A 2-Stage Approach for Joint Noise Reduction and Dereverberation by means of Multi-Channel Equalization and a Noise Post-Processor

Ashish Sharma, Stefan Goetze

Fraunhofer IDMT, Project Group Hearing Speech and Audio Technology, Marie-Curie-Str. 2, Oldenburg, Germany.

Email: {sharaj, s.goetze}@idmt.fraunhofer.de

Abstract

Noise reduction (NR) for audio signals has been extensively studied during the last decades. Various approaches ranging from single channel, multi-channel to hybrid NR techniques have been proposed in the literature. Within the last years, signal enhancement for dereverberation received much attention. The main challenge in dereverberation by means of robust acoustic channel equalization has been the problem of identifying the room impulse response (RIR) accurately.

In this paper, NR and dereverberation combining regularized multi-channel equalization (MCEQ-$\mathrm{F_{opt}}$) for dereverberation and hybrid NR based on sub-band filtering is proposed. Individual microphone signals are first equalized by designing the equalizer filters based on the regularized MCEQ-$\mathrm{F_{opt}}$ algorithm. Then, the equalized signals are post-processed using a hybrid filter for noise reduction. Although the noise part may be amplified after the equalization step, our results show that the proposed 2-stage approach is able to achieve dereverberation to a great extent even in presence of channel estimation errors as well as noise reduction after post-filtering.

System Model

Figure 1 shows the proposed block diagram for achieving joint noise reduction and dereverberation for an acoustic system with 4 microphones. The speech source is denoted by $s[k]$, the noise source by $n[k]$ and $h_m[k]$ is the RIR between the speech source and respective microphones which is of length L_h.

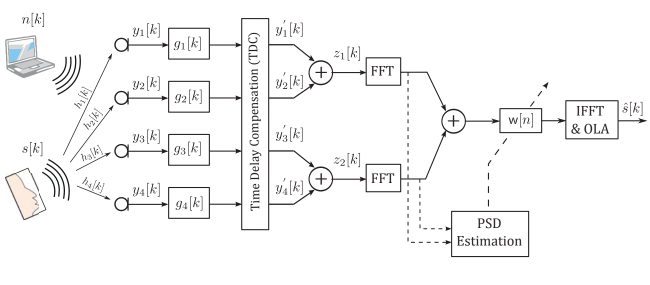

Figure 1: Block diagram of the proposed 2-stage approach.

The signal received by each microphone is the clean speech convolved with the respective RIR and perturbed by the noise. As shown in Figure 1, individual microphone signals are first equalized using multichannel equalization. The TDC block is used to steer the beam towards the direction of the speech source as well as to compensate for the delay introduced by the equalizers. Here, $z_i[k]$ is the equalized signal in which noise might have been amplified due to filtering. The noise is later reduced by using hybrid noise reduction technique utilizing subband filtering as proposed in [1]. After the 2-stage filtering, we expect the output signal $\hat{s}[k]$ to be dereverberated with reduced noise.

The MCEQ-$\mathrm{F_{opt}}$ Algorithm

The MCEQ-$\mathrm{F_{opt}}$ algorithm [2] was proposed by Rajan and Andy in 2013. They showed that choosing the equalizer filter with minimum l_2 norm gives optimum result. In this paper we use the regularized MCEQ-$\mathrm{F_{opt}}$ algorithm that helps us to design equalizer filters with low energy which is briefly summarized below

1. Construction of the multi-channel convolution matrix $\hat{\mathbf{H}}$ using erroneous RIR measurements,
2. Computation of the pseudo-inverse $\hat{\mathbf{H}}_{\mathrm{R}}^{\dagger} = (\hat{\mathbf{H}}^T\hat{\mathbf{H}} + \delta\mathbf{I})^{-1}\hat{\mathbf{H}}^T$ with an appropriate regularization parameter δ,
3. Selection of the equalizer filter $\mathbf{g}_{\mathrm{opt}} = \mathbf{h}_{\dagger k_{\mathrm{opt}}}$, for which $k_{\mathrm{opt}} = \min\limits_{i} ||\mathbf{h}_{\dagger i}||_2$, $1 \leq i \leq L_h$ and $\mathbf{h}_{\dagger i}$ is the i^{th} column of $\hat{\mathbf{H}}_{\mathrm{R}}^{\dagger}$.

Regularized MCEQ-$\mathrm{F_{opt}}$ is able to achieve dereverberation to a great extent with some pre and late echoes.

Hybrid Noise Reduction Approach

In Figure 1, after performing equalization, we use the hybrid noise reduction approach utilizing subband filtering as proposed in [1]. In this approach, frequency subbands are defined based on the inter-microphone distance and the uncorrelated microphone pairs are identified. Generally in the lowest subband, no microphone pairs are uncorrelated and hence, a single channel noise reduction approach based on Ephraim & Malah [3] in combination with Martin's minimum statistics [4] is used. For higher subbands, depending on microphone pairs that are uncorrelated, a modified Simmer's approach [1] is applied. Finally, the output signal after beamforming is filtered using a psychoacoustically motivated approach [5] in which the speech signal is obtained after pre-filtering utilizing the subband approach.

This work was partly funded by the EU ITN Dereverberation and Reverberation of Audio, Music, and Speech (DREAMS, project no. 316969) as well as the EU project Sounds for Energy-Efficient Buildings (S4EcoB, project no. 284628)

Simulation Results

Figure 2 shows the noise reduction (NR) in dB using the hybrid approach and the proposed 2-stage approach. It can be seen that the 2-stage approach is able to suppress noise to a similar level as compared to the hybrid approach. This shows that the 2-stage approach is also able to suppress noise to a great extent although noise might have been amplified after equalization.

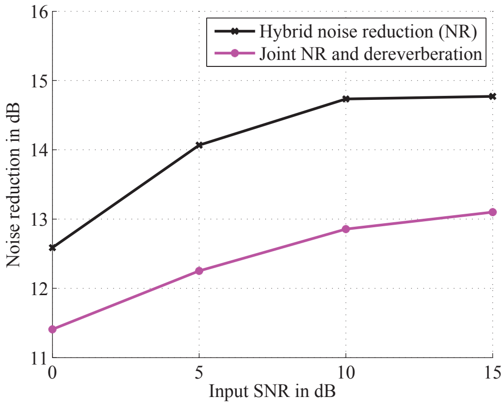

Figure 2: Comparison of noise reduction between hybrid noise reduction approach and proposed 2-stage approach.

Figure 3 shows the perceptual evaluation of speech quality (PESQ) scores for hybrid noise reduction approach and the proposed 2-stage approach. It can be clearly seen from Figure 3, that the 2-stage approach is able to achieve higher PESQ scores compared to the hybrid NR approach for all input SNR. This shows that the output signal is dereverberated to a great extent which was not possible in the hybrid noise reduction approach.

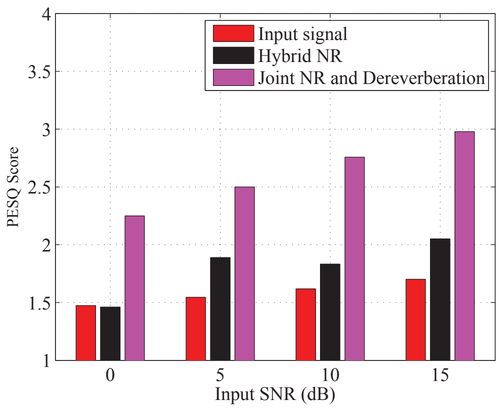

Figure 3: PESQ scores for hybrid noise reduction approach and the proposed 2-stage approach.

Figure 4 shows various signals in time domain. Comparing the top and the bottom panel, it can be seen that the 2-stage approach is able to recover the original speech signal to a great extent compared to the hybrid noise reduction approach. Furthermore, not much difference is seen in the noise reduction for both approaches.

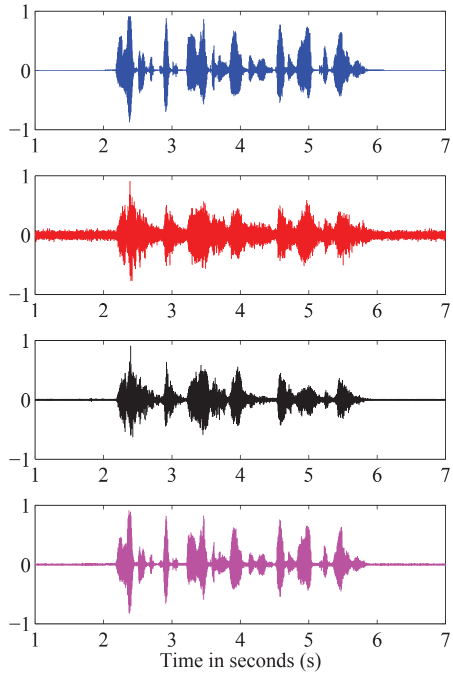

Figure 4: Top panel – clean speech, second panel – input signal at 10 dB input SNR, third panel – processed signal after hybrid noise reduction approach, bottom panel – processed signal after the 2-stage approach.

Conclusion

The proposed 2-stage approach for joint noise reduction and dereverberation is able to achieve both noise reduction and dereverberation to a great extent compared to the hybrid noise reduction approach. Although the noise reduction is slightly lower, there is clearly a great improvement in acoustic channel equalization resulting in higher PESQ score. Therefore, the 2-stage approach is a promising technique that can be used to achieve joint noise reduction and dereverberation in acoustic systems.

References

[1] V. Mildner, S. Goetze and K.D. Kammeyer, “Multi-Channel Noise-Reduction-Systems for Speaker Identification in an Automotive Acoustic Environment”, *Proc. AES*, Paris, France, May, 2006.

[2] R. S. Rashobh and A.W.H. Khong, “A multichannel time-domain subspace approach exploiting multiple time-delays for acoustic channel equalization”, *ICASSP*, Vancouver, Canada, May, 2013, pp. 418-422.

[3] 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.* Apr, 1985, pp. 443-445.

[4] Rainer Martin, “Noise power spectral density estimation based on optimal smoothing and minimum statistics”, *IEEE Trans. on Speech and Audio Processing, vol. 9.* July, 2001, pp. 504-512.

[5] S. Gustafsson, R. Martin, P. Jax and P. Vary, “A psychoacoustic approach to combined acoustic echo cancellation and noise reduction”, *IEEE Trans. on Speech and Audio Processing, vol. 10.* July, 2002, pp. 245-256.