Realtime Performance of Acoustic Echo Canceler and Postfilter for Residual Echo Suppression in the Car Environment*

Gerald Enzner[1], Dirk Mauler[2], Peter Vary[1]

[1] *Institute of Communication Systems (IND), Aachen University, D-52056 Aachen, Email: enzner@ind.rwth-aachen.de*

[2] *Now with: Institute of Communication Acoustics (IKA), Ruhr-Universität Bochum, D-44780 Bochum*

Introduction

There is a strong acoustic coupling between loudspeaker and microphone of car hands-free telephones. A power coupling factor between -10 dB and 10 dB must be expected. The exact value depends on the desired loudspeaker volume in the presence of background noise and on the installation of loudspeaker and microphone in the car. The impulse response of this electroacoustic loop has a few hundred coefficients at 8 kHz sampling frequency and is very sensitive to movements in the car. Acoustic echo and background noise at the microphone are very annoying for the remote talker.

A simple and robust signal processing solution to the acoustic echo control problem has been proposed recently [1]. The solution is based on the purely statistical adaptation of an acoustic echo canceler and a postfilter for residual echo suppression as shown in Figure 1. A sophisticated double talk detection is not required for the stability of this algorithm. We have realized a realtime prototype and in this paper we discuss the results of the performance evaluation in the car environment. It turns out that the main strengths of the proposed solution are the duplex ability (fast and robust adaptation in double talk), and the high degree of elegance and simplicity.

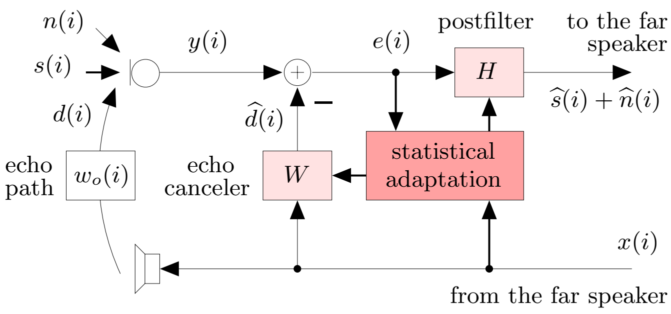

Figure 1: Hands-free telephone with time-varying echo path $w_o(i)$, echo canceler W, postfilter H, and purely statistical adaptation of both filters. The microphone signal $y(i)$ at sampling time index i is additively composed of clean near speech $s(i)$, local background noise $n(i)$, and acoustic echo $d(i)$. The output signal $\widehat{s}(i)+\widehat{n}(i)$ shall be an estimate of $s(i)+n(i)$.

Echo Canceler and Postfilter Facts

It is widely accepted that an acoustic echo canceler W cannot solve the acoustic echo control problem alone. It is very difficult to let the echo canceler coefficients follow the fast statistical echo path changes and to be robust against observation noise $s(i)+n(i)$ at the same time. Therefore, a postfilter H in the sending path of the hands-free telephone has been proposed in [2] for the statistical suppression of the residual echo $d(i)-\widehat{d}(i)$ which is part of the error signal $e(i)$. It has been shown that a frequency selective postfilter is able to reduce the residual echo and to preserve the duplex ability of the telephone. Moreover, such postfilter can be easily extended to perform combined residual echo and background noise suppression [3].

The quality of the output signal $\widehat{s}(i)+\widehat{n}(i)$ essentially depends on the fast and robust adaptation of the echo canceler and postfilter coefficients in the uncertain acoustic environment of the hands-free telephone. It has been claimed in [1] that this can be realized by a set of simple equations, at least when acoustic echo control is considered as a purely statistical problem:

For the realization of the echo canceler W, the frequency-domain adaptive filter (FDAF) [4] is recommended. The optimum step-size for the FDAF in the minimum mean-square error (MMSE) sense has been derived in [5] as

$$\mu(\ell,k)=\frac{|G(\ell,k)|^2\Phi_{xx}(\ell,k)}{\Phi_{ee}(\ell,k)}\ . \tag{1}$$

$\Phi_{ee}(\ell,k)$ and $\Phi_{xx}(\ell,k)$ are the power spectral densities of the error and excitation signal at the discrete frequency ℓ and frame-time index k. $|G(\ell,k)|^2$ is the time- and frequency-dependent system distance between echo canceler and echo path. If the time-varying echo path $w_o(i)$ can be modeled as a first order Markov process, it has been calculated in [1] that the convergence state $|G(\ell,k)|^2$ can be updated recursively by

$$|G(\ell,k+1)|^2=\big(1-\mu(\ell,k)\big)\cdot|G(\ell,k)|^2+\Delta(\ell,k)\ . \tag{2}$$

The statistical model parameter $\Delta(\ell,k)$ is the variance of the process noise in the Markov model.

It has been shown in [6] that a very close relation between the optimum statistical adaptation of echo canceler and postfilter exists. The MMSE postfilter for residual echo suppression is thus given in the frequency domain by

$$H(\ell,k)=1-\mu(\ell,k)\ . \tag{3}$$

Test Setup & Realtime Prototype

In order to investigate the performance of echo canceler and postfilter in a realistic acoustic environment, we have realized a realtime prototype of the adaptive algorithm proposed in [1] and we have evaluated this prototype in different cars as sketched in Figure 2.

*This work is supported by Nokia Research Center, Tampere, Finland, and Nokia Mobile Phones, Bochum, Germany.

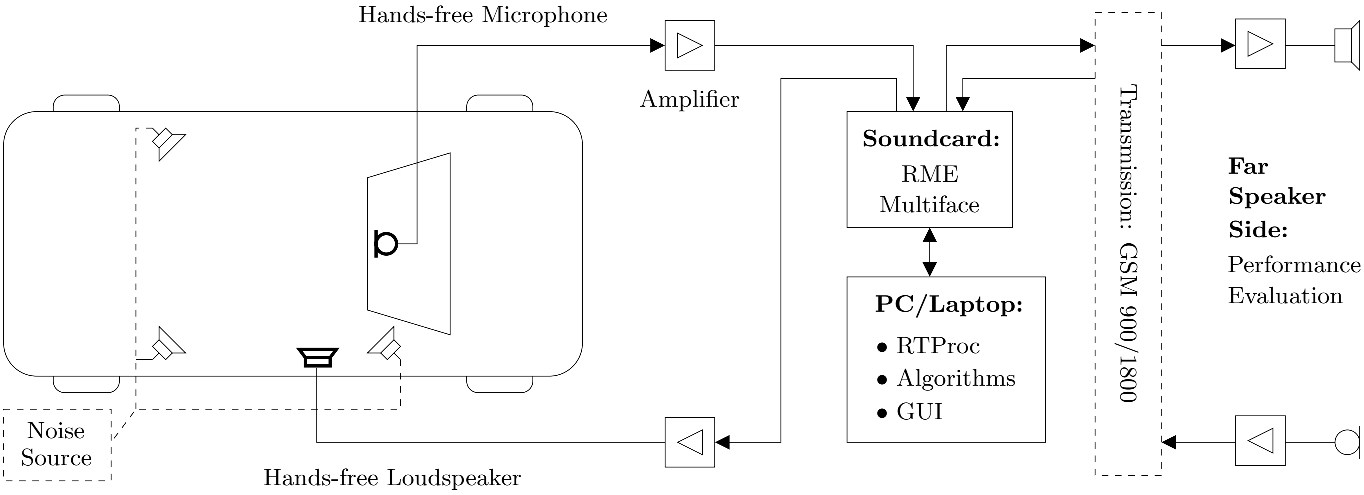

Figure 2: Test setup for the subjective performance evaluation of echo canceler and postfilter in the car environment. The exact positions of the hands-free loudspeaker and the hands-free microphone can vary in different cars. GSM is optional.

Inside the car, we have tested two different hands-free loudspeakers: the *Fostex Personal Monitor 6301B* which has a very low total harmonic distortion, and a small loudspeaker that is provided with the *Nokia Car Kit Cark-91*. In the sending direction of the system, we have always used the *Nokia HFM 8* hands-free mouse microphone. Car noise can be simulated with an external noise source and three additional loudspeakers in the car. To meet the typical (and also the extreme) requirements of automotive applications, we have tested sound pressure levels (SPL) in the car from 0 to 80 dB.

The signal processing part of the prototype has been implemented on the PC based realtime platform *RTProc* for *Microsoft Windows* [7]. The conversion between digital and analog signals is realized by a commercial soundcard. The realtime prototype can be connected by a *GSM* engine to a remote talker (far speaker) who evaluates the algorithm performance.

Subjective Performance Evaluation

A group of six expert listeners has evaluated the performance of the realtime prototype in various test situations: with double talk, with background noise, and with moving talkers in the car. Table 1 summarizes questions to the test persons and their corresponding performance rating. The achieved scores are explained in Table 2. It can be seen that the listeners have recognized a performance that is always better than the state-of-the-art. Especially the duplex ability of our system has been rated as excellent.

Algorithmic Complexity

An important advantage of the algorithm proposed in [1] is its structural elegance and simplicity, practically meaning that the implementation mainly requires simple vector arithmetics ($+/-/*/\div$) and FFT/IFFT. The estimated computational complexity for about 300 echo canceler coefficients is about 10 MIPS and the corresponding memory requirements are given by about 5 kWords static RAM.

Question	**Score**
Loudness and Sound Quality in General?	1.5
Acoustic Echo Attenuation?	1.66
Transmission Quality of the Background?	1.8
Duplex Ability in Double Talk Situations?	1.0

Table 1: Average score of six expert listeners.

Rating	**Score**
Excellent (Clearly Above the Average)	1
Good (State-of-the-Art)	2
Fair (Average Performance)	3
Poor (Below the Average)	4
Bad (Not Competitive)	5

Table 2: Explanation of the scores in Table 1.

References

[1] G. Enzner and P. Vary, "Robust and Elegant, Purely Statistical Adaptation of Acoustic Echo Canceler and Postfilter," Proc. IWAENC 2003, Kyoto (Japan), pp. 43-46, September 2003.

[2] R. Martin, J. Altenhöner, "Coupled Adaptive Filters for Acoustic Echo Control and Noise Reduction," Proc. IEEE Intl. Conference on Acoustics, Speech, and Signal Processing, pp. 3043-3046, May 1995.

[3] S. Gustafsson, R. Martin, and P. Vary, "Combined acoustic echo control and noise reduction for hands-free telephony," *Signal Processing*, vol. 64, pp. 21–32, 1998.

[4] E. R. Ferrara, "Frequency–domain adaptive filtering," *In: C.F.N. Cowan, P.M. Grant (Eds.), Adaptive Filters, Prentice–Hall, Englewood Cliffs, N.J.*, pp. 145–179, 1985.

[5] B. H. Nitsch, "A frequency-selective stepfactor control for an adaptive filter algorithm working in the frequency domain," *Signal Processing*, vol. 80, pp. 1733–1745, 2000.

[6] E. Hänsler, G. Schmidt, "Hands-free telephones - joint control of echo cancellation and postfiltering," *Signal Processing*, vol. 80, pp. 2295–2305, 2000.

[7] H. Krüger, T. Lotter, G. Enzner, P. Vary, "A PC based Platform for Multichannel Realtime Audio Processing," Proc. IWAENC 2003, Kyoto (Japan), pp. 195-198, September 2003.