ACOUSTIC ECHO CONTROL IN SUBBANDS - AN APPLICATION OF MULTIRATE SYSTEMS

Gerhard Schmidt
Fachgebiet Theorie der Signale, Darmstadt University of Technology,
Merckstr. 25, D-64283 Darmstadt
Tel: +49 6151 16-4142 ; Fax: +49 6151 16-3778
e-mail: schmidt@nesi.e-technik.tu-darmstadt.de

ABSTRACT

This paper discusses the application of multirate signal processing techniques for the design of a system to cancel acoustic echoes in subbands. After a description of the computational complexity and the filterbank design, the need of delaying the microphone signal is explained. Aliasing caused by nonideal filterbanks and especially the noncausal subband impulse responses limit the quality of the system. If certain restrictions (e.g. an upper threshold of delay or a finit amount of memory) have to be fulfilled, it is neccesary to "balance" the entire system. Hence, all components should be in the same "quality-range". Finally we point out the advantages gained by applying multirate processing.

1 Indroduction

Comfortable hands-free telephones reduce the acoustic coupling between the loudspeaker and the microphone by computing a (digital) replica $\underline{c}(k)$ of the loudspeaker-enclosure-microphone-system (LEMS) as shown in Fig. 1. The part of the microphone signal $y(k)$ which is caused by the far-end speaker $x(k)$ can be eliminated by subtracting the estimated signal $\hat{y}(k) = x(k) * c(k)$ from the signal $y(k)$.

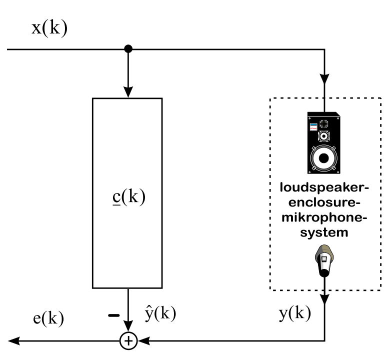

Figure 1: Structure of an acoustic echo control system operating in fullband

Acoustic echo control systems can be realized either in fullband or subband structures [5]. Both approaches show advantages and disadvantages, especially in computational complexity and delay.

2 Echo cancellation in subbands

By using a filterbank the signal $x(k)$ of the far-end speaker and the microphone signal $y(k)$ can be split up into several subbands (see Fig. 2). Depending on the abilities of the lowpass and bandpass filters, the sampling rate in the subbands can be reduced. According to this reduction the length of the adaptive filters can also be truncated. Instead of filtering the fullband signal and adapting one filter at the high sampling rate, M (number of subbands) convolutions and adaptions with subsampled signals are computed in parallel at reduced sample rates. The final error signal $e(k)$ is obtained by recombining and upsampling the subband signals $e_\mu(n)(\mu = 0 \dots \frac{M}{2})$. Operating with subsampled signals leads to a reduction of computational complexity, which is one of the main advantages of subband structures.

2.1 Computational complexity

In the fullband structure the number of multiplications for convolution and adaptation using the (e. g.) NLMS-algorithm and an adaptive filter of length L is approximately $N_{Mul,FB} \approx 2L$. Assuming that M subbands are used with equal bandwidth and equal subsampling factor $r \leq M$ the number of multiplications in the subband structure $\tilde{N}_{Mul,SB}$ can be denoted as $\tilde{N}_{Mul,TB} \approx 2L\frac{2M}{r^2}$. The factor $2M$ in the numerator results from the computation of $M/2$ complex-valued convolutions and adaptations. The subband signals are subsampled by a factor r, which reduces the complexity by r. The subsampled convolutions and adaptations have to be computed only every r sample periods, further reducing the complexity by the factor r.

$\tilde{N}_{Mul,SB}$ does not include the computational overhead caused by the filterbanks. By choosing a polyphase FIR-filterbank as described in [8] with $M = 16$ channels and a prototype lowpass filter of length $N = 256$, the additional complexity $\hat{N}_{Mul,SB}$ of two analysis- and one synthesis filterbank can be noted as $\hat{N}_{Mul,SB} \approx 3\frac{N}{r} +$

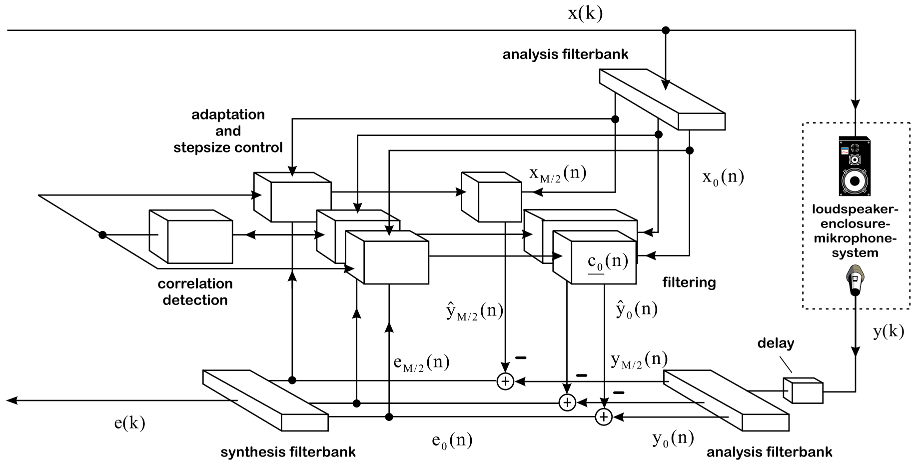

Figure 2: Structure of an acoustic echo control system operating in subbands

$3\frac{M}{r}\mathrm{ld}M$. Therefore, the total number of multiplications can be written as:

$$\begin{aligned} \text{Fullband:} \quad N_{Mul,FB} &\approx 2L, \\ \text{Subband:} \quad N_{Mul,SB} &\approx \underbrace{3\frac{N}{r} + 3\frac{M}{r}\mathrm{ld}M}_{filterbank} + \underbrace{4L\frac{M}{r^2}}_{\substack{adaptive \\ filtering}} \; . \end{aligned}$$

In Fig. 3 the number of multiplications per sample period versus the length of an adaptive fullband filter is shown. Depending on the subsampling factor r the effective reduction of computational load starts around a filter length of 100 coefficients.

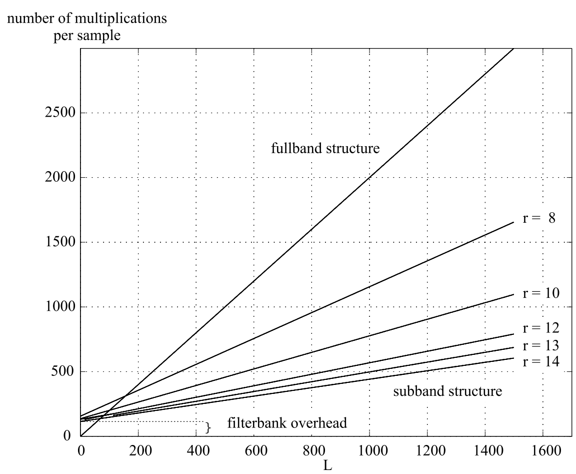

Figure 3: Number of multiplication per sample period (filterbank: 16 channels, lowpassfilter length 256)

2.2 Design of filterbanks

The quality of the entire echo control system relies strongly on the quality of the implemented filterbank. Therefore many papers on this topic have been published in recent years. Excellent tutorials about filterbanks and multirate systems can be found in [2] and [7].

In designing the filterbank a compromise between filter length, subsampling rate, stopband attenuation and aliasing has to be found. The filter should be as short as possible, which leads to a transition band of nonzero width. To avoid aliasing the subsampling factor has to be decreased or in critically subsampled filterbanks ($r = M$) a cascade of notch filters should be implemented. Decreasing the subsampling rate reduces the aliasing terms, but also distorts the flat filterbank transfer function in each subband and reduces the convergence speed. In the case of using notch filters the perfect reconstruction of the filterbank is lost. Nevertheless, these problems can be kept small by "fine tuning" after a rough startup design.

2.3 Speed of convergence

For white input signals, adapting with the NLMS-algorithm

$$\underline{c}(k+1) = \underline{c}(k) + \alpha(k)\frac{e(k)\,\underline{x}(k)}{\underline{x}^T(k)\,\underline{x}(k)}$$

and using a stepsize $\alpha(k) = 1$, the reduction of the mean square error (in fullband) can be approximated as [3]

$$\frac{\mathrm{E}\left\{|e(k)|^2\right\}}{\mathrm{E}\left\{|e(k-1)|^2\right\}} \approx 1 - \frac{1}{L}.$$

After r adaptations this quotient can be written as

$$\begin{aligned} \frac{\mathrm{E}\left\{|e(k)|^2\right\}}{\mathrm{E}\left\{|e(k-r)|^2\right\}} &\approx \left(1 - \frac{1}{L}\right)^r \\ &= 1 - \frac{r}{L} + \frac{r^2 - r}{2L^2} - \ldots \; . \end{aligned}$$

In subbands, assuming that there is no distortion caused by the filterbank, this ratio for the same period of time can be written as

$$\frac{\mathrm{E}\left\{|e_\mu(n)|^2\right\}}{\mathrm{E}\left\{|e_\mu(n-1)|^2\right\}} \approx 1 - \frac{r}{L}.$$

Especially for long filters both structures exhibit nearly the same speed of convergence. If speech is used as input signal, it can be shown [4], that the convergence of an adaptive filter depends on the eigenvalue spead of the correlationmatrix of the input signal, which is proportional to the quotient of the maximum and the minimum of the power spectral density. In ideal subband structures this quotient is reduced, leading to a faster convergence. This advantage is partially lost by nonideal filters, increasing the eigenvalue spread. For properly designed filterbanks both effects nearly compensate. Using speech as the input signal the speed of convergence is about the same for fullband and subband structures.

2.4 Artifical Delay

By transforming a causal fullband impulse response into a parallel system of subband impulse responses it can be shown, that the resulting subband impulse responses have noncausal tabs [5]. In order to reach a satisfying echo reduction, some of these noncausal tabs should also be modeled. This can be achieved by delaying the microphone signal. The artifical delay also increases the requirements on the echo attenuation [6].

In figure 4 two system identifications are presented. A measured impulse response of an office room (reverberation time ~300 ms) was used for the simulation. A filterbank with 16 channels and a subsampling factor of 13 was implemented and white noise was used as the exitation signal. The length of each adaptive subband filter was set to 100. Without any delay in the microphone path the echo could only be reduced by ~25 dB. With a delay of 64 (fullband) taps (~5 noncausal subband coefficients) an ERLE (echo return loss enhancement) of ~33 dB could be achieved.

2.5 Controlling the system

The subband structure offers the system designer an additional degree of freedom. Detectors and control mechanisms can be implemented separately for each channel. Also the length of the adaptive filters can be adjusted per channel to the statistical properties of the exitation signal (speech) and the room impulse response. For this reason the system in figure 2 has independent stepsize controls in each channel. The double-talk-detection (correlation detection) operates only in the band with the largest signalpower, which reduces the decetion error. It should be noted, that the the possibilty of a better system control increases the speed of convergence: An additional non-negligible advantage of subband structures.

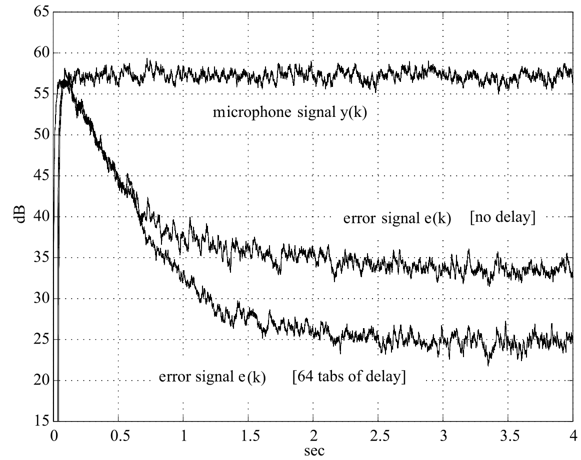

Figure 4: ERLE with 0 and 64 tabs of aritfical delay

3 "Balancing" the system

For real-time implementations the entire system has to comply with several restictions. If e.g. a certain amount of delay should not be exceeded, one has to optimize the ratio between the length of the filterbank and the number of taps for the artificial delay. Contributions to echo attenuation error caused by the aliasing terms, non-compensated taps and noncausal taps should be in the same order.

Another example for "balancing" a system is the optimal usage of the signal processor in terms of available memory and computational power. For a given amount of aliasing in the filterbank the system designer can vary the subsampling factor r according to the length of the filterbank N. By choosing a smaller r, the effect of noncausal impulse responses in the subbands is reduced, the length of the filterbank can be reduced and more memory for the adaptive filters is available. The disadvantage of this reduction is the need of more subband coefficients to compensate the same fullband length. Increasing the subsampling factor r also increases the length of the filterbank and therefore the delay. In this case the product of the available memory for the adaptive filters and the subsampling factor should be optimized.

In the Fig. 5 and 6 the aliasing terms of two different polyphase FIR-filterbanks are presented. Both were designed using the methods described in [8] and [1]. The prototyp lowpass filter in the systhesis part $g(k)$ was the reversed lowpass filter of the analysis part $g(k) = h(N-k)$. For examining the aliasing terms the absolute value of

$$H_{al}\left(e^{j\Omega}\right) = G\left(e^{j\Omega}\right) \sum_{n=1}^{r-1} H\left(e^{j\Omega - \frac{2\pi}{r}n}\right)$$

was printed for both filterbanks. The subsampling factor can be figured out by counting the number of notches in the magnitude spectra $H_{al}\left(e^{j\Omega}\right)$ of both filterbanks. In figure 5 a length of $N = 256$ was used. The aliasing terms were smaller than 47 dB, operating with a subsampling factor of $r = 13$. In figure 6 a filterbank with nearly the same amount of aliasing (50 dB), but with only half of the length of the first one ($N = 128$) is presented. The price of this memory reduction is a decreased subsampling factor of $r = 11$. Implementing the second filterbank an amount of 3x128 memory words can be used additionally for the adaptive subband filters. For short filters due to memory limitations the transformed fullband filter length can be increased by this form of "balancing" the system.

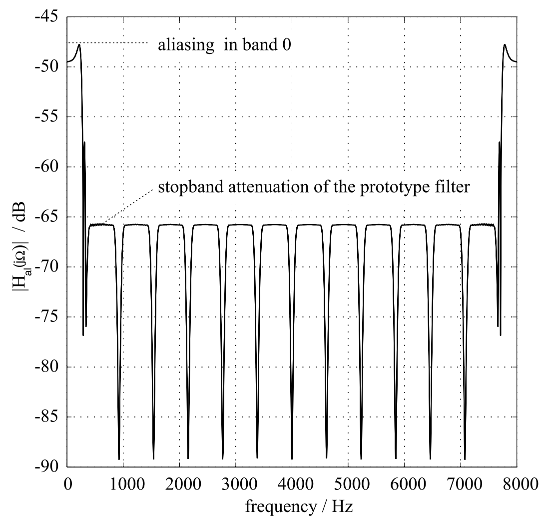

Figure 5: Properties of the filterbank (16 channels, N = 256, r = 13)

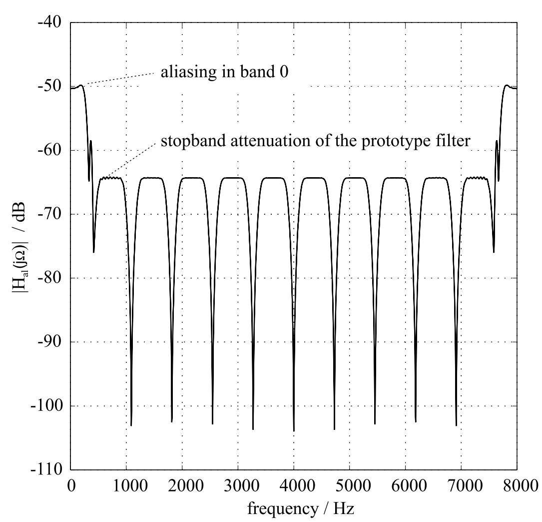

Figure 6: Properties of the filterbank (16 channels, N = 128, r = 11)

With these two examples for "balancing" a system it is explained, that the quality of the entire system will be best, if the contributions of all components are in the same order.

4 Conclusions

In this paper a short introduction in designing an acoustic echo control system operating in subbands is presented. After a comparison of complexity and convergence speed between fullband and subband structures, the design of filterbanks and the need of artifical delay in the microphone path are examined. For increasing the qualtity of the entire system, two possibilities of "balacing" a system are outlined.

Using multirate techniques in form of subband structures the computational complexity due to adaptive filtering can be reduced considerably. If a short delay is acceptable, the subband structure offers an additional degree of freedom in designing the entire system. All forms of control and detection can operate independent in each channel, which increases the speed of convergence. In summary, one can say that for many applications of acoustic echo control multirate processing is superior to single rate processing.

References

[1] R. Boite, H. Leich: *A New Procedure for the Design of High Order Minimum Phase Digital or CCD Filters*, Signal Processing 3, pp. 101 - 108, 1981

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

[3] R. Frenzel: *Freisprechen in gestörter Umgebung*, VDI Verlag, Reihe 10, Nr. 228, Düsseldorf, 1992

[4] S. Haykin: *Adaptive Filter Theory*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1986

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

[6] K. Shenoi: *Digital Signal Processing in Telecommunications*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1995

[7] P. P. Vaidyanathan: *Multirate Digital Filters, Filterbanks, Polyphase Networks, and Application: A Tutorial*, Proceedings of the IEEE, Vol. 78, No. 1, pp. 56 - 90, Jan. 1990

[8] G. Wackersreuther: *On the Design of Filters for Ideal QMF and Polyphase Filter Banks*, AEÜ, Band 39, Heft 2, Seite 123 - 130, 1985