

Instrumentelle Evaluierung von Innenraum-Kommunikationssystemen

Dissertation

zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften

(Dr.-Ing.)

der Technischen Fakultät

der Christian-Albrechts-Universität zu Kiel

vorgelegt von

Anne Theiß

Kiel 2016

Tag der Einreichung:

31.10.2016

Tag der Disputation:

03.04.2017

Berichterstatter:

Prof. Dr.-Ing. Gerhard Schmidt

Prof. Dr.-Ing. Sebastian Möller

Danksagung

„Wer nicht manchmal das Unmögliche wagt, der wird das Mögliche nie erreichen.“

Max von Eyth

Mit diesem Gedanken begann ich im Frühjahr 2011 die Arbeit als wissenschaftliche Mitarbeiterin der Arbeitsgruppe digitale Signalverarbeitung und Systemtheorie (DSS) an der Christian-Albrechts-Universität zu Kiel. Während dieser Zeit ist die hier vorgelegte Dissertation entstanden. Die Chance zu promovieren kam eher unverhofft, als ich gerade meine Auslandsstudium in Valencia absolvierte. Die Entscheidung, diese Chance zu nutzen, habe ich, wie so Vieles in meinem Leben, aus dem Bauch heraus getroffen und, um ehrlich zu sein, während meiner Promotionszeit mehr als nur einmal in Frage gestellt, aber niemals bereut. Genau aus diesem Grund geht mein erster Dank an meinen Doktorvater Prof. Dr.-Ing. Gerhard Schmidt, der mir diese einmalige Chance ermöglicht und die Promotion zugetraut hat. Die Gewissheit, dass seine Bürotür für meine Fragen und Probleme immer offen steht, hat mir stets das Gefühl geben, mit meiner Arbeit nie alleine zu sein. Danke Gerd!

Außerdem möchte ich mich auf diesem Wege bei Herrn Prof. Dr.-Ing. Sebastian Möller für das Interesse an meiner Arbeit und die Übernahme des Koreferats bedanken. Ebenso danke ich dem Vorsitzenden der Prüfungskommission Herrn Prof. Dr.-Ing. Jeffrey McCord, sowie Herrn Prof. Dr.-Ing. Ludger Klinkenbusch, der sich kurzfristig bereit erklärte, ein Teil der Prüfungskommission zu sein.

Ein weiterer Dank geht an die Kollegen der Arbeitsgruppe DSS und natürlich auch der Arbeitsgruppe Informations- und Codierungstheorie (ICT). Die vielen Kaffeeküchensgespräche und auch fachlichen bzw. nicht fachlichen Diskussionen haben den Arbeitsalltag zu etwas Besonderem gemacht. Ich war immer gerne ein Teil dieser außergewöhnlichen Gemeinschaft und freue mich sagen zu können, dass aus Kollegen viele Freunde geworden sind. Danke euch Allen für die tolle Zeit!

Ein mir sehr wichtiger Dank geht an dieser Stelle an meine Familie, die mich immer aufgefangen und unterstützt hat. Vor allen Dingen möchte ich meinen Eltern danken, die

meiner Schwester und mir immer das Gefühl gaben, dass alles möglich ist, man es nur in Angriff nehmen muss. Ohne ihre Rückhalt hätte ich wohl nicht das Selbstvertrauen gehabt, trotz aller Widrigkeiten, das Studium und auch die Promotion zu beginnen und erfolgreich abzuschließen. Auch meiner Schwester möchte ich hier meinen Dank ausdrücken. Sie hat jederzeit ein offenes Ohr für meine Problem und hat mich, trotz der großen räumlichen Distanz, stets unterstützt und ermutigt. Danke!

Den für mich schwierigsten Teil, das gesammelte Wissen zu Papier zu bringen, gelang mir in dieser Form nur durch die Hilfe von drei Personen: Alexander Meißner, Matthias Bachmann und Till Simon. Vielen Dank für eure Zeit und Ausdauer!

Flensburg, im Mai 2017

Anne Theiß

Abstract

The communication between the occupants of a speeding vehicle can be impaired. In particular, the conversation between the passengers of the front row and those of the rear seats is due to the occurring background noise and the arrangement of the passengers difficult. Usually the passengers change their position and their voice in order to compensate for this impairment. Another method to improve the communication situation is a so-called in-car communication system (ICC system). Such systems take the speech signal of a talking person using the already installed microphones, process and play back the amplified desired signal inside the passenger compartment through the loudspeakers nearby the listening passengers. Thus, the power of the desired signal inside the compartment increases as well as the intelligibility.

Following the successful implementation and development of such systems, the question arises to the quality of the ICC system. To answer this question, in this work, first, the term "quality" referring to an ICC system is defined. Subsequently, an evaluation strategy, which consists of a variety of individual instrumental methods, is presented. For a plausible classification of the methods and a transparent investigation of the overall quality, these methods are divided into three different evaluation groups. These evaluation groups examine the characteristics of the vehicle, the behavior of the ICC system itself, as well as the quality of the communication. Within each of these groups first instrumental methods are described and validated. Subsequently, the individual results of the instrumental methods are weighted by previously calculated attributes and combined into an overall result. The attributes of the vehicle and the ICC system are determined based on the determined characteristics of the vehicle. For the development of the instrumental methods a quality reference which reflects the subjective perception of passengers is required. For this purpose, some auditory evaluation methods are introduced in addition to the instrumental methods and the results analyzed. Therefore, not only listening tests, but also psychoacoustic perception experiments have been carried out and analyzed. The evaluation strategy implemented into a real-time framework is finally tested on a specific ICC system and an overall view of the quality of all its results is proposed.

Kurzfassung

Die Kommunikation zwischen den Insassen eines schnell fahrenden Fahrzeuges kann stark beeinträchtigt sein. Insbesondere die Gespräche zwischen den Passagieren der vorderen Sitzreihe und denen der Rücksitzbank sind, aufgrund der Geräuschkulisse und der Anordnung, erschwert. Für gewöhnlich verändern die Insassen ihre Position und ihre Sprache, um diese Beeinträchtigung zu kompensieren. Eine weitere Methode die Kommunikationssituation zu verbessern, ist ein sogenanntes Innenraumkommunikationssystem (engl. *In-Car Communication*-System, ICC-System). Solche Systeme nehmen das Sprachsignal einer Person mithilfe von Mikrofonen auf, verarbeiten dieses und geben das verstärkte Nutzsignal in der Fahrgastzelle, nahe dem Zuhörer wieder aus. Dadurch erhöht sich die Nutzsignalleistung im Fahrzeug und die Verständlichkeit.

Nach der erfolgreichen Umsetzung und Entwicklung solcher Systeme stellt sich die Frage nach der Qualität des ICC-Systems. Zur Beantwortung dieser Frage wird in dieser Arbeit zunächst der Begriff „Qualität“ im Sinne eines ICC-Systems eingeordnet. Anschließend wird eine Evaluierungsstrategie, welche aus einer Vielzahl an einzelnen, instrumentellen Verfahren besteht, vorgestellt. Für eine eingängige Klassifizierung der Verfahren und eine transparente Erhebung einer Gesamtqualität, werden diese in drei verschiedene Evaluierungsgruppen unterteilt. Es handelt sich hierbei um Verfahren, welche die Eigenschaften des Fahrzeuges erheben, das Verhalten des ICC-Systems selbst evaluieren, sowie Verfahren, welche die Kommunikationsqualität bewerten. Innerhalb jeder dieser Gruppen werden erste instrumentelle Verfahren beschrieben und validiert. Im weiteren Verlauf werden die einzelnen Ergebnisse der instrumentellen Verfahren mittels zuvor berechneter Attribute gewichtet und zu einem Gesamtergebnis kombiniert. Dabei werden die Attribute des Fahrzeuges und des ICC-Systems anhand der ermittelten Fahrzeugeigenschaften bestimmt. Für die Entwicklung der instrumentellen Verfahren zur Schätzung der Qualität wird eine Qualitätsreferenz, welche das subjektive Empfinden von Passagieren abbildet, benötigt. Zu diesem Zwecke werden neben den instrumentellen Verfahren auch einige auditive Evaluierungsverfahren vorgestellt und deren Ergebnisse analysiert worden. Dabei sind nicht ausschließlich Hörtests, sondern auch psychoakustische Perzeptionsexperimente durchgeführt und analysiert. Die in einem Echtzeitrahmenwerk umgesetzte Evaluierungsstrategie ist abschließend an einem konkreten ICC-System erprobt und eine Gesamtübersicht der Qualität mit all ihren Ergebnissen vorgeschlagen worden.

Inhaltsverzeichnis

Abkürzungen und Notation	xi
1 Einleitung	1
1.1 Gliederung	3
2 Innenraum-Kommunikationssystem	5
2.1 Funktionsprinzip	7
3 Evaluierungsstrategie	13
3.1 Was ist Sprachqualität?	13
3.2 Einteilung in Evaluierungsgruppen	16
4 Grundlagen Qualitätsbeurteilung	19
4.1 Auditive Evaluierungsverfahren	19
4.2 Instrumentelle Evaluierungsverfahren	22
4.2.1 Projektion auf eine auditive Qualitätsskala	23
4.3 Verifikation instrumenteller Verfahren	24
5 Evaluierungsumgebung	29
5.1 Umgebungsgeräuschsimulation	32
5.2 Lombardsprachsignale	35
5.3 Implementierung	37
5.4 Theoretische Modellbildung	42
6 Auditive Evaluierungsverfahren	49
6.1 Echoperzeption	50
6.1.1 Ergebnisse	53
6.2 Eigenechoperzeption	56

6.2.1	Ergebnisse	60
6.3	Diskrepanz zwischen SNR und Qualitätswahrnehmung	63
6.3.1	Ergebnisse	65
6.4	Gesamtqualitätsevaluierung	66
6.4.1	Ergebnisse	69
6.5	Zusammenfassung	71
7	Instrumentelle Evaluierungsverfahren	75
7.1	Fahrzeugeigenschaften	77
7.1.1	Distanzen zwischen den Komponenten	78
7.1.2	Hintergrundgeräusch	88
7.1.3	Fahrzeugattribute	91
7.2	Verhalten des ICC-Systems	97
7.2.1	Unterstützung durch das ICC-System	97
7.2.2	Systemverzögerung	105
7.2.3	Erhöhung des Hintergrundgeräusches	110
7.3	Kommunikationsqualität	115
7.3.1	ICC-Systemattribute	115
7.3.2	Modifizierter Sprachübertragungsindex	117
7.3.3	Modifizierter Signal-Rauschabstand	123
7.3.4	Spektrale Distanz	127
7.4	Gesamtqualitätsbewertung	137
7.5	Zusammenfassung	140
8	Evaluierung eines konkreten ICC-Systems	143
8.1	Beschreibung der Testumgebung	143
8.2	Evaluierungsergebnisse	146
8.2.1	Fahrzeugeigenschaften	146
8.2.2	Verhalten des ICC-Systems	151
8.2.3	Kommunikationsqualität	155
8.3	Evaluierungsergebnisse bei variabler ICC-Systemverzögerung	162
9	Zusammenfassung	167
9.1	Ausblick	169
Anhang		170

A Analyse der Regressionskoeffizienten	171
Literaturverzeichnis	173

Abkürzungen

ACR	absolute Kategorisierung (engl. <i>Absolute Category Rating</i>)
DAM	diagnostisches Akzeptanzmaß (engl. <i>Diagnostic Acceptability Measure</i>)
DFT	Diskrete Fourier-Transformation
DSS	Digitale Signalverarbeitung und Systemtheorie
FFT	schnelle Fourier-Transformation (engl. <i>Fast Fourier Transform</i>)
HTML	<i>Hypertext Markup Language</i>
ICC-System	Innenraumkommunikationssystem (engl. <i>In-Car Communication-System</i>)
ITU	<i>International Telecommunication Union</i>
KiRAT	<i>Kiel Real-time Application Toolkit</i>
LDS	Leistungsdichtespektrum
MOS	mittlere Beurteilungsnote (engl. <i>Mean Opinion Score</i>)
MTF	Modulationsübertragungsfunktion (engl. <i>Modulation Transfer Function</i>)
MUSHRA	<i>MUlti Stimulus test with Hidden Reference and Anchor</i>
NDGC	geräuschabhängige Verstärkungsregelung (engl. <i>Noise-Dependent Gain Control</i>)
NLMS	<i>Normalized Least Mean Square</i>
PKW	Personenkraftwagen
RMSE	mittlerer quadratischer Fehler (engl. <i>Root Mean Square Error</i>)
SNR	Signal-Rauschabstand (engl. <i>Signal-to-Noise Ratio</i>)
STI	Sprachübertragungsindex (engl. <i>Speech Transmission Index</i>)
TTS-System	<i>Text-to-Speech-System</i>

Notation

Konventionen und Operationen

$x(n)$	Folge von Abtastwerten
$x_w(k, l)$	Endliches, gefensterter Segment des Signals $x(n)$ im k -ten Rahmen
$X(\mu, k)$	Diskrete Fourier-Transformation des (endlichen) Signals $x_w(k, l)$
$X^{(m)}(\mu, k)$	Diskrete Fourier-Transformation des m -ten Kanals
\mathbf{x}	Vektor
\mathbf{X}	Matrix
\bar{x}	Mittelwert der Größe x
\hat{x}	Schätzwert bzw. Erwartungswert der Größe x
x^*	Konjugiert komplexer Wert der Größe x
$s_{xx}(\kappa)$	Autokorrelationsfolge des Signals $x(n)$
$s_{yx}(\kappa)$	Kreuzkorrelationsfolge der Signale $y(n)$ und $x(n)$
$S_{xx}(\Omega)$	Autoleistungsdichtespektrum des Signals $x(n)$
$S_{yx}(\Omega)$	Kreuzleistungsdichtespektrum der Signale $y(n)$ und $x(n)$
f_s	Abtastfrequenz
Ω	Normierte Frequenz ($2\pi f/f_s$)
$h(\cdot)$	Impulsantwort
$H(\cdot)$	Übertragungsfunktion
$E\{\cdot\}$	Operation zur Berechnung des Erwartungswertes
$\mathcal{F}\{\cdot\}$	Zeitdiskrete Fourier-Transformation
$\ \cdot\ _1$	Summennorm
$\max\{\cdot\}$	Maximalwertoperator
$\min\{\cdot\}$	Minimalwertoperator
$\arg \max_{\kappa}\{\cdot\}$	Operation zur Bestimmung des Argument des Maximums
$\arg \min_{\kappa}\{\cdot\}$	Operation zur Bestimmung des Argument des Minimums

Formelzeichen

a	Regressionskonstante des einfachen linearen Modells
-----	---

b	Regressionskoeffizient des einfachen linearen Modells
$b_E(n)$	Signal des Hintergrundgeräusches in der Fahrgastzelle am Zuhörerohr
$b_M(n)$	Signal des Hintergrundgeräusches in der Fahrgastzelle am Mikrofon des ICC-Systems
$f(\mu)$	Frequenz in Hz des μ -ten Frequenzbandindex
f_i	Oktavband i
$h_{LE}(n)$	Impulsantwort vom ICC-Lautsprecher zum Zuhörerohr
$h_{LM}(n)$	Impulsantwort vom ICC-Lautsprecher zum ICC-Mikrofon (<i>feedback path</i>)
$h_{NLE}(n)$	Impulsantwort vom Lautsprecher der Geräuschsimulation zum Zuhörerohr
$h_{NLM}(n)$	Impulsantwort vom Lautsprecher der Geräuschsimulation zum ICC-Mikrofon
$h_{TE}(n)$	Impulsantwort vom Sprechermund zum Zuhörerohr (Pfad des Direktschalls)
$h_{TM}(n)$	Impulsantwort vom Sprechermund zum ICC-Mikrofon
$h_{tot}(n)$	Impulsantwort des ICC-Systems inklusive des Rückkopplungspfades $h_{LM}(n)$
I_S	Anzahl der Wertepaare einer Regressionsanalyse
k	Rahmenindex
l_I	Kanalnummer des ICC-Lautsprecher
l_K	Kanalnummer der Lautsprecher der künstlichen Sprecher
m_K	Kanalnummer der Mikrofone der künstlichen Zuhörer
m_I	Kanalnummer des ICC-Mikrofons
\overline{MOS}_S	Qualitätsergebnis der auditiven Gesamtqualitätserhebung
\overline{MOS}_S	Normiertes Qualitätsergebnis der auditiven Gesamtqualitätserhebung
N_{sub}	Subbänder im Frequenzbereich
n	Diskreter Zeitindex
N_{FFT}	Dimension der Fouriertransformation bzw. Fenstergröße
O_i	Instrumentelle Evaluierungsergebnisse
$P_{BE}^{(l)}$	Geräuschleistung des l -ten Geräuschszenarios des Echoperzeptionsexperimentes
$P_{BS}^{(l)}$	Geräuschleistung des l -ten Geräuschszenarios des Eigenechoperzeptionsexperimentes
Q_{tot}	Gesamtqualität des evaluierten ICC-Systems
R	Rahmenversatz
RMSE	Mittlerer quadratischer Fehler

RMSE_{LOO}	Mittlerer quadratischer Fehler unter Auslassung eines Datensatzes bzw. Verwendung der <i>jackknife</i> -Methode
RSS	Summe der Abweichungsquadrate
s_b	Abstand der Geräuschsimulationslautsprecher zur Testperson
s_p	Abstand der zweiten Signalquelle zur Testperson
$s_v^{(i)}$	Abstand zwischen Torso und Testperson
S_i	Auditive Evaluierungsergebnisse
$\hat{s}_{xy}(\kappa)$	Geschätzte Kreuzkorrelationsfolge zur Distanzbestimmung
$s(n)$	Direktschallkomponente am Ohr des Zuhörers
$t(n)$	Geräuschsignal am Zuhörerohr aufgenommen durch das ICC-Mikrofon und übertragen über das ICC-System
$v(n)$	Klares Sprachsignal am Zuhörerohr übertragen über das ICC-System
$v_t(n)$	Signal des ICC-Systems am Zuhörerohr
$w_{\text{hann}}(n)$	Fensterfunktion eines Hann-Fensters
$x(n)$	Reaktionssignal bzw. Ohrsignal des künstlichen Zuhörers
$X(\mu, k)$	Diskrete Fouriertransformation des Signals $x(n)$ gefenstert mit $w(n)$
$y(n)$	Referenzsignal bzw. Mundsignal des künstlichen Sprechers
$Y(\mu, k)$	Diskrete Fouriertransformation des Signals $y(n)$ gefenstert mit $w(n)$
$\delta_S^{(n)}$	Verzögerungseinstellung des Eigenechoperzeptionsexperiments
Δ_{ICC}	Zusätzliche Verzögerung innerhalb des kooperativen ICC-Systems
κ	Diskreter Zeitindex
λ_1	Verstärkungsanpassung des Torsos
λ_2	Verstärkungsanpassung des Lautsprechersignals
$\lambda_N^{(i)}$	Verstärkungsanpassung des Geräuschsignals
$\lambda_{\text{TE}}^{(i,1)}$	Verstärkungseinstellung einer Testperson innerhalb eines Szenario des Echoperzeptionsexperiments
$\lambda_{\text{TS}}^{(n,k,f_i)}$	Verstärkungseinstellung einer Testperson innerhalb eines Szenario des Eigenechoperzeptionsexperiments
μ	Frequenzbandindex
ρ	Empirischer Korrelationskoeffizient
ρ_{LOO}	Empirischer Korrelationskoeffizient unter Auslassung eines Datensatzes bzw. Verwendung der <i>jackknife</i> -Methode

Fahrzeugeigenschaften

b_N	Gewicht zur Bestimmung von g_N
$b_S^{(i)}$	Gewicht zur Bestimmung von g_S

$b_T^{(i)}$	Gewicht zur Bestimmung von g_T
\mathbf{B}	Matrix mit Gewichten zur Bestimmung der Fahrzeugattribute
\mathbf{C}_V	Ergebnismatrix Fahrzeugeigenschaften
\mathbf{g}	Vektor der Fahrzeugattribute
g_N	Fahrzeugattribut Hintergrundgeräusch
g_S	Fahrzeugattribut Fahrzeuggröße
g_T	Fahrzeugattribut Wandler ICC-System

Distanzen zwischen den Komponenten

$c_D^{(i)}$	Einflussgröße der i -ten Distanzauswertungsgruppe
D_V	Verzögerung zwischen zwei Komponenten
$D_V^{(i)}$	Verzögerung zwischen zwei Komponenten der i -ten Distanzauswertungsgruppe
N_{Est}	Anzahl der analysierten Rahmen
$u(r, \nu)$	Inverse Fouriertransformation der Spektren $U(\mu, \nu)$
$U(\mu, \nu)$	Summe der multiplizierten Spektren
$\tilde{x}(n)$	Hochpass gefiltertes Mundsignal des künstlichen Sprechers
$\tilde{x}(r, k)$	Segment der Länge R des Signals $x(n)$, aufgefüllt mit Nullen bis zu einer Länge von N_{FFT}
$\tilde{X}(\mu, k)$	Diskrete Fouriertransformation des Segmentes $\tilde{x}(r, k)$
$\tilde{y}(n)$	Hochpass gefiltertes Ohrsignal des künstlichen Zuhörers
$\tilde{y}(r, k)$	Segment der Länge R des Signals $y(n)$, aufgefüllt mit Nullen bis zu einer Länge von N_{FFT}
$\tilde{Y}(\mu, k)$	Diskrete Fouriertransformation des Segmentes $\tilde{y}(r, k)$
$\delta_{\max}^{(i)}$	Obergrenze der Einflussgröße $c_D^{(i)}$ der i -ten Distanzauswertungsgruppe
$\delta_{\min}^{(i)}$	Untergrenze der Einflussgröße $c_D^{(i)}$ der i -ten Distanzauswertungsgruppe

Hintergrundgeräuschanstieg

$A_w(\mu)$	A-Gewichtung
c_N	Einflussgröße der Hintergrundgeräuschemessung
K_{BG}	Anzahl der Rahmen zur Schätzung der Geräuschemessung
N_{dBA}	Geräuschemessungsschätzung in dBA
$N_{\text{dBA}}^{\text{loud}}$	Geräuschemessungsschätzung in dBA für ein Fahrscenario bei 150 km/h
$N_{\text{dBA}}^{\text{loud}}(m_K)$	Geräuschemessungsschätzung in dBA für ein Fahrscenario bei 150 km/h des Mikrofonkanals m_K
$N_{\text{dBA}}^{\text{low}}$	Geräuschemessungsschätzung in dBA für ein Fahrscenario im Stillstand

$N_{\text{dBA}}^{\text{low}}(m_K)$	Geräuschleistungsschätzung in dBA für ein Fahrscenario im Stillstand des Mikrofonkanals m_K
N_D	Erhöhung der Hintergrundgeräuschleistung
$N_D(m_K)$	Erhöhung der Hintergrundgeräuschleistung des Mikrofonkanals m_K
N_I	Maximaler Geräuschanstieg
$\bar{N}_{\text{veh}}(\mu)$	Hintergrundgeräusch im Fahrzeug
$\bar{N}_{\text{veh}}^{m_K}(\mu)$	Mittleres Betragsspektrum des Signals $x(n)$ in Abhängigkeit des Mikrofonkanals m_K
δ_{max}^N	Obergrenze der Einflussgröße
δ_{min}^N	Untergrenze der Einflussgröße
μ_E^N	Frequenzbandindex der letzten analysierten Frequenz
μ_S^N	Frequenzbandindex der ersten analysierten Frequenz

Verhalten ICC-System

\mathbf{c}_I	Vektor der Einzelqualitätsergebnisse der Evaluierungsgruppe
$\check{\mathbf{c}}_I$	Vektor der vorläufigen Qualitätsergebnisse
c_R	Qualitätsergebnis des Hintergrundgeräuschanstiegs
c_S	Qualitätsergebnis der Systemunterstützung
c_T	Qualitätsergebnis der Systemverzögerung
Q_I	Gesamtqualitätsergebnis der Evaluierungsgruppe
$\mathbf{w}_I(\mathbf{g})$	Gewichte für die Berechnung der Gesamtqualität

Systemunterstützung

$a_x^G(\mu, k)$	Gewichte für die Bestimmung des Leistungsdichtespektrums $S_{xx}^G(\mu)$
$a_y^G(\mu, k)$	Gewichte für die Bestimmung des Leistungsdichtespektrums $S_{yy}^G(\mu)$
\check{c}_S	Vorläufiges Qualitätsergebnis
\hat{G}_S	Geschätzte ICC-Systemunterstützung
G_S	ICC-Systemunterstützung
c_S	Qualitätsergebnis der Systemunterstützung
$S_{bb}(\mu)$	Leistungsdichtespektrum des Hintergrundgeräusches
$S_{xx}(e^{j\Omega})$	Leistungsdichtespektrum des Eingangssignals
$S_{xx}^G(\mu)$	Mittleres Leistungsdichtespektrum des Reaktionssignals bzw. Ohrsignals
$S_{xx}^{G,(\text{off})}(\mu)$	Mittleres Leistungsdichtespektrum des Reaktionssignals bei deaktiviertem ICC-System

$S_{xx}^{G,(on)}(\mu)$	Mittleres Leistungsdichtespektrum des Reaktionssignals bei aktiviertem ICC-System
$S_{yy}(e^{j\Omega})$	Leistungsdichtespektrum des Ausgangssignals
$S_{yy}^G(\mu)$	Mittleres Leistungsdichtespektrum des Referenzsignales
$R^{(off)}(\mu)$	Leistungsverhältnis zwischen Aus- und Eingang bei deaktiviertem ICC-System
$R^{(on)}(\mu)$	Leistungsverhältnis zwischen Aus- und Eingang bei aktiviertem ICC-System
$R_S(\mu)$	Frequenzabhängige Systemunterstützung
μ_{End}	Frequenzbandindex der letzten analysierten Frequenz
μ_{Start}	Frequenzbandindex der ersten analysierten Frequenz

Systemverzögerung

$\check{\chi}_T$	Vorläufiges Qualitätsergebnis
D_{sys}	Verzögerung des ICC-Systems
\hat{D}_{sys}	Geschätzte Verzögerung des ICC-Systems
$\hat{D}_{\text{sys}}(l_K, m_K)$	Geschätzte Verzögerung des ICC-Systems in Abhängigkeit des Lautsprecherkanals l_K und des Mikrofonkanals m_K
$s_{\tilde{x}\tilde{y}}(\kappa)$	Kreuzkorrelationsfolge zwischen Reaktionssignal und Referenzsignal
$\hat{s}_{\tilde{x}\tilde{y}}(\kappa)$	Geschätzte Kreuzkorrelationsfolge
$\hat{s}_{\tilde{x}\tilde{y}}^{(off)}(\kappa)$	Geschätzte Kreuzkorrelationsfolge bei deaktiviertem ICC-System
$\hat{s}_{\tilde{x}_{m_K}\tilde{y}_{l_K}}^{(off)}(\kappa)$	Geschätzte Kreuzkorrelationsfolge bei deaktiviertem ICC-System in Abhängigkeit des Lautsprecherkanals l_K und des Mikrofonkanals m_K
$\hat{s}_{\tilde{x}\tilde{y}}^{(on)}(\kappa)$	Geschätzte Kreuzkorrelationsfolge bei aktiviertem ICC-System
$\hat{s}_{\tilde{x}_{m_K}\tilde{y}_{l_K}}^{(on)}(\kappa)$	Geschätzte Kreuzkorrelationsfolge bei aktiviertem ICC-System in Abhängigkeit des Lautsprecherkanals l_K und des Mikrofonkanals m_K
$\tilde{x}(n)$	Hochpassgefiltertes Reaktionssignal $x(n)$
$\tilde{y}(n)$	Hochpassgefiltertes Referenzsignal $y(n)$

Erhöhung des Hintergrundgeräusches

$\check{\chi}_R$	Vorläufiges Qualitätsergebnis
$a_x^N(k)$	Gewichte zur Mittelung des Leistungsdichtespektrums $S_{yy}^G(\mu)$
\hat{G}_N	Geschätzter Mittlerer Anstieg des Hintergrundgeräusches in dB
$\hat{G}_N(m_K)$	Geschätzter Mittlerer Anstieg des Hintergrundgeräusches in dB in Abhängigkeit des Mikrofonkanals m_K
G_N	Mittlerer Anstieg des Hintergrundgeräusches in dB
$N_{xx}(\mu)$	Mittleres Leistungsdichtespektrum des Reaktionssignals $x(n)$

$N_{xx}^{(\text{on})}(\mu)$	Mittleres Leistungsdichtespektrum des Reaktionssignals $x(n)$ bei aktiviertem ICC-System
$N_{xx}^{(\text{off})}(\mu)$	Mittleres Leistungsdichtespektrum des Reaktionssignals $x(n)$ bei deaktiviertem ICC-System
$R_N(\mu)$	Veränderung des Hintergrundgeräusches

Kommunikation

\mathbf{c}_C	Vektor der Qualitätsergebnisse
c_{MS}	Qualitätsergebnis der SNR-Messung
c_{SD}	Qualitätsergebnis der spektralen Abstandsmessung
c_{SI}	Qualitätsergebnis der STI-Messung
g_R	ICC-Systemattribut Notwendigkeit
Q_C	Gesamtqualitätsergebnis
\mathbf{w}_C	Gewichte für die Berechnung der Gesamtqualität

Modifizierter Sprachübertragungsindex

\check{c}_{SI}	Vorläufiges Qualitätsergebnis
d	Diskreter Zeitindex nach der Unterabtastung
D	Anzahl der gepufferten Blöcke
D_{sub}	Faktor der Unterabtastung
f_c^{LP}	Grenzfrequenz des Tiefpasses
$I_x^{(f_i)}(\mu, d)$	Spektrum der Intensitätseinhüllenden des Reaktionssignals $x(n)$
$I_y^{(f_i)}(\mu, d)$	Spektrum der Intensitätseinhüllenden des Referenzsignals $y(n)$
$M^{(f_i)}(\mu)$	Modulationsübertragungsfunktion
$\text{MTI}^{(f_i)}$	Modulationsübertragungsindex
N_{Oct}	Anzahl der Oktavbänder
N_{FFT}^s	Fenstergröße der Fouriertransformation
$\text{SNR}_{\text{app}}^{(f_i)}(\nu)$	Effektive Signal-Rauschabstand
STI	Sprachübertragungsindex
$\text{STI}(l_K, m_K)$	Sprachübertragungsindex in Abhängigkeit des Lautsprecherkanals l_K und des Mikrofonkanals m_K
$\widehat{\text{STI}}$	Geschätzter Sprachübertragungsindex
$\text{TI}^{(f_i)}(\nu)$	Übertragungsindex
$w^{(f_i)}$	Oktavbandabhängige Gewichte für die Berechnung des STI
$x_k(n)$	In Blöcke unterteiltes Reaktionssignal $x(n)$
$x_k^{(f_i)}(n)$	In Blöcke n und Oktavbänder f_i unterteiltes Reaktionssignal $x(n)$
$y_k(n)$	In Blöcke unterteiltes Referenzsignal $y(n)$
$y_k^{(f_i)}(n)$	In Blöcke n und Oktavbänder f_i unterteiltes Referenzsignal $y(n)$
μ_{end}^m	Frequenzbandindex der ersten analysierten Frequenz

μ_{start}^m Frequenzbandindex der ersten analysierten Frequenz

Modifizierter Signal-Rauschabstand

\check{c}_{MS} Vorläufiges Qualitätsergebnis
 K_{NE} Anzahl der verwendeten Rahmen für die Schätzung des Hintergrundgeräusches
 $\bar{N}_{\text{SNR}}(\mu)$ Mittleres Betragsspektrum des Hintergrundgeräusches
 $\widehat{S}_{\text{c}}(\mu, k)$ Geschätztes Betragsspektrum des Sprachsignals
 $\widehat{\text{SNR}}$ Signal-Rauschabstand
 $\widehat{\text{SNR}}_{\text{fw,seg}}$ Mittlerer, logarithmischer Signal-Rauschabstand
 $\widehat{\text{SNR}}_{\text{lin}}$ Geschätzter, linearer Signal-Rauschabstand
 $w(\mu, k)$ Gewichtungsfunktion

Spektrale Distanz

$a^{\text{D}}(\nu, k)$ Gewichte für die Bestimmung der segmentalen spektralen Distanz
 $a_{\text{act}}^{\text{D}}(k)$ Gewichte für die Detektion der aktiven Mel-Bänder
 $a_{\text{P}}^{\text{D}}(\nu, k)$ Gewichte für die Bestimmung des Betragsspektrums $\tilde{S}_{\text{P}}(\nu)$
 $a_{\text{T}}^{\text{D}}(\nu, k)$ Gewichte für die Bestimmung des Betragsspektrums $\tilde{S}_{\text{T}}(\nu)$
 \check{c}_{SD} Vorläufiges Qualitätsergebnis
 D_{spec} Spektrale Distanz
 \bar{D}_{spec} Mittlere spektrale Distanz
 $\bar{D}_{\text{spec}}(k)$ Segmentale spektrale Distanz
 $\hat{H}(\nu)$ Verstärkungsanpassung
 \bar{H}_{log} Mittlere logarithmische Verstärkungsanpassung
 $\hat{H}_{\text{log}}(\nu)$ Logarithmische Verstärkungsanpassung
 K_{SD} Anzahl der verwendeten Rahmen für die Schätzung des Hintergrundgeräusches
 $\bar{N}(\nu)$ Mittlere Leistung des Hintergrundgeräusches
 $\bar{S}_{\text{P}}(\mu, k)$ Geglättetes Betragsspektrum des verzögerten Reaktionssignals
 $S_{\text{P}}(\nu, k)$ Geglättetes Betragsspektrum des Reaktionssignals in Mel-Bänder aufgeteilt
 $\tilde{S}_{\text{P}}(\nu)$ Mittleres Betragsspektrum des Reaktionssignals unter ausschließlicher Berücksichtigung der Rauschpulse
 $S_{\text{R}}(\nu, k)$ Geglättetes Betragsspektrum des verzögerten Referenzsignals inklusive der Verstärkungsanpassung
 $\bar{S}_{\text{T}}(\mu, k)$ Geglättetes Betragsspektrum des verzögerten Referenzsignals
 $S_{\text{T}}(\nu, k)$ Geglättetes Betragsspektrum des verzögerten Referenzsignals in Mel-Bänder aufgeteilt
 $\tilde{S}_{\text{T}}(\nu)$ Mittleres Betragsspektrum des verzögerten Referenzsignals unter ausschließlicher Berücksichtigung der Rauschpulse

$\tilde{x}(n)$	Hochpassgefiltertes Reaktionssignal $x(n)$
$\tilde{y}(n)$	Hochpassgefiltertes Referenzsignal $y(n)$
$\check{y}(n)$	Um Δ verzögertes Referenzsignal $y(n)$
T_{act}	Schwellwert für die Bestimmung aktiver Frequenzbänder
Δ	Geschätzte Verzögerung
ν	Mel-Bandindex
ν_{end}	Mel-Bandindex der letzten analysierten Frequenz
ν_{start}	Mel-Bandindex der ersten analysierten Frequenz

Kapitel 1

Einleitung

Eine Schwierigkeit der ein jeder schon mindestens einmal in seinem Leben gegenüber Stand, ist das Führen eines Gespräches in einem fahrenden Fahrzeug. Besonders die Passagiere der Rücksitzbank haben große Probleme bei höheren Geschwindigkeiten den Fahrer oder den Beifahrer zu verstehen. Diese Beeinträchtigung der Kommunikation hat verschiedene Gründe. Zum einen sind alle Passagier der Windschutzscheibe und nicht den anderen Passagieren zugewandt, wie es in einer herkömmlichen Kommunikationssituation der Fall wäre. Zum anderen erhöht sich mit steigender Geschwindigkeit das Hintergrundgeräusch, welches das Sprachsignal des Sprechers überlagert und die Signalqualität verschlechtert.

Diese erschwerte Kommunikationssituation versuchen die Passagiere durch Veränderung ihrer Position oder durch lauterer Sprechen zu kompensieren, was auf Dauer anstrengend werden kann. Eine neuartige Idee diese Beeinträchtigung zu minimieren, ohne das die Insassen aktiv werden müssen, ist die Verwendung eines Innenraumkommunikationssystems (engl. *In-Car Communication-System*, ICC-Systems). Ein ICC-System verwendet unter Umständen bereits installierte Freisprechermikrofone im Fahrzeug, um das Sprachsignal des Sprechers aufzunehmen und gibt anschließend das aufbereitete und verstärkte Sprachsignal über die Lautsprecher nahe dem Zuhörer wieder aus. Dieses Vorgehen führt zu einer Erhöhung der Nutzsignalleistung am Zuhörerohr, welche quasi den zuvor erlittenen Verlust durch Orientierung der Insassen und der Hintergrundgeräuschleistung, kompensieren soll, sodass der Fahrer von den Zuhörern besser verstanden wird. Solche ICC-Systeme werden bereits seit einigen Jahren erforscht und entwickelt, wobei die Forschung noch nicht vollständig abgeschlossen ist. Insbesondere die Problematik der Rückkopplung ist noch nicht vollends gelöst. Dennoch sind erste stabile System entwickelt und bereits von ersten Automobilherstellern vermarktet worden.

Schon bei der Erforschung und der Entwicklung eines solchen ICC-Systems stellt sich die Frage nach dessen Qualität. Zum einen ist die Evaluierung hilfreich, um gezielt Schwachstellen zu identifizieren und Maßnahmen zur Behebung einzuleiten. Zum anderen können verschiedene ICC-Systeme qualitativ von einander abgegrenzt werden, was besonders für die Automobilindustrie sehr interessant ist. Dabei beschäftigen sich sowohl

die Forschung als auch die Industrie derzeit mit den gleichen Fragestellungen, welche bis heute nicht vollständig beantwortete sind.

Wie gut oder wie schlecht funktioniert ein ICC-System? Welches System funktioniert besser und welches schlechter? Was bedeuten die Begriffe „gut“ oder „schlecht“ im Bezug auf ein ICC-System? Lassen sich bereits entwickelte Verfahren zur Erhebung der Qualität eines ICC-Systems verwenden?

Diese Liste von Fragen lässt sich zweifelsohne beliebig fortführen. Dennoch ist die Kernfrage bei der Evaluierung eines sprachverbessernden bzw. sprachverarbeitenden Systems, die nach der Sprachqualität. Qualität ist jedoch das Resultat eines Bewertungsprozesses durch ein Individuum. Dieser Prozess ist hochgradig individuell, da jede Person unterschiedliche Anforderungen und Erwartung an die Qualität eines Sprachsignals stellt. Jedoch ist das Ziel einer jeden Evaluierung diese auditive Qualität zu bestimmen.

Befragt man nun eine größere Gruppe von Menschen, welche alle das gleiche Sprachsignal dargeboten bekommen, nach der Qualität des Selbigen, kann eine mittlere Qualitätsbewertung erhoben werden. Diese mittlere Qualitätsbewertung dient im Anschluss als eine Art Qualitätsreferenz, welche verwendet werden kann, um instrumentelle Verfahren zu entwickeln. Dabei werden instrumentelle Verfahren zur Analyse von Sprachsignalen entwickelt und die erlangten Ergebnisse mit der mittleren Qualitätsbewertung verglichen. Dieses Vorgehen ist ein Standardverfahren zur Erhebung der Qualität von beispielsweise Freisprechsystemen oder auch *Text-to-Speech*-Systemen (TTS-Systemen). In Bezug auf die Evaluierung eines ICC-Systems sind bis heute noch nicht viele solcher instrumentellen Evaluierungsverfahren publiziert und entwickelt worden.

In der vorgelegten Arbeit wird ein Evaluierungsverfahren vorgestellt, welches aus einer Vielzahl von einzelnen instrumentellen Evaluierungsverfahren besteht, mit deren Hilfe die Gesamtqualität eines ICC-Systems geschätzt wird. Die Grundidee ist es, die Einzelverfahren in Evaluierungsgruppen zu unterteilen, die wiederum unterschiedliche Einflüsse auf die Qualität eines ICC-Systems betrachten. Zunächst wird die Fahrzeugumgebung klassifiziert. Diese Klassifizierung stellt fest wie schwierig die Zielumgebung im Sinne der Signalverarbeitung eines ICC-Systems ist. Unter Verwendung dieses Wissens der Fahrzeugklassifizierung können nun die Ergebnisse der zweiten Evaluierungsgruppe, welche das eigentliche Verhalten des ICC-Systems bewerten soll, entsprechend gewichtet werden. Ein Beispiel: Im Falle einer sehr lauten Fahrzeugumgebung, beispielsweise während der Fahrt in einem Kleinwagen bei einer Geschwindigkeit von 150 km/h, ist das Mikrofon-signal des ICC-Systems stärker verrauscht als beispielsweise während der Fahrt in einer Limousine. Diese höhere Geräuschleistung führt dazu, dass die Qualität des Eingangssignals des ICC-Systems eines Kleinwagens als schlechter einzustufen und damit schwieriger zu verarbeiten ist. Folglich bedarf es einer aufwändigeren Signalverarbeitung innerhalb des ICC-Systems eines Kleinwagens, um eine effektive Geräuschunterdrückung durchzuführen können, als beispielsweise in einer Limousine. Aus diesem Grund sollte ein ICC-System,

welches in einer schwierigen Fahrzeugumgebung agiert und gute Qualitätsergebnis in den Einzelevaluierung erzielt, besser in seiner Gesamtqualität abschneiden, als ein ICC-System mit gleichen Qualitätsergebnissen in einer einfacheren Fahrzeugumgebung. In einer letzten Gruppe wird zusätzlich die Kommunikations- bzw. Sprachqualität in der Fahrgastzelle beurteilt. Die Ergebnisse der Einzelverfahren werden entsprechend gewichtet und kombiniert, um ein Gesamtergebnis der Qualität eines ICC-Systems zu erzielen.

Neben der Entwicklung dieser instrumentellen Verfahren mittels zuvor durchgeführter auditiver Experimente zur Erhebung der Qualitätsreferenz, sind alle Verfahren innerhalb des am Lehrstuhl entwickelten Echtzeit-Rahmenwerks implementiert worden. Dieses Rahmenwerk, welches beliebige Signale, auch Audiosignale, nahezu in Echtzeit verarbeiten kann, ist in C und C++ umgesetzt worden und dient allen wissenschaftlichen Mitarbeitern zur Entwicklung ihrer Algorithmen. Auch die ICC-Systemevaluierung, wie sie in dieser Arbeit vorgestellt wird, ist in diesem Rahmenwerk implementiert worden.

1.1 Gliederung

Kapitel 2 beschreibt zunächst die Gründe der erschwerten Kommunikationssituation innerhalb der Fahrgastzelle eines fahrenden Personenkraftwagens (PKWs). Zusätzlich wird die Funktionsweise eines ICC-Systems anhand eines Basissystems erläutert und einige Schwierigkeiten der Verwendung solcher ICC-Systeme kurz beschrieben.

Nach der Beschreibung des ICC-Systems wird die Frage der Qualität eines solchen Systems beleuchtet. In Kapitel 3 wird zu diesem Zweck zunächst der Begriff der Sprach- bzw. Kommunikationsqualität eingeordnet. Zusätzlich wird eine Betrachtung der Einflussfaktoren auf diese Qualität vorgenommen und im Anschluss eine Evaluierungsstrategie vorgestellt. Diese Strategie unterteilt eine Vielzahl von Einzelverfahren zur Erhebung der Qualität in drei Gruppen.

In Kapitel 4 werden einige grundlegende Verfahren zur Erhebung von Sprachqualität bzw. Kommunikationsqualität vorgestellt. Es folgt eine Unterteilung in auditive, durch Menschen bewertete und instrumentelle, durch Signalanalysen geschätzte Verfahren. Außerdem werden Methoden zur Entwicklung und Überprüfung instrumenteller Evaluierungsverfahren beschrieben.

Um ein ICC-System bewerten zu können, muss eine Umgebung geschaffen werden, die eine zielgerichtete Evaluierung ermöglicht. In Kapitel 5 wird eine solche Evaluierungsumgebung zur Erhebung der Qualität von ICC-Systemen vorgestellt. Zunächst werden einige Anforderungen, welche an das Evaluierungssystem gestellt werden, festgehalten und die resultierenden Konsequenzen diskutiert. Ein weiterer Baustein für eine gezielte Evaluierung, welche in diesem Kapitel erläutert wird, ist eine Umgebungsgeräuschsimulation. Diese hat die Aufgabe, ein ICC-System akustisch so anzuregen, wie es in einem fahrenden Fahrzeug der Fall wäre. Außerdem wird beschrieben, auf welche Weise die entsprechenden Sprachsignale innerhalb der betrachteten Geräuschszenarien erzeugt werden können, um diese

als Referenzsignale zu verwenden. Zusätzlich wird innerhalb dieses Kapitels ein kurzer Überblick über die Implementierung der Evaluierungsverfahren innerhalb eines in C und C++ umgesetzten Echtzeit-Rahmenwerks gegeben. Um die Ergebnisse der einzelnen instrumentellen Verfahren theoretisch einordnen zu können, wird abschließend die gesamte Evaluierungsumgebung mithilfe eines theoretischen Modells abgebildet.

Kapitel 6 stellt die eigenen auditiven Evaluierungsverfahren in Bezug auf die Qualität eines ICC-Systems dar. Hier werden neben einem klassischen Hörtest, welcher die Sprachqualität binauraler Signale eines ICC-Systems bestimmen soll, auch weitere Perzeptionsexperimente vorgestellt. Dabei sind zum einen die Echoperzeption durch die zuhörenden Insassen eines Fahrzeuges und zum anderen die Eigenechoperzeption des Sprechers in einem Fahrzeuge untersucht worden. Auch die Diskrepanz zwischen Qualitätswahrnehmung und Signal-Rauschabstand (engl. *Signal-to-Noise Ratio*, SNR) wird anhand eines Hörtests in diesem Kapitel analysiert. Abschließend werden die ermittelten Ergebnisse im Detail zusammengetragen und diskutiert.

Im Anschluss werden in Kapitel 7 die eigenen instrumentellen Verfahren, welche auch innerhalb des Echtzeit-Rahmenwerks implementiert worden sind, vorgestellt. Dabei werden die Einzelverfahren in den entsprechenden Evaluierungsgruppen aufgeteilt und in Bezug auf die Zielumgebung bzw. Evaluierungsumgebung entworfen. Es werden Verfahren zur Klassifizierung der Fahrzeugumgebung, des ICC-Systems und zur Bestimmung der Sprach- bzw. Kommunikationsqualität vorgestellt. Zusätzlich wird eine Kombination aller gefundenen Qualitätsergebnisse zu einer Gesamtqualität vorgestellt.

Kapitel 8 stellt die Ergebnisse einer konkreten Evaluierung anhand der beschriebenen Evaluierungsverfahren und der Kombination zu einem Gesamtergebnis vor. Dabei werden mithilfe des Echtzeit-Rahmenwerks die Ergebnisse der Evaluierungsverfahren erhoben und des Weiteren zu einem Gesamtergebnis zusammengeführt. Evaluierungsobjekt und auch Evaluierungsumgebung werden in diesem Zusammenhang gleichermaßen von eben jenem Echtzeit-Rahmenwerk bereitgestellt.

Im letzten Kapitel 9 wird eine Zusammenfassung der Arbeit vorgenommen. Zusätzlich wird ein Ausblick auf mögliche zukünftige Arbeiten im Rahmen der instrumentellen und auch auditiven ICC-Systemevaluierung gegeben.

Kapitel 2

Innenraum-Kommunikationssystem

Ein Kleinwagen fährt mit einer Geschwindigkeit von circa 130 km/h auf der Autobahn. Drei Insassen befinden sich im Fahrzeug: ein Fahrer, ein Beifahrer und eine Person auf der Rücksitzbank. Der Fahrer erzählt von seinem letzten Urlaub in Südamerika. Fahrer und Beifahrer diskutieren angeregt über die beste Route und die schönsten Sehenswürdigkeiten. Doch der Passagier auf der Rückbank bekommt nur Bruchstücke dieses Gespräches mit, wobei auch ihn die Details zu Machu Picchu interessieren würden. Um besser am Gespräch teilhaben zu können, lehnt er sich nach vorn und erkundigt sich wiederholt nach dem eben Gesagten. Bedingt durch dieses angestrengte Nachfragen seitens der dritten Person auf der Rückbank, fangen auch Beifahrer und Fahrer an ihr Position zu verändern und sprechen lauter. Diese Situation, wie in Abbildung 2.1 dargestellt, hat sicherlich jeder bereits mehr als einmal in seinem Leben erlebt und ebenso als störend, anstrengend und nicht zuletzt als mögliches Sicherheitsrisiko wahrgenommen.



Abbildung 2.1: Fahrzeuginsassen während der Fahrt, die Probleme bei der Verständigung haben und Sitzhaltung sowie Lautstärke an die Situation anpassen.

Zwei Faktoren verursachen im Wesentlichen diese beeinträchtigte Kommunikationssituation. Zum einen handelt es sich hierbei um das durch die Fahrt auftretende Hintergrundgeräusch. Zum anderen beeinflusst die Anordnung der Passagiere innerhalb der Fahrgastzelle die Kommunikation negativ. Bei einem fahrenden Fahrzeug steigt das Hintergrundgeräusch in Abhängigkeit der Geschwindigkeit und den Eigenschaften des Fahrzeuges, aufgrund von Motor-, Abroll- und/oder Windgeräuschen, stark an. Bei einer Beschleunigung eines Fahrzeuges von 0 auf 150 km/h können Anstiege des Hintergrundgeräusches von bis zu 30 dB gemessen werden [SH06]. Diese Zunahme des Hintergrundgeräusches führt zu einer Erniedrigung des SNR am Ohr des Zuhörers und so zu einer Verschlechterung der Verständlichkeit. Die Anordnung der Kommunikationsteilnehmer innerhalb der Fahrgastzelle schafft eine unnatürliche Kommunikationssituation, welche sich negativ auf die Verständigung auswirkt: In einer gewöhnlichen Kommunikationssituation stehen sich die Teilnehmer gegenüber oder sind zumindest einander zugewandt. Im Fahrzeug sitzen die Passagiere hintereinander und blicken alle in die Richtung der Windschutzscheibe. Diese Anordnung benachteiligt insbesondere die Passagiere auf der hinteren Sitzreihe. Denn das Sprachsignal des Sprechers auf der vorderen Sitzreihe ist, aufgrund der Richtcharakteristik des Kopfes, im hinteren Teil des Fahrzeuges deutlich stärker gedämpft als im Fall eines Sprechers auf der Rücksitzbank. In [SH06] ist der Einfluss der Dämpfung durch den Kopf beschrieben und quantitativ festgehalten worden. Es konnte eine Dämpfung des Sprachsignals zwischen Front und Hinterkopf von über 10 dB in einem Frequenzbereich von 1,4 bis 2 kHz ermittelt werden. Folglich wird durch die Anordnung der Sprecher und Hörer der SNR am Ohr der Zuhörers erniedrigt und die Kommunikation verschlechtert. Insgesamt können zwei Fälle innerhalb der Fahrgastzelle in Bezug auf die Kommunikationsqualität unterschieden werden: einerseits eine Kommunikation, welche von der vorderen Sitzreihe nach hinten gerichtet ist, andererseits eine Kommunikation, welche von hinten nach vorn stattfindet. Dabei ist die Kommunikationsrichtung von vorn nach hinten innerhalb der Fahrgastzelle, vornehmlich aufgrund der Personenkonstellation im Fahrzeug, als die schlechteste einzustufen. Dies trifft jedoch lediglich für Fahrzeuge mit zwei Sitzreihen zu. In Fahrzeugen mit mehreren Reihen, zum Beispiel ein Minivan oder ein Bus, ist auch die Richtung von vorn nach hinten, aufgrund der Distanz, stark beeinträchtigt.

Diese eingeschränkte bis desolante Kommunikationssituation versucht der Mensch, wie in jeder gestörten Kommunikationssituation, zu kompensieren. Die naheliegendste Anpassung, mit dem Ziel der Verbesserung des SNR, ist die Erhöhung der Leistung bzw. das lautere Sprechen des Sprechers. Bei einer Erhöhung des Hintergrundgeräusches beginnen Personen im Allgemeinen ihre Stimme bzw. ihr Sprachsignal an die Situation anzupassen. Diese Leistungsanpassung des Sprachsignales wird als sogenannter Lombard-Effekt bezeichnet, welcher zuerst von Étienne Lombard in 1911 beschrieben worden ist [Lom]. Zusätzlich können die Erhöhung der Grundfrequenz, eine Verringerung der Sprechrate sowie weitere Veränderungen der Sprache innerhalb dieses Effektes beobachtet werden [FS, JA90]. Wie bereits erwähnt, ist die Veränderung der Sitzposition eine weitere Metho-

de zur Kompensation der gestörten Kommunikationssituation. Meist wird diese Veränderung der Sitzposition als unangenehm und anstrengend von allen Insassen wahrgenommen. Wenn nun der Fahrer sich zusätzlich nach hinten dreht und somit nur noch einen eingeschränkten Blick auf die Fahrsituation hat, können solche Situationen augenscheinlich zu einem Sicherheitsrisiko durch die Unaufmerksamkeit des Fahrers führen [STWW12b].

Eine alternative Möglichkeit die Kommunikationssituation innerhalb der Fahrgastzelle zu verbessern, bietet ein sogenanntes ICC-System. Die Grundidee eines ICC-Systems ist es, das vom Sprecher ausgesendete Sprachsignal mithilfe von sprechernahen Mikrofonen aufzunehmen und über die zuhörernahen Lautsprecher innerhalb der Fahrgastzelle wieder auszugeben. Das aufgenommene Mikrofonsignal muss mithilfe von Signalverarbeitung aufbereitet und zusätzlich verstärkt werden. Dadurch wird die Leistung des Sprachsignals des Sprechers am Zuhörerohr erhöht und somit die Kommunikationssituation innerhalb der Fahrgastzelle verbessert. In [OLMG, LF, FP, LSTW14, SH06, LÖS⁺] sind bereits einige erste Umsetzungen und Algorithmen der Signalverarbeitung eines ICC-Systems vorgestellt worden. Durch die Verwendung und innerhalb der Entwicklung eines ICC-Systems entstehen Probleme, welche teilweise noch dem aktuellen Stand der Forschung entsprechen. Als Beispiel sei hier die Problematik der Rückkopplung genannt.

In diesem Kapitel wird die grundsätzliche Funktionsweise eines ICC-Systems erläutert. Dabei wird ein generelles Verständnis der notwendigen Signalverarbeitung angestrebt, jedoch nicht eine Darstellung aller derzeit erforschten und entwickelten Algorithmen eines solchen Systems.

2.1 Funktionsprinzip

Die größten Schwierigkeiten bei der Kommunikation in einem gängigen PKW haben die Insassen der vorderen Sitzreihe, wenn sie sich mit den Passagieren der Rücksitzbank unterhalten wollen. Dieser „schlechteste“ Fall in Bezug auf die Kommunikationsqualität wird zur Generierung eines Basis-ICC-Systems verwendet. In Abbildung 2.2 ist die Fahrgastzelle eines PKWs mit zwei Insassen dargestellt. Der Passagier auf der Fahrerposition repräsentiert den Sprecher, während der Passagier auf der Rückbank als Zuhörer fungiert. Neben den Insassen ist in Abbildung 2.2 ein Lautsprecher und ein Mikrofon des ICC-Systems beispielhaft als einkanaliger Ein- und Ausgang des Systems dargestellt. Diese Konfiguration der benötigten Wandler entspricht selbstverständlich nicht vollständig einer realen Konfiguration, wie sie zum Beispiel in der V-Klasse von Mercedes-Benz¹ erhältlich ist. Das von Mercedes-Benz vertriebene ICC-System besitzt mehrere Mikrofone im vorderen Teil der Fahrgastzelle und verwendet mehr als nur einen Lautsprecher, um das aufbereitete Sprachsignal wieder auszugeben [LÖS⁺]. Auch eine Unterstützung von hinten nach vorne wäre innerhalb einer realen Konfiguration denkbar und kann durch

¹Diese zusätzliche Ausstattung wird von Mercedes-Benz unter dem Namen Sprachverstärkung innerhalb des Burmester[®] *Surround-Soundsystem* vermarktet.

ein zusätzliches System, analog zu dem in Abbildung 2.2 dargestellten, realisiert werden [OLM05]. Zur grundlegenden Beschreibung der Funktionsweise eines ICC-Systems und der zugehörigen Basis-Algorithmen ist es ausreichend, ein einfaches Einweg-System² zu betrachten.

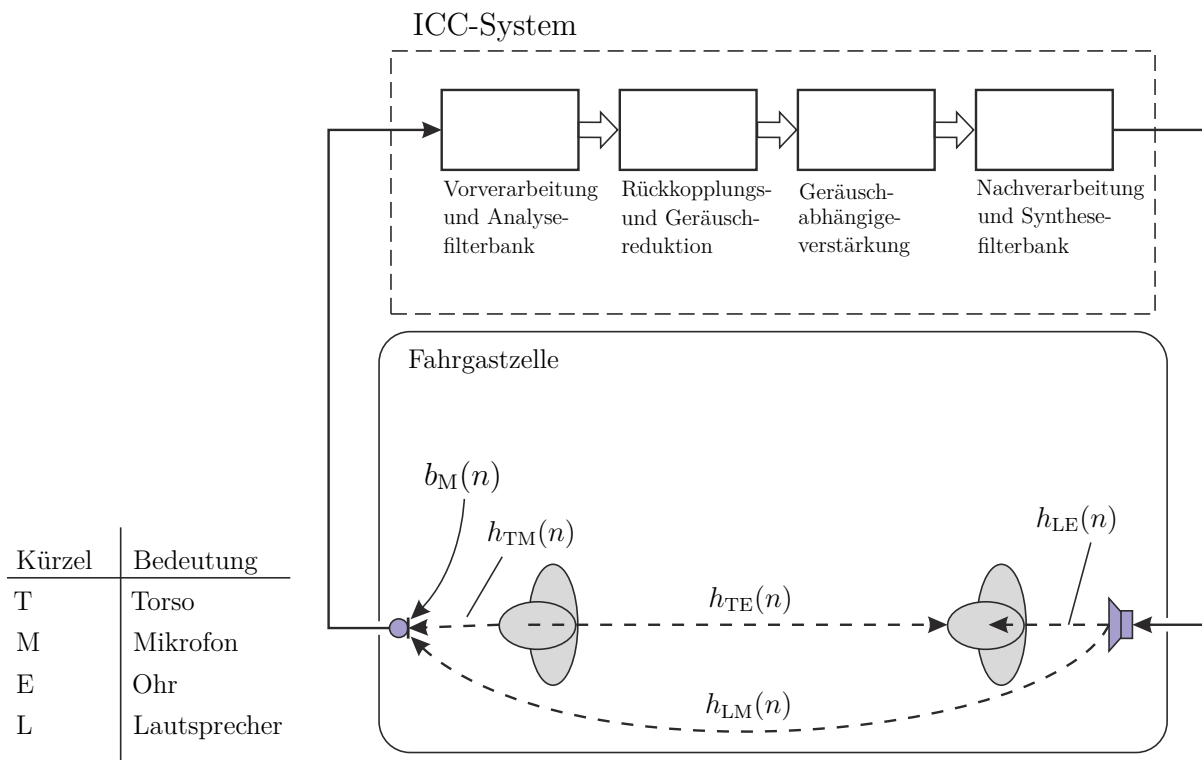


Abbildung 2.2: Basisstruktur eines ICC-Systems inklusive der fundamentalen Übertragungspfade zwischen den elektroakustischen Wandler und den Insassen innerhalb der Fahrgastzelle.

Für die Signalverarbeitung eines ICC-Systems gibt es zwei Hauptprobleme, die es gilt zu minimieren oder aufzuheben [OLM05].

1. Aufgrund der Wiedergabe des Sprachsignals innerhalb der Fahrgastzelle koppelt dieses Nutzsignal wieder in das Mikrofon ein und es entsteht eine elektroakustische Schleife. Dieser akustische Pfad entspricht der in Abbildung 2.2 dargestellten Impulsantwort $h_{LM}(n)$ und bedingt eine Rückkopplung des Nutzsignales. Über diese Schleife schwingt sich das System auf, bis es zu einem Pfeifen oder Heulen kommt und im schlechtesten Fall instabil wird.
2. Neben dem Nutzsignal nehmen die Mikrofone eines ICC-Systems auch das auftretende Hintergrundgeräusch $b(n)$ auf. Dieses Hintergrundgeräusch verringert den SNR

²Bei einem Einweg-System wird nur eine Kommunikationsrichtung, in diesem Fall die Kommunikation von vorn nach hinten, unterstützt. Ein Zweiwege-System würde auch den Zuhörer auf der Rückbank als Sprecher mitberücksichtigen und bräuchte hierfür die entsprechenden Mikrofone und Lautsprecher.

am Mikrofon und verschlechtert die Qualität des Eingangssignals. Falls keine entsprechende Signalverarbeitung zur Bereinigung des gestörten Nutzsignales durchgeführt wird, wird auch das Hintergrundgeräusch wieder über die Lautsprecher ausgegeben und der Geräuschpegel in der Fahrgastzelle erhöht. Die Kommunikationssituation wird nicht wesentlich verbessert.

Diese Problematik und die Anforderung ein verstärktes Nutzsignal auszugeben, gilt es nun mit der Hilfe von Signalverarbeitung zu lösen. In Abbildung 2.2 ist die Signalverarbeitung eines einfachen ICC-Systems dargestellt. Das auftretende Fahrzeuggeräusch $b_M(n)$ ist ein tieffrequenzdominantes Signal, aufgrund dessen wird innerhalb der Vorverarbeitung des ersten Blocks zunächst eine Hochpassfilterung durchgeführt. Dieser Hochpass mit einer 3dB-Grenzfrequenz von 200 Hz unterdrückt somit die Signalanteile, welche stark durch das Hintergrundgeräusch beeinträchtigt worden sind [LÖS⁺]. Um möglichst effizient und mit geringer Komplexität die nachfolgende Signalverarbeitung umsetzen zu können, wird das Eingangssignal in den Frequenzbereich überführt. Zusätzlich muss die Verarbeitung nahezu in Echtzeit durchgeführt werden. Zu diesem Zweck wird eine blockbasierte Analysefilterbank verwendet, die nur eine sehr geringe Verzögerung einfügen darf. Dies ist essentiell, denn bereits Verzögerung in kleinem Umfang werden von dem zuhörenden Passagier deutlich wahrgenommen, da beide Quellen des Nutzsignals, der Sprecher und die Ausgabe des ICC-Systems, sich im gleichen Raum befinden. Die Gesamtverzögerung eines ICC-Systems sollte zwischen 10-20 ms liegen [SH06]. Zusätzlich kann auch der Richtungseindruck durch eine zu große Verzögerung beeinflusst werden [Haa72]. Ebenso kann der Sprecher durch eine zu große Verzögerung zwischen den beiden Quellen im Fahrzeug irritiert oder immens gestört werden. Ansätze wie [WJST, MM] entsprechen diesen Anforderungen und werden für die Signalverarbeitung von ICC-Systemen verwendet.

Nach der Vorverarbeitung im Zeitbereich und der Analysefilterbank, wird zunächst das Mikrofonsignal im Frequenzbereich aufbereitet. Dazu wird in einem ersten Schritt das Geräusch des Mikrofonsignals reduziert [OLMG]. Einer der grundlegendsten Ansätze zur Reduktion des Geräusches ist das sogenannte Wiener-Filter, welches sowohl im Zeit- als auch im Frequenzbereich berechnet werden kann [Ros14]. Zur Bestimmung des Filters muss zunächst eine Schätzung des Umgebungsgeräusches durchgeführt werden. Dies kann mittels verschiedenster Verfahren geschehen. In [BRKS14] sind einige Vorschläge zur Schätzung des Hintergrundgeräusches speziell in einer Fahrzeugumgebung beschrieben. Diese Verfahren können ohne ein weiteres Referenzmikrofon das Geräusch unter Verwendung von rekursiven Algorithmen schätzen. Dazu wird die Leistung der Geräuschschätzung mithilfe von Anstiegs- und Abnahmekonstanten innerhalb eines jeden Rahmens und unter Zuhilfenahme der Kurzzeitleistung des Mikrofonsignals angepasst. Aus diesem Grund brauchen diese Verfahren immer, in Abhängigkeit des Geräuschanstiegs und der gewählten Konstanten, eine gewisse Zeit, um sich auf die Veränderung des Eingangssignals anzupassen. Darüber hinaus ist dies maßgeblich, um nicht versehentlich sich schnell ändernde Sprachanteile als Hintergrundgeräusch zu detektieren und zu reduzieren. Um die Rückkopplung des Nutz-

signals zu verringern, wird in einem nächsten Schritt das Leistungsdichtespektrum der Rückkopplung am Mikrofon geschätzt. Dazu muss zunächst die Übertragung zwischen Lautsprecher und Mikrofon geschätzt werden, zum Beispiel wie in [LÖS⁺] über ein Übertragungsmodell oder wie in [SH06, OLM05] mittels einer Echounterdrückung. Allerdings ist zur letztgenannten Methode anzumerken, dass sie sich, aufgrund der hohen Korrelation zwischen Mikrofon- und Lautsprechersignal [SH06], als komplexes Problem erwiesen hat und sich derzeit noch in der Erforschung befindet. Dennoch können schon erste, vielversprechende Ergebnisse unter Verwendung eines zuvor bestimmten Übertragungsmodells erzielt werden. Denn anhand dieses Wissens kann das Rückkopplungssignal am Mikrofon näherungsweise geschätzt und in Kombination mit der Geräuschschätzung mittels eines Wiener-Filteransatzes reduziert werden [LÖS⁺].

Nachdem das Mikrofonsignal aufbereitet worden ist, kann die geräuschabhängige Verstärkungsregelung (engl. *Noise-Dependent Gain Control*, NDGC) dieses Signal in seiner Leistung anpassen. Die Verstärkung des Sprachsignals wird in Abhängigkeit des auftretenden Hintergrundgeräusches durchgeführt. Das Prinzip dahinter ist relativ plausibel: Ein Fahrzeug im Stand mit laufendem Motor entwickelt ein geringes Hintergrundgeräusch und die Kommunikation der Insassen benötigen daher nur eine geringe Unterstützung durch das ICC-System. Im Fall eines mit 150 km/h fahrenden Fahrzeuges entsteht ein bis zu 30 dB erhöhtes Hintergrundgeräusch, was dazu führt, dass hier eine Verstärkung durch das ICC-System benötigt wird. Deswegen muss auch die NDGC über die zuvor durchgeführte Geräuschschätzung verfügen können, um anhand dieser eine Entscheidung über die nötige Verstärkung des Nutzsymbols durchführen zu können. Da bereits die Geräuschschätzung im Frequenzbereich vorliegt ist es nur folgerichtig, dass die NDGC einer frequenzabhängige Verstärkungskurve folgt, wie sie in [LÖS⁺] vorgestellt worden ist.

Im Anschluss wird das verarbeitete Mikrofonsignal mithilfe einer Synthesefilterbank wieder in den Zeitbereich transformiert. Diese Synthesefilterbank muss für eine perfekte Rekonstruktion des Zeitsignals exakt auf die Analysefilterbank angepasst sein [WJST]. In einem letzten Schritt wird das Zeitsignal im Zeitbereich nachverarbeitet. Dabei wird meist eine Anpassung des Zeitsignals in Bezug auf die verwendeten Lautsprecher und die Akustik im Fahrzeuginnenraum vorgenommen. Zu diesem Zweck wird eine Entzerrfilterstruktur verwendet, die das Signal entsprechend eines verbesserten Höreindrucks für den Zuhörer anpasst. Ebenso können dadurch Einflüsse der verwendeten elektroakustischen Wandler ausgeglichen und entsprechende, fahrzeugspezifische Anpassungen durchgeführt werden. Auch die Verteilung des Lautsprechersignals kann hier vorgenommen werden, denn das aufbereitete Signal soll ausschließlich für den Zuhörer wiedergeben werden. Deswegen wird in einem Szenario mit mehr als einem Lautsprecher ebendieses auf den entsprechenden Lautsprecher verteilt. Bei einem Einweg-System würde man zum Beispiel das Signal über die hinteren Lautsprecher ausgeben, um so nicht den Sprecher zusätzlich mit seinem Eigenecho zu belasten.

Zum Aufbau eines Grundverständnisses der Signalverarbeitung eines ICC-Systems, ist in diesem Kapitel ein entsprechendes Basissystem eingeführt worden. Wie bereits einleitend erwähnt, erhebt dies keineswegs den Anspruch auf Vollständigkeit. Da bis heute die Forschung und die Entwicklung in diesem Bereich nicht vollständig abgeschlossen sind, kann ein ICC-System in zukünftigen oder aktuellen Implementierung in der Anzahl und der Art der verwendeten Komponenten abweichen. Zur Entwicklung von Evaluierungsverfahren eines ICC-Systems ist es zunächst ausreichend, die gängigen Algorithmen und Komponenten, welche in diesem Kapitel beschrieben worden sind, zu betrachten. Folglich, wird das hier vorgestellte Basissystem als Grundlage zum Entwurf der einzelnen Qualitätsschätzverfahren verwendet.

Kapitel 3

Evaluierungsstrategie

Im Anschluss an die Entwicklung und Implementierung eines ICC-Systems stellt sich die Frage nach der Qualität eines solchen Systems. Im Speziellen, welcher Gewinn für die Fahrgäste durch die Verwendung eines ICC-Systems erzielt werden kann. Eine quantitative Kategorisierung der Qualität eines ICC-Systems ist nicht nur sinnvoll um ein einzelnes System zu bewerten und zu verbessern, sondern auch um verschiedene Systeme vergleichen zu können. Zu diesem Zweck werden instrumentelle Evaluierungsverfahren benötigt, welche wiederholbar, aussagekräftig und verlässlich die Qualität eines Systems schätzen können [QBC88]. Doch was bedeutet der Begriff Qualität in Bezug auf die Kommunikationssituation innerhalb einer Fahrgastzelle eigentlich?

Dieser, für die Evaluierung essentielle, Begriff wird zunächst, bezogen auf das im Fokus stehende, in Kapitel 2.1 bereits beschriebene ICC-System, eingeordnet. Zusätzlich wird in Kapitel 3.2 das zugrundeliegende Konzept der Evaluierung erläutert und im Detail beschrieben. In Kapitel 4 werden die Grundlagen auditiver und instrumenteller Verfahren und allgemein der Entwurf neuer Ansätze erläutert.

3.1 Was ist Sprachqualität?

Der Begriff der Qualität wird in allen wissenschaftlichen Bereichen, nicht nur in den Ingenieurwissenschaften, geprägt und verwendet. Der Grund hierfür liegt in der Notwendigkeit, die Güte von Systemen in vergleichbare und reproduzierbare Formulierungen überführen zu können. Denn nur durch die Quantifizierung von Qualität können Systeme miteinander verglichen, eingeordnet und Fehlfunktionen ermittelt werden. Das grundlegende Konzept von Qualität ist, auch wenn in verschiedenen Bereichen unterschiedlich formuliert und benannt, stets das gleiche. In den Wirtschaftswissenschaften wird der Begriff Qualität durch folgende Definition festgelegt:

„Übereinstimmung von Leistungen mit Ansprüchen. Ansprüche stellen Kunden, Verwender (Konsument/Produzent), Händler und Hersteller. Entscheidend

ist, was die Anspruchsteller vor dem Hintergrund ihrer Anforderungen wahrnehmen und für wichtig halten.“ [AADW04, S. 2460]

Es wird deutlich, dass zu Beginn der Qualitätseinschätzungen immer Anforderungen an eine Leistung stehen. Diese werden durch Ansprüche und andere Faktoren zuvor definiert. Die subjektive Qualitätseinschätzung entsteht schließlich, durch den Vergleich der Anforderungen mit den tatsächlich erhalten Leistungen.

Im Vergleich zu dieser wirtschaftswissenschaftlichen Formulierung wird der Begriff der Sprachqualität in [Jek05] wie folgt definiert:

„The result of assessing all the recognized and nameable features and feature values of a speech sample under examination, in terms of its suitability to fulfil the expectations of all the recognized and nameable features and feature values of individual expectations and/or social demands and/or demands.“ [Jek05, S. 6]

Beide Definitionen stellen das Individuum in das Zentrum der Qualitätsbestimmung, da sowohl sprachverarbeitende Systeme als auch Leistungen im Sinne einer wirtschaftlichen Projektierung, für eben diese entwickelt worden sind. Nur das Individuum kann im Anschluss eine Aussage über die Güte der Leistung bzw. des Systems treffen. Eine instrumentelle Evaluierung soll genau diese Qualitätsbeurteilung eines Menschen unter Verwendung von automatisierten, berechneten Methoden vorhersagen. Deswegen muss bei jeder instrumentellen Evaluierung eine Art Qualitätsreferenz bzw. „ground truth“ mithilfe einer Gruppe von Testpersonen für ein System bzw. Leistung gefunden werden. Eine Gruppe von Testpersonen ist notwendig, da Qualität als individueller bzw. subjektiver Vergleichs- und Bewertungsprozess angesehen wird [Möl10, S. 9]. Diese unterschiedlichen Einstufungen der Qualität verschiedener Systeme bzw. Szenarien führen dazu, dass eine einzige Testperson nicht ausreicht, um die Qualitätsreferenz zu bestimmen, sondern eine Vielzahl von Probanden notwendig ist. Anhand dieser Menge von Qualitätsaussagen lässt sich eine mittlere Beurteilungsnote (engl. *Mean Opinion Score*, MOS) bestimmen [ITU96a]. Einige mögliche auditive Evaluierungen für die Bestimmung einer Qualitätsreferenz werden in Kapitel 4.1 beschrieben.

Doch wie bereits in den Definitionen formuliert, handelt es sich bei Qualität um einen subjektiven Begriff, der von verschiedensten Attributen sowie den Erwartungen des Individuums abhängt. Im Detail bedeutet dies, dass die Qualität von ein und demselben Sprachsignal von einer Person in unterschiedlichen Situationen anders bewertet wird. Auch können sich Eindruck und Wahrnehmung einer Person und somit die Qualitätseinschätzung über die Zeit verändern. All diese Adaptionen erschweren die Erhebung einer „Qualitätsreferenz“ durch auditive Verfahren. Sie erfordern besondere Präzision und strenge Vorgaben, wie sie zum Beispiel in [ITU96a, KP95] beschrieben sind, um dennoch akkurate und reproduzierbare Ergebnisse zu erhalten.

Bei der Sprachqualität steht der Zuhörer im Fokus. Da ein ICC-System in erster Linie für ihn, den Zuhörer, entworfen worden ist, ist die Sprachqualität auch weiterhin ein ent-

scheidender Faktor in der Evaluierung. Betrachtet man die Situation innerhalb der Fahrgastzelle unter der Verwendung eines ICC-Systems, entstehen aufgrund der räumlichen Nähe von Sprecher und Zuhörer zusätzliche Qualitätskopplungen. Beispielhaft für eine solche Kopplung sei an dieser Stelle die Verstärkung des Sprachsignals angeführt. Für den Zuhörer ergibt sich ein Qualitätsgewinn, wenn die Leistung des Nutz- bzw. Sprachsignals des Sprechers erhöht wird. Der Sprecher hingegen kann, durch die zu laute Wiedergabe seines eigenen Sprachsignals, irritiert werden, da er dieses als Echo wahrnimmt. Dadurch verschlechtert sich für den Sprecher die Qualität des ICC-Systems bzw. die Kommunikationsqualität. Im Zuge eines auditiven Perzeptionsexperiments wird dieses spezielle Phänomen in Kapitel 6.2 genauer analysiert. Aus diesem Grund sollte bei der Evaluierung eines ICC-Systems zwar die Sprachqualität im Fokus stehen, jedoch die Kommunikationsqualität nicht vollständig ausgeklammert werden, denn die Qualität eines ICC-Systems ist sowohl vom Sprecher als auch vom Zuhörer abhängig.

Ähnlich wie in den beiden zuvor beschriebenen Qualitätsdefinitionen, hängt die Kommunikationsqualität von verschiedensten Attributen und Voraussetzungen ab. In [GK06] sind verschiedenste Faktoren, welche die Sprachqualität für den Zuhörer einer Telekommunikationssituation beeinträchtigen können, festgehalten. In Anlehnung an [GK06] sind in Abbildung 3.1 mögliche Attribute, die Einfluss auf die Kommunikationsqualität einer Fahrzeugsituation haben, beispielhaft dargestellt und in vier unterschiedliche Gruppen eingeteilt.

- Die erste Gruppe repräsentiert die Qualitätsfaktoren, welche im Zusammenhang mit dem Sprecher stehen. Ein Haupteinflussfaktor ist erfahrungsgemäß die Qualität des Sprachsignals selbst. Denn wenn schon das vom Sprecher, egal ob künstlich durch zum Beispiel ein TTS-System oder durch ein Individuum, ausgesendete Sprachsignal stark in seiner Qualität beeinträchtigt ist, ist auch die Kommunikationsqualität beeinträchtigt. Ein weiterer Faktor innerhalb dieser Gruppe ist der Aufwand, den ein Sprecher betreiben muss, um kommunizieren zu können: Muss der Sprecher sehr laut und deutlich reden, damit er verstanden wird oder ist dies nicht notwendig.
- Die zweite Gruppe beinhaltet Faktoren, welche den Zuhörer betreffen. Einflussfaktoren auf die Qualität sind zum Beispiel die Höranstrengung, die es dem Zuhörer erschweren an der Kommunikation teilzunehmen. Aber auch wie viel Aufwand der Zuhörer betreiben muss, um der Konversation zu folgen, die Konversationsanstrengung, fallen in diese Gruppe.
- Die dritte Gruppe fasst klassische Einflussfaktoren auf die Kommunikationsqualität, welche den Kanal betreffen, zusammen. Beispielhaft seien Signalverzerrung, aufgrund von Bandbreitenbegrenzungen, wie sie bei Telefonsprache entstehen, oder auch das auftretende Hintergrundgeräusch angeführt.
- Die vierte Gruppe entspricht den Qualitätsfaktoren, welche im Zusammenhang mit dem Umfeld stehen. Ein Faktor innerhalb dieser Gruppe ist zum Beispiel die Erwar-

tung des Zuhörers bzw. Sprechers an die Kommunikationssituation. Diese Erwartung kann zwischen einzelnen Individuen und auch Situation variieren, so wird bei einem Telefonat eine andere Qualität erwartet als bei einem Gespräch in einem Konferenzraum. In diese Gruppe fallen des Weiteren Einflüsse, welche durch die Erfahrungen der Teilnehmer einer Kommunikation entstehen.

Zweifelsohne sind die Listen der Einflussfaktoren der jeweiligen Gruppe beliebig erweiterbar. Alles in allem wird jedoch die Vielzahl der, die Kommunikationsqualität beeinflussenden, Faktoren und die Komplexität des Qualitätsbegriffes deutlich.

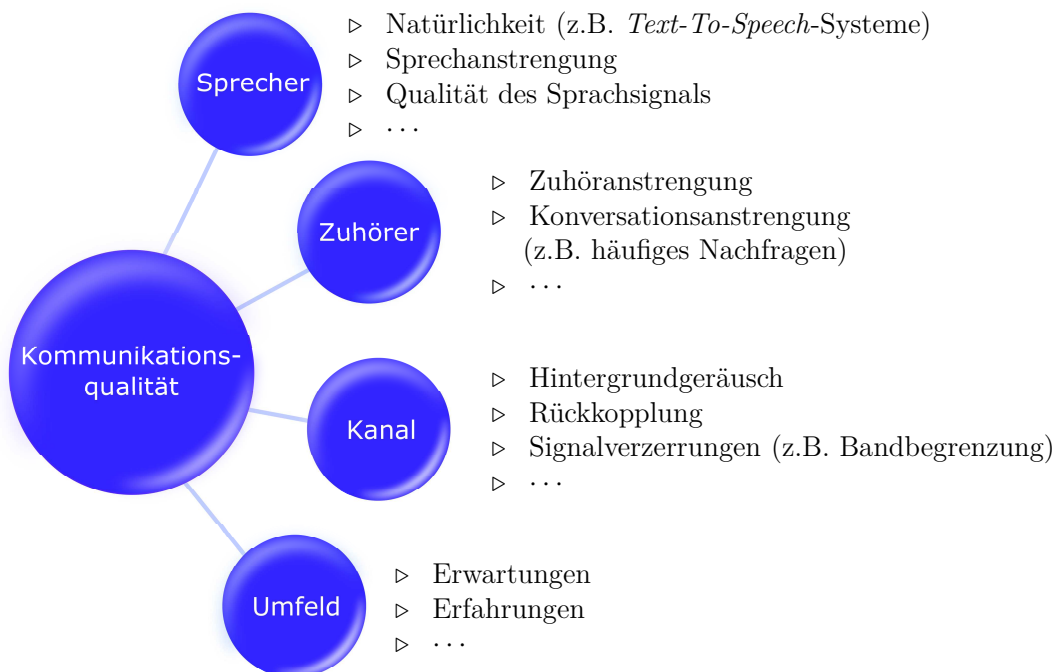


Abbildung 3.1: Überblick und Gruppierung einiger beispielhafter Einflussfaktoren auf die Kommunikationsqualität.

Um diese Vielzahl von Faktoren der Qualitätsbestimmung innerhalb eines komplexen Szenarios, wie die eines ICC-Systems in einem Fahrzeug, abdecken zu können, muss eine gut definierte Evaluierungsstrategie zugrunde liegen. Diese wird anschließend in Kapitel 3.2 im Detail beschrieben.

3.2 Einteilung in Evaluierungsgruppen

Da Qualität ein komplexes, durch verschiedenste Faktoren beeinflusstes, Konzept ist, wird im Folgenden eine differenzierte Evaluierungsstrategie zur Erhebung der ICC-Systemqualität entwickelt. Zu diesem Zweck werden verschiedene instrumentelle Evaluierungsmethoden, welche anhand von Signalanalysen die Qualität eines Systems schätzen, verwendet.

Die einzelnen Verfahren bestimmen dabei unterschiedlichste Parameter bzw. Maße und können in Evaluierungsgruppen zusammengefasst werden.

Die Aufteilung instrumenteller Evaluierungsmethoden in einzelne Evaluierungsgruppen, bietet zwei Vorteile bei der Evaluierung von ICC-Systemen. Zum einen kann über die Kombination mehrerer Verfahren der Einfluss auf die Gesamtqualität gesteuert und komplexe Zusammenhänge abgebildet werden. Zum anderen bietet ein solcher differenzierter Evaluierungsansatz die Möglichkeit einer einfachen Analyse der möglichen Ursachen einer Qualitätsbeeinträchtigung, da gezielt unterschiedlichste Qualitätsparameter evaluiert werden.

Für die Evaluierung eines ICC-Systems werden drei verschiedene Gruppen von instrumentellen Evaluierungsverfahren aufgestellt und im Folgenden definiert:

1. Fahrzeugeigenschaften

Diese Gruppe von Evaluierungsverfahren besitzt nur einen indirekten Einfluss auf die Gesamtqualität eines ICC-Systems. Innerhalb dieser Gruppe werden ausschließlich die Eigenschaften eines Fahrzeuges bestimmt. Dabei ist die Idee mittels verschiedenster Verfahren das Ausmaß der Komplexität der Fahrzeugsituation hinsichtlich eines ICC-Systems zu ermitteln. Zum Beispiel ist in einer Limousine die Leistung des auftretenden Hintergrundgeräuschs auch bei hohen Geschwindigkeiten gering, verglichen mit einem Kleinwagen. Auch die Abstände zwischen den vom ICC-System verwendeten elektroakustischen Wandler¹ sowie deren Frequenzgänge erschweren die Signalverarbeitung innerhalb eines ICC-Systems. Durch die Bestimmung dieser fahrzeugspezifischen Eigenschaften können die Ergebnisse der anderen Evaluierungsgruppen bewertet werden. Denn erwartungsgemäß sollte ein System, welches mit einer geringen Distanz zwischen Wiedergabelautsprecher und Mikrofon keine Rückkopplung erzeugt, bei einer Evaluierung besser abschneiden, als ein System mit einer großen Distanz und auftretender, hörbarer Rückkopplung. Zu diesem Zweck werden innerhalb dieser Gruppe die Eigenschaften des Fahrzeuges erhoben und in Fahrzeugattribute überführt, welche wiederum zur Gewichtung der weiteren Evaluierungsergebnisse verwendet werden. Wie dies im Detail erfolgt, wird in Kapitel 7.1.3 genauer erläutert.

2. Verhalten des ICC-Systems

Im Fokus dieser Evaluierungsgruppe steht das ICC-System mit seinen signalverarbeitenden Modulen. Dabei hat sich die Evaluierung der einzelnen Komponenten eines ICC-Systems, wie in Kapitel 2 beschrieben, als schwierig herausgestellt. Der Grund dafür liegt in der Tatsache, dass verschiedene Komponenten miteinander interagieren. So ist beispielsweise die Geräuschreduktion immer abhängig von der

¹Im weiteren Verlauf der Arbeit werden die Begriffe „Wandler“ und „elektroakustische Wandler“ synonym verwendet.

Qualität der verwendeten Geräuschschätzung. Eine effiziente und leistungsfähige Geräuschreduktion kann keine qualitativ hochwertigen Ergebnisse erzielen, wenn die Geräuschschätzung nicht korrekt durchgeführt worden ist. Eine Evaluierung der Geräuschschätzung oder auch anderer Module kann hingegen nur erfolgen, wenn ein direkter Zugriff auf die einzelnen Komponenten eines ICC-Systems möglich ist. Ist dies nicht der Fall, kann lediglich das Verhalten eines ICC-Systems evaluiert werden. Da innerhalb der verwendeten Evaluierungsumgebung, siehe Kapitel 5, beliebige ICC-Systeme ohne interne Zugriff bewertet werden sollen, werden in dieser Gruppe Verfahren zur Erhebung der ICC-Systemeigenschaften zusammengefasst. Mögliche Eigenschaften, die evaluiert werden können, sind zum Beispiel die Verstärkung des Nutzsignals und die Verzögerung des ICC-Systems.

3. Kommunikationsqualität

Diese letzte Evaluierungsgruppe beinhaltet vornehmlich bereits in anderen Bereichen, wie die Evaluierung von Freisprechsystemen, Telefonie oder Beschallungsanlagen, etablierte Maße zur Bestimmung der Qualität. Im Fokus der hier eingruppierten Evaluierungsverfahren steht die Erhebung der Gesamtqualität innerhalb der Fahrgastzelle. Ein Beispiel für ein Verfahren, welches zu dieser Gruppe gehört, ist die Schätzung des SNR am Zuhörerohr. Im Speziellen soll die Erhöhung des SNR durch die Verwendung eines ICC-Systems ermittelt werden, um so die Kommunikationsqualität des ICC-Systems abschätzen zu können.

Innerhalb all dieser Gruppen werden einzelne instrumentelle Verfahren, welche im Detail in Kapitel 7 beschrieben sind, ausgewertet und im Anschluss über eine Kombination in ein Gesamtqualitätsergebnis des betrachteten ICC-Systems überführt. Diese Gesamtqualität und die Kombination der Einzelergebnisse wird in Kapitel 7.4 beschrieben.

Kapitel 4

Grundlagen Qualitätsbeurteilung

Wie bereits in Kapitel 3.1 beschrieben, hängt die Qualität von dem Individuum ab. Es ist jene Instanz, welche die Qualität einordnen und bestimmen kann. Aus diesem Grund werden immer auditive Verfahren¹ benötigt, um die Qualität einer Zielumgebung im Allgemeinen oder die Sprachqualität im Speziellen definieren zu können. Die Gruppe der auditiven Verfahren zur Qualitätsbestimmung vereint eine Vielzahl von unterschiedlichen Vorgehensweisen. In Abschnitt 4.1 werden einige dieser Verfahren genauer beschrieben und eingeordnet. Meist gilt es aber auditive Verfahren zu vermeiden, da sie zeit- und kostenintensiv sind.

Abhilfe bei diesem Problem sollen instrumentelle Verfahren bringen, welche durch Analysen des verarbeitenden Signals und/oder des Originalsignals, Rückschlüsse auf die Sprachqualität geben. Einige Grundlagen und Klassifizierungen von instrumentellen Verfahren werden in Kapitel 4.2 vorgestellt. Wichtig ist, dass diese instrumentellen Verfahren mittels auditiver Experimente entworfen werden müssen, um so eine Qualitätsvorhersage durchführen zu können. Denn auch wenn instrumentelle Tests die auditiven ablösen sollen, brauchen die instrumentellen Verfahren gleichermaßen eine Qualitätsreferenz, welche als eine Art „Grundwahrheit“ behandelt wird. In Kapitel 4.3 wird beschrieben, wie genau die Verknüpfung zwischen instrumentellen und auditiven Verfahren erfolgt und wie die Güte eines instrumentellen Verfahrens bestimmt werden kann.

4.1 Auditive Evaluierungsverfahren

Qualität setzt sich immer aus dem Zusammenspiel einer Vielzahl von Einflussfaktoren zusammen, siehe Kapitel 3.1. Dabei ist die Kombination dieser Faktoren und das Wissen, um welches es sich tatsächlich handelt, zunächst unbekannt und stellt den Grund für

¹Diese Verfahren werden in der Literatur oft auch als „subjektive“ Verfahren bezeichnet, um so die Variation der Evaluierungsergebnisse aufgrund der individuellen Empfindung darzustellen. Im Gegenzug werden instrumentelle Verfahren in der Literatur auch als „objektive“ Verfahren bezeichnet, um so den Gegensatz darzustellen. Im weiteren Verlauf wird in der vorliegenden Arbeit der Begriff der auditiven bzw. instrumentellen Verfahren weiterverwendet werden.

die Erhebungen von Sprachqualität durch auditive Verfahren dar. Qualität wird immer durch den Vergleich zwischen Wahrnehmung und Erwartung des Individuums bestimmt. Ein Modell, welches diesen Evaluierungsprozess in Bezug auf akustische Signale beschreibt, ist in Abbildung 4.1 dargestellt.

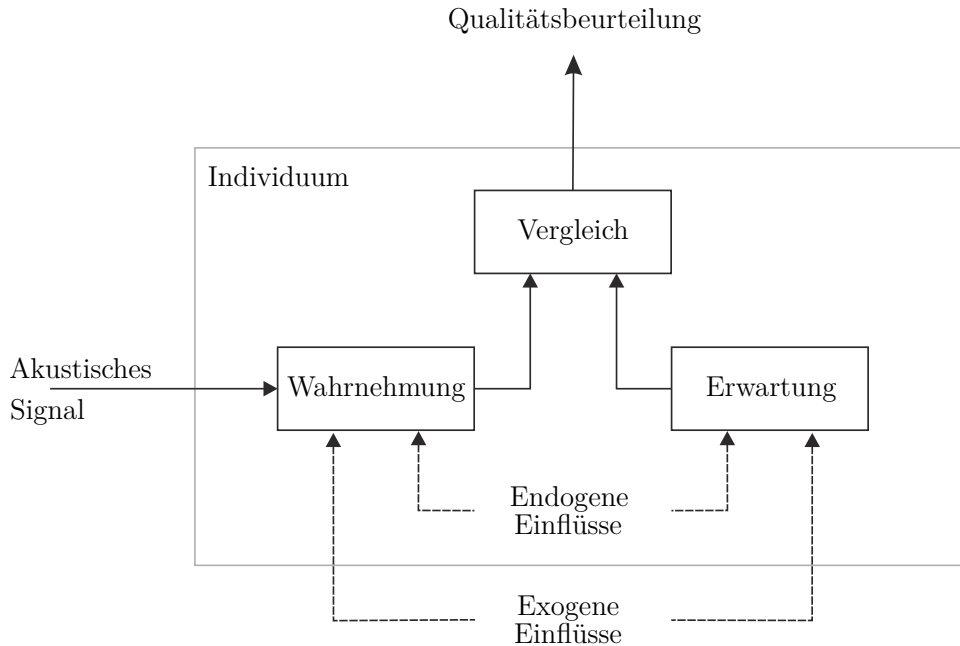


Abbildung 4.1: Modell der auditiven Qualitätsbeurteilungen von Individuen in Anlehnung an [Jek05, S. 52].

Im Mittelpunkt steht das Individuum, welches eine Aussage über die Qualität des dargebotenen akustischen Signals treffen soll. In der Literatur, zum Beispiel in [Bla97, Möl00, Jek05], wird meist nur der Zuhörer als zentrale Evaluierungsinstanz betrachtet, da in diesen Evaluierungen meist die Sprachqualität im Fokus der Evaluierung steht. Hingegen kann das selbe Vergleichsmodell, wie in Abbildung 4.1 dargestellt, ebenso für den Sprecher herangezogen werden, um auf diese Weise den Qualitätseindruck des Sprechers zu modellieren. In beiden Fällen muss das akustische Signal, welches auch als Schaller-eignis² bezeichnet wird, zunächst von dem Individuum wahrgenommen werden. Dabei zählt zu der Wahrnehmung nicht nur das reine Hören über das Sinnesorgan Ohr, sondern auch die kognitive Eigenschaften des Individuums³. Diese Wahrnehmung wird nicht nur durch endogene Einflüsse, wie zum Beispiel bereits erlernte kognitive Fähigkeiten, sondern auch durch exogene Einflüsse, wie durch ein auftretendes Hintergrundgeräusch, verändert. Mithilfe der Wahrnehmung wird ein Hörereignis⁴ erzeugt, welches wiederum mit den Erwartungen des Individuums verglichen wird, um so eine Beurteilung der Qua-

²In der Literatur auch als *sound event* bezeichnet [Bla97, Möl00].

³Für nähere Details siehe [Jek05].

⁴In der Literatur auch als *auditory event* bezeichnet [Bla97, Möl00].

lität durchzuführen. Ebenso können die Erwartungen durch endogene, wie bereits erlebte akustische Situationen, und exogene Einflüsse, beispielsweise durch die Fragestellung des auditiven Tests, beeinflusst werden. Wichtig ist, dass auch hier eine Vielzahl von Einflüssen auf das Individuum wirken. Deswegen müssen auditive Tests und Experimente sehr exakt geplant, durchgeführt und dokumentiert werden. Das vorgestellte Modell gewährt einen generellen Einblick in die Entstehung von Qualität, die es nun gilt durch auditive Verfahren zu erheben. Dabei stellt dieses Kapitel nicht den Anspruch der Vollständigkeit aller möglichen auditiven Verfahren zur Bestimmung von Sprachqualität, sondern dient lediglich dazu einen kurzen Einblick in das Thema und die verwendeten Methoden zu geben.

Im Bereich der auditiven Evaluierung von Telefonie gibt es zwei wichtige Testverfahren, den Konversationstest (engl. *conversation test*) und den reinen Hörtest (engl. *Listening-only tests*). Dabei wird ein Konversationstests dahingehend entworfen, dass die Testperson innerhalb einer Laborumgebung eine reale Telefonsituation mit einem fernem Teilnehmer erlebt und diese bewertet [ITU96a]. Da Konversationstests aufwändig und kostenintensiv sind, wird für die Beurteilung der Sprachübertragungsqualität meist auf reine Hörtests zurückgegriffen. Diese können zusätzlich in zwei Gruppen unterteilt werden, in integrale Verfahren und analytische Verfahren [Sch]. Integrale Verfahren beurteilen anhand von Skalen, beispielsweise der absoluten Kategorisierung (engl. *Absolute Category Rating*, ACR) oder kontinuierlicher Skalen, die Sprachqualität. Dabei haben die kontinuierlichen Skalen den Vorteil, dass Testpersonen ein wesentlich differenziertes Qualitätsergebnis abgeben und somit negative Sättigungseffekte vermieden werden können [Sch]. Eine detaillierte Diskussion dieser Schwachstellen einer ACR-Skala kann in [Möl00] gefunden werden. Die Mittelung dieser Einstufung der Gesamtqualität, über die Menge aller Testpersonen, ergibt die mittlere Beurteilungsnote (engl. *Mean Opinion Score*, MOS). Analytische Verfahren hingegen versuchen die Attribute von Qualität zu ermitteln, um auf diese Weise die Sprachqualität zu definieren. Verfahren wie das diagnostische Akzeptanzmaß (engl. *Diagnostic Acceptability Measure*, DAM) erheben die Gesamtqualität und darüber hinaus die Qualität des Hintergrundgeräusches und des Sprachsignals mittels verschiedener Parameter [QBC88, S. 67].

Neben den bereits in der Telefonie standardisierten und oft verwendeten Verfahren zur auditiven Qualitätserhebung, können auch sogenannte psychoakustische Experimente helfen, die Qualität eines ICC-Systems einzustufen. Dabei liegt der Hauptunterschied dieser Tests in der Aufgabenstellung der Testpersonen, die aktiv das vorgegebene Szenario unter Zuhilfenahme von Schaltern oder Reglern verändern können. Die Resultate dieser Tests entsprechen für gewöhnlich keiner direkten Qualitätseinstufung eines Systems, wie sie in Hörtests bestimmt werden können, sondern Einstellungen von Parametern. Dennoch können diese Einstellungen helfen, die Kommunikationssituation besser zu verstehen. Zudem lassen sie Rückschlüsse auf die Qualität systemspezifischer Eigenschaften zu.

Psychoakustische Perzeptionsexperimente, die genau diese Rückschlüsse auf die Qualität eines ICC-Systems zulassen, werden in Kapitel 6 vorgestellt.

4.2 Instrumentelle Evaluierungsverfahren

Wenn unterschiedliche Systeme, welche ein gleiches Ziel verfolgen, relativ zueinander ausgewertet werden sollen, müssen Verfahren gefunden werden, die wiederholbar, aussagekräftig und zuverlässig sind [QBC88, Seite 2 ff.]. Um dieses Ziel zu erreichen, können entweder die bereits erwähnten auditiven Verfahren eingesetzt werden oder instrumentelle Verfahren, welche durch Signalanalysen eine Schätzung der Qualität bestimmen. Diese Verfahren bieten den Vorteil einer gewissen Wiederholbarkeit und sind zudem auch noch kosten- und zeiteffizient.

Die instrumentellen Verfahren lassen sich nach [Raa06, MR, Sch] anhand von fünf Kriterien kategorisieren:

Anwendungsbereich: Je nachdem in welcher Entwicklungsstufe eines Systems (Planungsphase, Betriebsphase) die Qualität bestimmt werden soll, können unterschiedliche Ansätze zur instrumentellen Evaluierung verfolgt werden.

Qualitätsergebnis: Ähnlich wie bei den auditiven Verfahren, kann bei instrumentellen Verfahren zwischen analytischen und integralen Ergebnissen unterschieden werden. Zusätzlich können auch Ergebnisse wie Verständlichkeit oder Konversationsaspekte erhoben werden.

Betrachtete Systeme: Dieses Kriterium unterscheidet die Struktur des zu evaluierenden Systems. Zum einen existieren instrumentelle Verfahren die Qualität der einzelnen Komponenten eines Systems erheben und zum anderen Verfahren, welche die Qualität des Gesamtsystems feststellen.

Eingangsgrößen: Unterschiedliche Verfahren können unterschiedliche Eingangsgrößen verwenden, um die Qualität eines Systems zu bestimmen. Dabei können sowohl die verarbeiteten und/oder auch die originalen Sprachsignale, als auch andere Kenngrößen des Systems verwendet werden, die einen Einfluss auf die Qualität ebendieses haben.

Modellierung der Hörwahrnehmung: Einige Verfahren verwenden Daten aus empirischen Studien, um die Hörwahrnehmung innerhalb eines instrumentellen Verfahrens zu modellieren. Andere verwenden komplexe Modelle, welche das menschliche Gehör nachbilden sollen, um so die Güte des instrumentellen Verfahrens zu steigern.

In Kapitel 7 werden die instrumentellen Evaluierungsverfahren eines ICC-Systems vorgestellt. Auch diese lassen sich mithilfe der beschriebenen Kriterien einordnen, wobei

einige sich nicht verändern. Zum Beispiel bleibt bei allen Verfahren der Anwendungsbereich der gleiche. Die beschriebenen Verfahren sollen eine Evaluierung nach Abschluss der Entwicklung innerhalb der Betriebsphase vornehmen. Zusätzlich werden allen Verfahren sowohl die Signale nach der Verarbeitung, als auch die Originalsignale bzw. Referenzsignale zur Analyse der Qualität bereitgestellt. Dadurch wird bei der Evaluierung von ICC-Systemen immer ein referenzbasierter Ansatz⁵ verwendet, im Gegensatz zur Evaluierung von beispielsweise TTS-Systemen. Diese können zumeist nur mit einem nicht-referenzbasierten Ansatz evaluiert werden, da in dem Sinne kein Originalsignal vorhanden ist, das als Referenz dienen könnte.

4.2.1 Projektion auf eine auditive Qualitätsskala

Das Ergebnis eines instrumentellen Verfahrens führt normalerweise nicht direkt zu einer Schätzung der Sprachqualität. Dennoch soll das Ergebnis eines instrumentellen Maßes innerhalb des Wertebereiches bzw. der Qualitätsskala eines auditiven Verfahrens liegen. Unter der Annahme, dass eine gewisse Abhängigkeit zwischen auditiven und instrumentellen Qualitätsergebnissen existiert, kann mittels einer Regressionsanalyse diese Abhängigkeit bestimmt und als Vorhersage des Qualitätsergebnisses innerhalb der auditiven Skala verwendet werden. Ein gängiges und zudem eingängiges Modell ist die einfache lineare Regression, welche durch

$$S_i = a + b \cdot O_i + \epsilon_i. \quad (4.1)$$

beschrieben wird. Dabei entspricht O_i dem vorläufigen Ergebnis des instrumentellen Verfahrens, a und b den Geradenparametern, S_i der durch auditive Verfahren ermittelten Qualität und ϵ_i der zufälligen Abweichung der Größe S_i . Die Regressionskoeffizienten lassen sich mithilfe der I_S Wertepaare (S_i, O_i) durch Minimierung der Summe der Abweichungsquadrate

$$\text{RSS} = \sum_{i=1}^{I_S} (S_i - \hat{S}_i)^2 \rightarrow \min \quad (4.2)$$

schätzen. Während \hat{S}_i der geschätzten Qualität unter Verwendung der instrumentellen ermittelten Werte O_i und der Regressionsfunktion

$$\hat{S}_i = \hat{a} + \hat{b} \cdot O_i \quad (4.3)$$

entspricht. Durch Nullsetzen der partiellen Ableitung von (4.2) ergeben sich die Regressionskoeffizienten durch

$$\hat{b} = \frac{\sum_{i=1}^{I_S} (S_i - \bar{S})(O_i - \bar{O})}{\sum_{i=1}^{I_S} (O_i - \bar{O})^2} \quad (4.4)$$

⁵In der Literatur werden Verfahren ohne Referenzsignal auch als *non-intrusive* und Verfahren mit der Möglichkeit das Originalsignal als Referenz zu berücksichtigen als *intrusive* bezeichnet.

und

$$\hat{a} = \bar{S} - b \cdot \bar{O}. \quad (4.5)$$

Hierbei entspricht $\bar{S} = \frac{1}{I_S} \sum_i S_i$ dem arithmetischen Mittel aller auditiven Qualitätsergebnisse über alle I_S Szenarien und $\bar{O} = \frac{1}{I_S} \sum_i O_i$ dem entsprechenden arithmetischen Mittel aller vorläufigen Qualitätsergebnisse des betrachteten instrumentellen Verfahrens [RBW14, S. 177 ff.]. Diese Regressionsanalyse ermöglicht, neben der Projektion auf die gleiche Qualitätsskala, einen weiteren entscheidenden Vorteil. Innerhalb dieses Schrittes kann auch die Güte eines instrumentellen Verfahrens erhöht werden [QBC88, S. 6].

Um ein vollständiges Vorhersagemodell unter Verwendung einer Regressionsanalyse generieren zu können, müssen die erhobenen Daten der auditiven Untersuchung S_i ebenso wie die Daten der instrumentellen Erhebung O_i möglichst den vollständigen Qualitätsbereich bzw. die vollständige Qualitätsskala abdecken. Eine Vorhersage mit einer in (4.1) dargestellten Regressionsgeraden, wird immer einen gewissen Fehler aufweisen. Dieser Schätzfehler ϵ_i muss in jedem Fall bestimmt und optimiert werden, wenn eine solche Herangehensweise für instrumentelle Verfahren gewählt worden ist. Deswegen ist eine Validierung des Modells immer notwendig, um die Verlässlichkeit des instrumentellen Verfahrens bzw. des Modells zu ermitteln und gegebenenfalls zu optimieren.

4.3 Verifikation instrumenteller Verfahren

Bei dem Entwurf neuer instrumenteller Methoden zur Schätzung der Qualität eines sprachverarbeitenden Systems muss abschließend eine Überprüfung der Güte des neuen Maßes erfolgen. Bereits in [QBC88, S. 5] ist ein System, wie es in Abbildung 4.2 dargestellt ist, vorgestellt sowie die Wichtigkeit der Überprüfung von instrumentellen Maßen festgestellt worden.

In Abbildung 4.2 ist ein Modell zur Überprüfung der Güte eines potentiellen instrumentellen Evaluierungsmaßes dargestellt. Zunächst wird das zu testende System mit einem klaren Sprachsignal über einen Lautsprecher⁶ angeregt, welches wiederum das verarbeitete Sprachsignal über einen zweiten Lautsprecher ausgibt. Mithilfe eines Mikrofons kann dieses prozessierte Sprachsignal als Eingangssignal, sowohl dem zu überprüfenden instrumentellen Verfahren, als auch dem auditiven Verfahren zur Verfügung gestellt werden. Die Erhebung der Qualitätseinstufung durch auditive Verfahren muss nur einmal durchgeführt werden und kann im weiteren Verlauf wiederholt mit anderen instrumentellen Verfahren verglichen werden [QBC88, S. 6]. Die instrumentellen Verfahren können je nach Eingangsgrößen-Klassifizierung (siehe Kapitel 4.2) des Verfahrens, sowohl auf das klare als

⁶Bei der Evaluierung, beispielsweise von Sprachkodierungsverfahren, kann hier ein direkter Zugriff auf das klare Sprachsignal erfolgen, ohne den Umweg einer Wiedergabe über einen Lautsprecher zu nehmen. Diese Arbeit fokussiert die Evaluierung von beliebigen ICC-Systemen, weswegen eine Einspeisung des klaren Sprachsignals auf das zu testende System durch einen Lautsprecher und eine Mikrophon angenommen wird.

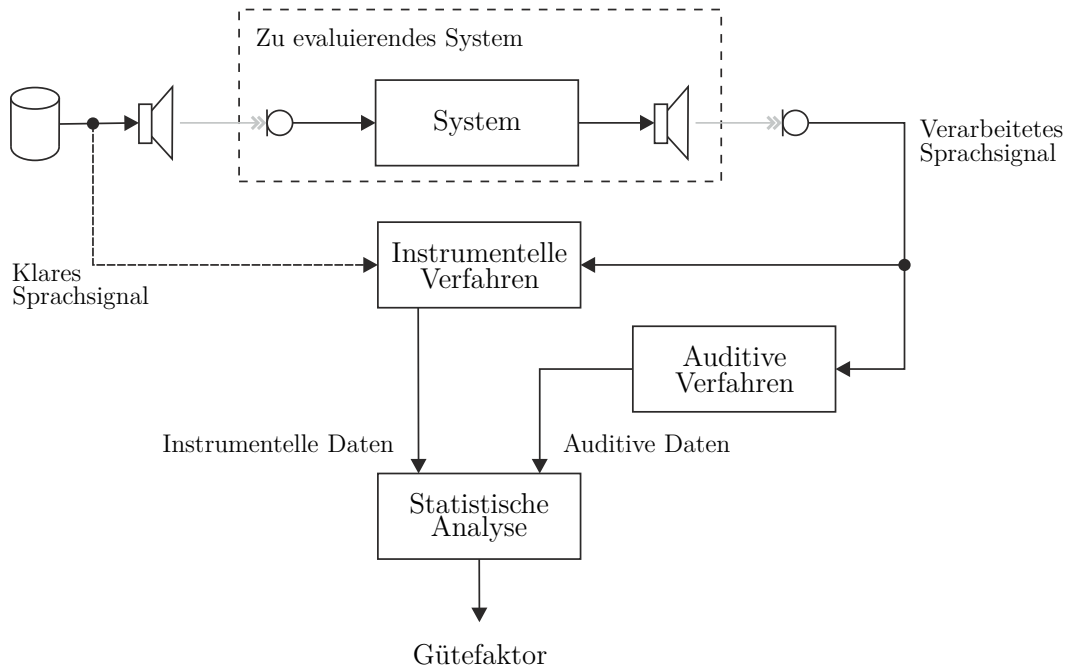


Abbildung 4.2: System zur Überprüfung der Güte von instrumentellen Evaluierungsverfahren in Anlehnung an [QBC88, S. 14].

auch auf das verarbeitete Sprachsignal zur Analyse zurückgreifen oder nur auf das verarbeitete Sprachsignal. Im Anschluss werden beide Ergebnisse, der instrumentellen und der auditiven Evaluierung, durch eine statistische Analyse verglichen und die Güte des Verfahrens berechnet. Hierzu werden verschiedene Szenarien und unterschiedliche Systeme, von einer Vielzahl von Testpersonen evaluiert.

In Kapitel 4.2.1 ist bereits der Entwurf eines instrumentellen Evaluierungsverfahrens mittels einer Regressionsanalyse beschrieben. Die Validierung des entworfenen instrumentellen Verfahrens erfolgt unter Verwendung der in Abbildung 4.2 dargestellten statistischen Analyse. Diese ermittelt die Güte des Verfahrens. Gebräuchlich in diesem Zusammenhang ist es die Korrelation zwischen den instrumentellen und den auditiven Daten zu bestimmen. Der empirische Korrelationskoeffizient⁷ wird durch

$$\rho = \frac{\sum_{i=1}^{I_S} (\hat{S}_i - \hat{\bar{S}})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^{I_S} (\hat{S}_i - \hat{\bar{S}})^2 \sum_{i=1}^{I_S} (S_i - \bar{S})^2}} \quad (4.6)$$

bestimmt [LW06, S. 17 ff.]. Es entspricht $\hat{\bar{S}} = \frac{1}{I_S} \sum_i \hat{S}_i$ dem arithmetischen Mittel der Qualitätsschätzung bestimmt durch (4.3). Ein weitere Metrik der Güte eines Modells ist

⁷In der Literatur, zum Beispiel in [Möl00, S. 152], wird dieser auch als Pearson-Korrelationskoeffizient bezeichnet.

der mittlerer quadratischer Fehler (engl. *Root Mean Square Error*, RMSE), er ist definiert durch:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{I_S} (\hat{S}_i - S_i)^2}{I_S}}. \quad (4.7)$$

Beim Entwurf und bei der Validierung eines neuen instrumentellen Maßes sollten die Daten in zwei Datensätze unterteilt werden, in Trainings- und Validierungsdaten. Diese Aufteilung wird in der Literatur auch als allgemeines Kreuzvalidierungsverfahren bezeichnet. Durch diese Vorgehensweise kann erreicht werden, dass die Bewertung des neuen Verfahrens unabhängig vom Entwurf durchgeführt wird und so eine faire Aussage über die Güte der Qualitätsvorhersage gegeben werden kann. Um statisch ähnliche Datensätze zu erzielen, muss die Gesamtdatenbasis bereits zu Beginn relativ groß sein und die Aufteilung in Trainings- und Validierungsdaten zufällig geschehen [QBC88, S. 181].

Ein weiteres Verfahren zur Validierung von instrumentellen Verfahren ist die *leave-one-out*-Kreuzvalidierung, welche auch als *jackknife*-Verfahren bezeichnet wird [QBC88, S. 181 ff.]. Dazu wird die Datenbasis ebenfalls in eine Trainingsdatenbasis und eine Validierungsdatenbasis aufgeteilt, wobei in diesem speziellen Fall nur ein Datensatz innerhalb der Validierungsdatenbasis verwendet wird. Anhand aller anderen Daten wird ein Vorhersagemodell, wie es in Kapitel 4.2.1 beschrieben worden ist, entworfen und zur Schätzung des Validierungsdatensatzes verwendet. Dies wird für jeden der I_S Datensätze durchgeführt, wodurch I_S Modelle und deren entsprechenden Qualitätsvorhersagen ermittelt werden. Im Anschluss kann mithilfe des empirischen Korrelationskoeffizienten (4.6) die Güte des Modells bestimmt werden. Hierbei ist anzumerken, dass der Gütefaktor unter Verwendung des *jackknife*-Verfahrens generell höher ist als bei dem allgemeinen Kreuzvalidierungsverfahren. Dennoch kann dieses Verfahren verwendet werden, um eine erste Einschätzung der Güte des Modells zu erhalten [QBC88, S. 181].

Um die Güte bzw. ein Regressionsmodell auswählen zu können, können ebenfalls die Regressionskoeffizienten analysiert werden. Dazu wird das Regressionsmodell, wie es in 4.2.1 beschrieben worden ist, unter Verwendung aller Datensätze entworfen sowie ein Konfidenzintervall für die Regressionskoeffizienten mit einem Konfidenzniveau von 95% entworfen. Wenn nun alle Koeffizienten der durch das *jackknife*-Verfahren bestimmten Modelle, innerhalb dieses Konfidenzintervalls liegen, kann davon ausgegangen werden, dass der Unterschied zwischen allen Modellen statistisch vernachlässigbar ist [QBC88, S. 181 ff.]. Als Konsequenz kann als repräsentatives Regressionsmodell jenes ausgewählt werden, welches im Training alle Datensätze berücksichtigt.

In Abbildung 4.3 sind beispielhaft die Koeffizienten \hat{a} und \hat{b} für alle erhobenen Datensätze und auf der Datenbasis ohne den aktuellen Datensatz dargestellt. Zusätzlich ist das 95%-Konfidenzintervall beider Koeffizienten mit aufgetragen. Dabei entstammen die auditiven Daten dem in Kapitel beschriebenen auditiven Verfahrens und die instrumentellen Daten sind mit Hilfe des modifizierten Sprachübertragungsindex (engl. *Speech Transmission Index*, STI) aus Kapitel 7.3.2 bestimmt worden. Zu sehen ist, dass die einzelnen Ko-

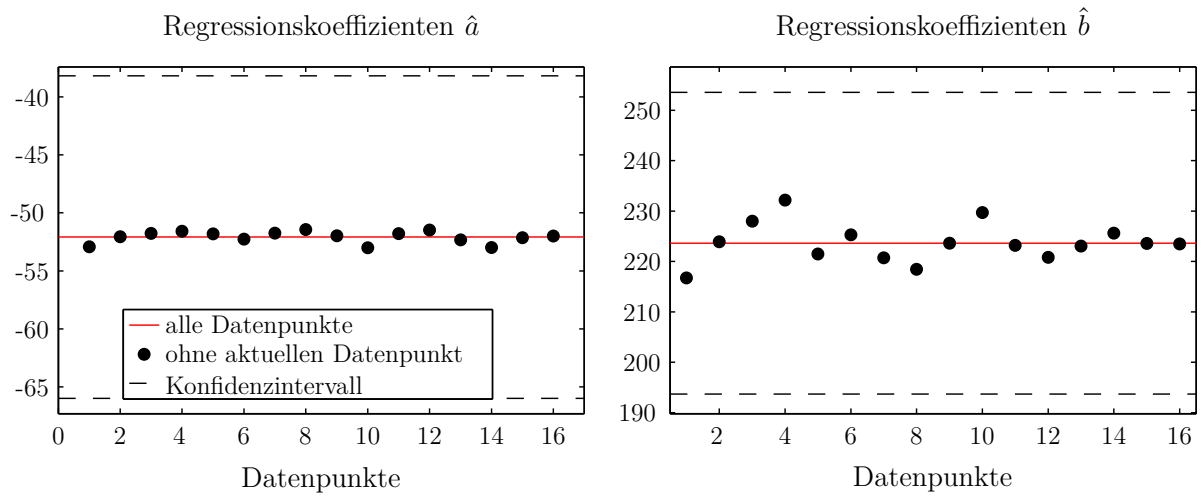


Abbildung 4.3: Regressionskoeffizienten \hat{a} und \hat{b} unter Verwendung aller erhobenen Datensätze (schwarze Linie) und mit dem *jackknife*-Verfahren (schwarze Kreise) berechnet. Zusätzlich ist das zugehörige 95%-Konfidenzintervall (grau gestrichelt) abgebildet.

effizienten des Regressionsmodells unter Verwendung des *jackknife*-Verfahren innerhalb des 95%-Konfidenzintervall liegen und somit die Unterschiede zwischen den einzelnen Modellen vernachlässigbar ist. Als Konsequenz kann als repräsentatives Regressionsmodell jenes ausgewählt werden, welches im Training alle Datensätze berücksichtigt.

Kapitel 5

Evaluierungsumgebung

In Kapitel 3.2 ist eine Strategie zur instrumentellen Ermittlung der Qualität von ICC-Systemen vorgestellt und im Detail erläutert worden. In einem nächsten Schritt wird diese abstrakte Evaluierungsstrategie mithilfe von instrumentellen Einzelverfahren innerhalb einer Evaluierungsumgebung umgesetzt. Neben den Anforderungen, welche sich aus der Evaluierungsstrategie ergeben, werden zusätzlich noch allgemeine Anforderungen an das Evaluierungssystem gestellt. Die Evaluierung eines ICC-Systems innerhalb der hier vorgestellten Evaluierungsumgebung, soll

- automatisch,
- reproduzierbar und
- mit einem beliebigen ICC-System

durchgeführt werden können. All diese Anforderungen an das Evaluierungssystem erfordern eine spezielle Umsetzung der Umgebung und der instrumentellen Verfahren zur Berechnung der Qualität eines ICC-Systems. Dabei liegt der Fokus in diesem Kapitel auf der Umsetzung und dem Entwurf der Evaluierungsumgebung mit ihren einzelnen Komponenten.

In Abbildung 5.1 ist eine Basisstruktur der Evaluierungsumgebung zur Bewertung eines ICC-Systems grafisch dargestellt. Diese Darstellung erhebt keinen Anspruch auf Vollständigkeit, sie soll lediglich einen Überblick über die Funktionsweise der Evaluierung eines ICC-Systems geben. Ein vollständiges ICC-System sowie Evaluierungssystem würde selbstverständlich mehr als die in Abbildung 5.1 beispielhaft dargestellten elektroakustischen Wandler benötigen. Dennoch ist dieser reduzierte Darstellungsansatz ausreichend, um die Funktionsweise der Evaluierungsumgebung entsprechend zu beschreiben.

Um zunächst ein beliebiges ICC-System evaluieren zu können, muss dieses als ein unabhängiges System in die Evaluierungsumgebung eingebracht werden können. Ein ICC-System wird immer in Kombination mit einem Fahrzeug zur Evaluierung bereitstehen, da

Einige Bestandteile der in diesem Kapitel beschriebenen Evaluierungsumgebung sind bereits in ersten Ansätzen in [TSWL] vorgestellt worden.

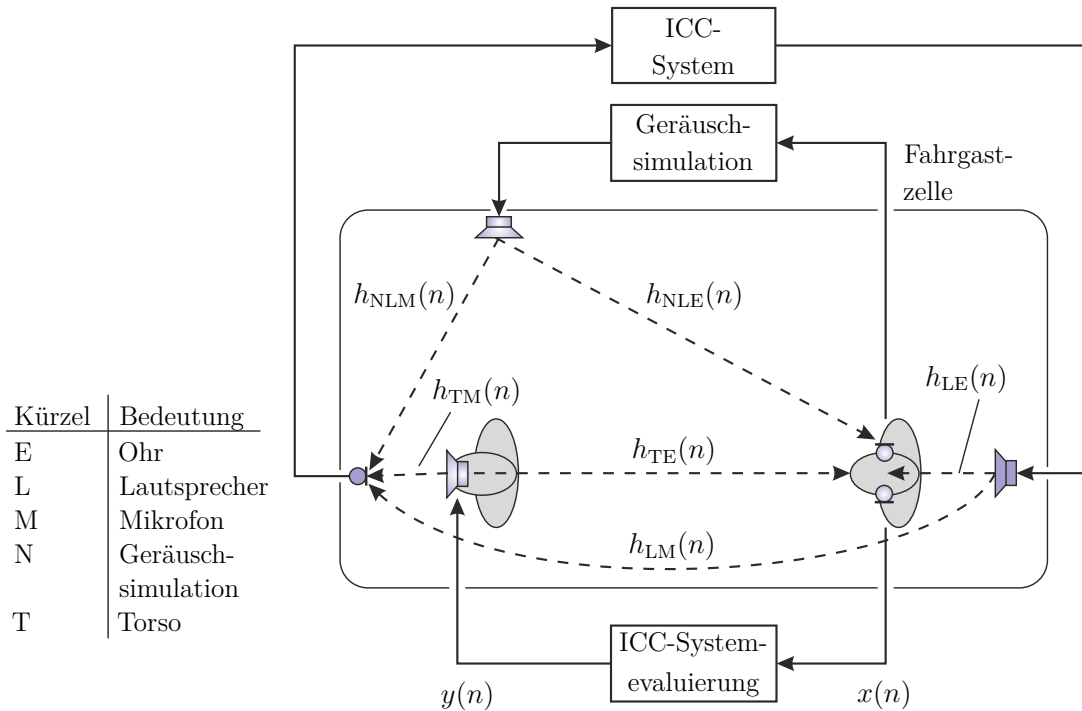


Abbildung 5.1: Basisstruktur der ICC-Systemevaluierung und der Evaluierungsumgebung innerhalb der Fahrgastzelle, mit den zugehörigen Übertragungspfaden (nach [TSWL]).

solche Systeme auf diese eine Umgebung speziell abgestimmt worden sind. Deswegen ist in Abbildung 5.1 neben dem ICC-System auch die Fahrgastzelle mit dem entsprechenden Ein- und Ausgang des ICC-Systems dargestellt. Um nun ein beliebiges ICC-System evaluieren zu können, muss davon ausgegangen werden, dass kein direkter Zugriff auf die Signale und andere Komponenten eines ICC-Systems besteht. Im Detail bedeutet dies, dass die elektrischen ICC-Eingangs- und Ausgangssignale, interne Signale sowie interne Systemparameter der eigentlichen Evaluierung nicht zu Verfügung stehen. Das ICC-System muss als eine *black box* betrachtet werden, welche nur akustisch angeregt und die Reaktion akustisch überwacht werden kann. Das Konzept der Qualitätsprüfung einer *black box* wird als *black-box testing* bezeichnet. Die Grundidee ist, ein Testverfahren zu entwerfen, welches nicht das interne Verhalten sondern die Struktur eines Programms bzw. eines Systems evaluiert [MSB12]. Um das ICC-System durch ein *black-box testing* zu evaluieren, müssen zum einen Lautsprecher zur Anregung des Systems im Fahrzeug platziert werden und zum anderen Mikrofone zur Überwachung der Reaktion des Systems. Diese sind beispielhaft in Abbildung 5.1 in Form eines Lautsprecher und eines Mikrofon als Ein- und Ausgang des Evaluierungssystems dargestellt. Aufgrund dieser Anordnung, müssen nicht nur die Verfahren zur Qualitätsmessung (siehe Kapitel 7), sondern auch die Testumgebung entsprechend entworfen werden.

In Kapitel 2.1 ist bereits beschrieben worden, dass ein ICC-System sich auf die ändernde Fahrzeugumgebung anpasst. Zum Beispiel verändert sich die Verstärkung des Nutzsignals anhand der Variation des Hintergrundgeräusches. Um nun trotzdem innerhalb einer Laborumgebung und unter Einhaltung der gegebenen Anforderungen, beispielsweise die Verstärkung des Nutzsignales ermitteln zu können, muss das ICC-System akustisch in den gewünschten Betriebszustand überführt werden. Dabei muss die akustische Anregung in gleicher Weise wie bei einer realen Fahrt mit dem Fahrzeug erfolgen. Da eine reale Fahrt nicht vollständig reproduzierbar und so die Vergleichbarkeit der Ergebnisse nicht gegeben ist, wird eine Geräuschsimulation wie sie in Abschnitt 5.1 beschrieben ist, verwendet. In Abbildung 5.1 ist eine solche Simulation mithilfe von Lautsprechern beispielhaft dargestellt. Dabei wird nicht nur der Eingang des ICC-Systems mit dem entsprechenden Geräusch angeregt, sondern auch der Eingang des Evaluierungssystems, welcher beim Entwurf der Einzelverfahren berücksichtigt werden muss. Die Geräuschsimulation ermöglicht es, nicht nur ein beliebiges ICC-System in den gewünschten Betriebszustand zu überführen, sondern auch eine reproduzierbare Umgebung zu erzeugen und so verschiedenste Systeme bei unterschiedlichen Umgebungsgeräuschen miteinander zu vergleichen.

Nicht nur das ICC-System passt sich an das Geräuschszenario innerhalb der Fahrgastzelle an, sondern auch die Passagiere. Dieses Phänomen der Veränderung des eigenen Sprachsignals, wird im Allgemeinen als Lombard-Effekt bezeichnet. Um die Situation sich unterhaltender Passagiere während einer Fahrt vollständig reproduzieren zu können, muss neben der Geräuschsimulation auch das entsprechende Sprachsignal innerhalb der Fahrgastzelle reproduziert werden. Zu diesem Zweck wird ein künstlicher Sprecher, wie in Abbildung 5.1 dargestellt, innerhalb der Fahrgastzelle platziert, welcher ein Testsignal mit den entsprechenden, zum Hintergrundgeräusch passenden Sprachpassagen wiedergibt. Dabei gilt es in einem ersten Schritt diese Sprachsignale bei unterschiedlichen Hintergrundgeräuschen aufzuzeichnen, um sie im Anschluss als Testsignal verwenden zu können. Hierzu ist eine Lombardsprachdatenbank zusammengestellt worden, welche in Kapitel 5.2 genauer beschrieben ist.

Neben der Reproduzierbarkeit der Evaluierung und der Möglichkeit ein beliebiges System bewerten zu können, ist die Anforderung einer automatischen Durchführung der Qualitätsbewertung eines ICC-Systems gestellt worden. Die Idee dabei ist es, dass die einzelnen Verfahren möglichst unabhängig von jeglicher Interaktion durch eine Person das Qualitätsergebnis eines ICC-Systems bestimmen. Hierfür werden alle Verfahren innerhalb eines Echtzeit-Rahmenwerks (*Kiel Real-time Application Toolkit* (KiRAT)), welches am Lehrstuhl Digitale Signalverarbeitung und Systemtheorie (DSS) entwickelt worden ist, umgesetzt. Innerhalb des in C und C++ realisierten Rahmenwerkes, werden Audio- und viele weitere Schnittstellen (zum Beispiel Sonar- oder EEG-Signale) zur Verfügung gestellt, welche in Echtzeit blockweise verarbeitet werden können. Dieses Rahmenwerk kann entsprechend auf eine beliebige Plattform transferiert werden und somit auch innerhalb eines Fahrzeuges verwendet werden. Weitere Details zur Umsetzung des Evaluierungssys-

tems innerhalb dieses Rahmenwerkes, werden in Kapitel 5.3 gegeben. Eine Besonderheit von KiRAT ist die gemeinschaftliche Arbeit und Weiterentwicklung dieses Rahmenwerks. So kann jeder wissenschaftliche Mitarbeiter des Lehrstuhls DSS seine eigenen Arbeiten dort implementieren und zusätzlich gängige Algorithmen den anderen Wissenschaftlern zugänglich machen. Dadurch können Entwicklungszeiten reduziert und die Identifikation von Fehlern optimiert werden.

5.1 Umgebungsgeräuschsimulation

Wie zuvor erwähnt, soll ein beliebiges ICC-System unter Verwendung der in Abbildung 5.1 dargestellten Evaluierungsumgebung ausgewertet werden. Zu diesem Zweck muss der Innenraum des Fahrzeuges, entsprechend dem gewünschten Szenario akustisch angeregt werden, wobei sich das Szenario in Abhängigkeit der Geschwindigkeit des Fahrzeug und der Umgebungssituationen verändern kann. Die nächstliegende Möglichkeit dies zu bewerkstelligen, ist die Durchführung von Testfahrten mit dem Fahrzeug, in welchem das ICC-System installiert ist. Diese Methode birgt den Nachteil, dass unterschiedliche Szenarien, zum Beispiel eine Stadtfahrt bei 50 km/h, nicht exakt reproduziert werden können und so die Vergleichbarkeit unterschiedlicher ICC-Systeme nicht gegeben ist. Des Weiteren sind solche Testfahrten kostenintensiv, da immer ein Fahrer und die entsprechenden Anfahrtszeiten zum gewünschten gleichen Straßenabschnitt benötigt werden.

Um vergleichbare und reproduzierbare Evaluierungsergebnisse erzielen zu können, müssen die Fahrscenarien, in einer Nicht-Fahrssituation, akustisch reproduzierbar sein. Hierzu werden akustische Geräuschsimulationen benötigt, die durchgängig alle möglichen Geräuschszenarien einer Testfahrt nahezu exakt wiedergeben können. Dabei wird zusätzlich der Anspruch gestellt, dass beliebige Fahrzeuge verwendet und verschiedene Geschwindigkeiten bzw. Fahrscenarien entsprechend angeregt werden können.

In dieser Beziehung, nicht nur zwecks einer Evaluierung von ICC-Systemen sondern auch von Freisprechsystemen, sind bereits erste Simulationssysteme entwickelt und veröffentlicht worden. Die in [KHG00] beschriebene Simulation von Fahrzeuggeräuschen wird bereits von der Firma HEADacoustics kommerziell zur Evaluierung von Freisprechsystemen vertrieben. Überdies werden Verfahren wie die Wellenfeldsynthese zur Generierung von beliebigen akustischen Umgebungen bereits seit einiger Zeit innerhalb von Laborumgebungen eingesetzt [Baa]. Im Innenraum eines Fahrzeuges sind solche Verfahren eher schwierig umzusetzen, da bei einer Wellenfeldsynthese eine Vielzahl von einzelnen Lautsprechern mit geringen Abstand platziert werden müssen, um ein optimales Ergebnis zu erzielen. All diese Anforderungen sind in einem beliebigen Fahrzeug schwierig umzusetzen und sind zudem aufwendig und teuer. Jedoch lassen sich auch mit wenigen Lautsprechern sinnvolle und vielversprechende Richtungs- und Distanzwahrnehmungen erzielen, obwohl das wirkliche Schallfeld nicht optimal wiederhergestellt werden kann [GK06].

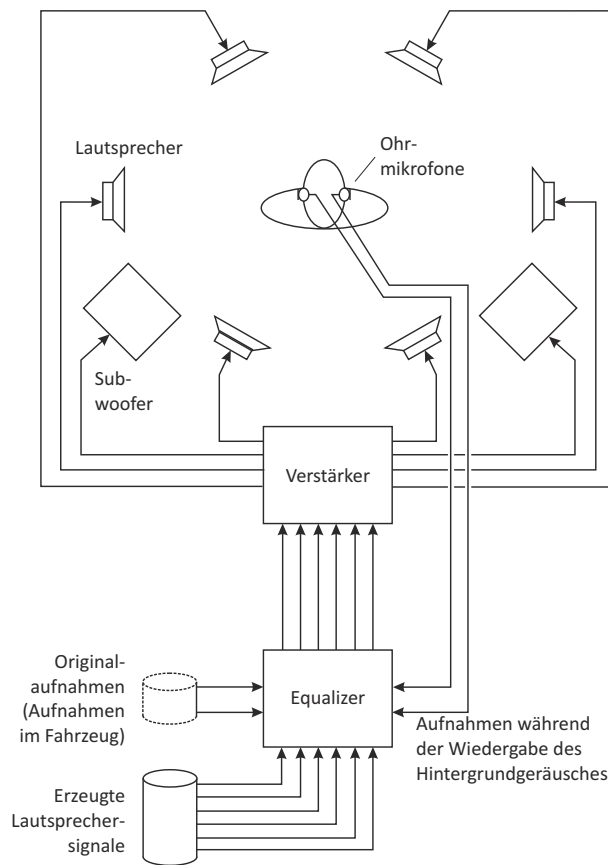


Abbildung 5.2: Schematischer Aufbau einer beliebigen Umgebungsgeräuschsimulation und deren Kalibrierung.

Innerhalb des Lehrstuhls DSS der Christian-Albrechts-Universität zu Kiel ist eine Geräuschsimulation entwickelt worden. In [LTS⁺] ist diese Simulation zum Beispiel verwendet worden, um eine Lombard-Sprachdatenbank zu erstellen. Diese Simulation zielt darauf ab, ein definiertes Ziel-Geräuschszenario innerhalb eines beliebigen Raumes, zum Beispiel eines Fahrzeuges oder eines reflexionsarmen Raumes, zu einem beliebigen Zeitpunkt mit einer geringen Anzahl von Lautsprechern wiederherstellen zu können. Die Simulation ist in zwei Phasen aufgeteilt: In der ersten Phase wird das Ziel-Geräuschszenario festgehalten; in der zweiten Phase diese Geräuschkulisse reproduziert.

Das Ziel-Geräuschszenario wird mittels einer beliebigen Anzahl von Mikrofonen aufgenommen. Im Falle der Simulation von verschiedenen Fahrzeuggeräuschen, werden diese Mikrofone in der Nähe der Ohren der Fahrzeuginsassen platziert. Um eine geeignete Referenz für die Geräuschsimulation zu erzielen, müssen dazu kalibrierte Mikrofone verwendet werden.

Im Anschluss wird die Geräuschsimulation kalibriert. Dazu werden zunächst unkorrelierte Lautsprechersignale benötigt. Diese bestehen aus einem weißen Rauschen, welches entsprechend für jeden Lautsprecher erzeugt wird. Anschließend wird den weißen Rausch-

signalen das Leistungsdichtespektrum der zuvor aufgenommenen Ohrsignale aufgeprägt, um so bereits ein ähnliches spektrales Verhalten der Lautsprechersignale zu erzeugen. Während der Wiedergabe dieser Signale über die, in der Laborumgebung platzierten, Lautsprecher, werden die Kurzzeitleistungsdichtespektren der Referenzmikrofone gemessen und entsprechend der Originalaufnahmen, die Lautsprechersignale in ihrer Leistung angepasst. Dies ist ein adaptiver Prozess, der die Aufnahme der aktuellen Ohrsignale, wie in Abbildung 5.2 dargestellt, voraussetzt. Diese Mikrofone müssen in gleicher Art und Weise wie zuvor bei den Aufnahmen der Referenzsignale kalibriert werden. Nähere Details zur Kalibrierung dieser Umgebungsgeräuschsimulation können [LTS⁺] entnommen werden.

Nach der erfolgreichen Kalibrierung wird, mithilfe der verwendeten Lautsprecher, ein zum aufgezeichneten Referenzszenario ähnliches Leistungsdichtespektrum an den Ohren der Fahrzeuginsassen erzeugt. In Abbildung 5.3 sind die entsprechenden mittleren Leistungsdichtespektren an den Ohrpositionen bei einer Geschwindigkeit von 100 km/h während der Fahrt, innerhalb der Simulationsumgebung und die zugehörige Abweichungen dargestellt.

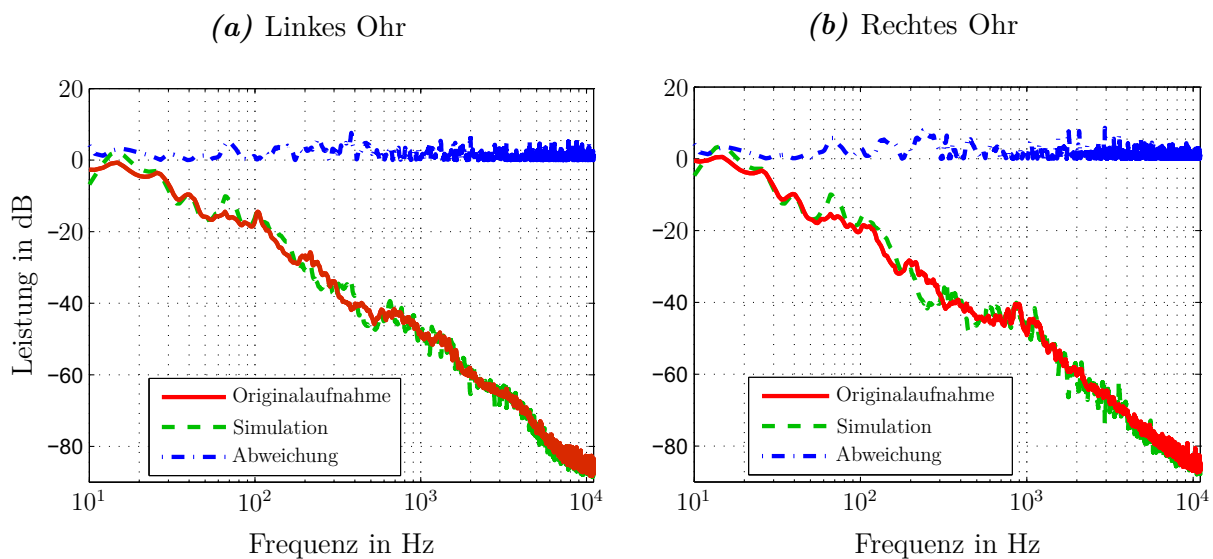


Abbildung 5.3: Leistungsdichtespektren der Ohrsignale bei einer Geschwindigkeit von 100 km/h während der Fahrt und im Labor unter Verwendung der Geräuschsimulation, mit entsprechender Abweichung. In (a) sind die Ergebnisse des linken Ohrs dargestellt, in (b) die des rechten Ohres (nach [LTS⁺]).

Dieses Verfahren, wie auch alle zuvor erwähnten Methoden zur Simulation einer Zielumgebung, hat speziell bei der Reproduktion von Fahrzeuggeräusch eine große Schwierigkeit. Wie in Abbildung 5.3 zu sehen ist, handelt es sich bei Fahrzeuggeräuschen um ein tieffrequent dominantes Geräusch. Um die Leistung in den tiefen Frequenzen bereitstellen zu können und hierdurch eine realistische akustische Umgebung zu erzeugen, wer-

den zusätzliche Lautsprecher mit einem tiefen Frequenzbereich und einer hohen Leistung benötigt. Innerhalb der Simulation, welche in Abbildung 5.2 dargestellt ist, werden zu diesem Zweck neben den Lautsprechern, welche einen mittleren Frequenzbereich anregen, zusätzliche zwei weitere Subwoofer verwendet. Dennoch ist zu sagen, dass auch bei der Verwendung von qualitativ hochwertigen Subwoofern nicht vollständig eine reale Testfahrt reproduziert werden kann. Allerdings bietet die Reproduzierbarkeit solcher Umgebungsgeräuschsimulationen den großen Vorteil, alle zu evaluierenden Systeme gleich anzuregen und somit vergleichbare Evaluierungsergebnisse zu erzielen.

5.2 Lombardsprachsignale

Ein ICC-System wird nicht nur durch das innerhalb der Fahrgastzelle auftretende Hintergrundgeräusch angeregt, sondern auch durch das entsprechende Sprachsignal. Dieses Sprachsignal besitzt in dieser speziellen Geräuschkulisse spezielle Eigenschaften. Neben der Erhöhung des Sprachpegels verändern die Insassen innerhalb eines fahrenden Fahrzeuges zusätzlich die Grundfrequenz und ihre Ausdrucksweise, um nur ein paar Beispiele zu nennen. Insgesamt bezeichnet man diese Anpassung des Sprachsignals an die aktuelle Umgebung durch den Sprecher als Lombard-Effekt [JA90, SPB⁺88]. Während der Fahrt bei unterschiedlichen Geschwindigkeiten verändert sich nicht nur das Hintergrundgeräusch, sondern auch die Sprachsignale der Insassen. Um dennoch ein ICC-System mit realistischen Signalen evaluieren zu können, wird neben dem Geräusch auch das zugehörige Sprachsignal benötigt. In [LTS⁺] ist die Erhebung einer Lombard-Sprachdatenbank beschrieben, die genau diese zur instrumentellen Evaluierung eines ICC-Systems benötigten Signale enthält.

Die Sprachsignale der Datenbank sind innerhalb der gleichen, in Kapitel 5.1 beschriebenen, Geräuschsimulation aufgenommen worden. In Abbildung 5.4 sind schematisch der Geräteaufbau und die verwendete Signalverarbeitung dargestellt. Neben einigen Lautsprechern, die einen mittleren Frequenzbereich abdecken, werden zusätzlich Subwoofer verwendet. Diese stellen sicher, dass ein tieffrequent dominantes Fahrzeuggeräusch nahezu realistisch nachgebildet werden kann, siehe Abschnitt 5.1. Zusätzlich zu den, zur Kalibrierung des Geräuschszenarios notwendigen Referenzmikrofonen, welche an den Positionen der Ohren des Sprechers platziert worden sind, ist ein Aufnahmemikrofon in einem Abstand von 10 cm fest zur der Testperson installiert [LTS⁺]. Dieses Mikrofon nimmt das benötigte, verrauschte Lombard-Sprachsignal auf. Nachdem die entsprechenden Geräuschszenarien kalibriert worden sind, ist eine Testperson an der Sprecherposition platziert worden. Im Anschluss ist das kalibrierte Geräuschszenario aktiviert und die Testperson dazu angehalten worden, erst nach einer Wartezeit von 30 Sekunden einen vorgegebenen Text zu reproduzieren. Die Testperson ist darüber hinaus gebeten worden so vorzulesen, dass eine fiktive, zuhörende Person im Raum den Text verstehen kann. Sowohl die Wartezeit als auch der fiktive Zuhörer sollen ein besseres Einfühlen der Testperson in die Geräuschsitua-

tion ermöglichen und auf diese Weise einen realistischeren Lombard-Effekt aufzeichnen.

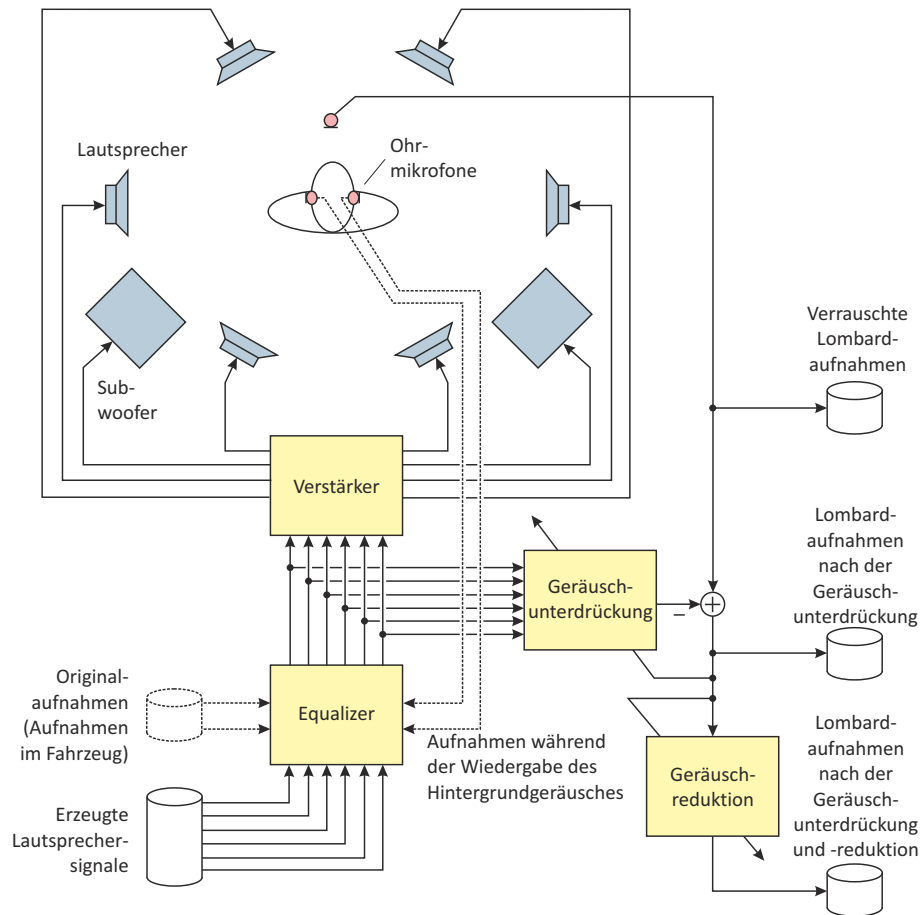


Abbildung 5.4: Schematischer Aufbau der Geräuschsimulation zur Aufnahme von Lombardsprachsignalen. Zusätzlich ist die Signalverarbeitung der Mikrofon-signale dargestellt, welche zur Erzeugung klarer Lombardsprachsignale benötigt wird (nach [LTS⁺]).

Da das Aufnahmемikrofon nicht nur das Sprachsignal aufnimmt, sondern auch das Umgebungsgeräusch, wird das aufgezeichnete verrauschte Mikrofon-signal mittels digitaler Signalverarbeitung aufbereitet. In einem ersten Schritt wird das künstlich erzeugte Fahrzeuggeräusch mithilfe der Lautsprechersignale und einer sogenannten Echo-Kompensation unterdrückt. Diese Geräuschunterdrückung wird im Frequenzbereich mittels eines adaptiven Filters basierend auf einem *Normalized Least Mean Square* (NLMS) Algorithmus durchgeführt [Hay01, S. 320-340].

Da die Geräuschunterdrückung lediglich die Signale der Umgebungssimulation unterdrücken kann, allerdings dennoch ein auftretendes Umgebungsgeräusch im Signal enthalten sein kann, wird im zweiten Schritt eine Geräuschreduktion durchgeführt. Diese Reduktion wird mithilfe eines Filters mit einer sogenannten Wiener-Charakteristik erzielt

[Hay01, S. 94-120]. Beide Verfahren sind zusätzlich in [LTS⁺] erläutert und entsprechende Zwischenschritte beschrieben. Ein Vorteil dieses Verfahrens ist, dass klare Sprachsignale mit enthaltenem Lombard-Effekt aufgezeichnet werden können, jedoch die Testperson nicht durch das Tragen von geschlossenen Kopfhörern beeinträchtigt wird. In Abbildung 5.5 ist beispielhaft ein kurze Sprachpassage mit den entsprechenden signalverarbeitenden Schritten dargestellt. Die obere Grafik stellt die Zeit-Frequenzanalyse des verrauschten Mikrofonsignals dar, welche sowohl aus dem gewünschte Sprachsignal als auch dem Hintergrundgeräusch zusammengesetzt ist. Die mittlere Grafik entspricht der gleichen Passage des Mikrofonsignals nach der Geräuschreduktion. Hier ist zu erkennen, dass bereits dieser erste Signalverarbeitungsschritt eine Verringerung des Geräusches im Bereich zwischen 0 und circa 500 Hz mit sich bringt. Dennoch sind im Vergleich zu der unteren Grafik, welche die Zeit-Frequenzanalyse des Signals nach der Geräuschreduktion darstellt, noch Einflüsse eines auftretenden Hintergrundgeräusches zu sehen, welche abschließend im letzten Signalverarbeitungsschritt der Reduktion bereinigt werden. Dieser letzte Schritt lässt sich besonderes gut zwischen 0,2 und 0,3 s und in einem tiefen Frequenzbereich erkennen.

Mit dem beschriebenen Verfahren werden Sprachsignale für verschiedene Geräuschszenarien aufgenommen und in einer Datenbank dokumentiert. In [LTS⁺] sind zum Beispiel vier verschiedene Geräuschszenarien (Stillstand mit laufendem Motor, Fahrt bei 50 km/h, Fahrt bei 100 km/h und ein Referenzszenario ohne Geräusch) mit drei männlichen und drei weiblichen Testpersonen aufgezeichnet. Diese Datenbank kann zum einen als Anregungssignale für die Evaluierung von ICC-Systemen verwendet werden und zum anderen um das Verhalten von sprechenden Personen in einem fahrenden Fahrzeug besser zu verstehen. In [LNS⁺] ist zum Beispiel die Geräuschsimulation aus Kapitel 5.1 verwendet worden, um die Sprachproduktion und Kommunikation zu analysieren und mit der Situation innerhalb eines fahrenden Fahrzeuges zu vergleichen. Es konnte festgestellt werden, dass sich der Lombard-Effekt auch innerhalb dieser simulierten Geräuschsituation in einem Labor reproduzieren lässt und somit die hier verwendeten Sprachsignale tauglich für eine realistische Anregung eines ICC-Systems sind.

5.3 Implementierung

Eine weitere Anforderung an die ICC-Systemevaluierung ist, dass sie eine automatisch Qualitätserhebung durchführen soll. Dazu wird ein Rahmenwerk benötigt, welches die Möglichkeit der autonomen Verarbeitung von Audio- bzw. Sprachsignalen bereitstellt. Der Lehrstuhl DSS der Christian-Albrechts-Universität zu Kiel besitzt zu einer solchen Verarbeitung, einer Vielzahl von Sensoren bzw. Wandlern, zum Beispiel Mikrofone oder auch medizinische Sensoren, ein Echtzeit-Rahmenwerk namens KiRAT. Dieses in C und C++ umgesetzte Rahmenwerk ist auch für die Implementierung der, innerhalb dieser Arbeit

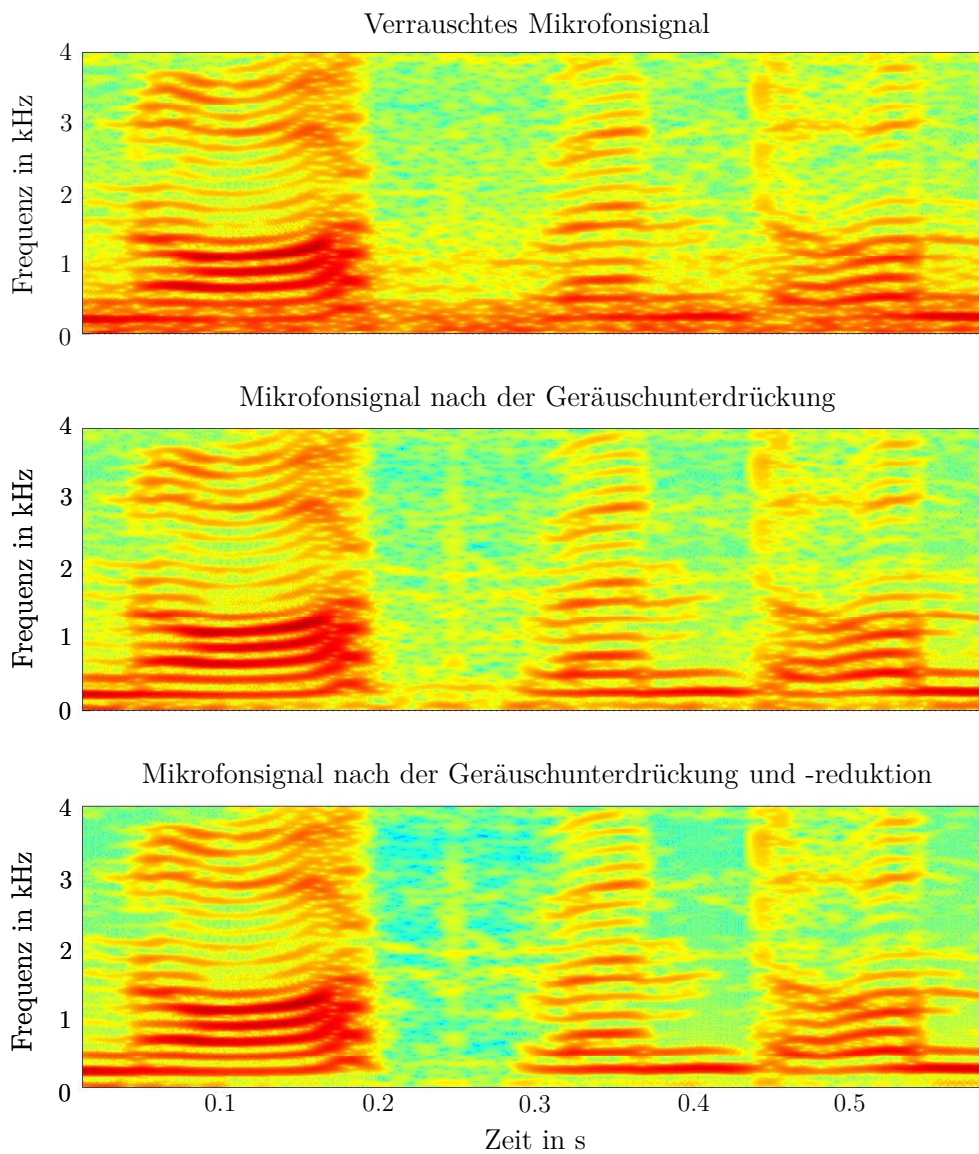


Abbildung 5.5: Zeit-Frequenzanalyse des aufgenommenen verrauschten Sprachsignals vor der Geräuschunterdrückung (obere Grafik), nach der Geräuschunterdrückung (mittlere Grafik) und nach der Geräuschunterdrückung als auch der Geräuschreduktion (untere Grafik) (nach [LTS⁺]).

beschriebenen, Evaluierung von ICC-System verwendet, um so eine automatische Evaluierung des Systems zu ermöglichen.

Zusätzlich bietet das Rahmenwerk KiRAT die Möglichkeit, Eingangssignale, zum Beispiel Audiosignale, in Echtzeit aufzunehmen, zu verarbeiten und wieder auszugeben. Diese Echtzeitfähigkeit wird mittels einer Blockverarbeitung der zuvor digitalisierten Signale gewährleistet, wobei die Größe eines Rahmens R sowie die Abtastrate f_s der Verarbeitung frei wählbar sind. Durch die Blockverarbeitung wird immer eine gewisse Verzögerung der Signal eingefügt, die in Zusammenhang mit dem gewählten Rahmenversatz und der Abtastrate steht. Dadurch wird eine Quasi-Echtzeitverarbeitung durchgeführt, welche für den Zweck einer ICC-Systemevaluierung ausreichend ist. Eine Standardparametrierung ist zum Beispiel eine Rahmengröße von 64 Abtastwerten bei einer Abtastrate von 44,1 kHz. Zur Effizienzsteigerung der Umsetzung von einzelnen Algorithmen, können die Zeitsignale durch eine Analyse- bzw. Synthesefilterbank im Frequenzbereich verarbeitet werden. Dazu wird ein zeitdiskretes Eingangssignal, wie etwa das Eingangssignal der ICC-Systemevaluierung $x(n)$, mithilfe einer Analysefilterbank in den Frequenzbereich überführt. Das zu transformierende Signal wird unter Verwendung eines Hann-Fensters $w_{\text{hann}}(n)$ zunächst gewichtet und im Anschluss mit einer diskreten Fourier-Transformation der Länge N_{FFT} in den Frequenzbereich überführt:

$$X(\mu, k) = \sum_{n=0}^{N_{\text{FFT}}-1} w_{\text{hann}}(n) x(n + kR) e^{-j \frac{2\pi}{N_{\text{FFT}}} \mu n} . \quad (5.1)$$

Dabei entspricht R dem Rahmenversatz, k dem Rahmenindex und μ dem Index des Frequenzbandes. Außerdem kann das Ausgangssignal der ICC-Systemevaluierung $y(n)$ analog zum vorherigen Zusammenhang (5.1) in den Frequenzbereich durch einer Analysefilterbank überführt werden. Das Spektrum des Ausgangssignals ist durch

$$Y(\mu, k) = \sum_{n=0}^{N_{\text{FFT}}-1} w_{\text{hann}}(n) y(n + kR) e^{-j \frac{2\pi}{N_{\text{FFT}}} \mu n} \quad (5.2)$$

gegeben.

Eine Echtzeitsignalverarbeitung, wie sie in KiRAT implementiert ist, bewirkt besondere Ansprüche an die Entwicklung von Evaluierungsverfahren. Neben den Anforderungen ergeben sich jedoch auch neue Möglichkeiten und Vorteile. Ein Beispiel hierfür ist der sogenannte sequentielle SNR, welcher innerhalb eines jeden Blocks bestimmt wird und anschließend entsprechend ausgewertet werden kann. Durch die blockweise Bestimmung des Maßes, können viel präzisere Aussagen über die Veränderung des SNR über die Zeit getätigt werden. Obendrein können als positiver Nebeneffekt bereits zur Laufzeit erste Eindrücke der Qualität eines ICC-Systems gewonnen werden. Diese Informationen in Echtzeit können darüber hinaus zusätzlich für die Entwicklung eines ICC-Systems von Bedeutung sein und entsprechend genutzt werden. Dennoch ist der hauptsächlichste Vorteil und damit der vorrangige Grund der Umsetzung der ICC-Systemevaluierung im Rah-

menwerk KiRAT, die plattformunabhängige, automatische Verarbeitung der Algorithmen. Der Entwurf und die Entwicklung der einzelnen Evaluierungsverfahren werden in Kapitel 7 eingehender beschrieben.

Ein weiterer Vorteil einer Implementierung innerhalb von KiRAT ist die Möglichkeit der Parallelisierung von Verarbeitungsschritten. Die genaue Vorgehensweise und Umsetzung dieser Parallelisierung soll hier nicht vorgestellt werden, lediglich die Möglichkeit erwähnt und die Konsequenzen für die ICC-Systemevaluierung erläutert werden. Bei dieser Parallelisierung werden neben dem Echtzeit-Thread, in dem sich die Hauptverarbeitung des Rahmenwerkes befindet, weitere Threads bereitgestellt. Diese Parallelthreads können anschließend beliebig mit weiteren Teilverarbeitungsschritten befüllt werden, um so die Rechenlast zu verteilen. Um diese Verteilung innerhalb der Evaluierung bewerkstelligen zu können, werden alle Evaluierungsverfahren in drei Phasen aufgeteilt. In der Beschreibung der Verfahren in Kapitel 6 wird nicht explizit auf die Phasen eingegangen, da diese lediglich einen Vorteil in dieser speziellen Implementierung haben und die Verfahren allgemeingültig erläutert werden. Innerhalb der Implementierung in KiRAT durchlaufen alle Evaluierungsverfahren immer die im Folgenden erläuterten drei Phasen:

- **Vorbereitung:** Innerhalb dieser ersten Phase werden vorbereitende Maßnahmen durchgeführt. So wird zum Beispiel die Umgebungsgeräuschsimulation auf das gewünschte Szenario oder auch die Ausgabe des Testsignals kalibriert. Ebenso findet das Anlegen des Evaluierungsprotokolls in dieser ersten Phase statt. Dabei können nur Teile der durchgeführten Maßnahmen in einen der Parallelthreads verschoben werden, wie etwa das Anlegen des Protokolls. Die Kalibrierung der Umgebungsgeräuschsimulation zum Beispiel muss innerhalb des Echtzeit-Threads durchgeführt werden, da sie die Audiodaten der Schnittstelle in Echtzeit benötigt.
- **Durchführung:** Während der Durchführung werden essentielle Daten innerhalb des Echtzeit-Threads gesammelt. Gleichmaßen können dabei erste vorverarbeitende Schritte wie zum Beispiel eine Hochpassfilterung oder eine Analysefilterbank innerhalb dieser Phase durchgeführt werden. Wichtig ist, nicht zu aufwendige Algorithmen zu implementieren, da solche den Echtzeitthread blockieren und dadurch einzelne Rahmen verloren gehen könnten.
- **Nachbereitung:** Diese letzte Phase berechnet mithilfe der Daten aus der Durchführungsphase das eigentliche instrumentelle Qualitätsmaß. Hier ist der Aufwand der Berechnung als zweitrangig zu betrachten, da auch diese Phase in einen der Parallelthreads ausgelagert wird. Abschließend werden in dieser Phase auch alle gewonnenen Erkenntnisse im einem Protokoll festgehalten.

Diese drei Phasen, Vorbereitung, Durchführung und Nachbereitung, werden nacheinander für jedes einzelne Verfahren abgearbeitet, bis alle Ergebnisse feststehen und eine Gesamtqualität wie sie in 7.4 beschrieben ist innerhalb des Rahmenwerkes berechnet werden kann.

Neben der Erhebung der Qualität durch die entsprechenden Evaluierungsverfahren wird ein Protokoll aller Ergebnisse im Detail und im Überblick generiert. Dieses Protokoll wird durch das Rahmenwerk vollautomatisch in *Hypertext Markup Language* (HTML) erzeugt und gespeichert. Dadurch können verschiedene ICC-Systeme unabhängig von einander evaluiert und mittels der Protokolle verglichen werden. Ein weiterer Vorteil ist die Möglichkeit des Informationsaustausches, da durch das detaillierte Protokoll das Evaluierungsergebnis mit den Entwicklern eines ICC-Systems ausgetauscht und so qualitätsfördernde Schritte eingeleitet werden können.

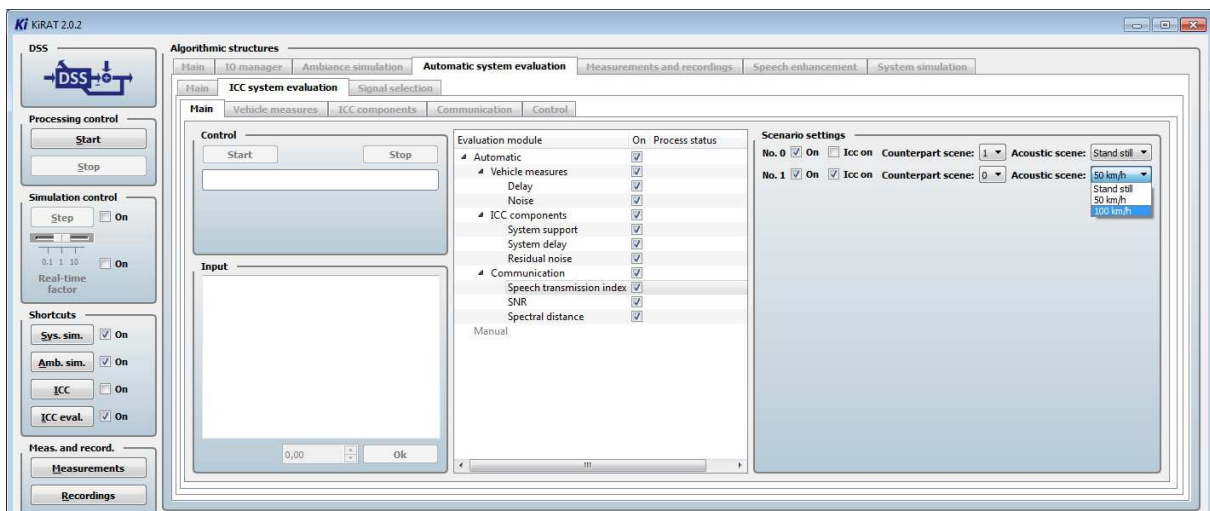


Abbildung 5.6: Hauptbenutzeroberfläche der Evaluierungsumgebung für ICC-Systeme, entwickelt innerhalb des Echtzeit-Rahmenwerks KiRAT unter Verwendung von Qt.

In Abbildung 5.6 ist die in Qt¹ entwickelte Benutzeroberfläche der ICC-Systemevaluierung in KiRAT beispielhaft dargestellt. Zu sehen ist die Hauptebene zur Einstellung der ICC-Systemevaluierung. Im linken Bereich ist der Kontrollbereich (*Control*) und ein Interaktionsfenster (*Input*) zu sehen. Im Kontrollbereich wird die Evaluierung gestartet bzw. gestoppt und der Fortschritt entsprechend angezeigt. Das Interaktionsfenster kann bei Bedarf genutzt werden, falls eines der Evaluierungsverfahren zusätzlich Daten benötigt. Im mittleren Bereich (*Evaluation module*) sind alle Einzelverfahren innerhalb der Gruppenzuordnung aus Kapitel 3.2 dargestellt. Bei Bedarf können bewusst einzelne Verfahren aktiviert und evaluiert werden, um zum Beispiel gezielt nur die Laufzeit eines ICC-Systems zu bestimmen. Der rechte Bereich (*Scenario settings*) dient zur Einstellung der Geräuschsimulation, siehe Kapitel 5.1. Ein weiterer Vorteil der Umsetzung der ICC-Systemevaluierung in KiRAT ist die Möglichkeit der Steuerung anderer im Rahmenwerk umgesetzter Komponenten. Im Fall der Evaluierung kann diese selbstständig ein neues Geräuschszenario kalibrieren und innerhalb einer Messung aktivieren. So kann zum Beispiel

¹Bei Qt handelt es sich um eine C++ -Klassenbibliothek mit deren Hilfe man entsprechende Benutzeroberflächen entwerfen kann.

der SNR innerhalb von verschiedenen Geräuschkennszenarien mit und ohne aktiviertem ICC-System evaluiert und ausgewertet werden. Diese Möglichkeit der Interaktion zwischen der Geräusksimulation und der ICC-Systemevaluierung kann, muss jedoch nicht zwingend genutzt werden. Dadurch bleibt die Möglichkeit bestehen, eine beliebige Geräusksimulation für die Evaluierung eines ICC-Systems zu verwenden.

5.4 Theoretische Modellbildung

Um die Zusammenhänge, die in Kapitel 5 dargestellten Evaluierungsumgebung beschreiben und analysieren zu können, wird in diesem Kapitel ein theoretisches Modell der ICC-Systemevaluierung hergeleitet. Dieses Modell bietet darüber hinaus die Möglichkeit, einer theoretischen Grundlage für den Entwurf und die Interpretation der einzelnen Evaluierungsverfahren.

Die in Abbildung 5.1 dargestellte Basisstruktur der ICC-Systemevaluierung und der Evaluierungsumgebung wird zunächst in einen theoretischen Signalflussgraphen überführt. Zudem wird die Geräusksimulation auf zwei Signale $b_E(n)$ und $b_M(n)$ reduziert. Diese Signale entsprechen dem Geräusch, welches direkt am Ohrmikrofon des künstlichen Zuhörers bzw. direkt am ICC-Mikrofon einkoppelt. Dadurch wird die Geräusksimulation verallgemeinert, sodass sowohl ein reales als auch ein simuliertes Fahrgeräusch angenommen werden können. Das Ergebnis dieser Abstraktionsschritte ist in Abbildung 5.7 dargestellt und bildet die Basis zur weiteren Modellbildung.

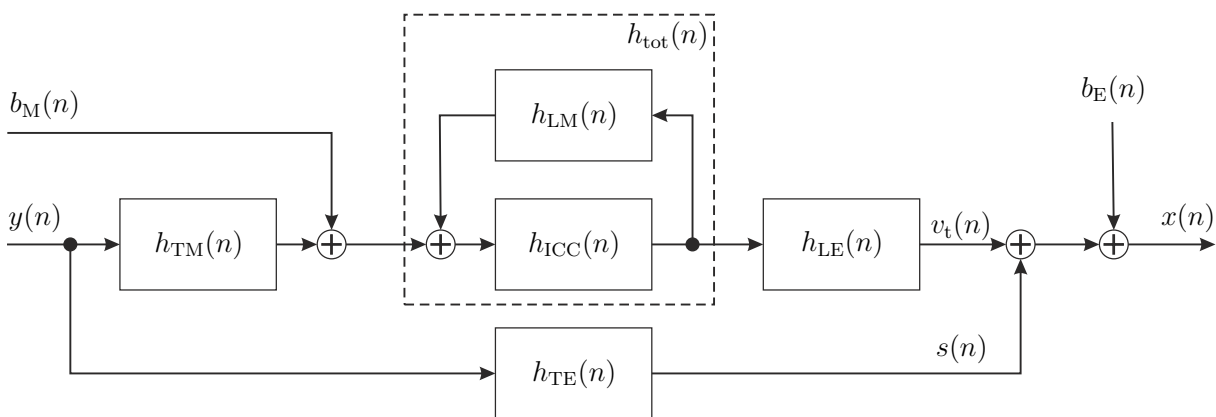


Abbildung 5.7: Theoretischer Signalflussgraph der ICC-Systemevaluierung innerhalb einer Fahrgastzelle mit zugehörigem Hintergrundgeräusch (nach [TSWL]).

Die in Abbildung 5.7 verwendeten Indizes sind in Tabelle 5.1 mit ihrer zugehörigen Bedeutung beschrieben. In Abbildung 5.7 sind die verschiedenen Übertragungspfade innerhalb der Fahrgastzelle dargestellt. Die Impulsantwort $h_{TM}(n)$ entspricht der akustischen Übertragung des Nutzsignals $y(n)$ vom Lautsprecher des künstlichen Sprechers zum ICC-

Abkürzung	Bedeutung
N	Geräuschsimulationspfade
L	Lautsprecher des ICC-Systems
M	Mikrofon des ICC-Systems
E	Ohrmikrofon (künstlicher Zuhörer)
T	Torsolautsprecher (künstlicher Sprecher)

Tabelle 5.1: Liste der in Abbildung 5.7 verwendeten Indizes und die entsprechende Bedeutung.

Mikrofon, $h_{TE}(n)$ hingegen der Übertragung vom Lautsprecher des künstlichen Sprechers zum Zuhörerohr, $h_{LM}(n)$ vom ICC-Lautsprecher zum ICC-Mikrofon (Rückkopplungspfad) und $h_{LE}(n)$ der Übertragung vom ICC-Lautsprecher zum Zuhörerohr. Auch die Übertragung über das ICC-System $h_{ICC}(n)$ ist mitberücksichtigt und wird zusammen mit der Rückkopplung $h_{LM}(n)$ des Nutzsingals über den Raum als $h_{tot}(n)$ beschrieben. Dabei wird das ICC-System als lineares zeitinvariantes System angenommen, was in Realität nicht der Fall ist, jedoch hier für die Modellentwicklung als Näherung angenommen wird. Der Frequenzgang dieser Übertragung $h_{tot}(n)$ ist durch die Kombination der Frequenzgänge der Rückkopplung $H_{LM}(e^{j\Omega})$ und des ICC-Systems $H_{ICC}(e^{j\Omega})$ durch

$$H_{tot}(e^{j\Omega}) = \frac{H_{ICC}(e^{j\Omega})}{1 - H_{ICC}(e^{j\Omega})H_{LM}(e^{j\Omega})} \quad (5.3)$$

gegeben [SH06]. Hier wird deutlich, warum der Rückkopplungspfad eine kritische Komponente der Verarbeitung eines ICC-Systems ist. Wenn $H_{LM}(e^{j\Omega})$ nicht durch das ICC-System $H_{ICC}(e^{j\Omega})$ kompensiert wird, kann sich das System aufschwingen und instabil werden, siehe Kapitel 2.1.

Als Ansatz zur Modellbildung wird die Autokorrelationsfolge des Signals am Ohr bzw. am Eingang der ICC-Systemevaluierung $x(n)$ berechnet. Im Allgemeinen kann eine Autokorrelationsfolge für reellwertige Signal durch

$$s_{xx}(\kappa) = E\{x(n)x(n + \kappa)\} \quad (5.4)$$

bestimmt werden [Mer13, S. 201]. Hierbei beschreibt der Operator $E\{\cdot\}$ die Berechnung des Erwartungswertes.

Aus Abbildung 5.7 wird ersichtlich, dass das Eingangssignal $x(n)$ einer additiven Überlagerung der Signale $v_t(n)$, $s(n)$ und $b_E(n)$ entspricht. Des Weiteren kann das Signal $v_t(n)$ in zwei weitere Signale $t(n)$ und $v(n)$ aufgeteilt werden, siehe Abbildung 5.8. Es ergibt sich, dass das Ausgangssignal wie folgt definiert ist:

$$x(n) = t(n) + v(n) + s(n) + b_E(n). \quad (5.5)$$

Hierbei entspricht das Signal $t(n)$ der Übertragung des Hintergrundgeräusches, aufgenommen durch das ICC-Mikrofon $b_M(n)$, über das ICC-System inklusive der Rückkopplung und über den Lautsprecher des ICC-Systems zum Zuhörerohr. In Abbildung 5.8 ist dieser Zusammenhang im oberen Signalpfad dargestellt. Das Signal $v(n)$ entspricht dem Nutzsignal $y(n)$, welches durch das ICC-Systems verarbeitet und übertragen worden ist, was dem mittleren Signalpfad entspricht. Der unterste Signalpfad beschreibt die Übertragung der Direktschallkomponente innerhalb der Fahrzeugumgebung und wird durch Signal $s(n)$ dargestellt. Abschließend koppelt additiv das Hintergrundgeräusch, welches am Zuhörerohr auftritt $b_E(n)$ in das Eingangssignal der ICC-Systemevaluierung, dem Ohrsignal $x(n)$, ein.

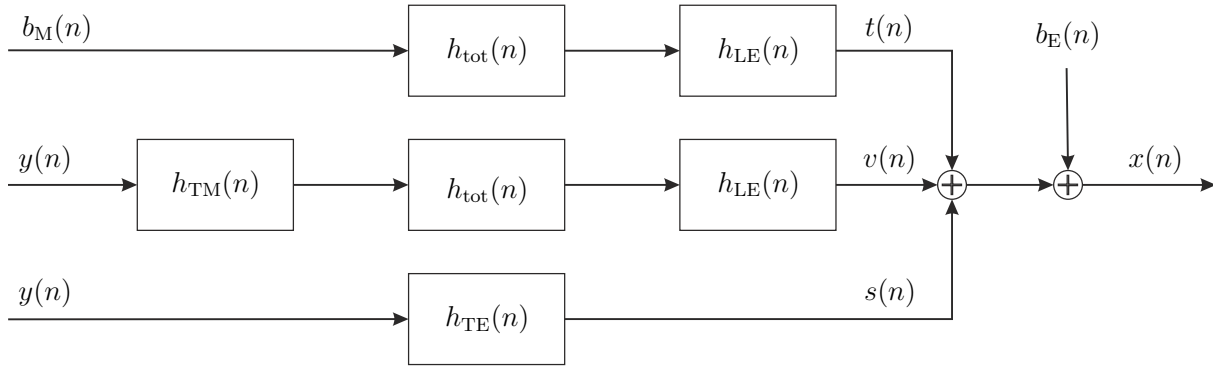


Abbildung 5.8: Weiterentwicklung des theoretischen Signalflussgraphen der ICC-Systemevaluierung aus Abbildung 5.7, durch die Aufteilung der einzelnen Signale in entsprechende Signalflusspfade.

Durch diese Aufteilung des Signales $x(n)$ in verschiedene Einzelsignale in Gleichung (5.5), kann nun der Ansatz aus (5.4) verwendet werden, um eine detaillierte Beschreibung der Autokorrelationsfolge vorzunehmen. Durch Einsetzen von (5.5) in (5.4) ergibt sich die Autokorrelationsfolge des Signals $x(n)$ zu

$$s_{xx}(\kappa) = E \left\{ \left(t(n) + v(n) + s(n) + b_E(n) \right) \cdot \dots \right. \\ \left. \left(t(n + \kappa) + v(n + \kappa) + s(n + \kappa) + b_E(n + \kappa) \right) \right\}. \quad (5.6)$$

Nach Ausklammern und einigen weiteren Umformungen, kann die Autokorrelationsfolge aus den entsprechenden Kreuzkorrelations- und Autokorrelationsfolgen der Teilsignale beschrieben werden. Es folgt, dass

$$s_{xx}(\kappa) = s_{tt}(\kappa) + s_{tv}(\kappa) + s_{ts}(\kappa) + s_{tb_E}(\kappa) + s_{vt}(\kappa) + s_{vv}(\kappa) + s_{vs}(\kappa) + s_{vb_E}(\kappa) + \dots \\ s_{st}(\kappa) + s_{sv}(\kappa) + s_{ss}(\kappa) + s_{sb_E}(\kappa) + \dots \\ s_{b_E t}(\kappa) + s_{b_E v}(\kappa) + s_{b_E s}(\kappa) + s_{b_E b_E}(\kappa) \quad (5.7)$$

ist.

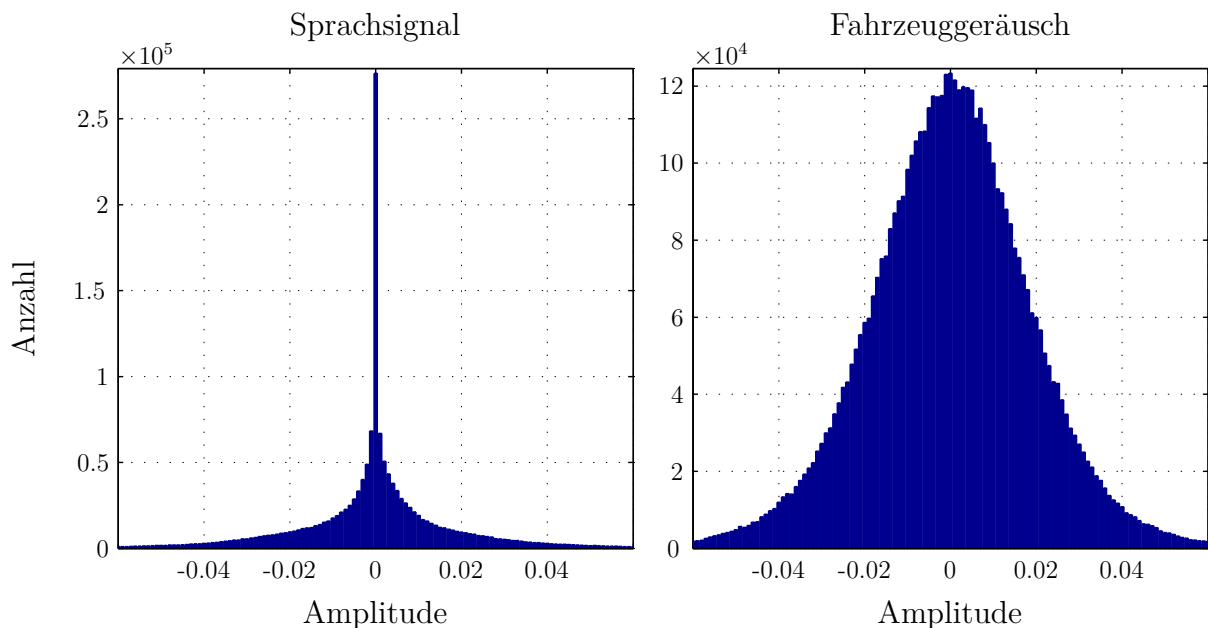


Abbildung 5.9: Amplitudenverteilung des Sprachsignals eines männlichen Sprechers und des Fahrzeuggeräusches einer Mercedes-Benz E-Klasse bei einer Geschwindigkeit von 150 km/h.

Bei einer genaueren Betrachtung des Sprachsignals und des auftretenden Hintergrundgeräusches können nun einige Annahmen getroffen werden. In Abbildung 5.9 ist links das Histogramm der Amplituden des Sprachsignals eines männlichen Sprechers und rechts das eines Fahrgeräusches innerhalb einer Mercedes-Benz E-Klasse bei einer Geschwindigkeit von 150 km/h dargestellt. Es wird deutlich, dass beide eine unterschiedliche Verteilung besitzen und mithilfe von Zufallsprozessen nachgebildet werden können. Für ein Sprachsignal wird meist eine Laplaceverteilung angenommen [GZ03] und offensichtlich kann ein Fahrzeuggeräusch mittels einer Gaußverteilung beschrieben werden. Gemeinsam ist beiden Verteilungen, dass sie mittelwertfrei sind. Da das Sprach- und das Geräuschsignal keinem gemeinsamen Ursprung entstammen, können sie als unkorreliert angenommen werden. Zwei Zufallsprozesse können als orthogonal angesehen werden, wenn beide unkorreliert und mindestens ein Prozess mittelwertfrei ist [Hän01, S. 63 ff.]. Bei Orthogonalität zweier Zufallsprozesse ist die Kreuzkorrelationsfolge dieser beiden Prozesse gleich null [Hän01, S. 63 ff.]. Es folgt für die entsprechenden Kreuzkorrelationsfolgen, dass

$$s_{tv}(\kappa) = s_{vt}(\kappa) = s_{ts}(\kappa) = s_{st}(\kappa) = s_{b_E v}(\kappa) = s_{v b_E}(\kappa) = s_{b_E s}(\kappa) = s_{s b_E}(\kappa) = 0 \quad (5.8)$$

ist. Durch Einsetzen des neuen Zusammenhanges (5.8) in (5.7) folgt:

$$\begin{aligned} s_{xx}(\kappa) &= s_{tt}(\kappa) + s_{tb_E}(\kappa) + s_{vv}(\kappa) + s_{vs}(\kappa) + s_{tb_E}(\kappa) + s_{sv}(\kappa) + \dots \\ &\quad + s_{ss}(\kappa) + s_{b_E t}(\kappa) + s_{b_E b_E}(\kappa). \end{aligned} \quad (5.9)$$

Des Weiteren wird das auftretende Hintergrundgeräusch als diffuses Schallfeld angenommen. Innerhalb eines diffusen Schallfeldes, ist die Energiedichte bzw. die Leistung an jedem Punkt identisch [NN97, Kut09]. Da die Leistungsdichte eines stochastischen Signals über die Fourier-Transformation im Zusammenhang mit der Autokorrelationsfolge steht, kann folgende Annahme für die Autokorrelationsfolgen getroffen werden:

$$s_{b_E b_E}(\kappa) = s_{b_M b_M}(\kappa) = s_{bb}(\kappa). \quad (5.10)$$

Zusätzlich wird angenommen, dass die Signale $t(n)$ und $b_E(n)$ zusätzlich orthogonal zueinander sind. Diese Annahme kann durch die Verarbeitung des Geräuschsignals $b_M(n)$ durch das ICC-System begründet werden, im besten Fall wird durch diese Verarbeitung das Signal $t(n) = 0$. Es folgt für die entsprechenden Kreuzkorrelationsfolgen, dass

$$s_{b_E t}(\kappa) = s_{t b_E}(\kappa) = 0 \quad (5.11)$$

ist. Durch diesen Zusammenhang kann die Gleichung (5.9) weiter angepasst werden. Die Autokorrelationsfolge des Eingangssignals der ICC-Systemevaluierung vereinfacht sich weiter zu:

$$s_{xx}(\kappa) = s_{tt}(\kappa) + s_{vv}(\kappa) + s_{vs}(\kappa) + s_{sv}(\kappa) + s_{ss}(\kappa) + s_{bb}(\kappa). \quad (5.12)$$

Unter Verwendung dieser theoretischen Autokorrelationsfolge $s_{xx}(\kappa)$ kann daraufhin das entsprechende Leistungsdichtespektrum bestimmt werden. Dazu wird $s_{xx}(\kappa)$ aus Gleichung (5.12) mithilfe einer diskreten Fourier-Transformation in den Frequenzbereich überführt. Das Auto-Leistungsdichtespektrum des Ausgangssignals $x(n)$ entspricht

$$S_{xx}(\Omega) = \mathcal{F}\{s_{xx}(\kappa)\} = \sum_{\kappa=-\infty}^{\infty} s_{xx}(\kappa) e^{-j\Omega\kappa}, \quad (5.13)$$

wobei $\mathcal{F}\{\cdot\}$ der zeitdiskreten Fourier-Transformation entspricht und Ω der normierten Frequenz [Mer13, S. 201 ff.]. Setzt man (5.12) in (5.13) ein und teilt die Summe entsprechend auf, folgt, dass

$$S_{xx}(\Omega) = S_{tt}(\Omega) + S_{vv}(\Omega) + S_{vs}(\Omega) + S_{sv}(\Omega) + S_{ss}(\Omega) + S_{bb}(\Omega) \quad (5.14)$$

ist. Die Kreuz- und Autoleistungsdichtespektren können auf die Eingangsleistungsdichtespektren $S_{bb}(\Omega)$ und $S_y(\Omega)$ weiter herunter gebrochen werden. Dazu werden die entsprechenden Betragsfrequenzgänge mit dem Eingangsleistungsdichtespektrum multipliziert.

Mithilfe von Abbildung 5.8 ergeben sich die folgenden Kreuz- und Autoleistungsdichtespektren:

$$S_{tt}(\Omega) = S_{bb}(\Omega) \left| H_{\text{tot}}(e^{j\Omega}) H_{\text{LE}}(e^{j\Omega}) \right|^2 \quad (5.15)$$

$$S_{vv}(\Omega) = S_{yy}(\Omega) \left| H_{\text{TM}}(e^{j\Omega}) H_{\text{tot}}(e^{j\Omega}) H_{\text{LE}}(e^{j\Omega}) \right|^2 \quad (5.16)$$

$$S_{vs}(\Omega) = S_{yy}(\Omega) H_{\text{TM}}(e^{j\Omega}) H_{\text{tot}}(e^{j\Omega}) H_{\text{LE}}(e^{j\Omega}) H_{\text{TE}}(e^{j\Omega})^* \quad (5.17)$$

$$S_{sv}(\Omega) = S_{yy}(\Omega) H_{\text{TM}}^*(e^{j\Omega}) H_{\text{tot}}^*(e^{j\Omega}) H_{\text{LE}}^*(e^{j\Omega}) H_{\text{TE}}(e^{j\Omega}) \quad (5.18)$$

$$S_{ss}(\Omega) = S_{yy}(\Omega) \left| H_{\text{TE}}(e^{j\Omega}) (e^{j\Omega}) \right|^2. \quad (5.19)$$

Diese können entsprechend in Gleichung (5.14) eingesetzt werden, um das Autoleistungsdichtespektrum am Eingang der ICC-Systemevaluierung zu ermitteln. Es folgt, dass

$$\begin{aligned} S_{xx}(\Omega) &= S_{bb}(\Omega) \left| H_{\text{tot}}(e^{j\Omega}) H_{\text{LE}}(e^{j\Omega}) \right|^2 + S_{yy}(\Omega) \left| H_{\text{TM}}(e^{j\Omega}) H_{\text{tot}}(e^{j\Omega}) H_{\text{LE}}(e^{j\Omega}) \right|^2 + \dots \\ &S_{yy}(\Omega) H_{\text{TM}}(e^{j\Omega}) H_{\text{tot}}(e^{j\Omega}) H_{\text{LE}}(e^{j\Omega}) H_{\text{TE}}(e^{j\Omega})^* + \dots \\ &S_{yy}(\Omega) H_{\text{TM}}^*(e^{j\Omega}) H_{\text{tot}}^*(e^{j\Omega}) H_{\text{LE}}^*(e^{j\Omega}) H_{\text{TE}}(e^{j\Omega}) + \dots \\ &S_{yy}(\Omega) \left| H_{\text{TE}}(e^{j\Omega}) (e^{j\Omega}) \right|^2 + S_{bb}(\Omega) \end{aligned} \quad (5.20)$$

ist. Des Weiteren können die Faltungen mit den Leistungsdichtespektren des Eingangssignals $y(n)$ zu einem Term zusammen gefasst werden.

$$\begin{aligned} S_{xx}(\Omega) &= S_{bb}(\Omega) \left| H_{\text{tot}}(e^{j\Omega}) H_{\text{LE}}(e^{j\Omega}) \right|^2 + S_{bb}(\Omega) + \dots \\ &S_{yy}(\Omega) \left| H_{\text{TM}}(e^{j\Omega}) H_{\text{tot}}(e^{j\Omega}) H_{\text{LE}}(e^{j\Omega}) + H_{\text{TE}}(e^{j\Omega}) \right|^2. \end{aligned} \quad (5.21)$$

Das Ergebnis aus (5.21) zeigt, dass das Eingangs-Leistungsdichtespektrum $S_{xx}(\Omega)$ sich aus drei Hauptkomponenten zusammen setzt.

1. **Zuhörer-Mikrofongeräusch:** $S_{bb}(\Omega)$

Dies ist der Anteil des Hintergrundgeräusches, welches direkt in das Eingangsmikrofon der ICC-Systemevaluierung einkoppelt.

2. **ICC-Mikrofongeräusch:** $S_{bb}(\Omega) \left| H_{\text{tot}}(e^{j\Omega}) H_{\text{LE}}(e^{j\Omega}) \right|^2$

Hierbei handelt es sich um den Anteil des Hintergrundgeräusches, welcher direkt vom ICC-Mikrofon aufgenommen wird. Dementsprechend muss an dieser Stelle die Übertragung des ICC-Systems in der Abschätzung hinzugefügt werden. Dieser Anteil ist zweifelsohne stark abhängig von dem ICC-System selbst, denn das System kann diese Komponente durch die Verwendung entsprechender Algorithmen fast vollständig kompensieren.

3. **Referenzsignalanteil:** $S_{yy}(\Omega) \left| H_{\text{TM}}(e^{j\Omega}) H_{\text{tot}}(e^{j\Omega}) H_{\text{LE}}(e^{j\Omega}) + H_{\text{TE}}(e^{j\Omega}) \right|^2$

Dieser dritte spektrale Anteil entspricht der Leistungsdichte, welche das Test- bzw.

Referenzsignal zum Gesamtleistungsdichtespektrum $S_{xx}(\Omega)$ beiträgt. Auch dieser Teil steht in direkter Abhängigkeit von der Übertragungsfunktion des ICC-Systems, jedoch ebenso auch von der Übertragungsfunktion $H_{TE}(e^{j\Omega})$ des Direktschalls.

Mithilfe dieses theoretischen Modells können nun die einzelnen Evaluierungsverfahren theoretisch beleuchtet werden. Das Modell ermöglicht eine Identifizierung der qualitätsrelevanten Signalkomponenten und den Einfluss des ICC-Systems auf diese Komponenten. Des Weiteren wird deutlich, dass alle Evaluierungsverfahren auch durch andere Anteile, wie das auftretende Hintergrundgeräusch, beeinflusst werden und dies in der Entwicklung der Verfahren berücksichtigt werden muss. Dennoch handelt es sich um einen theoretischen Ansatz, welcher einige Annahmen beinhaltet, die bei einer realen Fahrt unter Umständen nicht erfüllt sind. So kann zum Beispiel die Diffusität des Hintergrundgeräusches mitunter während der Fahrt nicht gewährleistet werden. Dies würde dazu führen, dass die Annahme aus Gleichung (5.10) nicht mehr gilt und die Anteile des Hintergrundgeräusches am ICC-Mikrofon und am Zuhörerohr separat betrachtet werden müssten. Innerhalb einer Laborumgebung mit Geräuschsimulation kann die Diffusität des Hintergrundgeräusches gezielt erzeugt und somit Zusammenhang (5.10) angenommen werden. Deshalb werden alle im weiteren Verlauf vorgestellten instrumentellen Evaluierungsverfahren stets innerhalb dieser Geräuschsimulation betrachtet und entwickelt.

Kapitel 6

Auditive Evaluierungsverfahren

Ein üblicher Ansatz zur Erhebung der Qualität von signal- bzw. sprachverarbeitenden Systemen sind auditive Evaluierungsverfahren. Wie bereits in Kapitel 4.1 beschrieben, beurteilen in diesem Fall Testpersonen die Qualität der dargebotenen Signale. Dabei wird davon ausgegangen, dass die betrachtete Testgruppe eine allgemeingültige Aussage über die Qualität in Bezug auf die gegebene Fragestellung trifft. Da diese Verfahren meist kosten- und aufwandsintensiv sind, gilt es, sie auf ein notwendiges Minimum zu reduzieren und die Qualität von Systemen anhand von instrumentellen Verfahren zu erheben. Dennoch sind die Ergebnisse auditiver Evaluierungsverfahren ein wichtiger Bestandteil im Entwurf instrumenteller Evaluierungsverfahren: Sie bilden die Qualitätsreferenz, die es zu schätzen gilt.

Für die Erhebung der Sprachqualität von signal- bzw. sprachverarbeitenden Systemen, anhand von auditiven Verfahren, gibt es bereits einige veröffentlichte Ansätze. So sind von der *International Telecommunication Union* (ITU) bereits Empfehlungen bzw. standardisierte Verfahren zu diesem Thema, zum Beispiel in [ITU96a, ITU96b, ITU94], publiziert worden. Diese Verfahren sind zum Teil speziell für Telekommunikationssysteme oder im Allgemeinen für sprachverbessernde Systeme entwickelt worden, jedoch nicht gezielt für die Evaluierung von ICC-Systemen. Über die Erhebung der Qualität eines ICC-Systems mithilfe auditiver Verfahren gibt es derzeit, nach Wissen der Autorin, keine Empfehlungen oder Standards. Dennoch bieten die bereits veröffentlichten Verfahren eine, aufgrund ihrer methodischen Vorgehensweise, einwandfreie Basis für die Entwicklung neuer auditiver Verfahren zur Evaluierung von ICC-Systemen.

In den folgenden Kapiteln werden unterschiedliche Hörtests sowie weitere interaktive psychoakustische Experimente vorgestellt, welche die Qualität von ICC-Systemen durch die Bewertung von Testpersonen erheben sollen. Es wird eine Qualitätsreferenz geschaffen, welche im Anschluss in Kapitel 7 zum Entwurf der einzelnen instrumentellen Verfahren weiterverwendet wird. Neben dieser wichtigen Qualitätsreferenz offenbaren die zunächst betrachteten auditiven Evaluierungsverfahren erste Erkenntnisse über das Verhalten von Passagieren in einem fahrenden Fahrzeug.

6.1 Echoperzeption

Aufgrund der Verwendung eines ICC-Systems in einem Fahrzeug entstehen während der Kommunikation für den Zuhörer zwei Signalquellen des gewünschten Sprachsignals, siehe Abbildung 6.1. Grund hierfür ist die zusätzlich Wiedergabe durch das ICC-System, welche bereits in Abschnitt 2.1 beschrieben worden ist. Bei den beiden Quellen handelt es sich zum einen, um

- den Direktschall, welcher direkt vom Sprecher ausgesendet und vom Zuhörer empfangen wird, und zum anderen, um
- den Schall, welcher vom ICC-System über die installierten Lautsprecher ausgesendet und vom Zuhörer empfangen wird.

Aufgrund von akustischen Übertragungszeiten und der Verarbeitungszeit des ICC-Systems selbst, werden der Direktschall und die zusätzliche Quelle nicht synchron von dem Zuhörer wahrgenommen. Um die korrekte Lokalisierung der Signalquelle, dem Sprecher, aufrechtzuerhalten, ist diese Verzögerung auch erwünscht, da so das „Gesetz der ersten Wellenfront“ oder der Präzedenz-Effekt erfüllt wird [MS52, Bla97]. Die Lokalisierung einer Quelle ist allerdings nicht ausschließlich von dem Laufzeitunterschied der beiden Signale abhängig, sondern auch von dem Leistungsunterschied. So wird, im Falle zweier Quellen mit einer gleichen Laufzeit und einem unterschiedlichen Schalldruckpegel die Quelle als Ursprung identifiziert, welche eine höhere Leistung besitzt [ZF99, S. 311]. Um die Lokalisierung auch bei der Verwendung eines ICC-Systems zu erhalten, muss eine gewisse Verzögerung und ein bestimmter Leistungsunterschied zwischen diesen beiden Signalquellen, dem Direktschall und der ICC-Systemwiedergabe, vorhanden sein. Bei der Entwicklung eines ICC-Systems werden die Parameter Verzögerung und Verstärkung, siehe Kapitel 2.1, im Rahmen des Machbaren so gewählt, dass ein stabiles und für den Entwickler wohlklingendes System entsteht. Nun stellt sich die Frage nach den theoretischen und den allgemeinen Grenzen dieser Parameter. Um dies zu klären, wird in einem psychoakustischen Experiment die Echowahrnehmung des zuhörenden Passagiers innerhalb eines Fahrzeugs, ausgestattet mit einem ICC-System, genauer beleuchtet.

Um einzelne Einflüsse auf den zuhörenden Passagier genauer untersuchen zu können, müssen zunächst die nicht zu analysierenden Einflüsse auf diese Person minimiert werden. Aus diesem Grund ist ein vom Fahrzeug abstrahiertes Experiment entworfen worden, welches in Abbildung 6.1 dargestellt ist. Da die Kommunikation zwischen einem Fahrer und einem Passagier hinter dem Fahrer am stärksten beeinträchtigt ist, wird innerhalb des Experiments der Fahrer als Sprecher und die Testperson hinter diesem Sprecher als zuhörender Passagier ausgewählt. Als Sprecher wird ein künstlicher Sprecher, ein Torso mit integriertem Lautsprecher, verwendet. Dadurch kann das Sprachsignal des Sprechers für alle Testpersonen in gleicher Weise angeregt werden. Zusätzlich sitzen der künstliche Sprecher sowie die Testperson nicht in einem Fahrzeug, da folglich das Experiment

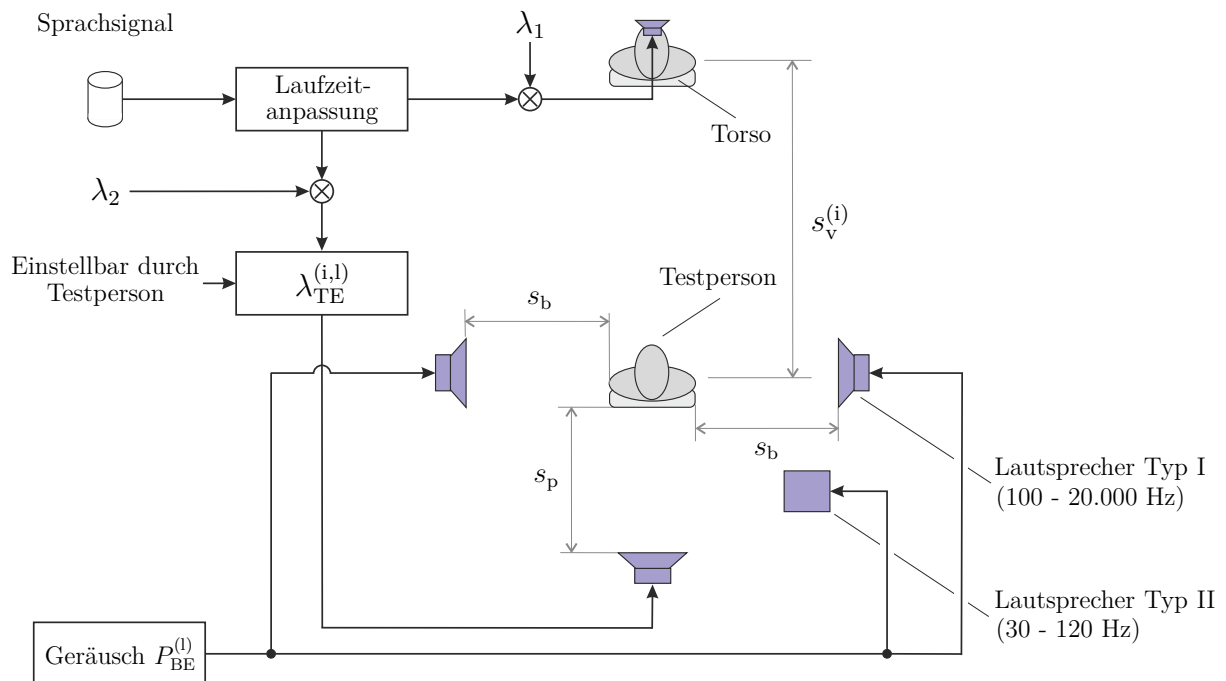


Abbildung 6.1: Schematischer Aufbau des Experiments zur Bestimmung der Echoperzeption des zuhörenden Passagiers in einem reflexionsarmen Raum. Des Weiteren ist der Signalflussgraf, welcher zur Anregung der Testsignale benötigt wird, abgebildet.

auf die eine Anordnung in diesem speziellen Fahrzeug beschränkt wäre. Innerhalb des Experiments sollen jedoch verschiedene Abstände, wie sie in unterschiedlichen Fahrzeugen auftreten, analysiert werden. Dazu wird der Abstand $s_v^{(i)}$ zwischen dem Zuhörer und dem Sprecher im Laufe des Experiments verändert. In Tabelle 6.1 sind die verwendeten Abstände aller Komponenten zusammengetragen. Insgesamt werden drei verschiedene Distanzen, die unterschiedlich große Fahrzeuge repräsentieren, zwischen dem Sprecher und dem Zuhörer $s_v^{(i)}$, betrachtet.

Um nun die zweite Quelle innerhalb der Fahrgastzelle, welche durch die Verwendung eines ICC-Systems auftritt, möglichst kontrolliert anregen zu können, wird ein Lautsprecher als zweite Signalquelle verwendet. Dadurch kann diese Wiedergabe des Nutzsignals für jede Testperson in gleicher Weise erfolgen und so eine Vergleichbarkeit der einzelnen Testergebnisse erzielt werden. Zusätzlich wird lediglich ein Lautsprecher verwendet, um auf diese Weise zunächst die Einflüsse durch eine mehrfache Wiedergabe zu minimieren. Der Lautsprecher, welcher die Wiedergabe des ICC-Systems simuliert, ist hinter der Testperson in einem Abstand von $s_p = 156$ cm positioniert. Zusätzlich ist in Abbildung 6.1 zu sehen, dass sowohl der Torso als auch der Lautsprecher durch das gleiche klare Sprachsignal angeregt werden. Diese beiden Signale werden in ihrer Laufzeit dahingehend angepasst, dass sie gleichzeitig den Zuhörer bzw. die Testperson erreichen. Zu diesem

Zweck müssen für die unterschiedlichen Distanzen $s_v^{(i)}$ zwischen dem Torso und der Testperson unterschiedliche Laufzeitanpassungen durchgeführt werden. Zum Beispiel muss bei einer Distanz $s_v^{(1)} = 80$ cm das Sprachsignal des Torsos um 2,3 ms verzögert werden, damit beide Schallereignisse gleichzeitig bei der Testperson eintreffen. Auch die Leistung der beiden Quellen wird zu Beginn eines Tests für jede Testperson auf einen Zielwert kalibriert. Dazu wird in Anlehnung an den ANSI-Standard S3.5-1997 in einem Abstand von 50 cm ein Sprachpegel von 62 dB bei beiden Lautsprechern mithilfe der Verstärkungsfaktoren λ_1 und λ_2 eingestellt [ANS97].

Bezeichnung	Abstand	Verzögerung	Ergänzungen
s_b	163 cm		
s_p	156 cm	4,5 ms	
$s_v^{(1)}$	80 cm	2,3 ms	Mercedes-Benz A-Klasse
$s_v^{(2)}$	100 cm	2,9 ms	Mercedes-Benz S-Klasse (Langversion)
$s_v^{(3)}$	210 cm	6,1 ms	Mercedes-Benz V-Klasse (Rückbank)

Tabelle 6.1: Innerhalb des Echoperzeptionsexperiments verwendete Distanzen, die zugehörigen akustischen Laufzeiten und einige ergänzende Informationen.

Zusätzlich zu dem Abstand zwischen dem Torso und der Testperson, soll das Hintergrundgeräusch variiert werden. Des Weiteren erfolgt der komplette, bisher beschriebene Aufbau des Experiments innerhalb einer Laborumgebung in einem reflexionsarmen Raum. Dieser Raum minimiert die durch Wände oder Scheiben, wie beispielsweise in einem Fahrzeug, auftretenden Reflexionen. Dadurch kann zum einen gezielt ein definiertes Geräuschszenario mithilfe von Lautsprechern erzeugt und zum anderen Einflüsse durch die Reflexion der Signale minimiert werden. Dies ist von Bedeutung, da Reflexionen des Sprachsignals als zusätzliche Quellen wahrgenommen werden könnten und in diesem Fall den Einfluss der zu untersuchenden zwei Quellen verändern würden. Zur Anregung des Hintergrundgeräusches, siehe Abbildung 6.1, sind zwei Lautsprecher vom Typ I, links und rechts der Testperson, sowie ein Lautsprecher vom Typ II platziert worden. Diese Wahl der Lautsprecher ermöglicht es, auch ein tieffrequent dominantes Geräusch, wie das in einem Fahrzeug auftretende, anzuregen. Da das Fahrzeuggeräusch abhängig vom gewählten Fahrzeug ist, wird in diesem Experiment ein künstliches Geräusch verwendet. Dieses ähnelt dem eines Fahrzeuges und ermöglicht daher eine Fahrzeug unabhängige Analyse der Ergebnisse. Um die Auswirkung verschiedener Geräusche analysieren zu können, wird neben einem Szenario ohne Geräusch, auch ein Szenario mit Anregung eines farbigen Rauschens und ein weiteres mit einem weißen Rauschen durchgeführt. In Abbildung 6.2 ist der allgemeine Verlauf der mittleren Leistungsdichtespektren der verwendeten Signale dargestellt. Ähnlich wie ein Fahrzeuggeräusch, regt das verwendete farbige Rauschen große Teile der Leistung in den tiefen Frequenzen an. Die Leistungen der einzelnen Geräuschsze-

narien, gemessen an der Position der Testperson, sind in Tabelle 6.2 zusammengetragen.

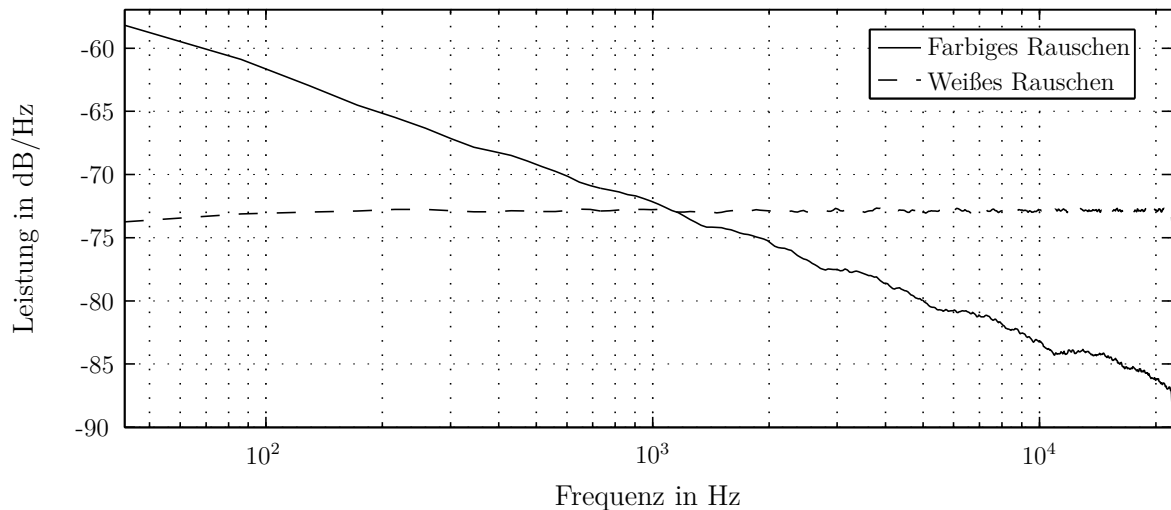


Abbildung 6.2: Allgemeine Leistungsdichtespektren der innerhalb des Experiments verwendeten Geräuschsignale. Die breitbandige Leistung der Geräuschszenarien an der Position der Testperson ist nach Tabelle 6.2 gewählt worden.

Während des eigentlichen Experiments werden die Testpersonen aufgefordert, die Verstärkung $\lambda_{\text{TE}}^{(i,l)}$ des hinteren Lautsprechers so einzustellen, dass sie den Text der abgespielten Sätze problemlos und ohne Mühe verstehen können. Die Testpersonen haben dabei keine visuelle Rückmeldung welche Verstärkung sie eingestellt haben, sondern können nur mithilfe von Plus- bzw. Minus-Knöpfen innerhalb einer Benutzeroberfläche die Verstärkung erhöhen und auch erniedrigen. Dadurch können sich die Testpersonen langsam an die gewünschte Verstärkung herantasten, bevor sie zum nächsten Szenario übergehen. Bei der Testgruppe handelte es sich um 14 Personen, 4 weibliche und 10 männliche, welche zwischen 22 und 31 Jahre alt waren. Zum Zeitpunkt des Hörtests war bei keiner der Testpersonen eine Einschränkung der Hörfähigkeit bekannt. Zwölf Testpersonen und damit ein Großteil der Testgruppe können als naiv eingestuft werden, da es sich um Studierende der Fakultät handelt. Zwei der Probanden sind zum Zeitpunkt des Experiments Doktoranden des Lehrstuhls DSS gewesen, welche sich im Zuge ihrer Tätigkeit mit Sprach- und Audiosignalverarbeitung beschäftigen.

6.1.1 Ergebnisse

In Abbildung 6.3(a) sind die Mittelwerte der Verstärkungseinstellung, in dB, und die zugehörige Standardabweichung über alle Testpersonen in Abhängigkeit des Geräuschszenarios und der Distanz dargestellt. Die rote Kurve entspricht einer Distanz zwischen dem künstlichen Sprecher und der Testperson von 210 cm, die grüne einer Distanz von 100 cm und

	Ohne Geräusch $P_{BE}^{(1)}$	Farbiges Geräusch $P_{BE}^{(2)}$	Weißes Geräusch $P_{BE}^{(3)}$
dBA	31,1	66,9	74,3
dB	51,3	82,0	78,1

Tabelle 6.2: Mittlere Geräuschleistungen in dBA bzw. dB der einzelnen Geräuschkategorien, gemessen mit einem Pegelmessgerät an der Position der Testperson.

die blaue einer Distanz von 80 cm. Die unterschiedlichen Geräuschkategorien sind entsprechend über die Abszisse zugeordnet. Es wird deutlich, dass mit dem Anstieg der Leistung des Hintergrundgeräusches ebenfalls die Verstärkungseinstellung steigt. Dies ist eine plausible Beobachtung, da sich mit zunehmender Geräuschleistung die Verständlichkeit des Sprachsignals verschlechtert und diese Beeinträchtigung durch die Erhöhung der Leistung des Lautsprechersignals kompensiert werden kann. Bei einem Abstand $s_v^{(1)} = 80$ cm steigt die Verstärkung zwischen einem fehlenden und dem farbigen Geräusch um 31,3 dB; bei einer Distanz von $s_v^{(2)} = 100$ cm um 30,1 dB und bei $s_v^{(3)} = 210$ cm um 39,5 dB. Zwischen dem farbigen und dem weißen Rauschen verändern sich die mittleren Verstärkungseinstellungen nur noch marginal. Dies deutet darauf hin, dass es sich in beiden Fällen um so stark beeinträchtigende Szenarien handelt, dass der spektrale Verlauf der Geräusche keinen Unterschied mehr auf die Verständlichkeit erzeugt und immer eine hohe Verstärkung benötigt wird. Dennoch kann eine erste maximale Verstärkungshypothese in Abhängigkeit des Geräusches anhand der in Abbildung 6.3(a) dargestellten Ergebnisse erzielt werden. Im Mittel wird ein Anstieg der Verstärkungseinstellung um circa 33,6 dB von keinem Geräusch auf eines der lauten Geräuschkategorien erzielt. Dadurch ergibt sich ein äquivalenter Dauerschallpegel in einem Abstand von 50 cm von circa 95 dB was der Lautstärke einer schreienden Person entspricht. Betrachtet man diesen Anstieg in Bezug auf den Leistungsanstieg des Hintergrundgeräusches, so lässt sich ein erster Eindruck der benötigten Verstärkung erzielen. Bei einem farbigen Geräusch, welches spektrale Ähnlichkeit zum Fahrzeuggeräusch besitzt, würde ein Anstieg der Verstärkung um 0,94 dB pro 1 dBA Anstieg des Hintergrundgeräusches benötigt, um das Sprachsignal ohne einen extra Aufwand verstehen zu können. Um die gefundene Hypothese zu verifizieren und den detaillierten Verlauf zu bestimmen, müssten weitere Experimente mit feineren Geräuschabstufungen durchgeführt und analysiert werden.

In Abbildung 6.3(b) sind die Mittelwerte der Verstärkungseinstellung und die zugehörige Standardabweichung über alle Testpersonen über die möglichen Distanzen aufgetragen, wobei alle Kurven von ihrem Mittelwert bereinigt worden sind. Dies ermöglicht einen Fokus auf die Veränderung der Verstärkungseinstellung über die gegebenen Freiheitsgrade. Dabei repräsentiert die rote Kurve ein weißes, die grüne Kurve ein farbiges Geräusch und die blaue ein Szenario ohne Hintergrundgeräusch. Hier ist nun deutlich zusehen, dass die Verstärkungseinstellung innerhalb des farbigen und weißen Geräuschkategorien sich ledig-

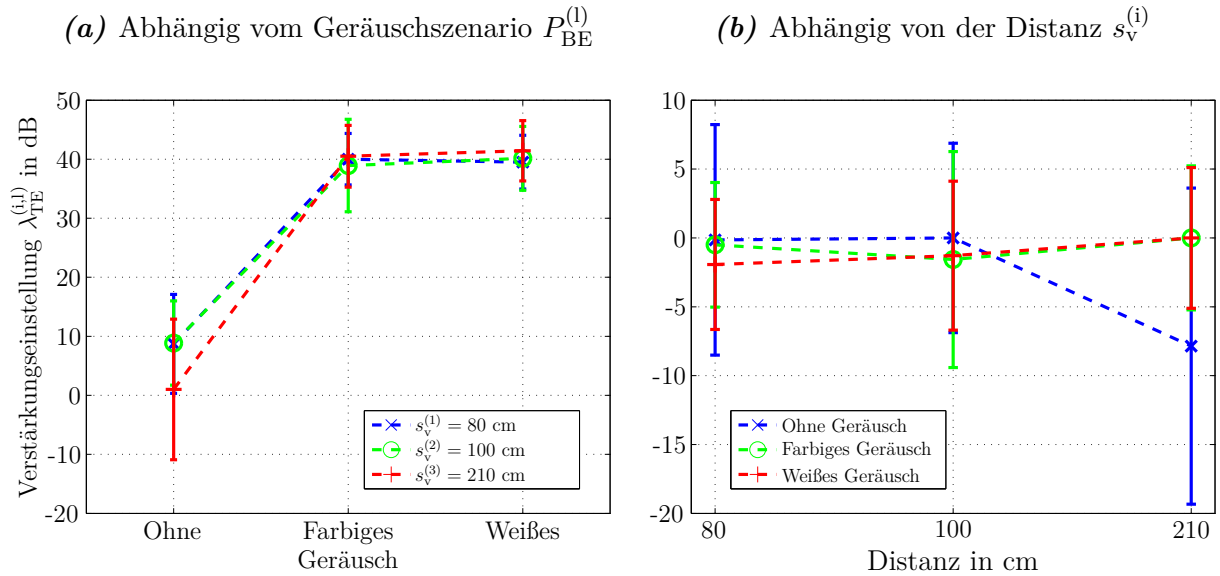


Abbildung 6.3: Mittlere Verstärkungseinstellung und zugehörige Standardabweichung über alle Testpersonen.

lich geringfügig verändert. Dies bestärkt die Einschätzung, dass bei diesen Szenarien die Lautstärke des Geräusches zu hoch gewählt worden ist. Dadurch ist das Signal des künstlichen Sprechers nahezu vollständig durch das Geräusch verdeckt worden und von der Testperson nur sehr schwer wahrgenommen worden. Um das Sprachsignal dennoch verstehen zu können, hat die Testperson die Verstärkung des zweiten Lautsprechers und somit die Leistung des Nutzsignals erhöht. Dieses Vorgehen der Testperson kann durch Aufnahmen, die während des Experiments angefertigt worden sind, auch bestätigt werden. In dem Fall des Szenarios ohne ein zusätzliches Hintergrundgeräusch verringert sich die Verstärkung mit steigender Distanz. Ein Grund für diesen Effekt könnte der Laufzeitausgleich der Signale sein, da so die beiden Schallereignisse von der Testperson synchron wahrgenommen werden. Wenn nun der künstliche Sprecher als Signalquelle erhalten werden soll, muss die Leistung des vom Torso ausgesendeten Sprachsignals am Ohr der Testperson größer sein als die vom hinteren Lautsprecher. Dies haben die Testpersonen erzeugt, indem sie die Leistung des Nutzsignals, wiedergegeben durch den hinteren Lautsprecher, reduzieren und so das Leistungsverhältnis zwischen Direktschall und zweiter Signalquelle aufrecht erhalten haben. Dadurch bleibt der Eindruck, das Signal werde vornehmlich vom Torso ausgesendet, bestehen.

Abschließend ist zu erwähnen, dass dieses Experiment zunächst nur den Fokus auf die Verstärkung des zweiten Schallereignisses gelegt hat. Wie allerdings bereits erwähnt, wird durch das ICC-System eine Verzögerung eingefügt. Aus diesem Grund sind weitere Betrachtungen in Richtung der Laufzeit und der Kombination von Laufzeit und Lautheit notwendig, um die Echoperzeption innerhalb der Fahrgastzelle genauer verstehen zu können. Trotzdem können erste Hypothesen und Erkenntnisse mittels des vorgestellten Experiments gefunden werden. Zum einen ist die Schlussfolgerung möglich, dass bei ei-

nem Anstieg des Hintergrundgeräusches von 1 dBA eine Verstärkung von circa 0,94 dB des Nutzsymbols benötigt wird, um das Sprachsignal angenehm wahrnehmen zu können. Zum anderen wird eine geringere Verstärkung des Nutzsymbols gewählt, wenn der Sprecher weiter von dem zuhörenden Passagier bzw. Testperson entfernt ist.

6.2 Eigenechoperzeption

Die Qualität eines ICC-Systems ist in erster Linie abhängig von der Qualitätsbeurteilung des zuhörenden Passagiers innerhalb der Fahrgastzelle. Dieser Passagier wird, im Vergleich zum Sprechenden Passagier auf dem Fahrersitz am stärksten durch die verschlechterte Kommunikationssituation beeinträchtigt und ist somit der Hauptgrund für den Einsatz eines ICC-Systems. Dennoch kann durch die zusätzliche Wiedergabe des Sprachsignals die Situation für den Sprechenden Passagier zusätzlich verschlechtert werden, wenn das eigene Sprachsignal zum Beispiel als starkes Echo wahrgenommen wird. Bereits bei einer Verzögerung von 50 ms zwischen dem Direktschall und einer zweiten Quelle, angeregt durch das ICC-System, können Echos wahrgenommen werden. Diese Echos werden stark störend, wenn eine Verzögerung von mehr als 100 ms zwischen beiden Quellen existiert. Bei geringeren Verzögerungen, zum Beispiel 50 ms, hängt die Schwelle, ab wann ein Echo als störend wahrgenommen wird, zusätzlich stark von der Leistung der beiden Signalquellen ab [ZF99, S. 311]. Zudem muss berücksichtigt werden, dass diese Werte innerhalb einer nicht-fahrzeugspezifischen Testanordnung, in welcher nicht der Sprecher selbst als Quelle fungiert, ermittelt worden sind. Dies erschwert die Verwendung der in der Literatur erwähnten Werte innerhalb einer ICC-Systemevaluierung. Aus diesem Grund soll durch ein psychoakustisches Experiment herausgefunden werden, in welchem Umfang der Sprecher durch die Wahrnehmung seines eigenen Sprachsignals bzw. Eigenechos gestört wird.

Innerhalb dieses Experiments wird die Wahrnehmung des Sprechers auf die Wiedergabe seines eigenen Sprachsignals während unterschiedlichster Szenarien festgehalten und analysiert. Im Detail werden die Veränderungen des Hintergrundgeräusches, des Frequenzbereiches der Wiedergabe und verschiedene Verzögerungen zwischen den beiden Nutzsignalquellen betrachtet.

In Abbildung 6.4 ist der Aufbau des Experiments und die Signalverarbeitung, welche zur Umsetzung der verschiedenen Testszenarien benötigt wird, schematisch dargestellt. Die dargestellte Signalverarbeitung wird mittels des Echtzeitrahmenwerks KiRAT umgesetzt. Als Zielumgebung für das Experiment wird eine Anordnung innerhalb einer Fahrgastzelle eines Kabrioletts betrachtet, wobei ausschließlich die vordere Sitzreihe des Fahrzeuges berücksichtigt wird. Da sich in keiner anderen Anordnung der Wiedergabelautsprecher eines ICC-Systems so nahe am Sprecher befindet, wie in dieser, stellt diese Anordnung quasi den gravierendsten Fall für die Wahrnehmung des eigenen Echos des

Einige Bestandteile des in diesem Kapitel vorgestellten Experiments, sind bereits in [TSa] veröffentlicht worden.

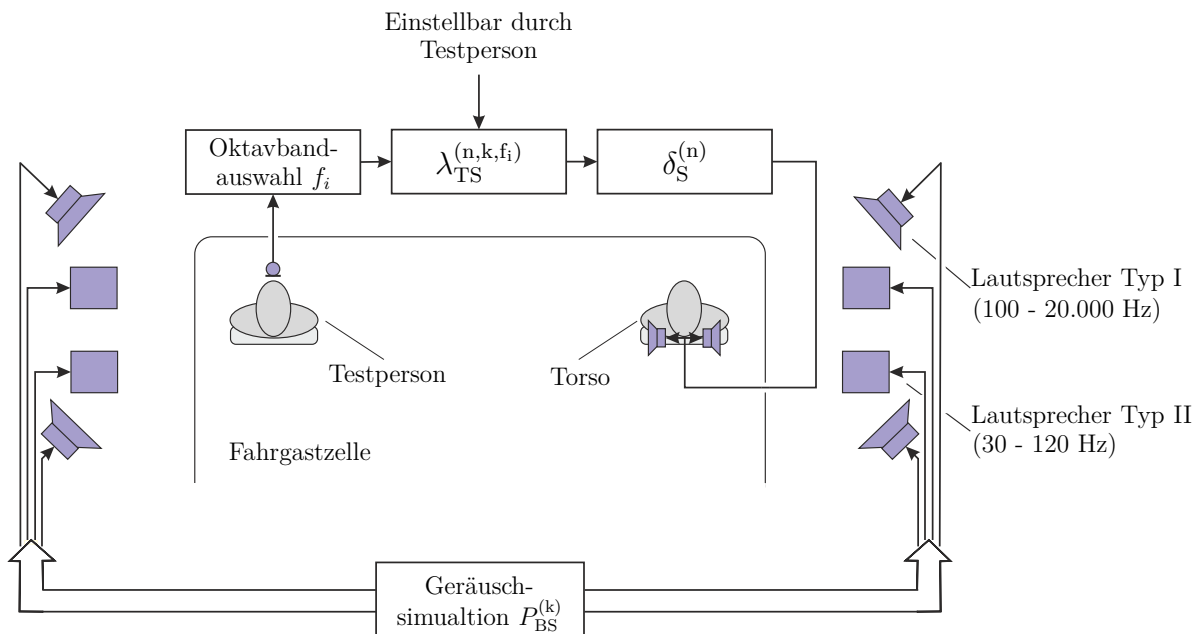


Abbildung 6.4: Schematischer Aufbau des Experiments zur Bestimmung der Eigenechoperzeption des sprechenden Passagiers. Des Weiteren ist der Signalflussgraf, welcher zur Anregung der Testsignale benötigt wird, abgebildet (nach [TSA]).

Sprechers dar. Aus diesem Grund wird die Testperson des Experiments auf der Fahrerposition und eine Zuhörerattrappe auf der Beifahrerposition platziert. Durch diese Attrappe werden auftretende Reflexionen und akustische Abschattungen einer realen Person simuliert. Die Testperson wird mit einem Headset-Mikrofon ausgestattet, um damit das Nutzsignal aufnehmen zu können. Im Anschluss können die Freiheitsgrade Verzögerung und Frequenzbereich des Ausgangssignals ausgewählt werden. Der Frequenzbereich des Ausgangssignals kann mittels einer Oktavfilterbank in 7 Bänder f_i , mit den Mittenfrequenzen 125 Hz, 250 Hz, 500 Hz, 1 kHz, 2 kHz, 4 kHz, 8 kHz, aufgeteilt werden. In Abbildung 6.5 sind die Betragsfrequenzgänge der verwendeten Oktavfilter im Detail dargestellt. Zusätzlich kann die Beschränkung des Frequenzbereiches durch die Deaktivierung der Oktavfilterbank aufgehoben werden, um so eine breitbandige Anregung durch das Nutzsignal zu ermöglichen.

Im Anschluss kann durch ein Verzögerungsmodul die Verzögerung $\delta_S^{(n)}$ entsprechend eingestellt werden. Als Verzögerungszeiten werden 0, 5, 7,5, 10, 15, 25 ms ausgewählt. Nachdem alle Anpassungen durchgeführt worden sind, wird das Sprachsignal über zwei Lautsprecher, welche sich in der Kopfstütze des Beifahrers befinden, wieder ausgegeben und auf diese Weise die Testperson mit seinem Eigenecho beschallt. Um den Freiheitsgrad des unterschiedlichen Umgebungsgeräusches innerhalb eines reflexionsarmen Raumes durchzuführen, wird eine Geräuschsimulation mit Lautsprechern, welche um das Kabriolett positioniert werden, verwendet. Damit die Testperson vollständig beschallt werden kann, werden sowohl das Verdeck als auch alle Fenster des Fahrzeuges geöffnet. Die ver-

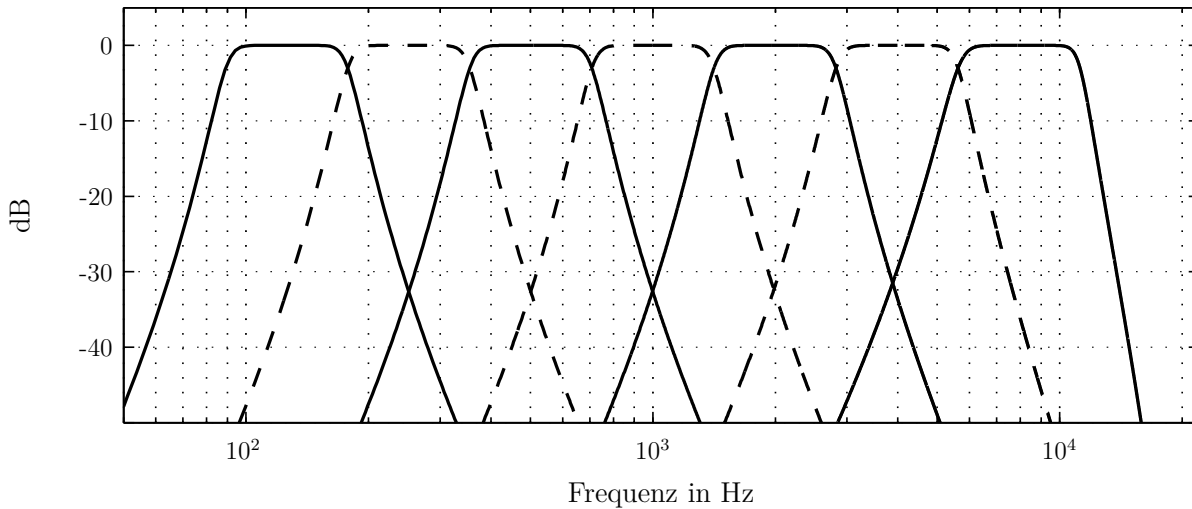


Abbildung 6.5: Betragsfrequenzgang der verwendeten Oktavfilter mit den entsprechenden Mittenfrequenzen von 125 Hz bis 8 kHz.

wendete Simulation ist bereits in Kapitel 5.1 beschrieben worden. Als Geräuschszenarien werden die in Tabelle 6.3 dargestellten Einstellungen verwendet. Die mittlere Geräuschleistung ist durch den äquivalenten Dauerschallpegel L_{eq} über eine Dauer von 30 s mit einer A-Gewichtung gemessen worden.

	Ohne Geräusch $P_{BS}^{(1)}$	Geräusch bei 50 km/h $P_{BS}^{(2)}$	Geräusch bei 100 km/h $P_{BS}^{(3)}$	Geräusch bei 150 km/h $P_{BS}^{(4)}$
L_{eq} [dBA]	31,1	67,6	71,4	79,6

Tabelle 6.3: Äquivalenter Dauerschallpegel der vier unterschiedlichen Hintergrundgeräuschszenarien in einer Fahrgastzelle, gemessen über eine Dauer von 30 s.

Die Kombination aller Freiheitsgrade würde zu einer Anzahl von 192 verschiedenen Testszenarios, die jede Testperson absolvieren müsste, führen. Unter der Annahme, dass die Beurteilung eines Testszenarios circa 1 Minute dauert, ergäbe sich eine Gesamtzeit des Experimentes für jede Testperson von 3 Stunden und 12 min. Diese Dauer ist nicht durchführbar, da eine Testperson sich wohl kaum so lange auf das Experiment konzentrieren könnte. Aus diesem Grund sind die in Tabelle 6.4 beschriebenen Testgruppen entworfen worden, welche lediglich eine Auswahl der Freiheitsgrade für eine Testperson kombinieren. Bei der Wahl der Testgruppen ist versucht worden, eine möglichst homogene Verteilung der Verzögerungen innerhalb einer Gruppe zu erzeugen, um so eine Interpolation der fehlenden Werte zu ermöglichen.

Test- gruppe	Verzögerung	Frequenzbereich	Hintergrundgeräusch
1	0, 7,5 und 15 ms	breitbandige Anregung	Ohne Geräusch, bei 50, 100 und 150 km/h
2	5, 10 und 25 ms	breitbandige Anregung	Ohne Geräusch, bei 50, 100 und 150 km/h
3	0, 7,5 und 15 ms	Jeweils Oktavband 1-7	Ohne Geräusch und bei 50 km/h
4	5, 10 und 25 ms	Jeweils Oktavband 1-7	Ohne Geräusch und bei 100 km/h
5	0, 5 und 7.5 ms	Jeweils Oktavband 1-7	Ohne Geräusch und bei 150 km/h

Tabelle 6.4: Mittlere Geräuschleistungen der vier unterschiedlichen Hintergrundgeräuschszenarien (nach [T_{Sa}]).

Die Testpersonen bekommen die Aufgabe, innerhalb eines jeden Testszenarios drei verschiedene Schwellen durch Einstellen der Verstärkung $\lambda_{\text{TS}}^{(n,k,f_i)}$ zu bestimmen. Die Schwellen sind in Anlehnung an die *detectability opinion scale* aus Anhang C von [ITU96a] definiert worden:

1. Die Wiedergabe wird gerade noch wahrgenommen.
2. Die Wiedergabe wird gerade noch als nicht störend wahrgenommen.
3. Die Wiedergabe wird definitiv als störend wahrgenommen.

Zur Erzeugung des Sprachsignals ist es den Testpersonen freigestellt, einen freien Text zu rezitieren oder aus einem bereitgestellten Buch vorzulesen. Entscheidend war lediglich, dass der Text so reproduziert wird, dass ein fiktiver Zuhörer bzw. die Zuhörerattrappe auf dem Beifahrersitz den Inhalt verstehen kann. Zusätzlich bekommen alle Testpersonen in einer Testphase die Chance, sich mit dem Experiment vertraut zu machen. Dazu können sie in einem Szenario ohne Geräusch frei die Verstärkung mal sehr hoch und mal sehr gering einstellen und die Auswirkungen entsprechend wahrnehmen.

Die Gruppe aller Testpersonen bestand aus 10 Probanden, darunter 2 Frauen und 8 Männer. Das Alter der Probanden lag zwischen 24 und 26 Jahren. Zum Zeitpunkt des psychoakustischen Experiments war bei keiner der Testpersonen eine Einschränkung der Hörfähigkeit bekannt. Fünf der zehn Testpersonen können als naiv angesehen werden. Bei ihnen hat es sich um Studierenden des Lehrstuhls DSS gehandelt. Die anderen fünf Testpersonen sind als nicht naiv einzustufen, da es sich um Doktoranden des Lehrstuhls handelt, welche sich mit Themen der Sprach- und Audiosignalverarbeitung beruflich beschäftigen.

6.2.1 Ergebnisse

Um die durch das Experiment erhobenen Daten vollständig analysieren zu können, ist zu berücksichtigen, dass neben der Verzögerung durch das entsprechende Modul in Abbildung 6.4, auch das Rahmenwerk selbst sowie die verwendeten Apparaturen eine gewisse Verzögerung des Sprachsignals erzeugen. Diese Verzögerung wird einmalig bestimmt und kann danach für die Auswertung der Daten genutzt werden. Zur Bestimmung dieser Verzögerung wird zunächst der Aufbau aus Abbildung 6.4 um zwei Mikrofone und einen künstlichen Sprecher erweitert. Der zusätzliche künstliche Sprecher, das Headsetmikrofon tragend, wird auf der Fahrerposition platziert und eines der beiden Mikrofone vor dessen Mund angebracht. Das zweite Mikrofon wird vor einem der beiden Kopfstützenlautsprecher positioniert. Im Anschluss wird ein Rauschsignal über den künstlichen Sprecher ausgegeben und mithilfe der zusätzlichen Mikrofone werden zwei Signale aufgezeichnet. Diese Signale werden durch eine Kreuzkorrelationsanalyse ausgewertet, wobei im Detail die Kreuzkorrelationsfolge geschätzt und der Zeitpunkt des Maximums dieser Folge bestimmt wird. Die so bestimmte Zeit entspricht der durch das Rahmenwerk eingefügten Verzögerung. In diesem speziellen Fall ist eine Verzögerung von 7.1 ms ermittelt worden, die auf die eingestellte Verzögerung aufaddiert werden muss.

Zusätzlich muss, um vergleichbare Ergebnisse des Experiments zu erzielen, für jede Testperson eine Kalibrierung der Ausgabe des Sprachsignals in seiner breitbandigen Leistung vorgenommen werden. Dazu wird in Stille der Sprachpegel der Testperson in einem Abstand von 10 cm mittels eines Pegelmessers gemessen. Dieser Pegelmesser bestimmt über eine Dauer von 20 ms den mittleren äquivalenten Dauerschallpegel in dBA. Zusätzlich wird das gemessene Sprachsignal, unter Verwendung des Headsetmikrofons, aufgezeichnet und im Anschluss über die Kopfstützenlautsprecher, siehe Abbildung 6.4, wieder ausgegeben. In einem Abstand von 10 cm zum Lautsprecher wird erneut der mittlere äquivalente Dauerschallpegel in dBA bestimmt. Die Differenz, der so bestimmten mittleren äquivalenten Dauerschallpegel, wird zur Kalibrierung des Lautsprechersignals verwendet. Dazu werden die gewählten Verstärkungseinstellungen entsprechend auf diesen Differenzwert normiert.

Die beiden Anpassungen in der Verzögerung und der Leistung der Daten ermöglichen den Vergleich der erzielten Ergebnisse zwischen den einzelnen Testpersonen. Da nicht alle Probanden alle Verzögerungen während des Tests durchlaufen haben, werden diese Daten mithilfe einer kubischen Interpolation zwischen den erhobenen Datenpunkten erweitert. Die so interpolierten Daten der einzelnen Testpersonen werden im Anschluss in ihren Mittelwerten und Standardabweichungen analysiert. Durch diese Interpolation der Daten können erste Tendenzen innerhalb dieses speziell auf eine Fahrzeugumgebung entworfenen Experiments abgelesen und interpretiert werden.

Zunächst soll die Veränderung der Verstärkungseinstellung über die einzelnen Geräuschenarien analysiert werden. Dazu werden lediglich die erste und die zweite Testgruppe aus Tabelle 6.4 für eine Analyse herangezogen. In Abbildung 6.6(a) ist das Ergeb-

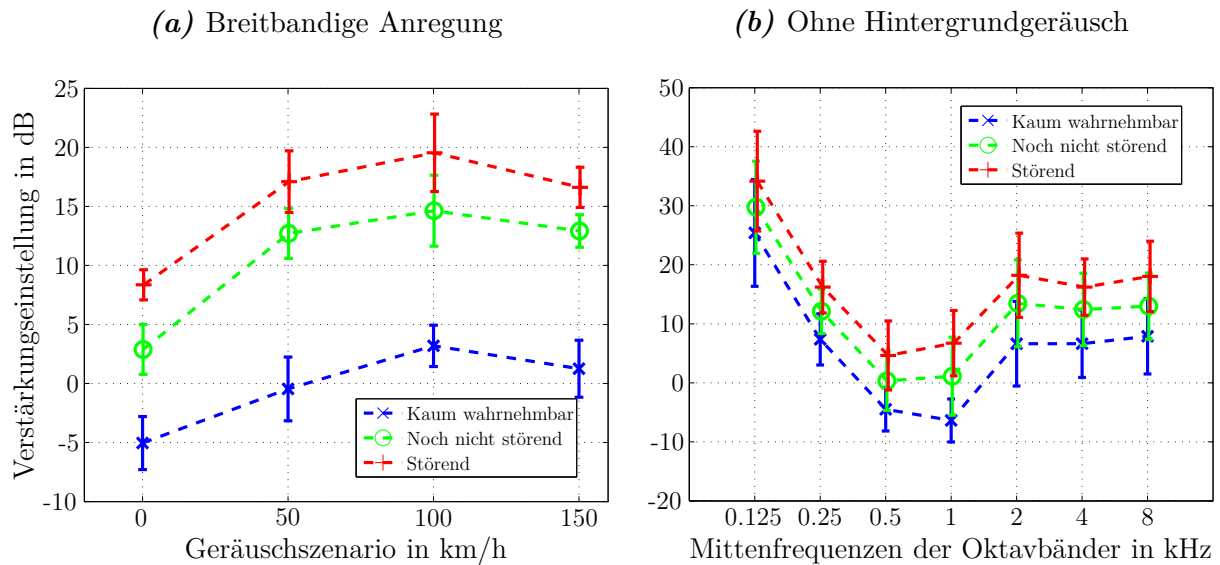


Abbildung 6.6: Mittelwert und Standardabweichung der Verstärkungseinstellung bei einer Gesamtverzögerung von 17.1 ms.

nis für eine Gesamtverzögerung von 17.1 ms bei einer breitbandigen Anregung durch das Sprachsignal der Testperson in Abhängigkeit der betrachteten Geräuschszenarien dargestellt. Die Markierungen in den Kurven entsprechen den Mittelwerten der Verstärkungseinstellung $\lambda_{\text{TS}}^{(n,k,f_i)}$ und die zugehörigen vertikalen Balken der Standardabweichung über alle Testpersonen. Die rote Kurve steht für die Schwelle, bei welcher die Wiedergabe definitiv als störend wahrgenommen wird. Die grüne zeigt die Schwelle an, bei der die Wiedergabe gerade noch nicht als störend wahrgenommen wird und der blaue Graph steht für eine leise, kaum noch wahrgenommen Wiedergabe des Sprachsignals. Diese Kurven sind in Abbildung 6.6(a) über das ansteigende Hintergrundgeräusch aufgetragen. Zusätzlich ist zu erkennen, dass die Kurven von blau über grün nach rot bei einem Anstieg der Verstärkungseinstellung verteilt sind. Dies ist kein verwunderliches Ergebnis, sondern eher ein Indikator für die Plausibilität der Daten, da nur ein vorhandenes Signal wahrgenommen und gegebenenfalls als störend empfunden werden kann. Somit ist die Wahrnehmungsschwelle eine notwendige Bedingung, um ein Signal als störend wahrzunehmen. Zusätzlich steigt mit der Leistung des Hintergrundgeräusches von 0 auf 100 km/h die Verstärkungseinstellung aller drei Kurven. Dies ist zunächst intuitiv und schlüssig, da das Sprachsignal durch das auftretende Hintergrundgeräusch maskiert wird und so mehr Sprachsignalleistung benötigt wird, um das Sprachsignal von dem Geräusch abgrenzen zu können. Eine Ausnahme ist dennoch in Abbildung 6.6(a) zu beobachten. Im letzten Szenario, einem Fahrzeuggeräusch bei 150 km/h, fallen die Verstärkungseinstellungen aller drei Kurven leicht ab. Dieser Abfall der Einstellung der Verstärkung durch die Testpersonen kann möglicherweise durch den sogenannten Lombard-Effekt begründet werden. Wenn das Sprachsignal, aufgenommen durch das Headsetmikrofon, bereits eine höhere Leistung als im leiseren Geräuschszenario aufweist, wird im Rückschluss eine geringere Verstär-

kung benötigt, um das Eigenecho wahrnehmen zu können. Der Lombard-Effekt könnte ebenso der Grund für die nichtlineare Veränderung der Verstärkungseinstellung bei einem Anstieg des Hintergrundgeräusches sein. Dennoch bestätigt dieses Ergebnis die Notwendigkeit einer geräuschabhängigen Verstärkungseinstellung (NDGC) durch das ICC-System. Zusätzlich lässt dieses Ergebnis Rückschlüsse auf einen möglichen Verlauf der NDGC zu, da in keinem Fall der Sprecher die Wiedergabe seines eigenen Sprachsignals als störend wahrnehmen sollte.

In Abbildung 6.6(b) sind die Mittelwerte und Standardabweichungen der Verstärkungseinstellung aller Testpersonen der Testgruppe 3-5 in Abhängigkeit der Frequenzbänder bei einer Gesamtverzögerung von 17.1 ms und ohne Hintergrundgeräusch dargestellt. Auch hier sind die Kurven von blau über grün nach rot bei ansteigender Verstärkungseinstellung verteilt. Dies ist, wie bereits erörtert, ein plausibles Ergebnis. Zusätzlich fällt die Verstärkungseinstellung innerhalb der ersten drei Bänder, stagnieren anschließend innerhalb des dritten und vierten Bandes bevor sie in Band fünf erneut ansteigen und sich innerhalb der letzten zwei Bänder kaum verändern. Es entsteht eine nach unten gewölbte Kurve, die stark, an die Hörkurve in [ZF99] erinnert. Der Abfall der Verstärkungseinstellung innerhalb der ersten drei Bänder liegt an der Verteilung der Sprachleistung über die verschiedenen Bänder. Nach [ZF99, S. 17 ff.] liegt der Anregungsbereich von Sprache in einem Bereich von 0,1 bis nahe 7 kHz. Das bedeutet, dass innerhalb des ersten Bandes mit einer Mittenfrequenz von 125 Hz nur wenig Sprache angeregt wird, was wiederum dazu führt, dass zur Wahrnehmung des Sprachsignals eine höhere Verstärkung benötigt wird als im nächsthöheren Band. Zusätzlich liegt innerhalb des dritten und vierten Bandes (0,5 und 2 kHz Mittenfrequenz) die Hauptenergie des Sprachsignals. Dadurch wird eine geringere Verstärkungseinstellung durch die Testpersonen ausgewählt als in den beiden untersten Bändern.

Der nächste Freiheitsgrad, welcher genauer untersucht wird, ist die Verzögerung der Echoquelle. Dazu werden die Ergebnisse des Szenarios ohne Hintergrundgeräusch der ersten zwei Testgruppen verwendet und zusätzlich mittelwertbereinigt. Das Ergebnis ist in Abbildung 6.7(a) dargestellt. Alle drei Kurven besitzen eine leicht nach oben gekrümmte Form, deren Maximum zwischen 12 und 18 ms liegt. Dieser spezielle Verlauf lässt sich auch innerhalb der einzelnen Oktavbänder, welche mit einer ausreichend Leistung des Sprachsignals angeregt worden sind, beobachten. Dazu sind beispielhaft die Ergebnisse der Testgruppen drei bis fünf ohne Hintergrundgeräusch für das vierte Oktavband mit einer Mittenfrequenz von 1 kHz in Abbildung 6.7(b) dargestellt. Auch hier ist eine Mittelwertbereinigung aller Kurven durchgeführt worden, um so den Fokus auf den Verlauf der Kurven zu legen.

Abschließend ist festzuhalten, dass sich bereits während der Durchführung des Experimentes herauskristallisierte, dass die Wahrnehmung des Eigenechos ein sehr individuelles Phänomen ist. Dies lässt sich zudem an den teilweise sehr großen Standardabweichungen der Ergebnisse ablesen. Natürlich ist dies keine verwunderliche Feststellung, da jeder

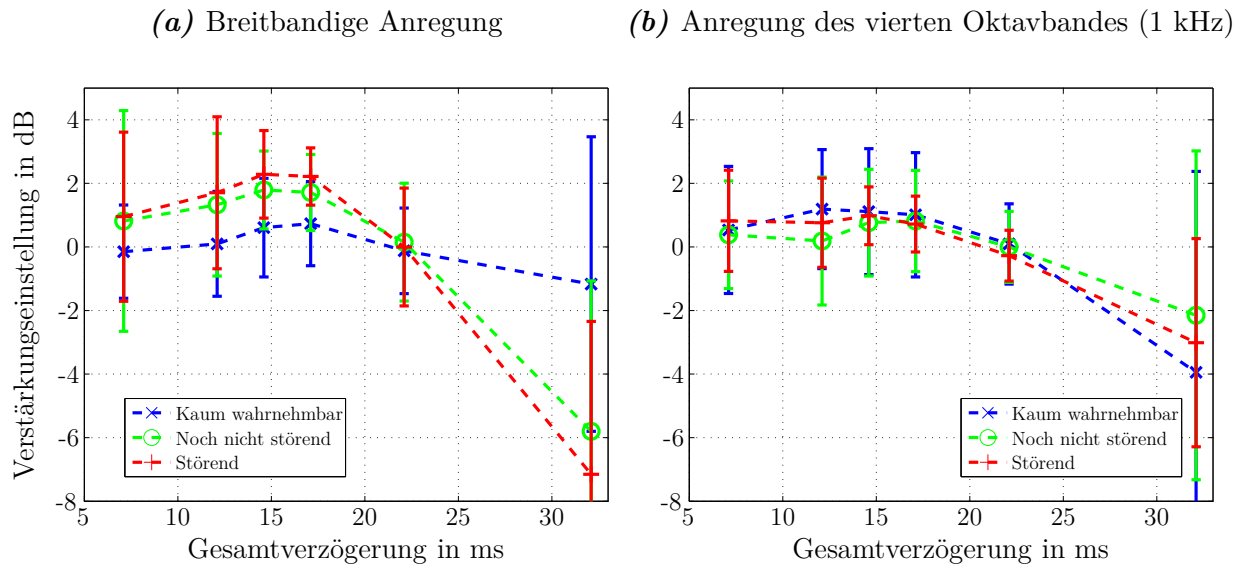


Abbildung 6.7: Mittelwert und Standardabweichung der Verstärkungseinstellung in Abhängigkeit der Gesamtverzögerung ohne die Einspeisung eines Hintergrundgeräusches. Alle Kurven sind von ihrem Mittelwert bereinigt.

Mensch unterschiedliche Erwartungen und physiologische Voraussetzungen mit in das Experiment einbringt. Dies ist auch der Grund für die kritische Betrachtung der hier dargestellten Ergebnisse. Da lediglich eine kleine Gruppe von Testpersonen betrachtet worden sind, ermöglichen diese Ergebnisse eine erste Hypothese. Diese gilt es in weiteren Betrachtungen zu konsolidieren. Ein für die Evaluierung eines ICC-Systems essentielles Fazit dieses Experiments ist, dass innerhalb einer Verzögerung zwischen 12 und 18 ms höhere Verstärkungseinstellungen verwendet werden können als bei anderen Verzögerungen. Zusätzlich sollte die Verstärkung des Nutzsignals in Verbindung mit einem zunehmenden Hintergrundgeräusch ansteigen, wobei der Grad des Anstiegs in Abhängigkeit von der Leistung des Hintergrundgeräusches gewählt werden sollte. Das ICC-System sollte, bezogen auf die Evaluierung, optimalerweise eine Verzögerung zwischen 12 und 18 ms aufweisen. In diesem Bereich können, im Vergleich zu anderen Verzögerungen, höhere Verstärkungsfaktoren verwendet werden ohne dass der Sprecher dadurch gestört werden würde.

6.3 Diskrepanz zwischen SNR und Qualitätswahrnehmung

Ein klassisches und dennoch umstrittenes Maß zur Erhebung der Qualität von sprachverarbeitenden System ist der SNR. Der klassische Ansatz, die Rauschleistung ins Verhältnis zur Signalleistung zu setzen, ist bereits in einigen Publikation als ein eher ungeeignetes Maß zur Bestimmung der Sprachqualität beschrieben worden [GTHS, QBC88, MSG78].

Weiterentwicklungen des SNR sind zum Beispiel in [Loi13, LS] veröffentlicht und haben dort bereits erste vielversprechende Resultate geliefert. Der hier beschriebene Hörtest, soll die Diskrepanz zwischen der durch die Testpersonen empfundenen Qualität und einem klassisch bestimmten SNR feststellen. Zu diesem Zweck werden die Qualitätsempfindungen verschiedener Testpersonen bei unterschiedlichen Geräuschen erhoben und mit dem SNR verglichen.

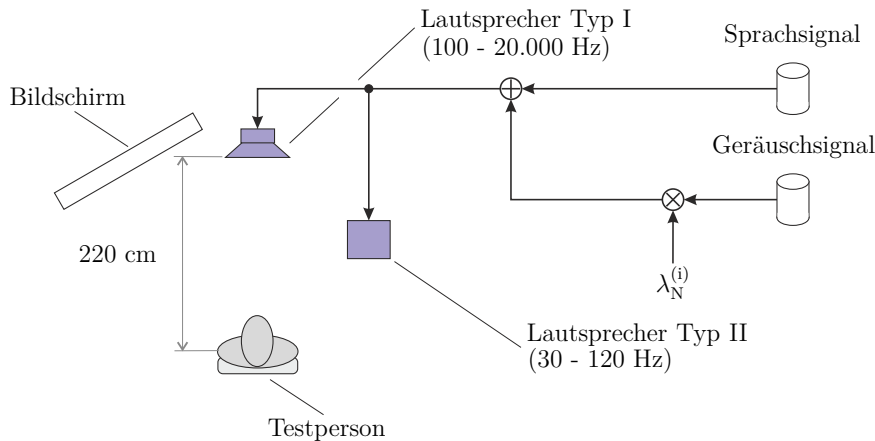


Abbildung 6.8: Schematischer Aufbau des Hörtests zur Bestimmung der Diskrepanz zwischen SNR und Qualitätswahrnehmung.

In Abbildung 6.8 ist schematisch der Aufbau des Hörtests dargestellt. Im Zentrum des Experimentes steht die Testperson, welche auf einer festen Position in einem Abstand von 220 cm zum Lautsprecher des Typs I positioniert ist. Neben diesem Lautsprecher wird ein zweiter Lautsprecher des Typs II, es handelt sich dabei um einen Subwoofer, verwendet. Dadurch können auch Signale mit einer hohen Leistungsdichte in den tiefen Frequenzen, wie zum Beispiel bei einem Fahrzeuggeräusch (siehe Abbildung 5.3(a)), angeregt werden. Im Hörtest werden zur Bestimmung der Diskrepanz zwischen Hörwahrnehmung und SNR zwei verschiedene Geräusche, ein weißes Rauschen und ein Fahrzeuggeräusch, verwendet. Diese beiden Signale werden abhängig von dem betrachteten Szenario bzw. dem betrachteten SNR mit $\lambda_N^{(i)}$ verstärkt, anschließend ein klares Sprachsignal aufaddiert und über die Lautsprecher ausgegeben. Diese einzelnen Szenarien werden den Testpersonen dargeboten und eine Skala zur Einstufung der Qualität bereitgestellt. Bei der Skala handelt es sich um die von der ITU empfohlene fünfstufige Skala, die Abstufungen zwischen schlecht und ausgezeichnet vorsieht. Die Testperson wählt über ein mobiles Eingabegerät die Stufe, die ihr auf einer Benutzeroberfläche auf einem Bildschirm gezeigt wird, aus. An dem Hörtest haben insgesamt 20 Testpersonen, 7 weibliche und 13 männliche, teilgenommen. Das Alter der Probanden lag zwischen 23 und 42 Jahren. Es handelt sich sowohl um naive Testpersonen als auch um Personen, welche sich mit Audio- bzw. Sprachsignalverarbeitung beruflich beschäftigen.

6.3.1 Ergebnisse

In Abbildung 6.9 sind die Ergebnisse des Hörtestes in Abhängigkeit des verwendeten Geräuschsignals und des breitbandig eingestellten SNR dargestellt. Die Kreise entsprechen den resultierenden MOS-Werten der Sprachqualität und die vertikalen Balken der zugehörigen Standardabweichung über alle Testpersonen. An dieser Stelle ist bereits die Diskrepanz zwischen dem SNR und der Einschätzung der Testpersonen zu sehen. Während bei einem Fahrzeuggeräusch schon bei einem SNR von -5 dB ein ordentliches Qualitätsurteil gegeben worden ist, wird diese Qualitätsstufe bei einem weißen Geräusch erst bei einem SNR von 10 dB erreicht.

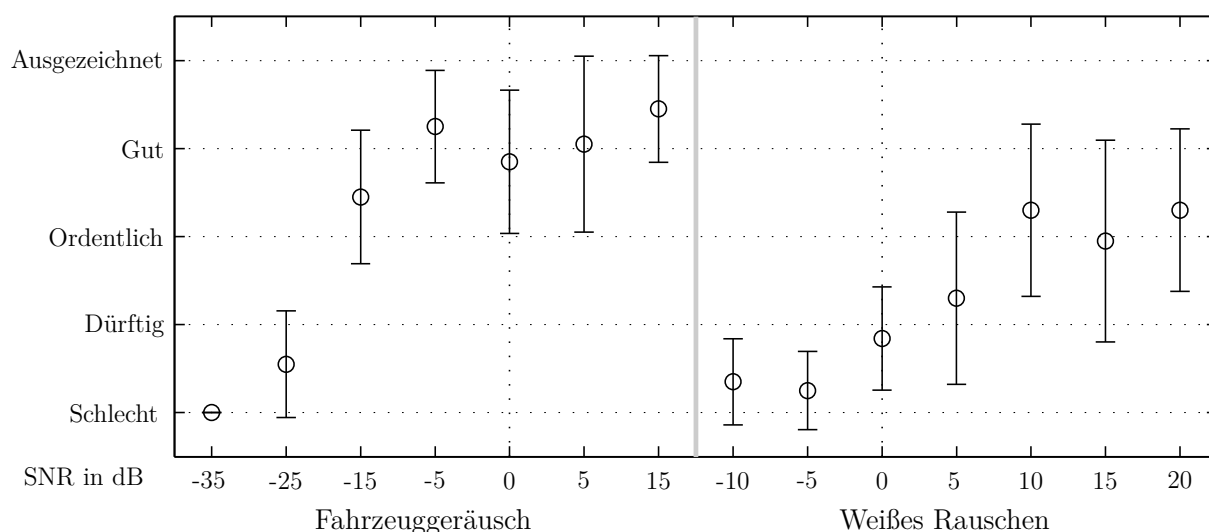


Abbildung 6.9: Dargestellt ist die mittlere Beurteilungsnote (engl. Mean Opinion Score, MOS) und die zugehörige Standardabweichung über alle Testpersonen in Abhängigkeit des untersuchten SNR. Das entsprechende Szenario ist erzeugt worden durch ein Fahrzeuggeräusch einer Mercedes-Benz E-Klasse (links) und ein weißes Rauschen (rechts).

Dieses Phänomen entsteht aufgrund der spektralen Leistungsverteilung beider Geräuschsignale und des Sprachsignals. In Abbildung 6.10 sind für ein fest vorgegebenes SNR von -15 dB sowohl das mittlere Leistungsdichtespektrum des Sprachsignals einer weiblichen Sprecherin als auch die verwendeten Geräuschsignale beispielhaft dargestellt. Das Fahrzeuggeräusch ist ein tieffrequent dominantes Signal. Dies führt dazu, dass ein Großteil der Geräuschleistung innerhalb eines Frequenzbereiches zwischen 0 und 100 Hz liegt, in dem kaum Sprachleistung vorhanden ist. Dadurch entsteht in dem für Sprache relevanten Frequenzbereich, dieser liegt nach [ZF99, S. 17 ff.] zwischen 0,1 und circa 7 kHz, ein deutlich höherer, frequenzabhängiger SNR. Bei einem weißen Geräusch wird das Sprachsignal aus Abbildung 6.10 vollständig vom Geräusch überlagert und so ein sehr geringer, frequenzabhängiger SNR erzielt. Folglich lässt sich das Sprachsignal nur schlecht verstehen. Diese führt zur Hypothese, dass bei einer Evaluierung der Sprachqua-

lität anhand eines SNR-Ansatzes, die spektralen Eigenschaften des Geräuschsignals mit berücksichtigt werden müssen.

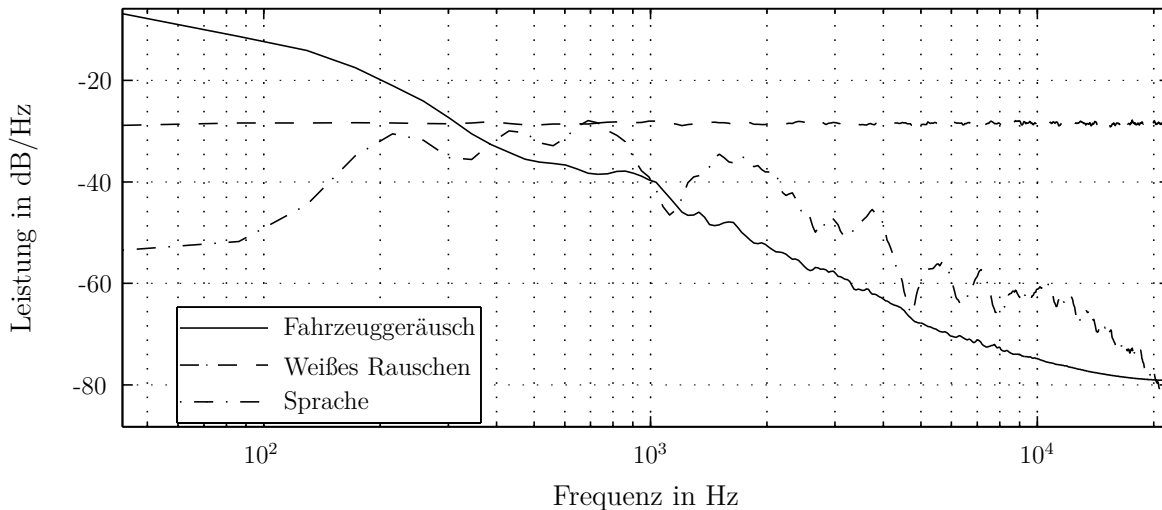


Abbildung 6.10: Mittleres Leistungsdichtespektrum des verwendeten Fahrzeuggeräusch einer Mercedes-Benz E-Klasse, des weißen Rauschen und des Sprachsignals einer weiblichen Sprecherin. Sprache und Geräuschsignale sind in ihrer Leistung so angepasst worden, sodass sich für das weiße Rauschen und das Fahrzeuggeräusch ein SNR in Bezug auf das Sprachsignal von -15 dB ergibt.

Abschließend ist festzuhalten, dass ein konventionell bestimmter SNR nicht zur Schätzung der Sprachqualität verwendet werden sollte. Die Diskrepanz zwischen dem SNR und der Wahrnehmung ist mit einem ähnlichen psychoakustischen Experiment bereits in [GTHS] diskutiert worden, wobei dieses Experiment den quantitativen Unterschied zwischen dem wahrgenommenen und dem realen SNR bestimmt hat. Auch hier konnte festgestellt werden, dass das Qualitätsempfinden einer Testperson und eine klassische SNR-Bestimmung voneinander abweichen, wenn sich der spektrale Verlauf des betrachteten Geräusches verändert. Neben dieser Analyse zwischen dem subjektiv empfundenen SNR und dem berechneten SNR, sind in [GTHS] weitere vielversprechende, instrumentelle Verfahren zur Evaluierung der Sprachqualität auf der Basis eines SNR publiziert worden. Diese Verfahren haben auch innerhalb einer Fahrzeugumgebung aussichtsreiche Ergebnisse erzielt, sodass sie ebenso zur Evaluierung von ICC-Systemen geeignet sein könnten. Die Erhebung eines dieser modifizierten SNRs in Bezug auf die Evaluierung eines ICC-Systems wird in Kapitel 7.3.3 genauer beschrieben.

6.4 Gesamtqualitätsevaluierung

Um eine Qualitätsreferenz für den Entwurf von instrumentellen Verfahren speziell zur Evaluierung von ICC-Systemen zu erhalten, muss ein Hörtest durchgeführt werden, welcher eine



Please listen to the audiofiles and rate the signal-to-noise ratio of each example by the quality scale on the left and the corresponding slider.

With the 'next'-button you can access the next test.

Quality Scale	38	59	10	83	54	84	18	62
Excellent								
Good								
Fair								
Poor								
Bad								
0								
Ref	1	2	3	4	5	6	7	8

Comment

Add Comment

<< back Test 2/3 next >>

Abbildung 6.11: Benutzeroberfläche des MUSHRA-Tests.

Qualitätsreferenz, zum Beispiel mithilfe des MOS, für verschiedene Szenarien und/oder Systeme bereitstellt. Zu diesem Zweck gibt es verschiedene, von der ITU standardisierte Verfahren. Innerhalb der hier vorgestellten ICC-Systemevaluierung wird ein sogenannter *Multi Stimulus test with Hidden Reference and Anchor* (MUSHRA) verwendet.

Der MUSHRA-Test ist in [ITU14] detailliert beschrieben und ist als Empfehlung von der ITU zur auditiven bzw. subjektiven Evaluierung von Audiosystemen mit mittlerem Qualitätsniveau veröffentlicht worden. Die Grundidee dieses Hörtests ist es, der Testperson mehrere Signale zum Vergleich bereitzustellen, die innerhalb einer fortlaufenden Qualitätsskala (engl. *continuous quality scale*) von 0 (schlecht) bis 100 (ausgezeichnet) beurteilt werden können. Dazu wird jedem Stimuli, siehe Abbildung 6.11, ein eigener Regler sowie eine eigene Schaltfläche zum Start der Wiedergabe zugewiesen. Neben diesen zu evaluierenden Signalen, bekommt die Testperson des Weiteren ein Referenzsignal als Stimulus, welches dem Fall einer ausgezeichneten Qualität entsprechen soll. Auch dieses Signal kann

mit einer entsprechenden Schaltfläche jederzeit abgespielt werden. Zur Beurteilung sollen der Testperson, nach der ITU-Empfehlung, maximal 12 Stimuli zur Evaluierung dargeboten werden. Alle Signale können immer wieder angehört und erneut abgespielt werden. Die maximal 12 Audiosignale beinhalten neben dem versteckten Referenzsignal bzw. *hidden reference*, einen versteckten niedrigen und einen versteckten mittleren Anker. Der niedrige Anker und das Referenzsignal spannen den Qualitätsraum zwischen „schlecht“ und „ausgezeichnet“ auf. Der niedrige Anker ist innerhalb der Empfehlung [ITU14] eindeutig definiert, als Originalsignal welches mithilfe eines Tiefpasses mit einer Grenzfrequenz von 3,5 kHz gefiltert worden ist. Der mittlere Anker entspricht dem tiefpassgefilterten Originalsignal mit einer Grenzfrequenz von 7 kHz. Diese klare Definition ermöglicht es, die Ergebnisse verschiedener Hörtest miteinander zu kombinieren und Daten auszutauschen. Für die Evaluierung eines ICC-Systems sind bisher noch keine Definitionen für die beiden Anker veröffentlicht worden. Sie müssen daher neu bestimmt werden. Bevor der eigentliche Hörtest beginnt, haben alle Testpersonen die Möglichkeit sich in einer Testphase mit dem Versuch vertraut zu machen. Diese Testphase wird nicht in die spätere Auswertung mit aufgenommen, sondern dient lediglich zum Training der Testpersonen.

Der Vorteil dieses Verfahrens ist, neben dem direkten Vergleich unterschiedlicher Systeme mit einer Referenz und somit der teilweisen Kompensation der internen Referenz einer Testperson, die Zuverlässigkeit der erhobenen Ergebnisse. Denn anhand der Anker und der versteckten Referenz kann die Zuverlässigkeit einer jeden Testperson bestimmt und diese unter Umständen aus dem Gesamtergebnis ausgeschlossen werden. So ist, zum Beispiel, in der Empfehlung [ITU14] festgelegt, dass Testpersonen, welche die Referenz in 15% der Fälle mit einem Wert geringer als 90 bewerten, von der Auswertung auszuschließen sind. Auch wenn der mittlere Anker in 15% der Fälle mit einem Wert größer 90 bewertet worden ist, ist diese Testperson von der Auswertung auszuschließen.

Auf Basis dieses MUSHRA-Tests soll ein Hörtest konzipiert werden, welcher die Sprachqualität eines ICC-Systems erhebt. Zunächst werden dazu binaurale Signale benötigt, welche den Testpersonen zum Vergleich vorgespielt werden. Dazu wird als Fahrzeug eine Mercedes-Benz S-Klasse verwendet, in der zunächst alle Impulsantworten und Geräusche aufgezeichnet worden sind. Innerhalb der Fahrgastzellen sollen zwei Passagiere, ein Fahrer und ein Passagier auf der Rückbank hinter dem Fahrer, positioniert werden. Dies entspricht einer Anordnung wie sie in Abbildung 5.1 dargestellt ist. Auch die zur Simulation notwendigen Impulsantworten ($h_{TM}(n)$, $h_{TE}(n)$, $h_{LE}(n)$ und $h_{LM}(n)$) sind dort beispielhaft dargestellt. Das dort dargestellte ICC-System entspricht lediglich einem zur Vereinfachung reduzierten System, siehe Kapitel 5. Für die vollständige Simulation eines ICC-Systems werden fünf Lautsprecher als Ausgang des ICC-Systems und ein Mikrofonsignal als Eingang verwendet. Bei den Lautsprechern handelt es sich um die beiden Tür Lautsprecher in den hinteren Türen und drei Lautsprecher auf der Hutablage. Als Eingang wird eines der am Gurt positionierten Mikrofone verwendet. Zusätzlich werden zwei Mikrofone an den Ohren des Zuhörers positioniert, um so binaurale Signale mithilfe der

zugehörigen Impulsantworten zu simulieren. Die Geräuschsignale, aufgezeichnet bei einer Geschwindigkeit von 100 km/h, und die Impulsantworten werden zusammen mit dem Echtzeitrahmenwerk KiRAT des Lehrstuhls DSS verwendet, um das akustische Übertragungssystem innerhalb der Fahrgastzelle zu simulieren. Zusätzlich wird innerhalb dieses Rahmenwerks auch das am Lehrstuhl DSS entwickelte ICC-System aktiviert. Dieses wird durch die akustische Systemsimulation mit den entsprechenden Signalen gespeist, und die entsprechenden Ausgangssignale werden der Systemsimulation bereitgestellt. Im Anschluss können die binauralen Signale unter Verwendung des Rahmenwerks aufgezeichnet werden. Als Szenarien wird eine Situation ohne Hintergrundgeräusch ausgewählt, eine mit dem aufgezeichnetem Hintergrundgeräusch bei einer Geschwindigkeit von 100 km/h und einem leisen Sprecher, mit einem um 5 dB lauteren Sprecher und einem Sprecher, der nochmals um 5 dB verstärkt wird. Alle vier Szenarien werden mit und ohne ICC-System und für einen weiblichen und einen männlichen Sprecher simuliert. Auf diese Weise ergeben sich 16 verschiedene Stimuli, welche in zwei Tests aufgeteilt werden. Dabei werden die Signale des weiblichen Sprechers und die des männlichen Sprechers getrennt von einander verglichen. Die zugehörige Benutzeroberfläche eines solchen Tests ist in Abbildung 6.11 dargestellt. Als Referenzsignal ist das Szenario ohne Hintergrundgeräusch und deaktiviertem ICC-System ausgewählt worden, welches einem SNR von circa 11 dB am Ohr des Zuhörers entspricht. Für den Hörtest wird nur ein unterer Anker definiert. Hierfür wird ein leiser Sprecher innerhalb des 100 km/h Geräuschszenarios ausgewählt, was einem SNR am Ohr des Zuhörers von circa -40 dB entspricht. Diese Auswahl von Referenz- und Anker-Stimuli ist ein wichtiger Punkt innerhalb der Evaluierung mittels des MUSHRA-Tests, welcher in jedem Fall kritisch betrachtet und diskutiert werden muss. An dieser Stelle wird zunächst ein erster Vorschlag zur Wahl dieser Stimuli gegeben und zur Diskussion dargeboten.

Der Hörtest wird unter Verwendung von Kopfhörern durchgeführt. Es haben insgesamt 11 Testpersonen, 4 weibliche und 6 männliche Probanden, im Alter zwischen 25 und 31 Jahren teilgenommen. Zum Zeitpunkt des Hörtests war bei keiner der Testpersonen eine Einschränkung der Hörfähigkeit bekannt. Bei den Probanden handelt es sich sowohl um naive als auch um Personen, welche sich beruflich mit Audio- bzw. Sprachsignalverarbeitung beschäftigen.

6.4.1 Ergebnisse

Bevor die Daten des Hörtests ausgewertet werden, wird die Zuverlässigkeit der Testpersonen anhand des Referenzsignals analysiert. Innerhalb dieses Hörtests haben alle Testpersonen alle Referenzsignale mit einem Wert größer 90 bewertet und werden somit als zuverlässige Testpersonen in der Analyse der Ergebnisse berücksichtigt. In Abbildung 6.12 sind die Mittelwerte der Qualitätsbewertung über alle Testpersonen entsprechend als Kreise dargestellt. Im weiteren Verlauf werden diese Werte als MOS_S bezeichnet, auch wenn diese Werte nicht der klassischen fünfstufigen MOS-Skala entsprechen. Die vertikalen Bal-

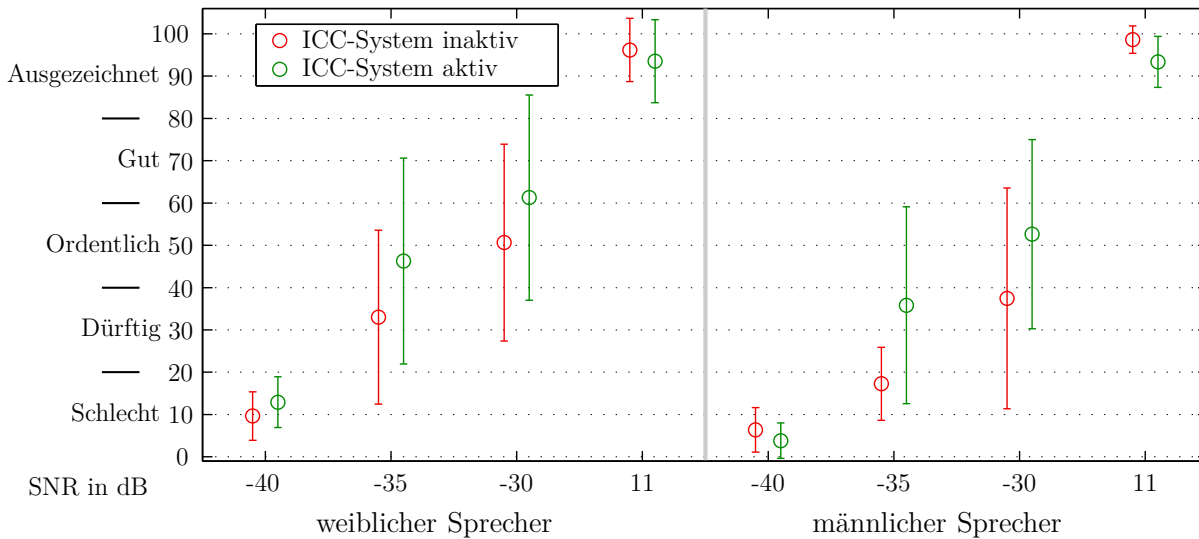


Abbildung 6.12: Ergebnis der auditiven Gesamtqualitätsevaluierung eines ICC-Systems mit einer Testgruppe von 11 Testpersonen. Die mittlere Qualitätsbeurteilungen (MOS) MOS_S über alle Testpersonen sind als Kreise und die zugehörige Standardabweichung als vertikalen Balken. Auf der Abszisse sind alle Stimuli mit ihren unterschiedlichen Parametrierungen dargestellt, wobei die Unterscheidung eines aktiven ICC-Systems in grün und eines deaktivierten ICC-Systems in rot erfolgt.

ken stellen die zugehörige Standardabweichung der Daten dar. Auf der Abszisse sind alle betrachtete Sprach-zu-Geräuschleistungsszenarios aufgetragen. Dabei ist der SNR am Zuhörerohr ohne ein aktives ICC-System angegeben. Des weiteren ist die Abszisse in einen weiblichen und männlichen Sprecher, gemäß der Trennung innerhalb des Hörtests, aufgeteilt worden. Das Ergebnis der Evaluierung ohne ein aktives ICC-System ist in rot und mit einem ICC-System ist in grün dargestellt.

In Abbildung 6.12 wird deutlich, dass ein großer Bereich der Qualitätsskala von den MOS-Werten abgedeckt wird. Dies ist für den Entwurf eines instrumentellen Evaluierungsverfahrens, wie er in Kapitel 4.2.1 dargestellt ist, wichtig, denn nur so kann die vollständige Qualitätsskala im Modell mit aufgenommen und später geschätzt werden. Einzig innerhalb des Bereiches „Gut“ ist die Datenbasis nicht stark besetzt, hier kann nicht mehr als ein MOS-Wert gefunden werden. Dennoch bieten diese Daten eine erste Basis für den Entwurf einiger instrumenteller Evaluierungsverfahren in Kapitel 7. Betrachtet man nun jedes einzelne Sprach-zu-Geräuschleistungsszenario für sich, kann festgestellt werden, dass die MOS-Werte mit der Aktivierung eines ICC-Systems bei einem SNR von -35 dB und -30 dB ansteigen. Das bedeutet, dass innerhalb dieser Szenarien eine Verbesserung der Qualität im Sinne des empfundenen SNRs durch die Verwendung des ICC-Systems erzielt werden kann. Eine Ausnahme dieses Verhaltens tritt im Referenzfall bzw. im Fall

eines Szenarios ohne Hintergrundgeräusch auf: Hier fallen die MOS-Werte leicht für beide Sprecher. Dies deutet auf ein weniger gut parametrisiertes ICC-System hin, welches ohne Hintergrundgeräusch das Eingangssignal verarbeitet und eine Ausgabe erzeugt, welche die Testpersonen im vorliegenden Experiment im Qualitätsempfinden beeinträchtigt hat. Auch bei dem unteren Anker bei einem SNR von -40 dB und einem männlichen Sprecher tritt eine Reduktion der MOS-Werte von 2,5 auf. Der Grund für die Verschlechterung der Qualität bei der Verwendung eines ICC-Systems ist möglicherweise auch in der weniger guten Parametrierung des ICC-Systems zu suchen. Denn innerhalb dieses Szenarios scheint das verwendete ICC-System das Nutzsignal zu gering zu verstärken, sodass keine deutliche Erhöhung des Sprachsignals innerhalb der Stimuli erfolgt. Quantitativ können die in Tabelle 6.5 aufgelisteten Veränderungen innerhalb der verschiedenen Geräusch-zu-Sprachleistungsszenarien und den beiden Geschlechtern der Sprecher bestimmt werden.

Geschlecht des Sprechers	SNR des Sprach-zu-Geräuschleistungsszenarios			
	-40 dB	-35 dB	-30 dB	11 dB
Weiblich	3.3	13.3	10.6	-2.6
Männlich	-2.5	18.5	15.2	-5.3

Tabelle 6.5: Differenzen zwischen den MOS-Werten mit und ohne aktivem ICC-System innerhalb der verschiedenen Sprach-zu-Geräuschleistungsszenarios in Abhängigkeit des Geschlechtes des für die Erzeugung des Stimuli verwendeten Sprechers.

Vergleicht man nun die Differenzen der MOS-Werte zwischen weiblichen und männlichen Stimuli, so kann besonders in den Sprach-zu-Geräuschleistungsszenarios mit -35 dB und -30 dB SNR gesehen werden, dass für einen männlichen Sprecher ein höherer Anstieg erzielt werden konnte als für einen weiblichen Sprecher. Dies liegt an der spektralen Leistungsverteilung der Sprachsignale: Meist sind weibliche Sprecher schon von Anfang an in einem durch Fahrzeuggeräusch beeinträchtigten Szenario besser zu verstehen als männliche. Weibliche Sprecher regen mehr Leistung in den höheren Frequenzen an und haben damit den Vorteil, weniger stark durch das tieffrequent dominante Geräuschsignal beeinträchtigt zu werden. Nichtsdestotrotz kann auch bei einem weiblichen Sprecher eine Verbesserung durch die zusätzliche Wiedergabe des Nutzsignals eines ICC-Systems erzielt werden.

6.5 Zusammenfassung

Bei der Evaluierung von ICC-Systemen unter Verwendung auditiver Verfahren, steht die Forschung derzeit noch am Anfang. Bisher sind nur wenig psychoakustische Experimente und Hörtests in Bezug auf ein Fahrzeug, ausgestattet mit einem ICC-System, durch-

geführt worden. Innerhalb dieses Kapitels sind erste auditive Evaluierungsverfahren in dieser speziellen Umgebung durchgeführt und die zugehörigen Ergebnisse vorgestellt worden. Insgesamt ist festzustellen, dass die Evaluierung durch Menschen nahezu immer zu sehr individuellen Ergebnissen führt. Nicht zuletzt aus diesem Grund sind für zukünftige auditive Evaluierungen von ICC-Systemen standardisierte Verfahren, wie sie in anderen Bereichen existieren, sinnvoll und notwendig.

In diesem Kapitel sind vier verschiedene auditive Evaluierungsverfahren, mit deren Hilfe erste Erkenntnisse und Hypothesen gefunden werden konnten, vorgestellt. So konnte zum Beispiel, anhand des Echoperzeptions-Experimentes, eine erste Hypothese in Bezug auf die notwendige Verstärkung des Nutzsignals durch das ICC-System aufgestellt werden. Der Fokus dieses Experimentes lag zunächst auf der Verstärkungseinstellung der zweiten Signalquelle. Aus anderen psychoakustischen Experimenten und aus der Literatur [ZF99, S.311] ist bekannt, dass die Echowahrnehmung auch durch die Verzögerung zweier Signalquellen zueinander beeinflusst wird. Außerdem sind die Geräuschkonstellationen bewusst mit großer Leistung gewählt worden, dennoch sollte in zukünftigen Betrachtungen zusätzlich eine stärkere Variation der gewählten Leistung des Hintergrundgeräusches gewählt werden. Aus diesen Gründen sind weitere und detailliertere Experimente in zukünftigen Analysen in diesem Bereich von großem Forschungsinteresse.

Neben dem Zuhörer ist auch der Sprecher selbst ein qualitätsrelevanter Faktor innerhalb der Evaluierung von ICC-Systemen. Deswegen ist zusätzlich die Eigenechoperzeption des Sprechers genau untersucht worden. Hier konnten erste Auswirkungen des Lombard-Effektes auf die, durch die Testpersonen, ausgewählte Verstärkung des Eigenechos beobachtet werden. Mit dem Anstieg der Leistung des Geräuschkonstellations steigt der Sprachpegel und gleichzeitig verringert sich die Verstärkung, welche erforderlich ist, um das Sprachsignal wahrzunehmen. Diese Erkenntnis legt die Vermutung nahe, dass für eine geräuschabhängige Verstärkung (NDGC) des ICC-Systems, ein monoton ansteigender Zusammenhang zwischen der Verstärkung und der Geräuschleistung eher ungeeignet ist. Des Weiteren konnte eine Empfehlung der, durch das ICC-System eingefügte, Verzögerung (zwischen 12 und 18 ms) mit diesem psychoakustischen Experiment gefunden werden. Dieses Ergebnis deckt sich mit den Werten aus der einschlägigen Literatur, wie zum Beispiel in [SH06] oder auch in einigen der in [LB64] veröffentlichten Experimente.

Im weiteren Verlauf werden die gefundenen Ergebnisse der Hörtests und der psychoakustischen Experimente, zur Entwicklung der Qualitätsschätzer in Kapitel 7 verwendet. Dabei bildet die Gesamtqualitätsbeurteilung eines ICC-Systems die Qualitätsreferenz, welche für den Entwurf der Verfahren aus der Evaluierungsgruppe Kommunikationsqualität verwendet wird. Die anderen psychoakustischen Experimente helfen bei der Qualitätsschätzung der Verfahren innerhalb der Evaluierungsgruppe Verhalten des ICC-Systems.

Letztendlich bieten all diese auditiven Verfahren einen ersten Einblick in das Verhalten und das Qualitätsempfinden von Probanden innerhalb eines, mit einem ICC-System

ausgestatten, Fahrzeuges. Um weitere und detaillierte Ergebnisse zu erlangen, bedarf es gerade innerhalb der auditiven Verfahren weiterer Forschung. Nicht nur die Anzahl der Probanden innerhalb der Experimente sollte erhöht, sondern auch weitere qualitätsrelevante Parameter eines ICC-Systems, wie zum Beispiel die Veränderung des Hintergrundgeräusches, sollten untersucht werden. Zudem sollte eine Empfehlung ähnlich zu [ITU14] speziell für die Evaluierung von ICC-Systemen entworfen werden. Denn mithilfe dieser Empfehlung könnten verschiedene Datenbanken unterschiedlicher Urheber zu einer großen Datenbasis kombiniert werden. In diesem Zusammenhang sollte nicht nur die Qualität des Zuhörers in einer ICC-Qualitätsdatenbasis mit aufgenommen sondern auch das Qualitätsempfinden des Sprechers mit berücksichtigt werden. Diese Kombination würde zu einer vollständigen und umfangreichen Datenbasis führen, die das Verständnis für das gegebene Szenario erhöhen und zusätzlich die Weiter- und Neuentwicklung von instrumentellen Evaluierungsverfahren voranbringen würde.

Kapitel 7

Instrumentelle Evaluierungsverfahren

Die instrumentelle Evaluierung von ICC-Systemen analysiert die Qualität solcher Systemen anhand von Verfahren, welche auf der Basis von Signalanalysen umgesetzt werden. Dabei assoziiert der Qualitätsbegriff im schlimmsten Fall die reine Sprachverständlichkeit und im besten Fall lediglich die Sprachqualität eines ICC-Systems. Diese Begriffe sind bereits in Abschnitt 3.1 einsortiert worden. Nach dem Wissen der Autorin gibt es derzeit noch keine Standards oder Empfehlungen zur Evaluierung von ICC-Systemen. Ursächlich hierfür ist, dass sich ICC-Systeme erst seit kurzem auf dem Automobilmarkt etablieren und die Forschung noch nicht vollständig abgeschlossen ist. Gegenwärtig ist noch nicht vollständig klar, welche Qualitätsbestandteile für die Evaluierung von ICC-Systemen wesentlich sind oder wie die Interaktion verschiedener Parameter die Qualität beeinflusst. Erste Ideen und Ansätze für die Evaluierung mithilfe von instrumentellen Verfahren sind beispielsweise in [MK, MKL, TSb, STWW12a, SH06] veröffentlicht worden. All diesen Veröffentlichungen ist gemeinsam, dass sie bereits etablierte Verfahren verwenden, um die Qualität eines ICC-Systems zu bestimmen. Die meisten Verfahren zur Evaluierung der Sprachqualität sind für Codecs, Telefonie oder Beschallungsanlagen konzipiert und entwickelt worden. Dennoch liefern sie einen ersten Ansatz, um das Problem der instrumentellen Evaluierung von ICC-Systemen zu lösen, wobei der Begriff instrumentelle Evaluierungsverfahren Methoden beschreibt, welche unabhängig von Personen, lediglich mithilfe von entsprechenden Signalen und Algorithmen, die Qualität eines Systems vorhersagen können. Jedoch benötigt auch ein solches Verfahren als Qualitätsreferenz die Evaluierung von Menschen. Deswegen wird für den Entwurf solcher Verfahren immer ein subjektiver Hör- und/oder Perzeptionstest, wie in Kapitel 6 beschrieben, benötigt.

Ferner wird bei diesen Evaluierungsmethoden meist davon ausgegangen, dass Zuhörer und Sprecher räumlich voneinander getrennt sind, beispielsweise bei einem Telefonat. Dadurch ergeben sich andere Anforderungen und Erwartungen an die Sprachqualität. Bei einem ICC-System sind Sprecher und Zuhörer in der gleichen räumlichen Umgebung. Dies

erfordert eine Anpassung der Evaluierungsmethoden und der Evaluierungsstrategie an die Umgebung, siehe Abschnitt 3.

Generell ist anzumerken, dass die Evaluierung eines ICC-Systems im Labor sich von der gewöhnlichen Verwendung unterscheidet. Ein entscheidender Unterschied zwischen der üblichen Verwendung eines ICC-Systems während der Fahrt und innerhalb einer Laborumgebung ist die langsame Anpassung des Hintergrundgeräusches. In der Laborumgebung kann von einem Szenario, zum Beispiel im Stillstand, auf ein anderes Szenario, beispielsweise während der Fahrt bei einer Geschwindigkeit von 150 km/h, instantan umgeschaltet werden. Dabei muss beachtet werden, dass das ICC-System eine gewisse Zeit benötigt, um sich an die neue Situation anzupassen. Aus diesem Grund ist es für die Evaluierung essenziell, um akkurate Ergebnisse zu bekommen, eine Anpassungsphase einzuhalten. In dieser Phase muss die Zielumgebung im Sinne des betrachteten Umgebungsgeräusches aktiv sein und ein entsprechendes Sprachsignal vom künstlichen Sprecher reproduziert werden. Die Dauer dieser Phase hängt stark vom Geräuschszenario und dem Wechsel, wie auch vom ICC-System selbst ab. Erfahrungsgemäß hat sich eine Zeitspanne zwischen 30 und 40 ms als ausreichend erwiesen.

Qualitätsklasse	Bedeutung
1	Innerhalb dieser Klasse wird eine außerordentlich gute Umsetzung in Bezug auf das betrachtete Systemverhalten bzw. Qualitätsmaß erreicht. Aus diesem Grund entspricht diese Klasse einem exzellenten Ergebnis. Systeme, welche ein Ergebnis dieser Klasse erzielen, sind in der Regel unter besonderen Anstrengungen entwickelt und implementiert worden.
2	Ergebnisse innerhalb dieser Klasse entsprechen einem gutem Ergebnis.
3	In dieser Klasse werden Ergebnisse klassifiziert, welche zwar die Anforderungen eines guten Systems nicht mehr erfüllen und einzelne kleine Schwächen aufweisen, diese jedoch noch tolerierbar sind. Dies kann zum Beispiel der Fall sein, wenn, in Bezug auf die Gesamtqualität, einzelne Schwächen durch andere Stärken kompensiert werden können.
4	Umsetzungen eines ICC-Systems, welche durch auditive und empirisch ermittelte Verfahren innerhalb eines mangelhaften Bereiches liegen, werden in diese Klasse eingestuft.

Tabelle 7.1: Definition der in den Evaluierungsgruppen „Verhalten des ICC-Systems“ und „Kommunikationsqualität“ verwendeten Qualitätsklassen in Anlehnung an [ITU13].

Innerhalb eines jeden Evaluierungsverfahrens entstehen unterschiedlichste Ergebnisse mit unterschiedlichen Einheiten, welche jedoch in eine Gesamtqualität überführt werden sollen. Zu diesem Zweck wird eine einheitliche Klassifizierung der Ergebnisse aus den

Evaluierungsgruppen „Verhalten des ICC-Systems“ und „Kommunikationsqualität“ verwendet, welche in Tabelle 7.1 im Detail beschrieben ist. Diese vierstufige Klassifizierung ist in Anlehnung an einen Leistungsklassensatz der ITU für die Klassifizierung von Teilmodulen einer Autofreisprecheinrichtung nach [ITU13] vorgenommen worden.

Im folgenden Kapitel sind in Abhängigkeit der Evaluierungsgruppen aus Kapitel 3.2 erste instrumentelle Evaluierungsverfahren dargestellt. Diese werden im Detail beschrieben. Des Weiteren wird ein Vorschlag zur Einstufung der Ergebnisse der instrumentellen Verfahren in Qualitätsklassen, auf Basis der gewonnen Erkenntnisse aus Kapitel 6, gegeben. Zusätzlich wird ein erster Ansatz zur Kombination der einzelnen Verfahren zu einer Gesamtqualität in Abschnitt 7.4 beschrieben.

7.1 Fahrzeugeigenschaften

Diese erste Gruppe von instrumentellen Maßen zur Evaluierung von ICC-Systemen bietet eine Einstufung der eigentlichen Umgebung des betrachteten Systems. Sie nimmt daher keinen direkten Einfluss auf die Gesamtqualität, kann jedoch indirekt über die Fahrzeugattribute die Qualitätsbestimmung beeinflussen. Ziel ist es, herauszufinden, wie komplex die Fahrzeugumgebung im Sinne der in einem ICC-System verwendeten Algorithmen ist. Ein ICC-System, welches in einer komplexen Umgebung bereits eine hinreichende Verbesserung, zum Beispiel durch die Verstärkung des Sprachsignals am Zuhörerohr, erzielt, sollte besser bewertet werden, als ein System, welches das selbe Ergebnis in einer eher unproblematischen Umgebung erzielt.

Zur Einstufung dieser Fahrzeugumgebung werden drei Attribute mithilfe der im folgenden beschriebenen Evaluierungsverfahren bestimmt. Es handelt sich dabei um die Attribute „Hintergrundgeräusch“, „Wandler des ICC-Systems“ und „Fahrzeuggröße“. Eine detaillierte Beschreibung der Attribute und ihrer Verknüpfung mit den Ergebnissen der Evaluierungsverfahren wird in Kapitel 7.1.3 gegeben. Einzig diese Fahrzeugattribute haben einen direkten Einfluss, in Form einer Gewichtung, auf das Qualitätsergebnis eines ICC-Systems.

Um die Fahrzeugumgebung anhand der Fahrzeugattribute klassifizieren zu können, muss zunächst eine Auswahl von Fahrzeugen vermessen und analysiert werden. Dabei sollten die ausgewählten Fahrzeuge möglichst unterschiedliche Fahrzeugtypen abdecken, um sowohl komplexe als auch einfache Umgebungen bestimmen zu können. Zu diesem Zweck ist eine Fahrzeuggruppe bestehend aus

- einer Mercedes-Benz E-Klasse (Kabriolett), wobei nur die Sitzplätze in der vorderen Reihe betrachtet werden,
- einer Mercedes-Benz S-Klasse in der Langversion¹,

¹Die Langversion der S-Klasse besitzt einen größeren Radabstand (130 mm größer) und somit auch einen längeren Innenraum der Fahrgastzelle.

- einem Multivan T5 von Volkswagen und
- einem Corsa D von Opel

ausgewählt worden. Diese vier Fahrzeuge decken alle gängigen Fahrzeugklassen bzw. Fahrzeuggrößen ab und werden zur Bestimmung der Einflussgrößen C_V aller Evaluierungsmaße der Fahrzeugeigenschaften verwendet. Für die Positionen einzelner Komponenten oder auch Insassen werden die in der Automobilbranche üblichen Anordnungen verwendet. Dieses Wissen über die Eigenschaften eines Fahrzeuges ermöglicht eine gewichtete Kombination der Ergebnisse der Einzelverfahren zu einer Gesamtqualität.

7.1.1 Distanzen zwischen den Komponenten

Zur Bestimmung der Eigenschaften eines Fahrzeuges in Bezug auf die Qualität eines ICC-Systems, ist die Anordnung der verwendeten Komponenten, wie Mikrofone und Lautsprecher, und der Kommunikationspartner ein wichtiger Baustein. Im Allgemeinen nimmt die Leistung einer Schallwelle über die Distanz ab. Daher wird die Verarbeitung des Mikrofonsignals eines ICC-Systems einfacher, wenn dieses zum Beispiel nahe am Sprecher positioniert ist.

Diese Informationen helfen abzuschätzen, wie komplex die gegebene Fahrzeugumgebung ist. Denn je nach Position der Komponenten müssen die Algorithmen eines ICC-Systems mit einem größeren Aufwand zur Überwindung der in Kapitel 2 dargestellten Anforderungen entwickelt und implementiert werden.

Entwurf des Verfahrens

Die Messung der Distanzen erfolgt unter Verwendung einer Kreuzkorrelationsanalyse zweier Signale, beispielsweise dem Lautsprecher des künstlichen Mundes $y(n)$ und einem der Mikrofone des künstlichen Zuhörers $x(n)$, siehe Abbildung 5.1. Selbstverständlich existieren mehr Eingangs- und auch Ausgangskanäle während der Evaluierung, diese sollen allerdings zunächst für die Herleitung des Maßes vernachlässigt werden. Die zeitliche Position des Maximums der Kreuzkorrelationsfolge gibt die Verzögerung der beiden Signal zueinander an. Unter Berücksichtigung der Schallgeschwindigkeit und der gemessenen Verzögerung kann die Distanz zwischen den beiden Signalquellen abgeschätzt werden. Dabei ist die Kreuzkorrelationsfolge durch

$$s_{xy}(\kappa) = E\{x^*(n)y(n + \kappa)\} \quad (7.1)$$

gegeben [Mer13, S. 201]. Entsprechen $x(n)$ und $y(n)$ zwei endlichen Folgen, für die gilt, dass $n \in [0, N - 1]$ ist, kann die Kreuzkorrelationsfolge mit

$$\hat{s}_{xy}(\kappa) = \frac{1}{N} \sum_{n=0}^{N-1} x^*(n)y(n + \kappa) \quad (7.2)$$

geschätzt werden. Aufgrund der mathematischen Ähnlichkeit zwischen der Korrelation zweier Folgen und der Faltung, kann die Kreuzkorrelationsfolge durch eine Faltung der beiden Folgen berechnet werden, wobei die Spiegelung einer der Folgen durch die Faltung, für die Kreuzkorrelationsanalyse aufgehoben werden muss [PM07, S. 115 ff.]. Für eine effiziente Faltung wird die schnelle Fourier-Transformation (engl. *Fast Fourier Transform*, FFT) zur Überführung der Zeitsignale in den Frequenzbereich verwendet. Die Schätzung der Kreuzkorrelationsfolge soll innerhalb, des in Kapitel 5.3 beschriebenen, Echtzeit-Rahmenwerks umgesetzt werden. Dazu muss in diesem speziellen Fall, neben der Transformation in den Frequenzbereich, auch die Blockverarbeitung bei der Herleitung des Verfahrens mit berücksichtigt werden.

Hierfür wird zunächst das Referenzsignal $y(n)$, welches vom künstlichen Sprecher reproduziert wird, ausgewählt. Für die Distanzmessung zwischen den verwendeten Komponenten wird ein weißer, mittelwertfreier Gaußprozess verwendet. Die beiden zu korrelierenden Signale werden mittels eines Hochpassfilters zunächst von unerwünschten Hintergrundgeräuschen befreit. Dafür wird ein Filter mit einer Grenzfrequenz von circa 100 Hz entworfen. Damit die Ähnlichkeit der zu korrelierenden Signale nicht durch die Filterung zerstört wird, werden sowohl das Referenzsignal $y(n)$, welches ungestört ist, als auch das Ohrsignal mit dem gleich entworfenen Filter gefiltert. Zusätzlich wird das Referenzsignal $y(n)$ zeitlich umsortiert, um so die Spiegelung der Folge durch die Faltung rückgängig zu machen. Im Anschluss werden die Abtastwerte eines Blocks, wobei R dem Rahmenversatz und somit der Anzahl von Abtastwerten innerhalb eines Blocks entspricht, mit Nullen aufgefüllt bis die gewünschte Länge der Diskreten Fourier-Transformation (DFT) N_{FFT} erreicht ist. Es ergibt sich folgender Zusammenhang:

$$\tilde{y}(r, k) = \begin{cases} \bar{y}((k+1)R - (r+1)), & 0 \leq r < R \\ 0, & R \leq r < N_{\text{FFT}}, \end{cases} \quad (7.3)$$

und

$$\tilde{x}(r, k) = \begin{cases} \bar{x}(r+kR), & 0 \leq r < R \\ 0, & R \leq r < N_{\text{FFT}}, \end{cases} \quad (7.4)$$

wobei $\bar{y}(n)$ und $\bar{x}(n)$ dem hochpassgefilterten Referenzsignal $y(n)$ und dem Ohrsignal $x(n)$ entsprechen, k den aktuellen Rahmenindex repräsentiert und $r \in [0, N_{\text{FFT}})$ für den Zeitindex innerhalb eines betrachteten Fensters steht.

Im Anschluss werden beide Signale mithilfe einer DFT in den Frequenzbereich überführt. Es ergibt sich für beide Signale die spektrale Darstellung

$$\tilde{X}(\mu, k) = \sum_{n=0}^{N_{\text{FFT}}-1} \tilde{x}(r, k) e^{-j \frac{2\pi}{N_{\text{FFT}}} \mu r} \quad (7.5)$$

und

$$\tilde{Y}(\mu, k) = \sum_{n=0}^{N_{\text{FFT}}-1} \tilde{y}(r, k) e^{-j \frac{2\pi}{N_{\text{FFT}}} \mu r}, \quad (7.6)$$

in Abhängigkeit des Frequenzbandes μ . Diese Spektren werden weiterverwendet, um die Faltung beider Signale durchzuführen. Da aber mehr als nur ein Rahmen eines jeden Signals verwendet werden muss, um die gesuchte Verzögerung in der Kreuzkorrelationsfolge abbilden zu können, müssen mehr Rahmen als nur der aktuelle betrachtet werden. Für die Faltung bzw. Multiplikation der Spektren ergibt sich aus diesem Grund:

$$U(\mu, \nu) = \sum_{k=0}^{K-1} \tilde{X}(\mu, k - \nu) \cdot \tilde{Y}(\mu, k) \quad l \in [0, N_{\text{Est}}), \quad (7.7)$$

wobei N_{Est} der Anzahl der maximal analysierbaren Rahmen entspricht. Diese Spektren werden im Anschluss mittels einer inversen Fourier-Transformation zurück in den Zeitbereich überführt. Es ergibt sich für jeden betrachteten Block l , dass

$$u(r, l) = \frac{1}{N_{\text{FFT}}} \sum_{\mu=0}^{N_{\text{FFT}}-1} U(\mu, l) e^{j \frac{2\pi}{N_{\text{FFT}}} \mu r} \quad (7.8)$$

ist. Diese Zeitbereichsdarstellung wird im Anschluss überlappend auf summiert, um eine erste Schätzung der Kreuzkorrelationsfolge zu erhalten.

$$\hat{s}_{xy}(\kappa) = \sum_{l=0}^{N_{\text{Est}}-1} u(\kappa - lR, l). \quad (7.9)$$

In Abbildung 7.1 ist das Ergebnis der vorgestellten, blockweisen Schätzung der Korrelationsfolge in Schwarz dargestellt, in Rot die zur Simulation verwendete Impulsantwort, welche aus drei unterschiedlich gewichteten Dirakimpulsen (an den Stellen $\kappa = [31, 332, 403]$) besteht. In Grün ist die Schätzung der Kreuzkorrelationsfolge dargestellt, berechnet mithilfe von Matlab.

Deutlich wird, dass die Schätzung von Matlab sehr ähnlich der zuvor festgelegten Impulsantwort ist. Die beiden Kurven überlagern sich, was dazu führt, dass die Schätzung mit Matlab hinter der eigentlichen Impulsantwort in Rot verschwindet. Diese Methode in Matlab kann jedoch nur unter Verwendung der vollständigen Signale, also nicht in einer blockweisen Verarbeitung, durchgeführt werden. Die vorgestellte Methode zur blockweisen Schätzung der Kreuzkorrelationsanalyse weist einen ähnlichen Verlauf wie die Impulsantwort als auch die Schätzung von Matlab auf. Sie ist lediglich um die Anzahl der Abtastwerte eines Rahmens verschoben. Dieses Verhalten liegt in der blockweisen Verarbeitung begründet, kann aber durch eine einfache Verschiebung kompensiert werden.

Die geschätzte Impulsantwort kann im Anschluss verwendet werden, um die Zeit, die der Schall von der Quelle zur Senke benötigt bzw. die Verzögerung des Direktschalls, zu bestimmen. Die Verzögerung entspricht der Zeit zwischen dem Ursprung bis zum ersten

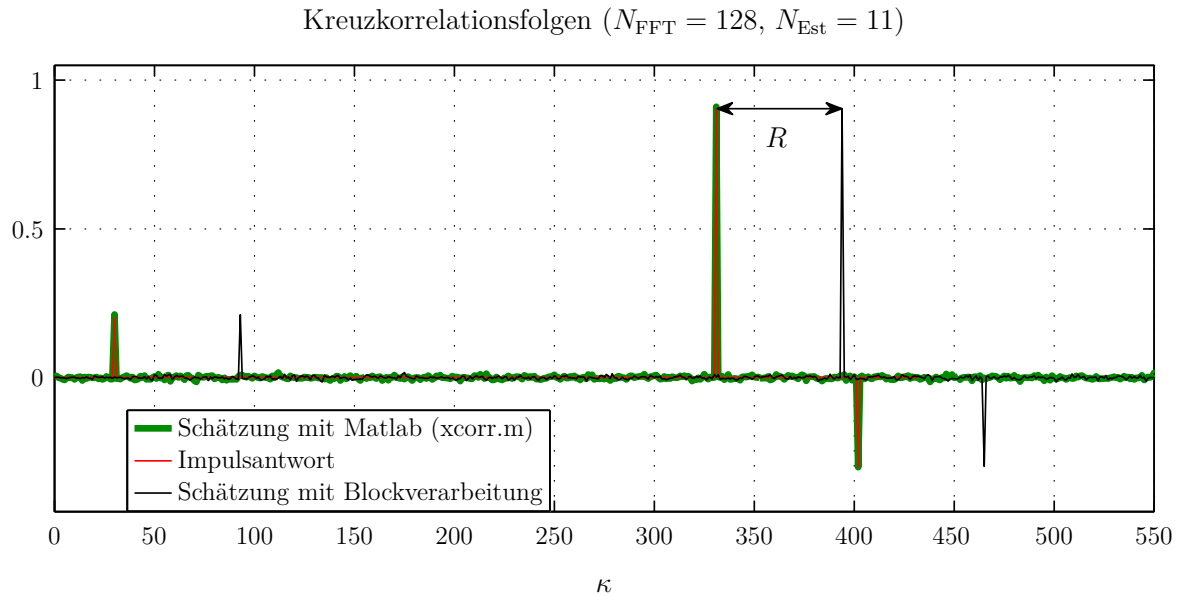


Abbildung 7.1: Vergleich zwischen der vorgegebenen Impulsantwort (rot) und der Kreuzkorrelationsfolgen geschätzt mit der Funktion `xcorr.m` von Matlab (grün) und unter Verwendung des beschriebene, blockweisen Verfahrens (schwarz).

drastischen Anstieg der Impulsantwort. Im Fall einer idealen Impulsantwort könnte man diesen Anstieg mithilfe einer Schwelle, welche der Betrag der Impulsantwort überschreiten muss, detektieren. Im Anwendungsfall hat sich diese Methode als wenig akkurat herausgestellt, so dass ein neues Verfahren zur Detektion der Verzögerung verwendet wird. Dazu wird zunächst die Standardabweichung der Impulsantwort in einem Bereich gemessen, in dem noch keine Anteile des Nutzsignals die Senke erreicht haben. Aufgrund der Verschiebung der Kreuzkorrelationsfolge, durch das beschriebene Verfahren, kann dies innerhalb der ersten R Werte der Folge durchgeführt werden. Diese Standardabweichung σ_n wird verwendet, um den ersten starken Anstieg bzw. die Direktschallkomponenten identifizieren zu können. Zunächst wird die Differenz der einzelnen Elemente der Schätzung der Kreuzkorrelationsfolge mithilfe von Gleichung (7.10) analysiert:

$$\hat{s}'_{xy}(\kappa) = \begin{cases} \frac{1}{\kappa}, & \text{wenn } |\hat{s}_{xy}(\kappa + 1) - \hat{s}_{xy}(\kappa)| > 4\sigma_n \\ 0, & \text{sonst.} \end{cases} \quad (7.10)$$

Die Verzögerung der Direktschallübertragung ist gegeben durch:

$$D_V = \left(\arg \max_{\kappa} \{ \hat{s}'_{xy}(\kappa) \} - R \right) / f_s. \quad (7.11)$$

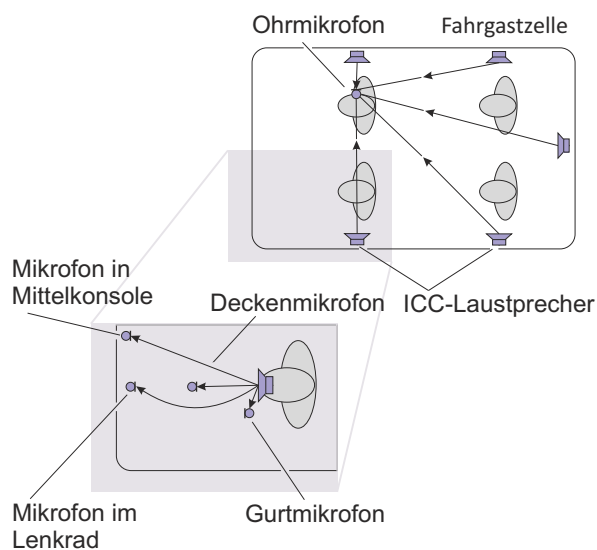
Um diese Verzögerung in Distanzen auszudrücken, müsste das Ergebnis D_V noch mit der Schallgeschwindigkeit multipliziert werden. Da dies lediglich einem linearen Zusammen-

hang, einer Multiplikation mit der Schallgeschwindigkeit, entspräche, wird dieser Schritt ausgelassen und die Verzögerungen als das Ergebnis betrachtet.

Mithilfe dieser Messung lassen sich nun alle denkbaren Distanzen bzw. Verzögerungen zwischen einzelnen, für ein ICC-System relevanten, Komponenten bestimmen. Da innerhalb dieser Gruppe keine direkte Qualität bestimmt wird, sondern ausschließlich die entsprechende Gewichte die zur weiteren Bestimmung ermittelt werden, wird hier auf eine Qualitätseinstufung verzichtet. Dennoch müssen die Grenzen ermittelt werden, die den Ergebnisraum der Einflussgrößen aufspannen. Im folgenden Abschnitt wird genauer auf den Fall der Distanzbestimmung innerhalb der Fahrgastzelle und der Bestimmung der Einflussgrößen eingegangen.

Spezialfall Fahrzeuginnenraum

(a) Ausbreitungswege zwischen Zuhörer bzw. Sprecher und ICC-System



(b) Ausbreitungswege zwischen den Insassen und den elektroakustischen Wandlern des ICC-Systems

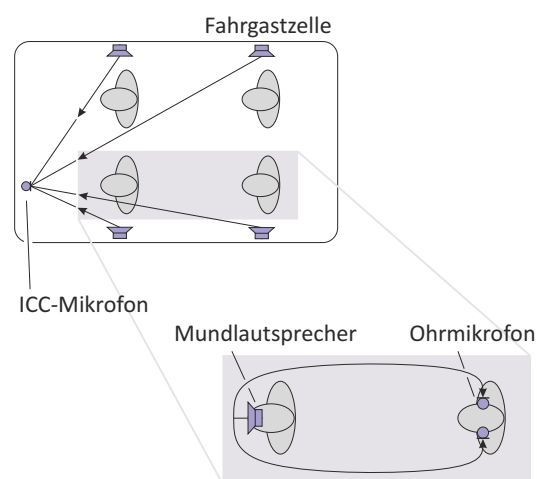


Abbildung 7.2: Verschiedene Ausbreitungswege von und zum Zuhörer bzw. Sprecher zu den einzelnen elektroakustischen Wandlern eines ICC-Systems und untereinander. Zusätzlich sind unterschiedliche Wandlerpositionen mit berücksichtigt.

Da die Distanzen zwischen unterschiedlichen Komponenten unterschiedlich die Gesamtqualität eines ICC-Systems beeinflussen, werden diese getrennt voneinander betrachtet und ausgewertet. Innerhalb dieser Messung werden vier mögliche Pfade unterschieden, welche in Abbildung 7.2 schematisch dargestellt sind. Die Darstellung erhebt keinen Anspruch auf Vollständigkeit, sie soll lediglich einen Einblick in die Möglichkeiten der Positionen von Bauteilen und Insassen geben. Um eine Analyse des ICC-Systems anhand eines *black-box testing* durchzuführen, müssen die ICC-Komponenten mit zusätzlichen Hilfsmikrofonen und -lautsprechern ausgestattet werden. Dies bedeutet im Detail, dass an allen

Positionen der ICC-Mikrofone zusätzliche, direkt mit der ICC-Systemevaluierung verbundene Mikrofone platziert werden. Gleiches wird ebenso mit den ICC-Lautsprechern und den zusätzlichen Hilfslautsprechern durchgeführt. Dadurch ist kein elektrischer Eingriff in das ICC-System mit seinen eigenen Komponenten notwendig. Außerdem werden die zuvor zur Herleitung des Maßes vernachlässigten, unterschiedlichen Kanäle der Lautsprecher l und der Mikrofone m wieder eingeführt. Dadurch ergibt sich für die mithilfe von (7.11) bestimmte Verzögerung eine Abhängigkeit von $x_m(n)$ für die Mikrofonsignale und $y_l(n)$ für die Lautsprechersignale. Die Berechnung wird analog wie bereits beschrieben durchgeführt. Das Ergebnis der Verzögerungsschätzung ist durch $D_V(l, m)$ gegeben.

Sprechermond zu ICC-Mikrofon Dieser Pfad bestimmt die Qualität des Sprachsignals, welches dem ICC-System zur Verfügung steht. Da der Schalldruck einer Quelle proportional zum reziproken Abstand von der Quelle abnimmt [Kut09], nimmt auch die Leistung des Sprachsignals, aufgenommen vom ICC-Mikrofon, mit Anstieg des Abstandes zum Sprecher ab. Ein weit entferntes Mikrofon stellt damit ein Signal zur Verfügung, welches schwieriger zu verarbeiten ist als das von einem nahen Mikrofon. Im unteren Bereich von Abbildung 7.2(a) sind realisierbare Mikrofonpositionen zur eingehenderen Betrachtung dargestellt. Dabei sind die Positionen unabhängig vom Fahrzeugtyp und können generell betrachtet werden. Die wichtigsten Mikrofonpositionen in der Automobilindustrie sind die im Fahrzeughimmel und in der Mittelkonsole. Mikrofone in den Gurten, aber auch in der A-Säule, stellen bis heute noch die Ausnahme dar, wobei gerade die Position in der A-Säule meist für Nachrüstlösungen von Freisprecheinrichtungen verwendet wird. Um einen vollständigen Überblick über der Qualität zu erhalten, sollen ebenfalls die nicht gängigen Positionen mitberücksichtigt werden.

Zur Bestimmung der Einflussgröße $c_D^{(0)}$ sind die verschiedenen Positionen in den einzelnen Fahrzeugen ausgemessen und analysiert worden. Das Ergebnis ist in Tabelle 7.2 zusammengestellt. Neben den Distanzen sind zusätzlich die entsprechend Verzögerungen angegeben.

Mit diesen Verzögerungen kann nun ein Raum für die Einflussgröße $c_D^{(0)}$ aufgespannt werden. Dabei wird als Minimalwert $\delta_{\min}^{(0)} = 0,8$ ms und damit als einfachster Fall für die Signalverarbeitung eines ICC-Systems angenommen. Als schwierigster Fall wird der maximal Wert $\delta_{\max}^{(0)} = 1,7$ ms als Obergrenze für den aufgespannten Raum verwendet. Im Anschluss muss für die betrachtete Gruppe ein Gesamtergebnis ermittelt werden. Dazu werden alle Verzögerungen zwischen allen Lautsprechern l_K , die einem künstlichen Sprechermond repräsentieren, und allen ICC-Mikrofonen m_I mithilfe des beschriebenen Verfahrens bestimmt und das Minimum ermittelt:

$$D_V^{(0)} = \min_{l_K, m_I} \{ D_V(l_K, m_I) \}. \quad (7.12)$$

Verzögerung	Abstand	Beschreibung
0,0 - 0,8 ms	0 - 27 cm	Sehr nahe Mikrofone, zum Beispiel Mikrofone in den Gurten.
0,8 - 1,2 ms	27 - 41 cm	Nahe Mikrofone, die zum Beispiel im Himmel oder am oberen Bereich der A-Säule positioniert sind.
1,2 - 1,7 ms	41 - 58 cm	Entfernte Mikrofone, zum Beispiel in der Mittelkonsole oder auch im Lenkrad.
> 1,7 ms	> 58 cm	Weit entfernte Mikrofone zum Beispiel im unteren Teil der A-Säule positioniert.

Tabelle 7.2: Empirisch ermittelte Distanzen bzw. Verzögerungen des Ausbreitungsweges: Sprechermund zum ICC-Mikrofon, bei unterschiedlichen Mikrofonpositionen.

Dadurch wird der Wandler in die Berechnung mit aufgenommen, welcher die geringste Distanz aufweist und von dem ICC-System verwendet werden kann. Unter Verwendung der minimalen Verzögerung wird das Endergebnis dieser Distanzgruppe als Einflussgröße mit

$$c_D^{(0)} = \min \left\{ \max \left\{ 0, \frac{1}{\delta_{\max}^{(0)} - \delta_{\min}^{(0)}} \left(D_V^{(0)} - \delta_{\min}^{(0)} \right) \right\}, 1 \right\} \quad (7.13)$$

bestimmt. Die Einflussgröße $c_D^{(0)}$ entspricht 1, wenn sehr weit entfernte Mikrofone verwendet werden und 0, wenn sehr nahe Mikrofone als ICC-Mikrofone eingesetzt werden. Diese Einflussgröße $c_D^{(0)}$ wird für die Berechnung der Fahrzeugattribute, siehe Abschnitt 7.1.3, weiter genutzt.

ICC-Lautsprecher zu Zuhörerohr Auch bei dem zweiten Pfad vom ICC-Lautsprecher zum Zuhörerohr, siehe oberer Teil der Abbildung 7.2(a), entsteht ein Leistungsverlust auf Grund der Distanz zwischen Quelle und Senke. Wenn nun der Lautsprecher zu weit entfernt ist, muss mehr Leistung des Sprachsignales auf den Lautsprecher durch das ICC-System ausgegeben werden. Dies verstärkt jedoch auch die Rückkopplung des Signals in das ICC-Mikrofon und erschwert somit die Kompensation dieser Rückkopplung. Es entsteht eine weitaus komplexere Fahrzeugumgebung als bei einem zuhörer-nahen Lautsprecher.

Für diese Gruppe sind ebenfalls die unterschiedlichen Wege zwischen den Lautsprecherpositionen l_I und den Zuhörerohren m_K innerhalb der verschiedenen Fahrzeuge gemessen und analysiert worden. Dabei sind die in Tabelle 7.3 beschriebenen Distanzen bzw. Verzögerungen ermittelt worden.

Verzögerung	Abstand	Beschreibung
0,0 - 0,4 ms	0 - 12 cm	Sehr naher Lautsprecher, zum Beispiel bei Lautsprechern in den Kopfstützen des Zuhörers.
0,4 - 1,2 ms	12 - 41 cm	Nahe Lautsprecher, zum Beispiel in der Hutablage direkt hinter dem Zuhörer eingebaut.
1,2 - 2,9 ms	41 - 100 cm	Entfernte Lautsprecher, zum Beispiel in der Tür nahe des Zuhörers.
> 2,9 ms	> 100 cm	Weit entfernte Lautsprecher, zum Beispiel im Sitz einer Limousine vor dem Zuhörer eingebaut.

Tabelle 7.3: Empirisch ermittelte Distanzen bzw. Verzögerungen des Ausbreitungsweges: ICC-Lautsprecher zu Zuhörerohr, bei unterschiedlichen Lautsprecherpositionen.

Auch hier werden die maximale $\delta_{\max}^{(1)} = 2,9$ ms und die minimale Verzögerung $\delta_{\min}^{(1)} = 0,4$ ms verwendet, um einen Wertebereich zwischen einer komplexen und einer einfacheren Fahrzeugumgebung zu definieren. Die kleinste Distanz bzw. Verzögerung zwischen allen ICC-Lautsprechern l_I zu allen Ohrmikrofonen m_K bestimmt das Ergebnis dieser Distanzgruppe und ist gegeben durch:

$$D_V^{(1)} = \min_{l_I, m_K} \{ D_V(l_I, m_K) \}. \quad (7.14)$$

Im Anschluss wird die Einflussgröße als Endergebnis dieser Gruppe anhand von

$$c_D^{(1)} = \min \left\{ \max \left\{ 0, \frac{1}{\delta_{\max}^{(1)} - \delta_{\min}^{(1)}} \left(D_V^{(1)} - \delta_{\min}^{(1)} \right) \right\}, 1 \right\} \quad (7.15)$$

bestimmt. Die Einflussgröße $c_D^{(1)} = 1$ entspricht dem komplexesten Fall und korrespondiert zu einem weit entfernten Lautsprecher. Weniger schwierig für die Signalverarbeitung eines ICC-Systems stellt sich eine Situation mit einem nahen Lautsprecher dar, was einem Ergebnis $c_D^{(1)} = 0$ entspricht.

ICC-Lautsprecher zu ICC-Mikrofon Der Ausbreitungsweg zwischen ICC-Lautsprecher und ICC-Mikrofon stellt wohl aus Sicht der Forschung und Entwicklung eines ICC-Systems den wahrscheinlich interessantesten Pfad dar. Über diesen Weg können Signalanteile, welche über den ICC-Lautsprecher ausgegeben worden sind, wieder in das ICC-Mikrofon einkoppeln und aufschwingen. Dies kann zu einem störenden Pfeifton führen. Dieses Aufschwingen des Gesamtsystems ist hochgradig qualitätsrelevant und muss in jedem Fall mitberücksichtigt werden. In Bezug auf die Länge des Rückkopplungspfades, gibt es selbstverständlich starke Abhängigkeiten zum betrachteten Fahrzeugtyp. Ein Corsa D von Opel weist geringere Distanzen zwischen dem ICC-Lautsprecher und dem ICC-Mikrofon auf als zum Beispiel ein T5 von

Volkswagen. Folglich koppelt das durch das ICC-System wiedergegebene Sprachsignal stärker wieder in das ICC-System ein. Ein ICC-System, welches dennoch gut mit einem kurzen Rückkopplungspfad zurecht kommt, sollte eine bessere Beurteilung bekommen, da es mit einem weitaus schwierigeren Szenario zurecht gekommen ist, als ein ICC-System, welches gut in einem großen Fahrzeug funktioniert.

Zunächst sind die unterschiedlichen Distanzen in den Beispielfahrzeugen ausgemessen und zu einer Übersicht in Tabelle 7.4 zusammengestellt worden.

Verzögerung	Abstand	Beschreibung
> 9,9 ms	>340 cm	Sehr große Distanz, zum Beispiel der Multivan T5 von Volkswagen.
9,9 - 7,3 ms	340 - 250 cm	Große Distanz, zum Beispiel die S-Klasse in der Langversion von Mercedes-Benz.
7,3 - 4,1 ms	250 -140 cm	Kleine Distanz, zum Beispiel in einem Corsa von Opel.
4,1 - 1,9 ms	140 - 65 cm	Sehr kleine Distanz, zum Beispiel in einem E-Klasse Kabrio von Mercedes-Benz.

Tabelle 7.4: Empirisch ermittelte Distanzen bzw. Verzögerungen des Ausbreitungsweges: ICC-Lautsprecher zu ICC-Mikrofon, bei unterschiedlichen Fahrzeuggrößen.

Dabei erschließt sich leicht, dass eine kurze Distanz die Unterdrückung der Rückkopplung erschwert und eine große Distanz, aufgrund der entstehenden Dämpfung des Wiedergabesignals, die Situation vereinfacht. Der Ergebnisraum wird mithilfe von $\delta_{\max}^{(2)} = 9,9$ ms und $\delta_{\min}^{(2)} = 1,9$ ms aufgespannt. Es wird aus allen möglichen Rückkopplungspfaden, zwischen allen ICC-Lautsprechern l_I und allen möglichen ICC-Mikrofonen m_I , die geringste Verzögerung bestimmt. Das Ergebnis der Abschätzung des kritischsten Rückkopplungspfades ist gegeben durch:

$$D_V^{(2)} = \min_{l_I, m_I} \{ D_V(l_I, m_I) \}. \quad (7.16)$$

Für die Bestimmung der Einflussgröße dieser Distanzgruppe wird der folgende Zusammenhang

$$c_D^{(2)} = \min \left\{ \max \left\{ 0, \frac{1}{\delta_{\max}^{(2)} - \delta_{\min}^{(2)}} \left(\delta_{\max}^{(2)} - D_V^{(2)} \right) \right\}, 1 \right\} \quad (7.17)$$

verwendet. Es ergibt sich, dass $c_D^{(2)} = 1$ im Fall des kürzesten und gleich 0 im Falle des längsten Rückkopplungspfades ist.

Sprechermund zu Zuhörerohr Dieser Ausbreitungsweg zwischen Sprecher und Zuhörer, dargestellt in Abbildung 7.2(b), ist ein Indikator für die Basiskommunikationssituation in Bezug auf die Sprachübertragung zwischen Sprecher und Zuhörer. Er beschreibt, wie eingeschränkt die Situation innerhalb der Fahrzeugkabine ohne ICC-System ist. Denn je näher sich die Passagiere sind, desto einfacher wird die Kommunikation, da das Sprachsignal am Zuhörerohr eine höhere Leistung besitzt als bei einer größeren Distanz. Es handelt sich hier um eine reine Distanzmessung, welche die Orientierung der Passagiere zum derzeitigen Zeitpunkt vernachlässigt. Dies könnte in zukünftigen Betrachtungen analysiert und zur Fahrzeugattributebestimmung verwendet werden, da diese auch die Basiskommunikationssituation innerhalb der Fahrgastzelle beeinflusst.

Erneut sind in den unterschiedlichen Fahrzeugen, Distanzmessungen durchgeführt worden. Die resultierenden Werte sind in vier Stufen eingeteilt worden und im Folgenden in Tabelle 7.5 dargestellt worden. Da wie bereits in den Fällen zuvor eine

Verzögerung	Abstand	Beschreibung
0 - 1,8 ms	0 - 60 cm	Sehr kleine Distanz, zum Beispiel in einem E-Klasse Kabrio von Mercedes-Benz.
1,8 - 2,6 ms	60 - 90 cm	Kleine Distanz, zum Beispiel in einem Corsa von Opel.
2,6 - 3,5 ms	90 - 120 cm	Große Distanz, zum Beispiel die S-Klasse in der Langversion von Mercedes-Benz.
> 3,5 ms	>120 cm	Sehr große Distanz, zum Beispiel der Multivan T5 von Volkswagen.

Tabelle 7.5: Empirisch ermittelte Distanzen bzw. Verzögerungen des Ausbreitungsweges: Sprecher zu Zuhörer, bei unterschiedlichen Fahrzeuggrößen.

größere Distanz zu einer stärkeren Dämpfung des Nutzsignals führt, wird der folgende durch $\delta_{\max}^{(3)} = 3,5$ ms und $\delta_{\min}^{(3)} = 1,8$ ms aufgespannte Ergebnisraum verwendet. Zunächst wird die größte Verzögerung als Ergebnis dieser Gruppe durch

$$D_V^{(3)} = \max_{l_K, m_K} \left\{ D_V(l_K, m_K) \right\} \quad (7.18)$$

bestimmt. Dieser Fall bildet den Zuhörer ab, welcher den Sprecher am schlechtesten innerhalb der Fahrgastzelle versteht. Für die Bestimmung der Einflussgröße dieser Distanzgruppe wird der folgende Zusammenhang

$$c_D^{(3)} = \min \left\{ \max \left\{ 0, \frac{1}{\delta_{\max}^{(3)} - \delta_{\min}^{(3)}} \left(D_V^{(3)} - \delta_{\min}^{(3)} \right) \right\}, 1 \right\} \quad (7.19)$$

verwendet. Es folgt aus dieser Definition, dass $c_D^{(3)} = 0$ einer Situation entspricht in der Zuhörer und Sprecher nahe zueinander sitzen und bei einem Ergebnis $c_D^{(3)} = 1$ weiter voneinander entfernt sind.

7.1.2 Hintergrundgeräusch

Das Hintergrundgeräusch, welches während der Fahrt innerhalb der Fahrgastzelle auftritt, hängt stark von dem betrachteten Fahrzeug ab. Die Lautstärke und die spektralen Eigenschaften dieses Geräusches werden durch verschiedenste Faktoren beeinflusst. Die Hauptfaktoren sind Motorgeräusch, Windgeräusche, Abrollgeräusche der Reifen und Geräusche von Bauteilen wie zum Beispiel der Lüftung [SH06]. Dadurch entsteht, je nach Fahrzeug, eine eher komplexe oder eine einfachere Fahrzeugumgebung, in der ein ICC-System operieren muss. Zum Beispiel ist es in einem leiseren Fahrzeug, unter Umständen, einfacher die Rückkopplung zu kompensieren, da das Sprachsignal weniger stark verstärkt wird als in einer lauten Umgebung. Aus diesem Grund wird im Anschluss ein Verfahren vorgestellt, welches die Leistungsdifferenz des Hintergrundgeräusches in einem Szenario mit und ohne ICC-System bestimmt.

Entwurf des Verfahrens

Zur Bestimmung des Hintergrundgeräusches wird das Ohrsignal des Kunstkopfes $x(n)$, siehe Abbildung 5.1, verwendet. Natürlich kann mehr als ein Kanal am Zuhörerohr analysiert werden. Aus Gründen der Einfachheit wird zur Herleitung der Schätzung des Hintergrundgeräusches lediglich ein Kanal betrachtet. Dieses Ohrsignal wird mittels einer Analysefilterbank in den Frequenzbereich überführt, siehe Abschnitt 5.3. Das Mikrofonspektrum $X(\mu, k)$, bestimmt durch (5.1), wird im Anschluss betragsmäßig über einen vordefinierten Zeitraum gemittelt. Das mittlere Betragsspektrum ist durch

$$\bar{N}_{\text{veh}}(\mu) = \frac{1}{K_{\text{BG}}} \sum_{k=0}^{K_{\text{BG}}-1} |X(\mu, k)| \quad (7.20)$$

gegeben. Die Dauer dieser Mittelung wird über die Anzahl der betrachteten Rahmen K_{BG} angegeben. Um eine hörgerechte Analyse des Hintergrundgeräusches durchzuführen, wird dieses mittlere Betragsspektrum unter Benutzung einer A-Bewertungsfilterkurve gewichtet und ein Ergebnis in dBA bestimmt. Zusätzlich wird nur ein definierter Frequenzbereich (von μ_S^N bis μ_E^N) zur Mittelung betrachtet. Es folgt, dass die mittlere Leistung in dBA

$$N_{\text{dBA}} = \frac{1}{\mu_E^N - \mu_S^N + 1} \sum_{\mu=\mu_E^N}^{\mu_S^N} 20 \cdot \log_{10} \left(\bar{N}_{\text{veh}}(\mu) \cdot A_w(\mu) \right) \quad (7.21)$$

ist. Die A-Gewichtung ist definiert durch den folgenden mathematischen Zusammenhang,

$$A_w(\mu) = \frac{12200^2 \cdot f(\mu) \cdot 10^{0,1}}{(f(\mu)^2 + 20,6^2)(f(\mu)^2 + 12200^2)\sqrt{(f(\mu)^2 + 107,7^2)(f(\mu)^2 + 737,9^2)}} \quad (7.22)$$

wobei $f(\mu)$ der Frequenz an der Stützstelle μ entspricht und durch

$$f(\mu) = \mu \cdot f_S / N_{\text{FFT}} \quad (7.23)$$

definiert ist. Diese Gewichtung bezeichnet man auch als A-Bewertung des Schallpegels und soll das menschliche Gehör, welches unterschiedliche Frequenzen unterschiedlich laut wahrnimmt, nachbilden [DIN]. In Abbildung 7.3 ist die entsprechende A-Bewertungsfilterkurve in dB dargestellt.

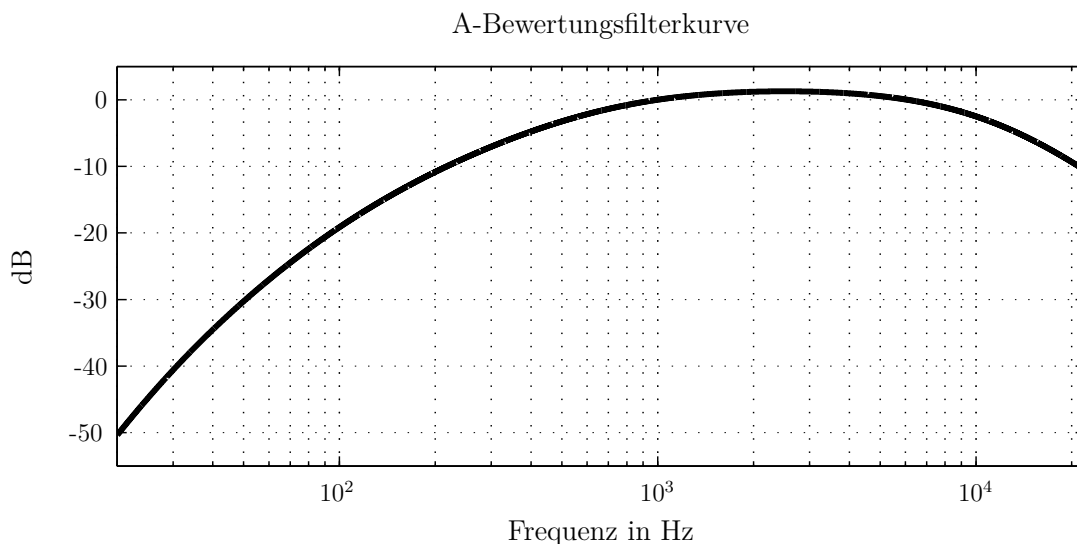


Abbildung 7.3: Übertragungskurve des A-Bewertungsfilters.

Anhand des Zusammenhanges aus (7.21) kann nun der Anstieg des Hintergrundgeräusches innerhalb der Fahrgastzelle bestimmt werden.

Spezialfall Fahrzeuginnenraum

Um das Fahrzeug kategorisieren zu können, muss innerhalb der Fahrgastzelle der Anstieg der Leistung des Hintergrundgeräusches geschätzt werden. Dazu muss zunächst das zu untersuchende Fahrzeug in ein beliebiges Umgebungsszenario versetzt werden. Dies kann auf zwei Arten durchgeführt werden: Zum einen mithilfe einer Umgebungssimulation, wie sie in Abschnitt 5.1 beschrieben worden ist, und zum anderen durch reale Testfahrten mit dem Fahrzeug. Das Hintergrundgeräusch wird dann unter Verwendung des zuvor beschriebenen Zusammenhanges (7.21) während zwei unterschiedlichen Geräuschkennszenarien ermittelt. Das erste Szenario entspricht dabei einem Fahrzeug im Stillstand mit laufendem

Motor. Innerhalb dieser Fahrzeugumgebung wird die Geräuschleistung $N_{\text{dBA}}^{\text{low}}$ geschätzt. Im Anschluss wird eine zweite Geräuschleistung $N_{\text{dBA}}^{\text{loud}}$ innerhalb eines Szenarios, welches einer Fahrt bei einer Geschwindigkeit von circa 150 km/h entspricht, geschätzt. Die Differenz

$$N_{\text{D}} = N_{\text{dBA}}^{\text{loud}} - N_{\text{dBA}}^{\text{low}} \quad (7.24)$$

bestimmt den Geräuschleistungsanstieg innerhalb der Fahrgastzelle. Durch diese Vorgehensweise wird der relative Anstieg der Geräuschleistung innerhalb der Fahrgastzelle bestimmt, welcher unabhängig von den Verstärkungseinstellungen der Aufnahmekomponenten, zum Beispiel Mikrofonvorverstärker ist. Dadurch müssen die Mikrofone nicht zwangsläufig vor der Messung in ihrer Leistung kalibriert werden. Um allerdings aussagekräftige Ergebnisse erzielen zu können, sollten die verwendeten Mikrofone einen möglichst tieffrequenten Bereich, zum Beispiel bis zu 50 Hz, abdecken. Der Grund hierfür ist das tieffrequent dominante Fahrzeuggeräusch, welches analysiert werden soll.

Um die Einflussgröße c_{N} mittels dieses Geräuschleistungsanstieges bestimmen zu können, sind zwei der bereits erwähnten Beispielfahrzeuge ausgewählt und während der Fahrt Geräuschaufnahmen durchgeführt worden [Pri14]. Für diese Zwecke sind Mikrofone verwendet worden, welche einen Übertragungsbereich von 50 Hz bis 17 kHz und eine Kugelcharakteristik besitzen. Als Beispielfahrzeuge ist die Mercedes-Benz S-Klasse, mit einer guten Dämmung der Fahrgastzelle und dadurch einem verhältnismäßig ruhigen Innenraum, und ein Multivan T5 von Volkswagen, welcher ein lautes Hintergrundgeräusch während der Fahrt aufweist, ausgewählt worden. Die aufgenommenen Signale sind mittels des Zusammenhanges aus (7.21) ausgewertet worden und die Ergebnisse in Tabelle 7.6 dargestellt.

Fahrzeug	Geräuschleistung bei Stillstand	Geräuschleistung bei 150 km/h	Geräuschleistungsanstieg
Mercedes-Benz S-Klasse	-72,3 dBA	-62,1 dBA	10,2 dB
Multivan T5 Volkswagen	-72,5 dBA	-52,4 dBA	20,1 dB

Tabelle 7.6: Empirisch ermittelter Leistungsanstieg des Hintergrundgeräusches in dBA bei verschiedenen Fahrzeugen.

Zur Berechnung dieser Werte sind in beiden Fahrzeugen Signale an den Ohrpositionen der Insassen auf der Rücksitzbank aufgenommen und analysiert worden. Diese Positionen sind ausgewählt worden, da dort im Stillstand der geringste Geräuscheinfluss durch den Motor entsteht und so die maximale Leistungsdifferenz bestimmt werden kann [Pri14]. Zusätzlich ist lediglich der Frequenzbereich von 50 Hz bis 11 kHz betrachtet worden, da dieser Bereich ausreicht den dominanten Anteil des Fahrzeuggeräusches zu bestimmen und gängige ICC-Systeme meist mit einer Abtastrate von 22,05 kHz arbeiten.

Zur Ermittlung der Einflussgröße auf das entsprechende Fahrzeugattribut, werden zunächst die, für die Herleitung vernachlässigten Mikrofonkanäle, wieder eingeführt. In diesem speziellen Fall werden einzig und allein die Ohrsignale aller Insassen betrachtet, sodass nur eine Abhängigkeit von m_K innerhalb von (7.24) entsteht. Damit folgt für (7.20), dass das Betragsspektrum

$$\bar{N}_{\text{veh}}^{m_K}(\mu) = \frac{1}{K_{\text{BG}}} \sum_{k=0}^{K_{\text{BG}}-1} |X^{(m_K)}(\mu, k)| \quad (7.25)$$

ist und $X_{m_K}(\mu, k)$ dem Spektrum des m_K -ten Mikrofons entspricht. Für die Berechnung des Leistungsanstieges ergibt sich unter Verwendung von Gleichung (7.21), dass

$$N_D(m_K) = N_{\text{dBA}}^{\text{loud}}(m_K) - N_{\text{dBA}}^{\text{low}}(m_K) \quad (7.26)$$

ist. Da der vermeidlich schlechteste Fall, bezogen auf den Anstieg der Geräuschleistung, gefunden werden soll, kann das Endergebnis dieses Verfahrens

$$N_I = \max_{m_K} \{N_D(m_K)\} \quad (7.27)$$

durch das Maximum aller Mikrofonkanäle ermittelt werden. Um einen Ergebnisraum für die Einflussgröße aufspannen zu können, werden die Grenzen $\delta_{\text{max}}^N = 20$ dB und $\delta_{\text{min}}^N = 10$ dB festgelegt. Im Anschluss wird die Einflussgröße als Endergebnis dieser Gruppe mithilfe von

$$c_N = \min \left\{ \max \left\{ 0, \frac{1}{\delta_{\text{max}}^N - \delta_{\text{min}}^N} (N_I - \delta_{\text{min}}^N) \right\}, 1 \right\} \quad (7.28)$$

berechnet.

7.1.3 Fahrzeugattribute

Zur Einstufung der Fahrzeugumgebung sind bereits zu Beginn dieses Kapitels, siehe 7.1, drei Attribute vorgeschlagen worden: Das Hintergrundgeräusch, die Wandler des ICC-Systems und die Fahrzeuggröße. Mithilfe der Ergebnisse der beiden zuvor beschriebenen Verfahren, werden anschließend die entsprechenden Fahrzeugattribute ermittelt. In einem weiteren Schritt werden diese Attribute verwendet, um die Gesamtqualität des ICC-System zu bestimmen. Die Gesamtqualität wird gegen Ende dieses Kapitels unter Abschnitt 7.4 beschrieben. Im folgenden Abschnitt wird der Zusammenhang zwischen den Ergebnissen der bereits beschriebenen Verfahren und den Fahrzeugattributen genauer beleuchtet. Zu Beginn wird auf die Bedeutung der einzelnen Fahrzeugattribute eingegangen. Abschließend wird die Bestimmung der einzelnen Attribute vorgestellt.

Hintergrundgeräusch

Das Fahrzeugattribut „Hintergrundgeräusch“ repräsentiert im Allgemeinen die Eigenschaften des, während der Fahrt auftretenden, Geräusches. Dabei ist die grundlegende Idee, dass, je lauter das Hintergrundgeräusch während der Fahrt, desto schwieriger die Situation für die Algorithmen eines ICC-Systems. Dabei beeinflusst das auftretende Hintergrundgeräusch nicht ausschließlich einen Teil der Signalverarbeitung, sondern mehrere. So ist zum Beispiel die gewählte Verstärkung eines ICC-Systems immer abhängig von dem Geräusch, denn je lauter das Geräusch, desto mehr Nutzsignalleistung wird benötigt, um den SNR am Zuhörerohr zu erhöhen, siehe Abschnitt 2. Durch die höhere Leistung des Geräusches bzw. der Verstärkung des Nutzsignales, wird ebenfalls die Rückkopplung des Lautsprecher Signals in das Mikrofon erhöht. Dadurch wird die Problematik der Rückkopplungsunterdrückung verschärft, welche algorithmisch wieder gelöst werden muss. Ein unter Umständen weitaus ausgeklügelter Algorithmus zur Rückkopplungsunterdrückung muss verwendet werden. In diesem Fall sollte dieser, falls er effektiv funktioniert, stärker bei einer Berechnung der Gesamtqualität ins Gewicht fallen. Zusätzlich verringert ein Hintergrundgeräusch mit höherer Leistung den SNR am ICC-Mikrofon. Dies wiederum kann die Extraktion des Nutzsignales deutlich erschweren. Bei einem geringen SNR am Eingang des ICC-Systems müssen auch die Geräuschreduktionsalgorithmen deutlich effektiver gestaltet werden.

Wandler des ICC-Systems

Die Komplexität einer Fahrzeugumgebung in Bezug auf die Signalverarbeitung eines ICC-Systems ist auch abhängig von den verwendeten Wandlern und ihrer Position innerhalb der Fahrgastzelle. Beide Eigenschaften liefern Informationen über die Qualität des Eingangs- und des Ausgangssignals eines ICC-Systems. Um die Charakteristika, zum Beispiel den Frequenzgang der Wandler, der verwendeten Mikrofone und Lautsprecher ermitteln zu können, müsste in das ICC-System elektrisch eingegriffen werden. In Abschnitt 5 ist bereits erwähnt worden, dass es Zielsetzung dieser Evaluierung ist, das ICC-System als *black box* zu betrachten. Deswegen sind nicht die Frequenzgänge der Wandler mitberücksichtigt, sie könnten jedoch eventuell in zukünftigen Betrachtungen eine Rolle spielen. Nichtsdestotrotz bietet bereits der Abstand zwischen dem Sprechermund und dem ICC-Mikrofon bzw. der Abstand zwischen dem ICC-Lautsprecher und dem Zuhörerohr ein gutes Indiz für die Komplexität der Umgebung und die unter Umständen entstehende Probleme eines ICC-Systems. So benötigt beispielsweise ein ICC-System, welches weit entfernte Mikrofone verwendet, deutlich effektivere Algorithmen, um das Nutzsignal zu bestimmen. Auch die Distanz zwischen den ICC-Lautsprechern und dem Zuhörerohr kann Rückschlüsse auf die Komplexität der Fahrzeugumgebung geben. Je größer diese ist, desto mehr Verstärkung durch das ICC-System wird benötigt. Dadurch erhöht sich jedoch auch die Gefahr einer Rückkopplung, welche wiederum vom ICC-System aufgefangen werden muss. Deswegen

ist eine solche Situation für ein ICC-System komplexer und erfordert weitaus fundiertere und komplexere Algorithmen zur Signalverarbeitung.

Fahrzeuggröße

Innerhalb dieses Attributes soll die Eigenschaft der Größe eines Fahrzeuges festgehalten werden. Die Größe eines Fahrzeuges bildet einen weiteren Grund für die Entwicklung und Implementierung eines ICC-Systems ab. Je weiter der Zuhörer vom Sprecher entfernt ist, desto geringer wird der SNR ab Zuhörerohr und somit auch die Sprachverständlichkeit. Deswegen ist diese Distanz zwischen dem Mund des Sprechers und dem Ohr des Zuhörers ein Indikator für die Notwendigkeit eines ICC-Systems, aber auch ein Indikator für die Komplexität der Verarbeitung eines ICC-Systems. Denn je weiter weg der Zuhörer sich befindet, desto mehr muss der Verlust der Sprachsignalleistung durch das ICC-System kompensiert werden. Des Weiteren gibt die Distanz zwischen dem ICC-Lautsprecher und dem ICC-Mikrofon einen Eindruck der Größe eines Fahrzeuges. Zusätzlich kann eine Einschätzung des Rückkopplungspfades mit dieser Messung vorgenommen werden. Je geringer diese Distanz ist, desto einfacher entstehen Rückkopplungen, welche für die einwandfreie Verwendung eines ICC-Systems kompensiert oder unterdrückt werden müssen.

Berechnung der Attribute

Zur Ermittlung der Fahrzeugattribute werden zunächst alle Ergebnisse der einzelnen Verfahren aus der Evaluierungsgruppe „Fahrzeugeigenschaften“ zusammengetragen. Im rechten Teil der Abbildung 7.4 sind alle betrachteten Verfahren, welche Einfluss auf die Fahrzeugattribute haben, dargestellt. Das Ergebnis der Hintergrundgeräuschschätzung geht direkt in das Fahrzeugattribut „Hintergrundgeräusch“ ohne eine Umrechnung ein. Die Ergebnisse der Distanzmessung zwischen dem Spechermund und dem dem ICC-Mikrofon $c_D^{(0)}$ und dem ICC-Lautsprecher sowie dem Zuhörerohr $c_D^{(1)}$ werden verwendet, um das Attribut „Wandler des ICC-Systems“ zu bestimmen. Das letzte Attribut Fahrzeuggröße wird mithilfe der Distanzmessungen zwischen dem ICC-Lautsprecher und dem ICC-Mikrofon, sprich der Messung des Rückkopplungspfades, sowie zwischen dem Spechermund und dem Zuhörerohr, dem Direktschallpfad, bestimmt.

Die entsprechenden Einflussgrößen werden zu einer Matrix

$$\mathbf{C}_V = \begin{bmatrix} 0 & c_N & c_D^{(0)} & c_D^{(1)} & c_D^{(2)} & c_D^{(3)} & 1 \\ \gamma & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (7.29)$$

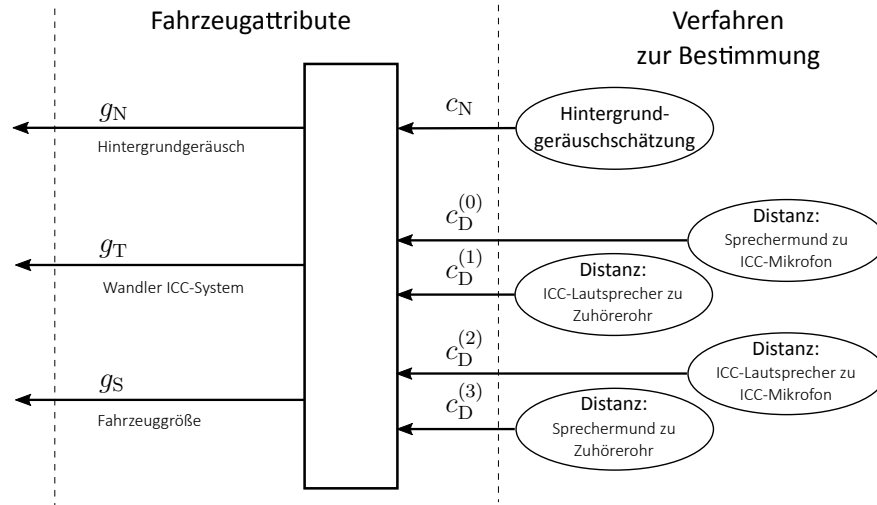


Abbildung 7.4: Verknüpfung der Ergebnisse der Evaluierungsgruppe Fahrzeug-eigenschaften zu den einzelnen Fahrzeugattributen.

kombiniert. Zur Begrenzung der resultierenden Fahrzeugattribute auf einen minimal Wert ungleich Null wird $\gamma = 0,1$ gewählt. Anschließend soll mittels der folgenden Gewichtungs-matrix

$$\mathbf{B} = \begin{bmatrix} 1 & 1 & 1 \\ b_N & 0 & 0 \\ 0 & b_T^{(0)} & 0 \\ 0 & b_T^{(1)} & 0 \\ 0 & 0 & b_S^{(0)} \\ 0 & 0 & -b_S^{(1)} \\ 0 & 0 & b_S^{(1)} \end{bmatrix}, \quad (7.30)$$

die einzelnen Fahrzeugattribute bestimmt werden. Dabei muss das Ergebnis der Distanz zwischen Sprechermund und Zuhörerohr $c_D^{(3)}$ invertiert werden, um so sicherzustellen, dass ein kleines Fahrzeug einem Attribut „Fahrzeuggröße“ $g_S = 1$ entspricht. Die Gewichte müssen dabei einige Kriterien erfüllen. Zum einen müssen sie alle

$$b_N, b_T^{(i)}, b_S^{(i)} \in [0,1] \quad (7.31)$$

erfüllen und zum anderen muss die Summe der einzelnen fahrzeugattributspezifischen Gewichte gleich eins sein:

$$b_N = \sum_{i=0}^1 b_T^{(i)} = \sum_{i=0}^1 b_S^{(i)} = 1. \quad (7.32)$$

Dadurch wird gewährleistet, dass die entsprechenden Einflussgrößen als gewichtete Summe zu einem Fahrzeugattribut kombiniert werden. Im Falle des Attributes „Fahrzeuggeräusch“ ergibt sich eine Gewicht $b_N = 1$, da keine weiteren Faktoren dieses Attribut beeinflussen.

Die beiden anderen Fahrzeugattribute werden durch die entsprechenden Gewichte und die Einflussgrößen ermittelt.

Für die Bestimmung der einzelnen Fahrzeugattribute ergibt sich nun folgende Matrixdarstellung:

$$\mathbf{g} = [g_N, g_T, g_S]^T = M_{\max}\{\mathbf{C}_V \cdot \mathbf{B}\}^T. \quad (7.33)$$

Dabei entspricht die Operation $M_{\max}\{\cdot\}$ dem Spaltenmaximum, welches im Allgemeinen wie folgt definiert ist:

$$M_{\max}\{\mathbf{A}\} = \max_i\{a_{ij}\}. \quad (7.34)$$

Diese Operation ermöglicht eine Begrenzung des Ergebnisraums der Fahrzeugattribute „Hintergrundgeräusch“ g_N , „Fahrzeuggröße“ g_S und „Wandler des ICC-Systems“ g_T von 0,1 bis 1. Die Obergrenze kann durch die Begrenzung der einzelnen Einflussgrößen zwischen Null und Eins gewährleistet werden. Die Begrenzung auf einen Mindestwert von 0,1 garantiert, dass bei einer späteren Gewichtung der einzelnen Qualitätsergebnisse alle Ergebnisse mitberücksichtigt werden und nicht durch eine Multiplikation mit Null vernachlässigt werden. Dabei steht der Wert 0,1 für eine eher einfache Fahrzeugumgebung in Bezug auf das entsprechende Attribut wohingegen 1 einer komplexen Umgebung entspricht. Im Detail bedeutet dies für das Fahrzeugattribut „Hintergrundgeräusch“, dass $g_N = 1$ einem sehr lauten und $g_N = 0,1$ einem sehr leisen Fahrzeug gleichkommt. Ein sehr kleines Fahrzeug würde, unter Verwendung der hier dargestellten Definition, dem Attribut „Fahrzeuggröße“ $g_S = 1$ und ein sehr großes Fahrzeug $g_S = 0$ entsprechen. Das bedeutet, dass ein kleineres Fahrzeug eine deutlich komplexere Situation, beispielsweise für die Rückkopplungskompensation eines ICC-Systems, darstellt als ein großes Fahrzeug. Das Fahrzeugattribut „Wandler des ICC-Systems“ entspricht $g_T = 1$, wenn sich die Wandler sehr weit weg von den Passagieren im Fahrzeug befinden und $g_T = 0$ bei einer sehr kleinen Distanz.

In Abbildung 7.5 ist beispielhaft ein zu evaluierendes Fahrzeug im linken Teil der Abbildung dargestellt. Es handelt sich hierbei um ein sehr lautes Fahrzeug, welches jedoch über sehr nahe Wandler, zum Beispiel über Gurtmikrofone und Kopfstützenlautsprecher, verfügt und zudem als eher klein gemessen worden ist. Als Beispiel zum Vergleich dieser Ergebnisse sind im rechten Teil der Abbildung 7.5 drei der vier Beispielfahrzeuge dargestellt. In der obersten Grafik ist ein Opel Corsa D mit seinen Fahrzeugattributen gezeigt. Dieses Fahrzeug verfügt über Wandler in einer mittleren Distanz und ist als mittellautes, kleines Fahrzeug klassifiziert worden. In der mittleren Grafik sind die Fahrzeugattribute eines, im Vergleich zum Corsa, deutlich größeren und leiseren Fahrzeuges - einer Limousine - dargestellt. Die unterste Grafik zeigt das Ergebnis eines Volkswagen T5, der als sehr laut, groß und mit weit entfernten ICC-System Wandlern klassifiziert worden ist. Es wird deutlich, dass das zu evaluierende Fahrzeug in keine der drei dargestellten Beispielfahrzeugklassen passt. Es handelt sich bei diesem Fahrzeug um das vierte Beispielfahrzeug einer Mercedes-Benz E-Klasse (Kabriolett).

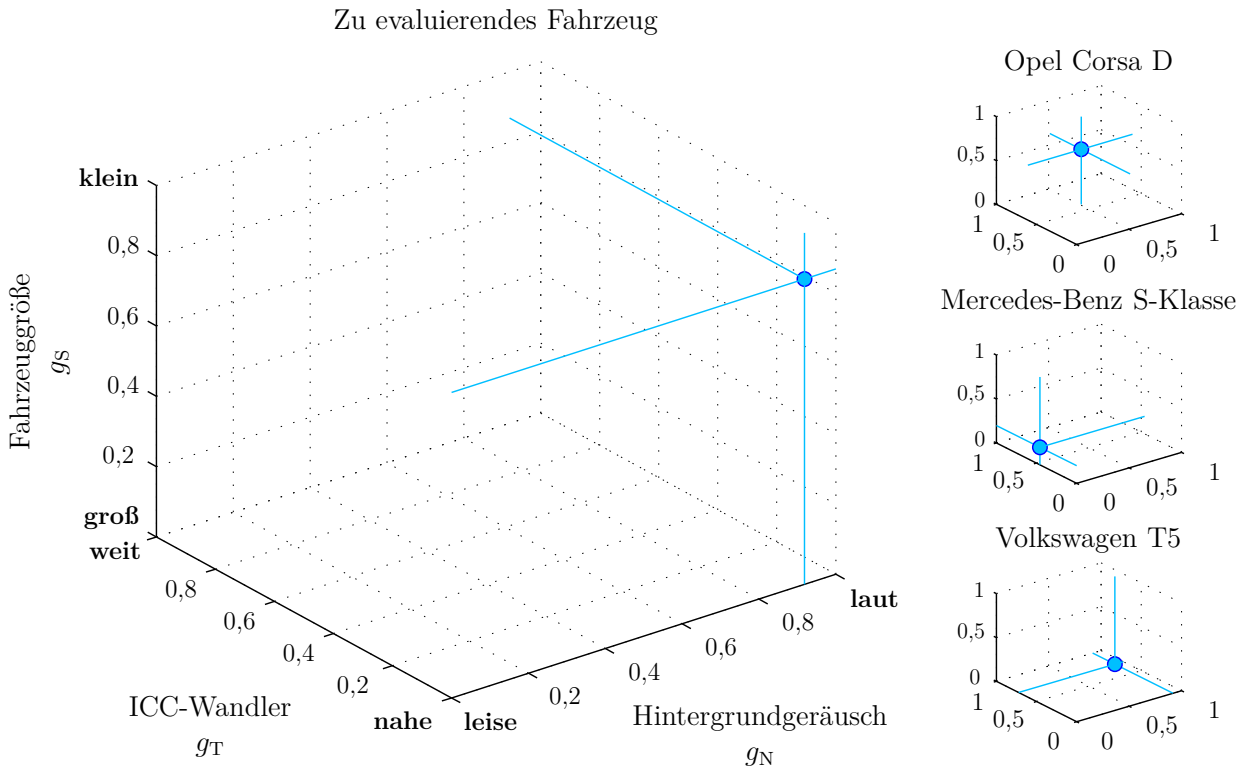


Abbildung 7.5: Beispielhaftes Ergebnis der Bestimmung der drei Fahrzeugattribute von eines beliebigen Fahrzeuges und drei unterschiedlichen Beispielfahrzeugen.

Bei den hier aufgeführten Verfahren handelt es sich lediglich um eine erste Auswahl für die Bestimmung der Attribute. In zukünftigen Betrachtungen könnten auch andere Verfahren und auf eine andere Weise realisierte Methoden Anwendung finden. Selbstverständlich müssten diese entsprechend in die Fahrzeugattribute integriert werden, um so zu einem Klassifizierungsergebnis des Fahrzeuges zu gelangen.

7.2 Verhalten des ICC-Systems

Die Qualität eines sprachverarbeitenden oder allgemein eines komplexen signalverarbeitenden Systems, ist natürlich stark mit dem Verhalten der Einzelkomponenten und deren Algorithmen verknüpft. Dabei wird angenommen, dass Sprachverbesserungssysteme, wie zum Beispiel ein ICC-System, aus einer Verkettung und Kombination einer Vielzahl von Einzelmodulen bestehen. Die Einzelmodule erfüllen unterschiedlichste Aufgaben innerhalb des Gesamtsystems. Dies führt zu der Annahme, dass ein Gesamtsystem nur so gut funktionieren kann wie das qualitativ schlechteste Einzelmodul. Ein einfaches Beispiel hierfür ist die Geräuschunterdrückung, wie sie in Abschnitt 2.1 beschrieben ist. Eine sehr gut umgesetzte Geräuschunterdrückung, kann mit einer qualitativ weniger hochwertigen Geräuschschätzung in der Gesamtbetrachtung zu einem eher schlechten System führen.

Innerhalb eines ICC-Systems können weitere, ähnliche Verknüpfungen der Qualität einzelner Module bzw. Verarbeitungsschritte mit der Gesamtqualität gefunden werden. Da jedoch für die ICC-Systemevaluierung der Anspruch eines *black-box testing* zu Beginn gestellt worden ist, siehe Abschnitt 3, können die Einzelmodule nicht evaluiert werden. Lediglich die Reaktion eines beliebigen ICC-Systems auf ein spezielles Anregungssignal kann ausgewertet werden. Dennoch lässt sich anhand des Wissens über Anregungs- und Ausgangssignal eines ICC-Systems das Verhalten des selbigen im Detail beobachten und Rückschlüsse auf die Qualität der ICC-Implementierung durchführen.

Im folgenden Kapitel werden zu diesem Zweck verschiedene Verfahren, welche unterschiedliche Verhaltensweisen ermitteln, vorgestellt und in Zusammenhang zur Qualität eines ICC-Systems gestellt. Zur Bestimmung der Qualität fließen nicht nur die auditiven Evaluierungsergebnisse aus Kapitel 6 ein, sondern auch die Erkenntnisse über die Fahrzeugeigenschaften aus Kapitel 7.1.3. Innerhalb aller im folgenden vorgestellten Evaluierungsverfahren, wird die in Kapitel 5 beschriebene Evaluierungsumgebung verwendet. Zur Herleitung der Verfahren wird das in Abbildung 5.1 dargestellte Evaluierungssystem angenommen.

7.2.1 Unterstützung durch das ICC-System

Hauptziel eines ICC-Systems ist es, den Pegel des Sprachsignales, ausgesendet vom Sprecher, am Zuhörerohr zu erhöhen. Diese Grundidee ist bereits in Abschnitt 2.1 beschrieben und weitere Details zur Umsetzung beschrieben worden. Die Erhöhung des Signalpegels wird erreicht, indem das aufgenommene Sprachsignal verarbeitet und verstärkt wieder ausgegeben wird. Diese Verstärkung geschieht in Abhängigkeit des aktuellen Hintergrundgeräusches und ist für gewöhnlich abhängig von der Frequenz. Die Systemunterstützung bzw. Verstärkung ist hochgradig qualitätsrelevant, da dies die Hauptaufgabe eines ICC-Systems ist. Zur Bestimmung der frequenzabhängigen Verstärkung des Nutzsignals durch

Teile der in diesem Kapitel beschriebenen Methoden sind bereits in einer ersten Grundidee in [TSWL] publiziert worden.

das ICC-System wird im Folgenden ein Verfahren basierend auf der Schätzung des quadratischen Betragsfrequenzgangs vorgestellt. Abschließend wird das ermittelte Ergebnis in eine entsprechende Qualitätsschätzung überführt.

Entwurf des Verfahrens

Im allgemeinen Fall eines linearen zeitinvarianten Systems, wie es in Abbildung 7.6 dargestellt ist, bildet der Betragsfrequenzgang $|H(e^{j\Omega})|$ das frequenzabhängige Dämpfungs- bzw. Verstärkungsverhalten des betrachteten Systems ab.

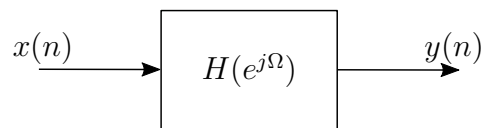


Abbildung 7.6: Einfaches lineares zeitinvariantes System $H(e^{j\Omega})$ mit Ein- und Ausgang $x(n)$ bzw. $y(n)$) dessen Übertragungsverhalten geschätzt werden soll.

Unter Verwendung der Leistungsdichtespektren des Ausgangssignals $S_{yy}(e^{j\Omega})$ und des Eingangssignals $S_{xx}(e^{j\Omega})$, kann der quadratische Betragsfrequenzgang des Systems durch

$$|\hat{H}(e^{j\Omega})|^2 = \frac{S_{yy}(e^{j\Omega})}{S_{xx}(e^{j\Omega})} \quad (7.35)$$

geschätzt werden [KK98, S. 64]. Zu berücksichtigen ist, dass nur innerhalb jener Frequenzbereiche, welche ausreichend vom Eingangssignal angeregt worden ist, eine Aussage über den Betragsfrequenzgang getätigt werden kann. Deswegen bietet sich zur Schätzung des quadratischen Betragsfrequenzgangs weißes Rauschen als Anregungssignal an.

Die Anregung eines ICC-Systems mit ausschließlich weißem Rauschen würde zu einem ungünstigen und nicht erwartungsgemäßen Verhalten des selbigen führen. Beispielsweise würde die Geräuschschätzung stetig ansteigen und das Nutzsignal bzw. Testsignal als unerwünschten Signalanteile ansehen und diesen immer weiter dämpfen. Die Schätzung des Leistungsdichtespektrums des Eingangssignals $x(n)$ der ICC-Systemevaluierung, würde unter Umständen zu wenig Leistung besitzen, um überhaupt eine Aussage über das Übertragungsverhalten tätigen zu können. Zudem wäre bei einer solchen Anregung das ICC-System in keinem Fall zeitinvariant, da sich die internen Parameter immer weiter verändern. Diese Invarianz des Systems gegenüber der Zeit ist eine Voraussetzung, wenn der in Gleichung (7.35) gegebene Zusammenhang verwendet wird. Um dennoch ein nahezu zeitinvariantes System annehmen zu können, soll die Trägheit des ICC-Systems ausgenutzt werden. Bekanntermaßen reagieren die Hauptkomponenten eines ICC-Systems, wie beispielsweise die Geräuschreduktion, nicht instantan auf eine Änderung des Eingangssignals, sondern benötigen eine gewisse Anpassungszeit, siehe Abschnitt 2.1. Innerhalb dieser Anpassungszeit kann das ICC-System als nahezu zeitinvariant angenommen und der zuvor

beschriebene Ansatz (7.35) dennoch verwendet werden. Um dies zu bewerkstelligen, wird ein spezielles Testsignal zur Messung der Systemunterstützung verwendet.

In Abbildung 7.7 ist das Testsignal für einen ausgewählten Zeitabschnitt dargestellt. Zu sehen ist, dass das Signal zunächst aus Sprachpassagen und kurzen (circa 100 ms) Rauschpulsen zusammengesetzt ist. Die Rauschpulse sind in Abbildung 7.7 durch eine grüne Markierung grafisch hervorgehoben. Die Sprachsignale sind breitbandig aufgenommen worden und enthalten den auftretenden Lombard-Effekt, siehe Abschnitt 5.2. Die Sprachanteile ermöglichen es, das betrachtete ICC-System, zusätzlich neben der Geräuschsimulation, in einen gewünschten operativen, realistischen Zustand ohne direkten, internen Eingriff zu bringen. Die Dauer der Rauschpulse ermöglicht die Annahme der Zeit-Invarianz des ICC-Systems.

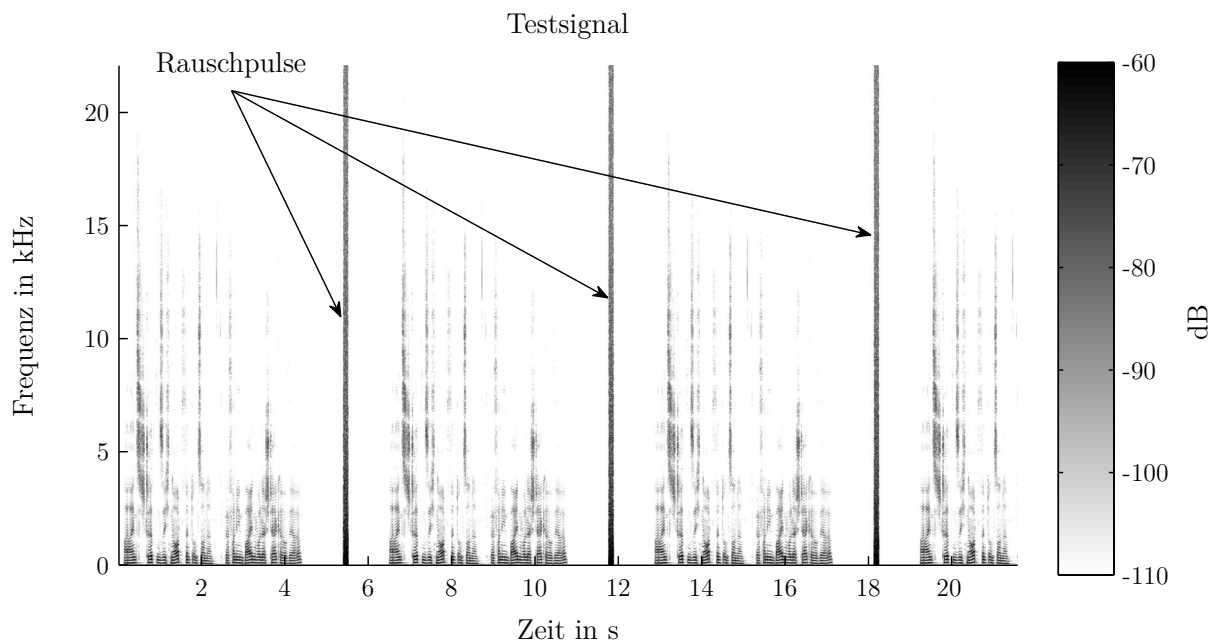


Abbildung 7.7: Für die Bestimmung der Systemunterstützung verwendetes Testsignal. Dieses besteht aus Sprachanteilen und einzelnen, bandbegrenzten Rauschpulsen mit einer Dauer von circa 100 ms.

Um nun die Verstärkung des ICC-System schätzen zu können, müssen zunächst die Leistungsdichtespektren am Ein- und Ausgang des ICC-Systems während der Rauschpulse geschätzt werden. In Abbildung 7.7 sind nur drei Rauschpulse abgebildet, zur Bestimmung des Leistungsdichtespektren werden mindestens 9 Rauschpulse angeregt, um so eine bessere Näherung zu erreichen. Das Leistungsdichtespektrum des Referenzsignales $y(n)$, ausgegeben durch den künstlichen Sprecher, wird durch

$$S_{yy}^G(\mu) = \frac{\sum_{k=0}^{K-1} |Y(\mu, k)|^2 \cdot a_y^G(\mu, k)}{\sum_{k=0}^{K-1} a_y^G(\mu, k)} \quad (7.36)$$

bestimmt, wobei

$$a_y^G(\mu, k) = \begin{cases} 1, & \text{wenn } Y(\mu, k) \text{ einem Rauschpuls entspricht,} \\ 0, & \text{sonst.} \end{cases} \quad (7.37)$$

In Abschnitt 5.3 ist bereits die Blockverarbeitung innerhalb des Frequenzbereiches eingeführt worden und die Berechnung des Spektrums $Y(\mu, k)$ in Abhängigkeit des Frequenzbandes μ und des Rahmens k in (5.2) festgehalten. Analog kann das Leistungsdichtespektrum des Ohrsignals $x(n)$ (siehe Abbildung 5.1) mit

$$S_{xx}^G(\mu) = \frac{\sum_{k=0}^{K-1} |X(\mu, k)|^2 \cdot a_x^G(\mu, k)}{\sum_{k=0}^{K-1} a_x^G(\mu, k)}, \quad (7.38)$$

wobei

$$a_x^G(\mu, k) = \begin{cases} 1, & \text{wenn } X(\mu, k) \text{ einem Rauschpuls entspricht,} \\ 0, & \text{sonst,} \end{cases} \quad (7.39)$$

bestimmt werden. Im Anschluss werden nun beide Schätzungen der Leistungsdichtespektren ins Verhältnis gesetzt, um so eine Abschätzung des allgemeinen Systemverhaltens zu gewinnen. Zunächst wird das Leistungsverhältnis zwischen Aus- und Eingang des Gesamtsystems bei deaktiviertem ICC-System mit

$$R^{(\text{off})}(\mu) = \frac{S_{xx}^{G,(\text{off})}(\mu)}{S_{yy}^G(\mu)} \quad (7.40)$$

bestimmt. Unter Berücksichtigung des in Abschnitt 5.4 hergeleiteten theoretischen Modells, im Speziellen des Zusammenhanges (5.21), kann eine theoretische Betrachtung der Gleichung (7.40) vorgenommen werden. Die hergeleiteten Zusammenhänge (5.21) für die normierte Frequenz Ω sind auf die einzelne Frequenzbänder μ durch Diskretisierung mit $\Omega_\mu = 2\pi\mu/N_{\text{FFT}}$, übertragen worden. Zusätzlich findet keine Übertragung über das ICC-System statt, so dass $H_{\text{tot}}(\mu) = 0$ angenommen werden kann. Das theoretische Leistungsverhältnis zwischen Ein- und Ausgangsleistungsdichtespektrum ist definiert durch:

$$R^{(\text{off})}(\mu) = \frac{S_{yy}^G(\mu) \cdot |H_{\text{TE}}(\mu)|^2 + S_{bb}(\mu)}{S_{yy}^G(\mu)} = |H_{\text{TE}}(\mu)|^2 + \frac{S_{bb}(\mu)}{S_{yy}^G(\mu)}. \quad (7.41)$$

Es wird deutlich, dass das Leistungsverhältnis $R^{(\text{off})}(\mu)$ vom Übertragungsverhalten des direkten Weges zwischen Sprecher und Zuhörer $H_{\text{TE}}(\mu)$ und dem Verhältnis der Leistungsdichteschätzung des Hintergrundgeräusches $S_{bb}(\mu)$ und der Rauschpuls des Testsignals $S_{yy}^G(\mu)$ abhängt.

Im nächsten Schritt wird das Leistungsverhältnis zwischen Aus- und Eingang der ICC-Systemevaluierung mit aktivem ICC-System bestimmt. In gleicher Weise wie im Fall eines deaktivierten ICC-Systems werden die zuvor beschriebenen Zusammenhänge des mathematischen Modells (5.21) aus Kapitel 5.4 herangezogen, um das Übertragungsverhalten zu analysieren. Der theoretische Zusammenhang des Leistungsverhältnisses zwischen Aus- und Eingang lässt sich wie folgt beschreiben:

$$R^{(\text{on})}(\mu) = |H_{\text{TE}}(\mu) + H_{\text{TM}}(\mu)H_{\text{tot}}(\mu)H_{\text{LE}}(\mu)|^2 + |H_{\text{tot}}(\mu)H_{\text{LE}}(\mu)|^2 \frac{S_{bb}(\mu)}{S_{yy}^{\text{G}}(\mu)} + \frac{S_{bb}(\mu)}{S_{yy}^{\text{G}}(\mu)}. \quad (7.42)$$

Neben der Übertragung des Testsignals $S_{yy}^{\text{G}}(\mu)$ über den Weg des Direktschalls $H_{\text{TE}}(\mu)$, wird es zusätzlich über das ICC-System inklusive des Rückkopplungspfads $H_{\text{tot}}(\mu)$ übertragen. Zusätzlich muss das Signal vom Sprecher zum ICC-System $H_{\text{TM}}(\mu)$ und vom ICC-System zum Zuhörer $H_{\text{LE}}(\mu)$ übertragen werden, um den Übertragungspfad des ICC-Systems zu vervollständigen. Der zweite Summand setzt den Anteil des Hintergrundgeräusches $S_{bb}(\mu)$, welcher über das ICC-System übertragen wird, ins Verhältnis zum Leistungsdichtespektrum des Testsignals. Der letzte additive Anteil entspricht dem Verhältnis der Leistungen zwischen Hintergrundgeräusch $S_{bb}(\mu)$ und dem geschätzten mittleren Leistungsdichtespektrum der Rauschpulse des Testsignals $S_{yy}^{\text{G}}(\mu)$. Mittels der Systemunterstützung wird der Unterschied zwischen Direktschall und dem Schallereignis, welches vom ICC-System ausgesendet wird, bestimmt. Die Systemunterstützung kann nur ermittelt werden, wenn der Einfluss des Verhältnisses zwischen der Leistungsdichte des Hintergrundgeräuschs und des Referenzgeräuschpulses in Zusammenhang (7.41) wie auch in (7.42) minimiert wird. Dazu muss das Verhältnis

$$\frac{S_{bb}(\mu)}{S_{yy}^{\text{G}}(\mu)} \rightarrow 0 \quad (7.43)$$

gehen. Diese Annahme kann getätigt werden, wenn gilt dass

$$S_{yy}^{\text{G}}(\mu) \gg S_{bb}(\mu) \quad (7.44)$$

ist. Da das Hintergrundgeräusch innerhalb eines vorgegeben Szenarios nicht veränderbar ist, muss zur Erfüllung der Annahme das Testsignal entsprechend angepasst werden.

In Abbildung 7.8 ist beispielhaft das mittlere Leistungsdichtespektrum eines Rauschpulses des Testsignals aus Abbildung 7.7 und eines Fahrzeuggeräusches dargestellt. Bei dem Fahrzeuggeräusch handelt es sich um ein Geräusch, aufgenommen in einer Mercedes-Benz S-Klasse bei einer Geschwindigkeit von 150 km/h. Zunächst ist zu sehen, dass es sich bei den verwendeten Rauschpulsen um ein farbiges Rauschen handelt. Dieses farbiges Geräusch bietet den Vorteil, dass es, ähnlich wie das Fahrzeuggeräusch, tieffrequent dominant ist und somit die Annahme aus Gleichung (7.44) auch dort realisierbar ist. Dennoch wird im Bereich zwischen 0 und circa 100 Hz nur eine logarithmische Differenz von maxi-

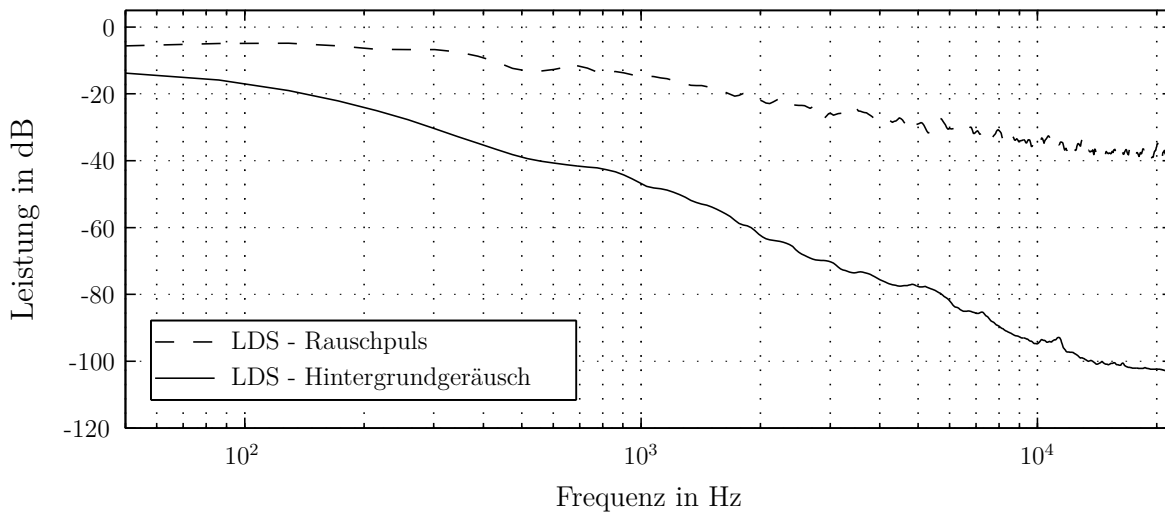


Abbildung 7.8: Mittleres Leistungsdichtespektrum (LDS) eines Rauschpulses und eines Fahrzeuggeräusches während der Fahrt bei einer Geschwindigkeit von 150 km/h mit einer Mercedes-Benz S-Klasse.

mal 10 dB erzielt, was bedeutet, dass in diesem Bereich die Annahme nicht erfüllt ist. Da jedoch ein gängiges ICC-System nur jenen Frequenzbereich verarbeitet, in dem Sprache angeregt wird, kann und sollte dieser Bereich zwischen 0 und 100 Hz bei der Auswertung der Systemunterstützung außen vor gelassen werden. Hierdurch kann die Annahme für alle betrachteten Frequenzen als erfüllt angesehen und die Beiträge des Hintergrundgeräusches in (7.41) und (7.42) können somit vernachlässigt werden.

Zur Bestimmung der Systemunterstützung des ICC-Systems wird das Leistungsverhältnis zwischen $R^{(\text{on})}(\mu)$ und $R^{(\text{off})}(\mu)$ gebildet. Es folgt für die frequenzabhängige Systemunterstützung, dass

$$R_S(\mu) = \frac{R^{(\text{on})}(\mu)}{R^{(\text{off})}(\mu)} \quad (7.45)$$

ist. Durch Einsetzen von (7.42) und (7.41) in (7.45) und unter der Annahme von (7.43), ergibt sich für die logarithmische Systemunterstützung, in Abhängigkeit des Frequenzbandes μ , folgender theoretischer Zusammenhang:

$$R_S(\mu) = 10 \cdot \log_{10} \left(\frac{|H_{\text{TE}}(\mu) + H_{\text{TM}}(\mu)H_{\text{tot}}(\mu)H_{\text{LE}}(\mu)|^2}{|H_{\text{TE}}(\mu)|^2} \right). \quad (7.46)$$

Hier wird nun deutlich, dass sich die beiden quadratischen Betragsfrequenzgänge, mathematisch nur durch den zusätzlichen Pfad des ICC-Systems unterscheiden, wobei, wie bereits erwähnt, auch die Pfad von den Insassen zum ICC-System und vice versa mit berücksichtigt werden. Im Detail bedeutet dies, dass die Systemunterstützung auch von den verwendeten Wandlern und der akustischen Übertragung über den Raum abhängt.

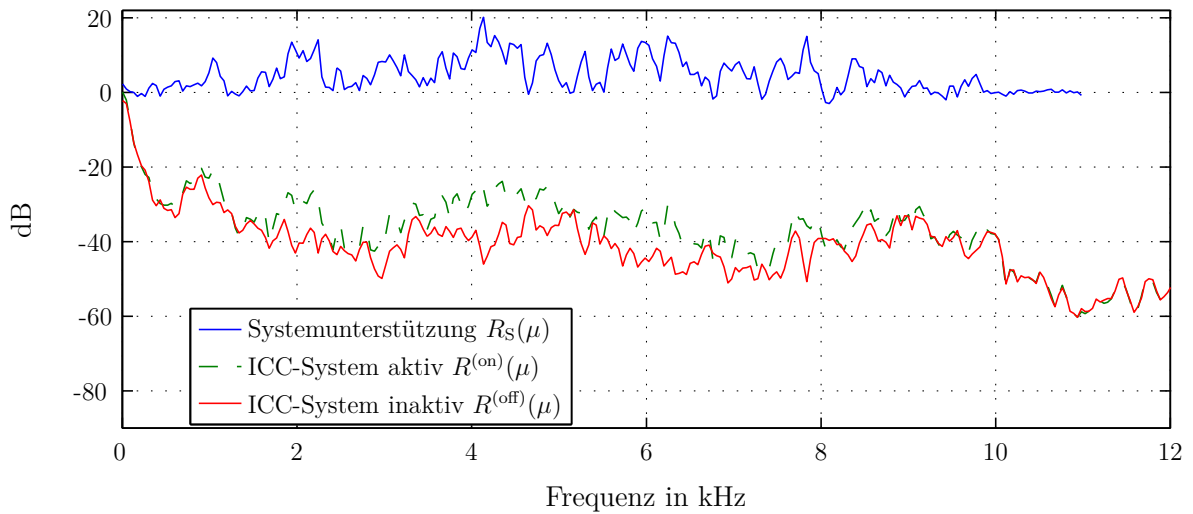


Abbildung 7.9: Die geschätzten logarithmischen Leistungsverhältnisse mit (grün gestrichelt) und ohne (rot) die Unterstützung eines ICC-Systems und die resultierende, frequenzabhängige Systemunterstützung (blau).

In Abbildung 7.9 sind die geschätzten logarithmischen Leistungsverhältnisse mit (grün gestrichelt) und ohne ICC-System (rot) dargestellt. Die Differenz dieser beiden Kurven entspricht der blauen Kurve und der frequenzabhängigen Systemunterstützung $R_S(\mu)$. Hier ist nun deutlich zu sehen, dass die Unterstützung durch das ICC-System nur innerhalb eines vordefinierten Frequenzbereiches durchgeführt wird. In den tiefen Frequenzen so wie in den höheren Frequenzen, sind die beiden Kurven $R^{(\text{on})}(\mu)$ und $R^{(\text{off})}(\mu)$ quasi identisch. Nur in einem Bereich zwischen 100 Hz und 10 kHz lässt sich eine Differenz der beiden Kurven feststellen. Wie bereits in Kapitel 2.1 erwähnt, liegt dieser Bereich aufgrund einer gängigen Parametrierung zwischen 100 Hz und 11 kHz. Die 11 kHz werden aufgrund der verwendeten Abtastrate von 22,05 kHz festgelegt und die 100 Hz als Untergrenze durch das in der Vorverarbeitung verwendete Hochpassfilter. Die Untergrenze in diesem Beispiel liegt ebenfalls bei circa 100 Hz, jedoch liegt die Obergrenze der Systemunterstützung bei 10 kHz. Grund hierfür sind möglicherweise die verwendeten Wandler, welche nur Frequenzen bis zu 10 kHz verarbeiten.

Da nur dieser, vom ICC-System verarbeitete, Bereich für die Evaluierung relevant ist, werden die logarithmischen Ergebnisse aus (7.45) über den vom ICC-System verarbeiteten Frequenzbereich von μ_{Start} bis μ_{End} mithilfe des folgenden Zusammenhangs bestimmt:

$$\hat{G}_S = \frac{1}{\mu_{\text{Start}} - \mu_{\text{End}} + 1} \sum_{\mu=\mu_{\text{Start}}}^{\mu_{\text{End}}} \max \left\{ 0, 10 \cdot \log_{10} \left(R^{(\text{on})}(\mu) \right) - 10 \cdot \log_{10} \left(R^{(\text{off})}(\mu) \right) \right\}. \quad (7.47)$$

Zusätzlich werden nur die Differenzen berücksichtigt, welche eine Verstärkung erzeugen, jedoch eine Normierung über alle betrachteten Frequenzen durchgeführt. Dadurch schneidet ein ICC-System, welches nur einen kleinen Frequenzbereich mittelmäßig verstärkt, schlech-

ter ab als ein System, welches den vollen Frequenzbereich mit einer gleichen Verstärkung unterstützt. Dieses Ergebnis der mittleren Verstärkung, zwischen dem Leistungsverhältnis mit und ohne ICC-System, gilt es nun in ein Qualitätsergebnis zu überführen.

Qualitätsabschätzung

Zunächst muss für eine vollständige Qualitätsabschätzung innerhalb der Fahrgastzelle, die Abstraktion auf nur je ein Lautsprecher- und Mikrofonsignal rückgängig gemacht werden. Dadurch ergibt sich für das Ergebnis der Systemunterstützung eine Abhängigkeit von den Lautsprecherkanälen der verwendeten künstlichen Sprecher l_K und der Mikrofonkanäle der verwendeten, künstlichen Zuhörer m_K . Unter Verwendung des Zusammenhanges aus (7.47) ergibt sich die mittlere Systemunterstützung zu:

$$\hat{G}_S(l_K, m_K) = \frac{1}{\mu_{\text{Start}} - \mu_{\text{End}} + 1} \sum_{\mu=\mu_{\text{Start}}}^{\mu_{\text{End}}} 10 \cdot \log_{10} \left(\frac{R_{l_K, m_K}^{(\text{on})}(\mu)}{R_{l_K, m_K}^{(\text{off})}(\mu)} \right). \quad (7.48)$$

Um ein eindeutiges Resultat der Evaluierung zu erzielen, wird nun der kleinste Wert der Systemunterstützung bestimmt, um so den schlechtesten Fall abzubilden. Die Systemunterstützung eines ICC-Systems ist gegeben durch:

$$G_S = \min \{ \hat{G}_S(l_K, m_K) \}. \quad (7.49)$$

Zusätzlich wird die Messung innerhalb eines lauten Geräuschszenarios, beispielsweise einer Fahrt mit 150 km/h, durchgeführt, um so eine maximale Systemunterstützung durch das ICC-System abschätzen zu können. Zur Bestimmung der Qualität in Bezug auf die Systemunterstützung muss zunächst eine Abschätzung des Leistungsanstiegs innerhalb der Fahrgastzelle erfolgen, denn auch ein ICC-System passt die gewählte Verstärkung an das Umgebungsgeräusch an. Dazu wird zunächst eine allgemeine Abschätzung des Leistungsanstieges zwischen einem Szenario bei einer Geschwindigkeit von 150 km/h und einem bei Stillstand mit angeschaltetem Motor erhoben. Zu diesem Zweck sind in zwei unterschiedlichen Fahrzeugen, einer Mercedes-Benz S-Klasse und einem Volkswagen T5, die Hintergrundgeräuschleistungen in den beiden Szenarien gemessen worden. Diese beiden Fahrzeuge sind ausgewählt worden, da sie sowohl den Fall eines eher lauten als auch eines eher leisen Fahrzeuges abbilden. Bei der Mercedes-Benz S-Klasse ist ein Anstieg von 10,1 dBA zwischen Stillstand und Fahrt mit 150 km/h ermittelt worden. Im Fall des T5 konnte ein Anstieg von 20,1 dBA gemessen werden. Im Mittel ergibt sich ein Anstieg der Leistung des Hintergrundgeräusches von circa 15,2 dBA. Die gemessenen Werte decken sich mit bereits in [SH06] publizierten Werten. Zusätzlich ist in Kapitel 6.1 bereits ermittelt worden, dass ein Anstieg der Verstärkung um 0,94 dB pro dBA Anstieg des Hintergrundgeräusches zu einem angenehmen Ergebnis für den Zuhörer führt. Damit ergibt sich für den mittleren Leistungsanstieg des Hintergrundgeräusches eine Verstärkung von circa 14,2 dB. Bei diesem Wert ist ein, für den Zuhörer, angenehmes ICC-System

erreicht und stellt damit die höchste Qualitätsstufe dar. Im schlechtesten Fall ist keine Verstärkung trotz Anstieg des Hintergrundgeräusches zu detektieren. Es werden vier diskrete Qualitätsstufen mit einer äquidistanten Verteilung der Schwellen festgelegt, welche in Tabelle 7.7 dargestellt sind.

	Exzellent	Gut	Mäßig	Mangelhaft
\check{c}_S	1	2	3	4
G_S	> 12 dB	12 - 8 dB	8 - 4 dB	< 4 dB

Tabelle 7.7: Einstufung der Qualitätsklassen \check{c}_S für die Systemunterstützungsmessung.

Unter Verwendung dieser Tabelle und dem Ergebnis aus Gleichung (7.49), lässt sich ein vorläufiges Qualitätsergebnis \check{c}_S bestimmen. Wichtig für die Bestimmung des Qualitätsergebnisses ist, dass innerhalb der Fahrgastzelle eine Erhöhung der Hintergrundgeräuschleistung zwischen dem Stillstand des Fahrzeuges und einer Fahrt von 150 km/h von circa 10 dBA erreicht wird. Bei einer solchen Festlegung der Geräuschleistung können nach Beendigung der Evaluierung die resultierenden Ergebnisse mit anderen Fahrzeugen verglichen werden. Das so ermittelte vorläufige Qualitätsergebnis \check{c}_S wird im Anschluss mithilfe der bereits ermittelten Fahrzeugattribute gewichtet. Da sowohl die Wandler als auch die Fahrzeuggröße Einfluss auf die Systemunterstützung haben, werden zur Gewichtung des finalen Qualitätsergebnisses dieses Evaluierungsverfahrens die entsprechenden Fahrzeugattribute verwendet. Das finale Qualitätsergebnis der Systemunterstützung ist gegeben durch:

$$c_S = \check{c}_S \cdot \frac{g_T + g_S}{2}. \quad (7.50)$$

Dabei entspricht g_T dem Fahrzeugattribut „Wandler des ICC-Systems“ und g_S dem Attribut „Fahrzeuggröße“, welche in Kapitel 7.1.3 beschrieben und bestimmt worden sind.

Die gegebene Qualitätseinstufung der Systemunterstützung ist eine erste Abschätzung, welche es in zukünftigen Betrachtungen und Forschung zu verifizieren und zu bestätigen gilt. Zudem könnte die Qualitätseinstufung in Abhängigkeit der Fahrzeugumgebung erfolgen, denn unterschiedliche Fahrzeuge führen eventuell zu unterschiedlichen Qualitätseinstufungen der Systemunterstützung. Des Weiteren stellt sich die Frage, ob eine Gewichtung der frequenzabhängigen Systemunterstützung zur Qualitätseinschätzung notwendig oder empfehlenswert ist. Um diese Frage akkurat zu beantworten, sind weitere Experimente und Erhebungen der auditiven Perzeption von Individuen notwendig.

7.2.2 Systemverzögerung

In Abschnitt 2.1 ist bereits erwähnt worden, dass aufgrund der Verwendung eines ICC-Systems künstlich eine zweite Signalquelle innerhalb der Fahrgastzelle entsteht (siehe

Abbildung 5.1). In einem gängigen Fahrzeug, beispielsweise in einem PKW, wird die zweite Signalquelle, aufgrund der Signalverarbeitung eines ICC-Systems, im Vergleich zum Direktschall, verzögert vom Zuhörer wahrgenommen. Je nach Leistung und Umfang der Verzögerung wird diese zweite Signalquelle durch den Zuhörer als Echo wahrgenommen. Ein Echo innerhalb der Fahrgastzelle kann den Zuhörer irritieren, wenn die Echschwelle deutlich überschritten wird sogar stören [BB, S. 104]. Daraus resultierend ist auch die Verzögerung der Wiedergabe des Sprachsignals durch ein ICC-System ein qualitätsrelevantes Merkmal. Diese Verzögerung wird mittels einer Kreuzkorrelationsanalyse geschätzt und ausgewertet.

Entwurf des Verfahrens

Eine Methode um die Verzögerung zweier „identischer Signale“ zu ermitteln ist eine Kreuzkorrelationsanalyse. Die resultierende Kreuzkorrelationsfolge bildet die Ähnlichkeit zweier Signale und deren Verzögerung zueinander ab [Ros14, S. 527]. Für die Durchführung der Kreuzkorrelationsanalyse wird zunächst ein Testsignal $y(n)$ ausgewählt, welches über den künstlichen Mund ausgegeben wird (siehe Abbildung 5.1). Da ein ICC-System dafür konzipiert ist Sprachsignale zu verarbeiten, wird auch ein Sprachsignal als Testsignal verwendet. Die verwendeten Sprachsignale sind aus der Lombard-Datenbank aus Kapitel 5.2 entnommen. Zusätzlich muss vorausgesetzt werden, dass das zu evaluierende ICC-System innerhalb der gewählten Evaluierungsumgebung auch Signale verarbeitet und so eine Ausgabe erzeugt. Dazu wird das ICC-System mithilfe eines Fahrgeräusches, beispielsweise bei einer Geschwindigkeit von 150 km/h, entsprechend angeregt. Da durch die Verwendung eines Geräuscheszenarios, das Eingangs- bzw. Ohrsignal des Zuhörers $x(n)$ (siehe Abbildung 5.1) auch gestört ist, wird zunächst eine Hochpassfilterung des Ohrsignals $x(n)$ durchgeführt. Bedingt durch den Anregungsbereich von Sprache, dieser liegt zwischen 100 Hz bis nahezu 7 kHz [ZF99], kann bereits ein Teil des Hintergrundgeräusches mittels eines Hochpasses mit einer Grenzfrequenz von 100 Hz unterdrückt werden.

In Abbildung 7.10 ist der Betragsfrequenzgang des verwendeten Filters und das mittlere Leistungsdichtespektrum (LDS) des Fahrzeuggeräusches vor und nach der Filterung dargestellt. Um die Ähnlichkeit der beiden Signale zu erhalten wird nicht nur das gestörte Ohrsignal $x(n)$ gefiltert, sondern auch das Testsignal $y(n)$. Dies gewährleistet eine korrekte Durchführung der Kreuzkorrelationsanalyse. Die resultierenden, hochpassgefilterten Signale werden mit $\tilde{y}(n)$, für das Testsignal, und $\tilde{x}(n)$, für das Ohrsignal des Zuhörers, bezeichnet. Die Kreuzkorrelationsfolge dieser beiden Signale ist definiert durch:

$$s_{\tilde{x}\tilde{y}}(\kappa) = E\{\tilde{x}(n)\tilde{y}(n + \kappa)\}. \quad (7.51)$$

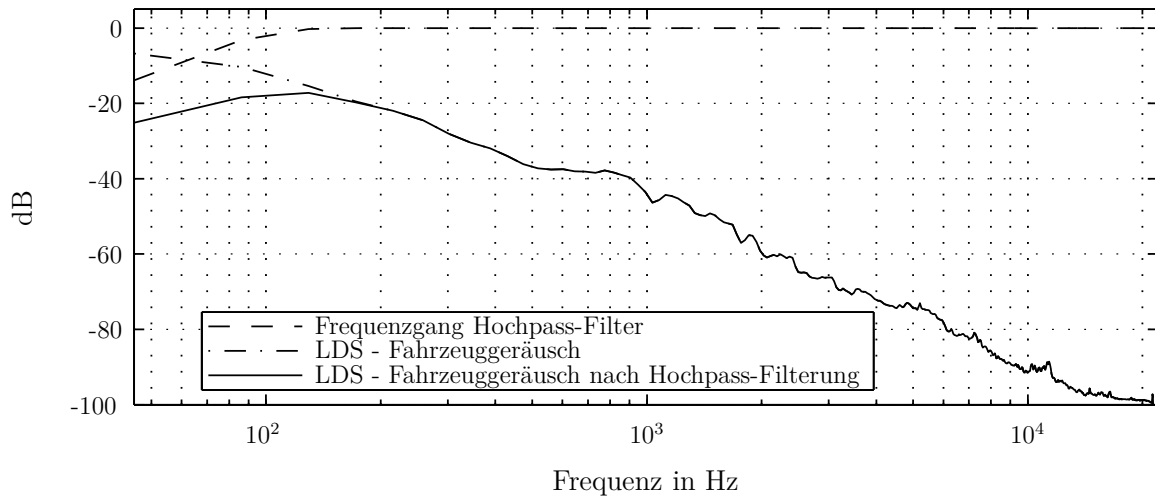


Abbildung 7.10: Mittleres Leistungsdichtespektrum (LDS) des Fahrzeuggeräusches vor und nach der Hochpassfilterung. Zusätzlich ist der Betragsfrequenzgang des verwendeten Hochpass-Filters (Grenzfrequenz gleich 100 Hz) dargestellt.

Dabei entspricht $E\{\cdot\}$ der Operation zur Berechnung des Erwartungswertes. Da lediglich eine Schätzung der Kreuzkorrelationsfolge für endliche Signale durchgeführt wird, kann folgender Zusammenhang genutzt werden:

$$\hat{s}_{\tilde{x}\tilde{y}}(\kappa) = \frac{1}{N} \sum_{n=0}^{N-1} \tilde{x}(n)\tilde{y}(n + \kappa). \quad (7.52)$$

Die Bestimmung der Kreuzkorrelationsfolge $\hat{s}_{\tilde{x}\tilde{y}}(\kappa)$ erfolgt mit dem gleichen, in Kapitel 7.1.1 beschriebenen, blockweisen Verfahren zur Schätzung der Kreuzkorrelationsfolge. Um möglichst viele Frequenzanteile innerhalb der Korrelationsanalyse anzuregen und entsprechend auswerten zu können, wird ein Testsignal $y(n)$ verwendet, welches aus unterschiedlichen Sprachpassagen eines weiblichen und eines männlichen Sprechers besteht.

Zunächst wird die Kreuzkorrelationsfolge ohne ein aktives ICC-System $\hat{s}_{\tilde{x}\tilde{y}}^{(\text{off})}(\kappa)$ und im Anschluss mit ICC-System $\hat{s}_{\tilde{x}\tilde{y}}^{(\text{on})}(\kappa)$ geschätzt. In Abbildung 7.11 sind diese beiden Folgen beispielhaft dargestellt. Auffällig ist zunächst, dass sowohl mit als auch ohne aktivem ICC-System, die beiden Kurven der oberen Grafik bis zu einem Zeitpunkt von 15 ms quasi identisch sind. Bei diesem Anteil handelt es sich um den Direktschall, welcher sowohl bei eingeschaltetem als auch bei ausgeschaltetem ICC-System auftritt. Zusätzlich ist bei der Kreuzkorrelationsfolge, geschätzt mit aktivem ICC-System, zu sehen, dass bei etwa 20 ms

die zweite Signalquelle, angeregt durch das ICC-System, auftritt. Um nun die Verzögerung des ICC-Systems zu bestimmen wird der folgende Ansatz verfolgt:

$$\hat{D}_{\text{sys}} = \left(\arg \max_{\kappa} \left\{ \left| \hat{s}_{\hat{x}\hat{y}}^{(\text{on})}(\kappa) - \hat{s}_{\hat{x}\hat{y}}^{(\text{off})}(\kappa) \right| \right\} - \arg \max_{\kappa} \left\{ \left| \hat{s}_{\hat{x}\hat{y}}^{(\text{off})}(\kappa) \right| \right\} \right) / f_{\text{S}} . \quad (7.53)$$

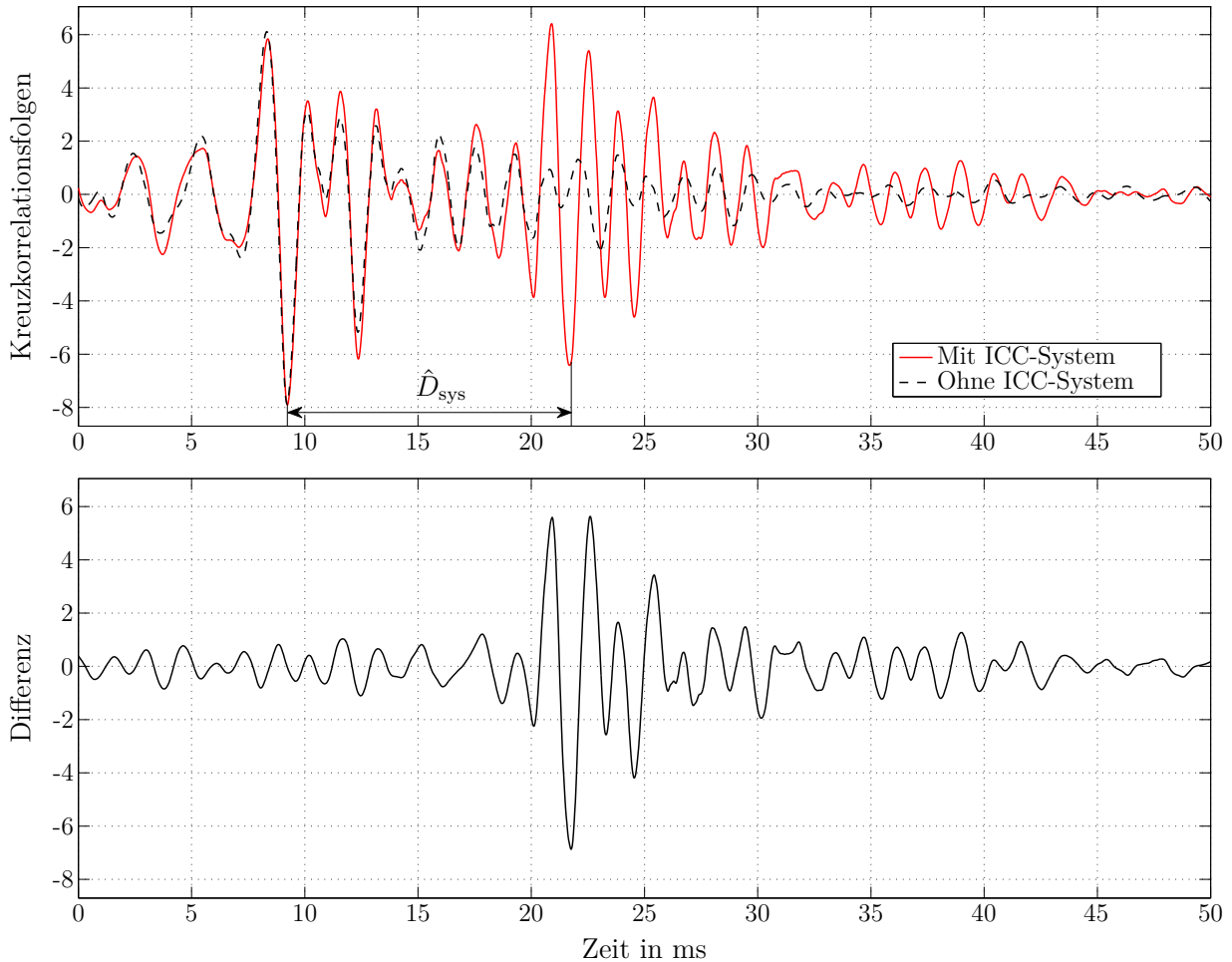


Abbildung 7.11: Die geschätzten Kreuzkorrelationsfolgen mit $\hat{s}_{\hat{x}\hat{y}}^{(\text{on})}(\kappa)$ (rot) und ohne ICC-System $\hat{s}_{\hat{x}\hat{y}}^{(\text{off})}(\kappa)$ (schwarz gestrichelt) sind im oberen Teil der Abbildung und die entsprechende Differenz im unteren Teil dargestellt. Zusätzlich ist die Systemverzögerung, ermittelt durch den Zusammenhang aus (7.53), im oberen Teil abgebildet.

Dabei wird davon ausgegangen, dass sich die Direktschallkomponente durch die Aktivierung des ICC-Systems nicht verändert. Somit kann die Schätzung der Kreuzkorrelationsfolge mit aktiviertem ICC-System $\hat{s}_{\hat{x}\hat{y}}^{(\text{on})}(\kappa)$ durch die Subtraktion der zuvor gemessenen Schätzung der Kreuzkorrelationsfolge ohne ICC-System $\hat{s}_{\hat{x}\hat{y}}^{(\text{off})}(\kappa)$ vom Direktschalleinfluss bereinigt werden. Die Differenz der beiden Kreuzkorrelationsfolgen ist beispielhaft in der

unteren Grafik in Abbildung 7.11 dargestellt. Ohne die Subtraktion der beiden Kreuzkorrelationsfolgen würde bei der Operation zur Bestimmung des Arguments des Maximums der falsche Zeitpunkt ermittelt werden. Im oberen Teil der Abbildung 7.11 wird ersichtlich, dass bei einer reinen Bestimmung des Arguments des Maximums von $\hat{s}_{\tilde{x}\tilde{y}}^{(\text{off})}(\kappa)$ ein weiteres Mal der Direktschall als Maximum ermittelt würde. Jedoch sollte in diesem Fall der vom ICC-System ausgesendete Anteil und dessen Verzögerung bestimmt werden. Die Differenz zwischen der Verzögerung der ICC-System- und der Direktschallkomponente, ergibt die Verzögerung des ICC-Systems. Dabei beinhaltet diese nicht nur die Zeit der Verarbeitung des untersuchten ICC-Systems, sondern auch die akustischen und elektrischen Laufzeiten zu bzw. von den Wandlern zu den Insassen.

Qualitätsabschätzung

Auch in diesem Fall muss zur Bestimmung der Qualität anhand des vorgestellten Verfahrens, zunächst die Abstraktion auf nur einen Ein- und Ausgangskanal der ICC-Systemevaluierung aufgehoben werden. Dadurch ergibt sich eine geschätzte Systemverzögerung

$$\hat{D}_{\text{sys}}(l_K, m_K) = \left(\arg \max_{\kappa} \left\{ \left| \hat{s}_{\tilde{x}_{m_K} \tilde{y}_{l_K}}^{(\text{on})}(\kappa) - \hat{s}_{\tilde{x}_{m_K} \tilde{y}_{l_K}}^{(\text{off})}(\kappa) \right| \right\} - \arg \max_{\kappa} \left\{ \left| \hat{s}_{\tilde{x}_{m_K} \tilde{y}_{l_K}}^{(\text{off})}(\kappa) \right| \right\} \right) / f_s, \quad (7.54)$$

in Abhängigkeit des Kanals des verwendeten künstlichen Sprecher l_K und des verwendeten künstlichen Zuhörers m_K . Zur Bestimmung der Gesamtqualität in Bezug auf die Systemverzögerung wird hier die maximale Verzögerung

$$D_{\text{sys}} = \max_{l_K, m_K} \left\{ \hat{D}_{\text{sys}}(l_K, m_K) \right\} \quad (7.55)$$

über alle möglichen Wege zwischen Sprecher und Zuhörer bestimmt. Die Einstufung der Qualität sollte neben der Wahrnehmung des Zuhörers, auch die des Sprechers abbilden. Denn auch der Sprecher kann durch die Wiedergabe seines eigenen Sprachsignals gestört werden. In Kapitel 6.2 ist ein Experiment, welches die Eigenechoperzeption des Sprechers untersucht, vorgestellt worden. Hier konnte ein Bereich zwischen 12 - 18 ms Verzögerung ermittelt werden, in welchem die maximale durch den Sprecher tolerierte Verstärkung liegt. Zusätzlich ist in [SH06] eine Verzögerung zwischen 10 - 20 ms als Bereich ermittelt worden, in dem die besten Ergebnisse in Bezug auf den Zuhörer erzielt werden konnten. Dies lässt darauf schließen, dass ein optimal funktionierendes ICC-System maximal eine Verzögerung von 12 ms benötigen sollte. Ein mangelhaft funktionierendes ICC-System würde Echos einfügen, die von Zuhörer und Sprecher wahrgenommen werden. Es ist zu beachten, dass in dieser speziellen Umgebung bereits leichte Echos einen deutlichen Qualitätsverlust hervorrufen. Bei zwei unterschiedlich lauten Schallquellen kann bereits eine Verzögerung von mehr als 30 ms nach [BB] als ein Echo wahrgenommen werden. Innerhalb der Fahrgastzelle ist diese Verzögerung damit als störend einzustufen und entspricht somit die

unteren Qualitätsgrenze. Mithilfe der in Tabelle 7.8 dargestellten äquidistanten Schwellen, kann nun die Qualitätseinstufung in Bezug auf die Verzögerung des ICC-Systems erfolgen.

	Exzellent	Gut	Mäßig	Mangelhaft
\check{c}_T	1	2	3	4
D_{sys}	< 12 ms	12 - 16 ms	16 - 20 ms	> 20 ms

Tabelle 7.8: Einstufung der Qualitätsklassen \check{c}_T für die Systemunterstützungsmessung.

Die Verzögerung innerhalb eines kleinen Fahrzeuges wird, aufgrund der Leistungsabnahme der Schallereignisse über den Raum, kritischer von den Insassen bewertet als bei einem großen Fahrzeug. Deswegen ist eine geringe Verzögerung des ICC-Systems bei einem kleineren Fahrzeug als wichtiger zu erachten als bei einem großen. Aus diesem Grund wird nach Bestimmung der Qualität, unter Verwendung der in Tabelle 7.8 angegebenen Qualitätsstufen, das vorläufige Qualitätsergebnis \check{c}_T zusätzlich mit dem Fahrzeugattribut „Fahrzeuggröße“ g_S gewichtet. Es folgt für das finale Ergebnis der Systemverzögerung:

$$c_T = \check{c}_T \cdot g_S. \quad (7.56)$$

Das vorgestellte Verfahren und die zugehörigen Qualitätsstufen bieten eine erste fundierte Abschätzung der Qualität im Sinne der Verzögerungen der beiden Schallereignisse in einem Fahrzeug, ausgestattet mit einem ICC-System. Eine spannende Fragestellung in diesem Zusammenhang ist allerdings die fahrzeugspezifische Echschwelle des Zuhörers. Die hier verwendeten Werte zur Einstufung der Qualität mittels der Systemverzögerung, sind anhand eines Fahrzeuges und dessen spezifischer Anordnung von Wandlern etc. erhoben worden. Für die weitere Forschung wäre jedoch interessant, ob die Veränderung des Fahrzeuges auch eine Veränderung der Echschwellen, sowohl des Eigenechos als auch des herkömmlichen Echos, hervorruft. Eine weitere offene Fragestellung in diesem Zusammenhang, ist: Was passiert, wenn die zweite Nutzsquelle, angeregt durch das ICC-System, vor dem Direktschall vom Zuhörer wahrgenommen wird? Ein solches Phänomen könnte beispielsweise in sehr großen Fahrzeugen wie Bussen mit mehr als drei Sitzreihen auftreten. Dabei sollte sicherlich auch der Einfluss der Leistung und der Verzögerung zwischen beiden Nutzsquellen in zukünftigen Betrachtungen untersucht werden.

7.2.3 Erhöhung des Hintergrundgeräusches

Wie bereits in Abschnitt 2 beschrieben, verwendet ein ICC-System die innerhalb der Fahrgastzelle installierten Mikrofone. Diese zeichnen neben dem gewünschten Sprachsignal, auch das auftretende Hintergrundgeräusch auf, welches mittels einer Geräuschreduktion unterdrückt werden soll. Das vom ICC-System aufbereitete Sprachsignal wird im

Anschluss verstärkt und über die zuhörer-nahen Lautsprecher innerhalb der Fahrgastzelle ausgegeben. Bei einer nicht optimalen Umsetzung der Geräuschunterdrückung, können dadurch Anteile des aufgenommenen Hintergrundgeräusches verstärkt ausgegeben werden. Dies würde zu einer Erhöhung des Hintergrundgeräusches und somit zu einer Verschlechterung der Kommunikationssituation für die Insassen führen. Diese Erhöhung des Hintergrundgeräusches soll mithilfe des im Folgenden vorgestellten Verfahrens bestimmt werden. Zunächst wird auch hier die vereinfachte Testanordnung aus Abbildung 5.1 mit nur einem Ein- und Ausgangssignal der ICC-Systemevaluierung und nur einem ICC-Mikrofon und ICC-Lautsprecher verwendet.

Entwurf des Verfahrens

Zur Ermittlung der Erhöhung des Hintergrundgeräusches, erzeugt durch ein ICC-System, wird ein spezielles Testsignal verwendet. Dieses Signal, welches beispielhaft als Spektrogramm in Abbildung 7.12 dargestellt ist, besteht neben den Sprachpassagen aus einzelnen Passagen in denen keine Anregung stattfindet. Die Sprachpassagen sind der, in Kapitel 5.2 beschriebenen, Lombard-Datenbank entnommen worden. Innerhalb der Pausen zwischen den Sprachpassagen, wird die Leistung des Hintergrundgeräusches, sowohl mit als auch ohne aktiviertem ICC-System, gemessen und anschließend verglichen.

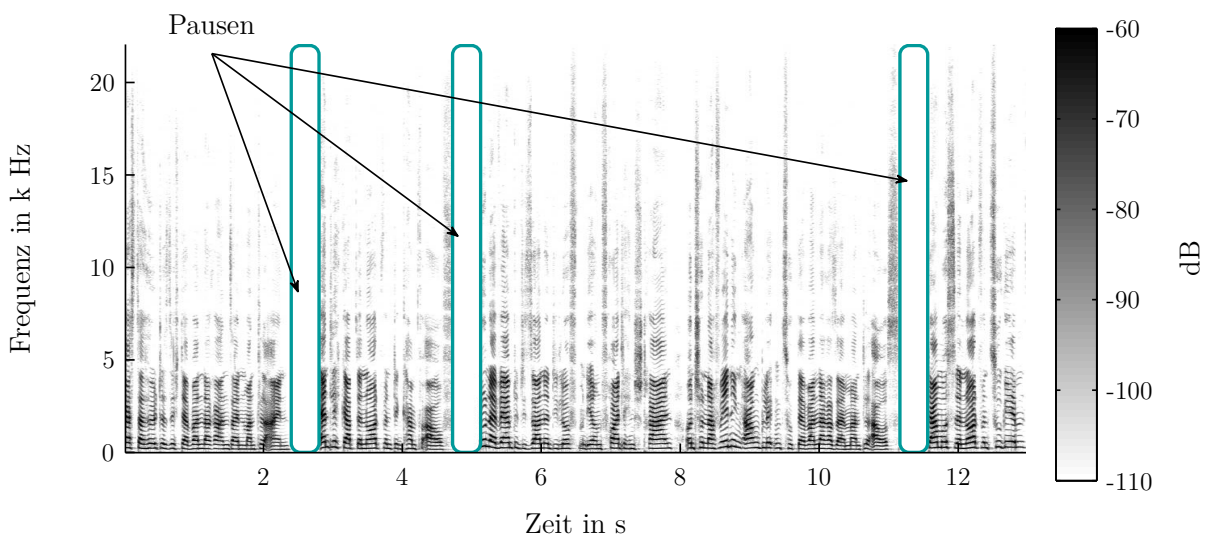


Abbildung 7.12: Spektrogramm des Testsignals zur Bestimmung der Erhöhung des Hintergrundgeräusches.

Für die Schätzung der frequenzabhängigen Leistung des auftretenden Hintergrundgeräusches wird das Mikrofonsignal am Zuhörerrohr $x(n)$ mittels einer Analysefilterbank, welche in Gleichung (5.1) bereits definiert worden ist, in den Frequenzbereich transfor-

miert. Mit dieser Frequenzbereichsdarstellung, in Abhängigkeit des Rahmens k und des Frequenzbandes μ , kann nun das mittlere Leistungsdichtespektrum durch

$$N_{xx}(\mu) = \frac{\sum_{k=0}^{K-1} |X(\mu, k)|^2 \cdot a_x^N(k)}{\sum_{k=0}^{K-1} a_x^N(k)}, \quad (7.57)$$

bestimmt werden, wobei

$$a_x^N(k) = \begin{cases} 1, & \text{wenn } X(\mu, k) \text{ einer Pause entspricht,} \\ 0, & \text{sonst.} \end{cases} \quad (7.58)$$

Da das Leistungsdichtespektrum nur während der Pausen geschätzt wird, wird ausschließlich die Leistung des auftretende Geräusches am Zuhörerohr gemessen. Mit Hilfe des existierenden Wissens über die Struktur des Testsignals $y(n)$ während der Ausgabe über den künstlichen Sprecher und der Verzögerung zwischen Sprecher und Zuhörer, können die Zeitpunkte der Pausen am Zuhörerohr gezielt bestimmt werden. Zusätzlich werden längere Pausen, zum Beispiel mit einer Dauer von 2 s, verwendet, um eine ausreichende Schätzung zu erzielen. Dieses geschätzte Leistungsdichtespektrum wird sowohl mit aktiviertem ICC-System $N_{xx}^{(\text{on})}(\mu)$ als auch bei deaktiviertem System $N_{xx}^{(\text{off})}(\mu)$ bestimmt und das Verhältnis

$$R_N(\mu) = \frac{N_{xx}^{(\text{on})}(\mu)}{N_{xx}^{(\text{off})}(\mu)} \quad (7.59)$$

gebildet. Dieses Verhältnis gibt den Umfang der Veränderung des Hintergrundgeräusches unter Verwendung eines ICC-Systems an. Unter Berücksichtigung des theoretischen Modells (5.21), welches in Kapitel 5.4 hergeleitet worden ist, kann der zuvor ermittelte Zusammenhang ferner durch den folgenden theoretischen beschrieben werden:

$$R_N(\mu) = \frac{S_{bb}(\mu) \cdot |H_{\text{tot}}(\mu)H_{\text{LE}}(\mu)|^2 + S_{bb}(\mu)}{S_{bb}(\mu)}. \quad (7.60)$$

Die normierte Frequenz Ω wird mithilfe einer Diskretisierung auf die Frequenzbänder μ übertragen. Zusätzlich ist zu berücksichtigen, dass die Schätzung der Leistungsdichte nur während der Sprachpausen durchgeführt worden ist und so das Eingangsspektrum $S_{yy}(\mu) = 0$ angenommen werden kann. Anhand dieser theoretischen Betrachtung kann bestätigt werden, dass durch das Verhältnis der beiden Leistungsdichten aus (7.59), der Anteil des durch das ICC-System hinzugefügten Hintergrundgeräusches bestimmt werden kann.

Diese Schätzung des Anstiegs des Hintergrundgeräusches setzt voraus, dass es sich um ein näherungsweise stationäres Geräusch handelt, welches sich nur geringfügig in seinen stochastischen Eigenschaften über die Zeit verändert. Nur so kann ein Vergleich der Leistungsdichtespektren zu verschiedenen Zeitpunkten aussagekräftige Ergebnisse liefern.

Aufgrund von Messungenauigkeiten können geringe negative logarithmische Werte des Verhältnisses zwischen dem Leistungsdichtespektrum mit und ohne ICC-System entstehen. Da lediglich der Anstieg des Geräusches von Interesse ist, wird das logarithmische Ergebnis von (7.57) auf einen Bereich größer Null, durch die verwendete Maximalwertoperation, beschränkt. Die Schätzung des mittleren Geräuschanstiegs, hervorgerufen durch ein ICC-System, ist gegeben durch:

$$\hat{G}_N = \frac{1}{\mu_{\text{Start}} - \mu_{\text{End}} + 1} \sum_{\mu=\mu_{\text{Start}}}^{\mu_{\text{End}}} \max \left\{ 0, 10 \cdot \log_{10} \left(\frac{N_{xx}^{(\text{on})}(\mu)}{N_{xx}^{(\text{off})}(\mu)} \right) \right\}. \quad (7.61)$$

Auch in diesem Fall der Mittelung über die frequenzabhängigen Geräuschanstieg, wird nur der Bereich in dem eine Verarbeitung durch das ICC-System durchgeführt wird, betrachtet. Der mittlere Geräuschanstieg wird in einem Bereich zwischen μ_{Start} und μ_{End} bestimmt.

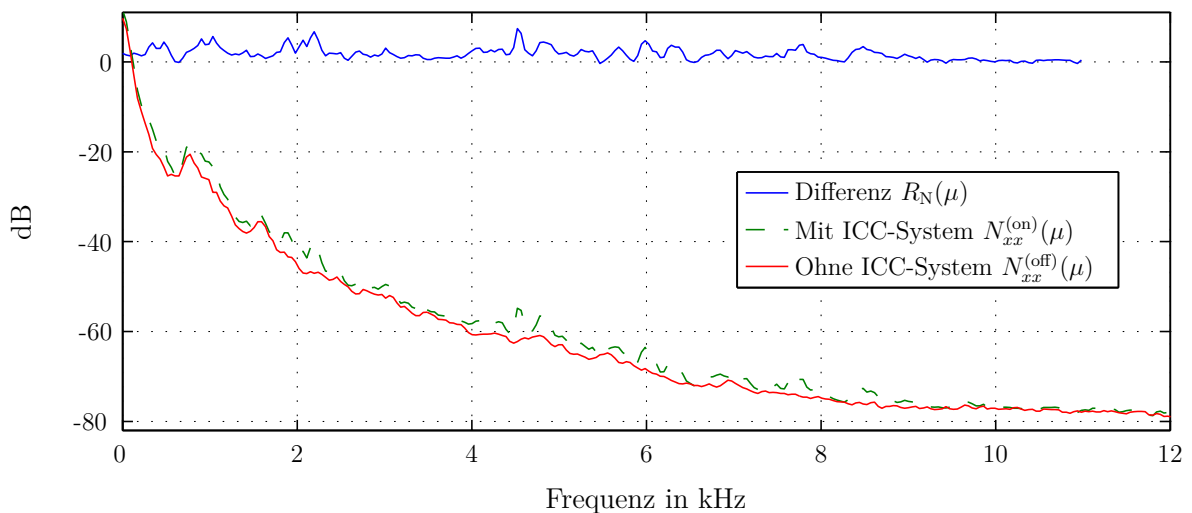


Abbildung 7.13: Die geschätzten logarithmischen Leistungsdichtespektren mit (grün gestrichelt) und ohne (rot) die Unterstützung eines ICC-Systems und der resultierende, frequenzabhängige Anstieg des Hintergrundgeräusches (blau).

In Abbildung 7.13 sind die geschätzten Leistungsdichtespektren mit (grün gestrichelt) und ohne ICC-System (rot) beispielhaft für ein beliebiges ICC-System dargestellt. Zusätzlich ist das Verhältnis bzw. die logarithmische Differenz der beiden Leistungsdichtespektren mit aufgetragen. Da das verwendete ICC-System nur Frequenzen innerhalb von 100 Hz und 11 kHz verarbeitet, ist lediglich in diesem Bereich die Erhöhung des Hintergrundgeräusches ermittelt worden. Diese Erhöhung fällt in diesem speziellen Fall auch eher gering aus, jedoch wird deutlich, dass nur innerhalb des durch das ICC-System verarbeiteten Frequenzbereiches eine Erhöhung zu erkennen ist.

Qualitätsabschätzung

Wie in den zuvor beschriebenen Verfahren, muss auch hier berücksichtigt werden, dass nur zur Herleitung des Verfahrens ein einzelner Eingangskanal der Evaluierung betrachtet worden ist. Innerhalb einer realen Evaluierungsumgebung gibt es mindestens zwei Zuhörerohren. Für jeden dieser Eingangskanäle m_K der Zuhörerohren wird die Erhöhung des Hintergrundgeräusches bestimmt durch:

$$\hat{G}_N(m_K) = 10 \cdot \log_{10} \left(\frac{N_{x_{m_K} x_{m_K}}^{(\text{on})}(\mu)}{N_{x_{m_K} x_{m_K}}^{(\text{off})}(\mu)} \right). \quad (7.62)$$

Zur Erhebung eines Gesamtergebnisses des Hintergrundgeräuschanstieges wird das Maximum über alle Kanäle bestimmt:

$$G_N = \max \{ \hat{G}_N(m_K) \}. \quad (7.63)$$

Dieses Ergebnis wird im nächsten Schritt einer Qualitätsstufe zugewiesen, welche in Tabelle 7.9 definiert sind.

	Exzellent	Gut	Mäßig	Mangelhaft
\check{c}_R	1	2	3	4
D_{sys}	< 2 dB	2 - 7 dB	7 - 12 dB	> 12 dB

Tabelle 7.9: Einstufung der Qualitätsklassen \check{c}_T für die Systemunterstützungsmessung.

Zur Bestimmung der Schwellen zwischen den einzelnen Stufen sind zunächst erste empirisch ermittelte Werte verwendet worden. Bereits eine mittlere Erhöhung des Hintergrundgeräusches von 12 dB könnte durch einzelne Personen als störend wahrgenommen werden. Eine mittlere Erhöhung von circa 2 dB ist als nicht wahrnehmbar eingestuft worden, was zur Folge hat, dass dieser Wert einem exzellent arbeitenden ICC-System entspricht.

Zum Abschluss wird das vorläufige Qualitätsergebnis \check{c}_R mit dem Fahrzeugattribut „Hintergrundgeräusch“ gewichtet. Dadurch werden die Qualitätsergebnisse eines eher lauten Fahrzeuges stärker in der Gesamtqualität berücksichtigt, als eines eher leisen Fahrzeuges. Denn innerhalb eines leisen Szenarios, enthält bereits das Eingangssignal eines ICC-Systems eine geringere Geräuschleistung als in einem lauten Szenario. Das Qualitätsergebnis der Bestimmung des Hintergrundgeräuschanstieges ist gegeben durch:

$$c_R = \check{c}_R \cdot g_N. \quad (7.64)$$

Abschließend kann festgehalten werden, dass das vorgestellte Verfahren offensichtliche Zusammenhänge kombiniert, welche eine erste plausible Evaluierung des Verhaltens ei-

nes ICC-Systems in Bezug auf die Verarbeitung des auftretenden Hintergrundgeräusches liefert. Dennoch gilt es in zukünftigen Forschungen gezielt den Einfluss, des durch das ICC-System hinzugefügten Hintergrundgeräusches weiter zu analysieren. So ist zum Beispiel noch nicht klar, welchen Einfluss eine Modulation des additiven Geräusches auf das Qualitätsempfinden einer Testperson hat. Um dieser Frage zu beantworten sind weitere psychoakustische Experimente notwendig, welche auch zur Verifikation und Weiterentwicklung der Qualitätsklassen verwendet werden können.

7.3 Kommunikationsqualität

Die Evaluierungsgruppe „Kommunikationsqualität“ vereinigt Verfahren deren Ziel es ist, die Qualität, empfunden durch die Insassen, zu schätzen. Klassische Verfahren, wie die Bestimmung des SNR oder spektrale Distanzansätze, fallen in diese Gruppe. Diese Evaluierungsverfahren ermitteln mithilfe von Signalanalysen beispielsweise die Sprachqualität eines Signals nach der Übertragung. Innerhalb der ICC-Systemevaluierung, soll unter Verwendung solcher Verfahren die Verbesserung der Kommunikations- bzw. Sprachqualität durch die Verwendung eines ICC-Systems erhoben werden. Aus diesem Grund werten alle im Folgenden vorgestellten Verfahren zunächst die Qualität eines Szenarios ohne, und anschließend unter Zuhilfenahme eines ICC-System aus. Mittels dieser Qualitätsergebnisse kann eine Abschätzung der relativen Verbesserung durch ein ICC-System erfolgen. Dabei macht es einen Unterschied, ob die gegebene Situation, in einem fahrenden Fahrzeug, die Verwendung eines ICC-Systems erfordert oder nicht. Diese Erfordernis eines ICC-Systems wird mittels des Attributes „Notwendigkeit“, welches in Kapitel 7.3.1 definiert ist, festgehalten und zur Bestimmung der Verbesserung der Qualität verwendet werden. Im Folgenden soll zunächst die Definition und die Bestimmung der zusätzlichen Abstraktionsebene der ICC-Systemattribute und das Attribut „Notwendigkeit“ beschrieben werden. Im weiteren Verlauf werden drei verschiedene Verfahren der Evaluierungsgruppe „Kommunikationsqualität“ und deren Umsetzung im Detail erläutert.

7.3.1 ICC-Systemattribute

Neben der Klassifizierung eines Fahrzeuges mit den in Kapitel 7.1.3 vorgestellten Fahrzeugattributen, können auch für ICC-Systeme spezifische Attribute ermittelt werden. Diese Attribute beschreiben das System und haben Einfluss auf die Evaluierungsverfahren der Gruppe „Kommunikationsqualität“. In dieser Arbeit wird nur ein Attribut vorgestellt und beschrieben. Dennoch sind weitere Attribute denkbar und sollten auch in zukünftigen Untersuchungen in Betrachtung gezogen werden.

Notwendigkeit

Das Attribut „Notwendigkeit“ definiert, ob die Fahrzeugumgebung die Kommunikationssituation der Insassen in dem Maße einschränkt, dass die Verwendung eines ICC-Systems als sinnvoll einzustufen ist. Zu diesem Zweck muss zunächst eine Einschätzung des Szenarios innerhalb der Fahrgastzelle erfolgen. Dabei sind Faktoren wie das auftretende Hintergrundgeräusches (je lauter es wird, desto schlechter verstehen sich die Passagiere) und die Distanz zwischen den Insassen (je weiter entfernt die Passagiere sind, desto stärker ist die Beeinträchtigung der Kommunikationssituation) für die Bestimmung der Notwendigkeit eines ICC-Systems relevant. Diese beiden Faktoren sind bereits in Bezug auf die Fahrzeugattribute in Kapitel 7.1.3 beschrieben und bestimmt worden. In Abbildung 7.14 sind die Zusammenhänge zwischen den Verfahren aus der Evaluierungsgruppe Fahrzeugeigenschaften und die daraus resultierenden Fahrzeugattribute noch einmal zur besseren Übersicht dargestellt.

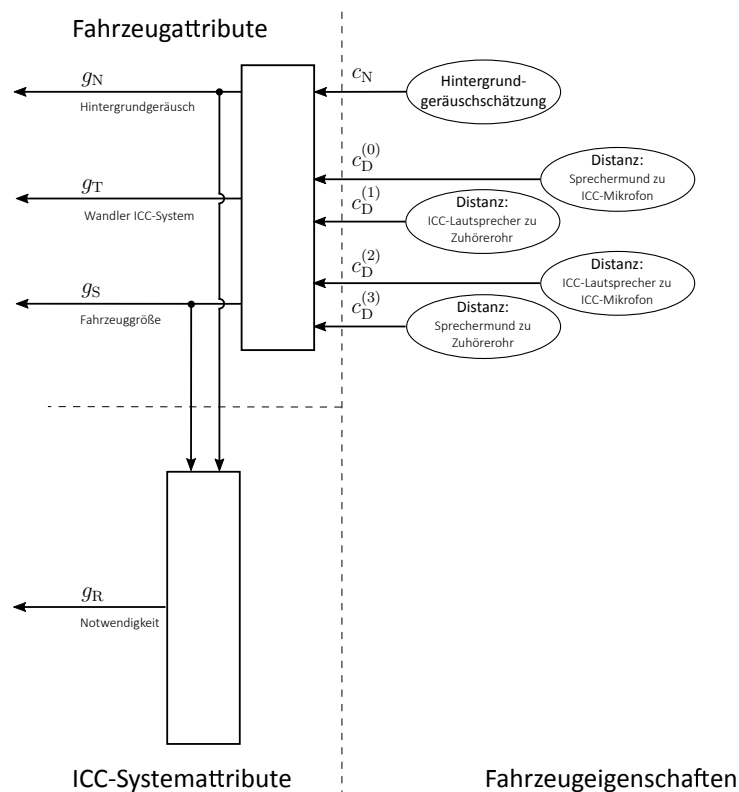


Abbildung 7.14: Verknüpfung der Ergebnisse der Evaluierungsgruppe Fahrzeugeigenschaften zu den einzelnen Fahrzeugattributen. Zusätzlich sind die einzelnen Verbindungen der Fahrzeugattribute zu dem ICC-Systemattribut „Notwendigkeit“ dargestellt.

Die Fahrzeugattribute werden, neben der Bestimmung der Qualitätsergebnisse der Evaluierungsverfahren aus der Gruppe „Verhalten des ICC-Systems“, auch zur Bestimmung

des ICC-Systemsattributs „Notwendigkeit“ verwendet. Das Attribut „Notwendigkeit“ ist definiert durch:

$$g_R = \frac{g_N + (1 - g_S)}{2}. \quad (7.65)$$

Dabei werden ausschließlich das Fahrzeugattribut „Hintergrundgeräusch“ g_N und die „Fahrzeuggröße“ g_S zur Bestimmung verwendet, da diese maßgeblich die Dringlichkeit der Verwendung eines ICC-Systems beeinflussen.

Im Zusammenhang der Systemattribute wären weitere Attribute und Kombinationen der Einflussgrößen denkbar. Besonders wenn weitere Fahrzeugattribute zur Berechnung bereitgestellt werden, sollten weitere ICC-Systemattribute erforscht werden. Dennoch kann dieses erste Attribut helfen, die Qualität eines ICC-System in Bezug auf die Notwendigkeit des Systems, relativ einzustufen und den Einfluss des Fahrzeuges auf die Erhebung der Qualität mittels der Evaluierungsverfahren der Gruppe „Kommunikationsqualität“ mitzuberechnen.

7.3.2 Modifizierter Sprachübertragungsindex

Der Sprachübertragungsindex (engl. *Speech Transmission Index*, STI) ist zur instrumentellen Bestimmung der Sprachverständlichkeit von akustischen oder auch elektroakustischen Kanälen entwickelt worden. Dabei wird innerhalb dieser Qualitätsmetrik davon ausgegangen, dass die Erhaltung der Intensitätseinhüllenden eines Sprachsignals von größter Wichtigkeit für die Sprachverständlichkeit ist. Die Veränderung der Modulation der Intensitätseinhüllenden durch den Übertragungskanal, die durch das Auftreten von Echos, Nachhall und/oder Geräusch entsteht, soll mittels des STI bestimmt und in einen Qualitätsschätzer überführt werden [IEC].

Die konventionelle, direkte Methode [SH80] basiert auf der sogenannten Modulationsübertragungsfunktion (engl. *Modulation Transfer Function*, MTF), welche in Abhängigkeit der Modulationsfrequenz ν , die Veränderung des Modulationsindex zwischen Ein- und Ausgang eines Übertragungssystems beschreibt. Zur Bestimmung der MTF wird weißem Rauschen das mittlere Langzeitspektrum von Sprache aufgeprägt, entsprechend der aktuellen Modulationsfrequenz in seiner Intensität moduliert und im Anschluss in Oktavbänder aufgeteilt.

In Abbildung 6.5 sind bereits die sieben verwendeten Oktavfilter mit dem entsprechenden Betragsfrequenzgang dargestellt. Sie decken die Mittenfrequenzen von 125 Hz bis 8 kHz ab. Jedes Oktavbandsignal bekommt eine oder mehrere Modulationsfrequenzen in seiner Intensität aufgeprägt und wird im Anschluss über den zu evaluierenden Kanal übertragen. Dabei können Modulationsfrequenzen in einem 1/3-Oktavabstand zwischen 0,63 und 12,5 Hz verwendet werden. Am Ausgang des Kanals wird der Modulationsindex bestimmt und in das Verhältnis zum Eingangsmodulationsindex gesetzt. Die so bestimmte MTF wird im weiteren Verlauf des Verfahrens zur Berechnung des STI verwendet. Wei-

tere detaillierte Informationen zur Bestimmung des STI können in [IEC, SH80, SH85] gefunden werden.

Der konventionelle Ansatz ist für die Evaluierung von Übertragungskanälen, wie zum Beispiel in einem Konzertsaal, entwickelt und optimiert worden. Dort können bereits qualitativ hochwertige Aussagen über die Übertragungsqualität von gestörten Kanälen getroffen werden. Dabei werden Störungen wie Hintergrundgeräusch, Nachhall, Echos und auch nichtlineare Verzerrungen mitberücksichtigt [IEC]. Bei der Evaluierung von sprachverarbeitenden Systemen, stellt sich die konventionelle Methode der STI-Erhebung als schwierig umsetzbar dar. Das Testsignal, welches einem Rauschsignal entspricht, würde das System stark verändern und somit die Qualitätsaussage verfälschen. Bereits in [SH85, IEC] sind verschiedenen Ansätze zur Bestimmung der MTF anhand von Sprachsignalen und der Differenz des Einhüllendenspektrums des Originalsignals und des Zuhörersignals beschrieben worden.

Ein Sprachsignal als Testsignal würde ein sprachverarbeitendes System, wie zum Beispiel ein ICC-System, mit dem vorgesehenen Signal anregen und so das System innerhalb des gewünschten operativen Zustandes auswerten. Aus diesem Grund wird bei der instrumentellen Evaluierung von ICC-Systemen ein Verfahren basierend auf dem STI unter Verwendung eines Sprachsignals verwendet. Das nachfolgend beschriebene Verfahren ist in einigen Teilen der Berechnung, zum Beispiel bei der endgültigen Bestimmung der MTF, aufgrund der besonderen Anforderungen leicht verändert und wird deswegen im weiteren Verlauf als modifiziertes STI-Verfahren bezeichnet.

Entwurf des Verfahrens

Zur Evaluierung eines ICC-Systems wird ein Verfahren vorgestellt, welches auf der Erhebung des STI mittels eines Sprachsignals basiert. Dazu ist die in Abbildung 7.15 dargestellte Signalverarbeitung zur Bestimmung der MTF angewandt und innerhalb des Rahmenwerkes, welches bereits in Kapitel 5.3 beschrieben worden ist, umgesetzt worden. Als Testsignal wird ein Sprachsignal aus der Lombard-Datenbank, beschrieben in Kapitel 5.2, verwendet. Dieses Testsignal $y(n)$ wird über den künstlichen Sprecher, siehe Abbildung 5.1, ausgegeben und durch das Ohrmikrofon $x(n)$ aufgenommen. Diese beiden Signale stehen für die Erhebung des STIs in Blöcken zur Verfügung. Diese Signale sind durch $y_k(n)$ und $x_k(n)$, wobei k dem Rahmenindex entspricht, entsprechend definiert. Zweifelsohne können mehr als nur je ein Mikrofon- und Lautsprecherkanal bei der eigentlichen Evaluierung berücksichtigt werden. Diese werden jedoch für die Beschreibung des Verfahrens zunächst aus Gründen der Anschaulichkeit vernachlässigt.

Diese beiden Signale, $y_k(n)$ und $x_k(n)$, werden zu Beginn in sieben Oktavbänder aufgeteilt. Dazu werden die in Abbildung 6.5 dargestellten Oktavfilter verwendet und die Signale in Abhängigkeit des Oktavbandes f_i durch $y_k^{(f_i)}(n)$ und $x_k^{(f_i)}(n)$ angeben. In einem nächsten Schritt wird in jedem Band die Intensitätseinhüllende beider Signale bestimmt. Dazu wird der Betrag der Signale gebildet und eine Tiefpassfilterung durchgeführt. Das

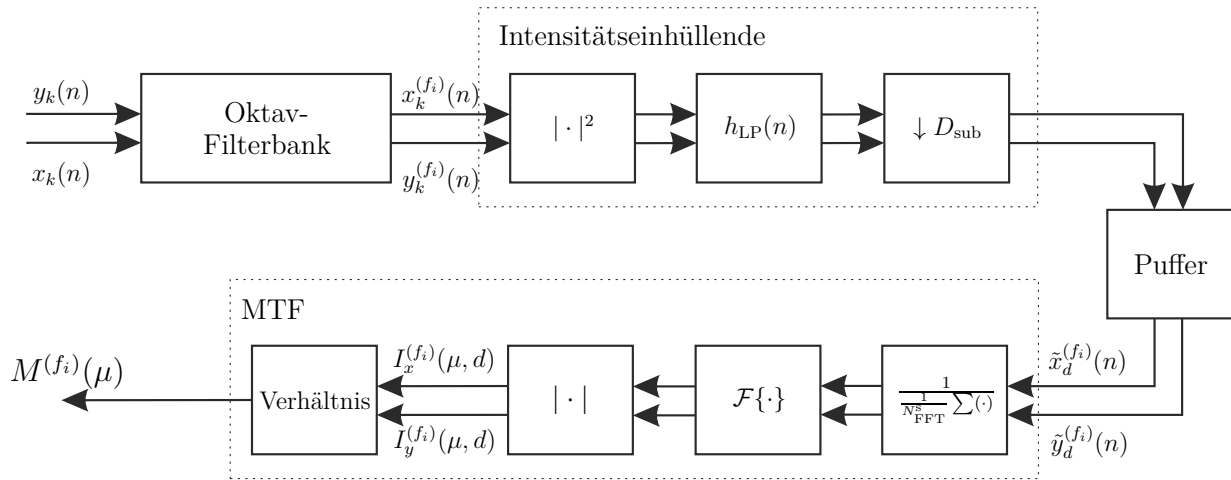


Abbildung 7.15: Signalflussgraph zur Bestimmung der MTF mithilfe eines Referenzsignals.

Tiefpassfilter besitzt eine Grenzfrequenz von $f_c^{\text{LP}} = 100$ Hz. Im letzten Schritt werden die Abtastwerte um $D_{\text{sub}} = \lfloor f_s/2f_c^{\text{LP}} \rfloor$ reduziert, da aufgrund der Tiefpassfilterung eine geringere Abtastrate verwendet werden kann.

Nach der Ermittlung der Intensitätseinhüllenden, siehe Abbildung 7.15, werden die Abtastwerte gepuffert bis eine Anzahl von N_{FFT}^s zusammengeführt worden ist. Dadurch kann die Intensitätseinhüllende über einen längeren Zeitraum zur Bestimmung der MTF betrachtet werden. Zum Beispiel kann bei einer originalen Abtastrate $f_s = 44,1$ kHz eine Reduktion der Abtastung von $D_{\text{sub}} = 220$ erfolgen und bei einer Puffergröße von $N_{\text{FFT}}^s = 512$ Abtastwerten würde sich ein Betrachtungszeitraum von 5,12 s ergeben. Diese Länge ist ausreichend, um die Modulationsindizes bzw. die Modulationsveränderung der betrachteten Modulationsfrequenzen auswerten zu können.

Im Anschluss wird die zur STI-Bestimmung benötigte MTF ermittelt, siehe Abbildung 7.15. Dazu wird zunächst eine Normierung der Intensitätseinhüllenden auf die mittlere Leistung vorgenommen. Diese Normierung wird bei einem sprachbasierten STI-Verfahren angewendet, um die zugehörigen Spektren der beiden betrachteten Signale ins Verhältnis setzen zu können [GG04]. Die normierten Zeitsignale werden im Anschluss durch eine diskrete Fourier-Transformation in den Frequenzbereich überführt und im Anschluss, wie in Abbildung 7.15 dargestellt, das Betragsspektrum berechnet. Für die beiden Signale ergeben sich somit die oktavbandabhängigen, normierten Spektren der Intensitätseinhüllenden durch:

$$I_x^{(f_i)}(\mu, d) = \left| \sum_{n=0}^{N_{\text{FFT}}^s-1} \frac{\tilde{x}_d^{(f_i)}(n)}{\frac{1}{N_{\text{FFT}}^s} \sum_{n=0}^{N_{\text{FFT}}^s-1} |\tilde{x}_d^{(f_i)}(n)|} e^{-j \frac{2\pi}{N_{\text{FFT}}^s} \mu n} \right| \quad (7.66)$$

und

$$I_y^{(f_i)}(\mu, d) = \left| \sum_{n=0}^{N_{\text{FFT}}^s - 1} \frac{\tilde{y}_d^{(f_i)}(n)}{\frac{1}{N_{\text{FFT}}^s} \sum_{n=0}^{N_{\text{FFT}}^s - 1} |\tilde{y}_d^{(f_i)}(n)|} e^{-j \frac{2\pi}{N_{\text{FFT}}^s} \mu n} \right|. \quad (7.67)$$

Diese beiden Spektren werden nun verwendet, um die MTF der entsprechenden Modulationsfrequenzen zu bestimmen. Es folgt, dass die mittlere MTF

$$M^{(f_i)}(\mu) = \frac{1}{D} \sum_{d=0}^{D-1} \frac{I_x^{(f_i)}(\mu, d)}{I_y^{(f_i)}(\mu, d) + \epsilon} \quad (7.68)$$

ist, wobei D der Anzahl von gepufferten Blöcken und ϵ dem kleinsten möglichen Wert innerhalb des umgesetzten Systems entspricht. Der Wert ϵ ermöglicht eine Vermeidung der Division durch Null und entsprechende Fehler, ohne das eigentliche Ergebnis drastisch zu verfälschen. Diese Bestimmung der MTF ist ein entscheidender Unterschied zur konventionellen Berechnung, da alle Frequenzen und nicht nur die Modulationsfrequenzen im 1/3-Oktavband betrachtet werden. Diese modifizierte MTF kann nun verwendet werden, um entsprechend des im Standard [IEC] beschriebenen Verfahrens, den STI zu ermitteln. Dazu wird zunächst der sogenannte „scheinbare“ oder effektive SNR mit:

$$\text{SNR}_{\text{app}}^{(f_i)}(\mu) = \min \left\{ \max \left\{ 10 \cdot \log_{10} \left(\frac{M^{(f_i)}(\mu)}{1 - M^{(f_i)}(\mu)} \right), -15 \right\}, 15 \right\} \quad (7.69)$$

bestimmt. Da die Bestimmung des effektiven SNR zu Werten im Unendlichen führen kann, werden die Werte auf einen Bereich zwischen -15 und 15 dB beschränkt [IEC]. Im nächsten Schritt wird der Übertragungsindex mit:

$$\text{TI}^{(f_i)}(\mu) = \frac{\text{SNR}_{\text{app}}^{(f_i)}(\mu) + 15}{30} \quad (7.70)$$

bestimmt. Anschließend wird der Modulationsübertragungsindex pro Oktavband berechnet. Dazu wird nicht der gesamte, ermittelte spektrale Bereich ausgewertet, sondern nur die relevanten Frequenzen zwischen 0,63 und 12,5 Hz. Der Modulationsübertragungsindex ist gegeben durch:

$$\text{MTI}^{(f_i)} = \frac{1}{\mu_{\text{end}}^m - \mu_{\text{start}}^m + 1} \cdot \sum_{\mu=\mu_{\text{start}}^m}^{\mu_{\text{end}}^m} \text{TI}^{(f_i)}(\mu), \quad (7.71)$$

wobei μ_{start}^m dem ersten und μ_{end}^m dem letzten zu betrachtenden Frequenzband entspricht. Der STI kann nun mittels einer gewichteten Summe der Modulationsübertragungsindizes zu

$$\text{STI} = \sum_{i=0}^{N_{\text{oct}}} \text{MTI}^{(f_i)} \cdot w^{(f_i)} \quad (7.72)$$

bestimmt werden, wobei $w^{(f_i)}$ den einzelnen, oktavbandabhängigen Gewichtungsfaktoren entspricht. Diese Gewichte sind so gewählt, dass die Summe aller Gewichte gleich Eins ist und so der resultierende STI-Wert zwischen 0 und 1 liegt. Im weiteren Verlauf sind die folgenden Gewichte aus [SH85] zur Bestimmung des STI verwendet worden: $w^{(f_0)} = 0,13$, $w^{(f_1)} = 0,14$, $w^{(f_2)} = 0,11$, $w^{(f_3)} = 0,12$, $w^{(f_4)} = 0,19$, $w^{(f_5)} = 0,17$ und $w^{(f_6)} = 0,14$.

Qualitätsabschätzung

Innerhalb der Fahrgastzelle werden in einem realistischen Evaluierungsszenario mehr als nur ein Kanal des künstlichen Zuhörers und auch des künstlichen Sprechers verwendet. Das bedeutet, dass eine Erhebung des STI über alle möglichen Kanäle des künstlichen Sprechers l_K und des künstlichen Zuhörers m_K erfolgt. Diese Einzelergebnisse sollen jedoch in einem einzelnen Qualitätsergebnis des Verfahrens münden. Zu diesem Zweck wird der „schlechteste“ Fall zur weiteren Qualitätsbestimmung herangezogen. Das Ergebnis der STI-Berechnung ist definiert durch:

$$\widehat{\text{STI}} = \min \{ \text{STI}(l_K, m_K) \} \quad (7.73)$$

Diesen geschätzten STI-Wert gilt es nun mithilfe des in Kapitel 4.2.1 beschriebenen Verfahrens auf eine auditive Qualitätsskala zu projizieren. Dazu werden die aus Kapitel 6.4 erhobenen auditiven Ergebnisse MOS_S verwendet und die zugehörigen geschätzten STI-Werte berechnet. Zunächst ist ein lineares Modell nach dem in Kapitel 4.2.1 beschriebenen Verfahren hergeleitet und überprüft worden. Hierzu werden die Regressionskoeffizienten, unter Ausschluss eines Datenpunktes ermittelt und überprüft, ob diese innerhalb des 95%-Konfidenzintervalls liegen. Das Ergebnis ist bereits beispielhaft in Kapitel 4.3 in Abbildung 4.3 gezeigt worden. Es kann festgehalten werden, dass alle Koeffizienten im Konfidenzintervall liegen und somit die Unterschiede zwischen den einzelnen Modellen nicht signifikant sind. Schließlich sind jene Koeffizienten verwendet worden, welche mithilfe aller Datenpunkte ermittelt worden sind. Das so erzeugte Modell ist zusätzlich auf eine Qualitätsskala zwischen 1 und 4 transformiert worden, wobei 1 einem sehr guten und 4 einem sehr schlechten Ergebnis entspricht. Dadurch ergibt sich die gleiche Skala wie innerhalb der Verfahren aus Kapitel 7.2 und eine Kombination zu einer Gesamtqualität ist möglich. Das resultierende lineare Modell ist gegeben durch:

$$\check{c}_{\text{SI}} = 5,56 - 6,7 \cdot \widehat{\text{STI}}. \quad (7.74)$$

Zusätzlich werden auch die Ergebnisse des auditiven Tests aus Kapitel 6.4 entsprechend auf die vorgegebene Qualitätsskala normiert, so dass sich die auditiven, normierten Ergebnisse $\widehat{\text{MOS}}_S$ ergeben.

In Abbildung 7.16 ist das zugehörige Streudiagramm des instrumentellen Maßes \check{c}_{SI} in Abhängigkeit der auditiven Ergebnissen $\widehat{\text{MOS}}_S$ dargestellt. Zusätzlich ist die Winkelhalbierende mit gegeben, welche bei der optimalen Schätzung durch das instrumentelle

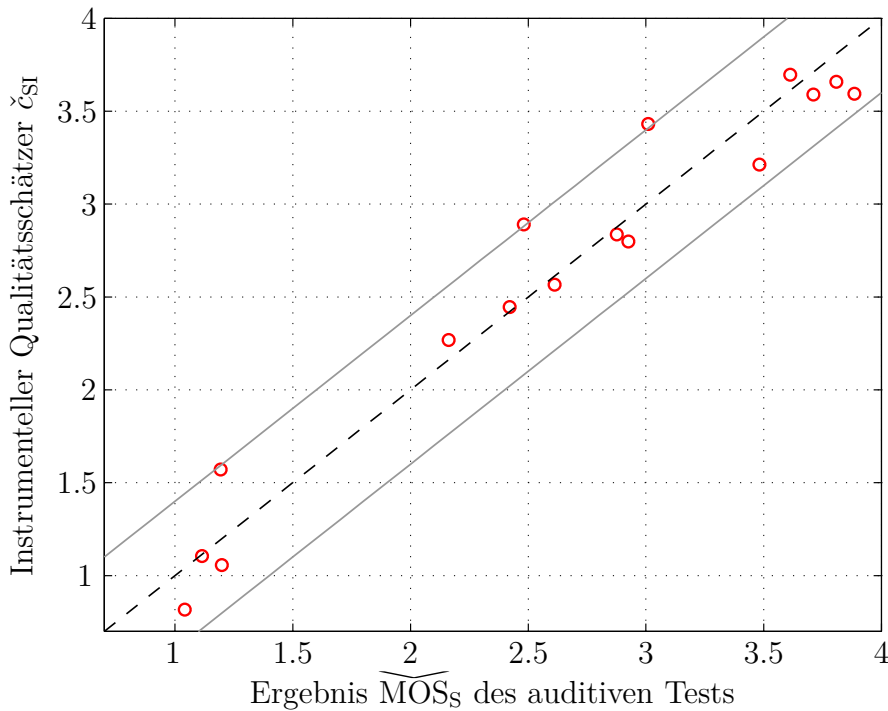


Abbildung 7.16: Streudiagramm in Abhängigkeit der normierten auditiven Qualitätsergebnisse und der Ergebnisse ermittelt durch den modifizierten STI \check{c}_{SI} . Dabei entspricht 1 einem außerordentlich gutem und 4 einem mangelhaften Qualitätsergebnis.

Verfahren entspricht. Die beiden grauen Kurven bilden ein 10%-Intervall in Bezug auf diese Winkelhalbierende ab. Zusehen ist, dass das Modell eine annehmbare Schätzung der Qualität erzielt, da die Ergebnisse zum größten Teil nahe der Winkelhalbierenden liegen.

Des Weiteren sind in Tabelle 7.10 der mittlere quadratische Fehler nach Gleichung (4.7) und die Korrelation nach Gleichung (4.6) für alle Datensätze und für die Ergebnisse der *jackknife*-Methode, siehe Kapitel 4.3, gegeben. Es wird deutlich, dass in beiden Betrachtungen bereits hohe Korrelationen erzielt werden können. Die RMSE-Werte liegen in einem akzeptablen Bereich und entsprechen circa 7,4% und 8,5% der verwendeten Qualitätsskala. Zusätzlich kann bei einem $\widehat{MOS}_S \approx 1$ ein Ausreißer aus dem betrachteten Qualitätsbereich entdeckt werden. Diese Ergebnisse deuten darauf hin, dass das vorgestellte Verfahren eine erste erfolgversprechende Schätzung der Qualität erzielen kann. Es sind jedoch weitere auditive Qualitätserhebungen notwendig, um Modellfehler auszuschließen und zusätzlich eine aussagekräftigere Validierung des Modells zu erzeugen.

Da je nach betrachtetem Szenario innerhalb der Fahrgastzelle eine weniger oder stärker beeinträchtigte Kommunikationssituation vorherrschen kann, soll nicht die absolute Qualität eines ICC-Systems mithilfe des modifizierten STI bestimmt werden, sondern eine relative. Dazu wird sowohl das Qualitätsergebnis des modifizierten STI mit $\check{c}_{SI}^{(on)}$, als

ρ	ρ_{LOO}	RMSE	RMSE _{LOO}
0,97	0,97	0,22	0,26

Tabelle 7.10: Empirischer Korrelationskoeffizient und RMSE aller Datensätze (ρ , RMSE) und der jackknife-Methode (ρ_{LOO} , RMSE_{LOO}).

auch ohne ein ICC-System $\check{c}_{\text{SI}}^{(\text{off})}$, ermittelt. Wichtig bei dieser Bestimmung der Ergebnisse ist es, ein Geräuschkennszenario mit hoher Leistung, zum Beispiel bei einer Geschwindigkeit von 150 km/h, zu verwenden, um so das ICC-System ausreichend anzuregen und eine Ausgabe des Nutzsymbols durch das ICC-System zu garantieren. Unter Verwendung des Systemattributes „Notwendigkeit“, siehe Kapitel 7.3.1, wird der relative Qualitätsgewinn gewichtet und im Gesamtergebnis des Verfahrens mit berücksichtigt. Das Qualitätsergebnis des modifizierten STI ist gegeben durch:

$$c_{\text{SI}} = \min \left\{ \max \left\{ 1, \check{c}_{\text{SI}}^{(\text{on})} + g_{\text{R}} \cdot \left(\check{c}_{\text{SI}}^{(\text{on})} - \check{c}_{\text{SI}}^{(\text{off})} \right) \right\}, 4 \right\}. \quad (7.75)$$

Diese relative Betrachtung der Qualitätsergebnisse ermöglicht es zum einen, den relativen Qualitätsgewinn durch die Verwendung eines ICC-Systems, und zum anderen, die Fahrzeugumgebung, mit zu berücksichtigen. Dadurch schneiden ICC-Systeme, welche nur wenig Qualitätsgewinn im Sinne des STI erzielen, schlechter ab als solche mit einem hohem Gewinn. Aber auch Systeme, welche in einer Umgebung operieren, die nicht beeinträchtigt ist und somit bereits hohe Sprachqualität besitzen, erzielen ein geringeres Qualitätsergebnis. Zusätzlich wird das Ergebnis auf den definierten Qualitätsbereich begrenzt, da eine höhere Einstufung als exzellent, oder eine geringere als mangelhaft, nicht sinnvoll ist.

7.3.3 Modifizierter Signal-Rauschabstand

Ein häufig verwendetes Maß für die Evaluierung von Sprachqualität, beispielsweise von Wellenform-Kodierern, ist der klassische Signal-Rauschabstand (engl. *Signal-to-Noise Ratio*, SNR). Dieser schätzt über das Verhältnis zwischen Nutzsymbols- und Geräuschleistung, die Qualität eines Sprachsignals. Ein ICC-System soll genau dieses Verhältnis erhöhen, um so die Kommunikationssituation innerhalb der Fahrgastzelle zu verbessern. Deswegen darf eine SNR-Berechnung im Grunde bei der Evaluierung von ICC-Systemen nicht fehlen. Unter der Annahme, dass es sich bei $y(n)$ um das nicht veränderte Eingangssymbol eines Systems und bei $x(n)$ um das Ausgangssymbols des zu evaluierenden Systems handelt, wobei beide Folgen der Länge N entsprechen, ist der klassische SNR gegeben durch:

$$\text{SNR} = 10 \log_{10} \left(\frac{\sum_{n=0}^{N-1} y^2(n)}{\sum_{n=0}^{N-1} (y(n) - x(n))^2} \right). \quad (7.76)$$

Damit diese Berechnung durchführbar ist, müssen beide Signale zeitlich synchron zueinander laufen und das Geräusch exakt bestimmbar sein [QBC88, S. 44]. Bereits in Kapitel 6.3 ist dargestellt worden, dass der klassische SNR nur einen dürftigen Schätzer für die Qualität von Sprachsignalen darstellt.

Dennoch können durch einige Modifikationen des klassischen Ansatzes, neue Maße wie der segmentale oder der frequenzgewichtete segmentale SNR bestimmt werden, welche bereits vielversprechende Ergebnisse bei der Evaluierung, zum Beispiel von Wellenform-Kodierern, erzielen konnten [QBC88, TNMC78]. Zusätzlich ist in Kapitel 6.3 die Diskrepanz zwischen konventioneller Bestimmung des SNRs und der auditiven Wahrnehmung diskutiert worden. Des Weiteren konnte in [GTHS] festgestellt werden, dass der frequenzgewichtete segmentale SNR auch innerhalb einer Fahrzeugumgebung erfolgversprechende Ergebnisse erzielt. Ein solches Verfahren wird im weiteren Verlauf diese Kapitels zur Schätzung der Qualität eines ICC-Systems verwendet.

Entwurf des Verfahrens

Zu Beginn des Verfahrens wird das Betragsspektrum des aktuellen Hintergrundgeräusches geschätzt. Dazu wird das Mikrofonsignal $x(n)$, siehe Abbildung 5.1, mittels der in (5.1) definierten Analysefilterbank in den Frequenzbereich überführt und der Betrag des Spektrums über eine bestimmte Anzahl von Rahmen K_{NE} gemittelt. Während dieser Schätzung darf kein Signal durch den künstlichen Sprecher ausgegeben werden, da nur die Leistung des aktuellen Geräusches geschätzt werden soll. Das mittlere Betragsspektrum $\bar{N}_{\text{SNR}}(\mu)$ ist definiert durch:

$$\bar{N}_{\text{SNR}}(\mu) = \frac{1}{K_{\text{NE}}} \sum_{k=0}^{K_{\text{NE}}-1} |X(\mu, k)|. \quad (7.77)$$

Da die Evaluierung eines ICC-Systems innerhalb einer Laborumgebung durchgeführt wird, in welcher das Hintergrundgeräusch mithilfe einer Geräuschsimulation erzeugt wird, kann man von einem annähernd stationären Geräuschsignal ausgehen. Dadurch kann dieses mittlere Betragsspektrum des Geräusches $\bar{N}_{\text{SNR}}(\mu)$ als eine gute Näherung des tatsächlichen Geräusches angesehen werden. Dies ermöglicht eine Schätzung des Betragsspektrums des Sprachsignales, angeregt durch den künstlichen Sprecher, am Eingang der ICC-Systemevaluierung $x(n)$. Dieses geschätzte Betragsspektrum des Sprachsignals ist gegeben durch:

$$S_c(\mu, k) = \begin{cases} |X(\mu, k)| - \bar{N}_{\text{SNR}}(\mu) & \text{wenn } |X(\mu, k)| > \bar{N}_{\text{SNR}}(\mu), \\ \bar{N}_{\text{SNR}}(\mu) & \text{sonst.} \end{cases} \quad (7.78)$$

Auch hier wird das Eingangssignal der ICC-Systemevaluierung $x(n)$ mithilfe der Analysefilterbank (5.1) in den Frequenzbereich überführt. In diesem Fall muss ein Sprachsignal durch den künstlichen Sprecher $y(n)$, welcher in Abbildung 5.1 dargestellt ist, angeregt werden. Zusätzlich werden jene Frequenzbänder, in welchen das Betragsspektrum vollständig von Geräusch überlagert wird, durch Nullsetzen aus der Berechnung herausgenom-

men. Unter Verwendung der Schätzung des Betragsspektrums des Sprachsignals $S_c(\mu, k)$ und des Geräuschsignals $\bar{N}_{\text{SNR}}(\mu)$, kann nun eine Schätzung des SNR für das betrachtete Frequenzband im aktuellen Rahmen durchgeführt werden. In [GTHS] ist ein Verfahren vorgestellt worden, welches den frequenzgewichteten segmentalen SNR bestimmt. Diese Bestimmung des SNRs hat nicht nur dort, sondern auch in [HL08] vielversprechende Ergebnisse bei der Evaluierung von Sprachverbesserungssystemen erzielen können. Zunächst wird der lineare SNR innerhalb des aktuellen Rahmens in Abhängigkeit des Frequenzbandes durch

$$\widehat{\text{SNR}}_{\text{lin}} = \frac{S_c(\mu, k)^2}{\bar{N}_{\text{SNR}}(\mu)^2} \quad (7.79)$$

bestimmt. Die Fallunterscheidung ermöglicht, dass jene Frequenzbänder, welche durch die spätere Gewichtung ausgeblendet werden, im logarithmischen Bereich keinen Beitrag bei der Mittelung leisten. Unter Verwendung des aktuellen linearen SNR $\widehat{\text{SNR}}_{\text{lin}}$ und dem Betragsspektrum des Sprachsignals $S_c(\mu, k)$, kann nun die Frequenzgewichtung und die Mittelung berechnet werden. Dazu wird eine vom Sprachsignal abhängige Gewichtungsfunktion, welche durch

$$w(\mu, k) = \left(S_c(\mu, k) \right)^{0,2} \quad (7.80)$$

gegeben ist, verwendet [HL08]. Die finale Mittelung des logarithmischen SNRs über alle Bänder und alle betrachteten Rahmen K wird bestimmt durch:

$$\widehat{\text{SNR}}_{\text{fw,seg}} = \frac{1}{K} \sum_{k=0}^{K-1} \frac{\sum_{\mu=0}^{N_{\text{FFT}}/2} 10 \cdot \log_{10} \widehat{\text{SNR}}_{\text{lin}} \cdot w(\mu, k)}{\sum_{\mu=0}^{N_{\text{FFT}}/2} w(\mu, k)}. \quad (7.81)$$

Diese Bestimmung des frequenzgewichteten segmentalen SNR ist bereits in [HL08] veröffentlicht worden, wobei dort das Sprach- und das Geräuschsignal separat vorlagen und nicht wie hier vorgestellt, geschätzt werden müssen.

Qualitätsabschätzung

Auch hier ist für die Herleitung des Evaluierungsverfahrens bisher nur ein Mikrofonsignal betrachtet worden, diese Vereinfachung wird nun aufgehoben, um so die verschiedenen Mikrofone in der Evaluierung zu berücksichtigen. Deswegen wird die SNR-Schätzung um die Abhängigkeit des Eingangskanals m_K erweitert. Unter Verwendung des Zusammenhangs (7.81) und der Annahme, dass ein pessimistischer Schätzer ermittelt werden soll, folgt für das Endergebnis dieses Verfahrens:

$$\widehat{\text{SNR}} = \min \left\{ \widehat{\text{SNR}}_{\text{fw,seg}}(m_K) \right\}. \quad (7.82)$$

Wie auch im Verfahren zuvor, muss im nächsten Schritt dieses Ergebnis auf die verwendete Qualitätsskala projiziert werden. Zu diesem Zweck wird ein lineares Modell entwickelt und die zugehörigen Koeffizienten in Bezug auf ihr 95%-Konfidenzintervall ausgewertet. Diese lagen alle innerhalb des Konfidenzintervalls, sodass alle Daten zur endgültigen Bestimmung des Modells verwendet worden sind. Die Ergebnisse der Koeffizientenanalyse sind im Anhang A dargestellt. Folgendes Modell ist für die Schätzung der Qualität verwendet worden:

$$\check{c}_{\text{MS}} = -1,95 - 0,63 \cdot \widehat{\text{SNR}}. \quad (7.83)$$

In Abbildung 7.17 sind die normierten auditiven Ergebnisse $\widehat{\text{MOS}}_S$ und die durch das beschriebene Verfahren ermittelten instrumentellen Ergebnisse \check{c}_{MS} , in einem Streudiagramm entsprechend dargestellt.

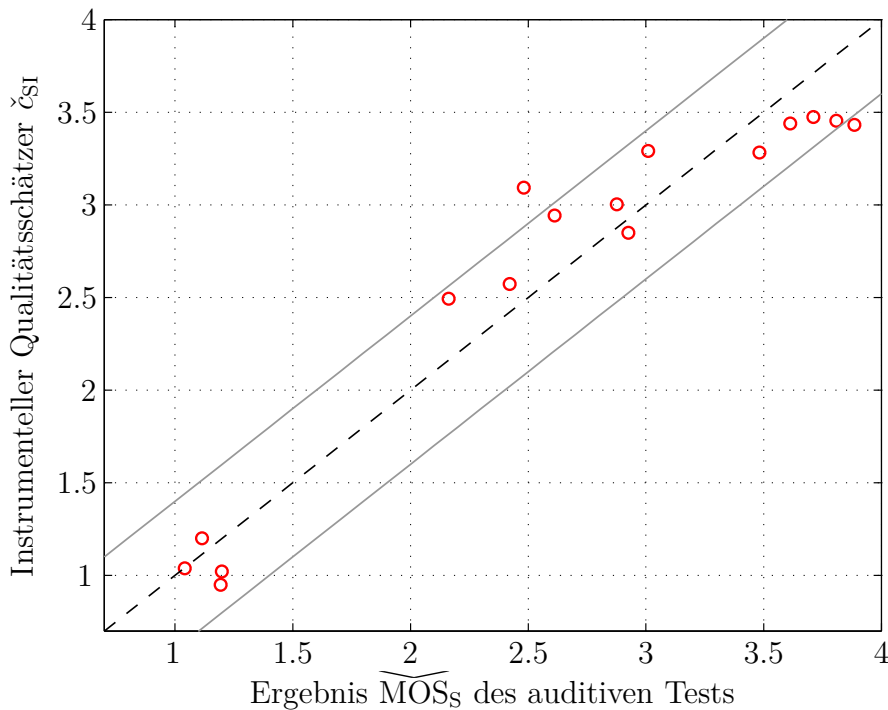


Abbildung 7.17: Streudiagramm in Abhängigkeit der normierten auditiven Qualitätsergebnisse und der Ergebnisse ermittelt durch den modifizierten SNR \check{c}_{MS} . Dabei entspricht 1 einem außerordentlich gutem und 4 einem mangelhaften Qualitätsergebnis.

Auch in diesem Fall gibt es einen Ausreißer, welcher einen Wert \check{c}_{MS} kleiner 1 annimmt, bei einem normierten auditiven Ergebnis $\widehat{\text{MOS}}_S$ von circa 1,2, zu verzeichnen. Dennoch konnten sowohl für das *jackknife*-Verfahren als auch in der Betrachtung aller Datenpunkte, hohe Korrelationen, siehe Tabelle 7.11, erzielt werden. Es wird deutlich, dass die mittleren Fehler bei circa 9,3% und 12% der vorgegebenen Qualitätsskala liegen und somit größere

Fehler als bei der Bestimmung des STI aufgetreten sind. Dennoch können auch hier erste vielversprechende Ergebnisse zur Schätzung der Qualität erzielt werden. Als kritisch ist die Anzahl der vorliegenden Datensätze von 16 zu erwähnen, diese sollte in jedem Fall bei weiteren Betrachtungen erhöht werden. Die Gesamtqualitätsbestimmung dieses Evaluie-

ρ	ρ_{LOO}	RMSE	RMSE _{LOO}
0,96	0,93	0,28	0,36

Tabelle 7.11: Empirischer Korrelationskoeffizient und RMSE aller Datensätze (ρ , RMSE) und der jackknife-Methode (ρ_{LOO} , RMSE_{LOO}).

rungsverfahrens erfasst, ebenso wie in der Bestimmung des Endergebnisses des STI, die relative Veränderung des SNR durch die Verwendung eines ICC-Systems. Zu diesem Zweck wird innerhalb eines vordefinierten Geräuschszenarios, welches eine ausreichend hohe Geräuschleistung besitzen sollte, beispielsweise bei einer Geschwindigkeit von 150 km/h, der SNR mit $\check{c}_{\text{MS}}^{(\text{on})}$ und ohne aktivem ICC-System $\check{c}_{\text{MS}}^{(\text{off})}$ ermittelt. Das Qualitätsergebnis dieser relativen Betrachtung ist durch folgenden Zusammenhang definiert:

$$c_{\text{MS}} = \min \left\{ \max \left\{ 1, \check{c}_{\text{MS}}^{(\text{on})} + g_{\text{R}} \cdot \left(\check{c}_{\text{MS}}^{(\text{on})} - \check{c}_{\text{MS}}^{(\text{off})} \right) \right\}, 4 \right\}. \quad (7.84)$$

Dabei entspricht g_{R} dem Systemattribut „Notwendigkeit“ aus Kapitel 7.3.1. Zusätzlich ist auch hier der Qualitätsbereich zahlenmäßig auf einen Bereich zwischen 1 und 4 begrenzt worden.

7.3.4 Spektrale Distanz

Wie bereits in einigen Kapiteln, beispielsweise in 6.1, erwähnt, entstehen durch die Verwendung eines ICC-Systems mehrere Signalquellen des Nutzsignals. Durch diese erneute Wiedergabe des Nutzsignals, verzögert zu dem Direktschall, kann die empfundene Halligkeit des Raumes erhöht werden, was einen Zuhörer stören kann. Diese zweite Signalquelle hat deswegen sowohl einen positiven als auch einen negativen Einfluss auf die Kommunikationsqualität innerhalb der Fahrgastzelle. In der oberen Grafik von Abbildung 7.18 ist neben Impulsantwort des Direktschalls (rot) zusätzliche eine theoretische Impulsantwort des ICC-Systems (hellblau) dargestellt. Diese Impulsantwort der Übertragung über das ICC-System ist nur theoretisch, da sich die Übertragungseigenschaften für gewöhnlich über die Zeit und das gegebene Szenario verändert. Dessen ungeachtet, wird deutlich, dass die zusätzliche Wiedergabe zeitlich verschoben und mit einer Verstärkung, im Vergleich zum Direktschall, erfolgt.

Teile des in diesem Abschnitt vorgestellten Evaluierungsverfahrens sind bereits in [TSb] veröffentlicht worden.

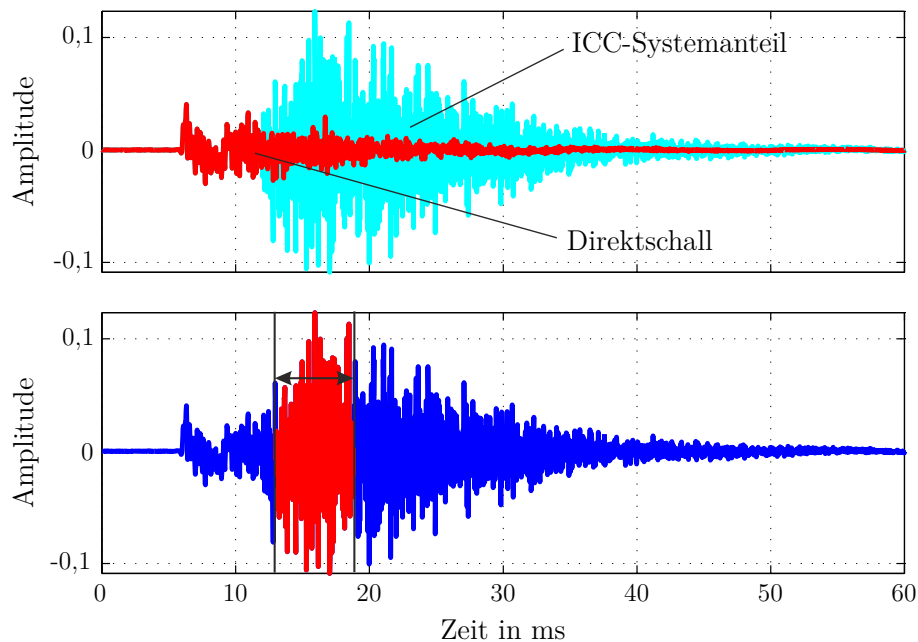


Abbildung 7.18: In der oberen Grafik sind die Impulsantworten des Direktschalls vom Sprechermund zum Zuhörerohr (rot) und eine theoretischen Impulsantwort eines beliebigen ICC-Systems (hellblau) dargestellt. Die untere Grafik veranschaulicht die Übertragung des Nutzsignals innerhalb der Fahrgastzelle im Ganzen. Dabei entspricht die resultierende Impulsantwort einer Aufsummierung der beiden Einzelimpulsantworten. Zusätzlich ist der für die Qualitätsschätzung essentielle Teil der Gesamtimpulsantwort in der unteren Grafik in rot markiert.

Diese beiden Komponenten, additiv überlagert, bilden das Signal, welches vom Zuhörer wahrgenommen und überdies bewertet wird. Das im Folgenden vorgestellte Evaluierungsverfahren soll mittels eines spektralen Abstands diese Qualität schätzen. Im Allgemeinen vergleichen Ansätze, basierend auf einem spektralen Abstand, immer das Spektrum des klaren Sprachsignals mit dem eines, beispielsweise durch einen Kodierer, verzerrten Sprachsignals. Dabei verzerrt ein gutes System das klare Sprachsignal nur minimal und der spektrale Abstand fällt dadurch ebenso gering aus. Die Referenz, beispielsweise im Falle eines Kodierers, ist das klare Sprachsignal am Eingang des Systems. Bei einem ICC-System ist die Bestimmung eines Referenzsignals nicht trivial, da das Sprachsignal des Sprechers immer über den Raum übertragen wird und so eine spektrale Veränderung des Sprachsignals erfolgt. Um trotzdem ein spektrales Abstandsmaß für die Qualitätsschätzung verwenden zu können, muss zunächst die Referenz bestimmt werden, um im Anschluss die Distanz zwischen Referenz und Ohrsignal zu ermitteln. Dabei wird innerhalb des Verfahrens der Fokus auf den dominanten bzw. leistungsstärksten Teil des Signals gelegt. In der unteren Grafik von Abbildung 7.18 ist dieser Bereich durch die rote Einfärbung der Impulsantwort illustriert. Die genaue Bedeutung dieses Bereichs wird im späteren Verlauf dieses Kapitels im Detail erläutert.

Entwurf des Verfahrens

Zu Beginn der Erhebung der spektralen Distanz zweier Signale, müssen diese zunächst zeitlich synchron zueinander sein. Nur durch diese Synchronisation wird gewährleistet, dass gleiche Sprachanteile des Signals miteinander verglichen werden. Hierzu wird zunächst die Verzögerung zwischen künstlichem Sprecher und künstlichem Zuhörer ermittelt. Wie in den Evaluierungsverfahren zuvor, wird auch in diesem von dem in Abbildung 5.1 dargestellten Evaluierungsaufbau ausgegangen. Die Verzögerung wird mit einem Sprachsignal, bestehend sowohl aus Sprachpassagen eines männlichen Sprechers, als auch einer weiblichen Sprecherin, welches über den künstlichen Mund $y(n)$ ausgegeben wird und am Ohr des künstlichen Zuhörers $x(n)$ aufgenommen wird, gemessen. Diese beiden Signale werden mit dem gleichen Filter aus Kapitel 7.2.2, welches eine Grenzfrequenz von 100 Hz besitzt, zunächst hochpassgefiltert. Diese beiden hochpassgefilterten Signale $\tilde{y}(n)$ und $\tilde{x}(n)$ werden im Anschluss mit einer Kreuzkorrelationsanalyse verglichen und die Verzögerung geschätzt. Die Bestimmung des Argumentes des Maximums der Kreuzkorrelationsfolge ergibt die Verzögerung und ist bestimmt durch:

$$\Delta = \arg \max_{\kappa} \left\{ \left| \mathbb{E} \{ \tilde{x}(n) \tilde{y}(n + \kappa) \} \right| \right\}. \quad (7.85)$$

Diese Berechnung der Verzögerung legt den Fokus auf jenen Signalanteil am Zuhörerohr, welcher die größte Leistung besitzt. Dadurch wird beispielsweise bei der Übertragung eines Signals mit den Impulsantworten aus Abbildung 7.18 nicht die Verzögerung des Direktschalls ermittelt, sondern die zwischen dem Sprecher und jenem Schallereignis, welches vom ICC-System ausgesendet wird. Es würde eine Verzögerung von $\Delta = 17$ ms berechnet, welche zur zeitlichen Anpassung der beiden Signale verwendet würde. Dadurch wird auch der Fokus der spektralen Distanzberechnung auf im Hinblick auf die Leistung dominante Anteile des Sprachsignals, angeregt durch das ICC-System, gelegt.

Die Verzögerung Δ wird nun für alle weiteren Berechnungen verwendet, in dem, wie in Abbildung 7.19 dargestellt, das Testsignal des künstlichen Sprechers um diesen Wert Δ verzögert wird.

Neben der Wiedergabe des Testsignals über den künstlichen Sprecher und dem Eingang des übertragenen Signals am Ohr des künstlichen Zuhörers, ist in Abbildung 7.19 die für die Erhebung der spektralen Distanz verwendete Signalverarbeitung dargestellt. Nach der Bestimmung und der Anpassung der Verzögerung der beiden Signale wird eine Verstärkungsanpassung des Referenzsignals ermittelt. Diese soll die erwartete betragsmäßige Veränderung des Spektrums des Sprachsignals vom Sprechermund zum Zuhörerohr abbilden. Dabei wird davon ausgegangen, dass die Signalverarbeitung eines ICC-Systems im Mittel keine drastischen qualitätsmindernden spektralen Veränderungen vornimmt. Auch die Verstärkungsanpassung wird innerhalb einer Blockverarbeitung realisiert, die zunächst, wie in Abbildung 7.19 dargestellt, das Testsignal $y(n)$ um Δ verzögert und im Anschluss sowohl das verzögerte Testsignal $\check{y}(n)$, als auch das Signal des Ohrmikrofons

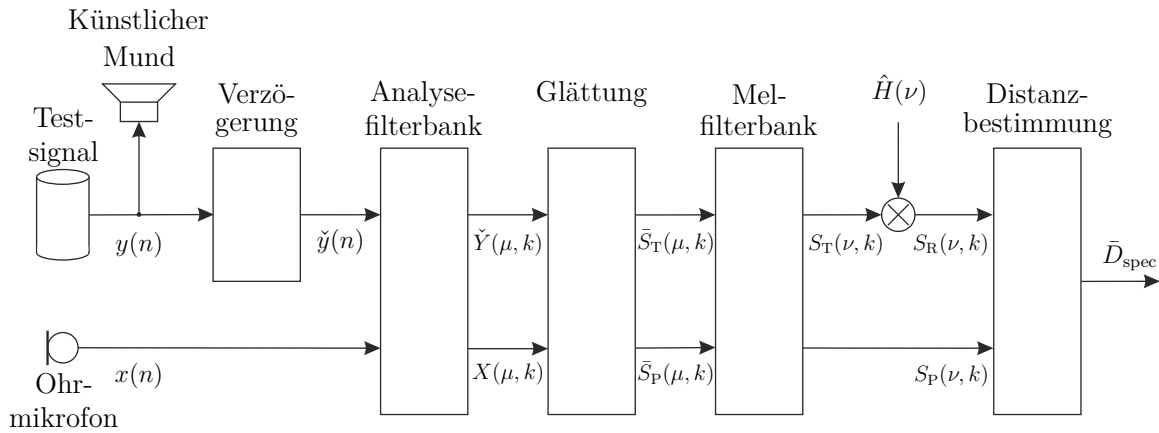


Abbildung 7.19: Signalverarbeitung der spektralen Distanzbestimmung innerhalb der Blocksignalverarbeitung (nach [TSb]).

$x(n)$ mithilfe einer Analysefilterbank in den Frequenzbereich überführt, siehe Gleichung (5.1) und (5.2). Von beiden Spektren wird der Betrag gebildet, der wiederum in auf- und absteigender Frequenzrichtung geglättet wird. Die Glättung verringert den Einfluss einzelner Ausreißer auf das Ergebnis der Schätzung der Spektren. Das resultierende geglättete Betragsspektrum für das verzögerte Testsignal $\check{y}(n)$ in aufsteigender Frequenzrichtung ist gegeben durch:

$$\hat{\check{S}}_T(\mu, k) = \begin{cases} |\check{Y}(\mu, k)|, & \mu = 0, \\ \alpha \cdot \hat{\check{S}}_T(\mu - 1, k) + (1 - \alpha) \cdot |\check{Y}(\mu, k)|, & \text{sonst,} \end{cases} \quad (7.86)$$

und in abfallender Frequenzrichtung durch:

$$\bar{\check{S}}_T(\mu, k) = \begin{cases} |\hat{\check{S}}_T(\mu, k)|, & \mu = N_{\text{FFT}} - 1, \\ \alpha \cdot \hat{\check{S}}_T(\mu + 1, k) + (1 - \alpha) \cdot |\hat{\check{S}}_T(\mu, k)|, & \text{sonst.} \end{cases} \quad (7.87)$$

Die Glättungskonstante α kann hier frei gewählt werden, wobei $0 \leq \alpha \leq 1$ eingehalten werden muss. Eine mögliche Wahl wäre $\alpha = 0,6$, bei einer Abtastrate $f_s = 44,1$ kHz und einer Fenstergröße $N_{\text{FFT}} = 1024$. In einem nächsten Schritt werden beide geglätteten Spektren mit einer Mel-Filterbank gefiltert.

In Abbildung 7.20 sind die Filtergewichte einer Mel-Filterbank mit 24 Bändern beispielhaft dargestellt. Ein jedes Dreieck wird verwendet, um ein Mel-Band zu bestimmen. Zu diesem Zweck werden die einzelnen Frequenzbänder μ mit einem der in Abbildung 7.20 dargestellten Dreiecke multipliziert und aufsummiert. Die betrachteten Bänder reduzieren sich nun, in diesem Fall auf $N_{\text{mel}} = 24$ und werden durch das Symbol ν beschrieben. Die Mel-Filterbank ermöglicht eine Aufteilung in kritische Bänder, welche die Auflösung des menschlichen Gehörs nachbilden soll [Zwi61].

Die Verstärkungsanpassung wird mithilfe eines speziellen Testsignals geschätzt. Dieses Testsignal, welches in Abbildung 7.21 dargestellt ist, besteht aus Sprachpassagen, welche

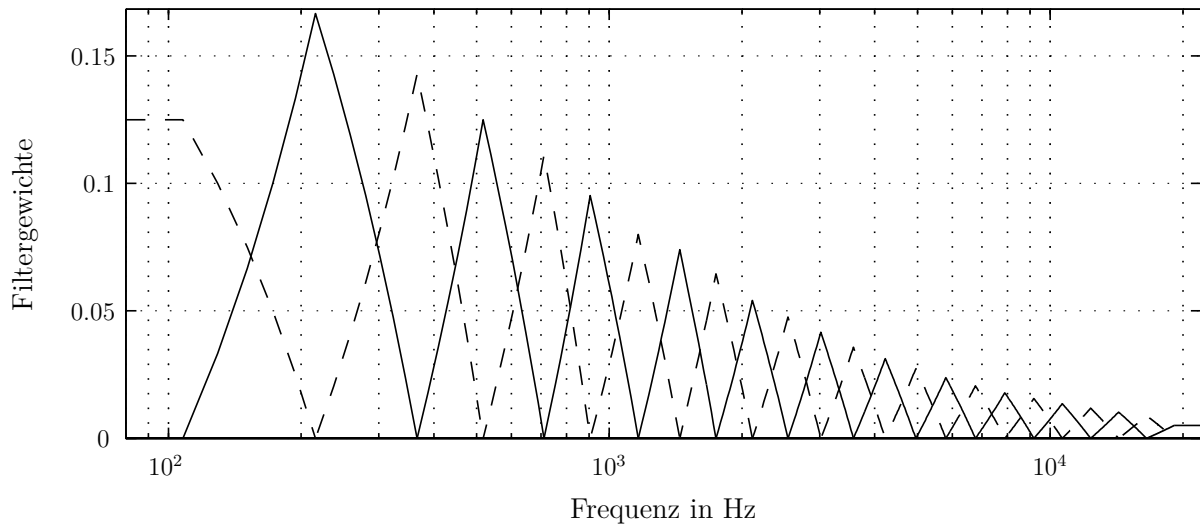


Abbildung 7.20: Filtergewichte einer Melfilterung mit 24 Bändern in Abhängigkeit der Frequenz.

durch sehr kurze Rauschpassagen unterbrochen werden. Diese Rauschpulse sind, wie in Kapitel 7.2.1, so kurz gewählt worden, dass von einem nahezu zeit-invarianten ICC-System ausgegangen werden kann. Zur zusätzlichen Sicherheit und um das ICC-System nicht zu stark zu belasten sind die Rauschpassagen in ihrem Frequenzbereich begrenzt worden. Wie in Abbildung 7.21 zusehen ist, werden drei Bänder nacheinander angeregt.

Diese Rauschpulse werden nun verwendet, um die Verstärkungsanpassung $\hat{H}(\nu)$ zu schätzen. In jedem Rauschpuls wird deswegen das Leistungsdichtespektrum des Testsignals $y(n)$ und des Ohrsignals $x(n)$ geschätzt. Das Leistungsdichtespektrum des Testsignals in dB ist gegeben durch:

$$\tilde{S}_T(\nu) = 20 \cdot \log_{10} \left(\frac{\sum_{k=0}^{K-1} S_T(\nu, k) \cdot a_T^D(\nu, k)}{\sum_{k=0}^{K-1} a_T^D(\nu, k)} \right). \quad (7.88)$$

Die Gewichtungsfunktion ist gegeben durch:

$$a_T^D(\nu, k) = \begin{cases} 1, & \text{wenn } S_T(\nu, k) > T_{\text{act}} \wedge S_T(\nu, k) \text{ einem Rauschpuls entspricht,} \\ 0, & \text{sonst.} \end{cases} \quad (7.89)$$

Dadurch werden innerhalb eines jeden Rahmens nur die Mel-Bänder berücksichtigt, welche ausreichend Leistung anregen bzw. den Schwellwert T_{act} überschreiten. Die mittlere

Leistungsdichte des Ohrsignals wird ähnlich berechnet und ist durch den folgenden Zusammenhang bestimmt:

$$\tilde{S}_P(\nu) = 20 \cdot \log_{10} \left(\frac{\sum_{k=0}^{K-1} S_P(\nu, k) \cdot a_P^D(\nu, k)}{\sum_{k=0}^{K-1} a_P^D(\nu, k)} \right). \quad (7.90)$$

Die Gewichtungsfunktion ist gegeben durch:

$$a_P^D(\nu, k) = \begin{cases} 1, & \text{wenn } S_P(\nu, k) > \bar{N}(\nu) \quad \wedge \quad S_P(\nu, k) \text{ einem Rauschpuls entspricht,} \\ 0, & \text{sonst.} \end{cases} \quad (7.91)$$

In diesem Zusammenhang werden nur Mel-Bänder zur Bestimmung der Verstärkungsanpassung herangezogen, welche mehr Leistung als die mittlere Leistung des Hintergrundgeräusches $\bar{N}(\nu)$ besitzen und einem Rauschpuls entsprechen. Das Hintergrundgeräusch

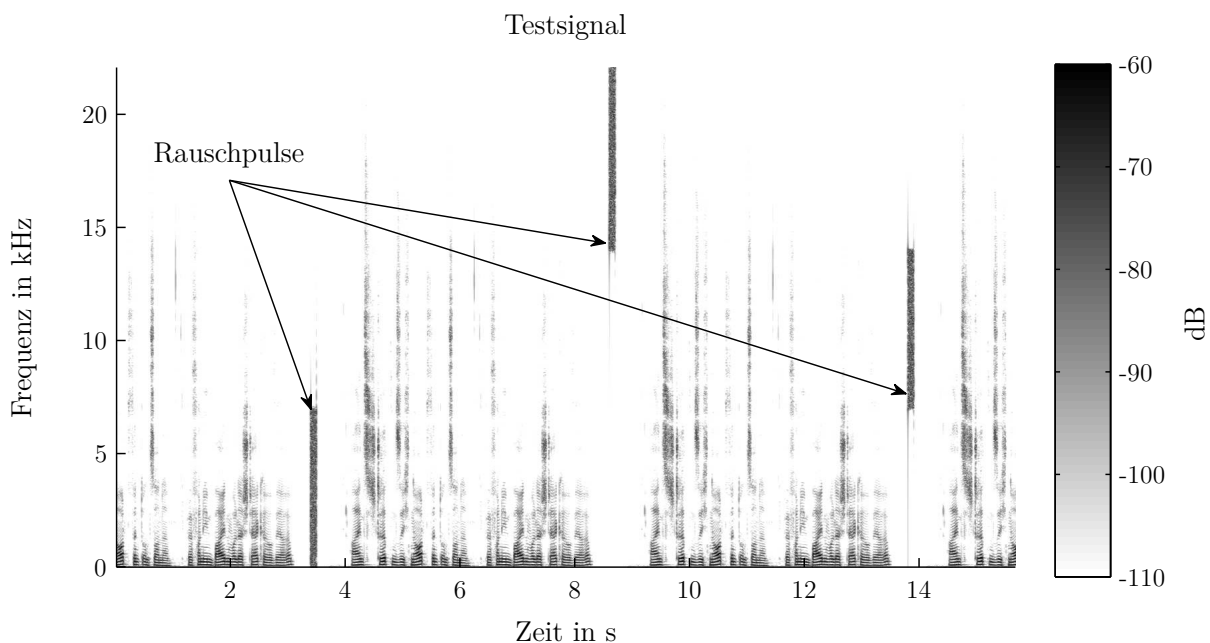


Abbildung 7.21: Zur Bestimmung der Verstärkungsanpassung verwendetes Testsignal. Dieses besteht aus Sprachanteilen und einzelnen, bandbegrenzten Rauschpulsen mit einer Dauer von circa 100 ms und einer über den Frequenzbereich gleichbleibenden Leistung.

wird zu Beginn in einer Phase in welcher kein Signal durch den künstlichen Sprecher angeregt wird unter Verwendung des Ohrsignals bestimmt. Es ergibt sich, dass

$$\bar{N}(\nu) = \frac{1}{K_{SD}} \sum_{k=0}^{K_{SD}-1} S_P(\nu, k) \quad (7.92)$$

ist, wobei K_{SD} der Anzahl der betrachtende Rahmen entspricht. Schließlich kann die Verstärkungsanpassung durch

$$\hat{H}_{\log}(\nu) = \tilde{S}_T(\nu) - \tilde{S}_P(\nu) \quad (7.93)$$

bestimmt werden. Diese Verstärkungsanpassung wird in dB ermittelt und muss, um sie entsprechend Abbildung 7.19, multiplikativ auf das geglättete Spektrum des Testsignals $S_T(\nu, k)$ anwenden zu können, in lineare Werte zurück gerechnet werden. Zusätzlich sollte die Verstärkungsanpassung in ihrer Dämpfungs- und Verstärkungseigenschaften limitiert werden. Die lineare Verstärkungsanpassung ist gegeben durch:

$$\hat{H}(\nu) = 10 \left(\min \left\{ \bar{H}_{\log} + H_T, \max \left\{ \bar{H}_{\log} - H_T, \hat{H}_{\log}(\nu) \right\} \right\} / 20 \right). \quad (7.94)$$

Die mittlere logarithmische Verstärkungsanpassung

$$\bar{H}_{\log} = \frac{1}{N_{\text{mel}}} \sum_{\nu=0}^{N_{\text{mel}}-1} H_{\log}(\nu) \quad (7.95)$$

und der festgelegte, logarithmische Grenzwert H_T geben die maximale Verstärkung bzw. Dämpfung an. Bereits bei der Bestimmung der Verstärkungsanpassung wird die Laufzeit zwischen Testsignal und Zuhörerohr ausgeglichen. Da die Schätzung der Laufzeit nur den leistungsmäßig dominanten Signalanteil berücksichtigt, wird auch bei der Verstärkungsanpassung nur dieser Teil betrachtet. Dadurch und aufgrund der Tatsache, dass die Spektren eines jeden Rahmens aufsummiert werden, wird nur ein Teil der Übertragung bzw. Impulsantwort geschätzt. In Abbildung 7.18 ist in der unteren Grafik die Gesamtimpulsantwort, Direktschall plus Übertragung über das ICC-System, in blau dargestellt. Zusätzlich ist der Anteil der Impulsantwort, welcher zur Bestimmung der Verstärkungsanpassung berücksichtigt wird, in rot markiert. Dieses Fenster entspricht der Fenstergröße N_{FFT} , welche zur Transformation der Signale in den Frequenzbereich verwendet wird. Nur dieser Anteil wird zur Schätzung des Referenzspektrums berücksichtigt, da davon ausgegangen wird, dass nur der Anteil des Signals mit der höchsten Leistung für das Qualitätsempfinden eines Passagiers entscheidend ist. Da die Verstärkungsanpassung über viele verschiedene Rauschpulse ermittelt wird, entsteht eine über die Zeit gemittelte Anpassung. Durch diese Anpassung kann die segmentale spektrale Distanz einzelne Veränderungen des Übertragungsverhaltens bestimmen und entsprechend im Gesamtergebnis berücksichtigen.

Im letzten Schritt wird nun die eigentliche spektrale Distanz mithilfe eines ausschließlich aus Sprache bestehenden Testsignals geschätzt. Auch hier wird die Signalverarbeitung, wie sie in Abbildung 7.19 gezeigt ist, verwendet, um die Leistungsdichtespektren in Abhängigkeit des Mel-Bandes zu bestimmen. Zu sehen ist, dass das geglättete Leistungsdichtespektrum $S_T(\nu, k)$ mit der ermittelten Verstärkungsanpassung $\hat{H}(\nu)$ nun multipliziert wird:

$$S_R(\nu, k) = S_T(\nu, k) \cdot \hat{H}(\nu). \quad (7.96)$$

Dieses geglättete Leistungsdichtespektrum des Referenzsignals $S_R(\nu, k)$ wird weiter verwendet, um die segmentale spektrale Distanz wie folgt zu berechnen:

$$\tilde{D}_{\text{spec}}(k) = \frac{\sum_{\nu=\nu_{\text{start}}}^{\nu_{\text{end}}} \left| 20 \cdot \log_{10} \left(\frac{S_R(\nu, k)}{S_P(\nu, k)} \right) \right| \cdot a^D(\nu, k)}{\sum_{\nu=\nu_{\text{start}}}^{\nu_{\text{end}}} a^D(\nu, k)}. \quad (7.97)$$

Da ein ICC-System nur in einem vorgegebenen Frequenzbereich operiert, kann mithilfe von ν_{start} und ν_{end} dieser Bereich zur Bestimmung der spektralen Distanz ausgewählt werden. Zusätzlich werden nur jene Mel-Bänder berücksichtigt, welche ausreichend Sprachleistung im Referenzspektrum anregen und zusätzlich das Ohrsignal mehr Leistung als das Hintergrundgeräusch besitzen. Somit ergibt sich die verwendete Gewichtungsfunktion zu:

$$a^D(\nu, k) = \begin{cases} 1, & \text{wenn } S_R(\nu, k) > T_{\text{act}} \quad \wedge \quad S_P(\nu, k) > \bar{N}(\nu) \quad \wedge \\ & S_P(\nu, k) \text{ einer Sprachpassage entspricht,} \\ 0, & \text{sonst.} \end{cases} \quad (7.98)$$

Die segmentale spektrale Distanz ist immer noch abhängig von dem betrachteten Rahmen. Um ein eindeutiges Gesamtqualitätsergebnis zu erzielen, werden die einzelnen spektralen Distanzen über alle betrachteten Rahmen K gemittelt.

$$\bar{D}_{\text{spec}} = \frac{1}{\sum_{k=0}^{K-1} a_{\text{act}}^D(k)} \sum_{k=0}^{K-1} \tilde{D}_{\text{spec}}(k) \quad (7.99)$$

Dabei werden nur die Rahmen in der Mittelung mitberücksichtigt, welche einen Beitrag zur segmentalen spektralen Distanz geleistet haben. Die aktiven Rahmen sind gegeben durch:

$$a_{\text{act}}^D(k) = \begin{cases} 1, & \text{wenn } \sum_{\nu=0}^{N_{\text{mel}}-1} a^D(\nu, k) > 0 \\ 0, & \text{sonst.} \end{cases} \quad (7.100)$$

Durch diese spezielle Gewichtung werden nur Rahmen bzw. Frequenzbänder bei der Distanzberechnung berücksichtigt, in welchen Sprache durch das Referenzsignal angeregt

worden ist. In Sprachpausen kann keine sinnvolle Aussage über die spektrale Distanz getätigt werden und somit müssen diese Bereiche ausgespart werden, um das Gesamtergebnis nicht zu verfälschen.

Qualitätsabschätzung

Zur Abschätzung der Qualität eines ICC-Systems mithilfe des spektralen Abstandes, muss zunächst die Vereinfachung auf nur einen Ein- und einen Ausgangskanal revidiert werden. Dadurch wird das Ergebnis aus Gleichung (7.99) abhängig von dem Lautsprecherkanal l_K und von den Mikrofonkanälen m_K der positionierten künstlichen Sprecher. Aus all diesen Kombinationen wird das Ergebnis der spektralen Distanz ausgewählt, welches am größten ist und dadurch dem qualitativ „schlechtesten“ Fall entspricht. Das Ergebnis der spektralen Distanz unter Berücksichtigung aller Kanäle ist gegeben durch:

$$D_{\text{spec}} = \max \left\{ \bar{D}_{\text{spec}}(l_K, m_K) \right\}. \quad (7.101)$$

Im nächsten Schritt wird ein lineares Qualitätsschätzungsmodell unter Verwendung der auditiven Ergebnisse aus Kapitel 6.4 und den Ergebnissen der zugehörigen spektralen Distanz nach dem Verfahren aus Kapitel 4.2.1 bestimmt. Für das verwendete Modell sind alle Datenpunkte verwendet worden, da kein signifikanter Unterschied zwischen den Einzelmodellen gefunden werden konnte, siehe Kapitel 4.3. Dazu ist die Verteilung der einzelnen Koeffizienten unter Berücksichtigung des 95%-Konfidenzintervalls untersucht worden, das Ergebnis ist in Anhang A dargestellt. Das verwendete lineare Modell ist gegeben durch:

$$\check{c}_{\text{SD}} = 0,22 + 0,147 \cdot D_{\text{spec}}. \quad (7.102)$$

In Abbildung 7.22 ist das zugehörige Streudiagramm aller erhobenen Datenpunkte in Abhängigkeit der normierten MOS-Werte $\widetilde{\text{MOS}}_S$ der auditiven Ergebnisse und der zugehörigen Qualitätsschätzung \check{c}_{SD} , unter Verwendung der spektralen Distanz, dargestellt. Neben der Winkelhalbierenden, welche der optimale Schätzung entspricht, ist von dieser aus ein 10%-Intervall aufgetragen. Es wird deutlich das gerade im Bereich der Qualitätsstufe 1 eine gute Schätzung erzielt werden konnte. Bei steigender Qualitätsklasse bzw. sinkender Qualität (4 entspricht einem mangelhaftem Ergebnis, siehe Tabelle 7.1), streuen die Schätzungen und liegen auch außerhalb des 10%-Intervalls.

Dennoch konnten sowohl in der Gesamtbewertung als auch bei der *Jackknife*-Methode hohe Korrelationen, siehe Tabelle 7.12, erzielt werden. Der RMSE liegt unter Betrachtung aller Datenpunkte bei circa 11% und für die *jackknife*-Methode bei circa 13,6% der Qualitätsskala.

Abschließend wird ein eindeutiges Ergebnis dieser Evaluierungsmethode berechnet, welches neben der absoluten Qualitätseinstufung auch die relative Veränderung der Qualität zwischen einem ein- und einem ausgeschalteten ICC-System berücksichtigt. Ebenso

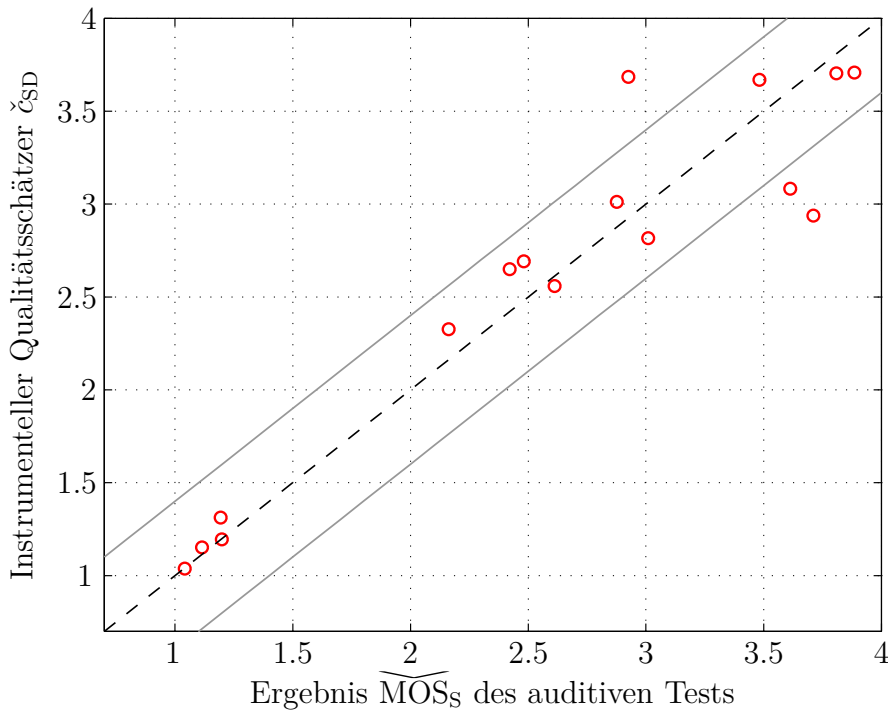


Abbildung 7.22: Streudiagramm in Abhängigkeit der normierten auditiven Qualität und des Qualitätsergebnis ermittelt mit dem spektralen Abstandsmaß.

ρ	ρ_{LOO}	RMSE	RMSE_{LOO}
0,94	0,91	0.33	0.41

Tabelle 7.12: Empirischer Korrelationskoeffizient und RMSE aller Datensätze (ρ , RMSE) und der jackknife-Methode (ρ_{LOO} , RMSE_{LOO}).

wie in den Evaluierungsverfahren zuvor, muss das ICC-System mit einem ausreichend leistungsstarkem Hintergrundgeräusch angeregt werden. Innerhalb dieses Szenario wird die spektrale Distanz mit $\check{c}_{\text{SD}}^{(\text{on})}$ und ohne aktivem ICC-System $\check{c}_{\text{SD}}^{(\text{off})}$ bestimmt. Das finale Ergebnis dieses Evaluierungsverfahrens ist gegeben durch:

$$c_{\text{SD}} = \min \left\{ \max \left\{ 1, \check{c}_{\text{SD}}^{(\text{on})} + g_{\text{R}} \cdot \left(\check{c}_{\text{SD}}^{(\text{on})} - \check{c}_{\text{SD}}^{(\text{off})} \right) \right\}, 4 \right\}. \quad (7.103)$$

Dabei entspricht g_{R} dem Systemattribut „Notwendigkeit“ aus Kapitel 7.3.1 und der Qualitätsbereich ist auf einen Bereich zwischen eins und vier begrenzt.

7.4 Gesamtqualitätsbewertung

Die in Kapitel 3.2 eingeführten Evaluierungsgruppen sind mit den entsprechenden Verfahren zur Abschätzung der Qualität eines ICC-Systems vorgestellt worden. Ein jedes dieser Verfahren liefert ein entsprechendes Qualitätsergebnisse, welches final in eine eindeutige Gesamtqualitätsbewertung münden soll. Zu diesem Zweck muss eine geeignete Kombination und Gewichtung aller Qualitätsschätzer durchgeführt werden. In Abbildung 7.23 sind zum einen die Qualitätsschätzer aller vorgestellten Evaluierungsverfahren und zum anderen die für die Gewichtung verwendeten Attribute und deren Einfluss auf andere Qualitätsschätzer dargestellt.

Der erste Schritt zur Bestimmung der Gesamtqualität ist die Berechnung der Fahrzeugattribute. Diese sind bereits in Kapitel 7.1.3 im Detail beschrieben und die Berechnung in 7.1.3 vorgestellt worden. Der Fahrzeugattributvektor \mathbf{g} wird bestimmt durch die einzelnen Ergebnisse der Gruppe „Fahrzeugeigenschaften“ unter Verwendung einer gewichteten Summe, siehe Abbildung 7.23. Zusätzlich werden die einzelnen Fahrzeugattribute auf einen Wertebereich zwischen 0,1 und 1 beschränkt. Dadurch wird eine Gewichtung der vorläufigen Qualitätsschätzer mit Null und somit ein Ausschluss des Schätzers aus der Bestimmung der Gesamtqualität verhindert.

Nach dem die Fahrzeugattribute ermittelt und die vorläufigen Qualitätsschätzer der Evaluierungsgruppe „Verhalten des ICC-Systems“ bestimmt worden sind, diese sind gegeben durch:

$$\check{\mathbf{c}}_I = [\check{c}_R, \check{c}_S, \check{c}_T]^T, \quad (7.104)$$

kann das Qualitätsergebnis dieser Evaluierungsgruppe durch die Gewichtung der Qualitätsschätzer mit den Fahrzeugattributen

$$\mathbf{w}_I(\mathbf{g}) = \left[g_N, \frac{g_S + g_T}{2}, g_S \right]^T \quad (7.105)$$

bestimmt werden. Dazu werden die Qualitätsschätzer entsprechend gewichtet, aufsummiert und normiert. Das so bestimmte Gruppenqualitätsergebnis der Evaluierungsgruppe ist gegeben durch:

$$Q_I = \frac{\check{\mathbf{c}}_I^T \cdot \mathbf{w}_I(\mathbf{g})}{\|\mathbf{w}_I(\mathbf{g})\|_1}. \quad (7.106)$$

Für eine bessere Veranschaulichung, ist die Kombination aller Ergebnis zu einer Gruppenqualität im oberen Teil von Abbildung 7.23 grafisch aufbereitet. Es wird deutlich, dass das Fahrzeugattribut „Hintergrundgeräusch“ direkt in die Bestimmung der Erhöhung des Hintergrundgeräusches einfließt. Im Detail bedeutet dies im Falle eines sehr lauten Fahrzeuges, dass $g_N = 1$ und auch das erste Element des Vektors $\mathbf{w}_I(\mathbf{g})$ gleich Eins entspricht. Dadurch wird der Fokus auf das Ergebnis der Schätzung der Hintergrundgeräuscherhöhung gelegt, da eine Unterdrückung des Hintergrundgeräusches innerhalb einer lauten

Erste Ansätze der hier vorgestellten Gesamtqualitätsbewertung sind bereits in [TSWL] vorgestellt worden.

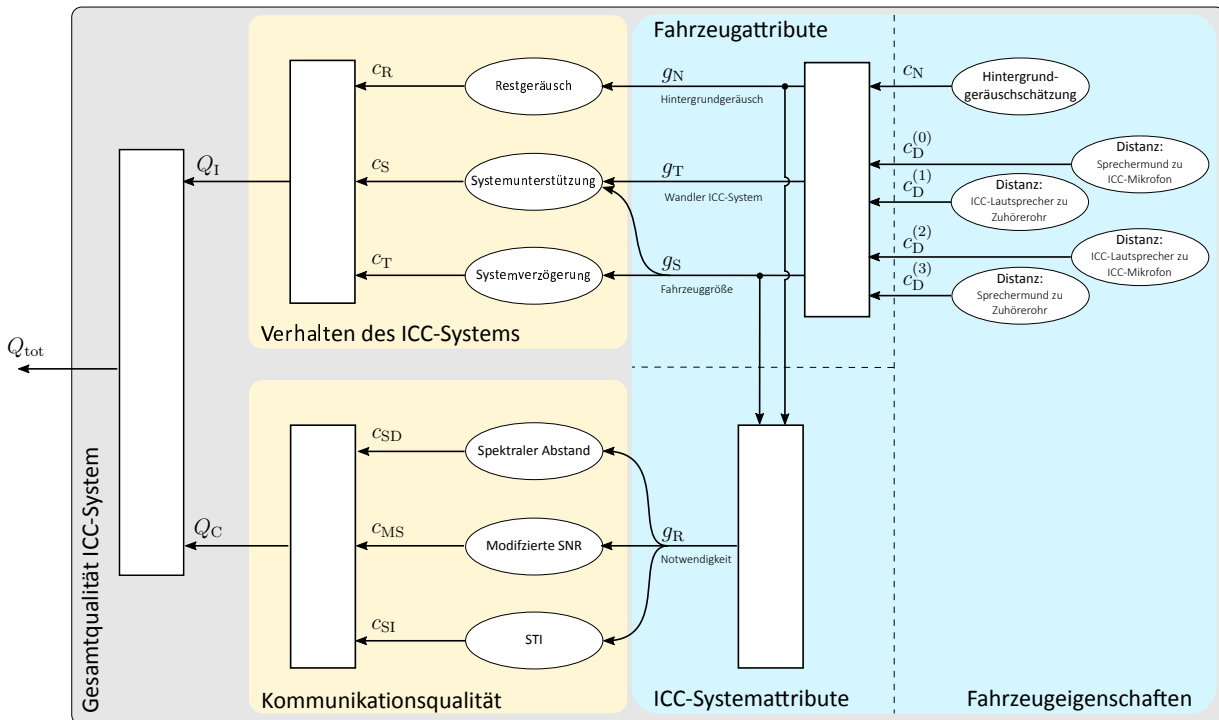


Abbildung 7.23: Bestimmung der Gesamtqualität eines ICC-Systems unter Verwendung aller zuvor beschriebenen Evaluierungsverfahren und den zugehörigen Attributen.

Umgebung sich als schwieriger darstellt als in einer leisen. In einer leisen Umgebung würde das Fahrzeugattribut $g_N = 0,1$ entsprechen und so der Einfluss des Qualitätsschätzers c_R geringer werden. Der vorläufige Qualitätsschätzer \check{c}_S der Systemunterstützung wird zu gleichen Teilen mit dem Fahrzeugattribut „Fahrzeuggröße“ und „Wandler des ICC-Systems“ gewichtet, da zum Beispiel ein Fahrzeug mit weit entfernten elektroakustischen Wandlern eine größer Verstärkung des Nutzsignals benötigt, um den SNR am Zuhörerohr zu erhöhen. Die Fahrzeuggröße gibt einen Eindruck, wie stark die Rückkopplung innerhalb des Fahrzeuges ist. In einem kleinen Fahrzeug ist die Distanz zwischen ICC-Lautsprecher und ICC-Mikrofon geringer und somit die mögliche Rückkopplung stärker. In diesem Fall muss eine Rückkopplungsunterdrückung sehr zuverlässig arbeiten, um keine Qualitätseinbußen zu erhalten. In einem großen Fahrzeug ist die Rückkopplungsproblematik als weniger kompliziert einzustufen, da lediglich ein geringer Teil des Nutzsignales wieder einkoppelt. Somit müssen die Algorithmen eines ICC-Systems weniger Leistung des Rückkopplungssignals kompensieren. Der vorläufige Qualitätsschätzer \check{c}_T der Systemverzögerung wird mit dem Fahrzeugattribut „Fahrzeuggröße“ gewichtet, denn je weiter die ursprüngliche Quelle des Schallereignisses, der Sprecher, vom Zuhörer entfernt ist, um so geringer wird die Nutzsignalleistung am Zuhörerohr. Bei einer zusätzlichen Quelle des Nutzsignales durch die Verwendung eines ICC-System, wird, durch die Leistungsunterschiede, die Verzögerung zwischen beiden weniger stark wahrgenommen. Wohingegen in

einem Szenario, in welchem beide Passagiere nahe zusammen sitzen, beispielsweise auf dem Fahrer und Beifahrersitz in einem Kabriolett, die beide Passagiere schneller durch auftretende Echos gestört werden können. In beiden Fällen kann der Richtungseindruck des zuhörenden Passagiers verändert sein.

Im nächsten Schritt wird das Gruppenqualitätsergebnis der Evaluierungsgruppe Kommunikationsqualität bestimmt. Die Einflüsse der Attribute auf die Bestimmung der gewichteten Qualitätsschätzer dieser Evaluierungsgruppe sind im unteren Teil der Abbildung 7.23 dargestellt. Zunächst wird mithilfe der Fahrzeugattribute „Hintergrundgeräusch“ und „Fahrzeuggröße“ das ICC-Systemattribut „Notwendigkeit“ durch

$$g_R = \frac{g_N + (1 - g_S)}{2} \quad (7.107)$$

bestimmt. Details über dieses Attribute können in Kapitel 7.3.1 eingesehen werden. Unter Verwendung der vorläufigen Qualitätsschätzer, zum Beispiel der STI-Bestimmung \check{c}_{SI} , und des Attributes „Notwendigkeit“ wird der gewichtete Qualitätsschätzer bestimmt durch:

$$c_{SI} = \min \left\{ \max \left\{ 1, \check{c}_{SI}^{(on)} + g_R \cdot (\check{c}_{SI}^{(on)} - \check{c}_{SI}^{(off)}) \right\}, 4 \right\}. \quad (7.108)$$

In gleicher Weise, durch Schätzung der vorläufigen Qualität mit und ohne einem aktivierten ICC-System, werden auch die gewichteten Qualitätsschätzer der anderen beiden Evaluierungsverfahren c_{MS} und c_{SD} berechnet. Durch diese Berechnung der Qualität wird nicht das absolute Qualitätsergebnis eines ICC-System in der Gesamtqualität berücksichtigt, sondern zusätzlich die Steigerung des Qualitätsschätzers zwischen dem gleichen Szenario mit und ohne ICC-System. Die Ergebnisse aller drei gewichteten Qualitätsschätzer der Evaluierungsgruppe, werden zu folgendem Vektor kombiniert:

$$\mathbf{c}_C = [c_{SD}, c_{MS}, c_{SI}]^T. \quad (7.109)$$

Zur Bestimmung des Qualitätsergebnisses dieser Evaluierungsgruppe werden die bereits gewichteten und relativ ermittelten Qualitätsschätzer gemittelt. Es folgt, dass

$$Q_C = \frac{\mathbf{c}_C^T \cdot \mathbf{w}_C}{\|\mathbf{w}_C\|_1} \quad (7.110)$$

ist, wobei der Gewichtungsvektor durch:

$$\mathbf{w}_C = [1, 1, 1]^T \quad (7.111)$$

gegeben ist.

Im letzten Schritt werden die Gruppenergebnisse gemittelt und geben die Gesamtqualität des ICC-System an:

$$Q_{\text{tot}} = \frac{Q_C + Q_I}{2}. \quad (7.112)$$

7.5 Zusammenfassung

In Kapitel 7 sind die verschiedenen instrumentellen Evaluierungsverfahren, der drei Evaluierungsgruppen aus Kapitel 3.2 beschrieben. Neben der eigentlichen Erhebung des Maßes, ist innerhalb eines jeden Evaluierungsverfahrens auch die zugehörige Qualität bestimmt worden, um so einen Qualitätsschätzer des betrachteten Verfahrens zu gewinnen. Zu diesem Zweck sind die Daten aus Kapitel 6 herangezogen worden. Zusätzlich sind die Ergebnisse der Evaluierungsgruppe „Verhalten des ICC-Systems“ mithilfe der zuvor bestimmten Fahrzeugattribute gewichtet worden, um so den Einfluss der Fahrzeugumgebung mit in der Evaluierung zu berücksichtigen. Auch die Ergebnisse der Evaluierungsgruppe „Kommunikationsqualität“ werden nicht direkt zur Bestimmung der Gesamtqualität verwendet, sondern die relative Verbesserung durch die Verwendung eines ICC-Systems mitberücksichtigt. Insgesamt sind eine Vielzahl von instrumentellen Verfahren vorgestellt worden, welche im Anschluss eine detaillierte Ursachenanalyse zulassen. Außerdem ist eine Kombination der Qualitätsschätzer im Detail beschrieben worden.

Die vorgestellten Evaluierungsverfahren innerhalb der drei betrachteten Gruppen stellen eine erste Zusammenstellung von Einzelverfahren dar. Innerhalb einer jeden Gruppe sind weitere Verfahren bzw. Qualitätsschätzer denkbar. So wäre zum Beispiel die Erhebung des Frequenzganges der einzelnen ICC-Wandler innerhalb der Gruppe „Fahrzeugeigenschaften“ eine mögliche Weiterentwicklung, um so die Qualität des Eingangsbzw. des Ausgangssignals eines ICC-Systems detaillierter zu klassifizieren. Diese Evaluierung würde einen elektrischen Eingriff in das ICC-System bedeuten und die Anforderung, die Evaluierung mithilfe eines *blackbox*-Ansatzes durchzuführen, nicht erfüllen. Zusätzlich könnten mit weiteren Messungen, zum Beispiel einer Messung der Abstrahlcharakteristik des Sprachsignals der Insassen, innerhalb der Gruppe der Fahrzeugeigenschaften weitere Fahrzeugattribute erhoben werden. Anhand dieser Messung könnte beispielsweise eine Klassifizierung der Kommunikationspfade zwischen den Insassen erfolgen. Auch die anderen beiden Gruppen könnten mit weiteren Evaluierungsverfahren erweitert werden. So wäre zum Beispiel die Evaluierung der Rückkopplungsunterdrückung innerhalb der Evaluierungsgruppe Verhalten des ICC-Systems eine mögliche Erweiterung. Die Evaluierungsgruppe „Kommunikationsqualität“ bietet die Möglichkeit, neben den klassischen Evaluierungsverfahren wie SNR, STI, etc., auch neue, innovative Evaluierungsverfahren zu integrieren. Zu diesem Zweck könnten neue und detaillierte Erkenntnisse über die Kommunikationssituation mit und auch ohne ICC-System, entscheidende Hinweise geben. Erste Schritte in Bezug auf linguistische Parameter sind zum Beispiel in [LNS⁺] getätigt worden. Auch eine erste Analyse des Lombard-Effektes in einem Fahrzeug ist in [LKKLS]

durchgeführt worden. Hier konnte festgestellt werden, dass sich bei der Verwendung eines ICC-Systems sowohl die für den Lombard-Effekt klassische Erhöhung der Grundfrequenz, als auch die Intensität des Sprachsignals reduzieren. Die durch solche Analysen erhobenen Hypothesen könnten neben der Entwicklung neuer Evaluierungsverfahren zusätzlich helfen, die bereits Umgesetzten in ihrer Qualitätsschätzung zu verbessern. So ist zum Beispiel derzeit nicht exakt geklärt, welchen Einfluss die Wiedergabe eines durch das ICC-System spektral veränderten oder gar modulierten Hintergrundgeräusches auf das Qualitätsempfinden der Passagiere hat. Derzeit wird nur die Erhöhung der Hintergrundgeräuschleistung ermittelt. Falls zum Beispiel spektrale Einflüsse qualitätsrelevant wären, sollten diese im Evaluierungsverfahren mit berücksichtigt werden.

Abschließend ist zu sagen, dass die innerhalb der Evaluierungsgruppe „Kommunikationsqualität“ entwickelten linearen Modelle, aufgrund der eingeschränkten Datenbasis nur als erster Startpunkt zur Evaluierung von ICC-Systemen angesehen werden können. Für ein detaillierteres Training der Modelle und eine exaktere Validierung, wäre eine größere Datenbasis notwendig. Dazu müssten nicht nur die Anzahl der einzelnen Szenarien, in denen ein ICC-System auditiv evaluiert werden soll, erhöht werden, sondern auch die Anzahl der Testpersonen. Zusätzlich müsste eine größere Anzahl an ICC-Systemen untersucht werden, um so eine größere Variabilität innerhalb der Qualitätsbeurteilung zu erzielen. Um eine solche Datenbasis generieren zu können, müsste zunächst ein standardisiertes Verfahren, ähnlich wie in [ITU14, ITU96a], definiert werden. Denn nur so kann garantiert werden, dass an verschiedenen Standorten von verschiedenen Gruppen kombinierbare und vergleichbare Daten erhoben werden. Eine solche Datenbasis würde die Erforschung und Neuentwicklung innovativer, akkurater Evaluierungsverfahren erleichtern und sollte deswegen Bestandteil weiterer Forschung im Bereich der ICC-Systemevaluierung sein.

Kapitel 8

Evaluierung eines konkreten ICC-Systems

Nach der Beschreibung aller Evaluierungsverfahren sowie der Kombination der Ergebnisse zu einem Gesamtergebnis, siehe Kapitel 7, wird diese Evaluierungsstrategie zur Ermittlung der Qualität eines konkreten ICC-Systems verwendet. Dazu wird zunächst die vorgestellte ICC-Systemevaluierung innerhalb des Echtzeit-Rahmenwerks KiRAT implementiert. Diese Umsetzung wird im Anschluss für die Evaluierung eines konkreten ICC-Systems verwendet. Als Evaluierungsobjekt dient dabei das am Lehrstuhl DSS entwickelte ICC-System. Genau wie die ICC-Systemevaluierung, ist auch das ICC-System innerhalb des Rahmenwerks KiRAT umgesetzt worden. Dies bietet den Vorteil, dass beide Hauptsysteme innerhalb des gleichen Rahmenwerks operieren und alle vorhandenen Funktionalitäten des Rahmenwerks KiRAT, wie zum Beispiel die akustische Systemsimulation, verwendet werden können.

Zunächst soll die Testumgebung innerhalb des Rahmenwerks mit all seinen Komponenten und den Übergängen zu den Hauptkomponenten, ICC-System und ICC-Systemevaluierung, erläutert werden. Im letzten Teil dieses Kapitels sind die Ergebnisse der Evaluierung aufgeführt und erläutert.

8.1 Beschreibung der Testumgebung

Innerhalb des Rahmenwerks KiRAT sind die beiden Hauptkomponenten, die ICC-Systemevaluierung und das zu evaluierende ICC-System, eingebettet. Neben der Hauptkomponenten können auch Hilfsmodule, wie beispielsweise eine akustische Systemsimulation verwendet werden, welche eine vollständigen Simulation der Übertragung von akustischen Signalen innerhalb des Rahmenwerks ermöglicht. Grundsätzlich funktioniert sie durch die Faltung zuvor ermittelter Impulsantworten mit den entsprechenden Ausgangssignalen der einzelnen Komponenten. Eine detaillierte Beschreibung der Umsetzung dieser Simulation ist nicht Bestandteil der vorgelegten Arbeit. Diese Faltung ist, ebenso wie das betrach-

tete ICC-System und das Evaluierungssystem, in Echtzeit umgesetzt. Dadurch können alle Übertragungspfade zwischen allen Wandlern innerhalb der Fahrgastzelle theoretisch reproduziert und das ICC-System realistisch angeregt werden. Zusätzlich können mithilfe der akustischen Systemsimulation Signale den Eingangssignalen bzw. Mikrofonsignalen additiv hinzugefügt werden. Diese Funktionalität wird verwendet, um das auftretende Hintergrundgeräusch innerhalb der Fahrgastzelle zu simulieren. Zu diesem Zweck müssen zunächst alle Mikrofonsignale während einer realen Fahrt aufgezeichnet werden. Die akustische Systemsimulation erzeugt somit ein beliebiges akustisches Szenario, beispielsweise das eines fahrenden Fahrzeuges.

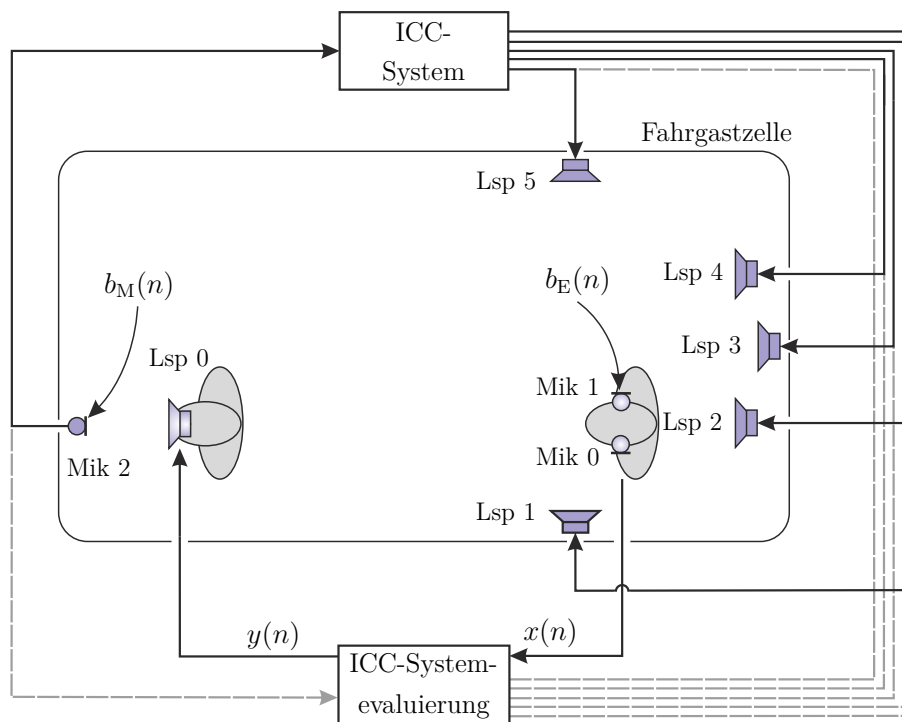


Abbildung 8.1: Wandleranordnung der ICC-Systemevaluierung und des ICC-Systems innerhalb der Fahrgastzelle. Die grau gestrichelten Linien deuten die Hilfslautsprecher bzw. das Hilfsmikrofon der ICC-Systemevaluierung an, siehe 7.1.1. Diese Hilfswandler werden in unmittelbarer Nähe der ICC-Systemwandlern platziert, um so ein black-box testing, ohne elektrischen Eingriff in das ICC-System, zu ermöglichen.

In Abbildung 8.1 ist die Fahrzeuginnenraumumgebung des zu testenden ICC-Systems dargestellt. Das ICC-System ist nur für einen Kommunikationsweg, von vorne nach hinten, parametrisiert. Deswegen stehen dem ICC-System lediglich ein Mikrofonkanal und 5 Lautsprecherkanäle zur Verfügung. Das Mikrofon ist im Himmel oberhalb des Fahrersitzes in einer Mercedes-Benz S-Klasse positioniert. Die 5 Lautsprecher befinden sich in der Hutablage (Lsp 2 bis 4) und in der linken (Lsp 1) und rechten (Lsp 5) hinteren Tür im Fußraum. Zusätzlich sind innerhalb des ICC-Systems, siehe Kapitel 2.1, nur die Basismo-

dule implementiert. Bei diesen Basismodulen handelt es sich um ein Hochpassfilter mit einer Grenzfrequenz von 100 Hz, eine Geräuschreduktion auf Basis eines Wiener-Filters, eine geräuschabhängige Verstärkungskontrolle und eine Entzerrfilterstruktur die in Bezug auf die Übertragung des Nutzsignales über den Raum entworfen worden ist.

Für die Evaluierung sind zwei Torsos in der Fahrgastzelle, einer auf dem Fahrersitz und einer auf der Rückbank hinter dem Fahrer, positioniert. Dabei simuliert der Torso auf dem Fahrersitz den Sprecher und der andere den Zuhörer. Neben diesem einen Aus- und den zwei Eingängen, hat die ICC-Systemevaluierung Zugriff auf die Lautsprecher- und Mikrofonpositionen der ICC-Systemwandler. Da kein direkter elektrischer Zugriff auf die ICC-Systemwandler aufgrund des *black-box*-Ansatzes möglich ist, wird dieser Zugriff durch zusätzliche Hilfslautsprecher und ein Hilfsmikrofon bewerkstelligt. Diese Hilfswandler werden zusätzlich, möglichst in unmittelbarer Nähe der elektroakustischen ICC-Systemwandler, platziert und der ICC-Systemevaluierung als Ein- bzw. Ausgänge bereitgestellt. Es ist zu beachten, dass diese zusätzlichen Hilfswandler ausschließlich für die Messung der Fahrzeugeigenschaften verwendet werden. Ein Nachteil dieses Vorgehens ist, dass lediglich die Positionen der ICC-Systemwandler während der Evaluierung berücksichtigt werden, jedoch nicht die Eigenschaften der ICC-Systemwandler. In Abbildung 8.1 sind diese zusätzlichen Hilfswandler nicht explizit dargestellt, sondern durch die gestrichelten Linien zwischen der ICC-Systemevaluierung und den ICC-Systemwandler angedeutet worden, um eine bessere Übersicht zu gewährleisten. Zur akustischen Simulation dieser Umgebung sind alle Impulsantworten zwischen allen Lautsprechern und allen Mikrofonen, auch den Hilfswandlern, innerhalb einer Mercedes-Benz S-Klasse gemessen worden. Dabei ist der künstliche Mund in seiner Wiedergabelautstärke so kalibriert worden, dass ein, durch den künstlichen Sprecher wiedergegebenes Sprachsignal, Signalpegel von 62 dB in einem Abstand von 50 cm erzeugt wird. Für die Kalibrierung ist eines der Sprachsignale aus der Lombard-Datenbank aus Kapitel 5.2 entnommen, welches von einem männlichen Sprecher in einer Umgebung ohne Hintergrundgeräusch generiert worden ist. Dadurch wird gewährleistet, dass dem ICC-System und den anderen Komponenten ein Sprachsignal zugeführt wird, welches in seiner Leistung dem eines realen Sprechers entspricht.

Neben den, mittels einer Faltung der Lautsprechersignale mit den entsprechenden Impulsantworten generierten, Nutzsignalen der Mikrofonkanäle, müssen diese zusätzlich mit dem entsprechenden Hintergrundgeräusch additiv belegt werden. Zu diesem Zweck sind zuvor die Hintergrundgeräusche während einer realen Fahrt mithilfe der in Abbildung 8.1 dargestellten Mikrofone aufgezeichnet worden. Diese Signale werden durch die akustische Systemsimulation auf die entsprechenden Mikrofonensignale aufsummiert und sowohl dem ICC-System als auch dem Evaluierungssystem zur Verarbeitung bereitgestellt.

8.2 Evaluierungsergebnisse

Nach erfolgreicher Implementierung des ICC-Systems innerhalb der vorgestellten Testumgebung, werden im Anschluss alle in Kapitel 7 beschriebenen Evaluierungsverfahren innerhalb dieser Testumgebung ausgewertet. Alle im Folgenden dargelegten Evaluierungsergebnisse sind mittels der Evaluierungsverfahren, welche innerhalb des Echtzeit-Rahmenwerk KiRAT umgesetzt sind, erhoben und kombiniert worden.

8.2.1 Fahrzeugeigenschaften

In einem ersten Schritt werden die Evaluierungsverfahren der „Fahrzeugeigenschaften“ berechnet und entsprechend ausgewertet. Diese können im Anschluss verwendet werden, um die Fahrzeugattribute entsprechend zu bestimmen.

Distanzen zwischen den Komponenten

Zunächst werden die Distanzen bzw. die Verzögerungen zwischen den einzelnen Wandlern und Passagieren ermittelt. Dabei werden vier unterschiedlichen Übertragungsgruppen mittels des in Kapitel 7.1.1 beschriebenen Verfahrens bestimmt und entsprechend ausgewertet. In Tabelle 8.1 sind die einzelnen Ergebnisse und die zugehörigen Gruppen der Übertragung dargestellt. Innerhalb einer jeden Gruppe wird ein Ergebnis nach der Definition aus Kapitel 7.1.1 bestimmt, dieses Gruppenergebnis ist in Tabelle 8.1 grau hinterlegt. Je nach Gruppe entspricht das Ergebnis dem maximalen oder dem minimalen Wert innerhalb aller Ergebnisse einer Gruppe.

Aus diesen Einzelergebnissen lassen sich nun die Einflussgrößen zur Bestimmung der Fahrzeugattribute berechnen. Dabei wird die Einflussgröße $c_D^{(0)}$ der Gruppe Sprechermund zu ICC-Mikrofon nach Gleichung (7.13) bestimmt und ist gegeben durch:

$$c_D^{(0)} = \min \left\{ \max \left\{ 0, \frac{1}{\delta_{\max}^{(0)} - \delta_{\min}^{(0)}} (1,47 \text{ ms} - \delta_{\min}^{(0)}) \right\}, 1 \right\} = 0,74, \quad (8.1)$$

wobei $\delta_{\min}^{(0)} = 0,8 \text{ ms}$ und $\delta_{\max}^{(0)} = 1,7 \text{ ms}$ gilt. Diese Werte spannen den Wertebereich der Einflussgröße auf und sind bereits in Kapitel 7.1.1 entsprechend für jede Gruppe eingeführt worden. Die Einflussgröße $c_D^{(1)}$ der Gruppe ICC-Lautsprecher zu Zuhörerohr wird nach Gleichung (7.15) bestimmt und ist gegeben durch:

$$c_D^{(1)} = \min \left\{ \max \left\{ 0, \frac{1}{\delta_{\max}^{(1)} - \delta_{\min}^{(1)}} (2,22 \text{ ms} - \delta_{\min}^{(1)}) \right\}, 1 \right\} = 0,73, \quad (8.2)$$

Übertragungsgruppe	Pfad	Verzögerung	
Sprechermund zu ICC-Mikrofon	Lsp 0 zu Mik 2	1,47 ms	
	Lsp 1 zu Mik 0	2,90 ms	
ICC-Lautsprecher zu Zuhörerohr	Lsp 1 zu Mik 1	3,15 ms	
	Lsp 2 zu Mik 0	2,49 ms	
	Lsp 2 zu Mik 1	2,22 ms	
	Lsp 3 zu Mik 0	2,81 ms	
	Lsp 3 zu Mik 1	3,43 ms	
	Lsp 4 zu Mik 0	3,20 ms	
	Lsp 4 zu Mik 1	2,74 ms	
	Lsp 5 zu Mik 0	4,85 ms	
	Lsp 5 zu Mik 1	4,15 ms	
	ICC-Lautsprecher zu ICC-Mikrofon	Lsp 1 zu Mik 2	3,29 ms
		Lsp 2 zu Mik 2	5,85 ms
Lsp 3 zu Mik 2		6,01 ms	
Lsp 4 zu Mik 2		6,08 ms	
Lsp 5 zu Mik 2		4,42 ms	
Sprechermund zu Zuhörerohr	Lsp 0 zu Mik 0	3,74 ms	
	Lsp 0 zu Mik 1	3,79 ms	

Tabelle 8.1: Ergebnisse des Evaluierungsverfahrens zur Erhebung der Distanz zwischen den einzelnen Komponenten nach Kapitel 7.1.1. Zusätzlich ist das finale, eindeutige Ergebnis dieses Verfahrens der jeweiligen Analysegruppe in grau hinterlegt.

wobei $\delta_{\min}^{(1)} = 0,4$ ms und $\delta_{\max}^{(1)} = 2,9$ ms gilt. Die Einflussgröße $c_D^{(2)}$ der Gruppe ICC-Lautsprecher zu ICC-Mikrofon wird nach Gleichung (7.17) bestimmt und ist gegeben durch:

$$c_D^{(2)} = \min \left\{ \max \left\{ 0, \frac{1}{\delta_{\max}^{(2)} - \delta_{\min}^{(2)}} (\delta_{\max}^{(2)} - 3,29 \text{ ms}) \right\}, 1 \right\} = 0,83, \quad (8.3)$$

wobei $\delta_{\min}^{(2)} = 1,9$ ms und $\delta_{\max}^{(2)} = 9,9$ ms gilt. Die Einflussgröße $c_D^{(3)}$ der Gruppe Sprechermund zu Zuhörerohr wird nach Gleichung (7.19) bestimmt und ist gegeben durch:

$$c_D^{(3)} = \min \left\{ \max \left\{ 0, \frac{1}{\delta_{\max}^{(3)} - \delta_{\min}^{(3)}} (3,79 \text{ ms} - \delta_{\min}^{(3)}) \right\}, 1 \right\} = 1, \quad (8.4)$$

wobei $\delta_{\min}^{(3)} = 1,8$ ms und $\delta_{\max}^{(3)} = 3,5$ ms gilt.

Hintergrundgeräusch

Dieses Evaluierungsverfahren schätzt mithilfe des in Kapitel 7.1.2 vorgestellten Evaluierungsverfahrens den Anstieg des Hintergrundgeräusches innerhalb der Fahrgastzelle. Da-

bei werden lediglich die Mikrofone an den Ohrpostionen (Mik 0 und Mik 1) berücksichtigt. Zunächst wird das mittlere Betragsspektrum des Geräusches an beiden Ohren in dBA bei Stillstand des Fahrzeuges mit laufenden Motor und im Anschluss bei einer Fahrt von circa 150 km/h bestimmt. Wie bereits zuvor in Kapitel 7.1.2 erwähnt wird ausschließlich der Frequenzbereich zwischen 0 Hz bis 11 kHz zur Bestimmung der mittleren Leistung in dBA verwendet.

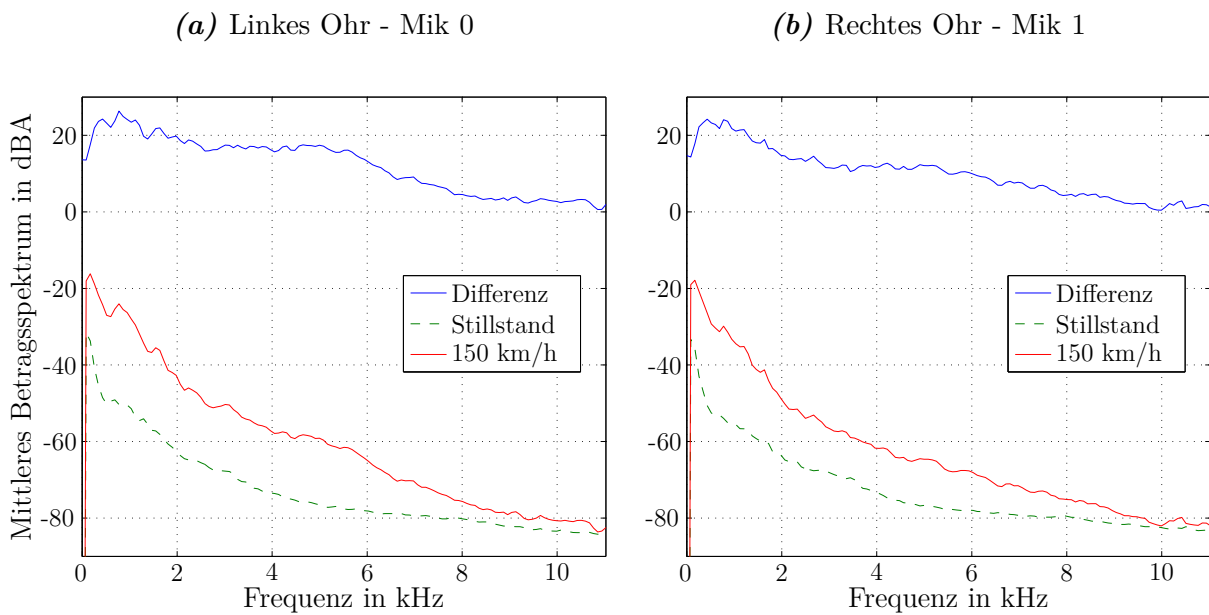


Abbildung 8.2: Mittlere Betragsspektren des Hintergrundgeräusches am Ohr des künstlichen Zuhörers in dBA bei Stillstand des Fahrzeuges mit laufendem Motor (grün) und bei einer Fahrt von circa 150 km/h (rot). Zusätzlich ist die frequenzabhängige Differenz (blau) beider Szenarien dargestellt.

In Abbildung 8.2 sind die mittleren Betragsspektren für beide Szenarien und für beide Kanäle des künstlichen Zuhörers dargestellt. Dabei entspricht die grüne Kurve dem Szenario eines, mit laufendem Motor, stehenden Fahrzeuges und die Rote einem mit einer Geschwindigkeit von 150 km/h fahrenden Fahrzeuges. Die blaue Kurve stellt die frequenzabhängige, logarithmische Differenz der beiden anderen Kurven dar. Zusätzlich ist in Tabelle 8.2 der relative Anstieg des Hintergrundgeräusches in dB gegeben und das Ergebnis dieses Evaluierungsverfahrens nach Gleichung (7.27) in grau hinterlegt.

Im Anschluss wird die Einflussgröße der Hintergrundgeräuschanstiegsschätzung c_N unter Verwendung durch Gleichung (7.28) bestimmt, sie ist gegeben durch:

$$c_N = \min \left\{ \max \left\{ 0, \frac{1}{\delta_{\max}^N - \delta_{\min}^N} \left(12,31 \text{ dB} - \delta_{\min}^N \right) \right\}, 1 \right\} = 0,21. \quad (8.5)$$

wobei $\delta_{\max}^N = 20,1 \text{ dB}$ und $\delta_{\min}^N = 10,2 \text{ dB}$ gilt.

Kanal	Mittlerer Geräuschanstieg in dB
Mik 0	12,31
Mik 1	9,94

Tabelle 8.2: Ergebnisse des Evaluierungsverfahrens zur Erhebung des Geräuschanstieges innerhalb der Fahrgastzelle nach Kapitel 7.1.2. Zusätzlich ist das finale, eindeutige Ergebnis dieses Verfahrens in grau hinterlegt.

Fahrzeugattribute

Nach der Bestimmung aller Ergebnisse aus den Messungen des Hintergrundgeräusches und der Distanzen zwischen den einzelnen Komponenten, können anschließend die einzelnen Fahrzeugattribute berechnet werden. Die Zusammenhänge der diversen Attribute zu den ermittelten Ergebnissen sind bereits in Kapitel 7.1.3 beschrieben worden. Für die hier betrachtete Fahrzeugumgebung einer Mercedes-Benz S-Klasse ergeben sich die aufgeführten Fahrzeugattribute:

- Das Fahrzeugattribut „Hintergrundgeräusch“ ist gegeben durch

$$g_N = \max \left\{ c_N b_N, 0,1 \right\} = 0,21, \quad (8.6)$$

wobei als Gewicht $b_N = 1$ verwendet worden ist, da nur eine Einflussgröße derzeit zur Bestimmung des Fahrzeugattributes verwendet wird.

- Das Fahrzeugattribut „Wandler des ICC-Systems“ ist gegeben durch

$$g_T = \max \left\{ c_D^{(0)} b_T^{(0)} + c_D^{(1)} b_T^{(1)}, 0,1 \right\} = 0,74, \quad (8.7)$$

wobei als Gewichte $b_T^{(0)} = b_T^{(1)} = 0,5$ verwendet worden sind, um so eine einfache Mittelung der betrachteten Einflussgrößen zur Bestimmung des Fahrzeugattributes durchzuführen.

- Das Fahrzeugattribut „Fahrzeuggröße“ ist gegeben durch

$$g_S = \max \left\{ c_D^{(2)} b_S^{(0)} + (1 - c_D^{(3)}) b_S^{(1)}, 0,1 \right\} = 0,42, \quad (8.8)$$

wobei als Gewichte $b_S^{(0)} = b_S^{(1)} = 0,5$ verwendet worden sind, um eine einfache Mittelung zu realisieren.

In Abbildung 8.3 sind die Ergebnisse der Fahrzeugattribute nochmals grafisch dargestellt. Diese Darstellung hilft bei der Klassifizierung des Fahrzeuges, welche nicht nur eine Relevanz bei der weiteren Evaluierung, sondern auch bei einer späteren Ursachenanalyse

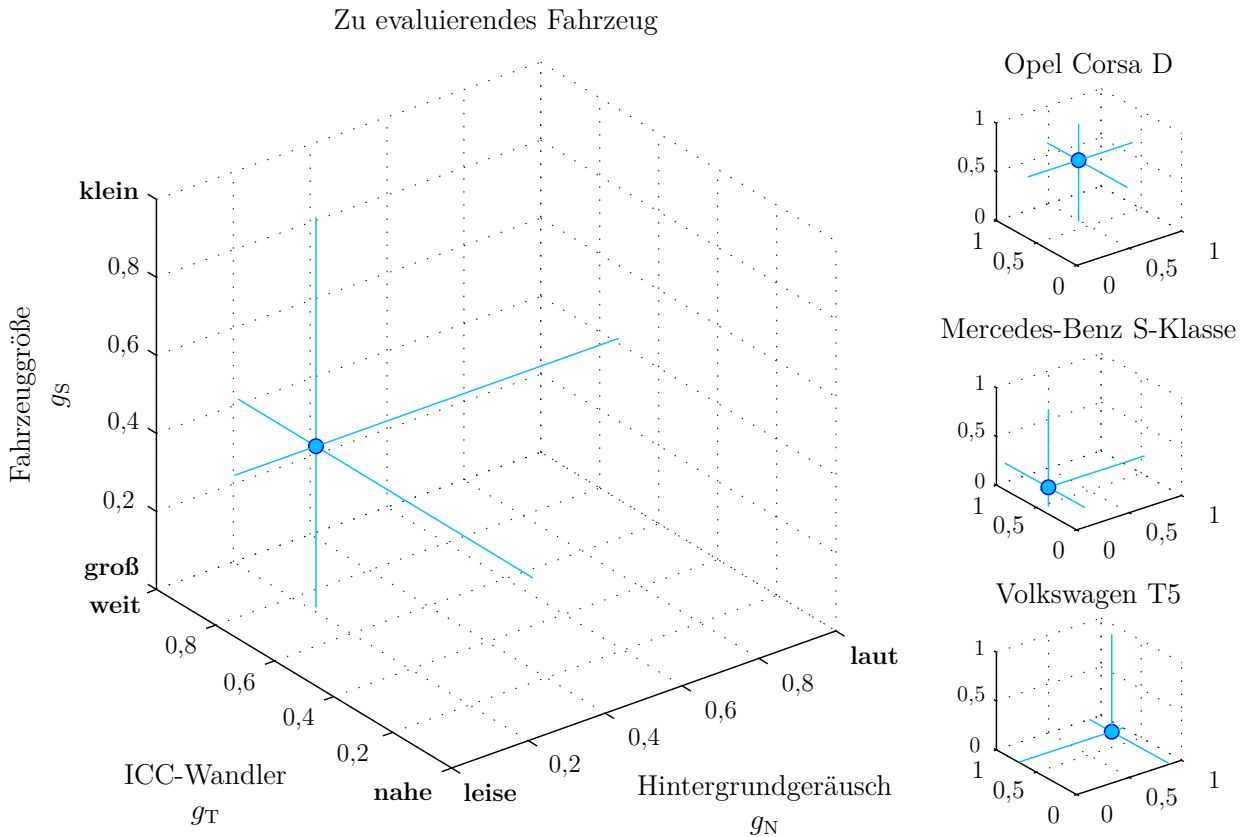


Abbildung 8.3: Fahrzeugattribute des zu untersuchenden Fahrzeuges. Zusätzlich sind drei Fahrzeuge als Vergleich mit dargestellt.

besitzt. Im vorliegenden Fall handelt es sich um ein eher leises Fahrzeug, welches nur als mittelgroß beurteilt worden ist. Da es sich bei dem untersuchten Fahrzeug jedoch um eine Mercedes-Benz S-Klasse handelt, ist dieses Ergebnis verwunderlich. Diese Diskrepanz kann auch unter Betrachtung der in Abbildung 8.3 bereitgestellten Beispielgrafik festgestellt werden. Bei genauerer Betrachtung der zuvor bestimmten Ergebnisse der Distanzen zwischen den Komponenten wird deutlich, dass der Abstand zwischen Sprecher und Zuhörer $\delta_{\min}^{(3)}$ als groß eingestuft worden ist, jedoch der Rückkopplungspfad $\delta_{\min}^{(2)}$ als eher klein. Aus diesem Grund ist das Ergebnis auch wenn ein anderes erwartet wird, plausibel. Die Wandler der betrachteten Anordnung werden als weit entfernt eingestuft. Betrachtet man das Ergebnis der Distanzmessung zwischen Sprechermund und ICC-Mikrofon, siehe Tabelle 8.1, und vergleicht dieses mit den empirisch ermittelten Distanzklassen aus Tabelle 7.2 wird deutlich, dass das Mikrofon mit einem Abstand von 50 cm nicht nahe am Sprecher positioniert ist, obwohl es sich um das Mikrofon im Himmel auf der Fahrerseite des Fahrzeuges handelt. Ein Grund hier für kann die Größe des künstlichen Sprechers sein, welche die Distanz zum Mikrofon im Himmel erhöht. Auch die Lautsprecher befinden sich

in einem eher großem Abstand, vergleicht man das Ergebnis von circa 76 cm bzw. 2,22 ms, mit den empirisch ermittelten Daten aus Tabelle 7.3.

8.2.2 Verhalten des ICC-Systems

Nach erfolgreicher Bestimmung der Fahrzeugattribute mittels der erhobenen Größen aus der Evaluierungsgruppe „Fahrzeugeigenschaften“, kann nun das Verhalten des ICC-Systems analysiert werden. Dazu werden die Evaluierungsverfahren der Gruppe „Verhalten des ICC-Systems“ eingesetzt und die Ergebnisse vorgestellt.

Systemunterstützung

Die Systemunterstützung bestimmt die durch das ICC-System eingebrachte Verstärkung des Nutzsignales am Ohr der Zuhörers. Dazu wird das in Kapitel 7.2.1 beschriebene Evaluierungsverfahren verwendet und das Leistungsverhältnis zwischen Ein- und Ausgang mit und ohne ein ICC-System bestimmt. Diese beiden frequenzabhängigen Leistungsverhältnisse des betrachteten ICC-Systems sind in Abbildung 8.4 für beide Zuhörerohren dargestellt. Zusätzlich sind die frequenzabhängigen Systemunterstützungen für beide Ohren visualisiert worden.

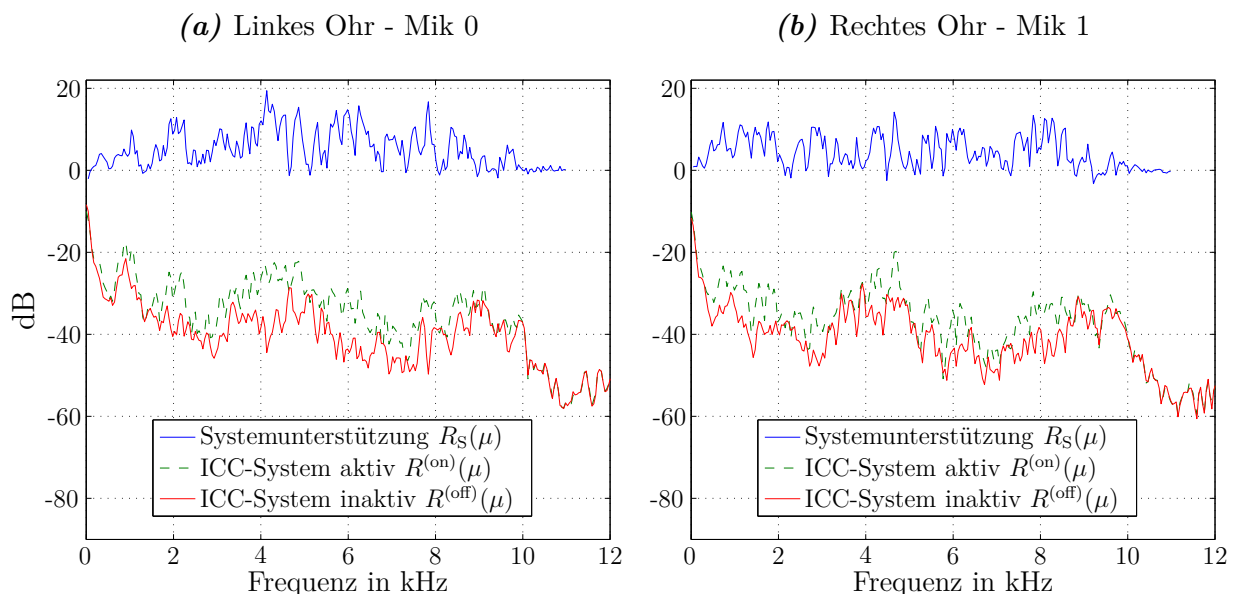


Abbildung 8.4: Die geschätzten logarithmischen Leistungsverhältnisse mit (grün gestrichelt) und ohne (rot) die Unterstützung eines ICC-Systems und die resultierende, frequenzabhängige Systemunterstützung (blau).

Zu erkennen ist, dass die beiden Leistungsverhältnisse sowohl in Abbildung 8.4(a) als auch in Abbildung 8.4(b) nur bis zu einer Frequenz von circa 10 kHz unterschiedlich sind und somit nur eine Unterstützung durch das ICC-System bis circa 10 kHz erfolgt. Gründe für diesen begrenzten Bereich der Unterstützung durch das ICC-System sind zum einen der

begrenzte Frequenzbereich der verwendeten elektroakustischen Wandler und zum andern die verwendete Abtastrate von 22,05 kHz mit welcher das ICC-System arbeitet. Da jedoch eine Verarbeitung der Signalanteile bis zu circa 11 kHz von dem betrachteten ICC-System vorausgesetzt wird, wird auch genau dieser Frequenzbereich zur Ermittlung der Systemunterstützung verwendet, die Ergebnisse sind in Tabelle 8.3 aufgeführt. Auch hier ist das Gesamtergebnis dieses Evaluierungsverfahrens nach Gleichung (7.49) in grau hinterlegt.

Kanal	Systemunterstützung in dB
Mik 0	5,23
Mik 1	4,31

Tabelle 8.3: Ergebnisse des Evaluierungsverfahrens zur Erhebung der Systemunterstützung eines ICC-Systems nach Kapitel 7.2.1. Zusätzlich ist das finale, eindeutige Ergebnis dieses Verfahrens in grau hinterlegt.

Für beide Kanäle des künstlichen Zuhörers kann nur eine mittlere Verstärkung von 4,7 dB erzielt werden, dies kann mehrere Gründe haben. Zum einen ist in diesem ICC-System keine Rückkopplungskompensation durchgeführt worden, was dazu führt das bei bereits geringen Verstärkungen ein Aufschwingen des Systems erfolgt, siehe Kapitel 2.1. Dadurch kann keine große Verstärkung durch ICC-System erfolgen, ohne ein instabiles System zu erzeugen. Zum anderen kann die große Distanz der Wandler ein Leistungsverlust des Nutzsignales darstellen, welcher nicht ausreichend bei der Parametrierung des ICC-Systems berücksichtigt worden ist.

Das Ergebnis aus Tabelle 8.2 wird im Anschluss mithilfe der Daten aus Tabelle 7.7 in eine Qualitätsklasse überführt. Die Qualitätsklasse für das vorliegende ICC-System in Bezug auf die Systemunterstützung entspricht:

$$\check{c}_S = 3. \quad (8.9)$$

Nach der Definition der Qualitätsklassen aus Tabelle 7.1 entspricht dieses Ergebnis einem System, welches die Anforderungen eines guten Systems nicht trifft, jedoch lediglich kleinere tolerierbare Mängel aufweist. In einem letzten Schritt wird dieses Qualitätsergebnis mithilfe des Zusammenhangs aus Gleichung (7.50) in das endgültige Qualitätsergebnis dieses Evaluierungsverfahrens überführt. Das Qualitätsergebnis der Systemunterstützung ist gegeben durch:

$$c_S = \check{c}_S \cdot \frac{g_T + g_S}{2} = 3 \cdot \frac{0,74 + 0,42}{2} = 1,74. \quad (8.10)$$

Systemverzögerung

Anschließend wird die Verzögerung, eingefügt durch das ICC-System, geschätzt. Hierfür wird das Evaluierungsverfahren aus Kapitel 7.2.2 verwendet, um die Systemverzögerung zu bestimmen. Dieses Verfahren schätzt die Kreuzkorrelationsfolge zwischen Sprechermund

und Zuhörerohren, mit und ohne ein aktives ICC-System. Diese beiden Folgen werden anschließend in Bezug auf ihren maximalen Ausschlag ausgewertet, um die entsprechende Verzögerung zwischen dem Direktschall und des Schallereignisses, ausgesendet vom ICC-System, zu schätzen. In Abbildung 8.5 sind die Kreuzkorrelationsfolgen mit aktiviertem ICC-System in Rot und ohne die Verwendung eines ICC-System in Schwarz gestrichelt dargestellt. Zusätzlich sind die nach dem Verfahren aus Kapitel 7.2.2 gefundenen Maxima als grüne Kreuze dargestellt. Der Zeitpunkt dieser Maxima und deren Differenz sind für beide Ohren in Tabelle 8.4 gegeben.

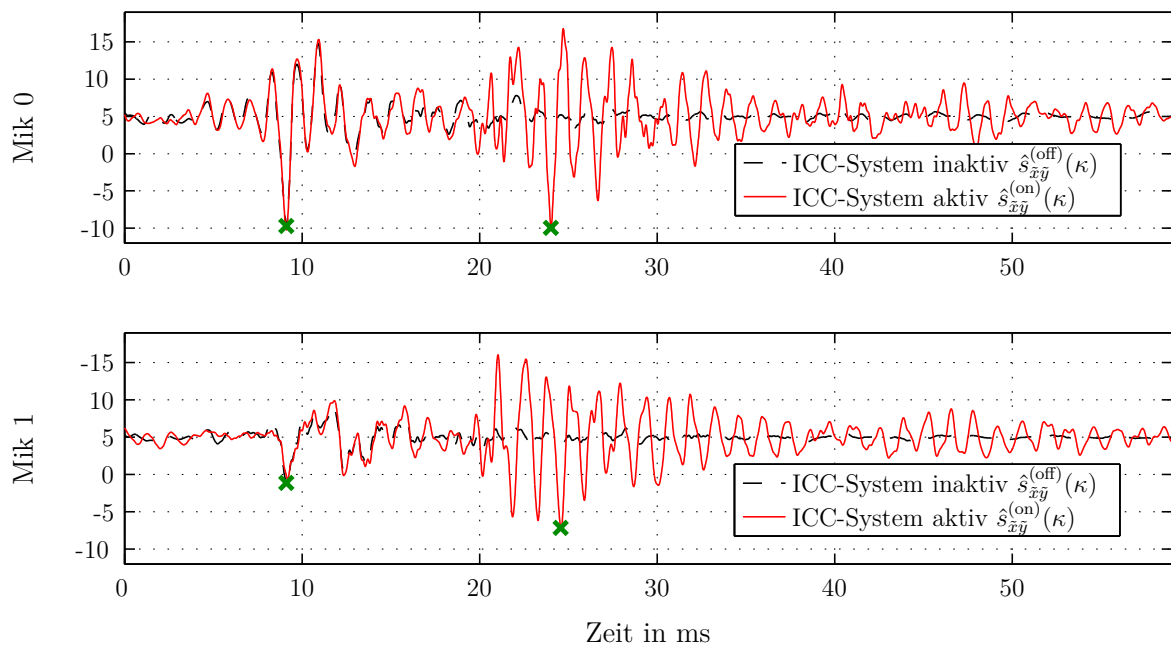


Abbildung 8.5: Die geschätzten Kreuzkorrelationsfolgen mit $\hat{s}_{\bar{x}\bar{y}}^{(\text{on})}(\kappa)$ (rot) und ohne die Unterstützung durch das ICC-System $\hat{s}_{\bar{x}\bar{y}}^{(\text{off})}(\kappa)$ (gestrichelt schwarz). Die grünen Kreuze repräsentieren die gefundenen maximalen Werte, siehe Tabelle 8.4.

Das Ergebnis der Systemverzögerungsschätzung nach Gleichung (7.55) ist in Tabelle 8.4 in grau hinterlegt. Dieses Ergebnis wird im Anschluss mithilfe der Daten aus Tabelle 7.8 in eine Qualitätsklasse überführt. Es ergibt sich für das evaluierte ICC-System in Bezug auf die Systemverzögerung die Qualitätsklasse:

$$\check{c}_T = 2. \quad (8.11)$$

Das bedeutet, dass es sich hier um ein, in Bezug auf die Systemverzögerung, gutes ICC-System handelt, welches eine geringe Verzögerung in das wiedergegebene Nutzsinal einfügt. Im Anschluss wird das Gesamtergebnis der Systemverzögerung unter Verwen-

Kanal	Verzögerung in ms		Systemverzögerung in ms
	ICC-System inaktiv	ICC-System aktiv	
Mik 0	9,12	24,01	14,91
Mik 1	9,13	24,65	15,52

Tabelle 8.4: Ergebnisse des Evaluierungsverfahrens zur Erhebung der Systemverzögerung eines ICC-Systems nach Kapitel 7.2.2. Zusätzlich ist das finale, eindeutige Ergebnis dieses Verfahrens in grau hinterlegt.

derung des Zusammenhanges von Gleichung (7.56) berechnet. Das Qualitätsergebnis der Systemverzögerung ist gegeben durch:

$$c_T = \check{c}_T \cdot g_S = 2 \cdot 0,42 = 0,84. \quad (8.12)$$

Hintergrundgeräuschanstieg

Das letzte Evaluierungsverfahren innerhalb der Evaluierungsgruppe „Verhalten des ICC-Systems“ bestimmt den Anstieg des Hintergrundgeräusches in der Fahrgastzelle bei Einschalten des ICC-Systems. Dazu wird das in Kapitel 7.2.3 beschriebene Evaluierungsverfahren verwendet, um den Geräuschanstieg zu schätzen und zu bewerten. Dabei wird das Hintergrundgeräusch während einiger Sprachpausen mit und ohne aktives ICC-System an den Ohren des Zuhörers geschätzt und ins Verhältnis gesetzt. In Abbildung 8.6 sind die geschätzten mittleren Leistungsdichtespektren der Sprachpausen mit aktivierten ICC-System in Grün gestrichelt und ohne die Unterstützung durch ein ICC-System in Rot dargestellt. Zusätzlich ist der frequenzabhängige Hintergrundgeräuschanstieg in Blau illustriert. Ähnlich wie in der Bestimmung des Systemunterstützung wird auch hier nur ein gewisser Frequenzbereich analysiert. In diesem Fall wird der Bereich von 100 Hz bis 11 kHz zur Bestimmung der Ergebnisse nach Gleichung (7.61) verwendet. Die Ergebnisse der beiden Kanäle sind in Tabelle 8.5 aufgeführt und das Gesamtergebnis nach Gleichung (7.63) in grau hinterlegt. Diesem Ergebnis wird im Anschluss mithilfe der Qualitätsstu-

Kanal	Hintergrundgeräuschanstieg in dB
Mik 0	0,35
Mik 1	0,23

Tabelle 8.5: Ergebnisse des Evaluierungsverfahrens zur Erhebung des Anstieges des Hintergrundgeräusches durch die Verwendung eines ICC-Systems nach Kapitel 7.2.3. Zusätzlich ist das finale, eindeutige Ergebnis dieses Verfahrens in grau hinterlegt.

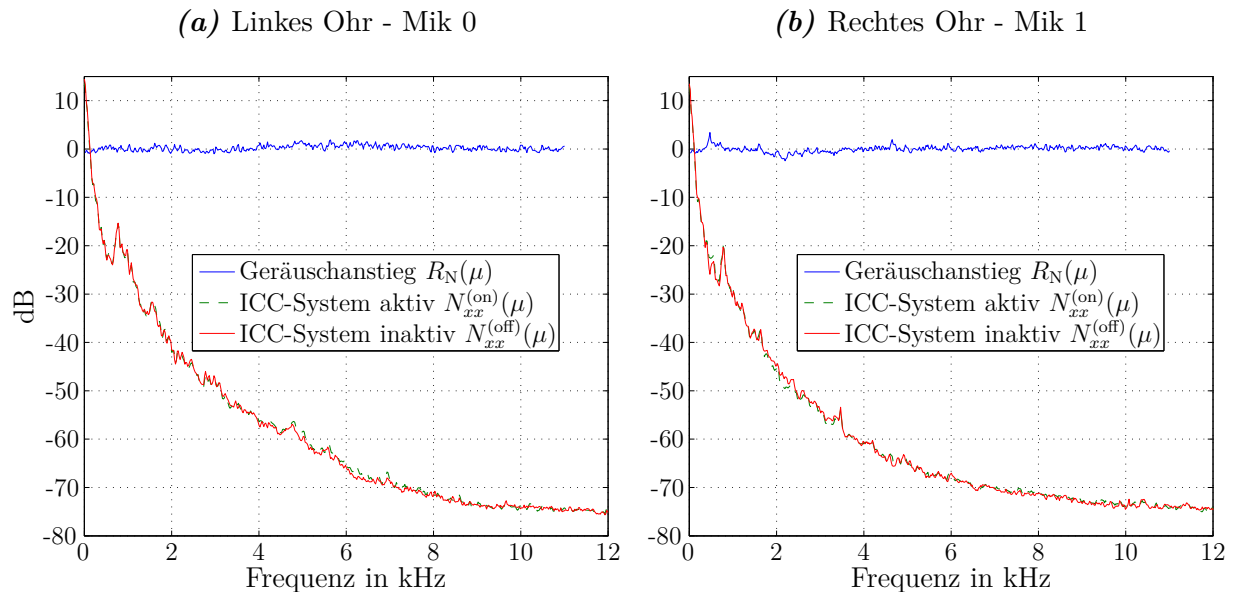


Abbildung 8.6: Die geschätzten mittleres Leistungsdichtespektrum mit (grün gestrichelt) und ohne (rot) die Unterstützung eines ICC-Systems und die resultierende, frequenzabhängige Erhöhung des Hintergrundgeräusches (blau).

fen aus Tabelle 7.9 und den zugehörigen Schwellen ein Qualitätsergebnis zugewiesen. Es ergibt sich ein vorläufiges Qualitätsergebnis von:

$$\check{c}_R = 1. \quad (8.13)$$

Es wird deutlich, dass nur in sehr geringem Maße das Hintergrundgeräusch durch das gegebene ICC-System erhöht wird. Diese geringe Erhöhung fällt durchaus in den Bereich der Messungenauigkeit des Verfahrens, sodass es sich hier, um ein überdurchschnittliches ICC-System im Sinne des Hintergrundgeräuschanstieges handelt. Abschließend wird das endgültige Qualitätsergebnis der Evaluierungsmethode anhand von Gleichung (7.64) bestimmt und ist gegeben durch:

$$c_R = \check{c}_R \cdot g_N = 1 \cdot 0,21 = 0,21. \quad (8.14)$$

8.2.3 Kommunikationsqualität

Die dritte und letzte Evaluierungsgruppe bestimmt mithilfe einzelner Evaluierungsverfahren die Kommunikations- bzw. Sprachqualität des betrachteten ICC-Systems. Um die Ergebnisse der einzelnen Verfahren wie in Kapitel 7.3 beschrieben, gewichten zu können, muss zunächst das ICC-Systemattribut „Notwendigkeit“ bestimmt werden.

ICC-Systemattribute

Das Attribut „Notwendigkeit“ wird anhand der zuvor in Kapitel 8.2.1 ermittelten Fahrzeugattribute bestimmt. Dazu wird der Zusammenhang aus Gleichung (7.65) verwendet und das Attribut „Notwendigkeit“ berechnet durch:

$$g_R = \frac{g_N + (1 - g_S)}{2} = \frac{0,21 + 0,58}{2} = 0,4. \quad (8.15)$$

Dieses Ergebnis entspricht einer Fahrzeugumgebung, welche keine starke Notwendigkeit eines ICC-System hervorruft. Hauptfaktor dieser Aussage ist der geringe Geräuschanstieg zwischen Stillstand des Fahrzeuges und einer Fahrt von 150 km/h. Dies führt zu einem geringen Wert des Fahrzeugattributes „Hintergrundgeräusch“ und somit auch zu einem eher geringen Wert des ICC-Systemattributs „Notwendigkeit“.

Modifizierter Sprachübertragungsindex

Der STI bestimmt die Übertragungsqualität durch die Beobachtung der Modulation eines Referenzsignals vor und nach der Übertragung. Eine detaillierte Beschreibung des Evaluierungsverfahrens ist in Kapitel 7.3.2 zu finden. Für die Erhebung des STI innerhalb der Fahrgastzelle werden die Übertragungen zwischen Sprechermund und den beiden Zuhörerohren untersucht. Das Ergebnis dieser Bestimmung der STI-Werte $STI(l_K, m_K)$ in Abhängigkeit der Kanäle des künstlichen Zuhörers ist in Tabelle 8.6 gegeben.

Pfad	STI		Differenz
	ICC-System inaktiv	ICC-System aktiv	
Lsp 0 zu Mik 0	0,50	0,60	0,10
Lsp 0 zu Mik 1	0,52	0,64	0,12

Tabelle 8.6: STI-Werte der untersuchten Pfade innerhalb der Fahrgastzelle zur Erhebung der Qualität eines ICC-Systems nach Kapitel 7.3.2. Zusätzlich ist das finale, eindeutige Ergebnis dieses Verfahrens in grau hinterlegt.

Mittels dieser Werte soll anschließend ein eindeutiges Qualitätsergebnis dieser Evaluierungsmethode erzielt werden. Zu diesem Zweck wird zunächst das Minimum der STI-Werte mit und ohne aktives ICC-System bestimmt. Diese Werte sind in Tabelle 8.6 entsprechend grau hinterlegt worden. Anschließend wird das lineare Modell aus Gleichung (7.74) zur Projektion auf die entsprechende Qualitätsskala verwendet. Das Ergebnis dieser Projektion ist in Tabelle 8.7 dargestellt. Beide Ergebnistabellen zeigen, dass das ICC-System eine Steigerung des STI hervorruft. Abschließend wird unter Verwendung der Ergebnisse aus Tabelle 8.7 und dem ICC-Systemattribute „Notwendigkeit“ g_R , welches zuvor be-

	ICC-System inaktiv	ICC-System aktiv
Qualitätsergebnis	$\check{c}_{\text{SI}}^{(\text{off})} = 2,21$	$\check{c}_{\text{SI}}^{(\text{on})} = 1,54$

Tabelle 8.7: Qualitätsergebnis der STI-Werte aus Tabelle 8.6, welche mithilfe des linearen Modells aus Gleichung (7.74) auf die auditive Qualitätsskala projiziert worden sind.

stimmt worden ist, das Gesamtergebnis der STI-Ermittlung berechnet. Nach Gleichung (7.75) folgt, dass das finale Ergebnis des Evaluierungsverfahrens des STI gegeben ist durch:

$$c_{\text{SI}} = \min \left\{ \max \left\{ 1, \check{c}_{\text{SI}}^{(\text{on})} + g_{\text{R}} \cdot \left(\check{c}_{\text{SI}}^{(\text{on})} - \check{c}_{\text{SI}}^{(\text{off})} \right) \right\}, 4 \right\} \quad (8.16)$$

$$= \min \left\{ \max \left\{ 1, 1,27 \right\}, 4 \right\} \quad (8.17)$$

$$= 1,27 . \quad (8.18)$$

Es wird deutlich, dass durch die Verwendung des ICC-Systems eine Verbesserung der Qualität bzw. Erhöhung des STI erzielt werden konnte, was sich im endgültigen Ergebnis der Evaluierung widerspiegelt.

Modifizierter Signal-Rauschabstand

Der modifizierte SNR bestimmt die Qualität des ICC-Systems mit dem in Kapitel 7.3.3 vorgestellten Evaluierungsverfahren. Dazu müssen zunächst die SNR-Werte für beide Ohrkanäle bei aktiven und inaktiven ICC-System ermittelt werden. Diese Ergebnisse sind in Tabelle 8.8 zusammen gefasst worden und zusätzlich ist das Minimum beider Kanäle für den aktiven und den inaktiven Fall in grau hinterlegt worden. Auch bei der Bestim-

Pfad	SNR in dB		Differenz
	ICC-System inaktiv	ICC-System aktiv	
Mik 0	-5,27	-3,21	2,06
Mik 1	-4,63	-2,72	1,91

Tabelle 8.8: SNR-Werte in dB der untersuchten Pfade innerhalb der Fahrgastzelle, für die Erhebung der Qualität eines ICC-Systems nach Kapitel 7.3.3. Zusätzlich ist das finale, eindeutige Ergebnis dieses Verfahrens in grau hinterlegt.

mung des SNR schneidet das linke Ohr schlechter als das rechte ab. Dies war bereits in der Bestimmung der Qualität mithilfe des STI zu beobachten. Der Grund hierfür ist möglicherweise der stärkere Anstieg des Hintergrundgeräusches zwischen Stillstand des Fahrzeuges und einer Fahrt mit 150 km/h. In Tabelle 8.2 sind die Geräuschleistungen beider Ohren quantitativ festgehalten und es kann eine Differenz von 2,4 dB zwischen linkem

und rechtem Ohr gefolgert werden. Zwar ist die Unterstützung durch das ICC-System am linken Ohr 0,9 dB (siehe Tabelle 8.3) höher als am rechten, dies reicht offenbar nicht aus, um das lautere Hintergrundgeräusch zu kompensieren.

Im Anschluss werden die gefundenen SNR-Werte auf die vorgegebene, auditive Qualitätsskala unter Verwendung des in Gleichung (7.83) definierten linearen Modells, projiziert. Das Ergebnis dieser Überführung in den vorgegebenen Qualitätsraum ist in Tabelle 8.9 dargestellt. Hier wird nun deutlich, dass der von den Trainingsdaten aufgespannte

	ICC-System inaktiv	ICC-System aktiv
Qualitätsergebnis	$\check{c}_{\text{MS}}^{(\text{off})} = 1,37$	$\check{c}_{\text{MS}}^{(\text{on})} = 0,07$

Tabelle 8.9: Qualitätsergebnis der SNR-Werte aus Tabelle 8.8, welche mithilfe des linearen Modells aus Gleichung (7.83) auf die auditive Qualitätsskala projiziert worden sind.

Qualitätsraum für die Bestimmung eines linearen Modells nicht ausreichend ist. Dadurch können Werte ermittelt werden, welche außerhalb des eigentlichen Qualitätsraums liegen. Dennoch ist eine Verbesserung der der Qualität durch die Verwendung eines ICC-Systems auch mit dem hier vorliegenden Evaluierungsverfahrens festzustellen.

Abschließend wird ein eindeutiges Ergebnis des Evaluierungsverfahrens mithilfe des Zusammenhanges aus Gleichung (7.84) bestimmt. Die Qualität eines ICC-Systems in Bezug auf den modifizierten SNR ist gegeben durch:

$$c_{\text{MS}} = \min \left\{ \max \left\{ 1, \check{c}_{\text{MS}}^{(\text{on})} + g_{\text{R}} \cdot \left(\check{c}_{\text{MS}}^{(\text{on})} - \check{c}_{\text{MS}}^{(\text{off})} \right) \right\}, 4 \right\} \quad (8.19)$$

$$= \min \left\{ \max \left\{ 1, -0,48 \right\}, 4 \right\} \quad (8.20)$$

$$= 1. \quad (8.21)$$

Spektrale Distanz

Das letzte Verfahren zur Erhebung der Qualität eines ICC-Systems innerhalb der Evaluierungsgruppe „Kommunikationsqualität“ bestimmt den spektralen Abstand zwischen dem durch die Ohrmikrofone aufgezeichneten Signalen zu dem entsprechenden Referenzsignal. Das genaue Vorgehen zur Erzeugung der Referenzsignale und die entsprechende Bestimmung des spektralen Abstandes ist in Kapitel 7.3.4 genau erläutert worden. Die Ergebnisse beider Ohrkanäle des künstlichen Zuhörers, mit und ohne ein aktives ICC-System, sind in Tabelle 8.10 aufgeführt. Da auch hier der, in Bezug auf das Evaluierungsverfahren, schlechteste Fall zur weiteren Berechnung der Qualität betrachtet wird, ist der maximale spektrale Abstand für jedes Szenario ausgewählt und in Tabelle 8.10 in grau hinterlegt. Deutlich wird, dass erneut das linke Ohr als schlechtester Fall ermittelt worden ist. Grün-

Pfad	Spektraler Abstand in dB		Differenz
	ICC-System inaktiv	ICC-System aktiv	
Mik 0	14,28	11,75	-2,53
Mik 1	13,57	10,81	-2,76

Tabelle 8.10: Spektraler Abstand in dB der untersuchten Pfade innerhalb der Fahrgastzelle zur Erhebung der Qualität eines ICC-Systems nach Kapitel 7.3.4. Zusätzlich ist das finale, eindeutige Ergebnis dieses Verfahrens in grau hinterlegt.

de hierfür sind bereits in den Abschnitten zuvor geschildert worden. Im Anschluss werden diese Ergebnisse mithilfe des linearen Modells aus Gleichung 7.102 in die vorgesehene Qualitätsskala projiziert. Die Ergebnisse für das entsprechende Szenario sind in Tabelle 8.11 aufgelistet. Eine Erhöhung der Qualität in Sinne der spektralen Distanz konnte durch die

	ICC-System inaktiv	ICC-System aktiv
Qualitätsergebnis	$\check{c}_{SD}^{(off)} = 2,32$	$\check{c}_{SD}^{(on)} = 1,95$

Tabelle 8.11: Qualitätsergebnis der spektralen Distanzwerte aus Tabelle 8.10, welche mithilfe des linearen Modells aus Gleichung (7.102) auf die auditive Qualitätsskala projiziert worden sind.

Verwendung eines ICC-Systems in dieser speziellen Fahrzeugumgebung erzielt werden.

Abschließend wird das endgültige Ergebnis des vorgestellten Evaluierungsverfahrens mittels des Zusammenhangs aus Gleichung 7.103 berechnet. Das Evaluierungsergebnis der spektralen Distanz für das vorliegende ICC-System ist gegeben durch:

$$c_{SD} = \min \left\{ \max \left\{ 1, \check{c}_{SD}^{(on)} + g_R \cdot (\check{c}_{SD}^{(on)} - \check{c}_{SD}^{(off)}) \right\}, 4 \right\} \quad (8.22)$$

$$= \min \left\{ \max \left\{ 1, 1,8 \right\}, 4 \right\} \quad (8.23)$$

$$= 1,8 \quad (8.24)$$

Gesamtqualitätsbewertung

Im Anschluss an die Erhebung der Ergebnisse der Einzelverfahren, kann die Gesamtqualität eines ICC-System bestimmt werden. Hierzu werden die Ergebnisse nach dem in Kapitel 7.4 vorgestellten Vorgehensweise gewichtet und kombiniert. Zunächst wird das Ergebnis der Evaluierungsgruppe „Verhalten des ICC-Systems“ nach dem in Gleichung

7.106 beschriebenen Zusammenhangs berechnet. Das Ergebnis dieser Evaluierungsgruppe ist gegeben durch:

$$Q_I = \frac{\check{\mathbf{c}}_I^T \cdot \mathbf{w}_I(\mathbf{g})}{\|\mathbf{w}_I(\mathbf{g})\|_1} \quad (8.25)$$

$$= \frac{c_T + c_R + c_S}{g_N + \frac{g_S + g_T}{2} + g_S} \quad (8.26)$$

$$= 2,31. \quad (8.27)$$

Dabei entspricht c_T dem Qualitätsergebnis der Bestimmung der Systemverzögerung, c_R dem Ergebnis der Evaluierung des Hintergrundgeräuschanstieges und c_S dem Qualitätsergebnis der Unterstützung durch das ICC-System. Nach Gleichung (7.105) sind die entsprechenden Gewichte durch folgenden Zusammenhang definiert:

$$\mathbf{w}_I(\mathbf{g}) = \left[g_N, \frac{g_S + g_T}{2}, g_S \right]^T. \quad (8.28)$$

Die vorläufigen Qualitätsschätzer der Einzelverfahren nach Gleichung (7.104) sind in dem folgenden Vektor,

$$\check{\mathbf{c}}_I = [\check{c}_R, \check{c}_S, \check{c}_T]^T, \quad (8.29)$$

zusammengefasst. Das Ergebnis der Evaluierungsgruppe „Kommunikationsqualität“ ist nach Gleichung (7.110) gegeben durch:

$$Q_C = \frac{\mathbf{c}_C^T \cdot \mathbf{w}_C}{\|\mathbf{w}_C\|_1} \quad (8.30)$$

$$= \frac{c_{SI} + c_{MS} + c_{SD}}{3} \quad (8.31)$$

$$= 1,36. \quad (8.32)$$

Dabei entspricht c_{SD} dem Qualitätsergebnis des Verfahrens zur Bestimmung der spektralen Distanz, c_{MS} dem modifizierten SNR und c_{SI} dem Qualitätsergebnis der STI-Bestimmung. Die einzelnen Gewichte des Vektors \mathbf{w}_C sind alle gleich eins gewählt worden, sodass eine einfache Mittelung der Einzelergebnisse erfolgt. Es ist zu beachten, dass die Einzelergebnisse bereits bei der Berechnung mit dem ICC-Systemattribut „Notwendigkeit“ gewichtet worden sind.

In einem letzten Schritt wird die Gesamtqualität des betrachteten ICC-Systems, unter Verwendung des Zusammenhanges aus Gleichung 7.112, berechnet. Die Mittelung der Ergebnisse der Evaluierungsgruppen „Kommunikationsqualität“ und „Verhalten des ICC-Systems“ ergibt sich die Gesamtqualität

$$Q_{\text{tot}} = \frac{Q_C + Q_I}{2} = 1,85. \quad (8.33)$$

Alle Ergebnisse und Attribute sind aus Gründen der Anschaulichkeit in Abbildung 8.7 zusätzlich grafisch aufbereitet worden. Diese Darstellung ist in Anlehnung an die grafische Aufbereitung, der Evaluierungsergebnissen von Freisprecheinrichtungen nach [ITU05, Qua08], entworfen worden. Dabei repräsentiert der Kreis in der Mitte die Qualität des ICC-Systems im Ganzen. Zum einen gibt die Füllung eines jeden Kreises Auskunft über die Gesamtqualität und zum anderen die gewählte Farbe. Die Farben verändern sich von Rot, was einem mangelhaften Ergebnis entspricht, zu Orange und über Gelb zu Grün, was einem ausgezeichneten Qualitätsergebnis entspricht. Des Weiteren gilt für die oberen drei Kreise je ausgefüllter sie sind bzw. je grüner desto besser hat das zu evaluierende ICC-System qualitativ abgeschnitten. Dabei entspricht der linke Kreis den Ergebnissen der Evaluierungsgruppe Verhalten des ICC-Systems und der rechte der Gruppe „Kommunikationsqualität“. Die unteren beiden, in Grau hinterlegten Kreise, stellen lediglich die Ergebnisse der entsprechenden Attribute dar, keine Qualitätsergebnisse. Dennoch ist hier eine ähnliche Farbkodierung vorgenommen: Je grüner die Kreise ausgefüllt sind, desto einfacher ist die Fahrzeugumgebung für die Verarbeitung eines ICC-Systems.

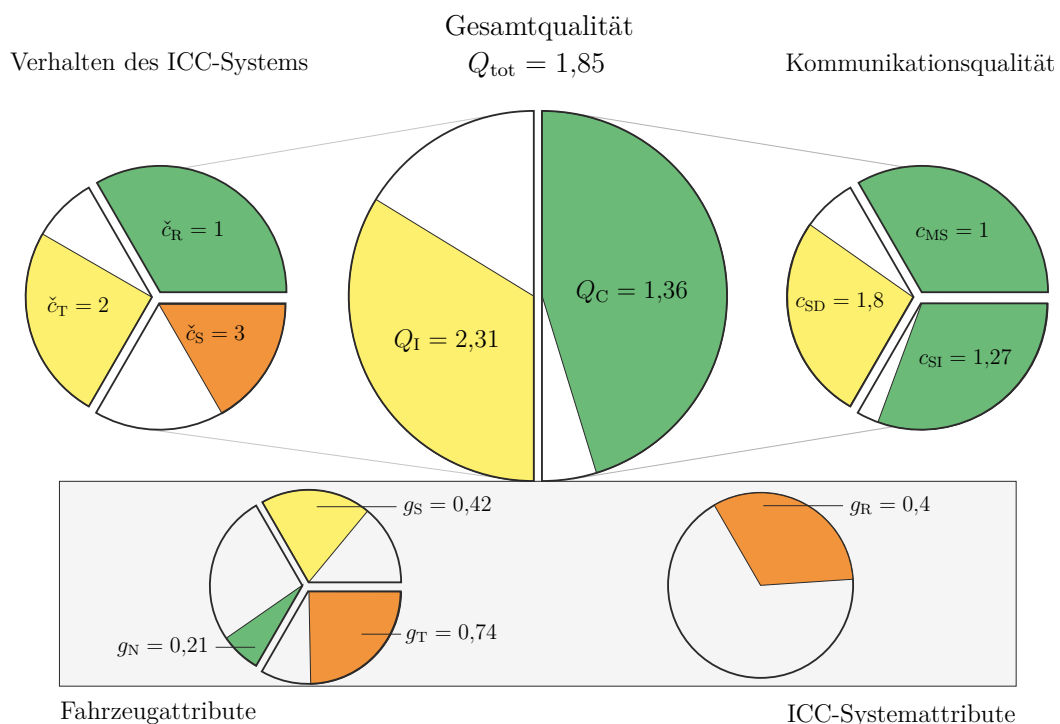


Abbildung 8.7: Grafische Auswertung der Gesamtqualität des betrachteten ICC-Systems und dem zugehörigen Fahrzeug. Neben den Einzelergebnissen sind die verwendeten Attribute des Fahrzeuges und des ICC-Systems mit aufgeführt.

Das zu evaluierende ICC-System schneidet mit $Q_{\text{tot}} = 1,85$ ab, dies entspricht einer guten Umsetzung eines ICC-Systems. Die Definition der Qualität kann Tabelle 7.1 entnommen werden. Eine eins entspricht in dieser Definition einer außergewöhnlich guten

Umsetzung und eine vier einem mangelhaften System. Dieses gute Evaluierungsergebnis des betrachteten ICC-Systems, entsteht aufgrund der vornehmlich ausgezeichneten Ergebnisse der Evaluierungsgruppe „Kommunikationsqualität“. Die Einzelverfahren dieser Gruppe schneiden mit einem ausgezeichneten bis guten Ergebnis ab, da in jedem der drei Evaluierungsverfahren eine Qualitätsverbesserung durch die Verwendung des ICC-Systems erzielt werden konnte. Diese Verbesserung aller drei Verfahren wird mit dem ICC-Systemattribut „Notwendigkeit“ gewichtet, welches zugegeben lediglich einem Wert von 0,32 entspricht, jedoch dazu führt, dass das bereits gute Qualitätsergebnis ohne ein ICC-System weiter aufgewertet wird.

Einschränkungen in der Qualität entstehen vornehmlich durch die Evaluierungsverfahren der Gruppe „Verhalten des ICC-Systems“. So ist zum einen die Systemverzögerung und zum anderen die Systemunterstützung für eine Verschlechterung der Gesamtqualität verantwortlich, wobei diese Qualitätsergebnisse lediglich gewichtet in das Gesamtergebnis eingehen. Sie werden mit den Fahrzeugattributen, welche im unteren Teil der Abbildung 8.7 zusehen sind, gewichtet und besitzen somit nur eingeschränkt Einfluss auf die Gesamtqualität. Insgesamt kann festgehalten werden, dass das verwendete ICC-System eine Verbesserung der Kommunikationsqualität erzielen konnte, obwohl die gegebenen Kommunikationssituation innerhalb einer Mercedes-Benz S-Klasse ohne die Verwendung eines ICC-Systems als ausreichend gut eingestuft werden konnte. Diese Verbesserung konnte mit einer geringen Unterstützung durch das betrachtete ICC-System erzielt werden, wie im Ergebnis der Systemunterstützung c_S deutlich zu erkennen ist. Das ICC-System scheint gut an die gegebene Situation angepasst zu sein und kann so, in einer als weniger komplex eingestuften Fahrzeugsituation, dennoch eine Verbesserung der Kommunikationssituation für die Insassen erzielen.

8.3 Evaluierungsergebnisse bei variabler ICC-Systemverzögerung

Die bereits im Abschnitt zuvor beschriebene Simulationsumgebung der ICC-Evaluierung innerhalb von KiRAT ermöglicht den Eingriff in die Funktionsweise des zu evaluierenden ICC-Systems. Es handelt sich um ein kooperatives System, welches zur Analyse der Evaluierungsstrategie in seiner Verzögerung angepasst werden soll. Dazu werden die Lautsprecher-signale vor der Ausgabe um Δ_{ICC} verzögert. Im Anschluss werden alle Daten, wie sie in Abbildung 8.7 dargestellt sind, neu berechnet. Da ausschließlich das Verhalten des ICC-Systems verändert wird, bleiben alle Fahrzeugattribute und auch das ICC-Systemattribut unverändert und werden nicht zusätzlich aufgeführt.

In Abbildung 8.8 sind die einzelnen Detailergebnisse der Evaluierungsgruppe „Verhalten des ICC-Systems“ und „Kommunikationsqualität“ in Abhängigkeit der zusätzlichen Verzögerung dargestellt. Abbildung 8.8(a) stellt die einzelnen Qualitätsergebnisse (\check{c}_T , \check{c}_S und \check{c}_R) der Evaluierungsgruppe „Verhalten des ICC-Systems“ dar. Die Daten bei keiner

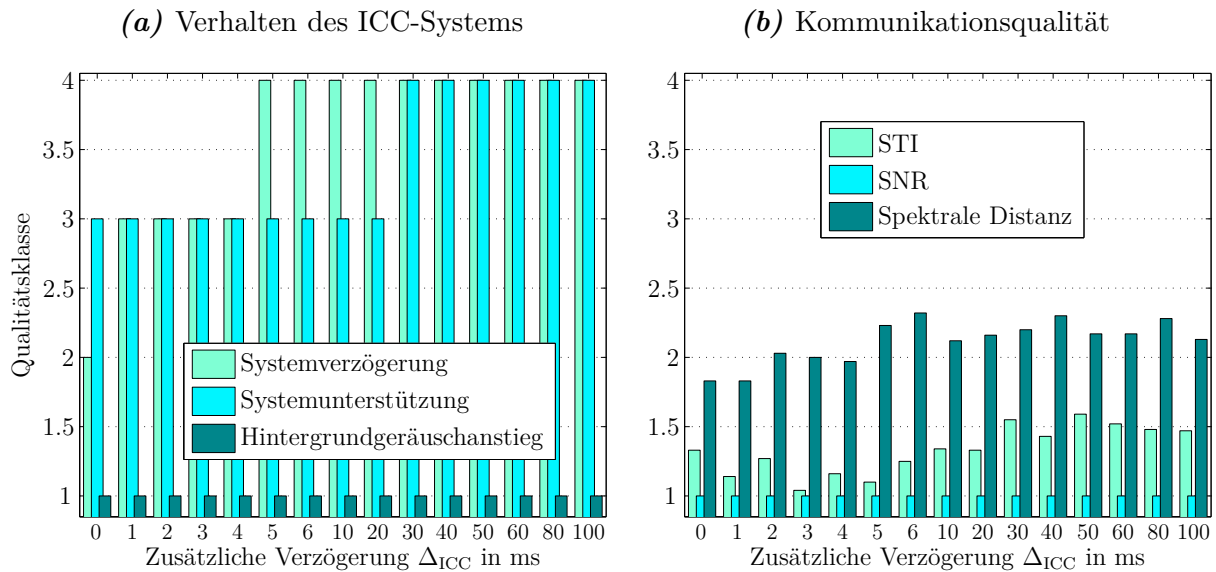


Abbildung 8.8: Detaillierte Qualitätsergebnisse der Evaluierungsgruppe „Verhalten des ICC-Systems“ und „Kommunikationsqualität“ in Abhängigkeit der zusätzlichen Verzögerung Δ_{ICC} in ms. Auf der Ordinatenachse sind die Qualitätsklassen aufgetragen, wobei 1 einem exzellenten und 4 einem mangelhaften Ergebnis entspricht, siehe auch Tabelle 7.1.

zusätzlichen Verzögerung $\Delta_{ICC} = 0$ ms entsprechen den bereits vorgestellten Ergebnissen aus Abbildung 8.7. In diesem speziellen Fall ist eine ICC-Systemverzögerung von 15,52 ms ermittelt worden, welche der 2. Qualitätsklasse entspricht. Die Schwelle zur 3. Klasse liegt bei 16 ms, sodass nun verständlich wird, warum bei einer zusätzlichen Verzögerung von einer Millisekunde bereits die 3. Qualitätsklasse erreicht wird. Des Weiteren wird deutlich, dass sich die Qualität im Bezug auf die ICC-Systemverzögerung mit ansteigender Verzögerung Δ_{ICC} verschlechtert. Das Qualitätsergebnis der Systemunterstützung bleibt bis zu einer zusätzlichen Verzögerung von 20 ms konstant bei Klasse drei. Erst bei einer Verzögerung von $\Delta_{ICC} = 30$ ms wird das zu evaluierende System in die 4. Qualitätsklasse eingestuft, da hier die Unterstützung des ICC-Systems zu einem zu späten und damit unvorteilhaften Zeitpunkt erfolgt. Lediglich die Ergebnisse der Messung des Hintergrundgeräuschanstieges bleiben über alle gewählten Verzögerungen konstant, da weiterhin kein zusätzliches Geräusch durch das ICC-System selbst eingefügt wird.

In Abbildung 8.8(b) sind entsprechend die Ergebnisse der Evaluierungsgruppe „Kommunikationsqualität“ dargestellt. Die Qualität im Sinne des modifizierten SNR wird nicht von der Erhöhung der ICC-Systemverzögerung beeinflusst. Dieser Wert bleibt über die ansteigende zusätzliche Verzögerung Δ_{ICC} quasi konstant. Der modifizierte STI ist in den geringen Verzögerung, bis 20 ms, nahezu unbeeinflusst von der Veränderung der Verzögerung. Dies ist zunächst verwunderlich, da der STI auch den Nachhall bei der Qualitätsbestimmung mitberücksichtigt. Eine Erhöhung des Nachhalls aufgrund der Veränderung

der ICC-Systemverzögerung von 20 ms, was einer Gesamtverzögerung von circa 35 ms entspricht, fällt bei den betrachteten Modulationsfrequenz kaum ins Gewicht. Folglich wird keine großen Veränderungen der Qualitätseinstufung des zu evaluierenden ICC-Systems mithilfe des modifizierten STI ermittelt. Jedoch kann ab einer Verzögerung von 30 ms ein leichter Anstieg der Qualitätsklasse verzeichnet werden, welcher wohl aufgrund der Nachhallerhöhung entsteht. Zusätzlich kann bei der Qualitätsbestimmung unter Verwendung der spektralen Distanzmethode eine leichte Verschlechterung der Qualität festgestellt werden. Dies liegt an der Erhöhung des Nachhalls innerhalb der Fahrgastzelle, da nur ein gewisser Bereich der dominanten Signalquelle (siehe Abbildung 7.18) zur Erhebung der spektralen Distanz verwendet wird. Durch den Nachhall entstehen weitere Signalkomponenten, welche die Distanz ansteigen lassen und so mit einer schlechteren Einstufung der Qualität abschneiden.

In Abbildung 8.9 sind die Gesamtqualität und die Ergebnisse der Evaluierungsgruppen in Abhängigkeit von der zusätzlich Verzögerung Δ_{ICC} dargestellt. Da innerhalb jeder

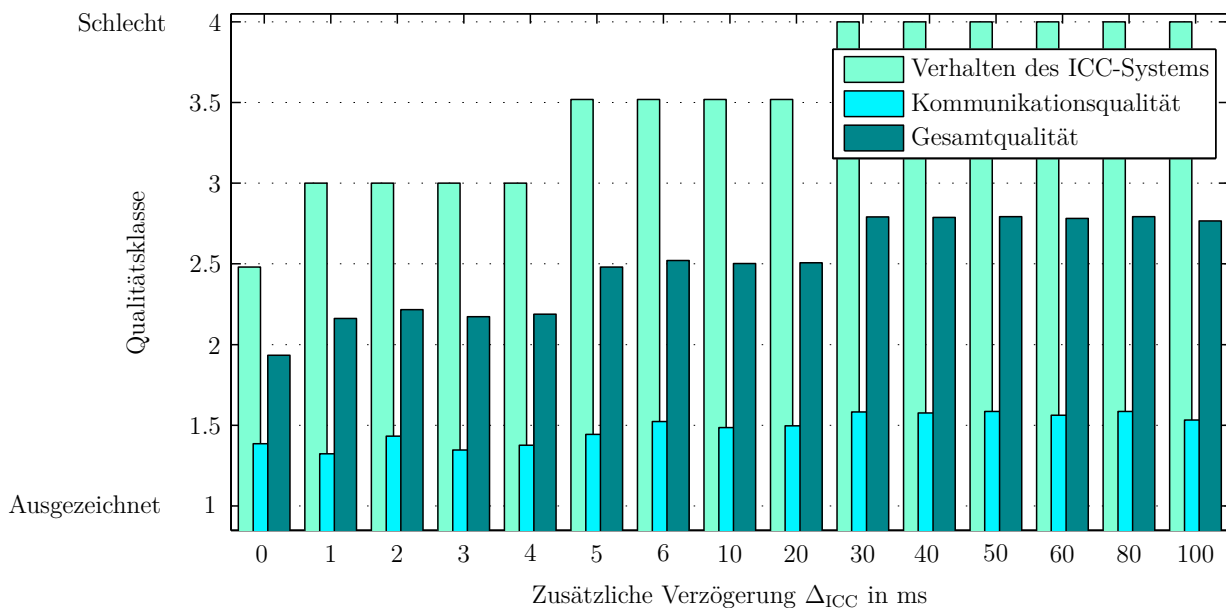


Abbildung 8.9: Qualitätsergebnisse der Evaluierungsgruppen und der Gesamtqualität in Abhängigkeit der zusätzlichen Verzögerung Δ_{ICC} in ms. Auf der Ordinate sind die Qualitätsklassen aufgetragen, wobei 1 einem exzellenten und 4 einem mangelhaften Ergebnis entspricht, siehe auch Tabelle 7.1.

Evaluierungsgruppe eine Verschlechterung der Qualität eines Evaluierungsverfahrens aufgetreten ist, verschlechtern sich auch die Qualitätsergebnisse der Evaluierungsgruppen mit ansteigender Verzögerung. Die Veränderung des Ergebnisses der Evaluierungsgruppe „Kommunikationsqualität“ fällt lediglich gering aus, da auch die Veränderung der Qualität im Sinne der einzelnen Evaluierungsverfahren nur gering ist. Die Veränderung der Eva-

luierungsgruppenqualität Verhalten des ICC-Systems, führt infolgedessen auch zu einer Verschlechterung der Gesamtqualität. Dieses Ergebnis zeigt, dass die vorgestellte Evaluierungsstrategie, mit allen Evaluierungsverfahren, in der Lage ist, eine Veränderung des Verhaltens des ICC-Systems zu detektieren und dies entsprechend zu bewerten.

Kapitel 9

Zusammenfassung

Das Bestreben der vorgelegten Arbeit ist es, eine Methode zur automatischen und reproduzierbaren Erhebung der Qualität eines beliebigen ICC-Systems zu entwickeln. Zu diesem Zweck ist zunächst eine Einordnung des Begriffs der Qualität vorgenommen worden. In dieser Definition ist, neben der allgemeingültigen Klassifikation, zwischen Sprachqualität und Kommunikationsqualität unterschieden worden. Während bei der Sprachqualität der Zuhörer im Mittelpunkt steht, werden bei der Kommunikationsqualität alle Teilnehmer einer Konversation als qualitätsrelevant betrachtet. Im Falle eines ICC-Systems werden sowohl Sprecher als auch Zuhörer durch die Verwendung eines ICC-Systems beeinflusst, da sich beide innerhalb der Fahrgastzelle befinden. Aus diesem Grund darf der Sprecher bei der Evaluierung eines ICC-Systems nicht vollständig außer Acht gelassen werden. Bereits die Erhebung der Sprachqualität für sich, ist kein einfaches Unterfangen, was dazu führt, dass die Kommunikationsqualität weitaus vielschichtiger und schwieriger zu erheben ist.

Um den komplexen Qualitätsbegriff besser abschätzen zu können, ist die grundlegende Idee der vorgestellten Evaluierung, die Gesamtqualität in einzelne Qualitätsergebnisse einer Vielzahl von unterschiedlichen instrumentellen Verfahren aufzusplitten. Dieses Vorgehen bietet den Vorteil, einen tieferen Einblick in die Entstehung der Qualität zu erlangen, da unterschiedlichste Eigenschaften und qualitätsrelevante Faktoren ermittelt werden. Für die Umsetzung einer solchen Evaluierungsstrategie wird neben dem eigentlichen Evaluierungssystem eine Umgebung benötigt, welche das ICC-System ohne direkten elektrischen oder algorithmischen Eingriff in einen gewünschten operativen Zustand versetzt. Diese Gesamtheit einer Evaluierungsumgebung inklusive eines ICC-Systems, ist im Detail betrachtet und ein theoretisches Modell für ein besseres Verständnis des Entwurfes der instrumentellen Verfahren vorgestellt worden.

Für die Übersichtlichkeit und eine sinnvolle Kombination der gefundenen einzelnen Qualitätsergebnisse, werden die Evaluierungsverfahren in drei Evaluierungsgruppen aufgeteilt. In der ersten Gruppe „Fahrzeugeigenschaften“ wird eine Klassifizierung des Fahrzeuges, anhand der Anordnung der Wandler bzw. der Insassen, sowie des auftretenden Hintergrundgeräusches, vorgenommen. Die zweite Gruppe analysiert das Verhalten eines

ICC-Systems, welches Rückschlüsse auf dessen Qualität zulässt. Die dritte Evaluierungsgruppe erhebt die Kommunikationsqualität innerhalb eines mit einem ICC-System ausgestatteten Fahrzeuges. Die Ergebnisse der ersten Evaluierungsgruppe, welche das Fahrzeug klassifizieren, werden anschließend in Attribute, welche das Fahrzeug beschreiben, überführt. Bei den Fahrzeugattributen handelt es sich um das Hintergrundgeräusch, die Größe des Fahrzeuges und die vom ICC-System verwendeten Wandler. Nach der Bestimmung der Fahrzeugattribute, werden diese verwendet, um die Qualitätsergebnisse der Evaluierungsgruppe, welche das Verhalten des ICC-Systems bewerten, zu gewichteten. Entscheidend ist, dass die Ergebnisse der Gruppe „Fahrzeugeigenschaften“ keinen direkten Einfluss auf die Gesamtqualität ausüben. Die Evaluierungsgruppe „Verhalten eines ICC-Systems“ bestimmt die Systemverzögerung, die Unterstützung durch das ICC-System und den Anstieg des Hintergrundgeräusches. Die Gruppe „Kommunikationsqualität“ ermittelt die Qualität anhand der Berechnung eines modifizierten SNR-, sowie STI-Ansatzes. Zusätzlich ist ein spektrales Abstandsmaß vorgestellt worden. Dieses Verfahren ermittelt anhand des leistungsmäßig dominanten Anteils des Nutzsignals innerhalb der Fahrgastzelle ein Referenzsignal und vergleicht dieses mit dem vom Zuhörer empfangenen Signal. Alle drei Verfahren konnten erste plausible und vielversprechende Ergebnisse zur Erhebung der Qualität eines ICC-Systems erzielen.

Zur Entwicklung der vorgestellten Qualitätsschätzer waren Erhebungen des Qualitätsempfinden durch Probanden notwendig. Zu diesem Zweck sind einige psychoakustische Experimente und auditive Hörtests durchgeführt, und die Ergebnisse hier vorgestellt worden. Neben der Qualitätsreferenz konnten auch einige erste Erkenntnisse in Bezug auf das Verhalten der Insassen und der Eigenschaften eines ICC-Systems gesammelt werden. So konnte beispielsweise durch eines der Experimente festgestellt werden, dass ein stetiger Anstieg der Verstärkung des Nutzsignales durch ein ICC-System mit dem Anstieg des Hintergrundgeräusches nicht empfehlenswert ist. Der vermeintliche Grund für dieser Hypothese ist, dass mit steigendem Hintergrundgeräusch die Sprachsignalleistung durch den Lombard-Effekt ebenso ansteigt und dadurch eine geringere Verstärkung durch das ICC-System notwendig ist. Des Weiteren konnte festgestellt werden, dass für den Sprecher eine durch das ICC-System eingefügte Verzögerung von 12 bis 18 ms im Sinne der möglichen Verstärkung am vorteilhaftesten ist. Die erlangten Erkenntnisse, auch die der Hörtests, sind im Anschluss für die Abschätzung der Qualität durch die einzelnen Verfahren angewendet worden.

Alle vorgestellten Verfahren sind nach der Entwicklung innerhalb des lehrstuhleigenen Echtzeit-Rahmenwerks KiRAT implementiert worden. Dadurch sind zusätzliche Anforderungen an die Verfahren entstanden, beispielsweise die Blockverarbeitung der Signale. Dennoch bietet diese Umsetzung den Vorteil einer automatischen Erhebung der Qualität eines ICC-Systems und die Flexibilität, beliebige ICC-Systeme evaluieren zu können. Nach erfolgreicher Implementierung des vorgestellten Evaluierungssystems ist eine konkrete Evaluierung eines ICC-System durchgeführt worden. Als Evaluierungsobjekt ist das

lehrstuhleigene ICC-System, welches ebenfalls in KiRAT realisiert worden ist, verwendet worden. Dadurch konnte eine praktische Anwendung aller Evaluierungsverfahren, die Kombination der Ergebnisse ermöglicht und plausible Ergebnisse erzielt werden. Auch ein Ansatz für eine Visualisierung der Ergebnisse der instrumentellen Verfahren ist vorgestellt worden.

Insgesamt ist eine erste vielversprechende Evaluierungsstrategie, welche in einem Evaluierungssystem innerhalb eines Echtzeit-Rahmenwerks umgesetzt worden ist, vorgestellt worden. Diese Strategie, sowie das System an sich, sind ein erster Anfang in einem noch nicht breit erforschten Gebiet. Aus diesem Grund ergeben sich einige weitere Ansätze zur Erforschung und Entwicklung, welche nachfolgend beleuchtet werden.

9.1 Ausblick

Zunächst sei zu erwähnen, dass die Forschung in Bezug auf die Evaluierung von ICC-Systemen derzeit noch am Anfang steht. Gegenwärtig sind zwar erste einzelne instrumentelle Verfahren, sowie auditive Erhebungen zur Evaluierung untersucht und publiziert worden, jedoch noch keine vollständige Strategie einer Evaluierung vorgestellt worden. Eine solche Strategie mit einer entsprechenden Umsetzung ist Bestandteil der vorgelegten Arbeit und kann als erster, vielversprechender Ansatz zur ganzheitlichen Evaluierung eines ICC-Systems angesehen werden. Dessen ungeachtet, ist die Erforschung weiterer instrumenteller Verfahren, sowie die Weiterentwicklung bestehender, ein wichtiger Punkt, welcher in zukünftigen Analysen berücksichtigt werden sollte. Des Weiteren ist die Frage nach der Kombination der erlangten Ergebnisse von großer Bedeutung. Dabei ist gegenwärtig nicht klar, in wie weit welches Verhalten oder welcher Qualitätsschätzer, welchen Einfluss auf die Gesamtqualität besitzt. Dies sollte mit weiteren psychoakustischen Experimenten ermittelt werden, die einen tieferen Einblick in die Kommunikationssituation innerhalb der Fahrgaszelle unter verschiedensten Gegebenheiten ermöglichen. Diese Ergebnisse ermöglichen es zusätzlich, ein besseres Verständnis für die Umsetzung eines ICC-Systems zu erlangen und somit auch die Qualität dieser Systeme zu erhöhen.

Im Speziellen ist eine Erweiterung der dargelegten Evaluierungsstrategie der drei Evaluierungsgruppen mit weiteren alternativen Evaluierungsverfahren, eine vorstellbare Fortführung der Forschung. Zu diesem Zweck sollten die Gruppen, welche zum einen die Qualität aufgrund des Verhaltens des ICC-Systems und zum anderen die Kommunikationsqualität bewerten, mit neuen instrumentellen Verfahren erweitert werden. Zusätzlich sollte die Bestimmung der Fahrzeugattribute mithilfe von weiteren Erhebungen der Eigenschaften eines Fahrzeuges weiter verfeinert werden. So ist beispielsweise derzeit nicht klar, wie sich das Frequenzverhalten der verwendeten Wandler eines ICC-Systems auf die Gestaltqualität des selbigen auswirkt und wird derzeit noch nicht mit berücksichtigt. Außerdem könnten die beschriebenen, instrumentellen Verfahren in Bezug auf die Qualitätsschätzung weiter verfeinert und stärker validiert werden, um somit eine noch

zuverlässigere Evaluierung erzielen zu können. Dazu sind jedoch weitere und erweiterte auditive Hörtests von Nöten, welche somit eine größere Datenbasis erheben. Solche Tests sind jedoch sehr kosten- und zeitintensiv und stellen einen eigenen, eigenständigen Forschungsbereich dar.

Gleichwohl ist dieser Punkt jedoch essentiell, da er die Grundlage aller Evaluierungsvorhaben eines ICC-System legt. Aufgrund dessen sind weitere Forschungsarbeiten in Bezug auf die auditiven Evaluierungsverfahren sehr wichtig, wenn nicht sogar zwingend notwendig. Hier sollten in jedem Fall internationale Standards und Empfehlungen eingeführt werden, welche es ermöglichen, die gefundenen Ergebnisse und Daten gemeinsam zu nutzen. Dadurch könnte auch einer Verbesserung und insbesondere eine Vergrößerung der Datenbasis auditiver Qualitätsbewertungen erfolgen, welche wiederum, wie bereits erwähnt, zu einer Verfeinerung der instrumentellen Verfahren führt. Denn nur mit einer ausreichend großen Datenbasis, lassen sich zuverlässige instrumentelle Verfahren trainieren und auch entsprechenden validieren.

Ein erster vielversprechender Schritt in die Richtung einer automatischen, reproduzierbaren und instrumentellen Evaluierung von ICC-Systemen ist gemacht worden. Dieser Startpunkt sollte in zukünftigen Forschungsprojekten genutzt werden, um einen Fortschritt der Evaluierung und somit auch der Umsetzung von ICC-Systemen zu erzielen.

Anhang A

Analyse der Regressionskoeffizienten

Bei den folgenden Abbildungen handelt es sich um die Betrachtung der mit dem *jackknife*-Verfahren bestimmten Regressionskoeffizienten und den zugehörigen 95%-Konfidenzintervallen. In Kapitel 4.3 sind die Details des Verfahrens beschrieben. Hier sind nun die entsprechenden Abbildungen der Regressionskoeffizienten der spektralen Distanzmessung erhoben nach Kapitel 7.3.4 und der SNR-Messung nach Kapitel 7.3.3 der Vollständigkeit dargestellt.

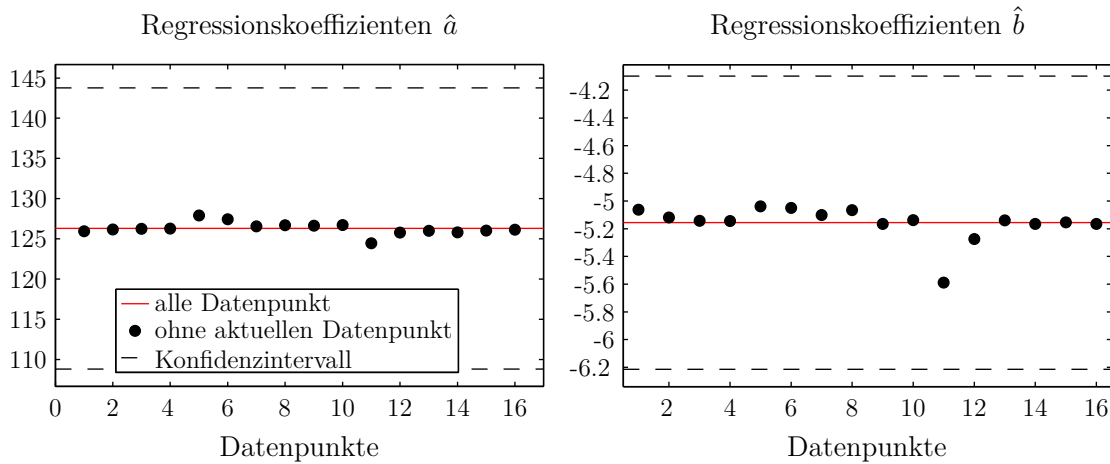


Abbildung A.1: Regressionskoeffizienten \hat{a} und \hat{b} unter Verwendung aller erhobenen Datensätze (schwarze Linie) und mit dem *jackknife*-Verfahren (schwarze Kreise) berechnet. Zusätzlich ist das zugehörige 95%-Konfidenzintervall (gestrichelt) abgebildet. Verwendet worden sind die instrumentellen Daten der spektralen Distanzmethode.

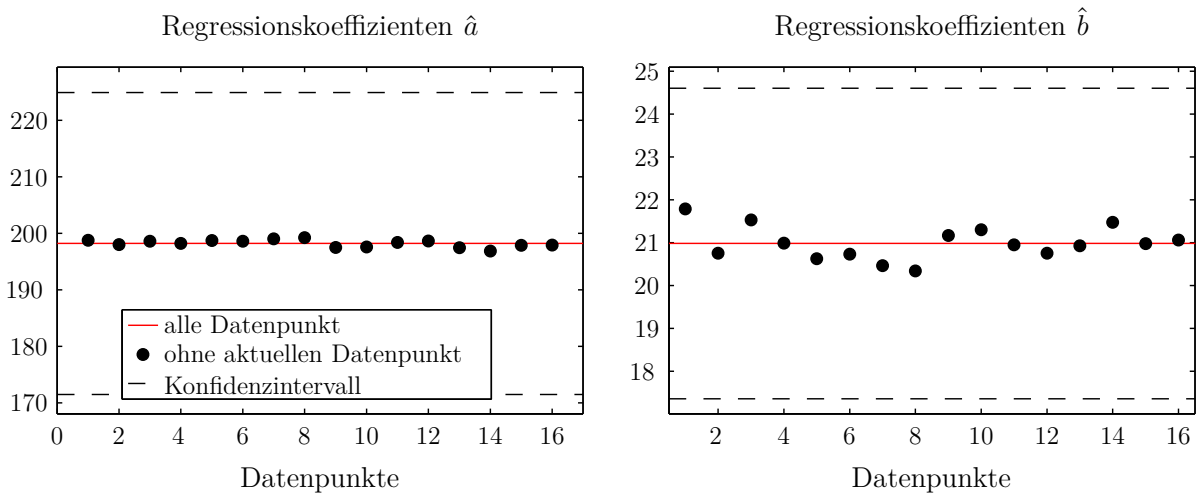


Abbildung A.2: Regressionskoeffizienten \hat{a} und \hat{b} unter Verwendung aller erhobenen Datensätze (schwarze Linie) und mit dem jackknife-Verfahren (schwarze Kreise) berechnet. Zusätzlich ist das zugehörige 95%-Konfidenzintervall (gestrichelt) abgebildet. Verwendet worden sind die instrumentellen Daten der Erhebung des modifizierten SNR.

Literaturverzeichnis

- [AADW04] ALISCH, KATRIN, UTE ARENTZEN und EGGERT DR. WINTER: *Gabler Wirtschaftslexikon*. Gabler, 16., vollst. Überarb. und aktualisierte Aufl. Auflage, 2004.
- [ANS97] ANSI/ASA S3.5-1997 (R2012): *Methods for Calculation of the Speech Intelligibility Index*, Juni 1997.
- [Baa] BAALMAN, MARIJE A.J.: *On wave field synthesis and electro-acoustic music - State of the Art 2007*. International Computer Music Conference, Kopenhagen, Dänemark, 2007.
- [BB] BLAUERT, JENS und JONAS BRAASCH: *Handbuch der Audiotechnik*. Springer-Verlag Berlin Heidelberg.
- [Bla97] BLAUERT, JENS: *Spatial Hearing: The Psychophysics of Human*. The MIT Press, Harvard MA, 1997.
- [BRKS14] BAASCH, C., V.K. RAJAN, M. KRINI und G. SCHMIDT: *Low-complexity noise power spectral density estimation for harsh automobile environments*. Seiten 218–222, September 2014. in Proc. International Workshop on Acoustic Signal Enhancement (IWAENC), Antibes, Frankreich, 2014.
- [DIN] DIN EN 61672-1: *Elektroakustik - Schallpegelmesser - Teil 1: Anforderungen*. Deutsches Institut für Normung, Berlin, Deutschland, 2003.
- [FP] FREUDENBERGER, JÜRGEN und JOHANNES PITTERMANN: *Noise and Feedback Suppression for In-car Communication Systems*. in Proc. 8. ITG-Fachtagung Sprachkommunikation, Aachen, Deutschland, 2008.
- [FS] FOLK, LAURA und FLORIAN SCHIEL: *The Lombard Effect in Spontaneous Dialog Speech*. in Proc. 12th Annual Conference of the International Speech Communication Association (Interspeech 2011), S. 2701-2704, Florenz, Italien, 2011.

- [GG04] GOLDSWORTHY, RAY L. und JULIE E. GREENBERG: *Analysis of speech-based speech transmission index methods with implications for nonlinear operations*. The Journal of the Acoustical Society of America, 116(6):3679–3689, 2004.
- [GK06] GIERLICH, H.W. und F. KETTLER: *Advanced speech quality testing of modern telecommunication equipment: An overview*. Signal Processing, 86(6):1327 – 1340, 2006. Applied Speech and Audio Processing.
- [GTHS] GRAF, SIMON, ANNE THEISS, TOBIAS HERBIG und GERHARD SCHMIDT: *Listening Test to Determine the Mismatch Between Signal-To-Noise Ratio and Human Perception*. in Vortragsband „Fortschritte der Akustik“ der 41. Jahrestagung für Akustik (DAGA), Nürnberg, Germany, 2015.
- [GZ03] GAZOR, S. und WEI ZHANG: *Speech probability distribution*. IEEE Signal Processing Letters, 10(7):204–207, July 2003.
- [Haa72] HAAS, HELMUT: *The Influence of a Single Echo on the Audibility of Speech*. J. Audio Eng. Soc, 20(2):146–159, 1972.
- [Hay01] HAYKIN, SIMON: *Adaptive Filter Theory*. Prentice-Hall, 4. Auflage, 2001.
- [HL08] HU, Y. und P. C. LOIZOU: *Evaluation of Objective Quality Measures for Speech Enhancement*. IEEE Transactions on Audio, Speech, and Language Processing, 16(1):229–238, Jan 2008.
- [Hän01] HÄNSLER, EBERHARD: *Statistische Signale*. Springer, 3. Auflage, 2001.
- [IEC] IEC 60268-16:2011: *Objective rating of speech intelligibility by speech transmission index*. International Electrotechnical Commission, 2011.
- [ITU94] ITU-T RECOMMENDATION P.85: *A method for subjective performance assessment of the quality of speech voice output devices*, Juni 1994.
- [ITU96a] ITU-T RECOMMENDATION P.800: *Methods for subjective determination of transmission quality*, August 1996.
- [ITU96b] ITU-T RECOMMENDATION P.830: *Subjective performance assessment of telephone-band and wideband digital codecs*, Februar 1996.
- [ITU05] ITU-T RECOMMENDATION P.505: *One-view visualization of speech quality measurement results*, Nov. 2005.
- [ITU13] ITU-T FOCUS GROUP ON CAR COMMUNICATION: *Final Report*, November 2013. Version 1.0.

- [ITU14] ITU-R RECOMMENDATION BS.1534-2: *Method for the subjective assessment of intermediate quality level of audio systems*, June 2014.
- [JA90] JUNQUA, JEAN-CLAUDE und Y. ANGLADE: *Acoustic and Perceptual Studies of Lombard Speech: Application to Isolated-Words Automatic Speech Recognition*. Seiten 841–844 vol.2, April 1990. in Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Albuquerque, New Mexico, USA1990.
- [Jek05] JEKOSCH, U: *Voice and Speech Quality Perception*. Springer, Berlin, Heidelberg, 2005.
- [KHG00] KREBBER, WINFRIED, HANS-WILHELM HIERLICH und KLAUS GENUIT: *Auditory Virtual Environments: Basics and Applications for Interactive Simulation*. Signal Processing, 80(11):2307–2322, 2000.
- [KK98] KAMMEYER, KARL-DIRK und KRISTIAN KROSCHEL: *Digitale Signalverarbeitung Filterung und Spektralanalyse mit MATLAB-Übungen*. Teubner Studienbücher Elektrotechnik, 4 Auflage, 1998.
- [KP95] KLEIJN, W. B. und K. K. PALIWAL (Herausgeber): *Speech Coding and Synthesis*. Elsevier Science Inc., New York, NY, USA, 1995.
- [Kut09] KUTTRUFF, HEINRICH: *Room Acoustics*. Spon Press, Milton Park, Abingon, Oxon, 2009.
- [LB64] LOCHNER, J.P.A. und J.F. BURGER: *The influence of reflections on auditorium acoustics*. Journal of Sound and Vibration, 1(4):426 – 454, 1964.
- [LF] LINHARD, KLAUS und JÜRGEN FREUDENBERGER: *Passenger In-Car Communication Enhancement*. in Proc. EUSIPCO, Vienna, Austria, 2004.
- [LKKLS] LANDGRAF, RABEA, JOHANNES KÖHLER-KAESS, CHRISTIAN LÜKE und GERHARD SCHMIDT: *Can you hear me now? Reducing the Lombard effect in a driving car using an In-Car Communication system*. in Proc. International Conference of Speech Prosody, Boston, MA, USA, 2016.
- [LNS⁺] LANDGRAF, RABEA, OLIVER NIEBUHR, GERHARD SCHMIDT, TINA JOHN, CHRISTIAN LÜKE und ANNE THEISS: *Von der Straße ins Labor: Die Modifikation der Sprachproduktion bei lauten Fahrgeräuschen*. in Vortragsband „Fortschritte der Akustik“ der 41. Jahrestagung für Akustik (DAGA), Nürnberg, Germany, 2015.
- [Loi13] LOIZOU, PHILIPPOS C: *Speech enhancement: theory and practice*. CRC press, 2013.

- [Lom] LOMBARD, ÉTIENNE: *Le Signe de l'Élévation de la Voix*. Annales Maladiers Oreille, Larynx, Nex, Pharynx, vol. 37, S. 101-109, 1911.
- [LS] LINHARD, KLAUS und HARALD SCHNEPP: *Modified SNR for Evaluation of Speech Quality*. in Proc. 8. ITG-Fachtagung Sprachkommunikation, Aachen, Germany, 2008.
- [LSTW14] LÜKE, CHRISTIAN, GERHARD SCHMIDT, ANNE THEISS und JOCHEN WITHOPF: *In-Car Communication*. In: SCHMIDT, GERHARD, HUSEYIN ABUT, KAZUYA TAKEDA und JOHN H.L. HANSEN (Herausgeber): *Smart Mobile In-Vehicle Systems*, Seiten 97–118. Springer New York, Heidelberg, Germany, 2014.
- [LTS+] LÜKE, CHRISTIAN, ANNE THEISS, GERHARD SCHMIDT, OLIVER NIEBUHR und TINA JOHN: *Creation of a Lombard Speech Database using an Acoustic Ambiance Simulation with Loudspeakers*. 6th Biennial DSP Workshop for In-Vehicle Systems, Seoul, Korea, 2013.
- [LW06] LEHN, J. und H. WEGMANN: *Einführung in die Statistik*. Lehrbuch : Mathematik. Vieweg+Teubner Verlag, 2006.
- [LÖS+] LÜKE, CHRISTIAN, HALIL ÖZER, GERHARD SCHMIDT, ANNE THEISS und JOCHEN WITHOPF: *Signal Processing for In-Car Communication Systems*. 5th Biennial Workshop on DSP for In-Vehicle Systems, Kiel, Germany, 2011.
- [Mer13] MERTINS, ALFRED: *Signaltheorie: Grundlagen der Signalbeschreibung, Filterbänke, Wavelets, Zeit-Frequenz-Analyse, Parameter- und Signalschätzung*. Springer, 3. Auflage, 2013.
- [MK] MUESCH, UWE und FRANK KETTLER: *Instrumental Testing of In-Car Communication Systems*. in Vortragsband „Fortschritte der Akustik“ der 41. Jahrestagung für Akustik (DAGA), Nürnberg, Germany, 2015.
- [MKL] MUESCH, UWE, FRANK KETTLER und MARC LEPAGE: *Speech Quality Tests of In-Car Communication Systems*. in Proc. 10. ITG-Fachtagung Sprachkommunikation, Braunschweig, Germany, 2012.
- [MM] MAULER, D. und R. MARTIN: *Optimization of Switchable Windows for Low-Delay Spectral Analysis-Synthesis*. in Proc. ICASSP, Dallas, Texas, USA, 2010.
- [MR] MÖLLER, SEBASTIAN und ALEXANDER RAAKE: *Telephone speech quality prediction: Towards network planning and monitoring models for modern network scenarios*. Speech Communication, 38(1-2):47-75, 2002.

- [MS52] MEYER, E und GR SCHODDER: *Über den Einfluss von Schallrückwürfen auf Richtungslokalisierung und Lautstärke bei Sprache*. Nachr Akad Wiss Göttingen, Math Phys Klasse II, a(6):31–42, 1952.
- [MSB12] MYERS, GLENFORD J., COREY SANDLER und TOM BADGETT: *The Art of Software Testing*. John Wiley & Sons, Inc., Hoboken, New Jersey, 3. Edition Auflage, 2012.
- [MSG78] MCDERMOTT, B., C. SCAGLIOLA und D. GOODMAN: *Perceptual and objective evaluation of speech processed by adaptive differential PCM*. Band 3, Seiten 581–585, 1978. in Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Tulsa, Oklahoma, USA, 1978.
- [Möl100] MÖLLER, SEBASTIAN (Herausgeber): *Assessment and Prediction of Speech Quality in Telecommunications*. Springer Science+Business Media Drodrecht, 2000.
- [Möl110] MÖLLER, SEBASTIAN (Herausgeber): *Quality Engineering. Qualität kommunikationstechnischer Systeme*. Springer-Verlag Berlin Heidelberg, 2010.
- [NN97] NÉLISSE, H. und J. NICOLAS: *Characterization of a Diffuse Field in a Reverberant Room*. The Journal of the Acoustical Society of America, 101(6):3517–3524, 1997.
- [OLM05] ORTEGA, A., E. LLEIDA und E. MASGRAU: *Speech Reinforcement System for Car Cabin Communications*. IEEE Transactions on Speech and Audio Processing, 13(5):917–929, Sept 2005.
- [OLMG] ORTEGA, ALFONSO, E. LLEIDA, ENRIQUE MASGRAU und FERNANDO GALLEGRO: *Cabin Car Communication System to Improve communications Inside a Car*. in Proc. ICASSP, Orlando, Florida, 2002.
- [PM07] PROAKIS, JOHN G. und DIMITRIS G. MANOLAKIS: *Digital Signal Processing: Principles, Algorithms, and Applications*. Pearson International Edition, 4. Auflag. Auflage, 2007.
- [Pri14] PRIGGE, LARS: *Untersuchung qualitätsrelevanter Fahrzeugeigenschaften für Innenraumkommunikationssysteme*. Bachelorarbeit, Christians-Albrechts-Universität zu Kiel, Kiel, Deutschland, 2014.
- [QBC88] QUACKENBUSH, SCHULYER R., THOMA P. BARNWELL und MARK A. CLEMENTS: *Objective Measure of Speech Quality*. Prentice Hall, Engelwood Cliffs, New Jersey, 1988.

- [Qua08] *Speech and Audio Processing in Adverse Environments*. Springer-Verlag Berlin Heidelberg, 2008.
- [Raa06] RAAKE, ALEXANDER (Herausgeber): *Speech Quality of VOIP*. John Wiley and Sons, 2006.
- [RBW14] RUGE, PETER, CAROLIN BIRK und MANFRED WERMUTH: *Das Ingenieurwissen: Mathematik und Statistik*. Springer-Verlag Berlin Heidelberg, 2014.
- [Ros14] ROSSING (Herausgeber): *Handbook of Acoustics*. Springer-Verlag Berlin Heidelberg, 2. Aufl. Auflage, 2014.
- [Sch] SCHOLZ, KIRSTIN: *Instrumentelle Qualitätsbeurteilung von Telefonbandsprache beruhend auf Qualitätsattributen*. Doktorarbeit, Christian-Albrechts-Universität zu Kiel. Arbeiten über Digitale Signalverarbeitung Nr. 32, Shaker Verlag, Aachen, 2008.
- [SH80] STEENEKEN, H. J. M. und T. HOUTGAST: *A physical method for measuring speech transmission quality*. The Journal of the Acoustical Society of America, 67(1):318–326, 1980.
- [SH85] STEENEKEN, H. J. M. und T. HOUTGAST: *A review of the MTF concept in room acoustics and its use for estimating speech intelligibility in auditoria*. The Journal of the Acoustical Society of America, 77(3):1069–1077, 1985.
- [SH06] SCHMIDT, GERHARD und TIM HAULICK: *Signal Processing for In-Car Communication Systems*. In: HÄNSLER, EBERHARD und GERHARD SCHMIDT (Herausgeber): *Topics in Acoustic Echo and Noise Control*, Seiten 547–598. Springer-Verlag Berlin Heidelberg, 2006.
- [SPB⁺88] SUMMERS, W. VAN, DAVID B. PISONI, ROBERT H. BERNACKI, ROBERT I. PEDLOW und MICHAEL A. STOKES: *Effects of noise on speech production: Acoustic and perceptual analyses*. Band 84, Seiten 917–928, 1988.
- [STWW12a] SCHMIDT, GERHARD, ANNE THEISS, JOCHEN WITHOPF und ARTHUR WOLF: *Evaluation of In-Car Communication Systems*. In: HANSEN, JOHN H.L., PINAR BOYRAZ, KAZUYA TAKEDA und HÜSEYİN ABUT (Herausgeber): *Digital Signal Processing for In-Vehicle Systems and Safety*, Seiten 73–107. Springer New York, 2012.
- [STWW12b] SCHMIDT, GERHARD, ANNE THEISS, JOCHEN WITHOPF und ARTHUR WOLF: *Signal Processing for In-Car Communication Systems and its Evaluation*. In: HANSEN, JOHN H.L., P. BOYARZ, KAZUYA TAKEDA und

- HUSEYIN ABUT (Herausgeber): *Digital Signal Processing for In-Vehicle Systems and Safety*, Seiten 73–107. Springer New York, Heidelberg, Germany, 2012.
- [TNMC78] TRIBOLET, J., P. NOLL, B. MCDERMOTT und R. CROCHIERE: *A study of complexity and quality of speech waveform coders*. Band 3, Seiten 586–590, 1978. in Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Tulsa, Oklahoma, USA, 1978.
- [TSa] THEISS, ANNE und GERHRAD SCHMIDT: *Investigation of Self-Masking Effects for the Evaluation of In-Car Communication Systems*. in Proc. International Workshop on Acoustic Signal Enhancement (IWAENC), Antibes, France, 2014.
- [TSb] THEISS, ANNE und GERHRAD SCHMIDT: *Spectral Distance Analysis for Quality Estimation of In-Car Communication Systems*. 7th Biennial DSP Workshop for In-Vehicle Systems, San Francisco, USA, 2015.
- [TSWL] THEISS, ANNE, GERHARD SCHMIDT, JOCHEN WITHOPF und CHRISTIAN LÜKE: *Instrumental Evaluation of In-Car Communication Systems*. in Proc. 11. ITG-Fachtagung Sprachkommunikation, Erlangen, Germany, 2014.
- [WJST] WITHOPF, JOCHEN, LAILA JASSOUME, GERHARD SCHMIDT und ANNE THEISS: *A Modified Overlap-Add Filter Bank With Reduced Delay*. in Vortragsband „Fortschritte der Akustik“ der 38. Jahrestagung für Akustik (DAGA), Darmstadt, Germany, 2012.
- [ZF99] ZWICKER, EBERHARD und HUGO FASTL: *Psycho-acoustics, Facts and Models*. Springer-Verlag Berlin Heidelberg, 2. Auflag. Auflage, 1999.
- [Zwi61] ZWICKER, E.: *Subdivision of the Audible Frequency Range into Critical Bands (Frequenzgruppen)*. The Journal of the Acoustical Society of America, 33(2), 1961.