

# Modellbasierte Entwicklung adaptiver Beamforming-Algorithmen für die Implementierung auf einem Professional Audio Development Kit (PADK)

## Projektleitung

Prof. Dr.-Ing.  
Thomas Felderhoff

## Kompetenzplattform

Kommunikationstechnik  
und Angewandte  
Signalverarbeitung

## Forschungsschwerpunkt

Process Improvement  
& CAQ

## Zeitraum

2006 – 2008

## Förderung

Fachhochschule  
Dortmund  
Forschungsbudget

## Kontakt

Prof. Dr.-Ing.  
Thomas Felderhoff  
Fachbereich  
Informations- und  
Elektrotechnik  
Institut für  
Informationstechnik  
Fachhochschule  
Dortmund  
Sonnenstraße 96  
44139 Dortmund  
Tel.: (0231) 9112-386  
E-Mail: felderhoff  
@fh-dortmund.de

## Abstract

Die Sprache ist das wichtigste Kommunikationsmedium des Menschen. Dieses Signal wird oft von Störeinflüssen beeinträchtigt, die die Verständlichkeit des Gesprochenen stark reduziert oder gar unmöglich macht. Eine wesentliche Aufgabe der Signalverarbeitung kann es hier sein, das Nutzsignal von den überlagerten Störanteilen zu befreien.

Der Erfolg möglicher Verfahren hängt bei vielen Ansätzen von der Beschaffenheit der Störung ab. Akzeptiert man hingegen, dass Störungen vorhanden sind, so fokussieren sich Beamforming-Ansätze darauf, die Quelle des Nutzsignals zu lokalisieren und diesen Bereich zu verstärken. Vom menschlichen Hören sind wir dieses Phänomen von Kindheit an gewöhnt. Über die beiden an unseren Ohren ankommenden Schallsignale können wir ausgesprochen gut auf die Raumrichtung der Signalquelle schließen.

Durch die Auswertung von Signalen an mindestens zwei räumlich getrennten Orten bildet ein Beamforming-Ansatz diese Lokalisierung nach. Da nicht davon ausgegangen werden kann, dass die Quelle des Nutzsignals, wenn es sich um einen Menschen handelt, statisch an einem Ort verweilt, ist der Beamforming-Algorithmus adaptiv auszulagern, um so die Quelle mit ihren räumlichen Bewegungen verfolgen zu können.

Mathematisch und wissenschaftlich gehören die Adaptionsvorschriften zu den anspruchsvollsten Aufgaben der Digitalen Signalverarbeitung. Entsprechend komplex ist eine praxistaugliche Umsetzung.

## I. Einleitung

Gerade dieser Herausforderung stellt sich der Bereich der Eingebetteten Signalverarbeitung im Institut für Informationstechnik. Hier besteht einerseits der Anspruch, jeden signalverarbeitungstechnischen Ansatz verstehen zu können, und andererseits aus diesem Verständnis heraus eine Implementierung auf Digitalen Signalprozessoren zu entwickeln, die in Echtzeit die Verarbeitung der Signale erlebbar macht.

Motiviert durch den Einsatz von Spracherkennungstechnologie bei der Multimodalen Sitzkiste [3] sollte das einem solchen Erkennen zur Verfügung gestellte Sprachsignal so gut wie eben möglich von allen Störanteilen befreit sein.

In Abbildung 1 ist die Überlagerung eines Nutzsignals mit einem Störsignal schematisch dargestellt.

Dabei wird hier eine eindeutig lokalisierbare Störquelle angenommen. Dies ist im Kraftfahrzeug für die Umgebungsgeräusche keinesfalls gegeben [8], vereinfacht aber die Darstellung eines Störanteils.

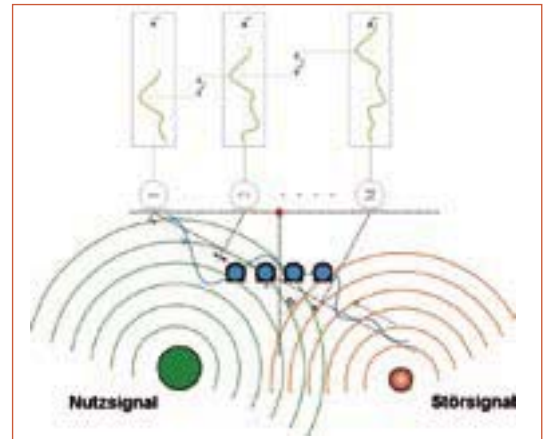


Abb. 1: Szenario der Überlagerung von Nutz- und Störsignal

Der Aspekt lokaler Störsignale an einem bekannten Ort, z.B. die Lautsprechersignale mit dem bekannten bzw. abgreifbaren Audiosignal, wird hier nicht weiter betrachtet. Dort kommen für die Signalverbesserung Echokompensationsalgorithmen [2,7,12] zum Einsatz.

Ferner ist in Abb.1 erkennbar, dass die Schallwellen des Nutzsignals die vier Mikrophone zu unterschiedlichen Zeiten, aufgrund der unterschiedlich langen Ausbreitungswege, erreichen. Solche Zeitverschiebungen werden der Mensch beim Hören zum Orten einer Schallquelle aus, und genau so bestimmt ein Beamforming-Ansatz [1] die Richtung aus der das Schallsignal kommt.

## II. Adaptive Beamforming-Algorithmen

Die Idee der Beamforming-Ansätze ist dem menschlichen Hören und Orten nachempfunden. An den Mikrophenen kommt das Nutzsignal zu unterschiedlichen Zeiten an. Werden die ähnlichen aber zeitlich verschobenen Mikrophonsignale nun jeweils derart verzögert, dass die ursprünglich vorhandenen Laufzeitunterschiede ausgeglichen werden, dann sind die Anteile des Nutzsignals optimal korreliert und eine Überlagerung, also Addition, der einzelnen Anteile führt zu einer Verstärkung.

Zusätzlich sind die Anteile eines Störsignals unkorreliert zum Nutzsignal und bleiben dies auch

nach dem oben beschriebenen Ausgleich der Laufzeitunterschiede, da das Störsignal ja in der Regel andere Laufzeitunterschiede besitzt und diese somit nicht kompensiert werden. Dann findet bei der Überlagerung nicht nur eine Verstärkung des Nutzsignals statt sondern gleichzeitig eine Verbesserung des Nutzsignals zu Störsignal Abstandes.

### 1. Delay & Sum – Beamformer

Aus den gerade beschriebenen Überlegungen zur grundsätzlichen Funktionsweise eines Beamformers kann der in Abb. 2 dargestellte Signalflussgraph direkt abgeleitet werden.

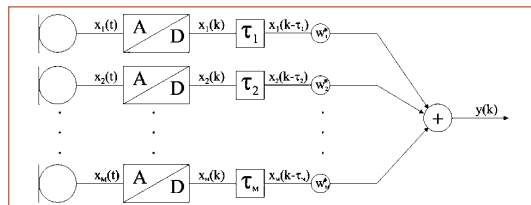


Abb. 2: Signalflussgraph eines Delay & Sum - Beamformers

Die unterschiedlichen Verzögerungen um  $\tau_i$  kompensieren die an den Mikrofonen vorhandenen Laufzeitunterschiede. Nach diesen Verzögerungen sind die Nutzsinalanteile korreliert. Die Gewichtungsfaktoren  $w_i^*$  erlauben für den allgemeinen Fall, die einzelnen Anteile unterschiedlich stark zu gewichten. Da es aufgrund der Ausrichtung eines Mikrophons sein kann, dass dort ein größerer Störanteil enthalten ist, den man in das überlagerte Summensignal nicht übernehmen möchte, kann der entsprechende Gewichtungsfaktor zu null gewählt werden.

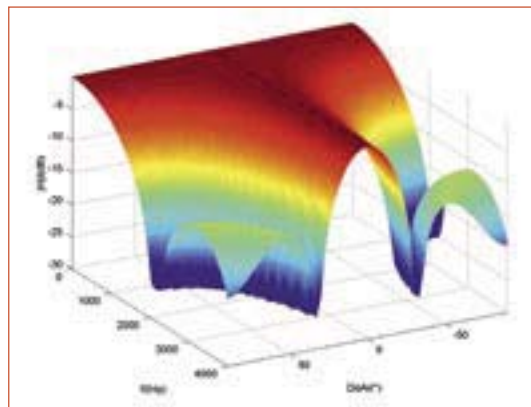


Abb. 3: Auswertung der Richtungsselektion eines Delay & Sum – Beamformers [9]

Eine Analyse, Abbildung 3, des augenscheinlich sehr sinnvollen Ansatzes zeigt in der Tat eine deutliche Richtungsselektion. D.h., über den gesamten Frequenzbereich eines Signals kann die Richtung der Quelle geortet werden. Nachteilig ist aber, dass für niedrige Frequenzanteile keine richtungsabhängige Dämpfung möglich ist. Somit werden alle niederfrequenten Signalanteile ebenso wie das Nutzsinal verstärkt.

Dieses Ergebnis ist für eine reine Ortung einer nicht niederfrequenten Quelle durchaus akzeptabel. Aber für den Einsatz im Kraftfahrzeug ist die nicht vorhandene richtungsabhängige Dämpfung bei niedrigen Frequenzen absolut unbrauchbar. Denn die diffusen Störanteile aus den Fahr-, Abroll- und Motorgeräuschen sind gerade im niederfrequenten Bereich besonders leistungsfähig vorhanden und könnten so nicht geeignet gedämpft werden.

### 2. Filter & Sum – Beamformer

Der grundsätzlich sinnvolle Ansatz eines Laufzeitausgleichs durch die Signalverarbeitungsstruktur bleibt erhalten, wenn anstelle der gewichteten Verzögerungen nichtrekursive Digitalfilter eingesetzt werden. In Abbildung 4 ist der geänderte Signalflussgraph eines Filter & Sum Beamformers dargestellt. Durch die verwendeten Filter ist neben einem Laufzeitausgleich auch die gewünschte Filterwirkung für niedrige Frequenzen möglich.

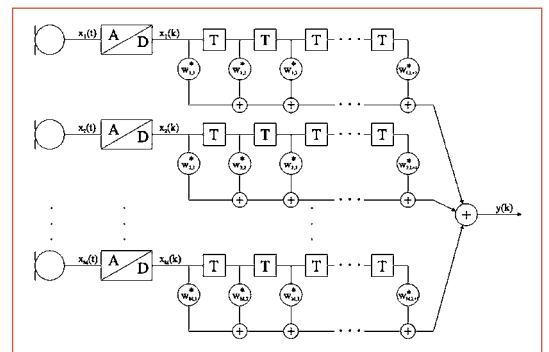


Abb. 4: Signalflussgraph eines Filter & Sum - Beamformers

Abbildung 5 zeigt einen richtungsselektiven Durchlassbereich über einen breiten Frequenzbereich. Damit werden niederfrequente Störanteile, wie gewünscht, aus allen Raumrichtungen gedämpft. Dieser Nutzen wird erkauft durch die Dämpfung des Nutzsignals im niederfrequenten Bereich.

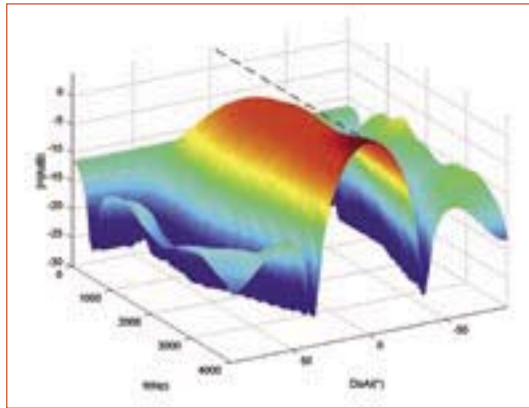


Abb. 5: Auswertung der Richtungsselektion eines Filter & Sum – Beamformers [9]

Insgesamt ist eine Verbesserung des Nutzsinal zu Störsignal Abstandes zu verzeichnen. Dies bedeutet ein weniger gestörtes Nutz-, also Sprachsignal nach der Beamforming-Verarbeitung.

Für die Beurteilung der praktischen Einsatzmöglichkeiten sind zwei weitere Aspekte zu bedenken. Der aufwendigere Filter & Sum Beamformer erfordert eine entsprechend höhere Rechenleistung des zu verwendenden Digitalen Signalprozessors, um eine Verarbeitung in Echtzeit gewährleisten zu können. Außerdem muss der Einfluss auf die Erkennungsrate eines Spracherkenners evaluiert werden, da ja das augenscheinlich verbesserte Sprachsignal niederfrequente Anteile des Sprachsignal nur noch in gedämpfter Form enthält.

Die bei beiden Beamformern verwendeten Gewichtsfaktoren bzw. Filterkoeffizienten werden zeitlich veränderbar ausgelegt, so dass darüber eine Adaption an Richtungsänderungen der Nutzsinalquelle erfolgen kann.

### III. Modellbasierte Entwicklung

Um bei der komplexen Thematik des Beamformings einfacher auf bereits erarbeitete Ergebnisse aufsetzen zu können, wurde die modellbasierte Entwicklungsmethodik konsequent angewendet. Die Firma The MathWorks stellt mit der Entwicklungsumgebung MATLAB/Simulink ein leistungsfähiges Tool zur Verfügung, dass für Aufgaben der Digitalen Signalverarbeitung bestens geeignet ist und zusätzlich die Möglichkeit einer automatischen Codegenerierung für ausgesuchte Mikroprozessoren bietet.

Ausgehend von wissenschaftlichen Veröffentlichungen, z.B. [1,2,7,8,12] wurden entsprechende Modelle der Beamformer in Simulink entwickelt. Abb. 6 stellt das Simulink-Modell eines vierkanaligen Delay & Sum Beamformers dar. Alle zuvor diskutierten Ergebnisse sind aus diesen Simulationen entstanden, wobei die studentischen Arbeiten [4-6,9,10] aufeinander aufbauend eine stetige Weiterentwicklung und Verbesserung zeigen.

Die modellbasierte Entwicklungsmethodik unterstützt die raschere Einarbeitung in bereits erzielte Ergebnisse. Separierte Aufgabenteile können zu einer übergeordneten Lösung zusammengefügt werden.

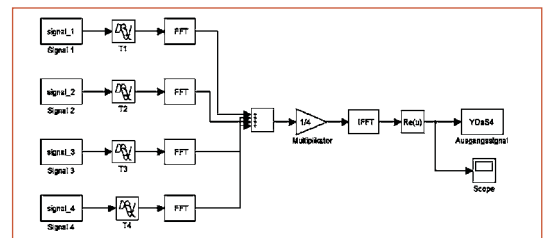


Abb. 6: Modell eines vierkanaligen Delay & Sum – Beamformers in Simulink [6]

### IV. Zielhardware – Target

Die TMS320C6000 Signalprozessorfamilie von Texas Instruments wird bei der automatischen Codegenerierung aus der Entwicklungsumgebung MATLAB/Simulink mit entsprechenden Erweiterungspaketen unterstützt. Die Firma Lyrttech bietet ein Professional Audio Development Kit (PADK) auf Basis des C6727 Signalprozessors an. Durch die vierkanaligen Ein- und Ausgänge ist dieses Board, Abbildung 7, bestens für die Beamforming-Aufgabe geeignet.



Abb. 7: Die Zielhardware: das Professional Audio Development Kit der Firma Lyrttech

Für die vollständige Integration des PADKs in die Methodik der modellbasierten Entwicklung müssen alle Eigenschaften, insbesondere die verfügbaren Schnittstellen, in MATLAB/Simulink verfügbar sein. Entsprechende Modell-Blöcke stehen am Institut für Informationstechnik zur Verfügung [11].

#### **V. Zusammenfassung und Ausblick**

An dem anspruchsvollen Thema des Beamformings kann der Kernkompetenzbereich Eingebettete Signalverarbeitung des Instituts für Informationstechnik durchgehend die modellbasierte Entwicklung verdeutlichen. Dabei werden unterschiedliche Lösungsansätze im Hinblick auf ihre konkrete Praxistauglichkeit, den erforderlichen Aufwand und die Qualität eines Gesamtsystems evaluiert.

Für solch ein Gesamtsystem kann mit der Multimodalen Sitzkiste [3] das Zusammenspiel aus Signalverbesserungsalgorithmen und Spracherkennern optimiert werden.

#### **VI. Danksagung**

Die Ausstattung des Instituts für Informationstechnik mit zwei PADKs ist dank der finanziellen Förderung im Rahmen der hochschulinternen Forschungsförderung möglich gewesen. Hierdurch gewinnen die gesamten wissenschaftlich anspruchsvollen Untersuchungen ein angemessenes Gewicht. Neben überzeugenden Simulationen können so die Ergebnisse auch praxisnah und in Echtzeit demonstriert werden.

#### **VII. Literatur**

- [1] Brandstein, M. und Ward, D., *Microphone Arrays: Signal Processing Techniques and Applications*, Springer, 2001
- [2] Davis, G.M., *Noise Reduction in Speech Applications*, CRC Press, 2002
- [3] Felderhoff, T. und Nennstiel, N., *Multimodale Sitzkiste – Ideenträger in Lehre und Forschung*, Forschungsbericht der FH Dortmund, 2008
- [4] Hör, I. und Klesing, I., *MATLAB-Implementierung eines adaptiven Transversal-Digitalfilters zur lokalen Echo- und Störgeräuschkompensation*, Projektarbeit, FH Dortmund, 2002
- [5] Klysz, S., *Untersuchungen mehrkanaliger Sprachdaten zur robusten Sprachaktivitätserkennung in geräuschvoller Umgebung*, Diplomarbeit, FH Dortmund, 2006
- [6] Kohn, A., *Modellbildung und Simulation am Beispiel mehrkanaliger Audiosignalverarbeitungsmethoden*, Projektarbeit, FH Dortmund, 2006
- [7] Loizou, P.C., *Speech Enhancement – Theory and Practice*, CRC Press, 2007
- [8] Lungwitz, T., *Untersuchungen zur mehrkanaligen adaptiven Geräuschreduktion für die Spracherkennung im Kraftfahrzeug*, Dissertation, 1999
- [9] Pirchner, A., *Mehrkanalige Signalverarbeitung am Beispiel datenunabhängiger Beamformer für richtungselektives Übertragungsverhalten*, Projektarbeit, FH Dortmund, 2003
- [10] Pirchner, A., *Simulation und Untersuchung von adaptiven Beamforming-Algorithmen zur Sprecherlokalisierung im Kraftfahrzeug*, Diplomarbeit, FH Dortmund, 2003
- [11] Shen, J., *Entwicklung eines „Embedded Target for the Professional Audio Development Kit“ zur automatischen Codegenerierung*, Master-Thesis, FH Dortmund, 2008
- [12] Vary, P., Heute, U. und Hess, W., *Digitale Sprachsignalverarbeitung*, Teubner Verlag, 1998