Vernachlässigt man die Filterwirkung des Außenohrs (Pinna), so ergeben sich für einen

abstrahierten, kugelförmigen Kopf für die Positionen a (vorne) und b (hinten) der nachfolgenden Abbildung 2.5 die gleichen ITDs und ILDs. In Analogie zu dieser vorne-hinten Ambiguität gibt es gleichermaßen auch die sog. Elevations-Ambiguität für die Schallquellenpositionen x (oben) und y (unten).

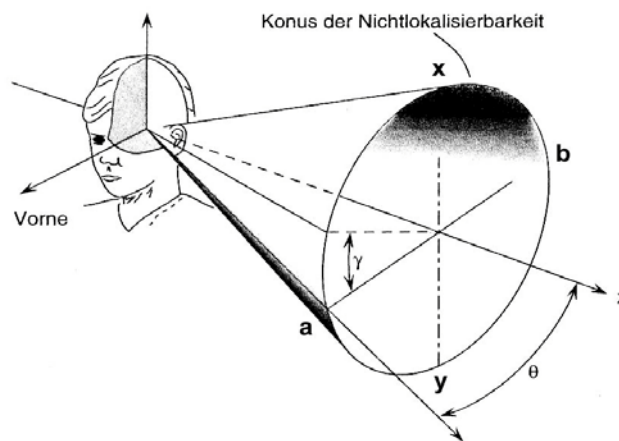


Abb. 2.5 „Konus der Nichtlokalisierbarkeit“ (engl. „cone of confusion“):

Eine nur laterale, d.h. seitliche Lokalisation der Schallquelle erfordert lediglich die Bestimmung des Winkels θ ; für eine eindeutige Lokalisation ist auch die Bestimmung des Winkels γ erforderlich (nach [von Hornbostel und Wertheimer, 1920], Graphik übernommen aus [Hall, 1980], S.342)

Beim ‚natürlichen‘ Hören (im Gegensatz zum Hören z.B. mit Kopfhörer) kommen zwei hilfreiche Mechanismen hinzu:

3) Der Frequenzgang monauraler Signale (Pinna-Frequenzgang) hilft normalerweise dabei, die Schallquelle innerhalb des ‚Cone of confusion‘ richtungsgemäß genauer einzuordnen. (siehe Abb. 2.5)

Ursprünglich hatte bereits Mach vermutet, dass die Außenohren eine gewisse ‚Richtwirkung‘ hätten und deshalb für die Lokalisation von Bedeutung wären. D. W. Batteau veröffentlichte in den 1960-er Jahren seine Untersuchungen über die Effekte der Pinnae-Reflexionen im Zeitbereich (siehe [Batteau]). Er sah in den Pinnae richtungsabhängige Reflektoren, die zum einfallenden Schall in Abhängigkeit von der

Einfallsrichtung (Azimut, Elevation) jeweils ein charakteristisches Muster an Echogenerieren. Bei seinen Messungen stellte er zeitliche Verzögerungen im Bereich von 10 bis $300\mu\text{s}$ fest. Die Aufgabe des Innenohrs wäre es demnach, diese Echomuster zu detektieren und dadurch die Schalleinfallsrichtung abzuleiten. Dafür wäre allerdings eine höhere Empfindlichkeit in der zeitlichen Domäne notwendig, als de facto für das menschliche Ohr nachgewiesen werden konnte.

[Hebrank und Wright, 1974] fanden heraus, dass die von Batteau gefundenen Pinnae-Reflexionen spektrale Entsprechungen haben, die tatsächlich vom menschlichen Gehör dekodiert werden können.

Die später erfolgten Untersuchungen, welche von verschiedenen Forschern durchgeführt wurden, hatten zum Ergebnis, dass die durch die Pinnae auftretende Filterung besonders für die Lokalisation in der Medianebene und die Vorne-/Hinten Unterscheidung wichtig sind.

4) Weiters helfen kleine Kopfdrehungen und die damit verbundenen ITD-Änderungen dem Menschen, bei Unsicherheit Klarheit über die Position einer Schallquelle zu erhalten [Blauert, 1997].

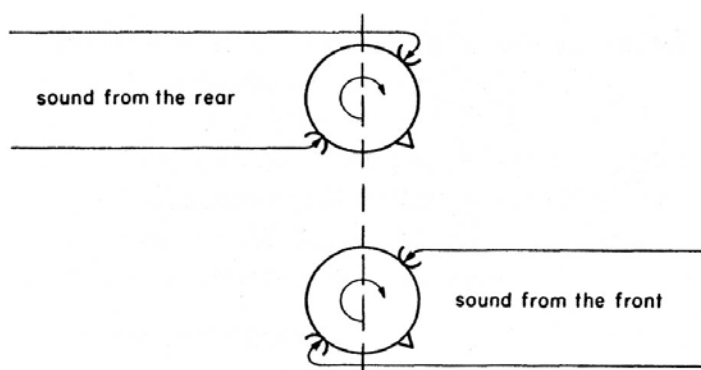


Abb. 2.6: Die Änderung der interauralen Differenzen der Ohrsignale im Verlauf von Kopfdrehungen sind für von vorwärts bzw. rückwärts einfallende Schallwellen entgegengesetzt (aus [Blauert, 1997], S.180)

5) Außerdem sind Menschen bezüglich interauraler Kohärenz sehr empfindlich (siehe hierzu z.B. [Boehnke et al., 2002] und die darin zitierten Referenzen, sowie [Tohyama and Suzuki, 1989]), welche vor kurzem auch als wichtiger Anhaltspunkt für Lokalisation in halligen Räumen und Umgebungen mit mehreren, gleichzeitigen Schallquellen eingestuft wurde [Faller and Merimaa, 2004].

Alle oben genannten Merkmale sind bei den Menschen individuell sehr unterschiedlich ausgeprägt, da sie von der Form des Kopfes, der Außenohren und dem Oberkörper des jeweiligen Hörers abhängig sind. Diese intersubjektiven Unterschiede können mit Hilfe der Messung der personenbezogenen HRTF's (head-related-transfer-functions) ermittelt werden.

Die zeitliche und spektrale Auflösung des menschlichen Gehörs wurde für monaurale Signale ausgiebig untersucht (siehe z.B. [Moore, 1997]). Für binaurales Hören scheint die Frequenzauflösung ähnlich der bei monauralem Hören zu sein [Kohlrausch, 1988], [van der Heijden and Trahiotis, 1998], wenngleich für einige Testsignale auch etwas größere Bandbreiten gefunden wurden [Kollmeier and Holube, 1992], [Holube et al., 1998]. Somit kann im Allgemeinen die für das monaurale Hören gefundene ‚Equivalent Rectangular Band‘ (ERB) Auflösung [Glasberg and Moore, 1990] verwendet werden (siehe Abb. 2.7). An dieser Stelle sollte auch erwähnt werden, dass offenbar nicht alle Frequenzen gleiches Gewicht bei der Bestimmung der Lokalisation von breitbandigen Klangreizen haben: das ‚Weighted-Image Lateralization Model‘ von Stern et al. [Stern et al., 1988] zeigt eine dominante Region um 600Hz, und Wightman und Kistler [Wightman and Kistler, 1992] haben die diesbezüglich große Bedeutung von ITDs im LF-Bereich nachgewiesen.

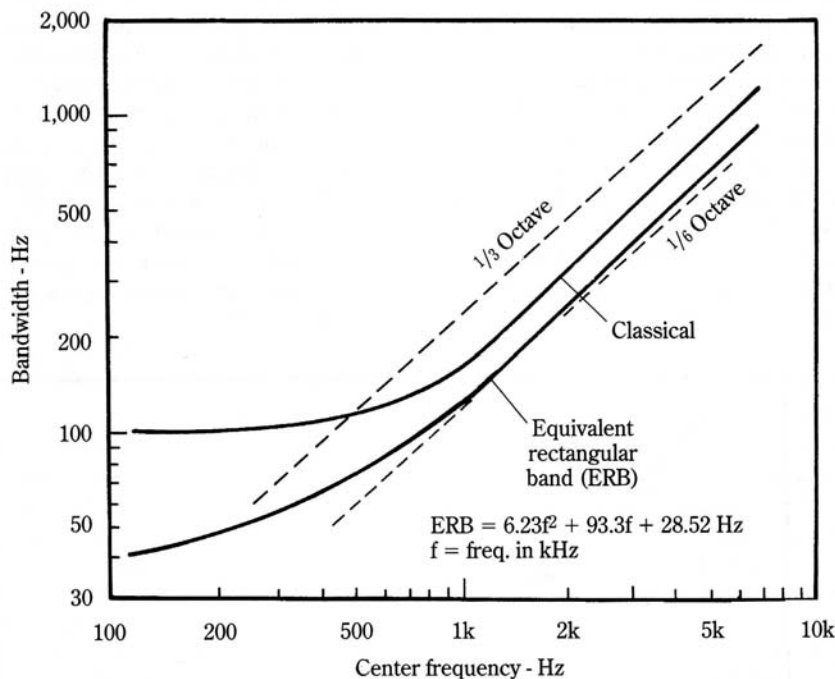


Abb. 2.7 ERB – ‚Equivalent Rectangular critical Band‘:

Ein Vergleich der Bandbreiten von 1/3- und 1/6-Oktavbändern, der ‚Kritischen Bandbreite‘ des menschlichen Gehörs, sowie des ‚Equivalent Rectangular critical Band‘ (ERB), welches nach der in der Abbildung angeführten Formel berechnet wurde (aus [Everest, 1994])

Die zeitliche Auflösung des menschlichen Gehörs zu bestimmen ist schwieriger. Es konnte nachgewiesen werden, dass menschliche Hörer in der Lage sind, die räumliche Bewegung von Klangquellen nachzuvollziehen, die in Form von sinusförmigen Fluktuationen der ITDs und ILDs ihren Niederschlag finden, allerdings nur bis zu einer Fluktuationsfrequenz von ca. 2.4 bis 3.1 Hz (siehe hierzu [Blauert, 1972]). Nichtsdestoweniger konnten [Granthman und Wightman, 1978] beobachten, dass ihre Testhörer in der Lage waren, ITD-Fluktuationen bis 500Hz zu registrieren, allerdings primär basierend auf der daraus resultierenden Verbreiterung der Schallquellenbreite (ASW – Apparent Source Width).

Mehrere Studien zu diesem Thema hatten zum Ergebnis, dass die Binaural-Wahrnehmung offenbar ähnlich einer Signalverarbeitung mit einem doppelseitig exponentiellen, gerundet exponentiellen, oder Gauß’schen Zeitfenster funktioniert (siehe hierzu [Holube et al., 1998], [Grantham and Wightman, 1979], [Kollmeier and Gilkey, 1990], [Culling and Summersfield, 1999], [Akeroyd and Summersfield, 1999],

sowie [Breebaart et al., 2002]). Die für diese Zeitfenster angegebenen Längen sind teilweise sehr unterschiedlich, meist liegt die durchschnittliche ‚Equivalent Rectangular Duration‘ (ERD), d.h. die durchschnittliche Integrationszeit bzw. ‚Zeitkonstante‘ des menschlichen Ohrs in der Größenordnung von ca. 100ms.

Allerdings wurden auch wesentlich kürzere Zeitfenster von einigen Forschern gefunden. (siehe hierzu [Akeroyd und Bernstein, 2001] mit einem ERD von 10ms, sowie [Bernstein et al., 2001] mit 13.8ms). Als mögliche Erklärung hierfür wurde angegeben, dass die unterschiedlichen Zeitkonstanten verschiedenen Teilbereichen des binauralen Gehörmechanismus entsprechen. ([Kollmeier and Gilkey, 1990], [Akeroyd und Bernstein, 2001], [Bernstein et al., 2001])

Der sog. ‚Präzedenz‘-Effekt (Haas Effekt; sog. ‚Gesetz der ersten Wellenfront‘, siehe [Haas, 1951]) spielt ebenfalls im Zusammenhang mit der Lokalisation eine bedeutende Rolle. Bei zwei knapp aufeinander folgenden Reizen aus unterschiedlichen Richtungen dominiert der erste von beiden die Lokalisationsrichtung, sofern die beiden Schallereignisse einen Zeitversatz von nur wenigen Millisekunden haben [Blauert, 1997], [Litovsky et al., 1999]. Bei Wiedergabe eines Testsignals (mono) über Kopfhörer, wandert dieses für ITDs von 0 – 0.6ms scheinbar im Kopf vom Zentrum zu jenem Ohr, bei dem der Schall zuerst eintrifft. Für den Zeitbereich von 0.6 bis ca.35ms wird die Schallquelle nur an dem ‚vorausgehenden‘ Ohr wahrgenommen; bei ITDs, die größer als 35ms sind, werden zwei getrennte Schallereignisse in Form eines Echos wahrgenommen.

HRTF Phasen-Charakteristik

Zeitunterschiede können auch als Phasenunterschiede interpretiert bzw. dargestellt werden, wenn es sich um rein sinusförmige Signale handelt. Das Ohr ist in der Lage Phasenunterschiede bei tiefen Frequenzen wahrzunehmen; ab einer Frequenz von ca. 700 Hz entspricht der Kopfdurchmesser etwa der halben Wellenlänge des Schalls und somit wird auch der ‚Phasendetektionsmechanismus‘ des menschlichen Gehörs außer Kraft gesetzt, der bis dahin feststellen konnte, an welchem Ohr das Schallereignis früher eintrifft. An den Ohren auftretende Phasendifferenzen können außerdem dann

irreführend sein, wenn sie durch in Räumen auftretende Raum-Moden, oder auftretende Reflexionen bedingt sind.

Die untenstehende Abb.2.8 zeigt die interaurale Phasendifferenz für eine Person von 0° bis 150° in 30° Schritten in der Horizontalebene. Die dickeren Linien zeigen die Umsetzung der Phasenverzögerung in interaurale Zeitverzögerung bei 0.25ms, 0.5ms und 1.0ms.

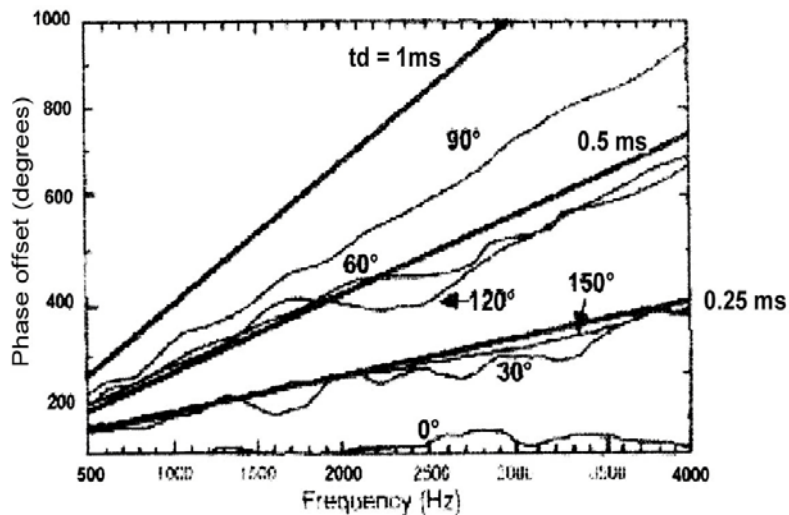


Abb. 2.8 Kontinuierlich fortgesetzte interaurale Phasendifferenz (nach [Begault, 1994])

Wie bereit weiter oben erwähnt, funktioniert Lokalisation auf Basis der ITDs gut bis ca. 1.5kHz. Oberhalb ist die Phasenlage nicht mehr eindeutig, unser Gehörmechanismus ist jedoch in der Lage, bei höheren Frequenzen auch ITDs der sog. ‚Einhüllenden‘ (Amplituden-Hüllkurve) des Schallereignisses auszuwerten.

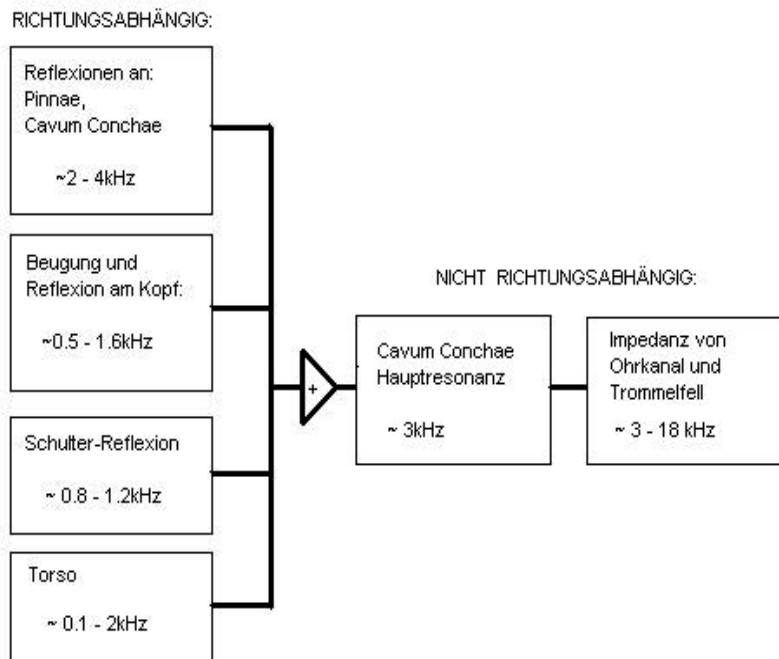


Abb. 2.9 HRTF-Blockdiagramm nach Blauert (aus [Fellner und Höldrich, 1998 b])

Neben dem Außenohr, das den größten spektralen Einfluss hat, führen auch Kopf, Schultern und Torso zu leichten Färbungen. In der obenstehenden Abb. 2.9 ist die Wichtigkeit der richtungsabhängigen und unabhängigen Komponenten in fallender Reihenfolge (weiter unten = weniger wichtig) angeordnet.

Das ‚Cavum Conchae‘ ist die zentrale Höhle und der größte Resonanzbereich der Ohrmuschel am Eingang zum Ohrkanal. Der äußere Gehörgang reicht vom Cavum Conchae bis zum Trommelfell. Der Ohrkanal ist etwa 2.5cm lang mit 7-8mm Durchmesser. Durch die Abmessungen bedingt besteht eine starke Resonanz um ca. 4kHz.

Man nimmt heute an, dass die Gehörgangs-Resonanz auch dafür verantwortlich ist, dass die Externalisation (im Gegensatz zur ‚Im-Kopf-Lokalisation‘) eines Schalleindrucks entstehen kann (vgl. [Rumsey, 2001], S. 26). Aus diesem Grund gibt es auch Bemühungen, den Gehörgang (z.B. bei Schallwiedergabe mittels Kopfhörer) gezielt aus bestimmten Richtungen von seinem Eingang anzuregen, um dem am Trommelfell eintreffenden Signal somit auch bestimmte individuelle, für den jeweiligen Hörer charakteristische Merkmale aufzuprägen, die eine ‚Außer-Kopf-Lokalisation‘ (Externalisation) erleichtern. [Tan und Gan, 2000]

Für die Schultern liegt die spektrale Beeinflussung im wirksamen Bereich bei ca. +/- 5dB, für den Torso bei etwa +/-3dB.

Lokalisation mit HRTFs

Eines der wichtigsten Ergebnisse der Untersuchungen von Blauert in den 1970-er Jahren war die Entdeckung der Richtungsbestimmenden Frequenzbänder (engl. directional bands) [Blauert, 1997]

Wie bereits aus der obenstehenden Abb. 2.3 der einzelnen HRTF-Kurven zu ersehen war, gibt es in Abhängigkeit von der Position der Schallquelle mehrere charakteristische Pegel-Überhöhungen und Einbrüche im Frequenzgang. Diese stellen eine Art ‚Frequenzcodierung‘ von Schallquellenpositionen dar, die das menschliche Gehirn offenbar in jungen Jahren erlernt. Typischerweise sind z.B. Schallquellen hinter dem Hörer in der Regel durch geringeren Pegel bei hohen Frequenzen gekennzeichnet, was auf die Richtwirkung der Außenohren zurückzuführen ist.

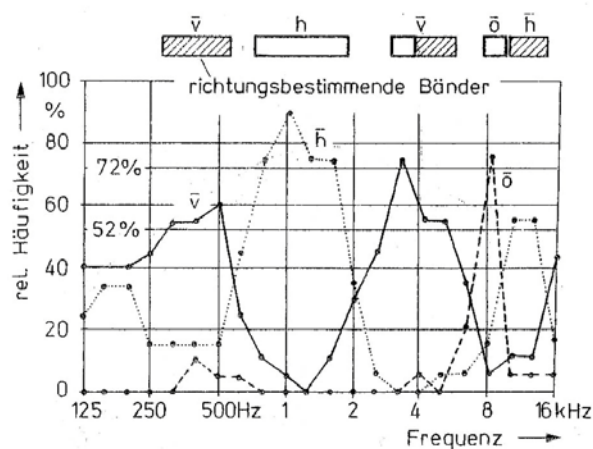


Abb. 2.10 Richtungsbestimmende Bänder nach [Blauert,1974]:

Relative Häufigkeiten der Versuchspersonen, die mit 95% Sicherheit eine der drei Antworten h, o, v häufiger abgeben, als die beiden anderen zusammen. Oben im Bild sind Frequenzbänder angegeben, in denen die absolute Mehrheit einer Grundgesamtheit von Versuchspersonen die gleiche Antwort häufiger abgibt, als die beiden anderen möglichen Antworten zusammen; schwarz umrandete Bereiche: 90% Sicherheit, schraffierte Bereiche: wahrscheinlichster Fall (aus [Blauert, 1974])

Wie aus der obenstehenden Abbildung ersichtlich, ist der Frequenzbereich um 8kHz für

die ‚über Kopf‘ Lokalisation verantwortlich, und ‚vorne‘-Lokalisation wird durch die Frequenzbänder von 300-600Hz und 3000-6000Hz bestimmt. ‚Hinten‘-Lokalisation ist hingegen mit dem Frequenzbereich um 1200Hz und 12000Hz verknüpft.

Ähnliche Ergebnisse wurden auch von [Hebrank und Wright, 1974], sowie [Begault, 1994] gefunden.

2.2 Mechanismen der Distanzwahrnehmung

Nachfolgende Erklärungen stellen eine Übersetzung eines Ausschnitts aus [Nielsen, 1993], mit Ergänzungen aus [Fellner und Höldrich, 1998a], sowie [Rumsey, 2001] dar.

2.2.1 Schallintensität

Der wohl wichtigste Anhaltspunkt für die Distanzwahrnehmung ist die Schallintensität, die vom menschlichen Gehör in das subjektive Empfinden der Lautheit umgesetzt wird. Die Lautheit stellt die empfundene Amplitude der Schallintensität dar, welche frequenzabhängig ist, entsprechend der Empfindlichkeit des Gehörs (vgl. Fletcher-Munson Kurven). Die Schallintensität nimmt bei Verdopplung der Entfernung von der Schallquelle im Freifeld mit dem Quadrat der Entfernung ab (was einer Pegelreduktion von 6dB_{SPL} entspricht); die Halbierung der subjektiv empfundenen Lautheit – welche in Sone angegeben wird – entspricht einer Abnahme des Schalldruckpegels um ca. 10dB (dies gilt für den Frequenzbereich von 400Hz bis 5kHz im Schallpegelbereich von 40 bis ca. 100dB). Anzumerken ist hierbei, dass die Abschätzung der Distanz einer Schallquelle in einer ‚natürlichen‘ Umgebung normalerweise in Verbindung mit dem Sehen erfolgt, weshalb die Beurteilung der Distanz einer bekannten Schallquelle in einer bereits bekannten Umgebung leichter fällt.

Die Aussagekraft der Lautheit als Gradmesser für die Entfernung hängt auch davon ab, ob dem Hörer die ‚normale‘ Lautstärke der Schallquelle bekannt ist [Nielsen, 1993].

2.2.2 Diffusschallanteil

Im Gegensatz zum schalltoten Raum (oder bei Freifeldausbreitung) ist die Abnahme des Schalldrucks mit steigender Entfernung von der Schallquelle in einem ‚normalen‘ Raum aufgrund der auftretenden Hallanteile geringer.

Das Verhältnis von Diffusschall (Hall) zu Direktschall stellt offenbar einen noch stärkeren Gradmesser zur Distanzbestimmung dar, als die beim Hörer auftretende Schallintensität. Nach einer Hypothese von Peter Craven wird die empfundene Distanz einer Schallquelle durch die relative Zeitverzögerung und die relative Amplitude des frühen reflektierten Schalls im Verhältnis zum Direktschall bestimmt (siehe hierzu [Gerzon, 1992]).

Ist der Hörraum nicht bekannt (wie dies z.B. in Testsituationen der Fall sein kann), so ist der menschliche Gehörmechanismus offenbar in der Lage, sich die benötigte Information über die Beschaffenheit des Raumes - mehr oder weniger bewusst - aus dem Gehörten zu extrahieren [Nielsen, 1993]. In ihren Experimenten stellten [Mershon und Bowers, 1979] fest, dass es keinen Unterschied ausmachte, ob ihren Testpersonen der Hörraum vor Präsentation der Stimuli bereits bekannt war, oder nicht. Allerdings stellt der durch die Schallquelle hervorgerufene Diffusschall offenbar eine weitere, wichtige Informationsquelle für den menschlichen Hörer dar: so zeigte sich, dass im schalltoten Raum - vor allem aufgrund des Fehlens von Reflexionen / Diffusschall – von den Testhörern kein eindeutiger Bezug zwischen der empfundenen und der tatsächlichen physikalischen Distanz der Schallquelle hergestellt werden konnte [Nielsen, 1993].

2.2.3 Frequenzgang

Da hochfrequente Schallanteile bei der Ausbreitung in Luft stärker bedämpft werden, als tieffrequente, enthält auch das Frequenzspektrum eines Schallereignisses Informationen über dessen Entfernung zum Hörer. Hierfür ist natürlich die Kenntnis über die spektrale Zusammensetzung des Klanges auf Hörerseite eine Voraussetzung. In einem Raum mit Diffusschallanteil ändert sich die spektrale Zusammensetzung des Halls sowohl aufgrund des frequenzabhängigen Absorptionsverhaltens der Begrenzungsflächen, sowie auch aufgrund der vom Schall in der Luft zurückgelegten Strecke.

Der Absorptionskoeffizient der Luft ist sowohl von der Temperatur, als auch der Luftfeuchtigkeit abhängig, was vor allem bei Freiluft-Beschallungsanlagen zu relevanten klanglichen Änderungen führen kann. Auch andere lokale klimatische Faktoren (Wind, etc.) können zu einer Veränderung des Frequenzganges in Form einer spektralen Färbung beitragen.

2.2.4 Binaurale Differenzen

Vor allem im Zusammenhang mit Schallquellen, die sich sehr nahe zum Hörer befinden, treten relevante Änderungen der HRTF-Frequenzgänge auf. So u. a. auch eine Verstärkung der interauralen Pegelunterschiede vor allem bei tiefen und hohen Frequenzen (vgl. [Huopaniemi, 1999]).

2.3 Raumeindruck

Der bei einem Hörer entstehende Raumeindruck (engl. ‚spatial impression‘ oder ‚spaciousness‘) setzt sich primär aus zwei Komponenten zusammen: der sog. ‚Apparent Source Width‘ (ASW), d.h. der empfundenen Breite oder Ausdehnung einer Schallquelle, sowie ‚Listener Envelopment‘ (LEV), der akustische ‚Einhüllung‘ des Hörers.

Die akustischen Eigenschaften eines Raumes lassen sich im Zeitbereich durch die sog. Impulsantwort (engl. ‚Impulse Response‘ (IR)) bzw. das daraus abgeleitete Reflektogramm darstellen.

Traditionsgemäß unterscheidet man bei der Untersuchung des in einem Raum auftretenden Diffusschalls (Halls) zwischen dem frühen Anteil an Reflexionen, die innerhalb von 80ms nach dem Schallereignis auftreten (engl. ‚early reflections‘, (ER)), und den späten Reflexionen (engl. ‚late reflections‘, (LR)). Während die ersten (Einzel-) Reflexionen an den Begrenzungsflächen des Raumes dem Hörer einen Eindruck von seiner Position im Raum und dessen Abmessungen geben, sind die später auftretenden Reflexionen (Nachhall) bereits wesentlich diffuser und vom Pegel her schwächer, da sie

von mehrfachen Reflexionen an den Begrenzungsflächen des Raumes bzw. darin befindlichen Objekten herrühren.

Die auf Basis psychoakustischer Untersuchungen getroffene Unterscheidung der reflektierten Schallanteile früher oder später als 80ms nach dem Schallereignis, findet ihren Niederschlag u. a. in der im Bereich der Raumakustik üblichen, messbaren Größe des Klarheitsmaßes (engl. ‚clarity‘ (C)):

$$C = 10 \lg \frac{\int_0^{80 \text{ ms}} p^2(t) dt}{\int_{80 \text{ ms}}^{\infty} p^2(t) dt} \quad [\text{dB}] \quad (2.1)$$

welches durch [Reichardt und Kussev, 1972] definiert wurde. (mit: p ... Schalldruck)

Die oben genannte ASW ist primär beeinflusst durch Schallreflexionen, die innerhalb der ersten 80ms beim Hörer eintreffen, wobei in diesem Zusammenhang vor allem der Frequenzbereich von 1- 10kHz von Bedeutung ist (siehe die untenstehende Abb. 2.11). Zur Ausprägung von LEV hingegen tragen vor allem tieffrequente Schallanteile (von unterhalb 200Hz bis ca. 500Hz) bei, welche erst nach 80ms beim Hörer eintreffen, und dies vorzugsweise aus verschiedenen Richtungen um den Hörer. (siehe ebenfalls untenstehende Abb.)

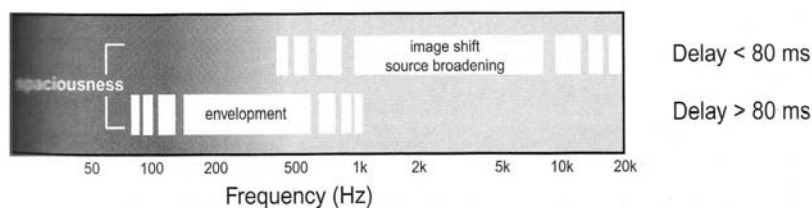


Abb. 2.11 Frequenz- und Zeitbereiche, in denen Reflexionen des Direktschalls zu ASW and LEV beitragen (aus [Toole, 2008]).

Gemäß Untersuchungen von [Hidaka et al., 1997] ist die ASW am stärksten beeinflusst durch Reflexionen in den Oktav-Frequenzbändern bei 500, 1000 und 2000Hz, welche in diesen Frequenzbereichen eine geringe IACC (Inter-Aural Cross Correlation; für nähere Details zur Definition des IACC siehe Kapitel 3) zur Folge haben. In der von Hidaka et

al. definierten Größe $IACC_{E3}$ wird der Mittelwert der IACC dieser 3 Oktavbänder unter Beschränkung auf die ersten 80ms (E = early) berechnet, welcher in direkte Relation zur ASW gesetzt werden kann.

[Ando, 1977] hatte bereits gezeigt, dass es hinsichtlich des Einfallswinkels einer Einzel-Reflexion eine Hörer-Präferenz für 60 Grad gibt, was messtechnisch in einem reduzierten IACC für diesen Bereich seinen Niederschlag findet. In einer späteren Untersuchung von [Barron and Marshall, 1981] finden diese Erkenntnisse Bestätigung.

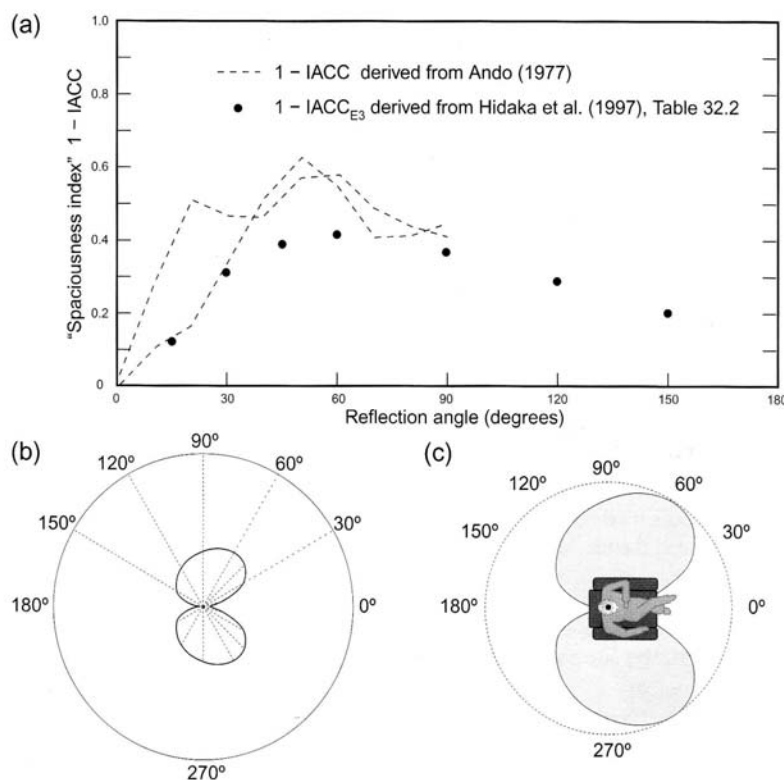


Abb. 2.12: ‚Räumlichkeits-Index‘ (1-IACC) in Abhängigkeit vom Winkel einer einfallenden Reflexion; verschiedenen Darstellungsformen (a, b, c) (Abbildung 7.4 a aus: [Toole, 2008], S. 105)

Untersuchungen von Griesinger haben gezeigt, dass die idealen Einfallswinkel zur Erzielung des Eindrucks der ‚Einhüllung‘ beim Hörer frequenzabhängig sind: während tieffrequente Komponenten unter 700Hz idealerweise seitlich einfallen sollten (d.h. +/- 90° vom Hörer), so lässt sich bei höheren Frequenzen LEV auch bereits mit Schallanteilen hervorrufen, die der Medianebene wesentlich näher sind. (siehe hierzu [Griesinger, 1999])

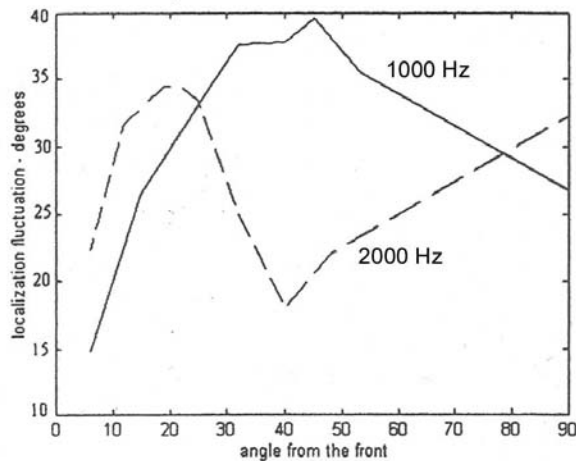


Abb. 2.13: Fluktuation der ITD für eine einzelne Reflexion in Abhängigkeit vom Einfallswinkel für zwei Frequenzen (Computer-Simulation). Dies zeigt zugleich, dass der sich ergebende Raumeindruck von Einfallswinkel und Frequenz des Schallreizes abhängig ist. (aus [Griesinger, 1997])

Wie aus der obenstehenden Abbildung 2.13 ersichtlich ist, tritt beispielsweise für einen Schallreiz (Einzel-Reflexion) mit einer Frequenz von 1000Hz bei einem Einfallswinkel um 45 Grad eine hohe Fluktuation der ITD, und somit ein starkes Gefühl der akustischen Einhüllung (Envelopment) auf. Für Frequenzen nahe 2000Hz hingegen tritt im selben Winkelbereich (40-50 Grad) nur wesentlich geringere Fluktuation auf, sodass sich eine geringere Einhüllung ergibt.

Bei [Toole, 2008] werden die englischen Begriffe ‚spaciousness‘ und ‚envelopment‘ gleichgesetzt (siehe ebendort S. 99). Bei [Rumsey, 2001] findet sich in diesem Zusammenhang auch den Begriff ‚room impression‘, und es wird eine Verbindung zur sog. ‚externalisation‘ (d.h. ‚Außer-Kopf-Lokalisation‘) des Klangreizes hergestellt (siehe ebendort S. 38).

Bei [Saroff und Bello, 2008] wurden zur Abschätzung des in einer stereophonen Aufnahme enthaltenen Räumlichkeitseindrucks zwei quantitative Modelle zur Erfassung a) der stereophonen Basisbreite (für ASW) mit Hilfe unterschiedlich breit panoramierter Einzelquellen (Stützmikrofonsignale), wie auch b) des Gesamtanteils des Hallsignals (für LEV) vorgeschlagen und getestet.

In [Rafaely and Avni, 2009] wird untersucht, inwiefern sich der Räumlichkeitseindruck über den Zusammenhang zwischen IACC und der räumlichen Korrelation in diffusen Schallfeldern mittels ‚Spherical Harmonics‘ darstellen bzw. erfassen lässt.

2.4 Physikalische Größen im Zusammenhang mit dem Räumlichkeits- eindruck

Im letzten Abschnitt wurde bereits dargestellt, dass der subjektive Eindruck der ASW am besten durch den $IACC_{E3}$ erfasst wird (siehe [Hidaka et al., 1997]).

Bereits 1975 wurde von [Lehmann, 1975] das ‚Raumeindrucksmaß‘ R definiert (siehe hierzu auch [Reichardt et al., 1975]). Um dieses messtechnisch erfassen zu können, wird ein 40° -Kegel um die 0° -Achse der Bezugsrichtung festgelegt; Schallenergieanteile, die mit mehr als 80ms Verzögerung innerhalb des Kegels, sowie jene, die seitlich (also außerhalb des Kegels) mit einer Verzögerung zwischen 25ms und 80ms eintreffen, werden dabei als räumlichkeitsfördernd betrachtet.

Nicht die Räumlichkeit fördernd sind alle innerhalb der ersten 25ms eintreffenden Schallanteile, sowie jene Schallreflexionen, die von vorne innerhalb 25-80ms eintreffen. Wird die räumlichkeitsfördernde Schallenergie mit E_R bezeichnet und die nicht förderliche mit E_{NR} so ergibt sich das Raumeindrucksmaß zu:

$$R = 10 \lg \frac{E_R}{E_{NR}} \text{ [dB]} \quad (2.2)$$

Die messtechnische Erfassung des Raumeindrucksmaßes erfordert zumindest den Einsatz eines Richtmikrofons zusätzlich zum Druckempfänger (Kugelcharakteristik), sodass sich in der praktischen Anwendung R wie folgt ermitteln lässt:

$$R = 10 \lg \frac{\int_{25ms}^{\infty} p_{Kugel}^2(t) dt - \int_{25ms}^{80ms} p_{Richtrohr.vorne}^2(t) dt}{\int_0^{25ms} p_{Kugel}^2(t) dt + \int_{25ms}^{80ms} p_{Richtrohr.vorne}^2(t) dt} \text{ [dB]} \quad (2.3)$$

Der Vollständigkeit halber sei an dieser Stelle auch noch das von [Beranek und Schultz, 1965] definierte ‚Hallmaß‘ H genannt:

$$H = 10 \lg \frac{\int_0^{\infty} p^2(t) dt}{\int_0^{50ms} p^2(t) dt} \text{ [dB]} \quad (2.4)$$

Betreffend die von einem Hörer im Konzertsaal empfundene Einhüllung (LEV) wurden von verschiedenen Forschern die messtechnisch erfassbaren Größen ‚Lateral Fraction‘ (LF), sowie ‚Lateral Gain‘ (LG₈₀) vorgeschlagen.

Untersuchungen von [Marshall, 1968] und [Barron, 1971] belegen die Wichtigkeit der seitlichen Reflexionen für den Räumlichkeitseindruck im Konzertsaal. Die ‚Lateral Energy Fraction‘ (LF) wurde von [Barron und Marshall, 1981] als Maß für die ASW verwendet, und wie folgt definiert:

$$LF = 10 \log \left[\frac{\int_0^{0.08} p_F^2(t) dt}{\int_0^{0.08} p_O^2(t) dt} \right] \text{ [dB]} \quad (2.5)$$

wobei $p_F(t)$ dem Schalldruck der Impulsantwort entspricht, der an einem Punkt im Konzertsaal mittels eines Mikrofons mit 8-er Charakteristik gemessen wird, dessen Nullachse in Richtung der Schallquelle zeigt, und $p_O(t)$ dem Schalldruck am selben Punkt entspricht, allerdings mit einem omnidirektionalen Mikrofon (Schalldruckempfänger) gemessen. Normalerweise werden die beiden Größen zur Berechnung von LF über die ersten 80ms der Impulsantwort integriert.

Bei [Bradley und Soloudre, 1985] findet sich die Definition des ‚Lateral Gain‘, welcher die Schallenergie des Anteils des späten (d.h. nach 80ms), seitlich einfallenden Diffusschalls in Relation setzt zu dem an der Hörerposition mit einem omnidirektionalen Mikrofon gemessenen. In [Soulodre et al., 2003] ist LG₈₀ definiert, wie folgt:

$$LG_{80} = 10 \log \frac{\int_{0.08}^{\infty} p_F^2(t) dt}{\int_0^{\infty} p_A^2(t) dt} \text{ [dB]} \quad (2.6)$$

wobei $p_F(t)$ dem Schalldruck in Abhängigkeit von der Zeit entspricht, gemessen mit einem seitlich ausgerichteten Mikrofon mit Achter-Charakteristik, und $p_A(t)$ der Schalldruck gemessen mit einem omnidirektionalen Mikrofon, beide in einer Entfernung von 10m unter ‚Freifeld‘-Bedingungen.

In diesem Paper wird auch erörtert, dass die Beschränkung auf ein festes Zeitintervall von 80ms (unabgänglich vom Frequenzbereich, um den es geht) der frequenzabhängigen

‚Trägheit‘ des menschlichen Gehörs nur ungenügend Rechnung trägt. Soulodre et al. schlagen deshalb für Frequenzen bis 500Hz eine Integrationszeit von 160ms vor, bei 1000Hz etwa die Hälfte davon (75ms) und für Frequenzen oberhalb von 2000Hz eine Zeit von 45ms.

| Octave Bands (Hz) | Integration Limit x (ms) |
|----------------------|-------------------------------|
| 63 | – |
| 125 | 160 |
| 250 | 160 |
| 500 | 160 |
| 1000 | 75 |
| 2000 | 55 |
| 4000 | 45 |
| 8000 | 45 |

Tab. 2.1: frequenzabhängige Integrationszeiten, der menschlichen Psychoakustik angepasst
(nach [Soulodre et al., 2003])

Weiters argumentieren sie überzeugend für ein neues, objektives Maß für LEV, welches sie mit GS_{perc} bezeichnen:

$$GS_{\text{perc}} = 0.5 G_{\text{perc}} + S_{\text{perc}} \quad [\text{dB}] \quad (2.7)$$

wobei G_{perc} eine Pegel-Komponente (engl. Gain) darstellt, und S_{perc} (engl. spaciousness) eine Komponente der räumlichen Verteilung.

G_{perc} entspricht dem aus der Konzertsaal-Akustik bekannten Maß der ‚Stärke‘ (engl. 'strength', oder 'strength factor' (G) (siehe auch [Beranek, 2004], S. 617), welcher erstmals von [Lehmann, 1976] definiert wurde.

$$G_x = 10 \log \left[\frac{\int_x^\infty p_O^2(t) dt}{\int_0^\infty p_A^2(t) dt} \right] \quad [\text{dB}] \quad (2.8)$$

Gemäß dem Vorschlag von Soulodre et al. ist G_x ein Maß für den relativen Schallpegel der späten Energie, wobei in Abhängigkeit vom jeweiligen Oktavband, in dem gemessen wird, die entsprechenden Integrationszeiten x nach Tabelle 2.1 verwendet werden. Mit diesem neuen – gegenüber LG_{80} verfeinerten – Maß für die Hörer-Einhüllung gelang es, die Korrelation zu LEV um weitere ca. 4% auf 0.98 zu steigern. Auf der Suche nach einem physikalischen Maß für die Komponente der räumlichen Verteilung S_{perc} argumentiert Soulodre weiter, dass hierfür ev. ein Maß basierend auf der IACC in Frage kommen könnte, da diese gemeinhin verwendet wird,

um die räumliche Verteilung eines Schallfeldes zu messen. (siehe [Soulodre et al., 2003], S. 838, 2.Absatz).

Als interessant ist in diesem Zusammenhang anzumerken, dass IACC also nicht nur als Maß im Zusammenhang mit der ersten Komponente des Räumlichkeitseindrucks, ASW dienen kann (wie bereits weiter oben festgestellt), sondern auch ein Teil von LEV (nämlich jener, der die räumliche Verteilung S_{perc} betrifft), ebenfalls vermutlich durch ein Maß im Zusammenhang mit IACC dargestellt werden kann. Dies sind beides starke Argumente dafür, die IACC auch im Rahmen der vorliegenden Arbeit zu untersuchen und ggf. als Maß für den beim Hörer entstehenden Raumeindruck heranzuziehen.

[Griesinger, 1996] schlägt – vor allem auch in Zusammenhang mit der Akustik kleiner Räume - die sog. 'Lateral Early Decay Time' (deutsch sinngemäß etwa ‚frühe Abklingzeit seitlicher Reflexionen‘), abgekürzt LEDT, als Maß für Räumlichkeit vor:

$$LEDT_{350} = \frac{60 * 350ms}{(S(0) - SD(350)) * 1000ms / sec} \quad (2.9)$$

wobei $S(t)$ das Schröder-Integral der Impulsantwort ist, und $SD(t)$ das Schröder-Integral der Interauralen Pegel-Differenz, also des ILD.

Eine Vielzahl von Faktoren trägt zum Räumlichkeitseindruck bei: Untersuchungen von [Griesinger, 1997] zeigen, dass *Fluktuationen* von ITD und ILD ganz wesentlich zum Raumeindruck eines Klangereignisses beitragen. Untersuchungen von [Mason und Rumsey, 2002] scheinen dies zu bestätigen. Im Rahmen der letztgenannten Untersuchung zeigte sich, dass Messungen basierend auf der Bestimmung der IACC (genauer: $IACC_{E3}$) die höchste Korrelation zur subjektiv empfundenen Stereo-Breite einer Klangquelle (source width) bzw. räumlichen Ausdehnung des Aufführungsortes (environmental width) aufweisen.

ASW und IACC sind zwar im Zusammenhang mit Messungen der Akustik von Konzertsälen nützlich, eignen sich allerdings nicht immer gleichermaßen für kleine Räume. Aus diesem Grund hat [Griesinger, 1998] zusätzlich zwei neue Maße vorgeschlagen:

Die 'Diffusefield Transfer Function' (DTF) als Maß für die Einhüllung, welche sich gleichermaßen für kleine und große Räume eignet, sowie das 'Average Interaural Time Delay' (AITD, etwa ‚durchschnittliche interaurale Zeitverzögerung‘), als Maß für die 'Externalisation' (also ‚Externalisierung‘ / ‚Außer-Kopf Ortung‘), eine Klangeigenschaft, die – laut Griesinger - nur bei kleinen Räumen auftritt.

Wie aus dem obenstehenden Abschnitt hervorgeht, setzt sich der subjektive Höreindruck der Räumlichkeit aus den Elementen ASW und LEV zusammen, welche durch Reflexionen des Direktschalls hervorgerufen werden, die in unterschiedlichen Zeitabschnitten und aus verschiedenen Raumwinkeln eintreffen. Beide bedingen, dass der am Kopf des Hörers auftretenden IACC (aufgrund des seitlich einfallenden Schalls) verringert wird, wobei sich im Zusammenhang mit der Qualitätsbewertung der Akustik von Konzertsälen gezeigt hat, dass die Hörerpräferenz mit abnehmendem IACC (d.h. wachsender Dekorrelation) steigt.

In diesem Zusammenhang wurde auch der sog. 'Binaural Quality Index' BQI definiert, welcher sich erstmals bei [Keet, 1968] findet. Dieser ist laut [Beranek, 2004] „... one of the most effective indicators of the acoustic quality of concert halls.“ (ebendort, S. 506) und definiert, wie folgt:

$$\text{BQI} = (1 - \text{IACC}_{\text{E3}}) \quad (2.10)$$

Der Subindex ‚E3‘ steht hierbei für den frühen Anteil der Schallenergie im Zeitbereich von 0 bis 80ms in den Oktavfrequenzbändern mit Mittenfrequenzen bei 500Hz, 1kHz und 2kHz.

Die nachfolgende Abbildung 2.14 zeigt eindrucksvoll den Zusammenhang zwischen BQI und der subjektiven Präferenz:

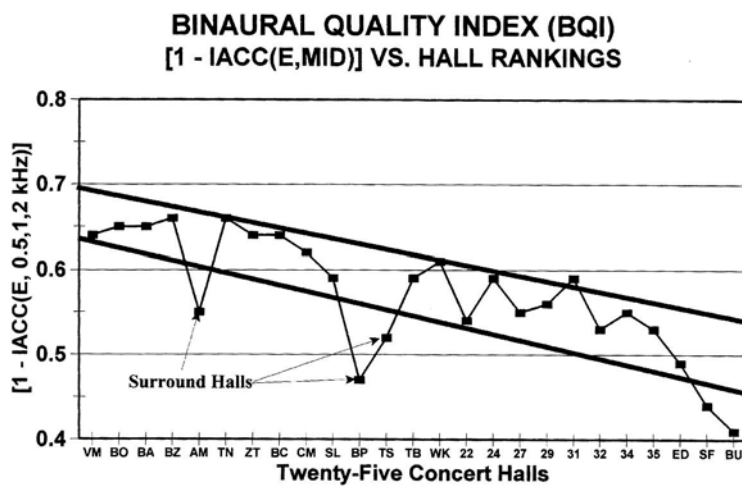


Abb. 2.14: Binaural Quality Index (BQI) von 25 Konzertsälen, gemessen in unbesetztem Zustand, aufgetragen über der subjektiven Rangordnung gemäß ihrer akustischen Qualität (durchschnittliche Standard-Abweichung 0,11 sec) (übersetzt aus [Beranek, 2004])

2.5 Einflüsse der Eigenschaften von Lautsprecher und Akustik des Abhörtraumes auf die Hörerpräferenz

Wenn im Rahmen eines Hörvergleichs die Qualität und unterschiedlichen Aspekte von Mikrofonierungstechniken beurteilt werden sollen, kann davon ausgegangen werden, dass auch die Eigenschaften der zur Signalwiedergabe verwendeten Lautsprecher, sowie die akustischen Eigenschaften des Wiedergaberaumes Einfluss auf das Ergebnis haben. Veränderungen bei der Signalübertragung von der Informations-Quelle (z.B. Musikinstrument) zur Informations-Senke (Ohr des menschlichen Hörers) ergeben sich gezwungenermaßen u. a. auch durch die Beschränkungen des technischen Übertragungsweges, z.B. in Form qualitativer Grenzen der verwendeten elektroakustischen Wandler (Mikrofone, Lautsprecher) sowie der dazwischen liegenden Speicher oder Übertragungskanäle.

Im Zusammenhang mit der Wahl der Lautsprecher stellt sich beispielsweise die Frage, ob deren Richtcharakteristik einen Einfluss auf die von den Testhörern abgegebene Bewertung hat. Unterschiedliche Forscher scheinen hier zu abweichenden Ergebnissen zu kommen: [Kates, 1960] ist ein Befürworter von stärker gerichtet abstrahlenden Lautsprechern, da er die Entstehung von Raumreflexionen möglichst vermeiden möchte. Zu einer ähnlichen Schlussfolgerung kommt [Zacharov, 1998], was allerdings auf starken Widerspruch bei anderen gewichtigen Vertretern der Audio-Branche stößt (siehe hierzu u.a. [Holman, 2000])

Letzterer verweist hierbei u.a. auf die Ergebnisse seiner Untersuchungen, die er in [Holman, 1991] erläutert und aus denen hervorgeht, dass eine stärkere Richtwirkung von Lautsprechern sich zugunsten der Lokalisationsgenauigkeit äußert, geringere Richtwirkung allerdings zur Erzielung einer besseren Einhüllung beim Hörer (vor allem für Surround-Lautsprecher) von Vorteil ist.

Ebenso kommt [Toole, 2008] (S. 137) - basierend auf einer erneuten Analyse der von ihm bereits früher durchgeführten Hörvergleiche mit Testhörern [Toole, 1985, 1986] – zum Schluss, dass die Mehrzahl der Hörer offenbar Lautsprecher mit breiterer Richtcharakteristik bevorzugt. Dies hat allerdings vor allem auch damit zu tun, dass in den von ihm verwendeten Wiedergaberräumen das breitere Abstrahlverhalten Reflexionen an den (akustisch nicht bedämpften) Seitenwänden hervorrief, welche zu

einer Verringerung des IACC und einer Erhöhung der ASW bei den Testhörern führte. Dies schlug in Form einer gesteigerten Hörerpräferenz für diese Lautsprecher zu Buche.

Dass größere Basisbreite von Hörern bevorzugt wird, ging auch aus einer Untersuchung von [Berg and Rumsey, 2001] hervor, und findet ebenso Bestätigung bei [Bech, 1998]. Im Falle der Untersuchung von Bech wurde die Variation der Basisbreite der Signalquelle rein wiedergabeseitig durch unterschiedlich große Abstände der (Stereo-) Frontlautsprecher realisiert (siehe Abb. 2.15 unten). Weitere Untersuchungen, die zu ähnlichen Schlussfolgerungen kommen, finden sich bei [Moulton et al., 1986], [Moulton, 1995], sowie [Augspurger, 1990].

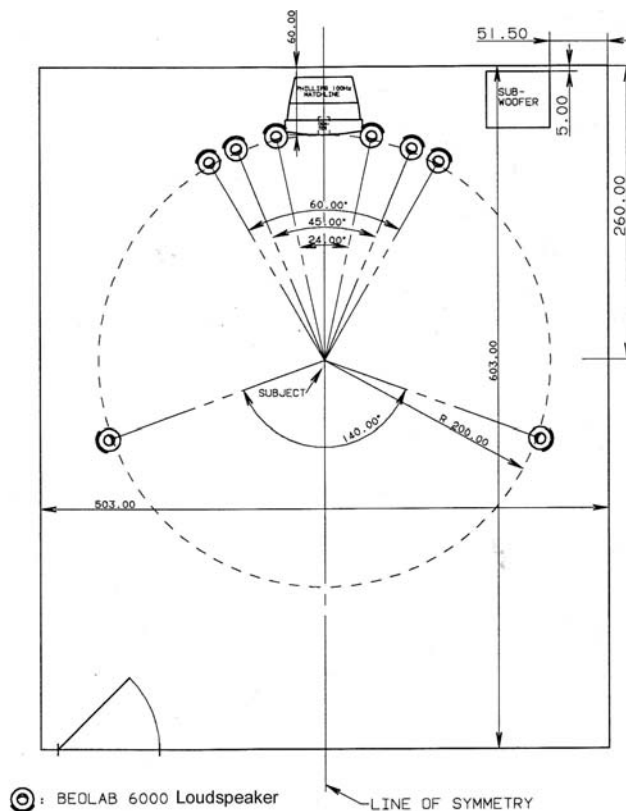


Abb. 2.15: Lautsprecher- und Fernseher-Setup im Testraum aus der Untersuchung zum Einfluss der Basisbreite auf die empfundene Qualität von [Bech, 1998]

Solche durch Abstrahlcharakteristik der Lautsprecher und Raumeigenschaften bedingten Effekte (betreffend Veränderung der vom Hörer wahrgenommenen

Basisbreite ASW) sollten in dem für die vorliegende Untersuchung verwendeten Wiedergaberaum (‚Produktionsstudio‘ = Regieraum des IEM) nicht in Erscheinung treten, da der Raum akustisch adaptiert ist: die Decke ist mit sog. ‚Akustik-Elementen‘ bestückt, es befinden sich Absorberpanele an strategischen Punkten an den Seitenwänden und mobile Absorber-Stellwände wurden für den Hörvergleich so positioniert, dass Reflexionen im Raum - für die Abhörposition - möglichst unterdrückt wurden.

Abgesehen davon würden sich eventuelle, durch die Wechselwirkung von Abstrahlverhalten der Lautsprecher mit der Raumakustik bedingte Änderungen in der empfundenen Basisbreite für alle zu vergleichenden Surround-Mikrofontechniken gleichermaßen auswirken, sodass es dadurch also nicht zur Bevorzugung einer bestimmten Mikrofontechnik kommen dürfte.

Nähere Details zu den messtechnisch ermittelten Eigenschaften des Regieraums IEM finden sich im Anhang. Aus diesen geht unter anderem hervor, dass dieser hinsichtlich Raumgröße konform mit der ITU-R Recommendation BS.1116-1 (siehe [ITU-R BS.1116-1, 1997]) ist, und auch seine Nachhallzeit bis zu einer unteren Grenzfrequenz von 90 Hz im vorgeschriebenen Toleranzbereich liegt.

2.6 Psychoakustische Effekte bei Signalwiedergabe über Lautsprecher betreffend Lokalisation und Räumlichkeit

Bei Klangwiedergabe über eine Standard Stereo-Lautsprecheranordnung (bei -30° und $+30^\circ$) führen entsprechend große Pegel- oder Zeitdifferenzen zwischen den Signalen des linken und rechten Kanals zur Lokalisation eines Mono-Stimulus in lediglich einem der beiden Lautsprecher. Hierfür ist entweder ein Pegelunterschied von ca. 15dB bzw. ein Zeitunterschied von 1.1ms notwendig (siehe Abb. 2.16; nach [Simonsen, 1984]). Erste Untersuchungen hierzu – allerdings mit davon abweichenden Werten - finden sich bereits bei [de Boer, 1940], in späteren Untersuchungen bei [Theile et al. 1988]], sowie [Gernemann, 1994]. [Sengpiel, 1992] kommt aufgrund seiner Untersuchungen zu Differenzen von 18dB im Pegel bzw. 1.5ms betreffend Laufzeit.

Bei Vergleich der Ergebnisse der verschiedenen Untersuchungen lässt sich feststellen, dass die Abweichung der gefundenen Werte bei laufzeitbasierter Lokalisation wesentlich größer ist, als bei pegeldifferenzenbasierter. Die großen Diskrepanzen für laufzeitbasierte Stereophonie gehen auch aus einer vergleichenden Studie von [Wittek und Theile, 2002] hervor.

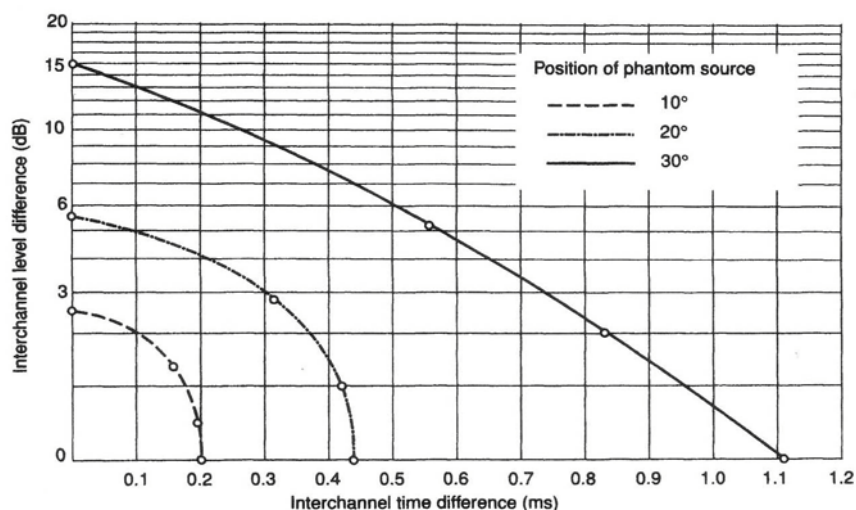


Abb. 2.16: Lokalisationskurven nach Williams/Simonsen (aus [Rumsey, 2001])

Wie bereits anhand der am Beginn der Einleitung gezeigten Tabelle 1.1 gezeigt werden konnte, betragen diese Unterschiede teilweise mehr als 100% (!) (man vergleiche die Ergebnisse von Wittek und Sengpiel für das Klein-AB System mit einem Kapselabstand von 50cm).

Bei [Knothe und Plenge, 1978] werden verschiedenen Studien genannt, welche belegen, dass der Effekt der Frequenzabhängigkeit der Schallquellen-Lokalisation zumindest seit 1934 bekannt ist. Untersuchungsergebnisse jüngerer Datums zu diesem Thema finden sich u.a. bei [Griesinger, 2002].

Ursachen für die großen Unterschiede zwischen den in der Studie [Wittek and Theile, 2002] einander gegenübergestellten Ergebnissen mögen folgende sein¹:

- Unterschiede bei den Testsignalen,
- unterschiedliche akustischen Charakteristiken der bei den Tests verwendeten Lautsprecher (siehe zu dieser Problematik [Gernemann, 1998]),
- große interindividuelle Unterschiede in der Hörwahrnehmung der Testpersonen, gepaart mit Lokalisationsverzeichnungen, die im Bereich mittlerer Frequenzen auftreten (siehe [Benjamin and Brown, 2007], sowie Abb. 2.17 unten)

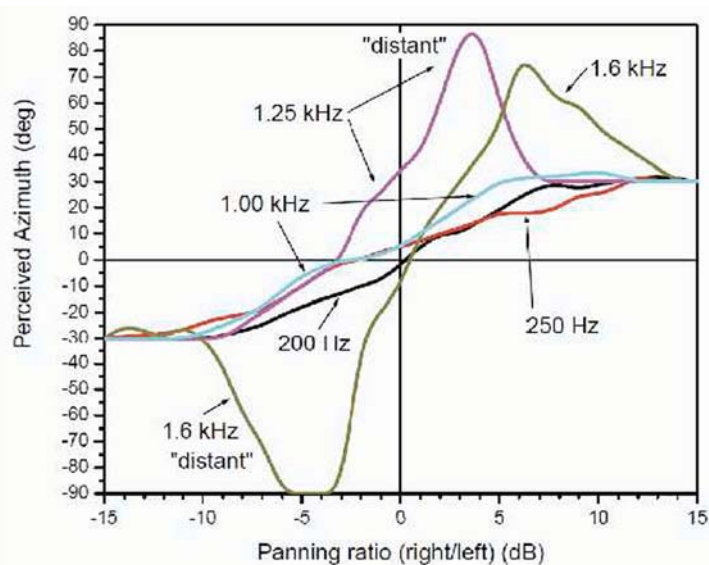


Abb. 2.17: Wahrgenommener Azimut bei verschiedenen Frequenzen für Signale (schmalbandige Gauß'sche Sinus-Bursts), die mittels unterschiedlicher Pegel zwischen den Stereo-Kanälen panoramisiert wurden (aus [Benjamin, 2006])

¹ Möglicherweise ist eine Kombination aus mehreren der oben genannten Faktoren für die Diskrepanz in den Ergebnissen verantwortlich, was ein lohnender Gegenstand weiterer Forschungsarbeit sein könnte.

Betrachtet man die in der Studie [Wittek and Theile, 2002] aufgezeigten Unterschiede von mehr als 100%, so scheint es klar, dass ‚Klein-AB‘ basierte Aufnahmetechniken nicht über Lokalisationseigenschaften verfügen, die stabil genug sind, um interindividuelle Unterschiede bei den Testhörern (wie auch qualitative Unterschiede bei den Lautsprechern) vernachlässigen zu dürfen.

In Experimenten, die die Lokalisation von Signalquellen in der nach [ITU-R BS.775-1, 1994] definierten 5.1 Surround Lautsprecher-Anordnung betrafen, fanden [Martin et al., 1999] heraus, dass bei den Rear-Lautsprechern bereits ein Laufzeitunterschied von nur 0.6ms ausreicht, um die Signalquelle gänzlich in einem der beiden Lautsprecher zu lokalisieren. Dies ist vermutlich auf den größeren Abstand bzw. Öffnungswinkel von 120° der Lautsprecher (im Gegensatz zu 60° bei den Frontlautsprechern) relativ zum Hörer zurückzuführen (siehe hierzu [Rumsey, 2001], S. 32). Eine weitere Schlussfolgerung aus dieser Arbeit war, dass pegeldifferenzbasierte Lokalisation stabilere Eindrücke beim Hörer hervorruft, als laufzeitbasierte.

2.6.1 Frequenzabhängige Lokalisationsverzerrungen in der horizontalen Ebene

Wie u.a. in einer Untersuchung von [Knothe und Plenge, 1978] festgestellt wurde, ist die Lokalisation einer mittels Panoramareglers (Richtungsstellers) entlang der Basisbreite einer Stereo-Lautsprecheranordnung panoramisierten Mono-Signalquelle von deren Frequenz abhängig. (siehe hierzu Abb. 2.18):

Die durch den Panoramaregler auf dem Summenbus erzielte Kanalpegel-Differenz ΔS muss zur Erzielung einer bestimmten Richtungsempfindung des Signalstimulus für tiefe Frequenzen größer sein, als für hohe Frequenzen. Je weiter die seitliche Auslenkung des Richtungsreizes auf der +/- 30 Grad Basis der Standard Stereo-Lautsprecheraufstellung sein soll, desto größer wird auch der - für eine korrekte Lokalisation über den gesamten Frequenzbereich - notwendige Unterschied in der Größe von ΔS :

bei einer gewünschten Auslenkung des Richtungseindrucks von 10 Grad liegt der Unterschied von ΔS zwischen hohen und tiefen Frequenzen bei etwa 1-2 dB, bei einer Auslenkung von 25 Grad beträgt die notwendige Abweichung bereits - je nach Frequenz - bis zu 5 dB (vergleiche ΔS bei 400Hz und 7 kHz)

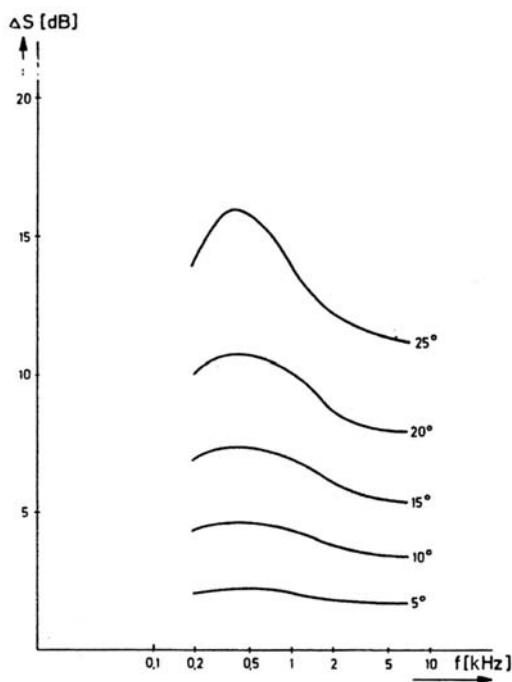


Abb. 2.18: frequenzabhängige Lokalisation eines Klangreizes in Abhängigkeit vom Pegelunterschied ΔS zwischen linkem und rechtem Stereokanal (aus [Knothe und Plenge, 1978])

[Griesinger, 2002] kommt bei einer Untersuchung dieser Problematik hinsichtlich des Verhaltens von 2-Kanal-Stereo und Surround-Signalen ebenfalls zu dem Ergebnis, dass die Lokalisation von Klangreizen stark von deren Frequenzspektrum abhängig ist. Dabei wandert die Lokalisation für Spektralanteile, die höher als 600Hz sind überproportional rasch in Richtung des Lautsprechers, auf den hin panoramisiert wird. Dies ist gemäß seinen Untersuchungen auf die Interferenz zwischen dem Direktsignal aus einem Lautsprecher, das sich mit dem am Kopf des Hörers gebeugten Signal aus dem anderen Lautsprecher überlagert, zurückzuführen.

2.6.2 Frequenzabhängige Lokalisationsverzerrungen in der vertikalen Ebene

Wie die Untersuchungen von [Ferguson and Cabrera, 2005] gezeigt haben, tritt auch in der vertikalen Ebene eine Verzerrung bezüglich der Lokalisation von Klangreizen bei hohen und tiefen Frequenzen auf: während hochfrequente Schallquellen generell an ihrer tatsächlichen physikalischen Position lokalisiert werden, trifft dies für tieffrequente Schallquellen nicht zu. Diese werden normalerweise unterhalb ihrer tatsächlichen physikalischen Position wahrgenommen. Im Falle eines breitbandigen Signals wird die Lokalisation der Schallquelle primär durch die hohen Frequenzanteile bestimmt (siehe hierzu auch [Morimoto et al., 2003]).

2.6.3 Effekte betreffend die Wiedergabe des Räumlichkeitseindrucks

In der Praxis führt die im obenstehenden Abschnitt angesprochene frequenzabhängige Lokalisation von Klangreizen dazu, dass somit unter Umständen auch die räumliche Abbildung eines nicht nur von seinen physikalischen Abmessungen, sondern auch vom Frequenzspektrum her breiten Instruments (z.B. Klavier) verzerrt erfolgt: die sich für den menschlichen Hörer ergebende Stereoabbildung erfolgt für tiefen Frequenzen tendenziell enger, als für hohe, sofern aufgrund der verwendeten Aufnahmetechnik rein pegelbedingte Panoramisierung erfolgt; was z.B. bei koinzidenten Mikrofontechniken wie dem ‚Blumlein-Paar‘, bestehend aus zwei gekreuzten Mikrofonen mit Achter-Charakteristik, der Fall ist.

In diesem Zusammenhang wurden Methoden zur ‚räumlichen Filterung‘ (eng. ‚spatial equalisation‘) von verschiedenen Forschern (siehe u.a. [Gerzon, 1986], sowie [Griesinger, 1986]) vorgeschlagen.

Die Änderung in punkto Raumabbildung zu den tiefen Frequenzen hin kann überprüft werden, indem man selektiv isolierte Frequenzbänder abhört, um zu analysieren, welchen Raumeindruck ein Hauptmikrofonsignal in den verschiedenen Frequenzbereichen erzeugt. Besonders wichtig ist in diesem Zusammenhang der Frequenzbereich unterhalb von ca. 700-800 Hz, da der menschliche Kopf für tiefe Frequenzen noch nicht als Trennkörper wirksam ist und diese Schallanteile um ihn herum gebeugt werden. Oberhalb von ca. 800 Hz wirkt sich die Abschattung von Schallquellen zunehmend stärker aus, und der menschliche Gehörmechanismus basiert hauptsächlich auf interauralen Differenzen in Bezug auf Signalpegel, während er bei tiefen Frequenzen hauptsächlich auf einer Analyse der Phasen- und Zeitdifferenzen beruht (siehe auch die entsprechende Abbildung A1 im Anhang, aus der Untersuchung von [Steinberg and Snow, 1934]).

In diesem Zusammenhang seien die Untersuchungen von [Yost et al., 1971] genannt, welche gezeigt haben, dass die tieffrequenten Anteile von transienten Binaural-Signalen von größter Bedeutung für die Lokalisation sind: die Hochpass-Filterung von Clicks (ähnlich Dirac-Impulsen) mit einer Grenzfrequenz von 1500 Hz brachte eine eindeutige Verschlechterung der Lokalisation, während eine Tiefpass-Filterung mit der selben Grenzfrequenz nur eine geringe Änderung mit sich brachte.

Untersuchungen von [Hirata, 1983] beschäftigen sich mit dem Phänomen der Lokalisationsverschlechterung der tieffrequenten Signalkomponenten eines Stereosignals, sobald es über Lautsprecher wiedergegeben wird. Er schlägt die Begriffsbestimmung eines sog. ‚Perzeptuellen Interauralen Kreuzkorrelations Koeffizienten‘ (engl. ‚perceptual interaural cross-correlation coefficient‘ - PICC) vor:

$$\text{PICC} = D R_0 + (1-D) R_E \quad (2.8)$$

mit D ... Deutlichkeit
 R_0 ... der interaurale Kreuzkorrelations-Koeffizient des Schall-Direktanteils (1 für Einfall von vorne)
 R_E ... der interaurale Kreuzkorrelations-Koeffizient des Diffusschall-Anteils,

ausgedrückt durch:

$$R_E = [\sin k r(f)] / k r(f) \quad (2.9)$$

wobei:

$$k = 2 \pi f / c \quad \dots \text{ der Wellen-Vektor,} \quad (2.10)$$

c ... die Schallgeschwindigkeit und

$r(f)$... die akustisch wirksame Distanz zwischen beiden Ohren ist,
welche etwa 30cm beträgt. (siehe [Yanagawa et al., 1976],
sowie [Suzuki and Tohyama, 1981])

Weiters definiert er einen ‚Index des akustischen Raumeindrucks‘ (Index of Acoustic Spatial Impression - ASI), wie folgt:

$$ASI = (1 - D) * 100 \quad [\%] \quad (2.11)$$

Ein vollkommener Raumeindruck wird als ASI=100% gekennzeichnet, das komplette Fehlen eines Raumeindrucks hingegen mit ASI=0%.

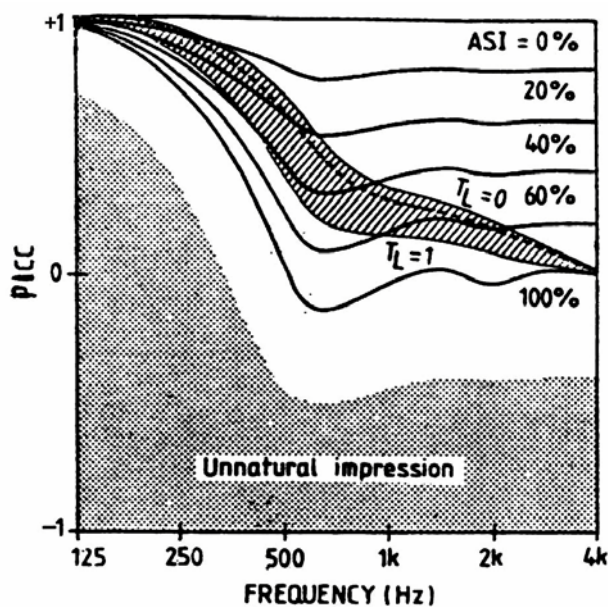


Abb. 2.19: Die PICC-Kurven für Stereo-Klangwiedergabe in einem Hörraum mit einer Nachhallzeit T_L (0 bis 1s) zeigen kleine ASI-Werte bei tiefen Frequenzen, verglichen mit einem ASI=60% für die Sitze im mittleren Bereich eines Konzertsaaes. Die unterbrochene Kurve steht für den Wert $T_L=0.3s$. (aus [Hirata, 1983])

Die obenstehende Abbildung zeigt, dass in einem Standard-Abhörraum ($RT_{60}=0.3s$) der ASI im stereophonen Klangfeld für Frequenzen unter 800Hz gering, und für

Frequenzen oberhalb von 800Hz groß im Vergleich zum Eindruck an einem Platz im Konzertsaal ist, bei dem ASI z.B. bei 60% liegt.

Untersuchungen von Griesinger zur Erzielung guter Hörer-Einhüllung [Griesinger, 1999] haben ergeben, dass sich zu hohen Frequenzen hin die ideale Lautsprecherposition zur Median-Ebene verschiebt. Für Frequenzen unterhalb von 700Hz hingegen ist eine Lautsprecheraufstellung ideal, welche eine möglichst große Trennung zwischen den Schallwandlern bietet, d.h. links und rechts vom Hörer bei +/- 90 Grad.

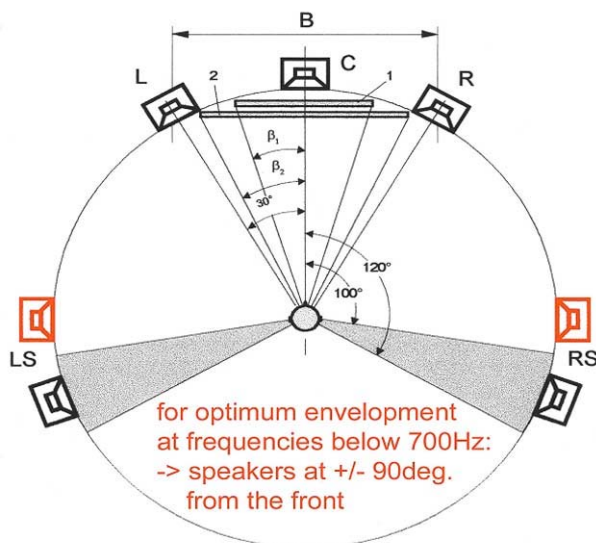


Abb. 2.20: ‚5.2‘ Anordnung mit Subwoofer-Lautsprechern (rot) an der Seite bei +/-90 Grad, optimiert für räumliche Wiedergabe tiefer Frequenzen, gemäß Empfehlung [Griesinger, 1999]

Die Untersuchungen von Hirata und Griesinger zeigen, dass die Standard-Lautsprecheraufstellung für Stereo (sowie auch für 5.1 Surround, siehe ITU-Spezifikation gemäß Richtlinie [BS.775-1]) mit Lautsprecherpositionen bei +/- 30 Grad in Bezug auf die Räumlichkeitsabbildung alles andere als ideale Bedingungen für getreue Signalwiedergabe tieffrequenter Anteile bietet.

Aus diesem Grund ist es wichtig, bereits bei der Aufnahme eine (Surround-) Mikrofontchnik zu wählen, welche die Signale über den gesamten Frequenzbereich möglichst dekorreliert erfasst, da wiedergabeseitig diesbezüglich ohnehin mit einer Verschlechterung zu rechnen ist.

3 Korrelation und Kohärenz

Eine detaillierte Analyse betreffend die Unterschiede zwischen Signal-Korrelation und Signal-Kohärenz findet sich - neben anderen - bei [Cremer, 1976], eine kurze Zusammenfassung bei [Martin, 2005].

Aus diesem Grund soll an dieser Stelle nur kurz die Definitionen beider Begriffe angegeben werden:

Mathematisch ist die Kreuz-Korrelation zwischen zwei Signalen $s_x(t)$ und $s_y(t)$ definiert als:

$$\Phi_{xy}(\tau) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} s_x(t) s_y(t + \tau) dt \quad (3.1)$$

wobei $s_x(t)$ und $s_y(t)$ die Schalldrücke zum Zeitpunkt (t) an zwei Messpunkten sind (vgl. [Hesselmann, 1993], S. 174).

Für verschiedene Anwendungen kann es zweckmäßig sein, anstelle der Funktion $\Phi_{xy}(\tau)$ auf die normierte Kreuzkorrelationsfunktion $\varphi_{xy}(\tau)$ zurückzugreifen:

$$\varphi_{xy}(\tau) = \frac{\Phi_{xy}(\tau)}{\sqrt{\Phi_{xx}(0) \Phi_{yy}(0)}} \quad (3.2)$$

Per Definition kann $\varphi_{xy}(\tau)$ nur Werte zwischen -1 und 1 annehmen: eine Korrelation von 1 bedeutet, dass beide Signale ident sind, bei einer Korrelation von -1 hingegen ident, aber von entgegengesetzter Polarität (180° oder 'gegenphasig'), und 0 bedeutet, dass die beiden Signale nicht korreliert sind.

Für Anwendungen im Bereich der digitalen Signalverarbeitung ist es wichtig, dass es eine eindeutige Beziehung der Kreuzkorrelationsfunktion zu spektralen Systemfunktionen gibt. Im Frequenzbereich ist diese Systemfunktion das sogenannte Kreuzleistungsdichtespektrum $\underline{W}_{xy}(\omega)$ (siehe [Hesselmann, 1993], S. 175):

$$\underline{W}_{xy}(\omega) = \int_{-\infty}^{\infty} \Phi_{xy}(\tau) e^{-j\omega\tau} d\tau \quad (3.3)$$

Formal ist die von der Frequenz abhängige Kohärenzfunktion definiert als:

$$\underline{\gamma}_{xy}(\omega) = \frac{|\underline{W}_{xy}(\omega)|}{\sqrt{W_{xx}(\omega) W_{yy}(\omega)}} \quad (3.4)$$

Die Kohärenz $\underline{\gamma}_{xy}(\omega)$ ist eine Funktion der spektralen Leistungsdichte $W_{xx}(\omega)$ und $W_{yy}(\omega)$ der Signale x und y und der spektralen Kreuzleistungsdichte $\underline{W}_{xy}(\omega)$ von x und y. (vgl. [Hesselmann, 1993], S. 176)

Für die Messungen im Rahmen dieser Arbeit wurde die 'mscohere'-Funktion in MATLAB verwendet, welche eine Kohärenz-Schätzung unter Anwendung der Methode nach Welch (gemittelte modifizierte Periodogramm-Methode) verwendet. Das Resultat dieser auf $|\underline{\gamma}_{xy}(\omega)|^2$ basierenden Berechnung liefert Werte zwischen 0 und 1, die angeben, wie ähnlich sich x zu y bei einer gegebenen Frequenz verhält.

3.1 Signalkorrelation bei Stereo-Mikrofonsystemen

Die Signale der beiden Kanäle eines Mikrofonsystems für Stereoaufnahme ergeben sich aus der Summierung des an den entsprechenden Raumpunkten der Mikrofone auftretende Direkt- und Diffusschalls. Demgemäß ist die Zusammensetzung des Stereosignals u.a. abhängig vom Abstand zwischen der Schallquelle und dem Mikrofonsystem (Verhältnis Direkt-/Diffusschall), der akustischen Beschaffenheit des Raumes (Absorptionskoeffizienten der Begrenzungsflächen, sowie deren Oberflächenbeschaffenheit; Stichwort: Absorption, Reflexion, Diffusion), sowie der relativen Position von Schallquelle und Mikrofon im Raum.

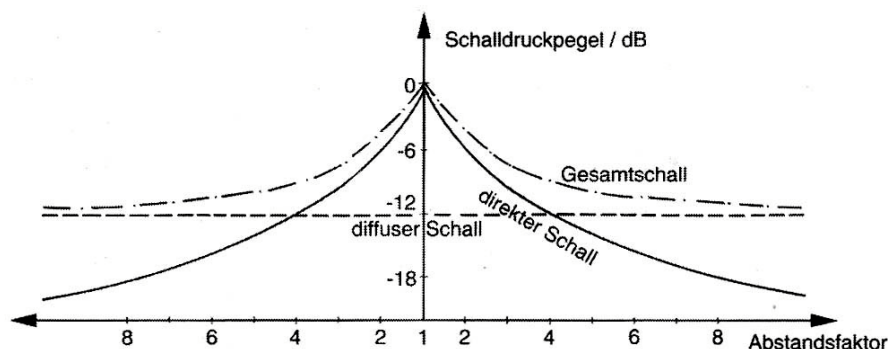


Abb. 3.1: Zusammenhang zwischen Gesamtschallpegel, Direkt- und Diffusschall (aus [Wuttke, 1981])

Eine frühe Untersuchung zur Korrelation der Signale zweier Druckempfänger im diffusen Schallfeld (Hallkammer) findet sich bei [Cook et al., 1955], also in etwa zu der Zeit, als auch Plattenfirmen seit kurzem begonnen hatten, sich der neuen Herausforderung der Stereo-Tonaufnahmetechnik zu widmen.

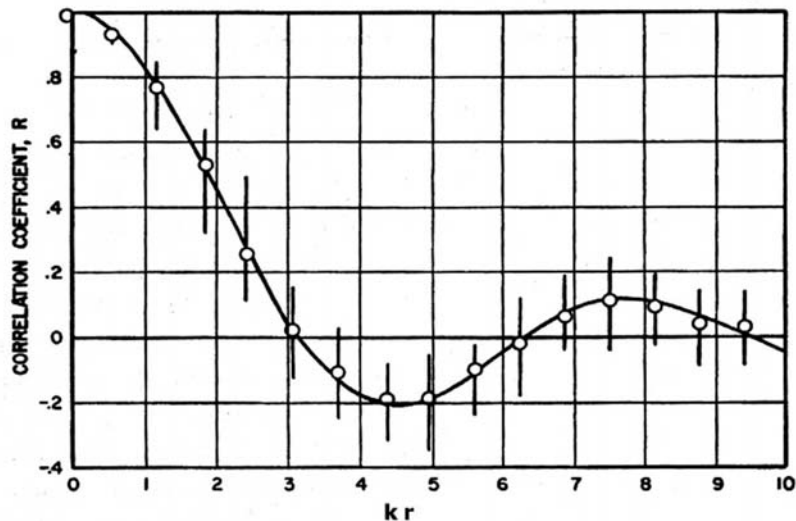


Abb. 3.2: Korrelations-Koeffizient R aufgetragen über dem Kapselabstand zweier Mikrofone; ___ = Berechnung; o = Messung; (nach [Cook et al., 1955])

Zur Erklärung der in der obenstehenden Abbildung verwendeten Symbole:

$$R = (\sin k r) / k r \quad (3.3)$$

mit: $k = 2 \pi / \lambda$ (k = Wellen-Vektor; λ = Wellenlänge)

und somit

$$k r = 2 \pi r / \lambda \quad (3.4)$$

bzw. $k = \omega / c$, da gilt $\lambda = c / f$ (c ... Schallgeschwindigkeit, d.h. ca. 340m/s)

Die komplette Herleitung der Ergebnisse ist zu finden bei [Elko, 2001].

Ebendort finden sich auch Berechnungen zur Signal-Kohärenz verschiedener Mikrofon-Arrays, die durch die Kombination von jeweils zwei Kapseln mit unterschiedlicher Richtcharakteristik und verschiedener Ausrichtung (0° , 90° , 180° zueinander) charakterisiert sind. Einige Beispiele sind nachfolgend dargestellt:

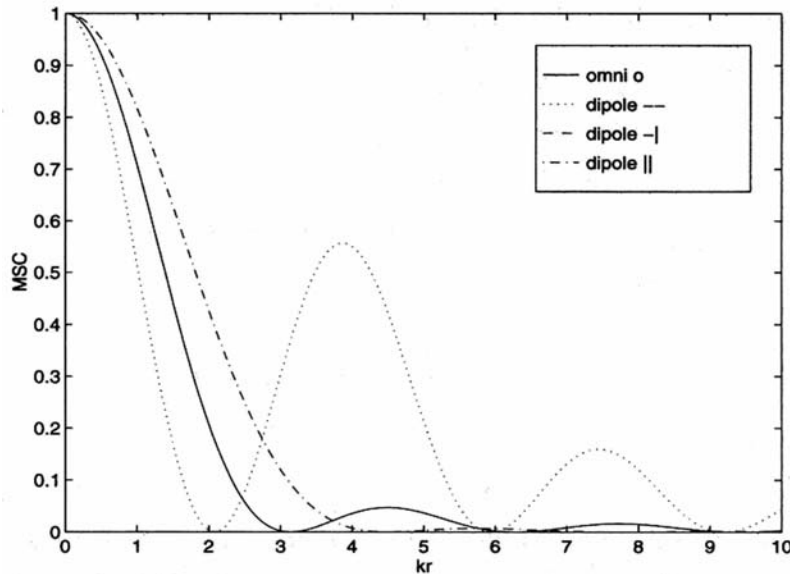


Abb. 3.3: Magnitude-squared coherence (MSC) für omnidirektionale (Kugeln) und bidirektionale (Achten) Mikrofone in einem kugelförmig isotropen Schallfeld. Man beachte, dass die Kurve für die orthogonalen Achten (Dipole) mit der Abszisse zusammenfällt. (aus [Elko, 2001])

Während die orthogonal zueinander stehenden Achten (Dipole) einer Blumlein-Paar Anordnung entsprechen, sind die zueinander parallel stehenden Achten (symbolisch dargestellt als ||) in der Praxis als ‚Faulkner Phased Array‘ bekannt. Eine Anordnung von um 90° gedrehten, nebeneinander angeordneten Achten (quasi ‚seriell‘), stellt eine Art halben ‚Hamasaki Square‘ dar (in der Abbildung oben symbolisch dargestellt als --).

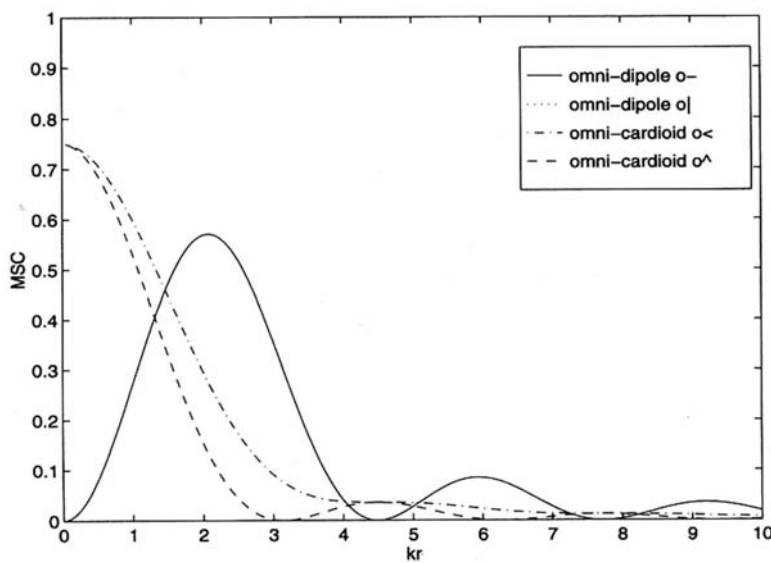


Abb. 3.4: Magnitude-squared coherence (MSC) für omnidirektionale, bidirektionale (dipole) und Nieren-Kapseln in einem kugelförmig isotropen Schallfeld. Die Kurven für die 2 Kombinationen ‚omni-dipole‘ fallen zusammen. (aus [Elko, 2001])

Die in Abbildung 3.4 dargestellte Kombination von Omni und Acht entspricht beispielsweise einem MS-Stereomikrofon mit Kugelcharakteristik für das Mittelmikrofon.

In [Pfanzagl-Cardone, 2002] konnte anhand einer empirischen Untersuchung gezeigt werden, dass sich bei Verwendung derselben Mikrofontechnik in akustisch unterschiedlichen Räumen dennoch für die jeweilige Mikrofontechnik charakteristische Muster in der Signalkorrelation über der Frequenz feststellen ließen.

In einer späteren Arbeit wurde bei [Muruoka et al., 2007] die frequenzabhängige Kreuz-Korrelation des Signals verschiedener Stereo-Mikrofontechniken (AB, Wavefront [=ähnlich ‚Mikrofonvorhang‘, bestehend aus einem inneren und einem äußeren AB-Mikrofonpaar], ORTF, XY und MS) für unterschiedliche Kapselabstände berechnet:

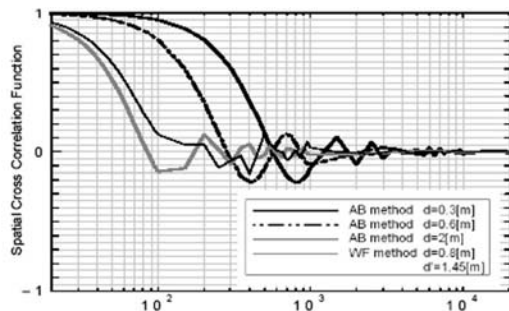


Fig. 7 FSCC-patterns of AB-microphones and WF microphone

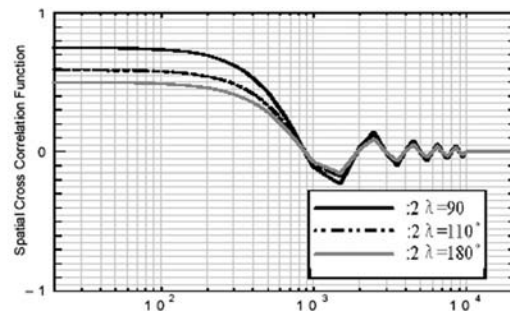


Fig. 9 FSCCs of ORTF-microphones (directional azimuth dependent)

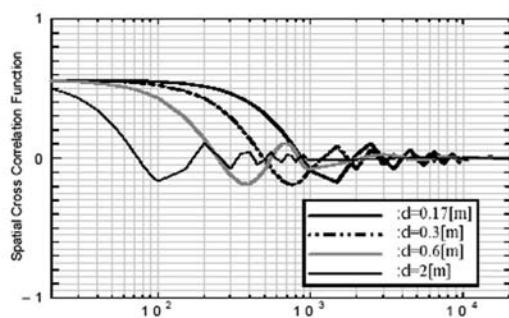


Fig. 8 FSCCs of ORTF-microphones (microphone spacing dependent)

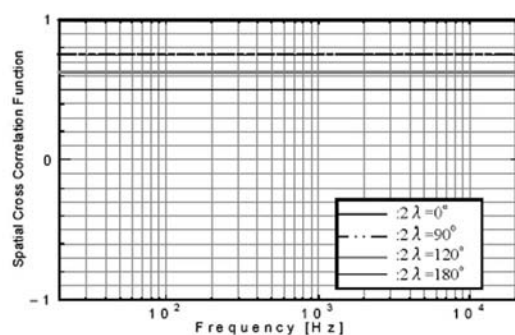


Fig. 10 FSCCs of XY-microphone

Abb. 3.5: Fig 7-10 aus [Muraoka et al., 2007] – Kreuz-Korrelationsfunktion verschiedener Stereo-Mikrofontechniken basierend auf Berechnung (Anm.: FSCC = Frequency-dependent Spatial Cross Correlation)

Die durch den Autor bei Messung von realen Aufnahmen ermittelten Kreuz-Korrelationen diverser Stereo-Mikrofontechniken (siehe Abschnitt 10.7: Frequenzabhängige Kreuz-Korrelation (FCC) bei Stereo-Hauptmikrofonsignalen) weichen von den auf rein theoretischer Basis berechneten vor allem deshalb ab, weil:

- 1) Richtcharakteristiken realer Mikrofone niemals ideal sind (d.h. sich in Abhängigkeit von der Frequenz praktisch immer eine Abweichung von der gewünschten Richtcharakteristik ergibt; siehe in diesem Zusammenhang z.B. die steigende Gerichtetheit von Kapseln mit Omni-Charakteristik zu höheren Frequenzen), und
- 2) die akustische Beschaffenheit des jeweiligen Raumes die Charakteristik des Diffusschalls bestimmt und somit ebenfalls einen Einfluss auf die sich ergebende Signal-Korrelation hat.

Die Berechnung der effektiven Aufnahmewinkel bei Stereo- (und Surround-) Mikrofonensystemen beziehen sich naturgemäß in erster Linie auf den Direktanteil des Schalls. Die Frage, wie sich das jeweilige Mikrofonensystem in Bezug auf die Aufnahme von Diffusschall verhält, wurde bisher weit seltener in die Überlegungen miteinbezogen.

Während Direktschall also relativ leicht durch den Einsatz von gerichteten Mikrofonen mit dazwischen liegendem Öffnungswinkel in (frequenzabhängig mehr oder weniger) ‚dekorrelierter‘ Weise aufgenommen werden kann, ist es für eine (im interessierenden Frequenzbereich) in ähnlicher Weise ‚dekorrelierte‘ Aufnahme von Diffusschall notwendig, dass die Mikrofonkapseln entweder starke Richtcharakteristik (Hyperniere, Acht) in Kombination mit einem entsprechenden Öffnungswinkel aufweisen (z.B. 90° bei Blumlein-Paar), oder - bei Verwendung reiner Druckempfänger – der Abstand der beiden Kapseln größer ist als der Hallradius.

In der Praxis verändert sich die Richtcharakteristik vieler Mikrofone mit der Frequenz (‚Nieren‘ werden zu tiefen Frequenzen hin oftmals eher zu ‚Kugeln‘), sodass auch die Aufnahme des Diffusschalls bei tiefen Frequenzen nicht mehr dekorreliert erfolgt. Für die Ausbildung eines guten Raumeindrucks ist es erforderlich, dass Signalanteile unterhalb 500Hz (und besonders unterhalb 200Hz) möglichst dekorreliert sind. Somit kann es passieren, dass Stereo-Mikrofonensysteme mit geringen Kapselabständen im Bereich von einigen kHz eine befriedigende stereophone Abbildung (ASW) einer Schallquelle liefern, jedoch bei tiefen Frequenzen (vor allem im Diffusschallanteil !)

weitgehend monophon sind, was das Entstehen eines überzeugenden Raumeindrucks verhindert. Überlegungen zu diesem Thema finden sich bei [Griesinger, 1997], S. 7.

Unter dem Blickwinkel, wie dekorreliert eine Mikrofontechnik Diffusschallanteile aufnimmt, stellen jene zwei oben beschriebenen Systeme (Blumlein-Paar und Breit-AB mit einem Kapselabstand größer als der Hallradius) – im Zusammenhang mit der Erzielung eines guten Raumeindrucks als positiv zu betrachtende - Extremfälle dar, bei denen komplette De-Korrelation besteht (ideale Richtcharakteristiken der Mikrofone vorausgesetzt). Der Großteil der gängigen Stereo-Mikrofontechniken erfüllt diese Forderung hinsichtlich des Diffusschalls nicht.

In [Julestrom, 1991] findet sich eine Analyse, welche die unterschiedliche Abbildung von Direkt- und Diffus-Schallanteilen auf der Stereo-Basisbreite durch verschiedene koinzidente Mikrofonssysteme (XY bzw. MS) zum Gegenstand hat. Hierzu hat Julstrom die jeweilige XY-Technik in die MS-Technik mit äquivalenter Richtcharakteristik umgewandelt: das virtuelle Seitenmikrofon hat dabei stets die Richtcharakteristik ‚Acht‘, das virtuelle Mittenmikrofon nimmt unterschiedliche Richtcharakteristiken an; die Pegelverhältnisse zwischen Seiten- und Mittenmikrofon werden entsprechend dem zu erzielenden Aufnahmewinkel variiert, wobei in den untenstehenden Abbildungen ein ‚Spread‘ (engl. für ‚Verteilung‘, ‚Aufspreizung‘) von z.B. 25% - je nach Richtcharakteristik – einem Relativpegel des Seitensignals von -11dB (Kugel) bis -14dB (Acht), sowie ein Spread von 100% einem Relativpegel von $+3\text{dB}$ (Kugel) bis 0dB (Acht) entspricht.

In der untenstehenden Abbildung ist die Verteilung von Direkt- und Diffusschall für ‚Stereo-Surround‘ Wiedergabe für einige der gebräuchlichsten Mikrofontechniken zu sehen.

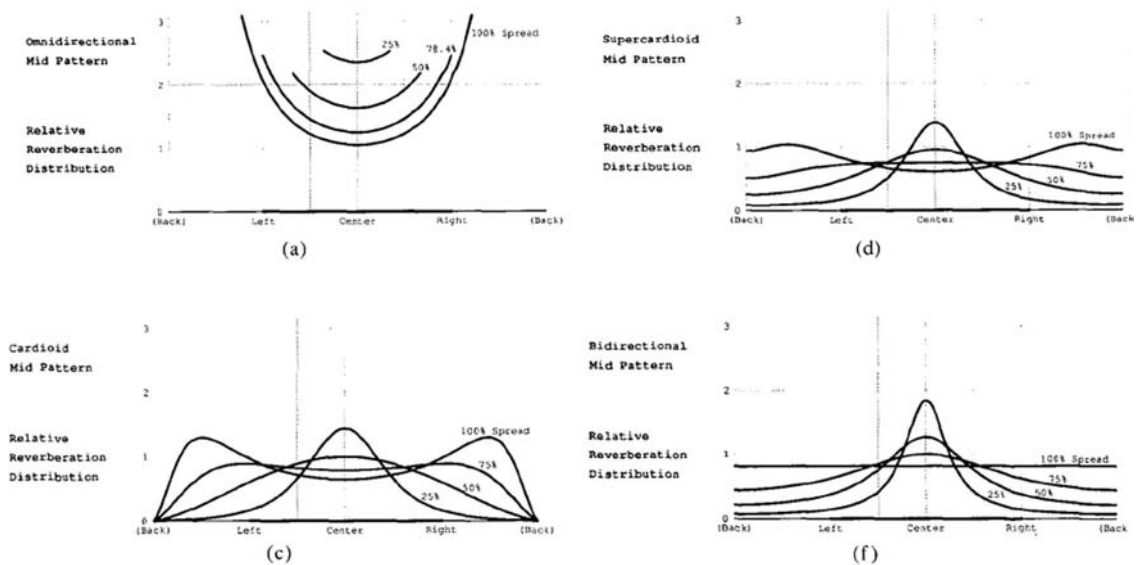


Abb. 3.6: Hallverteilung für verschiedene (MS-)Stereo-Codierungen mit Pegel der Seitensignale entsprechend der angegebenen Basisbreite („Spread“ in [%]); (Fig.15 a, c, d, f aus [Julstrom, 1991])

(Anm.: Im Falle von reiner 2-Kanal Stereowiedergabe existiert natürlich nur der innere Teil der horizontalen Skala Left-Center-Right, außen liegende Abschnitte werden entsprechend in diesen ‚hineingespiegelt‘; im Falle einer Mehrkanal-Surroundwiedergabe und dafür passender Decodierung können die entsprechenden Signalanteile auch tatsächlich in den Rear-Quadranten wiedergegeben werden)

In der obenstehenden Abb. 3.6 fallen als zwei Extreme der Hallverteilung auf:

- 1) bei der Verwendung eines Mitten-Mikrofons mit Kugelcharakteristik fallen Direktschall wie Diffusschall – aufgrund des Fehlens einer Richtcharakteristik – gemeinsam innerhalb der Frontal-Basis L, C, R an.
- 2) bei Verwendung einer Acht als Mittenmikrofon verteilt sich hierbei (als einzige Technik) der Diffusschall für den Fall des gleichen Pegels zwischen Seiten- und Mittenmikrofon ($C = 0\text{dB}$) gleichmäßig über 360 Grad. Diese Technik (Blumlein Paar) ist somit die einzige, die – für den Fall der Surround-Wiedergabe – eine naturgetreue, gleichmäßige Verteilung der Diffusschallanteile erlaubt.

Die Kombination aus zwei Achsen in Form eines MS-Paars mit einem Spread von 100% lässt sich in die traditionelle Anordnung des ‚Blumlein-Paars‘ (mit einer Kapselausrichtung bei $-45^\circ/+45^\circ$) konvertieren. Diese stellt auch insofern einen Spezialfall unter den Techniken mit starker Richtcharakteristik dar, da bei ihr das Verhältnis von gegenphasigen Schallanteilen zu gleichphasigen Schallanteilen bis zur maximalen Pegelgleichheit von Seiten- und Mittensignal stets mehr gleichphasige als gegenphasige Diffusschall-Anteile enthält (siehe Abb. 3.7 unten)

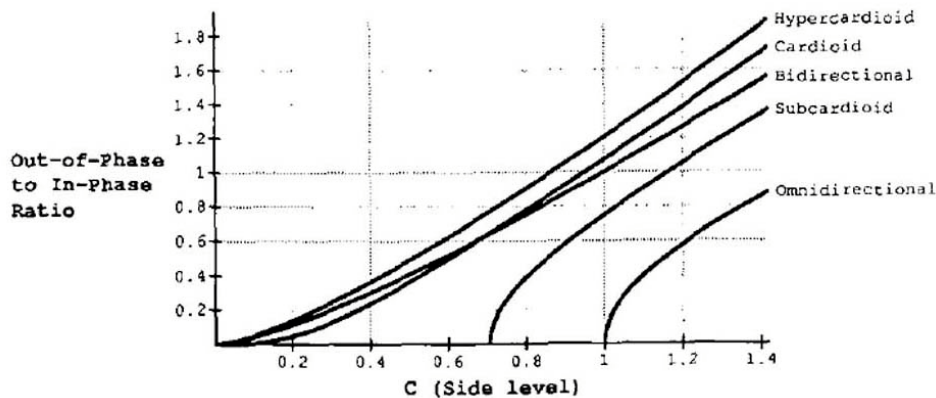


Abb. 3.7: ‚Fig 14‘ - Das Verhältnis der Aufnahme des gegenphasigen Hallanteils im Vergleich zur Aufnahme des gleichphasigen Schallanteils. (aus [Julstrom, 1991])

Nach [Julstrom, 1991] lässt sich anhand der ‚Difference-to-Sum Ratio‘ eines MS-Signals „... ein Aspekt der stereophonen Räumlichkeit der (jeweils verwendeten) Mikrofon-Richtcharakteristik vorhersagen. Bei kleinen Werten der ‚Difference-to-Sum Ratio‘ geht der Klangeindruck Richtung mono. [Griesinger, 1987] empfiehlt ein Verhältnis von ≥ 1 um geringe Korrelation zwischen den Stereo-Lautsprechersignalen zu erhalten. Bei Surround-Wiedergabe werden Signalanteile, die im Summensignal vorkommen frontal abgebildet, während Differenz-Signale tendenziell eher hinten abgebildet werden, sodass das die ‚Difference-to-Sum Ratio‘ praktisch die vorne/hinten Hall-Balance bestimmt. Es scheint in jedem Fall empfehlenswert, Verhältnisse zu vermeiden, die wesentlich größer als 1 sind ...“ (übersetzt aus [Julstrom, 1991], S. 644). (Anm.: Julstrom geht bei diesem Ansatz vermutlich noch von einer Surround-Wiedergabe aus, die – entsprechend dem alten Dolby-Surround Codierungs Algorithmus – nur einen (mono rear) Surround-Kanal vorsieht, dessen Information ja als Differenzsignal in das Dolby-Stereo Gesamtsignal codiert war.)

Die untenstehende Betrachtung des Verhältnisses von Differenz- und Summensignal zeigt, dass sich mit Richtcharakteristiken, die zwischen Niere und Acht liegen, zwar die Forderung von Griesinger nach einer ‚Difference to Sum Ratio‘ von ≥ 1 zwar erfüllen lässt, wie wir aus der obenstehenden Abbildung betreffend den ‚Outphase-to-Inphase Reverberation Pickup‘ gesehen haben, überwiegen jedoch bei allen Charakteristiken, die einen über die Niere hinausgehenden Richtfaktor aufweisen (außer Acht), für $C=1$ die Outphase-Signalanteile, was für den Höreindruck nicht von Vorteil ist, da sich zwischen dem L und R-Kanal auftretende gegenphasige Diffusschallanteile vermutlich negativ auf den Natürlichkeitseindruck einer Tonaufnahme auswirken.

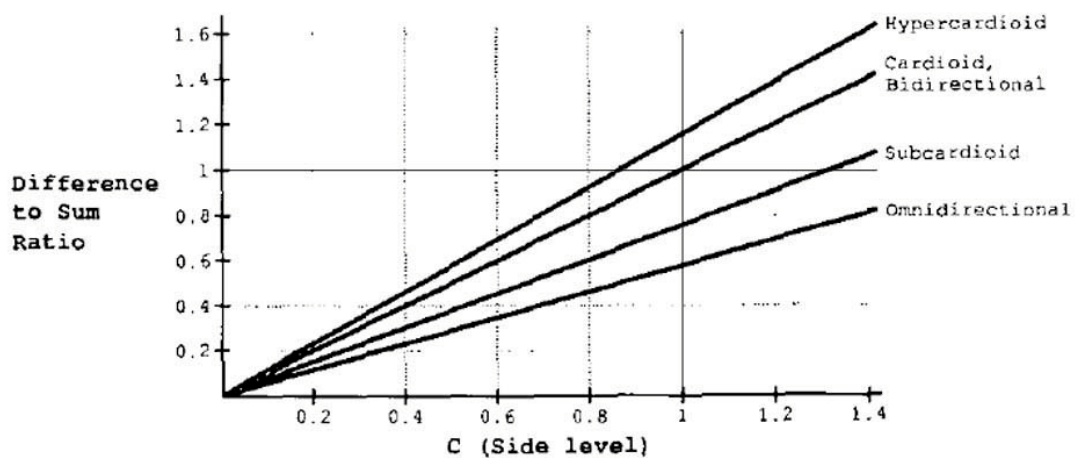


Abb. 3.8: Verhältnis des Diffusschall-Anteils (Hall) im Stereo-Differenzsignal (Seiten-Signal) zum Summen-Signal (Mitten-Signal, mono); (Fig.13 aus [Julstrom, 1991])

Entsprechend der obenstehenden Analyse geht die Mikrofontechnik ‚Blumlein-Paar‘, bestehend aus zwei unter 90 Grad gekreuzten Achten, als ein von Julstrom untersuchter Sonderfall hervor, der – insbesondere für die Aufnahme von Diffusschall – über ideale (weil naturgetreue, d.h. ‚360 Grad Panoramisierung‘) Abbildungseigenschaften verfügt.

3.2 Optimale Signal-Kohärenz bei Surround-Mikrofonssystemen

In verschiedenen Veröffentlichungen (siehe z.B. [Theile, 2000]) vermittelt Günther Theile die Idee, dass eine minimale Signal-Kohärenz von $\gamma \geq 0.35$ für ein (quadrophones) Surround-Mikrofonssystem erstrebenswert sei, um „... eine natürliche Abbildung von Raum und Einhüllung ...“ zu erreichen. (ebendort S. 419)

Er kommt zu dieser Schlussfolgerung aufgrund seiner Interpretation der Forschungsergebnisse von [Damaske, 1967]. Damaske hatte bei der Mehrzahl seiner Experimente bandbegrenzte Rausch- und Musiksignale im Frequenzbereich von 250Hz bis 2000Hz mit unterschiedlichen Kohärenzgraden verwendet. Dies diente dem Zweck, die sich für die einzelnen Testhörer ergebende ‚Diffusität‘ in Abhängigkeit von der räumlichen Aufstellung der Lautsprecher, wie auch des Kohärenzgrads zu ermitteln.

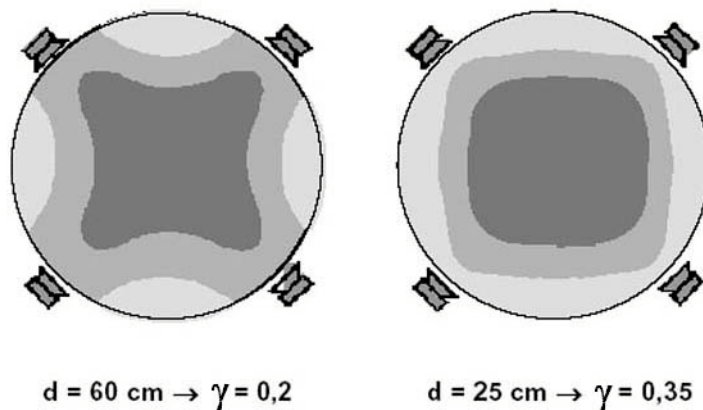


Abb. 3.9: Wahrgenommene Ereignisrichtungen in einem diffusen Schallfeld, generiert durch vier Lautsprecher, die Rauschen (0.25-2.5kHz) mit Kohärenz-Faktor γ abstrahlen. Das Rauschen wurde im Hallraum mit vier Mikrofonen abgenommen, die einen Kapselabstand d zueinander hatten (nach [Theile, 2000]).

In seinem Paper schreibt Damaske: „... einen Kohärenzgrad für Musik anzugeben ist problematisch, da dieser sicherlich von der jeweiligen momentanen Bandbreite des Musiksignals abhängt.“ Aufgrund der Ergebnisse seiner Experimente nimmt er an, dass die Kohärenz von Musiksignalen nicht weit von den Werten für bandbegrenzte Rauschsignale abweichen sollte (siehe [Damaske 1967], S. 203).

In einem anderen Teil seiner Veröffentlichung gibt Damaske an, dass die Testhörer Phantomschallquellen in inkohärenten Schallfeldern als „...aus verschiedenen Richtungen kommend ...“ beschrieben, welche nicht auf die Verbindungslinie zwischen den Lautsprechern beschränkt waren. Außerdem gaben sie an, dass der sich dabei ergebende Raumeindruck unabhängig von Kopfbewegungen und angenehmer als in einem kohärenten Schallfeld war (siehe [Damaske, 1967], S. 204).

In Anbetracht dieser Schlussfolgerungen von Damaske überrascht es, dass Theile die von Damaske für bandbegrenztetes Rauschen gefundenen ‚optimalen‘ Kohärenzwerte auf breitbandige Musiksignale überträgt.

Wie in den nachfolgenden Abschnitten gezeigt werden wird, konnte durch Experimente im Rahmen der vorliegenden Dissertation nachgewiesen werden, dass die von Theile vorgeschlagenen minimale Kohärenz von $\gamma \geq 0.35$ nicht dazu geeignet ist, einen ‚natürlichen Raumeindruck‘ hervorzurufen; die Untersuchungsergebnisse lassen darauf schließen, dass - ganz im Gegenteil - gute Dekorrelation (oder Inkohärenz) in allen Frequenzbereichen für die Wiedergabe eines überzeugenden Raumeindrucks zu bevorzugen ist.

In diesem Zusammenhang drängt sich auch die Frage auf, ob die zweifelhafte Interpretation von Damaskes Ergebnissen durch Theile (gemeinsam mit anderen Lehrmeinungen wie u.a. [Sengpiel, 1998] und [Wuttke, 1998]) nicht evt. dazu geführt hat, dass eine ganze Generation von Tonmeistern und Forschern in eine falsche Richtung gewiesen wurde: nach Surround-Mikrofosystemen mit hoher Korrelation zu suchen, statt nach solchen, die sich durch möglichst geringe Korrelation (bzw. Kohärenz) in allen Frequenzbändern auszeichnen.

3.3 Optimaler interauraler Korrelationsgrad

Die obenstehend genannte Angabe von Theile betreffend eine optimale Kohärenz mit $\gamma \geq 0.35$ steht auch im Widerspruch zu den Erkenntnissen von [Gottlob, 1973]:

Ähnlich wie für die Wiedergabe von klassischer Musik im Konzertsaal eine von den meisten Hörern bevorzugte Nachhallzeit gefunden werden konnte, welche – natürlich auch in Anhängigkeit von der Saalgröße - in der Regel zwischen 1,8 bis ca. 2,1 s liegt (vgl. [Beranek, 2004], S. 494, 495, sowie [Pierce, 1983], S.151 und [Gottlob, 1973]; Anm.: für eine genauere Berechnung der optimalen Nachhallzeit in Abhängigkeit vom Raumvolumen siehe [Watson, 1923]), kommt Gottlob in seiner Arbeit zum Schluss, dass es auch einen optimalen interauralen Korrelationsgrad (für Musikkwiedergabe im Konzertsaal) gibt, welcher 0,23 beträgt. (siehe hierzu auch [Cremer und Müller, 1978], S. 482)

In diesem Zusammenhang ist auch die Untersuchung von [Nakahara, 2005] interessant, aus der hervorgeht, dass einen geringe Korrelation der Lautsprecher-signale (L/R vs. C vs. LS/RS) eine bessere Kompatibilität zwischen unterschiedlichen Wiedergabe-Umgebungen und eine Verbreiterung des ‚sweet spot‘ bewirkt (siehe hierzu auch [Prokofieva, 2007]).

Auch aus diesem Grund sollten vorzugsweise Mikrofontechniken verwendet werden, die möglichst über den gesamten Frequenzbereich eine niedrige Signalkorrelation aufweisen.

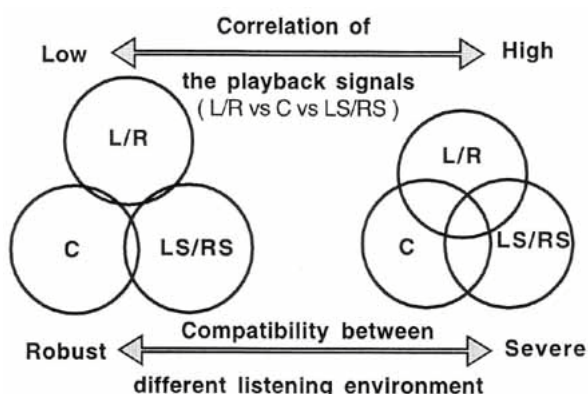


Abb.3.10: Zusammenhang zwischen Signalkorrelation und Kompatibilität zu unterschiedlichen Hörumgebungen (aus [Nakahara, 2005])

3.4 Gedanken zur Wechselwirkung zwischen Lautsprecher-Richtwirkung und Hörer-Envelopment

Viele praktisch arbeitende Tonmeister kennen das ‚Phänomen‘, dass eine Surround-Mikrofontechnik eine gute ‚akustische Einhüllung‘ sowie einen guten Raumeindruck erzeugt, solange es um die Aufnahme/Wiedergabe von musikalischen Signalen geht, sobald aber z.B. Publikumsapplaus zu hören ist, wird plötzlich für den (kritischen) Hörer evident, dass er die Musik nicht ‚live‘, sondern über Lautsprecher zu hören bekommt. Der Hauptgrund hierfür liegt vermutlich in der grundlegend unterschiedlichen spektralen Zusammensetzung und Bandbreite von Musik- und Rausch-Signalen (Applaus). Während (akustische) Musik normalerweise geringere Pegel in den höheren Frequenzbändern aufweist, lässt sich Applaus mit gefärbtem Rauschen vergleichen. (ähnlich ‚Rosa Rauschen‘, das eine spektrale Leistungsdichte proportional $1/f$ aufweist.)

Lautsprecher haben in der Regel eine stark frequenzabhängige Abstrahlcharakteristik, wobei tiefe Frequenzen annähernd omnidirektional abgestrahlt werden, mittlere und hohe Frequenzen hingegen mit gerichteter Charakteristik (siehe Abb.3.11). Dieses Faktum ist vermutlich auch verantwortlich dafür, dass man als Hörer den Lautsprecher selbst (statt des Nutzsignals) wahrnimmt, wenn das Nutzsignal – wie im oben beschriebenen Fall - einen großen Anteil an hohen Frequenzen enthält.

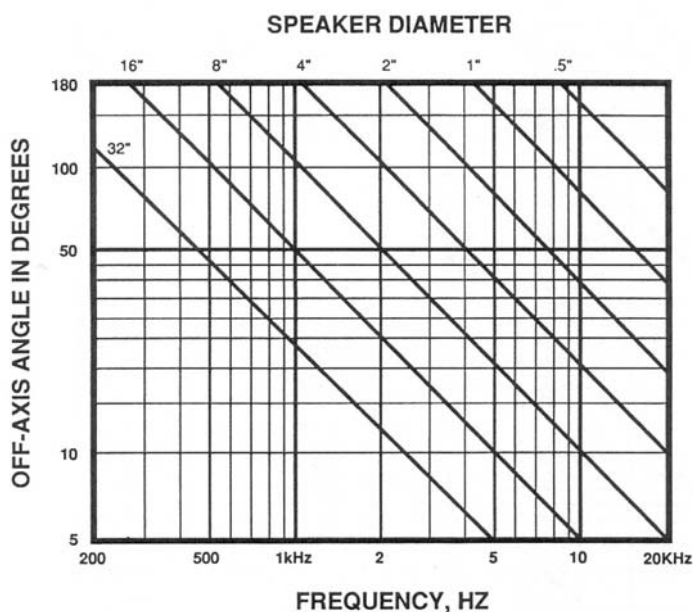


Abb. 3.11: Abstrahlwinkel einer Lautsprecher-Membran (Diaphragma, dynamisches Prinzip) in Abhängigkeit von Treiber-Durchmesser und Frequenz (aus [Dickason, 2000])

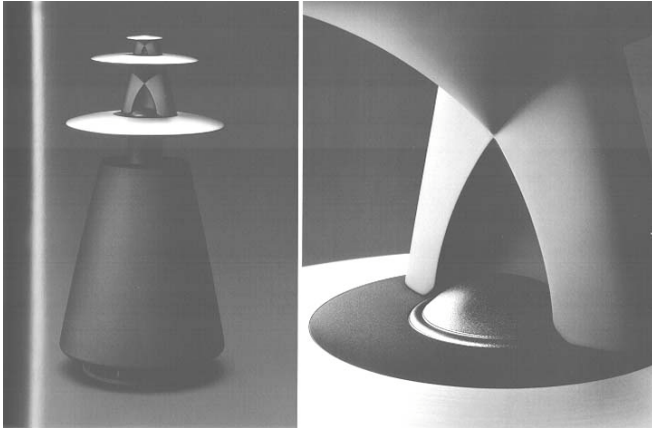


Abb. 3.12: ‚BeoLab 5‘ –
3-Weg Lautsprecher mit Acoustic
Lens Technology (ALT)
(aus Bang & Olufsen Prospekt)

In diesem Zusammenhang darf die Vermutung geäußert werden, dass Lautsprecher mit einer wohldefinierten, aber weitgehend frequenzunabhängigen Richtcharakteristik (siehe z.B. jene Lautsprecher der Firma Bang und Olufsen (Beolab 5, Beolab 3, u.a.), welche die von Sausalito Audio patentierte ‚Acoustic Lens Technology‘ (abgekürzt ALT) verwenden ([Moulton, 1998], siehe Abb. 3.12-3.14) beim Einsatz im Rahmen eines Experiments wie dem obengenannten von Damaske möglicherweise andere Ergebnisse zur Folge gehabt hätten.

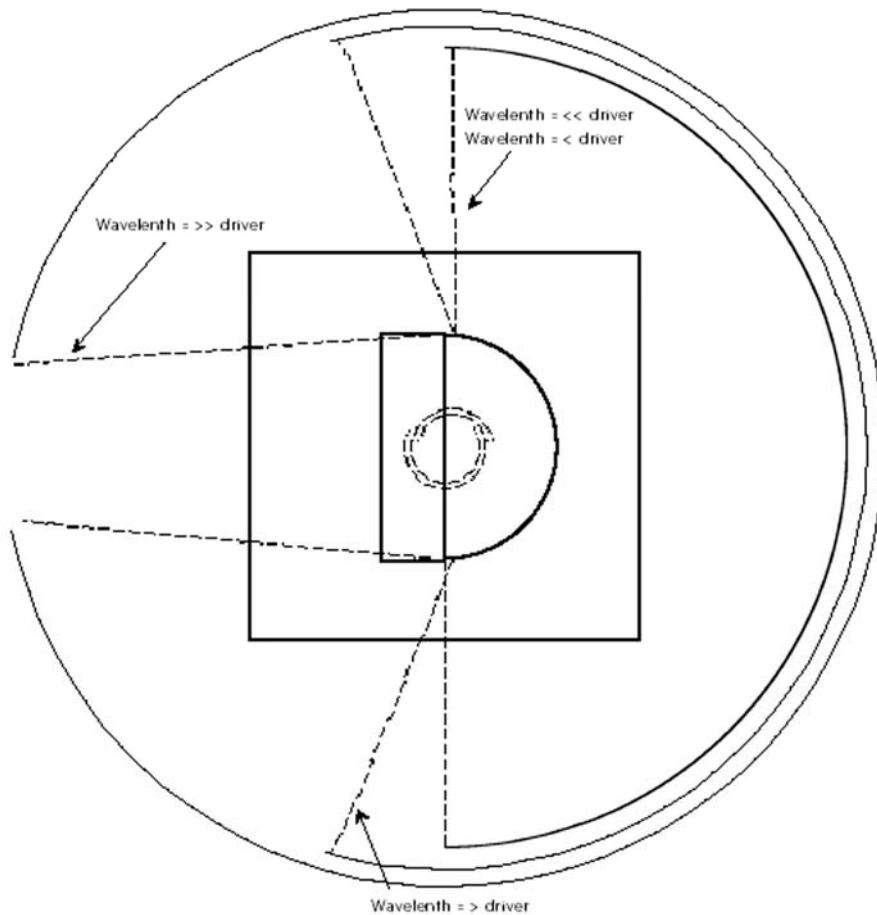


Abb. 3.13: Schematische Darstellung der horizontalen Abstrahlcharakteristik der ‚Acoustic Lens Technology‘, in Abhängigkeit von der Relation Wellenlänge / Treiber-Durchmesser

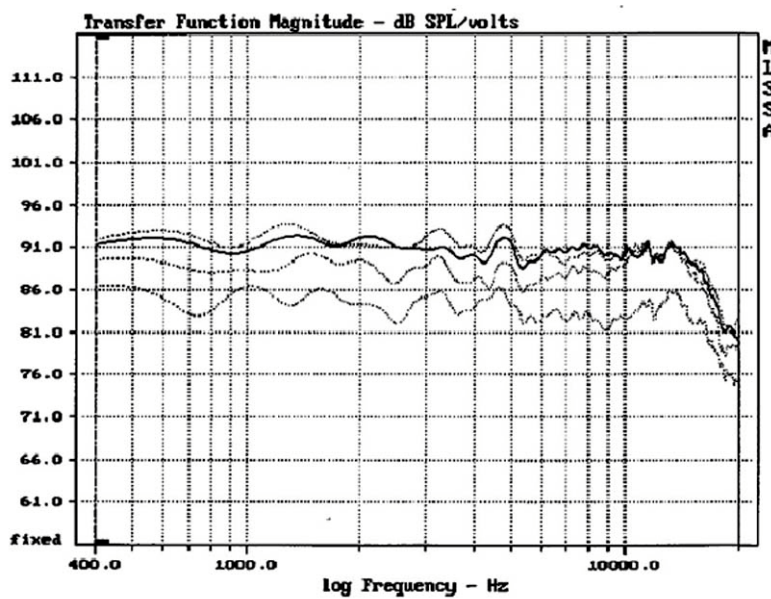


Abb. 3.14: Übertragungsfunktion (quasi-anechoic) des 3-Weg Lautsprecher-prototypen gemessen bei 0°, 30° (= Design-Axis), 60° und 90°, unter Verwendung der Acoustic Lens Technology für Mittel- und Hochton-Bereich (aus [Moulton, 1998])

Mit traditionellen n-Weg (mit $n \geq 2$) dynamischen Lautsprechersystemen (nach dem ‚direct radiator‘ Prinzip, d.h. mit direkter, gerichteter Frequenzabstrahlung) wird das Nutzsignal vor Abstrahlung in mehrere Frequenzbereiche aufgespalten. Bei obengenannten Rauschsignalen oder Applaus müssen die physikalisch in der Regel vergleichsweise kleinen Hochton-Wandler wesentlich mehr akustische Energie pro Quadratcentimeter Wandlerfläche abstrahlen, als Mittel- oder Tieftonwandler. Dies, gemeinsam mit dem - rein physikalisch/akustisch bedingt - stärker gerichteten Abstrahlverhalten von dynamischen Hochfrequenz-Wandlern (sofern diese ohne Horn oder Waveguide eingesetzt werden) verstärkt den Effekt, dass der Hörer den HF-Wandler wesentlich rascher als scharf abgegrenzte Punktquelle wahrnimmt, als dies beim Tiefton-Wandler der Fall wäre (immer unter der Voraussetzung eines entsprechend beschaffenen Nutzsignals).

Dieser akustische Effekt ist vermutlich bei Lautsprechern mit weniger gerichteter Abstrahlcharakteristik (bzw. mit von Frequenzänderungen unabhängiger Richtcharakteristik) nicht gegeben oder zumindest weit weniger ausgeprägt; hierzu gehören alle Lautsprecher, die große Membranflächen zur Abstrahlung des gesamten Nutzfrequenzspektrums nützen (oder sich mittels ‚indirekter Abstrahlung‘ / Reflexion größerer Flächen bedienen (siehe hierzu z.B. [Pfanzagl, 2007])), wie elektrostatische und magnetostatische Lautsprecher, sowie auch jene, die das oben genannte Prinzip der ‚Akustischen Linse‘ nützen.

(Anm.: siehe hinsichtlich der Interaktion von Lautsprecher-Abstrahlcharakteristik und reflektierenden Begrenzungsflächen auch die Informationen in Kapitel 2.5 ‚Einflüsse der Eigenschaften von Lautsprecher und Akustik des Abhörtraumes auf die Hörerpräferenz‘ bzw. die Arbeiten von [Olive et al., 1994], sowie [Toole, 2008])

Wenn Theile also im Rahmen seiner Argumentation (siehe [Theile, 2000]) für eine erwünschte Mindest-Kohärenz von $\gamma \geq 0.35$ für (4-Kanal) Surround-Signale plädiert, weil sich nur bei Erfüllung dieser Vorgabe auch für Applaus das Gefühl der ‚Einhüllung‘ für den Hörer ergibt (statt die Lautsprecher isoliert wahrzunehmen), dann ignoriert er dabei die spektral völlig andere Energieverteilung eines normalen Musiksignals - welches das eigentlich relevante Nutzsignal für den Hörer darstellt - und die damit verbundene, eigentlich anzustrebende Signal- Kohärenz.

4 Stereo-Mikrofontechniken

Die nachfolgende Beschreibung der gebräuchlichsten Stereo-Mikrofontechniken basiert in erster Linie auf [Streicher and Dooley, 1985]

4.1 Koinzidente Techniken

Hierbei sind die Mikrofone in der Regel auf einer vertikalen Achse übereinander und mit einem horizontalen Öffnungswinkel zueinander angeordnet; die Unterschiede im Stereo-Signal gehen dabei auf die, durch die verwendete Richtcharakteristik bedingten Pegelunterschiede zwischen dem L und R Kanal zurück. Die hierbei verwendeten Richtcharakteristiken reichen von Halbnieren über Nieren und Hypernieren bis zur Acht (Blumlein-Paar).

4.1.1 XY mit Nieren und Hypernieren

Normalerweise werden hier Öffnungswinkel zwischen den Mikrofonen von 60° bis 120° verwendet, welche wiederum den sich daraus für die Schallquelle ergebenden sog. ‚Aufnahmewinkel‘ bedingen. Da die Einzelmikrofone den Großteil einer ausgedehnten Schallquelle eigentlich ‚off-axis‘ erfassen, ist ein möglichst linearer, winkelnunabhängiger Frequenzgang wünschenswert, um möglichst geringe Klangfärbungen hervorzurufen. Aus diesem Grund werden für Stereoanwendungen meist bevorzugt sog. ‚matched pairs‘ verwendet, die einander sowohl im Frequenzgang, wie auch Polar-Diagramm möglichst ähnlich sind.

Bei der Verwendung von Nierencharakteristik liegen die verwendeten Öffnungswinkel meist zwischen 90° bis 120° , wobei die Achsen der Mikrofone mitunter auf den jeweils linken und rechten Randbereich der Schallquelle ausgerichtet werden.

Bei der Verwendung von Hypernieren wird in der Regel ein geringerer Öffnungswinkel bevorzugt, um eine stabile Abbildung der Schallquellen-Mitte zu gewährleisten. Aufgrund der stärkeren Richtcharakteristik können Hypernieren weiter von der Schallquelle positioniert werden als Mikrofone mit Nierencharakteristik (Stichwort: ‚direct-to-reverberant sound ratio‘ oder D/R-Ratio).

4.1.2 XY mit Achten (Blumlein-Paar)

Die Anordnung mit zwei im Winkel von 90° zueinander stehenden Achten geht auf Alan Dower Blumlein's Erfindung in den 1930-er Jahren zurück [Blumlein, 1931]. In seinem Patent beschreibt er zwei mögliche Konfigurationen der gekreuzten Achten:

- a) mit einer Ausrichtung von $-45^\circ/+45^\circ$ relativ zur Schallquelle, wobei die so gewonnenen Signale direkt auf den L und R Lautsprecher einer Stereo-Abhöreranlage geführt werden, sowie
- b) als MS-System angeordnet, d.h. eine Acht wird direkt auf die Mitte der Schallquelle ausgerichtet, die zweite im Winkel von 90° seitlich dazu. Für die Stereo-Wiedergabe geeignete Signale werden durch entsprechende MS-Matrizierung gewonnen.

Anzumerken wäre eine systemimmanente Eigenschaft der unter a) beschriebenen Anordnung, welche dem Tonmeister – unter Umständen - Probleme bereiten kann: das nach links ausgerichtete Mikrofon nimmt – gemäß seiner Richtcharakteristik - nicht nur Schall aus dem linken vorderen Quadranten auf, sondern auch Schallanteile aus dem Quadranten rechts hinten, welche dann (mit invertierter Polarität) über den Lautsprecher links vorne wiedergegeben werden. Dies kann – in Abhängigkeit von der Raumakustik – unter Umständen problematisch sein; um dies zu entschärfen bietet es sich ev. an, die bidirektionale Richtcharakteristik der ‚Achten‘ durch den Einsatz von Absorber- oder Trennpanelen auf rein akustischem Wege einzuschränken.

Starke seitliche Reflexionen in einem Raum können sich in diesem Zusammenhang mitunter ungünstig auswirken.

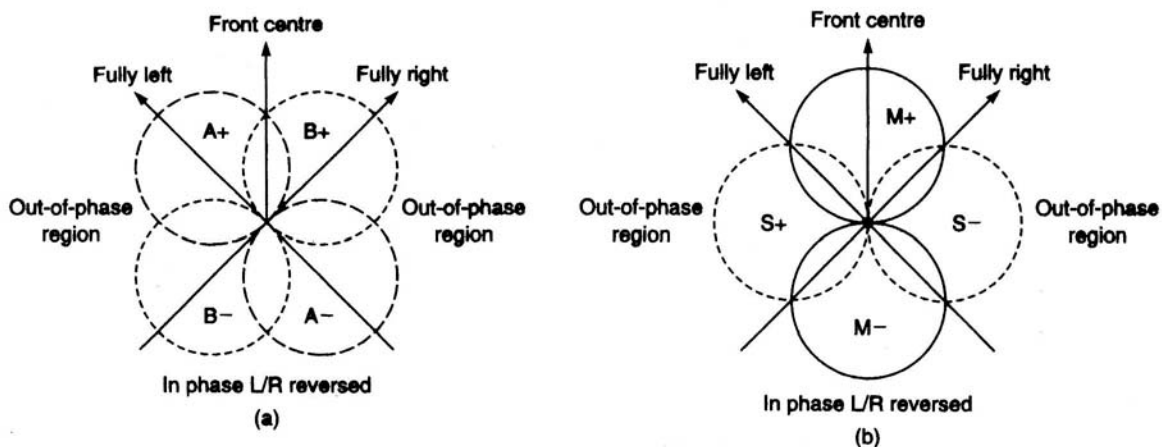


Abb. 4.1 ‚Blumlein-Paar‘ Mikrofon-Anordnung: a) XY-Ausrichtung der Achten ($X=-45^\circ/Y=+45^\circ$); b) MS-Ausrichtung der Achten ($S=-90^\circ/M=0^\circ$) (aus [Rumsey, 2001])

Vor allem die unter a) beschriebene Blumlein-Technik wird von vielen Hörern als besonders natürlich klingende Mikrofonierungstechnik geschätzt [Streicher and Dooley, 1985]

4.1.3 MS-Stereo

Die MS-Anordnung besteht – wie bereits als Spezialfall des Blumlein-Paars angesprochen - aus einer seitlich ausgerichteten Acht und einem nach vorne in die Schallquellenmitte ausgerichteten Mikrofon, welches verschiedene Richtcharakteristiken – von Kugel bis Hyperniere – aufweisen kann. Der sich ergebenden Aufnahmewinkel ist durch das Pegelverhältnis der Signale M und S in der Matrizierung zur Gewinnung der L- und R-Signalinformation des Stereosignals bestimmt. Aus diesem Grunde wird die MS-Technik von vielen Tonmeistern sehr geschätzt, da sich dadurch auch noch in der Postproduction große Gestaltungs-Möglichkeit bieten (sofern die Signale M und S noch getrennt vorliegen). Im Zusammenhang mit Musikaufnahmen ist bei traditioneller Positionierung der Schallquellen im S-Signal vor allem Hallanteil enthalten, während das M-Signal primär Direktschall enthält. Dies wirkt sich günstig auf die Monokompatibilität aus, da für das Monosignal entsprechend nur der Direktschall-Anteil wirksam wird.

4.1.4 Koinzidente Omnidirektionale Mikrofone

Diese zunächst sinnwidrig erscheinende Mikrofonierungsmethode bezieht ihre Berechtigung aus folgender Gegebenheit: da alle omnidirektionale Mikrofone mit steigender Frequenz rein physikalisch bedingt ebenfalls eine Richtwirkung aufweisen, läßt sich diese zur Gestaltung eines unaufdringlichen ‚Stereo-Eindrucks‘ (Stichwort: ‚Apparent source width‘) einer Klangquelle nützen. Die Signale der beiden Omni-Mikros sind relativ hoch korreliert, was u.a. den Vorteil hat, dass z.B. ein Solist zwar in Stereo abgebildet wird, sich ev. Bewegungen seinerseits jedoch nicht störend in Form von ‚Image Shifts‘ (Lokalisationsverzerrungen) bemerkbar machen. ([Streicher and Dooley, 1985], [Swedien, 1997]).

4.2 Äquivalenz-Stereophonie

Bei dieser Gruppe von Mikrofon-Techniken beruht die Stereoabbildung auf Pegel- und Laufzeitdifferenzen, die zwischen den Signalen der beiden Kapseln auftreten.

4.2.1 ORTF

Entwickelt bei und benannt nach dem französischen Rundfunk (Office de Radiodiffusion-Télévision Française), besteht diese Konfiguration aus zwei Nierenmikrofonen mit einem Kapselabstand von 17cm (in Anlehnung an die durchschnittliche Distanz der Ohren eines Menschen) und einem physikalischen Öffnungswinkel von 110 Grad. Gemäß den Williamskurven weist diese Anordnung einen Aufnahmewinkel von 150 Grad auf.

4.2.2 NOS

Entstanden beim Holländischen Rundfunk (Nederlandsche Omroep Stichting), verwendet diese Technik zwei Mikrofone mit Nierencharakteristik, mit einem Kapselabstand von 30cm und einem Öffnungswinkel von 90 Grad, was – nach Williams – einen effektiven Aufnahmewinkel von 70 Grad ergibt.

4.2.3 Faulkner ‚Phased Array‘

Entwickelt vom britischen Tonmeister Tony Faulkner [Faulkner, 1981]: zwei parallel nach vorne ausgerichtete Mikrofone mit Achtercharakteristik im Abstand von 20cm. Diese Technik kombiniert die hohe Kohärenz des – ähnlich aufgebauten – Blumlein-Paars mit einer gewissen ‚Offenheit‘ des Klang ([Streicher and Dooley, 1985], S. 552), wie er durch die Laufzeitunterscheide bei AB-Mikrofonierung entsteht. Faulkner empfiehlt außerdem eine von der Schallquelle weiter als üblich entfernte Aufstellung seines ‚Phased Array‘, was in einer besseren Lautstärken-Balance zwischen den vorderen und hinteren Elementen einer (ausgedehnten) Schallquelle resultiert.

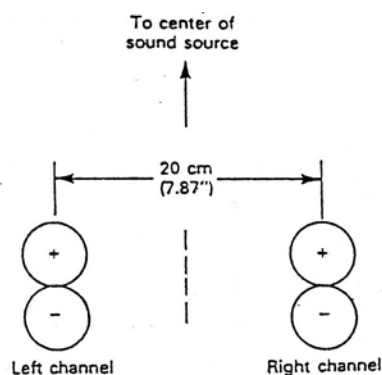


Abb. 4.2 Faulkner ‚Phased Array‘
(aus [Streicher and Dooley, 1985])

4.2.4 OSS – Jecklin Disk

Von seinem Erfinder, dem Schweizer Rundfunk-Tonmeister Jürg Jecklin als ‚Optimal Stereo Signal‘ (OSS) Technik bezeichnet [Jecklin, 1981], ist sie den meisten besser bekannt unter dem Schlagwort ‚Jecklin-Scheibe‘ (engl. Jecklin-Disk): zwei diffusfeldentzerrte Mikrofone mit Omni-Charakteristik, mit Kapselabstand von 16,5cm, getrennt durch eine auf beiden Seiten akustisch bedämpfte Scheibe (Absorberpanel) mit 28cm Durchmesser.

Die Mikrofone sollen dabei so ausgerichtet sein, dass sie die Schallquelle leicht ‚überschießen‘ ([Jecklin, 1998]).

Jecklin beschreibt die Abbildungseigenschaften seines Systems wie folgt (sinngemäß, leicht gekürzt übersetzt aus [Jecklin, 1981]):

„ ...1) Im Frequenzbereich unterhalb 200 Hz ist die Scheibe (akustisch) unwirksam. Beide Mikrofone erhalten das gleiche Signal.

2) Mit steigender Frequenz nimmt auch die Trennkörperwirkung der Scheibe zu und hat einen stärkeren Unterschied der Mikrofon-signale zur Folge.

3) Die Richtcharakteristik des OSS Mikrofons ist omnidirektional, was in Räumen mit normaler oder kurzer Nachhallzeit befriedigende Ergebnisse liefert. In Räumen mit langer Nachhallzeit lässt sich die kugelförmige Charakteristik in eine mit Nieren-Charakteristik verwandeln, indem man ein Absorberpanel hinter dem OSS-Mikrofon platziert.“

Eigentlich wurde die Erfindung eines ‚Klein AB‘ Mikrofon-Paars (Omni-Charakteristik), die durch ein akustisches Panel getrennt sind, bereits von A. D. Blumlein in seinem Patent aus den 1930-er Jahren vorweggenommen. (siehe [Blumlein, 1931]).

4.3 AB-Systeme (Laufzeit Stereophonie)

Bereits anlässlich der Internationalen Ausstellung der Elektrizität in Paris im Jahre 1881 stellten Besucher beim Anhören der via Kabel von der Pariser Oper übertragenen Signale von (jeweils) zwei mit Abstand zueinander entlang der Bühnenkante platzierten Mikrofonen beim Abhören über Kopfhörer einen ‚Räumlichkeits-Effekt‘ fest [Engineering, 1881]. Dieser Effekt beruht auf der unterschiedlichen Laufzeit des Schalls zu den beiden Schallwandlern.

Man unterscheidet in der Tontechnik zwischen sog. ‚Klein-AB‘ und ‚Groß-AB‘ Mikrofon-systemen. Erstere verwenden Kapselabständen von ca. 50cm (bis max. 100cm), welche Laufzeitunterschiede von ca. 1.5ms für seitlich einfallenden Schall hervorrufen und auf den damit

verbundenen Effekt der Richtungsempfindung beim menschlichen Hörer abzielen. ‚Groß-AB‘ Anordnungen hingegen verwenden Kapselabstände von einem bis mehreren Metern und nehmen somit – aufgrund der bedeutend längeren Schalllaufzeiten zwischen linkem und rechtem Kanal – keine Rücksicht auf gehörphysiologisch ‚richtige‘ ITD-Werte, die beim menschliche Hörer auftreten. Die stereophone Ausdehnung einer Schallquelle bzw. Lokalisation einer punktförmigen Schallquelle innerhalb der abgebildeten Stereo-Basis erfolgt demzufolge nicht nur aufgrund der Laufzeitunterschiede, sondern – bei wachsendem Kapselabstand – zunehmend mit den damit einhergehenden Zuwachs an Pegeldifferenz zwischen den beiden Kanälen. Man spricht in diesem Zusammenhang auch von ‚Localisation Zones‘, die sich um die Mikrofon ausbilden. Mit steigendem Abstand der Mikrofone zueinander wächst auch die Gefahr der Ausbildung des sog. ‚Mittenlochs‘, d.h. die Schallinformationen werden vom Hörer primär im Umfeld der Lautsprecher wahrgenommen, während sich zwischen diesen eine Art ‚tote Zone‘ ausbildet.

Der Ausbildung des Mittenlochs kann durch den Einsatz eines geeigneten sog. ‚Centerfill-Systems‘, z.B. im einfachsten Fall durch die Hinzufügung eines dritten (omnidirektionalen) Mikrofans in der Mitte abgeholfen werden.

Eine solche Technik unter Verwendung von 3 gleichwertigen, ungerichteten Mikrofanen wurde erstmals bei den Bell Laboratories in den 1930-er Jahren angewendet [Keller, 1981]. Ein ähnliches System wurde später bei den legendären ‚Living Presence‘ Aufnahmen des Plattenlabels ‚Mercury‘ eingesetzt [Valin, 1994].

Eine erste Untersuchung zu Lokalisationsverzeichnungen, die sich bei 2 und 3-Kanal Stereo-Systemen (allerdings nur unter Berücksichtigung der Pegelverhältnisse zwischen den Mikrofansignalen) ergeben, stammt bereits aus den 1930-er Jahren [Steinberg and Snow, 1934]. In eine späteren Untersuchung berücksichtigte einer der beiden Autoren [Snow, 1953] zusätzlich auch die dabei auftretenden Lautzeit-Effekte und kam aufgrund dessen zu einer Korrektur der Ergebnisse der ersten Veröffentlichung.

Bereits in der Untersuchung aus den 1930-er Jahren wurde festgestellt, dass – im Falle des 3-Kanal Systems – der zusätzliche Center-Kanal eine stabilere Schallquellen-Abbildung ermöglicht.

Bei ‚Groß-AB‘ Systemen gilt die Richtlinie, als Kapselabstand ein Drittel bzw. die Hälfte der Breite der Schallquelle zu wählen (d.h. bei großen Orchestern ergeben sich somit Kapselabstände bis zu 10m und darüber, was u. U. Laufzeitunterschiede mit mehr als 30ms mit sich bringt).

Bei 3-Kanal Mikrofon-Systemen kann es bei Summierung der Signale für 2-Kanal Stereo am Stereo-Bus zu Kammfiltereffekten kommen, die sich für den Hörer als Klangfärbung bemerkbar machen, weshalb hier besondere Vorsicht geboten ist.

Im Prinzip sind AB-Anordnungen mit kleinen, wie großen Kapselabständen auch mit Mikrofonen mit Richtcharakteristik vorstellbar und werden in der Praxis auch (teilweise in Kombination mit anderen Techniken) angewendet. Die sich ergebenden Aufnahmewinkel lassen sich auf Basis der sog. ‚Williamskurven‘ [Williams, 1987] ermitteln.

AB-Polycardioid Centerfill (AB-PC)

Als ein Spezialfall einer Breit-AB + Centerfill Anordnung sei das sog. ‚AB-Polycardioid Centerfill System‘ [Pfanzagl-Cardone, 2002] genannt, welches auch im Rahmen der vorliegenden Arbeit zum Einsatz kam. Nähere Details hierzu (inkl. der Erweiterung für Surround-Anwendung) finden sich im Abschnitt über Surround-Mikrofonssysteme. Das AB-PC System geht dabei in seinem Grundprinzip auf die ‚AB-Centerfill‘ Systeme der RCA ‚Living Stereo‘-Serie, wie auch der DECCA-Aufnahmen (Decca-Tree + Outriggers) zurück.

5 Surround-Mikrofontechniken

5.1 Koinzidente und ‚Quasi-koinzidente‘ Arrays

5.1.1 Soundfield Mikrofon

Die Kombination von 4 Sub-Cardioid Kapseln, die in Form eines Tetraeders angeordnet sind, geht auf Michael Gerzon zurück (siehe [Gerzon, 1973]). Die Orientierung der Kapseln ist hierbei:

(Front): Left-up, Right-down, (Rear): Right-up, Left-down

Mittels Matrizierung dieser 4 sog. ‚A-Format‘ Signale lassen sich die 4 ‚B-Format‘ Komponenten gewinnen, welche den Signalen eines Omni- und 3 orthogonal stehender 8-er Mikrofone entsprechen:

$W = \text{Druck-Komponente} = (L_u + R_d + R_u + L_d)$

$X = \text{vorne/hinten Schallschnelle Komponente} = (L_u + R_d) - (R_u + L_d)$

$Y = \text{links/rechts Schallschnelle Komponente} = (L_u + L_d) - (R_d + R_u)$

$Z = \text{oben/unten Schallschnelle Komponente} = (L_u + R_u) - (R_d + L_d)$

Aus diesen 4 B-Format Signalen lassen sich beliebige Richtwirkungen erster Ordnung in jede gewünschte Richtung bilden.

Mit Hilfe des sog. MKV-Prozessors bzw. dessen Emulation als Software-Plugin der Fa. Soundfield können die A-Signale in B-Format Signale gewandelt und Klangeinstellungen (virtuelle Rotation der Ausrichtung des Mikrofons in 3 Raumachsen, Richtcharakteristiken der Kapseln, etc.) vorgenommen werden. Eine an das Soundfield-Mikrofon angelehnte Mikrofonanordnung, allerdings diskret aufgebaut aus Einzelmikrofonen, wurde von [Sokol, 2003] vorgeschlagen.

Das Soundfield-Mikrofon ist u.a. aufgrund seines ‚zukunftssicheren‘ Aufnahmeformats interessant, da sich mit passenden Algorithmen Mehrkanal-Wiedergabesignale für eine beliebige Anzahl an Kanälen realisieren lassen. So bestechend dieses Ambisonic-basierte System in der Theorie ist, unterlag es dennoch in zwei Hörvergleichen ([Camerer und Sodl, 2001], [Hildebrandt und Braun, 2000]) den Konkurrenz-Systemen.

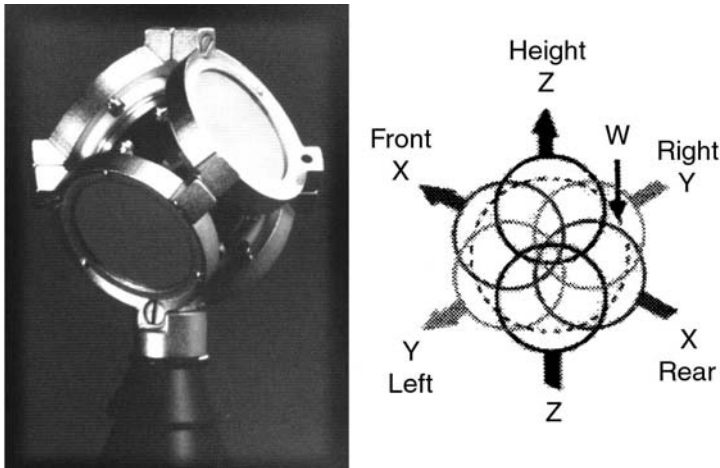


Abb.5.1: Soundfield MK V Surround Mikrofon, Tetraeder-Kapselanordnung, B-Format Schema

5.1.2 DMS (Double-MS Technik)

Ein erster Vorschlag für eine Double-MS Anordnung (bestehend aus zwei back-to-back angeordneten Nieren-Mikrofonen, in Kombination mit - einem - seitlich ausgerichteten Achter-Mikrofon) geht auf Curt Wittig und Neil Muncy Mitte der 1980-er Jahre zurück, damals allerdings nicht in einer koinzidenten Anordnung, sondern mit dem MS-Paar für die Rear-Kanäle im Abstand des Hallradius' nach hinten versetzt. [Mitchell, 1999]. Bereits 1980 wurde allerdings von Hiroyuki et al. eine Patentschrift betreffend eine back-to-back MS-Anordnung beim japanische Patentamt eingereicht. [Hiroyuki et al., 1980]

In den letzten Jahren sind von zwei deutschen und einem japanischen Mikrofonhersteller Spezialanfertigungen für Double-MS Anordnungen auf den Markt gebracht worden (siehe [Schoeps Double-MS, 2010], [Sennheiser MKH800 Twin, 2010] und [Sanken WMS-5, 2010]).

Im Rahmen eines Hörvergleichs [Kamekawa et al., 2007], bei dem das DMS-System mit 7 anderen Systemen verglichen wurde, erhielt dieses bei Orchesteraufnahmen die schlechteste Einstufung (siehe Abb. 5.2)

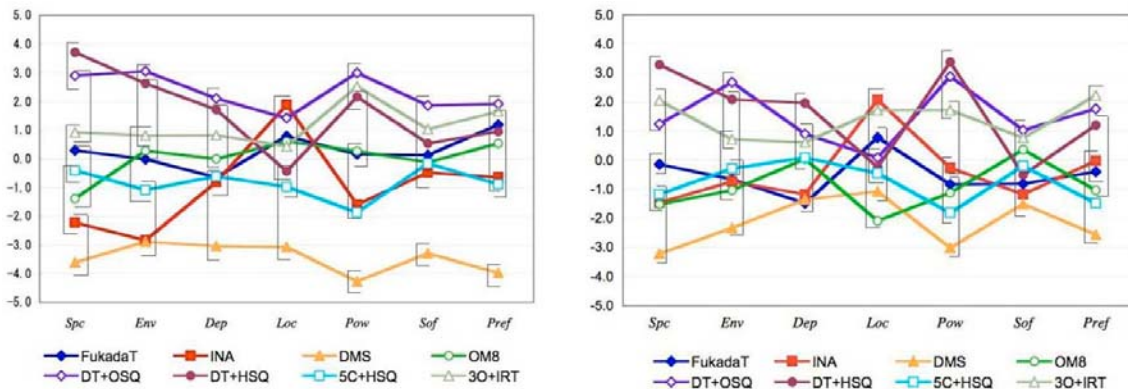


Abb. 5.2: Beurteilung von 8 Surroundmikrofon-Systemen (Fig. 3 und 5 aus [Kamekawa et al. 2007]); links: Musikausschnitt Oper „Die Hochzeit des Figaro“, rechts: Musikausschnitt Orchester „Wellington’s Sieg“

DMS=DoubleMS, DT=Decca-Tree, OSQ=Omni-Square, HSQ=Hamaski-Square, 5C= 5 Cardioid Vorhang, 3O= 3 spaced Omni-Vorhang

5.1.3 Double-MS + Kunstkopf

Michael Bishop, Cheftonmeister der amerikanische Fa. TELARC, wendet u.a. eine um 90 Grad gedrehte ‚Doppel MS‘ Anordnung in Kombination mit dem Neumann Kunstkopf KU 100 an (siehe Abb.5.3). Die MS-Systeme sind zur Seite gerichtet und haben einen Abstand von ca. 2m zueinander; sie werden 1 bis 2,5m hinter dem Kunstkopf aufgestellt. Die richtige Balance der MS-Signale muss - lt. Aussage Bishop - hierbei nach Gehör eingestellt werden, um Lokalisation und Räumlichkeitseindruck zu optimieren. Unter Umständen kann es auch notwendig sein, die Mitten-Signale der Nierenmikrofone jeweils zwischen den Front- und Rear-Lautsprechern der linken und rechten Seite zu panoramisieren.

Bishop macht mit dieser seitlich gedrehten ‚Doppel MS-Technik‘ etwas, das dem Schoeps Kugelflächenmikrofonmikrofon, ergänzt mit seitlichen Achten - gemäß Vorschlag von Jerry Bruck [Bruck, 1996] - nicht unähnlich ist. Auch dort werden zwei MS-Paare gebildet, bei Michael Bishop allerdings ohne die Kugel als Trennkörper dazwischen, mit Nierencharakteristik des Mittenmikrofons und räumlich weiter auseinander gezogen.

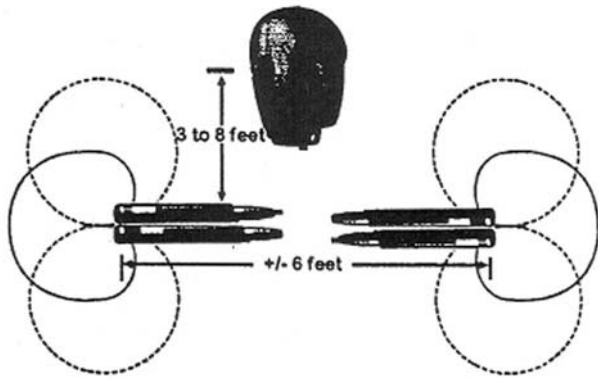


Abb. 5.3: „Double-MS mit Kunstkopf“ Anordnung nach Bishop (aus [Mitchell, 1999])

5.1.4 BPT (Blumlein-Pfanzagl-Triple)

Das BPT-Mikrofon [Pfanzagl-Cardone, 2005] wurde vor allem für die Aufnahme von kleinen Schallquellen (Solo-Instrumente, kammermusikalische Besetzungen) entwickelt und besteht aus drei entlang einer vertikalen Achse angeordneten Kapseln mit 8-er Charakteristik, welche – idealerweise – in der Horizontalen zueinander drehbar sind. Die Normalausrichtung der Kapseln beträgt -45° , 0° , $+45^\circ$ relativ zur Schallquelle, d.h. die bekannte Blumlein-Paar Anordnung aus zwei gekreuzten 8-en wird um eine 8 in der Mitte erweitert. Im Betriebsmodus ‚BPT 3.0‘ werden die Signale dieser 3 Kapseln direkt auf die Kanäle L, M, R aufgeschaltet. Das BPT kann auch als ‚Centerfill‘-System einer Breit-AB Hauptmikrofonanordnung verwendet werden (ggf. gemischt mit Signalen von Stützmikrofonen) und findet somit auch Anwendung bei der Aufnahme von großen Schallquellen, wie z. B. Orchester.

Der (bei Doppelmembran-Kondensatormikrofonen) baubedingt bassarme Frequenzgang der 8-er Kapseln sollte mittels Filterung (z.B. Shelving-Type, +2 bis +6dB mit Eckfrequenz zwischen 120 und 180Hz) kompensiert werden. Um das Gesamtsignal der Front-Anordnung L,M,R allerdings bei tiefen Frequenzen möglichst dekorreliert zu halten, empfiehlt es sich, diese Filterung nur bei den Signalen der Kapseln L, R vorzunehmen und bei der M-Kapsel ggf. die Bassfrequenzen stattdessen weiter zu bedämpfen (z.B. mittels Hochpassfilter, oder abgesenktem Shelving Filter).

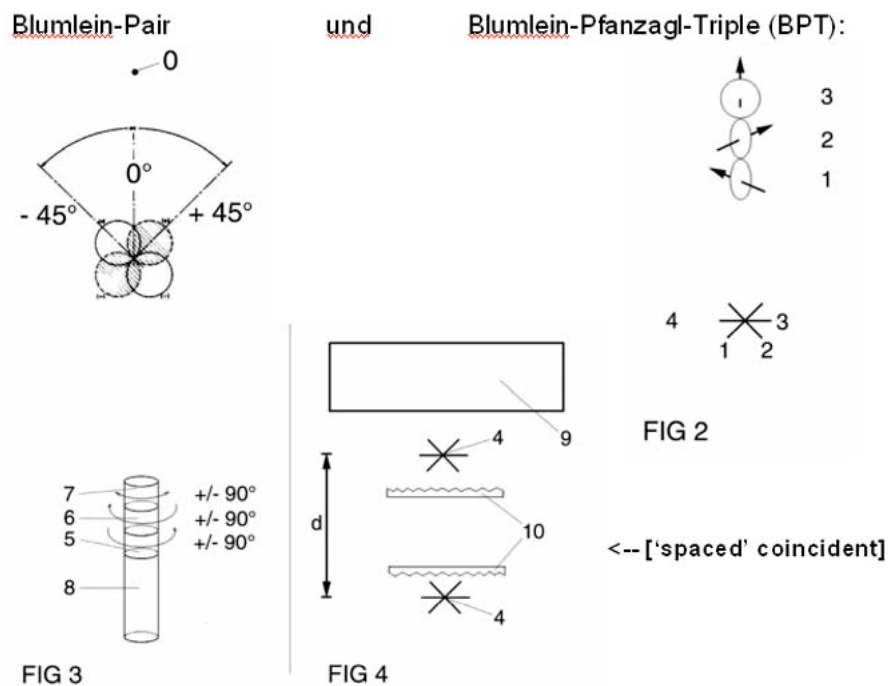


Abb. 5.4: BPT-Surroundmikrofon
 oben: Blumlein-Paar; Blumlein-Pfanzagl-Triple
 Mitte: Mikrofonaufbau schematisch; 2x BPT (4) back-to-back mit Schallquelle (9) und Absorberpanelen (10)
 unten: Grafik-Simulation von 2 x BPT back-to-back auf Videokamera

Im Betriebsmodus ‚BPT 5.1‘ wird das Mikrofon um 45° gedreht, sodass nun eine Kapsel nach vorne und die zweite nach links weist, und die dritte Kapsel wird auf Omni-Charakteristik umgestellt. Somit ergibt sich die Möglichkeit, mittels MS-Dematrixierung Signale für Surround-Wiedergabe zu generieren, wobei mittels geeigneter Decoder-Software auch Wiedergabe im Format 5.1 oder höher möglich ist (siehe Abb.5.5).

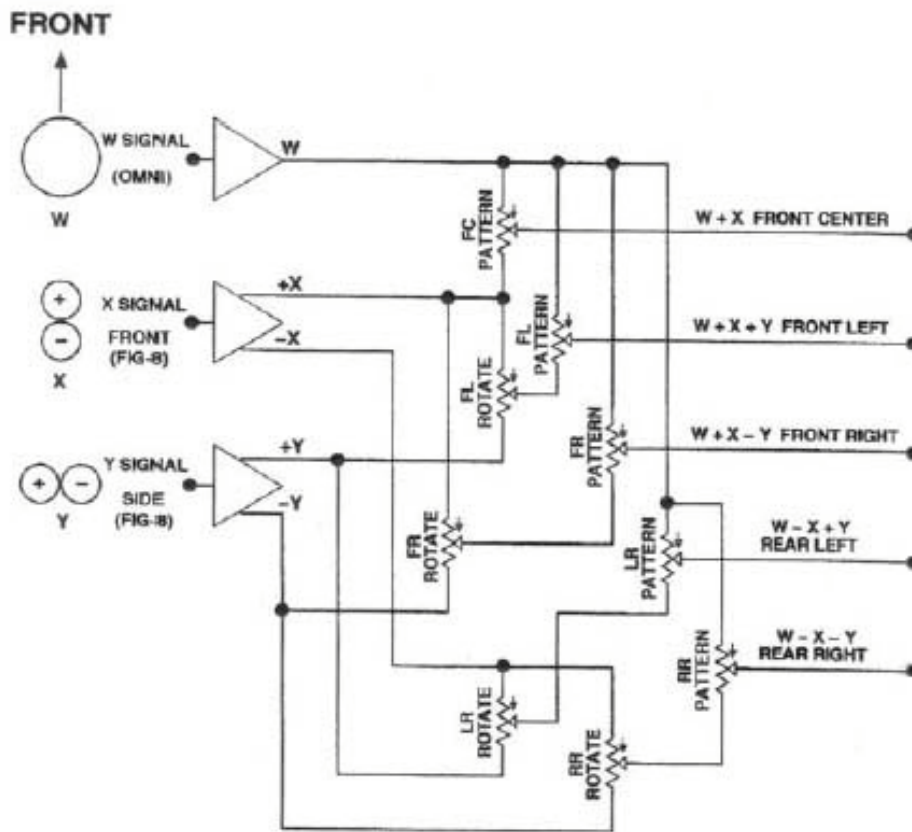


Abb. 5.5: 5.0 Decoding für BPT-Mikrofon (nach Josephson, aus [Eargle 2004])

Theoretische Überlegungen zu Kombinationen von drei vertikal koinzident angeordneten Kapseln finden sich bereits bei [McKinnie and Rumsey, 1997]. Allerdings wird dort weder die Kombination von 3 Kapseln mit 8-er Charakteristik untersucht, noch die Möglichkeit der Gewinnung von 5.1 Surroundsignalen mittels MS-Matrizierung (nach Josephson, siehe [Eargle 2004]).

Bereits davor wurde von [Cohen and Eargle, 1995] eine koinzidente Kapselanordnung zur Gewinnung dreier Front-Signale beschrieben, allerdings wurde hierbei auf eine Nieren-Richtcharakteristik 2. Ordnung abgezielt, was eine Realisierung zurzeit (aufgrund verschiedener technisch-qualitativer Überlegungen) nicht sinnvoll erscheinen lässt.

Wie sich mit Hilfe des sog. "Microphone Assistant"-Tools [Wittek, 2002] für die Front-Anordnung von Mikrofonen berechnen lässt, ergibt sich bei gleicher Empfindlichkeit bzw. Verstärkung der Signale der drei Kapseln (8-er Charakteristik) des BPT-Mikrofons für Schallquellen auf der Mittelachse des Gesamtsystems eine Pegelüberhöhung von ca. 3dB (bei einem Aufnahmewinkel von 164Grad). Eine Reduktion der Empfindlichkeit

des Center- Mikrofons um 10dB bewirkt eine weitgehend gleichmäßige Pegelverteilung entlang der Basisbreite L-M-R. (siehe untenstehende Abb. 5.6), sowie einen Aufnahmewinkel von 110 Grad.

Die Plausibilität dieser auf Basis theoretischer Berechnungen gefundenen Werts konnte durch den Erfinder bei praktischen Anwendungen rein gehörmäßig bestätigt werden. Allerdings konnte in der Praxis auch festgestellt werden, daß die richtige PegelEinstellung des mittleren BPT-Mikrofons stark von dessen Entfernung zur Schallquelle anhängt: bei relativ geringen Distanzen zur Schallquelle ist der obenstehend angegebene Wert von -10dB relativ zum L und R Mikrofon angebracht, bei sehr großen Distanzen zur Schallquelle (welche z.B. dann Sinn machen, wenn die Schallquelle selbst über große Abmessungen verfügt, aber trotzdem in ihrer Gesamtheit nur durch das BPT als Hauptmikrofon alleine abgebildet werden soll; z.B. bei großem Symphonieorchester kann somit der Abstand zum BPT-Mikrofon zwischen 10-15m betragen) kann für eine gehörrichtige Signalwiedergabe eine gleich laute Pegelung aller drei Mikrofone des BPT-Systems erforderlich sein. Dem Tonmeister ist somit ein flexibles Werkzeug in die Hand gegeben, mit welchem er sich an die jeweilige Aufnahme- und Raumsituation anpassen kann.

Erwähnenswert ist in diesem Zusammenhang, dass auch Geoff Martin [Martin, 2005] betreffend sein untenstehend beschriebenes System (siehe ‚Martin-Triple‘) einen Empfindlichkeits-Unterschied von zumindest 9 dB (für von vorne einfallenden Schall) zwischen der Center-Kapsel und den Kapseln L,R als erstrebenswert erachtet, und eine entsprechend veränderte Ausrichtung der Center-Kapsel vorschlägt, um die Signal-Korrelation im Front-System zu verringern. (Details siehe den Abschnitt über das ‚Martin-Triple‘ unten).

Das BPT-System ist in den Grundzügen seiner Anwendung für Surround- (und Stereo-) Aufnahmen in folgenden Gebrauchsmuster bzw. Patentschriften beschrieben:

Australien: Innovation Patent No. AU2005100255 ‚3.0 Microphone for Surround-Recording‘; Österreich: Gebrauchsmuster Nr. 9616 ‚Hauptmikrofon für Surroundaufnahmen‘; Deutschland: Gebrauchsmuster Nr. 20 2005 017 198.9 ‚Hauptmikrofon für Surroundaufnahmen‘; Japan: Utility Model Registration No. 3128257 ‚Microphone for Surround-Recording‘; USA: Utility Model Application No. 11279154 ‚Microphone for Surround-Recording‘ (Patent pending)

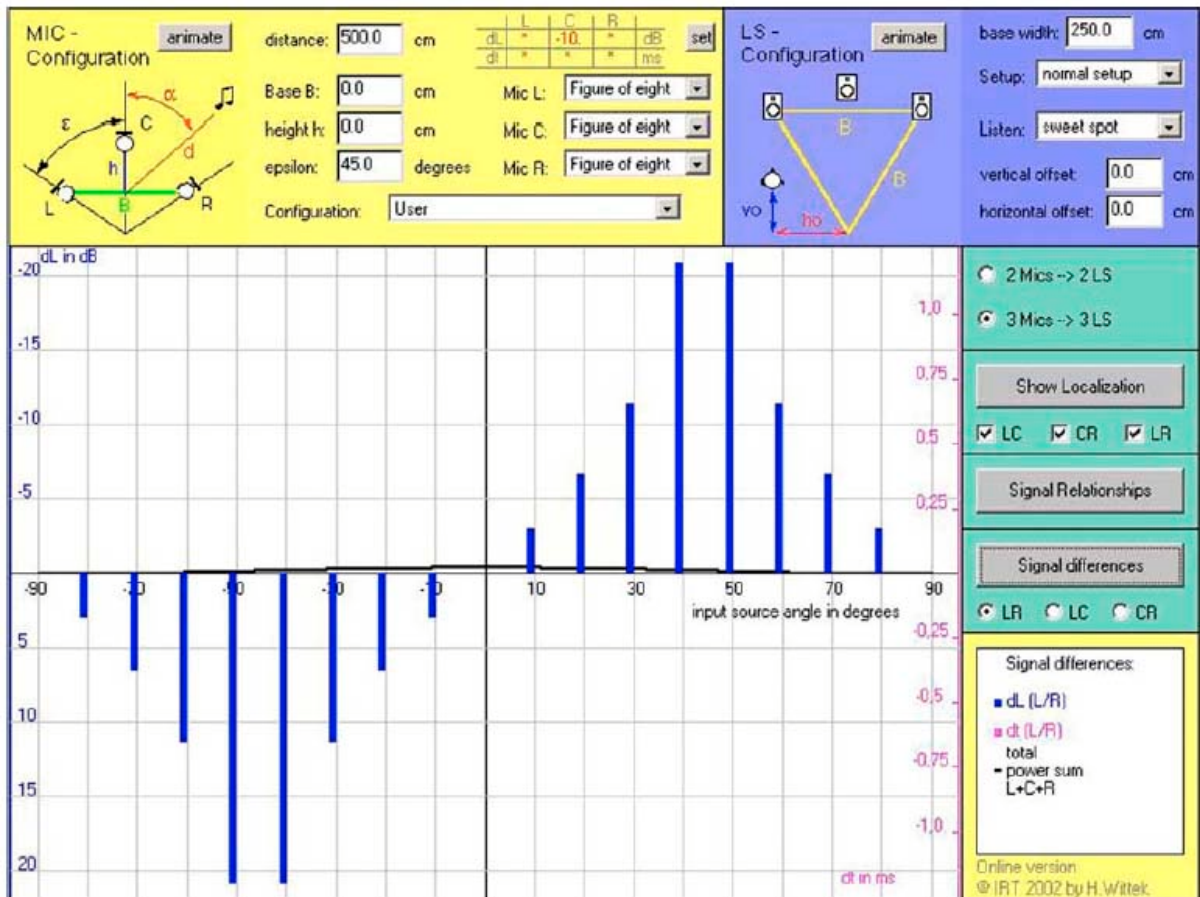


Abb. 5.6: Pegelverteilung über das Gesamtsystem BPT 3.0 bei um 10dB bedämpfitem Centersignal (generiert mittels "Microphone Assistant V2.0" nach [Wittek, 2002])

5.1.5 Martin-Triple

Das ‚Martin-Triple‘ wurde erstmals auf der 118. AES Convention in Barcelona vorgestellt [Martin, 2005] und besteht aus einer Anordnung von 3 Mikrofonen mit 8-er Charakteristik, die entlang einer vertikalen Achse angeordnet sind. Im Unterschied zum BPT ist beim Hauptbetriebsmodus die ‚Center-Kapsel‘ nach unten (oder zumindest diagonal nach unten weisend) ausgerichtet, da Martin dadurch zu hohe Korrelation zwischen dem L bzw. R-Kanal und dem C-Kanal ausschließen möchte. Ein weiterer Vorschlag seinerseits hierzu sieht vor, die C-Kapsel in der Vertikalen zumindest 6cm von der L- und R-Kapsel versetzt anzuordnen, womit die ansonsten hohe Korrelation zwischen L/R und C ebenfalls vermieden wird. Für die Rear-Signale sieht er ein Blumlein-Paar vor, dessen Öffnungswinkel zwischen 45° und 90° betragen kann. Um ein zu starkes Einsprechen von frontalen Schallquellen zu vermeiden, empfiehlt er ggf. die beiden Mikrofone diagonal nach oben zu richten (damit würde die Schallquelle also

quasi ‚überschossen‘). Die wesentlich effizientere Methode, ein Absorberpanel zwischen Front- und Rear-System einzubringen, wird von ihm offenbar nicht in Erwägung gezogen. Ebenso spezifiziert er einen maximalen Abstand von 70cm zwischen den Systemen, was der für den Raumeindruck so wichtigen hohen Dekorrelation zwischen den Front- und Rear-Kanälen ebenfalls nicht zuträglich sein dürfte.

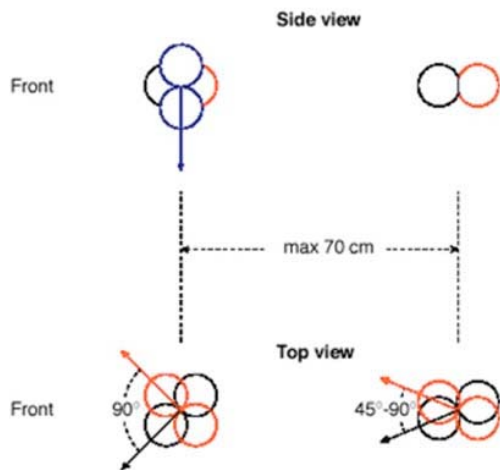


Fig. 19: Schematic view of configuration for Version 1 of the technique. The L, R, LS and RS microphones are forward-facing. The microphone for the centre channel information is coincident with the front array and is pointing towards the floor.

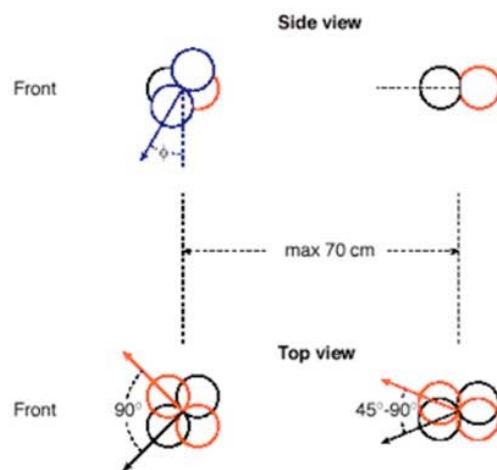


Fig. 20: Schematic view of configuration for Variation 1 of the technique. This can be used to increase the level of frontal direct sound sources in the centre microphone.

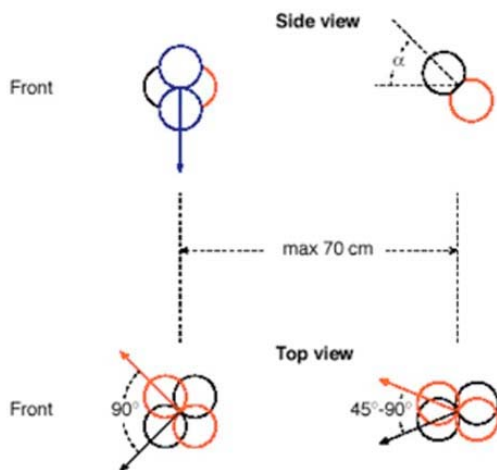


Fig. 21: Schematic view of configuration for Variation 2 of the technique. This can be used to decrease the level of frontal sources in the rear microphone array if a problem occurs with front images „pulling“ into the surrounds.

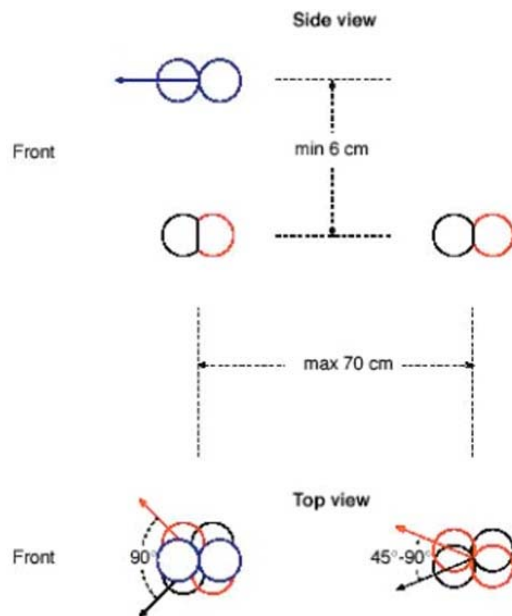


Fig. 22: Schematic view of configuration for Version 2 of the technique. Not drawn to scale.

Abb. 5.7: Martin-Triple ‚Fig. 19-22‘ aus [Martin, 2005]

5.1.6 XY-Nierenanordnung (Back-to-Back)

Doppelte XY-Anordnungen mit Nierencharakteristik und Öffnungswinkeln zwischen 90 und 120 Grad wurden erstmals von [Yamamoto, 1973] vorgeschlagen. In den letzten Jahren hat ein japanischer Hersteller diese Idee wieder aufgegriffen und Kapseln in entsprechender Anordnung in seinen Compact-Recorder integriert (ZOOM H2).

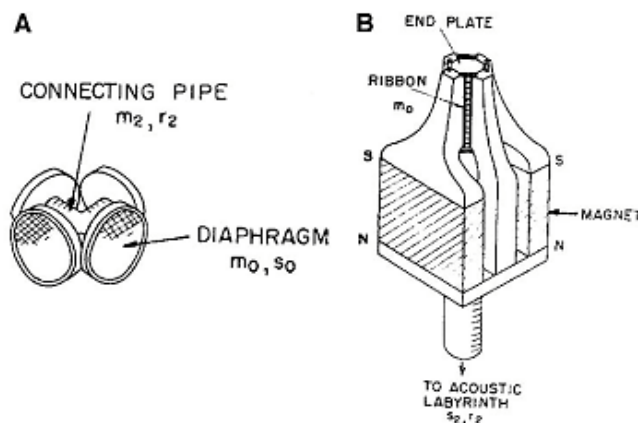


Abb. 5.8: ‚One-Point‘ Quad-Mikrofon mit vier Nierenkapseln
(A = Kondensator; B = Bändchen)

5.1.7 MILAB 360 (360° Surround-Mikrofon)

Das Surroundmikrofon SRND 360 der Fa. MILAB enthält 3 Kapseln mit Nierencharakteristik mit zwischenliegenden Winkeln von 120° und verwendet eine spezielle Matrizierung, um Signale virtueller dazwischen liegender Nierenkapseln zu bilden, mit deren Hilfe sich Ausgangssignale für eine Vielzahl von Surround-Formaten generieren lassen., u. a. 6 (virtuelle) Nierenmikrofone im Winkelabstand von 60°.

5.1.8 Compact 4-3-4 Phasor Array

Die von [Bauer, 1981] vorgeschlagene Anordnung verwendet zwei Druckgradienten Mikrofone (40,42 in Abb.5.9), 0° und 90° relativ zur Schallquelle, sowie einen Druckempfänger (44) und bedient sich div. Signalprozessings, um Ausgangssignale für Quadrophonie-Sound zu erzeugen.

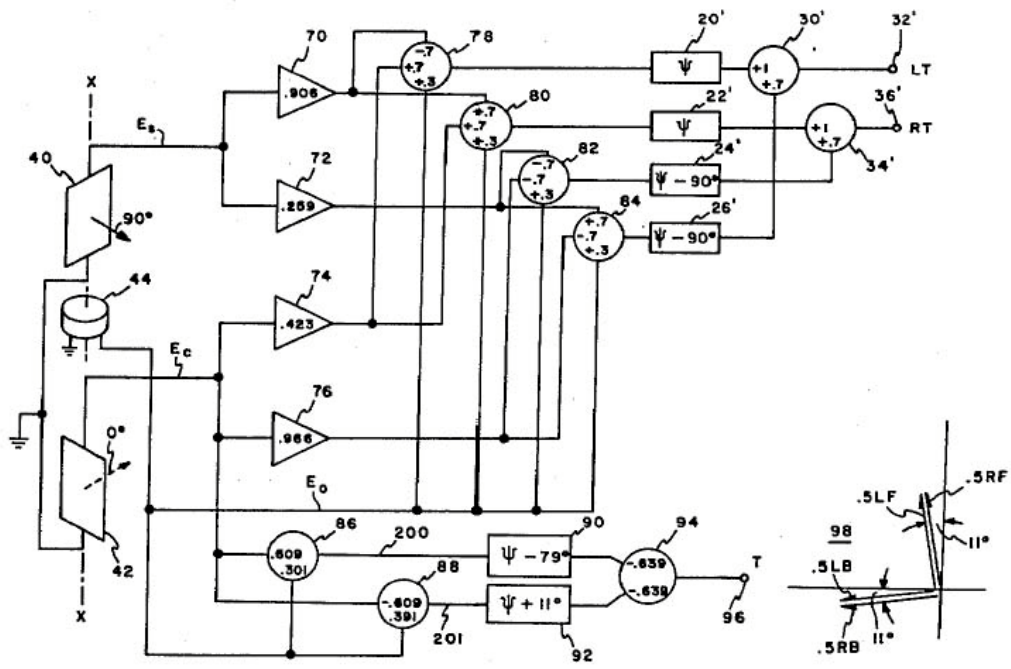


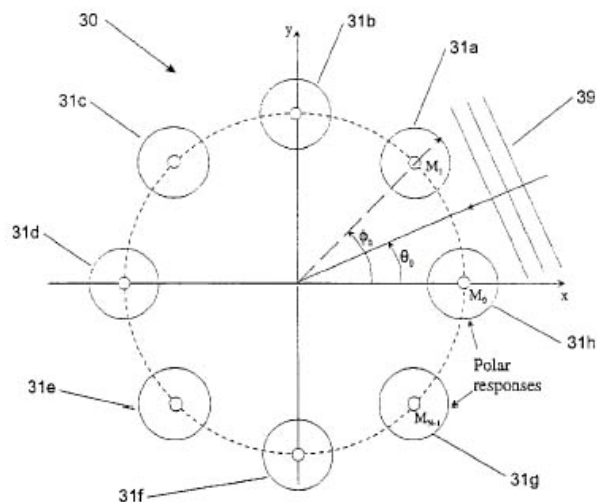
Abb. 5.9: Blockschaltbild des ‚Phasor-Arrays‘ nach Bauer

5.2 Spaced Arrays (2-dimensionale Arrays)

5.2.1 Circular Microphone Array

Mark Poletti stellt das Mikrofon-Array in Abb.5.10 - basierend auf kreisförmig angeordneten Wandlern - vor, welches mittels digitalem Signal-Prozessing, Frequenzgangs-Kompensationsfiltern, Summen- und Differenzbildung, etc. dazu dient, wiedergabeseitig möglichst das ursprünglich aufgenommene Schallfeld nachzubilden. (siehe hierzu [Poletti, 2000], sowie die internat. Patentschrift WO 01/ 58209 A1)

(54) Title: MICROPHONE ARRAYS FOR HIGH RESOLUTION SOUND FIELD RECORDING



(57) Abstract: A circular transducer array (30) is provided for use in recording a sound field. The array (30) comprises a plurality of microphones (31a-31h), a digital signal processor (33), frequency compensation filters (34) and a sum and difference network (35). The digital signal processor calculates the Fourier transform of sampled output signals from the transducers to produce a plurality of sound wave components specifying the sound field. The frequency compensation network (34) equalises each component using Bessel functions to flatten the apparent response of the array (30) and the sum and difference network (35) then combines the equalised components to provide a plurality of audio signals which represent the sound field.

Abb. 5.10: Circular Microphone Array (nach [Poletti, 2000])

5.2.2 Trinnov-Array

In [Laborie et al., 2004] beschreiben die Autoren ein 2-dimensionales Array, welches auf ca. 20 x 20cm untergebracht ist, aus 8 Mikrofonen besteht und auf dem Prinzip der ‚High-Resolution Spherical Harmonics‘ beruht. Sie schlagen darin vor, dass eine Lösung mit Richtcharakteristiken 5.Ordnung ausreichend wäre, um damit die vom 5.0 Standard (ITU-R BS. 775) geforderte Richtwirkung / Signaltrennung zu erreichen.

Ein solches System 5. Ordnung reicht gemäß ihrer Aussage aus, um den irregulären (in Bezug auf vorne/hinten unsymmetrische) Aufbau der genormten 5.1 Lautsprecher-Anordnung mit einem geeigneten Panning-Law gerade noch technisch korrekt realisieren zu können.

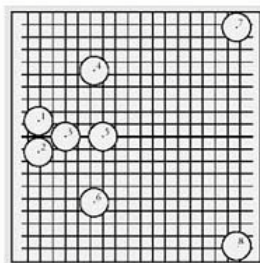
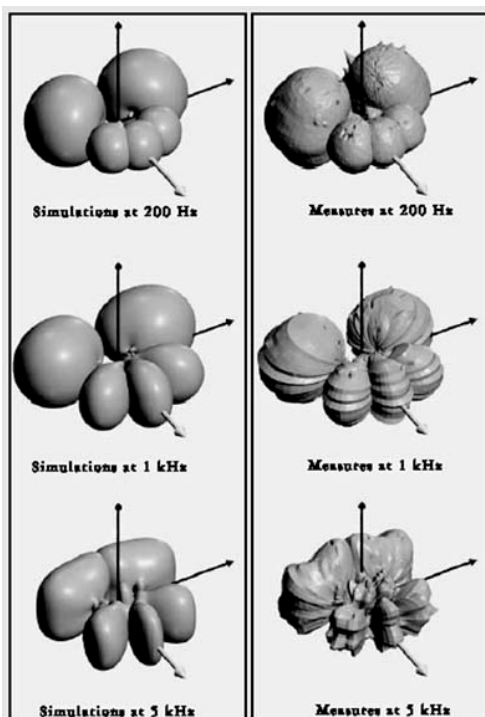


Abb. 5.11: TRINNOV-Surround Mikrofon Array (schematische Anordnung; bauliche Ausführung)

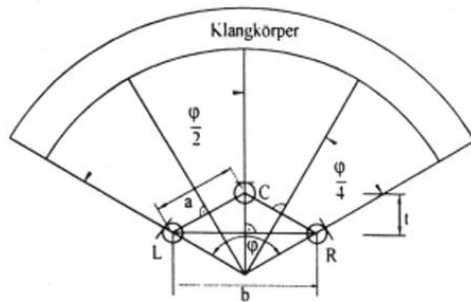


Die Abb. 5.12 zeigt jedoch relevante Unterschiede zwischen berechneter und gemessener Richtcharakteristik: einerseits lässt sich die Richtcharakteristik 5. Ordnung in der Praxis für tiefe Frequenzen nicht mehr erzielen, andererseits bilden sich bei hohen Frequenzen störende ‚secondary lobes‘ aus (siehe das ‚zerfranzte‘ Bild der gemessenen Richtcharakteristiken bei 1kHz und 5kHz). (Anm. des Autors: diese unerwünschten Nebeneulen haben natürlich klangliche Auswirkung in Form von richtungsabhängigen Klangfarben-Unterschieden).

Abb. 5.12: TRINNOV Surround-Mikrofon Array – gerechnete und gemessene Richtcharakteristik in Abhängigkeit von der Frequenz (aus: [Laborie et al., 2004])

5.2.3 ABC 5 (Omni Array)

Diese in [Hermann, Henkels und Braun, 1998] vorgestellte Anordnung geht letztlich auf die DECCA-Dreieck Anordnung zurück und ergänzt diese mit zwei Druckempfängern für die Mikrofone LS, RS (nicht dargestellt). Die Abbildungseigenschaften des Arrays hinsichtlich Lokalisation basieren dabei vor allem auf ITDs, welche beim Hörer durch die unterschiedlichen Laufzeiten des Schalls zu den Mikrofonen hervorgerufen werden.



$\Delta t \dots 1,1 \text{ms}$
(nach Blauert, Williams)

Stereobasis L-C (C-R):
 $a = 37 \text{cm} / \sin(\varphi/4)$
 $b = 2a \cos(\varphi/4)$
 $t = a \sin(\varphi/4) = 37 \text{cm}$

| Gesamtaufnahme- winkel φ in Grad | Mikrofon abstand a in cm | Mikrofon abstand b in cm |
|---|-----------------------------|-----------------------------|
| 100 | 87,5 | 158,5 |
| 120 | 74 | 128 |
| 140 | 64,5 | 105,5 |
| 160 | 57,5 | 88 |

Abb. 5.13: ABC-Mikrofonarray: schematischer Aufbau der Front-Kanäle, Berechnung der Aufnahmewinkel (Graphiken teilweise übernommen aus [Hermann et al., 1998])

5.2.4 INA 5 (Ideale Nieren Anordnung)

Diese ebenfalls in [Hermann, Henkels und Braun, 1998] beschriebene Anordnung ist die Umsetzung der obenstehenden ‚ABC 5‘ Technik auf 5 Kapseln mit Nierencharakteristik (INA – ‚Ideale Nieren Anordnung‘). Für die Lokalisation ausschlaggebend ist hier die Kombination aus den interauralen Laufzeitdifferenzen mit den aufgrund der Richtcharakteristik der Kapseln auftretenden Pegeldifferenzen in den einzelnen Wiedergabekanälen.

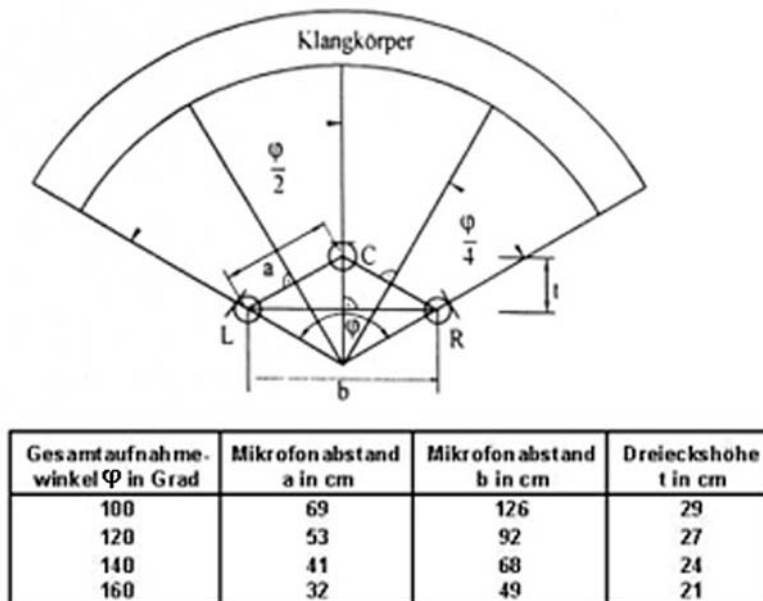


Abb 5.14: ‚INA 5‘ – schematischer Aufbau der Frontkanäle und Angabe von Aufnahmewinkeln basierend auf Berechnung nach den ‚Williamskurven‘ (nach [Hermann et al., 1998])

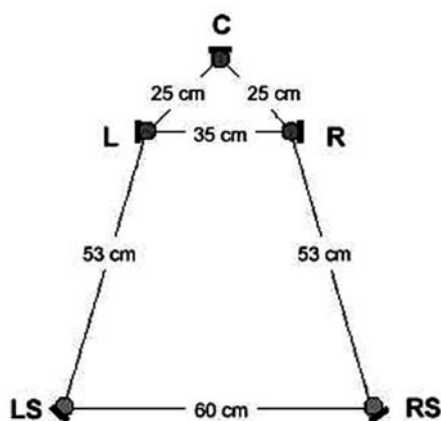


Abb. 5.15: INA5 – Aufbau und Kapselabstände für Aufnahmewinkel 180° (aus [Theile, 2000])

5.2.5 OCT-Surround (Optimal Cardioid Triangle – Surround)

Günther Theile vom IRT München kam aufgrund von Berechnungen zum Schluss, dass die Kanaltrennung einer nur mit Nierenkapseln realisierten Front-Anordnung nicht ausreichend sei, um den psychoakustischen Gegebenheiten des menschlichen Gehörs in optimaler Form gerecht zu werden. Aus diesem Grund schlug er die ‚Optimal Cardioid Triangle‘ Anordnung vor, welche zunächst nur für die Richtcharakteristik und Abstände der 3 Front-Mikrofone spezifiziert wurde [Theile, 2000]: linkes und rechtes Front-Mikrofon als Supernieren, das Mikrofon für die Signale des Centerkanals als Niere. Später erweiterte Theile das Konzept um zwei nach hinten gerichtete Nierenmikrofone für die Kanäle LS und RS. Aufgrund des geringen Abstands untereinander, sowie zum Front-System (siehe hierzu [Pfanzagl-Cardone und Höldrich 2008]) ist die Korrelation der Signale des OCT-Systems bei tiefen Frequenzen (unterhalb 200Hz) ungünstig hoch, was den beim Hörer entstehenden Raumeindruck negativ beeinflusst. (siehe hierzu die Emails von Cornelis van der Gragt und Jan Korte im Anhang)

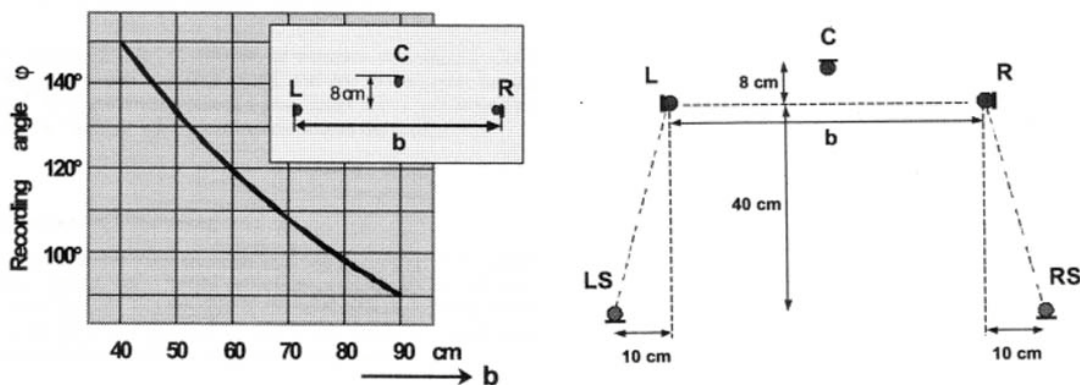


Abb. 5.16: OCT-Surround - Aufnahmewinkel in Abhängigkeit des Kapselabstands, sowie schematischer Aufbau (nach [Theile, 2000] und [Theile, 2001])

Um den generell bassarmen Frequenzgang der für L, R verwendeten Supernieren zu kompensieren schlägt Theile als eine mögliche Ergänzung vor, das (mit Eckfrequenz 100Hz) tiefpass-gefilterte Signal zweier zusätzlich eingesetzter Druckempfänger beizumischen, die unmittelbar neben den Supernieren positioniert werden (nicht abgebildet). (Anm.: Er vernachlässigt dabei allerdings, dass die Korrelation der Signalanteile unter 100Hz somit sehr hoch wird, was dem Räumlichkeitseindruck ebenfalls abträglich ist.)

5.2.6 OCT V.2

Offenbar war das ‚Optimale Nieren Dreieck‘ (OCT) – entgegen der selbstbewussten Namensgebung – doch noch weiter verbesserungsfähig, da Wittek und Theile im Jahre 2004 eine Modifikation desselben veröffentlichten [Wittek, Theile 2004].

Expected Properties of the OCT2 (2004, Wittek, Theile):

- Optimised **Downmix**-Properties in comparison to the OCT1 with respect to the **sound colour** and the **spatial impression** (due to the reduced correlation of the reverb tail between L/C and C/R)
- The precise and stable **directional image** of the OCT1 remains unchanged
- The imaging characteristics with respect to the display of **depth** and **distance** remain unchanged
- The imaging characteristics with respect to the display of **reverb** and **spaciousness** are improved (in 3/2 as well), the sonic image is now a bit more like the well-known Decca-Tree.

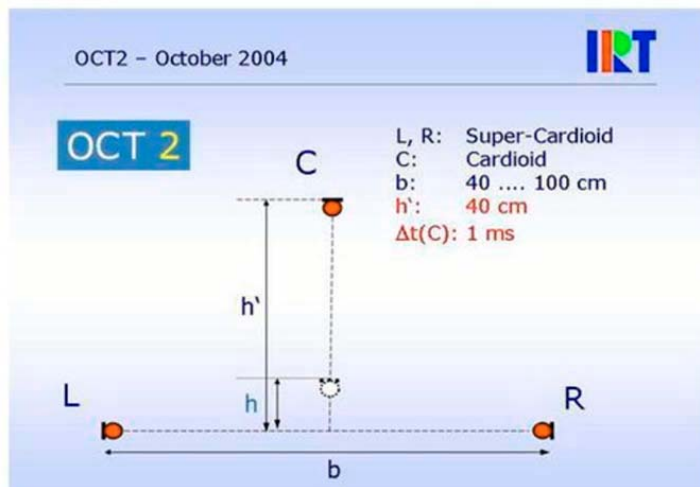


Abb. 5.17: OCT-Version 2 (nach [Wittek, Theile 2004])

Wie die beiden Autoren feststellen, sollte sich durch die um 40cm nach vorne Richtung Schallquelle versetzte Mitten-Niere nun eine natürlichere Klangfarbe und ein besserer Räumlichkeitseindruck ergeben; das Gesamtsystem ähnele nun mehr dem bekannten DECCA-Dreieck.

5.2.7 Klepko-Array

Das Klepko-Array greift Grundprinzipien der ORTF-Technik auf: der Abstand zwischen den außen gelegenen Supernieren und der Mittenniere beträgt jeweils 17,5cm. Mit den physikalischen Öffnungswinkeln von jeweils 30Grad ergibt sich nach Klepko ein Aufnahmewinkel (von L n. R) von insgesamt 150°. Für die Gewinnung der Rear-Signale kommt ein Kunstkopf zum Einsatz, der 124cm hinter der Front-Anordnung aufgebaut wird. (siehe [Klepko, 1997])

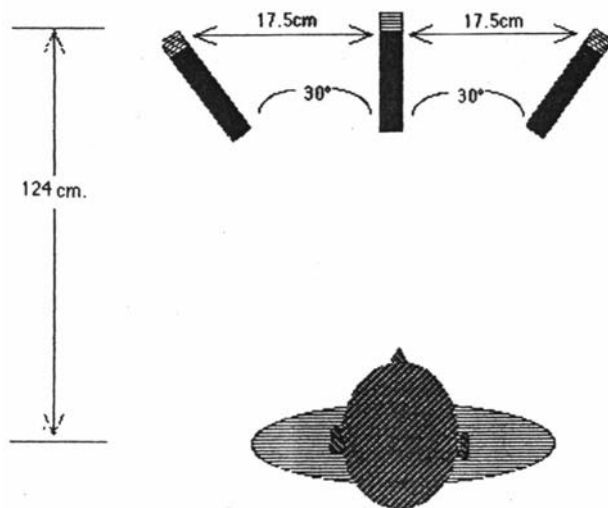


Abb. 5.18: Klepko-Array (nach [Klepko, 1997])

5.2.8 Atmo-Kreuz

Das Atmo-Kreuz [Theile, 2000] besteht aus einer Anordnung von 4 Nierenkapseln, deren Abstände untereinander (21-25cm) auf eine Dimensionierung gemäß der sog. ‚Williamsskurven‘ zurückgehen. (siehe [Williams 1987]). Gemäß Angabe der Erfinder eignet sich das Atmo-Kreuz (auch als ‚IRT-Cross‘ bekannt) sehr gut für die Surround-Aufnahme von Atmosphären, Geräuschen, etc. oder im Bereich der (Musik-) Aufnahme auch als Ergänzung zu einem Front-System zur Abbildung des Diffusschallanteils im Raum. Hierbei werden die nach vorne ausgerichteten Nieren auf die vorderen Wiedergabekanäle aufgeschaltet um eine gute ‚Einhüllung‘ des Hörers zu gewährleisten.

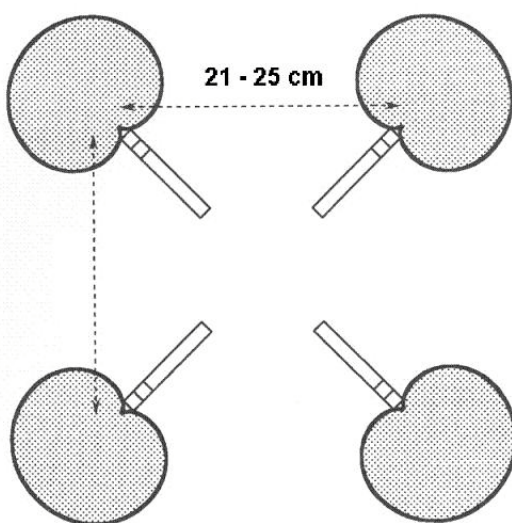


Abb. 5.19: IRT- od. Atmo-Kreuz, bestehend aus 4 Nieren (jeweils 21-25cm Abstand, 90°); (aus [Eargle, 2004])

5.2.9 Hamasaki-Square

Hierbei handelt es sich um die rechteckige Anordnung von 4 seitlich ausgerichteten Mikrofonen mit 8-er Charakteristik, die ebenfalls in erster Linie für die Aufnahme von Atmosphären bzw. Diffus-schall konzipiert ist. Ähnlich wie beim IRT-Kreuz kann die Information der vorderen Mikrofone entweder direkt auf die Front-Kanäle L,R aufgeschaltet bzw. zugemischt werden, oder frei auf den seitlichen Lautsprecherbasen L-LS bzw. R-RS ‚gepannt‘ werden.

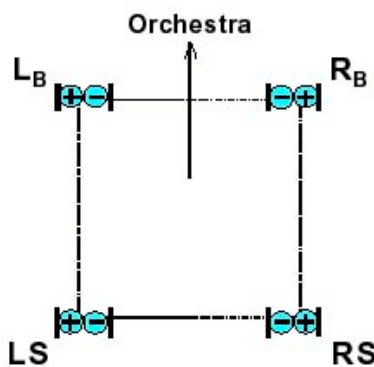


Fig. 38:

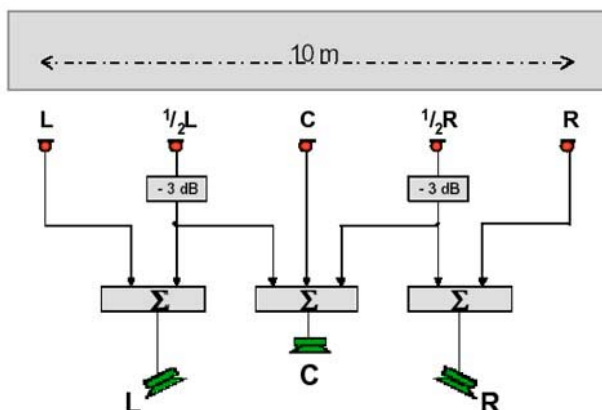
“Hamasaki-Square” for spatial imaging

In [54], [63] the square arrangement is proposed to be configured with figure of eight microphones (spacing $d = 1$ m). The null point of the directional characteristic of each microphone is facing the stage in order to reduce the energy of direct sound as much as possible. For lateral sound a pure A/B-microphone is effective. The array has been placed “at a very high position in the concert hall where the sound is very diffused and energy from direct sound is diminished” [54]. The microphones L_R and R_R are routed to channels L and R or panned between L-LS and R-RS, the microphones LS and RS are routed to channels LS and RS. See also Fig. 41.

Abb. 5.20: Hamasaki-Square (,Fig. 38‘ aus [Theile 2000])

Betrag der ursprünglich vorgeschlagene Kapselabstand noch 1m, so wurde dieser – in Hinblick auf die bessere Dekorrelation der Signale (und damit einhergehend verbesserten Räumlichkeitseindruck) - bei einer späteren Modifikation [Hamasaki et al., 2003] auf 2-3m erweitert. Außerdem schlagen Hamasaki und Hiyama als eine mögliche Alternative nun vor, die rückwärtigen 8-en durch nach hinten gerichtete Nieren zu ersetzen bzw. zu ergänzen.

5.2.10 Multiple-AB (,Mikrofonvorhang‘)

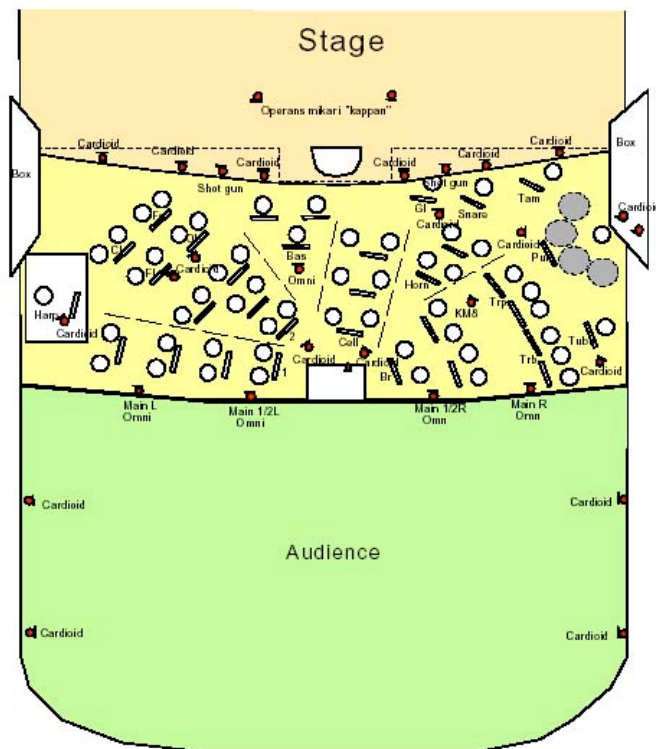


Five microphones in line and widely spaced (“Multiple-A/B”)

The microphones are distributed across the stage width, providing five negligibly correlated signals to produce three stable sources plus two phantom sound sources for directional imaging. The R/D-ratio and the balance of the orchestra elements can be controlled in certain limits by microphone positioning. Cardioid microphones may be used in order to reduce the indirect sound energy in the front channels.

Abb.5.21: ,Mikrofonvorhang‘ – schematischer Aufbau (aus [Theile, 2000])

Die obenstehende Anordnung von 5 Omni-Mikrofonen (siehe [Theile, 2000]) ist gut für die Gewinnung von Front-Channel Information geeignet und bei manchen Tonmeistern sehr beliebt. Zu diesen zählt u. a. auch Eberhard Sengpiel, der als Tonmeister für die TELDEC arbeitet. Untenstehend ist die Anwendung eines ‚Mikrofonvorhangs‘ für die Aufnahme einer Opernproduktion zu sehen, ergänzt durch Stützmikrofone im Orchester und an der Bühnenkante (Nierencharakteristik).



Für die Aufnahme von Surround-Signalen sind im Zuschauerraum 4 Nierenmikrofone bei den Seitenwänden platziert.

Abb. 5.22: Mikrofonierung für Opernaufnahme (Mikrofonvorhang, Stützmikrofone, Rampenmikrofone) (aus [Evers, 2003])

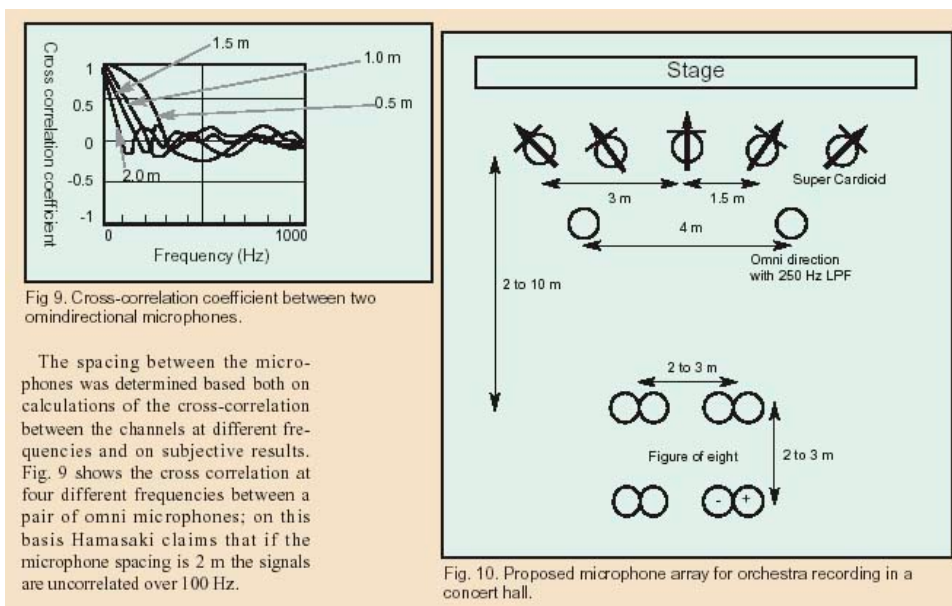


Fig 9. Cross-correlation coefficient between two omnidirectional microphones.

The spacing between the microphones was determined based both on calculations of the cross-correlation between the channels at different frequencies and on subjective results. Fig. 9 shows the cross correlation at four different frequencies between a pair of omni microphones; on this basis Hamasaki claims that if the microphone spacing is 2 m the signals are uncorrelated over 100 Hz.

Fig. 10. Proposed microphone array for orchestra recording in a concert hall.

Abb. 5.23: Mikrofonvorhang aus Supernieren, in Kombination mit Breit-AB für tiefe Frequenzen und Hamasaki-Square (nach [Hamasaki et al., 2003])

Tree-Anordnungen

Vorbemerkung: Letztlich gehen alle weiter unten dargestellten Tree-Anordnungen auf das historische Vorbild des DECCA-Trees oder DECCA-Triangles (Dreieck) zurück, das oftmals mit zwei weiteren Mikrofonen in Breit-AB Aufstellung (sog. ‚Outriggers‘) ergänzt wird, um - für den Raumeindruck so wichtige - stärker dekorrelierte Signalanteile einzufangen. Während für die Outriggers in aller Regel Mikrofone mit Omni-Charakteristik zum Einsatz kamen, wurde bei DECCA für die inneren Mikrofone des ‚Dreiecks‘ im Lauf der Jahre mit unterschiedlichsten Richtcharakteristiken experimentiert. Eine Zeitlang wurden - der besseren Kanaltrennung wegen - sogar akustische Absorberpanele zwischen die Mikrofone des Dreiecks gesetzt, welche deren ‚akustische Einzugsbereiche‘ dann ähnlich Tortenstücken in Segmente aufteilte. (siehe hierzu [Valin, 1994])

(Die untenstehend gewählten Bezeichnungen stammen fallweise nicht von den jeweiligen Erfindern, sondern wurden vom Autor - der Übersichtlichkeit halber - mit deren Namen versehen)

5.2.11 Fukada-Tree

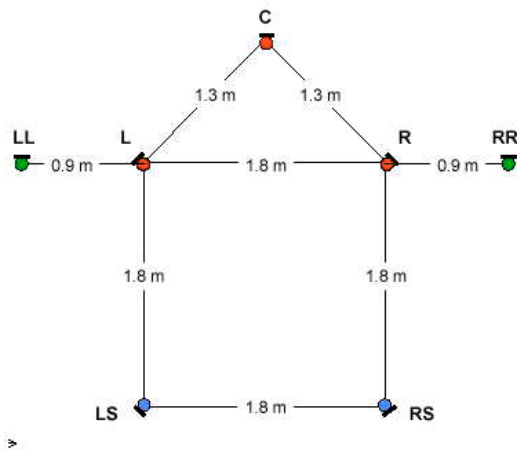
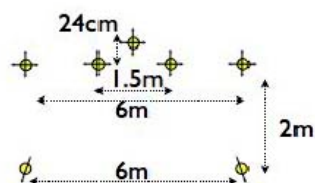


Abb. 5.24: Fukada-Tree (Grafik aus [Theile, 2000])

Der Fukada-Tree wurde 1999 erstmals vorgestellt [Sawaguchi et al., 1999] und ist – zumindest betreffend die Front-Anordnung – stark an das historische Vorbild des DECCA-Trees angelehnt: die Mikrofone L,C,R,LS,RS sind Nieren, die links und rechts flankierenden sog. ‚Outriggers‘ haben Omni-Charakteristik.

Bei späteren Versionen des Fukada-Trees sind – ähnlich wie beim Hamasaki-Square – teilweise größere Mikrofonabstände anzutreffen, welche vermutlich ebenfalls in Hinblick auf bessere räumliche Abbildung gewählt wurden.



1. Fukada Tree

Five omni-directional microphones for front and two uni-directional (cardioid) microphones for rear.

Abb 5.25: Fukada-Tree (aus [Kamekawa et al., 2007])

In der Literatur und Papers div. Fachtagungen (AES, VDT) sind bei Tree-Anordnungen Vorschläge für die Kombinationen unterschiedlichster Mikrofontechniken anzutreffen. Dabei finden sich Anwendungen unterschiedlichster Richtcharakteristiken und Kapselabstände, die in vielen Fällen mehr mit der Experimentierfreude von Tonmeistern, denn mit dahinterstehenden theoretisch-psychoakustischen Überlegungen zu tun haben scheinen.

5.2.12 Diverse Tree-Anordnungen

[Corey und Martin, 2003]

Front: 3 Halbnieren mit einer L-R Basisbreite zwischen 120-160cm, die Center-Halbniere 8-15cm nach vorne versetzt

Rear: zwei parallel nach oben oder nach hinter gerichtete Nieren mit Kapselabstand von (nur !) 30cm, Abstand Front-Rear: 60-90cm

[Christensen, 2003]:

Front: 3 Omni-Mikrofone angeordnet im gleichseitigen Dreieck, Kapselabstände 60-120cm

Rear: ORTF-Anordnung, oder ‚Breit AB‘ mit parallel ausgerichteten Nieren (Kapselabstand 2m)

Abstand Front-Rear: 8-10m

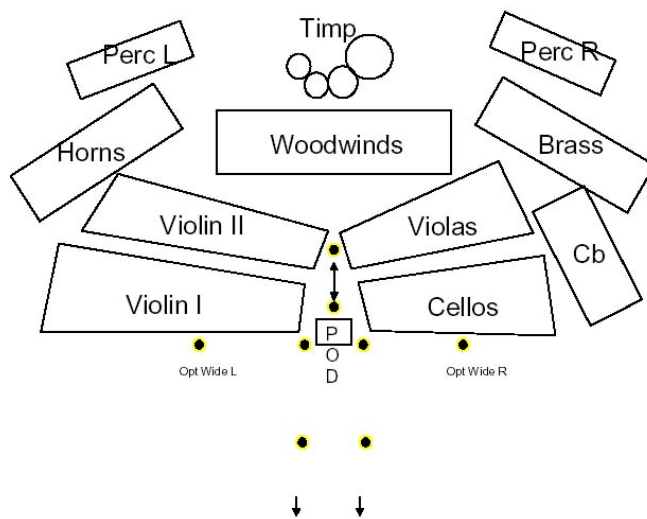
‚Polyhymnia-Pentagon‘, Philips-Klassik [Peters et al., 2007]:

Front: 3 Omni-Mikrofone im Dreieck angeordnet, L-R Basisbreite ca. 2.5m, Center-Mic ca. 60cm nach vorne Richtung Schallquelle versetzt

Rear: 2 Omni-Mikrofone, Breit-AB mit Kapselabstand von ca. 2.5m

Abstand Front-Rear: ca. 3.5m

[King, 2003]:



Bei R. King (Sony-Studios) findet sich der Vorschlag, für die Rear-Mics (LS, RS) diffusfeldentzernte Omnis zu verwenden, sowie das Center-Mikrofon um 20cm tiefer zu hängen, um den Präzedenz-Effekt zu nützen. Alle in der nebenstehenden Abbildung verwendeten Mikrofone sind Omnis.

Abb. 5.26: ‚Sony Studios‘-Tree (nach [King, 2003])

5.2.13 AB-PC (AB-Polycardioid Centerfill)

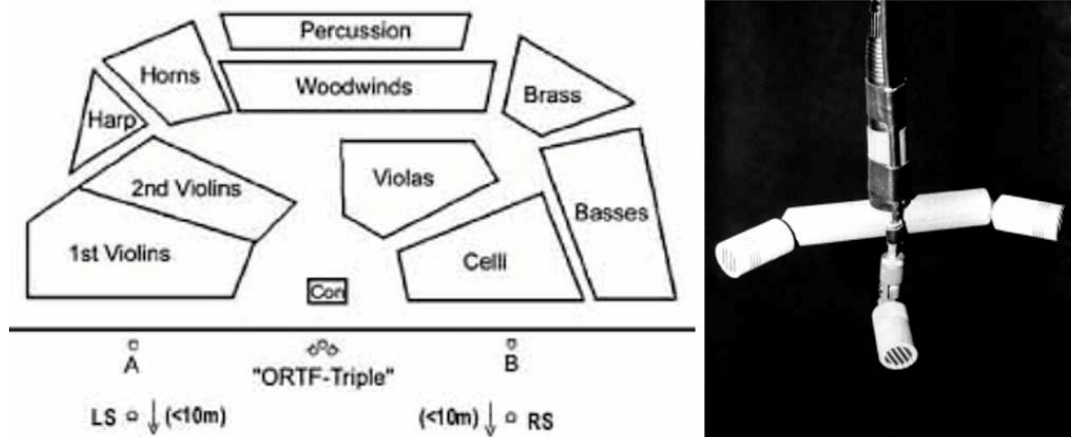


Abb. 5.27: links: AB-PC System; rechts: ORTF-Triple (aus [Pfanzagl-Cardone, 2002])

Das AB-PC System besteht aus einer ‚Breit AB‘ Anordnung ([diffusfeldentzernte] Omnis), mit einem Centerfillsystem z.B. in Form des sog. ORTF-Triple (siehe Abb. 5.27 rechts, welche ein Schoeps CMC6 ORTF-Mikrofon mit MK4 Mittenkapsel zeigt): bei diesem wird zum regulären ORTF-Hauptmikrofon eine ‚Center-Niere‘ hinzugefügt, deren Signal direkt auf den Center-Kanal der 5.1 Lautsprecheranordnung geführt wird. Die L, R Signale des ORTF-Systems werden bei ca. 10:30 (bzw. 13:30) Uhr in den L/R-Bus Mix ‚gepannt‘ und dienen somit quasi als Stütze für die ersten Pulte der Streicher links und rechts vom Dirigenten.

Um die Signalkorrelation bei tiefen Frequenzen gering zu halten, wird der Frequenzgang des ORTF-Triples nach unten hin beschnitten; alle drei Mikrofonsignale des Triples werden bei ca. 150Hz hochpassgefiltert um gute Dekorrelation im Bassbereich über die Front-Kanäle L-M-R zu erhalten.

Als Mikrofone zur Gewinnung der Signalanteile LS, RS kommen Nieren zum Einsatz, welche Richtung Saalrückwand gerichtet sind, und bis zu 10m vom Orchester entfernt platziert werden können (bei größerer Entfernung würde das Delay zwischen Front-Signal und Rear-Signal mehr als 30ms betragen, was psychoakustisch kritisch sein könnte (Ausprägung einer getrennten Wahrnehmung der Schallereignisse vorne/hinten; ‚Echo-Effekt‘) Weiter Details siehe [Pfanzagl-Cardone, 2002]. Ähnliche Mikrofonierungsvorschläge für Orchester (betreffend die Front-Anordnung) finden sich bei [Holman, 2001] und [Eargle, 2004], allerdings ohne die Erweiterung des ORTF-Systems zum ORTF-Triple mittels Mitten-Niere.

Als Alternative zum ORTF-Triple kommt für das Centerfill-System auch das bereits

weiter oben beschriebene BPT-Mikrofon in Frage. Bei der Kombination dieser beiden Systeme (BPT 3.0 und ‚Breit AB‘) ergänzen sich deren Eigenschaften in günstiger Weise: während das Breit-AB System aufgrund der hohen Signal-Dekorrelation durch eine große Räumlichkeit auszeichnet, bringt das koinzidente BPT-System eine hohe Lokalisations-Schärfe (siehe hierzu [Muraoka et al., 2007], S.2). Aus diesem Grund ist die Kombination beider Systeme von Vorteil für das sich ergebende Stereo- oder auch Surround-(Front)Signal.

In [Williams, 1992] wird vorgeschlagen, die Signale zweier Mikrofonpaare in folgender Weise zu kombinieren: Frequenzanteile oberhalb 700Hz werden von einem koinzidenten Mikrofonarray aufgenommen, Frequenzanteile unterhalb 700Hz über ein AB-Paar omnidirektionaler Mikrofone. In einer auf der Untersuchung von Williams aufbauenden Studie kommt [Preston, 1998] (aufgrund theoretischer Berechnungen) zum Ergebnis, dass die Lokalisationseigenschaften dieses „Frequency Dependent Hybrid Microphone Arrays“ tatsächlich den Lokalisationseigenschaften jedes der Einzelsysteme überlegen ist.

Die oben genannte Kombination aus ORTF-Triple bzw. BP-Triple geht prinzipiell in die gleiche Richtung, auch wenn dabei auf eine Splittung in Frequenzbänder und selektive Summierung der Signalanteile verzichtet wird. Anzumerken wäre in diesem Zusammenhang allerdings, dass Williams in seiner Veröffentlichung ‚Extremfälle‘ (wie Breit-AB und Stereosonic-System (=Blumlein-Paar) – also genau jene oben genannten Systeme) für seine streng an psychoakustischen Kriterien orientierten Ansatz ausschließt.

5.3 Trennkörper und ,3D'-Techniken

5.3.1 KFM 360 (Kugelflächenmikrofon)

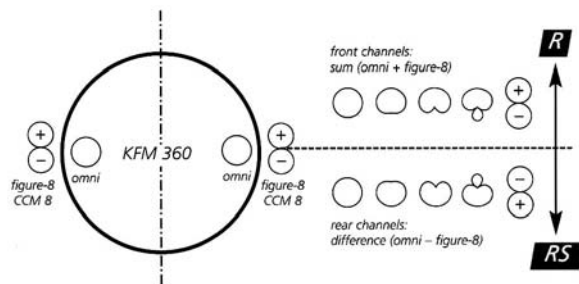


Abb. 5.28: KFM360 - Gewinnung der Signale R und RS mittels MS-Matrixierung

Zurückgehend auf eine Erfindung von [Theile, 1986], hatte das KFM 6U, die erste Version des Kugelmikrofon von Schoeps, einen Durchmesser von 20cm und war für Stereo-Aufnahmen konzipiert. Bei der Weiterentwicklung des KFM6 zum KFM 360 wurde sein Durchmesser auf 18cm reduziert um den Aufnahmewinkel auf ca. 120 Grad zu erweitern, weiters werden – für Surround - die links und rechts außen eingebauten Druckempfänger mit zwei Achten ergänzt, um mittels MS-Matrixierung die Gewinnung von Surround-Signalen möglich zu machen. Mit Hilfe einer Implementation der sog. Gerzon-Matrix in den zugehörigen ,DSP-4' Prozessor von Schoeps ist es möglich, auch ein Signal für den Center-Kanal eines 5.1 Surround-Signals zu erzeugen. Darüber hinaus bietet der DSP-4 noch verschiedene weitere Möglichkeiten zur Signalbearbeitung, so u.a. variabel einstellbare Richtcharakteristiken für die dematrixierten Mikrofonensignale, Aufschaltung eines Delays auf die Rear-Signale (um – bei Bedarf - die Lokalisation möglichst in der Front-Ebene zu halten [Gesetz der ersten Wellenfront]), etc.

5.3.2 Diverse Trennkörper-Surroundmikrofonensysteme

Ähnlich dem Kugelflächenmikrofon, jedoch – dank mehrerer diskreter Mikrofone - meist ohne Matrixierung bei der Signalgewinnung auskommend, sind die nachfolgenden Vorschläge, die in der Regel aus dem Bereich der Film- und Videoaufnahmetechnik stammen bzw. für diesen Bereich optimiert sind:

- Modulares Mic-Array für Surround Recording [Fox und McGregor, 2002]
- 4-Kapsel Surround-Anordnung nach [Iredale und Keller, 2003]

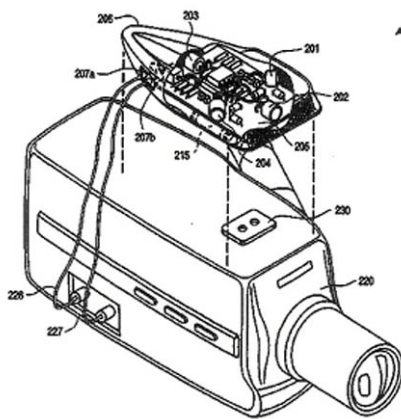


Abb. 5.29: 4-Kapsel Trennkörper Surround-Mikrofon nach [Iredale und Keller, 2003]

Die nebenstehend von [Iredale und Keller 2003] exemplarisch dargestellte Anordnung basiert auf 4 Kapseln (Elemente 201, 203, 204, 205): 3 für die Gewinnung der Front-Signale L,C,R, sowie eine Kapsel für (mono) Surround. Um die Aufzeichnung der Signale zu vereinfachen, wird eine Matrizierung auf 2 Signale vorgeschlagen, um mit ‚standard decoders‘ (hier: Dolby Surround) kompatibel zu bleiben.

5.3.3 Holophone ‚H2-Pro‘ und ‚PortaMic 5.1‘

Bei diesen beiden Systemen der Fa. Holophone sind mehrere Einzelmikrofone in die Oberfläche des Trennkörpers integriert, um ohne Matrizierung Surroundsignale bis 7.1 inkl. einem zusätzlichen Mikrofon für die Gewinnung eines ‚Höhensignals‘ (engl. ‚height-signal‘; d.h. von oben kommende Signalinformation) (H2-Pro) bzw. 5.1 (PortaMic) erzeugen zu können. Während sich das erste Modell von den Abmessungen (ca. 19x15cm) her am menschlichen Kopf orientiert, ist das kleinere PortaMic für Anwendungen auf Consumer-Videocameras konzipiert und deswegen vom Durchmesser her weit kleiner, wodurch die Wirksamkeit als Trennkörper allerdings nur für hohe Frequenzen gegeben sein dürfte; d.h. bei der Wiedergabe ist geringe bis schwache Räumlichkeit zu erwarten, da für die tiefen Frequenzen vermutlich hohe Signalkorrelation besteht. [Holophone, 2007]

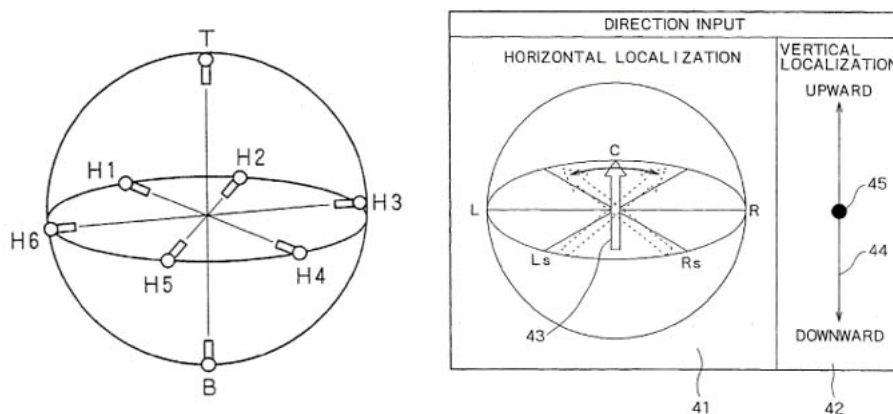


Abb. 5.30: Kugelanordnung; Patent der Fa. SONY [Ito, 2001]

5.3.4 Sony-Kugelanordnung

Dem Holophone ‚H2-Pro‘ nicht ganz unähnlich ist eine von Akitaka Ito im Rahmen eines Patents der Fa. SONY [Ito, 2001] vorgeschlagene System, bestehend aus 8 Mikrofonen, die auf einer Kugeloberfläche angeordnet sind. Davon sind 6 in der horizontalen, und je eines oben und unten für die Lokalisation in der vertikalen Ebene positioniert. In einem nachgeschalteten Prozessor kann mittels Signalbearbeitung horizontales und vertikales ‚Panning‘ der Schallquelle ad libitum durchgeführt werden.

5.3.5 Multichannel Microphone Array

Dem SONY Kugelflächen-Array sehr ähnlich ist das von Johnston und Lam vorgeschlagene 7-Kanal Array, bei dem 5 Hypernieren-Mikrofone mit zwischenliegenden Öffnungswinkeln von je 72 Grad in der horizontalen Ebene angeordnet sind. Gemeinsam mit zwei vertikal ausgerichteten Mikrofonen (up - down) liegen die die Kapseln aller Mikrofone auf einer virtuellen Kugeloberfläche mit einem Durchmesser von ca. 29cm. Normalerweise wird diese Mikrofonanordnung in einer typischen ‚best seat‘ Position im Aufnahmezimmer, in ca. 3m Höhe platziert. [Eargle 2004, S. 261]

Gemäß Angabe der Autoren Johnston-Lam wurde die Anordnung so gewählt, um Laufzeit- und Pegeldifferenzen in optimaler Weise für Hörer im (oder nahe) dem Zentrum der Lautsprecher-Wiedergabeanordnung zu reproduzieren. Laut ihrer Angabe sind die so gewonnenen Signale gut mit der 5.1 Surround-Wiedergabeanordnung nach ITU-R 775.1 kompatibel, idealerweise sollten allerdings die Lautsprecher - so wie die Mikrofone - jeweils 72 Grad voneinander positioniert sein.



5.3.6 Eigenmike

In [Meyer und Agnello, 2003] wird eine Anordnung von 24 Miniatur-Mikrofonen, die in die Oberfläche einer Kugel mit nur ca. 7,5cm Durchmesser integriert sind, beschrieben. Mittels Postprocessing der Signale lassen sich Richtcharakteristiken bis 3. Ordnung realisieren.

Abb. 5.31: ‚Eigenmike‘ (nach [Meyer und Agnello, 2003])

Beim Eigenmike bestimmt die Anzahl der Mikrofone die realisierbare Ordnung:

Anzahl der Mics = $(N+1)^2$ mit N=Mikrofon Ordnung (Richtcharakteristik) (vgl.: mit den 4 Kapseln des Soundfield-Mikrofons lassen sich nur Richtcharakteristiken 1. Ordnung realisieren)

Aufgrund akustisch-physikalische Gegebenheiten sind mit den 24 Mikrofonen des ‚Eigenmike‘ oberhalb von ca. 1.5kHz Richtcharakteristiken 3. Ordnung möglich, darunter 2. Ordnung, sowie unterhalb ca. 700 Hz nurmehr 1. Ordnung.

Eine vergleichbare Anordnung ist unter dem internationalen Patent WO 03/061636 A1 registriert: Gary ELKO, Robert KUBLI, Jens MEYER (2003) „Audio system based on at least second-order Eigenbeams“

5.3.7 OSIS-System (Optimal Sound Image Space)

Abgeleitet von seinem für 2-Kanal Stereo-Aufnahme vorgeschlagenen ‚OSS‘ (‚optimales Stereo Signal‘) System, schlägt Jecklin nun für OSIS (‚Optimal Sound Image Space‘) [Jecklin, 2002] die Erweiterung des Systems auf eine Trennscheibe für das Front-Signal sowie einer Trennscheibe für das Rear-Signal vor.

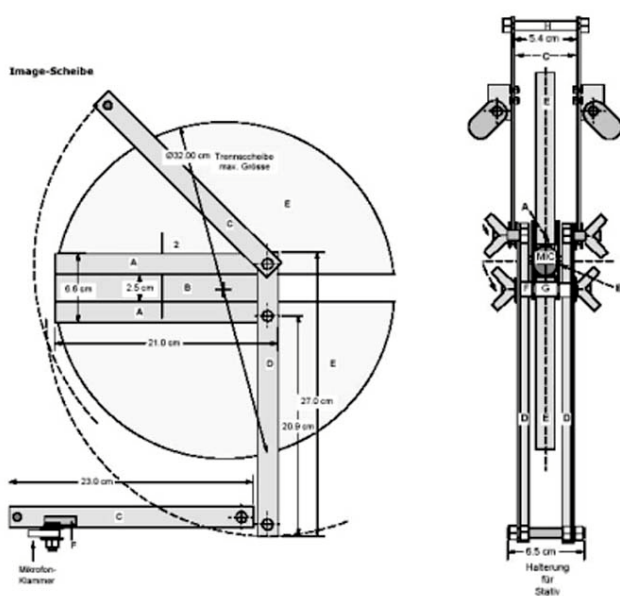


Abb. 5.32: Image-Scheibe des ‚OSIS-System‘ (nach [Jecklin, 2002])

Wie schon beim OSS System wird für die ‚Image‘-Scheibe (Front) ein Durchmesser von 32cm gewählt und die beiden diffusfeld-entzerrten Druckempfänger mit Ohrenabstand (ca. 17cm) links und rechts von dieser angebracht und

so ausgerichtet, dass sie die Schallquelle ‚überschießen‘ (d.h. ggf. leicht nach oben gewinkelt). Außerdem ist für die Gewinnung eines Center-Signals ein Richtmikrofon in die Image-Scheibe integriert.

Der Durchmesser der ‚Space‘-Scheibe (Rear) wird mit 28cm vorgeschlagen und statt der Omnis kommen zwei nach hinten gerichtete Nieren-Mikrofone zum Einsatz.

Gemäß Jecklin’s Motto „No Image in Space and no Space in Image“ zielt die oben beschriebene Anordnung darauf ab, mit dem Rear-System möglichst nur Raumanteil (Diffusschall) einzufangen („no image in space“) und mit der Front-Anordnung („Image‘-Scheibe) nur Direktschall-Anteile („no space in image“).

Wie bereits beim OSS System vom Erfinder angegeben [Jecklin, 1981], werden tiefe Frequenzen um die Trennkörper-Scheiben gebeugt, sodass Signalanteile unter ca. 200Hz praktisch mono wiedergegeben werden. Ähnliches gilt auch zumindest für die Front-Anordnung des OSIS Systems, was der Ausbildung eines überzeugenden Raumeindrucks beim Hörer abträglich ist, welcher gute Dekorrelation tiefer Frequenzen unterhalb von ca. 500Hz und mehr noch unterhalb von 200Hz erfordert (siehe hierzu [Hidaka et al.,1995], sowie [Morimoto and Meakawa, 1988])

5.3.8 Pan Ambiophonic 2D/3D System



Fig. 9 - New design of the PanAmbiophone. Two pairs of small pressure microphone elements are coincident at 'ear' positions. A discrete soundfield array is mounted atop – a total of eight recording channels for PerAmbio 3D (with height).

Abb. 5.33: ‚Pan Ambiophonic 2D/3D System‘

Miller kombiniert die Signale einer diskret aus Einzelmikrofonen aufgebauten Ambiophonie-Anordnung mit denen zweier Druckempfänger, die an den Ohrpositionen des nebenstehend gezeigten (dem menschlichen Kopf nachempfundenen) Trennkörper-Objekts [ergänzt mit einem schallabschattenden ‚Trichter‘] positioniert sind.

Die so gewonnene Signal-Mischung wird für ‚Ambiophonic 2.0‘ über zwei Lautsprecher mit einer Basisbreite von lediglich 10-20° wiedergegeben und liefert gemäß Angabe des Erfinders eine 120° weite Abbildung. Das Stereo-System kann durch zusätzliche ‚Ambience‘-Lautsprecher erweitert werden, über welche Signale abgestrahlt werden, die mittels Impulse-Response Convolution gewonnen wurden.

Im Rahmen des PerAmbio3D/2D Systems werden die oben angesprochenen Signale über insgesamt 10 Lautsprecher (angeordnet gemäß Abb. 5.34, Fig. C) wiedergegeben, wobei auf die genaue Platzierung der oberhalb des/der Hörer positionierten Lautsprecher durch entsprechende Programmierung im Decoder Rücksicht genommen werden kann.

Miller merkt weiters an, dass ein Verrücken aus dem ‚sweet spot‘ um 26% nach hinten in einer Lautsprecheranordnung relativ zum Hörer gemäß dem Standard 5.1-Surround Lautsprecher-Layout resultiert. [Miller, 2004a]

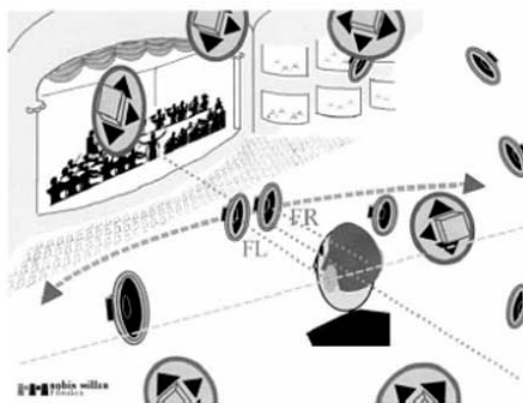


Fig.A - **Ambiophonic 2.0** turns stereo "inside-out," reproducing a 120° wide, natural sounding front stage using 10-20°-spaced speakers. Add 3D ambience speakers using impulse response convolution. Stereo recordings play compatibly in Ambio.

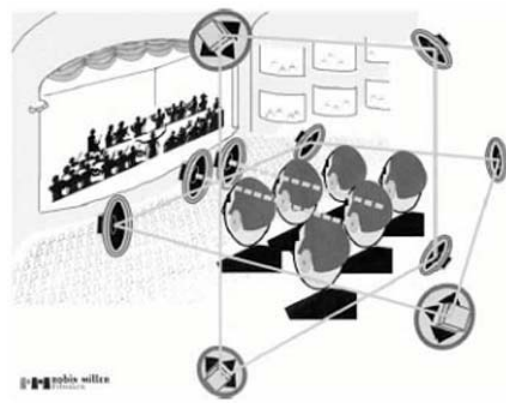


Fig.C – **PerAmbio 3D/2D** (Pat.pend.) plays both Ambiophonic and 3D (with height) recordings using 10 speakers. Height speakers may be flexibly positioned and their coordinates programmed in the decoder. Sitting back 26% of the speaker diameter results in standard ITU 5.1 angles for compatibility.

Abb. 5.34: ‚Pan Ambiophonic 2D/3D System‘ – Fig. A und Fig C (aus [Miller, 2004a])

5.4 Allgemeine Betrachtungen über Mikrofon-Arrays

Bei [Woszczyk, 1992] finden sich Gedanken zu Mikrofon-Arrays, die für Musikaufnahme optimiert sind:

Sowohl koinzidente als auch Arrays mit großen Mikrofonabständen sind geeignet, um unerwünschte Interferenzen (Phasenauslöschungen im Spektrum) zu vermeiden. Die Signale, die ein ‚spaced‘ Mikrofonarray (d.h. mit räumlich voneinander getrennten Kapseln) aufnimmt, können oberhalb jener Frequenz als statistisch unabhängig betrachtet werden, deren halbe Wellenlänge dem Abstand zwischen den Mikrofonkapseln entspricht. In diesem Zusammenhang werden auch die Fachausdrücke ‚correlation length‘ und ‚coherence interval‘ verwendet.

Stereomikrofon-Arrays können auch als Mono-Arrays aufgefasst werden, deren Einzelmikrofone hart L und R gepannt sind; erweitert werden kann ein solches Array z.B. auch durch die Einfügung weiterer Mono-Arrays, die zwischen den Bussen L, R gepannt werden.

Mit einer Kombination aus koinzidenten und ‚spaced‘ Arrays deren Verhalten jeweils über schmale Frequenzbänder optimiert wurde, lassen sich Ergebnisse in punkto gewünschter Richtcharakteristik und/oder gleichmäßiger Abdeckung erzielen, die mit einem konventionellen Einzelmikrofon nicht zu erreichen wären. Weiters stellt Woszczyk fest, dass Arrays z.B. auch für die Aufnahme großer Schallquellen – wie Orchester – besonders gut geeignet sind.

In [Burkhard et al., 1991] wird analysiert, wie sich die Beschaffenheit des für die Wiedergabe verwendeten Lautsprecher-Arrays auf den entstehenden Raumeindruck, inklusive Richtungseindruck, Einhüllung und Deutlichkeit der Abbildung der Schallquellenmitte auswirkt.

5.5 Zusammenfassung

Wie bereits im Zusammenhang mit Stereo-Mikrofontechniken festgestellt wurde, zeichnen sich koinzidente („one-point“) Aufnahmetechniken oftmals durch gute Lokalisationseigenschaften aus, das von ihnen erfasste Signal enthält jedoch – systembedingt – keine Laufzeitinformation. Das in Abschnitt 5.1.1 beschriebene Soundfieldmikrofon ist zwar in der Lage, das in einem Raumpunkt auftretende Schallsignal ‚3-dimensional‘ zu erfassen und – mittels geeigneter Matrizierung und Dematrizierung – über frei wählbare Mehrkanal-Lautsprecheranordnungen optimiert wiederzugeben, allerdings führt dies auch – im Idealfall – zu einer perfekten Rekonstruktion des ursprünglichen Schallsignals wiederum nur an einem Punkt des Wiedergaberaumes.

Da der Hörer jedoch im Normalfall mit beiden Ohren hört, ist somit so gut wie ausgeschlossen, dass das Signal an beiden Ohren ‚korrekt‘ sein kann. Dies erklärt möglicherweise auch die Diskrepanz zwischen der theoretisch ‚perfekten‘ Signalreproduktion, die das Ambisonic-Verfahren gemäß des zugrunde liegenden Prinzips der Signalerfassung zu liefern im Stande ist, und der schlechten Beurteilung, welche die auf diesem Verfahren beruhenden Surround-Mikrofonssysteme bei Hörvergleichen erfahren haben (vgl. [Camerer und Sodl, 2001], [Hildebrandt und Braun, 2000]).

Allerdings sind koinzidente Mikrofonssysteme immer dann anderen Mikrofontechniken überlegen, wenn es darum geht, Schallquellen mit geringer Ausdehnung (z. B. bei Einzelinstrumenten) und/oder komplexer Abstrahlcharakteristik möglichst unverfälscht abzubilden: bei Systemen, die mehrere Einzelmikrofone verwenden läuft man hierbei nämlich Gefahr, dass durch die frequenzabhängige Abstrahlcharakteristik und damit einhergehenden unterschiedlichen Pegeln an verschiedenen Mikrofonen Lokalisationsverzeichnungen in Form eines ‚Wanderns‘ oder ‚Springens‘ der Schallquelle auftreten.

Der Großteil der in Abschnitt 5.2 beschriebenen ‚Spaced Arrays‘ basiert auf der Anwendung des psychoakustischen Prinzips der Schallquellenlokalisierung gemäß Laufzeitunterschieden, wie sie am menschlichen Kopf auftreten (siehe Abschnitt 5.2.3. das ‚ABC 5‘ System) - in den meisten Fällen zusätzlich unterstützt durch Pegelunterschiede, die durch den Einsatz von Mikrofonen mit Richtcharakteristik erzielt werden (siehe 5.2.4 ‚INA 5‘ System, 5.2.5 ‚OCT‘-System, 5.2.11 Fukada-Tree, sowie

einige der Tree-Anordnungen, die in Abschnitt 5.2.12 beschrieben sind). Manche von den vorgestellten Systemen sind eher für die Aufnahme des Diffusschallfeldes gedacht (siehe 5.2.8 Atmo-Kreuz und 5.2.9 Hamasaki-Square), andere hingegen verwenden große Mikrofonabstände und zielen darauf ab, die Schallquelle in ihrer gesamten Breite möglichst ‚komplett‘ zu erfassen (siehe 5.2.10 Multiple-AB (Mikrofonvorhang), wie auch 5.2.13 ‚AB-PC‘ (AB-Polycardioid Centerfill)) und sind somit durchaus auch für die Anwendung als Surround-Mikrofon für die Aufnahme der Frontsignale geeignet.

Im Gegensatz zu den ‚Spaced Arrays‘ mit kleinen Kapselabständen treten bei den beiden zuletzt genannten Systeme aufgrund der großen Kapselabstände bei den Signalen der Druckempfänger nicht nur wesentlich größere Laufzeitunterschiede, sondern auch relevante Pegelunterschiede auf, die – im Falle z.B. einer Orchesteraufnahme - durch die relative Nähe zu unterschiedlichen Instrumentengruppen bedingt sind. Dies kommt in der Regel der Eindeutigkeit der Signallokalisierung zugute, da hier Laufzeit- und Pegelunterschiede positiv zusammenwirken. In Fällen einer ‚Breit-AB‘ Anordnung muss - um das Auftreten eines ‚Mittenlochs‘ zu vermeiden – zusätzlich ein Centerfill-System eingesetzt werden (Details siehe Abschnitt 5.2.13 AB-PC). Somit zielt dieser Ansatz der ‚Breit-AB‘ Mikrofonierung (oder auch des Mikrofonvorhangs) darauf ab, das Schallfeld entlang einer Linie zu erfassen, welches dann im Hörraum (aufgrund geringerer Lautsprecher-Basisbreite) in der Regel mit ‚geschrumpften‘ Dimensionen wiedergegeben wird.

Das in Abschnitt 5.2.1 beschriebene ‚Circular Microphone Array‘ hingegen verfolgt einen grundlegend anderen Ansatz und zielt darauf ab, mit einer vergleichsweise hohen Anzahl an Einzelmikrofonen das Schallfeld an einem geeigneten Ort im Aufnahmeraum so zu erfassen, dass die Einzelsignale mit einem Prozessor in einer Weise nachbearbeitet werden können, die es ermöglicht, das ursprüngliche Schallfeld bei (Mehrkanal-)Wiedergabe möglichst authentisch zu reproduzieren. Für die Anwendung im Bereich der Musikaufnahme scheint dieses System – u.a. aufgrund der relativ großen Anzahl an benötigten Wandlern für das Hauptsystem - jedoch nur beschränkt praktikabel.

Das geometrisch komplexer aufgebaute ‚Trinnov-Array‘ (Abschnitt 5.2.2) benötigt ebenfalls einen Prozessor, um die Mikrofonensignale für eine Wiedergabe z. B. im 5.1 Surround-Format aufzubereiten.

In Abschnitt 5.3 sind diverse Trennkörper- und ‚3D‘-Techniken beschrieben, von denen die meisten darauf abzielen, das Schallfeld in einer Weise zu erfassen, die auch die getrennte Aufnahme von Schallanteilen von oben und unten ermöglicht. Einige von diesen verwenden hierzu Schallwandler, die – zur Ausnutzung von Abschattungseffekten - in eine Kugeloberfläche oder einen der Form des menschlichen Kopfes nachempfundenen Trennkörper integriert sind (siehe 5.3.6 Eigenmike, sowie 5.3.3 Holophone ‚H2-Pro‘); andere beschränken sich auf die Anordnung mehrerer Mikrofone auf einer (virtuellen) Kugeloberfläche (siehe 5.3.4 ‚Sony-Kugelanordnung‘ und 5.3.5 Multichannel Microphone Array). Bei diesen Techniken werden zur Schallquellen Lokalisation Laufzeiteffekte und mitunter auch die Richtcharakteristik der verwendeten Mikrofone genutzt (siehe 5.3.5 Multichannel Microphone Array), fallweise auch unterstützt durch Signalprozessing (siehe 5.3.4 Sony-Kugelanordnung, sowie 5.3.6 Eigenmike).

Eine Reduktion des 3-dimensionalen kopfähnlichen Trennkörpers auf eine quasi 2-dimensionale ‚Trennebene‘ findet sich bei dem in Abschnitt 5.3.7 beschriebenen OSIS-System (Optimal Sound Image Space), welches allerdings aus zwei getrennten Systemen für die Aufnahme der von vorne und von hinten einfallenden Schallinformation aufgebaut ist. Ein ähnlich individueller aufnahmetechnischen Ansatz findet sich beim ‚Pan Ambiophonic 2D/3D System‘ (Abschnitt 5.3.8), welches die Signale einer diskret aus Einzelmikrofonen aufgebauten Ambiophonie-Anordnung mit denen zweier Druckempfänger kombiniert, die an den Ohrpositionen eines (dem menschlichen Kopf nachempfundenen) Trennkörper-Objekts - ergänzt mit einem schallabschattenden Trichter - positioniert sind.

6 Kunstkopf-Aufnahmen

6.1 Verwendung eines Kunstkopfs als ‚Bezugsnormal‘

Wie bereits früher erwähnt, war im Rahmen der Untersuchung ein Kunstkopf zum Einsatz gekommen, um das mit diesem im Konzertsaal aufgenommene binaurale Musiksignal mit dem im 5.1 Wiedergaberaum auftretenden Binaural-Signal vergleichen zu können und zwar sowohl im A/B-Hörvergleich, wie auch rein messtechnisch.

Da zu Beginn dieser Untersuchung nicht klar war (bzw. von verschiedenen Fachleuten unterschiedliche Lehrmeinungen vertreten werden: Günther Theile plädiert für hohe Korrelation, David Griesinger hingegen tritt für Dekorrelation der vorderen und hinteren Surround-Signalen, sowie der hinteren Surround-Signale untereinander ein), welcher Korrelationsgrad zwischen den einzelnen Kanälen des Surroundsignals anzustreben wäre, um eine möglichst ‚natürliche‘ Abbildung eines Klangkörpers zu erzielen, schien es sinnvoll, einen Kunstkopf als ‚Bezugsnormal‘ – zumindest für die messtechnische Betrachtung - heranzuziehen.

Dieser Gedanke drängt sich auch deshalb auf, weil die 5.1 Lautsprecher-Signale letztlich ‚am‘ menschlichen Kopf (bzw. dessen Hörmechanismus) wirksam werden sollen - der Mensch also quasi als ‚Maß der Dinge‘. (Dieser Ansatz sollte im Grunde genommen richtig sein, basiert doch die Argumentation bzw. Berechnung des passenden Kapselabstandes bei ‚Klein-AB‘ Techniken auf der Berücksichtigung des Ohrenabstands beim Menschen; somit müssten bei Elefanten oder Vögeln für eine ‚gehörriichtige‘ Abbildung entsprechend davon drastisch abweichende Kapselabstände gewählt werden ...)

Der Einsatz des Kunstkopfs (abgekürzt: KK) bringt außerdem eine starke Verringerung des messtechnischen Aufwands mit sich: statt die frequenzabhängige Korrelation zwischen jeweils zwei der fünf ‚Fullrange‘-Kanäle bestimmen zu müssen, genügt es nun, die frequenzabhängige Korrelation der Binaural-Signale zu vergleichen.

Beispiele aus der Literatur zur Verwendung eines Kunstkopfs als ‚Bezugsnormal‘ finden sich u.a. bei [Peters et al., 2007], [Braun et al., 1996], [Olive et al., 1994], [Toole, 1991].

Im Prinzip könnte die Kunstkopfaufnahme auch als ‚Bezugsnormal‘ oder Referenzsignal beim Hörvergleich abwechselnd mit Lautsprecherwiedergabe verwendet werden, allerdings tritt dabei die Problematik auf, daß der Testhörer versuchen müsste, möglichst objektiv zwei (bzw. im Tripel-Vergleich 3) auf deutlich unterschiedlichen technischen Wegen reproduzierte Signale miteinander zu vergleichen: nämlich jeweils zwischen Referenz-Stimulus (Kunstkopf-Aufnahme dargeboten über Kopfhörer) und zu beurteilendem Stimulus (dargeboten über die Lautsprecheranlage) abzuwechseln.

Das wäre entweder mit ständigem Auf- und Absetzen eines Kopfhörers verbunden gewesen (ein zusätzlicher Ermüdungs- oder Ablenkungsfaktor für die Testpersonen), oder man hätte das binaurale KK-Signal entsprechend entzerren müssen, um es für die Wiedergabe auf einer 5.1 Lautsprecher Abhöranlage kompatibel zu machen.

Geräte bzw. Signalbearbeitungs-Algorithmen dafür gibt es (siehe z.B. LEXICON „Logic 7“-Prozessor bzw. Literatur [Griesinger, 1988], allerdings ist davon auszugehen, daß die hierbei vorgenommene Signal-Verarbeitung eine Veränderung und somit leichte Verfälschung des Nutzsignales mit sich bringt.

Aus den oben angeführten Gründen entschied der Autor, die Testpersonen einen rein kopfhörerbasierten A/B/REF-Hörvergleich durchführen zu lassen und auf den Wechsel zwischen Kopfhörer- und Lautsprecherwiedergabe innerhalb desselben Hörvergleichs zu verzichten.

6.2 Studien betreffend die Qualität verschiedener Kunstköpfe

In einer Studie des Acoustics Laboratory der dänischen Aalborg Universität [Moller et al., 1999] wurden 8 Kunstköpfe in Hinblick auf ihre Lokalisationseigenschaften unter Einsatz von 20 Testhörern bewertet.

Der Testteil A (‚Real –Life‘) wurde als ‚Identifikations Experiment‘ durchgeführt, bei dem die Testhörer das Schallereignis einem von 19 Lautsprechern zuordnen sollten (14 davon waren auf einer (virtuellen) Kugeloberfläche in 1 Meter Abstand vom Hörer positioniert, 7 davon auf der Median-Ebene; die restlichen 5 Lautsprecher befanden sich in Entfernungen bis zu 5 Meter).

Als Testsignal diente weibliche Sprache.

In den Testteilen B und C wurde den Hörern über einen Beyerdynamic DT 990 (bzw. STAX für die HMS Kunstköpfe) das in der oben beschriebenen Testanordnung mittels der verschiedenen Kunstköpfe aufgenommene Binaural-Signal zur Beurteilung vorgespielt. Für den Testteil B erfolgte die Aufnahme – soweit vorhanden - mit den in den Kunstköpfen eingebauten Mikrofonen, für den Testteil C wurden Miniaturmikrofone unmittelbar außerhalb des Gehörgangs platziert.

Um Verfälschungen bei der Beurteilung über Kopfhörer möglichst gering zu halten, wurde die kombinierte Transferfunktion der Signalübertragungsstrecke Kopfhörer-Kunstkopfmikrofon vorher vermessen (von Kopfhörer-Eingang zu Kunstkopfmikrofon-Ausgang), und bei der Wiedergabe die Testsignale entsprechend amplitudenkorrigiert.

Für den Test kamen folgende Kunstköpfe bzw. Kunstkopf-Torso Kombinationen zum Einsatz:

- KEMAR (Knowles Electronics Mannikin for Acoustic Research) Head-and-Torso Simulator; (bestückbar mit 4 unterschiedlichen Außenohr-Größen; ergänzt mit Brüel & Kjaer 4134 Mikrofonen)
- Neumann KU 80i [nur Kunstkopf]
- Neumann KU 81i (diffusfeld-entzerrt) [nur Kunstkopf]
- HMS I (Fa. HEAD Acoustics / Techn. Hochschule Aachen) [Kopf-und-Torso]
(ergänzt mit B&K 4166)
- HMS II (Fa. HEAD Acoustics / Techn. Hochschule Aachen) [Kopf-und-Torso]
(erg. mit B&K 4165)
- Bruel & Kjaer ‚4128‘ (ergänzt mit 4134 Mikrofonen) [Kopf-und-Torso]
- Bruel & Kjaer ‚5930‘ (ergänzt mit 4009 Mikrofonen) [Kopf-und-Torso]
- Artificial Head der University of Toronto (auf KEMAR-Torso, ergänzt mit Sennheiser MKE 4-211-2 Mikrofonen)

Im Rahmen der Experimente wurden die Kunstköpfe hinsichtlich der Lokalisationsverfälschung in der Horizontale, wie auch Vertikale untersucht. Teststimuli wurden dabei auf die Hauptbereiche (‚cones‘) LINKS, VORNE, OBEN, HINTEN, RECHTS aufgeteilt. Wurde ein Stimulus von einem Hörer innerhalb eines Hauptbereichs falsch lokalisiert, so wurde dies als ‚within cone error‘ erfasst, wurde der Stimulus sogar in einem falschen Hauptbereich geortet, so galt dies als ‚out of cone error‘.

Gesondert ausgewiesen wurden außerdem die Fehlerraten hinsichtlich der geschätzten Entfernung, sowie die – nicht nur beim realen Hören, sondern vor allem auch - bei Kunstkopfwiedergabe häufig auftretenden Lokalisationsfehler im Zusammenhang mit der vorne/hinten Vertauschung von Klangreizen („Median-Plane Errors“).

Um der Komplexität bei der Qualitätsbewertung von Kunstköpfen gerecht zu werden, wurde als Ergebnis nicht einfach eine Reihung aufgrund eines ev. errechenbaren numerischen Mittelwertes angegeben, sondern die Resultate der Hörtestauswertung im Detail für die verschiedenen erfassten Parameter angeführt.

Auffallend scheinen dabei folgende Resultate:

Obwohl der HMS I als Sieger mit dem geringsten Prozentsatz an „Median Plane Errors“ hervorging, und auch hinsichtlich der Entfernungsfehler an zweiter Stelle liegt, hatte er den höchsten Prozentsatz an „out of cone“ Fehllokalisationen, zugleich aber auch den geringsten Prozentsatz an „within cone errors“.

Der HMS II andererseits lag in Hinblick auf die „Median Plane Errors“ in den verschiedenen Testteilen an vorletzter bzw. letzter Stelle und ging in Punkto Entfernungsfehler – zumindest in einem Teil des Tests (Abschnitt B) als Sieger hervor.

Diese großen Unterschiede im Verhalten der beiden Modelle HMS I und II lassen sich teilweise damit begründen, dass für den HMS I bei einem realen Menschen „Maß genommen“ wurde, während es sich beim HMS II um ein Kopfmodell mit stark vereinfachter Geometrie handelt. Die überproportional hohe „out of cone“ Fehlerrate beim HMS I wurde denn auch auf einen überdurchschnittlich großen Ohrenabstand (interaural time delay-Messung) zurückgeführt, der offenbar bei der Mehrheit der Testpersonen zu einer Fehlortung führt. (siehe [Moller et al., 1999] S. 93)

Die „Out of cone“ und „within cone“ Fehlerrate liegt sonst bei allen Kunstköpfen in der selben Größenordnung, mit Ausnahme für die Neumann Kunstköpfe KU80i und KU81i: bei beiden liegt sie deutlich höher als bei den Konkurrenten, allerdings bietet sich dafür - laut Studie – folgende Erklärung an: die Neumann Kunstköpfe sind die einzigen Testobjekte, die über keinen Torso verfügen; die Abwesenheit von Beugungs-, Brechungs- und Reflexionserscheinungen an den Schultern dürfte dafür verantwortlich sein (Stimuli aus den Richtungen LINKS UNTEN sowie RECHTS UNTEN wurden von den Testpersonen oft als LINKS OBEN und RECHTS OBEN angegeben).

In Hinblick auf Entfernungsfehler ging der KU80 allerdings – für die zweite Hälfte des Tests (Testteil C) - als Sieger hervor, der KU81 lag dabei an 4. Stelle. (Das sollte sich positiv auf die mit dem Kunstkopf erzielbare ‚Tiefenstaffelung‘ z.B. bei einer Orchesteraufnahme auswirken.)

In Bezug auf ‚Median Plane Errors‘ wurde der KU80 letzter (im Testteil B), in der selben Kategorie belegt er allerdings Platz 4 im Testteil C.

Zum Vergleich: der KU81 belegt in dieser Kategorie die Plätze 4 (Testteil B) und 2 (Testteil C).

Nun zu einem Vergleich von KU80 und 81 mit den verbliebenen 4 Kunstköpfen: KEMAR 2 (bestückt mit Außenohr-Variante 2), sowie den B&K 4128 bzw. 5930 (wobei der Kunstkopf identisch ist; der Unterschied besteht darin, daß die Mikrofone beim 4128 eingebaut sind, während sie beim 5930-System direkt außen an der Eintrittsöffnung des Gehörgangs positioniert werden) und dem ‚Toronto‘-Kopf:

Im Testteil B, ‚Median Plane Error‘ sieht die Reihung wie folgt aus: (von ‚besser‘ nach ‚schlechter‘)

HMS I, Kemar 2, Toronto, KU81, B&K, HMS II, KU80

Im Testteil C, ‚Median Plane Error‘ ergibt sich folgende Reihung:

HMS I, KU81, Kemar 2, KU80, Toronto, B&K, HMS II

Im Testteil B, ‚Distance Error‘ sieht die Reihung wie folgt aus:

HMS II, HMS I, B&K, Toronto, Kemar 2, KU80, KU81

Im Testteil C, ‚Distance Error‘ sieht die Reihung wie folgt aus:

KU80, HMS I, Toronto, KU81, HMS II, Kemar 2, B&K

Betrachtet man nur die Qualitäts-Kriterien ‚Median Plane Error‘ und ‚Distance Error‘, so schneidet der HMS I vergleichsweise am besten ab: er belegt zwei Mal den 1., sowie zweimal den 2. Platz. Allerdings ist er auch der Kunstkopf mit einer ‚out of cone‘ Error-Rate, die etwa doppelt so hoch ist wie bei allen Konkurrenz-Produkten.(Ursachen: siehe Erklärung weiter oben)

Der HMS II belegt – außer im Testteil B, Distance Error – immer hintere Plätze.

Der B&K-Kunstkopf schafft es – ebenfalls außer im Testteil B, Distance Error – nie über den 5. Platz hinaus.

Der Toronto-Kopf schneidet durchschnittlich ab: zwei 3. und zwei 4. Plätze.

Der KU80 zählt in Hinblick auf ‚Median Plane Errors‘ eher zu den schlechteren Kunstköpfen. Neben dem Fehlen des Torsos, mag auch ein weiterer Grund hierfür darin liegen, dass der KU80 – laut Angabe der Autoren der Studie – im Frequenzgang für hohe Frequenzen eine starke Bedämpfung aufweist. Er wurde deshalb nur mit Mess-Stimuli bis zu einer maximalen Frequenz von 6.5kHz getestet.

Trotzdem belegt er aber im Testteil C, ‚Distance Error‘, den ersten Platz.

Der KU81 belegt zweimal den 4., einen 2. und einen 7. Platz. Der KU81 ist von den genannten Kunstköpfen der einzige, der laut Herstellerangabe diffusfeld entzerrt ist, und dessen Signal deshalb beim Abhören mit ebenfalls diffusfeld-entzerrten Kopfhörern keine weitere Bearbeitung erfordert.

Die Köpfe HMS I und II sind prinzipiell für Freifeld-Bedingungen entzerrt und müssen laut Hersteller-Angabe mit STAX SR Lambda und in Verbindung mit den mitgelieferten Filterbänken abgehört werden.

Aus der dänischen Studie geht eigentlich keiner der Kunstköpfe als eindeutiger Sieger hervor; eine Schlußfolgerung der Autoren lautet, daß eine gezielte Verbesserung des Designs bei Kunstköpfen eine wesentliche Qualitätssteigerung bringen könnte, da mit binauralen Aufnahmen, die in den Ohren ausgesuchter Personen angefertigt wurde, wesentlich bessere Ergebnisse erzielt werden konnten.

Im Rahmen einer Untersuchung zum „Binaural Room Scanning“ [Rathbone et al., 2000] wurden am IRT mit einer Gruppe von 18 Hörern folgende 5 Kunstköpfe hinsichtlich Lokalisation und Klangfarbe getestet :

- HMS III (HEAD Acoustics) (Kopf-und-Torso)
- HUGO (Inst. für tech. Akustik der techn. Hochschule Aachen) [Kopf-und-Torso]
- MANIKIN MK1 (Neutrik-Cortex Instruments)[Kopf-und-Torso]
- KU 81 (Neumann) [nur Kopf]
- KU 100 (Neumann) [nur Kopf]
- KU 100 (Neumann) mit Torso des MK1

Die Studie kommt zu dem Schluss, daß in Hinblick auf die Lokalisationsgenauigkeit keine signifikanten Unterschiede zwischen den Köpfen festzustellen sind; weiters zeigten alle Kunstkopf-Systeme im Mittel eine Elevation von etwa 7°.

Auch in Bezug auf die Klangfarbe zeigten alle Kunstköpfe – mit Ausnahme des KU100 (ohne Torso) – zufriedenstellende Ergebnisse für die tonalen Testbeispiele (Klassik, Pop, Sprache). Im Falle eines rauschähnlichen Signals (Applaus) wurden die Hörergebnisse – unabhängig vom Kunstkopfsystem – als nicht zufriedenstellend bewertet (allerdings schnitt auch hier der KU100 ohne Torso am schlechtesten ab).

Während der KU100 in Kombination mit dem Torso des MK1 als durchaus gleichwertig mit den anderen Kunstkopfsystemen zu betrachten ist, schneidet er alleine für sich genommen deutlich schlechter ab (also auch deutlich schlechter als der KU81, der ebenfalls ohne Torso verwendet wurde)

Der Neumann KU100 stellt eigentlich eine Weiterentwicklung der Kunstköpfe KU80 bzw. KU81 dar, weist im Gegensatz zu diesen jedoch eine starke Vereinfachung (Abstrahierung, teilweise relativ scharfe Kanten) in seiner Oberflächengeometrie auf. Dies dürfte sich negativ auf die ‚Natürlichkeit‘ seiner ‚Performance‘ auswirken.

6.3 Zusammenfassung und Schlussfolgerung zu den Kunstkopf-Studien

Ein Vergleich der Ergebnisse der beiden ausgewählten Studien, deren Ergebnisse im obenstehenden Abschnitt 6.2 näher erläutert wurden:

Die am IRT durchgeführte Studie erfasst weniger Aspekte als die Studie an der Aalborg Universität, schließlich werden dabei keine Fehler in der Medianebene erfasst und auch kein Versuch unternommen, die Qualität der Entfernungswahrnehmung der Kunstköpfe zu testen.

Dass verschiedene Köpfe der dänischen Studie in der vertikalen Ebene hinsichtlich der Lokalisation Schwächen aufweisen, ist für den im Rahmen dieser Dissertation geplanten Verwendung eines Kunstkopfs nicht von primärer Bedeutung, da sich alle zu erfassenden Schallobjekte (Orchestermusiker) in etwa auf einer horizontalen Ebene vor dem Kunstkopf befinden.

Relevanter scheint in diesem Zusammenhang, wie gut sich mit den Köpfen lokalisieren lässt, und wie gut Entfernungen über die einzelnen Kunstköpfe abgebildet werden. Bezüglich des zweiten Kriteriums hatten sich die Neumann Kunstköpfe KU80 und

KU81 in einem Teil der Tests (Aalborg Universität) als Favoriten herauskristallisiert. Die Köpfe HMS I und II scheiden meines Erachtens aufgrund ihrer ‚uneinheitlichen‘ Gesamt-Performance aus (HMS I höchste ‚out-of cone‘ Error-rate, HMS II in den Testteilen B, C sehr hohe Median-Plane Error-Rate)

Der Toronto-Kopf ist hinsichtlich seiner Performance guter Durchschnitt, wird allerdings nur als Einzelstück auf Bestellung produziert. Seiner Verfügbarkeit (für Anmietung) ist deshalb kaum gegeben.

Der B&K Kunstkopf liegt hinsichtlich seiner Performance eher unter dem Durchschnitt.

Im Prinzip wären folgende Kunstköpfe zur Verfügung gestanden:

Fa. Müller-BBM, München: Neutrix-Cortex ‚MK1‘, Neumann ‚KU81‘

Fa. AUDI, Ingolstadt: HEAD-Acoustics ‚HMS‘

Technischen Universität, Graz: HEAD-Acoustics ‚HMS II‘

Aufgrund der Schlussfolgerungen, die sich aus den beiden oben genannten Studien ergeben, entschied ich mich, aus den im Branchenumfeld zur Verfügung stehenden Kunstköpfen, den Neumann KU 81 auszuwählen. Trotz der Tatsache, daß der KU81 über keinen Torso verfügt, konnte er sich in der Studie des IRT gegen die Konkurrenten behaupten. Das Fehlen eines Torsos und die damit eventuell verbundenen qualitativen Beeinträchtigungen sind außerdem vermutlich aufgrund der Positionierung des KU81 inmitten von Zuschauerreihen mit Holzlehnen vermutlich weit weniger relevant, als in der in der Studie beschriebenen Testsituation.

Der Neumann KU81 verfügte darüber hinaus als einziger über den Vorteil, bereits herstellerseitig diffusfeldentzerrt zu sein (Mitarbeiter der Fa. Müller-BBM geben ihm vermutlich deshalb für Musikaufnahmen den Vorzug gegenüber dem MK1), und außerdem über längere Zeit zur Verfügung zu stehen. (Schließlich musste der identische Kunstkopf nicht nur während der Aufnahmen im Konzertsaal in Salzburg, sondern auch noch einige Tage danach für die Aufnahme der dann bereits in 5.1 Surround abgemischten verschiedenen Mikrofon-Arraysignale im Wiedergaberaum am IEM in Graz zur Verfügung stehen.)

7 Hörvergleich

7.1 Grundlegendes zur Problematik von Hörvergleichen

Eine Übersicht zur Problematik der Durchführung von Tests betreffend Hörwahrnehmung findet sich u. a. bei [Bech, 1999], sowie eine umfassende Analyse im Buch von [Bech und Zacharov, 2006].

In einem Übersichts-Artikel mit dem Titel ‚New Horizons in Listening Test Design‘ [Staff-Writer, 2004] gibt der Autor einen Überblick über den aktuellen Stand und neue Tendenzen im Bereich des Designs von Hörversuchen. Egal ob die Testpersonen die Höreindrücke in eigenen Worte frei beschreiben sollen, oder eine Bewertung anhand vorgegebener Skalen erfolgt – in jedem Fall ist die Erarbeitung eines klar definierten gemeinsamen Vokabulars erforderlich. Die im Rahmen der Hörvergleiche verwendeten Methoden stammen aus verschiedensten Bereichen, wie der Psychologie oder der Qualitätsbewertung im Bereich der Nahrungsmittelindustrie.

Angewandte Methoden sind beispielsweise ‚Multidimensionale Skalierung‘ (MDS), ‚Preference Mapping‘, und – im Zusammenhang mit der Entwicklung eines der Aufgabenstellung angepassten Wortschatzes – ‚Descriptive Analysis‘ (DA) (siehe [Bech, 1999]), ‚Repertory Grid Technique‘ (RTG) [Berg and Rumsey, 1999] und ‚Free-Choice Profiling‘ [Bech, 1999].

Die Methode der Multidimensionalen Skalierung bietet die Möglichkeit, relativ einfachen Datensätzen zugrundeliegende Strukturen herausarbeiten zu können. Die Testpersonen müssen dabei nur die simple Aufgabe erfüllen, Stimuli paarweise zu vergleichen, wobei sie diese hinsichtlich ihrer Unterschiedlichkeit bewerten, oder ihre subjektive Präferenz des einen vor dem anderen angeben.

Eine Untersuchung zur Frage der ev. unterschiedlichen Resultate bei Verwendung von ‚Expert Listeners‘ oder untrainierten Testhörern [Olive, 2003], hatte zum Ergebnis, dass generell von Übereinstimmung in der Bewertung ausgegangen werden kann, wobei allerdings die ‚Expert Listeners‘ tendenziell stärker differenzieren und generell weniger hohe Präferenzbewertungen abgeben.

Ebenfalls mit der Beurteilung der Qualität räumlicher Tonwiedergabe beschäftigen sich die nachfolgend aufgezählten Untersuchungen von [Berg und Rumsey, 2006], [Choisel und Wickelmaier, 2006], [Mason und Rumsey, 2000] sowie [Camerer und Sodl, 2001], auf die untenstehend eingegangen wird.

7.2 Anforderungen an die Wiedergabeanordnung

Für die vorgenommenen Hörtests relevante Qualitätsanforderungen hinsichtlich der akustischen Beschaffenheit des Wiedergaberaums und der 5.1 Surround-Abhöranlage finden sich u.a. in folgenden Dokumenten:

- ITU-R (1992-1994) Recommendation ITU-R BS.775-1: „Multichannel stereophonic sound system with and without accompanying picture“ (Geneva)
- SMPTE (1991) Recommended Practice SMPTE RP-173: „Loudspeaker placement for audio monitoring in high definition electronic production“ (SMPTE N 15.04/152-300B)
- EBU (2000) Recommendation R22: „Listening conditions for the assessment of sound programme material“ (Details see EBU Tech 3276 with Supplement 1)
- ITU (1997) Recommendation ITU-R BS.1116-1: „Methods for the subjective assessment for small impairments in audio systems including multichannel sound systems“ (Geneva, Supplement 1 to Volume 1997)
- SSF (Surround Sound Forum) Empfehlungen für die Praxis SSF – 01: „Hörbedingungen und Wiedergabeanordnungen für Mehrkanal Stereophonie“
- SSF (Surround Sound Forum) Empfehlungen für die Praxis SSF – 02: „Mehrkanalton-Aufzeichnungen im 3/2 Format – Parameter für Programmaustausch und Archivierung, Einstellung von Wiedergabeanlagen“

Im Zusammenhang mit der Schallaufzeichnung und Wiedergabe im Rahmen der Dissertation ist das Nutzsignal einer mehrfachen Wandlung unterworfen, die diverse Verzerrungen des Nutzsignales mit sich bringen:

Mittels Mikrofon wird der Schall zunächst in ein analoges elektrisches Signal umgewandelt, welches nach Digitalisierung als linear-PCM Signal (16bit/48kHz) gespeichert wird. Im Rahmen der Abmischung und Aufbereitung der Hörbeispiele sind weitere Bearbeitungsschritte (z.B. Pegelanpassung) in der analogen bzw. digitalen

Ebene notwendig. Zur Wiedergabe der Signale ist eine Rückwandlung von der elektrischen Domäne in die mechanische durch den elektroakustischen Wandler Lautsprecher erforderlich.

Während es heutzutage möglich ist, Mikrofone herzustellen, die äußerst frequenzlinear sind und nur sehr geringe Klirrfaktoren aufweisen, sind die meisten Lautsprecher nach wie vor z.B. mit Klirrfaktoren im Prozentbereich behaftet.

Bezüglich qualitativer Parameter, die zu einer signaltechnischen Verfälschungen des Nutzsignals insbesondere in Hinblick auf die Abbildung der Räumlichkeit führen (und somit auch für die Ergebnisse von Hörvergleichen relevant sein können) sind nachfolgend nur einige wenige Arbeiten – beispielhaft – angeführt:

- Gernemann, A., Rösner, T. (1998): „Die Abhängigkeit der stereophonen Lokalisation von der Qualität der Wiedergabelautsprecher“, 20. Tonmeistertagung des VDT, Karlsruhe, S. 828-846, 1998

- Braun, D., Gutzke, K., Wönicker, C. (1996): „Ein objektives Verfahren zur Beurteilung der räumlichen Abbildung von Lautsprechern“, 19. Tonmeistertagung des VDT, Karlsruhe, S. 630-641, 1996

- Olive S. E., Schuck P. L. and Sally S. L. and Bonneville M. E. (1994): “The Effects of Loudspeaker Placement on Listener Preference Ratings”, *J. Audio Eng. Soc.*, Vol.42, No.9, pg. 651-669, Sept. 1994

Abgesehen vom physikalischen Aufstellungsort der Wiedergabelautsprecher gibt es - betreffend den Bereich der Psychoakustik - weitere Faktoren, die den hervorgerufenen Höreindruck (betr. die beiden nachfolgenden Papers genauer: den Lokalisationseindruck) beeinflussen; hierzu zählt z.B. die frequenzmäßige Zusammensetzung des Testsignals (vgl. hierzu S.45, Abb. 2.17 und S. 32, Abb. 2.12):

- Knothe, J., Plenge G. (1978): „Panoramaregler mit Berücksichtigung der frequenzabhängigen Pegeldifferenzbewertung durch das Gehör“, Bericht zur 11. Tonmeistertagung des VDT, Berlin, S. 136-143, 1978

- D. Griesinger „Stereo and Surround-Panning in Practice“, 112th AES-Convention, München, Paper 5564, 2002

7.3 Einflussfaktoren auf die Hörerbewertung

Generell lässt sich zur Problematik von Hörvergleichen feststellen, dass der Test so aufgebaut sein sollte, dass

- a) sich die Reliabilität und Validität der Hörerfragen überprüfen lässt,
- b) die Abfolge der Klangbeispiele so organisiert sein muss, dass sich ausschließen lässt, dass beim Hörer ev. auftretende Gewöhnungseffekte, die im Zusammenhang mit der Reihenfolge der Klangbeispiele stehen, einen Einfluss auf das Endergebnis haben (Stichwort: Randomisierung), sowie
- c) ein Hörvergleich (ohne Referenz-Stimulus) nach Möglichkeit so gestaltet ist, dass sich die mögliche Beeinflussung der Hörerbeurteilung eines Klangbeispiels (Sample A) durch das Vergleichs-Beispiel (Sample B) über den gesamten Hörvergleich gesehen kompensiert.

Gegenstand der vorliegenden Hörvergleiche ist die Bewertung von vier unterschiedlichen Surround-Mikrofontechniken (DECCA, KFM, OCT und AB-PC bzw. BPT), wobei es prinzipiell 3 Faktoren gibt, die einen Einfluss auf die Bewertung der Hörbeispiele haben können:

- 1) subjektive Hörerpräferenz (Geschmacksrichtungen), sowie unterschiedliche Fähigkeit der Testhörer bei der Unterscheidung von Klangreizen (Stimuli) sowie der Beurteilung der verschiedenen Aspekte, die ihren Gesamtklang bestimmen
- 2) der zum jeweiligen Stimulus (A) angebotenen Vergleichsstimulus (B),
- 3) die unterschiedlichen klanglichen Eigenschaften der 4 Mikrofon-Systeme (um deren Bewertung es eigentlich geht).

Ad 1)

Es kann davon ausgegangen werden, dass die zu beurteilenden Klangbeispiele durch die einzelnen Personen einer Gruppe von Testhörern unterschiedlich bewertet werden. Das letztendliche Resultat der Beurteilung der Klangbeispiele ergibt sich aus der Mittelung der Hörerantworten.

Durch die Verwendung einer hinlänglich großen und passend ausgewählten Gruppe an Testhörern wird versucht, ein möglichst objektives Ergebnis zu erzielen.

Interessant ist in diesem Zusammenhang die Frage, ob die Gruppe der Testhörer - aus Gründen des subjektiven Geschmacks oder von Hörgewohnheiten (Stichwort: ‚Industrie-Standardklangbild‘) - ev. in zwei (oder mehrere) Untergruppen zu teilen wäre, wie z.B. Anhänger eines Klangbildes, das durch starken Einsatz vieler Stützmikrofone geprägt ist (d.h. ein ‚direkter‘, ‚naher‘ Klangeindruck wird bevorzugt, Tiefenstaffelung ist weniger ausgeprägt) im Gegensatz zu einer anderen Gruppe, die vielleicht eher puristische Aufnahmen mit Einsatz lediglich eines Hauptmikrofons bevorzugt (d.h. vermutlich ein durch mehr ‚Raumanteil‘ [Diffusschall] gekennzeichnetes Klangbild, das ev. auch durch eine stärkere Tiefenstaffelung gekennzeichnet ist).

Zur Beantwortung dieser Frage wird an anderer Stelle dieser Untersuchung eine Analyse der Hörerpräferenz bezogen auf die vier Mikrofontechniken mit Hilfe von Histogrammen vorgenommen.

Ad 2)

Da bei der vorliegenden Art des A/B Hörvergleichs (ohne Referenz-Stimulus) anzunehmen ist, dass die Beurteilung eines Reizes A von der Beschaffenheit des im Vergleich angebotenen B-Stimulus beeinflusst wird, wurde darauf geachtet, dass jedes zu beurteilende Klangbeispiel jedem anderen im Vergleich gegenübergestellt wird. Nur so ist gewährleistet, dass sich eine eventuelle Verfälschung in der Beurteilung (als Beispiel: ein ‚trockenes‘ Klangbeispiel wird einem sehr ‚halligen‘ Klangbeispiel gegenübergestellt, womit die Gefahr besteht, dass ersteres als übermäßig trocken bewertet wird) nicht auf die Gesamtbewertung auswirkt, da sich die Verfälschung von Einzelantworten über den kompletten Datensatz der Hörerbeurteilungen gemittelt wieder neutralisiert.

Aufgrund des so gewonnenen Datensatzes kann die Frage der Beeinflussung durch den jeweiligen B-Stimulus eines Paarvergleichs vernachlässigt werden.

Ad 3)

Die eigentlich interessierende Frage bei der vorliegenden Untersuchung besteht in der Bewertung der unterschiedlichen Abbildungs-Eigenschaften der vier zur Diskussion gestellten Mikrofontechniken (OCT, DECCA, KFM, AB-PC) durch die Testhörer.

7.4 Auswahl der zu bewertenden Klangattribute

Um eine Entscheidung hinsichtlich der Auswahl passender Attribute für den Hörvergleich treffen zu können, wurde eine Reihe von Studien, welche ebenfalls Hörvergleiche zum Gegenstand hatten, herangezogen:

Eine japanische Studie zur Bewertung der mehrkanaligen Wiedergabe von Surround-Aufnahmen [Nakayama and Miura, 1969] wurde mit 10 Testhörern mittels Präferenz- und Ähnlichkeits-Beurteilung von Musikaufnahmen durchgeführt. Die Bewertung der dargebotenen Klangbeispiele erfolgte dabei mittels ‚Multidimensional Scaling‘:

Im Rahmen der Präferenz-Beurteilung mussten die Testhörer die Aufnahmen gemäß einer 7 stufigen Skala (von ‚very good‘ bis ‚very bad‘) bewerten.

Im Rahmen der Ähnlichkeits-Beurteilung mussten je zwei Klangbeispiele ebenfalls gemäß einer 7-stufigen Skala (von ‚just the same‘ bis ‚quite different‘) bewertet werden.

Aufgrund der mehrdimensionalen Auswertung (multidimensional analysis) der ‚Ähnlichkeits-Beurteilung‘ zeigte sich, dass die Mehrkanal-Klangwiedergabe vor allem durch die drei Attribute

- fullness
- clearness, und
- depth of the image sources

gekennzeichnet ist. ‚Fullness‘ dürfte etwa ‚Klangfülle‘ entsprechen, charakterisiert durch einen ausgewogenen Frequenzgang, angenehme stereophone Basisbreite, passendes Direkt-/Diffusschall-Verhältnis, etc. ; ‚Clearness‘ entspricht ‚Klarheit‘ im Klangeindruck, charakterisiert durch gute Lokalisation einzelner Schallquellen, ausreichend hohen Direktsignalanteil im Gesamtsignal, etc.; ‚Depth of the Image Sources‘ entspricht der ‚Tiefenstaffelung‘.

Als ein Ergebnis der Untersuchung zeigte sich, dass das Klangattribut ‚Fullness‘ in direkter Verbindung mit der subjektiven Präferenz eines Hörers für ein dargebotenes Klangbeispiel steht.

Weiters konnte nachgewiesen werden, dass ‚clearness‘ eine Funktion des akustisch messbaren Parameters ‚Deutlichkeit‘ (engl. definition) ist.

Eine Studie zur Beurteilung der klanglichen Eigenschaften von Stereo-Hauptmikrofontechniken aus dem Jahre 1972 [Coen, 1972] mit 64 Testhörern (Besucher einer AES-Convention in Köln, 1971) unterschied bei den klanglichen Attributen zwischen:

- Liveness (der in der Aufnahme enthaltene Raumeindruck / ‚diffuse sound field‘)
- Intimacy (Direktheits-Eindruck; Deutlichkeit)
- Perspective (stereophone Perspektive [Stereo-Abbildungswinkel, Abstand vom Klangkörper, etc.])
- Stage Continuity (Kontinuität in der Lokalisation [keine Dehnung, Stauchung, oder

„Mittenloch“, etc.], ev. auch hinsichtlich der Tiefenstaffelung]

- Extra-Width (Auftreten einer *extra-breiten* Stereobasis)
- Dynamic Range
- Warmth (subjektive ‚Wärme‘ im Klangeindruck; hoher Bassanteil, ev. fehlende Höhen)
- Brilliance (hervorragende Auflösung bei hohen Frequenzen; ‚brillanter Klang‘)

Eine Studie, die sich im Rahmen des EU-Forschungsprojekts MEDUSA („Multichannel Enhancement of Domestic User Stereo Applications“, EUREKA 1653) (siehe [Rumsey, 2000]) mit der qualitativen Bewertung von „Virtual Home Theatre Algorithms“ beschäftigte [Mason and Rumsey, 2000] unterschied bei den klanglichen Attributen zwischen folgenden Parametern:

- source width, depth, envelopment und naturalness.

(auf deutsch etwa: „Basisbreite“ oder „Schallquellen-Ausdehnung“, „Tiefenstaffelung“ oder „räumliche Tiefe“, „Einhüllung“ und „Natürlichkeit“)

Darüber hinaus wurde auch die subjektive Präferenz gemessen.

Weitere Attribute, die im Rahmen von Studien des MEDUSA Projekts regelmäßig in den Beschreibungen von Testhörern auftauchen sind ‚source focus‘ (Schärfe in der Abbildung der Schallquelle, also quasi ‚Lokalisationsgenauigkeit‘) und ‚room impression‘ (‚Raumeindruck‘).

Die bei weitem größten Unterschiede in der Bewertung hinsichtlich der Qualität der Klangwiedergabe der einzelnen Algorithmen traten beim Punkt ‚Natürlichkeit‘ (naturalness) auf; dieser Parameter schien auch größten Einfluss auf die subjektive Präferenz eines Hörers für einen Algorithmus zu besitzen.

Eine Untersuchung von 5 Surround-Mikrofonssystemen [Hildebrandt und Braun, 2000] stellte in A/B-Hörvergleichen die Mikrofon-signale einander gegenüber. An den Hörtests nahmen 59 Hörer teil, die in Paarvergleichen eine Wertung anhand einer 5-stufigen Skala durchführten.

Die Hörbeispiele bestanden aus Orchester- und Streichquartettaufnahmen, die

hinsichtlich der Parameter

- Raumeindruck (einhüllende Wirkung),
- Lokalisation und
- Klangfarbe (bzw. Natürlichkeit/ Gesamtklang) bewertet werden mussten.

Im Rahmen eines beim österreichischen Rundfunk ORF im Mai 2001 durchgeführten Hörvergleichs von 7 unterschiedlichen 5.1 Surround-Mikrofonsystemen mit 18 Testhörern (später erfolgte ein zweiter Testdurchgang mit 14 Testhörern an der FH Düsseldorf) sollten die Klangbeispiele gemäß einer 5-stufigen Skala hinsichtlich folgender Kriterien bewertet werden [Camerer und Sodl, 2001]:

- Expansion of the orchestra image: wide / narrow
- Distance of the orchestra image: close / distant
- Depth of the orchestra image: deep / flat
- Stability of the image: stable / instable
- Precision of the image: precise / blurred
- Sound colour: satisfactory / unsatisfactory
- Room impression: perfect / imperfect
- Presence of room information: too much / too little
- Surround signals: identifiable / not identifiable

Die Ergebnisse der Hörerbefragung wurden unter Angabe der Mittelwerte und des 95% Konfidenz-Intervalls getrennt für die Testgruppe des ORF und der FH Düsseldorf ausgewiesen. (siehe Anhang S. 295 ff)

In Hinblick auf die vorliegende Untersuchung schienen dem Autor die im Rahmen der ORF-Hörvergleiche verwendeten Attribute gut für eine Bewertung der klanglichen Eigenschaften von 5.1 Surround-Mikrofonsystemen geeignet. Lediglich bei der Fragestellung ‚Surround Signals: identifiable / not identifiable‘, ist unklar, ob nun eine Hörbarkeit/Identifizierbarkeit des Vorhandenseins von Surround-Signalen eigentlich als positiv oder negativ zu bewerten ist: im Prinzip sollten die Surroundsignale einfach dazu beitragen, die Plastizität / den Räumlichkeitseindruck der Tonaufnahme - und damit die Natürlichkeit des Klangeindrucks - zu erhöhen. Das bedingt aber andererseits, dass das Vorhandensein von ‚Surround-Signalen‘ dem Hörer nicht speziell bewusst wird – schließlich sitzt der normale Konzertbesucher auch nicht im Saal und denkt an

die von hinten oder von der Seite einfallenden Schallreflexionen (es sei denn, er ist raumakustisch vorgebildet und es handelt sich um einen Saal mit schlechter Akustik ...)

Zusätzlich zu den im ORF-Hörtest erfassten Klang-Attributen sollten meines Erachtens noch Bewertungen hinsichtlich der ‚Natürlichkeit‘ des Gesamtklangbilds und der Präferenz der Testperson in Bezug auf die dargebotenen Hörbeispiele erhoben werden. (Gemäß den Ergebnissen des MEDUSA Projekts, siehe oben)

Damit ließe sich u.a. auch die Frage untersuchen, ob das am natürlichsten empfundene Klangbild auch das von den Hörern bevorzugte ist. Dies sollte Rückschlüsse auf Erwartungshaltungen bzw. Hörgewohnheiten (‚Industrie-Standardklangbild‘) auf Konsumentenseite zulassen.

Gemäß den obenstehenden Überlegungen wurden die nachfolgend erklärten 16 Attribute für den Fragebogen ausgewählt:

1) Unterschied zwischen den Hörbeispielen (sehr groß – kein Unterschied)

Hier war der subjektiv von den Hörern empfundene Klangunterschied zwischen dem jeweils angebotenen A- und dem B-Klangbeispiel zu bewerten. Es handelte sich hierbei also um einen klassischen AB-Vergleich, wohingegen die untenstehenden Attribute für jedes Klangbeispiel alleine – also ohne den Vergleichspartner – in direkter Größenschätzung bewertet werden konnten.

2) subjektive Präferenz der Aufnahme (sehr gut – gar nicht):

Die Beurteilung dieses Aspekts wurde bewusst an den Beginn des Fragebogens gestellt, um eine ganz spontane, ‚rein gefühlsmäßige‘ Antwort des Hörers (quasi ‚aus dem Bauch heraus‘) zu erhalten, die noch nicht durch ev. Zweifel aufgrund des nachfolgend gewünschten, wesentlich analytischeren Hörens bei der Beurteilung der anderen Aspekte getrübt sein sollte.

3) Natürlichkeit (sehr natürlich – wenig natürlich):

auch dieser Aspekt wurde bewusst vorangestellt, um ebenfalls einen spontanen ‚Ersteindruck‘ beim Anhören des jeweiligen AB-Vergleichspaares zu erhalten.

4) Klangfarbe (des Orchesters) - Natürlichkeit:

Zunächst war eine Beurteilung der Natürlichkeit der Orchesterklangfarbe gefragt (Semantisches Differenzial: natürlich – unnatürlich).

5) Klangfarbe (des Orchesters) – Balance Höhen/Tiefen:

Hier ging es darum zu beurteilen, ob der Klang der Aufnahme tendenziell höhen- oder tiefenbetont, oder als ausgewogen empfunden wurde (also ein – in der Auswertung - zunächst primär als ‚quantitativ‘ anzusehendes Urteil, da die ‚Helligkeit‘ oder aber ‚Dumpfheit/Wärme‘ eines Klangbildes von verschiedenen Probanden – da geschmacksabhängig - durchaus qualitativ unterschiedlich empfunden werden kann).

6) Lokalisation (sehr gut – nicht gut):

Die hier gestellte Frage lautete: „Wie gut lassen sich die Instrumente bzw. Instrumentengruppen im Orchester orten?“, wobei als Orientierungshilfe deren ungefähre Anordnung entlang der Basisbreite von L nach R angegeben war (siehe Fragebogen im Anhang: Abb. A28).

7) Balance (mehr – weniger):

Gemeint ist hierbei die Lautstärken-Balance zwischen den Instrumenten; die gestellte Frage lautete: „Sind die einzelnen Instrumente(ngruppen) im Orchester lautstärkenmässig ausgewogen?“

Entsprechend dem Schwerpunkt der Dissertation folgten mehrere Fragen betreffend den Aspekt Räumlichkeit:

8) Räumlichkeit des Orchesters (breit – schmal)

Zunächst war die ‚Stereo-Basisbreite‘ (entsprechend ‚Width‘, oder ‚Apparent Source Width‘ – ASW) zu bewerten.

9) Räumlichkeit des Orchesters (nahe – entfernt)

‚Entfernungseindruck‘: d.h. wie weit entfernt empfindet der Hörer die Klangquelle? Dass hier das semantische Differenzial so angelegt wurde, dass ‚nahe‘ links steht (also auf der ‚positiv‘ besetzten Seite, und ‚entfernt‘ auf der ‚negativ‘ besetzten Seite, hat vor allem damit zu tun, dass eine Tonaufnahme, bei der der Diffusschallanteil über den

Direktschallanteil überwiegt, in der Regel eher negativ bewertet wird, als umgekehrt. Dies lässt sich argumentativ u. a. damit begründen, dass sich aus dem bei [Beranek und Schulz, 1965] gefundenen optimalen Hallmaß ein optimales Deutlichkeitsmaß ableiten lässt.

10) Räumlichkeit des Orchesters (tief – flach):

„Tiefenstaffelung“: was entlang der X-Achse die empfundene Basisbreite oder Ausdehnung der Schallquelle ausmacht, findet in der Y-Achse in Form der empfundenen Tiefenstaffelung seine Entsprechung.

11) Räumlichkeit des Orchesters (stabil – instabil):

Hier war eine Beurteilung der Stabilität der Abbildung gefordert.

12) Räumlichkeit des Orchesters (präzise – verschwommen):

„Differenziertheit (der Abbildung)“: bei einem Überwiegen des Diffusschallanteils in der Tonaufnahme ist anzunehmen, dass die Aufnahme als „verschwommen“ empfunden wird. Dies müsste sich dementsprechend auch auf die Lokalisation (bzw. genauer: Lokalisationsschärfe) negativ auswirken.

13) Abbildung des Raumes (überzeugend – unbefriedigend)

14) Abbildung des Raumes (zu viel Raum/zu hallig – zu wenig Raum/zu trocken)

Hier beurteilt der Testhörer eigentlich das von ihm subjektiv empfundene Verhältnis von Direkt- zu Diffusschall.

Gemeinsam mit der unmittelbar davor gestellten Frage lassen sich Rückschlüsse über die Abbildung der Diffusschallanteile durch das jeweilige Surround-Mikrofonverfahren ziehen. Die Beantwortung dieser Frage wird vermutlich auch in engem Zusammenhang mit dem vom Hörer empfundenen Entfernungseindruck stehen.

15) Hörbarkeit der Surroundkanäle (fügen sich unauffällig ein – sind herauszuhören):

Hier geht es um die Auffälligkeit oder Unauffälligkeit der Surround-Kanäle im entstehenden Gesamtklangbild.

16) Hörbarkeit der Surroundkanäle (sind zu leise – sind zu laut):

Gemeinsam mit der unmittelbar zuvor gestellten Frage lassen sich daraus Rückschlüsse

darüber ziehen, ob es der *Signal-Inhalt*, und/oder die *Signal-Lautstärke* ist, welche bedingt, dass die Surroundkanäle als auffällig oder unauffällig empfunden werden.

Zusätzlich zur Beantwortung der anhand der 7-stufigen Skala zu bewertenden Aspekte hatten die Testhörer auch die Möglichkeit, Kommentare zur besseren Klärung oder als Ergänzung ihrer Antworten hinzuzufügen.

Im Rahmen des Fragebogens wurde vom Autor bewusst darauf verzichtet, die musikalisch / akustische Einhüllung (engl. *envelopment*), welche beim Hörer entsteht, bewerten zu lassen. Diese trägt zwar – gemeinsam mit der ‚*Apparent Source Width*‘ – relevant zum Entstehen des Räumlichkeitseindrucks bei (Details siehe Kapitel 2, ‚*Räumliches Hören*‘), steht aber bewusst nicht im zentralen Fokus der Dissertation:

Ein starkes Gefühl von *Envelopment* beim Hörer zu erzielen hängt nicht nur vom Korrelationsgrad der Signale ab, sondern zweifellos auch von den absoluten Pegeln, die z.B. von den hinteren Kanälen abgegeben werden. Da sich nun aber bei der Hörerbefragung herausgestellt hat, dass der Signalinhalt der Surround-Kanäle aller 4 Mikrofontechniken (OCT, DECCA, KFM, ABPC) als unauffällig (vgl. S. 143, Abb. 8.2, Aspekt: ‚*Rear Ch.: un-/auffällig*‘), sowie auch deren Lautstärke als tendenziell leise eingestuft wurde (siehe ebendort, ‚*Rear Ch.: leise/laut*‘), bietet das im Rahmen dieses Hörvergleichs verwendete Material keine passende Ausgangsbasis für eine Untersuchung hinsichtlich *Envelopment*.

Anzumerken wäre in diesem Zusammenhang noch, dass die Pegel der Rear-Signale – sofern nicht im Rahmen der jeweiligen Mikrofonierungstechnik genau definiert – möglichst ‚*unaufdringlich*‘ eingestellt wurden.

7.5 Art des Hörvergleichs

Für die vorliegende Versuchsreihe wurde entschieden, dass die Bewertung der verschiedenen Aspekte der Surround-Mikrofontechniken durch die Testhörer mittels einer direkten Größenabschätzung anhand eines semantischen Differenzials, verbunden mit einem Paarvergleich (und Bewertungsskala) erfolgen sollte:

7.5.1 Bewertungs-Skala

Im Interesse einer ausreichenden Differenzierung der Hörerantworten wurde für die Bewertung der Aspekte eine 7-stufige Skala gewählt.

7.5.2 ‚Kalibrierung‘ der Testhörer

Da gewünscht war, dass die Testhörer nach Möglichkeit die gesamte Skala für die Bewertung nutzen, war es notwendig, ihnen zu Beginn des Tests einen akustischen ‚Überblick‘ über alle vier Mikrofontechniken zu geben. Zu diesem Zweck wurde ihnen vorab eine ‚Einhörsektion‘ bestehend aus jeweils ca. 10 Sekunden langen Abschnitten (desselben musikalischen Ausschnitts) mit den vier unterschiedlichen Mikrofontensystemen vorgespielt. Die Ausschnitte folgten ohne Pause unmittelbar aufeinander, um dem Hörer einen direkten Vergleich zu ermöglichen. Diese Einhörsektion wurde dem Probanden mindestens zweimal vorgespielt, er hatte aber auch die Möglichkeit sie – zwecks besserer Orientierung – mehrfach hintereinander zu hören.

7.5.3 Vorbemerkungen zur Anwendbarkeit der Varianzanalyse

Vorab ist anzumerken, dass für eine korrekte Anwendung der Varianzanalyse generell folgende Bedingungen bezüglich des auszuwertenden Datensatzes gegeben sein müssen:

- 1) die durch die Testhörer erfolgten Bewertungen sind normalverteilt,
- 2) die Bewertungen weisen für jeden Testhörer die selbe Varianz auf,
- 3) die Bewertungen durch die Testhörer sind unabhängig voneinander.

Der ANOVA Test ist relativ robust gegen eine leichte Verletzung der ersten beiden Annahmen.

Ad 1) Vor allem betreffend den Aspekt Präferenz scheint es von Interesse, zu überprüfen, wie die Verteilung der Hörerantworten bzgl. aller vier Mikrofontechniken aussieht. Dies besonders in Hinblick darauf, ob (ev. bzgl. auch nur einer der Mikrofontechniken) eine Spaltung der Hörschaft in zwei Gruppen vorliegt, von denen die eine z.B. eine eindeutige Präferenz für eine der Techniken hat, die andere diese

jedoch eindeutig ablehnt. Im numerischen Mittel ergäbe sich in diesem Fall z.B. ein unauffälliger Mittelwert, das Histogramm hingegen würde aufzeigen, dass in Wahrheit keiner der Testhörer diese Technik tatsächlich als ‚mittelgut‘ bewertet. Im Falle einer solchen Polarisierung der Hörerschaft wäre der Aspekt Präferenz – bei Fehlen einer annähernden Normalverteilung der Hörerbewertungen – für die Durchführung einer Varianzanalyse ungeeignet. Wie man anhand der nachfolgenden Abbildung 7.1 erkennen kann, ist dies allerdings nicht der Fall.

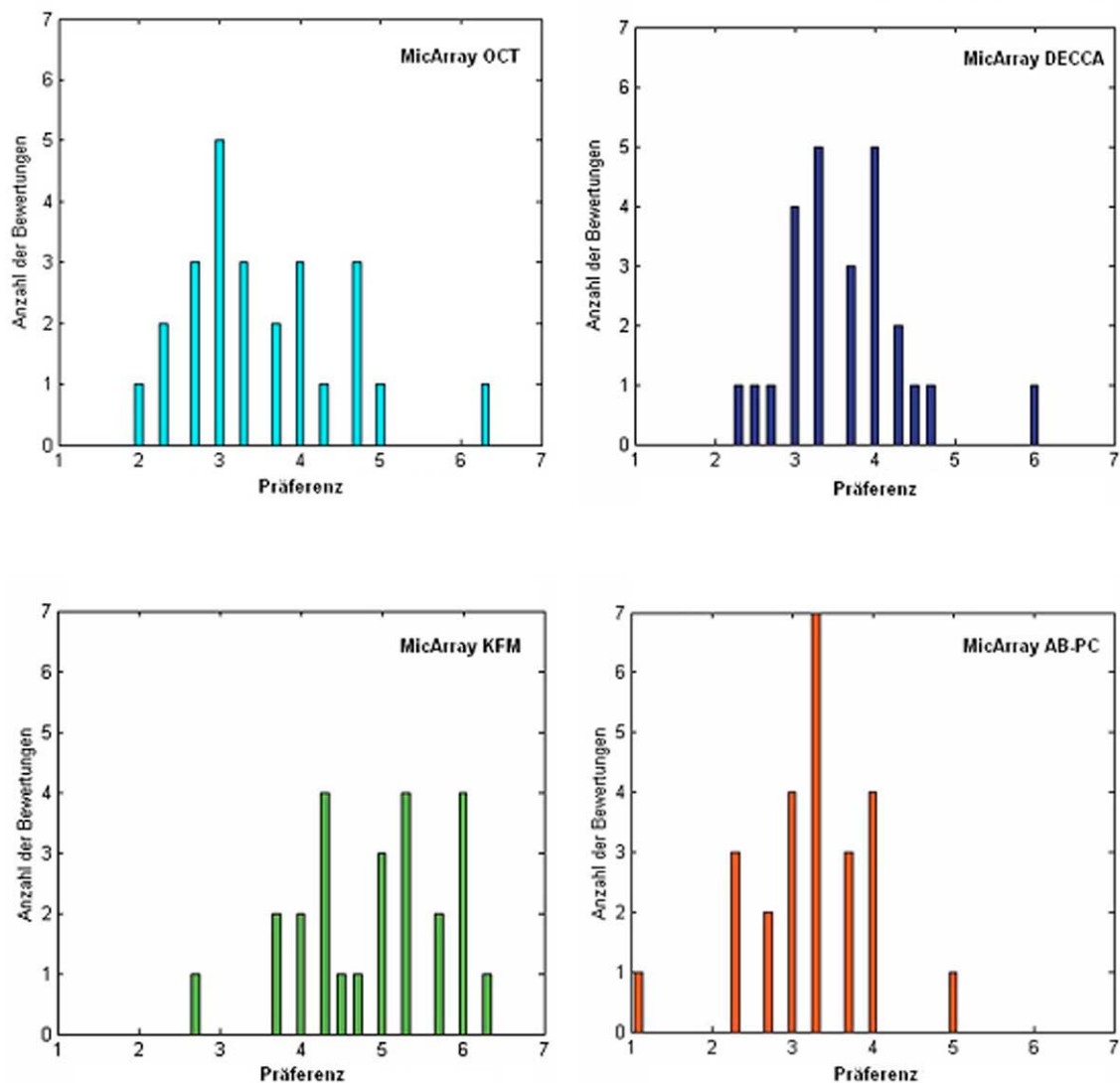


Abb. 7.1: Histogramme der Hörerantworten betr. Aspekt Präferenz für 4 Mikrofontechniken (Stimulus: Orchester 5.1; 25 Testhörer)

Sicherheitshalber wurde hinsichtlich der Frage des Vorliegens einer Normalverteilung der Hörerantworten bei den verschiedenen Aspekten eine Untersuchung mittels Lilliefors-Test durchgeführt, deren Ergebnis in Tab. 7.1 zu sehen ist:

| | ABPC | DECCA | KFM | OCT |
|-----------------------|------|-------|-----|-----|
| Präferenz | | | | |
| Natürlichkeit | x | | x | |
| Klangfarbe | | | | |
| Höhen-/Tiefen-Balance | x | x | x | x |
| Lokalisation | | | | |
| Balance | | x | | |
| Basisbreite | | | | x |
| nahe/entfernt | | | | |
| Tiefenstaffelung | x | | | x |
| Stabilität | | | | |
| Differenziertheit | | | x | |
| Räumlichkeit | x | | | |
| hallig/trocken | | | | |
| Surr_un/auffällig | | x | | |
| Surr_leise/laut | | | x | |

Tab.7.1: Ergebnis des Lilliefors-Tests (Stimulus: Orchester 5.1; 25 Testhörer): wenn mit x gekennzeichnet, liegt keine Normalverteilung der Hörerantworten hinsichtlich eines Aspekts vor

Betrachtet man die in Tab.7.1 dargestellten Ergebnisse, so fällt auf, dass bei den Testhörern offenbar eine sehr große Uneinigkeit darüber besteht, wie die Höhen-/Tiefen Klangbalance bei den Aufnahmen der verschiedenen Mikrofontechniken beschaffen ist, da hier für *keine* der Techniken eine Normalverteilung der Hörerantworten vorliegt. Dies ist umso erstaunlicher, da für den damit zweifellos in Verbindung stehenden Aspekt Klangfarbe bei allen Mikrofontechniken Normalverteilung besteht. Da allerdings die Frage der Klangbalance nicht im zentralen Fokus der vorliegenden Arbeit liegt, wurde dieser Sachverhalt nicht näher untersucht.

Bedauerlich ist das Fehlen einer Normalverteilung beim Aspekt Räumlichkeit für die Mikrofontechnik ABPC; hier wäre es sicherlich wünschenswert, wenn im Rahmen zukünftiger Untersuchungen die somit bestehenden Vorbehalte hinsichtlich der Richtigkeit der Auswertung der Hörerantworten mittels Varianzanalyse beseitigt werden könnten.

Auf die in Tab. 7.1. gekennzeichneten Aspekte, bei denen keine Normalverteilung der Hörerantworten vorliegt, ist die Varianzanalyse streng genommen nicht anwendbar bzw. die Ergebnisse derselben nur als beschränkt richtig anzusehen.

In Abb. 7.2 sind Histogramme der Hörerantworten betreffend den Aspekt ‚Surround un-/auffällig‘ dargestellt. Gemäß der Ergebnisse des Lilliefors-Tests liegt für die DECCA-Technik keine Normalverteilung vor, d.h. dass für diese sowohl die Berechnung des p-Werts, als auch des Konfidenzintervalls keinen brauchbaren Wert liefert.

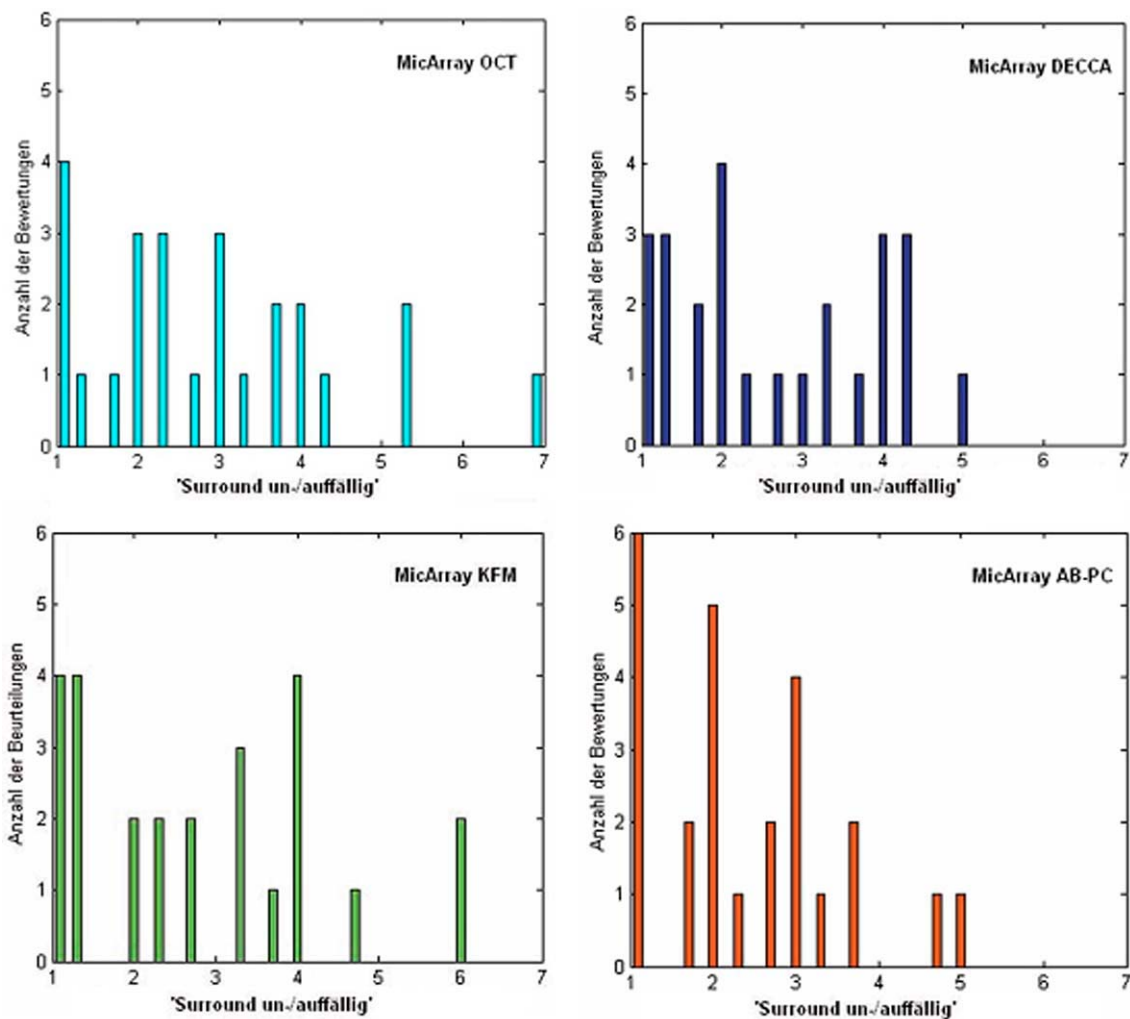


Abb. 7.2: Histogramme der Hörerantworten betr. den Aspekt ‚Surround un-/auffällig‘ für 4 Mikrofontechniken (Stimulus: Orchester 5.1; 25 Testhörer)

Ad 3) Betreffend die Frage der Erfüllung des dritten Punkts sind folgende Gedanken anzubringen:

In den Datensatz, der zur Varianzanalyse herangezogen wird, gehen die bereits gemittelten Antworten der einzelnen Hörer ein; d.h. pro Aspekt und Mikrofontechnik liegt von jedem Hörer eine (bereits gemittelte) Bewertung vor.

Für den Aspekt ‚Natürlichkeit‘ gibt es also z.B. im Falle der 25 kompletten Datensätze insgesamt $25 \text{ (Hörer)} \times 4 \text{ (Mikrofontechniken)} = 100$ Hörerantworten. Das gleiche gilt für alle anderen Aspekte einschließlich der Qualitätsbewertung ‚Präferenz‘, die ebenfalls gemäß einer 7-stufigen Skala vom Testhörer anzugeben war.

Da für die Hörer keine ‚Expert listeners‘ verwendet werden konnten und auch nicht die Möglichkeit bestand, die Testhörer zuvor in irgendeiner Weise (gemäß ihrer Hörfähigkeit, oder nach anderen Kriterien) zu selektieren, sollte man davon ausgehen können, dass die Hörerbewertungen als unabhängig voneinander anzusehen sind. Anders würde es sich u.U. verhalten, handelte es sich um eine Gruppe von Testhörern, die durch wiederholtes Hörtraining auf eine bestimmte ‚ästhetische Richtung‘ in Hinblick auf Tonaufnahmen ‚getrimmt‘ worden wären. Dies wäre der Objektivität der Bewertung zweifelsohne abträglich.

Da die Testhörer allerdings nicht nur aus unterschiedlichen geographischen Regionen (Österreich, Deutschland, USA), sondern auch aus unterschiedlichen beruflichen Umfeldern (Studenten, Praktikanten, Tontechniker und Tonmeister, sowie ein Lautsprecher-Konstrukteur) und Altersklassen (zwischen 20 und 60) stammen, darf eine Normalverteilung der Hörer-Bewertungen erwartet werden. Darüber hinaus scheint die Wahl der Hörergruppe ein guter Kompromiss zwischen ‚Expert listeners‘ und ‚Normalbevölkerung‘ zu sein.

8 Auswertung des Hörvergleichs ,ORCH 5.1‘ (Lautsprecher)

8.1 Durchführung der Varianzanalyse

Die Datenberechnung und graphische Darstellung der Ergebnisse der deskriptiven Statistik-Auswertung (Mittelwerte, Konfidenzintervalle) erfolgte mittels EXCEL-Programm. Aufgrund der einfachen Struktur des Programms konnte hinsichtlich der Berechnung der Konfidenzintervalle auf die Einschränkungen, die bezüglich kleiner Stichproben ($n < 30$) zu berücksichtigen sind (siehe hierzu [Bortz, 2005] S. 103), keine Rücksicht genommen werden.

Die in EXCEL erstellten Datensätze wurden in das MATLAB-Programm (Version 7; Fa. ,The Mathworks‘) importiert und für jeden einzelnen Aspekt mittels Varianzanalyse auf folgende Hypothesen getestet:

- 1) ,Nullhypothese‘: es bestehen keine statistisch signifikanten Unterschiede zwischen den Mikrofontechniken.
- 2) ,Alternative Hypothese‘: zumindest eine der Mikrofontechniken unterscheidet sich von den übrigen statistisch signifikant.

Für die vorliegende Untersuchung wurde als Schwellwert für das Signifikanz-Niveau $p=0.05$ gewählt.

Zunächst wurde – der Vollständigkeit halber - eine 2-faktorielle Varianzanalyse durchgeführt, um die Abhängigkeit vom Hörer sowie von der Mikrofontechnik auf die Bewertung des jeweils angebotenen Klangreizes (Stimulus, Hörbeispiel) ermitteln zu können. Dabei konnten keine Wechselwirkungen zweiter Ordnung Hörer*Mikrofontechnik nachgewiesen werden.

Danach wurden die p-Werte und Konfidenzintervalle für beide Gruppen IEM und SF zunächst getrennt betrachtet. (Details hierzu siehe Anhang). Aufgrund der Ähnlichkeit der Ergebnisse für die Mehrzahl der Aspekte scheint es vertretbar, die Antworten beider Gruppen zusammenzufassen und als einen gemeinsamen Datensatz

der Varianzanalyse zu unterziehen.

Mit 25 Testhörern wurde beim vorliegenden, kombinierten Datensatz für 13 von 15 Aspekten Signifikanz-Niveau erreicht ($p < 0.05$), wie der untenstehenden Tabelle zu entnehmen ist: Lediglich die Ergebnisse der Aspekte ‚Differenziertheit‘ und ‚Surround un-/auffällig‘ lagen unterhalb des Signifikanz-Niveaus.

| 'Dependent Variable' | Sum of squares | df | Mean square | F | p | signifikante Differenz bei den Mittelwerten der Mic-Array Beurteilungen |
|----------------------|----------------|----|-------------|---------|---------------|---|
| Präferenz | 38.5897 | 3 | 12.8632 | 16.7695 | 0 | 1 & 3, 2 & 3, 4 & 3 |
| Natürlichkeit | 20.7831 | 3 | 6.9277 | 8.4185 | 0 | 1 & 3, 2 & 3, 4 & 3 |
| Klangfarbe | 14.5719 | 3 | 4.8573 | 7.5008 | 0.0001 | 1 & 3, 4 & 3 |
| Höhen/Tiefen | 9.7675 | 3 | 3.2558 | 6.3025 | 0.0006 | 1 & 4, 2 & 4 |
| Lokalisation | 14.9475 | 3 | 4.9825 | 6.3767 | 0.0005 | 1 & 4, 2 & 4, 3 & 4 |
| Balance | 21.6322 | 3 | 7.2107 | 9.4913 | 0 | 1 & 3, 2 & 3, 4 & 3, 1 & 4 |
| Basisbreite | 79.2342 | 3 | 26.4114 | 30.7490 | 0 | 1 & 2, 1 & 3, 1 & 4, 2 & 3, 4 & 3 |
| Entfernungseindr. | 35.5233 | 3 | 11.8411 | 16.8673 | 0 | 1 & 2, 2 & 3, 2 & 4 |
| Tiefenstaffelung | 17.6275 | 3 | 5.8758 | 5.9987 | 0.0009 | 1 & 3, 2 & 3 |
| Stabilität | 8.4431 | 3 | 2.8144 | 2.8690 | 0.0405 | 3 & 4 |
| Differenziertheit | 7.6853 | 3 | 2.5618 | 2.3699 | 0.0754 | keine |
| Raumabbildung | 43.7697 | 3 | 14.5899 | 13.4332 | 0 | 1 & 3, 2 & 3, 4 & 3 |
| hallig/trocken | 43.0233 | 3 | 14.3411 | 22.2303 | 0 | 1 & 3, 2 & 3, 4 & 3 |
| Surr.un-auffällig | 4.0756 | 3 | 1.3585 | 0.7353 | 0.5335 | keine |
| Surr. leise/laut | 7.2100 | 3 | 2.4033 | 2.7409 | 0.0475 | keine |

Anm.: 1=OCT, 2=DECCA, 3=KFM, 4=ABPC

Tab. 8.1: 1-faktorielle Varianzanalyse betreffend Einfluss der Variable ‚Mikrofon-Array‘ auf die Bewertung der 15 Klangaspekte (25 Hörer)

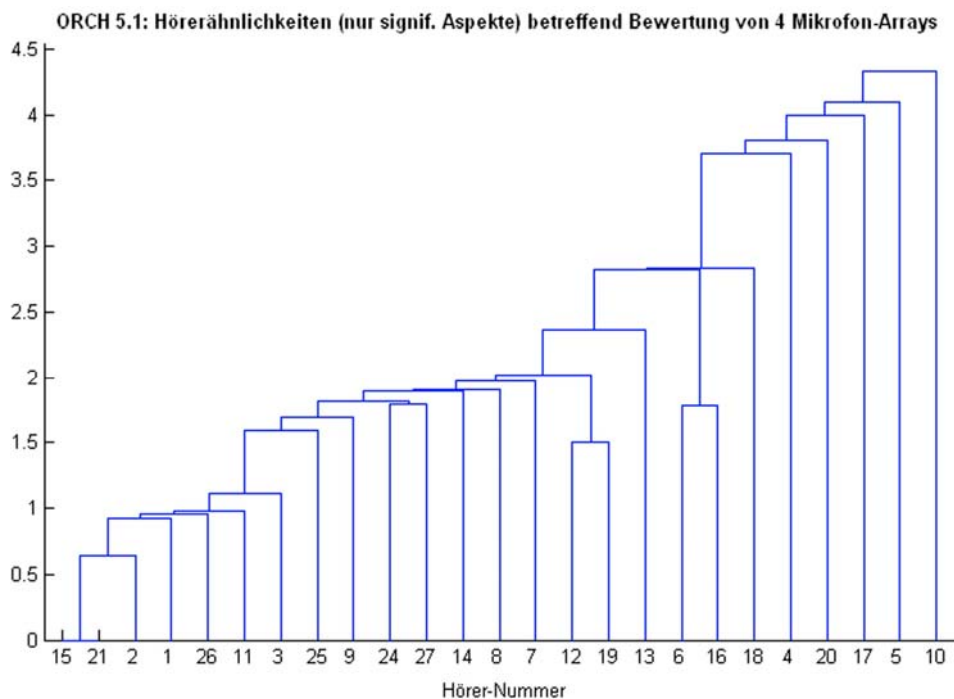


Abb. 8.1: Dendrogramm der Hörerähnlichkeiten betreffend Hörvergleich ORCH 5.1 (single linkage)

Differenziertheit: mit $p=0.0754$ liegt der Wert relativ nahe dem Schwellenwert von 0.05.

Bei Weglassung der Hörerantworten der Testpersonen 4,5,10,17 und 20 (siehe Abb. 8.1, Dendrogramm ORCH 5.1, oben) konnte mit einem Wert von 0.041 auch hier Signifikanz erreicht werden.

Der Autor konnte bereits während der Durchführung des Hörvergleichs feststellen, dass einige Versuchspersonen Schwierigkeiten hatten, dem Begriff Differenziertheit (trotz Begriffsklärung im Rahmen des vorbereitenden Einführungsvortrags, an dem alle Testpersonen hätten teilnehmen sollen) bestimmte klangliche Aspekte zuzuordnen. In solchen Fällen wurde durch den Versuchsleiter individuell versucht, dem jeweiligen Probanden während des Hörvergleichs klarzumachen, was damit gemeint war.

Vermutlich hat diese relativ weit verbreitete Unsicherheit hinsichtlich des Begriffs auch in der großen Varianz der Hörerbewertung (siehe auch 95% Konfidenzintervall) ihren Niederschlag gefunden.

Surround un-/auffällig: Wie bereits weiter oben erläutert, kann die Varianzanalyse bzgl. dieses Aspekts aufgrund des Abweichens der Werte-Population von der

Normalverteilung nicht angewendet werden.

Betrachtet man also nur die in der Tab. 8.1 dargestellten Mittelwerte, so sieht man, dass - über alle Hörerantworten gemittelt - die Signale der hinteren Surroundkanäle LS, RS offenbar bei allen vier Mikrofontechniken als unauffällig empfunden wurden.

Surround leise/laut: Auch hier war offenbar keine Einstimmigkeit unter den Hörern gegeben, allerdings lag der p-Wert mit 0.0475 knapp über die Signifikanz-Schwelle.

Stabilität: Mit einem p-Wert von 0.0405 ist statistische Signifikanz der sich für die Mikrofon-Arrays ergebenden Mittelwerte gerade erreicht. Bei Weglassung der Hörerantworten der Testpersonen 4,5,10,17 und 20 ergibt sich eine höhere Signifikanz bei einem p-Wert von 0.0187.

8.1.1 Signifikante Unterschiede bei den Mittelwerten der 4 Mikrofontechniken

Wie sowohl aus Abbildung 8.2, wie auch der Tab. 8.2 zu ersehen ist, liegen statistisch signifikante Unterschieden bei den Mittelwerten der Hörerbeurteilung bei folgenden Aspekten vor:

Präferenz

Mikrofontechnik Nr. 3, das Kugelfächenmikrofon KFM, wird bei Gegenüberstellung mit allen anderen Mikrofontechniken eindeutig schlechter bewertet. Bezogen auf die 7-stufige Skala schnitt diese Technik um 1.5 Skalenpunkte schlechter ab.

Natürlichkeit und Klangfarbe

Bei diesen Aspekten sieht die Situation ganz ähnlich aus, wie bei der Präferenz-Bewertung: KFM wird eindeutig schlechter bewertet, als alle anderen Systeme.

Höhen/Tiefen Balance

Bei der Bewertung dieses quantitativen Aspektes gibt es signifikante Unterschiede in den Mittelwerten zwischen den Techniken OCT und AB-PC, sowie DECCA und AB-PC.

Lokalisation

Bei diesem Aspekt unterscheidet sich lediglich AB-PC signifikant von OCT und KFM.

Balance

Hier ist KFM signifikant schlechter als alle anderen, AB-PC signifikant besser als OCT und KFM.

Basisbreite

Bei diesem Aspekt ist für den vorliegenden Datensatz die größte Trennschärfe bei der Differenzierung zwischen den verschiedenen Mikrofontechniken gegeben: am besten bewertet wird DECCA, knapp gefolgt von AB-PC. KFM schneidet deutlich am schlechtesten ab.

Entfernungseindruck (nahe/entfernt)

Hier unterscheidet sich lediglich DECCA signifikant von den anderen Techniken.

Die DECCA-Aufnahme wird als eindeutig 'weiter entfernt' klingend eingestuft, was insofern nicht verwundert, kommen doch für alle Front-Mikrofone nur Druckempfänger (Omni-Charakteristik) zum Einsatz, die Direkt- wie auch Diffusschall aus allen Raumrichtungen aufnehmen, während alle anderen Mikrofontechniken für die Gewinnung des Front-Signals (zumindest auch) Mikrofone mit Richtcharakteristik einsetzen, woraus sich zwangsläufig ein anderes Mischungsverhältnis Direkt-/Diffusschall ergibt.

Tiefenstaffelung

Bezüglich der Tiefenstaffelung unterscheiden sich nur OCT und DECCA signifikant von KFM, welches als das System mit der geringsten Tiefenstaffelung eingestuft wird. Rein numerisch liegt der Mittelwert für das KFM um zumindest 0.6 Skaleneinheiten unter denen der anderen Systeme.

Stabilität

Statistisch signifikant unterscheiden sich diesbezüglich nur die Techniken KFM und AB-PC.

Differenziertheit

Wie bereits weiter oben beschrieben, lieferten die Bewertungen des Aspekts Differenziertheit beim vorliegenden Datensatz keine statistisch signifikanten Resultate. ($p > 0.05$)

Raumabbildung

Hierbei unterscheidet sich KFM signifikant von allen anderen bewerteten Mikrofontechniken und wird um mehr als einen Skalenpunkt schlechter bewertet als die anderen.

Hallig/trocken

Ebenso wie bei der Raumabbildung unterscheidet sich auch bei diesem – allerdings quantitativen – Aspekt nur das KFM-System signifikant von allen anderen: es wird offenbar als deutlich trockener bewertet.

Surround un-/auffällig

Wie bereits oben beschrieben, lässt sich die Varianzanalyse auf diesen Aspekt nicht anwenden. Eine Betrachtung der Mittelwerte zeigt, dass bei den vorliegenden Aufnahmen die Signale der hinteren Surroundkanäle im Mittel als unauffällig (der Mittelwert aus den Mittelwerten aller vier Mikrofontechniken liegt bei 2,68) bewertet wurden.

Surround leise/laut

Mit 0.0475 liegt dieser Wert zwar gerade noch oberhalb des Signifikanzniveaus, allerdings liegen keine signifikanten Unterschiede bezüglich der Mittelwerte der verschiedenen Mikrofontechniken vor.

| | OCT | DECCA | KFM | AB-PC |
|---------------------|------|-------|------|-------|
| Präferenz | 3,55 | 3,64 | 4,90 | 3,23 |
| Natürlichkeit | 3,50 | 3,69 | 4,52 | 3,25 |
| Klangfarbe | 3,52 | 3,73 | 4,35 | 3,25 |
| Höhen/Tiefen | 3,99 | 4,29 | 3,91 | 3,48 |
| Lokalisation | 3,92 | 3,68 | 3,95 | 3,05 |
| Balance | 3,61 | 3,54 | 4,35 | 2,97 |
| Basisbreite | 3,96 | 2,63 | 4,87 | 2,84 |
| Entfernungseindruck | 3,55 | 4,68 | 3,31 | 3,35 |
| Tiefenstaffelung | 3,58 | 3,39 | 4,57 | 3,84 |
| Stabilität | 3,25 | 3,37 | 3,78 | 3,01 |
| Differenziertheit | 3,57 | 3,75 | 3,81 | 3,16 |
| Raumabbildung | 3,67 | 3,53 | 4,96 | 3,27 |
| hallig/trocken | 3,87 | 3,39 | 5,18 | 3,96 |
| Surr. un-/auffällig | 2,89 | 2,67 | 2,79 | 2,37 |
| Surr. leise/laut | 3,85 | 3,79 | 3,31 | 3,34 |

Tab. 8.2: Beurteilung von 4 Mikrofonarrays bezüglich 15 Aspekten (gemittelt über 25 Hörer) bei 7-stufiger Skala

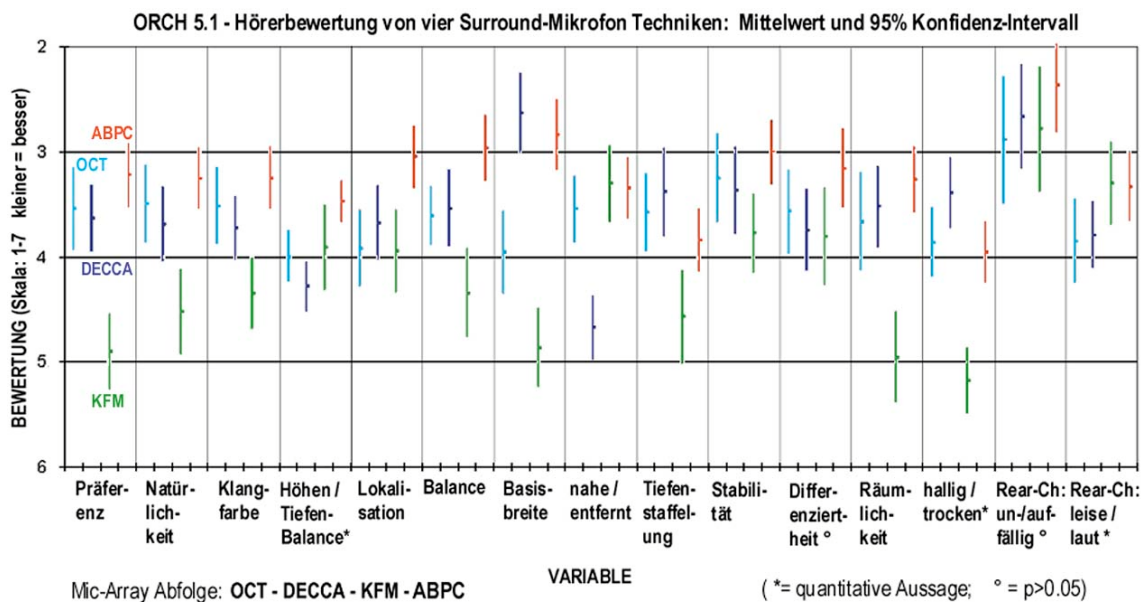


Abb. 8.2: ORCH 5.1 – Darstellung von Mittelwert und 95%-Konfidenzintervall für 4 Mikrofontechniken betreffend 15 Aspekte (25 Testhörer)

8.1.2 Reliabilität und Validität der Hörerbewertungen

Im Zusammenhang mit der Reliabilität der Testergebnisse, d.h. deren Freiheit von zufälligen Fehlern im Sinne von gleichbleibenden Ergebnissen im Falle wiederholter Erhebungen (Hörerbefragungen) kann im Zusammenhang mit der vorliegenden Untersuchung folgendes festgestellt werden:

Um die Reliabilität der Hörerbewertungen gewährleisten zu können, bewerteten alle Testhörer (mit wenigen Ausnahmen) auch ein Vergleichspaar, das zweimal denselben (identischen) Stimulus enthielt. Dies war am allerersten Testtag die Paarung CC (DECCA) danach die Paarung EE (AB-PC).

Die Resultate jener Hörer, die das identische Hörbeispiel nicht als solches erkannten, wurden ausgeschieden.

Als Maßstab für die Reliabilität der Hörerantworten kann generell die Tatsache der hohen Signifikanz (p -Werte ≤ 0.0001 , siehe Tab. 8.1) ins Treffen geführt werden, was für etwa die Hälfte aller Aspekte zutrifft.

Das Heranziehen der redundanten Vergleichspaare für eine Bestimmung der Reliabilität der Höreraussagen ist leider nur beschränkt möglich, da im vorgegebenen Zeitrahmen – wie bereits weiter oben erläutert – die Hörer der Gruppe IEM nur zum Teil zur Bewertung des ‚Wiederholungsbeispiels‘ kamen. (Dies war nur für 9 Testhörer der Fall)

Weitere Aufschlüsse über die Reliabilität der Höreraussagen lassen sich nach Meinung des Autors durch die Auswertung der Spezialfälle (6 und 9, nachfolgend) gewinnen, wobei für die Rear-Kanäle (LS, RS) bei einem der Stimuli der Pegel verändert wurde.

Die externe Validität der Untersuchung (vgl. [Bortz, 2005], S. 8), d.h. die Anwendbarkeit bzw. Gültigkeit der Ergebnisse auch für die ‚Normalbevölkerung‘ scheint insofern zumindest ansatzweise gegeben, da die Gruppe der Testhörer nicht aus ‚Expert Listeners‘ bestand, sondern sich aus zwar ‚hörerfahrenen‘, aber nicht speziell trainierten Hörern zusammensetzte. (Nähere Details hierzu siehe S. 136)

8.1.3 ‚Spezialfall 6‘

Paarvergleich OCT – KFM, wobei bei den Kanälen LS, RS des KFM-Mikrofonsystems der Pegel um 6dB angehoben wurde.

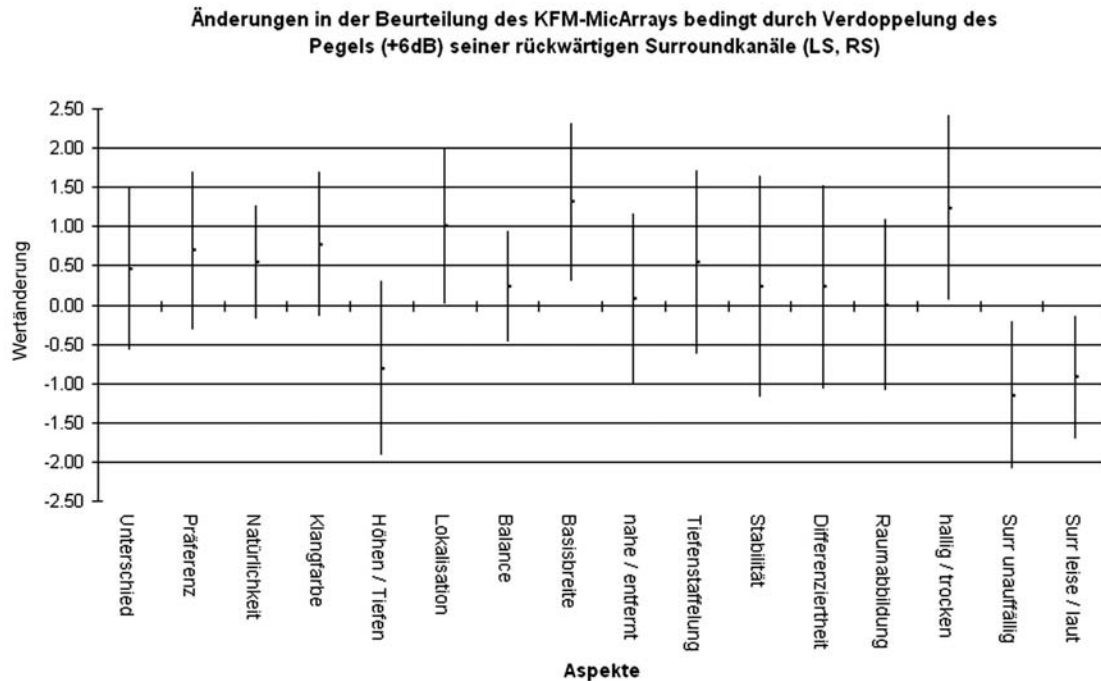


Abb. 8.3: Mittelwertänderung und 95% Konfidenzintervall beim KFM-MicArray, für 15 Aspekte (KFM ‚normal‘ – KFM ‚Rear +6db‘)

In der obenstehenden Abbildung 8.3 sind die Auswirkungen dieser Änderung auf die Bewertung des KFM-Mikrofonsystems dargestellt.

Betrachtet werden hier zunächst nur jene Änderungen, die ungefähr in der Größenordnung eines Skalenwertes liegen:

Die Lokalisation wurde nun besser bewertet, ebenso ist ein Anstieg der empfundenen Basisbreite um mehr als einen Skalenwert zu verzeichnen. Durch die Verdopplung des Pegels der hinteren Kanäle wird die Aufnahme nun allerdings auch als deutlich halliger empfunden.

Als Bestätigung für die Richtigkeit der hier analysierten Bewertung durch die Testhörer darf die Tatsache herangezogen werden, dass bei der vorliegenden Aufnahme nun auch die Surround-Kanäle als auffälliger und lauter bewertet wurden. Die Bewertungsänderung liegt in der Größenordnung eines Skalenwertes und darf

deshalb nach Meinung des Autors als signifikant betrachtet werden.

Weitere - wenngleich auch geringere - Bewertungsänderungen ergaben sich für die Aspekte Klangfarbe (Verbesserung), sowie die Balance Höhen/Tiefen (Verschiebung zugunsten der Tiefen, d.h. ‚dunkleres / dumpferes‘ Klangbild)

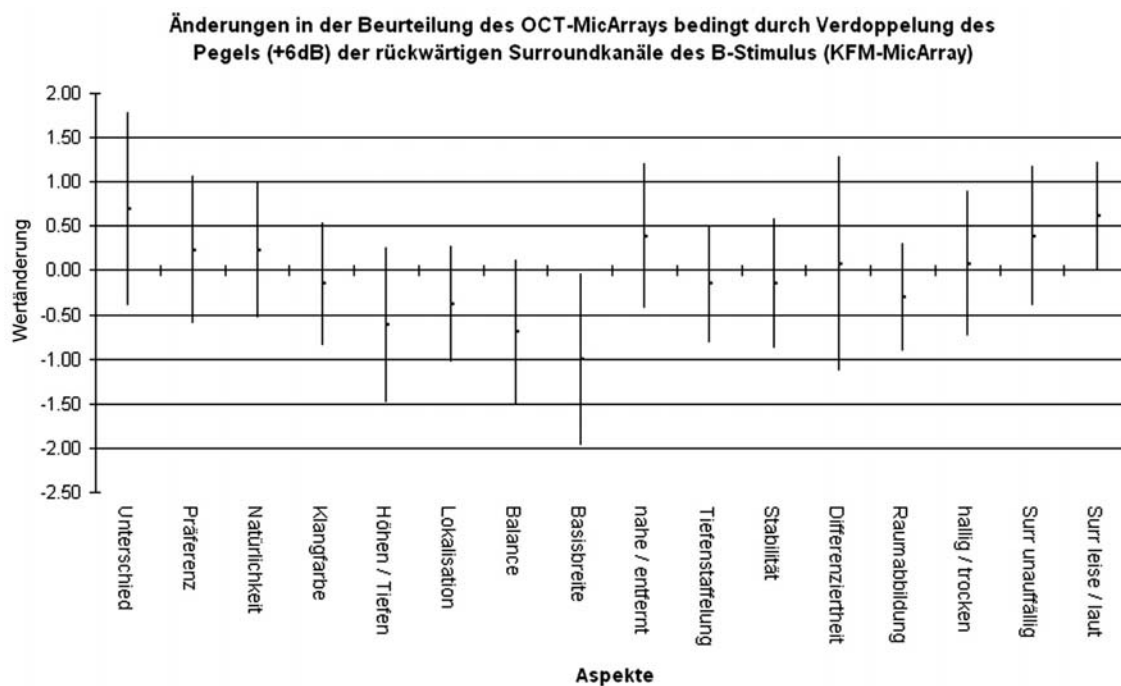


Abb. 8.4: Mittelwertänderung und 95% Konfidenzintervall beim OCT-MicArray, für 15 Aspekte (KFM ‚normal‘ – KFM ‚Rear +6db‘)

Wiederum sollen zunächst nur die Änderungen betrachtet werden, die in der Größenordnung eines Skalenwertes liegen:

Dies ist nur bei der ‚Basisbreite‘ der Fall, welche nun als geringer bewertet wurde. Dieses Resultat ist insofern schlüssig, als sich beim B-Stimulus (KFM) durch die vorgenommene Änderung die empfundene Basisbreite ja vergrößert hatte. Die indirekte Rückwirkung auf den A-Stimulus OCT scheint – aus psychoakustischer Sicht - somit zwingend.

Darüber hinaus scheint lediglich erwähnenswert, dass die Bewertung des Unterschieds zwischen A und B-Stimulus nun im Mittelwert um etwa 0.7 (Skalenwert) grösser geworden ist, d.h. dass sich das OCT-Signal nun stärker vom KFM-Signal unterscheidet.

Stimmig mit der durch die Testhörer am B-Stimulus wahrgenommenen Veränderung scheinen auch folgende Änderungen betreffend Wahrnehmung am A-Stimulus OCT: Der Entfernungseindruck verschob sich um beinahe einen halben Skalenwert Richtung ‚näher‘, und die Surround-Kanäle wurden nur als leiser empfunden (Bewertung ca. +0.6); allerdings liegen diese beiden Veränderungen nur in der Größenordnung eines halben Skalenwertes und sind somit von beschränkter Aussagekraft.

Generell lässt sich sagen, dass die an den beiden Stimuli A und B durch die Testhörer wahrgenommenen Änderungen durchaus in Einklang mit den Erwartungen stehen, die gemäß den Gesetzen der Psychoakustik hinsichtlich der Art der an den hinteren Kanälen des B-Stimulus KFM vorgenommenen Änderungen berechtigt scheinen.

8.1.4 ‚Spezialfall 9‘

Paarvergleich C-E (DECCA – ABPC), wobei die Signale der Kanäle LS, RS des ABPC-Mikrofonsystems komplett abgeschaltet wurden (Pegel = $-\infty$).

In Abbildung 8.5 sind die Auswirkungen dieser Änderung auf die Bewertung des ABPC-Mikrofonsystems dargestellt.

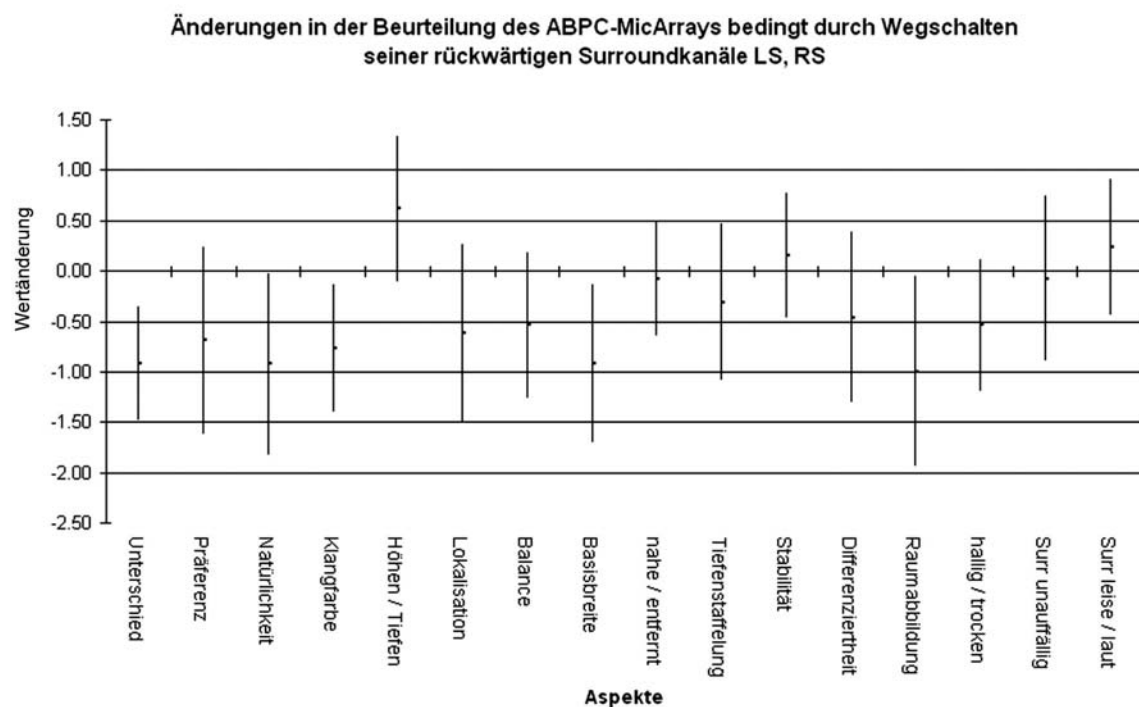


Abb. 8.5: Mittelwertänderung und 95% Konfidenzintervall beim ABPC-MicArray, für 15 Aspekte (ABPC ‚normal‘ – ABPC ‚Rear - ∞ ‘)

Zunächst sollen wieder jene Aspekte betrachtet werden, bei denen die Änderungen ungefähr in der Größenordnung eines Skalenwertes liegen:

Der Unterschied zum DECCA-System ist offenbar geringer geworden, wobei zugleich die Natürlichkeit und die empfundene Basisbreite abgenommen haben. Nicht verwunderlich ist, dass die Raumabbildung nun schlechter eingestuft wird, da ja das Signal der Rear-Kanäle komplett fehlt (fehlendes ‚Envelopment‘, siehe [Griesinger 1998])

Überraschend ist, dass sich in Hinblick auf die wahrgenommen Lautstärke und Auffälligkeit (Aspekte ‚Surr leise/laut, ‚Surr un-/auffällig‘), sowie den Entfernungseindruck (‚nahe/entfernt‘) bei den Hörern im Mittel offenbar kaum eine Änderung ergeben hat.

Die Mehrzahl der anderen Veränderungen sind vom Betrag her wohl zu gering um als signifikant gelten zu dürfen, stimmen allerdings von der Tendenz der Auswirkung her mit den aus der Psychoakustik zu erwartenden Konsequenzen einer Wegnahme der hinteren Surroundkanäle überein.

Erwähnenswert scheint auch, dass sich durch die Weglassung der Rear-Kanäle ausnahmslos alle nicht-quantitativen Aspekte in ihrer Bewertung verschlechtert haben; lediglich die Stabilität scheint sich marginal verbessert zu haben (nicht signifikant).

Dies bestätigt, wie wichtig das Vorhandensein eines (passend beschaffenen) Rear-Signals für das klangliche Gesamtergebnis ist.

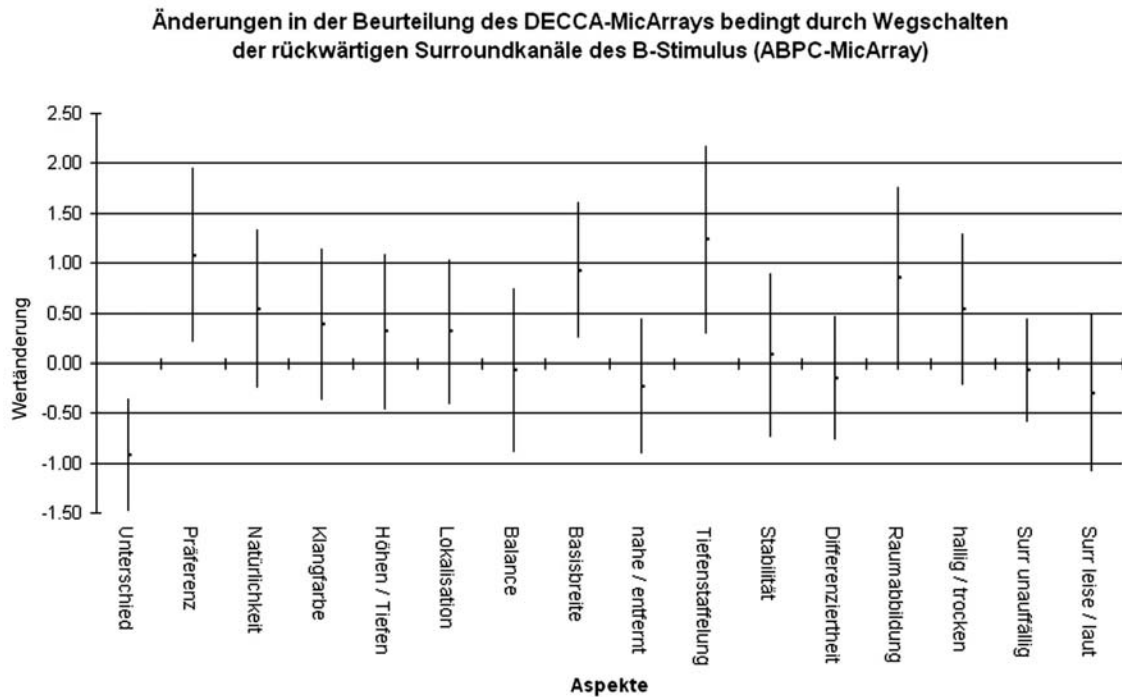


Abb. 8.6: Mittelwertänderung und 95% Konfidenzintervall beim DECCA-MicArray, für 15 Aspekte (ABPC ‚normal‘ – ABPC ‚Rear -∞‘)

Nachfolgend eine Besprechung lediglich der Aspekte, bei denen die Änderung ungefähr in der Größenordnung eines Skalenwertes liegt:

Das Signal des AB-PC Mic-Arrays ist durch das Wegschalten der hinteren Kanäle dem Signal des DECCA Systems ähnlicher geworden, das DECCA Mic-Array wird nun im Vergleich auch deutlich besser bewertet (Präferenz). Ebenso werden seine Basisbreite und Tiefenstaffelung als größer bewertet und ihm eine bessere Raumabbildung bescheinigt.

Von den insgesamt 12 Aspekten, die keine quantitative Bewertung betreffen, werden nun 8 besser eingestuft, als zuvor.

Generell lässt sich auch in diesem Fall bezüglich der vorliegenden Daten sagen, dass die an den beiden Stimuli A und B durch die Testhörer wahrgenommenen Änderungen – gemäß den Erkenntnissen der Psychoakustik - durchwegs erwartungsgemäß sind.

8.2 Präferenz-Abschätzung durch Mittelwertberechnung

(Abschnitt gestrichen)

8.3 Datenauswertung mittels Hauptfaktorenanalyse (Principal Component Analysis – PCA)

Eine gebräuchliche Methoden zur Auswertung von Daten, die durch Hörerbefragung gewonnen wurden, besteht darin, sie einer Faktoranalyse zu unterziehen (siehe hierzu [Bech, 1999], [Berg and Rumsey, 2001], [Gabrielsson, 1979], [Nakayama, 1971]). Innerhalb des Gebiets der Faktorenanalyse stellt die sog. Hauptfaktorenanalyse (Principal Component Analysis – PCA) einen Spezialfall dar, der auf die Ermittlung einer möglichst geringen Anzahl von Faktoren abzielt, die die vorgegebenen Daten („Ausgangsdatensatz“) hinreichend genau abbilden.

Vor Durchführung der Faktorenanalyse wird der Datensatz in der Regel normiert, womit alle Variablen den Mittelwert 0 und die Streuung 1 erhalten. (Durch die Normierung wird allen Aspekten a priori die gleiche Bedeutung zugesprochen.) Ausgehend von den Rohdaten wird zunächst – in unserem Fall: aspektbezogen - der Mittelwert für n Stichproben bestimmt:

$$\bar{x}_i = \frac{1}{n} \sum_{j=1}^n x_{ij} \quad (8.1)$$

mit x_{ij} als Bewertung, die für den i -ten Aspekt durch den j -ten Hörer abgegeben wurde. Die Standardabweichung s_i ergibt sich dann zu:

$$s_i = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (x_{ij} - \bar{x}_i)^2} \quad (8.2)$$

Der normiert Wert z_{ij} (bzw. für den gesamten Ausgangsdatensatz: die Z-Matrix [engl. z-score]) errechnet sich dann aus:

$$z_{ij} = \frac{x_{ij} - \bar{x}_i}{s_i} \quad (8.3)$$

Als Ergebnis der mittels MATLAB „princomp“-Function durchgeführten Faktorenanalyse ist in Abb. 8.7 die Relation aller 15 Aspekte inklusive der Hörerpräferenz in der Faktorenebene F1, F2 dargestellt.

Da diesen Ergebnissen unmittelbar die Resultate aus einer faktorenanalytischen Untersuchung über die subjektive Beurteilung der Akustik von Konzertsälen von [Wilkins, 1975] gegenübergestellt werden soll, sind nachfolgend die Gegensatzpaare aus jener Untersuchung aufgelistet, da sie zum besseren Verständnis der untenstehenden Abbildungen notwendig sind:

Gegensatzpaare aus der Untersuchung von Wilkins:

| | |
|-----------------|--------------------|
| 1 klein | groß |
| 2 angenehm | unangenehm |
| 3 undeutlich | deutlich |
| 4 weich | hart |
| 5 brillant | matt |
| 6 rund | spitz |
| 7 kräftig | gedämpft |
| 8 gefällt | gefällt nicht |
| 9 stumpf | scharf |
| 10 diffus | konzentriert |
| 11 aufdringlich | zurückhaltend |
| 12 hell | dunkel |
| 13 verschwommen | klar |
| 14 trocken | hallig |
| 15 schwach | stark |
| 16 höhenbetont | nicht höhenbetont |
| 17 tiefenbetont | nicht tiefenbetont |
| 18 schön | hässlich |
| 19 leise | laut |

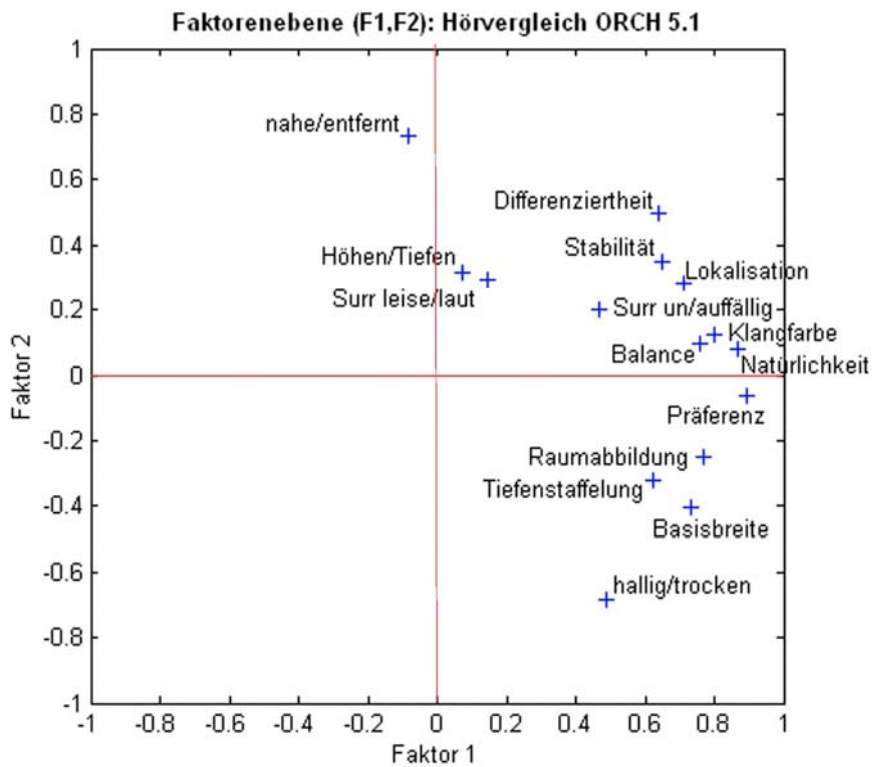


Abb. 8.7: Hörvergleich ORCH 5.1 – 15 Aspekte dargestellt in der Faktorenebene F1,F2

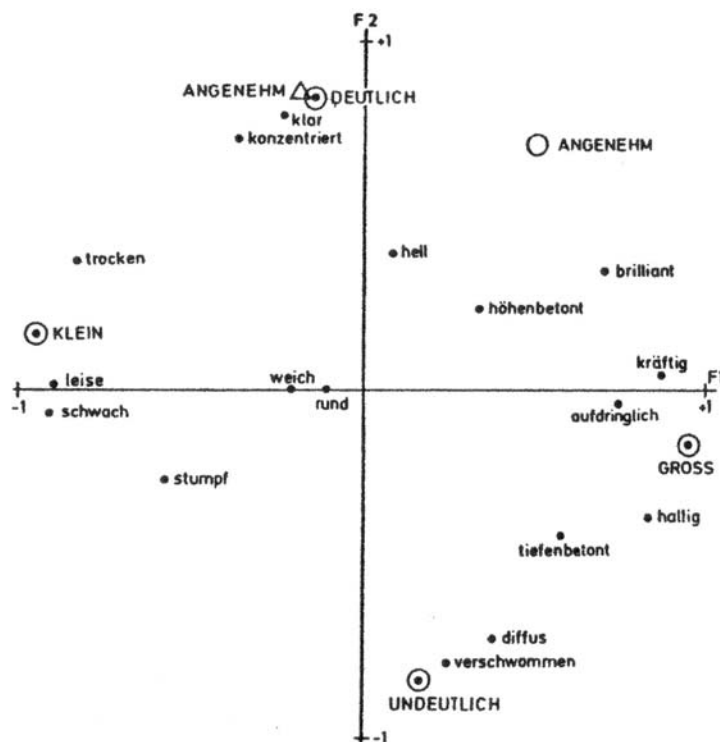


Abb. 8.8: Darstellung von 19 Aspekten in der Faktorenebene F1,F2 (aus [Wilkins, 1975b]; Gegensatzpaare und Sonderfälle extra gekennzeichnet)

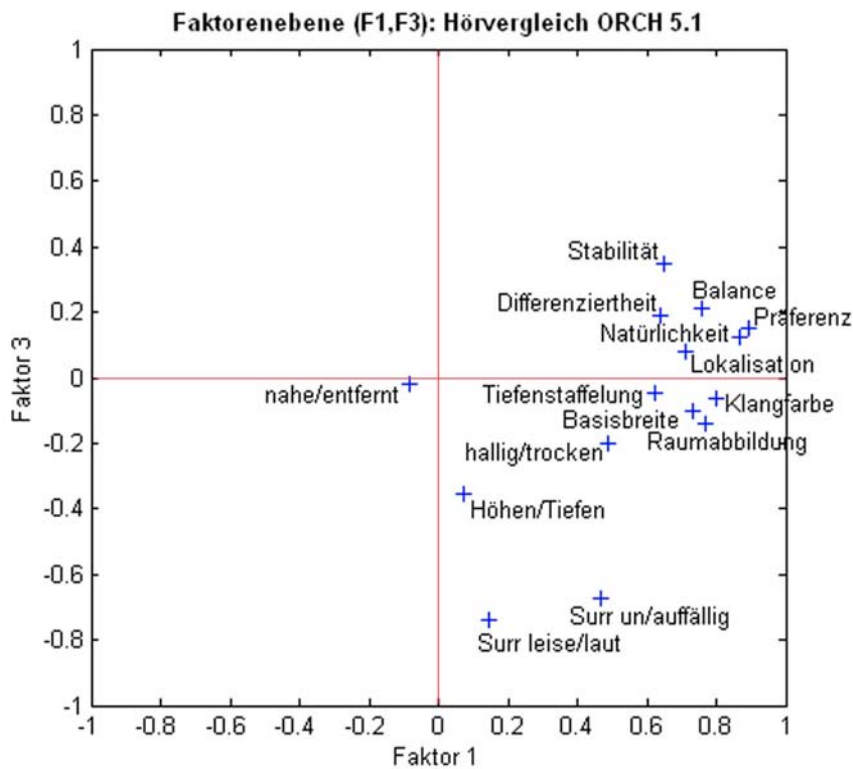


Abb. 8.9: Hörvergleich ORCH 5.1 – 15 Aspekte dargestellt in der Faktorenebene F1,F3

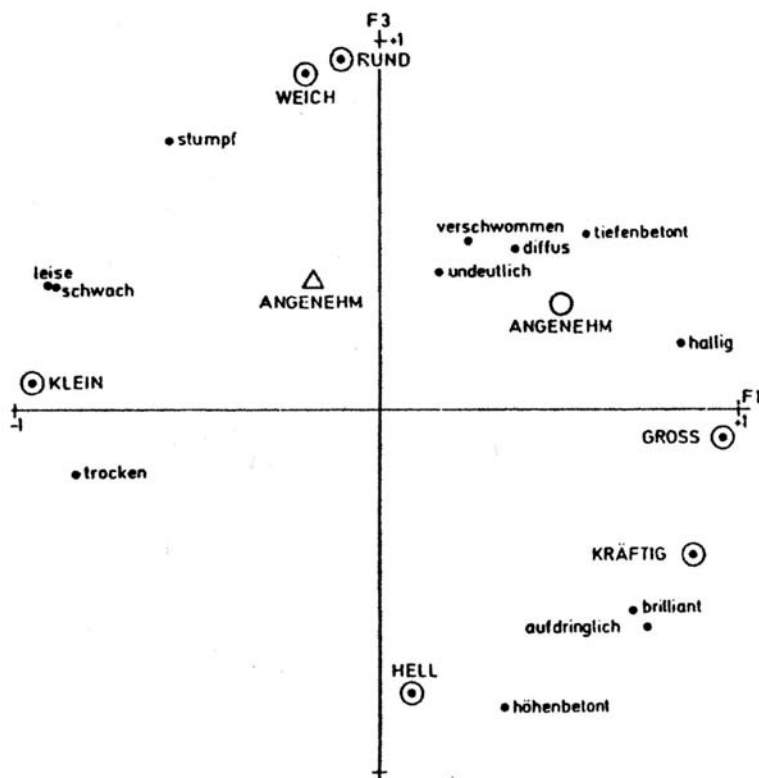


Abb. 8.10: Darstellung von 19 Aspekten in der Faktorenebene F1,F3 (aus [Wilkens, 1975 b])

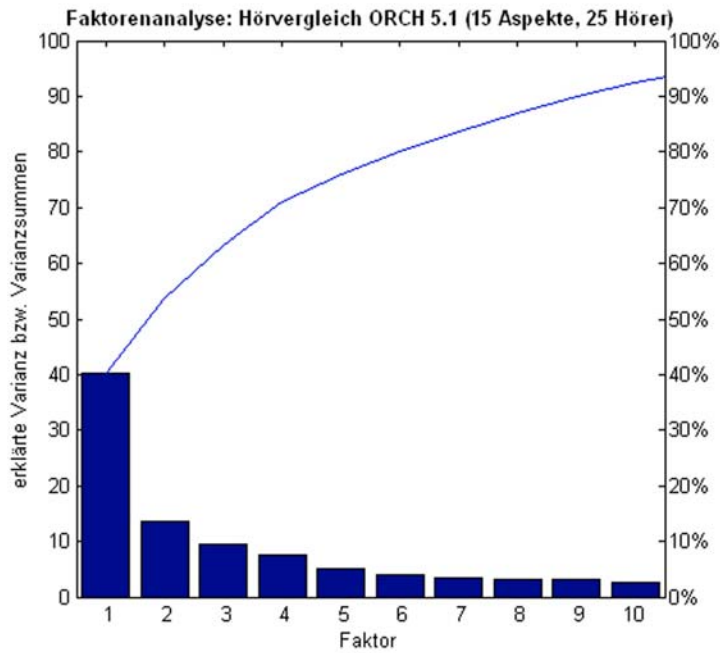


Abb. 8.11: Scree-Plot der Faktoren zu Hörvergleich ORCH 5.1

| Faktor Nr. | erklärte Varianz [%] | Eigenwert λ | Varianzsummen [%] |
|------------|----------------------|---------------------|-------------------|
| 1 | 40.3234 | 6.1096 | 40.3234 |
| 2 | 13.5133 | 2.0475 | 53.8367 |
| 3 | 9.5699 | 1.4500 | 63.4066 |
| 4 | 7.4842 | 1.1340 | 70.8908 |
| 5 | 5.1961 | 0.7873 | 76.0869 |
| 6 | 3.9616 | 0.6002 | 80.0485 |
| 7 | 3.5356 | 0.5357 | 83.5841 |
| 8 | 3.2438 | 0.4915 | 86.8279 |
| 9 | 3.1063 | 0.4706 | 89.9342 |
| 10 | 2.4793 | 0.3756 | 92.4135 |
| 11 | 1.9859 | 0.3009 | 94.3994 |
| 12 | 1.8424 | 0.2792 | 96.2418 |
| 13 | 1.7008 | 0.2577 | 97.9426 |
| 14 | 1.3452 | 0.2038 | 99.2878 |
| 15 | 0.7123 | 0.1079 | 100.0000 |

Tab. 8.4: Auflistung der den Faktoren zugehörigen Eigenwerten λ , sowie der durch sie erklärten Varianzen (sowie der aufsummierten Varianzen)

8.3.1 Anzahl der zu berücksichtigenden Faktoren

Eine Entscheidung über die in der Analyse zu berücksichtigende Anzahl von Faktoren wird in der Praxis in der Regel aufgrund folgender Kriterien getroffen:

- gemäß ‚Kaiser-Guttman Kriterium‘ sind Faktoren, deren Eigenwert λ kleiner oder gleich 1 ist, unberücksichtigt zu lassen (siehe hierzu [Bortz, 2005], [Guttman, 1954] und [Kaiser, 1959])
- gemäß ‚Scree-Test‘ von [Cattell, 1966] werden nur jene Faktoren als bedeutsam betrachtet, deren Werte höher als der ‚Knick‘ in der Kurvenfunktion liegen

Auch wenn der ‚Knick‘ in der Funktion ev. bereits zwischen erstem und zweitem Faktor gesehen werden kann, so legt eine Berücksichtigung des ‚Kaiser-Guttman Kriteriums‘ es doch nahe zumindest die ersten drei bis vier Faktoren in der Analyse zu berücksichtigen.

8.3.2 Interpretation der Faktoren

Wie aus Abb. 8.7 zu erkennen, ist der Faktor 1 (über dessen genaue Eigenschaft oder physikalisch-akustische Entsprechung wir vorerst noch keine Annahme getroffen haben) so gut wie in allen Aspekten vertreten, mit Ausnahme der Aspekte ‚nahe/entfernt‘, ‚Surr. leise/laut‘ und ‚Höhen/Tiefen‘, an welchen der Faktor 1 nur einen geringen Anteil hat.

Die Darstellung der Aspekte in der Faktorenebene lässt sich graphisch interpretieren: In Form eines Vektordiagramms werden die Aspekte (Variablen) so als Vektoren dargestellt, daß der von zwei Vektoren eingeschlossene Winkel dem Korrelationskoeffizienten dieser beiden Variablen entspricht.

Dabei entspricht ein Winkel von 90° einem Korrelationskoeffizienten von $r = 0$, ein Winkel von 0° entsprechend einem $r = 1$.

Aus Abb. 8.7 ist zu erkennen, daß die Variablen Präferenz, Natürlichkeit, Klangfarbe und Balance eine relativ große Korrelation untereinander aufweisen, da die zugehörigen Winkel der Vektorendiagramme klein sind. Außerdem weichen sie nur gering von der Achsenrichtung des Faktors 1 ab und die Längen ihrer Vektoren ist groß, was bedeutet, dass sie mit diesem Faktor nicht nur stark korreliert, sondern auch

anteilmäßig stark in ihm vertreten sind.

Abgesehen von der Variablen ‚Präferenz‘ (die streng genommen keinen (akustischen) Aspekt darstellt), sind die der Achse des 1. Faktors zunächst gelegenen Aspekte Natürlichkeit, Klangfarbe, Balance und Raumabbildung. Der 1. Faktor stellt weiters quasi die ‚Winkelhalbierende‘ zwischen den Vektoren der Aspekte Raumabbildung und Lokalisation dar, welche offenbar leichte Gegensatzpaare bilden.

Das lässt den Rückschluss zu, dass der Faktor 1 eine Empfindung sein könnte, die eine Art ‚Mittelding‘ zwischen konkret zu lokalisierendem Klangreiz und einer diffusen Art von ‚Einhüllung‘ darstellt. Die ‚richtige‘ Mischung zwischen diesen beiden Extremen – und da auch der Aspekt (Lautstärken-) Balance zur Achse des 1. Faktors stark korreliert ist – drängt sich die Frage auf, ob es sich bei diesem ev. um die in anderen Untersuchungen gefundene ‚Fullness‘ (also [Klang-]fülle, Größe) oder ‚Width‘ (‚Weite‘ in einem allgemeinen Sinn) handelt. Während Lokalisation positiv mit dem zweiten Faktor korreliert ist, liegt für den Aspekt Raumabbildung eine negative Korrelation mit diesem vor.

Bestärkt wird diese Vermutung durch Ergebnisse der Faktorenanalyse, die sich bei [Wilkens, 1975] finden. Wie in Abb. 8.8 zu sehen liegen jene Adjektive, zu denen es auch in der vorliegenden Arbeit Entsprechungen gibt, an durchaus vergleichbaren Positionen:

| Pfanzagl | Wilkens |
|-----------------|----------------|
| Höhen/Tiefen | hell |
| nahe/entfernt | deutlich, klar |
| hallig/trocken | hallig |

Aufgrund dieser Übereinstimmungen lässt sich nach Erachten des Autors der Schluss ziehen, dass es sich auch in der vorliegenden Untersuchung beim Faktor 1 um die *“Empfindung der Stärke und Ausdehnung der Quelle”* handelt, wie von Wilkens formuliert.

Annähernd orthogonal zu den oben genannten Aspekten stehen die Vektoren der Aspekte ‚nahe/entfernt‘ und ‚hallig/trocken‘, die einander gegenüberliegen. Die Tatsache, dass ersterer Aspekt sehr nahe der Achse des 2. Faktors liegt, lässt die Schlussfolgerung zu, dass der 2. Faktor mit dem ‚Entfernungseindruck‘ (oder einem damit eng verwandten (psycho-) akustischen Begriff) in Zusammenhang steht bzw. als raumakustische Größe (objektives Kriterium) interpretiert werden darf, bei der das Verhältnis von Direkt- zu Diffusschall eine große Rolle spielt. Genauer gesagt ist das Adjektiv ‚nahe‘ positiv mit dem 2. Faktor korreliert, sowie weiters die Aspekte Differenziertheit, Stabilität, ein etwas ‚Höhen‘-lastiger (‚präsender‘, direkter) Klang sowie ein ‚leises‘ Signal in den hinteren Surroundkanälen: mit einem Wort alle Aspekte, die zu einem eher direkten, deutlichen Klangeindruck gehören. Im Englischen würde wohl der Begriff ‚nearness‘, oder ‚definition‘ am besten dazu passen. Im Deutschen bietet sich der – auch messtechnisch aufzufassende – Begriff ‚Deutlichkeit‘ (D) an, welcher sich aus dem Verhältnis von Direkt-Schall inklusive der ersten Reflexionen (genauer: jenes Signal, das innerhalb der ersten 50ms am Hörerplatz eintrifft) zur Gesamtenergie (Integrationszeit theoretisch bis unendlich), also letztlich die Relation trocken/hallig ergibt und nach [Thiele, 1953] definiert ist, wie folgt:

$$D = \frac{\int_0^{50 \text{ ms}} p^2 dt}{\int_0^{\infty} p^2 dt} \quad (8.4)$$

mit: p ... Schalldruck

Allerdings wäre es auch durchaus möglich, den Faktor 2 einem eng verwandten Begriff, nämlich dem musikspezifischeren ‚Klarheitsmaß‘ (C) zuzuordnen, das nach [Reichardt, 1972] folgendermaßen definiert ist:

$$C = 10 \lg \frac{\int_0^{80 \text{ ms}} p^2 dt}{\int_{80 \text{ ms}}^{\infty} p^2 dt} \text{ dB} \quad (8.5)$$

Die Annahme, dass es sich beim zweiten Faktor um die empfundene Deutlichkeit handelt, wird dadurch erhärtet, dass auch bei der Faktorenanalyse von Wilkens der 2. Faktor mit den Adjektiven ‚deutlich‘ und ‚klar‘ korreliert ist.

Betrachten wir die Relationen der Aspekte gemäß Vektordiagrammen, so scheint es naheliegend, dass die Aspekte Raumabbildung und Tiefenstaffelung eng korreliert sind. Weiters fällt auf, dass der Aspekte Basisbreite offenbar mit dem Aspekt Tiefenstaffelung ebenfalls stark korreliert ist, da der von den zugehörigen Vektoren eingeschlossene Winkel nahe Null ist. Dies scheint insofern ganz plausibel, da die Erstreckung einer ausgedehnten Schallquelle in der Y-Achse (‚Tiefenstaffelung‘) dem entspricht, was in der X-Achse durch die Angabe der ‚Basisbreite‘ einer Schallquelle zum Ausdruck gebracht wird.

Im 3. Faktor ist von allen Aspekten die ‚Stabilität‘ am stärksten von allen vertreten, allerdings ist sein Anteil mit einem Wert von 0.33 nicht sehr hoch (siehe Abb. 8.9 bzw. Tab 8.5). Am stärksten – allerdings negativ – korreliert sind die Aspekte Surr. leise/laut und Höhen/Tiefen.

Aus der hohen negativen Korrelation ließe sich der Schluss ableiten, dass der 3. Faktor positiv mit dem Gegenteil der jeweiligen Aspekte assoziiert sein müsste, d.h. also z.B. ‚Surr. laut‘ (Signal der hinteren Lautsprecher laut, was – vor allem einen Diffusschallanteil in den Signalen LS, RS vorausgesetzt, wie es bei der vorliegenden Untersuchung der Fall ist – normalerweise auch in einer leichten Erhöhung des Entfernungseindrucks bzw. Verstärkung des Raumeindrucks resultiert ->), ‚entfernt‘ und ‚tiefenbetont‘ (wie bereits weiter oben beschrieben: Bedämpfung der Höhen mit steigender Entfernung aufgrund von Absorption)

Allerdings spielt beim 3. Faktor der Anteil des Diffusschalls (bzw. das Verhältnis Diffus-/Direktschall) offenbar keine große Rolle, da der Aspekt hallig/trocken mit einem Wert von -0.20 so gut wie keinen Anteil hat. Der Entfernungseindruck könnte somit u.a. indirekt – z.B. durch das subjektiv empfundene Fehlen von Höhen (ev. im Zusammenhang mit einem Übermaß an tiefen Frequenzen) – bedingt sein.

Mit einer Erhöhung des Pegels der hinteren Kanäle geht in der Regel auch eine bessere akustische ‚Einhüllung‘ einher [Griesinger, 1996], was den Schluss zuließe,

dass hinter dem 3. Faktor der psychoakustische Begriff ‚Envelopment‘ stehen könnte. Im Scree-Plot der Abb. 8.11 sind die Prozentsätze dargestellt, wie viel Prozent der Varianz durch den jeweiligen Hauptfaktor erklärt werden. In der Abbildung ist zu erkennen, dass der 1. Faktor ca. 40% ausmacht, der 2. Faktor 13,5 % und der 3. Faktor 9,6%; die ersten 3 Hauptfaktoren erklären also gemeinsam bereits rund 63% der Gesamtvarianz.

Aus diesem Grund soll auch auf einen Interpretationsversuch für den 4. Faktor verzichtet werden, da dieser mit einem Eigenwert λ von nur 1,13 kaum über dem Kaiser-Guttman Kriterium (siehe oben) liegt.

Untenstehend sind – der Vollständigkeit halber – die numerischen Werte der Faktorenladungen der 15 Aspekte bezüglich der ersten 4 Faktoren in Form einer tabellarischen Auflistung angeführt:

| | Fakt. 1 | Fakt. 2 | Fakt. 3 | Fakt. 4 |
|----------------------|---------|---------|---------|---------|
| 1 Präferenz | 0.8955 | -0.0623 | 0.1493 | -0.0076 |
| 2 Natürlichkeit | 0.8658 | 0.0770 | 0.1242 | 0.1397 |
| 3 Klangfarbe | 0.8016 | 0.1237 | -0.0616 | 0.0220 |
| 4 Höhen/Tiefen | 0.0723 | 0.3161 | -0.3536 | -0.7966 |
| 5 Lokalisation | 0.7092 | 0.2827 | 0.0776 | -0.1947 |
| 6 Balance | 0.7602 | 0.0948 | 0.2114 | 0.1479 |
| 7 Basisbreite | 0.7314 | -0.4046 | -0.1009 | -0.0000 |
| 8 nahe / entfernt | -0.0821 | 0.7345 | -0.0202 | -0.2527 |
| 9 Tiefenstaffelung | 0.6226 | -0.3221 | -0.0467 | -0.2701 |
| 10 Stabilität | 0.6490 | 0.3495 | 0.3502 | 0.1927 |
| 11 Differenziertheit | 0.6387 | 0.4968 | 0.1885 | -0.0130 |
| 12 Raumabbildung | 0.7700 | -0.2511 | -0.1427 | -0.0495 |
| 13 hallig/trocken | 0.4905 | -0.6838 | -0.2009 | -0.1625 |
| 14 Surr un/auffällig | 0.4656 | 0.2015 | -0.6743 | 0.2084 |
| 15 Surr leise/laut | 0.1436 | 0.2939 | -0.7392 | 0.4091 |

Tab. 8.5: Auflistung der Faktorenladungen (Hörvergleich ORCH 5.1)

8.3.3 Grenzen der Faktorenanalyse

Der Autor möchte die in den vorangehenden Abschnitten aus der Faktorenanalyse gezogenen Schlüsse vorsichtig interpretiert wissen, da es bei der FA einige Grundannahmen gibt, die nicht unbedingt bei allen Sachverhalten - welche man mittels der Faktorenanalyse untersuchen möchte - zutreffen müssen:

Die FA setzt unter anderem voraus, dass die ursprünglichen Ausgangsdaten (zumindest näherungsweise) als Linearkombination zweier Matrizen (Faktorenladungsmatrix A und Faktorenmatrix F) aufgefasst werden können.

Ein ev. Schwachpunkt der FA kann darin gesehen werden, dass das Ergebnis der FA wesentlich durch die Wahl der Gegensatzpaare beeinflusst sein kann, indem entweder eine wichtige Hörempfindung nicht oder nur ungenügend repräsentiert ist, oder ein Faktor durch viele eng verwandte Attribute (Aspekte) überbetont sein kann (siehe hierzu [Cremer, 1978], S.471)

Wie bei [Bortz, 2005] zu lesen, ist die PCA als ein “ ...Daten reduzierendes und Hypothesen generierendes Verfahren nicht dazu geeignet, inhaltliche Hypothesen über die Art einer Faktorenstruktur zu überprüfen. Die Uneindeutigkeit des Verfahrens, die auf der formalen Gleichwertigkeit verschiedener Rotationslösungen beruht, lässt es nicht zu, eine Lösung als richtig und eine andere als falsch zu bezeichnen. ... Man entscheidet sich letztlich für diejenige Lösung, die nach dem jeweiligen Stand der Theorienbildung am plausibelsten ist. ...” (S.522/523)

8.4 Clusteranalyse

Die Clusteranalyse des vorliegenden Datensatzes wurde nach Durchführung einer multivariaten Varianzanalyse (MANOVA) mittels der ‚manovacluster‘-Funktion in MATLAB berechnet und basierte auf verschiedenen Linkage-Methoden (Single-Linkage, Complete-Linkage). Die Clusteranalyse gibt Aufschlüsse über die ‚Ähnlichkeit‘ (Verwandtschaft) der Aspekte untereinander.

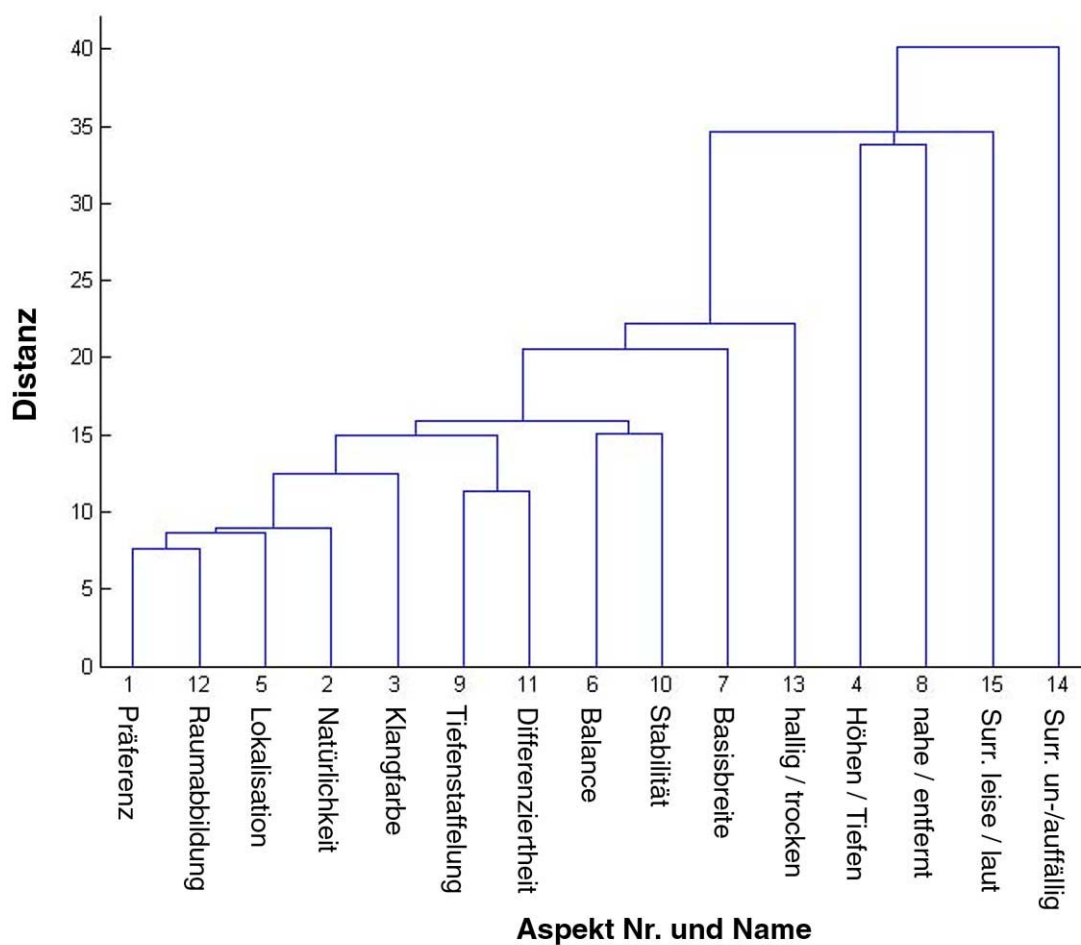


Abb. 8.12: Dendrogramm-Darstellung der 15 Aspekte des Hörvergleichs ORCH 5.1 Surround (Berechnung nach ‚Single Linkage‘ Methode)

Nachfolgend eine kurze Erklärung der Unterscheide zwischen ‚Single- und ‚Complete-Linkage‘ Clusteranalyse-Methode, jeweils gekürzt zitiert aus [Bortz, 2005]):

„Single Linkage“ Methode (auch Minimum-Methode genannt)

Hierbei richtet sich die Ähnlichkeit zweier Cluster nach den paarweisen Ähnlichkeiten der Objekte des einen Clusters zu den Objekten des anderen Clusters: es werden diejenigen zwei Cluster vereint, welche die zueinander am nächsten liegenden Nachbarobjekte („nearest neighbour“) besitzen.

Dadurch, daß jeweils nur zwei beieinander liegende Einzelobjekte über die Fusionierung zweier Cluster entscheiden, kann es zu Verkettungen bzw. kettenförmigen Clustergebilden kommen, in denen sich Objekte befinden, die zueinander eine geringere Ähnlichkeit aufweisen, als zu Objekten anderer Cluster.

„Complete Linkage“ Methode (auch Maximum-Methode genannt)

Dieses Cluster-Kriterium bestimmt auf jeder Fusionsstufe für alle Paare von Clustern die jeweils am weitesten entfernten Objekte („furthest neighbour“); es werden dann allerdings diejenigen Cluster fusioniert, für die diese Maximaldistanz minimal ist. Da das Kriterium auf diese Weise alle Einzelbeziehungen berücksichtigt, ist – anders als bei „single linkage“ – gewährleistet, dass alle paarweisen Objektähnlichkeiten innerhalb eines Clusters kleiner sind, als der Durchschnitt der paarweisen Ähnlichkeiten zwischen verschiedenen Clustern. In diesem Sinn resultiert „complete linkage“ in homogenen Clustern und ist damit für viele Fragestellungen geeignet.

Bei Betrachtung der Abb. 8.12 fällt auf, daß der Aspekt Raumabbildung die größte Nähe zum Aspekt Präferenz aufweist. Dies bestätigt, dass eine gute räumliche Abbildung des Klangereignisses von größter Wichtigkeit für das subjektive Gefallen ist. Wie in der Zusammenfassung (Abschnitt 8.7) festgestellt wird, steht dies auch in guter Übereinstimmung mit den dort näher erläuterten Ergebnissen der Arbeiten von [Toole, 1985], sowie [Bech et al., 1995].

Mit dem nächsten Verwandtschaftsgrad zur Präferenz folgen die Aspekte Lokalisation und Natürlichkeit.

Gemäß den Dendrogrammen in den Abbildungen 8.12 und 8.13 besteht Verwandtschaft zwischen Differenziertheit und Tiefenstaffelung, wie auch zwischen Balance und Stabilität.

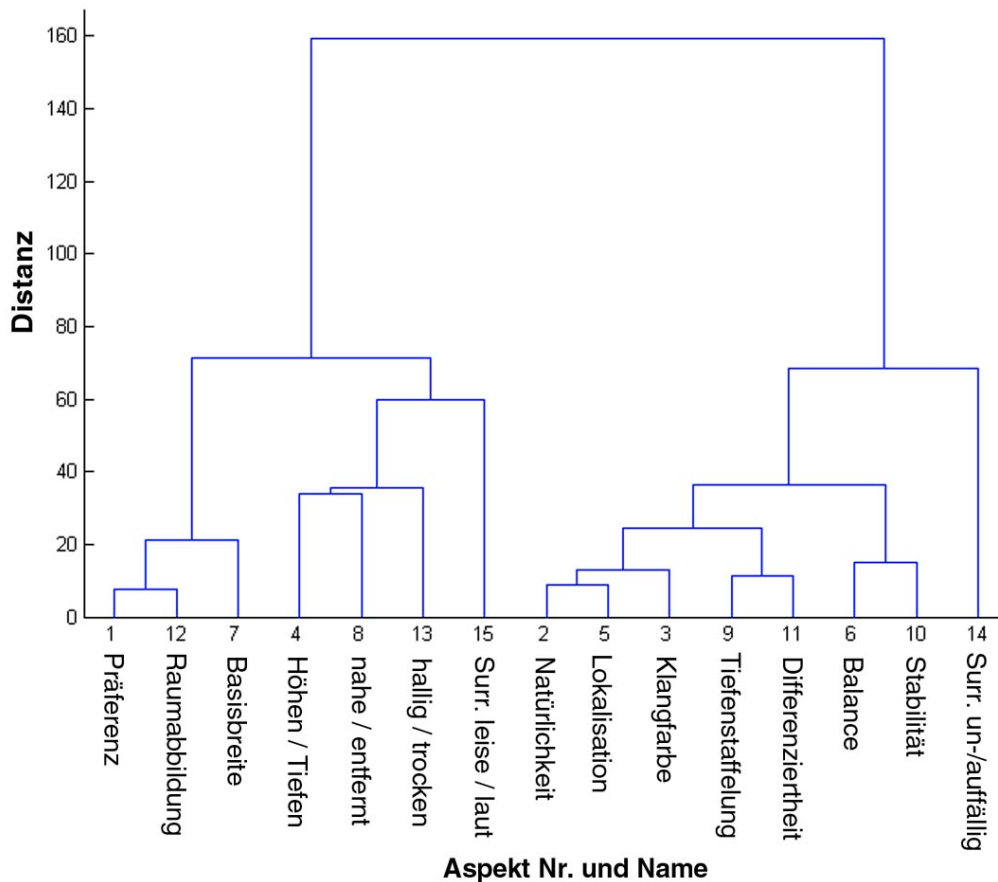


Abb. 8.13: Dendrogramm-Darstellung der 15 Aspekte des Hörvergleichs ORCH 5.1 Surround (Berechnung nach ‚Complete Linkage‘ Methode)

Aus der Dendrogramm-Berechnung nach der ‚Complete Linkage‘ Methode ist die enge Verwandtschaft zwischen den Aspekten Höhen/Tiefen, nahe/entfernt und hallig/trocken noch besser herauszulesen, als aus dem ersten Dendrogramm nach ‚Single Linkage‘ Methode. Die Verwandtschaft des Paares nahe/entfernt, hallig/trocken bedarf wohl keiner weiteren Begründung, dass auch der Aspekt ‚Höhen/Tiefen‘ in so enger Verbindung zu ihnen steht ist insofern nicht unplausibel, da der Eindruck von ‚Nähe‘ in der Regel mit erhöhter ‚Präsenz‘ der Klangquelle einhergeht.

In relativer Nähe zu diesen drei offenbar stark Raumklang-spezifischen Aspekten steht dann auch der Aspekt Surr. leise/laut, was von der Logik her stimmig erscheint.

Generell betrachtet scheinen die linken sechs Aspekte (neben Präferenz) insgesamt eher raumspezifische Aspekte zu sein, während die rechts stehenden Aspekte zu einer Klangquellen-spezifischen Gruppe zu gehören scheinen, deren ‚übergeordneter‘ Aspekt die Natürlichkeit ist.

Insofern ist es auch nicht verwunderlich, dass der Aspekt Surr. un-/auffällig eher zu dieser Gruppe gehört, kommt es doch (abgesehen natürlich vom Pegel) primär auf den Signalinhalt der Surroundkanäle an, ob diese als auffällig oder unauffällig empfunden werden. (Zu dieser wichtigen Problematik folgt in Abschnitt 8.6 noch eine genauere Analyse)

Lokalisation und Tiefenstaffelung sind die einzigen Aspekte in dieser Gruppe, die eigentlich eher zu den Räumlichkeits-Aspekten zählen. Stimmig ist allerdings, dass die Tiefenstaffelung mit relativ geringer Distanz (d.h. engem Verwandtschaftsgrad) eine Verbindung zum Aspekt Differenziertheit aufweist; ebenso scheint die Nähe von Balance und Stabilität sehr plausibel.

Die drei Aspekte mit größter Verwandtschaft in der rechten (*‚Klangquellen-‘*) Gruppe sind also Natürlichkeit, Lokalisation und Klangfarbe, während dies in der linken (*‚Räumlichkeits-‘*) Gruppe die Aspekte Präferenz, Raumabbildung und Basisbreite sind.

Betrachtet man die Abbildungen der Faktorenanalyse im vorhergehenden Abschnitt, so sind auch dort diese ‚Untergruppen‘-Bildungen durch die relative Nähe ihrer Koordinaten (sowie kleine Winkel zwischen den zugehörigen Vektoren) ersichtlich.

Es fällt auf, dass bei der durchgeführten Faktorenanalyse die Aspekte Natürlichkeit und Balance in jeder der dargestellten Faktorenebenen eng korreliert erscheinen. Dieses Verhältnis tritt in den Dendrogrammen der Clusteranalyse nicht in Erscheinung, was etwas überrascht.

Ebenfalls in den Faktorenebenen-Plots ersichtlich ist, dass zur Natürlichkeit außerdem die Aspekte Lokalisation und Klangfarbe in der Regel relativ eng korreliert sind.

Grenzen der Clusteranalyse (zitiert aus [Bortz, 2005], S. 565)

„Trotz der umfangreichen in diesem Bereich geleistete Forschungsarbeit basiert keine der heute verfügbaren Clustermethoden auf einer Theorie, die es gewährleistet, daß die beste Struktur der Objekte entdeckt wird. An diesem Faktum hat sich seit den Anfängen der Clusteranalyse nichts geändert, die mit einer Bewertung Tryons, die Clusteranalyse sei “die Faktorenanalyse der armen Leute” (vgl. [Tryon, 1939]), insoweit treffend beschrieben ist.“

8.5 Korrelationsanalyse

Die 15 Aspekte der vorliegenden Studie wurden auch einer Korrelationsanalyse unterworfen, für welche die ‚corr‘-Funktion in MATLAB verwendet wurde.

Ein Auszug aus der sich ergebenden Korrelationsmatrix ist in der untenstehenden Tabelle aufgelistet.

| | Präferenz | Natrl. | Klangf. | Raumabb. | Basisbr. | Lokal. |
|-------------------|---------------|---------------|---------------|---------------|---------------|---------------|
| Präferenz | 1 | 0.8383 | 0.6712 | 0.6829 | 0.6289 | 0.5875 |
| Natürlichkeit | 0.8383 | 1 | 0.7539 | 0.6398 | 0.5010 | 0.5967 |
| Klangfarbe | 0.6712 | 0.7539 | 1 | 0.5724 | 0.4978 | 0.5113 |
| Höhen/Tiefen | 0.0238 | -0.0377 | 0.0935 | 0.0083 | -0.0054 | 0.2035 |
| Lokalisation | 0.5875 | 0.5967 | 0.5113 | 0.4336 | 0.4410 | 1 |
| Balance | 0.6879 | 0.5964 | 0.5178 | 0.4867 | 0.4748 | 0.5292 |
| Basisbreite | 0.6289 | 0.5010 | 0.4978 | 0.6522 | 1 | 0.4410 |
| nahe/entfernt | -0.0964 | -0.0530 | 0.0392 | -0.0901 | -0.2470 | 0.1248 |
| Tiefenstaffelung | 0.4858 | 0.3996 | 0.4308 | 0.4892 | 0.4637 | 0.3497 |
| Stabilität | 0.5407 | 0.6025 | 0.4845 | 0.2910 | 0.3021 | 0.4676 |
| Differenziertheit | 0.5011 | 0.5017 | 0.4772 | 0.2661 | 0.2438 | 0.5234 |
| Raumabbildung | 0.6829 | 0.6398 | 0.5724 | 1 | 0.6522 | 0.4336 |
| hallig/trocken | 0.4288 | 0.2912 | 0.2808 | 0.5099 | 0.6483 | 0.1330 |
| Surr un/auffällig | 0.2988 | 0.3512 | 0.3836 | 0.3826 | 0.2987 | 0.2875 |
| Surr leise/laut | -0.0069 | 0.0827 | 0.1738 | 0.0632 | 0.0220 | 0.0718 |

Tab. 8.6: Auszugs aus der Korrelationsmatrix für die Aspekte des Hörvergleichs ‚ORCH 5.1‘

Von vorrangigem Interesse scheint die Korrelationen der verschiedenen Aspekte zur Präferenz und zur Natürlichkeit:

Dabei zeigt sich, dass zur Präferenz die Natürlichkeit (mit $r=0.8383$) die größte Korrelation hat, danach folgen Balance ($r=0.6879$) und Raumabbildung ($r=0.6829$). Ähnlich stark mit der Präferenz korreliert sind außerdem Klangfarbe ($r=0.6712$) und Basisbreite ($r=0.6289$).

Mit der Natürlichkeit sind (nach der Präferenz natürlich) vor allem die Aspekte Klangfarbe ($r=0.7539$) und Raumabbildung ($r=0.6398$) korreliert. Danach folgen Stabilität, Balance und Lokalisation alle mit einem Wert nahe $r=0.6$.

Generell interessieren vor allem Korrelationen mit einem Wert $r>0.6$:

Der Aspekt Klangfarbe ist (außer mit Präferenz und Natürlichkeit) – erstaunlicher Weise - am stärksten mit der Raumabbildung korreliert ($r=0.5724$). Danach folgen Lokalisation und Basisbreite mit einem Wert um 0.51.

Die Raumabbildung ist – neben Präferenz und Natürlichkeit – vor allem mit der Basisbreite korreliert ($r=0.6522$). Nicht weiter überrascht die Korrelation zum Aspekt hallig/trocken ($r=0.5099$).

Die Basisbreite ist zur Natürlichkeit ($r=0.5010$) weniger korreliert, als zur Präferenz ($r=0.6289$), stärker allerdings zu den Aspekten Raumabbildung ($r=0.6522$) und hallig/trocken ($r=0.6483$).

Die Lokalisation ist zur Balance ($r=0.5292$) und zur Differenziertheit (0.5234) etwas weniger stark korreliert, als zu Präferenz ($r=0.5875$) und Natürlichkeit ($r=0.5967$)

Die stärksten Korrelationen bestehen also zwischen:

| Aspekte-Kombination | Korr. Koeff. |
|-----------------------------|-------------------------|
| Präferenz – Natürlichkeit | 0.8383 |
| Natürlichkeit – Klangfarbe | 0.7539 |
| Präferenz – Balance | 0.6879 |
| Präferenz – Raumabbildung | 0.6829 |
| Präferenz - Klangfarbe | 0.6712 |
| Raumabbildung – Basisbreite | 0.6522 |

Tab. 8.7: Auflistung der 6 stärksten Korrelationen

Eine komplette Auflistung aller Korrelationen ist untenstehend in Tab. 8.8. zu finden. An dieser Stelle soll auch erwähnt werden, dass die Korrelation zwischen den Aspekten Surr. un-/auffällig und Surr. leise/laut sich mit $r=0.5091$ als statistisch signifikant ergab, da bezüglich dieser beiden Aspekte im nächsten Kapitel eine genauere Analyse vorgenommen werden soll.

Die in der vorliegenden Arbeit gefundene hohe Korrelation ($r=0.8383$) zwischen Präferenz und Natürlichkeit scheint durchaus erwartungsgemäß, allerdings war in der Studie von [Berg und Rumsey, 2001] für diese beiden Aspekte eine (ebenfalls relativ hohe) Korrelation von lediglich 0.595 festgestellt worden (wobei andere Aspekte höhere Korrelation aufgewiesen hatten).

Gute Übereinstimmung mit den Ergebnissen von [Berg und Rumsey, 2001] zeigt sich hingegen hinsichtlich der Korrelation zwischen Präferenz und Raumabbildung: bei der vorliegenden Studie ergibt sich der obenstehende Wert von 0.6829, bei Berg und Rumsey tritt für die Aspekt-Paarung Preference – Envelopment ein Wert von $r=0.674$ auf. Im obenstehenden Abschnitt betr. Clusteranalyse hat der Autor bereits dargestellt, welchen Zusammenhang er zwischen dem Aspekt ‚Raumabbildung‘ der vorliegenden Studie und dem Aspekt ‚Envelopment‘ der Studie Berg/Rumsey sieht.

DOKU Korrelation der Aspekte:
 >> [RHO,PVAL] = corr(Y,Y)

RHO =

| | Präf. | Natrl. | Klangf. | Hö./Tief | Lokal. | Bal. | Basisbr. | nahe/entf. |
|-----------------|---------|---------|---------|----------|--------|---------|----------|------------|
| Präferenz | 1.0000 | 0.8383 | 0.6712 | 0.0238 | 0.5875 | 0.6879 | 0.6289 | -0.0964 |
| Natürlichkeit | 0.8383 | 1.0000 | 0.7539 | -0.0377 | 0.5967 | 0.5964 | 0.5010 | -0.0530 |
| Klangfarbe | 0.6712 | 0.7539 | 1.0000 | 0.0935 | 0.5113 | 0.5178 | 0.4978 | 0.0392 |
| Höhen/Tiefen | 0.0238 | -0.0377 | 0.0935 | 1.0000 | 0.2035 | -0.0645 | -0.0054 | 0.2603 |
| Lokalisation | 0.5875 | 0.5967 | 0.5113 | 0.2035 | 1.0000 | 0.5292 | 0.4410 | 0.1248 |
| Balance | 0.6879 | 0.5964 | 0.5178 | -0.0645 | 0.5292 | 1.0000 | 0.4748 | -0.0219 |
| Basisbreite | 0.6289 | 0.5010 | 0.4978 | -0.0054 | 0.4410 | 0.4748 | 1.0000 | -0.2470 |
| nahe/entfernt | -0.0964 | -0.0530 | 0.0392 | 0.2603 | 0.1248 | -0.0219 | -0.2470 | 1.0000 |
| Tiefenstaffling | 0.4858 | 0.3996 | 0.4308 | 0.1067 | 0.3497 | 0.3813 | 0.4637 | -0.2565 |
| Stabilität | 0.5407 | 0.6025 | 0.4845 | -0.0402 | 0.4676 | 0.5312 | 0.3021 | 0.0675 |
| Diffmzrt.heit | 0.5011 | 0.5017 | 0.4772 | 0.1603 | 0.5234 | 0.5742 | 0.2438 | 0.1955 |
| Raumabbildung | 0.6829 | 0.6398 | 0.5724 | 0.0083 | 0.4336 | 0.4867 | 0.6522 | -0.0901 |
| hallig/trocken | 0.4288 | 0.2912 | 0.2808 | 0.0112 | 0.1330 | 0.2790 | 0.6483 | -0.4005 |
| Surr un-/auff. | 0.2988 | 0.3512 | 0.3836 | 0.1406 | 0.2875 | 0.2021 | 0.2987 | 0.0471 |
| surr leise/lt. | -0.0069 | 0.0827 | 0.1738 | 0.0777 | 0.0718 | 0.1135 | 0.0220 | 0.0777 |

| | Tiefenst. | Stabil. | diff. | raumabb. | hall/trk | s.un/auff. | s.leis/lt. |
|-----------------|-----------|---------|--------|----------|----------|------------|------------|
| Präferenz | 0.4858 | 0.5407 | 0.5011 | 0.6829 | 0.4288 | 0.2988 | -0.0069 |
| Natürlichkeit | 0.3996 | 0.6025 | 0.5017 | 0.6398 | 0.2912 | 0.3512 | 0.0827 |
| Klangfarbe | 0.4308 | 0.4845 | 0.4772 | 0.5724 | 0.2808 | 0.3836 | 0.1738 |
| Höhen/Tiefen | 0.1067 | -0.0402 | 0.1603 | 0.0083 | 0.0112 | 0.1406 | 0.0777 |
| Lokalisation | 0.3497 | 0.4676 | 0.5234 | 0.4336 | 0.1330 | 0.2875 | 0.0718 |
| Balance | 0.3813 | 0.5312 | 0.5742 | 0.4867 | 0.2790 | 0.2021 | 0.1135 |
| Basisbreite | 0.4637 | 0.3021 | 0.2438 | 0.6522 | 0.6483 | 0.2987 | 0.0220 |
| nahe/entfernt | -0.2565 | 0.0675 | 0.1955 | -0.0901 | -0.4005 | 0.0471 | 0.0777 |
| Tiefenstaffling | 1.0000 | 0.2660 | 0.3283 | 0.4892 | 0.5055 | 0.2367 | -0.0427 |
| Stabilität | 0.2660 | 1.0000 | 0.6547 | 0.2910 | 0.0415 | 0.2213 | 0.0217 |
| Diffmzrt.heit | 0.3283 | 0.6547 | 1.0000 | 0.2661 | 0.0052 | 0.2588 | 0.1391 |
| Raumabbildung | 0.4892 | 0.2910 | 0.2661 | 1.0000 | 0.5099 | 0.3826 | 0.0632 |
| hallig/trocken | 0.5055 | 0.0415 | 0.0052 | 0.5099 | 1.0000 | 0.1328 | 0.0108 |
| Surr un-/auff. | 0.2367 | 0.2213 | 0.2588 | 0.3826 | 0.1328 | 1.0000 | 0.5092 |
| surr leise/lt. | -0.0427 | 0.0217 | 0.1391 | 0.0632 | 0.0108 | 0.5092 | 1.0000 |

Tab. 8.8: Auflistung der Korrelationen zwischen 15 Aspekten (ORCH 5.1)

8.6 Analytische Betrachtung der Signale der Rear-Kanäle LS, RS

Da mittels der vorangegangenen Analysen (Varianzanalyse, Faktorenanalyse, Clusteranalyse, Korrelationsanalyse) und im Vergleich mit Ergebnissen einschlägiger Arbeiten (siehe den nachfolgenden Abschnitt 8.7) die Richtigkeit der Ergebnisse der vorliegenden Arbeit überprüft werden konnte, möchte der Autor nun eine genauere Betrachtung hinsichtlich der Beurteilung der Signale LS, RS vornehmen.

Wie bereits im Abschnitt ‚Clusteranalyse‘ kurz angesprochen, interessiert in Hinblick auf den Gegenstand der Dissertation im besonderen die Frage in welcher Weise die Korrelation der Kanäle L, C, R, LS, RS untereinander zum Raumeindruck beiträgt; in diesem Zusammenhang ist natürlich auch der Signalinhalt und die Lautstärke der beiden hinteren Kanäle LS, RS von großer Bedeutung.

Aus diesem Grund wurde auch darauf verzichtet, eine generelle Frage nach dem vom Hörer empfundenen Grad der ‚Einhüllung‘ zu stellen, da die Stärke des Gefühls der ‚Einhüllung‘ – nach Meinung des Autors - an sich noch kein Qualitätsurteil darstellt (siehe hierzu auch Abschnitt 7.4, S.131). Relevant ist hingegen die Frage, wie befriedigend/überzeugend der beim Hörer entstehenden Raumeindruck (Aspekt ‚Raumabbildung‘) ist und – zugleich – die Frage, ob die Surroundkanäle dabei als getrennte Signalquelle wahrgenommen werden, oder sich unauffällig in das Gesamtsignal einfügen. (Letzteres ist wohl zur Erzielung eines ‚gelungenen‘ Raumeindrucks anzustreben.)

Wenn die hinteren Kanäle dem Testhörer auffallen, so kann dies zwei Gründe haben:

- zu hoher Pegel der Kanäle LS, RS
- ‚unpassender‘ Signalinhalt der Kanäle LS, RS

Bei der Erstellung der Surround-Abmischungen für die vier zu bewertenden Mikrofon-Arrays wurde vom Autor darauf geachtet, die Pegelverhältnisse zwischen Front- und Rearkanälen entweder so zu wählen, wie es bei den jeweiligen Techniken vorgeschrieben ist, oder – im Falle des Fehlens einer verbindlichen Angabe – den Pegel der Signale der hinteren Kanäle so einzustellen, dass deren Existenz sich erst (indirekt) bei Wegschalten derselben bemerkbar macht.

Dass dies insgesamt offenbar gut gelungen ist, lässt sich daran erkennen, dass die vier Mittelwerte (der Mikrofon-Arrays) betreffend den Aspekt ‚Surr. un-/auffällig‘ alle deutlich über dem Skalenmittelwert von 4 liegen, d.h. die Surround-Abmischung bei allen System von der überwiegenden Mehrzahl der Testhörer als unauffällig eingestuft wurde.

Wie mittels Lilliefors-Test überprüft, weisen die Hörerantworten betreffend den Aspekt ‚Surr. un-/auffällig‘ bei drei der vier Mikrofontechniken Normalverteilung auf, außerdem waren die Signale der Kanäle LS, RS bei allen Mikrofontechniken als unauffällig eingestuft worden. Ebenso war auch die Lautstärke dieser Kanäle als tendenziell eher leise (und somit unauffällig) eingestuft worden (allerdings lagen hierbei die zugehörigen Mittelwerte nur leicht oberhalb des Skalenmittelwerts)

Nun ist es interessant, sich jene Fälle anzusehen, bei denen die Surround-Kanäle als auffällig eingestuft wurden, und davon insbesondere jene, bei denen diese ‚Auffälligkeit‘ nicht mit einem als zu hoch empfundenen Signalpegel einhergeht, da bei diesen dann offenbar der Signalinhalt alleine dafür verantwortlich ist, daß das Signal vom Testhörer als auffällig eingestuft wurde.

In der untenstehenden Abbildung 8.14 entspricht der Skalenmittelwert 4 einer neutralen ‚Mittelposition‘ in der Bewertung. Wie sich erkennen lässt, ist der Großteil der Bewertungen im Quadranten links oben angesiedelt, d.h. im Bereich ‚Signal unauffällig und leise‘.

Der rechte untere Quadrant enthält Bewertungen, bei denen das Signal als ‚auffällig‘ und ‚laut‘ eingestuft wurde. Die Frage also, ob diese Signale – von einigen offenbar kritischeren / ‚hellhörigeren‘ Testpersonen - als auffällig oder nicht eingestuft werden, ist vermutlich eben auch vom Signalpegel abhängig.

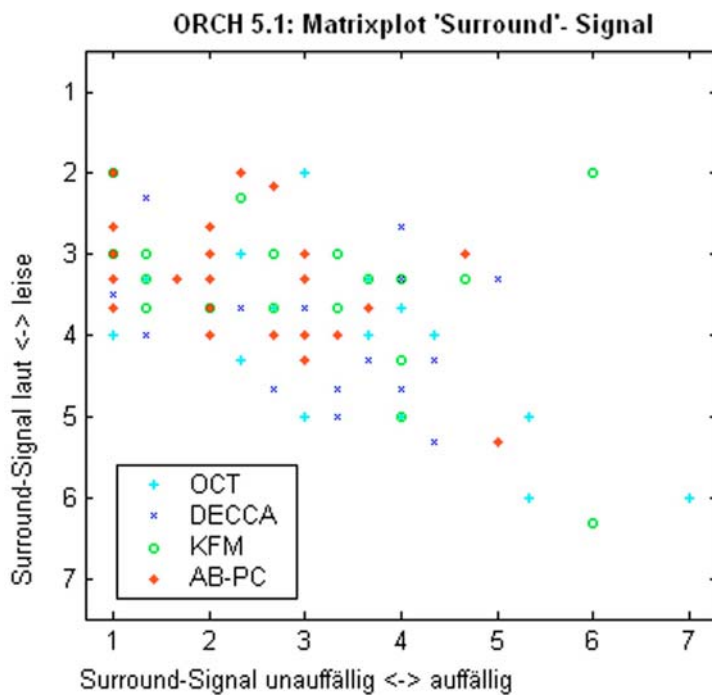


Abb. 8.14: Darstellung des Zusammenhangs in der Hörerbeurteilung von Signalinhalt und Signallautstärke betreffend die hinteren Kanäle LS, RS (25 Testhörer)

Im rechten oberen Quadranten befindet sich die Gruppe der Bewertungen, bei denen das Signal der Kanäle LS, RS zwar als leise, aber dennoch auffällig eingestuft wurde. (An dieser Stelle soll nochmals daran erinnert werden, dass in der Vorbesprechung mit den Testhörern festgelegt worden war, dass eine Auffälligkeit der Signale LS, RS unerwünscht und somit negativ zu bewerten ist.)

Betrachten wir nun den Bereich der Bewertungen, bei denen die x-Koordinaten > 4 und die y-Koordinate ≤ 4.00 ist, so bleiben insgesamt 5 Bewertungen übrig, die in die Gruppe ‚leise und auffällig‘ fallen: davon entfallen zwei Bewertungen auf das KFM-Mikrofonarray und je eine Bewertung auf die Techniken OCT, DECCA und AB-PC (wobei die negative Bewertung bzgl. der OCT Technik genau an der Grenze zwischen leise und laut angesiedelt ist).

Betrachtet man alle Bewertungen, die – unabhängig von ihrer Lautstärke – als auffällig empfunden wurden (d.h. beide Quadranten rechts, oben und unten; $x > 4$), so erhält man:

- OCT: 4 Bewertungen
- DECCA: 3 Bewertungen
- KFM: 3 Bewertungen
- AB-PC: 2 Bewertungen

In der Gruppe, die als stark auffällig bewertet wurde (bei denen $x \geq 6$) liegen 3 Bewertungen vor: zweimal KFM, einmal das OCT System.

Aus den soeben angestellten Betrachtungen drängt sich der Schluss auf, dass es vor allem das KFM-System und – an zweiter Stelle – das OCT-System ist, bei dem die Signale der Kanäle LS, RS so beschaffen sind, dass sie für einen Teil der Testhörer als auffällig (und damit unerwünscht, d.h. negativ) hervortreten.

In diesem Zusammenhang soll auch – quasi als ‚argumentum ad hominem‘ - auf die im Anhang angeführten Emails von J. Korte und C. van der Gragt bzgl. OCT-System hingewiesen werden, die sich beide auf die (negativen) klanglichen Konsequenzen der hohen Signalkorrelation bei tiefen Frequenzen beziehen.

Aufgrund der technischen Beschaffenheit der beiden Systeme (geringe, oder – im Fall des KFM - keine physikalische Trennung der für Gewinnung der vorderen und hinteren Signale verwendeten Mikrofonkapseln) darf dies wohl auf die hohe Korrelation der Signale – insbesondere bei tiefen Frequenzen – zurückgeführt werden, was sich (z.B. in Übereinstimmung mit der Arbeit von [Hidaka, Beranek und Okano, 1995]) negativ auf die empfundene Räumlichkeit in der Signalwiedergabe auswirkt.

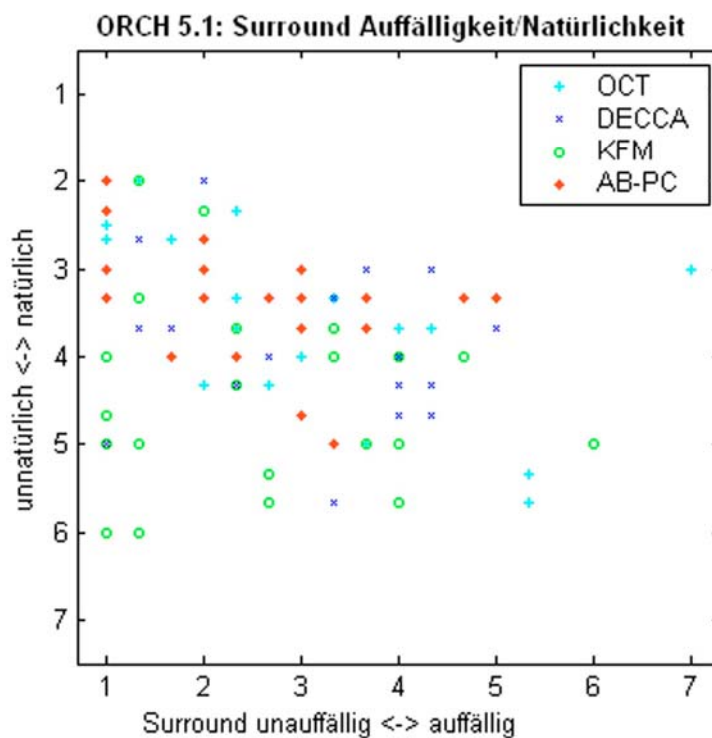


Abb. 8.15: Zusammenhang zwischen der empfundenen Natürlichkeit und der Auffälligkeit der Kanäle LS, RS

Dass diese hohe Signalkorrelation bei den genannten Mikrofonssystemen tatsächlich vorliegt, wird noch im zweiten Teil der Arbeit messtechnisch verifiziert.

Als weitere Bestätigung für die auf der vorhergehenden Seite geäußerte Vermutung lässt sich die obenstehenden Abb. 8.15 heranziehen: aus dieser ist ersichtlich, dass von allen vier Mikrofontechniken die KFM-Technik am öftesten als ‚unnatürlich‘ klingend eingestuft wurde.

Im den beiden rechten Quadranten (Surr. auffällig) der Abb. 8.15 sind folgende Mikrofontechniken vertreten:

(Anm.: Abweichungen von der in Abb. 8.14 auftretenden Anzahl hinsichtlich der Einstufung ‚Surround auffällig‘ erklären sich daraus, dass Bewertungen zu einer Mikrofontechnik die zwar durch unterschiedliche Hörer erfolgt, aber am gleichen Koordinaten-Punkt landen, optisch nicht getrennt ausgewiesen werden. Genauere Information liefert hier nur eine Betrachtung in verschiedenen Ebenen, bzw. die Histogramm-Darstellung).

- OCT: 4
- DECCA: 4
- KFM: 2
- AB-PC: 2

Abschließend soll noch der Zusammenhang zwischen der Hörer-Präferenz und der Auffälligkeit der Surround-Kanäle betrachtet werden (siehe Abb. 8.16).

Hierbei zeigt sich die Ablehnung der KFM-Mikrofontechnik noch deutlicher: die Bewertung des KFM-Systems fällt fast ausnahmslos in die unteren beiden Quadranten, d.h. den Bereich geringer Hörerpräferenz (bzw. – bei Annahme, dass der Skalenmittelwert von 4 eine ‚Neutralmarke‘ darstellt – sind die unteren beiden Quadranten als Bereich der ‚Ablehnung‘ durch den Hörer anzusehen).

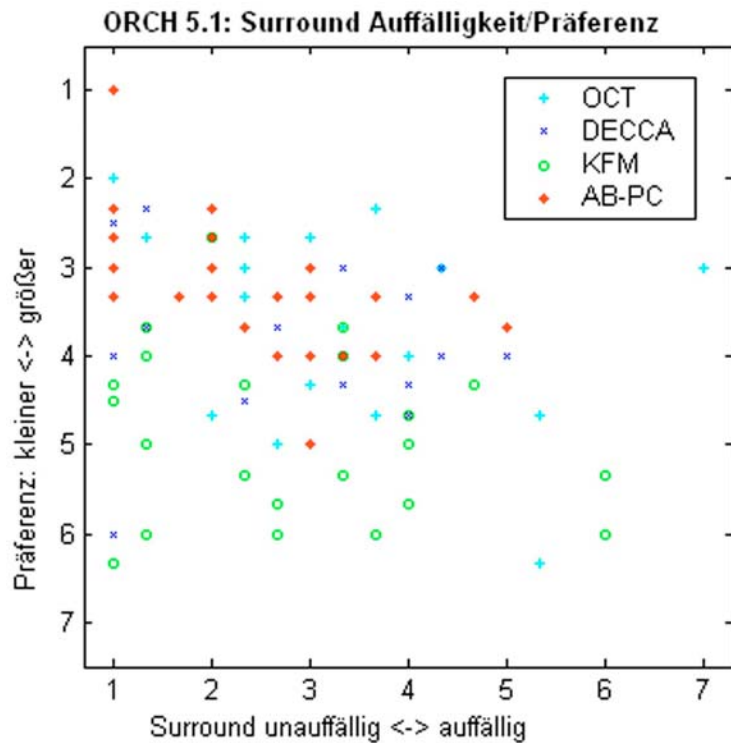


Abb. 8.16:
Zusammenhang zwischen Hörerpräferenz und der Auffälligkeit der Kanäle LS, RS

In diesem Bereich ($y > 4$) gibt es folgende Anzahl von Bewertungen:

- OCT: 6
- DECCA: 5
- KFM: 17
- AB-PC: 1

Im Quadranten rechts unten (Präferenz gering, Surround auffällig) mit $x > 4$, $y > 4$ treten folgende Bewertungen auf:

- OCT: 2
- KFM: 2
- DECCA und AB-PC: keine

Auch diese letzte Betrachtung des Quadranten rechts unten geht konform mit dem bisherigem Bild der Bewertung der vier Mikrofontechniken hinsichtlich Präferenz und Natürlichkeit unter besonderer Berücksichtigung der Auffälligkeit bzgl. der Rear-Kanäle LS, RS.

Schlußfolgerung

Gemäß der obenstehenden Auswertung ergibt sich hinsichtlich der Präferenz folgende Reihenfolge für die untersuchten Mikrofonssysteme:

1. AB-PC
2. DECCA
3. OCT
4. KFM

Der Autor vermutet – wie bereits oben angesprochen – eine Verbindung zwischen der Signalkorrelation (insbesondere im tieffrequenten Bereich) und der Hörerpräferenz, welche primär über den Einfluss der Raumabbildung (und ev. der Basisbreite) auf die Hörerpräferenz bedingt ist. Dies wird im zweiten Teil der Arbeit anhand der messtechnischen Ergebnisse (frequenzabhängige Korrelation der Signale) untersucht.

8.7 Zusammenfassung und Vergleich mit den Ergebnissen anderer Untersuchungen

8.7.1 Varianzanalyse

Diese hatte zum Resultat, dass für den vorliegenden Datensatz mit 25 Testhörern für 13 von 15 Aspekten (inkl. Präferenz) Signifikanz-Niveau erreicht wurde, indem zumindest eine der vier zu bewertenden Mikrofon-Techniken in der Beurteilung – aufgrund der errechneten Konfidenzintervalle - als statistisch signifikant unterscheidbar eingestuft wurde.

8.7.2 Faktorenanalyse

Gemäß den Ausführungen in Abschnitt 8.3.2 möchte der Autor die drei Hauptfaktoren des Hörvergleichs ‚ORCH 5.1‘ interpretieren, wie folgt:

- 1. Faktor: Empfindung der ‚*Größe und Ausdehnung*‘ der Klangquelle
- 2. Faktor: Empfindung der ‚*Deutlichkeit*‘
- 3. Faktor: Empfindung der musikalisch/raumakustischen ‚*Einhüllung*‘

Diese ersten 3 Faktoren erklären gemeinsam etwa 63% der Gesamtvarianz. Die gute Übereinstimmung – insbesondere mit der Arbeit von [Wilkins, 1975] zur subjektiven Beurteilung der Akustik von Konzertsälen, wie auch mit anderen facheinschlägigen Arbeiten - ist ein wichtiges Indiz für die Plausibilität der für die Faktoren ermittelten psychoakustischen Größen.

Vergleich mit anderen Arbeiten

Bei [Cremer, 1978] werden die Ergebnisse einiger einschlägiger Arbeiten betrachtet, welche sich in erster Linie mit der qualitativen Beurteilung von Konzertsälen durch Testhörer und anschließender faktorenanalytischer Auswertung, sowie der Zuordnung der subjektiven (psychoakustischen) Faktoren zu objektiven (raumakustischen) Kriterien befassen. (siehe dazu die Arbeiten auf den folgenden Seiten)

Die Ergebnisse dieser Arbeiten scheinen insofern für die vorliegende Untersuchung relevant, da auch bei der Beurteilung von (über eine 5.1 Surround-Lautsprecheranordnung) wiedergegebener Musik für jeden Testhörer wohl – im Falle des Fehlens eines angebotenen Vergleichs-Stimulus - nur die (Erinnerung an eine)

Konzertsituation (bewusst/unbewusst) als Referenz herangezogen werden kann, und somit wieder die gleichen (oder zumindest ähnliche) Aspekte für die Qualitätsbeurteilung ins Spiel kommen.

Es handelt sich hierbei um einen dreifachen Transformations- oder Projektionsvorgang der eigentlichen Schallquelle: der vom Orchester ausgehende Klangreiz trifft - mit der Akustik des Raumes ‚gefaltet‘ - am Hörerplatz (bzw. an der jeweiligen Mikrofonposition) ein, wird durch die anschließende Tonaufnahme in 5.1 Surround einer Transformation im Sinne einer Dateneinschränkung durch Festlegung auf 6 Übertragungskanäle mit definierter Bandbreite und Abtastrate unterworfen [Stichwort: ‚Kanalkapazität‘], anschließend im Abhör-Raum durch die Lautsprecherwiedergabe erneut einer Signalveränderung (Transformation) unterzogen, um letztendlich – mit der Akustik des Hörraums ‚gefaltet‘ - zum Ohr des Testhörers zu gelangen. In diesen mehrfachen Transformations-Vorgang fließen natürlich auch alle Signalverfälschungen ein, die durch lineare oder nichtlineare Verzerrungen der einzelnen Glieder der Übertragungskette zustande kommen. Insofern ist es bereits als Erfolg zu werten, wenn sich die ursprünglich qualitätsbestimmenden Attribute im Konzertsaal auch als ebenso relevant für die Tonwiedergabe im Hörraum herausstellen, wie das anhand der obenstehenden Ergebnisse gezeigt werden konnte.

Denn letztlich geht es – sowohl beim Live-Erlebnis im Konzertsaal, wie auch bei der Wiedergabe desselben im Wohnzimmer bzw. Studio-Abhörzimmer – darum, mit welcher Qualität bestimmte - für das ‚Qualitätsempfinden‘ ausschlaggebende Faktoren - vom Hörer wahrgenommen werden.

Im ersteren Fall wird anhand dieser die Qualität des Konzertsaals (ev. an verschiedenen Sitzplätzen, oder auch verschiedener Konzertsäle untereinander) beurteilt, im letzteren Fall die Qualität unterschiedlicher Mikrofonierungstechniken dahingehend, wie sehr sie die Qualität der einstmals im Original-Signal enthaltenen Klangeindrucks verschlechtern bzw. (signaltheoretisch) ‚verzerren‘.

Bei [Hawkes und Douglas, 1971] ließ sich die Zahl der unterscheidbaren Faktoren auf 6 reduzieren; für die Auswertung eines Datensatzes, der bei einem Konzert in der Royal Festival Hall gewonnen worden war, sogar auf 4, welche sie mit den Ausdrücken:

- Definition
 - Reverberance
 - Balance and Blend, sowie
 - Proximity
- beschrieben.

[Siebrasse, 1973] kam bei seinen Untersuchungen mit 13 Versuchspersonen und 25 Konzertsälen zu einer Anzahl von 4 Faktoren, die berücksichtigt werden müssten und mit denen sich 91% der Varianz erklären ließ.

[Wilkins, 1975] kommt für seine Arbeit zum Schluss, dass sich die Zahl der Faktoren auf 3 beschränken ließe, welche er der Empfindung der

- Stärke und Ausdehnung der Quelle,
- Deutlichkeit, und
- Klangfarbe

zuordnet, wie bereits im vorhergehenden Abschnitt 8.3.2 angeführt.

Im Rahmen weiterführender Untersuchungen gelang es [Lehmann, 1976], diesen subjektiven Empfindungen folgende objektive raumakustische Kenngrößen zuzuordnen:

- Stärkemass G (nach [Lehmann, 1976],
siehe hierzu auch das Kapitel 2.4, Formel 2.5)
- Schwerpunktszeit t_s (zur näheren Definition siehe [Cremer, 1978] S. 348), sowie
- Early Decay Time T_E

(Anm.: Die Zuordnung zu diesen raumakustischen Größen war möglich, da im Rahmen der Untersuchungen diverse raumakustische Parameter der Konzertsäle gemessen worden waren.)

Im Rahmen einer Untersuchung über die empfundene Qualität von Lautsprechern [Gabrielsson, 1985] wurden die nachfolgenden Aspekte von 18 Testpersonen bewertet:

Clarity (Clearness/distinctness)

Sharpness/hardness versus softness

Brightness vs. dullness

Fullness vs. thinness

Spaciousness (feeling of space)

Nearness

Absence of extraneous sounds

Loudness

Ein Ergebnis der Studie bestand darin, dass der Aspekt ‚Natürlichkeit‘ am stärksten mit den Aspekten ‚Clarity‘ (Klarheit, Deutlichkeit), ‚Spaciousness‘ (Räumlichkeit) und ‚Fullness‘ ([Klang-]Fülle, Ausdehnung) korreliert ist.

Eine weitere Untersuchung bzgl. der subjektiv empfundenen Qualität von Lautsprechern [Toole, 1985] verwendete folgende Aspekte zur Beurteilung von insgesamt 37 Lautsprechern durch 42 Testhörer:

Zur Bewertung der klanglichen Abbildung (sound quality) waren zu bewerten:

- Clarity / definition
- Softness
- Fullness
- Brightness
- Spaciousness / openness
- Nearness /presence
- Hiss / Noise distortions
- Loudness
- Spectral balance / frequency response
- Colouration / distortion

sowie:

- Pleasantness
- Fidelity

Im Rahmen der Bewertung der räumlichen Abbildung (spatial quality) war zu beurteilen:

- Definition of sound images
- Continuity of the sound stage
- Width of the sound stage
- Impression of distance / depth
- Abnormal effects
- Reproduction of ambience, spaciousness and reverberation
- Perspective
- Overall spatial rating

Ein Ergebnis der Studie bestand darin, dass der Aspekt ‚Natürlichkeit‘ (Fidelity) am stärksten mit der ‚(Overall) Spatial Quality‘ korreliert ist ($r=0.7$) (Anm.: Allerdings wurde keine Aufschlüsselung bezüglich der einzelnen – oben aufgeführten - Aspekte der ‚Spatial quality‘ gegeben.)

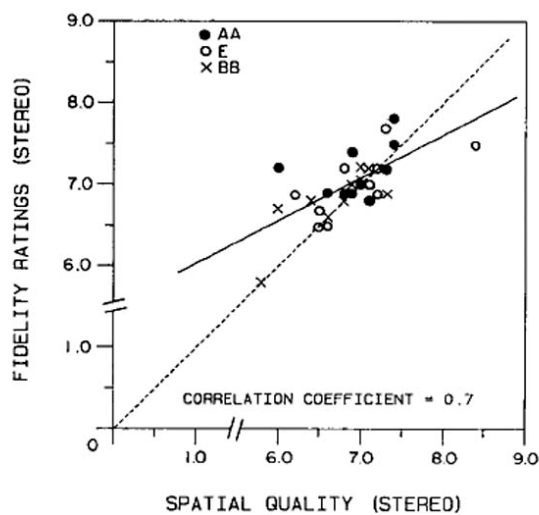


Fig. 21. Relationship between sound quality (fidelity) ratings and spatial quality ratings in stereophonic listening for the three loudspeakers in stereo/mono series II.

Abb. 8.17: Korrelation zwischen Natürlichkeits- und Räumlichkeits-Bewertung (nach [Toole, 1985])

Eine subjektive Bewertung der Klang-Qualität bei Mehrkanal-Lautsprecherwiedergabe [Nakayama et al., 1971] verwendete eine Anzahl von bis zu 8 Surround-Lautsprechern (mono bis 8-Kanal), deren Signalwiedergabe (Pop- und Folk-Musik mit akustischen Instrumenten) von 10 Hörern (Studenten) bewertet wurde:

Die Testhörer mussten die angebotenen Stimuli hinsichtlich Ähnlichkeit (Similarity-

Rating) und Präferenz (Preference-Rating) beurteilen.

Eine frühere Studie von [Yoshida et al., 1960] hatte gezeigt, dass die Klangaspekte bei Zwei-Kanal Stereo durch 2 Faktoren ausgedrückt werden können:

- Faktor 1: ‚Diffusion of the image sources‘
- Faktor 2: ‚Improved naturality of the perceived reverberation‘

Ein Ergebnis der Studie von [Nakayama et al., 1971] bestand darin, dass sich die in Form des damaligen Datensatzes vorliegende Hörer-Präferenz Q mittels multipler Regression aus der Linearkombination folgender Faktoren herleiten ließ:

$$Q = 0.253 D_1 + 0.365 D_2 + 0.161 D_3 - 0.066 \quad (8.6)$$

Die drei Faktoren konnten den subjektiven Aspekten D_1 : ‚Depth of the Image Sources‘, D_2 : ‚Fullness‘ und D_3 : ‚Clearness‘ zugeordnet werden, welche jeweils 45.1%, 16.7% und 15.3% der Gesamtvarianz erklärten. (Diese ersten 3 Faktoren erklärten somit eine Gesamtvarianz von 77.1 %)

Mit der Beschreibung des ersten Faktors als ‚Depth of the Image Sources‘ dürfte eigentlich eher ‚Nearness‘ oder ‚Closeness‘ gemeint gewesen sein, wie in [Rumsey, 1998] überzeugend dargestellt wird. Die beiden weiteren Faktoren konnten folgenden objektiven Kriterien zugeordnet werden:

- ‚Fullness‘ konnte mit dem ‚Öffnungswinkel‘ (Position) der seitlichen bzw. hinteren Lautsprecher, sowie der ‚subjektive diffuseness‘ in Verbindung gebracht werden, welche ihrerseits mit dem Inter-Aural Cross-correlation Coefficient IACC in Verbindung steht (siehe hierzu [Anazawa et al., 1967]); dies lässt nach Meinung des Autors die Interpretation zu, dass sich hinter dem damaligen Begriff ‚Fullness‘ – nach heutigem Sprachgebrauch – das sog. ‚Envelopment‘ (nach Griesinger) verbergen könnte. (siehe auch hierzu [Rumsey, 1998])
- ‚Clearness‘: nach messtechnischer Ermittlung der Deutlichkeit D im Konzertsaal, wo die Aufnahmen der Hörbeispiele stattgefunden hatten, kamen die Autoren zum Schluss, dass die ‚Clearness‘ der raumakustischen Größe Deutlichkeit (nach [Thiele, 1953]) zuzuordnen sei.

Eine aktuelle Studie [Berg und Rumsey, 2001] beschäftigte sich mit der Verifizierung und Korrelation von Attributen, die zur Beschreibung der Räumlichkeits-Aspekte bei der Wiedergabe von Musik ausgewählt wurden. An den Hörtests, die mit einer 5.0 Lautsprecheranordnung durchgeführt wurden, nahmen insgesamt 19 Testpersonen teil. Die gewonnenen Daten wurden unter anderem auch einer Hauptkomponentenanalyse (PCA), Faktorenanalyse, Clusteranalyse und Korrelationsanalyse unterzogen.

Die dabei untersuchten 12 Aspekte wurden von den Autoren folgenden Aspekt-Gruppen zugeordnet:

| Aspekt | Abk. | Gruppe |
|-------------------------|-------------|---------------|
| Naturalness | nat | General |
| Presence | psc | General |
| Preference | prf | General |
| Envelopment | env | General |
| Source width | swd | Source |
| Localisation | loc | Source |
| Source distance | dis | Source |
| Room width | rwd | Room |
| Room size | rsz | Room |
| Room spectral bandwidth | rsp | Room |
| Room sound level | rlv | Room |
| Background noise level | bgr | Other |

Tab. 8.9: Aspekte und Gruppen
aus der Studie von [Berg und Rumsey, 2001]

Die im Rahmen der PCA gefundenen 3 Komponenten (Hauptfaktoren) wurden von den Autoren interpretiert als:

- Width (in a general meaning) or ‚bigger event‘
(entspricht: Basisbreite oder ‚Größe‘/Einhüllung ?)
- distance to the sound (Entfernungseindruck nahe / entfernt bzw. Tiefenstaffelung)
- ‚source width versus localisation‘ (Gegensatz Basisbreite – Lokalisation)

Die Faktorenanalyse hatte zum Ergebnis, dass es eine starke Verbindung zwischen dem ersten gefundenen Faktor und den Attributen der ‚General Class‘ gab, während der 2. Faktor mit den ‚Room‘ Attributen in Verbindung stand. Der 3. Faktor wurde in der Interpretation den ‚Source Attributes‘ zugeordnet und könnte ev. als mit ‚Image Focus‘ in Verbindung stehend aufgefasst werden.

Die Ergebnisse der ebenfalls durchgeführten Clusteranalyse sollen erst im nächsten Abschnitt besprochen werden.

In einer Studie von [Gabrielsson und Sjören, 1979] fanden die Autoren, dass einer der Faktoren wie folgt beschaffen war: „...a general quality factor emphasizing clearness/distinctness, feeling of space and nearness in the reproduction ...“.

In [Bech et al., 1995] wurde eine Bewertung der ‚overall reproduction quality‘ durchgeführt und es konnte gezeigt werden, dass hierbei eine starke Korrelation zur Qualität der räumlichen Reproduktion besteht.

8.7.3 Clusteranalyse

Die Clusteranalyse nach ‚Complete Linkage‘ Methode zeigte, dass die Aspekte in eine ‚Räumlichkeits-Gruppe‘ (Raumabbildung, Basisbreite, Höhen/Tiefen, nahe/entfernt, hallig/trocken, Surr. leise/laut), die mit der Präferenz verbunden ist, sowie eine ‚Klangquellen-Gruppe‘ (Lokalisation, Klangfarbe, Tiefenstaffelung, Differenziertheit, Balance, Stabilität, Surr. un-/auffällig) geteilt werden können, welche mit der Natürlichkeit in Verbindung steht.

Von den mittels Faktorenanalyse gefundenen Hauptfaktoren fallen die ersten zwei in die Gruppe der ‚Klangquellen‘-spezifischen Aspekte: ‚Größe und Ausdehnung‘

(entsprechend ‚Balance‘), sowie ‚Deutlichkeit‘ (wenn man den Aspekt ‚Differenziertheit‘ als eine Spielart von ‚Deutlichkeit‘ gelten lässt).

Um die Ergebnisse der Clusteranalyse des Datensatzes aus dem Hörvergleich ORCH 5.1 mit den Ergebnissen einer ähnlich gelagerter Studie vergleichen zu können, ist untenstehend die Clusteranalyse aus der Arbeit von [Berg und Rumsey, 2001] abgebildet:

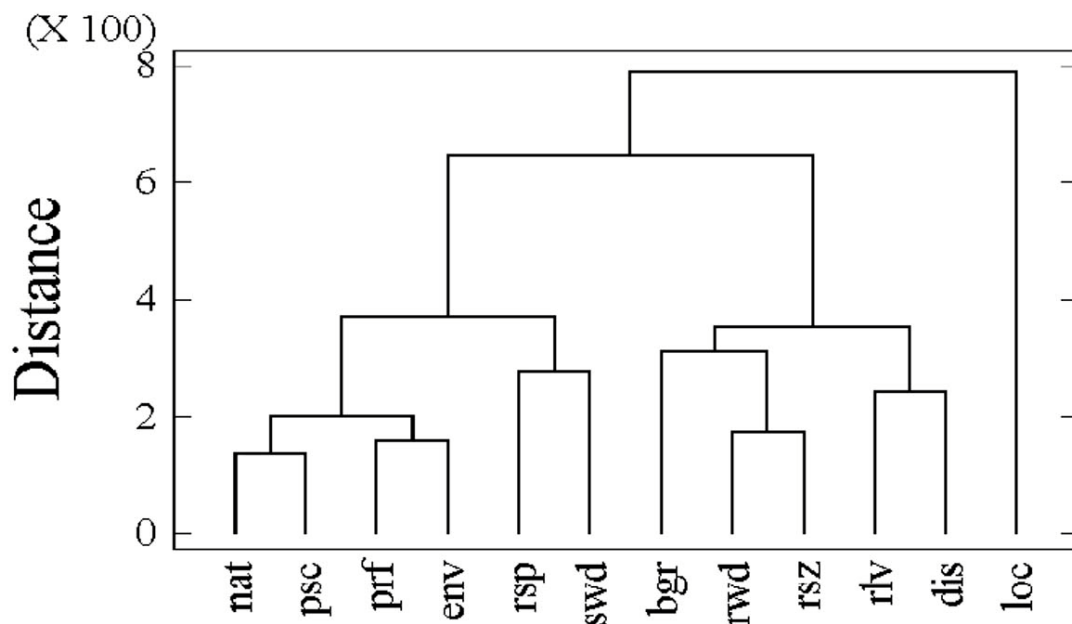


Abb. 8.18: Dendrogramm von 12 Aspekten (aus [Berg und Rumsey, 2001], S.11) (Berechnung nach ‚Complete Linkage‘ Methode, squared Euklidian)

| Aspekt | Abk. | Gruppe | Aspekt | Abk. | Gruppe |
|--------------|------|---------|-------------------------|------|--------|
| Naturalness | nat | General | Source distance | dis | Source |
| Presence | psc | General | Room width | rwd | Room |
| Preference | prf | General | Room size | rsz | Room |
| Envelopment | env | General | Room spectral bandwidth | rsp | Room |
| Source width | swd | Source | Room sound level | rlv | Room |
| Localisation | loc | Source | Background noise level | bgr | Other |

Tab. 8.10: Erklärung der Begriffe aus der Untersuchung von [Berg und Rumsey, 2001], inkl. Gruppen-Zuordnung

Vergleich der Clusteranalysen Pfanzagl und Berg/Rumsey

Beiden Analysen gemeinsam ist, dass die Aspekte Präferenz und Natürlichkeit sehr eng miteinander verwandt sind.

Bei genauerer Betrachtung zeigt sich, daß der Aspekt ‚Envelopment‘ (Einhüllung) aus der Studie von Berg/Rumsey durchaus mit dem Aspekt ‚Raumabbildung‘ (Studie Pfanzagl) verglichen werden darf:

Pfanzagl ließ die Testhörer bewusst keine Bewertung der ‚Einhüllung‘ vornehmen, da ‚gute‘ (bzw. ein hoher Grad der empfundenen) Einhüllung seines Erachtens für sich genommen noch kein Qualitätskriterium darstellt, vielmehr geht es (zumindest im Zusammenhang mit der vorliegenden Untersuchung) darum, ob diese ‚Einhüllung‘ in der Lage ist, einen natürlichen Raumeindruck beim Hörer hervorzurufen. Genau dieses wurde auch im Rahmen der Bewertung des Aspekts ‚Raumabbildung‘ erfragt. (Siehe zu dieser Problematik auch die näheren Erläuterungen gegen Ende des Abschnitts 7.4)

Da ‚Envelopment‘ in der Surround-Situation (im Gegensatz zur 2-Kanal Stereosituation, bei der sich allerdings – nach Griesinger – ebenfalls bis zu einem gewissen Grad des Gefühl der Einhüllung ergeben kann; siehe hierzu Abschnitt 2.3, Abb. 2.12) in erster Linie durch die Signale der hinteren Kanäle LS, RS erzielt werden kann oder soll, ist deren Signalinhalt und Pegel von eminenter Bedeutung.

Diese beiden letzteren Komponenten waren durch die Testhörer in Form der Aspekte ‚Surr. un-/auffällig‘ und ‚Surr. leise/laut‘ zu beurteilen. In Abschnitt 8.6 wurden die Ergebnisse aus den Analysen betreffend diese Teilaspekte näher betrachtet. Das über die Raumabbildung abgegebene Qualitätsurteil kann somit – nach Meinung des Autors – mit der Angabe über ‚Envelopment‘ in der Studie von Berg/Rumsey verglichen werden.

An vierter Stelle in punkto Ähnlichkeit steht bei der Studie Pfanzagl der Aspekt Basisbreite, bei Berg/Rumsey der Aspekt ‚Presence‘.

Die weiteren Ergebnisse der beiden Studien sind insofern vergleichbar, als bei Pfanzagl als nächster Aspekt (siehe Clusteranalyse mit ‚Complete Linkage‘ Methode) Höhen/Tiefen auftritt, während es bei Berg/Rumsey die Aspekte ‚Room spectral

balance' und ‚Source width' (also: Basisbreite) sind; d.h. daß vom 3. bis zum 6. Aspekt die Begriffe in beiden Studien einander durchaus ähneln und lediglich eine etwas andere Abfolge vorliegt.

Die weiteren bei Berg/Rumsey auftretenden Aspekte fallen bis auf zwei Ausnahmen (Source Distance, Background Noise Level) alle in die – von den Autoren definierte – Gruppe der ‚Room' Aspekte.

Ein großer Unterschied zwischen den Studien besteht darin, dass der Aspekt Lokalisation bei Berg/Rumsey isoliert als eigene Gruppe hervorgeht, während er bei Pfanzagl bestens integriert und in großer Nähe von Präferenz und Natürlichkeit angesiedelt ist. Diese Isoliertheit hatte sich auch bereits bei der von den Autoren durchgeführten Faktoren- und Komponenten-Analyse, in Form einer (allerdings mäßig schwachen) negativen Korrelation mit allen 3 (Haupt-) Faktoren gezeigt. Eine Erklärung für die Isolation des Aspekts Lokalisation bei Berg/Rumsey könnte ev. auch darin liegen, dass er im Rahmen der Varianzanalyse den zweitschlechtesten F-Wert (nach ‚Room Spectral Balance') erhalten hat, d.h. daß – in Relation zu den anderen Aspekten des vorliegenden Datensatzes - eine relativ große Varianz in der Bewertung (verursacht durch Hörerabhängigkeit) vorliegt.

Aus einem ähnlichen Grund ist in der Studie Pfanzagl der Aspekt ‚Surr. un-/auffällig' in der Cluster-Analyse ebenfalls an letzter Position gelandet, wenngleich er allerdings keine eigene Gruppe bildet.

Insgesamt lässt sich feststellen, dass beim Vergleich der Clusteranalysen Pfanzagl und Berg/Rumsey sehr große Übereinstimmungen der Ähnlichkeitsgrade der verschiedenen Aspekte untereinander bestehen:

In beiden Studien gehen die Aspekte Präferenz, Natürlichkeit, Raumabbildung und Basisbreite, sowie Klangfarbe, Höhen/Tiefen (siehe ‚Presence' bei Berg/Rumsey) als die wichtigsten Aspekte mit den höchsten Ähnlichkeitsgraden hervor. Diese gute Übereinstimmung kann nach Meinung des Autors als Indiz für die Richtigkeit der Ergebnisse der Clusteranalyse der Studie Pfanzagl gewertet werden.

8.7.4 Korrelationsanalyse

Auch bei der Korrelationsanalyse des vorliegenden Datensatzes ORCH 5.1 konnte gute Übereinstimmung mit den Ergebnissen einer vergleichbaren Arbeit (vgl. [Berg und Rumsey, 2001]) sowie anderen Arbeiten zur Problematik ‚Perceived Sound Quality‘ (vgl. [Gabrielsson, 1985]) gefunden werden.

Die Ergebnisse der Korrelationsanalyse decken sich auch gut mit der im vorhergehenden Abschnitt besprochenen Clusteranalyse: beide bestätigen die enge Korrelation zwischen Raumabbildung und Präferenz, (sowie auch zwischen Basisbreite und Präferenz), wie auch zwischen Klangfarbe und Natürlichkeit (sowie generell zwischen Präferenz und Natürlichkeit).

Die in der vorliegenden Arbeit gefundene hohe Korrelation ($r=0.8383$) zwischen Präferenz und Natürlichkeit scheint durchaus erwartungsgemäß, allerdings war in der Studie von [Berg und Rumsey, 2001] für diese beiden Aspekte eine (ebenfalls relativ hohe) Korrelation von lediglich 0.595 festgestellt worden (wobei andere Aspekte höhere Korrelation aufgewiesen hatten).

Gute Übereinstimmung mit den Ergebnissen von [Berg und Rumsey, 2001] zeigt sich hingegen hinsichtlich der Korrelation zwischen Präferenz und Raumabbildung: bei der vorliegenden Studie ergibt sich der obenstehende Wert von 0.6829, bei Berg und Rumsey tritt für die Aspekte Preference – Envelopment ein Wert von $r=0.674$ auf. Im obenstehenden Abschnitt betr. Clusteranalyse wurde bereits dargestellt, welcher Zusammenhang zwischen dem Aspekt ‚Raumabbildung‘ der vorliegenden Studie und dem Aspekt ‚Envelopment‘ der Studie Berg/Rumsey vermutet wird.

So wie bereits bei der Clusteranalyse decken sich auch die Ergebnisse der Korrelationsanalyse gut mit den durch die Faktorenanalyse gefundenen Beziehungen zwischen den Aspekten.

8.7.5 Schlussfolgerung

Aufgrund der obenstehenden Analysen der Hörerbefragungen kann davon ausgegangen werden, dass tatsächlich objektivierbare Unterschiede in der Bewertung der vier Surround-Mikrofontechniken bestehen, welche auf deren unterschiedliche Abbildungseigenschaften zurückzuführen sind.

Die generell gute Übereinstimmung der Ergebnisse der Analysen des Autors mit fach einschlägigen Arbeiten anderer Autoren spricht ebenfalls für die Gültigkeit der Ergebnisse der vorliegenden Untersuchung.

Wie bereits im Rahmen der Varianzanalyse festgestellt, wurde eine Mikrofontechnik (KFM) deutlich schlechter bewertet, und eine der Techniken (AB-PC) schnitt sowohl in der Einzel-Bewertung (bei 8 von insgesamt 10 qualitativen Aspekten), als auch hinsichtlich der Präferenz-Bewertung besser als die anderen ab.

Eine nähere statistische Betrachtung bzgl. der Beurteilung des Signalinhalts der hinteren Kanäle in Abschnitt 8.6 brachte als Ergebnis, dass es vor allem das KFM-System und – an zweiter Stelle – das OCT-System sind, bei dem die Signale der Kanäle LS, RS so beschaffen sind, dass diese für einen Teil der Testhörer als auffällig (und damit unerwünscht, d.h. negativ) hervortreten.

Aufgrund der technischen Beschaffenheit dieser beiden Systeme führt der Autor dies auf die hohe Korrelation der Signale – insbesondere auch bei tiefen Frequenzen – zurück, was sich (z.B. in Übereinstimmung mit der Arbeit von [Hidaka, Beranek und Okano, 1995]) negativ auf die empfundene Räumlichkeit in der Signalwiedergabe auswirkt, was weiter – durch die enge Korrelation zwischen dem Aspekt Raumabbildung und Präferenz bedingt – zu einer qualitativ schlechteren Einstufung in der Hörerbeurteilung führt.

Dass diese hohe Signalkorrelation bei den genannten Mikrofonssystemen tatsächlich vorliegt, wird im zweiten Teil der Arbeit messtechnisch verifiziert.

8.8 Qualitative Reihung der Surround-Mikrofonarrays – ein Vergleich mit anderen Untersuchungen

Nach dem im obenstehenden Abschnitt 8.7.5 erfolgten Abschluss der Analyse der Daten aus dem Hörvergleich ORCH 5.1 sollen die gewonnenen Ergebnisse jenen aus anderen Untersuchungen gegenübergestellt werden:

Von den Fragestellungen her vergleichbar, bietet sich in erster Linie der im Frühjahr 2001 beim ORF durchgeführte Hörvergleich von 7 Surround-Mikrofontechniken an, im Rahmen dessen u.a. auch die Techniken ‚OCT-Surround‘, ‚KFM‘, ‚DECCA‘ und ‚Soundfield‘ untersucht wurden (siehe [Camerer und Sodl, 2001])

Eine komplette graphische Präsentation der Ergebnisse aus [Camerer und Sodl, 2001] findet sich im Anhang.

Vorweg soll gesagt werden, dass es sich bei dem im Hörvergleich des ORF verwendeten DECCA-System um eine für Surround mit Hamasaki-Square erweiterte Technik handelt, die zusätzlich zum DECCA-Dreieck zwei Breit-AB Omni-Mikrofone an der Front verwendet. Aus diesem Grund ist die ORF-DECCA Aufstellung nicht mit der vom Autor verwendeten vergleichbar.

In Bezug auf die Techniken OCT-Surround und KFM zeigen sich große Parallelen in den Ergebnissen der Studie des ORF und der aktuell vorliegenden Untersuchung: so wurde z.B. das KFM-Mikrofon auch beim ORF-Hörvergleich generell wesentlich schlechter bewertet als das OCT-System, welches tendenziell sehr gute Bewertung erhielt.

Auch wenn an dieser Stelle auf einen detaillierten Aspekt-für-Aspekt Vergleich der vorliegenden Studie mit den Ergebnissen des ORF verzichtet werden soll, so kann doch festgestellt werden, dass die Konformität in den Ergebnissen (zusätzlich zur erfolgten Varianzanalyse, um die Signifikanz festzustellen) als ein weiteres Indiz für die Richtigkeit der Ergebnisse der vorliegenden Studie angesehen werden darf.

Die Ergebnisse der aktuellen Untersuchung stimmen weiters auch mit den Ergebnissen der Untersuchung von [Hermann et al., 1998] gut überein, bei der folgende 3/2 Surround-Mikrofontechniken im Hörvergleich beurteilt worden waren:

- ABC (3 Mikrofone mit Kugelcharakteristik), ergänzt mit 2 Raummikrofonen
- INA (,Ideale Nieren Anordnung'), ergänzt mit dem IRT-Kreuz (nach [Theile, 1996]) als Raummikrofon
- SK (Surround-Kugel = Schoeps KFM) in der Anwendung nach [Bruck, 1996], ergänzt mit zwei Mikrofonen mit 8-er Charakteristik zur Gewinnung des Surround-Signals, sowie
- ein auf ,Ambisonic' basiertes Verfahren (ähnlich Soundfield Mikrofon)

Vergleichbar mit den Surround-Mikrofontechniken der vorliegenden Untersuchung sind davon nur das ABC-System (ähnlich der DECCA-Anordnung des Autors, allerdings wurden die hinteren Surround-Signale anders gewonnen), sowie das KFM-basierte System.

Hinsichtlich einer qualitativen Reihung lagen bei der Untersuchung von [Hermann et al., 1998] ABC und INA vorne, das KFM belegte den dritten Platz und das Ambisonic Verfahren lag bei allen Fragestellungen an letzter Stelle.

9 Hörvergleich Orchesteraufnahme-Kunstkopf („ORCH-KK“)

9.1 Programm-Material

Während der Aufnahme des Orchesterstücks im Großen Saal des Salzburger Festspielhauses war – zugleich mit den verschiedenen 5.1 Surround Mikrofon-Arrays – auch eine Referenz-Aufnahme mittels Neumann Kunstkopf KU81i erfolgt, welcher aus Gründen des Verhältnisses von Direkt- zu Diffusschall in der sechsten Reihe des Parketts positioniert worden war.

Im Produktionsstudio des IEM wurden die mittels OCT, DECCA, KFM und AB-PC erstellten 5.1 Surroundaufnahmen über die zu Beginn der Analyse beschriebene Lautsprecheranordnung wiedergegeben und dieses Signal am sog. ‚Sweet-Spot‘ mittels desselben Neumann KU81 Kunstkopfs erneut aufgezeichnet. (siehe Abb. 9.1 unten) ¹

Die neu gewonnenen Kunstkopfsignale der vier Mikrofontechniken wurden gemeinsam mit dem Referenzsignal einer Bewertung in Form eines A/B/REF Hörvergleichs unterzogen.



Abb. 9.1: Re-Recording der vier 5.1 Surround-MicArray Signale im Produktionsstudio des IEM mittels Neumann KU81i-Kunstkopf (Anm.: Das Yamaha DM 2000 Mischpult wurde mit Kleidung behelfsmäßig abgedeckt, um Reflexionen an seiner schallharten Oberfläche – und die dabei ev.

entstehenden Kammfiltereffekte - zu vermeiden)

¹⁾ In Kapitel 10 wird untersucht, welches dieser neu gewonnen Kunstkopfsignale dem Original-Kunstkopfsignal hinsichtlich Korrelation am ähnlichsten ist, um daraus Schlüsse über die Qualität der vier untersuchten Mikrofonarrays ableiten zu können.

9.2 Kopfhörer-Wiedergabe

Für die Wiedergabe über Kopfhörer wurde der Beyerdynamic DT 990 Kopfhörer ausgewählt, da er diffusfeld-entzerrt ist und annähernd über FEC-Eigenschaften („free-air equivalent coupling“) verfügt (siehe hierzu [Moller et al., 1995]).

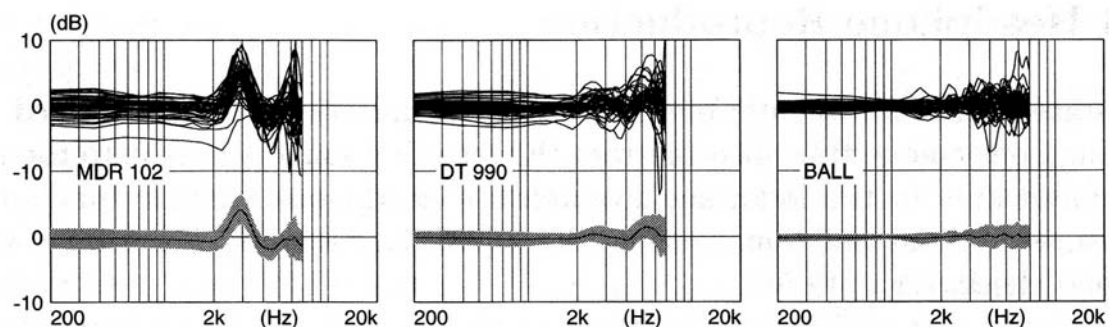


Abb. 9.2: PDR's von 3 verschiedenen Kopfhörern (aus [Moller et al., 1995], S. 212) „ ... Beispiele von Druckverteilungs-Verhältnissen (*pressure-division ratios, PDR's*) für den Sony MDR 102, Beyerdynamic DT990 und für einen freistehenden Ball-Lautsprecher. In der oberen Hälfte der Abbildungen sind individuelle Messkurven von 40 Testpersonen zu sehen; darunter der Mittelwert +/- Standardabweichung. Jede PDR wurde auf der Basis von vier Messungen errechnet: offener und geschlossener Gehörgang einmal gemessen im Freifeld, sowie einmal gemessen mit dem Kopfhörer als Schallquelle. Die Richtigkeit der Resultate für höhere Frequenzen ist zweifelhaft, weshalb in den Graphiken nur der Frequenzbereich bis 7kHz dargestellt ist. ...“ (aus dem Englischen übersetztes Zitat aus [Moller et al., 1995], S. 212)

Zur Erklärung der FEC-Eigenschaften ist bei [Hammershoi and Moller, 2005] zu lesen (in Übersetzung):

“ ... Wenn das Druckverteilungs-Verhältnis bei Kopfhörerwiedergabe dem des Druckverteilungsverhältnisses in der Freifeld-Situation gleicht, dann reduziert sich die PDR auf 1. ... Für Kopfhörer, die FEC Eigenschaften erfüllen, muss ein eingeschaltetes Kompensationsfilter nur die Kopfhörer Transfer-Funktion selbst kompensieren, selbst wenn die Aufnahmen dazu bei blockiertem Gehörgang gemacht wurden. Für Kopfhörer, die keine FEC-Eigenschaften haben, muss die Kompensationsfilterung auch die PDR berücksichtigen. ...”

Die Untersuchungen von [Moller und Hammershoi, 1995] haben gezeigt, dass die

PDRs für einige – hochqualitative – Kopfhörer vernachlässigbar sind, zu denen auch der ‚Beyerdynamic DT990 Professional‘ (,offene‘ Kopfhörer-Bauweise) zählt.

9.3 Ermittlung der Transfer-Funktion ‚Kunstkopf + Kopfhörer‘

Um für die Testhörer die mittels Kunstkopf KU81 gemachten Aufnahmen über den Beyerdynamic DT990 korrekt entzerrt wiedergeben zu können, wurde die kombinierte Übertragungsfunktion KU81+DT990 messtechnisch ermittelt.

Audio Precision

12/05/05 20:51:33

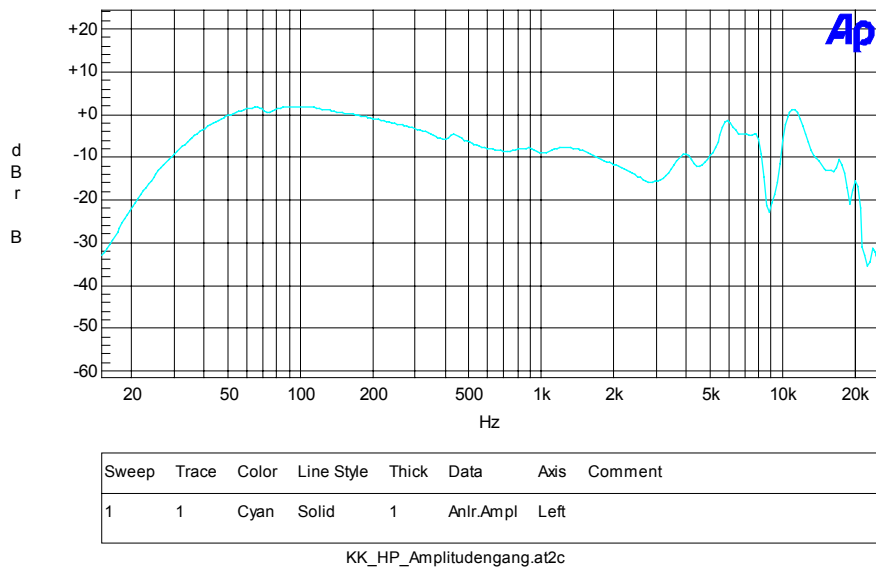


Abb. 9.3: kombinierter Amplitudengang Kunstkopf KU81i+Kopfhörer DT990, linkes Ohr

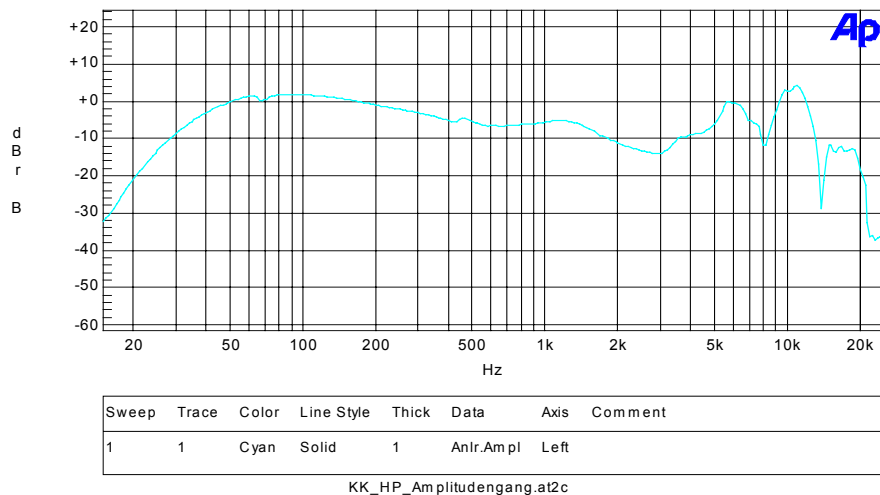


Abb. 9.4: kombinierter Amplitudengang Kunstkopf KU81i+Kopfhörer DT990, rechtes Ohr

Aus den in Abb. 9.3 und 9.4 dargestellten Übertragungsfunktionen für das linke und rechte Ohr des Kunstkopfs kann man erkennen, dass – bezogen auf den Pegel bei 1kHz – im tieffrequenten Bereich unter 200Hz eine breitbandige Überhöhung in der Größe von ca. 6-10dB besteht. Oberhalb von 1kHz gibt es einen – ebenfalls relativ breitbandigen - Pegel einbruch mit etwa 8dB Dämpfung bei 3kHz, danach wieder einen Pegelanstieg um ca. 6-8dB bis 7kHz.

Normalerweise sind die Frequenzgänge von HRTF's (Head-Related-Transfer-Functions) durch eine Resonanz im Bereich zwischen etwa 3-4kHz gekennzeichnet (Resonanzfrequenz des Gehörgangs), sowie durch den als ‚Pinnae Notch‘ bezeichneten Pegel einbruch im Amplituden-Frequenzgang, der durch das Außenohr bedingt ist. Dieser Einschnitt im Frequenzgang liegt in der Regel zwischen etwa 6-10kHz. (siehe Abb. 9.5. unten)

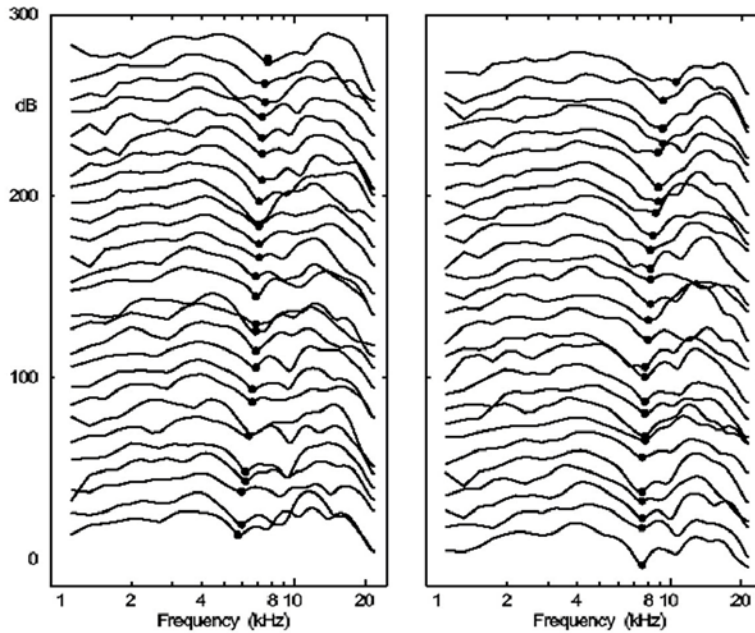


Abb. 9.5: HRTF-Amplitudenfrequenzgang ($\varphi=0^\circ$) von 54 Testpersonen (aus [Algazi et al., 2001]) (jeweils durch Punkt gekennzeichnet: ‚Pinnae Notch‘ Frequenz)

Wie in [Algazi et al. 2001] gezeigt werden konnte, ist die genaue Lage der Pinnae-Frequenz mit den Abmessungen des Außenohrs korreliert.

Die durch den Gehörgang bedingte Resonanzüberhöhung im Bereich zwischen 3 – 4kHz lässt sich durch die Analogie zu einer gedeckten Orgel-Pfeife (stehende Welle bei $\lambda/4$) erklären. Wie bei [Mehrgardt und Mellert, 1977] erklärt, tritt außerdem noch eine weitere Resonanzüberhöhung im Bereich um 9kHz auf.

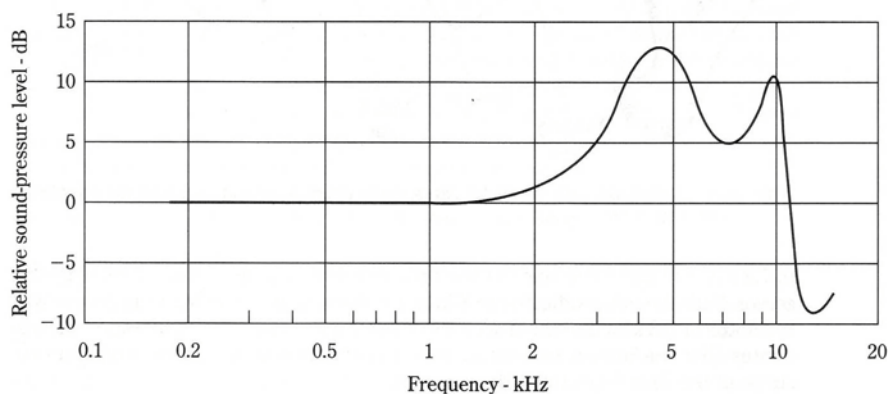


Abb. 9.6: Übertragungsfunktion des Gehörgangs mit Resonanzen bei 4 und 9kHz (nach [Mehrgardt und Mellert, 1977], aus [Everest, 1994])

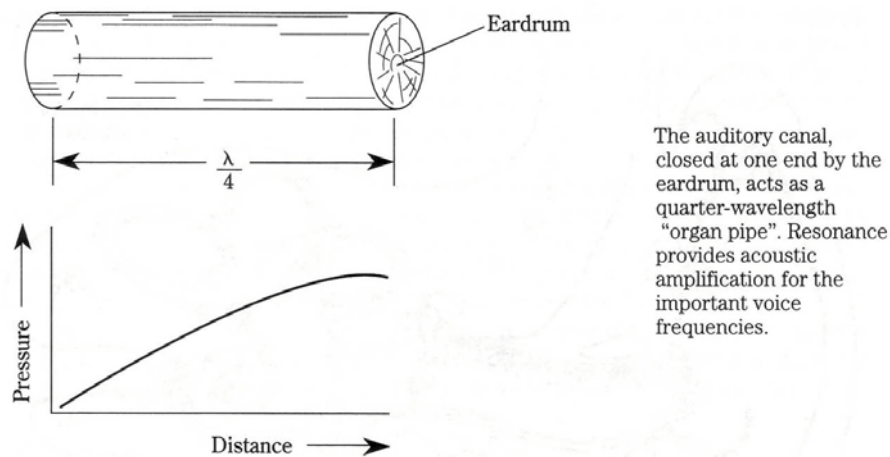


Abb. 9.7: Analogie Gehörgang (Cavum Conchae) – gedeckte Röhrenpfeife
(aus [Everest, 1994])

Das Zusammenwirken des durch das Außenohr bedingten Amplitudeneinbruchs (Pinnae Notch, im Bereich zwischen 6-10kHz) mit der oberen (durch den Gehörgang bedingten) Resonanzfrequenz, die in einem ähnlichen Frequenzbereich angesiedelt ist, ergibt sich für hohe Frequenzen (oberhalb ca. 7kHz) ein relativ ‚welliger‘ Amplituden-Frequenzgang siehe Abb.9.5) mit großen intersubjektiven Unterschieden.

Auf die Wiedergabe mittels individueller HRTFs wurde verzichtet. Im Prinzip könnte man die obenstehend ermittelte kombinierte Übertragungsfunktion (Kunstkopf + Kopfhörer) (siehe Abb. 9.3 und 9.4) durch Anwendung eines inversen Equalizers kompensieren, worauf allerdings bewusst verzichtet wurde. Dass die Ergebnisse der auf der Kopfhörerbeurteilung basierenden Hörvergleiche gut mit jenen der Surround-Lautsprecherwiedergabe übereinstimmen zeigt, dass durch diesen Verzicht offenbar keine relevante Verfälschung der Resultate eingetreten ist, was diese Entscheidung legitimiert.

Gemäß den bei [Moller et al., 1995] gewonnenen Erkenntnissen wäre ggf. von Frequenzgangkorrekturen oberhalb von 7kHz Abstand zu nehmen gewesen, zumal die (im oberen Frequenzbereich sehr komplexen) individuellen HRTF's der einzelnen Testpersonen ebenfalls nicht berücksichtigt werden. Somit wären im wesentlichen Entzerrungen im Bereich um 3 kHz (Anhebung um 6-8 dB), sowie eine breitbandige Pegelbedämpfung (ca. 6-8 dB) für den Bereich tiefer Frequenzen unter 200Hz

verblieben.

Die Tatsache, dass diese Entzerrungen nicht vorgenommen wurden, stellt allerdings nach Meinung des Autors keinen großen Nachteil dar, da sich die fehlende Anhebung der Frequenzen um 3 kHz für die Testpersonen ev. nur in einem teilweisen Fehlen der Präsenzen (bzw. ‚Direktheitseindruck‘ des Signals – siehe 2. Richtungsband betr. ‚Vorne-Ortung‘ (um 4kHz mit ca. 3kHz Bandbreite) nach [Blauert, 1974]) äußern, bzw. die Überhöhung bei den tiefen Frequenzen zu einer besseren Unterscheidung des Räumlichkeitseindrucks zwischen den verschiedenen Aufnahmen (= Mikrofontechniken) beitragen dürfte. (Betr. Wichtigkeit des Signalinhalts und Pegels bei tiefen Frequenzen für die Räumlichkeit siehe [Hidaka, Beranek und Okano, 1995].)

Außerdem sollte erwähnt werden, dass die Wahl für den Neumann KU81i Kunstkopf unter anderem deshalb getroffen worden war, weil dieser – gemäß Angaben des Herstellers – bei Verwendung eines diffusfeldentzerrten Kopfhörers keine zusätzliche Entzerrung benötigt. (siehe hierzu [Moller et al., 1999] S. 87), eine Bedingung, die durch die Verwendung des Beyerdynamic DT990 Pro erfüllt wurde.

Die unterschiedlichen Amplitudenfrequenzgänge der Abb. 9.3 und 9.4 erstaunen bei näherer Betrachtung, dürften allerdings eine Eigenheit des Neumann KU81i Kunstkopfs sein, der offenbar mit zwei leicht unterschiedlichen Ohren ausgestattet ist.

Die durch den Autor ermittelten Amplitudenfrequenzgänge stimmen jedenfalls relativ gut mit den Kompensationsfilter-Kurven aus der Arbeit von [Moller et al., 1999] überein (welche den invertierten Amplitudenfrequenzgängen entsprechen), wie untenstehend abgebildet:

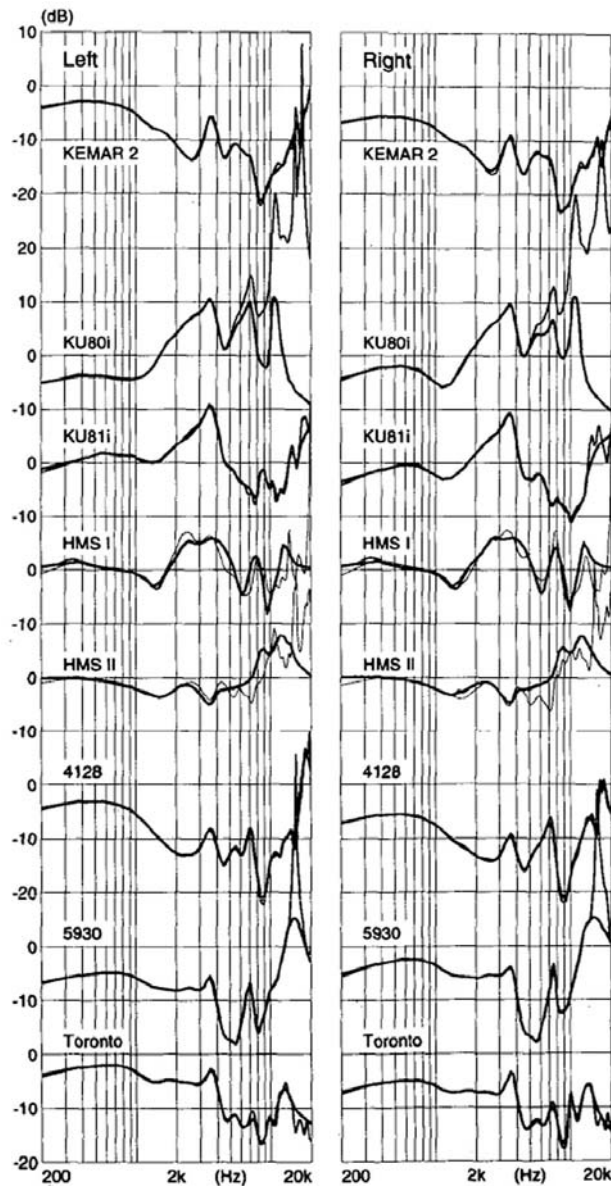


Fig. 1. Equalization filters for reproduction of recordings with built-in microphones. — target; — implemented IIR filter. Headphones were Beyerdynamics DT990 Professional (except for HMS I and HMS II) or Stax SR lambda professional (for HMS I and HMS II).

Abb. 9.8: Kompensationsfilter-Amplitudenfrequenzgänge unterschiedlicher Kombinationen aus Kunstkopf+Kopfhörer (Kopfhörer = Beyerdynamic DT990, sofern nicht anders angegeben). (aus [Moller et al., 1999], S. 88)

Hörvergleich KUNSTKOPF-ORCHESTERAUFNAHME

Name: _____ Datum: _____

Bitte markieren Sie jeweils durch Eintragung eines X in das passende Feld - Danke !

Ich bin Audiotechniker: Musiker:

Vergleichspaar #: _____ (bitte von der Testperson einzutragen)

1) Wie groß empfinden Sie den Unterschied zum Referenzbeispiel "REF" ?

| | | | | | | | | |
|--------------|---|---|---|---|---|---|---|------------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| A: sehr groß | | | | | | | | kein Unterschied |
| B: sehr groß | | | | | | | | kein Unterschied |

2) Wie groß empfinden Sie den Unterschied zwischen Beispiel A und B ?

| | | | | | | | | |
|-----------|--|--|--|--|--|--|--|------------------|
| sehr groß | | | | | | | | kein Unterschied |
|-----------|--|--|--|--|--|--|--|------------------|

3) Subjektives Gefallen der Aufnahme:

| | | | | | | | | |
|---------------|--|--|--|--|--|--|--|-----------|
| REF: sehr gut | | | | | | | | gar nicht |
| A: sehr gut | | | | | | | | gar nicht |
| B: sehr gut | | | | | | | | gar nicht |

4) Natürlichkeit:

| | | | | | | | | |
|-----------|--|--|--|--|--|--|--|-------|
| REF: sehr | | | | | | | | wenig |
| A: sehr | | | | | | | | wenig |
| B: sehr | | | | | | | | wenig |

5) Klangfarbe des Orchesters:

| | | | | | | | | |
|------------------|--|--|--|--|--|--|--|------------------|
| REF: natürlich | | | | | | | | REF: unnatürlich |
| A: natürlich | | | | | | | | A: unnatürlich |
| B: natürlich | | | | | | | | B: unnatürlich |
| REF: höhenbetont | | | | | | | | REF: bass-betont |
| A: höhenbetont | | | | | | | | A: bass-betont |
| B: höhenbetont | | | | | | | | B: bass-betont |

6) Lokalisation L-R:

(ungefähre Anordnung der Instrumente bzw. Instr.Gruppen entlang der Bühnenbreite von l nach r:
Klav, Schlagw, Hörner, Harfe, Viol., Xyloph., Holzbläser, Celli, Blechbläser, Kontra-Bässe)
Wie gut lassen sich die Instrumente bzw. Instrumentalgruppen im Orchester orten ?

| | | | | | | | | |
|---------------|--|--|--|--|--|--|--|----------------|
| REF: sehr gut | | | | | | | | REF: nicht gut |
| A: sehr gut | | | | | | | | A: nicht gut |
| B: sehr gut | | | | | | | | B: nicht gut |

7) Balance:

Sind die einzelnen Instrumente(ngruppen) im Orchester lautstärkenmässig ausgewogen ?
(siehe dazu auch die Liste der vorkommenden Instrumente(ngruppen) unter Punkt "Lokalisation")

| | | | | | | | | |
|-----------|--|--|--|--|--|--|--|--------------|
| REF: mehr | | | | | | | | REF: weniger |
| A: mehr | | | | | | | | A: weniger |
| B: mehr | | | | | | | | B: weniger |

8) Räumliche Abbildung des Orchesters:

| | | | | | | | | |
|---------------------------------|--|--|--|--|--|--|--|-------------------|
| "Stereo-Basisbreite" REF: breit | | | | | | | | REF: schmal |
| A: breit | | | | | | | | A: schmal |
| B: breit | | | | | | | | B: schmal |
| Entfernungseindruck REF: nahe | | | | | | | | REF: entfernt |
| A: nahe | | | | | | | | A: entfernt |
| B: nahe | | | | | | | | B: entfernt |
| Tiefenstaffelung REF: tief | | | | | | | | REF: flach |
| A: tief | | | | | | | | A: flach |
| B: tief | | | | | | | | B: flach |
| Stabilität der Abb. REF: stabil | | | | | | | | REF: instabil |
| A: stabil | | | | | | | | A: instabil |
| B: stabil | | | | | | | | B: instabil |
| Differenziertheit REF: präzise | | | | | | | | REF: verschwommen |
| A: präzise | | | | | | | | A: verschwommen |
| B: präzise | | | | | | | | B: verschwommen |

9) Abbildung des Raumes:

| | | | | | | | | |
|--------------------------------|--|--|--|--|--|--|--|------------------------------|
| REF: überzeugend | | | | | | | | REF: unbefriedigend |
| A: überzeugend | | | | | | | | A: unbefriedigend |
| B: überzeugend | | | | | | | | B: unbefriedigend |
| REF: zu viel Raum (zu hallig) | | | | | | | | R: zu wenig Raum(zu trocken) |
| A: zu viel Raum (zu hallig) | | | | | | | | A: zu wenig Raum(zu trocken) |
| B: zu viel Raum (zu hallig) | | | | | | | | B: zu wenig Raum(zu trocken) |

Kommentarzeile zu Punkt _____ : (bei Bedarf vom Testhörer zu verwenden)

Abb.9.9: Fragebogen Hörvergl. ,ORCHESTER-KK' (Kopfhörer-Wiedergabe)

9.4 Ablauf des Hörversuchs

Bei jedem Hörvergleich wurde zusätzlich zum Vergleichspaar von je zwei Mikrofontechniken auch der REFerenzstimulus angeboten. Die vergleichende Bewertung der Aufnahmen der Mikrofon-Arrays sollte also immer anhand der Referenz durchgeführt werden. Die Probanden hatten die Möglichkeit, jederzeit zwischen den drei angebotenen Aufnahmen umzuschalten. Das Umschalten wurde am digitalen Mischpult DM 2000 wiederum so gelöst, dass der Proband keine optische Rückmeldung über (LED-Meter basierte) Pegelanzeigen zu den Hörbeispielen bekam, um die Möglichkeit einer Beeinflussung durch visuelle Eindrücke auszuschließen.

Die Abhörlautstärke sollte dabei von den Probanden über den im Mischpult eingebauten Kopfhörer-Lautstärkeregler frei eingestellt, aber so gewählt werden, dass auch trotz der relativ langen Testdauer von 90min möglichst keine Ermüdung eintritt.

Wie in Abb. 9.9 ersichtlich, musste der Fragebogen für den Hörvergleich unter Kopfhörerwiedergabe etwas modifiziert werden: der Art der Abhörvorrichtung gemäß konnten ja keine ‚hinteren‘ Signale vom Testhörer unterschieden werden, weshalb die diesbezüglichen Fragen entfielen.

Als neuer Punkt kam hingegen die Frage hinzu, wie sehr sich die angebotenen Stimuli A,B nach Meinung des Hörers von der REFerenz (Original-Kunstkopfaufnahme aus dem Konzertsaal) unterschieden (Frage 1).

Zusätzlich sollte der Testhörer bei jeder einzelnen Frage auch den REF-Stimulus bewerten.

Aufgrund des Fehlens von ‚hinteren Kanälen‘ fielen im Rahmen der Abfolge der A/B/REF Hörvergleiche die Sonderbeispiele mit ‚+6dB Surround-Pegel‘ bzw. stummgeschalteten Surroundkanälen weg. Es blieben lediglich die Sonderfälle des Ident-Beispiels (REF/A/REF) und des Redundanz-Beispiels (Wiederholung eines Stimulustripels). Somit waren mit den normalen A/B/REF Vergleichen (=6 Stück) zzgl. der zwei Sonderfälle insgesamt 8 Tripel-Vergleiche zu bewerten.

Am 6. und 7.12.2005 wurden im Produktionsstudio des IEM – mit dem Autor als Versuchsleiter - die Hörvergleiche der Kunstkopfaufnahmen mit 10 Probanden („Gruppe KK“; alles Studenten des IEM, sowie ein fertig ausgebildeter Tonmeister; keine personelle Überschneidung mit Testhörern der anderen Gruppen) durchgeführt.

Der Hörvergleich wurde als ‚double-blind‘ Test mit Referenz-Stimulus durchgeführt, wobei den Testhörern bekannt war, welches der Signale der REFERENZ-Stimulus (Original Kunstkopf-Aufnahme aus dem Konzertsaal) ist.

Die Teststimuli des Kunstkopf-Hörvergleichs bestanden aus der originalen Kunstkopf-Orchesteraufnahme aus dem Konzertsaal (Grosses Festspielhaus, Salzburg) die als Referenz angeboten wurde, und den mit demselben Kunstkopf (Neumann KU81) aufgenommenen Re-Recordings der über die Surround-Lautsprecheranordnung (des Produktionsstudios IEM) wiedergegebenen Aufnahmen der zu bewertenden vier Mikrofon-Arrays.

Wie bereits oben angegeben wurde zur Wiedergabe der Kunstkopfaufnahmen – aus den erwähnten Gründen (FEC-Eigenschaften, etc.) – ein Beyerdynamic ‚DT990 Professional‘ Kopfhörer zum Einsatz gebracht.

Durch die mischpultinternen Beschriftungsmöglichkeiten (Channel Display) wurde dem Testhörer angezeigt, welches Signal (A, B oder REF) an welchem Kanal anliegt; die Reihenfolge der Kanal-Belegung war bei jedem Tripel-Vergleich dieselbe.

Aufgrund der Erfahrungen der Hörvergleiche IEM und SF wurde die Testdauer pro Proband von vornherein mit 90min angesetzt.

Vor Beginn des Hörvergleichs hörten die Probanden eine „Kalibrierungs-Sektion“ ab, die ähnlich wie beim Hörvergleich mittels Lautsprecherwiedergabe gestaltet war.

Bedingt durch die hohe Anzahl an Fragen, die an die Testhörer gestellt wurden, war eine volle Absolvierung des Hörtests nur einem der Probanden möglich; beinahe die Hälfte der Probanden schaffte in der zur Verfügung stehenden Zeit 5 von 8 Hörvergleichen.

Daraus lässt sich der Schluss ziehen, dass die vergleichende Bewertung der über Kopfhörer angebotenen Stimuli offenbar deutlich schwieriger ist, als die Bewertung der über Surround-Lautsprecheranlage angebotenen Stimuli.

9.5 Analyse: Hörvergleich Orchesteraufnahme unter Kopfhörerwiedergabe

Vorab soll gesagt werden, dass – aufgrund der nicht vollständigen Beantwortung der Fragen durch die Testhörer – kein wirklich ausgewogenes Datenmaterial vorliegt. Deshalb kann man nicht davon ausgehen, dass die - bereits weiter oben erläuterten - Grundbedingungen für die Anwendung einer Varianzanalyse streng eingehalten sind. Trotzdem wurde eine solche Analyse vorgenommen, deren Ergebnisse untenstehend aufgeführt sind.

| „Source“ | Sum of squares | Df | Mean square | F | p | signifikante Differenz in den Mittelwerten der Mic-Array Beurteilungen |
|-------------------|----------------|----|-------------|---------|---------------|--|
| Unterschied | 31.3394 | 2 | 10.4465 | 5.6872 | 0.0027 | 1&3, 2&3, 4&3 |
| Präferenz | 31.4602 | 3 | 10.4867 | 5.5144 | 0.0032 | 3&4 |
| Natürlichkeit | 35.4 | 3 | 11.8 | 10.5526 | 4.02E-05 | 1&3, 2&3, 4&3 |
| Klangfarbe | 37.8214 | 3 | 12.6071 | 14.0776 | 3.14E-06 | 1&3, 2&3, 4&3 |
| Höhen/Tiefen | 9.4583 | 3 | 3.1528 | 2.8363 | 0.0517 | keine |
| Lokalisation | 23.7339 | 3 | 7.9113 | 3.5448 | 0.0239 | 2&4, 3&4 |
| Balance | 26.4394 | 3 | 8.8131 | 7.1988 | 6.61E-04 | 3&4 |
| Basisbreite | 29.2201 | 3 | 9.74 | 8.2485 | 2.63E-04 | 1&3, 2&3, 4&3 |
| Nahe / entfernt | 10.1352 | 3 | 3.3784 | 2.0932 | 0.1183 | keine |
| Tiefenstaffelung | 14.5589 | 3 | 4.853 | 3.251 | 0.0329 | keine |
| Stabilität | 0.4451 | 3 | 0.1484 | 0.0775 | 0.9718 | keine |
| Differenziertheit | 7.6208 | 3 | 2.5403 | 1.4245 | 0.2516 | keine |
| Raumabbildung | 29.0688 | 3 | 9.6896 | 11.5494 | 1.89E-05 | 1&3, 2&3, 4&3 |
| hallig/trocken | 15.7089 | 3 | 5.2363 | 6.9765 | 8.08E-04 | 1&3, 2&3 |

Anm.: 1=OCT, 2=DECCA,
3=KFM, 4=ABPC

Tab. 9.1: ‚ORCH-KK‘ 1-faktorielle Varianzanalyse betreffend Einfluss der Variable ‚Mikrofon-Array‘ auf die Bewertung der Klangaspekte (Kopfhörerwdgb., 10 Hörer)

Wie man aus obenstehender Abbildung ersehen kann, liegen immerhin bei 10 von 14 Aspekten p-Werte mit $p < 0.05$ vor, sodass von einem statistisch signifikanten Unterschied bei zumindest jeweils einer der 4 Mikrofontechniken hinsichtlich des Mittelwerts des jeweiligen Aspekts ausgegangen werden kann.

Lediglich bei den Aspekten Höhen/Tiefen, nahe/entfernt, Stabilität und Differenziertheit konnte das Signifikanz-Niveau nicht erreicht werden. Wie auch aus der nachfolgenden Abbildung 9.10 zu erkennen ist, treten die signifikanten Unterschiede zwischen den Mikrofontechniken allerdings vor allem in Bezug auf die dritte Mikrofontechnik (KFM) auf, welche auch bereits bei der Hörerbeurteilung unter 5.1 Surround-Lautsprecherwiedergabe deutlich schlechter als die anderen abgeschnitten hatte.

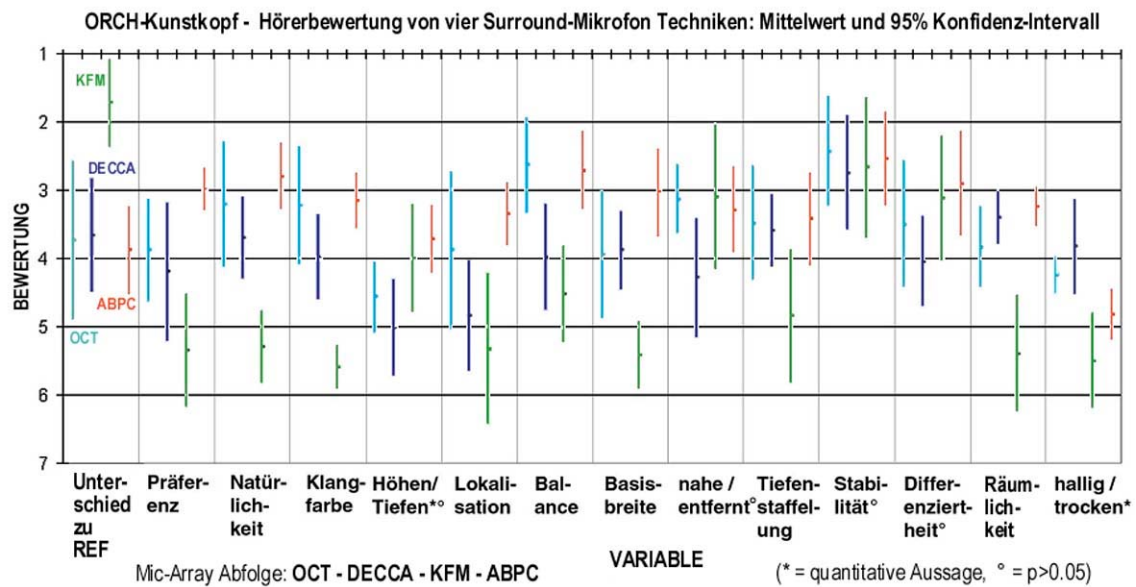


Abb. 9.10: ORCH-Kunstkopf - Mittelwert und 95%-Konfidenzintervall für 14 Aspekte und 4 Mikrofontechniken (10 Hörer) *

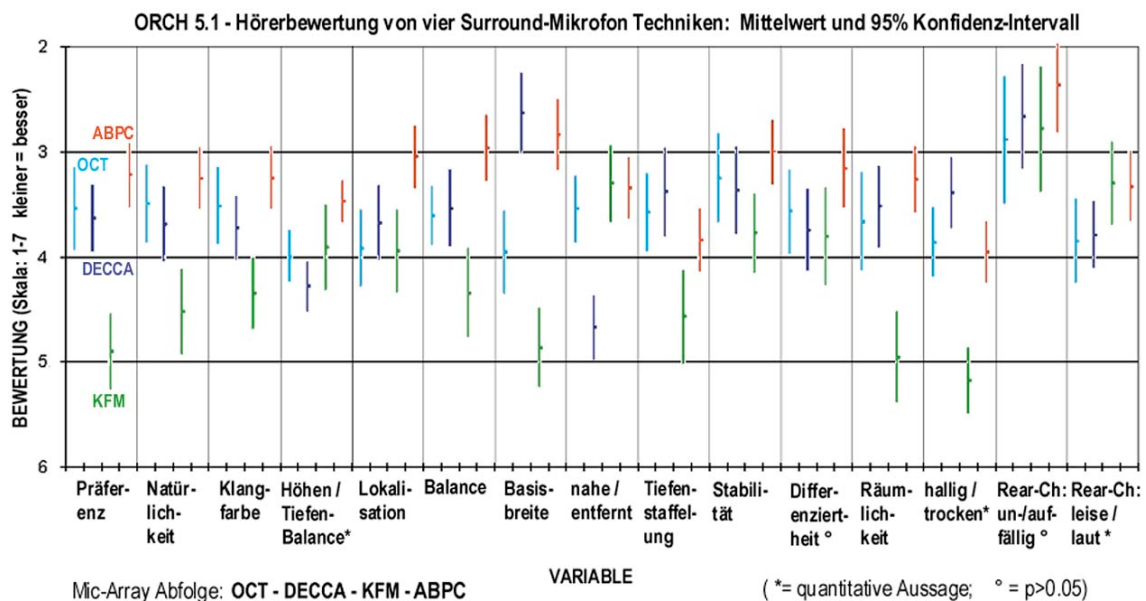


Abb. 9.11: ORCH 5.1 - Mittelwert und 95% Konfidenzintervall für 15 Aspekte und 4 Mikrofontechniken (25 Hörer) (Wiederholung)

* **Anm.:** die Darstellung der Ergebnisse betreffend ‚Unterschied zu REF‘ der Abb. 9.10 müsste korrekter Weise eigentlich um den Skalenmittelwert 4 gespiegelt sein. (Die aktuelle Darstellung resultiert aus einem Fehler in der betreffenden Formulierung des Fragebogens)

Aufgrund der geringeren Anzahl an Testpersonen (10 Probanden) und der Werteverteilung bezüglich des jeweiligen Aspekts im vorliegenden Datensatz ‚ORCH-Kunstkopf‘ sind die errechneten Konfidenzintervalle teilweise sehr breit und überlappend. Statistisch signifikante Unterschiede liegen primär nur zwischen der dritten Mikrofontechnik (KFM) und den anderen Mikrofon-Arrays vor.

| | Präferenz- Bewertung (durch Testhörer) |
|-------|---|
| OCT | 3,48 |
| DECCA | 4,20 |
| KFM | 5,35 |
| AB-PC | 2,99 |

Tab. 9.2: Angabe zur Präferenz von 4 Mikrofontechniken
(Kunstkopfaufnahme, Kopfhörerwiedergabe)

Angemerkt sei an dieser Stelle noch, dass prinzipiell auch die Möglichkeit bestanden hätte, dass die Signalinformation durch die Reduktion von 5 auf 2 Wiedergabekanäle (‚binaurales Re-Recording‘: 5.1 Surround → Kunstkopf) soweit verändert wird, dass dabei auch die hörbaren Unterschiede zwischen den vier Mikrofontechniken relevant geringer werden. Anstatt einer ‚Nivellierung‘, scheint aber im Gegenteil eine Art ‚Vergrößerung‘ (im positiven Sinne einer besseren Unterscheidbarkeit) des Höreindrucks stattgefunden zu haben.

Erwartungsgemäß hat sich – im Vergleich zum Hörvergleich mit Surround-Lautsprecherwiedergabe – die Bewertung diverser raumakustischer Aspekte verändert, da u. a. die Möglichkeit zur Diskriminierung von vorn/hinten Klangreizen (oder auch der Kopfdrehung generell, z.B. zwecks besserer Lokalisation, etc.) unter Kopfhörerwiedergabe entfällt.

Besonders betroffen von der Verschlechterung der Signalinformation durch Reduktion von Surround/Lautsprecher- auf Binaural/Kopfhörer- Wiedergabe ist die

DECCA Technik, da sich für diese die Einstufung (Reihung) in Relation zu den anderen drei Mikrofontechniken bei immerhin 5 von 15 Aspekten verschlechtert hat (dies sind die Aspekte: Lokalisation, Balance, Basisbreite, Tiefenstaffelung, Differenziertheit).

Als Erklärung bietet sich hier an, dass die diesbezüglichen Signalinformationen, welche – gemäß der omnidirektionalen Charakteristik der drei Frontmikrofone – in erster Linie auf Laufzeitunterschieden basieren, durch das ‚Downfolding‘ auf zwei Kanäle (mittels Re-Recording durch den Kunstkopf) stark beeinträchtigt werden. Die anderen Mikrofontechniken, bei denen die Unterschiede in den Signalinformationen primär durch die Richtcharakteristik der verwendeten Mikrofone bedingt sind, leiden nicht unter einer vergleichbaren Verschlechterung hinsichtlich ihrer qualitativen Bewertung.

Interessant scheint auch, dass sich durch die Reduktion auf zwei (binaurale) Kanäle beim Aspekt ‚Stabilität‘ für alle Mikrofontechniken eine bessere Bewertung ergeben hat. Das ‚Einfrieren‘ der - mittels Kunstkopf-Rerecording neu abgebildeten vier Tonaufnahmen - scheint sich somit generell auf die Eindeutigkeit (Stabilität) aller verwendeten Mikrofontechniken positiv auszuwirken. (Ein Grund dafür könnte im Fehlen der Möglichkeit der Kopfdrehung und somit dem Fehlen ev. Ambiguitäten in der Signalperzeption durch den Testhörer liegen.)

Abschließend lässt sich feststellen, dass generell eine gute Übereinstimmung der Ergebnisse aus dem Hörvergleich der Kunstkopfsignale mit Kopfhörerwiedergabe (ORCH-KK) mit den Ergebnissen aus dem Hörvergleich mit Surround-Lautsprecheranordnung (ORCH-5.1) vorliegt.

Darüber hinaus soll an dieser Stelle angemerkt werden, dass der Hörvergleich ORCH-KK auch deshalb durchgeführt worden war, um zu überprüfen, ob die für eine qualitative Bewertung durch den Testhörer relevanten Signalinhalte auch im Signal der Kunstkopf-Aufnahme nach wie vor ausreichend vertreten sind, um sicherzustellen, dass eine messtechnische Auswertung desselben (in unserem Fall hinsichtlich der Signal-Korrelation bzw. Kohärenz) sinnvoll ist. Auch dieses ließ sich anhand der in diesem Abschnitt obenstehend dokumentierten Ergebnisse bestätigen.

10 Untersuchung der Korrelation und Kohärenz bei Surround-Signalen

Vermutlich ist die technisch objektivste Methode zur Ermittlung, welche Mikrofontechnik ein akustisches Ereignis am natürlichsten erfasst, folgende:

1. Man nimmt das Klangereignis (z.B. Orch.-Konzert) mit der Mikrofontechnik seiner Wahl auf
2. Zur gleichen Zeit fertigt man mittels Kunstkopf eine zweite Aufnahme an einer möglichst guten Hörposition im Konzertsaal bzw. Aufnahmeraum an
3. während man die erste Aufnahme auf einer Stereo- oder Surround-Anlage abspielt, wird mittels am ‚sweet spot‘ (optimale Abhörposition) aufgestelltem Kunstkopf eine dritte Aufnahme angefertigt
4. man bestimmt für beide so entstandenen Kunstkopfaufnahmen die Korrelation ihrer Ohrensinnale (vorzugsweise nicht nur des Gesamtsignals, sondern über der Frequenz, da sich hiermit wesentlich genauere Aussagen über die Ähnlichkeit der Signale treffen lassen; d.h. die Ohrensinnale der Aufnahme aus dem Konzertsaal werden kreuz-korreliert und ebenso die Ohrensinnale des Re-Recordings im Abhörraum).

Quasi unvermeidbar ist bei dieser Art der Neu-Aufnahme (Re-Recording), daß eine Signalverfälschung in Form von Amplituden-, Frequenz- und/oder auch Phasen-Verzerrung durch die gesamte Signalkette von der verwendeten Aufnahmetechnik über das Wiedergabesystem sowie die Akustik des Abhörraum in das neu entstehende Binauralsignal einfließen.

Unabhängig davon sollte jene Mikrofontechnik, welche die größte Ähnlichkeit zwischen FIACC (Frequency-dependent Inter-Aural Cross-Correlation) des Originalsignals und FIACC der aufgenommenen Reproduktion aufweist, die Mikrofontechnik mit der natürlichsten Klangabbildung sein.

10.1 Auswahl der Surround-Mikrofonssysteme und des Aufnahmewinkels

Von den zum Zeitpunkt der Testaufnahmen (21.5.2003) bekannten Surround-Mikrofontechniken wurden die untenstehend beschriebenen fünf ausgewählt. Hierbei spielte die Anzahl der zur Verfügung stehenden Tonspuren (in unserem Fall $3 \times 8 = 24$

linear 16-bit PCM Spuren in Form von drei TASCAM DA-88 digital 8-Spur Recordern), als auch eine gewünschte Diversifizierung der Surround-Verfahren eine Rolle:

- Das OCT-Surround Verfahren wurde ausgewählt, weil es bereits im Rahmen einiger anderer Hörvergleiche gut abgeschnitten hatte und somit auch gut als Referenz herangezogen werden kann, um die erstmals zu bewertenden Surround-Mikrofontechnik AB-PC in Relation zu anderen bereits bekannten Techniken zu setzen.

(Anm.: der ‚OCT Version 2‘ Systemvorschlag wurde von Wittek und Theile erst im Oktober 2004 vorgestellt und stand deshalb zum Zeitpunkt der Durchführung der Aufnahme nicht zur Diskussion)

- Ähnliches lässt sich auch für das DECCA-System ins Treffen führen.

- Das Kugelflächenmikrofon KFM und das Soundfield-Mikrofon (SF) sind beide nicht aus Einzelmikrofonen zusammengesetzt, sondern speziell für die Surround- (bzw. auch Stereo-) Anwendung gebaut. Deshalb nehmen sie eine Sonderstellung ein und verdienen alleine deshalb Beachtung; außerdem ist der ‚one-point‘ Ansatz des Soundfield-Mikrofons, welcher sich auf die Erfassung von Pegelunterschieden in Abhängigkeit von der Einfallsrichtung des Schalls beschränkt ein grundlegend anderer, als bei den Vergleichssystemen. (Anm.: diesem steht der quasi rein ‚laufzeitbasierte‘ Systemansatz des DECCA-Systems gegenüber)

Alle Systeme wurden im Kapitel über Surround-Mikrofonsysteme in ihren Grundzügen bereits dargestellt, deshalb soll hier nur kurz auf die spezifische Anordnung (Aufnahmewinkel, etc.) im Rahmen der bei den Salzburger Festspielen gemachten Tonaufnahmen eingegangen werden.

Die Wahl des Aufnahmewinkels von 120 Grad lag einerseits aufgrund der Herstellerangabe von Schoeps, das Kugelflächenmikrofon hätte (in der Stereoanwendung) ebendiesen effektiven Aufnahmewinkel, nahe; zudem stimmte dieser auch gut mit der vor Ort an der Mikrofonposition tatsächlich auftretenden Schallquellenausdehnung des Orchesters überein.

- OCT-Surround: für den Aufnahmewinkel von 120° wurden die Mikrofonabstände des Front-Systems entsprechend gewählt; 60cm zwischen den Kapseln L und R (Hypernieren), die mittlere ist standardgemäß um 8cm nach vorne Richtung Schallquelle versetzt.

Für die seitwärts nach außen gerichteten Hypernieren kamen Schoeps MK5 zum Einsatz, für die Nieren des Center- und das LS, RS Mikrofons Schoeps MK4 Kapseln.

- die drei omnidirektionalen Mikrofone des DECCA-Dreiecks wurden mit Abständen gemäß den ITD-basierten Lokalisations-Prinzipien der Psychoakustik gewählt; nach [Hermann et al., 1998] wurden hierfür folgende Mikrofonabstände gewählt: Kapselabstand zwischen L und R 128cm, das mittlere Mikrofon war um 74cm nach vorne versetzt. Mikrofone: Schoeps MK2s (omnidirektional)

Für die Rear-Mikrofone wurde vom rein lauffzeit-basierten Prinzip abgewichen und jene zwei - Richtung Rückwand des Konzertsaals gerichteten - Nierenmikrofone verwendet, die Teil des unten beschriebenen AB-PC Systems bilden. Dies geschah vor allem auch deshalb, weil im Rahmen der 24 zur Verfügung stehenden Aufnahmespuren keine weiteren zwei für eigene Rear-Mikrofone zur Ergänzung des DECCA-Dreiecks frei gewesen wären.

- das Kugelflächenmikrofon ‚KFM 360‘, welches gemäß Herstellerangabe im Stereomodus einen Aufnahmewinkel von 120° hat, wurde für die Surround-Anwendung mit zwei seitlichen Mikrofonen mit Achtercharakteristik ergänzt. Die Dematrizierung der Audiosignale für 5.1 Surround erfolgte mittels des Schoeps DSP-4 Prozessors (in Form einer einfachen MS-Dematrizierung und Bildung einer Center-Information gemäß Gerzon-Matrix, allerdings ohne zusätzliches Processing wie Delays für die Rear-Kanäle, o. ä.)

- das Soundfield-Mikrofon (SF) mit dem zugehörigen SP451 Prozessor wurde einige Meter weiter als die anderen Mikrofone vom Orchester im Konzertsaal gehängt, da Untersuchungen gezeigt hatten, dass sein Signal von den Testhörern generell als trockener bewertet wird, als jenes anderer Surroundmikrofon-Systeme, die an vergleichbarer Position hängen.

(An dieser Stelle sollte nochmals angemerkt werden, dass die Aufnahme des SF-Mikrofons letztlich in den Hörvergleich nicht einbezogen wurde, da zu diesem System bereits mehrere Bewertungen in Form von Hörvergleichen vorliegen, und die umfangreiche Bewertung der vier anderen Systeme die Testhörer bereits an ihre Leistungsgrenze brachten. Da keine Hörerbewertung vorliegt, wurde auch darauf verzichtet, die Signale des SF einer messtechnischen Auswertung zu unterziehen.)

- AB-PC Surround-System (erstmalig vorgeschlagen von [Edwin Pfanzagl-Cardone, 2002]) bestehend aus einem ‚Breit-AB‘ Paar von omnidirektionalen Mikrofonen (bei der vorliegenden Aufnahme: Schoeps MK3, diffusfeldentzerrte Druckempfänger) mit einem Kapselabstand von 12m, sowie einem Centerfill-System in Form des sog. ORTF-Triples (ORTF-T), welches aus einem ORTF-Stereomikrofon mit einer zusätzlichen Center-Niere besteht, deren Signal direkt auf den Center-Kanal geschaltet wird. Die Signale des L und R ORTF-Mikrofons werden hingegen bei ca. 11 und 13 Uhr panoramisiert und auf den L und R Front-Bus dazugemischt. Gemäß ihrer Position relativ zum Orchester sind diese Mikrofone mehr als Stützen für die ersten Pulte der Streicher (in unmittelbarer Nähe des Dirigenten) zu verstehen, denn als eigenständiges Hauptmikrofonsystem. Das ORTF-Triple ist in seinem Verhalten durch die Richtcharakteristik der drei verwendeten Nieren charakterisiert und hilft somit ev. ‚Unsicherheiten‘, die durch das Breit-AB System alleine hinsichtlich der Lokalisation auftreten könnten, zu beseitigen.

Für die Rear-Mikrofone des AB-PC Systems kommen zwei Richtung Saalrückwand gerichtete Nierenmikrofone zum Einsatz, die zueinander und zum Orchester einen Abstand von zumindest ca. 10m haben.

Beim AB-PC-System besteht keine Möglichkeit, den effektiven Aufnahmewinkel gezielt in Richtung 120 Grad zu wählen; das System ist für die Abbildung mittlerer bis großer Schallquellen konzipiert, und trachtet danach, diese in ihrer Gesamtheit möglichst unverfälscht (hinsichtlich verschiedenster akustischer Aspekte) aufzunehmen.

- im Rahmen einer Duo-Aufnahme mit Cello und Klavier kam statt des AB-PC Systems das sog. ‚Blumlein Pfanzagl-Triple‘ (abgek. BPT) zum Einsatz, welches ebenfalls im Kapitel über Surround-Mikrofone im Detail beschrieben ist.

So wie das AB-PC System für große Schallquellen, verfolgt das BPT System für kleine Schallquellen das Ziel einer möglichst ‚dekorrelierten‘, mehrkanaligen Aufnahme des Schallfeldes.

10.2 Mikrofon- und Kunstkopfpositionen

Alle oben genannten Surround-Mikrofonsysteme (außer SF) wurden etwa oberhalb der 1. und 2. Parterre-Sitzreihe von der Decke des Konzertsaals abgehängt, sodass sie sich auf einer Höhe von ca. 3.5m über dem Orchesterpodium befanden. Um einen möglichst objektiven Vergleich der getesteten Mikrofonssysteme zu ermöglichen und nicht durch ev. bestehende mikrofon-/oder herstellerspezifische Unterschiede beeinflusst zu werden, wurden (für die Orchesteraufnahme) nur Mikrofone eines Herstellers (Fa. Schoeps, Deutschland) verwendet.

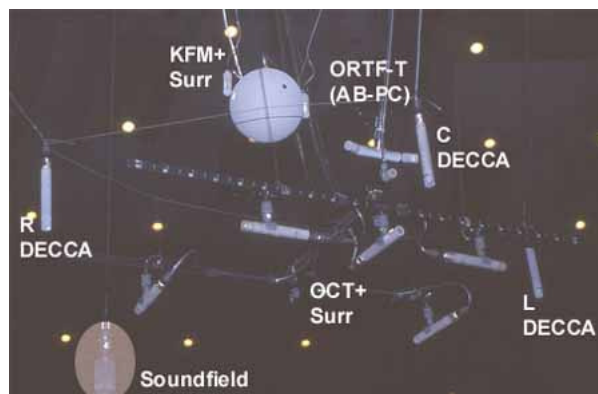


Abb. 10.1: Die bei der Aufnahme ‚ORCH 5.1‘ verwendeten Mikrofone (Platzierung: ca. 3.5m über dem Orchesterpodium, zwischen den ersten beiden Sitzreihen im Parterre)



Abb. 10.2: Platzierung der Surround-Mikrofon Arrays im Gr Festspielhaus Salzburg (nur Vorbühne ersichtlich, da sich der hintere Teil des Konzertsaals hinter dem [temporär] geschlossenen Eisernen Vorhang [im Bild rechts außen] befindet)

Zeitgleich mit den Surroundmikrofon-Systemen wurden auch die Signale eines Neumann KU81i Kunstkopfes aufgezeichnet, der mittig in der 6. Parterre-Reihe platziert war. Die Wahl des Aufstellungsortes erfolgte vor allem auch um für dessen Signale ein passendes Direkt-/Diffusschallverhältnis zu erzielen: der Hallradius des Großen Festspielhauses Salzburg liegt (für mittlere Frequenzen) bei 5.5m, somit befanden sich zumindest die ersten Reihen der Musiker für den Kunstkopf innerhalb desselben.



Abb. 10.3: Kunstkopf Neumann KU81i in der 6. Reihe, Mitte des Großen Festspielhauses Salzburg

| Oktavband: [Hz] | RT ₆₀ [s] | Hallradius r [m] |
|--------------------|-------------------------|---------------------|
| 8000 | - | - |
| 4000 | 1.54 | 5,72 |
| 2000 | 1.73 | 5,40 |
| 1000 | 1.75 | 5,36 |
| 500 | 1.85 | 5,22 |
| 250 | 1.88 | 5,18 |
| 125 | 1.78 | 5,32 |
| 63 | 1.53 | 5,74 |

| |
|---|
| Raumvolumen V [m ³] 15500 |
|---|

$$r = 0,057 * \sqrt{V / RT_{60}}$$

| |
|------------------------------|
| Raum: Grosses Haus |
| B x H x T in [m]: |

Anm.: ‚RT₆₀‘ Werte sind in diesem Fall eigentlich RT 5ms-35ms Messwerte aus der Studie von Takenaka Research & Development Institute, Japan "Acoustic Measurements of Grosses Festspielhaus, Salzburg" (1996)

Tab. 10.1: Berechnung des frequenzabhängigen Hallradius‘ (critical distance) für das Große Festspielhaus Salzburg auf Basis von Nachhallzeit RT₆₀ (ohne Publikum) und Raumvolumen

10.3 Programm-Material

Orchesteraufnahme

Gegenstand der Aufnahme, welche am 21.5.2003 im Großen Festspielhaus Salzburg stattfand, war ein symphonisches Werk für Orchester von Lutoslavsky. Das Große Festspielhaus hat ca. 2160 Sitzplätze, ein Luftvolumen von etwa 15500 m³ [Beranek, 2004] und ist durch eine mittlere Nachhallzeit von 1.5sec (mit Publikum) gekennzeichnet. Die akustische Qualität des Konzertsaals wurde im Rahmen einer internationalen Studie von [Hidaka et al., 1995] mit 'B+' bewertet, was ‚gut bis exzellent‘ entspricht.

Duo Aufnahme

Die Aufnahme eines Stücks für Duo Cello-Klavier wurde in einem Proberaum des Salzburger Festspielhauses durchgeführt. Dieser Raum ist raumakustisch adaptiert (Schallabsorber und Diffusoren) und hat die Abmessungen 20.5m x 11.35m x 7.8m und somit ein Raumvolumen von ca. 1815 Kubikmeter.



Abb. 10.4: Duo-Aufnahme im Orchesterproberaum 447

| Oktavband: [Hz] | RT ₆₀ [s] | Hallradius r [m] |
|--------------------|-------------------------|---------------------|
| 8000 | 0.58 | 3.19 |
| 4000 | 1.03 | 2.39 |
| 2000 | 1.43 | 2.03 |
| 1000 | 1.54 | 1.96 |
| 500 | 1.34 | 2.10 |
| 250 | 1.17 | 2.25 |
| 125 | 1.22 | 2.20 |
| 63 | 1.98 | 1.73 |

| |
|--|
| Raumvolumen V [m ³] 1815 |
|--|

$$r = 0,057 * \sqrt{V / RT_{60}}$$

| |
|--|
| Raum: Proberaum 447 L x B x H in [m]: 20,5 x 11,35 x 7,8 |
|--|

Tab 10.2: Berechnung des frequenzabhängigen Hallradius' des Proberaums 447 basierend auf Nachhallzeit RT₆₀ und Raumvolumen

(Anm.: die Messung der RT₆₀-Werte erfolgte mittels Neutrik ‚Acoustilizer‘ [Klasse 2 Gerät nach IEC 60651])

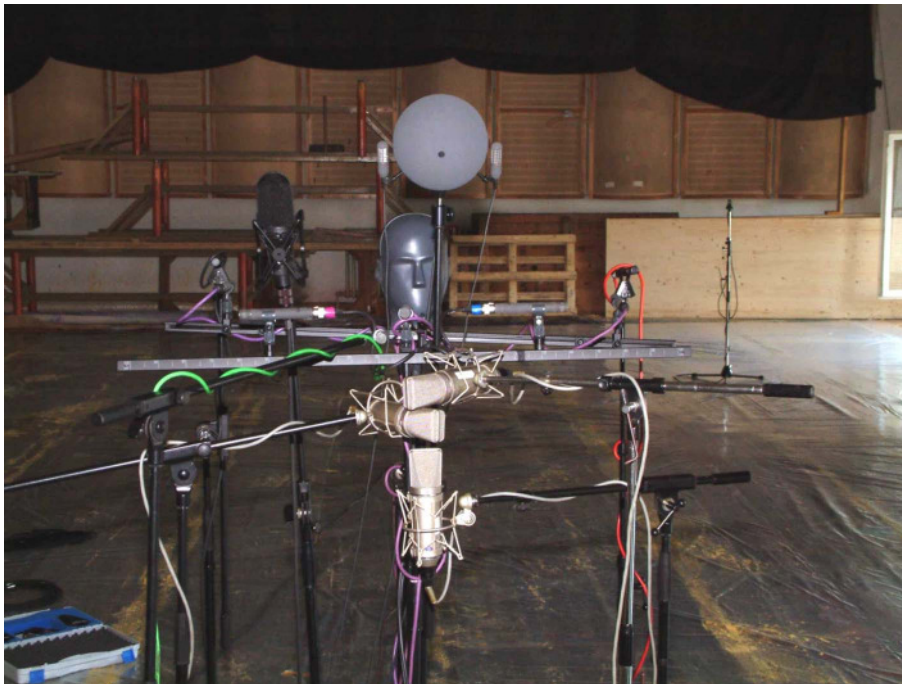


Abb. 10.5: Surround-Mikrofonsysteme (BPT, DECCA, KFM, OCT und Soundfield) sowie Kunstkopf bei der Duo-Aufnahme im Proberaum 447

10.4 Messung der frequenzabhängigen Kreuz-Korrelation (FCC) und frequenzabhängigen interauralen Kreuz-Korrelation (FIACC)

Um die frequenzabhängige Kreuzkorrelation von zwei Signalen zu bestimmen werden die Signale der beiden Kanäle in 31 Frequenzbänder gesplittet (mittels entsprechender Hoch- und Tiefpassfilterung; unterhalb von 59 Hz mit einem Filter 4. Ordnung [d.h. mit einer Flankensteilheit von 24dB/Oktave im Sperrbereich], darüber mit Filtern 6.Ordnung [d.h. 36dB/Oktave]) mit Mittenfrequenzen nach ISO-Norm, entsprechend 1/3-Oktavbandbreite.

Nach der Filterung werden die Bandpasssignale einer Korrelations-Analyse unterzogen, welche als Ergebnis den Korrelations-Koeffizienten zwischen -1 und 1 liefert.

Zusätzlich zu den Korrelationskoeffizienten der 31 Frequenzbänder wird auch der Korrelationskoeffizient des Gesamtsignals, sowie auch folgender Teilsignale ermittelt: das mit einer Grenzfrequenz von 400Hz tiefpass-gefilterten Signals (relevant vor allem in Hinblick auf den Raumeindruck), sowie des mit einer Frequenz von 1200Hz hochpassgefilterten Signals.

Für die Messung wurden geeignete Ausschnitte der für den Hörtest verwendeten Aufnahme ausgewählt:

ORCHester-Aufnahme (Saal 1):

- Sample (Ausschnitt) # 1, 7 sec (Dynamik: forte)
- Sample # 2, 16sec (piano)
- Sample # 4, 60sec (piano bis forte)

DUO-Aufnahme (Saal 2):

- Sample # 3, 16sec

In den Abbildungen 10.6, 10.8 und 10.9 ist der FIACC der originalen Kunstkopfaufnahme aus dem Konzertsaal gemeinsam mit der Kunstkopf Neu-Aufnahme der jeweiligen Surround-Mikrofontechnik dargestellt.

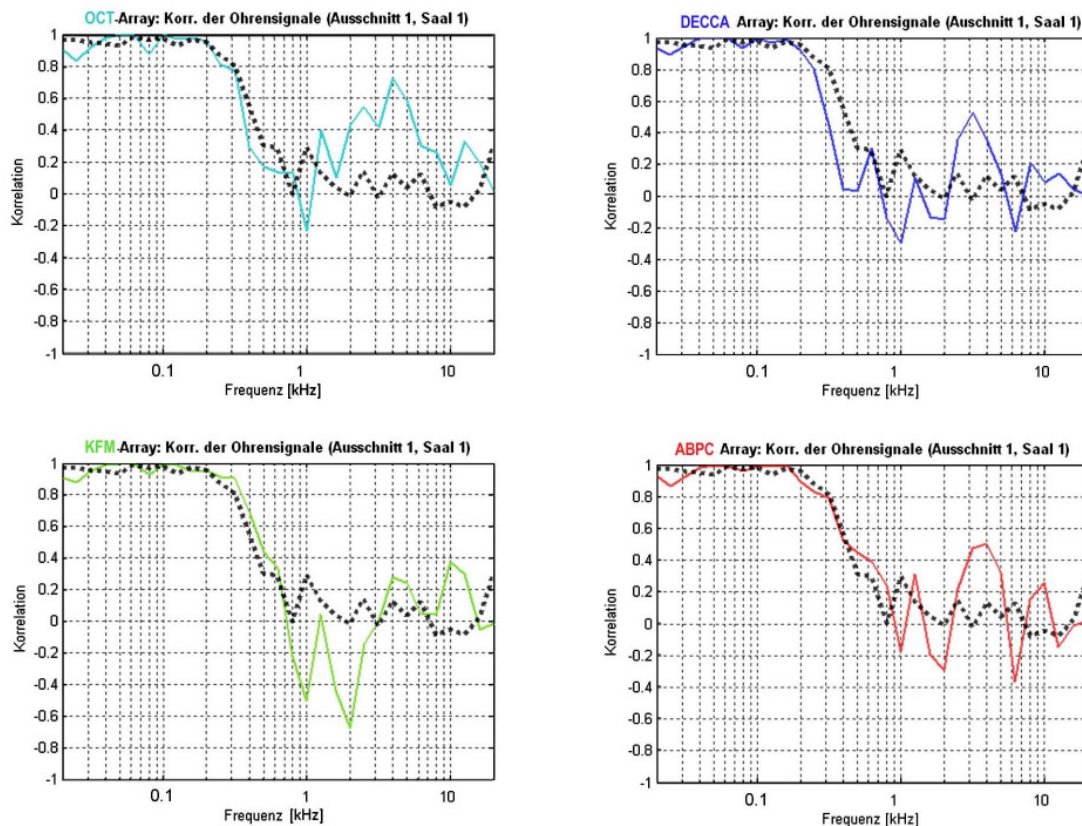


Abb. 10.6: FIACC des Mikrofon Arrays (durchgehende Linie), sowie des KU81 (gepunktete Linie); Musikausschnitt 1 (7sec, forte), Saal 1

Betreffend Abb.10.6 fällt auf, dass bei jenen Mikrofontechniken, die auch omnidirektionale Mikrofone verwenden (d.h. DECCA, AB-PC) eine moderate Korrelation von max. +0.5 im Frequenzbereich von ca. 2-3kHz auftritt. Wie bereits von [Hiyama et al., 2002] gezeigt, ist dies eine Konsequenz der standardisierten 5.1 Lautsprecheraufstellung, in Zusammenwirken mit den HRTF's eines (menschlichen oder Kunst-) Kopfs, wenn ausreichend dekorrelierte Signale über die Lautsprecher abgestrahlt werden.

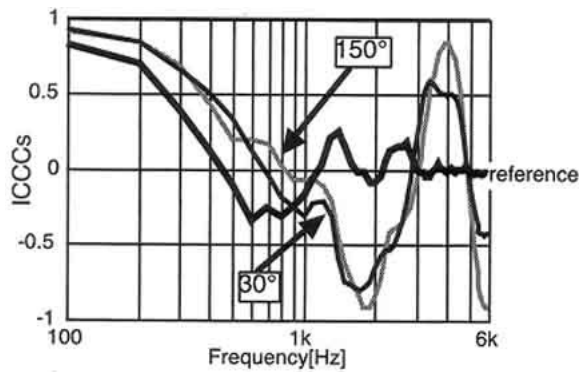


Abb. 10.7: ICCCs (=FIACCs) für die symmetrische Anordnung eines Lautsprechers bei den angegebenen Winkelpositionen (d.h. bei $\pm 30^\circ$ und $\pm 150^\circ$); (aus [Hiyama et al., 2002])

(Anm.: ICCCs = Interaural Cross-Correlation Coefficient)

Weitere Belege hierzu, sowie eine Berechnung des sich daraus ergebenden Frequenzgangs, basierend auf der Simulation eines 3-dimensionalen Kopfmodells, finden sich in [Shirley et al., 2007].

Im Vergleich zu den Systemen DECCA, AB-PC ist das OCT System in Abb. 10.6 durch eine wesentlich höhere Korrelation (max. +0.7) im gleichen Frequenzbereich gekennzeichnet, während hingegen das KFM System eine *negative* (d.h. gegenphasige) Korrelation (max. ca. -0.7) im Frequenzbereich von 700Hz bis 3kHz aufweist.

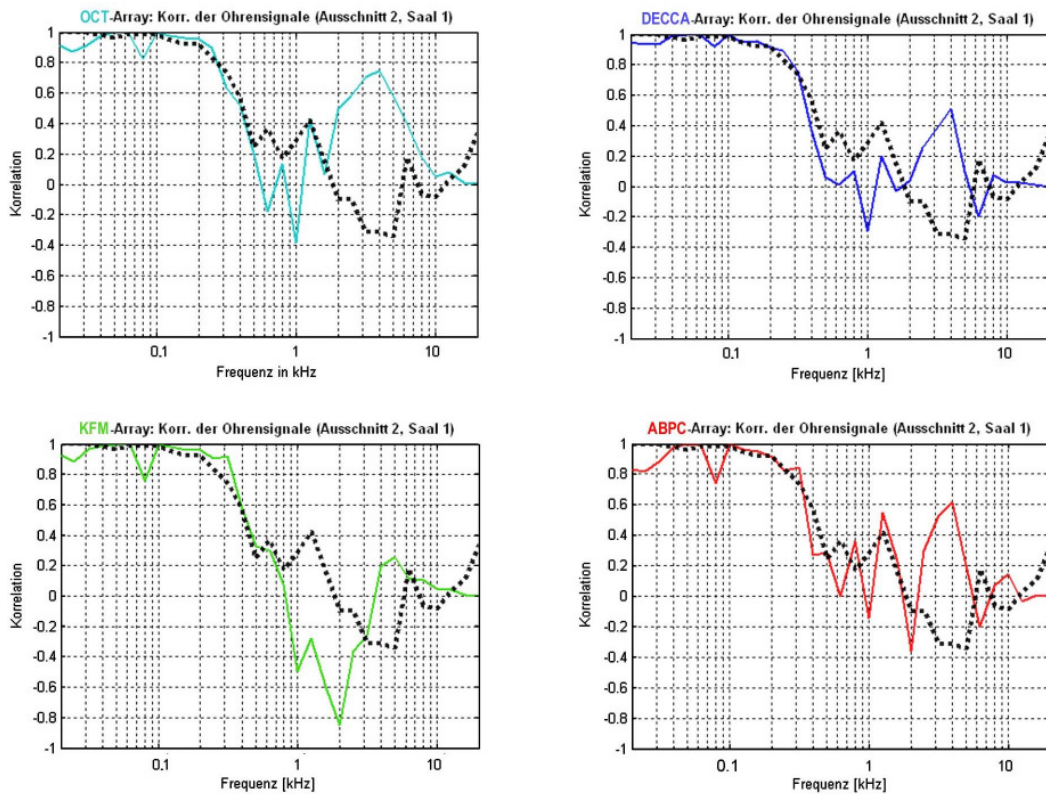


Abb. 10.8: FIACC des Mikrofon Arrays (durchgehende Linie), sowie des KU81 (gepunktete Linie); Musikausschnitt 2 (16sec, piano), Saal 1

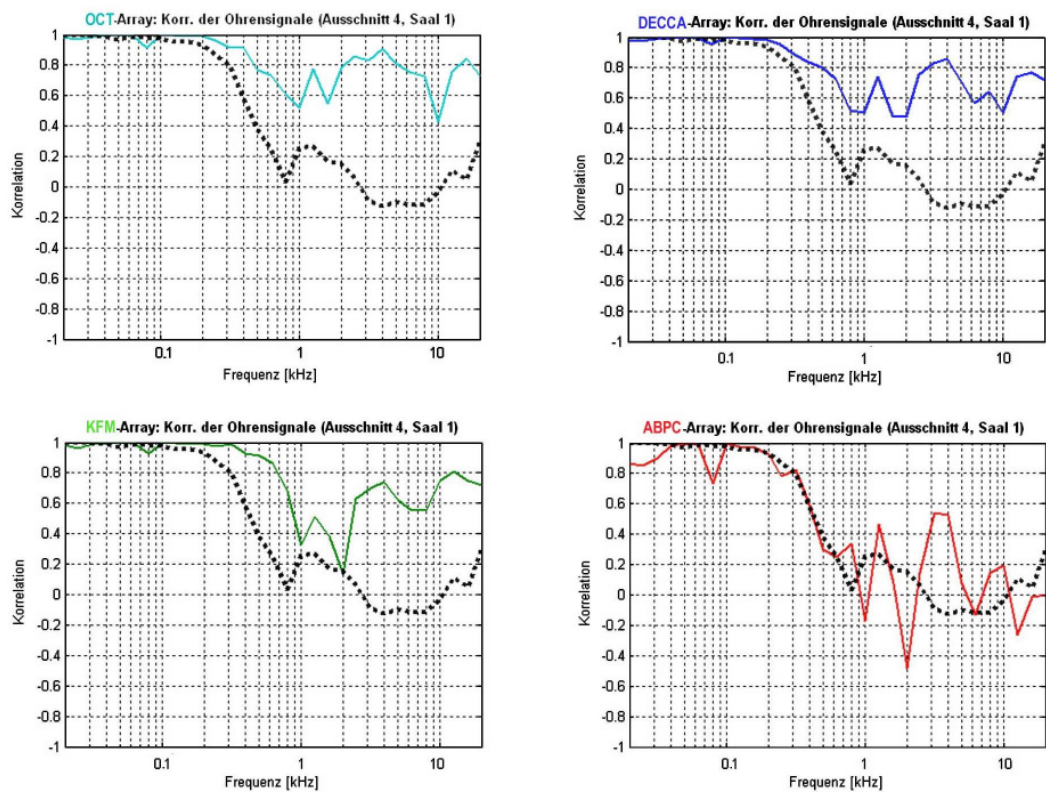


Abb. 10.9: FIACC des Mikrofon Arrays (durchgehende Linie), sowie des KU81 (gepunktete Linie); Musikausschnitt 4 (60sec, forte + piano), Saal 1

Bei Analyse des 16sec Ausschnitts (Sample #2, piano, Abb. 10.8) fällt auf, dass die Überhöhung der Korrelation bei OCT nun noch breiter, und auch beim KFM System die negative Korrelation noch stärker geworden ist (diese reicht nun von 800Hz bis 3.5kHz, mit einem negativen Maximum kleiner als -0.8).

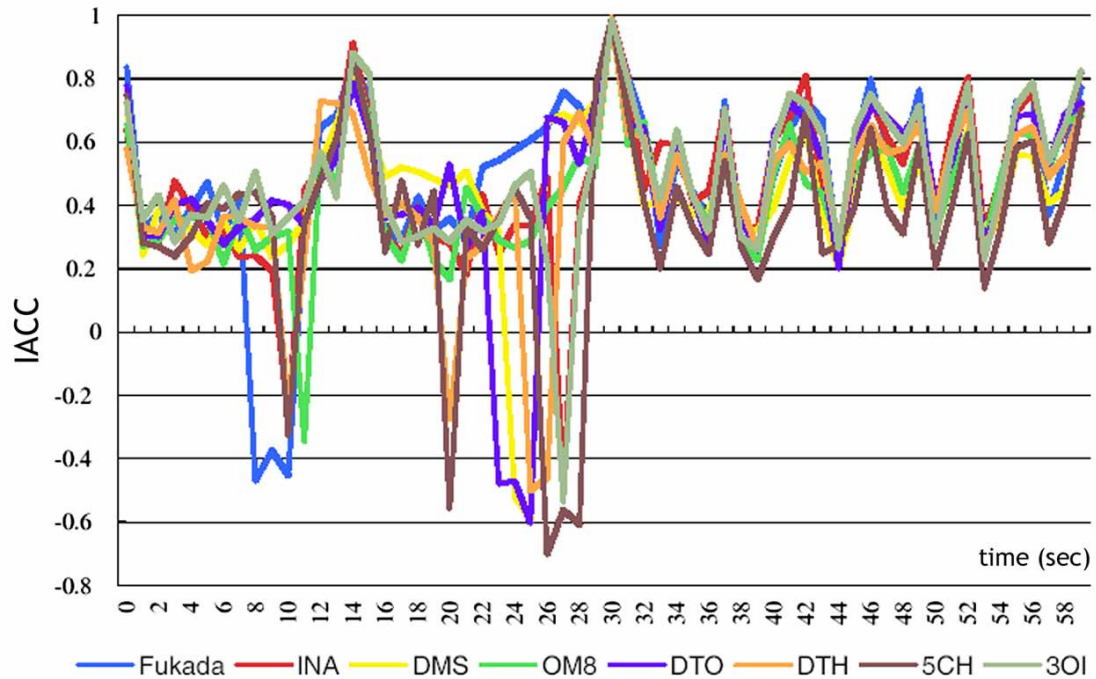
In Abb. 10.9 sind die Ergebnisse des mit 60sec wesentlich längeren Ausschnitts (Sample #4), welcher sowohl die forte-Passage (Sample #1), als auch die piano-Passage (Sample #2) enthält, dargestellt:

Es zeigt sich, dass sich bei den Systemen OCT, DECCA und KFM die Korrelation für den gesamten Frequenzbereich erhöht hat, während das AB-PC System das einzige ist, bei dem die frequenzabhängige Korrelations-Funktion (auch unterhalb von 700Hz) weiterhin nahe der Referenz (der Kunstkopfaufnahme aus dem Konzertsaal) bleibt.

Weiters fällt auf, dass das AB-PC System das einzige ist, welches der KU81-Referenzfunktion auch unterhalb von ca. 700Hz genau folgt. Der tieffrequente Bereich unterhalb 500Hz ist wichtig für das Entstehen des Raumeindrucks beim Hörer.

Gemäß den vorliegenden Messungen der FIACCs aus Abb. 10.9 ist also das AB-PC System jenes, welches die Korrelationscharakteristik des Original-Schallfelds aus dem Konzertsaal am besten nachbildet.

In diesem Zusammenhang sei auch auf eine Untersuchung von [Kamekawa et al., 2007] verwiesen, in deren Rahmen u. a. auch die zeitliche Abhängigkeit des interauralen Korrelationsgrades bei Wiedergabe von Surround-Aufnahmen, die mittels unterschiedlicher Mikrofonsysteme erstellt worden waren, gezeigt werden konnte. In der untenstehenden Abbildung 10.10 ist der IACC von insgesamt 8 unterschiedlichen Mikrofonsystemen über einen Zeitabschnitt von 60sec dargestellt. Man beachte, dass der sich am Kunstkopf ergebende IACC – in Abhängigkeit von der temporären Beschaffenheit des Musiksignals – für verschiedene Mikrofonsysteme äußerst unterschiedlich ausfallen kann (siehe hierzu z.B. die Systeme 5CH, DTO, DMS, DTH und 3OI im Abschnitt 20-30sec).



DMS, DTH und 3OI im Abschnitt 20-30sec)

Abb. 10.10: IACC Werte für einen 60-sekündigen Auschnitt aus dem Orchesterwerk ‚Wellington’s Sieg‘ (aus [Kamekawa et al., 2007])

(Erklärung einiger Abkürzungen: DMS = double MS, OM8= omni+8, DTO = Decca-Tree + Omni-square, DTH = Decca-Tree+Hamsaki-square, 5CH = 5 cardioids + Hamsaki-square, 3OI = 3 Omnis + IRT-cross)

Betrachtet man die zeitweise auftretende hohe negative interaurale Korrelation bei einigen der verwendeten Surround-Mikrofontechniken in Abbildung 10.10, so ist dies – gemeinsam mit den in den Abbildungen 10.6, 10.7 und 10.9 gefundenen Signalverläufen - ein weiteres Indiz dafür, dass es sinnvoll ist, auch bei einer frequenzabhängigen Korrelationsanalyse über einen längeren Zeitabschnitt (z.B. 60 sec) zu mitteln, wie dies auch für Ausschnitt #4 (Abb. 10.9) durchgeführt wurde.

In der untenstehenden Tab. 10.2 sind die numerischen Werte der Korrelation für die vier Surround-Mikrofontechniken und drei Ausschnitte dargestellt: das AB-PC System ist normalerweise der Referenz (Kunstkopf KU81) am nächsten, und zwar nicht nur betreffend die Korrelation des Gesamtsignals, sondern auch betreffend den separat analysierten LF und HF-Bereich. Wie auch im Vergleich mit den Korrelationswerten für das Gesamtsignal gesehen werden kann, sind die FIACC-

Werte beim Ausschnitt #1 (forte) geringer, als bei Ausschnitt #2 (piano), was vermutlich durch den höheren Anteil an Diffusschall bedingt ist, der im Konzertsaal bei Anregung durch eine lautere Schallquelle hervorgerufen wird.

Aufgrund der vorliegenden Ergebnisse darf die Feststellung gemacht werden, dass das AB-PC System nicht durch generell höhere oder tiefere Korrelationwerte ‚per se‘ gekennzeichnet ist, sondern es offenbar einfach besser in der Lage ist, die im Konzertsaal auftretenden Schallfelder (bzw. deren indirekte Repräsentation mittels der messtechnischen Erfassung des FIACC am KU81 Kunstkopf) zu replizieren.

| Corr.- Coeff.: | LF(400Hz) | HF(1200Hz) | ALL |
|--|-------------|-------------|-------------|
| ORCH-Ausschnitt 1: (forte, 7sec) | | | |
| OCT | 0.87 | 0.28 | 0.41 |
| DECCA | 0.85 | -0.01 | 0.23 |
| KFM | 0.88 | -0.30 | 0.39 |
| AB-PC | 0.89 | 0.15 | 0.48 |
| KU81 | 0.92 | 0.11 | 0.55 |
| ORCH-Ausschnitt 2: (piano, 16sec) | | | |
| OCT | 0.94 | 0.36 | 0.85 |
| DECCA | 0.96 | 0.11 | 0.89 |
| KFM | 0.98 | -0.46 | 0.88 |
| AB-PC | 0.92 | 0.37 | 0.78 |
| KU81 | 0.85 | 0.24 | 0.74 |
| ORCH-Ausschnitt 4: (piano + forte, 60sec) | | | |
| OCT | 0.98 | 0.71 | 0.88 |
| DECCA | 0.98 | 0.66 | 0.90 |
| KFM | 0.99 | 0.44 | 0.87 |
| AB-PC | 0.88 | 0.27 | 0.54 |
| KU81 | 0.87 | 0.20 | 0.57 |

Tab. 10.3: Korrelations-Koeffizienten für drei Musikausschnitte und vier Surround-Mikrofonsysteme, sowie Kunstkopf (Referenz)

Abschließend sollen noch die FIACCs der für die DUO-Aufnahme verwendeten 4 Surroundmikrofonarrays gezeigt werden:

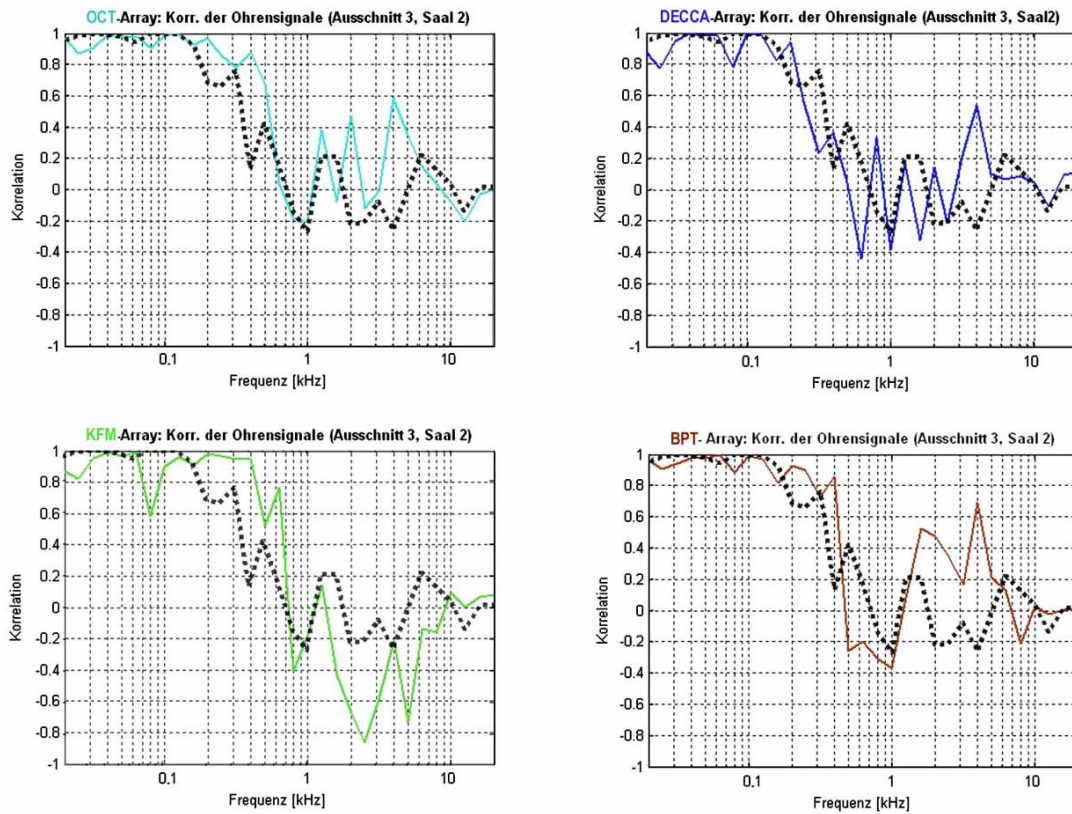


Abb. 10.11: FIACC des Mikrofon Arrays (durchgehende Linie), sowie des KU81 (gepunktete Linie); Musikausschnitt 3 (16sec), Saal 2 (DUO-Aufnahme: Mic-Arrays – OCT, DECCA, KFM, BPT)

Auch bei Musikausschnitt 3, der Aufnahme eines Duos Cello-Klavier in einem akustisch ganz anders beschaffenen Raum, lässt sich beim KFM-Array eine deutliche Gegenphasigkeit im mittleren bis hohen Frequenzbereich feststellen, wie dies auch bereits bei der Orchesteraufnahme der Fall war (vgl. Abb. 10.8 links unten).

Die Binauralsignale der anderen drei Surroundmikrofon-Arrays OCT, DECCA, BPT weisen hingegen nur in vergleichsweise schmalbandigen Frequenzbereichen geringe Gegenphasigkeiten bis zu einem Maximal-Wert ca. -0.4 auf und entsprechen der Referenzfunktion des KU81 generell besser.

10.5 Messung der frequenzabhängigen Kohärenz

Einige Forscher sind davon überzeugt, dass vorzugsweise die Kohärenz (statt der Korrelation) zweier Signale ermittelt werden sollte (siehe u. a. [Martin, 2005], [Jacobsen and Roisain, 2000]). Die grundlegenden Unterschiede zwischen Kohärenz und Korrelation wurden bereits in Abschnitt 3 kurz dargestellt.

Für die nachfolgend dokumentierten Kohärenzmessungen wurde die ‚mscohere‘-Funktion in MATLAB verwendet, welche die ‚magnitude squared‘ Kohärenz-Abschätzung der Eingangssignale errechnet. Da dieser Signalprocessing-Algorithmus auf der diskreten Fouriertransformation (DFT) basiert, muss eine ausreichende DFT-Länge gewählt werden, um die gewünschte Auflösung im Frequenzbereich realisieren zu können. Bei einer Sampling-Rate von 44.1kHz wurde eine 4096-Punkt DFT Auflösung gewählt (sofern nicht anders angegeben), um einen Frequenzraster mit ca. 10Hz Abstand zu erhalten und somit auch für den Tieffrequenzbereich noch eine hinlängliche Anzahl von Berechnungspunkten zu gewährleisten.

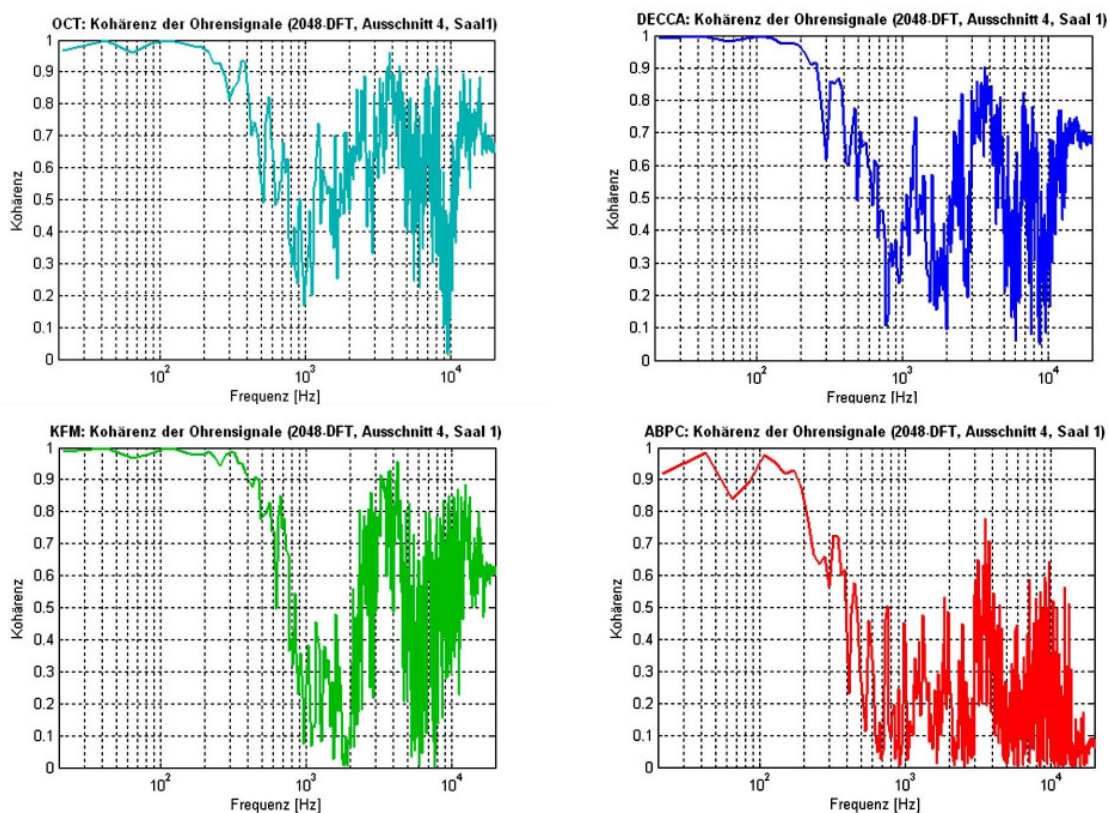


Abb. 10.12: Kohärenz über der Frequenz, für die Binauralsignale von 4 Surroundmikrofonarrays; Musikausschnitt 4 (60sec, forte + piano), Saal 1

Während im Kurvenverlauf bei fast allen Mikrofonssystemen (außer AB-PC) eine starke Überhöhung im Frequenzbereich von 3-4kHz zu sehen ist, ist beim KFM-System zusätzlich der Einbruch im Bereich von 1-2kHz zu erkennen. Dieser ist zweifelsohne durch die in diesem Bereich auftretende Phasenänderung in Richtung negativer Korrelation (sichtbar bei der Analyse kürzerer Signalausschnitte; siehe Sample #1 und #2) bedingt. Ein starke Gegenphasigkeit in einem bestimmten Frequenzbereich, wie sie z.B. beim KFM-Array bei Sample #2 auftritt (siehe Abb. 10.8 links unten) wäre anhand der Kohärenzfunktion des gleichen Ausschnitts nicht festzustellen.

Der besseren Übersichtlichkeit halber und damit der Verlauf der Kohärenzfunktion besser mit dem Verlauf der Korrelationsfunktion verglichen werden kann, wurden für Abb. 10.13 die aus der Kohärenz-Abschätzung gewonnenen Daten in Frequenzbändern mit 1/3-Oktav Bandbreite (mit Mittenfrequenzen nach ISO-Norm) zusammengefasst. Auch hier fällt auf, dass der Funktionsverlauf beim ABPC-Array der Referenzfunktion (KU81) am nächsten kommt, vor allem auch in dem – für die Räumlichkeit so wichtigen - tieffrequenten Bereich unter 1kHz.

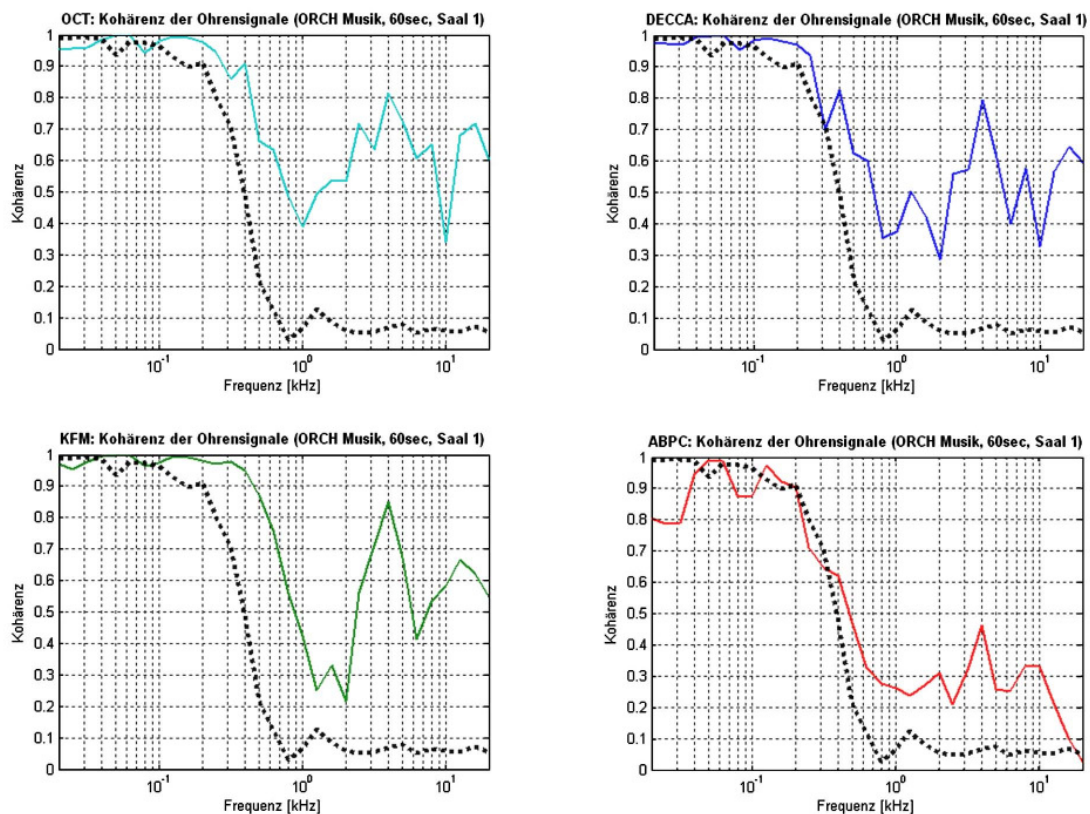


Abb. 10.13: Kohärenz (über der Frequenz) der Binauralsignale von 4 Surroundmikrofonen (durchgezogene Linie), sowie der originalen KU81 Kunstkopfaufnahme aus dem Konzertsaal (unterbrochene Linie); Auflösung 31-Band, 1/3 Oktav

Aus den oben angestellten Betrachtungen darf geschlossen werden, dass (zumindest für den gegebenen Anwendungsfall) die Bestimmung der Korrelation der beiden Signale eine stärkere Aussagekraft hat, als deren Kohärenz zu ermitteln, da bei dieser die konkrete Phasenbeziehung der beiden Signale keine Berücksichtigung findet. So sind z.B. auch die stark gegenphasigen Signalanteile beim KFM-System in einem Frequenzbereich (1-3kHz), in welchem das menschliche Gehör eine erhöhte Empfindlichkeit aufweist, höchstwahrscheinlich für die schlechte Bewertung hinsichtlich der Klangfarbe verantwortlich, die das System im Rahmen des Hörvergleichs erhalten hat. Eine genauere Betrachtung zur Problematik der Klangfärbung aufgrund von Kammfilter-Effekten bei der akustischen Summierung von Lautsprecher-Signalen findet sich in einem kürzlich erschienenen Paper zum Thema Wellenfeldsynthese [Wittek et al., 2007].

10.6 Paarweise Korrelation und Kohärenz

Auch wenn die Korrelations- und Kohärenz-Analyse der binauralen Kunstkopfsignale bereits sehr aussagekräftige Ergebnisse geliefert hat, welche es ermöglichen, die untersuchten Surround-Mikrofontechniken unterscheiden zu können, so scheint es doch wert, auch das Verhältnis zwischen den Surroundkanälen paarweise zu untersuchen, um ein besseres Verständnis über die genauen Eigenschaften der entstehenden Schallfelder bei der Surround-Wiedergabe der Aufnahmen der untersuchten Mikrofonssysteme zu erhalten.

Betrachtet man die Kohärenzen zwischen den Kanälen des OCT Systems in Abb. 10.14, so ist zu erkennen, dass die Kohärenz zwischen den Frontkanälen L und R über den gesamten Frequenzbereich eher gering ist, wie es für zwei in entgegengesetzte Richtungen weisende Hypernieren zu erwarten ist. Die Kohärenz zwischen dem L (bzw. R) Front-Kanal und dem Center-Kanal ist unterhalb 500Hz relativ hoch, mit Spitzenwerten, die bis 0.75 (bzw. R: über 0.9) reichen. Bei der Kohärenzdarstellung, die alle berechneten Einzelwerte umfasst, sind im Bereich um 10kHz auch schmalbandige Kohärenz-Spitzen bis zu einem Wert von max. 0.8 sichtbar. Die Kohärenz der Kanäle LS, RS steigt unter 200Hz stark an, mit Werten bis 0.8. Für die Kohärenz der Signale vorne-hinten (L, LS und R, RS) gilt ähnliches, mit einem Anstieg der Funktion bis zu Spitzenwerten um 0.6.

Die abweichenden Funktionsverläufe bei den Signalpaarungen L/C und C/R, sowie L/LS und R/RS sind vermutlich auf eine – unbeabsichtigt - nicht ganz symmetrische Ausrichtung der Mikrofone, sowie ev. auch auf individuelle (frequenzabhängige) Abweichungen von der gewünschten Richtcharakteristik zurückzuführen.

Ein ausgewogeneres Bild, welches auch leichter mit den Korrelationsmessungen vergleichbar ist, ergibt sich bei Betrachtung der ‚geglätteten‘ Darstellung mit Aufteilung auf 1/3-Oktav Bänder in Abb. 10.15, auf der übernächsten Seite.

OCT-System

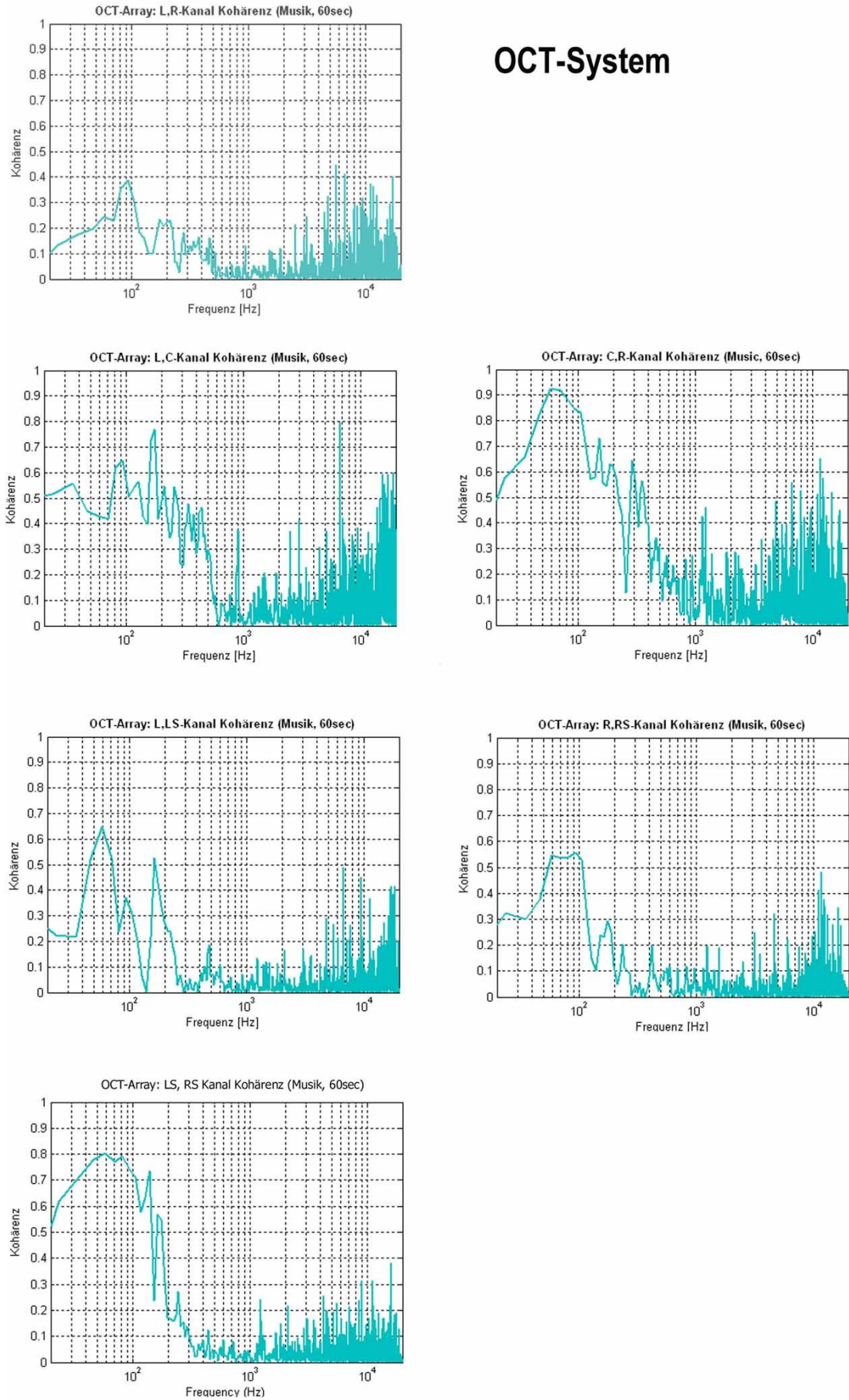


Abb. 10.14: OCT-Array, Kohärenz für je zwei Kanäle des Surround-Signals

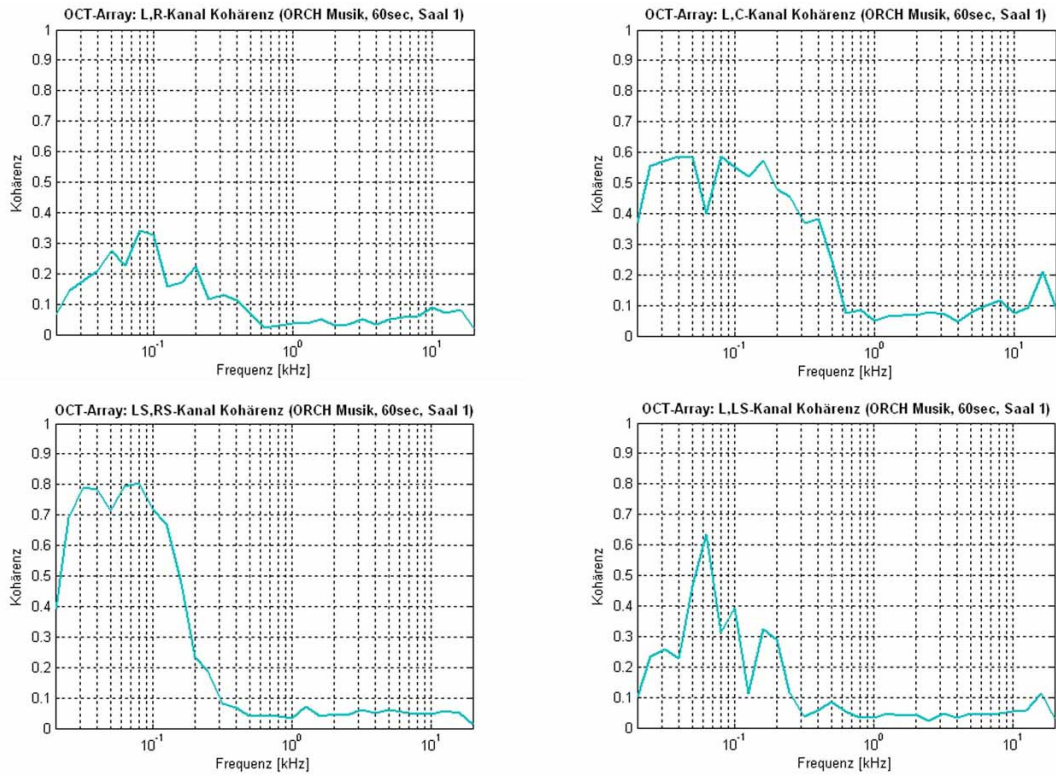


Abb. 10.15: OCT-Array, Kohärenz von je zwei Surroundkanälen (Auflösung: 1/3-Oktavbandbreite)

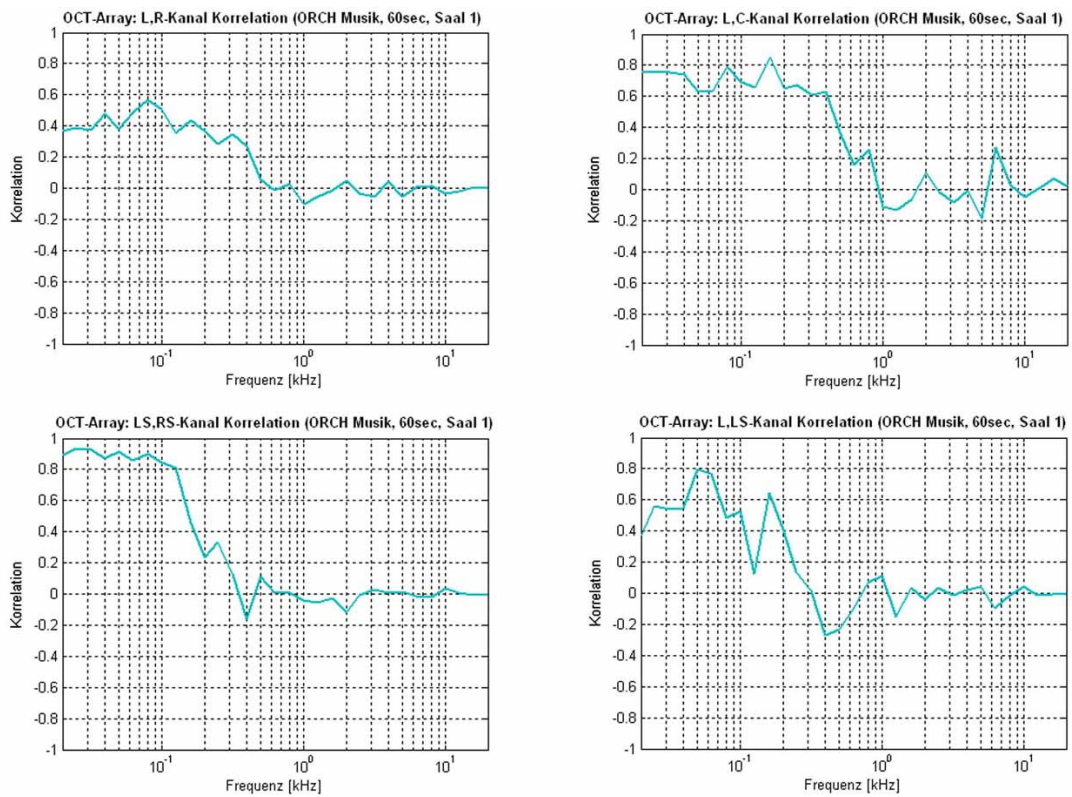


Abb. 10.16: OCT-Array, Korrelation von je zwei Surroundkanälen (Auflösung: 1/3-Oktavbandbreite)

Vergleicht man die in Abbildung 10.15 dargestellte Kohärenzfunktionen mit den Korrelationsfunktionen der Abb. 10.16, so ergeben sich daraus ähnliche Schlussfolgerungen.

Anhand der Korrelationsfunktion ist zusätzlich folgendes zu erkennen:

Die Signalinhalte des L und C-Kanals sind oberhalb von 1kHz stark dekorreliert; für die Signale LS/RS gilt dies bereits ab ca. 300Hz aufwärts; die Korrelation der benachbarten Kanäle vorne/hinten (L/LS) sind ebenso oberhalb von 300Hz dekorreliert; zwischen 400-500Hz tritt dabei eine leichte Gegenphasigkeit auf (mit ca. - 0.2 etwa in der gleichen Größenordnung, wie die maximal auftretende negative Korrelation von zwei omnidirektionalen AB-Mikrofonen im diffusen Schallfeld).

Für die anderen drei Surround-Mikrofonarrays (DECCA, KFM, ABPC) wurden ebenfalls die Kohärenzfunktionen aller Nachbarkanäle untersucht, allerdings traten hierbei keine relevanten Asymmetrien (wie beim OCT-Array) hinsichtlich L und R auf, weshalb auf eine Abbildung derselben verzichtet werden soll.

Weiters wollen wir uns auf die Darstellung der Funktionen mit einer Auflösung in 1/3-Oktavbandbreite beschränken:

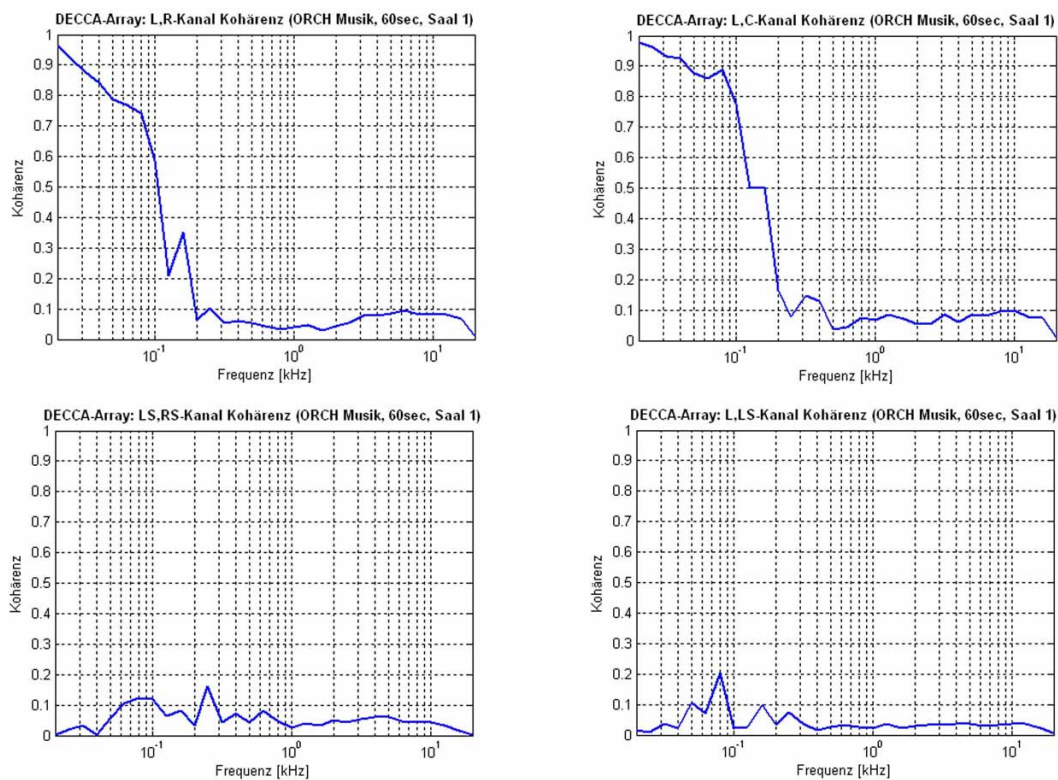


Abb. 10.17: DECCA-Array, Kohärenz von je zwei Surroundkanälen (Auflösung: 1/3-Oktavbandbreite)

Erwartungsgemäß ist beim DECCA-System die Kohärenz bei tiefen Frequenzen wesentlich höher, zumindest für die drei Frontkanäle, die aus Mikrofonen mit omnidirektionaler Charakteristik bestehen. Auch bei diesem System gibt es einen steilen Anstieg der Kohärenz unterhalb von 200Hz (Anm.: der L/R Kapselabstand war ähnlich groß wie beim OCT-System), die mit fallender Frequenz gegen 1 konvergiert. Aufgrund des großen Abstands (ca. 10m) der Rear-Mikrofone zueinander, als auch zum Frontsystem, gibt es so gut wie keine Kohärenz ihrer Signale zueinander, wie auch zur Front. Dies ist für die Ausbildung eines guten Raumeindrucks - zumindest zwischen Front und Rear - vorteilhaft.

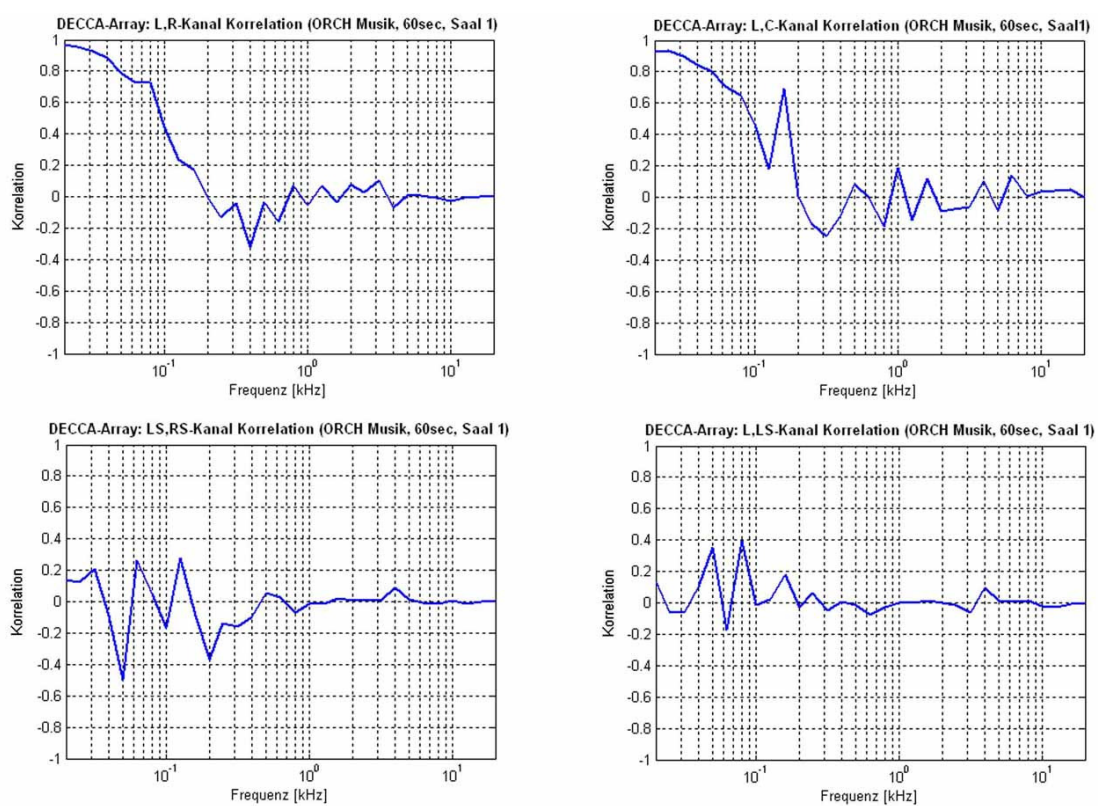


Abb. 10.18: DECCA-Array, Korrelation für je zwei Surround-Kanäle

Die Schlussfolgerungen, die sich aus der Betrachtung der Korrelationsfunktionen für das DECCA-Array in Abbildung 10.18 ergeben, sind mit den aus der Kohärenzfunktion gewonnenen konform: unterhalb von 200Hz ein Anstieg der Korrelation innerhalb der Signale des Front-Systems, ansonsten dekorreliert.

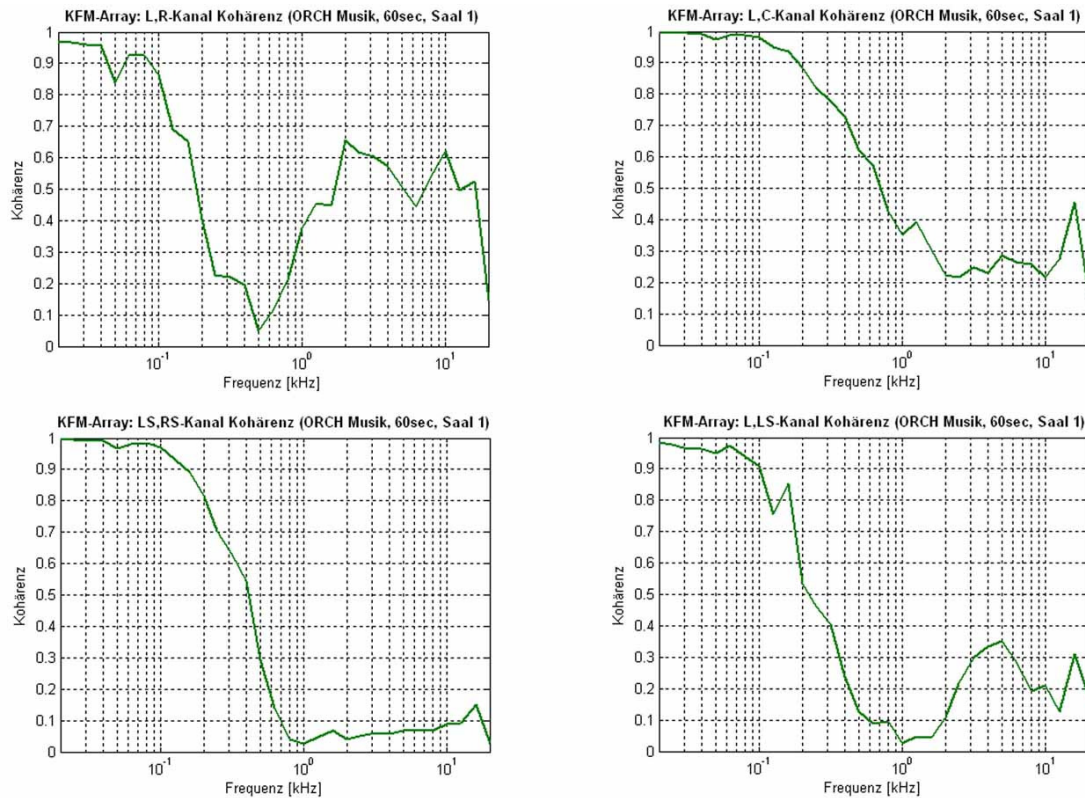


Abb. 10.19: KFM-Array, Kohärenz für je zwei Surround-Kanäle

Das KFM-System in Abb. 10.19 ist hingegen durch wesentlich höhere Kohärenz zwischen allen Kanälen gekennzeichnet; hauptsächlich bei tiefen Frequenzen, aber nicht nur: die L/R-Kohärenz sinkt bei 500Hz bis auf 0 ab und steigt dann steilflankig zu hohen Frequenzen wieder an, wobei in einzelnen Frequenzbändern Maximalwerte bis 0.65 auftreten. Die Kohärenz zwischen L und C-Kanal ist - im Vergleich mit den anderen Systemen - im Durchschnitt deutlich höher, was auch für höhere Frequenzen gilt.

Betreffend die Rear-Kanäle LS/RS ist hohe Kohärenz auf das Tieffrequenzband bis etwa 300Hz hinauf reduziert. Die Kohärenz zwischen L und LS ist für tiefe Frequenzen unterhalb von 200Hz ebenfalls sehr hoch, fällt im Bereich von ca. 400Hz bis 2kHz auf 0 ab und steigt danach zu den hohen Frequenzen hin wieder an.

In der untenstehenden Abbildung 10.20 sind zum Vergleich die Korrelationsfunktionen der gleichen Kanalpaarungen zu sehen.

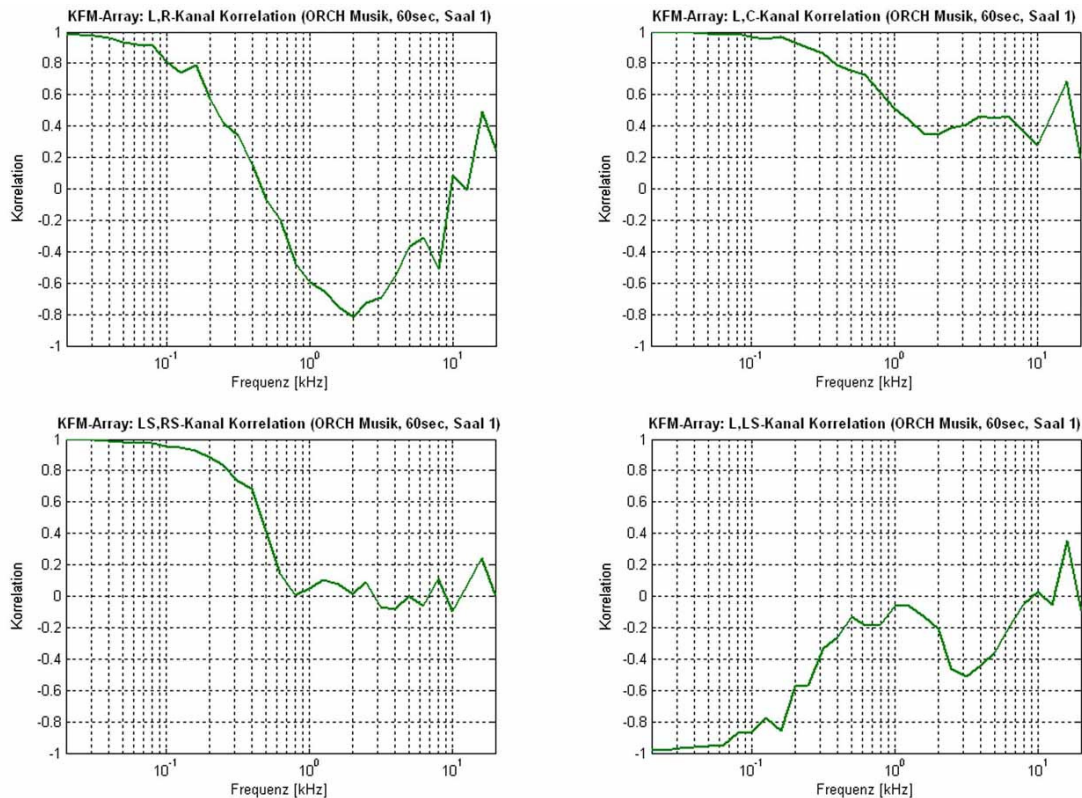


Abb. 10.20: KFM-Array, Korrelation für je zwei Surround-Kanäle

Bei Betrachtung der Korrelationsfunktionen des KFM-Arrays fällt auf, dass hier bei einigen Kanalkombinationen breitbandig Gegenphasigkeiten auftreten: im Frontkanal-Paar L/R tritt im Bereich von 500Hz-9kHz breitbandigst Gegenphasigkeit auf, ebenso zwischen den benachbarten Front- und Rear Kanälen L/LS (praktisch von 0Hz bis 10kHz, wobei bei ca. 1kHz die Signale – mit einem Wert nahe 0 – praktisch dekorreliert sind).

Die MS-Dematrixierung der mittels des KFM-Kugelflächenmikrofons + 2 seitlichen Kapseln mit 8-er Charakteristik gewonnenen Signale erfolgte unter Verwendung des zugehörigen DSP-4 Signalprozessors der Fa. Schoeps. Es ist nicht auszuschließen, dass die - wie obenstehend in Abb.10.20 ersichtlich – signaltechnisch ungünstige Phasenbeziehung auf eine nicht-optimalen Programmierung des Dematrixierungsalgorithmus zurückzuführen ist.

Auf jeden Fall erklärt das Resultat der Analyse der Korrelationen der Einzelkanäle, wieso es auch für das am Kunstkopf auftretenden Binauralsignal (Re-Recording) zu breitbandigen Gegenphasigkeiten kommen muss (vgl. Abb. 10.8).

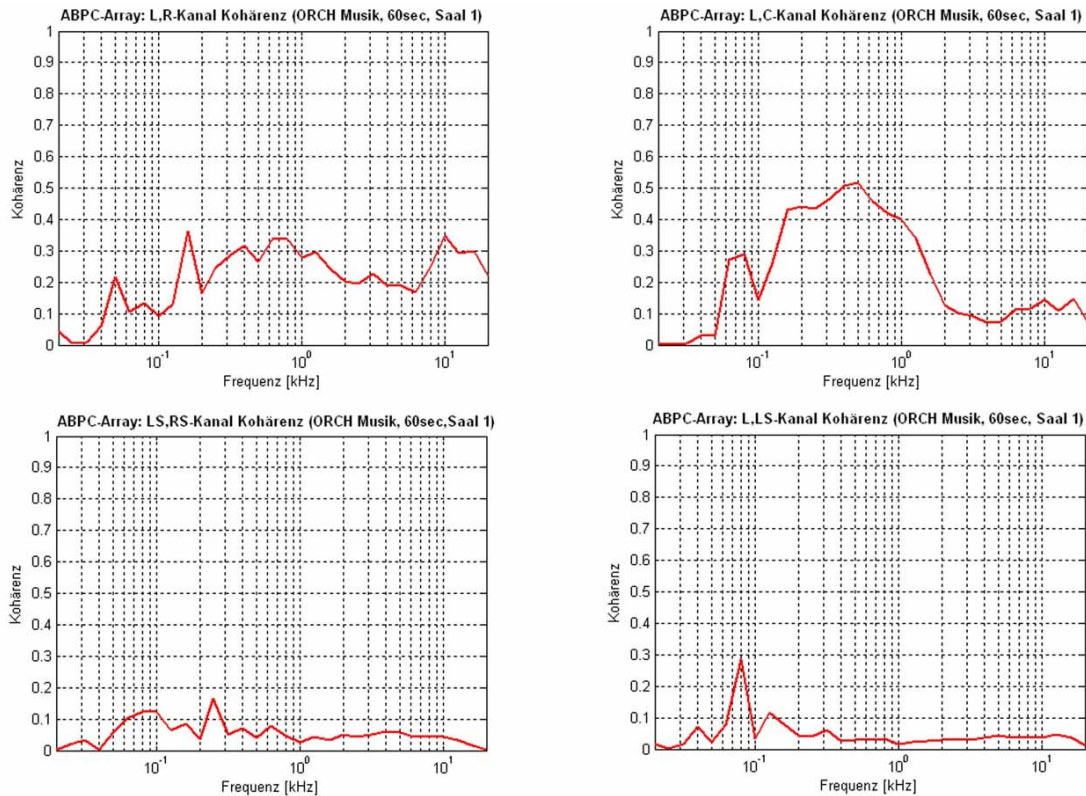


Abb. 10.21: ABPC-Array, Kohärenz für je zwei Surround-Kanäle

Beim AB-PC System tritt vergleichsweise ein gänzlich anderes Signalverhalten auf: wie in Abb. 10.21 zu sehen, ist die Kohärenz auch bei tiefen Frequenzen äußerst gering. Beim L/R Kanal-Paar besteht ein leichter Anstieg zu den hohen Frequenzen hin, wobei die Kohärenz bis ca. 10kHz stets unter 0.4 bleibt. Auch zwischen dem L und C-Kanal ist die Kohärenz generell niedrig: sie erreicht ein Maximum von ca. 0.5 im Bereich von 400-500Hz, und fällt danach zu höheren Frequenzen hin wieder ab. Die Signale der Rear-Kanäle und zwischen den Front- und Rear-Kanälen sind - abgesehen von einem schmalbandigen ‚Spike‘ - nahezu komplett inkohärent: die weitgehende Absenz von Kohärenz (und somit hohe Dekorrelation) bei tiefen Frequenzen bis ca. 400Hz hinauf ist auch der Grund für den guten Raumeindruck, den dieses Mikrofonsystem beim Hörvergleich hervorzurufen im Stande war.

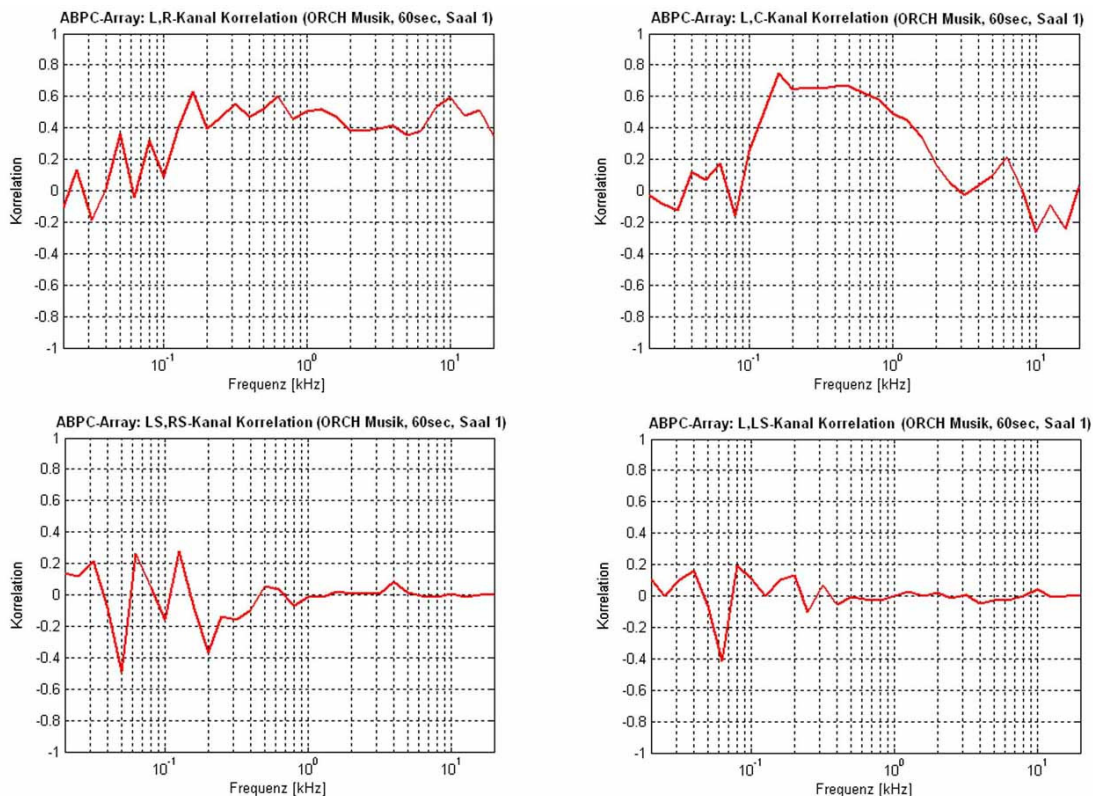


Abb. 10.22: ABPC-Array, Korrelation für je zwei Surround-Kanäle

Die Analyse der Korrelation der Surroundkanäle des ABPC-Systems in Abb. 10.22 liefert ein ähnliches Bild: unterhalb von 200Hz sinkt die L/R Korrelation mit fallender Frequenz gegen den Wert 0, darüber bewegt sie sich zwischen 0.4 und 0.6. Bei der Kombination L/C tritt im Bereich zwischen ca. 200Hz und 800Hz hohe Korrelation mit einem Wert von ca. 0.6 auf, oberhalb von 2kHz ist das Signal weitgehend dekorreliert. Die Signale der Rear-Mikrofone und der Kombination L/LS sind weitestgehend – bis auf schmalbandige Spikes im tieffrequenten Bereich unterhalb 200Hz - dekorreliert.

In der nachfolgenden Tab. 10.2 sind die numerischen Werte der Kohärenz zwischen den Surround-Kanälen aufgelistet, sowie auch - nur für die Orchesteraufnahme - auch die Werte für die Kohärenz zwischen den L- und R-Ohrsignalen der KU81 Kunstkopfaufnahmen. So wie bereits bei den numerischen Werten betreffend Korrelation, zeigt sich auch bei Auswertung der Kohärenz, dass das AB-PC System der Referenzaufnahme (KU81) aus dem Konzertsaal am nächsten kommt.

| ORCH - | Binaural | Front | | | | | Rear |
|---------------|-----------------|------------|------------|------------|-------------|-------------|--------------|
| Rec. | (Dummy) | L-R | L-C | C-R | L-LS | R-RS | LS-RS |
| OCT | 0.74 | 0.09 | 0.24 | 0.31 | 0.13 | 0.13 | 0.20 |
| DECCA | 0.68 | 0.19 | 0.24 | 0.24 | 0.03 | 0.02 | 0.05 |
| KFM | 0.70 | 0.49 | 0.55 | 0.45 | 0.37 | 0.32 | 0.38 |
| AB-PC | 0.44 | 0.21 | 0.24 | 0.25 | 0.03 | 0.03 | 0.04 |
| KU81 | 0.37 | | | | | | |
| | | | | | | | |
| DUO- | Binaural | Front | | | | | Rear |
| Rec. | (Dummy) | L-R | L-C | C-R | L-LS | R-RS | LS-RS |
| OCT | 0.56 | 0.16 | 0.24 | 0.36 | 0.21 | 0.19 | 0.22 |
| DECCA | 0.50 | 0.18 | 0.28 | 0.29 | 0.08 | 0.10 | 0.10 |
| KFM | 0.62 | 0.54 | 0.60 | 0.53 | 0.36 | 0.37 | 0.38 |
| BPT | 0.54 | 0.24 | 0.42 | 0.44 | 0.08 | 0.09 | 0.10 |
| KU81 | 0.48 | | | | | | |

Tab. 10.2: Kohärenz der Signale für jeweils 4 Surround-Mikrofonsysteme (ORCH und DUO-Aufnahme), sowie der Binauralsignale (nur ORCH) [Samplelänge: 60sec]

Bei Betrachtung der in Tab. 10.2 (ORCH-Rec.), angegebenen Kohärenzwerte für die verschiedenen Surround-Mikrofonarrays erkennt man, dass das KFM-Array erwartungsgemäß auch numerisch die höchsten Werte aufweist. Auffallend ist jedoch, dass die Kohärenz des Binauralsignals des OCT-Arrays trotzdem einen marginal höheren Wert hat, obwohl die Signale seiner Surroundkanäle weit geringere Kohärenz aufweisen. Dass dies möglich ist, liegt wohl an dem beim KFM-Array vorliegenden ‚Einbruch‘ in der Kohärenz (beim Wechsel auf Gegenphasigkeit) im mittleren Frequenzbereich, sowie auch an der hohen Kohärenz der OCT-Surroundsignale im tieffrequenten Bereich .

Wie man sehen kann, ist allerdings auch beim numerischen Vergleich das ABPC-Array mit einem Wert von 0.44 dem Binauralen Referenzsignal des KU81 von 0.37 am nächsten.

Zum Vergleich sind unmittelbar unter den Messwerten für die Surround Orchester-Aufnahme auch die Werte für die Surround Duo-Aufnahme aufgelistet: man sieht, dass sich die Kohärenzwerte – obwohl die Duo-Aufnahme ja aus einem akustisch komplett anders beschaffenen Saal stammt – in Übereinstimmung mit den bereits für die Orchesteraufnahme gefundenen Werten stehen. [Ausschnittlänge: jeweils 60sec] Auch hier tritt beim OCT-Array wieder eine leichte Asymmetrie im Front-System zwischen L und R auf (vgl. L/C, C/R), wie bereits zuvor bei der graphischen Auswertung der Kohärenz über der Frequenz (vgl. Abb. 10.14).

Die für das DECCA- und KFM-Array gefundenen Kohärenzwerte der DUO-Aufnahme in Tab. 10.2 sind denen der ORCH-Aufnahme sehr ähnlich, was den Schluss zulässt, dass es sich bei den gefundenen Charakteristiken quasi um einen messtechnischen ‚Fingerabdruck‘ des jeweiligen Surround-Mikrofonarrays handelt.

Die in der obenstehenden Tabelle für das BPT-Array gefundenen Werte belegen, dass eine hohe Kohärenz zwischen den Signalen der Frontmikrofone L (bzw. R) und C besteht. Die Kohärenz zwischen den Rear-Kanälen, sowie derselben zum Front System ist – aufgrund der Verwendung von zur Saalrückwand weisenden Nierenmikrofonen mit jeweils zwischenliegenden Abständen, die größer sind als der Hallradius – sehr gering. Der detaillierte Kohärenzverlauf über der Frequenz ist untenstehend zu sehen:

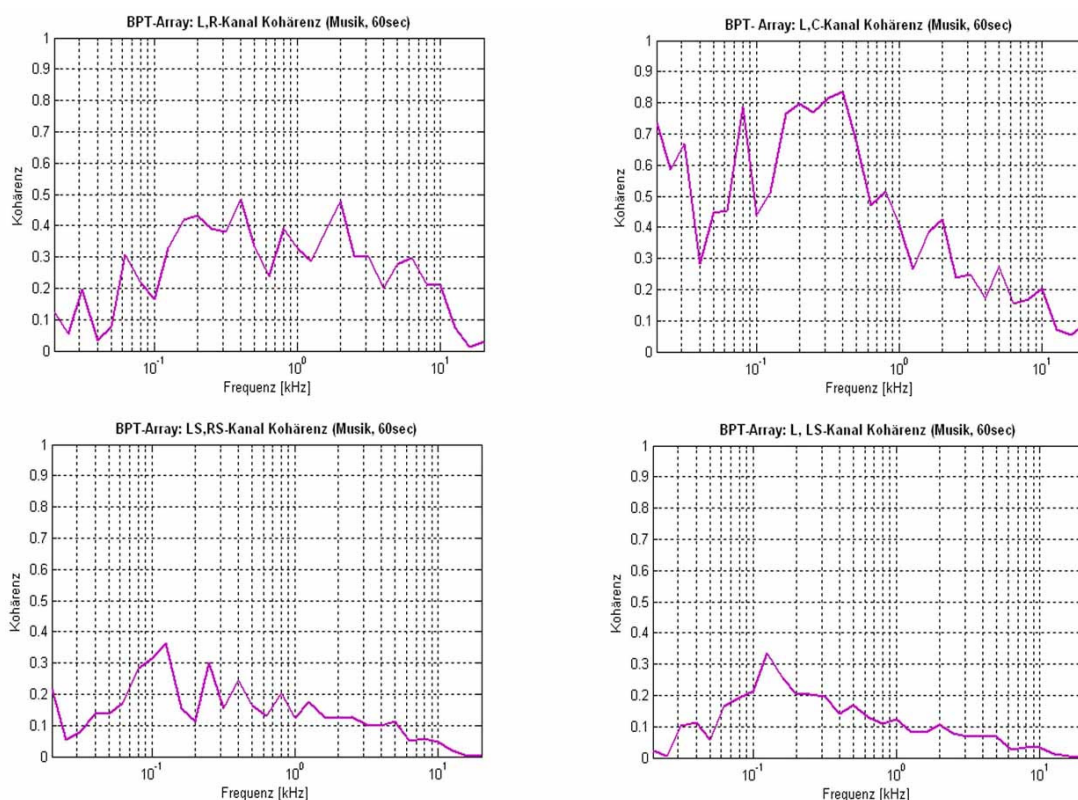


Abb. 10.23: BPT-Array, Kohärenz von je zwei Surround-Kanälen

In Abb.10.23 sind die Messungen der Kohärenz für das BPT-Surround-Mikrofonsystem dargestellt, welches anstatt des AB-PC Systems bei der Duo-Aufnahme (Cello+Klavier; Musik-Ausschnitt mit Länge 60sec [nicht identisch mit Ausschnitt #3], Saal 2 [Orchesterproberaum 447]) zum Einsatz kam:

Betreffend das L/R-Signal ist die Kohärenz bis ca. 100Hz gering, dann steigt sie an und oberhalb von ca. 200Hz erreicht sie mitunter Spitzenwerte bis ca. 0.5. Danach fällt sie zu den hohen Frequenzen hin wieder ab, wobei sie nie einen Wert von 0.3 überschreitet.

Aufgrund der starken Richtcharakteristik der Kapseln mit Achter-Charakteristik ist die Kohärenz zwischen dem L- und C-Kanal ausgeprägt: bei tiefen Frequenzen bis ca. 550Hz hinauf ist sie hoch mit Spitzenwerten bis knapp über 0.8, darüber geht sie mit Erhöhung der Frequenz zurück. Die Richtung Saal-Rückwand weisenden Rear-Mikrofone mit Nieren-Charakteristik waren in Art eines ‚Breit-AB‘ Systems positioniert. Da der Kapselabstand zueinander, wie auch zum Front-System größer war als der Hallradius, ist die Signal-Kohärenz zwischen den Kanälen LS/RS sowie zu den Front Kanälen (siehe LS/L) sehr gering.

Auch wenn die bei manchen Frequenzen auftretende hohe Signal-Kohärenz im Mitten- und Hochfrequenzband psychoakustisch nicht unbedingt von Vorteil ist, so sichert die generell niedrige Kohärenz im tieffrequenten Bereich (unter 200Hz), dass das BPT-System geeignet ist, um einen gute Räumlichkeitseindruck zu erzeugen.

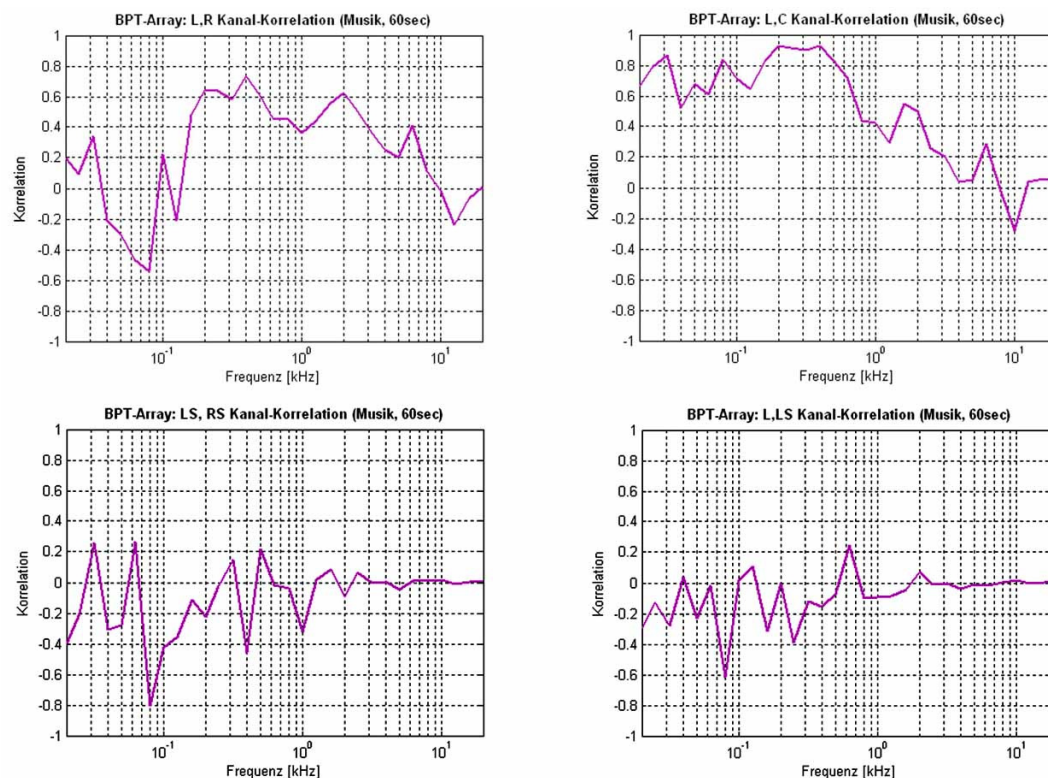


Abb.10.24: BPT-Array, Korrelation von je zwei Surround-Signalen

Die in der obenstehend in Abb. 10.24 dargestellte Korrelation über der Frequenz bestätigt dies ebenfalls: unterhalb von ca.150Hz ist die Korrelation der Surround-Signale generell gering (Ausnahme: L/C bzw. C/R), was günstig für den Raumeindruck ist. Abschließend sei nochmals auf den in der Abb. 10.11 unten rechts befindlichen Teil verwiesen, in der zum Vergleich das am Kunstkopf auftretende Binauralsignal des BPT-Arrays (allerdings eines kürzeren Ausschnitts in der Länge von nur 16sec) zu sehen ist.

Angemerkt werden darf in diesem Zusammenhang noch, dass die hohe Korrelation der Signale L/C bzw. R/C im tieffrequenten Bereich unter 500Hz durch die für die Praxis empfohlene Filterung zumindest der Signale L und R (siehe das Kapitel über Surround-Mikrofonierungstechniken, Abschnitt 5.1, Systembeschreibung des BPT-Mikrofons) zur Linearisierung des Frequenzgangs der Kapseln mit 8-er Charakteristik weitgehend entschärft wird, da durch die Filterung auch die Phasenbeziehung der betroffenen Signale in Richtung besserer Dekorrelation verändert wird.

10.7 Frequenzabhängige Kreuz-Korrelation (FCC) bei Stereo-Hauptmikrofonsignalen

Die ursprüngliche Motivation für die Durchführung der Stereo-Aufnahmen und frequenzabhängigen Kreuz-Korrelation der beiden Kanäle der Hauptmikrofonsignale lag in der bereits im Vorwort angesprochenen Frage nach der ‚Richtigkeit‘ der Schallquellenabbildung durch Klein-AB Mikrofonanordnungen im Vergleich zu Groß-AB Anordnungen.

Die erste Gruppe von Aufnahmen stammt von einem Kammermusikensemble, welches gleichzeitig mit 5 verschiedenen 2-Kanal Hauptmikrofonsystemen (Klein-AB, ORTF, XY (Nieren), MS, Schoeps-Kugelflächenmikrofon) aufgenommen wurde. (Jecklin, 1999); ‚Saal 3‘

Die zweite Gruppe von Aufnahmen wurde in den Jahren 2001/2002 im ‚Großen Festspielhaus‘ in Salzburg gemacht. Dabei kamen folgende Mikrofontechniken zur Anwendung:

ORTF, AB (0.2m, 0.4m, 0.8m, 1.3m, 3.2m, 7.2m und 12.0m) und AB-PC (12m)

Schallquellen: verschiedene Symphonie-Orchester

Verwendete Mikrofone: Schoeps MSTC64 (ORTF-Nieren) und Schoeps CMC5+MK3 (omnidirektionale Kapsel mit Diffusfeld-Entzerrung). ‚Saal 1‘

Zusätzliche Einzelaufnahmen wurden im Jahr 2000 in folgenden Räumen durchgeführt: eine kleine Kirche in London (relativ ‚hallige‘ Akustik) und ein großes TV-Studio in London (sehr ‚trockene‘ Akustik).

Klangquellen: kleines Orchester (Kirche), Instrumental-Ensemble (8 Musiker, TV-Studio). Verwendete Mikrofone: Neumann U-87 mit 8-er und Nieren-Charakteristik (ORTF, Blumlein-Paar; Kirche), Neumann U-87 mit 8-er und Nieren-Charakteristik (MS; TV-Studio)

Weiters wurde eine Gruppe von Studio-Aufnahmen (‚trockene‘ Akustik) eines Holzbläser-Trios zu Messungen herangezogen, um die klanglichen Eigenschaften folgender zwei Mikrofontechniken zu untersuchen: XY Nieren mit 90°

Öffnungswinkel und Blumlein-Paar. (Chesky, 1994) ‚Studio 1‘

(Anm.: über die verwendeten Mikrofone liegt leider keine Information vor)

Weiters wurde auch eine Orgel- Kunstkopfaufnahme aus dem ‚Alten Gewandhaus‘ in Leipzig im Rahmen der Messungen analysiert.

Da die untenstehend angeführten Stereo-Mikrofonssysteme in ihren Grundzügen und Eigenschaften bereits in Kapitel 4 beschrieben wurden, kann an dieser Stelle darauf verzichtet werden. Da wir auch im vorangehenden Kapitel 10.6 festgestellt haben, dass - in Hinblick auf die in dieser Arbeit interessierenden Parameter – eine Analyse der Korrelation statt der Ermittlung der Kohärenz zweier Signale vorzuziehen ist, wollen wir uns auf erstere beschränken.

In den Abb. 10.25-10.34 ist die Korrelation der elektrischen Signale der jeweiligen Aufnahme in Form einer schwarzen, durchgehenden Linie dargestellt (FCC), sowie - in Form einer violetten, unterbrochenen Linie - die frequenzabhängige interaurale Kreuz-Korrelation (FIACC). Dieser wurde mittels eines VST-Software Plug-Ins realisiert (Ausnahme: Kunstkopfaufnahme in Abb. 10.31), das ein 'normales' Stereo-Signal in ein entsprechendes Binaural-Signal umwandelt, welches den Signalen entspricht, die als Ohrensinalen an einem menschlichen oder Kunst-Kopf auftreten, wenn das elektrische Signal der jeweiligen Original-Aufnahme über eine Standard-Lautsprecheranordnung (Lautsprecherpositionen bei $\pm 30^\circ$, Radius 2m) wiedergegeben wird. Das Plug-In mit dem Namen ‚Pano 5‘ der Fa. WAVEARTS basiert auf den HRTF's des am MIT vermessenen ‚Kemar‘-Kunstkopfs der Fa. Knowles Electronic.

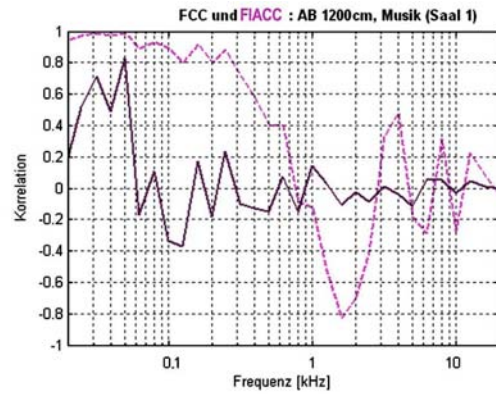
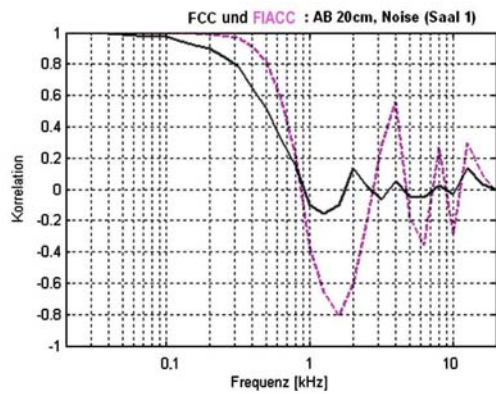


Abb. 10.25: Klein-AB 20cm (Rauschen, Saal 1) und Groß-AB 12m (Musik, Saal 1); FCC (durchgehende Linie) und FIACC (unterbrochene Linie)

Für das Klein-AB Stereo-Mikrofonsystem mit 20cm Kapselabstand kann man in Abb.10.25 das Maximum der Gegenphasigkeit (mit einem Wert von ca. -0.2) bei etwa 1100Hz erkennen. Bei Vergrößerung des Kapselabstandes auf 12m rutscht dieses – rein rechnerisch - entsprechend in den Bereich unter 40Hz. Der Grund für die davon abweichende Korrelationsfunktion dürfte an Wechselwirkungen mit der Raumakustik von Saal 1 (oder ev. störendem Einfluss von Körperschall) liegen.

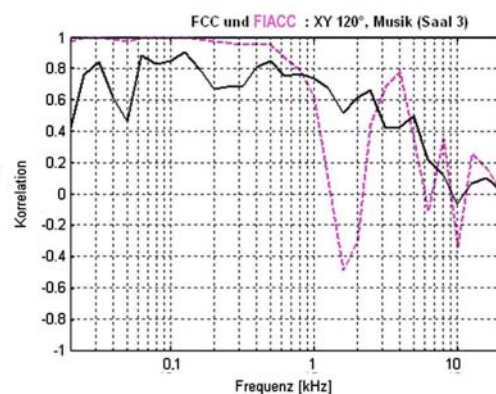
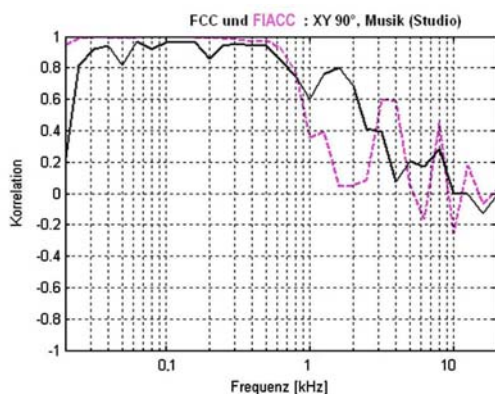


Abb. 10.26: XY 90° (Musik, Studio) und XY 120° (Musik, Saal 3)

Im Vergleich zu Abb. 10.25 erkennt man für die Stereo-Mikrofonsignale XY 90° und XY 120° (mit Nierenkapseln) von Abb.10.26 hohe Korrelation, die sich erst zu den höheren Frequenzen hin verliert. Für die XY-Kombination mit 90° physikalischem Öffnungswinkel ist bis 2kHz hohe Korrelation der Signale gegeben; ähnliches gilt für XY 120°, wobei sich erkennen lässt, dass hier auch unter 1kHz etwas geringere Korrelation auftritt, was aufgrund der stärkeren Signaltrennung L/R (größerer Öffnungswinkel) auch durchaus plausibel ist.

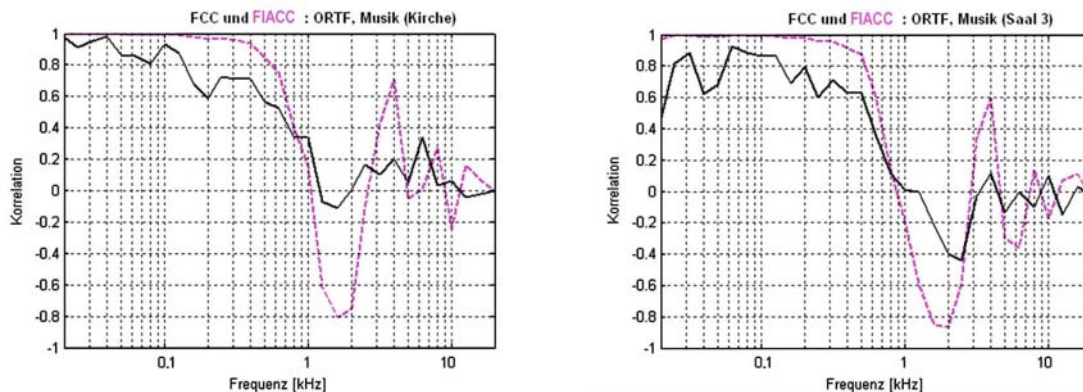


Abb. 10.27: ORTF (Musik, Kirche) und ORTF (Musik, Saal 3)

In Abbildung 10.27 sind die Funktionen für zwei Aufnahmen mit ORTF-Technik dargestellt: beide weisen unterhalb von ca.500Hz hohe Korrelation (≥ 0.6) auf und sind oberhalb von 1kHz so gut wie dekorreliert.

Die XY120°-Aufnahme und die ORTF-Aufnahme aus Abb.10.27 unterscheiden sich von der Aufnahmetechnik her in erste Linie durch den bei ORTF hinzukommenden Kapselabstand von 17.5cm, da beide ansonsten Kapseln mit Nierencharakteristik verwenden und die Öffnungswinkel mit 120° und 110° (bei ORTF) beinahe identisch sind.

Vergleicht man die Korrelationsfunktionen der XY 120°-Aufnahme aus Abb. 10.26 mit der rechten ORTF-Aufnahme aus Abb. 10.27, so fällt auf, dass sie im tieffrequenten Bereich bis ca. 500Hz ähnliche Charakteristik aufweisen: beide haben einen leichten Einbruch in der Korrelation im Bereich um ca. 40-50 Hz, der vermutlich auf raumakustische Ursachen zurückzuführen sein dürfte, da eine ähnliche Charakteristik bei praktisch allen Aufnahmen anzutreffen ist, die in Saal 3 gemacht wurden (einzige Ausnahme: KFM6U).

Für die darüber liegenden Frequenzen wirkt sich die Kapseltrennung beim ORTF-

System (Abb.10.27 rechts) dahingehend aus, dass ein Einbruch mit leichter Gegenphasigkeit (hier – entsprechend etwa dem Ohrenabstand - bei einer Frequenz zwischen 1-2kHz) auftritt, die für alle AB-basierten Mikrofonsystemen typisch ist. Während bei den koinzidenten XY-Aufnahmen (Abb.10.26) die Korrelation oberhalb von 1kHz nur langsam abfällt, ist diese bei den ORTF-Aufnahmen aufgrund des Kapselabstandes bereits ab dieser Frequenz gut dekorreliert.

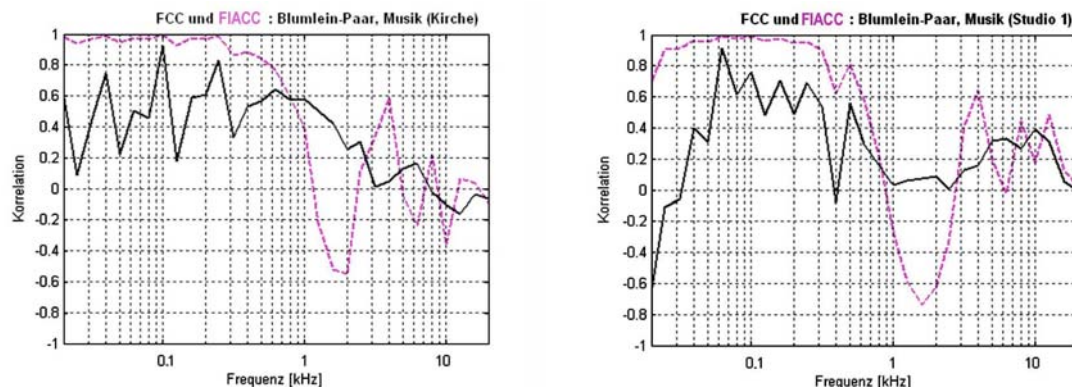


Abb. 10.28: Blumlein-Paar (Musik, Kirche) und Blumlein-Paar (Musik, Studio 1)

Die Signale des Blumlein-Paars sind aufgrund der Richtcharakteristik der Achter-Kapsel für einen Öffnungswinkel von 90° in Bezug auf Diffusschall komplett dekorreliert (siehe [Elko, 2001]). Die in der Abb. 10.28 gezeigten Funktionsverläufe resultieren aus der Kombination von Direkt- und Diffusschall in zwei verschiedenen Räumen, und zeigen über den gesamten Frequenzbereich eine deutliche Tendenz zur ‚Dekorrelation‘.

Ähnliches gilt auch für die MS-Aufnahmen in der untenstehenden Abbildung 10.29, die ebenfalls in zwei akustisch sehr unterschiedlichen Räumen entstanden sind (TV-Studio mit sehr trockener Raumakustik, sowie Saal 3).

(Anm.: für die linke MS-Aufnahme der Abb.10.29 ist das Pegelverhältnis von M- und S-Signal nicht bekannt, weshalb kein korrespondierender Aufnahmewinkel angegeben werden kann)

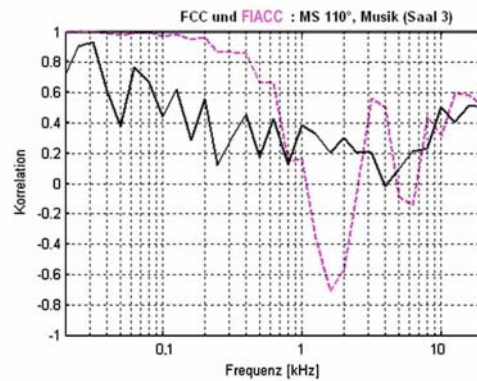
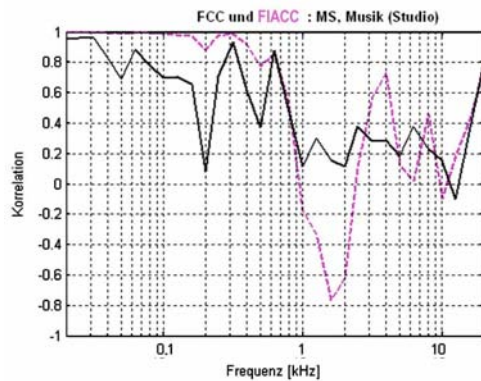


Abb. 10.29: MS (Musik, Studio) und MS 110° (Musik, Saal 3)

Es folgen nun einige etwas ausgefallener Stereo-Mikrofontechniken, deren Korrelations-Charakteristik näher betrachtet werden soll:

In Abb.10.30 sind die Funktionsverläufe des Faulkner ‚Phased Array‘ [Faulkner, 1981], sowie des Kugelflächenmikrofons KFM6U der Fa. Schoeps dargestellt.

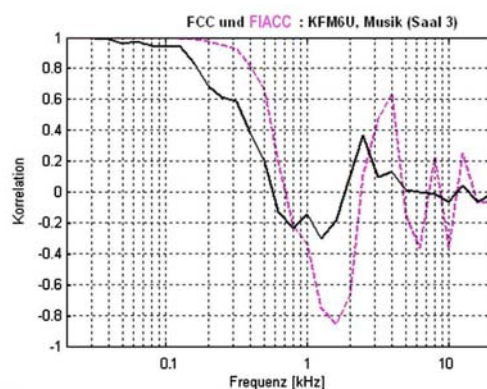
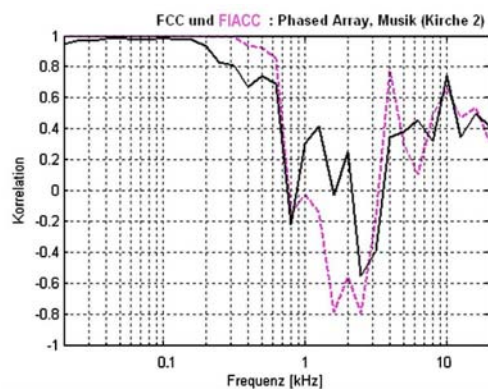


Abb. 10.30: Phased Array (Musik, Kirche 2) und KFM6U (Musik, Saal 3)

Für Details zu den beiden Mikrofonsystemen siehe die Abschnitte 4.2.3 (Faulkner Phased Array), sowie 5.3.1 (KFM6U). Beide Systeme verwenden eine Kapselabstand von 20cm; das Phased Array zwei parallel zur Schallquelle gerichtete ‚Achten‘, das KFM zwei Druckempfänger, die in die Oberfläche einer Kugel integriert sind.

Das Phased Array zeichnet sich durch hohe Korrelation unter ca. 750 Hz aus, eine relativ hohe, schmalbandige Gegenphasigkeit (bis beinahe -0.6) im Frequenzband zwischen ca. 2-3.5kHz), sowie darüber wiederum mittlere bis hohe Korrelation.

Die Korrelationsfunktion des KFM ist der eines Kunstkopfs sehr ähnlich, was nicht weiter überrascht. (siehe zum Vergleich die Korrelationsfunktion einer Orgelaufnahme in Abb.10.31, sowie auch die Korrelation am Kunstkopf in Abb.10.6)

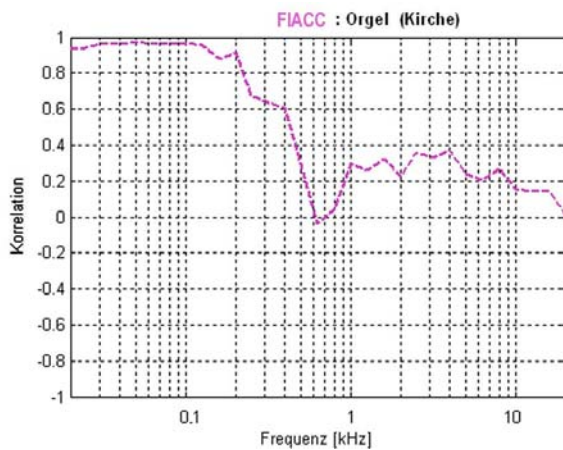


Abb. 10.31: Kunstkopfaufnahme (Modell unbekannt); Musik, Kirche 3

Als weiterer Spezialfall einer Trennkörper-Aufnahme wurde das OSS-System (Jecklin-Scheibe) analysiert. Wie bereits im Abschnitt über Stereo-Mikrofontechniken beschrieben, liefert das System für tiefe Frequenzen im Wesentlichen ein monophones Signal. In der Abb.10.32 links ist ersichtlich, dass die Korrelation bis 800Hz mit >0.6 sehr hoch ist, danach kommt es zur schmalbandigen, leichten Gegenphasigkeit, die für AB-Systeme typisch ist (zwischen ca. 1-2kHz) und im Frequenzbereich darüber ist die Trennscheibe wirksam und verändert den sonst gewohnt dekorrelierten Funktionsverlauf des AB-Mikrofonpaars, was in der vorliegenden Messung interessanterweise eine relativ breitbandige, leichte Gegenphasigkeit im Frequenz-bereich von 9-20kHz zur Folge hat.

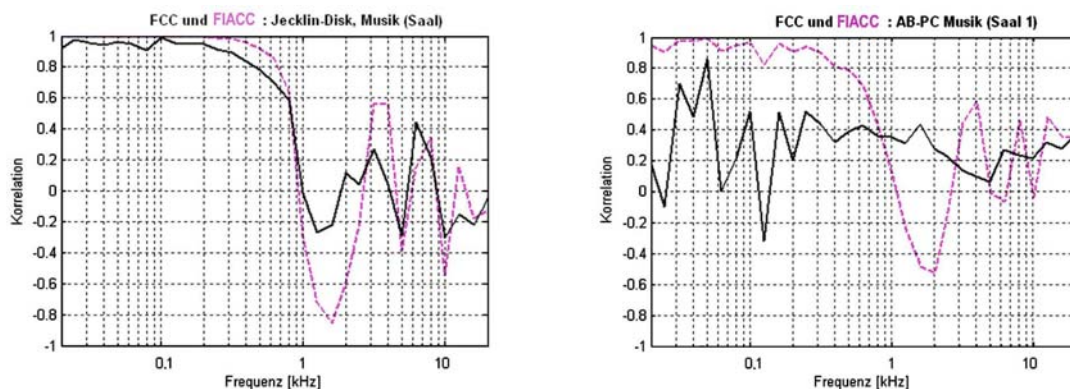


Abb.10.32: Jecklin-Disk ‚OSS‘ (Musik, Saal) und ABPC-Array (Musik, Saal 1)

Das ABPC-Array in Abb. 10.32 rechts – zusammengesetzt aus einem Breit-AB System mit 12m Kapselabstand und einem ORTF-Triple als Centerfill (nähere Details siehe Kapitel über Surround-Mikrofonsysteme) – zeichnet sich durch einen bis ca. 300Hz hinunter weitgehend dekorrelierten bzw. mit geringer Korrelation (≤ 0.4) charakterisierten Funktionsverlauf aus. Darunter kommt es zu vermutlich raumakustisch bedingten schmalbandigen Korrelationsspitzen, wie bereits obenstehend in Bezug auf das Breit-AB System alleine erörtert wurde (Abb.10.25 rechts).

Zum Vergleich sind untenstehend in Abb.10.33 die beiden Funktionsverläufe des Breit-AB Mikrofonpaars, sowie des ORTF-Stereomikrofons (nicht ORTF-Triple!) dargestellt, aus dem sich im Wesentlichen die Kombinationsfunktion des ABPC-Arrays in Abb. 10.32 rechts (allerdings ohne das Nieren-Mittelmikrofon des ORTF-Triples) zusammensetzt.

Daraus ist zu erkennen, dass die korrelierten Signalanteile des ABPC-Arrays im Frequenzbereich von ca. 60Hz – 2kHz von den Signalanteilen des ORTF (-Triples) herrühren, da das Breit-AB Signal bis 60Hz hinunter dekorreliert ist.

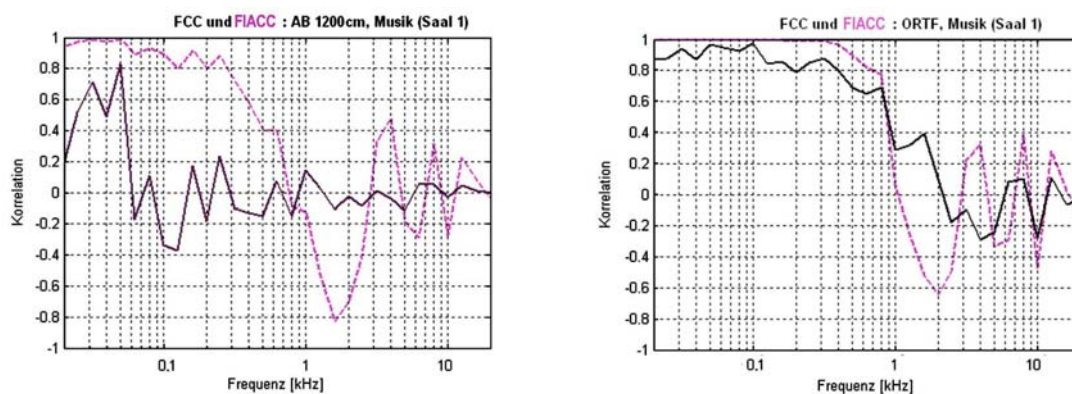


Abb. 10.33: Breit-AB 12m (Musik, Saal 1) und ORTF (Musik, Saal 1)

Zum Abschluss sollen die Korrelationsfunktionen von zwei Industrie-Aufnahmen betrachtet werden, die - stilistisch betrachtet – quasi zu zwei verschiedenen ‚ästhetischen Schulen‘ gehören und nicht – wie bisher – lediglich ein Stereo-Hauptmikrofonsystem verwenden:

Bei der ersten (Abb. 10.34, links) handelt es sich um eine Aufnahme der TELDEC (TELDEC, 1995), durchgeführt durch den deutschen Tonmeister Eberhard Sengpiel,

der auch als Dozent an der Universität der Künste in der Berlin im Bereich Tonmeisterausbildung unterrichtet. Bei der Mikrofonierungstechnik handelt es sich um Klein-AB mit Stützmikrofonen.

Die zweite Aufnahme (Abb. 10.34, rechts) stammt aus der Serie „Living Stereo“ (RCA, 1957) im Rahmen derer die amerikanische Plattenfirma RCA ab Mitte der 1950-er Jahre ihre ersten Stereo-Aufnahmen veröffentlichte. Die damals angewandten Mikrofontechniken dürfen als eher puristisch eingestuft werden, da sie ein Minimum an Mikrofonen (meist Breit-AB mit Centerfill basiert; siehe hierzu [Valin, 1994]) einsetzen, werden von Musikliebhabern weltweit geschätzt und wurden in den 1990-er Jahren erneut auf CD und ab 2003 auch auf SA-CD wiederveröffentlicht.

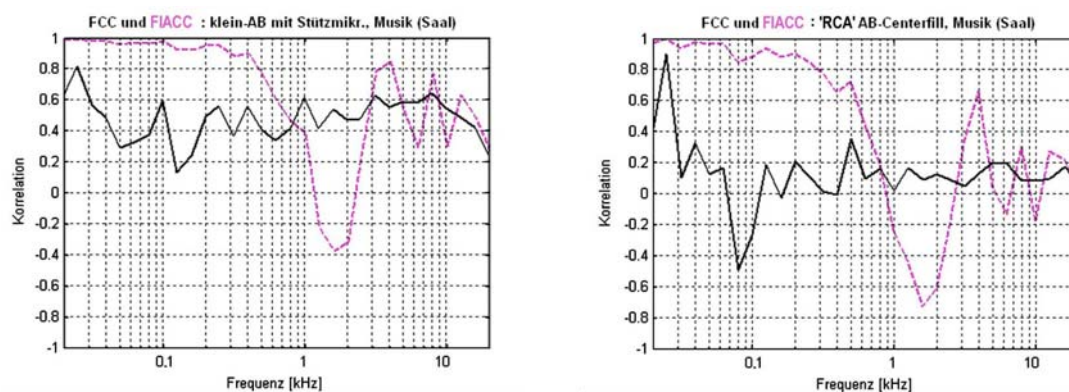


Abb.10.34: Klein-AB+Stützmikr. (Musik, Saal) und AB-Centerfill (Musik, Saal)

Wie sich anhand von Abb.10.34 deutlich erkennen lässt, ist die AB-Centerfillaufnahme praktisch über den gesamten Frequenzbereich dekorreliert, während die Klein-AB Aufnahme mit Stützmikrofonen zwischen geringer (0.4) und hoher (0.6) Korrelation schwankt. Während die Korrelation über den gesamten Frequenzbereich gemittelt bei der TELDEC Aufnahme ca. 0.5 beträgt, liegt diese bei der RCA-Aufnahme bei nur 0.1.

Abschließend soll angemerkt werden, dass die teilweise sehr unterschiedlichen Funktionsverläufe der FCC natürlich auch in entsprechend davon abgeleiteten FIACC-Funktionen ihren Niederschlag finden, wie in den obenstehenden Abbildungen ersichtlich.

Zusammenfassung

Aufgrund der bei den Surround-Aufnahmen erhaltenen Ergebnisse wurde für die im Rahmen der vorliegenden Arbeit untersuchten Stereo-Aufnahmen auf eine Ermittlung der Kohärenzfunktion verzichtet und lediglich die frequenzabhängige Korrelation (FCC und FIACC) bestimmt:

Sofern verfügbar, wurden pro Hauptmikrofontechnik jeweils zwei Aufnahmen aus unterschiedlichen Aufnahmeräumen zur Analyse herangezogen, um ein ausgewogeneres Bild über deren Eigenschaften zu erhalten. Wie anhand der in den Abb. 10.25-10.34 gezeigten Funktionsverläufe ersichtlich wurde, weisen so gut wie alle obenstehend betrachteten Stereo-Hauptmikrofonsysteme charakteristische FCC's auf, die in ihren Grundzügen von der jeweiligen Raumakustik relativ unabhängig sind. Auffallend ist weiters, dass sich viele der untersuchten Stereo-Mikrofontechniken durch hohe Signal-Korrelation (FCC) bei niedrigen Frequenzen (unterhalb ca. 500Hz) auszeichnen, was dem Entstehen eines guten Raumeindrucks beim Hörer abträglich ist.

Unter diesen befinden sich:

- Klein AB
- XY (mit Nierenkapseln)
- ORTF
- Faulkner Phased Array
- KFM und
- OSS („Jecklin-Scheibe“)

Von den in diesem Abschnitt untersuchten Stereo-Mikrofontechniken liefern - prinzipbedingt - lediglich die Mikrofontechniken MS, Blumlein-Paar, ABPC-Array und AB-Centerfill auch bei tiefen Frequenzen weitgehend dekorrierte Signale (vgl. Abb. 10.28, 10.29, 10.32 und 10.34).

10.8 Binaural Quality Index für reproduzierte Musik (BQI_{rep})

Wie bereits im Abschnitt ‚Physikalische Größen im Zusammenhang mit dem Raumeindruck‘ erläutert, wurde der ‚Binaural Quality Index‘ zur Bewertung der Qualität der Akustik von Konzertsälen definiert (vgl. [Beranek, 2004]).

Im Rahmen der vorliegenden Untersuchung stellte sich bei der messtechnischen Bestimmung eines neu definierten ‚Binaural Quality Index für wiedergegebene Musik‘ (engl. ‚Binaural Quality Index for reproduced music‘ – abgekürzt: BQI_{rep}) heraus, dass eine der zu bewertenden Mikrofontechniken (ABPC) in der Lage war, im Hörraum einen BQI_{rep} zu erzielen, der mit dem im Konzertsaal gemessenen BQI übereinstimmt, während der BQI_{rep} der anderen Mikrofontechniken davon weit abwich (siehe nächste Seite, Tab. 10.3). Der BQI_{rep} soll definiert werden, wie folgt:

$$BQI_{rep} = (1 - IACC_3) \quad (10.1)$$

wobei der $IACC_3$ definiert ist als der Mittelwert der Korrelations-Koeffizienten der in den Oktavbändern 500Hz, 1kHz und 2kHz enthaltenen L/R- Binauralsignale.

Die messtechnische Ermittlung des BQI_{rep} lehnt sich an den konzertsaal-spezifischen BQI insofern an, als dabei ebenfalls die Korrelation der beiden Ohren-Signale eines Kunstkopfs in den Oktav-Frequenzbändern mit den Mittenfrequenzen bei 500Hz, 1000Hz und 2000Hz gemessen wird.

Anders als im Konzertsaal kann hierbei allerdings nicht zwischen zeitlich frühen Komponenten reflektierter Schallenergie („early“) und späteren Rückwürfen („late“) unterschieden werden, da die Messung nicht auf einer Raumimpulsantwort basiert, sondern auf eine bereits bestehende Musikaufnahme zurückgegriffen werden muss, in der Direktschall und alle Raumanteile überlagert sind.

Deshalb entfällt auch der Subindex ‚E3‘, bei dem das E für den frühen Anteil der Schallenergie im Zeitbereich von 0 bis 80ms steht ; der besseren Unterscheidbarkeit zu den in der Literatur bereits definierten Größen wie $IACC_{E3}$ und $IACC_{L3}$ zuliebe wird die Bezeichnung BQI_{rep} vorgeschlagen, aus der unmittelbar hervorgeht, dass es

sich hierbei um die Bewertung bereits aufgenommener Musik (bzw. Schallsignale) handelt.

| MicArray | rho500 | rho1000 | rho2000 | BQI_{rep} |
|-----------------|---------------|----------------|----------------|--------------------------|
| OCT | 0.83 | 0.64 | 0.66 | 0.29 |
| DECCA | 0.77 | 0.61 | 0.57 | 0.35 |
| KFM | 0.90 | 0.61 | 0.39 | 0.37 |
| ABPC | 0.43 | 0.26 | 0.11 | 0.73 |
| KU81 | 0.41 | 0.19 | 0.20 | 0.73 |

Tab. 10.3: BQI_{rep} und rho (500, 1000, 2000), basierend auf einer messtechnischen Auswertung der Re-Recordings der Surroundaufnahme ‚ORCH-Sample #4‘ durch OCT, DECCA, KFM und ABPC mittels KU81-Kunstkopf; sowie Angabe dieser Werte für die Original-Kunstkopfaufnahme aus dem Konzertsaal [KU81] ¹

Die oben genannte Beschränkung auf bereits aufgenommene (Musik-) Signale bietet jedoch auch den Vorteil, dass im Prinzip jede Aufnahme – im Wiedergaberaum mittels Kunstkopf neu aufgenommen – einer qualitativen Bewertung auf Basis der Bestimmung des BQI_{rep} zugeführt werden kann. Als Länge für den zu verwendenden Ausschnitt der Tonaufnahme wird 1 Minute vorgeschlagen, da bei Verwendung kürzeren Ausschnitts (von z.B. nur 10 Sekunden) die Gefahr besteht, ev. nicht repräsentative Abschnitte der Gesamtaufnahme zu erfassen. (Siehe hierzu Abb. 10.10: die Surround-Technik „5CH“ (braun) ist im Zeitabschnitt von Sekunde 25-29 stark gegenphasig bis -0.6, aber im Zeitabschnitt Sekunde 44-48 gleichphasig um +0.5 !). Die in Abb. 10.10 dargestellten Ergebnisse sind ein Beleg dafür, dass es – zur Erzielung eines objektiveren Ergebnisses – sinnvoll ist, den Messzeitraum in der Regel auf zumindest eine Minute auszudehnen, um einen realistischeren Mittelwert zu erhalten.

Um die Gültigkeit des in der obenstehenden Tabelle ausgewiesenen BQI_{rep} für die mit dem KU81 im Konzertsaal durchgeführten Originalaufnahme erhärten zu können, ist untenstehend in Tab. 10.4 ein relevanter Auszug aus den Messungen von [Hidaka et al., 1996] wiedergegeben, aus dem ähnliche Werte hervorgehen.

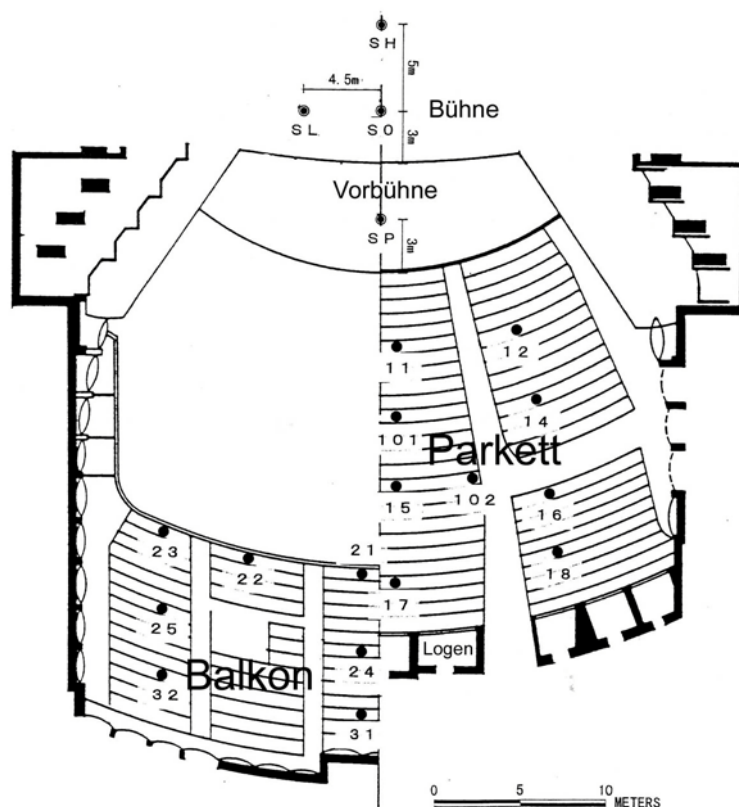
¹⁾ (Anm.: wir beschränken uns hier bewusst auf eine Auswertung des Samples #4, da die Samples #1 und #2 in diesem enthalten sind, während das Sample #3 aus der Duo-Aufnahme stammt, die an einem ganz anderen Ort aufgenommen wurde, nämlich im Proberaum)

Die bei den Orchesteraufnahmen verwendete Position des Neumann KU81 Kunstkopfs entsprach der in der nachfolgenden Abbildung dargestellten Position 11. Für diese traten bei den Messungen von Hidaka et al. bei Verwendung der Senderposition S0 außergewöhnlich hohe $IACC_{E3}$ -Werte von 0.90 auf; über die in der untenstehenden Tabelle aufgeführten Messpunkte gemittelt ergab sich allerdings ein $IACC_{E3}$ von nur 0.59. Betrachten wir außerdem den Wert der $IACC_{A3}$ -Werte ($A='all'$; Zeitfenster 0-3500ms), über alle Empfängerpositionen gemittelt („mean“), so liegt dieser für S0 bei 0.38.

Für die Senderposition Sp ergab sich vergleichsweise ein noch geringerer Mittelwert (über alle Empfängerpositionen) der $IACC_{A3}$ -Werte von 0.29. (Der Durchschnitt der $IACC_{A3}$ -Mittelwerte S0 und Sp läge also bei ca. 0,34; das ergäbe für $1-IACC_{A3}$ einen Wert von 0.66)

Berücksichtigt man nun, dass bei der Original-Kunstkopfaufnahme mittels KU81 als ursprünglich den Raum anregende Quelle nicht der kugelförmig abstrahlende Dodekaeder-Messlautsprecher des Forscherteams um Hidaka zur Anregung des Konzertsaals des Großen Festspielhauses gedient hatte, sondern eine wesentlich größere und komplex abstrahlende Schallquelle (Symphonie-Orchester), so scheint der Unterschied in den auftretenden Werten zwischen dem Wert für $1-IACC_3$ (des KU81) von 0.73 und dem oben errechneten durchschnittlichen $1-IACC_{A3}$ von 0.66

aus dem Messungen des Teams um Hidaka durchaus plausibel.



Großes Festspielhaus
Salzburg

S ... Senderpositionen
Zahlen ... Empfangs-
positionen

Abb.10.35: Schema der Messpositionen aus der Untersuchung von [Hidaka et al., 1996] (Grafik überarbeitet)

Festspielhaus, Salzburg
[IACC]

| | | IACCE | | | | | IACCL | | | | | IACCA | | | | | |
|----|-----|-------|-------|-------|---------|-------|-------|-------|-------|---------|-------|-------|-------|-------|---------|-------|-------|
| | | 500 | 1k | 2k | 3 bands | 4k | 500 | 1k | 2k | 3 bands | 4k | 500 | 1k | 2k | 3 bands | 4k | |
| SO | 101 | 0.851 | 0.850 | 0.861 | 0.85 | 0.748 | 0.257 | 0.185 | 0.140 | 0.19 | 0.049 | 0.444 | 0.535 | 0.530 | 0.50 | 0.541 | |
| | 102 | 0.550 | 0.625 | 0.475 | 0.55 | 0.545 | 0.209 | 0.203 | 0.099 | 0.17 | 0.071 | 0.322 | 0.417 | 0.275 | 0.34 | 0.335 | |
| | 11 | 0.883 | 0.877 | 0.945 | 0.90 | 0.823 | 0.287 | 0.126 | 0.102 | 0.17 | 0.054 | 0.509 | 0.493 | 0.735 | 0.58 | 0.688 | |
| | 14 | 0.702 | 0.620 | 0.624 | 0.65 | 0.416 | 0.223 | 0.105 | 0.080 | 0.14 | 0.053 | 0.405 | 0.353 | 0.372 | 0.38 | 0.236 | |
| | 17 | 0.511 | 0.537 | 0.669 | 0.57 | 0.491 | 0.206 | 0.115 | 0.100 | 0.14 | 0.114 | 0.325 | 0.327 | 0.445 | 0.37 | 0.362 | |
| | 18 | 0.491 | 0.308 | 0.211 | 0.34 | 0.328 | 0.090 | 0.139 | 0.108 | 0.11 | 0.039 | 0.258 | 0.223 | 0.143 | 0.21 | 0.245 | |
| | 21 | 0.797 | 0.729 | 0.714 | 0.75 | 0.463 | 0.162 | 0.089 | 0.064 | 0.11 | 0.056 | 0.542 | 0.418 | 0.395 | 0.45 | 0.291 | |
| | 23 | 0.587 | 0.329 | 0.405 | 0.44 | 0.352 | 0.180 | 0.067 | 0.065 | 0.10 | 0.058 | 0.478 | 0.236 | 0.282 | 0.33 | 0.253 | |
| | 24 | 0.371 | 0.487 | 0.534 | 0.46 | 0.306 | 0.175 | 0.063 | 0.085 | 0.11 | 0.054 | 0.273 | 0.317 | 0.350 | 0.31 | 0.242 | |
| | 25 | 0.619 | 0.410 | 0.239 | 0.42 | 0.283 | 0.188 | 0.145 | 0.054 | 0.13 | 0.033 | 0.471 | 0.324 | 0.187 | 0.33 | 0.207 | |
| | | mean | 0.636 | 0.577 | 0.568 | 0.59 | 0.476 | 0.198 | 0.124 | 0.090 | 0.14 | 0.058 | 0.403 | 0.364 | 0.371 | 0.38 | 0.340 |
| Sp | 101 | 0.844 | 0.719 | 0.283 | 0.62 | 0.550 | 0.161 | 0.167 | 0.075 | 0.13 | 0.061 | 0.398 | 0.410 | 0.180 | 0.33 | 0.405 | |
| | 102 | 0.487 | 0.538 | 0.548 | 0.52 | 0.504 | 0.112 | 0.072 | 0.078 | 0.09 | 0.072 | 0.252 | 0.246 | 0.314 | 0.27 | 0.318 | |
| | 11 | 0.938 | 0.838 | 0.791 | 0.86 | 0.842 | 0.101 | 0.088 | 0.094 | 0.09 | 0.067 | 0.582 | 0.520 | 0.569 | 0.56 | 0.723 | |
| | 14 | 0.400 | 0.488 | 0.554 | 0.48 | 0.459 | 0.056 | 0.084 | 0.056 | 0.07 | 0.073 | 0.222 | 0.293 | 0.284 | 0.27 | 0.300 | |
| | 17 | 0.630 | 0.574 | 0.610 | 0.60 | 0.517 | 0.119 | 0.106 | 0.144 | 0.12 | 0.051 | 0.390 | 0.346 | 0.408 | 0.38 | 0.415 | |
| | 21 | 0.180 | 0.321 | 0.431 | 0.31 | 0.366 | 0.289 | 0.111 | 0.084 | 0.16 | 0.065 | 0.158 | 0.154 | 0.269 | 0.19 | 0.260 | |
| | 23 | 0.551 | 0.228 | 0.206 | 0.33 | 0.328 | 0.233 | 0.084 | 0.136 | 0.15 | 0.055 | 0.321 | 0.168 | 0.075 | 0.19 | 0.203 | |
| | 24 | 0.585 | 0.231 | 0.247 | 0.35 | 0.156 | 0.249 | 0.089 | 0.092 | 0.14 | 0.059 | 0.370 | 0.127 | 0.119 | 0.21 | 0.102 | |
| | 25 | 0.488 | 0.187 | 0.271 | 0.32 | 0.312 | 0.167 | 0.083 | 0.118 | 0.12 | 0.060 | 0.275 | 0.091 | 0.161 | 0.18 | 0.205 | |
| | | mean | 0.567 | 0.458 | 0.438 | 0.49 | 0.448 | 0.165 | 0.098 | 0.097 | 0.12 | 0.063 | 0.330 | 0.262 | 0.264 | 0.29 | 0.326 |

Tab.10.4: IACC_E, IACC_L und IACC_A; ‚3 bands‘ = Mittelwert der Frequ. Bänder 500Hz, 1kHz, 2kHz (_E=0-80ms, _L=80-3500ms, _A=0-3500ms (aus [Hidaka et al., 1996])

Auch wenn die Bestimmung des BQI_{rep} alleine vermutlich kein ausreichendes Kriterium zur Qualitätsbewertung einer Surround- oder 2-Kanal Stereoaufnahme darstellt (sonst müssten sich bei der Bewertung der ‚ORCH-5.1‘ Aufnahme die Werte der Hörer-Präferenz für die vier einander gegenübergestellten Mikrofontechniken stärker voneinander unterscheiden, und die ABPC-Technik eine klare Spitzenposition einnehmen), so darf doch angenommen werden, dass die tendenzielle Bevorzugung der ABPC-Technik hinsichtlich der Hörerpräferenz und der empfundenen Räumlichkeit (siehe Abb. 9.10) damit in Zusammenhang steht. Wenn die starke Dekorrelation tieffrequenter Signalanteile im Konzertsaal einen so entscheidenden Einfluss auf die Hörerpräferenz hat (IACC_{E3} bzw. BQI; siehe [Beranek, 2004]), dann ist es plausibel, dass ähnliches auch für den Klangeindruck im Abhörraum/Wohnzimmer gelten sollte (auch wenn in diesem Zusammenhang natürlich angemerkt werden muss, dass die in einer Aufnahme fixierte Dekorrelation tiefer Frequenzen durch die mitunter sehr individuelle akustische Beschaffenheit kleiner Hörräumen überlagert bzw. außer Kraft gesetzt sein kann (siehe hierzu u. a. [Griesinger, 1997])).

Wie sich aus den obenstehend in Tabelle 10.3 aufgeführten Ergebnissen schließen lässt, wäre der neu definierte BQI_{rep} vermutlich ein lohnender Gegenstand weiterer Untersuchungen betreffend die hochwertige/naturgetreue Aufnahme und Wiedergabe von Musik bzw. – allgemeiner formuliert – von Schallsignalen.

10.9 Zusammenfassung

Im Rahmen dieses Kapitels wurde für insgesamt fünf Surround-Mikrofonsysteme (OCT, DECCA, KFM, AB-PC, BPT) die frequenzabhängige Korrelation, wie auch frequenzabhängige Kohärenz untersucht.

Als ‚Bezugsnormal‘ wurde hierbei das an einem Kunstkopf (Neumann KU81) auftretende Binauralsignal genommen, da für den beim Hörer auftretenden Klangeindruck die bei seinen Ohren einlangenden Signale ausschlaggebend sind. Für dieses Binauralsignale wurde die frequenzabhängige Korrelation (FIACC), wie auch die frequenzabhängige Kohärenz bestimmt und jeweils der Korrelations- (bzw. Kohärenz-) Funktion der Referenzaufnahme (Original-Kunstkopfaufnahme aus dem Konzertsaal) gegenübergestellt.

Es konnte gezeigt werden, dass es durch die Untersuchung des am Kunstkopf auftretenden Binauralsignals möglich ist, relevante (und teilweise für die jeweilige Mikrofontchnik charakteristische) Unterschiede aufzuzeigen.

Zugleich zeigte sich aber auch, dass es nicht möglich ist, aufgrund der Untersuchung der frequenzabhängigen Kohärenzfunktion alleine potentielle Schwächen eines Mikrofonsystems (z.B. hinsichtlich Phasenanomalien, wie beim KFM) so klar aufzuzeigen, wie durch Ermittlung der Korrelationsfunktion.

Um für die vorliegenden Surround-Aufnahmen genauere Informationen bezüglich der Ähnlichkeiten von Signalinhalten der Kanäle untereinander zu erhalten, wurden benachbarte Kanäle jeweils paarweise auf Korrelation und Kohärenz über der Frequenz untersucht, wobei auch hier klare und für die einzelnen Mikrofonsysteme charakteristische Unterschiede noch detaillierter aufgezeigt werden konnten.

Die im Hörvergleich (Aspekt Präferenz) am schlechtesten bewertete Surround-Mikrofontchnik KFM wies nicht nur zwischen dem L und R Front-Kanal breitbandige Gegenphasigkeiten auf, sondern auch zwischen den vorderen und hinteren Kanälen (Paarung L/LS; vgl. Abb. 10.20). Die bereits zuvor durchgeführte Untersuchung der frequenzabhängigen Korrelation des bei der Wiedergabe im Abhörraum am Kunstkopf auftretenden Binauralsignals hatte ebenfalls gezeigt, dass bei diesem System starke Gegenphasigkeiten im Frequenzbereich um 1-2kHz auftreten (siehe Abb. 10.6 und 10.8); es handelt sich dabei um einen Frequenzbereich, der auch für das Verstehen von Sprache

besonders wichtig, und in dem das menschliche Ohr entsprechend empfindlich ist.

Zugleich sind die Signale des KFM-Mikrofonsystems allerdings auch bis zu einer höheren Frequenz hinauf korreliert bzw. kohärent, als das bei den Vergleichssystemen der Fall ist. Dies bewirkt unter anderem, dass sich für das KFM-System ein vergleichsweise wesentlich geringerer Räumlichkeitseindruck ergibt (vgl. hierzu Ergebnisse des Hörvergleichs in Abb. 9.10 und 9.11).

11 Zusammenfassung, Diskussion, Schlussfolgerung und Ausblick

11.1 Einleitung

Den Anstoß für Untersuchungen zu diesem Thema gab die Frage nach der ‚Korrektheit‘ der Abbildung einer Schallquelle bei Verwendung von Klein- vs. Groß-AB Mikrofontechniken: wenn die Mikrofontechnik ‚Klein-AB‘ - im Gegensatz zu ‚Groß-AB‘ - aufgrund der Gesetze der Psychoakustik eine korrekte Lokalisation der Schallquelle ermöglicht, und Groß-AB hingegen das inhärente Problem des so genannten ‚Mittenlochs‘ aufweist, warum kommt es dann so häufig vor, dass Tonmeister oft ‚Groß-AB‘ basierte Techniken in Ergänzung zu den von ihnen favorisierten Hauptmikrofonsystemen einsetzen ... ? (siehe hierzu [La Grou, 1994], pg.32, [Valin, 1994], pg. 123/124, [Betz, 2000])

Eine Studie namhafter Fachleute brachte auch zu Tage, dass die Lokalisation von Klangquellen einer Musikaufnahme unter Anwendung von ‚Klein-AB‘ Mikrofontechniken mitunter stark unterschiedlich ausfallen kann, da hierbei Abweichungen von teilweise über 100 % gemessen wurden (vgl. [Wittek und Theile, 2002]).

Auch die Untersuchung von [Gernemann und Rösner, 1998]) belegt, dass ‚Klein-AB‘ Mikrofontechniken (mit minimalen ITD's entsprechend den Erfordernissen der Psychoakustik) nicht in der Lage sind, verlässliche Lokalisationseindrücke beim Hörer hervorzurufen. Darüber hinaus führt ein geringer Kapselabstand zweier omnidirektionaler Mikrofone unweigerlich zu einer hohen Signal-Korrelation bei tiefen Frequenzen, was für den Raumeindruck abträglich ist.

Teilergebnisse der vorliegenden Studie gehen in eine ähnliche Richtung: die DECCA-Surroundtechnik, basierend auf zwischen den Mikrofonen auftretenden (minimalen) Laufzeitdifferenzen litt unter einer stärkeren Signalverschlechterung bei ungünstigeren Wiedergabebedingungen (Wiedergabe des re-recordeten KU81-Binauralsignals via Kopfhörer), als die Konkurrenzsysteme (OCT, KFM, AB-PC) – (vgl. hierzu Kapitel 9, Abb. 9.10 und 9.11).

11.2 Zusammenfassung und Diskussion

In Kapitel 2 (Räumliches Hören) wurde der aktuelle Wissenstand zu den psychoakustischen Mechanismen der Lokalisation, Distanzwahrnehmung und des Raumeindrucks zusammengefasst. Weiters wurden physikalische Größen im Zusammenhang mit dem Raumeindruck näher betrachtet, sowie der Einfluss von Lautsprecher-Eigenschaften und Akustik des Abhörtraumes auf die Hörerpräferenz. In diesem Zusammenhang spielt vor allem die u. U. stattfindende Vergrößerung der ASW (Apparent Source Width) – in Abhängigkeit von der Breite der Lautsprecheraufstellung und ev. im Abhörtraum auftretenden starken Reflexionen – eine wichtige Rolle (siehe [Toole, 1986] und [Bech, 1998]).

In Kapitel 3 wurden die Grundlagen von Signal-Korrelation und Signal-Kohärenz erklärt, sowie der bisherige Wissensstand betreffend Signalkorrelation bei Stereo-Mikrofonsystemen kurz dargestellt. Weiters wurde die Frage nach der ‚optimalen‘ Signal-Kohärenz bei Surround- Mikrofonsystemen angerissen (vgl. [Theile, 2000]), sowie in Abschnitt 3.3 des ‚optimalen‘ interauralen Korrelationsgrades im Konzertsaal, wie auch bei Mehrkanal-Lautsprecherwiedergabe.

In Abschnitt 3.4 wurden einige Gedanken zur Wechselwirkung zwischen Lautsprecher-Richtwirkung und akustischer Hörer-Einhüllung eingebracht.

Kapitel 4 versucht einen kompakten Überblick über die etablierten Stereo-Mikrofontechniken zu geben, wobei eine vom Dissertanten (auch für Surround-Aufnahmen geeignete) neue Mikrofontechnik (AB-PC: ‚AB-Polycardioid Centerfill‘) vorgestellt wurde (siehe Abschnitt 4.3)

In Kapitel 5 wurde der Versuch unternommen, eine möglichst vollständige Übersicht der unterschiedlichen Surround-Mikrofonssysteme zu bieten. Hierbei wurden sowohl die unterschiedlichen Ansätze (Koinzidente und quasi-koinzidente Systeme, Spaced Arrays (2-dimensional), sowie Trennkörper und ‚3D‘-Techniken) betrachtet, wie auch auf Spezialanfertigungen (Soundfield-Mikrofon, Holophone ‚H2-Pro‘, Schoeps ‚KFM 360‘, etc.) eingegangen. Im Rahmen dieser Übersicht wurde eine weitere vom Dissertanten vor allem für die Aufnahme von kleinen Schallquellen konzipierte Surround-Mikrofontechnik (BPT – ‚Blumlein-Pfanzagl-Triple‘) vorgestellt. Das Kapitel schließt mit allgemeinen Betrachtungen über Mikrofon-Arrays.

Kapitel 6 ist einer Betrachtung über Kunstkopfaufnahmen und einer Analyse von Studien betreffend die Qualität verschiedener Kunstköpfe gewidmet. Außerdem wird die Verwendung eines Kunstkopfs (Neumann KU81i) als ‚Bezugsnormal‘ im Rahmen der vorliegenden Arbeit diskutiert (Abschnitt 6.1).

In Kapitel 7 wird zunächst Grundlegendes zur Problematik von Hörvergleichen erörtert, Anforderungen an die Wiedergabeanordnung, Einflussfaktoren auf die Hörerbewertung, sowie die Auswahl der zu bewertenden Klangattribute. Die Art des Signals zur ‚Kalibrierung‘ der Probanden des vorliegenden Hörvergleichs wurde beschrieben und Gedanken zur Anwendbarkeit der Varianzanalyse angebracht. Bezüglich der Frage des Vorliegens einer Normalverteilung bei den Hörerantworten wurde für alle Aspekte und untersuchten Mikrofontechniken ein Lilliefors-Test durchgeführt.

Kapitel 8 widmet sich schließlich der Auswertung des Hörvergleichs ‚ORCH 5.1‘, bei dem die Signalwiedergabe über Lautsprecher erfolgte. An diesem Teil des Tests nahmen 25 Probanden teil; die Beurteilung erfolgte mittels einer direkten Größenabschätzung anhand eines semantischen Differenzials, verbunden mit einem Paarvergleich (und Bewertungsskala). Es wurde eine Varianzanalyse für die insgesamt 14 zu beurteilenden Aspekte, sowie die Hörerpräferenz durchgeführt, und für die vier in Konkurrenz stehenden Mikrofontechniken (OCT, DECCA, KFM, ABPC) pro Aspekt jeweils der Mittelwert über die Hörerantworten mit 95% Konfidenzintervall dargestellt (vgl. Abb. 8.2). Außerdem wurde ein Dendrogramm der Hörerähnlichkeiten erstellt (siehe Abb. 8.1) sowie – in Zusammenhang mit der Auswertung der dafür erstellten Testteile - Überlegungen bezüglich Reliabilität und Validität der Hörerbewertungen angestellt.

Die Auswertung des Hörvergleichs ‚ORCH 5.1‘ zeigte, dass das KFM-Surroundmikrofon bei der Hälfte aller Aspekte statistisch signifikant schlechter bewertet wurde, als die Konkurrenzsysteme, auch hinsichtlich der ‚Hörerpräferenz‘.

Für den Datensatz des Hörvergleichs ‚ORCH 5.1‘ wurde außerdem eine Faktorenanalyse durchgeführt, deren Ergebnis nahe legt, dass die ersten drei Hauptfaktoren der Empfindung der ‚*Größe und Ausdehnung der Klangquelle*‘, der ‚*Deutlichkeit*‘ und der musikalisch/raumakustischen ‚*Einhüllung*‘ zugeordnet werden könnten.

Darüber hinaus wurden im Kapitel 8 eine Clusteranalyse und eine Korrelationsanalyse der Daten durchgeführt, deren Ergebnisse gut mit anderen Arbeiten übereinstimmen (vgl. hierzu

u.a. [Berg und Rumsey, 2001])

Eine nähere statistische Betrachtung bzgl. der Beurteilung des Signalinhalts der hinteren Kanäle in Abschnitt 8.6 brachte als Ergebnis, dass es vor allem beim KFM-System und – an zweiter Stelle – dem OCT-System die Signale der Kanäle LS, RS so beschaffen sind, dass diese für einen Teil der Testhörer als auffällig (und damit unerwünscht, d.h. negativ) hervortreten.

Aufgrund des technischen Aufbaus dieser beiden Systeme führt der Autor dies auf die hohe Korrelation der Signale – insbesondere auch bei tiefen Frequenzen – zurück, was sich (z.B. in Übereinstimmung mit der Arbeit von [Hidaka, Beranek und Okano, 1995]) negativ auf die empfundene Räumlichkeit in der Signalwiedergabe auswirkt, was weiter – durch die enge Korrelation zwischen dem Aspekt Raumabbildung und Präferenz bedingt – zu einer qualitativ schlechteren Einstufung in der Hörerbeurteilung führt.

Dass diese hohe Signalkorrelation im tieffrequenten Bereich bei den genannten Mikrofonssystemen tatsächlich vorliegt, wurde im zweiten Teil der Arbeit messtechnisch verifiziert.

In Abschnitt 8.8 wurde der Versuch einer qualitativen Reihung der untersuchten Surround-Mikrofonarrays – auch unter Zuhilfenahme der Ergebnisse anderer Hörvergleiche – unternommen.

Kapitel 9 beschäftigt sich mit den Ergebnissen des Hörvergleichs ‚ORCH-KK‘, bei dem die mittels Kunstkopf im Studio re-recordeten Signale der vier Surroundmikrofonssysteme der Original-Kunstkopfaufnahme aus dem Konzertsaal gegenübergestellt wurden. Die Wiedergabe erfolgte – sinngemäß – über Kopfhörer; die Hörerbeurteilung erfolgte in Form eines Tripel-Vergleichs (double-blind) mit Referenzstimulus, wobei den insgesamt 10 Testhörern bekannt war, welches der Signale der Referenz-Stimulus (Original Kunstkopf-Aufnahme aus dem Konzertsaal) ist.

Erwähnenswert scheint an dieser Stelle, dass sich durch die Reduktion auf zwei (binaurale) Kanäle beim Hörvergleich ‚ORCH-KK‘ für den Aspekt ‚Stabilität‘ bei allen Mikrofontechniken eine bessere Bewertung ergeben hat. Das ‚Einfrieren‘ der - mittels Kunstkopf-Rerecording neu abgebildeten vier Tonaufnahmen - scheint sich somit generell auf die Eindeutigkeit (Stabilität) aller verwendeten Mikrofontechniken positiv auszuwirken. (Ein Grund dafür könnte im Fehlen der Möglichkeit der Kopfdrehung und somit ev.

auftretender Ambiguitäten in der Signalperzeption durch den Testhörer liegen.)

Abschließend lässt sich feststellen, dass generell eine gute Übereinstimmung der Ergebnisse aus dem Hörvergleich der Kunstkopfsignale mit Kopfhörerwiedergabe (ORCH-KK) mit den Ergebnissen aus dem Hörvergleich mit Surround-Lautsprecheranordnung (ORCH-5.1) besteht.

Darüber hinaus soll an dieser Stelle angemerkt werden, dass der Hörvergleich ORCH-KK auch deshalb durchgeführt worden war, um zu überprüfen, ob die für eine qualitative Bewertung durch den Testhörer relevanten Signalinhalte auch im Signal der Kunstkopfaufnahme nach wie vor ausreichend vertreten sind, um sicherzustellen, dass eine messtechnische Auswertung desselben (in unserem Fall hinsichtlich der Signal-Korrelation bzw. Kohärenz) sinnvoll ist. Die Ergebnisse in Kapitel 9 bestätigen dieses.

Für eine weitere zu Testzwecken durchgeführte Surround-Aufnahme ‚DUO 5.1‘ wurde ebenfalls ein Hörvergleich mit 15 Probanden durchgeführt (Testbedingungen wie bei ‚ORCH 5.1‘), allerdings war diese Tonaufnahme – aus verschiedenen Gründen – unter (u.a. raumakustisch) ungünstigen Bedingungen entstanden, was auch in sehr breiten Konfidenzintervallen seinen Niederschlag findet. Lediglich der Vollständigkeit halber sind die Testbedingungen und Ergebnisse des Hörvergleichs im Anhang aufgeführt. Im Rahmen der DUO-Aufnahme wurde eine weitere neue Surround-Mikrofontechnik (BPT: ‚Blumlein-Pfanzagl-Triple‘) in Konkurrenz zu den bereits etablierten Techniken OCT, DECCA und KFM getestet.

Das vom Autor erstmals vorgeschlagene BPT Surround-Mikrofonsystem, welches in erster Linie für die koinzidente Aufnahme kleiner Schallquellen konzipiert ist (wenngleich es auch gut als Centerfill-System eines Groß-AB Systems zum Einsatz gebracht werden kann), basiert - wie auch das AB-PC-System - auf dem Grundprinzip möglichst großer Dekorrelation zwischen allen Kanälen. Anzumerken wäre in diesem Zusammenhang allerdings, dass beide Mikrofon-Arrays systembedingt eine relativ hohe Korrelation im mittleren Frequenzbereich zwischen dem jeweiligen Frontkanal L bzw. R und C aufweisen; zwischen den anderen Kanälen wird der Grundsatz der Dekorrelation eingehalten.

In Kapitel 10 wurde die frequenzabhängige Kohärenz und frequenzabhängige Korrelation für fünf Surround-Mikrofontechniken (OCT, DECCA, KFM, ABPC, BPT) untersucht. Die ersten vier Surroundsysteme kamen bei einer Orchesteraufnahme (ORCH 5.1) zum Einsatz, die letzte – ebenfalls in Konkurrenz mit den ersten drei Techniken- im Rahmen einer kammermusikalischen Aufnahme („DUO 5.1“).

Im Rahmen der vorliegenden Arbeit wurde eine der beiden neu vorgeschlagene Surround-Mikrofontechniken (AB-PC-Array), sowohl im Hörvergleich, als auch messtechnisch (FCC und FIACC, Kohärenz) drei bereits im Bereich der Tonaufnahme etablierten Systemen (OCT, DECCA, KFM) gegenübergestellt.

Aufgrund der Hörerbewertung darf diese als qualitativ mit den (bereits in anderen Untersuchungen – siehe z.B. [Camerer und Sodl, 2001] - sehr positiv bewerteten Mikrofonssystemen) OCT und DECCA auf gleicher Höhe eingestuft werden.

Darüber hinaus ist das AB-PC-Array unter den im Rahmen der vorliegenden Arbeit getesteten Surround-Systemen jenes, das die Korrelations-Eigenschaften des im Konzertsaal am Kunstkopf auftretenden Original-Schallfelds bei Wiedergabe über eine 5.1 Lautsprecher-Anlage gemäß ITU-R BS. 775-1 am besten approximiert (vgl. Tab. 10.3).

Somit lässt sich feststellen, dass die Ergebnisse der messtechnischen Auswertung in guter Übereinstimmung mit den Resultaten der Hörvergleiche in Kapitel 8 und 9 stehen, wenn man den engen Zusammenhang zwischen Hörerpräferenz, Natürlichkeit, Basisbreite und Raumeindruck berücksichtigt, der – wie in Abschnitt 8.7 dargestellt – bereits in anderen Arbeiten nachgewiesen werden konnte.

Die Gültigkeit der Ergebnisse der Hörvergleiche wurde in den Kapiteln 8 und 9 mit Verweis auf die gute Übereinstimmung mit anderen Untersuchungen belegt (vgl. hierzu u.a. [Camerer und Sodl, 2001]).

Die Ergebnisse der vorliegenden Untersuchung deuten darauf hin, dass statt Mikrofonssystemen, die durch hohe Signalkorrelation gekennzeichnet sind, solche mit geringer Korrelation verwendet werden sollten, sofern ein guter Raumeindruck erwünscht ist.

Vorliegende Untersuchungen zur Frage des optimalen Korrelationsgrades bei der Wiedergabe von Surround-Signalen (siehe Abschnitt 3.3) brachten das Ergebnis, dass ein geringerer

Korrelationsgrad eine bessere Kompatibilität zwischen unterschiedlichen Wiedergabe-Umgebungen und eine Verbreiterung des ‚sweet spot‘ bewirkt (siehe hierzu [Nakahara, 2005] und [Prokofieva, 2007]).

Für die in Abschnitt 10.8 betrachteten Stereo-Aufnahmen wurde kein Hörvergleich durchgeführt, allerdings lassen sich aus den im Rahmen des Surround-Hörvergleichs gewonnenen Erkenntnissen auch Schlüsse für Stereo-Hauptmikrofonsysteme ableiten. In diesem Zusammenhang stellt sich z.B. auch die Frage, wie eine ‚gesunde‘ FCC-Funktion eigentlich beschaffen sein sollte:

11.3 Schlussfolgerung

Um beim Hörer einen guten Raumeindruck erzielen zu können, ist Dekorrelation der Signale bei tiefen Frequenzen eine der Voraussetzungen. Zusätzlich sollten diese auch bei Frequenzen oberhalb von 1kHz weitgehend dekorreliert sein, da hohe Korrelation im Lautsprechersignal (FCC) auch zu einer ungünstig hohen Korrelation des FIACC führt, wie sich anhand der Abb. 10.25-10.34 nachvollziehen lässt.

Sowohl für die vorliegend untersuchten Surround-Mikrofonsysteme, als auch für diverse Stereo-Hauptmikrofonsysteme wurde gezeigt, welche von ihnen den im letzten Absatz formulierten Bedingung für eine gute räumliche Wiedergabe gerecht werden, und welche nicht (siehe Abschnitt 8.8, sowie die Zusammenfassung von Abschnitt 10.7).

Dies lässt auch die Schlussfolgerung zu, dass vor allem beim Einsatz vieler Stützmikrofone darauf geachtet werden muss, dass die Kombination ihrer Signale bei der Tonabmischung nicht zu einer zu hohen Signalkorrelation im Bereich tiefer Frequenzen führt. (Anm.: im tieffrequenten Bereich wird die Charakteristik von Richtmikrofonen oftmals nicht eingehalten und tendiert gegen ‚omni‘; weiters erfolgt die Schallausbreitung bei tiefen Frequenzen physikalisch bedingt omnidirektional, was ebenfalls ein erhöhtes ‚Übersprechen‘ zwischen den Stützmikrofonen in diesem Bereich zur Folge hat). Dem kann durch geeignete Filterung (Pegelabsenkung im tieffrequenten Bereich mittels Hochpass- oder Shelving-Filter) der Stützmikrofonsignale abgeholfen werden, was allerdings in der Praxis nicht von allen Tonmeistern berücksichtigt wird; ein Sachverhalt, der anhand vieler Tonaufnahmen mit ‚Standard-Industrieklangbild‘ – gehörmässig, wie auch messtechnisch - überprüft werden kann.

Im Rahmen der Untersuchung stellte sich weiters heraus, dass die Einführung eines ‚Binaural Quality Index for reproduced music‘ (BQI_{rep}) – in Anlehnung an den in der Konzertsaalakustik bekannten ‚Binaural Quality Index‘ (siehe [Beranek, 2004]) – auch im Zusammenhang mit der Wiedergabe von Musik in kleinen Räumen Sinn machen könnte.

11.4 Ausblick

Die Frage, inwiefern die frequenzabhängige Richtcharakteristik der verwendeten Lautsprecher den beim Hörer in einem (nicht schalltoten) Abhörraum auftretenden Eindruck musikalisch/akustischer ‚Einhüllung‘ bei Surround-Beschallung beeinflusst, wurde in Abschnitt 3.4 kurz angesprochen. Weiterführende Forschungsarbeit in diese Richtung brächte vermutlich wichtige Erkenntnisse, die auch auf das technische ‚Design‘ zukünftiger Surround-Mikrofontechniken Auswirkungen haben könnten.

Abschließend lässt sich feststellen, dass im Rahmen der vorliegenden Dissertation aufgezeigt wurde, welche signaltheoretischen Aspekte berücksichtigt werden müssen, um bei Surround- und Stereoaufnahmen einen besseren Raumeindruck zu erzielen. Diese Ergebnisse stehen teilweise im Widerspruch zur Lehrmeinung einzelner Fachleute (vgl. u. a. [Theile, 2000], S. 418/419), aber auch in guter Übereinstimmung mit den Ergebnissen einer Vielzahl anderer facheinschlägiger Arbeiten, wie ebenfalls dargestellt werden konnte.

Somit hätte die Tonträgerindustrie ein neues Hilfsmittel in Form von adäquateren Mikrofonierungstechniken und eine klare Richtlinie an der Hand, um in Zukunft qualitativ hochwertigere Tonaufnahmen im Bereich akustischer Musik realisieren zu können.

Literaturverzeichnis:

Akeroyd M. A., and Summerfield A. Q. (1999): "A Binaural Analog of Gap Detection", *J. Acoust. Soc. Am.*, vol. 105, pp. 2807-2820, 1999.

Akeroyd M. A., and Bernstein L. R. (2001): "The Variation Across Time of Sensitivity to Interaural Disparities: Behavioural Measurements and Quantitative Analyses", *J. Acoust. Soc. Am.*, vol. 110, pp. 2516-2526, 2001.

Augspurger G. L. (1990): "Loudspeakers in Control Rooms and Listening Rooms", 8th AES Int. Conference, 1990

Avni, A und Rafaely, B. (2009): "Inter-aural cross correlation in a sound field represented by spherical harmonics", *J. Acoust. Soc. Am.*, 125(4): 2545, April 2009.

Barron, M. (1971): "The Subjective Effects of First Reflections in Concert Halls – The need for Lateral Reflections", *J. Sound Vib.*, 15, pp. 475-494, 1971

Barron, M. and Marshall, A. H. (1981): "Spatial Impression Due to Early Lateral Reflections in Concert Halls: The Derivation of a Physical Measure", *J. Sound and Vibration*, vol. 77, pp.211-232 (1981)

Batteau, D. W. (1967): "The Role of the Pinna in Human Localisation", *Proceedings of the Royal Society*, B168 (1011), pp.158-180, 1967

Bauer, B. (1981): "Microphone system for producing signals for surround-sound transmission and reproduction", US-Pat. No. 4.262.170, (1981)

Bech S., et al. (1995): "Interaction between audio-visual factors in home theatre systems: definition of subjective attributes", Preprint 4133, AES 99th Convention New York (1995)

Bech S. (1998): "The Influence of Stereophonic Width on the Perceived Quality of an Audiovisual Presentation Using a Multichannel Sound System", *J. Audio Eng. Soc.*, vol.46,

No.4, pp. 314-322, April 1998

Bech S. (1999): "Methods for Subjective Evaluation of Spatial Characteristics of Sound", Paper 16-044, The AES 16th International Conference on Spatial Sound Reproduction, pp. 487-504, April 1999

Bech S., und Zacharov N. (2006): "Perceptual Audio Evaluation – Theory, Method and Application", Wiley and Sons, ISBN-13: 978-0-470-86923-0, (2006)

Begault D. (1994): "3-D Sound for Virtual Reality and Multimedia", *Academic Press*, USA (1994)

Benjamin, E. (2006): "An Experimental Verification of Localization in Two-Channel Stereo", Preprint 6968, 121st AES-Convention, 2006

Benjamin, E., Brown, R. (2007): "The effect of head diffraction on stereo localisation in the mid-frequency range", Preprint 7018, 122nd AES-Convention, Vienna, 2007

Beranek, L. L. und Schultz, Th. J. (1965): "Some recent experiences in the design and testing of concert halls with suspended panel arrays", *Acustica* vol. 15, S.307 (1965)

Beranek, L. (2004): "Concert Halls and Opera Houses: Music, Acoustics and Architecture", (Second Edition), Springer-Verlag, New York, 2004.

Berg, J. and Rumsey, F. (1999): "Identification of perceived spatial attributes of recordings by repertory grid technique and other methods", AES 106th Convention, Preprint #4929, May 1999

Berg, J. and Rumsey, F. (2001): "Verification and Correlation of Attributes used for Describing the Spatial Quality of Reproduced Sound", Preprint, AES 19th Int. Conference, (2001)

Berg, J. and Rumsey, F. (2006): "Identification of Quality Attributes of Spatial Audio by Repertory Grid Technique", *J. Audio Eng. Soc.*, vol. 54, No. 5 (2006)

Bernstein, L. R., Trahoitis, C., Akeroyd, M. A., and Hartung, K. (2001): "Sensitivity to Brief Changes of Interaural Time and Interaural Intensity", *J. Acoust. Soc. Am.*, vol. 109, pp. 1604-1615, 2001

Betz, G. (2000): "Surroundaufnahmen – Praktische Erfahrungen", Beitrag zur 21. Tonmeistertagung des VDT, Tagungsbericht (ISBN 3-598-20362-4), S.485-494, Hannover 2000

Blauert, J. (1972): "On the Lag of Lateralization Caused by Interaural Time and Intensity Differences", *Audiology*, vol. 11, pp.265-270, 1972.

Blauert, J. (1974): "Räumliches Hören", ISBN 3-7776-0250-7, S. Hirzel Verlag, Stuttgart, 1974.

Blauert, J. (1997): "Spatial Hearing", ISBN 0-262-02413-6, The MIT Press, 1997

Blumlein, A. D. (1931): "Improvements in and relating to Sound-transmission, Sound-recording and Sound-reproducing Systems", British patent No. 394325, Dec.14 1931, (reprinted in the "Anthology of Stereophonic Techniques", AES, 1986, pp.32-40)

Boehnke, S. E., Hall S. E., and Marquadt T. (2002): "Detection of Static and Dynamic Changes in Interaural Correlation", *J. Acoust. Soc. Am.*, vol. 112, pp. 1617-1626, 2002.

Bortz, J. (2005): "Statistik für Human- und Sozialwissenschaftler", 6. Auflage, Springer Verlag, ISBN 3-540-21271-X, 2005

Bradley, J. and Soulodre, G. (1995): "Objective Measures of Listener Envelopment", *J. Acoust. Soc. Amer.*, vol. 98, pp.2590-2597, (1995)

Braun, D., Gutzke, K., Wönicker, C. (1996): "Ein objektives Verfahren zur Beurteilung der räumlichen Abbildung von Lautsprechern", 19. Tonmeistertagung des VDT in Karlsruhe, S. 630-641, 1996

Breebart, J., van der Par, S., and Kohlrausch, A. (2002): "A Time-Domain Binaural Signal Detection Model and Its Predictions for Temporal Resolution Data", *Acta Acustica-Acustica*, vol. 88, pp.110-112, 2002.

Bruck, J. (1996): "Solving the Surround Dilemma", Paper zur 19. Tonmeistertagung des VDT, S.117-124, ISBN 3-598-20360-8 (1996)

Burkhard, M., Bray, W., Genuit, K., and Gierlich, H. W. (1991): "Binaural Sound for Television", *Proc. of the AES 9th Int. Conference "Television Sound Today and Tomorrow"*, Detroit, Feb. 1991

Camerer, F. und Sodl, C. (2001): "Classical Music in Radio and TV - a multichannel challenge", The IRT/ORF Surround Listening Test, <http://www.hauptmikrofon.de/orf.htm>, 2001 (Datum der Recherche: Sept. 2008)

Cattell, R.B. (1996): "The scree test for the number of factors", *Multivariate behav. Res.*, vol. 1, 245-276 (1966)

Ceoen, C. (1972): "Comparative Stereophonic Listening Tests", *J. Audio Eng. Soc.*, Vol. 20, no.1, Jan./Feb. 1972

Ceoen, C. (1974): "Correlation and Virtual Sound Sources", Preprint zur 47th Convention der Audio Engineering Society in Kopenhagen, 1974

Choisel, S. und Wickelmaier, F. (2006): "Extraction of Auditory Features and Elicitation of Attributes for the Assessment of Multichannel Reproduced Sound", *J. Audio Eng. Soc.*, vol. 54, No.9 (2006)

Christensen, L. (2003): "A Surround Microphone Setup for Classical Music", (DPA Microphones Workshop on Mic Techniques for Multichannel Audio), Audio Eng. Soc. 24th Int. Conference in Banff, Canada, 2003

Cohen, E. and Eargle, J. (1995): "Audio in a 5.1 Channel Environment", Preprint #4071, 99th AES-Convention in New York, Oct. 1995

Cook, R. K., Waterhouse, R.V., Berendt, R.D., Edelman, S. and Thompson Jr., M.C. (1955): "Measurement of Correlation Coefficients in Reverberant Sound Fields", *Journal Acoust. Soc. Amer.*, vol.27, no.6, Nov. 1955, pp.1072

Corey, J. and Martin, G. (2003): "Description of a five-channel Microphone Technique", Paper at the 'DPA Microphones Workshop on Mic Techniques for Multichannel Audio', Audio Eng. Soc. 24th Int. Conference in Banff, Canada, 2003

Cremer, L. (1976): „Zur Verwendung der Worte ‘Korrelationsgrad’ und ‘Kohärenzgrad’” (German), *ACUSTICA*, vol.35, pp.215-218, 1976

Cremer, L. und Müller, H. A. (1978): "Die wissenschaftlichen Grundlagen der Raumakustik", (Band 1), Hierzel-Verlag Stuttgart, (1978)

Culling J. F., and Summerfield A. Q. (1999): "Measurement of the Binaural Temporal Window Using a Detection Task", *J. Acoust. Soc. Am.*, vol. 103, pp. 3540-3553, 1999.

Damaske, P. (1967/68): "Subjektive Investigation on Sound Fields" (German) *ACUSTICA*, vol. 19 (1967/68), pg. 199-213

de Boer, K. (1940): "Plastische Klangwiedergabe", Philips Technische Rundschau, 5.Jahrgang, Heft 4, 1940

Dickason, V. (2000): "The Loudspeaker Design Cookbook", Audio Amateur Publ., 2000

Eargle, J. (2004): "The Microphone Book" (Second Edition), Focal Press (2004)

Elko, G. W. (2001): "Spatial Coherence Functions for Differential Microphones in Isotropic Noise Fields", in: 'Microphone Arrays' (M. Brandstein, D. Ward, Editors), Springer, 2001, (ISBN 3-540-41953-5), pp.61

Elko, G., Kubli, R., Meyer, J. (2003): "Audio system based on at least second-order Eigenbeams", International Patent Publication Number: WO 03/061636 A1, (2003)

Engineering (1981): "100 Years with Stereo: The Beginning", *Journ. Audio Eng. Soc.*, vol. 29, no.5, pp.368-372, May 1981

Everest, F. A. (1994): "The Master Handbook of Acoustics", 3rd. Edition, ISBN 0-8306-4437-7, TAB Books (McGraw-Hill), 1994

Evers, H. (2003): "Swan Lake in 5.1", presentation at the DPA Microphones Workshop on Mic Techniques for Multichannel Audio, AES 24th Int. Conference in Banff, Canada, 2003

Eyfert, K. und Baltes, P.B. (1969): "Über Normierungseffekte in einer Faktorenanalyse von Fragebogendaten", *Zeitschr. f. exp. u. angew. Psychol.* 16, 38-51 (1969)

Faller, C., and Merimaa, J. (2004): "Source Localization in Complex Listening Situations: Selection of Binaural Cues Based on Interaural Coherence", *J. Acoust. Soc. Am.*, vol. 116, pp. 3075-3089, 2004

Faulkner, T. (1981): "A Phased Array", *Hi Fi News Record Rev.*, pg. 44ff, July 1981

Fellner, M. und Höldrich, R. (1998a): "Physiologische und psychoakustische Grundlagen des räumlichen Hörens", IEM-Report 03/98 der KUG: Univ. für Musik und darst. Kunst, Graz, (1998)

Fellner, M. und Höldrich, R. (1998b): "Außenohr-Übertragungsfunktion – Messung und Datensätze", aus: IEM-Report 04/98 der KUG: Univ. für Musik und darst. Kunst, Graz (1998)

Ferguson, S., and Cabrera, D. (2005): "Vertical Localisation of Sound from Multiway Loudspeakers", *Journal of the Audio Eng. Soc.*, vol..53, No.3, pg.163-173, March 2005

Fox, C. and McGregor, W. (2002): "A Modular Microphone Array for Surround Sound Recording", Paper # 5566, 112th AES-Convention, Munich, May 2002

Gabrielsson, A. und Sjören, H. (1979): "Perceived Sound Quality of Sound Reproducing Systems", *Journal Acoust. Soc. Am.*, vol. 65, pg. 1019-1033 (1979)

Gabrielsson, A. und Lindström, B. (1985): "Perceived Sound Quality of High-Fidelity Loudspeakers", *JAES* vol. 33, No. 1/2, pg. 33 (1985)

Gernemann A., (1994): "Summenlokalisierung im Stereodreieck – Überlegungen zu psychoakustischen Untersuchungen mit dynamischem Testsignal und hochpräzisen Schallwandlern", Eigenverlag, Düsseldorf, 1994

Gernemann, A., und Rösner, T. (1998): "Die Abhängigkeit der stereophonen Lokalisation von der Qualität der Wiedergabelautsprecher", *Proceedings, 20. Tonmeistertagung*, pp.828, Karlsruhe 1998

Gerzon, M. (1973): "Periphery: With-Height Sound Reproduction", *J. Audio Eng. Soc.*, vol. 21, No. 1 (1973)

Gerzon, M. (1986): "Stereo shuffling: new approach, old technique", *Studio Sound*, July 1986

Glasberg, B. R., and Moore, B. C. J. (1990): "Derivation of Auditory Filter Shapes from Notched-Noise Data", *Hear. Res.*, vol.47, pp.103-138, 1990.

Gottlob, D. (1973): "Vergleich objektiver akustischer Parameter mit Ergebnissen subjektiver Untersuchungen an Konzertsälen", Dissertation an der Universität Göttingen, 1973

Grantham, D. W., and Wightman, F. L. (1978): "Detectability of Varying Interaural Temporal Differences", *J. Acoust. Soc. Am.*, vol. 63, pp. 511-523, 1978.

Grantham D. W., and Wightman F. L. (1979): "Detectability of a Pulsed Tone in the Presence of a Masker with Time-Varying Interaural Correlation", *J. Acoust. Soc. Am.*, vol. 65, pp. 1509-1517, 1979.

Griesinger, D. (1986): "Spaciousness and localization in listening rooms and their effects on

the recording technique“, *J. Audio Eng. Soc.*, vol. 34, no.4, pp.255-268, 1986

Griesinger, D. (1987): “New Perspectives on Coincident and Semi-Coincident Microphone Arrays“, Preprint 2464, 82nd 85th Int. Convention der Audio Eng. Society, May 1987

Griesinger, D. (1988): ”Equalization and Spatial Equalisation of Dummy Head Recordings for Loudspeaker Reproduction“, Preprint # 2704, 85th Int. Convention der Audio Eng. Society, Los Angeles, 1988

Griesinger, D. (1996): ”Spaciousness and Envelopment in Musical Acoustics“, 19th Int. Conf. on Sound Design VDT, Proceedings (ISBN 3-598-20360-8), S.375-391 (1996)

Griesinger, D.(1997): "Spatial Impression and Envelopment in Small Rooms", paper #4638 (H-2), 103rd AES Convention, 1997

Griesinger, D. (1998): ”General Overview of spatial impression, envelopment, localization and externalization“ Proc. 15th Int. Conf AES on Small Room Acoustics, pp. 136-149, Denmark Oct/Nov 1998

(Internetrecherche: siehe www.world.std.com/~griesngr am 22.8.2008)

Griesinger, D. (1999): “Objective Measures of Spaciousness and Envelopment“, Paper 16-003, 16th Int. AES Conference on Spatial Sound Reproduction, (1999)

Griesinger, D. (2000): "The Theory and Practice of Perceptual Modeling – How to use Electronic Reverberation to Add Depth and Envelopment Without Reducing Clarity", paper presented at the 21. Tonmeistertagung of the VDT, Hannover 2000, pp.766-795 (see also www.world.std.com/~griesngr)

Griesinger, D. (2002):"Stereo and Surround Panning in Practice", Preprint No. 5564, 112th AES-Convention, München, Mai 2002

Guttman, L. (1954): ”Some necessary conditions for common factor analysis“, *Psychometrika*, vol. 19, 149-161 (1954)

Hall, D. E. (1980): "Musical Acoustics ", Brooks/Cole Publ. Company, Pacific Grove, California, (1980); (Anm.: deutsche Übersetzung erschienen als "Musikalische Akustik – ein Handbuch", Schott-Verlag)

Hamasaki, K., and Hiyama, K. (2003): "Reproducing Spatial Impression with Multichannel Audio", AES 24th Int. Conference, Banff (Canada), 2003

Hammershoi, D., Moller, H. (2005): "Binaural Technique – Basic Methods for Recording, Synthesis, and Reproduction", from Communication Acoustics (Blauert, J., Editor), pg.232, Springer 2005, ISBN-13: 978-3-540-22162-3

Hawkes, R.J. und Douglas, H. (1971): "Subjective acoustic experience in concert auditoria", *Acustica*, vol. 24, S. 235 (1971)

Haas, H. (1951): "The Influence of a Single Echo on the Audibility of Speech" (German), *Acoustica*, vol. 1, no.2, (1951)

Hecker, P. (2000): "The Decision of the Microphone Spacing and its Creative Benefit", (German) Paper zur 21. Tonmeistertagung des VDT, Hannover 2000, pp.796-804, Proceedings (ISBN 3-598-20362-4)

Hebrank, J. and Wright, D. (1974): "Spectral cues in the localization of sound sources on the median plane ", *Journ. Acoust. Soc. Amer.*, vol. 56, no.3, pp. 1829-1834, (1974)

Herrmann, U., Henkels, V. und Braun, D. (1998): "Vergleich von 5 Surround-Mikrofonverfahren", Paper zur 20. Tonmeistertagung Karlsruhe, S. 508ff, ISBN 3-598-20361-6 (1998)

Hesselmann, N. (1993): "Digitale Signalverarbeitung – Rechnergestützte Erfassung, Analyse und Weiterverarbeitung analoger Signale", Vogel Verlag, Würzburg (1993)

Hidaka, T., Beranek L., and Okano, T. (1995): "Interaural cross-correlation, lateral fraction, and low- and high-frequency sound levels as measures of acoustical quality in concert halls", *J. Acoust. Soc. Am.* 98 (2), August (1995)

Hidaka, T., Masuda, S., Arai, T. and Nakajima, H. (1996): "Acoustical Measurement of Grosses Festspielhaus, Salzburg", Messdokumentation/Report des Takenaka Research & Development Institute, Japan; 1. Juli (1996)

Hidaka, T., Beranek; L. and Okano, T. (1997): "Some Considerations of Interaural Cross Correlation and Lateral Fraction as Measures of Spaciousness in Concert Halls", Chapter 32 im Buch 'Music and Concert Hall Acoustics', Y. Ando and D. Noson (Editors), Academic Press, London (1997)

Hildebrandt, A. und Braun D. (2000): "Untersuchungen zum Centerkanal im 3/2 Stereo-Format", Paper zur 21. Tonmeistertagung des VDT, ISBN 3-598-20362-4, S. 455ff (2000)

Hirata, Y. (1983): "Improving stereo at L.F.", *Wireless World*, pp.60, Oct. 1983

Hiroyuki, N., Michio, M., Katsunori, F., Isanaga, Y. (1980): "Microphone", JPO-Patent application number 55-117538 (Japanisches Patentamt), dated 26.08.1980

Hiyama, K., Komiyama, S., Hamasaki, K. (2002): "The minimum Number of Loudspeakers and Its Arrangement for Reproducing the Spatial Impression of Diffuse Sound Field", paper 5674, AES 113th Convention, Los Angeles, Oct. 2002

Holman, T. (2000): "Comments on the 'Subjective Appraisal of Loudspeaker Directivity for Multichannel Reproduction' ", *J. Audio Eng.Soc.*, vol.48, No. 4, pg.314-317, April 2000

Holman, T. (2001): "Mixing the Sound (Part 2): Perspective – where do the sounds go ?", *Surround Professional*, May/June 2001, pg. 35

Holophone (2008): Stand der Internet-Recherche: (Juni 2008), siehe www.holophone.com

Holube, I., Kinkel, M., and Kollmeier, B. (1998): "Binaural and Monaural Auditory Filter Bandwidths and Time Constants in Probe Tone Detection Experiments", *J. Acoust. Soc. Am.*, vol. 104, pp. 2412-2425, 1998.

Hugonnet, C. and Walder, P. (1998): "Stereophonic Sound Recording", John Wiley & Sons, 1998

Huopaniemi, J. (1999): "Virtual Acoustics and 3D sound in Multimedia Signal Processing", PhD Thesis, Helsinki University of Technology, 1999

Iredale, J. J. and Keller, R.S. (2003): "Microphone Apparatus for Producing Signals for Surround Reproduction", US-Patent # 6507659, Jan. 2003

Ito, A. (2001): "Surround sound field reproduction system and surround sound field reproduction method", European Patent No. EP 1 259 097 A2, 2001

ITU-R Recommendation BS.1116-1 (1997): "Methods for the Subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems", Int. Telecommunications Union Radiocommunication Assembly, 1997

ITU-R Recommendation BS.775-1 (1994): "Multichannel Stereophonic Sound Systems with and without Accompanying Picture", Int. Telecommunications Union Radiocommunication Assembly, 1994

Jackobsen; F. and Roisain, Th. (2000): "The coherence of reverberant sound fields", *J. Acoust. Soc. Am.*, vol. 108 (1), July 2000

Jecklin, J. (1981): "A Different Way to Record Classical Music", *J. Audio Eng. Soc.*, vol. 29, no. 5, pp.329-332, May 1981

Jecklin, J. (1998): persönliches Gespräch mit Jürg Jecklin am 'Inst. für Elektroakustik' der Univ. für Musik und darstl. Kunst, Wien im Jahre 1998

Jecklin, J. (2002): "Surround-Aufnahmetechnik OSIS 321", Bericht zur 21. Tagung des VDT, Hannover, Nov. 2002

Julstrom, S. (1991): "An Intuitive View of Coincident Stereo Microphones", *J. Audio Eng.Soc.*, vol. 39, no.9, Sept. 1991

Kaiser, H. F. und Dickman, K. (1959): "Analytic determination of common factors", *Amer. Psychol.*, vol. 14, 425ff (1959)

Kamekawa, T., Marui, A. and Irimajiri, H. (2007): "Correspondence Relationship between Physical Factors and Psychological Impressions of Microphone Arrays for Orchestra Recording", paper # 7233, 123rd AES-Convention, New York, Oct. 2007

Kates, J. M. (1960): "Optimum Loudspeaker Directional Patterns", *J. Audio Eng. Soc.*, vol.28, pp.787-794, 1960

Keet, W. de V. (1968): "The influence of early lateral reflections on spatial impression", 6th Int. Congress on Acoustics, Tokyo, (1968)

Keller, A.C. (1981): "Early Hi-Fi and Stereo Recording at Bell Laboratories (1931-1932)", *Journ. Audio Eng. Soc.*, vol. 29, pp.274-280, April 1981

Kelley, T. L. (1935): "Essential traits of mental life", *Harvard Stud. in Educ.* Vol. 26, Cambridge, Mass.: Harvard Univ. Press (1935)

King, R. (2003): "A Five-Microphone Technique for Music Recording in a Large Venue", DPA Microphones Workshop on Mic Techniques for Multichannel Audio, AES 24th Int. Conference in Banff, Canada, 2003

Kleiner, M. (1989): "A new way of measuring the lateral energy fraction", *Appl. Acoust.* 37, 1989, pp.321-327

Klepko, J. (1997): "5-Channel Microphone Array with Binaural Head for Multichannel Production", *J. Audio Eng. Soc.*, vol. 45, S.127 ff, Nov. 1997

Knothe, J., Plenge, G. (1978): "Panoramaregler mit Berücksichtigung der frequenzabhängigen Pegeldifferenzbewertung durch das Gehör", Bericht zur 11. Tonmeistertagung des VDT, Berlin, S.136 ff, 1978

Kohlrausch, A. (1988): "Auditory Filter Shape Derived from Binaural Masking Experiments", *J. Acoust. Soc. Am.*, vol. 84, pp. 573-583, 1988.

Kollmeier, B., and Gilkey, R. H. (1990): "Binaural Forward and Backward Masking: Evidence for Sluggishness in Binaural Detection", *J. Acoust. Soc. Am.*, vol. 87, pp. 1709-1719, 1990

Kollmeier, B., and Holube, I. (1992): "Auditory Filterbandwidths in Binaural and Monaural Listening Conditions", *J. Acoust. Soc. Am.*, vol. 92, pp. 1889-1901, 1992

Laborie, A., Montoya, S., Remy, B.(2004) : "High Spatial Resolution Multichannel Recording", Paper #6116, 116. AES-Convention, May 2004

La Grou, J. (1994): "Orchestral Recording ", *Mix*, Feb. 1994, pp. 32

Langen, C. (2000): "Signal Processing for the KFM 360", paper (German)
21. Tonmeistertagung des VDT, Hannover 2000, pg.329-346 (ISBN 3-598-20362-4)

Lehmann, U. (1975): "Untersuchung zur Bestimmung des Raumeindrucks bei Musikdarbietungen und Grundlagen der Optimierung", Dissertation an der TU Dresden, 1975

Lehmann, P. (1976): "Über die Ermittlung raumakustischer Kriterien und deren Zusammenhang mit subjektiven Beurteilungen der Hörsamkeit", Dissertation an der TU Berlin, 1976

Lindemann, W. (1986a): "Extension of a Binaural Cross-Correlation Model by Means of Colateral Inhibition. I. Simulation of Lateralization for Stationary Signals", *J. Acoust. Soc. Am.*, vol. 80, pp. 1623-1630, 1986

Lindemann, W. (1986b): "Extension of a Binaural Cross-Correlation Model by Means of Colateral Inhibition. II. The Law of the First Wavefront", *J. Acoust. Soc. Am.*, vol. 80, pp. 1623-1630, 1986.

Litovsky, R. Y., Colburn H. S., Yost W. A, and Guzman S. J. (1999): "The Precedence Effect", *J. Acoust. Soc. Am.*, vol. 106, pp. 1633-1654, 1999.

Marshall, A.H. (1968): "Acoustical Determinants for the Architectural Design of Concert Halls", *Arch. Sci. Rev.*, Australia 11, 1968, pp.81-87

Martin, G., Woszczyk, W., Corey, J., and Quesnel, R. (1999): "Sound source localisation in a five channel surround sound reproduction system", Preprint 4994, 107th AES-Convention, New York, Sept. 1999

Martin, G. (2005): "A New Microphone Technique For Five-Channel Recording", paper # 6427, 118th AES-Convention, Barcelona 2005

Mason, R. and Rumsey, F. (2000): "An assessment of the spatial performance of virtual home theatre algorithms by subjective and objective methods", Preprint #5137, 108th AES Convention, Paris, May (2000)

Mason, R. and Rumsey, F. (2002): "A comparison of objective measurements for predicting selected subjective spatial attributes", Preprint # 5591, 112th-AES Convention, München 2002

Maxfield, I. and Albasheim, W. (1947): "An Acoustic Constant of Enclosed Spaces Correlatable with Their Apparent Liveness", *J. Acoust. Soc. Amer.*, Vol. 19, S.71-79 (1947)

McKinney, D., and Rumsey, F. (1997): "Coincident Microphone Techniques for Three Channel Stereophonic Reproduction", Preprint #4429 (E2) zur 102. AES-Convention in München, März 1997

Mehrgardt, S. und Mellert, V. (1977): "Transformation Characteristics of the External Human Ear", *J. Acoust. Soc. Amer.*, vol. 61, 6 (1977), pg. 1567-1576

Meindl, M., Vette U. und Görne T. (2006): "Investigations on the Effect of Surround Microphone Setup on Room Perception", Paper zur 28. Int. Conf. der Audio Eng. Soc., in Pitea (Schweden), Juni/Juli 2006

Merimaa, J., and Pulkki, V. (2005): "Spatial Impulse Response Rendering I: Analysis and Synthesis", *J. Audio Eng. Soc.*, Vol. 53, No.12, December 2005

Mershon, D. H. and Bowers, J. N. (1979): "Absolute and Relative Cues for the Auditory Perception of Egocentric Distance", *Perception*, vol. 8, pp. 311-322, (1979)

Meyer, J. and Agnello, T. (2003): "Spherical Microphone Array for Spatial Sound Recording", Preprint # 5975, 115th AES Convention, New York (2003)

Miller, R. (2004a): "Spatial Definition and the PanAmbiophone microphone array for 2D surround & 3D fully periphonic recording", Convention Paper, 117th AES Convention San Francisco, 2004

Miller, R. E. (2004b): "System and Method for compatible 2D/3D (Full Sphere with Height) Surround Sound Reproduction", US-Patent US 2004/0247134, (2004)

Mitchell, D. (1999): "Tracking for 5.1 – Surround-Recording Techniques", *Audio Media*, Nov. 1999

Moller, H., Hammershoi, D., Jensen, C. B. and Sorensen, M. F. (1995): "Transfer Characteristics of Headphones Measured on Human Ears", *J. Audio Eng. Soc.*, vol. 43, pp. 203-217, April 1995

Moller, H., Hammershoi, D., Jensen, C. B. and Sorensen, M. F. (1999): "Evaluation of Artificial Heads in Listening Tests", *J. Audio Eng. Soc.*, vol. 47, No.3, pg.83-100, (1999)

Moore B. C. J. (1997): "An Introduction to the Psychology of Hearing", 4th ed., Academic Press, London UK, 1997.

Morimoto, M. and Meakawa, Z. (1988): "Effects of Low Frequency Components on Auditory Spaciousness", *Acustica*, Vol. 66, 1988

Morimoto, M., Yairi, M., Iida, K. and Itoh, M. (2003): "The Role of Low Frequency Components in Median Plane Localization", *Acoust. Sci. Technol.*, vol. 24, pp 76-82, 2003

Moulton D., Ferralli M., Hebrock S., and Pezzo M. (1986): “ The Localization of Phantom Images in an Omnidirectional Stereophonic Loudspeaker System“, Preprint 2371, 81st-AES Convention, 1986

Moulton, D. (1995): “ The Significance of Early High-Frequency Reflections from Loudspeakers in Listening Rooms“, Preprint 4094, 99th-AES Convention, 1995

Moulton, D. (1998): “The Use of an Acoustic Lens to Control the High Frequency Dispersion of Conventional Soft Dome Radiators”, Paper presented at the American Loudspeaker Manufacturers Association Symposium, Las Vegas, Jan. 1998

Muraoka, T., Miura, T., Ifukuba, T., (2007): “Ambience sound recording utilizing dual MS (Mid-Side) microphone systems based upon Frequency dependent Spatial Cross Correlation (FSCC)”, preprint 6997, 122nd AES-Convention, Vienna 2007

Nakahara, M. (2005): “Multichannel Monitoring Tutorial Booklet (M2TB) rev. 3.5.2, Yamaha Corp. 2005, SONA Corporation 2005, pg. 41

Nakayama, T., Miura, T. , Kosaka, O., Okamoto, M., Shiga, T. (1971): ”Subjective Assessment of Multichannel Reproduction”, *J. Audio Eng. Soc.*, vol. 19, no. 9, pp.744 –751, October (1971)

Nielsen, S. H. (1993): “Auditory Perception in Different Rooms“, *J. Audio Eng. Soc.*, vol. 41, no.10, October (1993)

Olive S.E., Schuck P. L. and Sally S. L. and Bonneville M. E. (1994): “The Effects of Loudspeaker Placement on Listener Preference Ratings”, *J. Audio Eng. Soc.*, Vol.42, No.9, pg. 651-669, Sept. 1994

Olive, S. (2003): “Differences in Performance and Preference of Trained versus Untrained Listeners in Loudspeaker Tests: A Case Study”, *J. Audio Eng. Soc.*, vol. 51, pp.808-825 September (2003)

Peters, N., McAdams, S., Braasch, J. (2007): "Evaluating Off-Center Sound Degradation in Surround Loudspeaker Setups for various Multichannel Microphone Techniques", Paper 7197, 123rd AES-Convention, New York 2007

Pfanzagl-Cardone, E. (2002): "In the light of 5.1 Surround: Why AB-PC is Superior for Symphony -Orchestra Recording", paper 5565, 112th AES-Convention Munich, 2002

Pfanzagl-Cardone, E. (2005): "3.0 Microphone for Surround-Recording", Australian patent No. AU2005100255, www.ipaustralia.gov.au (Datum der Internet-Recherche 2.1.2010; Link: <http://pericles.ipaustralia.gov.au/aub/pdf/nps/2005/0526/2005100255A4/2005100255.pdf>)

Pfanzagl, E. (2007): "Vom Punkt zum Flächenstrahler", *PROSPECT – Magazin der OETHG für Bühnen und Veranstaltungstechnik*, Dez. 2007, S. 32ff

Pfanzagl-Cardone, E. and Höldrich, R. (2008): "Frequency-dependent Signal-Correlation in Surround- and Stereo-Microphone Systems and the Blumlein-Pfanzagl-Triple (BPT)", Preprint 7476, 124th AES Convention, Amsterdam 2008

Pierce, J. R. (1983): "The Science of Musical Sound", Scientific American Books, 1983

Poletti, M. A. (2000): "A Unified Theory of Horizontal Holographic Sound Systems", *J. Audio Eng. Soc.*, vol. 48 Number 12 pp. 1155-1182; December 2000

Preston, C. (1998): "An Analysis of Frequency Dependent Hybrid Microphone Arrays for Stereophonic Sound Recording", Preprint # 4793, 105th AES-Convention, San Francisco 1998

Prokofieva, E. (2007): "Relation between correlation characteristics of sound field and width of listening zone", Paper 7089, 122nd Convention der Audio Eng. Society, Wien 2007

Rathbone, B., Fruhmann, M., Spikofski, G., Mackensen, P. und Theile, G. (2000): "Untersuchungen zur Optimierung des BRS-Verfahrens (Binaural Rooms Scanning)", 21. Tonmeistertagung Hannover, Tagungsband pp. 92-106, Nov. 2000

Reichardt, W., und Kussev, A. (1972), in: *Zeitschrift elektr. Inform. u. Energietechnik*, vol. 3, No.2, S. 66, Leipzig, 1972; ohne Titelangabe (siehe hierzu [Cremer und Müller, 1978], Fußnote 2, S. 345]

Reichardt, W., Abdel Alim, O. und Schmidt, W. (1975): „Definitionen und Messgrundlage eines objektiven Maßes zur Ermittlung der Grenze zwischen brauchbarer und unbrauchbarer Durchsichtigkeit bei Musikdarbietung“, *Acustica*, vol. 32, S.126 (1975)

Reichardt, W., Abdel Alim, O. und Schmidt, W. (1975): ”Zusammenhang zwischen Klarheitsmaß C und anderen objektiven raumakustischen Kriterien“, *Zeitschrift elektr. Inform. u. Energietechnik*, vol. 5, S.144, 1975

Rumsey, F. (1998): ”Subjective Assessment of the Spatial Attributes of Reproduced Sound”, Preprint, AES 15th Int. Conference, (1998)

Rumsey, F. (2000): “Surround sound: placing European Research in an international operational context“, 21. Tonmeistertagung des VDT, Hannover 2000, pg.44-62

Rumsey F., Segar P., (2001): "Optimisation and Subjective Assessment of Surround Sound Microphone Arrays", paper #5368, 110th AES Convention, Amsterdam, May 2001

Rumsey, F. (2001): “Spatial Audio“, Focal Press (Elsevier), (2001)

Sancken WMS-5 (2010): Weblink (Datum der Internet-Recherche 2.4.2010)

<http://www.sanken-mic.com/en/product/product.cfm/10.1001200>

Sarroff, A., and Bello, J. P. (2008): “Measurements of spaciousness for stereophonic music”, Preprint 7539, 125th Convention der Audio Eng. Soc., Oct. 2008

Sawaguchi, M., Fukada, A. (1999): “Multichannel Sound Mixing Practice for Broadcasting”, Paper zur IBC Conference, 1999

Schoeps Double-MS (2010): Weblink (Datum der Interner-Recherche: 2.4.2010)

<http://www.schoeps.de/de/products/categories/doublems>

Sengpiel, E. (1992): “Grundlagen der Hauptmikrofon-Aufnahmetechnik – Skripten zur Vorlesung“ (Musikübertragung), Hochschule der Künste, Berlin, 1992
siehe auch www.sengpielaudio.de, (Datum der Internetrecherche: 2004)

Sennheiser MKH 800 Twin (2010): Weblink (Datum der Internet-Recherche 2.4.2010)
http://www.sennheiser.de/sennheiser/home_de.nsf/root/professional_wired-microphones_studio-recording-mics_mkh-800-series_502431

Shirley, B., Kendrick, P. , Churchill, C. (2007): “The Effect of Stereo Crosstalk on Intelligibility: Comparison of a Phantom Stereo Image and Central Loudspeaker Source”, *J. Audio Eng. Soc.*, vol. 55 No. 10, Oct. 2007, pg. 825ff

Siebrasse, K.A., (1973): Dissertation Göttingen (1973) – ohne Titelangabe

Simonsen, G. (1984): Master’s Thesis, Technische Universität von Lyngby, Dänemark – ohne Titelangabe

Snow, W. B. (1953): “Basic Principles of Stereophonic Sound”, *SMPTE*, vol. 61, pp.567-589, Nov. 1953

Sokol, M. (2003): “FLRB-Array”, paper at the ‘DPA Microphones Workshop on Mic Techniques for Multichannel Audio’, AES 24th Int. Conference in Banff, Canada, 2003

Soulodre, G. A., Lavoie, M. C. and Norcross, S. G. (2003): “Objective measures of Listener Envelopment in Multichannel Surround Systems”, *J. Audio Eng. Soc.*, vol. 51, no.9, Sept. (2003)

Staff-Writer (2004): “New Horizons in Listening Test Design“ , *Journ. Audio Eng. Soc.*, vol.52, no. 1/1, Jan/Feb 2004

Steinberg, J.C. and Snow, W.B. (1934): “Auditory Perspective – Physical Factors”, *Electrical Engineering*, vol.53, no. 1, pp. 12-15, January 1934

Stern, R.M., Zeiberg, A. S., and Trahoitis, C. (1988): "Lateralization of Complex Binaural Stimuli: A Weighted Image Model", *J. Acoust. Soc. Am.*, vol. 84, pp. 156-165, 1988

Streicher, R. and Dooley, W. (1985): "Basic Microphone Perspectives – A Review", *J. Audio Eng. Soc.*, Vol. 33, No.7/8, July/August 1985

Suzuki, A., Tohyama, M. (1981): "Interaural cross-correlation coefficient of Kemar head and torso simulator", IECE Japan, Tech. Rep. EA80-78, (1981)

Theile, G. (1978): "Weshalb ist der Kammfilter-Effekt bei Summenlokalisierung nicht hörbar?", Paper zur 11. Tonmeistertagung des VDT, ISBN 3-598-20355-1, S. 200-214, (1978)

Theile, G. (1986): "Das Kugelflächenmikrofon", Paper zur 4. Tonmeistertagung VDT, S. 277, München, 1986

Theile G., et al. (1988): "Raumbezogene Stütztechnik – eine Möglichkeit zur Optimierung der Aufnahmequalität", Paper zur 15. Tonmeistertagung des VDT, 1988

Theile, G. (1996): "Möglichkeiten und Grenzen der 3/2-Stereo Aufnahme", Paper zur 19. Tonmeistertagung des VDT, S.145 (1996)

Theile, G. (2000): "Mikrofon- und Mischungskonzepte für 5.1 Mehrkanal-Musikaufnahmen" Bericht zur 21. Tagung des VDT, S. 348 ff, Hannover, 2000

Theile, G. (2001): "Multichannel Natural Music Recording Based on Psychoacoustic Principles", Paper, AES 19th Int. Conf., 2001

Thiele, R. (1953): "Richtungsverteilung und Zeitfolge der Schallrückwürfe in Sälen" *Acustica* vol. 3, S. 291-302 (1953)

Tohyama, M. and Suzuki, A. (1989): "Interaural cross-correlation coefficients in stereo-reproduced sound fields", *J. Acoust. Soc. Am.*, vol. 85, no.2, February 1989 (reprinted in: F. Rumsey (Editor), "An Anthology of Articles on 'Spatial Sound Techniques'

- Part 2: Multichannel Audio Techniques”, Audio Engineering Society, ISBN 0-937803-56-1, (2006))

Toole, F. E. (1985): ”Subjective Measurements of Loudspeaker Quality and Listener Performance”, *J. Audio Eng. Soc.*, vol. 33, No. 1/2, pg. 2-32 (1985)

Toole, F. E., (1986): “Loudspeaker Measurements and their Relationship to Listener Preferences”, *J. Audio Eng. Soc.*, vol. 34, pt.1, pp.227-235, pt.2, pp. 227-235, 1986

Toole, F. E. (1991): “Binaural Record/Reproduction Systems and Their Use in Psychoacoustic Investigations”, preprint # 3179, 91st Convention of the Audio Eng. Soc., 1991

Toole, F. E. (2008): “Sound Reproduction – Loudspeakers and Rooms”, Focal Press (Elsevier), 2008

Tryon, R.C. (1939): “ Cluster Analysis”, Ann Arbor: Edwards Brothers (1939)

Valin, J. (1994): “The RCA Bible – A Compendium of Opinion on RCA Living Stereo Records”, The Music Lovers Press (Cincinnati, Ohio), Second Edition, 1994

Van der Heijden, M., and Trahiotis, C. (1998): “Binaural Detection as a Function of Interaural Correlation and Bandwidth of Masking Noise: Implications for Estimates of Spectral Resolution“, *J. Acoust. Soc. Am.*, vol. 103, pp. 1609-1614, 1998

von Hornbostel, E.M. und Wertheimer, M. (1920): “Über die Wahrnehmung der Schallrichtung”, Sitzungsbericht der Akad. Wiss. Berlin, S.388-396, (1920)

Watson, F. R. (1923): “Acoustics of Auditoriums“, New York, 1923

Wightman, F. L. and Kistler, D. J. (1992): “The Dominant Role of Low-Frequency Interaural Time Differences in Sound Localization“, *J. Acoust. Soc. Am.*, vol. 91, pp. 1648-1661, 1992.

Williams, M. (1987): "Unified Theory of Microphone Systems for Stereophonic Sound Recording", Preprint 2466 (H-6), 82nd AES-Convention, 1987

Williams, M. (1992): "Frequency Dependent Hybrid Microphone Arrays for Stereophonic Sound Recording", Preprint # 3252, 92nd AES-Convention, 1992

Wilkens, H. (1975a): "Mehrdimensionale Beschreibung subjektiver Beurteilung der Akustik von Konzertsälen", Dissertation TU Berlin (1975)

Wilkens, H. (1975b): "Subjektive Beurteilung der Akustik von Konzertsälen", Paper zur 10. Tonmeistertagung des VDT, S. 166-171 (1975)

Willson, V. L. (1980): "Research techniques in AERJ articles: 1969-1975", *Educ. Researcher*, vol. 9, 5-10 (1980)

Wittek, H. (2002): "Microphone Assistant V2.0" Java-Applet auf Website www.hauptmikrofon.de; (Stand der Recherche: Juni 2008)

Wittek, H. and Theile, G. (2000): "Investigations into directional imaging using L-C-R stereo microphones" Paper zur 21. Tonmeistertagung (German), ISBN 3-598-20362-4, pg.432-454, 2000

Wittek H., and Theile G., (2002): "The Recording Angle – based on Localisation Curves", Paper 5568, 112th AES-Convention, München, 2002

Wittek, H., und Theile, G. (2004): siehe OCT V2-Info auf Website www.hauptmikrofon.de (Stand der Recherche: Juni 2008)

Woszczyk W. R. (1992): "Microphone Arrays Optimized for Music Recording", *J. Audio Eng. Soc.*, vol. 40, no. 11, Nov. 1992

Wuttke, J. (1981): "Mikrofondaten und ihre Bedeutung in der Praxis", Beitrag zur 12. Tonmeistertagung des VDT (Tagungsband), München 1981

Wuttke, J. (1998): "Das Mikrofon zwischen Physik und Emotion", Paper zur 20. Tonmeistertagung des VDT, Karlsruhe 1998, (ISBN 3-598-20361-6), S.460

Yamamoto, T. (1973): "Quadraphonic One-Point Pickup Microphone", JAES, No. 21, vol. 4, (1973)

Yanagawa, H., Higashi, H., Mori, S. (1976): "Interaural correlation coefficients of the dummy head and the feeling of widness", A.S. J. Tech. Rep. H-35-1, (1976)

Yoshida, T., Iwasaki, S. and Nagai, K. (1960): "What is the Basic Factor of Stereophonic Sound ?", *J. Acoust. Soc. Jap.*, Vol.16, pp.249-257 (1960)

Yost, W.A., Wightman, F.L., Green, D.M. (1971): "Lateralisation of filtered clicks", *Journal of the Acoust. Soc. of America*, 50, pp. 1526-1531, (1971)

Zacharov, N., (1998): "Subjective Appraisal of Loudspeaker Directivity for Multichannel Reproduction", *J. Audio Eng. Soc.*, vol.46, No.4, pp.288-303, April 1998

Zollner, M., und Zwicker, E. (2003): "Elektroakustik", 3. Auflage, Springer Verlag, 2003

Diskographie:

Chesky (1994): CD "Best of Chesky Classics & Jazz and Audiophile Test Disc" Volume 3, Tracks 15, 18; Chesky Records Nr. JD 111

Hyperion (1984): CD "Sacred Vocal Music of Claudio Monteverdi", Track 2 ("Confitebor tibi Domine"); Hyperion Records Ltd, London, Nr. CDA66021, 1984

Jecklin, J. (1999): "The Low Art of Recording", Referenz-CD mit Musikbeispielen und Testsignalen, zusammengestellt von Jürg Jecklin; Track 57 („Canadian Brass Ensemble“)

Pfanzagl-Cardone, E. (2003): "Romantic-Minimal", Komposition für Violoncello und Klavier; Manuskript

RCA (1957): CD Antonin Dvorak, Symphony No. 9 in E minor, op.95, 1. Satz, RCA-Victor (BMG Classics) Nr. 09026 62587 2, 1957 (CD-Release: 1995)

Teldec (1995): CD Gustav Mahler, Symphony No. 9 in D major, 3. Satz, TELDEC Nr. 4509-90882-2, 1995

Videoquellennachweis:

Swedien, B. (1997): "Recording with Bruce Swedien", VHS-Video (Producer: Phil Cochlin), Palefish Enterprises Inc., (P.O. Box 417, White Salmon, WA), AcuNet Corporation, 1997

ANHANG

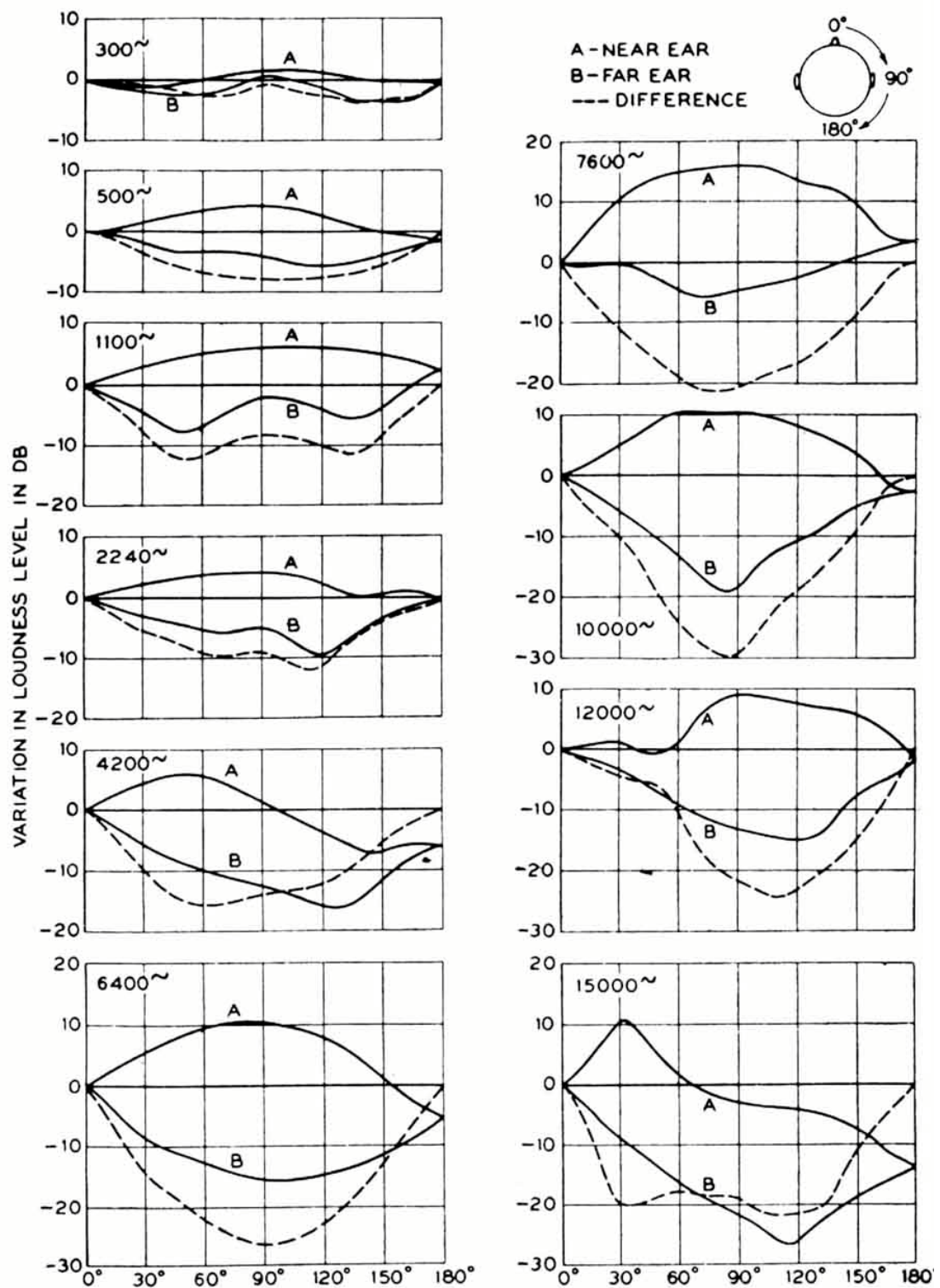


Abb. A1: Änderung der Lautstärke bei horizontaler Rotation einer Schallquelle um den Kopf (aus [Steinberg and Snow, 1934])

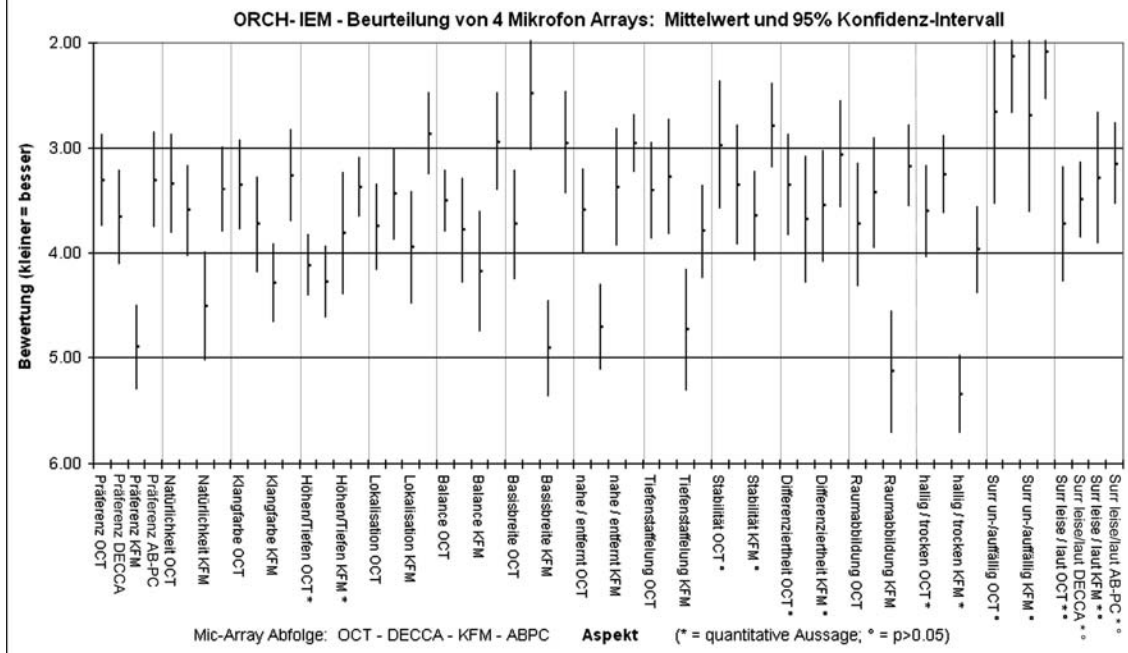


Abb. A2: Hörergruppe IEM - Mittelwert (über 15 Hörer) und 95%-Konfidenzintervall betr. 15 Aspekte von 4 Mikrofon-Arrays

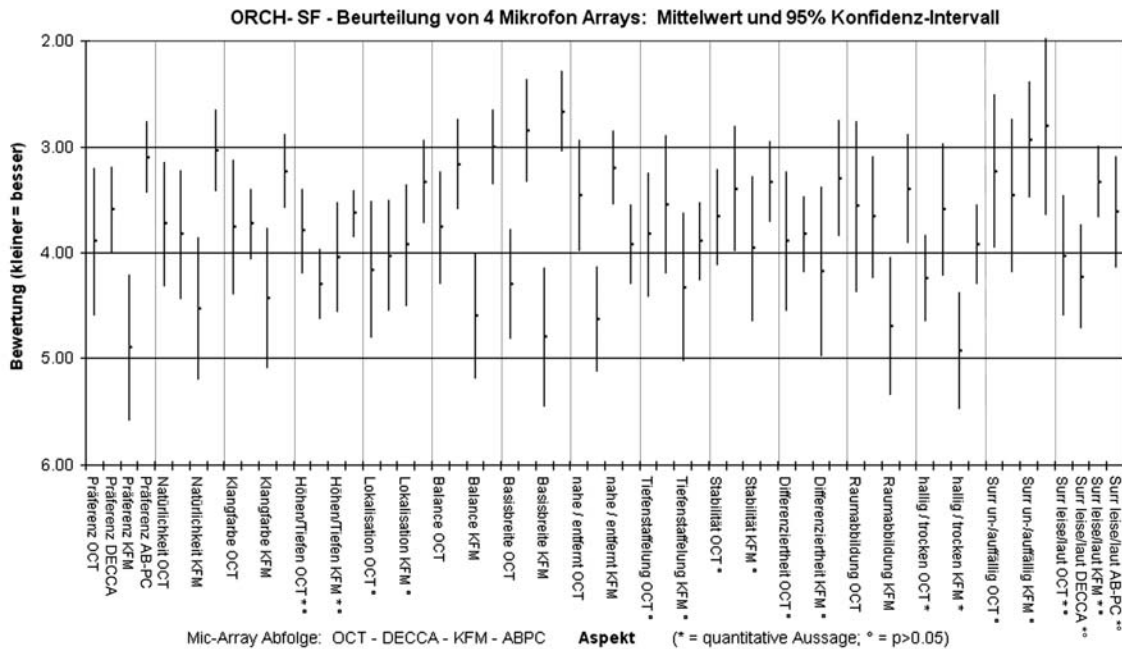


Abb. A3: Hörergruppe SF - Mittelwert (über 10 Hörer) und 95%-Konfidenzintervall betr. 15 Aspekte von 4 Mikrofon-Arrays

| | SF (10 Hörer) | IEM (15 Hörer) |
|---------------------|-------------------------|--------------------------|
| Aspekt: | | |
| Präferenz | 0.0007 | 0 |
| Natürlichkeit | 0.0097 | 0.0027 |
| Klangfarbe | 0.0268 | 0.0060 |
| Höhen/Tiefen | 0.1020 | 0.0139 |
| Lokalisation | 0.1301 | 0.0082 |
| Balance | 0.0002 | 0.0056 |
| Basisbreite | 0 | 0 |
| Entfernungseindruck | 0.0004 | 0 |
| Tiefenstaffelung | 0.3464 | 0.0010 |
| Stabilität | 0.3202 | 0.1003 |
| Differenziertheit | 0.2615 | 0.4110 |
| Raumabbildung | 0.0354 | 0 |
| hallig/trocken | 0.0057 | 0 |
| Surr. un-/auffällig | 0.5822 | 0.5118 |
| Surr. leise/laut | 0.0642 | 0.4000 |

Tab. AT1: Vergleich der p-Werte bezüglich Einfluss der Variablen ‚Mic-Array‘ auf die Beurteilung der Klang-Aspekte für die beiden Hörergruppen SF und IEM nach 1-faktorieller Varianzanalyse des Hörvergleichs ORCH 5.1

(Anm.: Werte kleiner als 0.0001 sind in der Tabelle mit 0 angegeben)

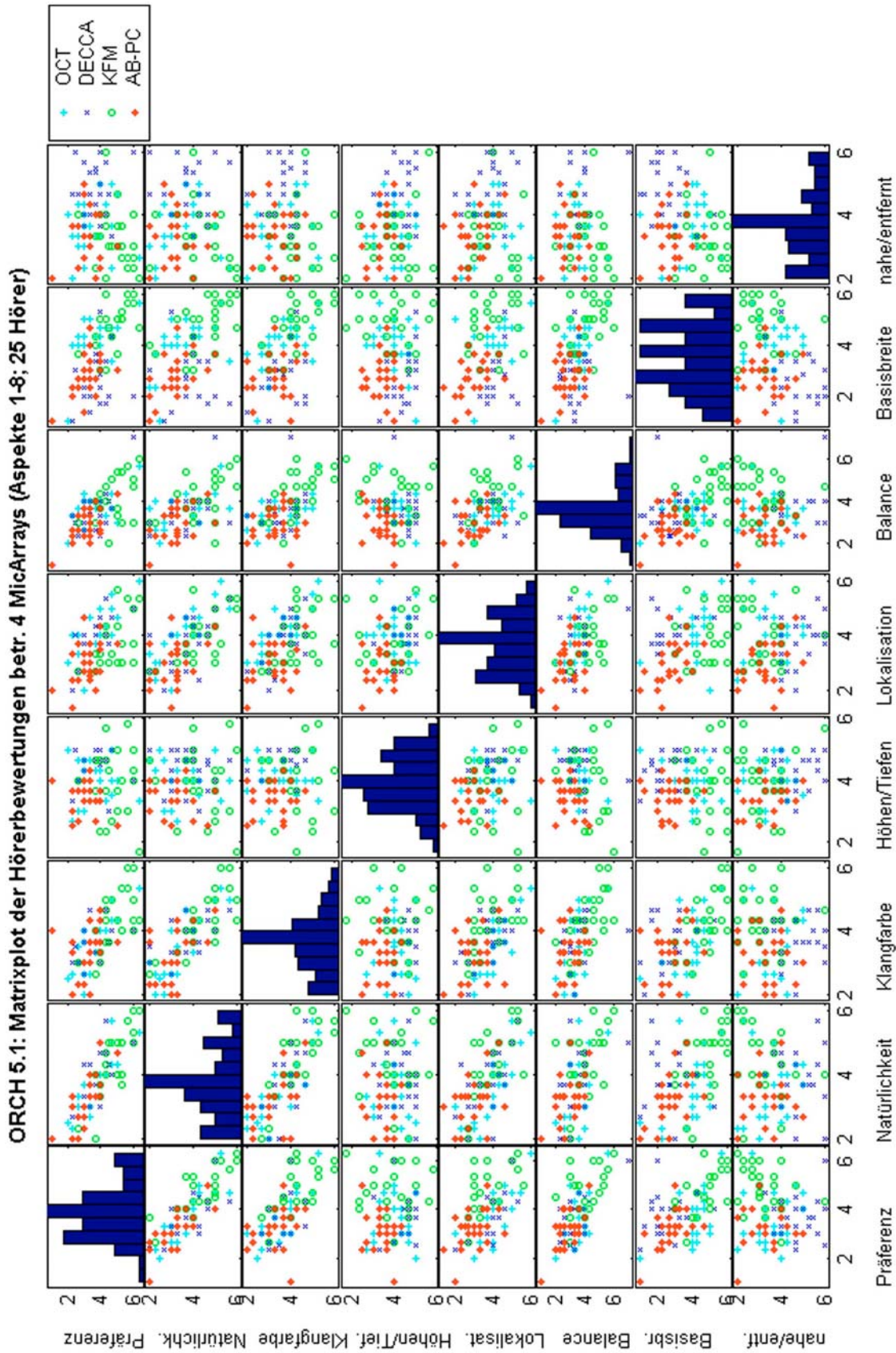


Abb. A4: Matrixplot der Hörerbewertungen betreffend 4 Mikrofon-Arrays (Aspekte 1-8, 25 Personen), [Stimulus: Orchester 5.1]

ORCH 5.1: Matrixplot der Hörerbewertungen betr. 4 MicArrays (Aspekte 9-15 zzgl. Präferenz; 25 Hörer)

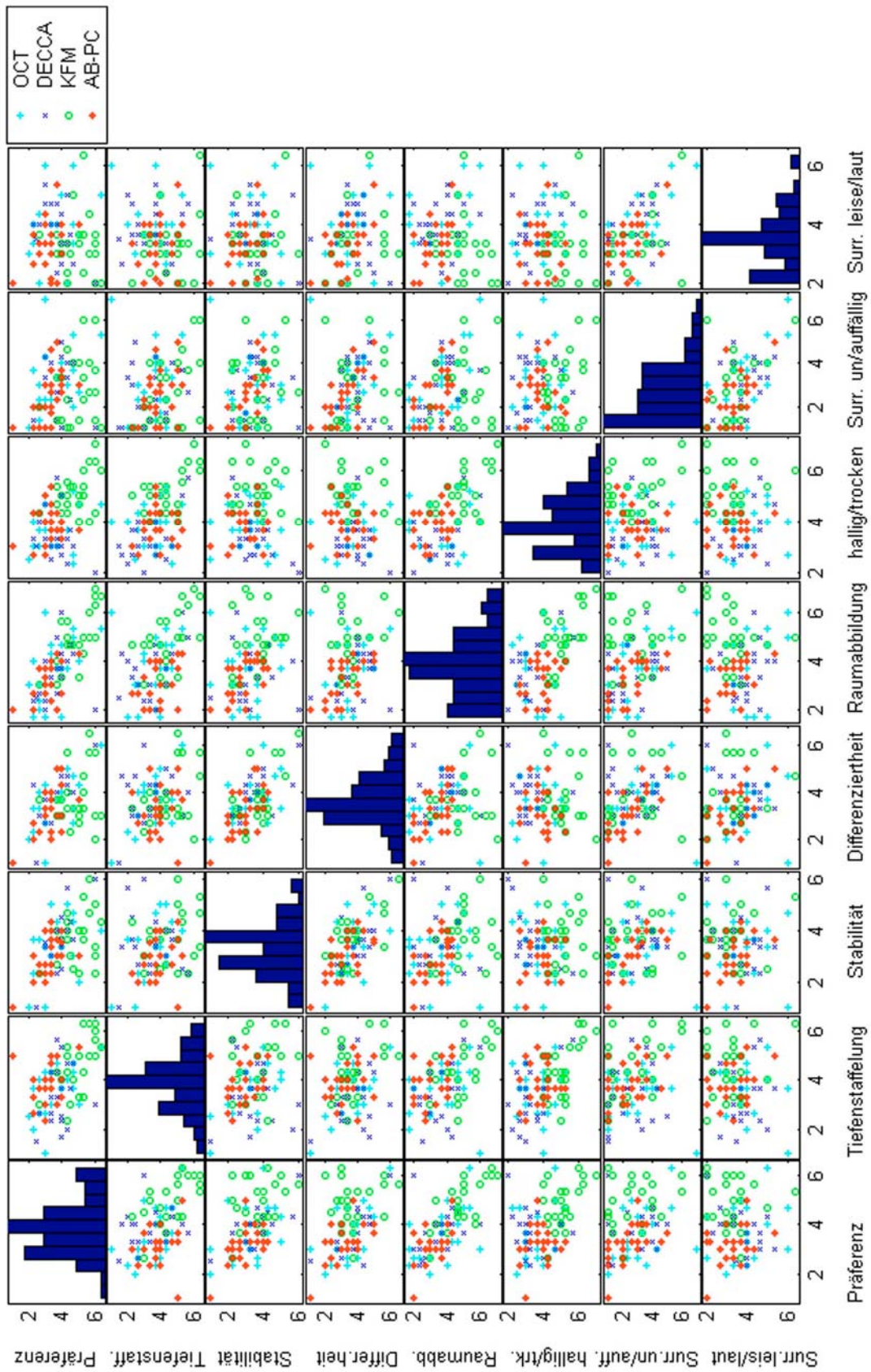


Abb. A5: Matrixplot der Hörerbewertungen betreffend 4 Mikrofon-Arrays (Aspekte 9-15, zzgl. Präferenz; 25 Personen), [Stimulus: Orchester 5.1]

Ergebnisse des Surround-Hörvergleichs des ORF (2001):

(Anm.: FH D = Hörvergleich an der FH Düsseldorf) (aus [Camerer und Sodl, 2001])

Question 1: Expansion of the orchestra image: wide / narrow

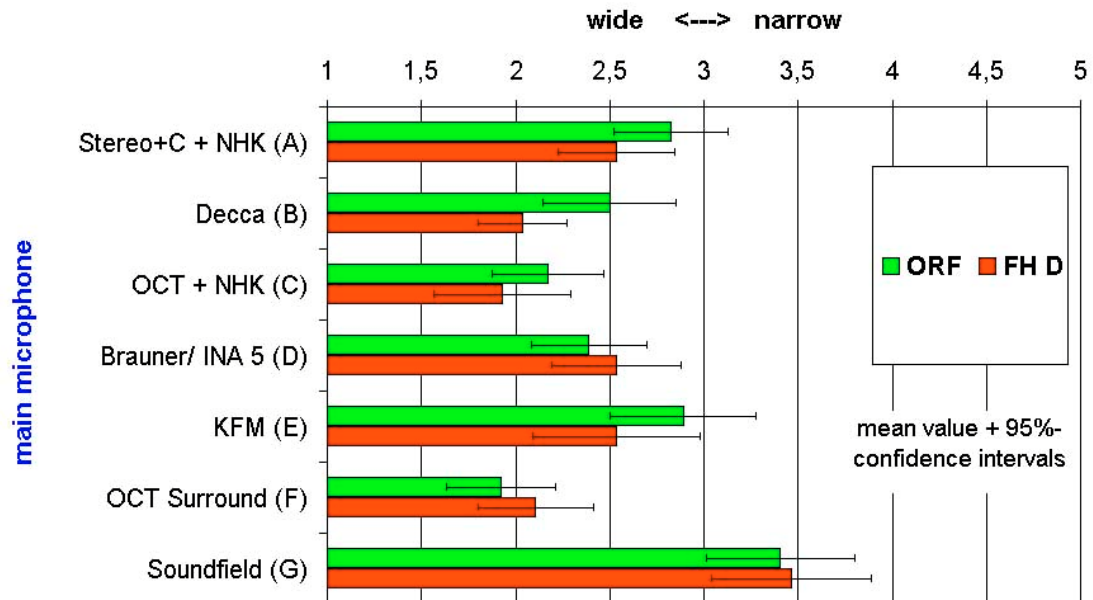


Abb. A6: Bewertung ‚Klangquellenausdehnung‘

Question 2: Distance of the orchestra image: close / distant

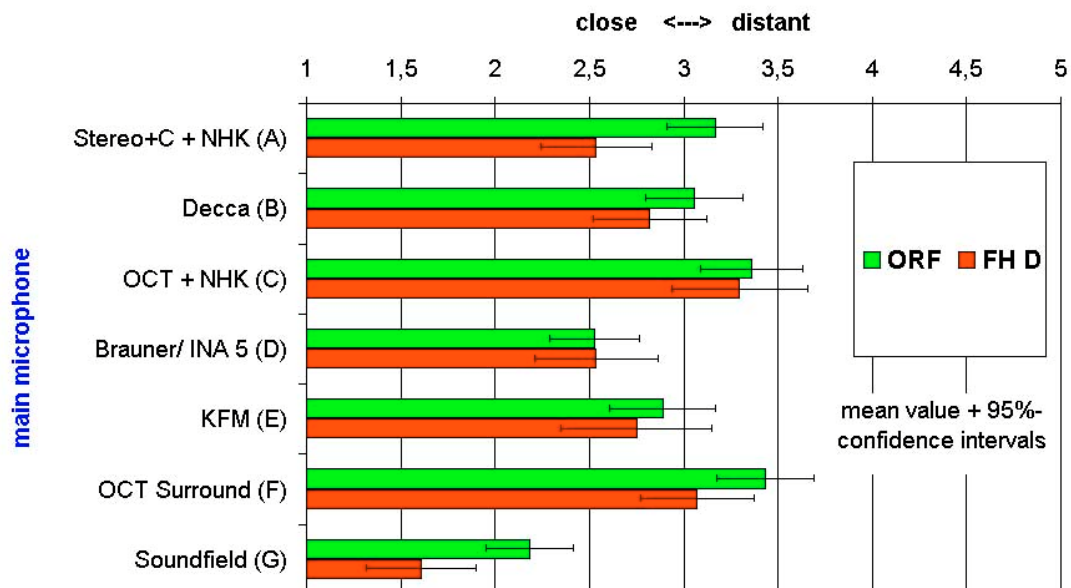


Abb. A7: Bewertung ‚Entfernungseindruck‘

Question 3: Depth of the orchestra image: deep / flat

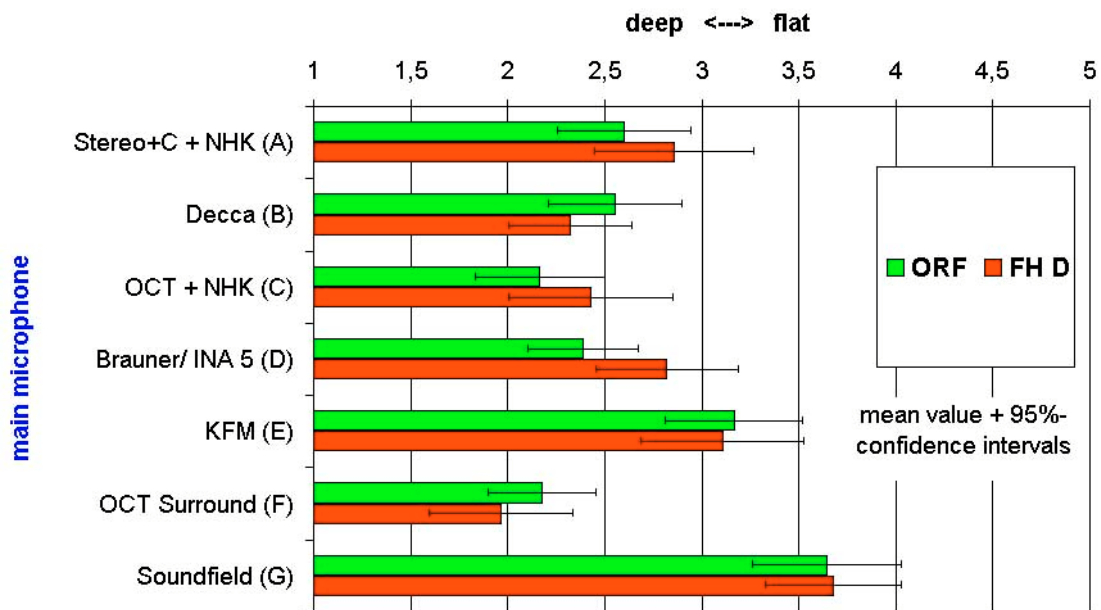


Abb. A8: Bewertung ‚Tiefenstaffelung‘

Question 4: Stability of the image: stable / instable

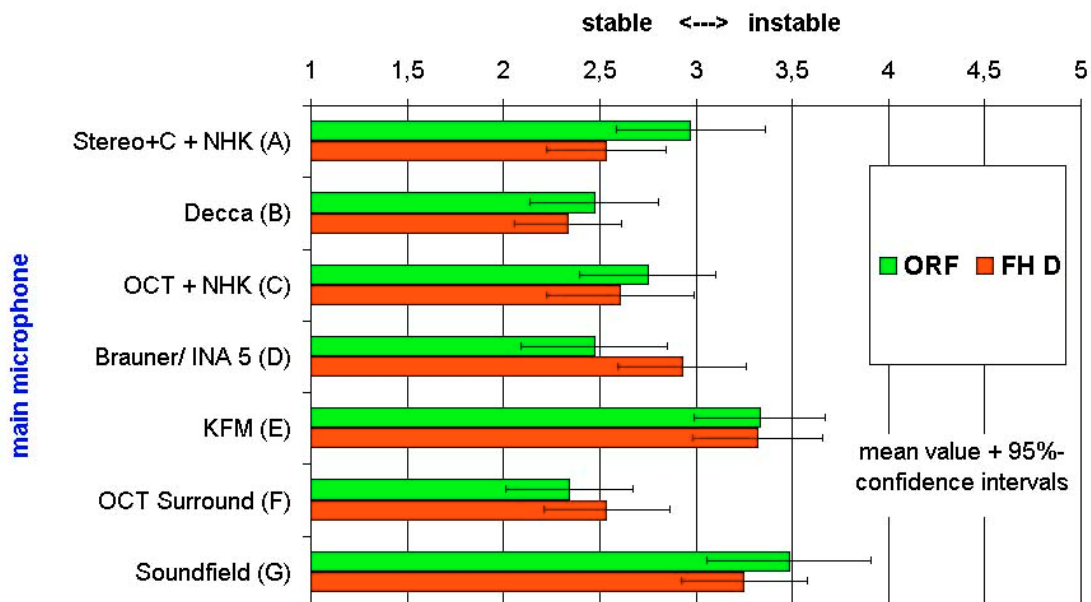


Abb. A9: Bewertung ‚Stabilität‘

Question 5: Precision of the image: precise / blurred

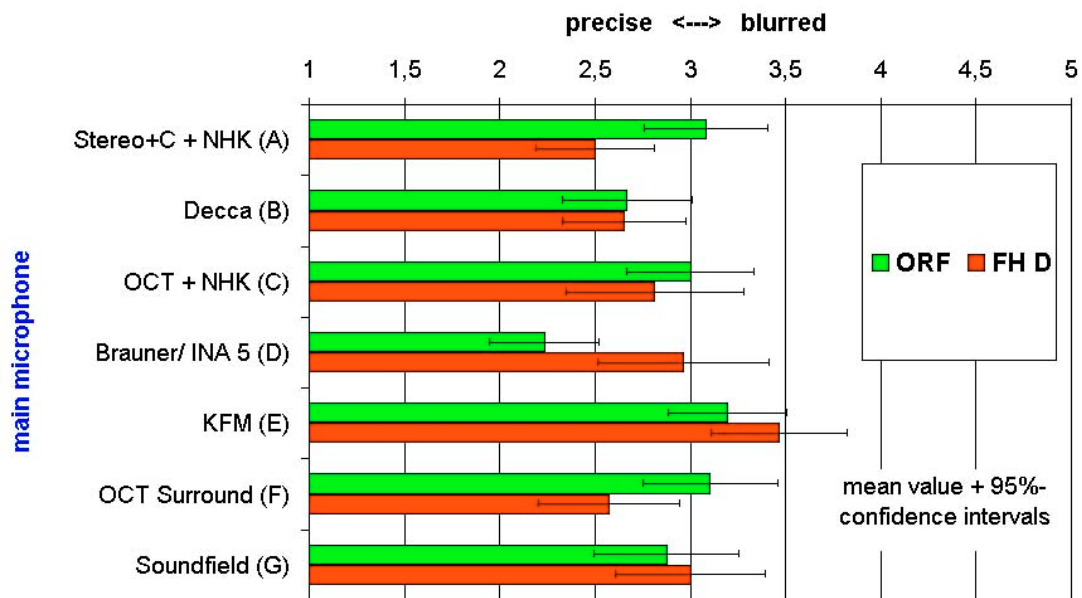


Abb. A10: Bewertung ‚Prazision‘ (Abbildungsscharfe des Klangquelleneindrucks)

Question 6: Sound colour: satisfactory / unsatisfactory

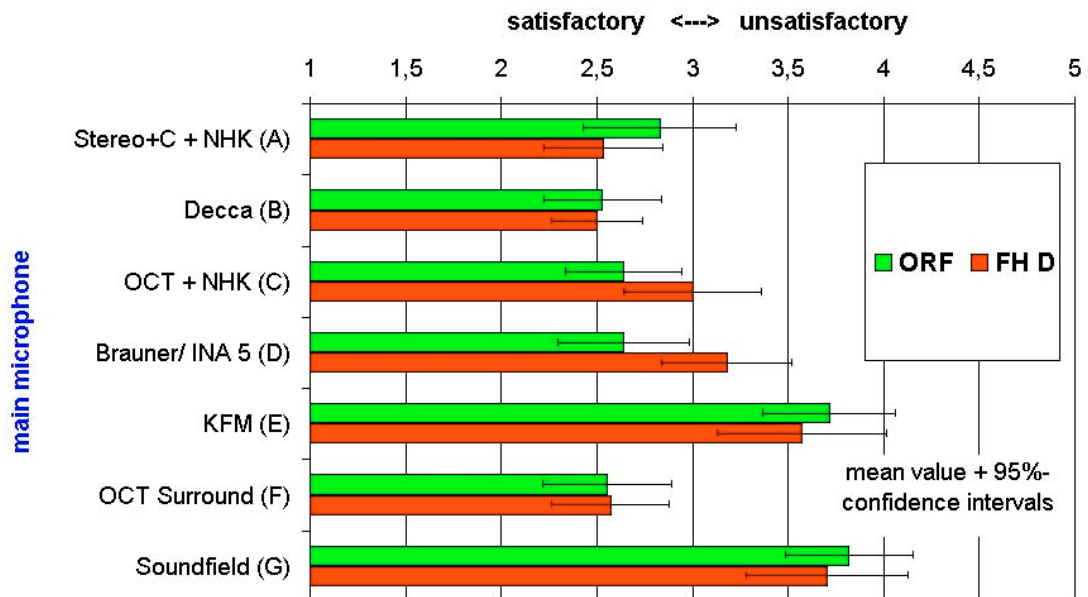


Abb. A11: Bewertung ‚Klangfarbe‘

Question 7: Room impression: perfect / imperfect

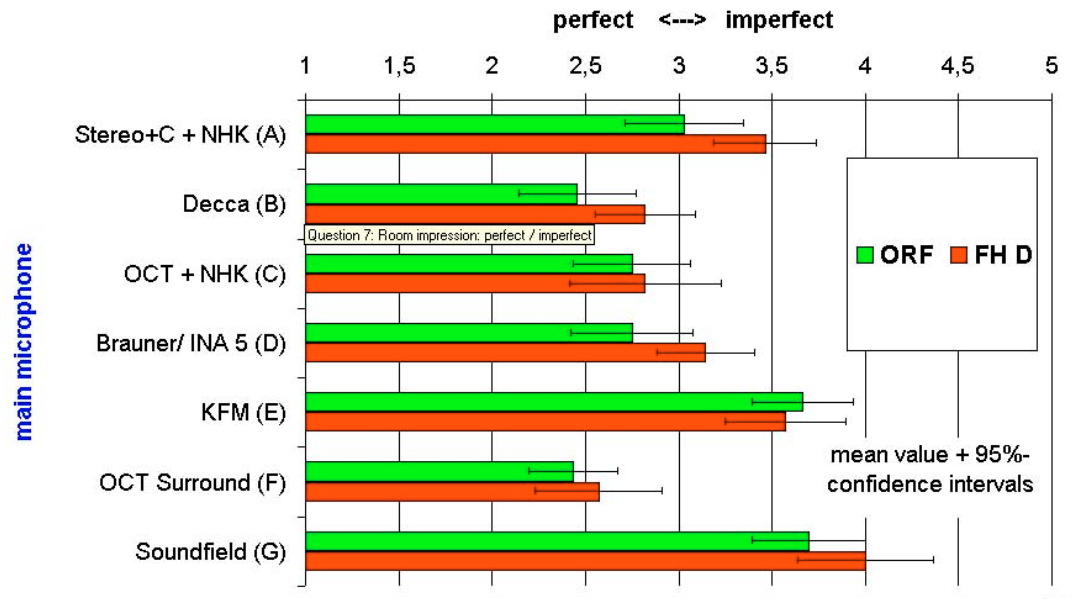


Abb. A12: Bewertung ‚Raumeindruck‘

Question 8: Presence of room information: too much / too little

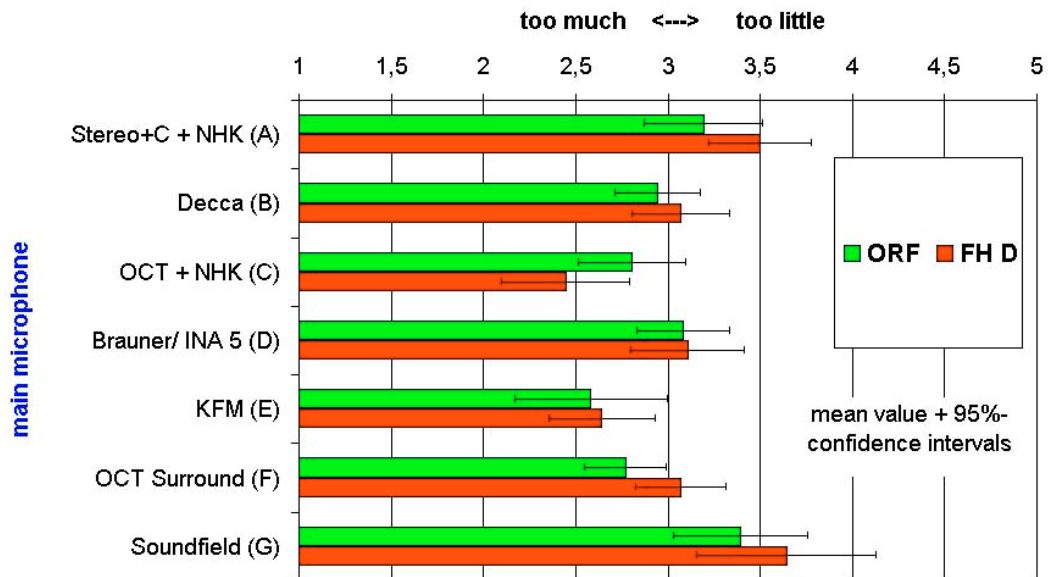


Abb. A13: Bewertung ‚Raumanteil‘

Question 9: Surround signals: identifiable / not identifiable

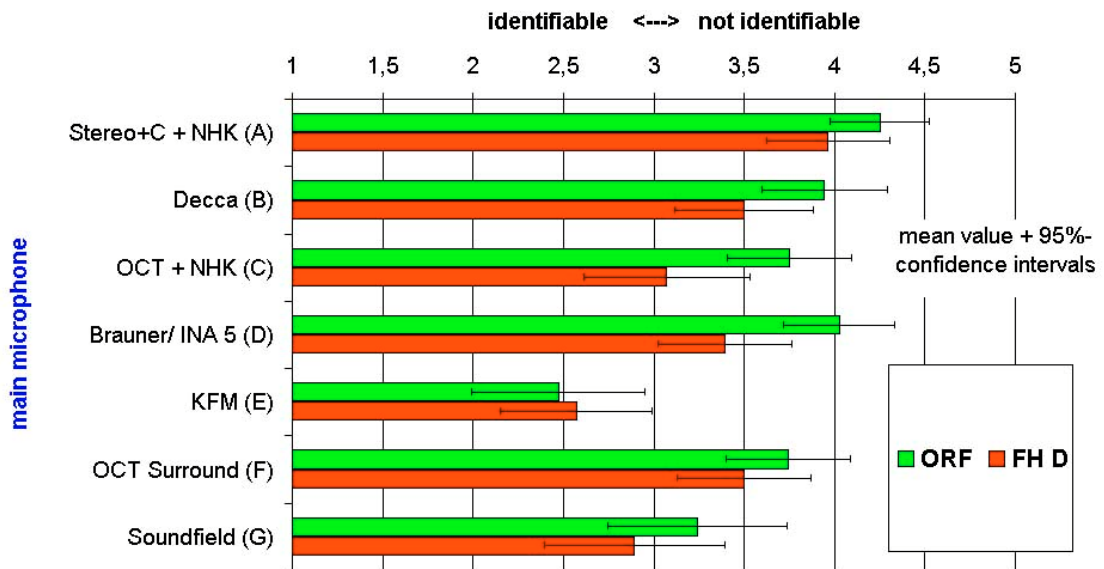


Abb. A14: Bewertung ‚Surround-Signale‘

Dokumentation zur raumakustischen Messung des IEM Produktionsstudios

Ergebnisse der Messungen im IEM Produktionsstudio
Stand: 15.11.2004 – Autor: Dr. Alois Sontacchi

Anschließend an den Hörversuch wurden eine Reihe von Messungen durchgeführt. Dabei wurden folgende Größen ermittelt:

- Frequenzgang
- Impulsantwort
- Integrierte Intensität
- Nachhallzeit

Die Frequenzgänge wurden räumlich als auch spektral gemittelt (vgl. dazu [Goertz, Wolff, 2002]). Für die räumliche Mittelung wurden die Raumimpulsantworten an 9 Positionen herangezogen. Die 9 Messpunkte spannen einen Raster um den idealen Abhörpunkt (blauer Kreis) auf siehe schematische Skizze in Abbildung 1. Die Kantenlängen zwischen 2 Positionen im flächigen Messgitter wurden mit 30cm festgelegt. An jedem Messpunkt wurde die Raumimpulsantwort durch 3 Messungen ermittelt. Als Messsignal wurde ein logarithmischer Sweep verwendet (vgl. dazu [Farina, 2000]).

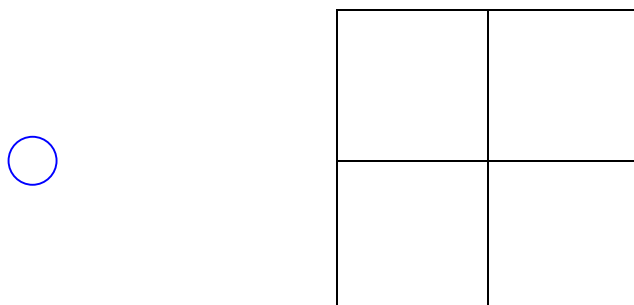


Abbildung A15: Schematische Skizze der gerasterten Messpunkte.

Bei der spektralen Mittelung wurde ein Terzfilterung der Frequenzgänge angewendet.

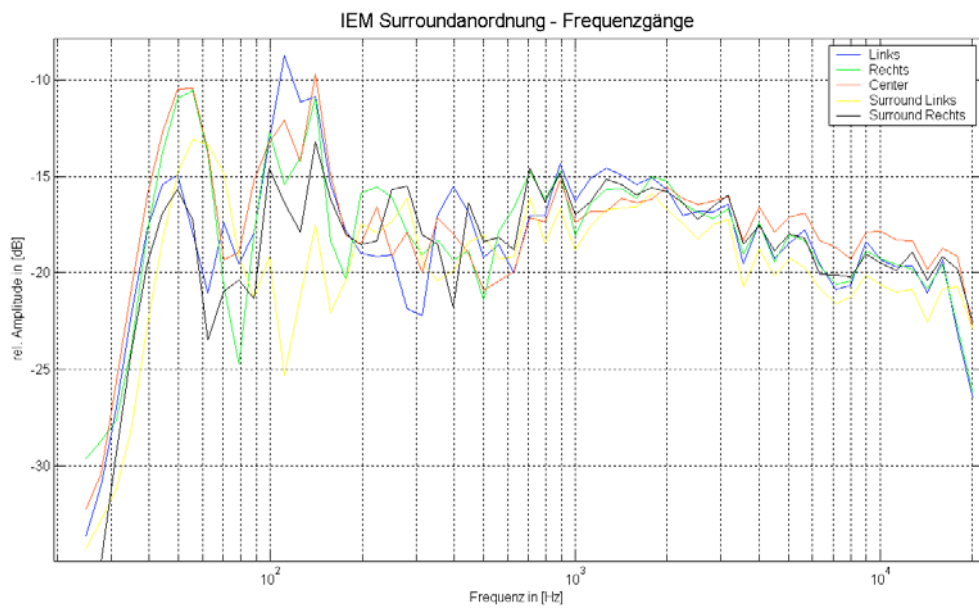


Abbildung A16: Frequenzgänge örtlich und spektral gemittelt.

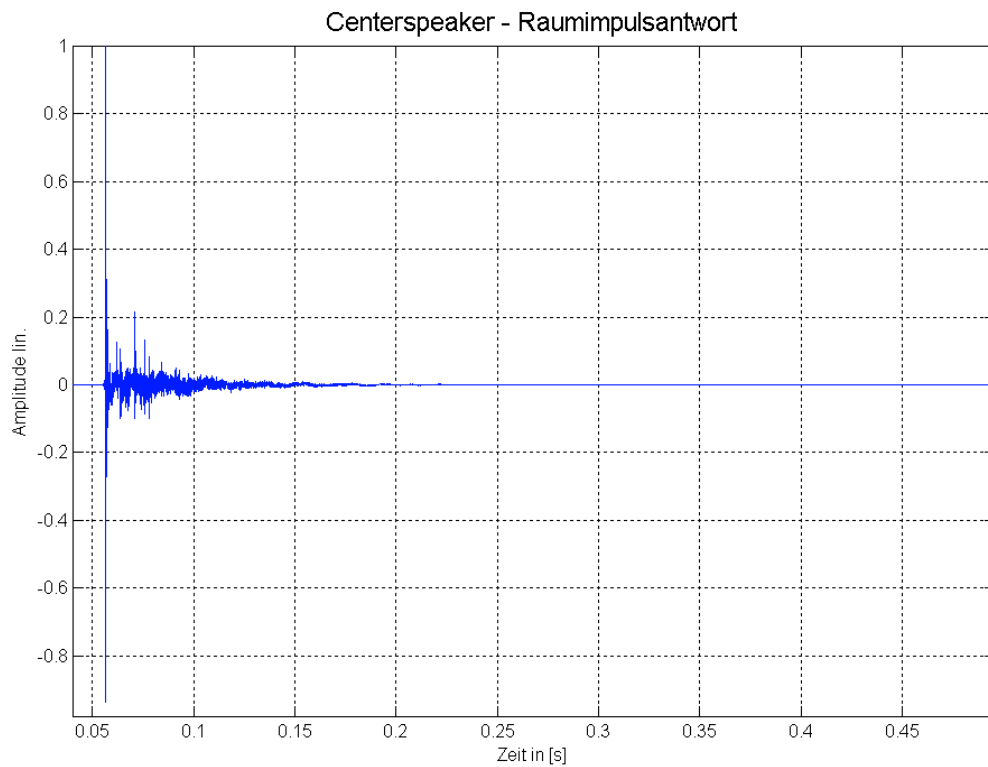


Abbildung A17: Normierte lineare Raumimpulsantwort des Centerlautsprechers am idealen Abhörpunkt. Darstellung auf 500ms beschnitten.

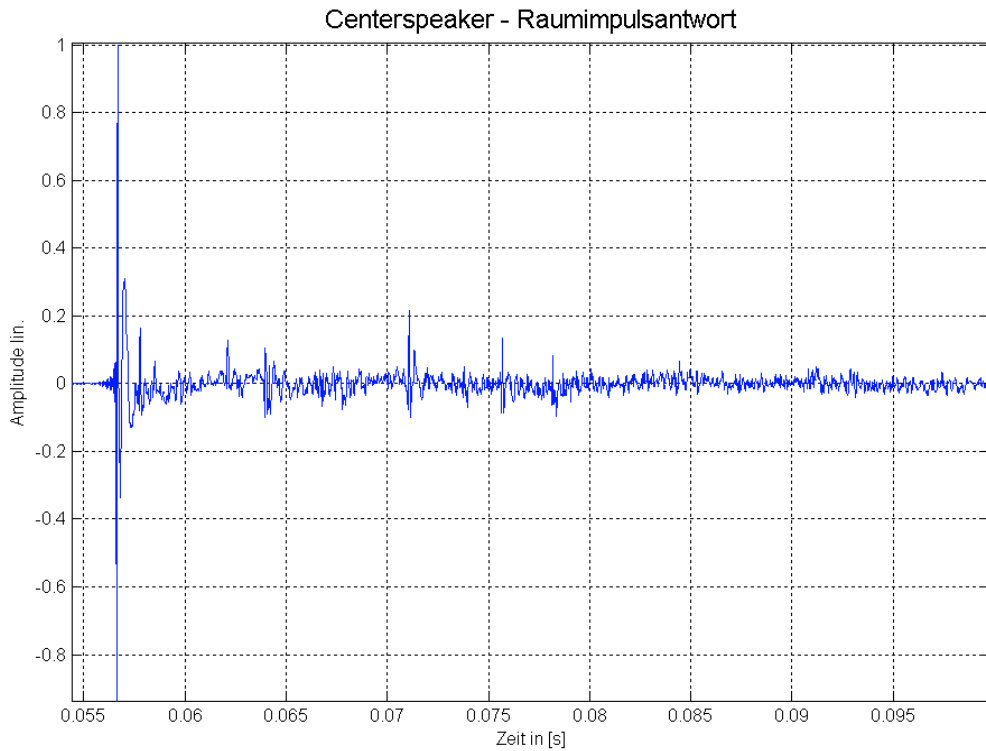


Abbildung A18: Normierte lineare Raumimpulsantwort des Centerlautsprechers am idealen Abhörpunkt. Darstellung auf 100ms beschnitten.

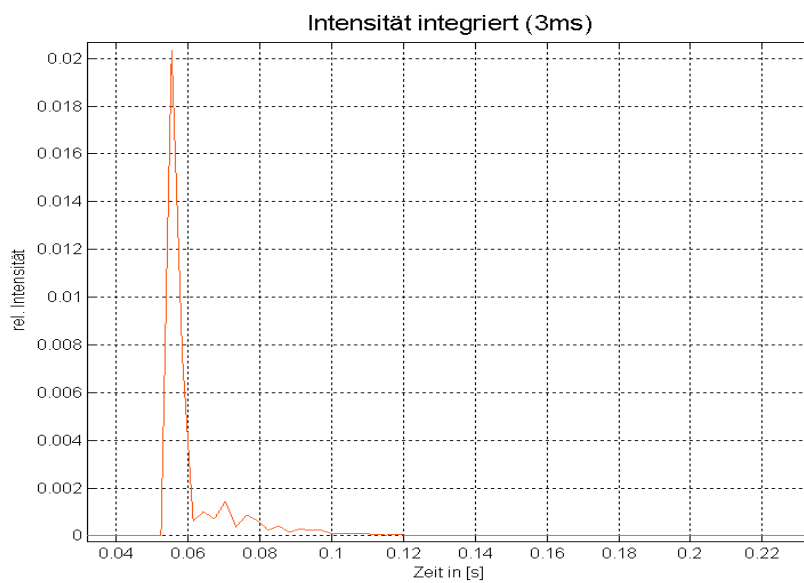


Abbildung A19: Integrierte Intensität mit Zeitkonstante 3ms.

Bestimmt aus der Raumimpulsantwort des Centerlautsprechers am idealen Abhörpunkt.

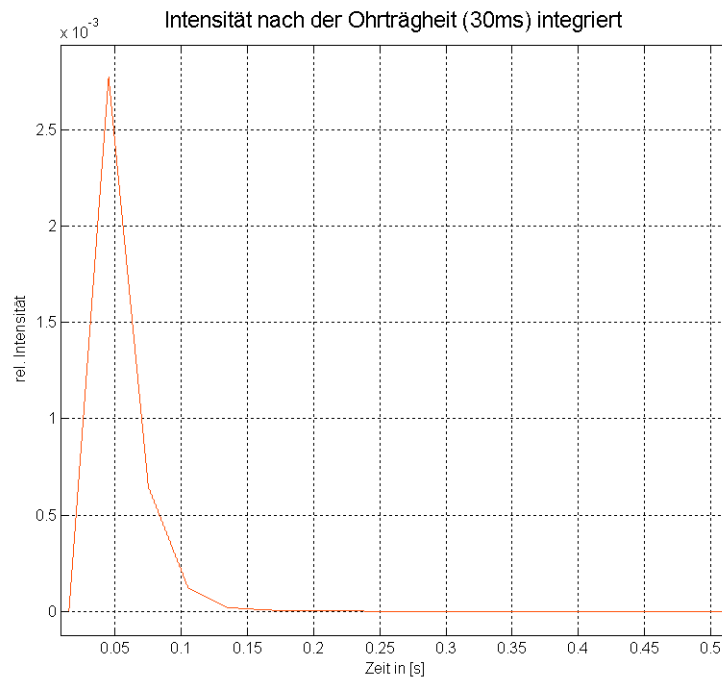


Abbildung A20: Integrierte Intensität mit Zeitkonstante 30ms (entspricht der Ohrträgheit). Bestimmt aus der Raumimpulsantwort des Centerlautsprechers am idealen Abhörpunkt.

Aus der integrierten Intensität, die sich aus der integrierten quadrierten Impulsantwort berechnet, kann man den Klangeinsatz als auch die Verschmelzung von Reflexionen zu einem Klangereignis erkennen. Bei einer Integrationskonstante von 30ms kann man überhaupt nur ein Ereignis ausmachen (vgl. dazu [Ahnert, Steffen, 1993]).

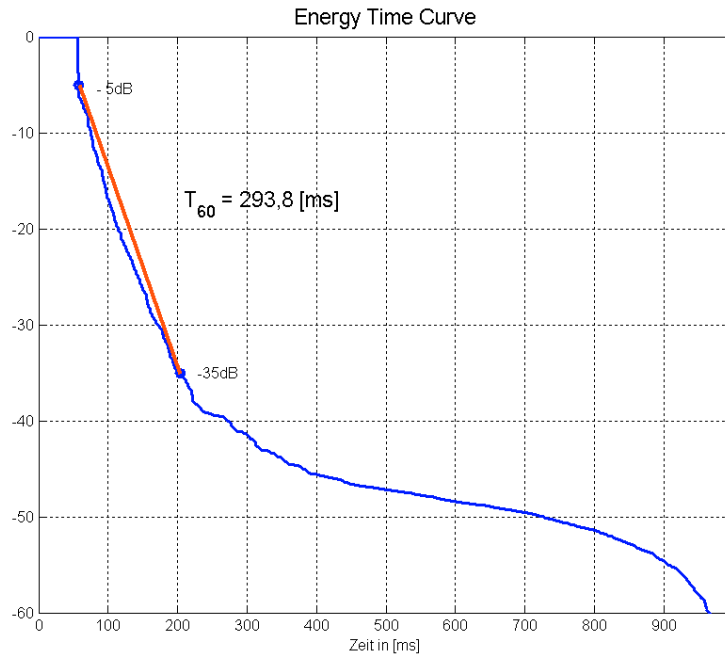


Abbildung A21: Energy Time Curve (ETC - Energiezeitkurve) bestimmt aus der Impulsantwort des Centerlautsprechers am idealen Abhörpunkt.

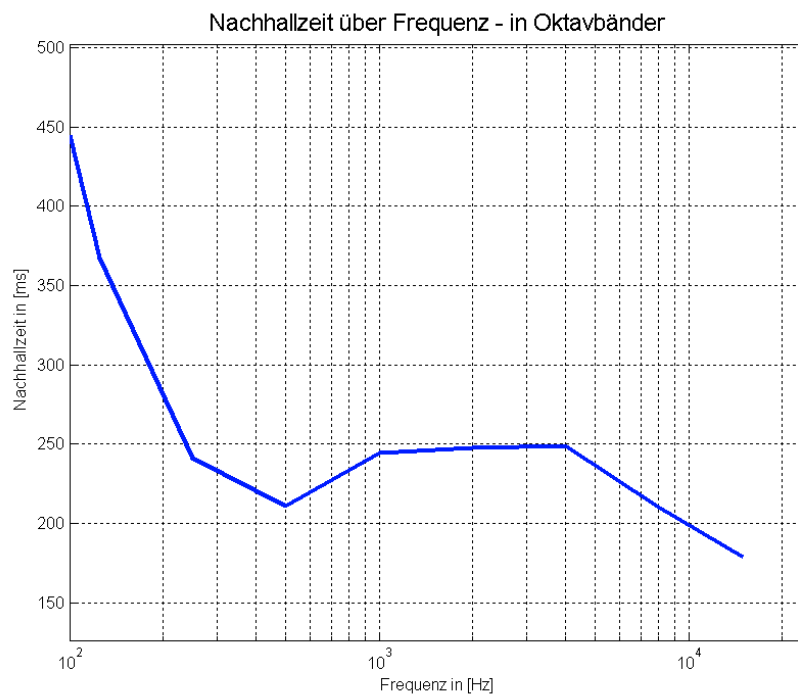


Abbildung A22: Nachhallzeiten für die Oktavbänder (125Hz bis 16kHz) bestimmt durch ETC der Oktavband gefilterten Impulsantwort des Centerlautsprechers am idealen Abhörpunkt.

Die Energy Time Curve (ETC) (siehe Abb. A22, oben) wird durch Rückwärtsintegration der gemessenen quadrierten Raumimpulsantwort berechnet (Rückwärtsintegration nach Schröder). Sie liefert ein Maß für den zeitlichen Energieabfall in einem angeregten Raumschallfeld.

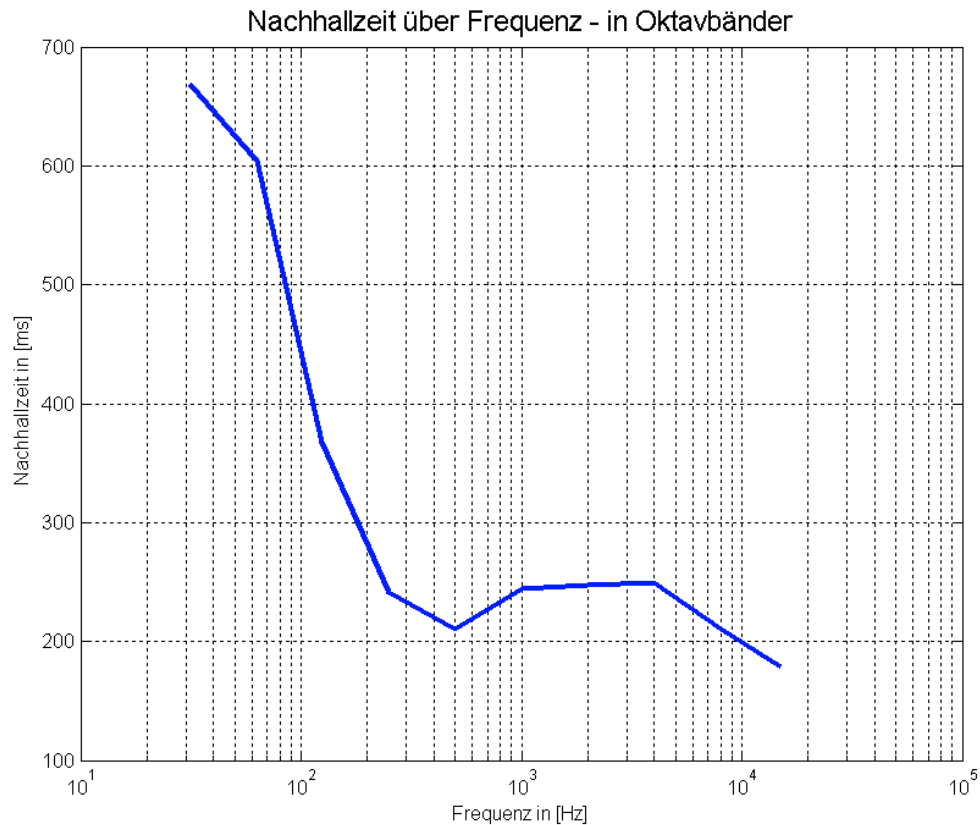


Abbildung A23: Nachhallzeiten für die einzelnen Oktavbänder (31Hz bis 16kHz) bestimmt durch die ETC der oktavbandgefilterten Impulsantwort des Centerlautsprechers am idealen Abhörpunkt.

Literaturquellen

Goertz, Anselm; Wolff, Markus (2002): "Optimierung der Wiedergabe von Surround Lautsprecheranordnungen in Tonstudios und Abhörräumen", 22. Tonmeistertagung, Hannover, November, 2002.

Farina A.(2000): “Simultaneous measurements of impulse response and distortion with a swept sine technique”, J. Audio Eng. Soc. (Abstracts), vol. 48, p. 350, April 2000.

Ahnert, Wolfgang; Steffen, Frank (1993): “Beschallungstechnik”, Grundlagen und Praxis, S. Hirzel Verlag Stuttgart, Leipzig 1993 (ISBN 3-7776-0496-8 Gb).

| Oktavband: [Hz] | RT60 [s] | Hallradius r [m] |
|--------------------|-------------|---------------------|
| 16000 | 0.17 | 1.50 |
| 10000 | 0.20 | 1.38 |
| 8000 | 0.22 | 1.32 |
| 4000 | 0.26 | 1.21 |
| 2000 | 0.25 | 1.24 |
| 1000 | 0.24 | 1.26 |
| 500 | 0.22 | 1.32 |
| 250 | 0.24 | 1.26 |
| 125 | 0.37 | 1.02 |
| 63 | 0.61 | 0.79 |

| |
|---|
| Raumvolumen V [m ³] 117.6 |
|---|

$$r = 0,057 * \sqrt{V / RT60}$$

| |
|--|
| Location: IEM-Studio |
| L x B x H in [m]: 4,8 (7,2) x 5,6 x 3,5 |

RT60-Werte wurden aus ETC-Messung von Dr.Sontacchi übernommen

Tab. AT2: Berechnung des frequenzabhängigen Hallradius' ('critical distance') des IEM-Produktionsstudios (Regieraum) auf Basis von Nachhallzeit RT60 und Raumvolumen

| Oktavband: [Hz] | RT60 [s] | Hallradius r [m] |
|--------------------|-------------|---------------------|
| 8000 | 0.28 | 0.81 |
| 4000 | 0.32 | 0.76 |
| 2000 | 0.33 | 0.75 |
| 1000 | 0.36 | 0.72 |
| 500 | 0.36 | 0.72 |
| 250 | 0.44 | 0.65 |
| 125 | 0.77 | 0.49 |
| 63 | 1.28 | 0.38 |

| |
|--|
| Raumvolumen V [m ³] 57 |
|--|

$$r = 0,057 * \sqrt{V / RT60}$$

| |
|------------------------------------|
| Location: SF-Studio |
| L x B x H in [m]: 6 x 3,8 x 2,5 |

Für RT60 angegebene Werte jeweils über 10 Messdurchgänge gemittelt; gemessen mit NTI Acoustilizer (Klasse 2-Gerät); 13.2.2009, E. Pfanzagl

Tab. AT3: Berechnung des frequenzabhängigen Hallradius' ('critical distance') des SF-Produktionsstudios (Regieraum) auf Basis von Nachhallzeit RT60 und Raumvolumen



Abb. A24: ein Proband beim Hörtest im Regieraum des IEM-Produktionsstudios



Abb. A25: Der Zuspielplatz (Pro-Tools) im IEM-Regieraum

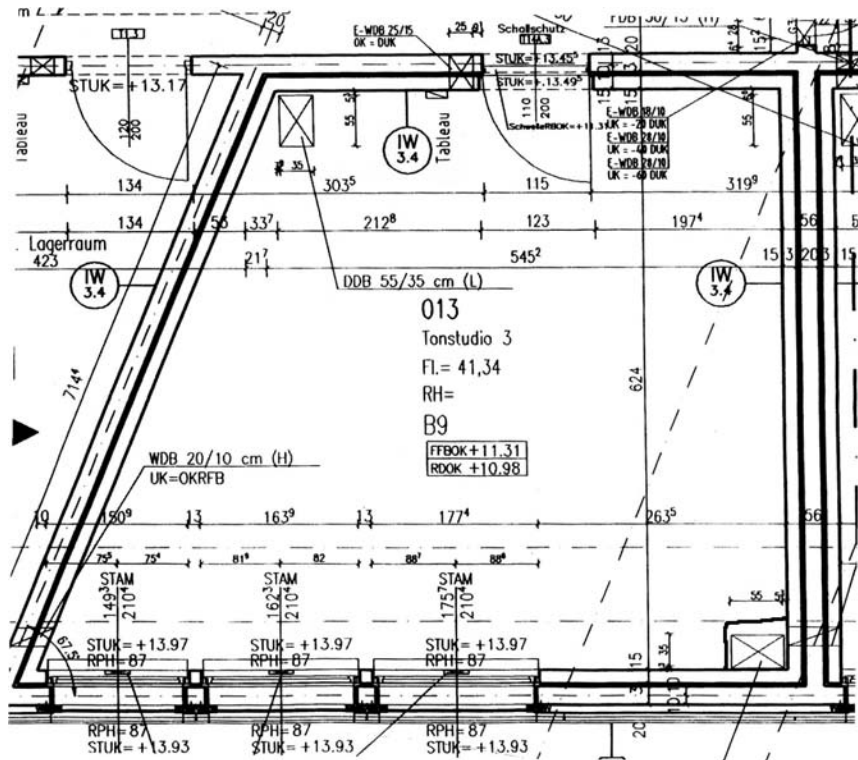


Abb. A26: Grundriss IEM-Produktionsstudio (Regieraum);
 Breite 4,80m (bzw. 7,20m), Länge 5,60m, Höhe 3,50m;
 Grundfläche 33,6m², Raumvolumen 117,6 m³

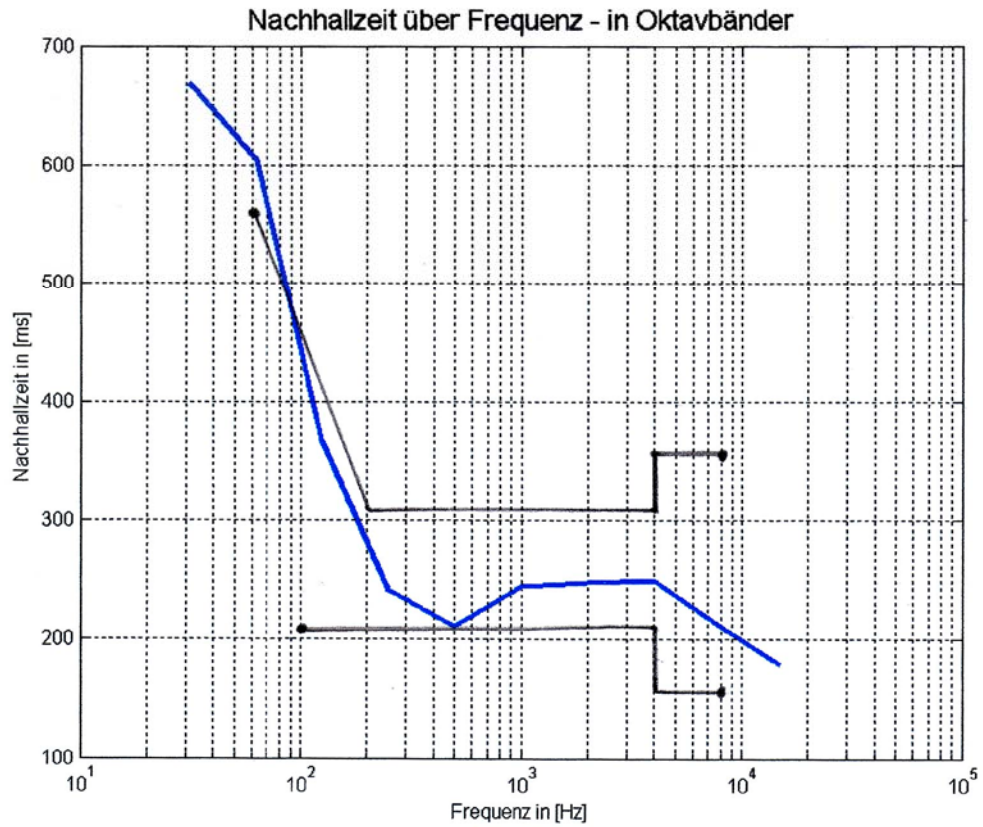


Abb. A27: Nachhallzeit des Regieraums des IEM-Produktionsstudios, mit oberem und unterem Grenzwert gemäß ITU-R Empfehlung BS.1116-1 (Messung übernommen aus Meßprotokoll von Dr. Sontacchi, basierend auf ETC der oktavbandgefilterten Impulsantwort des Centerlautsprechers am idealen Abhörpunkt)

Hörvergleich Surround-Hauptmikrofonsysteme "Orchesteraufnahme"

Name: _____

Datum: _____

Bitte markieren Sie jeweils durch Eintragung eines X in das passende Feld - Danke !

Ich bin -

Audiotechniker:

Musiker:

Laie:

Vergleichspaar #: **1**

(bitte von der Testperson einzutragen)

1) Wie groß empfinden Sie den Unterschied zwischen den beiden Hörbeispielen ?

sehr groß

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 kein Unterschied

2) Subjektives Gefallen der Aufnahme:

A: sehr gut

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 gar nicht

B: sehr gut

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 gar nicht

3) Natürlichkeit:

A: sehr

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 wenig

B: sehr

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 wenig

4) Klangfarbe des Orchesters:

A: natürlich

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: unnatürlich

B: natürlich

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: unnatürlich

A: höhenbetont

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: bass-betont

B: höhenbetont

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: bass-betont

5) Lokalisation L-R:

(ungefähre Anordnung der Instrumente bzw. Instr.Gruppen entlang der Bühnenbreite von l nach r: Klav, Schlagw, Hörner, Harfe, Viol., Xyloph., Holzbläser, Celli, Blechbläser, Kontra-Bässe)

Wie gut lassen sich die Instrumente bzw. Instrumentalgruppen im Orchester orten ?

A: sehr gut

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: nicht gut

B: sehr gut

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: nicht gut

6) Balance:

Sind die einzelnen Instrumente(ngruppen) im Orchester lautstärkenmässig ausgewogen ?

(siehe dazu auch die Liste der vorkommenden Instrumente(ngruppen) unter Punkt "Lokalisation")

A: mehr

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: weniger

B: mehr

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: weniger

7) Räumliche Abbildung des Orchesters:

"Stereo-Basisbreite" A: breit

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: schmal

B: breit

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: schmal

Entfernungseindruck A: nahe

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: entfernt

B: nahe

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: entfernt

Tiefenstaffelung A: tief

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: flach

B: tief

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: flach

Stabilität der Abb. A: stabil

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: instabil

B: stabil

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: instabil

Differenziertheit A: präzise

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: verschwommen

B: präzise

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: verschwommen

8) Abbildung des Raumes:

A: überzeugend

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: unbefriedigend

B: überzeugend

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: unbefriedigend

A: zu viel Raum (zu hallig)

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: zu wenig Raum (zu trocken)

B: zu viel Raum (zu hallig)

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: zu wenig Raum (zu trocken)

9) Hörbarkeit der Surroundkanäle:

"positiv"

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 "negativ"

A: fügen sich unauffällig ein

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: sind herauszuhören

B: fügen sich unauffällig ein

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: sind herauszuhören

A: sind zu leise

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: sind zu laut

B: sind zu leise

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: sind zu laut

Kommentarzeile zu Punkt _____ : (bei Bedarf vom Testhörer zu verwenden)

Abb. A28: Fragebogen Hörvergleich ,ORCH 5.1‘ (Lautsprecher)

Hörvergleich KUNSTKOPF-ORCHESTERAUFNAHME

Name: _____

Datum: _____

Bitte markieren Sie jeweils durch Eintragung eines X in das passende Feld - Danke !

Ich bin Audiotechniker: Musiker:

Vergleichspaar #: _____ (bitte von der Testperson einzutragen)

1) Wie groß empfinden Sie den Unterschied zum Referenzbeispiel "REF" ?

| | | | | | | | | |
|--------------|---|---|---|---|---|---|---|------------------|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| A: sehr groß | | | | | | | | kein Unterschied |
| B: sehr groß | | | | | | | | kein Unterschied |

2) Wie groß empfinden Sie den Unterschied zwischen Beispiel A und B ?

| | | | | | | | | |
|-----------|--|--|--|--|--|--|--|------------------|
| sehr groß | | | | | | | | kein Unterschied |
|-----------|--|--|--|--|--|--|--|------------------|

3) Subjektives Gefallen der Aufnahme:

| | | | | | | | | |
|---------------|--|--|--|--|--|--|--|-----------|
| REF: sehr gut | | | | | | | | gar nicht |
| A: sehr gut | | | | | | | | gar nicht |
| B: sehr gut | | | | | | | | gar nicht |

4) Natürlichkeit:

| | | | | | | | | |
|-----------|--|--|--|--|--|--|--|-------|
| REF: sehr | | | | | | | | wenig |
| A: sehr | | | | | | | | wenig |
| B: sehr | | | | | | | | wenig |

5) Klangfarbe des Orchesters:

| | | | | | | | | |
|------------------|--|--|--|--|--|--|--|------------------|
| REF: natürlich | | | | | | | | REF: unnatürlich |
| A: natürlich | | | | | | | | A: unnatürlich |
| B: natürlich | | | | | | | | B: unnatürlich |
| REF: höhenbetont | | | | | | | | REF: bass-betont |
| A: höhenbetont | | | | | | | | A: bass-betont |
| B: höhenbetont | | | | | | | | B: bass-betont |

6) Lokalisation L-R:

(ungefähre Anordnung der Instrumente bzw. Instr.Gruppen entlang der Bühnenbreite von l nach r:

Klav, Schlagw, Hörner, Harfe, Viol., Xyloph., Holzbläser, Celli, Blechbläser, Kontra-Bässe)

Wie gut lassen sich die Instrumente bzw. Instrumentalgruppen im Orchester orten ?

| | | | | | | | | |
|---------------|--|--|--|--|--|--|--|----------------|
| REF: sehr gut | | | | | | | | REF: nicht gut |
| A: sehr gut | | | | | | | | A: nicht gut |
| B: sehr gut | | | | | | | | B: nicht gut |

7) Balance:

Sind die einzelnen Instrumente(ngruppen) im Orchester lautstärkenmässig ausgewogen ?

(siehe dazu auch die Liste der vorkommenden Instrumente(ngruppen) unter Punkt "Lokalisation")

| | | | | | | | | |
|-----------|--|--|--|--|--|--|--|--------------|
| REF: mehr | | | | | | | | REF: weniger |
| A: mehr | | | | | | | | A: weniger |
| B: mehr | | | | | | | | B: weniger |

8) Räumliche Abbildung des Orchesters:

| | | | | | | | | |
|---------------------------------|--|--|--|--|--|--|--|-------------|
| "Stereo-Basisbreite" REF: breit | | | | | | | | REF: schmal |
| A: breit | | | | | | | | A: schmal |
| B: breit | | | | | | | | B: schmal |

| | | | | | | | | |
|------------------------------|--|--|--|--|--|--|--|---------------|
| Entfernungseindruck REF:nahe | | | | | | | | REF: entfernt |
| A: nahe | | | | | | | | A: entfernt |
| B: nahe | | | | | | | | B: entfernt |

| | | | | | | | | |
|----------------------------|--|--|--|--|--|--|--|------------|
| Tiefenstaffelung REF: tief | | | | | | | | REF: flach |
| A: tief | | | | | | | | A: flach |
| B: tief | | | | | | | | B: flach |

| | | | | | | | | |
|---------------------------------|--|--|--|--|--|--|--|---------------|
| Stabilität der Abb. REF: stabil | | | | | | | | REF: instabil |
| A: stabil | | | | | | | | A: instabil |
| B: stabil | | | | | | | | B: instabil |

| | | | | | | | | |
|--------------------------------|--|--|--|--|--|--|--|-------------------|
| Differenziertheit REF: präzise | | | | | | | | REF: verschwommen |
| A: präzise | | | | | | | | A: verschwommen |
| B: präzise | | | | | | | | B: verschwommen |

9) Abbildung des Raumes:

| | | | | | | | | |
|------------------|--|--|--|--|--|--|--|---------------------|
| REF: überzeugend | | | | | | | | REF: unbefriedigend |
| A: überzeugend | | | | | | | | A: unbefriedigend |
| B: überzeugend | | | | | | | | B: unbefriedigend |

| | | | | | | | | |
|--------------------------------|--|--|--|--|--|--|--|------------------------------|
| REF: zu viel Raum (zu hallig) | | | | | | | | R: zu wenig Raum(zu trocken) |
| A: zu viel Raum (zu hallig) | | | | | | | | A: zu wenig Raum(zu trocken) |
| B: zu viel Raum (zu hallig) | | | | | | | | B: zu wenig Raum(zu trocken) |

Kommentarzeile zu Punkt _____ : (bei Bedarf vom Testhörer zu verwenden)

Abb. A29: Fragebogen Hörvergleich ,ORCH-KK' (Kopfhörer)

Hörvergleich Surround-Hauptmikrofonsysteme "Duo Klavier-Cello"

Name: _____

Datum: _____

Bitte markieren Sie jeweils durch Eintragung eines X in das passende Feld - Danke !

Ich bin -

Audiotechniker:

Musiker:

Laie:

Vergleichspaar #: _____ (bitte von der Testperson einzutragen)

1) Wie groß empfinden Sie den Unterschied zwischen den beiden Hörbeispielen ?

sehr groß

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 kein Unterschied

2) Subjektives Gefallen der Aufnahme:

A: sehr gut

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 gar nicht

B: sehr gut

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 gar nicht

3) Natürlichkeit:

A: sehr

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 wenig

B: sehr

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 wenig

4) Klangfarbe der Instrumente:

A: natürlich

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: unnatürlich

B: natürlich

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: unnatürlich

A: höhenbetont

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: bass-betont

B: höhenbetont

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: bass-betont

5) Lokalisation L-R:

Wo lokalisieren Sie das Cello ?

Mitte

A: ganz L

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: ganz R

B: ganz L

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: ganz R

Wie stabil ist für Sie die Ortung der Instrumente ?

Kommt es zu Lokalisationsverzeichnungen (kurzzeitige Fehl-Ortung) ?

Stabilität A: stabil

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: instabil

B: stabil

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: instabil

6) Balance:

Stimmt für Sie das Lautstärkenverhältnis zwischen den Instrumenten ?

A: ausgewogen

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: unausgewogen

B: ausgewogen

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: unausgewogen

7) Räumliche Abbildung der Instrumente:

"Stereo-Basisbreite" A: breit

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: schmal

B: breit

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: schmal

Entfernungseindruck A: nahe

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: entfernt

B: nahe

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: entfernt

Tiefenstaffelung A: tief

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: flach

B: tief

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: flach

Differenziertheit A: präzise

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: verschwommen

B: präzise

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: verschwommen

8) Abbildung des Raumes:

A: überzeugend

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: unbefriedigend

B: überzeugend

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: unbefriedigend

A: zu viel Raum (zu hallig)

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: zu wenig Raum (zu trocken)

B: zu viel Raum (zu hallig)

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: zu wenig Raum (zu trocken)

9) Hörbarkeit der Surroundkanäle:

"positiv" "negativ"

A: fügen sich unauffällig ein

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: sind herauszuhören

B: fügen sich unauffällig ein

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: sind herauszuhören

A: sind zu leise

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 A: sind zu laut

B: sind zu leise

| | | | | | | | |
|--|--|--|--|--|--|--|--|
| | | | | | | | |
|--|--|--|--|--|--|--|--|

 B: sind zu laut

Kommentarzeile zu Punkt _____ : (bei Bedarf vom Testhörer zu verwenden)

Abb. A30: Fragebogen Hörvergleich ‚DUO 5.1‘ (Lautsprecher)

Referenz-Email 1 zu OCT-Array (Cornelis van der Gragt)

Von: sursound-bounces@music.vt.edu im Auftrag von ThomasChen@aol.com

Gesendet: Donnerstag, 07. Februar 2008 15:53

An: sursound@music.vt.edu

Betreff: Re: [Sursound] Multichannel rec techniques

I have used the OCT array however the hypercardoids are about 1 meter apart and instead of a cardioid in the center I use M/S Schoeps. I find it gives a better impression of size in the recording.

Also there is less of the shift between speakers when listening

ThomasChen

In a message dated 2/6/2008 3:32:57 A.M. Pacific Standard Time, cvdgragt@koncon.nl writes:

Hello List,

Indeed, from the recordings that my students and I made the last few years, we consider OCT to be one of the valuable tools for 5.0 recordings of classical music and larger musical formations (in our case symphonic, and wind orchestras, organs, choirs, etc.). However, when using it we should bear in mind that:

1. The supercardioids involved, which are responsible for the very good separation between the L and R signals, must have excellent off-axis frequency responses, normally not the case with directional mics. For this reason and specially for the OCT configuration, Schoeps developed the capsule MK41V.
2. the system only works well when the lowfrequency roll-off of supercardioids (60 percent pressuregradient driven) is compensated by 2 omnis that must be low-passed at, say, 100hz.
3. Never use the proposed backwards-directed 2 cardioids at about 1 meter from the OCT support. The crosstalk from the direct signal ruins the result. We had good results with a Hamasaki Square at considerable distance or simply a backwards directed 'traditional' X/Y(90) pair. Again, positioned at considerable distance, the latter with its more than 90 degrees coverage angle, very smoothly captures the ambient signals at the back of a hall

cheers

Cornelis H. Van der Gragt
Consultant for Acoustics, Audio and Music Recording
Jan Steenlaan 8
NL-1213 EL Hilversum
phone: int 31 35 624 61 64 (home)
mobile: int 31 62 551 77 23
E-mail: Cornelis van der Gragt <cvdgragt@hotmail.com>
or:

Royal Conservatoire
Music Recording Dept./Art-of-Sound Program
NL-2595 CA The Hague
phone: int 31 35 624 61 64 (home)
mobile: int 31 62 551 77 23
E-mail: Cornelis van der Gragt <cvdgragt@koncon.nl>

Referenz-Email 2 zu OCT-Array (Jan Korte)

Von: sursound-bounces@music.vt.edu im Auftrag von Jan Ola Korte
[mail@jan-korte.de]

Gesendet: Mittwoch, 06. Februar 2008 20:04

An: Surround Sound discussion group

Betreff: Re: [Sursound] Multichannel rec techniques

Hello,

Am 06.02.2008 um 12:31 schrieb Cornelis van der Gragt:

- > 3. Never use the proposed backwards-directed 2 cardioids at about 1
- > meter
- > from the OCT support. The crosstalk from the direct signal ruins
- > the result.

I used this setup recording a symphonic orchestra with choir in a big shoebox-hall and can absolutely confirm that. As the recording was part of a student paper about recording techniques for 3/2-Surround I had some different setups to compare. Setups that delivered best performance for the rear speakers were definitely the ones placed in the diffuse field a couple of meters from the reverberation radius. The one I considered best were two fig-of-eights facing the side walls with a base of 2 meters and about 6 meters from the reverberation radius (kind of a half Hamasaki-Square).

Best,

Jan

Sursound mailing list

Sursound@music.vt.edu <https://mail.music.vt.edu/mailman/listinfo/sursound>

Dokumentation zum Hörvergleich ‚DUO 5.1‘ (Lautsprecher)

Im Rahmen der vorliegenden Untersuchung wurde auch eine eigens angefertigte Surround-Tonaufnahme einer Duo-Besetzung sowohl einer Hörerbewertung wie auch einer messtechnischen Auswertung unterzogen.

Ursprünglich war nicht geplant gewesen, die Duo-Aufnahmen in die Dissertation einzubinden. Die Aufnahmen, Beurteilung durch die Testhörer und messtechnische Auswertung sollten dem Autor in erster Linie eine objektivere Beurteilung des von ihm vorgeschlagenen ‚Blumlein-Pfanzagl-Triple‘ Surroundmikrofons ermöglichen.

Gegenstand der Aufnahme war ein Stück im romantischen Stil für Violoncello und Klavier [Pfanzagl-Cardone, 2003]; der Aufnahmeort war ein größerer Proberaum im Festspielhaus Salzburg, mit einem Raumvolumen von ca. 1815 Kubikmetern. (L x B x H: 20.5 x 11.35 x 7.8 [m])

Dieser Aufnahmeraum weist eine mittlere Nachhallzeit von 1.54 s bei 1 kHz auf, was einem Hallradius von ca. 2m entspricht. Dieser wächst für eine Frequenz von 8kHz auf etwas über 3m an.

Die Position der Musiker relativ zu den Surround-Mikrofonsystemen ist aus dem nachfolgenden Dokumentationsphoto in Abb. A31 ersichtlich. Wie auch in den Tonaufnahmen zu hören war, ergab sich – aufgrund des relativ geringen mittleren Hallradius von ca. 2m - das Problem, dass das Cello sich innerhalb desselben befand, während das Klavier (auch durch sein – aufgrund der Deckelreflexion – teilweise indirektes Abstrahlverhalten) sich außerhalb desselben befand.

Bei einer der Mikrofontechniken, welche für die Front-Anordnung ausschließlich Mikrofone mit Kugelcharakteristik verwendet (DECCA), überwiegt der Diffusschallanteil gegenüber dem Direktschall, sodass eine Lokalisation der beiden Instrumente so gut wie unmöglich ist. Dies äußert sich bei der Auswertung der Hörerantworten dann auch dahingehend, dass diese Aufnahme als signifikant ‚halliger‘ empfunden wurde, als die Vergleichsaufnahmen.



Abb. A31: Aufnahme des Duos Cello-Klavier in Proberaum 447 des Festspielhauses Salzburg im Mai 2003 (Weitwinkelobjektiv)

Aufgrund des starken Hallanteils (bzw. geringen Hallradius⁶) hätten die meisten Tonmeister für eine kommerziell zu verwertenden Aufnahme in diesem Saal vermutlich zumindest für das Klavier Stützmikrofone zusätzlich zu einem Hauptsystem eingesetzt.

Im Rahmen der Untersuchung unterblieb dies natürlich, da es ja um die Ermittlung der Abbildungseigenschaften der unterschiedlichen Hauptsysteme alleine ging. Damit ist aber eine bei Tonaufnahmen zu erfüllenden Grundbedingung für die Erzielung eines – für die meisten Hörer – befriedigenden Klangeindruckes nicht erfüllt: das Überwiegen des Direktschallanteils gegenüber dem Diffusschall; eine Bedingung, die z.B. auch Voraussetzung für eine befriedigende Lokalisation ist.

Dennoch soll in der Folge - der Vollständigkeit halber – in stark eingeschränkter Form - kurz auf die Ergebnisse der Hörerbewertung der DUO-Aufnahme eingegangen werden.

Verwendete Surround-Mikrofonarrays

Wie bereits im Rahmen der ORCHester-Aufnahme kamen auch hier wieder die Surround-Mikrofonssysteme OCT, DECCA, KFM und Soundfield zum Einsatz. Statt des AB-PC Systems wurde bei der DUO-Aufnahme das neu vorgeschlagene BPT-System verwendet: im Modus BPT 3.0 für die Gewinnung der Front-Signale L, C, R und mit zwei zur Saalrückwand weisenden Mikrofonen mit Nierencharakteristik zur Gewinnung der Signale für LS, RS. Der Abstand dieser Rear-Mikrofone (deren Signale auch für das DECCA System als LS, RS zur Anwendung kamen) zum Front-System betrug ca. 8-10m. Die Aufnahmen des Soundfield-Mikrofons wurden – primär aus Zeitgründen – nicht in die Hörerbewertung miteinbezogen.

Ein weiteres Manko – in diesem Fall allerdings nur die BPT-Technik betreffend – bestand darin, dass sich erst nach der Aufnahme herausstellte, dass eines der für das BPT verwendeten Großmembranmikrofone (Neumann U87) elektrisch nicht mehr ganz einwandfrei funktionierte, wodurch sich eine leichte Verzerrung der verwendeten Richtcharakteristik ‚Acht‘ ergab. Weniger als die nicht 100%-ige Einhaltung der Richtcharakteristik stört dabei aber die Vermutung, dass dieser Defekt möglicherweise auch Auswirkungen auf den Gesamtklang dieser Kapsel gehabt haben dürfte (eine sich in Abhängigkeit von der Frequenz ändernde Richtcharakteristik hat zumindest sog. ‚Off-Axis‘ Klangfärbungen zur Folge).

Weiters wäre als Argument gegen die Verwendung der Hörerbewertung des BPT-Mikrofonsystems im Vergleich mit den anderen Systemen der Einwand gerechtfertigt, dass für dieses System Großmembranmikrofone eines anderen Herstellers (Fa. Neumann) verwendet wurden, während bei allen anderen Surround-Mikrofonen Kleinmembran-Kapseln des Herstellers Schoeps zum Einsatz kamen.

Als weiterer – allerdings geringerer Einwand – darf das Argument ins Treffen geführt werden, dass das für die Aufnahme verwendete BPT-System hybrid aufgebaut war, und somit in der Genauigkeit der Ausrichtung der Kapseln natürlich gegenüber den anderen Systemen benachteiligt ist (dieses Argument gilt zumindest im Vergleich mit dem KFM-System). Die Genauigkeit der Ausrichtung der Einzel-Mikrofone, die bei OCT und DECCA System zum Einsatz kamen, dürfte – nicht zuletzt aufgrund der

größeren Kapselabstände - weit weniger kritisch sein, als dies beim koinzidenten BPT-(Front)System der Fall ist.

Testhörer, Testdauer und modifizierter Fragebogen

Eine Gruppe von 15 Probanden (aus dem selben Personen-Pool, wie für die anderen Hörvergleiche stammend) musste die vier Surround-Mikrofontechniken wiederum unter double-blind Bedingungen mittels direkter Größenschätzung anhand eines semantischen Differenzials verbunden mit einem Paarvergleich und 7-stufiger Bewertungsskala beurteilen. Die Testdauer lag – je nach Proband – zwischen 60 und 90min.

Bei einem ersten, rein informellen Anhören der DUO-Aufnahmen schien aufzufallen, dass es beim OCT-System - aufgrund der Verwendung von Hypernieren und deren seitlicher Ausrichtung – zu stärkeren Lokalisationsverzeichnungen aufgrund von Bewegungen der Schallquelle (Bewegung des Cellos durch seinen Spieler) bzw. dem frequenzabhängigen Abstrahlverhalten des Instruments kommt, als bei den anderen Mikrofonssystemen. Aus diesem Grund wurde die Frage betreffend die Lokalisation dahingehend modifiziert, dass der Testhörer nun nicht mehr die Genauigkeit der Lokalisation bewerten sollte (welche ja – unabhängig von der vom Hörer wahrgenommenen Position des Instruments – u. U. stets gleich gut sein könnte), sondern statt dessen *wo* im Stereo-Panorama das Instrument vom Hörer wahrgenommen wurde.

Eine Kopie des Fragebogens ist der Dokumentation (vor den Emails von Cornelis van der Gragt und Jan Korte) vorangestellt (siehe S. 313).

Die ausgewählten Ausschnitte in der Dauer von ca. 1min. wurden den Testhörern über eine 5.1 Lautsprecheranlage vorgespielt. Anlage und Wiedergabebedingungen entsprachen dabei jenen des Hörvergleichs ‚ORCH 5.1‘, da wiederum die gleichen Abhörräume (Regieraum des IEM, Produktionsregieraum im Salzburger Festspielhaus) zum Einsatz kamen.

Ergebnisse der Hörerbewertung der DUO-Aufnahme

Für 13 der 15 Fragen ergab die Hörerbewertung statistisch aussagekräftige Resultate ($p > 0.05$). Ausnahmen hierbei waren die Bewertungen der Aspekte ‚Differenziertheit‘ und ‚Surr.-Kanäle nicht auffallend/ auffallend‘.

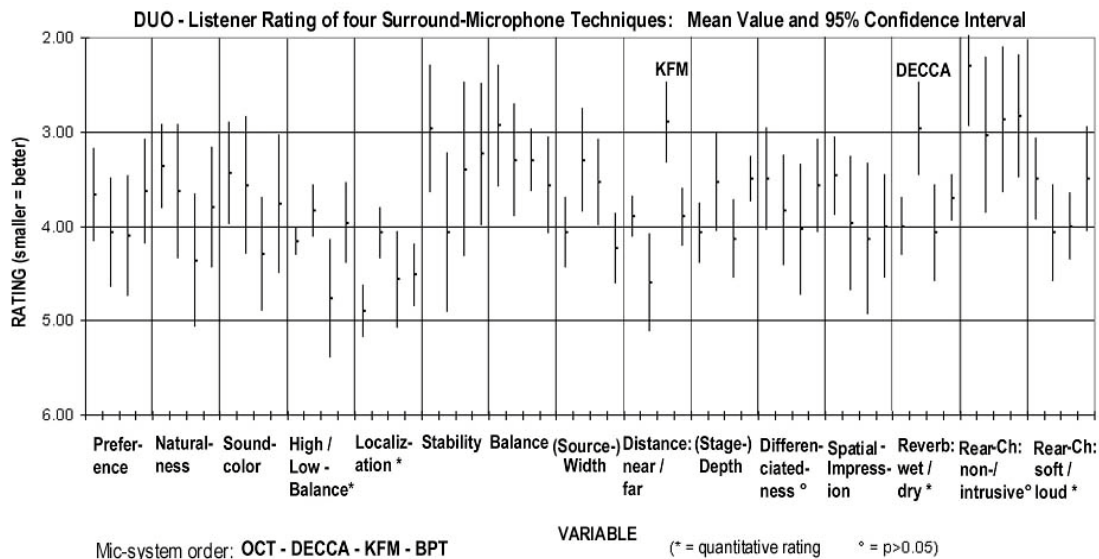


Abb. A32: ‚DUO 5.1‘ – Darstellung von Mittelwert und 95%-Konfidenzintervall für 4 Mikrofontechniken betreffend 15 Aspekte (15 Testhörer)

Bei Betrachtung der Abb. A32 fällt im Vergleich mit der Hörerbewertung der Aufnahme ‚ORCH 5.1‘ auf, dass für die meisten Aspekte die Konfidenzintervalle bei der ‚DUO 5.1‘ Aufnahme wesentlich breiter sind, und zwar unabhängig von der Mikrofontechnik.

Zwei Bewertungen fallen allerdings sofort ins Auge: das KFM-System wurde als deutlich ‚näher‘ empfunden, als die anderen Systeme, und das ‚DECCA‘-System als deutlich ‚halliger‘.

Betreffend den Aspekt Lokalisation stellte sich folgendes heraus:

Tatsächlich traten statistisch signifikante Unterschiede – wenngleich auch nur zwischen zwei der Techniken – auf: das Cello wird beim OCT-System (im Vergleich zum DECCA-System) fast einen ganzen Skalenwert weiter rechts lokalisiert. Dass der Unterschied zwischen gerade diesen beiden Systemen signifikant ist, hat vermutlich mit kurzzeitigen, rechtslastigen Lokalisationsverzeichnungen des OCT-Arrays genauso viel zu tun, wie mit der bereits oben angesprochenen Tatsache, dass das DECCA-Array aufgrund der Verwendung von ungerichteten Mikrofonen für das

Front-System und des daraus resultierenden Überwiegens von Diffusschall eine Lokalisation der Instrumente kaum zulässt.

Bezogen auf den Mittelwert wird das Cello also beim OCT-System weiter rechts abgebildet, als bei den anderen Systemen.

Das BPT-System zeichnet sich – zumindest im Vergleich mit dem OCT-System statistisch signifikant - durch eine höhere Tiefenstaffelung aus; der im Vergleich zum Cello geringere Pegel des Klaviers scheint dann allerdings auch in einer schlechteren Wiedergabe der Balance des Duos zu resultieren.