

# Echtzeit-Signalvorverarbeitung zur Störschallreduktion für technische Hörhilfen

## Noise Reduction by Real-time Signal Pre-Processing for Hearing Instruments

Wolfgang H. Döring\*, Joachim Müller-Deile\*\*, Jochen Nicolai\*\*\*/\*\*\*\*, Richard van Hoesel\*\*\*\*

\* Universitätsklinikum Aachen, \*\* Christians Albrecht Universität Kiel, \*\*\* Cochlear AG, Basel

\*\*\*\* CRC for Cochlear Implant and Hearing Aid Innovation, Melbourne. E-mail: wdoering@ukaachen.de

### Einleitung

Das binaurale Hören erlaubt beidohrig normalhörenden Personen ein besseres Sprachverstehen in störschallerfüllter Umgebung („Cocktail Party Effekt“). Für die meisten Hörgeräte- und Cochlear-Implant (CI) -Träger stellt diese Hörsituation jedoch eine große Herausforderung dar, auch wenn durch eine beidseitige Hörgeräteversorgung oder Cochlea-Implantation versucht wird, zumindest eine partielle Nutzung von Fähigkeiten des binauralen Hörens zu ermöglichen. Als Alternative wird in einer laufenden Studie untersucht, in welchem Maße eine Signalvorverarbeitung mit beidseits angebrachten Mikrofonen auch bei nur einseitiger Implantation die auditive Kommunikationsfähigkeit in schwierigen Hörsituationen verbessern kann.

In einem interdisziplinären DFG-Forschungsprojekt an der RWTH-Aachen (Institut für Nachrichtengeräte und Datenverarbeitung / HNO-Klinik) wurden Algorithmen für die Störschallreduktion entwickelt und in Simulationen getestet. Basierend auf diesen Ergebnissen wurde in einem Folgeprojekt ein tragbares Echtzeitsystem aus drei Signalprozessormodulen realisiert, das dem vom Probanden verwendeten Sprachprozessor vorgeschaltet wird. Im Folgenden werden die in der Studie verwendeten Signalvorverarbeitungsverfahren sowie erste Ergebnisse mit 16 einseitig versorgten CI-Trägern in Melbourne und Aachen vorgestellt.

### Grundlagen und Methodik

Die erste Stufe der Echtzeit-Signalvorverarbeitung besteht aus zwei parallel arbeitenden Signalprozessoren, mit denen durch geeignete Filterung und Summation eine 'begrenzte Superdirektivität' der beidseits am Kopf zu tragenden 'Zwei-Mikrofon Endfire-Arrays' erzielt wird (Abb. 1). Durch diese Mikrofonrichtcharakteristik kann die räumliche Verteilung von Nutz- und Störschallquellen zur Verbesserung des Signal-Rauschverhältnisses (SNR) der monauralen Signale der linken und der rechten Seite auch im tieffrequenten Bereich genutzt werden [1]. Die zweite Stufe ermöglicht eine weitere Störfreieung durch die gemeinsame Verarbeitung der beiden vorverarbeiteten monauralen Signale in Form einer zweikanaligen kohärenzbasierten adaptiven spektralen Subtraktion, die die diffusen Signalanteile als Störschall abschwächt [2,3].

Mit Satztests wurde die Verständlichkeit im diffusen Störgeräusch als Funktion des SNR bei CI-Trägern mit und ohne Vorverarbeitung gemessen (Aachen und Melbourne). Darüber hinaus erprobten die Probanden in Aachen das System auch in unterschiedlichen realen akustischen Umgebungen.

Die sprachaudiometrischen Untersuchungen wurden sowohl im Schallfeld (Melbourne, n=5) als auch unter virtuellen akustischen Bedingungen (Aachen, n=11) erprobt. In den untersuchten Hörsituationen wies das Nutzsignal stets ein gerichtetes Schallfeld, das Störsignal (Melbourne: sprachsimulierendes Rauschen, Aachen:

Cafeteria-Lärm) ein diffuses Schallfeld auf. Die Wirkung der Signalvorverarbeitung wurde als Veränderung des notwendigen SNR für 50% Satzverstehen (Melbourne) bzw. auch als prozentualer Verständlichkeitsgewinn (Aachen) gegenüber der Referenzsituation (Verständlichkeit ohne Vorverarbeitung der Signale) gemessen. In Melbourne wurde der HINT-Satztest adaptiv als Nutzsignal eingesetzt, in der virtuellen Cafeteria-Bedingung in Aachen der HSM-Satztest in einem SNR-Bereich von +7,5 dB bis -10,5 dB in 1,5 dB Schritten. In Melbourne wurden die beiden Signalverarbeitungsstufen getrennt und kombiniert bewertet; in Aachen wurde nur die zweikanalige spektrale Subtraktion und das kombinierte Verfahren mit dem Satztest gemessen, aber zusätzlich wurde das System in den realen Situationen 'Büro', 'Treppenhaus', 'im Freien' und 'Cafeteria' erprobt und subjektiv bewertet.

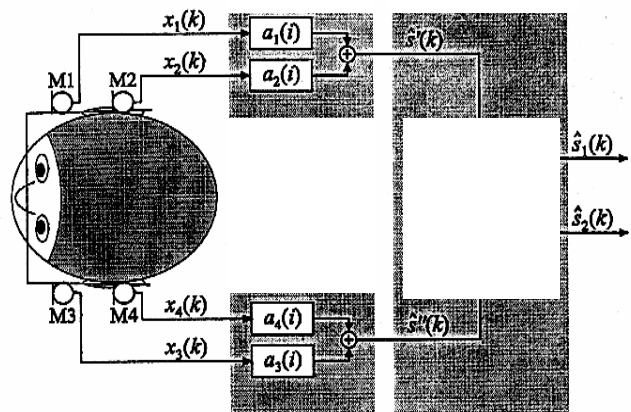


Abbildung 1: Schema der Echtzeit-Signalvorverarbeitung. Die Endfire-Mikrofonarrays M1/M2 und M3/M4 werden beidseits am Kopf getragen und ergeben mit den nachgeschalteten FIR-Filtern (Prozessoren SPEAR1 und SPEAR 2) eine superdirektive Charakteristik. Die anschließende 2-kanalige spektrale Subtraktion mit adaptivem Nachfilter (Prozessor SPEAR 3) hat 2 Ausgänge, die für die monaurale Weiterverarbeitung summiert und dem Sprachprozessor des CI-Trägers zugeführt werden.

Fig. 1: Structure of the real time pre-processing stages. Endfire microphone arrays M1/M2 and M3/M4 are placed on both sides of the head. They are connected to two FIR-filters to form superdirective characteristics using the processors SPEAR 1 and SPEAR 2. The following two-channel spectral subtraction and post-filter stage has 2 outputs. They were summed up for further monaural processing by the CI-user's speech processor.

### Ergebnisse

Die Messungen in Melbourne ergaben als Mittelwert über fünf Probanden eine signifikante Verbesserung des SNR durch jede der beiden Stufen der verwendeten Signalvorverarbeitung und sogar eine hochsignifikante Verbesserung durch deren Kombination. Sowohl durch die superdirektionale Mikrofoncharakteristik als auch durch die spektrale Subtraktion jeweils für sich alleine konn-

ten Verbesserungen von 2,5 bzw. 4,2 dB erreicht werden. Bei Kombination der beiden Verfahren ergab sich eine hoch signifikante Verbesserung von 7,8 dB. Keiner der Probanden verschlechterte sich durch die Aktivierung der Signalvorverarbeitung, wenn auch die erzielten Gewinne stark streuten (Abb. 2).

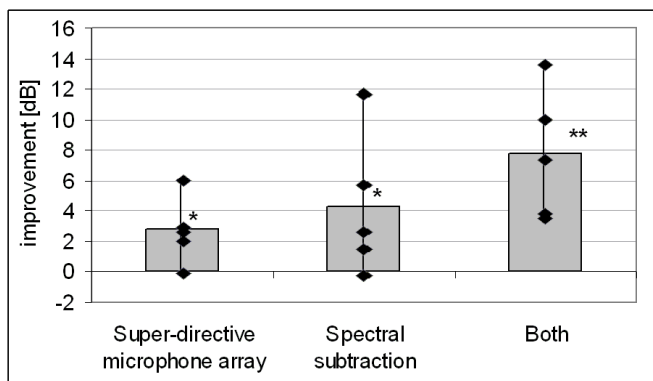


Abbildung 2: Verbesserung des SNR bei Einsatz der Signalvorverarbeitungsstrategien im diffusen Schallfeld. Dargestellt sind der Mittelwert (n=5), sowie die individuellen Messwerte (Melbourne). Als Referenz diente der klinisch genutzte Sprachprozessor ohne Signalvorverarbeitung.

Fig. 2: Improvement of SNR with the use of signal pre-processing strategies in diffuse sound field. The mean value over 5 subjects and the recorded values of the individuals are shown (Melbourne). The clinically used speech processor without preprocessing was used as reference.

Alle 11 Probanden, die in Aachen untersucht wurden, zeigten mit der aktivierten Vorverarbeitung eine Verbesserung der Sprachverständlichkeit in der simulierten Cafeteria-Umgebung von 19% bis 45% (Mittelwert: 30,7%) für die spektrale Subtraktion alleine und von 23% bis 73% (Mittelwert: 43,1%) durch den Einsatz der kombinierten Signalvorverarbeitung (Abb.3).

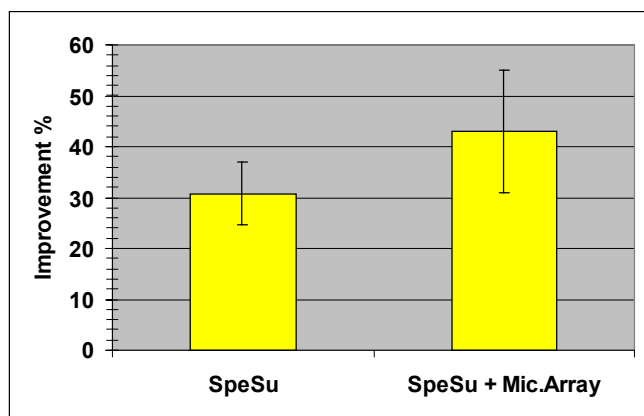


Abbildung 3: Verbesserung des Satzverstehens (Prozentpunkte) in virtueller Cafeteria-Umgebung bei Einsatz der spektralen Subtraktion (SpeSu) und der Kombination der beidseitigen 'Endfire Arrays' mit der zweikanaligen spektralen Subtraktion; Mittelwerte über 11 Probanden (Aachen). Als Referenz diente der klinisch genutzte Sprachprozessor ohne Signalvorverarbeitung.

Fig. 2: Improvement of speech perception in virtual cafeteria environment when using spectral subtraction alone and the combination of the bilateral endfire arrays with the two-channel spectral subtraction; mean values of 11 subjects (Aachen). The clinically used speech processor without preprocessing was used as reference.

Bei 6 der 11 Probanden konnten die vollständigen Diskriminationsfunktionen gemessen werden. Die hieraus ermittelten Verbesserungen des kritischen SNR betragen für die spektrale Subtraktion

alleine 1,4 bis 9,8 dB (Mittel: 4,8 dB, sd 2,9 dB) und für das kombinierte Verfahren 7,5 bis 12,6 dB (Mittel: 9,8 dB, sd 2,0 dB).

Die subjektive Bewertung der Signalvorverarbeitung in realen Hörsituationen (n=14) ergab eine Präferenz für das kombinierte Signalverarbeitungsverfahren in den Hörsituationen 'Cafeteria', 'im Freien' und 'Treppenhaus' (Abb. 4). Lediglich im 'Büro' wurde das superdirektionale Verfahren alleine bevorzugt. Die Ursache hierfür könnte das weniger diffuse Störschallfeld durch wenige diskrete Störquellen wie Pcs, Drucker, Kopierer etc. sein. Vereinzelt wurde hier auch die Einstellung ohne Signalvorverarbeitung favorisiert.

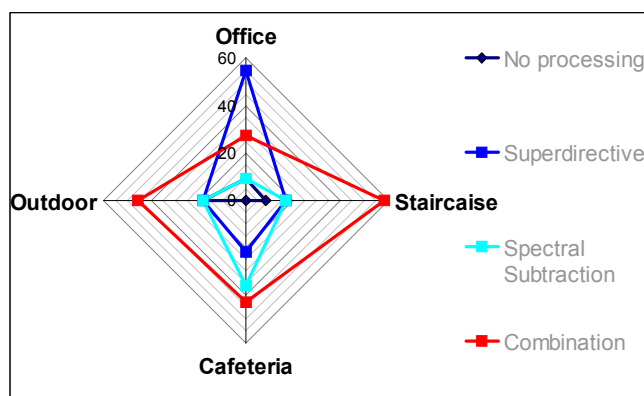


Abbildung 4: Subjektive Präferenz der CI-Träger (n=14) für die einzelnen Signalvorverarbeitungsverfahren in vier realen Hörsituationen. In radialer Richtung ist aufgetragen, welcher Prozentsatz der Probanden sich für das jeweilige Verfahren in der jeweiligen Hörsituation entschieden hat.

Fig. 4: Subjective preference for one of the pre-processing procedures in different real-life acoustical environment. The preferences of 14 CI-users are indicated on the radial percentage axes for each acoustical situation.

## Schlussfolgerung

Eine geeignete bilaterale Signal-Vorverarbeitung kann bei monaural versorgten CI-Trägern das Sprachverstehen in realen diffusen Störgeräusch-Situationen deutlich verbessern und vereinfachen.

In weiteren Untersuchungen soll die Wirkung der vorgestellten Signalvorverarbeitung auch bei bilateral implantierten CI-Trägern ermittelt werden. Die Effektivität der bilateralen Signalvorverarbeitung bei monaural implantierten CI-Trägern kann dann mit der Effektivität der bilateralen Cochlea-Implantation mit und ohne Signalvorverarbeitung verglichen werden.

Die Notwendigkeit einer Datenverbindung zwischen linkem und rechtem Sprachprozessor zur Synchronisation der Verarbeitung muss geprüft und mit den kommerziellen Anforderungen nach hohem Tragekomfort (keine störenden Verbindungskabel am Kopf) und geringen Betriebskosten (Stromverbrauch) in Einklang gebracht werden.

<sup>1</sup>] Dörbecker M (1998) Mehrkanalige Signalverarbeitung zur Verbesserung akustisch gestörter Sprachsignale am Beispiel elektronischer Hörhilfen. Dissertation, Verlag der Augustinus Buchhandlung, Aachen, ISBN 3-86073-439-3

<sup>2</sup>] Ephraim Y, Malah E (1985) Speech Enhancement Using a Minimum Mean-Square Error Lo-Spectral Amplitude Estimator; IEEE Transaction on Acoustics, Speech and Signal Processing, ASSP-33, no.2, 443-445

<sup>3</sup>] M.Dörbecker, S.Ernst; Combination of Two-Channel Spectral Subtraction and Adaptive Wiener Post-Filtering for Noise Reduction and Dereverberation; Proc. EUSIPCO, S.995-998, Sept.1996