Real-Time Implementation of an Adaptive Beamformer-Postfilter System

Jan Bajorat, Christian Lüke, Nicklas Eiselt, Gerhard Schmidt

Christian-Albrechts-Universität zu Kiel, 24143 Kiel, E-Mail: jba/cl/ne/gus@tf.uni-kiel.de

Introduction

The presented system consists of a beamformer and a postfilter. The beamformer is based on the well-known generalized side-lobe canceller structure and features several extensions. The postfilter is a modified Wiener filter. Special emphasis is on extracting information about interfering signals within the side-lobe path and on the frequency-selective control of the adaptive blocking matrix and the interference cancellation. A signal-to-interference ratio (SIR) estimation with maximum and minimum tracking was implemented to allow a frequency-selective adaptation of the blocking matrix. The absence of any desired speaker is reliably detected by a criterion based on a smoothed instantaneous SIR estimate. The postfilter features several extensions compared to the original Wiener filter such as comfort noise injection and adaptive overestimation of noise and interference components in order to improve its performance in real-world applications. Moreover, the modal subspace decomposition (MSD) technique was implemented to improve the control of the beamformer even further. The focus of this paper is on the impact of an implementation of the MSD technique in the fixed part of the beamformer on the control of the adaptive filters.

Overview

The overall system is depicted in figure 1 and is described in detail in [1]. Figure 2 shows an overview of the structure of the beamformer part.

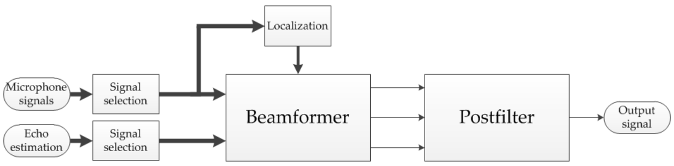

Figure 1: Beamformer-postfilter system.

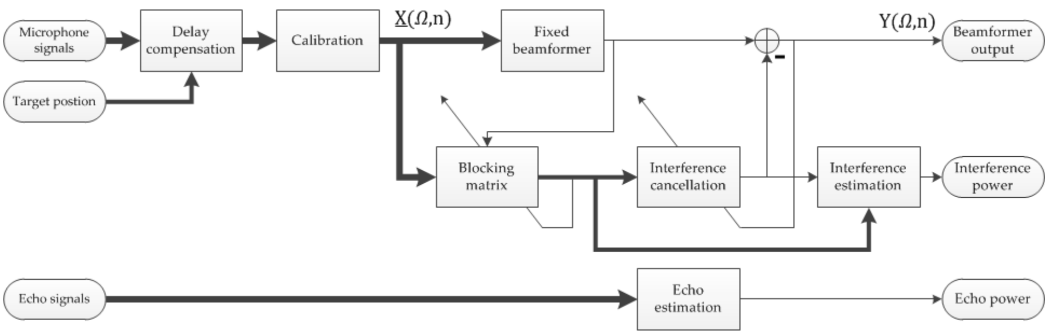

Figure 2: Beamformer.

The entire system is implemented in the frequency domain, which means that it is embedded in FFT-based overlap-add filterbanks. The beamformer system is based on the well-known general side-lobe canceller (GSC) structure. It features a microphone calibration and a frequency selective blocking matrix. The adaptation of both the blocking matrix and the interference cancellation is based on criteria which are described in [1] and [2]. The basic idea is to compare the signal power ratio of the fixed beamformer output and the mean of the blocking matrix outputs with a threshold. If the ratio is above this threshold, a source at the desired direction is assumed to be active. This fixed beamformer is commonly implemented as a simple sum beamformer or a linear constrained minimum variance (LCMV) beamformer. Both are optimum in a certain way: The sum beamformer is optimum regarding the white noise gain, the LCMV beamformer is optimum regarding the SNR of a noisy signal in case the statistics of the noise are known [6]. The implemented MSD beamformer is optimum in approximating a desired beampattern regarding the minimum squared distance [4]. Thus in theory, MSD is optimally suppressing noise coming from certain angles if a desired beampattern is chosen accordingly. Therefore it should improve the control of the adaptive filters in the overall beamformer structure.

Modal Subspace Decomposition (MSD)

An extensive description and analysis of the MSD technique and how to calculate the MSD microphone coefficients can be found in [4]. A brief summary of the design process for the weights can be found in [1] and [3]. Summarized, MSD is based on the idea of calculating a projection of the desired beampattern on the beampattern eigenspace given by the array geometry. The process essentially consists of two steps:

1. Determination of a desired beampattern.
2. Calculation of the fixed beamformer microphone weights using the MSD technique taking into account the desired beam-pattern and the array geometry.

These steps are performed for every frequency band in order to obtain the respective optimum microphone coefficients. These weights are optimum in a way that the resulting beampattern is the best approximation of the desired beampattern in terms of the mean square distance.

Evaluation

Microphone Array

A microphone array consisting of eight omnidirectional microphones was used (figure 4). Seven microphones were arranged as a heptapole array with a radius of 3 cm. The eighth microphone is an extension in a way that a second array, a linear array consisting of four microphones, is available as well. Focus is on the heptapole array because the heptapole geometry is the optimum 2D geometry given a number of seven microphones and a certain maximum distance between the microphones [5].

Matlab Simulation

A Matlab tool which performs the whole design process of the MSD coefficients has been created. Figure 3 depicts some comparison between the MSD and the delay and sum beampattern.

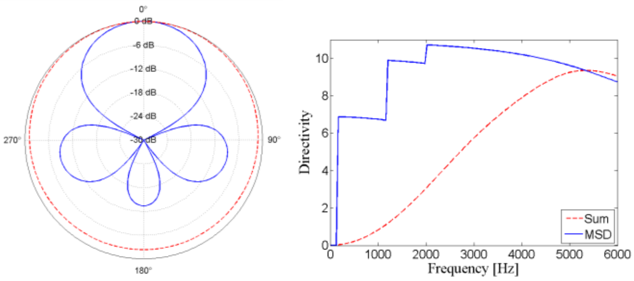

Figure 3: Delay and sum vs. MSD: power pattern at 991 Hz and directivity over frequency.

Real-World Results

The beamformer-postfilter system has been implemented in C and is part of the KiRAT (Kiel Real-Time Audio Toolkit) project. The test setup consists of one desired speaker positioned at an angle of 90° and three interfering speakers at 0°, 180° and 270° as depicted in figure 8.

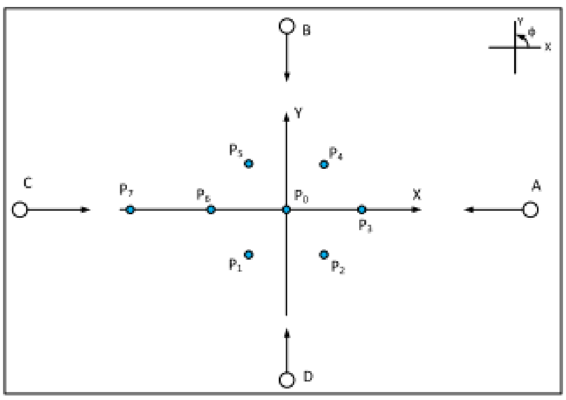

Figure 4: Test setup with four persons (denoted as A-D).

The measurement results shown in figure 5 and table 1 were taken during single-talk situations only. On one hand, informal listening tests show that there is little to no difference in single-talk situations. On the other hand, the tests clearly show that the quality of the output signal during multi-talk situations is improved if the sum beamformer is replaced by a MSD beamformer.

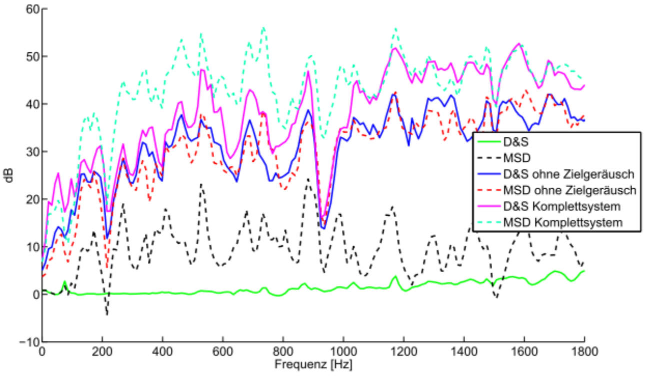

Figure 5: Signal-to-noise attenuation ratio.

Table 1: Noise attenuation for different interfering sources.

Noise attenuation	Position of interferer		
	0°	180°	270°
D&S	44.33 dB	31.06 dB	41.99 dB
MSD	41.34 dB	39.49 dB	41.20 dB

Conclusions

The beamformer-postfilter system is able to suppress the interference during single-talk situations very well. Thus in this scenario the usage of the MSD technique instead of a sum beamformer in the fixed part of the beamformer did not show any significant improvement. In contrast, informal listening tests show a higher quality of the output signal in case of double-talk situations in terms of less distortions of the desired speaker's speech signal as well as of a better suppression of interference. This is most likely due to improvements of the robustness of the control of the adaptive algorithms in the beamformer, especially at lower frequencies. We are working on an objective evaluation method for double-talk situations in order to measure the improvement.

Literature

[1] Bajorat, J.: *Implementierung eines Sprachverbes-serungssystemes mit mehreren Mikrofonen*, Diplomarbeit, CAU Kiel, 2011.

[2] Hänsler, E., Schmidt, G.: *Acoustic Echo and Noise Control – A Practical Approach*, Wiley, 2004.

[3] Eiselt, N.: *Entwicklung und Aufbau mehrdimensionaler Mikrofon-Arrays zur Echtzeit-Sprachverbesserung mittels Beamforming*, Bachelor thesis, CAU Kiel, 2012.

[4] Eichler, M.: *Breitbandige Beamforming-Algorithmen zur Erfassung von Audiosignalen mit kompakten Mikrofon-Arrays*, Dissertation, Goethe-Universität Frankfurt am Main, 2009.

[5] Bevelacqua, P.J.: *Antenna Arrays: Performance Limits and Geometry Optimization*, PhD thesis, Arizona State University, May 2008.

[6] Van Trees, H.L.: *Detection, Estimation, and Modulation Theory Part IV - Optimum Array Processing*, Wiley, 2002.