# Digitale Simulation der Farzeuginnenraumakustik zur Unterstützung der Entwicklung und Evaluierung von Innenraum-Kommunikationssystemen

Christian Lüke<sup>a</sup>, Arthur Wolf<sup>b</sup>, Michael Brodersen<sup>a</sup>, Gerhard Schmidt<sup>a</sup>

<sup>a</sup>Christian-Albrechts-Universität zu Kiel

<sup>b</sup>Daimler AG

cl@tf.uni-kiel.de

## Motivation

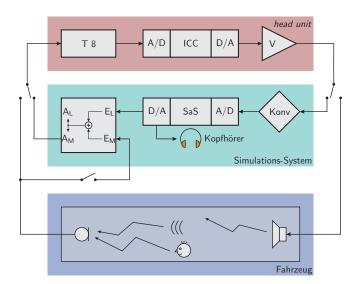
Sprachsignalverarbeitungssysteme wie z. B. ICC-Systeme arbeiten unter dem Einfluss von Hintergrundgeräuschen und in einer geschlossenen akustischen Schleife. Ihre Hauptaufgaben bestehen daher in einer Signalverbesserung durch Unterdrückung der Hintergrundgeräusche sowie die Kompensation und Unterdrückung von Rückkopplungen [1].

Oftmals sollen derartige Systeme bewertet werden, auch um Potential zur Optimierung erkennen zu können. Optimalerweise kann eine solche Bewertung erfolgen, ohne das System im Detail zu kennen (sog. black-box testing). Die Güte eines Sprachsignalverarbeitungssystems muss daher einerseits in einem rückgekoppelten System ermittelt werden; andererseits sollte die Ausgabe des Systems möglichst unverfälscht einer Evaluation zugeführt werden können. Durch ein Simulationssystem ist es möglich, das Nutzsignal, das Hintergrundgeräusch sowie Rückkopplungen zu simulieren und zugleich das reine Ausgangssignal des Sprachverarbeitungssystems zu analysieren.

Ein solches Simulationssystem kann entweder akustisch durch eine kalibrierte Lautsprecherwiedergabe ausgeführt werden [2] oder durch elektrische Signaleinspeisung erfolgen. Die rein elektrische Simulation kann durch eine PC-basierte Implementierung oder durch einen digitalen Signalprozessor (DSP) umgesetzt werden [3]. Die Umsetzung auf einem DSP bietet durch die Echtzeitfähigkeit des Prozessors den Vorteil einer geringen Latenz, Konfiguration und Wartung gestalten sich jedoch relativ umständlich. Umgekehrt bietet der PC die Möglichkeit der komfortablen Konfiguration durch eine graphische Benutzungsoberfläche, während die Latenz nach unten durch die nebenläufigen Betriebssysteme und die Treiberarchitektur begrenzt wird. Unsere PC-basierte Lösung erreichte in der praktischen Anwendung eine Latenz von 3,265 ms.

#### Gesamtsystem

Abbildung 1 zeigt einen möglichen Einsatz zur Simulation akustischer Systeme: oben das zu vermessende blackbox-System, beispielsweise eine head unit ("Autoradio") mit einem ICC-System, A/D- und D/A-Wandlung und Lautsprecherendstufe (V). Zur Spannungsversorgung der anzuschließenden Mikrophone wird üblicherweise eine Tonaderspeisung verwendet, die auf die Signaladern des Mikrophons eine Spannung von 8 V anlegt (T 8). Das



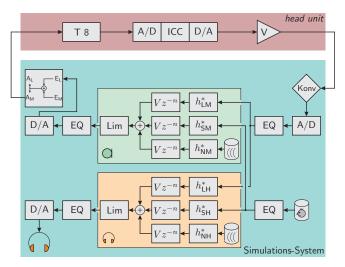
**Abbildung 1:** Komponenten zur Simulation akustischer Systeme

zu vermessende System kann nun entweder mit dem Simulationssystem oder mit der realen Umwelt, beispielsweise mit dem Fahrzeug, interagieren. Das Simulationssystem passt zunächst mittels eines Pegelkonverters (Konv) die Lautsprechersignale an den Aussteuerungsbereich des A/D-Wandlers an. Eine Anpass-Hardware fängt eine eventuell durch die  $head\ unit$  zur Verfügung gestellte Tonaderspeisung ab und kann Signale realer Mikrophone zu den simulierten Signalen analog addieren, um beispielsweise subjektive Tests vornehmen zu können.

#### Software

Die Algorithmen zur Simulation akustischer Systeme sind im Echtzeitrahmenwerk für Signalverarbeitung Kiel Real-time Application Toolkit (KiRAT) implementiert. KiRAT wird an der Arbeitsgruppe für Digitale Signalverarbeitung und Systemtheorie (DSS) an der Christian-Albrechts-Universität zu Kiel entwickelt und stellt unter anderem eine Schnittstelle zu Audiowandlern zur Verfügung, die mit geringer Latenz betrieben werden kann.

Die Struktur des Softwaremoduls zur Simulation akustischer Systeme (SaS) ist in Abbildung 2 gezeigt. Die Signalquellen auf der rechten Seite sind die Lautsprechersignale der head unit sowie Sprachdateien aus einer Datenbank, die genutzt werden können, um innerhalb eines Fahrzeuges verschiedene Sprecher zu simulieren. Die Signalsenken auf der linken Seite sind die simulierten Mikrophonsignale und die Kopfhörersignale. Die Hauptbe-



**Abbildung 2:** *Head unit* (oben) und Simulationssystem mit detaillierter Softwarestruktur (unten)

standteile der Software sind die beiden identischen Untermodule zur Simulation von Mikrophonsignalen (obere Box, grün hinterlegt) und von Kopfhörersignalen (untere Box, orange hinterlegt). Innerhalb dieser Untermodule werden die simulierten Signale aus drei Quellen zusammengefügt. Im oberen Pfad erfolgen die  $N_{\rm Lsp}$  ×  $N_{\rm Mic}$  Faltungen  $(h_{\rm LM}^*)$  der  $N_{\rm Lsp}$  Lautsprechersignale auf die  $N_{\rm Mic}$  Mikrophonsignale sowie die Anpassung von Verstärkungen und Latenzen  $(Vz^{-n})$ . Im mittleren Pfad werden  $N_{\rm Src}$  lokale Quellen mittels  $N_{\rm Src} \times N_{\rm Mic}$  Faltungen  $(h_{\text{SM}}^*)$  gefiltert und können ebenfalls in Verstärkung und Latenz angepasst werden. Im unteren Pfad werden Hintergrundgeräusche aus einer Datenbank gelesen, bei Bedarf gefiltert, verstärkt und verzögert, und dann auf die Mikrophonsignale aufaddiert. Zur realistischen Nachbildung von Rückkopplungen bildet schließlich ein Begrenzer (Lim) das nichtlineare Verhalten von Verstärkern und Lautsprechern nach.

Zur Berechnung der Faltungen wurden verschiedene Verfahren implementiert, die zur Laufzeit umgeschaltet werden können: Eine Overlap-save-basierte Faltung auf dem Hauptprozessor (CPU), eine Overlap-save-basierte Faltung auf einem CUDA-fähigen Graphikprozessor (GPU) sowie eine Overlap-save-basierte Faltung auf dem Hauptprozessor (CPU) mit nichtgleichförmigen Blockgrößen und paralleler Verarbeitung [4]. Innerhalb von KiRAT können Entzerrfilter direkt auf die Ein- und Ausgangssignale der Soundkarte angewendet werden (EQ). Ein graphisch gestützter Filterentwurf kann mithilfe eines Prädiktoransatzes in KiRAT ausgeführt werden. Dazu kann eine gemessene Impulsantwort g(n) als Ausgangspunkt genutzt werden, deren Betragsfrequenzgang  $|G(e^{j\Omega})|$  zu invertieren ist.

#### Hardware

Für die Simulation von Fahrzeuginnenräumen wurde ein Simulationsrack aufgebaut, das folgende Komponenten beinhaltet:

- 1. Lautsprechereingänge mit Leistungskonvertern nach line-Pegel,
- 2. eine Multiface-II-Soundkarte mit PC-Anschluss und

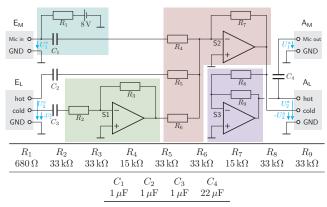


Abbildung 3: Schaltskizze und Wertetabelle des analogen Mischers im Simulations-Rack

3. eine analoge Mischerschaltung zum Mischen der symmetrischen Soundkartenausgänge zu asymmetrischen Signalen von realen Mikrophonen sowie Ausgabe der gemischten Signale als symmetrisches *line*-Signal und als asymmetrisches Mikrophonsignal.

Die analoge Mischerschaltung ist in Abbildung 3 dargestellt und wurde achtkanalig aufgebaut. Links befinden sich die grau hinterlegten Eingänge, oben der asymmetrische Mikrophoneingang  $E_{\rm M}$  zum Zumischen realer Mikrophone und unten ein symmetrischer Line-Eingang  $E_{\rm L}$  für die Ausgabe der Simulationssoftware. Links oben, türkis hinterlegt, findet sich die Tonaderspeisung für das hinzuzumischende Mikrophon. Die grün hinterlegte Inverterschaltung stellt die korrekte Signalpolarität des invertierenden Eingangs (cold) her. Mittels der Widerstände  $R_4$ ,  $R_5$  und  $R_6$  werden die drei Teilsignale aufsummiert und vom zweiten Operationsverstärker invertierend verstärkt (rot hinterlegt). Um die korrekte Signalpolarität wiederherzustellen, ist eine weitere, violett hinterlegte Inverterschaltung angefügt. Ihr Ausgang führt in den nichtinvertierten symmetrischen Ausgang (hot) sowie über den Blockkondensator  $C_4$  zum Mikrophonausgang. Auf der rechten Seite befinden sich die grau hintelegten Ausgänge, oben der asymmetrische Mikrophonausgang  $A_{\rm M}$  zum Anschluss einer head unit und unten ein symmetrischer Line-Ausgang  $A_{\rm L}$  zum Anschluss einer Soundkarte.

Der Widerstand  $R_1$  für die Tonaderspeisung wurde nach ITU-Empfehlung mit  $R_1=680\,\Omega$  gewählt. Für die Wahl der Größen der weiteren Widerstände sind die jeweils gewünschten Verstärkungsfaktoren der Operationsverstärkerschaltungen entscheidend. So soll der Inverter des symmetrischen Eingangssignals eine Verstärkung von  $v_1=-1$  besitzen, was durch das Verhältnis der Widerstände  $R_2$  und  $R_3$  festgelegt wird.

derstände 
$$R_2$$
 und  $R_3$  festgelegt wird,
$$-1 = v_{S1} = \frac{U_{S1}^a}{U_{S1}^e} = -\frac{R_3}{R_2} \quad \Leftrightarrow \quad R_2 = R_3. \quad (1)$$

Dieses Schema gilt analog für den Inverter am Ausgang  $(R_8=R_9)$ . Entsprechend dieser Vorgabe wurden die Widerstände  $R_2=R_3=R_8=R_9=33\,\mathrm{k}\Omega$  gewählt. Für den invertierenden Addierer war das Ziel, die Amplituden bzw. Spannungen an den Eingängen jeweils unverändert an den Ausgängen wieder auszugeben, es soll also

$$U_1^{\rm a} = U_2^{\rm a} = U_1^{\rm e} + U_2^{\rm e} \tag{2}$$

gelten, wobei  $U_2^{\rm e}$  durch die beiden Signalanteile hot und cold zusammensetzt. Um dies zu erfüllen, wird

$$R_4 = R_7 = 15 \,\mathrm{k}\Omega$$
 und  $2R_7 \approx R_5 = R_6 = 33 \,\mathrm{k}\Omega$  (3)

gewählt. Diese Wahl stellt sicher, dass die Eingangswiderstände des symmetrischen Eingangs, die durch  $R_5$  und  $R_2$  festgelegt werden, gleich groß sind. Statt des rechnerisch korrekten Wertes von  $30\,\mathrm{k}\Omega$  wurden Widerstände der Größe  $33\,\mathrm{k}\Omega$  aus der verbreiteten Reihe E12 (DIN IEC 60063) verwendet. Daraus ergibt sich dann eine rechnerische Verstärkung für die symmetrischen Pfade von

$$v_2 = \frac{U_2^{\rm a}}{U_2^{\rm e}} = -\left(\frac{R_7}{R_5} + \frac{R_3}{R_2} \frac{R_7}{R_6}\right) \left(-\frac{R_9}{R_8}\right) = 2 \frac{R_7}{R_6}$$
(4)  
=  $2 \frac{15 \,\mathrm{k}\Omega}{33 \,\mathrm{k}\Omega} \approx 0.909 \equiv -0.83 \,\mathrm{dB}.$ 

Diese Dämpfung muss in der Simulations-Software ausgeglichen werden.

Die in der Schaltung eingesetzten Kondensatoren sind nötig, um eventuelle Gleichspannungen an den Einund Ausgängen abzublocken. Sie stellen zusammen mit den ihnen nachgeschalteten Widerständen immer Hochpassfilter dar, deren Grenzfrequenz durch

$$f_{\rm c} = \frac{1}{2\pi RC} \tag{5}$$

gegeben ist. Für die Kondensatoren  $C_1$ ,  $C_2$  und  $C_3$  sind die nachfolgenden Widerstände wie oben beschrieben durch die eigene Schaltung festgelegt. Um eine maximale Grenzfrequenz von ca. 10 Hz zu erhalten, wurde  $C_1=C_2=C_3=1\,\mu\mathrm{F}$  gewählt. Der Mikrophonausgang wurde für Tonaderspeisungen mit einem Eingangswiderstand von  $680\,\Omega$  ausgelegt. Für den ungepolt auszuführenden Kondensator  $C_4=22\,\mu\mathrm{F}$  beträgt die Grenzfrequenz

$$f_{\rm e2}^{\rm c} = \frac{1}{2\pi R_{\rm o}^{\rm T} C_4} = \frac{1}{2\pi 680 \Omega 22 \,\mu\text{F}} \approx 10.6 \,\text{Hz}.$$
 (6)

Die Spannungsversorgung durch ein Schaltnetzteil wird intern zunächst durch einen Gleichspannungsregler stabilisiert, der auch die für die Operationsverstärker benötigten  $-12\,\mathrm{V}$  liefert. Die Ausgangsspannung kann dabei eine Restwelligkeit von  $60\,\mathrm{mV}$  (Peak-Peak) aufweisen. Diese Restwelligkeit muss durch einen weiteren Festspannungsregler entfernt werden. Dadurch konnte ein Störabstand von  $-95,6\,\mathrm{dBFS}$  bzw. ein Störpegel von  $-82,6\,\mathrm{dBu}$  erreicht werden (bei  $+4\,\mathrm{dBu}$  Referenzpegel,  $0\,\mathrm{dBFS} \equiv 13\,\mathrm{dBu}$ ). Ein Foto der aufgebauten Schaltung mit Platine zur Glättung der Spannungsversorgung für die Tonaderspeisung ist in Abbildung 4 gezeigt.

#### Latenzen

Um eine korrekte Wiedergabe im Simulationssystem zu erreichen, wird in einem ersten Schritt die Latenz des Messsystems ermittelt. Diese Latenz setzt sich aus diversen Pufferungen und den Latenzen der Analog-Digitalund Digital-Analog-Wandler zusammen. Die Messung der Latenz des Messsystems kann wie in Abbildung 5 erfolgen. Dabei wird ein weißes Rauschen generiert (engl. white noise generator, WNG), durch die Wandler geschickt und ihr Verhalten durch ein adaptives Filter  $\hat{h}_{\rm D}(n)$  nachgebildet. Diese gemessene Impulsantwort



Abbildung 4: Aufgebauter analoger Mischer für das Simulationsrack

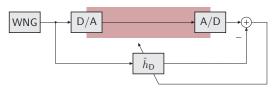


Abbildung 5: Messung der Laufzeit des Messsystems; analoge Signale sind in roter Farbe hinterlegt.

der D/A-A/D-Wandlung  $\hat{h}_{\rm D}(n)$  wird vereinfachend als Dirac-Impuls  $\delta(n-d_{\rm D})$  mit der Latenz  $d_{\rm D}=d_{\rm DA}+d_{\rm AD}$  aufgefasst. Soll die Impulsantwort des Systems h(n) aus Abbildung 6 gemessen werden, wird zunächst stets die Impulsantwort

$$h'(n) = h_{DA}(n) * h(n) * h_{AD}(n)$$
 (7)

des Systems aus D/A-Wandlung, h(n) und A/D-Wandlung gemessen. Durch die soeben berechnete Latenz  $d_{\rm D}$  entsteht dabei eine Totzeit

$$\hat{h}'(n) \approx 0 \quad \forall n < d_{\rm D}$$
 (8)

in der gemessenen Impulsantwort  $\hat{h}'(n)$ . Um eine Schätzung  $\hat{h}$  für die Impulsantwort des Systems h(n) zu erhalten, muss die Latenz  $d_{\rm D}$  des Messsystems abgezogen werden,

$$\hat{h}(n) = \hat{h}'(n + d_{\mathcal{D}}),\tag{9}$$

wobei wegen Gleichung 8 keine Information verlorengeht. Für die Simulation akustischer Systeme ist nun entscheidend, dass im Rückkopplungsweg erneut eine D/Aund A/D-Wandlung vorliegt, die im zu simulierenden Fahrzeuginnenraum nicht stattfinden würde (siehe Abbildung 2). Die durch diese Wandlung zusätzlich eingefügte Latenz muss aus der Impulsantwort  $\hat{h}_{\rm LM}(n)$  herausgerechnet werden, um die Gesamtlaufzeit originalgetreu nachbilden zu können. Die zu berechnende Faltung erfolgt daher mit der gekürzten Impulsantwort

$$h_{\text{LM}}^*(n) = \hat{h}_{\text{LM}}(n + d_{\text{D}})$$
 (10)

$$=\hat{h}'_{\rm LM}(n+2\cdot d_{\rm D}). \tag{11}$$

Hieraus ergibt sich die Forderung, dass die Latenz  $d_{\rm D}$ kleiner sein muss als die Totzeit der Impulsantwort

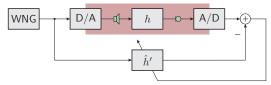
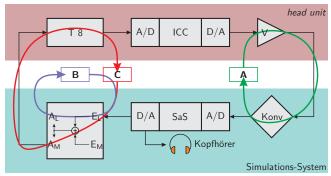


Abbildung 6: Messung einer Impulsantwort



**Abbildung 7:** Gesamtsystem aus *head unit* (oben) und digitaler Raumsimulation (unten) mit drei Teilsystemen (A, B, C)

 $h_{\rm LM}(n)$ , dass also

$$\hat{h}'_{\rm LM}(n) \approx 0 \quad \forall n < 2 \cdot d_{\rm D}$$
 (12)

gilt. Eine weitere Forderung ist, dass der Direktschall und die Ausgabe der head unit den Kopfhörer bzw. den Zuhörer synchron erreichen. Durch die D/A-A/D-Wandlungen im Rückkopplungspfad wird eine zusätzliche Latenz  $d_{\rm D}$  eingefügt. Diese zusätzliche Latenz muss ausgeglichen werden, indem sie der Impulsantwort  $h_{\rm SH}^*(n)$  aufaddiert wird. Damit ergeben sich im Vergleich zu den gemessenen Impulsantworten  $\hat{h}'(n)$  und den im Simulationssystem zu verwendenden Impulsantworten  $h^*(n)$  die Beziehungen

$$h_{\rm SM}^* = \hat{h}_{\rm SM}'(n+d_{\rm D}),$$

$$h_{\rm LM}^* = \hat{h}_{\rm LM}'(n+2\cdot d_{\rm D}),$$

$$h_{\rm SH}^* = \hat{h}_{\rm SH}'(n) \quad \text{und}$$

$$h_{\rm LH}^* = \hat{h}_{\rm LH}'(n+d_{\rm D}).$$
(13)

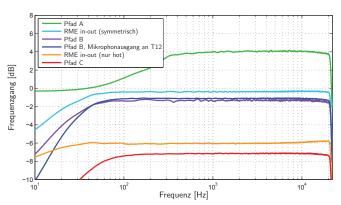
# Ergebnisse

In einem Aufbau zur Simulation von 2 Mikrophon- und 8 Kopfhörerkanälen auf der Basis von 4 Lautsprecherkanälen mit Filterlängen von 1024 und einer Abtastrate von 44,1 kHz konnte eine Latenz von

$$d_{\rm D}^{\rm min} = 144 \, \text{Abtastwerte} \approx 3,265 \, \text{ms}$$
 (14)

erreicht werden. Dazu wurde eine Rahmengröße von 32 Abtastwerten gewählt und die Prozesspriorität auf "Echtzeit" eingestellt, um Rahmenverluste zu vermeiden.

Für eine qualitative Messung wurde das Rack an eine head unit angeschlossen. Zur quantitativen Bewertung wurden drei mit A, B und C bezeichnete Teilsysteme des in Abbildung 7 gezeigten Gesamtsystems vermessen. Die jeweiligen Betragsfrequenzgänge sind in Abbildung 8 dargestellt. Das Teilsystem A besteht aus dem Lautsprecherverstärker der head unit und dem Leistungskonverter des Simulationsracks, der grün dargestellte Frequenzgang muss softwareseitig ausgeglichen werden. Zum Vergleich ist der Frequenzgang der zur Messung verwendeten RME-Soundkarte in türkis gezeigt, wobei Aus- und Eingang der Soundkarte symmetrisch verbunden wurden. Der Betragsfrequenzgang des Teilsystems B bei unbelegtem Mikrophonausgang  $A_{\rm M}$  ist in blau, bei einer an  $A_{\rm M}$  angeschlossene Tonaderspeisung  $(R_{\rm e}=680\,\Omega)$  in violett dargestellt. Der Vergleich dieser Kurven mit dem Frequenzgang der Soundkarte in



**Abbildung 8:** Betragsfrequenzgänge der Pfade A, B und C sowie Vergleich zur RME-Soundkarte

türkis zeigt die in Gleichung 4 bestimmte Dämpfung um etwa 0,8 dB. Wichtig für die Simulation ist das Teilsystem C, das den Signalweg vom Simulations- zum Sprachverarbeitungssystem beschreibt. Die rote Kurve in Abbildung 8 zeigt den softwareseitig auszugleichenden Betragsfrequenzgang des Teilsystems C. Die orangefarbene Kurve zeigt, dass die gemessene Dämpfung von  $-7,2\,\mathrm{dB}$  zum Großteil durch die asymmetrisch angeschlossenen RME-Soundkarte verursacht wird. Zum Ausgleich der Betragsfrequenzgänge der Teilsysteme A und C wurden mittels linearer Prädiktion minimalphasige Entzerrfilter entworfen.

#### **Fazit**

Es wurde ein System zur digitalen Innenraumsimulation für ICC-Systeme vorgestellt und auf die Aspekte der Simulationssoftware, der Anpasshardware und der Konfiguration eingegangen. Das vorgestellte Simulationssystem ist über eine graphische Benutzungsoberfläche komfortabel konfigurierbar und erreicht eine Latenz von 3,265 ms. Die eigens entworfene Hardware ermöglicht bei einem sehr gutem Störabstand und leicht zu entzerrenden Frequenzgängen den Anschluss von diverser blackbox-Systeme, insbesondere von head units.

### Literatur

- [1] Christian Lüke, G. Schmidt, A. Theiß, J. Withopf: In-Car Communication, in G. Schmidt, H. Abut, K. Takeda, J. Hansen (Hrsg.): Smart Mobile In-Vehicle Systems. Springer, Januar 2014.
- [2] Christian Lüke, A. Theiß, G. Schmidt, O. Niebuhr, T. John: Creation of a Lombard Speech Database using an Acoustic Ambiance Simulation with Loudspeakers. 6<sup>th</sup> Biennial DSP Workshop for In-Vehicle Systems 2013, Seoul, Korea.
- [3] Jürgen Freudenberger, S. Hermanutz, S. Stenzel: A Hardware-in-the-Loop Testing Approach for In-car Communication Systems. 5<sup>th</sup> Biennial Workshop on DSP for In-Vehicle Systems, Kiel, 2011.
- [4] William G. Gardner: Efficient Convolution without Input-Output Delay. J. Audio Engineering Society, Vol. 43, No. 3, März 1995.