# D6.6

## ADAPTIVE FILTERING IN SUB-BANDS

André Gilloire* and Martin Vetterli**

*CNET/LAA/TSS/CMC BP 40 F-22301 Lannion FRANCE
** Dept. of E.E. and CTR Columbia University New York NY 10027

ABSTRACT

Adaptive filtering in sub-bands, that is adapting several short filters in parallel on sub-sampled frequency sub-bands, is an attractive alternative to implement a long adaptive FIR filter, this both for computational efficiency and for faster convergence. The sub-sampling process introduces aliased versions of the original signals; it is shown how to annihilate these by adequate adaptive cross-terms. The computational complexity is improved nearly proportionally to the number of sub-bands, and the convergence speed is improved as well, making this scheme attractive for applications like acoustic echo cancellation.

### 1. INTRODUCTION

When the adaptive filter is a very long FIR filter, like in acoustic echo cancellation for teleconferencing where filters of several thousand taps are used [2], the computational complexity is very high and the performances of the algorithm (convergence speed and computational noise) are poor. It has therefore been suggested to divide the problem into sub-bands, and to perform an adaptive filtering in each of the sub-bands separately [2,3,5]. The problem of identifying a single very long filter has thus been changed into the task of identifying several smaller filters in parallel and at a lower speed. The computational gain is close to the number of bands, and convergence and computational noise are improved [1].

However, it has been noticed that undesirable aliased components appear in the output [1] when critically sub-sampled filter banks are used for the sub-band division [4]. Critical sub-sampling is desired because it achieves the lowest computational complexity. Schemes using non-overlapping bands [5] are less satisfactory because the reconstructed signal has spectral gaps which are noticeable when the number of bands is large.

In this paper we demonstrate that sub-band adaptive filtering is possible without aliasing or spectral reconstruction gaps. This is achieved by modifying the classical sub-band scheme used for speech coding and adding adaptive cross-terms between the channels. First, the two-channel case is analyzed in detail, showing how aliasing can be cancelled by using appropriate cross-terms. Then, the N-channel case is investigated, showing that cross-terms from adjacent bands are sufficient if the band-pass filters are selective enough. Both cases are simulated with and without cross-terms, showing the aliasing cancellation and the achieved performances (convergence and residual error).

### 2. ADAPTIVE IDENTIFICATION IN SUB-BANDS

The acoustic echo cancellation problem for teleconferencing is an adaptive identification problem, as shown in Figure 1 [6]:

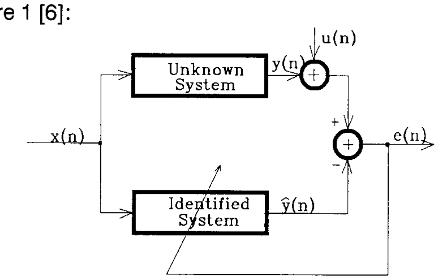


Figure 1: adaptive identification system

The signal x(n) coming from the distant teleconference room is convolved with the acoustic response of the local room; this results in an echo signal y(n). The useful signal u(n) is added to the echo at the microphone. Ideally, the signal e(n) which is sent to the distant room should be free of echo. Therefore, in the purpose of achieving echo cancellation, the room acoustic impulse response has to be identified and convolved with x(n) to produce an estimate $\hat{y}(n)$ of the echo which is subtracted from the microphone output to yield e(n), the useful signal to be sent.

In the case of a sub-band echo canceller, the identification process is made in sub-sampled sub-bands of the signals, as shown in Figure 2.

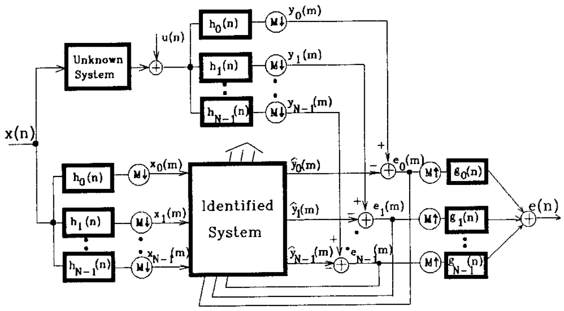


Figure 2: Adaptive identification in sub-bands

As can be seen from figure 2, the identification is made in a sub-sampled domain, and the final error signal is obtained by upsampling and recombining the error signals from the sub-bands. The general case will involve an analysis into M sub-bands followed by a sub-sampling by N, where N has to be smaller or equal to M [4]. Note that the identified system is general, and not restricted to parallel independent filters as it was done up to now.

## 3. THE TWO CHANNEL CASE

Let us first consider the two channel case with a sub-sampling by 2 (N=M=2). We will assume that u(n) is zero, and then we will try to obtain e(n) equal to zero by finding the appropriate identification structure in the sub-bands.
Assuming that all z-transforms of signals and filters exist, and with the notations of figure 2, one can verify that:

$$\begin{bmatrix} Y_0(z) \\ Y_1(z) \end{bmatrix} = 1/2 \begin{bmatrix} H_0(z^{1/2}) & H_0(-z^{1/2}) \\ H_1(z^{1/2}) & H_1(-z^{1/2}) \end{bmatrix} \begin{bmatrix} S(z^{1/2}) & 0 \\ 0 & S(-z^{1/2}) \end{bmatrix} \begin{bmatrix} X(z^{1/2}) \\ X(-z^{1/2}) \end{bmatrix} \quad (1)$$

$$\begin{bmatrix} X_0(z) \\ X_1(z) \end{bmatrix} = 1/2 \begin{bmatrix} H_0(z^{1/2}) & H_0(-z^{1/2}) \\ H_1(z^{1/2}) & H_1(-z^{1/2}) \end{bmatrix} \begin{bmatrix} X(z^{1/2}) \\ X(-z^{1/2}) \end{bmatrix} \quad (2)$$

where S(z) is the z-transform of the system to be identified. In (1-2), we used the fact that if a signal is sub-sampled by 2, that is if y(n) = x(2n), then [7]:

$$Y(z) = 1/2\ [X(z^{1/2}) + X(-z^{1/2})] \quad (3)$$

The output of the identified system is:

$$\begin{bmatrix} \hat{Y}_0(z) \\ \hat{Y}_1(z) \end{bmatrix} = \begin{bmatrix} C_{00}(z) & C_{01}(z) \\ C_{10}(z) & C_{11}(z) \end{bmatrix} \begin{bmatrix} X_0(z) \\ X_1(z) \end{bmatrix} \quad (4)$$

Now, in order to obtain a zero error signal, it is necessary and sufficient that both $e_0(m)$ and $e_1(m)$ are zero. The condition $e_0(m) = e_1(m) = 0,\ \forall m$, is equivalent to choosing $\mathbf{C}(z)$ as:

$$\mathbf{C}(z) = \mathbf{H}(z^{1/2}) \cdot \mathbf{S}(z^{1/2}) \cdot [\mathbf{H}(z^{1/2})]^{-1} \quad (5)$$

where we have used the following notations:

$$\mathbf{C}(z) = \begin{bmatrix} C_{00}(z) & C_{01}(z) \\ C_{10}(z) & C_{11}(z) \end{bmatrix}, \quad \mathbf{H}(z) = \begin{bmatrix} H_0(z) & H_0(-z) \\ H_1(z) & H_1(-z) \end{bmatrix}$$

$$\mathbf{S}(z) = \begin{bmatrix} S(z) & 0 \\ 0 & S(-z) \end{bmatrix} \quad (6)$$

For simplicity, we assume that the filters are choosen as $H_0(z) = H(z)$, $H_1(z) = H(-z)$, where H(z) is a half-band low-pass filter. This solution is usual in sub-band coding applications [7]. Furthermore, we assume that the determinant of $\mathbf{H}(z)$ is a simple delay, a good approximation when H(z) is a "good" half-band filter. Assuming that H(z) is a length-L linear phase filter, then $\mathrm{Det}\ [\mathbf{H}(z)] = \beta_{L-1}\ z^{-L+1}$. From the above, the identified system matrix $\mathbf{C}(z)$ becomes:

$$\mathbf{C}(z) = \beta_{L-1}^{-1}\ z^{(L-1)/2} \cdot$$

$$\begin{bmatrix} H^2(z^{1/2})S(z^{1/2})-H^2(-z^{1/2})S(-z^{1/2}) & H(z^{1/2})H(-z^{1/2})[S(-z^{1/2})-S(z^{1/2})] \\ H(z^{1/2})H(-z^{1/2})[S(z^{1/2})-S(-z^{1/2})] & H^2(z^{1/2})S(-z^{1/2})-H^2(-z^{1/2})S(z^{1/2}) \end{bmatrix} \quad (7)$$

Note that $\mathbf{C}(z)$ is diagonal only if i) or ii) are verified:
i) H(z)H(-z) = 0, that is H(z) is a perfect half-band filter;
ii) S(z) = S(-z), that is S(z) is an even function of z, which is not the case for any real system impulse response.

Therefore, in any real system, the identified system matrix $\mathbf{C}(z)$ is not diagonal, unlike what was assumed in previous studies [2,3,5]. Actually, assuming a diagonal matrix $\mathbf{C}(z)$ will lead to the apparition of aliased versions of the input signal that cannot be cancelled by the adaptive system, as noted in [1]. More precisely, it can be verified that if $\mathbf{C}(z)$ is choosen with diagonal terms as in (7) but with zero off-diagonal terms, then (assuming that $G_0(z) = H(z)$ and $G_1(z) = -H(-z)$ as usual in sub-band coding) the z-transform of the error signal e(n) becomes:

$$E(z) = C_{01}(z) \cdot [\ H(z)\,H(-z) \quad H^2(z) - H^2(-z)\ ] \begin{bmatrix} X(z) \\ X(-z) \end{bmatrix} \quad (8)$$

Since H(z)H(-z) is small compared to $H^2(z)-H^2(-z)$, the error signal contains a large part due to the aliased input X(-z). It is therefore clear that a diagonal approximation of $\mathbf{C}(z)$ is unsatisfactory.

From an implementation point of view, note that the elements of $\mathbf{C}(z)$ correspond to FIR filters of length $L_C = (2\,L_H + L_S - 1)/2$, where $L_H$ and $L_S$ stand for the lengths of the filter and system respectively. Note however that $C_{01}(z) = -C_{10}(z)$ has a factored form where only one part is dependent on the system S(z). Furthermore, the constant factor $F_0$ is obtained from H(z).H(-z), that is, the product of a half band low-pass with the corresponding high-pass, and can therefore be approximated by a shorter filter in general. The resulting identified system has the form depicted in Figure 3:

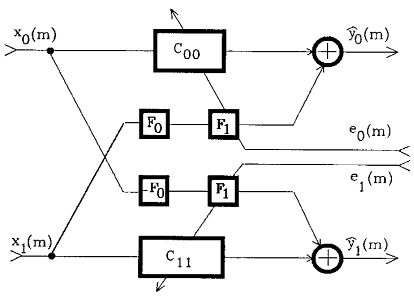


Figure 3: identified system with factorized cross-terms

## 4. THE N CHANNEL CASE

We consider the case of critical sampling, that is, the N channels are sub-sampled by M = N. As analysis filters, we take band-pass filters which have sufficient out of band rejection so that:

$$H_i(z).H_j(z) = 0, \quad |i-j| > 1 \tag{9}$$

That is, only adjacent filters will overlap. A convenient class of such filters are the so-called pseudo-QMF filters [8]. The matrix $\mathbf{H}(z)$ is of the form:

$$\mathbf{H}(z) = \begin{bmatrix} H_0(z) & H_0(wz) \ldots\ldots & H_0(w^{N-1}z) \\ H_1(z) & H_1(wz) \ldots\ldots & H_1(w^{N-1}z) \\ \cdot & \cdot \quad \ldots\ldots & \cdot \\ H_{N-1}(z) & H_{N-1}(wz) \ldots & H_{N-1}(w^{N-1}z) \end{bmatrix} \tag{10}$$

where $w = \exp(-j2\pi/N)$ and the $H_i(z)$'s are obtained from a low-pass prototype filter by DFT modulation. Similarly, $\mathbf{S}(z)$ is a diagonal matrix made up from $S(z)$, $S(wz)$ . . . $S(w^{N-1}z)$. The convenience with the pseudo-QMF filters is that a good approximation of the inverse of $\mathbf{H}(z)$ is known, namely the equivalent synthesis filter matrix transposed $\mathbf{G}(z)^T$ which is also made of modulated versions of a single low-pass prototype.

One can verify that the product $\mathbf{H}(z).\mathbf{S}(z).\mathbf{G}(z)^T$ is tridiagonal if one takes the constraint (9) into account. This means that cross-terms between neighbouring channels are sufficient to cancel aliasing terms, since the other terms are eliminated by the frequency selectivity of the analysis filters. The filter length of both the main terms and the cross-terms is of the order of:

$$L_C = (L_H + L_G + L_S) / N \tag{11}$$

where $L_H = L_G$ usually. The cross-terms are obtained from the product of adjacent filters and can be approximated by shorter filters in general.

## 5. COMPUTATIONAL COMPLEXITY

Consider the number of multiplications that are required to process N input samples. A direct, full band algorithm would use:

$$\mu(\text{full band}) = 2 \,.\, L_S \,.\, N \text{ mults.} \tag{12}$$

A N-band system with sub-sampling by N requires the evaluation of 3 filter banks every N input samples, that is $3(N\text{Log}N + L_H)$ mults. if we assume that the filter bank was implemented using a polyphase network and a fast transform. Assuming that the cross-term filters are a third of the length of the main filters and noting that there are about 2N of them, the adaptive filtering in the sub-bands uses about $10/3.N.L_C$ mults. With $L_C$ from (11) and $L_H = L_G$, the total complexity of the sub-band system is about:

$$\mu(\text{sub-band}) = 10/3\, L_S + 29/3\, L_H + 3N\, \text{Log}N \tag{13}$$

Since $L_S$ is much larger than $L_H$ and N, the gain of (13) over (12) is nearly proportional to the number of bands.

## 6. EXPERIMENTAL RESULTS

Simulations in real time were conducted to verify experimentally the validity of the proposed method. The unknown system was the acoustic coupling channel between a loudspeaker and a microphone installed in a small test room. The experimental conditions and the computer for real time implementation were identical to those used in [1]. The input and output signals were band-limited at 7 kHz and sampled at 16 kHz. USASI noise was used as the input x(n) to simulate speech.

Three echo cancellers named A232, A896 and AFB were simulated. The LMS algorithm with normalized gain [6] was used in each of them; the gain was scaled to half the value for which instability of the algorithm appeared.

A232 used 2 bands QMF filter banks with 32 taps (N = M = 2). The adaptive filters in the main branches had 750 taps; the filters for cross-term cancellation had 240 taps for the adaptive part and 32 taps for the fixed part.

A896 used pseudo-QMF filter banks with a low-pass prototype of 96 taps; it included 8 sub-bands of equal width (1 kHz), of which the highest one (7-8 kHz) was not processed. The decimation factor was M = 8. The adaptive filters in the main branches had 250 taps; the filters for cross-terms cancellation had 128 taps for the adaptive part and 12 taps for the fixed part. Due to limitations in the capacity of the computer, it was only possible to make one cross-term elimination; however, this is thought to be sufficient to demonstrate the validity of the method.

AFB was a classical full-band canceller; it included an adaptive filter of 1500 taps.

Two characteristics were measured:
1. the convergence rate, starting from all coefficients equal to zero in the adaptive filters (null state);
2. the asymptotic performance: residual echo spectrum.

6.1. Convergence from the null state

The decay rates of the short-term (16 ms) rms levels of the residual echo from the null state for AFB and A232 are shown in Table 1. For each of them, the global decay rate in the whole signal band is given first; the 7 values that follow correspond to frequency bands 1kHz wide.

| Canc | Global | 0-1 | 1-2 | 2-3 | 3-4 | 4-5 | 5-6 | 6-7 |
|---|---|---|---|---|---|---|---|---|
| AFB | 9.2 | 18.4 | 15.7 | 12.9 | 9.9 | 8.5 | 6.5 | 4.4 |
| A232 | 17.0 | 23.3 | 29.3 | 19.2 | 5.9 | 15.0 | 30.5 | 25.1 |

Table 1: decay rates in dB/s for AFB and A232
(Limits of bands are in kHz)

It appears that the convergence rate is generally much faster for the sub-band canceller than for the full-band one, especially for high frequency terms. For A232, the values around 4 kHz include the decay rates of aliasing terms, that appear smaller than the decay rates of the other components; this is probably due to the fact that cross-terms are narrow-band signals.

6.2. Asymptotic performances

The figures 4 and 5 show the averages over 64 spectra of the residual echo computed after 30 s from the beginning of convergence, respectively for A232 and A896. The spectrum of the echo y(t) (room signal) is shown in dashed lines. The aliased terms that exist when there is no cross-term cancellation are drawn in dotted lines. The ERLEs achieved (Echo Return Loss Enhancements) are reported on the figures.

For A232, the cross-terms around 4 kHz have been almost completely cancelled. Total cancellation would have been achieved by using a somewhat longer filter to synthesize the cross-terms.

For A896, the aliased terms around 1 kHz have been completely cancelled. In other experiments (not shown) the cross-terms filters have been shifted around the other aliased terms; then, the complete cancellation of the corresponding cross-terms has been observed.

## 7. CONCLUSION

A new adaptive filtering structure has been proposed to identify large impulse responses, offering reduced computational complexity while achieving good performances like fast convergence. The ability of this structure to solve the problem of the non-cancellation of aliased terms that appear when critical sub-sampling is used has been demonstrated theoretically and experimentally. This structure could be used successfully in applications such as acoustic echo cancellation.

## REFERENCES


[1] A. Gilloire, "Experiments with sub-band acoustic echo cancellers for teleconferencing", Proc. ICASSP-87, April 1987, Dallas, pp. 2141-2144

[2] Y. Itoh et al, "An acoustic echo canceller for teleconference", Proc. ICC-85, Chicago, 1985, pp. 1498-1502

[3] W.Kellermann, "Kompensation akustisher Echos in Frequenzteilbandern", Frequenz, Vol 39, 1985, No. 7-8, pp. 209-215

[4] M.Vetterli, "A Theory of Multirate Filter Banks", IEEE Trans. ASSP, Vol 35 No. 3, pp. 356-372, March 1987

[5] H.Yasukawa et al, "Acoustic echo canceller with high speech quality", Proc. ICASSP-87, April 1987, Dallas, pp. 2125-2128

[6] B. Widrow and S.D.Stearns, "Adaptive Signal Processing", Prentice-Hall, Englewood-Cliffs, NJ, 1985

[7] R.E.Crochiere and L.R.Rabiner, "Multirate Digital Signal Processing", Prentice-Hall, Englewood-Cliffs, NJ, 1983

[8] P.L.Chu, "Quadrature Mirror Filter design for an arbitrary number of equal bandwidth channels", IEEE Trans. ASSP, Vol. 33 No.1, pp. 203-218, Feb. 1985


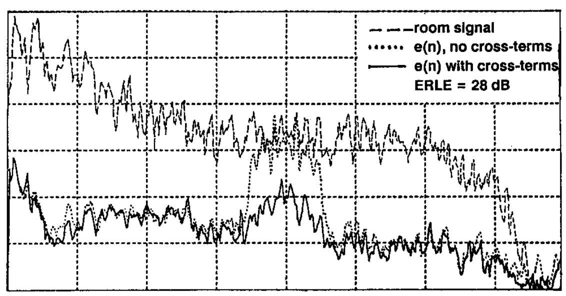


Figure 4: residual echo spectrum for A232
Scales: X: 1 kHz/div; Y: 10 dB/div

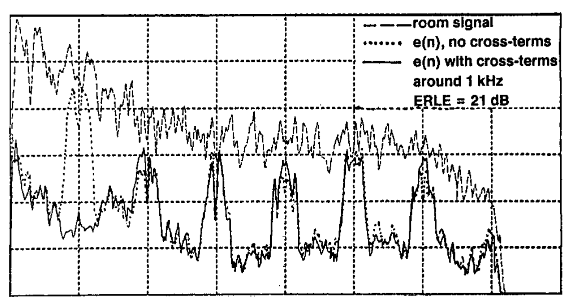


Figure 5: residual echo spectrum for A896
Scales: X: 1 kHz/div; Y: 10 dB/div