

SIMULTANEOUS ECHO CANCELLATION AND CAR NOISE SUPPRESSION EMPLOYING A MICROPHONE ARRAY

Mattias Dahl and Ingvar Claesson and Sven Nordebo

University of Karlskrona/Ronneby
Department of Signal Processing
S-372 25 Ronneby, Sweden
E-Mail: Array@isb.hk-r.se

ABSTRACT

This paper presents a method to simultaneously perform 20 dB acoustic echo cancellation and 15-20 dB speech enhancement using an adaptive microphone array combined with spectral subtraction. Primarily intended for handsfree telephones in automobiles, the microphone array system simultaneously emphasizes the near-end talker and suppresses the handsfree loudspeaker and the broadband car noise. The array system is based on a fast and efficient on-site calibration and can be used in other situations such as conventional speaker phones.

1. SUMMARY

The noisy environment in a car is known to severely degrade performance of handsfree mobile telephones and speech recognition devices relying on a priori knowledge of human speech, requiring a fair Signal-to-Interference-Noise Ratio. Methods to improve the quality using single microphone or array techniques are presented in [1]-[7]. Furthermore, acoustic echoes picked up by microphones are normally difficult to reduce and will often cause perceptual problems at the far-end side. To avoid these problems the proposed method employs a self-calibrating microphone array system [8]. This array adapts in each situation to the actual environment and the electronic equipment by using digital calibration data gathered on-site in advance. The background noise can be further reduced with minimal loss in signal quality by using a subsequent spectral subtraction algorithm Fig. 1.

This work was supported by The Swedish National Board for Industrial and Technical Development.

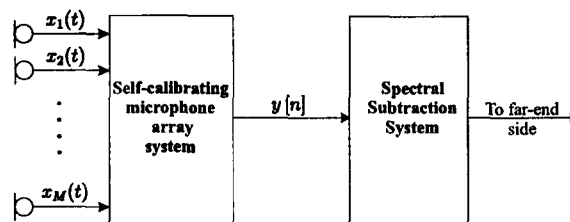


Figure 1: Operating phase, lower beamformer adaptation using the stored jammer and target calibration signals and the incoming car cabin noise. The upper beamformer producing the output to the spectral subtraction.

2. WORKING SCHEME

The system is aimed at handsfree telephones in automobiles and therefore it takes into account the near field in a small enclosure. A near field and enclosed situations is difficult to describe in an a priori model and this is the reason for employing gathered signals from the real jammer and target positions, i.e. the handsfree loudspeaker and near-end talker in the car. These signals contain useful information about the the acoustic environment in the car cabin and the electronic equipment, such as microphones, amplifiers, A/D-converters and antialiasing filters etc.

The system is based upon adaptive FIR-filters at each microphone controlled by the LMS algorithm. The adaptive filters are continuously reducing the influence of acoustic echoes and ambient car cabin noise e.g. road, tires, engine and fan. The cancellation of ambient car noise and the handsfree loudspeaker is thus both in the spatial and frequency domains by combining the linear array system and the non-linear spectral subtraction.

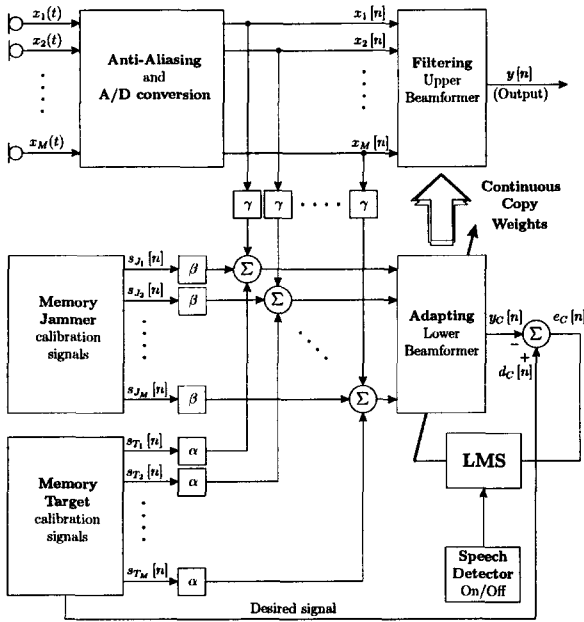


Figure 2: Operating phase, lower beamformer adaptation using the stored jammer and target calibration signals and the incoming car cabin noise. The upper beamformer producing the output to the spectral subtraction.

3. ON-SITE CALIBRATION

The on-site calibration for the microphone array system take place in the actual environment by emitting a representative sequence from the handsfree loudspeaker and from the driver's seat with the car parked. This can be performed by letting the handsfree loudspeaker, i.e. the unwanted source, emit colored noise and let the near-end talker read a representative sequence from the driver's seat in the car storing the signals in jammer and target memories, c.f. Fig 2. In this way array signals with a fair signal to noise ratio (SNR) for each sequence and channel are stored in a digital memory for subsequent use. The calibration signals should have approximately the same spectral content as the real far-end and near-end speaker. The number, placement and type of microphone elements can in practice be chosen arbitrarily, but should not be altered or moved unless a new calibration is made.

4. OPERATING PHASE

During the operation phase noise impinges on each of the microphone element in the array. This is a suitable time to obtain a good noise estimate and indirectly calibrate the microphone array system by means of the real incoming noise and the stored signals from the on-

site calibration, see Fig 2.

The incoming car cabin noise, the stored virtual handsfree signals and virtual talker signals are mixed to comprise the inputs to the adaptive filters in the lower beamformer. The stored virtual handsfree signal represents the handsfree loudspeaker in the cabin and the stored virtual near-end speaker signal represents a person sitting in the driver's seat. Note that the virtual stored handsfree and near-end speaker signals have passed through the same electronic equipment as the real impinging noise signals, one at a time, under conditions of almost no background noise.

The coefficients in the lower beamformer are continuously copied into a fixed upper beamformer that produces the output to the subsequent spectral subtraction algorithm. Note that the inputs to this beamformer contain only signals coming from the microphone elements. In general the amount of noise reduction by using spectral subtraction depends upon the type of background noise, the actual signal to noise ratio and the allowable loss in the quality of the signal. The output from the pre-processing microphone array system yields a fair signal to noise ratio and low distortion, the residual broadband cabin noise can therefore be dramatically reduced with minimal loss in signal quality by using spectral subtraction.

5. PERFORMANCE RESULTS

The performance evaluation of the noise and echo canceller has been carried out in a Volvo 940 GL station wagon, and recordings on a multichannel DAT-recorder with a sample rate at 12 kHz, and with 5 kHz bandwidth were gathered. The microphones were mounted as a linear microphone array below the visor with a fixture. We used six microphones and the distance between two adjacent microphone elements were 50 mm, which yields a total aperture of the array of 250 mm.

Typical results from the Volvo are illustrated in Figs. 3-4. Note that the name of the corresponding wav-sound files is in the upper right corner of each plot. Speech coming from the near-end speaker in the car is denoted "Speaker" and the unwanted far-end speech residual originating from the handsfree loudspeaker is denoted "Echo". The result is presented as a short time (20 ms) power estimate in dB and the outputs are restricted to telephone bandwidth 300-3400 Hz. The figure illustrates the improvement of the noise reduction improves the SNR with 10 dB and the echo cancelling is about 20 dB. By using spectral subtraction after the array system the influence of car cabin noise can be further reduced with about 10 dB which yields a total SNR improvement with 20 dB.

6. REFERENCES

- [1] Y. Kaneda J. Ohga, "Adaptive Microphone-Array System for Noise Reduction" **IEEE Transactions on Acoustics, Speech, and Signal Processing**, vol. ASSP-34, no. 6, pp. 1391-1400, December 1986.
- [2] S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction", **IEEE Trans. on Acoustics, Speech and Signal Processing**, vol. 27, Apr. 1979.
- [3] J. R. Deller jr., J. G. Proakis, J. H. L. Hansen, **Discrete-Time Processing of Speech Signals**, Macmillan, 1993.
- [4] J. L. Flanagan, D. A. Berkley, G. W. Elko, J. E. West, M. M. Sondhi, "Autodirective Microphone Systems", **ACOUSTICA**, vol. 73, pp. 1629-1636, March 1993.
- [5] G.W. Elko, T. C. Chou, R. J. Lustberg, M. M. Goodwin, "A Constant-Directivity Beamforming Microphone Array", **Proceedings of the 128th Meeting of the Acoustical Society of America**, Austin TX, in JASA, vol. 96, no. 5, pt. 2, p. 3244, November 1994.
- [6] Q. Lin, C. W. Che, J. L. Flanagan, "Robust Hands-Free Speech Recognition", **Proceedings of the 128th Meeting of the Acoustical Society of America**, Austin TX, in JASA, vol. 96, no. 5, pt.2, p. 3244, November 1994.
- [7] S. Nordholm, I. Claesson, B. Bengtsson, P. Eriksson, "A Multi-DSP Implementation of a Broad-Band Adaptive Beamformer for Use in a Hands-Free Mobile Radio Telephone", **IEEE Transactions on Vehicular Technology**, vol. 40, Feb. 1991.
- [8] I. Claesson, S. Nordebo, S. Nordholm, M. Dahl, "An "in situ" Calibrated Adaptive Microphone Array", submitted for publication in **The Journal of the Acoustical Society of America**, March 1995.
- [9] M. Dahl, I. Claesson, "Acoustic Echo Cancelling with Microphone Arrays", **Research Report**, 2/95, ISSN 1103-1581

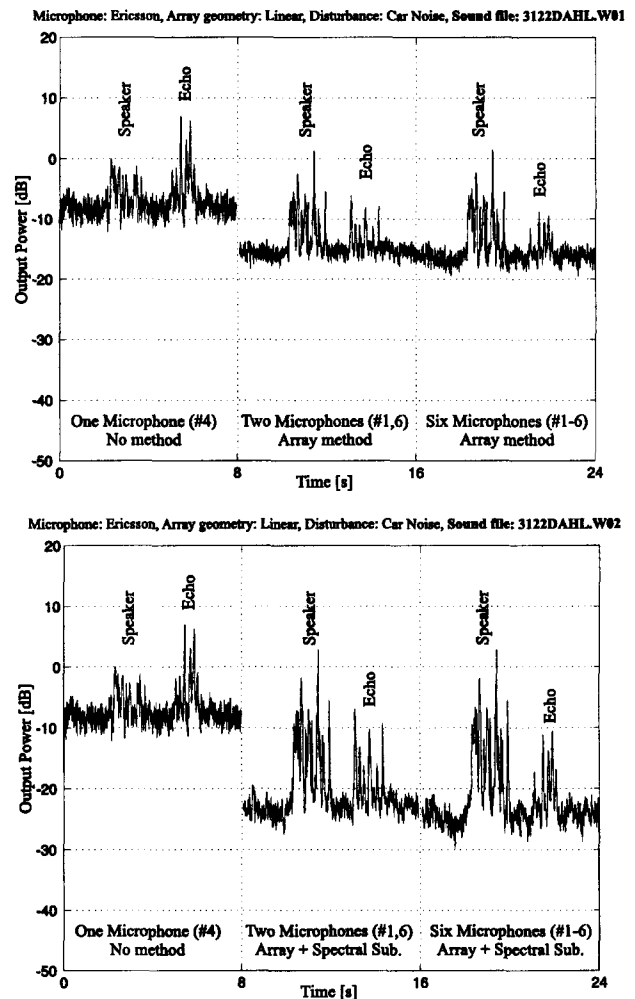
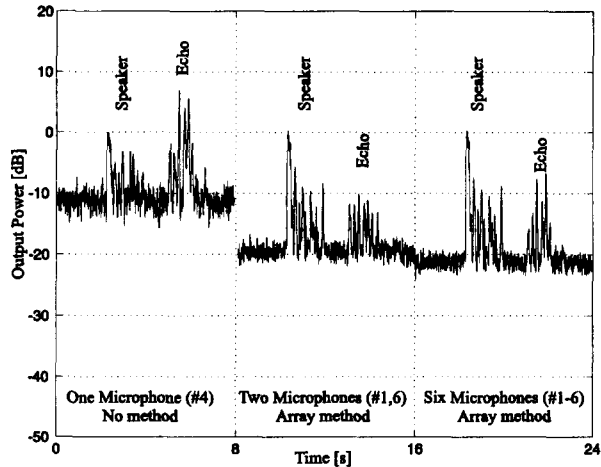


Figure 3: Microphone Ericsson: Echo and noise cancelling, using one microphone (0-8 sec) vs. using an on-site calibrated echo and noise cancelling array with two (8-16 sec) and six microphones (16-24 sec). The two figures illustrates how the system works with and without the subsequent spectral subtraction system.

Microphone: Sennheiser, Array geometry: Linear, Disturbance: Car Noise, Sound file: 3122DAHL.W83



Microphone: Sennheiser, Array geometry: Linear, Disturbance: Car Noise, Sound file: 3122DAHL.W84

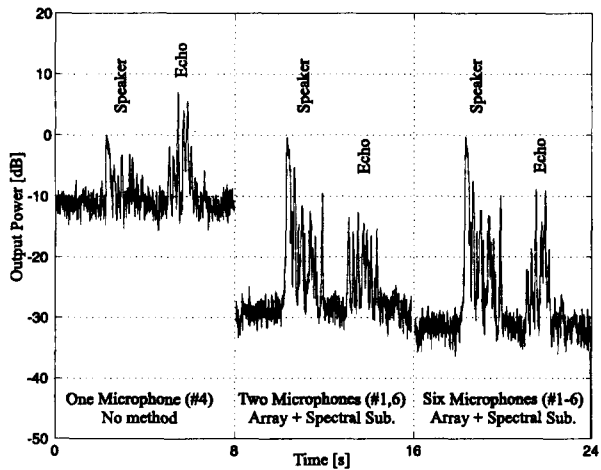


Figure 4: Microphone Sennheiser: Echo and noise cancelling, using one microphone (0-8 sec) vs. using an on-site calibrated echo and noise cancelling array with two (8-16 sec) and six microphones (16-24 sec). The two figures illustrates how the system works with and without the subsequent spectral subtraction system.