Reducing the Delay of an Acoustic Echo Canceller with Subband Adaptation

Matthias Dörbecker *Peter Vary*

Institute of Communication Systems and Data Processing (IND)
Aachen University of Technology, Templergraben 55, 52056 Aachen, Germany
Phone: +49 241 806963, Fax: +49 241 8888186, Email: matthias@ind.rwth-aachen.de

ABSTRACT

In this contribution a novel structure for acoustic echo cancellation is presented. On one side the new structure takes advantage of subband adaptation, which results in high convergence speed as well as low complexity in comparison to the time domain NLMS algorithm. On the other side the key problem of subband echo cancellers, i.e. the large delay of the signal transmitted to the far-end talker, is eliminated by carrying out parts of the compensation in the time domain. Consequently, the proposed approach can be interpreted as combination between subband and time domain echo cancellers. Experimental simulations confirm the theoretical results.

1. INTRODUCTION

Due to the long duration of acoustic impulse responses the application of the normalized least mean square (NLMS) algorithm to acoustic echo control leads to FIR filters with a large number of coefficients, which results in a substantial computational complexity. Furthermore, the convergence behaviour of the NLMS algorithm is affected by the correlation characteristics of the stimulation signal. Especially, in case of speech signals the rate of convergence is very slow.

To overcome these drawbacks the signal of the far-end talker $x(k)$ and the microphone signal $v(k)$ can be decomposed into M subsampled narrowband signals $x_\mu(\rho)$ and $v_\mu(\rho)$, respectively, where μ $(0 \leq \mu < M)$ denotes the index of the subband. Because of the possibility to reduce the sampling rate

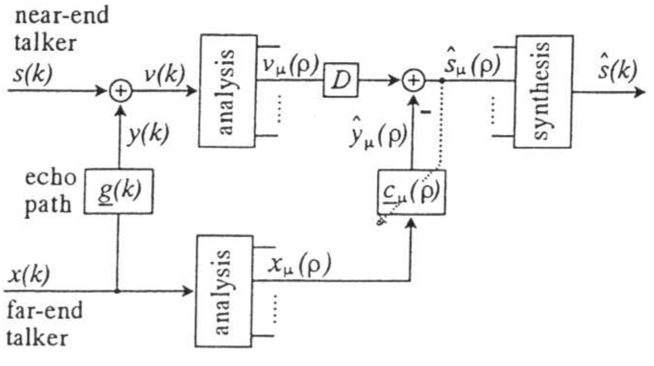

Figure 1. Conventional subband echo canceller

of the subband signals the compensation within each subband can be performed with shortened filters $\underline{c}_\mu(\rho)$ at a lower adaptation rate. The decomposition into subbands yields a reduced computational complexity as well as an effective decorrelation of the far-end speech signal $x(k)$. Therefore, the speed of convergence of the NLMS-driven filters in each subband is improved, as described in e.g. [1]. However, in the conventional subband echo canceller as shown in Fig. 1 the compensation takes place in the subband domain. Therefore, the near-end speech signal $s(k)$ is significantly delayed by the analysis-synthesis-filterbank. Furthermore, an additional delay D is necessary to realize the non-causal subband impulse responses $\underline{c}_\mu(\rho)$.

The structure proposed in this contribution uses subband adaptation to obtain a high rate of convergence at low computational complexity. The delay in the signal path can be eliminated, if the compensation takes place in the time domain using a single FIR filter. For this reason, an efficient transformation is required to determine the time domain impulse response $\underline{c}(k)$ which is equivalent to the analysis-synthesis-filterbank with subband convolution. In section 2 an efficient realisation of this transformation is derived.

However, the convolution with the time domain impulse response $\underline{c}(k)$ results in a high computational burden. Therefore, in section 3 a new structure is presented which determines only the early parts of the estimated echoes in the time domain. Aspects of computational complexity are discussed in section 4. In Section 5 the new concept is evaluated by simulation.

2. TRANSFORMATION OF SUBBAND IMPULSE RESPONSES INTO THE TIME DOMAIN

In this section it is shown how a system consisting of an analysis-synthesis-filterbank, a sampling rate reduction by r, and a subband convolution can be replaced by a filter in the time domain. Fig. 2 represents one channel of such a system. The bandpass filters of the analysis- and synthesis-filterbank are obtained by modulating (i.e. frequency-shifting) the prototype lowpass filters $\underline{h}_A$ and $\underline{h}_S$, respectively. The output of the subband filter $\underline{c}_\mu$, whose time variance is neglected

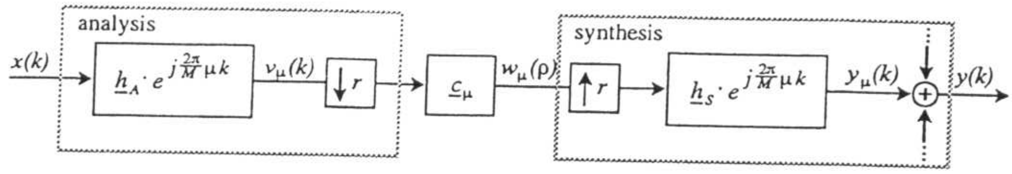

Figure 2. Analysis-synthesis-structure with subband filtering

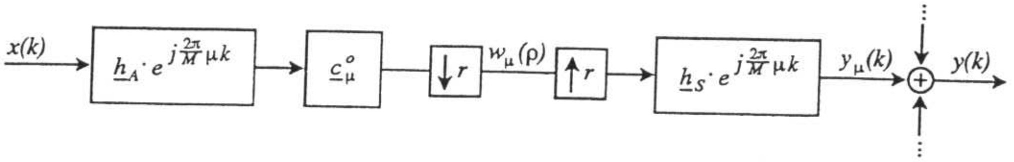

Figure 3. Equivalent system to Fig. 2

in this derivation, is given by

$$w_\mu(\rho) = \sum_{\sigma=0}^{N_c-1}\sum_{m=0}^{N_A-1} x(\rho r - \sigma r - m)\, h_{A,m}\, e^{j\frac{2\pi}{M}\mu m}\, c_{\mu,\sigma}\ ,$$

where $h_{A,m}$ denotes the m-th coefficient of the analysis prototype lowpass filter of length N_A. Similarly $c_{\mu,\sigma}$ represents the σ-th coefficient of the filter within the subband with index μ.

The subband impulse responses $\underline{c}_\mu$ are converted to the output sampling rate by inserting zeros according to:

$$c^o_{\mu,n} = \begin{cases} c_{\mu,\, n/r} & \text{if } n = \sigma r\ ,\ \sigma \in \mathbb{Z} \\ 0 & \text{else} \end{cases}$$

Using the substitution $n = \sigma r$ leads to:

$$w_\mu(\rho) = \sum_{n=0}^{r(N_c-1)}\sum_{m=0}^{N_A-1} x(\rho r - n - m)\, h_{A,m}\, e^{j\frac{2\pi}{M}\mu m}\, c^o_{\mu,n}$$

A realization of this relation is shown in Figure 3. In the following we assume that the analysis and synthesis lowpass filters are designed properly such that only negligible aliasing effects occur in the resynthesized signal $y(k)$. In this case the consecutive sampling rate converters in Figure 3 can be replaced by an attenuation by $1/r$. For this assumption the signal before the output summation point reads

$$y_\mu(k) = \frac{1}{r}\sum_{n=0}^{r(N_c-1)} c^o_{\mu,n}\sum_{m=0}^{N_A-1} h_{A,m}\, e^{j\frac{2\pi}{M}\mu m} \cdot \sum_{p=0}^{N_S-1} h_{S,p}\, e^{j\frac{2\pi}{M}\mu p} \cdot x(k-n-m-p)\ ,$$

where $h_{S,p}$ denotes the p-th coefficient of the synthesis lowpass filter. The substitutions $p' = m+p$, and $n' = n+m+p$ result in:

$$y_\mu(k) = \frac{1}{r}\sum_{n'=0}^{r(N_c-1)+N_A+N_S-2} x(k-n') \cdot \sum_{p'=0}^{N_A+N_S-2} c^o_{\mu,n'-p'}\, e^{j\frac{2\pi}{M}\mu p'}\sum_{m=0}^{N_A-1} h_{A,m}\, h_{S,p'-m}$$

The output signal of the analysis-synthesis-system is obtained by summation over M subbands:

$$\begin{aligned} y(k) &= \frac{1}{r}\sum_{\mu=0}^{M-1} y_\mu(k) \\ &= \frac{1}{r}\sum_{n'=0}^{r(N_c-1)+N_A+N_S-2} x(k-n') \cdot \\ &\quad \underbrace{\sum_{\mu=0}^{M-1}\sum_{p'=0}^{N_A+N_S-2} c^o_{\mu,n'-p'}\, e^{j\frac{2\pi}{M}\mu p'}\overbrace{\sum_{m=0}^{N_A-1} h_{A,m}\, h_{S,p'-m}}^{h_{A,p'}*h_{S,p'}}}_{c_{n'}} \end{aligned}$$

Consequently, the output signal can be calculated by the convolution

$$y(k) = \sum_{n'=0}^{r(N_c-1)+N_A+N_S-2} x(k-n')\, c_{n'}\ ,$$

with the impulse response $\underline{c}$ whose coefficients are given by:

$$c_{n'} = \frac{1}{r}\sum_{\mu=0}^{M-1}\sum_{p'=0}^{N_A+N_S-2} c^o_{\mu,n'-p'}\, e^{j\frac{2\pi}{M}\mu p'}\sum_{m=0}^{N_A-1} h_{A,m}\, h_{S,p'-m}$$

To obtain the impulse response in the time domain, the subband impulse responses must be increased in sampling frequency by zero insertions, convoluted by bandpass filters, and summed up over all channels, as it is shown in Figure 4. Consequently, this transformation is equal to a synthesis filterbank, whose prototype lowpass filter is given by $(\underline{h}_A * \underline{h}_S)/r$. An efficient realization of this synthesis is the use of an inverse FFT followed by a weighted overlap-add structure [2].

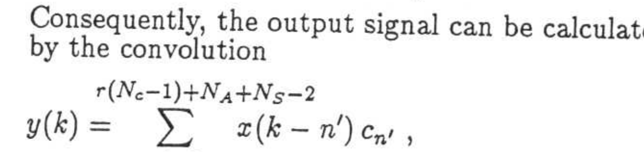

Figure 4. Transformation of subband impulse responses

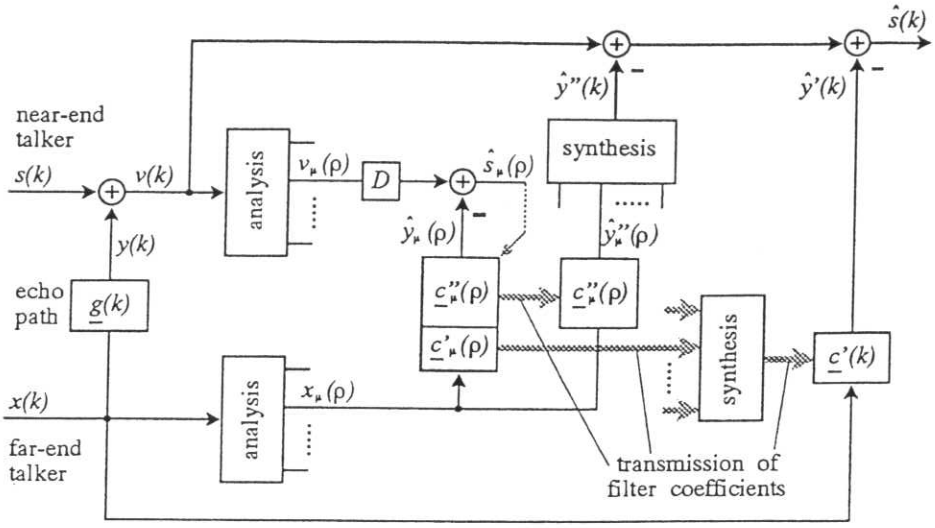

Figure 5. Proposed structure

If the prototype filters $\underline{h}_A$ and $\underline{h}_S$ are designed properly, the analysis-synthesis-filterbank is equivalent to a pure delay of $(N_A + N_S)/2$ samples. Therefore, the resynthesized impulse response $\underline{c}$ is shifted in time such that the impulse response starts with $(N_A + N_S)/2$ leading zero-valued samples.

For this derivation it was assumed that the subband filters are time invariant. The consequences resulting from the time variance of the NLMS-adapted subband filters are examined by a simulation in section 5.

3. AN EFFICIENT STRUCTURE WITH SUBBAND ADAPTATION

The application of the previously described transformation leads to an acoustic echo canceller which uses subband adaptation and time domain compensation. For the adaptation a conventional subband echo canceller is employed whose filter coefficients $\underline{c}_\mu(\rho)$ are transformed into an equivalent time domain impulse response $\underline{c}(k)$. The delay D_Σ of the conventional subband system depends on the length of the prototype filters N_A, N_S, and the value of the additional delay D in the subband signal path:

$$D_\Sigma = \frac{N_A + N_S}{2} + rD \ .$$

In the resynthesized impulse response $\underline{c}(k)$ this delay can be compensated by eliminating the D_Σ leading zero-valued coefficients. Therefore, the compensation according to

$$\hat{s}(k) = v(k) - \sum_{i=0}^{r(N_c-1)+N_A+N_S-2-D_\Sigma} c_{(i+D_\Sigma)}(k) \cdot x(k-i)$$

excludes any delay in the signal path. However, the convolution with the resynthesized impulse response is highly inefficient. For this reason, in the following a structure with reduced computational complexity is derived.

According to [3] a subband and a time domain echo canceller are connected in parallel. The subband echo canceller is modified to compensate only the late parts of the acoustic echoes. The late echoes correspond to that section of the acoustic impulse response whose coefficients have indices greater than the delay induced by the subband system. By this modification the delay can be eliminated. The remaining early echoes are compensated by the time domain echo canceller. In [3] it is proposed to adapt the time domain echo canceller by the NLMS algorithm. However, the employment of the time domain NLMS algorithm results in a low rate of convergence. Applying the previously described transformation the time domain NLMS algorithm can be replaced.

Figure 5 depicts the structure of the new echo canceller. The adaptation of the subband filter coefficients is equivalent to the conventional subband concept. Subsequently, each subband filter is split up into an early $\underline{c}'_\mu(\rho)$ and a late $\underline{c}''_\mu(\rho)$ part. The length N'_c of the early part is chosen according to

$$N'_c \geq \frac{D_\Sigma}{r} = \frac{N_A + N_S}{2r} + D \ .$$

Now only the early coefficients $\underline{c}'_\mu(\rho)$ are transformed into an equivalent time domain FIR filter, which is used to compensate the early echoes. Since the convolution in the time domain is constrained only to the early portion of the estimated acoustic impulse response, the computational complexity is reduced.

The compensation of the late echoes can be performed more efficiently in subbands. Since the late part of the subband echo is estimated by

$$y''_\mu(\rho) = \sum_{i=0}^{N_c-N'_c-1} x(\rho - i)\, c_{\mu,i+N'_c}(\rho) \ ,$$

it is accelerated by N'_c samples, which is equivalent to rN'_c samples in the time domain. Therefore, in this structure the delay caused by the filterbanks as well as the subband delay unit D is eliminated.

4. COMPLEXITY

In the following the computational complexity of the new structure is pointed out in terms of real-valued multiply-add-operations per time. It is assumed that every complex-valued operation requires four real-valued operations. The time domain signals are sampled at a frequency f_s.

Due to the symmetry of the analysis, only the subbands with indices $0 \leq \mu \leq M/2$ have to be processed. All subband signals are complex-valued except the subbands with indices 0 and $M/2$.

The computational complexity of the employed blocks amounts to:

- analysis/synthesis of signals $x(k)$, $v(k)$, and $\hat{s}(k)$ using polyphase filterbanks [2] with window lengths $N_A = N_S = N$
  $C_{AS} = 3\,(N + 2\,M\ \mathrm{ld}M)\,\frac{f_s}{r}$
- transformation of the early subband filter coefficients into the time domain
  $C_{Tr} = N_c'\,(2\,N + 2\,M\ \mathrm{ld}M)\,\frac{f_s}{r}$
- subband adaptation
  $C_{Ad} = 4(M-1)\,N_c\,\frac{f_s}{r}$
- subband compensation of the late echoes
  $C_C'' = 2(M-1)\,N_c''\,\frac{f_s}{r}$
- time domain compensation of the early echoes
  $C_C' = (r\,N_c' + N)\,f_s$

As an example, we regard an echo canceller with $M=16$ subbands and a sampling rate reduction of $r=10$. For the analysis and synthesis of the speech signals prototype low pass filters of length $N=96$ are used. The subband delay is set to $D = 9$. Therefore, the length of the subband filters has to be chosen as $N_c = 18 + N_g/r$, where N_g denotes the length of the acoustic impulse response. In comparison with the time domain NLMS-adapted echo canceller, the computational complexity of the new concept is lower, if the acoustic impulse response has a length greater than 1000 coefficients. If the echo canceller should be applied to a comfort telephone used in an office room, acoustic impulse responses with more than 2000 coefficients have to be modelled. For this application the computational load of the new approach is 30 % below the complexity of the time domain NLMS algorithm.

5. SIMULATION RESULTS

For the derivation in section 2 time invariant subband impulse responses have been assumed. For this reason, in this section a simulation is presented which examines the consequences resulting from the time variance of the NLMS-adapted subband filters. Figure 6 depicts the Echo Return Loss Enhancement

$$ERLE = 10\ \lg\ \frac{E\{y^2(k)\}}{E\{(y(k)-\hat{y}(k))^2\}}\ \mathrm{dB}$$

of the new approach in comparison with the performance of the conventional subband concept and of the time domain NLMS algorithm. The far-end

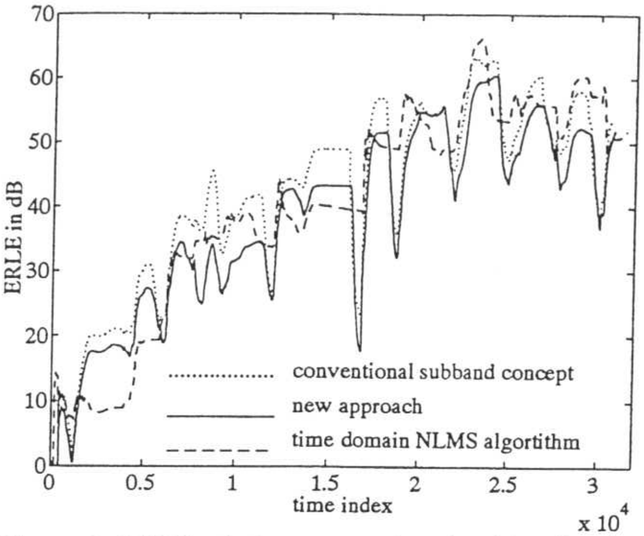

Figure 6. ERLE of the conventional subband echo canceller, time-domain NLMS algorithm, and of the new approach (sampling frequency 8 kHz)

talker $x(k)$ is modelled by a speech signal while the near-end talker keeps silent ($s(i) = 0$).

Figure 6 reflects that the best performance is reached by the conventional subband echo canceller. Due to the time variance of the subband filters the echo suppression of the new approach is slightly decreased in comparison to the conventional subband canceller. However, compared to the time domain NLMS algorithm, the new approach leads to comparable results.

6. CONCLUSION

In this contribution a new structure for an acoustic echo canceller with subband adaptation has been presented, which represents a compromise between conventional subband and time domain echo cancellers. The new approach eliminates the delay in the transmission path, which is usually a severe problem of conventional subband echo cancellers. In comparison with the conventional time domain NLMS algorithm the new structure results for long acoustic impulse responses in a reduced computational complexity and in a similar performance.

REFERENCES

[1] W. Kellermann, *Zur Nachbildung physikalischer Systeme durch parallelisierte digitale Ersatzsysteme im Hinblick auf die Kompensation akustischer Echos*, VDI-Fortschrittberichte, Reihe 10, Nr. 102. VDI-Verlag, Düsseldorf, 1989.

[2] R.E. Crochiere and L.R. Rabiner, *Multirate Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, New Jersey, 1983.

[3] J. Chen, H. Bes, J. Vandewalle, and P. Janssens, "A New Structure For Sub-Band Acoustic Echo Canceller", in *Proc. ICASSP*, 1988, pp. 2574–2577.