MULTI-MICROPHONE NOISE REDUCTION BY POST-FILTER AND SUPERDIRECTIVE BEAMFORMER

Joerg Bitzer†, *Klaus Uwe Simmer*‡ *and Karl-Dirk Kammeyer*†

† University of Bremen, FB–1, Dept. of Telecommunications
P.O. Box 330 440, D–28334 Bremen, Germany, `email: bitzer@comm.uni-bremen.de`
‡ Aureca GmbH,
Mozartstrasse 26, D–28203 Bremen, Germany, `email: uwe.simmer@aureca.com`

ABSTRACT

In this contribution a new noise reduction scheme for multi-microphone speech enhancement, called **A**daptive **P**ost-filter **E**xtension for **S**uperdirective beamformers (**APES**), will be introduced, The combination of the superdirective beamformer and the adaptive post-filter has a great potential to suppress diffuse noise. It also outperforms all related algorithms in terms of noise reduction, without increasing signal cancellation. Additionally, a new and easy-to-implement estimator of the post-filter transfer function will be given.

1. INTRODUCTION

Multi-microphone techniques are a growing research field, since beamformers and related techniques have a great potential for noise reduction by using spatial information in adverse environments. However, the array design is often restricted by the dimension of the array. Therefore, solutions with a small array aperture are important. One well-known solution is the superdirective design, but the potential of noise reduction is limited by the number of microphones and the steering direction [1]. On the other hand, post-filter structures have a great potential for noise reduction, provided that the inherent beamformer can reduce the noise [2]. In the following section we will present a new combination of these two algorithms called **A**daptive **P**ost-filter **E**xtension for **S**uperdirective beamformers (**APES**). In section three we will show that the new algorithm outperforms all related algorithms in terms of noise reduction. Additionally, our new scheme produces less signal degradation than other adaptive algorithms.

2. ALGORITHM

The standard implementation of a directivity-controlled version of the post-filter transfer function is given by [2] (for brevity, the frequency variable ω is omitted)

$$\hat{W} = \frac{1}{\Re\left\{\sum_{i=0}^{N-2}\sum_{j=i+1}^{N-1} a_i a_j^*\right\}} \frac{\Re\left\{\sum_{i=0}^{N-2}\sum_{j=i+1}^{N-1} a_i a_j^* P_{X_i X_j}\right\}}{P_{Y_b Y_b}} \qquad (1)$$

where $P_{X_i X_j}$, $P_{Y_b Y_b}$ and N denote the cross power spectral density (CPSD) of the input signal, the auto-PSD of the beamformer output, and the number of microphones, respectively. In small array apertures the delay-and-sum beamformer cannot suppress noise at lower frequencies and therefore, the post-filter cannot suppress noise either. Unfortunately, the coefficients of superdirective beamformers lead to unstable transfer functions W, as shown in [2].

This contribution addresses the design of post-filter structures, where the properties of superdirective beamformers can be used for further noise reduction in the post-filter transfer function. Our new algorithm consists of three parts (see figure 1):

- A standard delay-and-sum beamformer and a post-filter $W_0(\omega)$
- A superdirective extension (lower path)
- A second post-filter $W_1(\omega)$

The post-filter transfer function $W_0(\omega)$ is estimated according to equation (1) and $a_i = 1/N$.

$$\hat{W} = \frac{2}{N^2 - N} \frac{\Re\left\{\sum_{i=0}^{N-2}\sum_{j=i+1}^{N-1} P_{X_i X_j}\right\}}{P_{Y_b Y_b}} \qquad (2)$$

In order to reduce the computational complexity we can rewrite the nominator by using the beamformer output

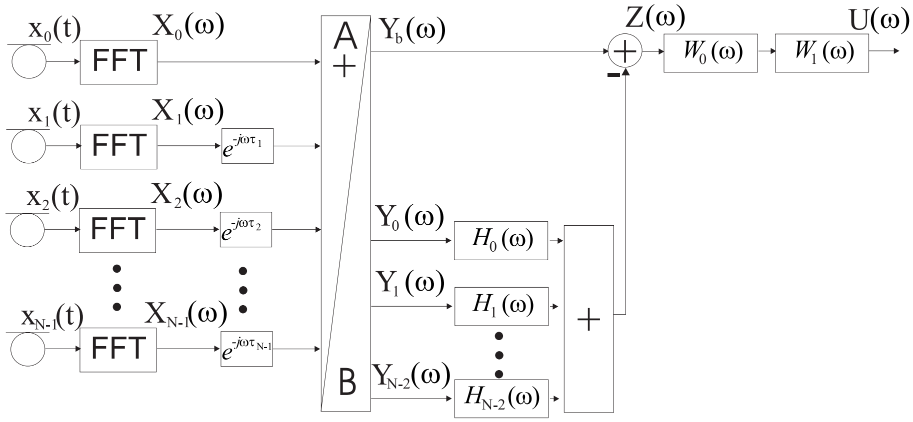

Figure 1: Block diagram of a GSC-like superdirective and post-filter beamformer in a frequency domain implementation

$Y_b(\omega)$. The PSD of $Y_b(\omega)$ is given by

$$P_{Y_bY_b}(\omega) = \frac{1}{N^2}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} X_i(\omega)X_j^*(\omega) \tag{3}$$

$$= \frac{1}{N^2}\sum_{i=0}^{N-1}|X_i(\omega)|^2 + \tag{4}$$

$$\frac{2}{N^2}\Re\left\{\sum_{i=0}^{N-2}\sum_{j=i+1}^{N-1} X_i(\omega)X_j^*(\omega)\right\}$$

and therefore, the double-sum of the numerator can be expressed in terms of the APSDs of the input signal $X_i(\omega)$ and the beamformer output $Y_b(\omega)$.

$$\hat{W}_0 = \frac{N}{N-1}\frac{P_{Y_bY_b} - \frac{1}{N^2}\sum_{i=0}^{N-1} P_{X_iX_i}}{P_{Y_bY_b}} \tag{5}$$

The computational power for the estimation process of the numerator reduces from $\frac{N^2-N}{2}$ to N and we save computational power starting from $N = 3$.
If the aperture of the sensor array is small, this part of the algorithm can only suppress higher frequencies in diffuse or reverberant noise fields [3]. In order to suppress noise at lower frequencies an extension is necessary.
The second part of the algorithm is a superdirective extension of the delay-and-sum beamformer. The mathematical background and the design procedure of this new scheme can be found in [4][1]. The idea can be explained as follows: Cox et al. [1] have shown that the design of superdirective beamformers and Frost's adaptive algorithm [5] are based on the same optimization criterion. Thus, the Frost-algorithm converges to the superdirective beamformer in an isotropic (diffuse) noise field. Furthermore, Buckley [6] has shown that the Frost-algorithm is equivalent to the generalized sidelobe canceller (GSC) [7] under special conditions. Therefore, it is possible to implement the superdirective beamformer in a GSC-like structure having fixed filters in the sidelobe path. These filters can be computed in advance by using the Wiener optimization criterion. The new structure has the advantage of reduced complexity, since the fixed filters H_i are real- or imaginary-valued only. Furthermore, the delay-and-sum beamformer output is available without the additional superdirective part.
The second post-filter can be estimated by

$$\hat{W}_1 = \frac{P_{ZZ}(\omega)}{P_{Y_bY_b}(\omega)} \tag{6}$$

where P_{ZZ} is the PSD of the superdirective beamformer. This transfer function leads to 1 at higher frequencies, as the superdirective beamformer and the delay-and-sum beamformer perform equally at higher frequencies. On the other hand, at lower frequencies the transfer function tends to zero, as the superdirective beamformer suppresses the spatially correlated diffuse noise field in contrast to the delay-and-sum beamformer. Furthermore, the estimation depends on the signal-to-noise ratio (SNR): if the SNR is high, the transfer function tends to 1, and at a low SNR it is close to zero.
The two post-filter transfer functions can be combined by multiplication and the result should be restricted to

$$0.05 \le \hat{W}_0 \cdot \hat{W}_1 \le 1$$

for better speech quality.
Therefore, the complete structure depicted in fig. 1 has three outputs for further extensions: a delay-and-sum beamformer signal, a superdirective beamformer output, and an adaptive broad-band noise reduced output.

3. SIMULATION AND RESULTS

Our simulation system consists of three parts. In a first step the original signals (speech and noise) will be convolved with the room impulse responses (RIR), computed

[1]You can download all papers from `http://www.comm.uni-bremen.de/pub/`

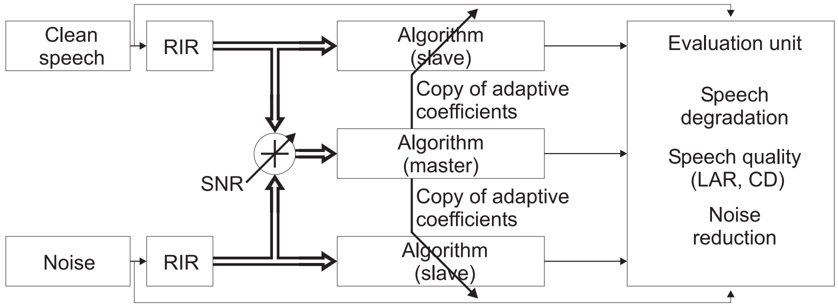

Figure 2: Simulation system

according to the image method by Allen and Berkley. The multichannel signals are mixed to get the desired SNR in a second step. The mixed signal controls the master algorithm, and all adaptive parts are copied to two slave algorithms, which filter only either speech or noise. Therefore, we can compute information on speech degradation and true noise reduction (see figure 2). The advantage of this structure is that signal degradation due to reverberation and artefacts of adaptive algorithms can be computed separately. For example, after reverberation the log-area-ratio distance (LAR) increases, but after processing with any non-adaptive algorithm the LAR will decrease, since the algorithms have some dereverberation effects. In the evaluation unit the following quantities can be computed: LAR, Cepstral Distance (CD), SNR, Noise Reduction (NR) and Critical Band NR (CBNR). All quantities have short-time and global averaged values.

In order to evaluate the performance of our new algorithm (APES) in different situations we simulated two reverberation times ($\tau_{60} = 150ms, 450ms$), two different look-directions (broadside, endfire), and different noises (white gaussian and speech-like). The speech-like noise was generated by white noise which was filtered by a Butterworth-filter with the coefficients:
$B = [0.045 \quad 0 \quad$ -0.045], $A = [1 \quad -1.893 \quad 0.91]$ and $f_s = 8kHz$.

We compared the new algorithm (APES) with

- a Delay-and Sum-Beamformer (D&S),
- a Standard Adaptive Post-Filter [8] (APF),
- a Superdirective-Beamformer (SB),
- and a one-channel solution named the Ephraim and Malah MMSE-logSTSA Estimator [9] (EM).

Figure 3 shows the noise reduction performance for $\tau_{60} = 450ms$, white noise, and endfire steering direction. We can see that APES outperforms all other algorithms. Furthermore, all algorithms improve the speech quality at low SNRs (see figure 4), but the EM-algorithm degrades the signals at higher SNRs.

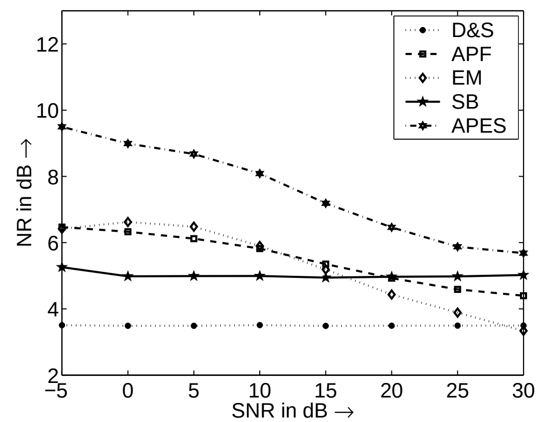

Figure 3: Noise reduction vs. SNR ($N = 4$, $\tau_{60} = 450ms$)

Figure 5 illustrates the signal degradation due to adaptive processing. The non-adaptive algorithms (D&S, SB) increase the speech signal quality due to their dereverberation properties. In contrast, all adaptive schemes (APF, EM, APES) degenerate the speech signal by signal cancellation, whereas our new algorithm produces less cancellation and less musical tones over a wide range of simulated SNRs.

In figure 6 some spectrograms are depicted to give an idea of the noise reduction over frequency.[2]

In a second experiment we changed the noise type to speech-like. In this case the noise reduction performance degrades. Especially, D&S and APF cannot reduce the main noise power at lower frequencies. On the other hand, the EM-algorithm is independent of the noise spectrum and has a good noise reduction behaviour. SB and APES are the only multi-microphone algorithms that reduce the noise and improve the speech quality for all SNRs. In order to evaluate the audible noise reduction we examined the noise reduction in the critical bands (CBNR) of the ear (16 bands) and averaged over all bands in order to get more realistic results (The exact results are not shown here, due to space constraints).

[2] Sound-files can be downloaded from `http://www.comm.uni-bremen.de/whomes/meyer`.

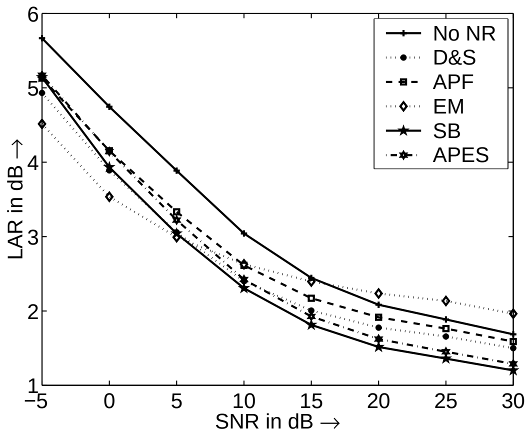

Figure 4: Log area ratio distance vs. SNR ($N = 4$, $\tau_{60} = 450ms$)

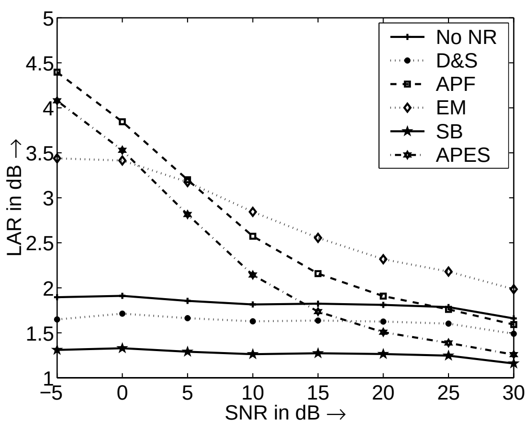

Figure 5: Signal degradation measured in LAR ($N = 4$, $\tau_{60} = 450ms$)

4. CONCLUSION

In this paper we have shown that it is possible to combine a superdirective beamformer with an adaptive post-filter. The new algorithm outperforms all related techniques without producing more artefacts than other adaptive techniques. Furthermore, we have introduced a new simulation scheme that allows to measure the perfor mance of adaptive algorithms in terms of SNR-enhancement, speech quality and signal degradation.

5. REFERENCES

[1] H. Cox, R. M. Zeskind, and M. M. Owen, "Robust adaptive beamforming," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 35, pp. 1365–1375, Oct 1987.

[2] C. Marro, Y. Mahieux, and K. U. Simmer, "Analysis of Noise Reduction and Dereverberation Techniques Based on Microphone Arrays with Postfiltering," *IEEE Trans. on Speech and Audio Processing*, vol. 6, pp. 240–259, May 1998.

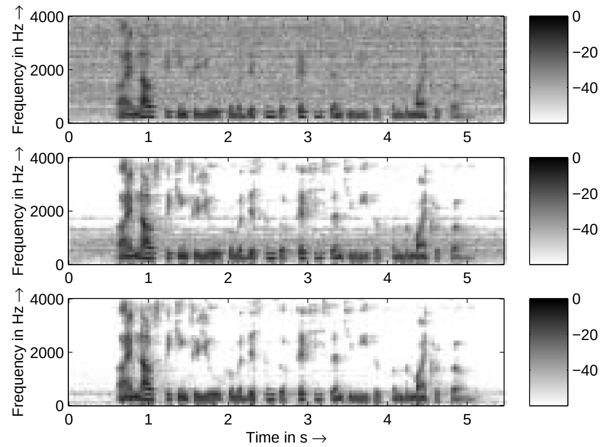

Figure 6: Example of the performance of the new algorithm (SNR = 10dB, N=4) (top: noisy signal, middle: delay and sum + post-filter, bottom: superdirective + W_0 + W_1)

[3] J. Bitzer, K. U. Simmer, and K. D. Kammeyer, "Multichannel Noise Reduction –Algorithms and Theoretical Limits–," in *Proc. EURASIP European Signal Proc. Conference (EUSIPCO)*, vol. 1, (Rhodes, Greece), pp. 105–108, Sep 1998.

[4] J. Bitzer, K. U. Simmer, and K. D. Kammeyer, "An alternative implementation of the superdirective beamformer," in *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, (New Paltz, New York), Oct 1999. accepted.

[5] O. L. Frost, "An algorithm for linearly constrained adaptive array processing," *Proceedings of the IEEE*, vol. 60, pp. 926–935, August 1972.

[6] K. M. Buckley, "Broad-band beamforming and the generalized sidelobe canceller," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 34, pp. 1322–1323, Oct 1986.

[7] L. J. Griffiths and C. W. Jim, "An Alternative Approach to Linearly Constrained Adaptive Beamforming," *IEEE Trans. on Antennas Propagation*, vol. 30, pp. 27–34, 1982.

[8] K. U. Simmer and A. Wasiljeff, "Adaptive Microphone Arrays for Noise Suppression in the Frequency Domain," in *Second Cost 229 Workshop on adaptive Algorithms in Communications*, (Bordeaux, France), pp. 185–194, October 1992.

[9] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean–square error log-spectral amplitude estimator," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 33, no. 2, pp. 443–445, 1985.