

NOISE CANCELLING FOR MICROPHONE ARRAYS

Jens Meyer¹

Carsten Sydow²

¹Institute of Telecommunications and Electroacoustics
Darmstadt University of Technology
64283 Darmstadt, Germany
J.Meyer@uet.th-darmstadt.de

²SIEMENS AG
P.O. Box 80 17 09, 81617 Munich, Germany

ABSTRACT

In this paper an application of the noise cancelling method for suppression of noise of a microphone array system is discussed. First an overview of the noise cancelling approach is given. This is followed by a description of the employment of the method in a realized microphone array system. The limiting factors are described and theoretical limits of the noise suppression are derived. Experimental results, which are obtained in a realistic environment, are presented. The results show that depending on the recording situation the noise cancelling approach applied to a microphone array system leads to a significant enhancement of the signal to noise ratio of the array output signal.

1. INTRODUCTION

With microphone arrays a good signal-to-noise ratio can be achieved when recording speech signals without the constraints of hand-held or body-worn microphones. Due to the directivity of the array, diffuse noise and reverberation are attenuated compared to the direct sound.

If the noise suppression is not sufficient, the directivity index can be increased to achieve a higher signal-to-noise ratio. An increase of the directivity index of a microphone array by 3 dB requires twice the number of microphones and signal processing hardware. Thus even for a small improvement of the noise suppression of 3 dB a great expense in additional hardware is required.

Noise suppression without increasing the number of microphones can be achieved by post-processing the output signal of the array with noise reduction methods like the noise cancelling method [1], the spectral subtraction method [2] or a frequency selective attenuation [3].

As mentioned above this paper will focus on the noise cancelling approach.

2. THE NOISE CANCELLING METHOD

The concept of the two channel noise cancelling approach described by Widrow [1] is depicted in Fig 1. The desired signal $s(i)$ is corrupted by additive noise $n_1(i)$ which is a filtered version of the original noise signal $n(i)$. The resulting signal $d(i)$ forms the primary input. The filter impulse response $h(i)$ is unknown and may be time variant. In addition to the noisy signal on the primary channel, the original noise signal is available at the 'reference input' of the noise cancelling system. Now the unknown system $h(i)$ is esti-

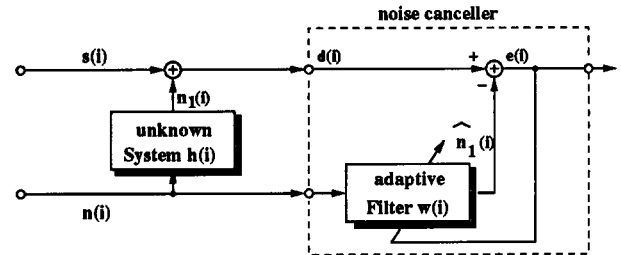


Figure 1. Concept of the noise cancelling method

mated with an adaptive FIR-filter $w(i)$. By filtering the noise signal $n(i)$ with this impulse response, an estimation of the signal $n_1(i)$ is generated. This estimated noise signal $\hat{n}_1(i)$ is subtracted from the signal $d(i)$. The better the estimation of $w(i)$, the closer $n_1(i)$ and $\hat{n}_1(i)$ will be and the less noise will remain in the output signal $e(i)$. If $s(i)$ and $n(i)$ are stationary and uncorrelated signals the filter coefficients of $w(i)$ are best adjusted if the output power is minimized [1]. This also yields the maximum achievable noise reduction. For a time varying unknown system $h(i)$ the filter $w(i)$ has to be adaptive. In the system presented in this paper an NLMS-algorithm (normalized least mean square) is employed to adapt the filter coefficients.

The following problems can arise with the noise cancelling approach:

1. If the filter $w(i)$ is shorter than $h(i)$ the noise reduction is limited [4].
2. If the unknown system $h(i)$ changes faster than the filter $w(i)$ can adapt, only little noise suppression is achieved.
3. If portions of the desired signal are picked up in the reference channel, the desired signal is treated as noise and the system tries to cancel it out.

3. EMPLOYMENT OF THE METHOD

A linear microphone array with 16 cardioid microphones is used to form a steerable directivity pattern to capture the signals for the noise canceller. The steering is performed by switching between different directivity patterns which have their maximum sensitivity at broadside (0°), $\pm 15^\circ$, $\pm 30^\circ$ and $\pm 46^\circ$ respectively. The patterns are formed after the method described in [6] with the request of a beam width between zeros of approximately 40° and a minimum side

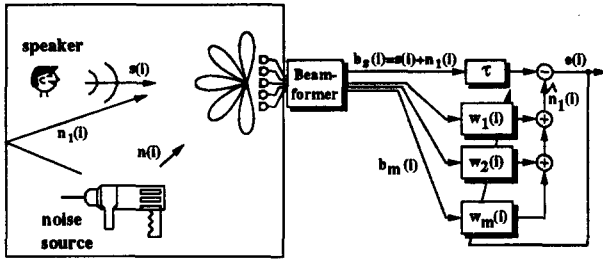


Figure 2. Use of the noise cancelling method in the microphone array system

lobe attenuation of about 25 dB. The pattern overlap in their 3 dB attenuation points. In the frequency range from 500 Hz to 4 kHz the directivity pattern is nearly frequency independent.

To get the signal for the primary channel the array is steered towards the signal source, usually a human speaker. Due to reflections in the room and the limited side-lobe level of the directional response of the array, noise generated by a noise source and filtered by a room impulse response is also picked up by this beam. The signal for the reference channel is picked up by the array steered towards the noise source. If there is more than one noise source, other beams can be used to form additional reference signals. The resulting setup is shown in Fig. 2.

In the signal path a delay τ is introduced to approximate the unconstrained solution for the noise canceller, which requires a noncausal filter ([1]).

4. THEORETICAL LIMITS

4.1. Limited Filter Length

The infinite impulse response $h(i)$ of a room is adapted by the finite length filter $w(i)$. This leads to an error in the output signal of

$$e(i, N) = \sum_{l=N}^{\infty} h(l)n(i-l)$$

assuming that the first N coefficients are adapted ideally.

For white noise as a noise source the influence of the filter length N to the signal-to-noise enhancement (SNRE) of the system can be estimated according to [5]:

$$\text{SNRE} = \frac{N}{T_{60} f_s} 60 \text{ dB} \quad (1)$$

(T_{60} : reverberation time of the room, f_s : sampling frequency). The SNRE is the ratio of the SNR at the output of the noise canceller to the SNR at the primary input expressed in dB. It is used as the quality factor of the system.

For a filter length of 500 taps and with a sampling frequency of 8 kHz a SNRE of 7.5 dB can be expected in a room with $T_{60} = 0.5$ s.

4.2. Time Variations of the Room Impulse Response

The achievable noise reduction is further limited by the time variance of the room impulse response. This variation is

caused for instance by the movement of the noise source or changes in the room environment (e.g. opening of a door or variations of the room temperature). The room impulse response is very sensitive even to small movements of the signal source. Measurements have shown that a displacement of the signal source of 5 cm causes significant changes in the room impulse response.

4.3. Cancellation of the Desired Signal

Mainly by room reflections the desired signal $s(i)$ is picked up by the beam pointing towards the noise source. This results in partial cancellation of the desired signal, thus reducing the SNRE. Therefore the delay τ in the primary channel should be chosen to be smaller than the time between the arrival of the direct sound and the first reverberations. Fig. 3 shows the compensation of the speech signal as a function of the delay τ for the conference room setup described later. The filter length was 4000. The compensation

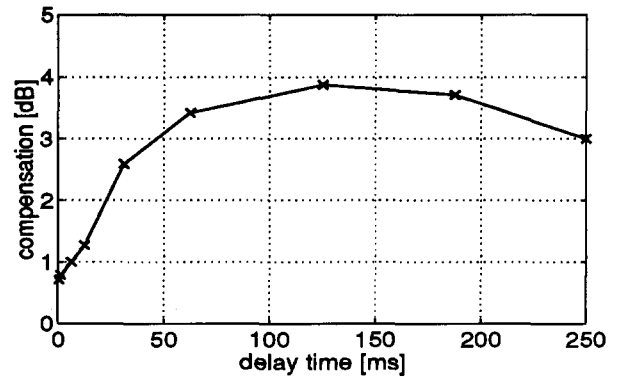


Figure 3. Measured compensation of the desired signal as a function of the delay time τ

decreases for large delay times because the speech components in the primary channel and in the reference channel will be increasingly decorrelated.

5. EXPERIMENTAL RESULTS

5.1. Conference Room Environment

For this experiment the array was set up in a conference room (reverberation time $T_{60} = 0.5$ s) as shown in Fig. 4. The achieved noise reduction as a function of the filter length is shown in Fig. 5. The desired signal is the signal of human speaker referred to as speaker 1. The experiment was repeated for three different noise sources: White noise, which was radiated by a loudspeaker, a second human speaker, referred to as speaker 2 and a drill.

The maximum enhancement of 5 dB is achieved with white noise as the noise signal. For the two other noise sources a SNRE of 2 dB and 1.5 dB is achieved. This difference can be explained by the time variance of the room impulse response for moving sources and the adaption ability of the NLMS algorithm for different input signals. While the white noise is radiated by a fixed loudspeaker the two other sources perform small movements, due to unavoidable natural hand or head motions. These movements cause the room impulse response $h(i)$ to vary with time. The time

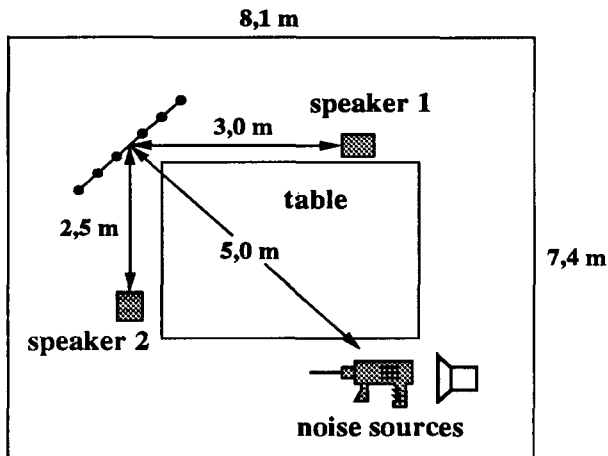


Figure 4. Position of array and sound sources in the conference room (height=2.7 m)

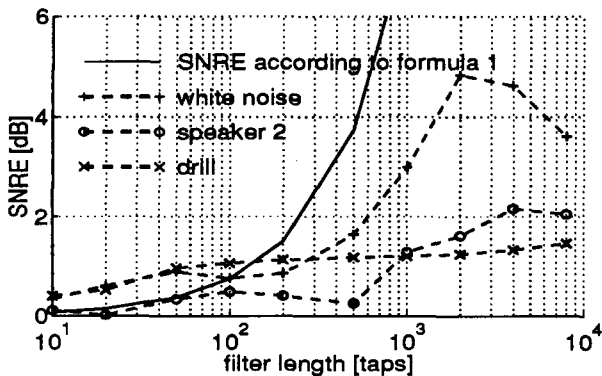


Figure 5. Noise suppression as a function of the filter length

variance is so fast, that the adaptive filter can not follow sufficiently to achieve a better SNRE. Furthermore the system adapts better to the excitation with white noise than to instationary signals like speech or the sound of a drill.

For the white noise signal and filter length below 2000 taps, the SNRE rises with rising number of taps as indicated by equation 1. Due to the fact that the recorded signals are of finite duration, for higher filter length the overall time is not sufficient for the filter to adapt and the SNRE is reduced. If longer recordings were made the SNRE is expected to increase with increasing filter length.

When the distance between the loudspeaker radiating white noise and the array is reduced to 80 cm a maximum enhancement of 8 dB is achieved. The noise-to-signal ratio increase at the reference channel improves the noise cancellation [1].

To evaluate the performance of the noise canceller to narrow band signals a sine wave of 1 kHz was radiated from the loudspeaker. The result was an enhancement of more than 15 dB even at a filter length of 100 taps.

6. SUMMARY AND CONCLUSIONS

By employing the noise cancelling method for the output signal of a microphone array, the signal-to-noise ratio of the output signal can be significantly increased without an increase in the number of microphones, with the improvements depending on the noise source.

For locally fixed noise sources radiating white noise an enhancement of the SNR of 5 dB to 8 dB is achieved, varying with the distance between the noise source and the array. For narrow band signals the SNR enhancement is more than 15 dB.

Only small enhancements could be achieved for noise sources performing small movements. The major reason for this is the time variance of the impulse responses.

REFERENCES

- [1] Widrow, B. et al.: Adaptive Noise Cancelling: Principles and Applications. *Proc. of the IEEE* 63 (1975), pp. 1692-1716
- [2] Peter Vary: Noise suppression by spectral magnitude estimation - mechanism and theoretical limits. *Signal Processing* 8 (1985), pp. 387-400
- [3] R. Zelinski: Noise reduction based on microphone array with LMS adaptive postfiltering. *Electronics Letters* 26 (1975), pp. 1692-1716
- [4] Rainer D. Poltmann: Stochastic gradient algorithm for system identification using adaptive FIR-filters with too low number of coefficients. *IEEE Trans. on Circuits and Systems* 35 (1988), pp. 247-250
- [5] Carsten Sydow: Selfsteering Microphone Array Systems. *PhD. Thesis, Darmstadt 1995*, VDI Forschungsberichte, Reihe 10, Nr.429
- [6] Carsten Sydow: Broadband beamforming for a microphone array. *JASA* 96 (1994), pp.845-849

