Least Squares Equalizer Design under Consideration of Tail Effects

Stefan Goetze[1], Markus Kallinger[2], Alfred Mertins[3], and Karl-Dirk Kammeyer[1]

[1] *University of Bremen, Dept. of Communications Engineering, D-28334 Bremen, Email: goetze@uni-bremen.de*

[2] *Fraunhofer Institute for Integrated Circuits, D-91058 Erlangen, Email: markus.kallinger@iis.fraunhofer.de*

[3] *University of Lübeck, Institute for Signal Processing, D-23538 Lübeck, Email: alfred.mertins@isip.uni-luebeck.de*

Abstract

Modern high-quality hands-free telecommunication systems have to cope with several real-world problems, such as corruption of the desired signal by additive noise, acoustic echoes and reverberation. This paper addresses the mutual impacts of the subsystems for Acoustic Echo Cancellation and Listening Room Compensation (LRC). In acoustic systems for LRC the equalizer is placed in front of the loudspeaker. An estimate of the room impulse response (RIR) is necessary for the equalizer to compensate for the influence of the RIR at the position of the reference microphone where the human user is located. Since the RIR is identified by the acoustic echo canceller (AEC) anyway, its estimate can be used to design the equalizer. The quality of equalization in dependence of the degree of system identification will be investigated in this contribution. Furthermore the influence of the equalizer on an echo canceller is analyzed.

Listening Room Compensation

Figure 1 shows the basic setup for an LRC filter $\mathbf{c}_{\mathrm{EQ}}$ preceding the RIR $\mathbf{h}$ whose influence has to be compensated.

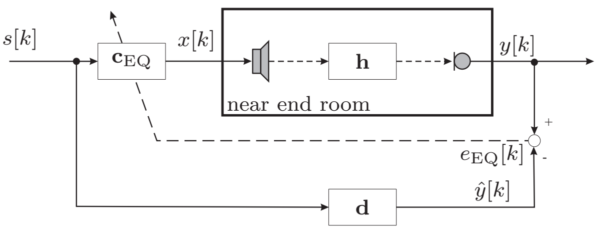

Figure 1: Least-squares equalizer for Listening Room Compensation.

By minimizing the mean square error of

$$e_{\mathrm{EQ}}[k] = \mathbf{s}^T[k]\mathbf{H}\mathbf{c}_{\mathrm{EQ}} - \mathbf{s}^T[k]\mathbf{d} \tag{1}$$

with the definitions

$$\mathbf{s}[k] = [\, s[k],\, s[k-1],\, \dots,\, s[k-L_h-L_{c,\mathrm{EQ}}+2]\,]^T \tag{2}$$

$$\mathbf{c}_{\mathrm{EQ}} = \left[\, c_{\mathrm{EQ},0},\, c_{\mathrm{EQ},1},\, \dots,\, c_{\mathrm{EQ},L_{c,\mathrm{EQ}}-1}\,\right]^T \tag{3}$$

$$\mathbf{d} = [\,\underbrace{0,\, \dots,\, 0}_{k_0},\, d[0],\, d[1],\, \dots,\, d[L_d-1],\, \underbrace{0,\, \dots,\, 0}_{L_h+L_{c_{\mathrm{EQ}}}-1-L_d-k_0}\,]^T \tag{4}$$

and the convolution matrix $\mathbf{H}$ of dimension $(L_h+L_{c,\mathrm{EQ}}-1 \times L_h)$ we get the well known least squares equalizer

$$\mathbf{c}_{\mathrm{EQ}} = \mathbf{H}^{+}\mathbf{d} \tag{5}$$

for a white noise input $s[k]$. In (5) $\mathbf{H}^{+}$ denotes the Moore-Penrose pseudoinverse of the channel matrix and $\mathbf{d}$ is the desired system which should be approximated by the concatenated system $\mathbf{c}_{\mathrm{EQ}}\mathbf{H}$. Here $\mathbf{d}$ is chosen as a 10^{th} order butterworth bandpass with band limits at 200Hz and 3700Hz for a sampling frequency of $f_s = 8$kHz. The lengths of the RIR, the LRC filter and the desired system $\mathbf{d}$ are denoted by L_h, $L_{c,\mathrm{EQ}}$ and L_d, respectively.

System Identification by an Acoustic Echo Canceller

For LRC an estimate of the RIR in eq. (5) is needed which can be delivered by the AEC since the estimate of the echo $\hat{\psi}[k]$ is obtained by system identification anyway.

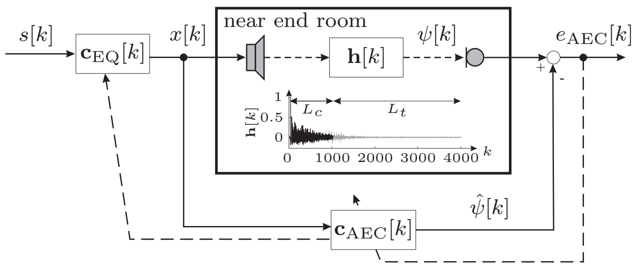

Figure 2: System for Listening Room Compensation with an Acoustic Echo Canceller for system identification.

Since the length L_h of the RIR which has to be identified is greater than the length L_c of the identification filter, the system identification will be biased for a nonwhite input $x[k]$. This is known from echo cancellation as the *tail effect* [1]. We split up the RIR into a part $\mathbf{h}_c[k]$ which can be modeled by the AEC and a tail $\mathbf{h}_t[k]$ according to Figure 2. By minimizing the power of the AEC error $\mathrm{E}\{e^2_{\mathrm{AEC}}[k]\}$ with the error signal

$$e_{\mathrm{AEC}}[k] = \mathbf{h}_c^T[k]\mathbf{x}_c[k] - \mathbf{c}_{\mathrm{AEC}}^T[k]\mathbf{x}_c[k] + \mathbf{h}_t^T[k]\mathbf{x}_t[k] \tag{6}$$

and the signal- and coefficient-vectors

$$\mathbf{x}_c[k] = [x[k]\,,\, x[k-1]\,,\, \dots\,,\, x[k-L_c+1]]^T \tag{7}$$

$$\mathbf{x}_t[k] = [x[k-L_c]\,,\, \dots\,,\, x[k-L_c-L_t+1]]^T \tag{8}$$

$$\mathbf{h}_c[k] = [h_0[k]\,,\, h_1[k]\,,\, \dots\,,\, h_{L_c-1}[k]]^T \tag{9}$$

$$\mathbf{h}_t[k] = [h_{L_c}[k]\,,\, h_{L_c+1}[k]\,,\, \dots\,,\, h_{L_h-1}[k]]^T \tag{10}$$

$$\mathbf{c}_{\mathrm{AEC}}[k] = [c_{\mathrm{AEC},0}[k],\, c_{\mathrm{AEC},1}[k],\, \dots,\, c_{\mathrm{AEC},L_c-1}[k]]^T \tag{11}$$

we obtain

$$\mathbf{c}_{\mathrm{AEC}}[k] = \mathbf{h}_c[k] + \mathrm{E}\left\{\mathbf{x}_c[k]\mathbf{x}_c^T[k]\right\}^{-1}\mathrm{E}\left\{\mathbf{x}_c[k]\mathbf{x}_t^T[k]\right\}\mathbf{h}_t[k]. \tag{12}$$

From equation (12) we see that the exact identification of the RIR is only possible for a white input signal since $\mathrm{E}\{\mathbf{x}_c[k]\mathbf{x}_t^T[k]\}$ is zero only for a white input. The more the early part of the input signal $\mathbf{x}_c[k]$ is correlated to the late part of the input signal $\mathbf{x}_t[k]$ the stronger the influence of the tail $\mathbf{h}_t[k]$ is. As we can see from Figure 2 further correlation is caused by the equalizer in the input path of the AEC.

Simulation Results

The RIR was simulated with a reverberation time of $\tau_{60} = 300$ ms. The filter orders of the AEC and the equalizer (EQ) were 1024 and 2048, respectively. As input signals white Gaussian noise and a recorded speech signal (male speaker) were used.

AEC convergence

The convergence of the AEC is influenced by the additional coloration introduced by the EQ.

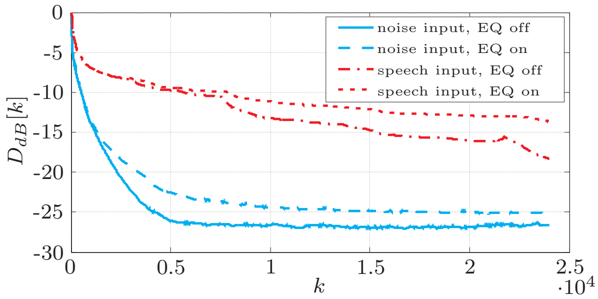

Figure 3: Relative System Misalignment $D_{\mathrm{dB}}[k]$

Figure 3 shows the relative system misalignment

$$D_{\mathrm{dB}}[k] = 10 \cdot \log_{10} \frac{||\mathbf{h}[k] - \mathbf{c}_{\mathrm{AEC}}[k]||^2}{||\mathbf{h}[k]||^2} \qquad (13)$$

with the quadratic vector norm $||\mathbf{h}[k]||^2 = \mathbf{h}^T[k]\mathbf{h}[k]$ for the two input signals $s[k]$ (white noise or speech) and for the cases of active and inactive EQ. If the EQ is switched off and the system input is white the AEC reaches the best system identification and the fastest convergence. If we switch on the EQ, both convergence and maximum system identification decrease. This is due to the correlation introduced by the EQ filter as we can see from (12). The same tendency can be observed for speech input.

Influence of the AEC on the EQ

For evaluation of the LRC subsystem we use the error criterion after [2]. The spectrum of the concatenated system of $\mathbf{c}_{\mathrm{EQ}}[k]$ and $\mathbf{h}[k]$ can be calculated by $E[m] = H[m] \cdot C_{\mathrm{EQ}}[m]$ with $H[m]$ and $C_{\mathrm{EQ}}[m]$ being the frequency-discrete room transfer function (RTF) and the LRC-filter, respectively. The variance of $E[m]$ gives a measure for the spectral flatness of the equalized system and thus for the quality of equalization:

$$\sigma_E^2 = \frac{1}{m_{max} - m_{min}} \sum_{m=m_{min}}^{m_{max}} \left(20 \cdot \log_{10}|E[m]| - \bar{E}_{\mathrm{dB}}\right)^2 .$$

Here the mean value of the logarithmic spectrum is given by $\bar{E}_{\mathrm{dB}} = 1/(m_{max} - m_{min}) \sum_{m=m_{min}}^{m_{max}} 20 \cdot \log_{10}|E[m]|$. The limits m_{min} and m_{max} are chosen to match the Discrete Fourier Transform (DFT) bins at 200Hz and 3700Hz respectively because this is our desired equalization area specified by the reference system $\mathbf{d}$.

Figure 4 shows the variance σ_E^2 in dependance of the system misalignment of the AEC for the white noise and the speech input signal. It should be mentioned that the x-axis is flipped so that high values for D_{dB}, which indicate *bad* convergence, are left and smaller values indicating *good* convergence are right. The two horizontal lines at $\sigma_E^2 = 1.05$ and $\sigma_E^2 = 14.34$ indicate the least-squares equalization with an ideally known impulse response and the unequalized case (EQ switched off) respectively.

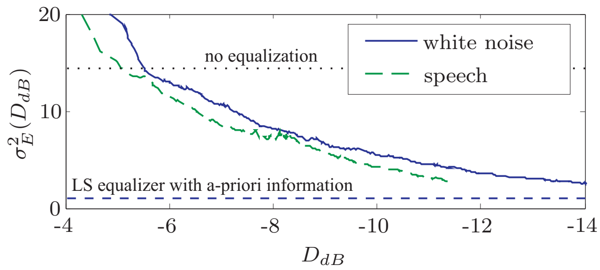

Figure 4: Variance σ_E^2 of the equalized system depending on the degree of system identification

The system shows the same behavior for both white noise and speech input. If the AEC shows poor convergence, which means that the system misalignment is high, the EQ introduces further distortion to the loudspeaker signal and should better be switched off in such periods. The *better* the system identification is the more the variance decreases which indicates a good equalization.

Conclusion

In this contribution we analyzed the mutual influences of the video-conferencing subsystems *Listening Room Compensation* and *Acoustic Echo Cancellation*. The quality of system identification and thus of echo reduction was shown in dependance of the coloration introduced by the equalizer. Furthermore the quality of equalization was analyzed in dependance of the degree of system identification. Using these results it is possible to influence the adaptation of one of the subsystems by analyzing the other to archive a better overall performance.

Literatur

[1] J. Benesty, D. R. Morgan, and M. M. Sondhi. A Better Understanding and an Improved Solution to the Specific Problems of Stereophonic Acoustic Echo Cancellation. *IEEE Trans. on Speech and Audio Processing*, 6(2):156–165, Mar 1998.

[2] J. N. Mourjopoulos. Digital Equalization of Room Acoustics. *Journal of the Audio Engineering Society*, 42(11):884–900, November 1994.