IMPLICIT DECIMATION FOR FIR SYSTEMS AND ITS APPLICATION TO ACOUSTIC ECHO CANCELLATION

Walter A. Frank and Imre Varga

Siemens AG, Mobile Phones Dept.
Munich, Germany
walter.frank@pn.siemens.de

ABSTRACT

This paper presents a filter structure which performs implicit decimation of the impulse response. As a result, the number of required operations is reduced or, equivalently, the impulse response length of the filter can be increased. Analysis in the frequency domain shows that this implicit decimation can be applied to systems that exhibit low-pass characteristics or have a smooth transfer function at high frequencies. Such behaviour can be assumed for many technical systems. For the determination of the optimal coefficients many well known algorithms for FIR systems can be used after a slight modification of the signal vector. The performance of implicit decimation is demonstrated for acoustic echo cancellation. Comparison with different algorithms shows that implicit decimation outperforms conventional FIR filtering.

1. INTRODUCTION

Digital signal processing techniques have a widespread use in most telecommunication devices. Especially in consumer products it is important to reduce cost, weight and power consumption. Hence, most devices have to cope with limited processing power. However, for effective acoustic echo cancellation, as it is well known, filter lengths of at least 150 to 200 taps are required, even at a quite moderate sampling frequency of 8 kHz. Hence, those algorithms are attractive which reduce the processing requirements or enable the use of filters with longer memory.

One possibility is to use Quadrature Mirror Filter banks (QMF), which split the signal in different frequency bands and perform subsampling [1]. Due to subsampling, they can operate with reduced filter lengths but require additional analysis and synthesis filtering to perform decimation.

In this paper implicit decimation of the filter coefficients is proposed. Because the resulting system operates in the fullband it does not require analysis and synthesis filters and hence, the complexity is further reduced. An analysis of the implicit decimation shows to what kind of systems this technique can be applied. The good performance is demonstrated in an acoustic echo compensation environment.

2. IMPLICIT DECIMATION

For convenience and simplicity, we illustrate below implicit decimation (ID) with a subsampling factor of two and a symmetric splitting of the impulse response length. Generalization to higher order decimation, i.e. subsampling factors 4, 8, etc., and unsymmetrical splitting is straightforward.

In a number of system identification problems it can be observed that for larger memory lags the difference between neighbouring values of the impulse response is small. The basic idea presented in this paper is the approximation by combining two or more coefficients in order to reduce the complexity of the modelling FIR filter.

In general, FIR filtering with coefficients $h[i]$ and filter length N performs the operation

$$y[n] = \sum_{i=0}^{N-1} h[i]x[n-i].$$

Combining always two coefficients in the 2nd half of the original FIR coefficients results in the approximation (without loss of generality we assume that N is a multiple of 4)

$$\begin{aligned}\tilde{y}[n] &= \sum_{i=0}^{N/2-1} h[i]x[n-i] \quad + \\ &\frac{1}{2}\sum_{i=0}^{N/4-1} \left(h[\tfrac{N}{2}+2i]+h[\tfrac{N}{2}+2i+1]\right)\times \\ &\qquad \left(x[n-\tfrac{N}{2}-2i]+x[n-\tfrac{N}{2}-2i-1]\right).\end{aligned}$$

This may be written in vector form as

$$\tilde{y}[n] = \mathbf{h_1}^T\mathbf{x_1}[n] + \mathbf{h_2}^T\mathbf{x_2}[n], \tag{1}$$

where

$$\begin{aligned}\mathbf{h_1} &= \left[\, h[0]\, h[1] \cdots h[\tfrac{N}{2}-1] \,\right]^T \\ \mathbf{x_1}[n] &= \left[\, x[n]\, x[n-1] \cdots x[n-\tfrac{N}{2}+1] \,\right]^T \\ \mathbf{h_2} &= \left[\, h[\tfrac{N}{2}]+h[\tfrac{N}{2}+1] \cdots h[N-2]+h[N-1] \,\right]^T \\ \mathbf{x_2}[n] &= \frac{1}{2}\left[\, x[n-\tfrac{N}{2}]+x[n-\tfrac{N}{2}-1] \cdots \right. \\ &\qquad \left. \cdots\, x[n-N+2]+x[n-N+1] \,\right]^T.\end{aligned}$$

The length of $\mathbf{x_1}$ and $\mathbf{h_1}$ equals $\frac{N}{2}$ whereas $\mathbf{x_2}$ and $\mathbf{h_2}$ are only length $\frac{N}{4}$ vectors. One observes that $\mathbf{x_2}[n]$ loses its

shifting property if it is updated at every time instant n. Updating $\mathbf{x_2}[n]$ only at even n, i.e.

$$\mathbf{x_2}[n] = \begin{cases} \frac{1}{2}\left[\, x[n-\frac{N}{2}]+x[n-\frac{N}{2}-1]\ \cdots\,\right] & n \text{ even} \\ \frac{1}{2}\left[\, x[n-\frac{N}{2}-1]+x[n-\frac{N}{2}-2]\ \cdots\,\right] & n \text{ odd} \end{cases} \tag{2}$$

one just has to calculate the most recent entry of $\mathbf{x_2}[n]$, the rest is shifted only. Additionally, this modified update has the advantage that the convolution $\mathbf{h_2}^T\mathbf{x_2}[n]$ also has to be performed only every other time instant. Thus, the average number of multiplications per sample is

$$M = \frac{N}{2} + \frac{1}{2}\frac{N}{4} = \frac{5}{8}N.$$

This compares to N multiplications for the conventional FIR filter. The complexity is further reduced when the number of decimated coefficients is increased or when higher order decimation is used.

The determination of the ID coefficients is straightforward. Nearly all algorithms that are known for conventional FIR systems can be used for ID systems. See Section 4 for a more detailed description.

3. ANALYSIS IN THE FREQUENCY DOMAIN

As it has been shown in the previous Section, ID systems allow a significant reduction in processing requirements by approximating the true impulse response, but what does this approximation mean in the frequency domain?

Fig.1 sketches the structure of ID filtering according to (1) with the update of $\mathbf{x_2}[n]$ every other time instant according to (2). $w_1[i]$ and $w_2[i]$ window the first and second part of the impulse response, respectively, i.e.

$$w_1[i] = \begin{cases} 1 & \text{for} \quad i = 0, \ldots, \frac{N}{2}-1 \\ 0 & \text{else} \end{cases}$$

$$w_2[i] = \begin{cases} 1 & \text{for} \quad i = \frac{N}{2}, \ldots, N-1 \\ 0 & \text{else.} \end{cases}$$

In the first part (I) of the decimation branch, non-causal filtering with $d[i]$ performs the combination of two coefficients:

$$\sum_{k=-1}^{N-1} d[k]\, h[i-k] w_2[i-k] \tag{3}$$

with

$$d[i] = \begin{cases} 0.5 & \text{for} \quad i = -1, 0 \\ 0 & \text{else.} \end{cases}$$

Note that (3) contains $\frac{N}{2}$ different undecimated coefficients and only every other one is the true ID coefficient. Hence, every other coefficient is picked in part (II) and part (III) performs a one-sample hold.

With the modulation theorem of the z-Transform

$$(-1)^n x[n] \quad \circ\!\!-\!\!\bullet \quad X(-z)$$

and

$$W_1(z) = \frac{1 - z^{-N/2}}{1 - z^{-1}}$$

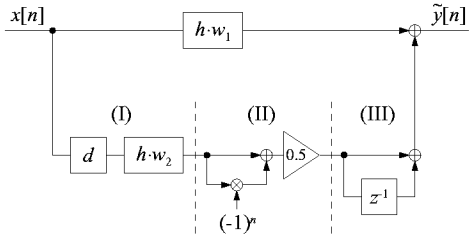

Figure 1: Structure of ID filtering according to (1) and (2): The sample vector $\mathbf{x_2}[n]$ which corresponds to the decimated coefficients is updated only every other time instant.

$$\begin{aligned} W_2(z) &= z^{-N/2} W_1(z) \\ D(z) &= 0.5 + 0.5z \end{aligned}$$

the transform of the ID coefficients is given by

$$\tilde{H}(z) = H(z) \otimes W_1(z) + \tfrac{1}{2}\left(1 + z^{-1}\right) \times$$
$$[D(z)(H(z) \otimes W_2(z)) + D(-z)(H(-z) \otimes W_2(z))]$$

where $\otimes$ denotes cyclic convolution.

To understand how $\tilde{H}(z)$ differs from $H(z)$ we inspect $\tilde{H}(z)$ at very low and very high frequencies. Since $D(z)$ is a low-pass filter it tends to 1 for very low frequencies and $D(-z)$ tends to 0. Hence, it is easily seen that

$$\tilde{H}(z)\Big|_{z\to 1} = H(z) \otimes W_1(z) + H(z) \otimes W_2(z) = H(z). \tag{4}$$

On the other hand, for very high frequencies, i.e. $z \to -1$, $(1 + z^{-1})$ tends to 0 and we get

$$\tilde{H}(z)\Big|_{z\to -1} = H(z) \otimes W_1(z) \tag{5}$$

which corresponds to the windowed undecimated first part of the coefficients. Windowing a time function in this way results in a "smoothing" in the frequency domain.

Eqs. (4) and (5) imply our conclusion that every system with low-pass characteristics and/or a flat transfer function at high frequencies may be approximated effectively by the ID filter. Many real-world systems meet at least one of these requirements. An example is shown in the next Section.

4. APPLICATION EXAMPLE: ACOUSTIC ECHO CANCELLATION

To demonstrate the good performance of the ID filter, an application example is analysed which arises in hands-free telephony. The echo which originates from the acoustic path from the loudspeaker to the microphone must be compensated at the microphone output. The setup for acoustic echo cancellation (AEC) is shown in Fig.2.

AEC was performed with a conventional FIR filter and with the ID filter described above. A 3rd order ID filter was used with N_1 coefficients for the undecimated part, N_2 coefficients for the decimation-by-2 part and N_3 coefficients for the decimation-by-4 path:

$$\tilde{y}[n] = \left[\, \mathbf{h_1}^T\ \mathbf{h_2}^T\ \mathbf{h_3}^T \,\right] \begin{bmatrix} \mathbf{x_1}[n] \\ \mathbf{x_2}[n] \\ \mathbf{x_3}[n] \end{bmatrix} = \tilde{\mathbf{h}}^T \tilde{\mathbf{x}}[n]. \tag{6}$$

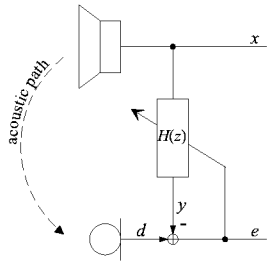

Figure 2: Setup for acoustic echo cancellation.

Hence, the effective filter length and the number of required multiplications is, respectively,

$$N_{\text{eff}} = N_1 + 2N_2 + 4N_3$$
$$M = N_1 + \frac{1}{2}N_2 + \frac{1}{4}N_3.$$

As can be seen from (6) the operation of the ID filter can be written in the same way as for conventional FIR filters. Hence, nearly all algorithms known for FIR systems can also be applied to ID systems, just by exchanging the FIR signal vector $\mathbf{x}[n]$ by $\tilde{\mathbf{x}}[n]$ as indicated in Eq. (6).

The performance of FIR and ID filtering was compared for Least Squares (LS) and Affine Projection (AP). The test signal was the sentence "Mein Name ist Bond, James Bond", Fig.5a, played and recorded at 8 kHz sampling frequency in an office room.

4.1. Least Squares Algorithm

The optimal ID coefficients in a least squares sense are determined in the same way as for FIR systems [2]:

$$\tilde{h} = \Phi^{-1}\theta$$

with

$$\Phi = \sum_n \tilde{\mathbf{x}}[n]\tilde{\mathbf{x}}[n]^T$$
$$\theta = \sum_n \tilde{\mathbf{x}}[n]d[n].$$

Fig.3 compares the Mean Squared Error (MSE) of the ID and the FIR filter with different number of coefficients, based on the same effective filter length and the same number of multiplications, respectively. The normalized MSE is defined as

$$\text{MSE} = 10\log\frac{\sum_n e^2[n]}{\sum_n d^2[n]} \quad \text{(dB)}.$$

Note that for FIR systems the effective filter length equals the number of multiplications. The total number N of ID coefficients was generally divided according to

$$N_1 \approx 0.4N, \quad N_2 \approx 0.4N, \quad N_3 \approx 0.2N$$

although this may not be the optimal choice for every N.

As can be seen in Fig.3, the performance of the ID filter, based on the same effective filter length, is comparable

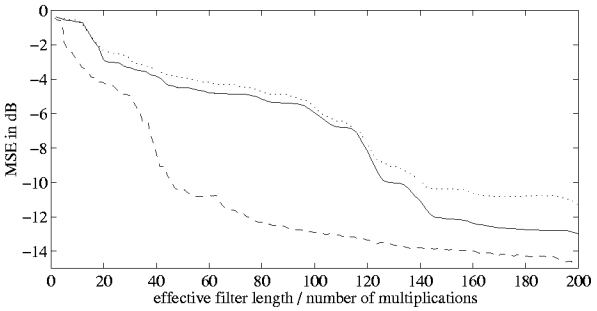

Figure 3: MSE of FIR and ID filters with different number of coefficients. The solid line represents the conventional FIR filter. The performance of ID is compared based on same effective filter length (dotted line) and on the same complexity (dashed line).

to the conventional FIR filter. This indicates the good approximation quality of ID.

More importantly, for the same number of multiplications, i.e. for the same complexity, the ID filter clearly outperforms the FIR filter. Note especially that for a given complexity the performance of the ID filter is *always* better than that of the corresponding FIR filter, regardless of the specific complexity.

As expected, ID is particularly valuable if the processing power is small compared to the required filter length for efficient AEC.

Fig.4 compares the impulse response and the magnitude of the transfer function of the FIR and ID filter, respectively. For both systems the effective filter length equals 200 taps. It can be seen that the echo path does not exhibit a clear low-pass characteristic although some energy is concentrated below 500 Hz. It definitely has no smoothed transfer function at high frequencies. Nevertheless, ID filtering performs quite well.

Generally, the ID filter provides nearly the same transfer function at low frequencies and a smoothed version at high frequencies as compared to the FIR filter. This correlates with the theoretical results of Section 3.

4.2. Affine Projection Algorithm

In the second example, AEC is performed with an adaptive algorithm which is more realistic for real-time processing. The Affine Projection (AP) algorithm was first proposed by Ozeki and Umeda [3]. In terms of complexity and performance AP is placed between Recursive Least Squares (RLS) and Normalized Least Mean Squares (NLMS). The coefficient update is performed according to

$$\tilde{\mathbf{h}}_{n+1} = \tilde{\mathbf{h}}_n + \mu\,\tilde{\mathbf{X}}[n]\left(\tilde{\mathbf{X}}[n]^T\tilde{\mathbf{X}}[n] - \delta\,\mathbf{I}\right)^{-1}\mathbf{e}[n]$$

with

$$\tilde{\mathbf{X}}[n] = [\,\tilde{\mathbf{x}}[n]\;\tilde{\mathbf{x}}[n-1]\;\cdots\;\tilde{\mathbf{x}}[n-D+1]\,].$$

D, the dimension of the projection, was set to 3. μ is the step size and the regulation parameter δ prevents the autocorrelation matrix from becoming singular. $\mathbf{e}[n]$ is a $D \times 1$

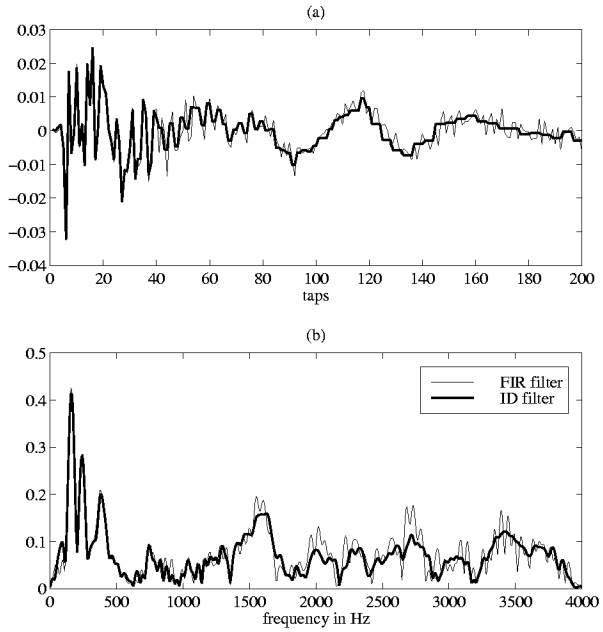

Figure 4: Impulse response (a) and frequency response (b) of the FIR and ID filter. They result from the LS algorithm with an effective filter length of 200.

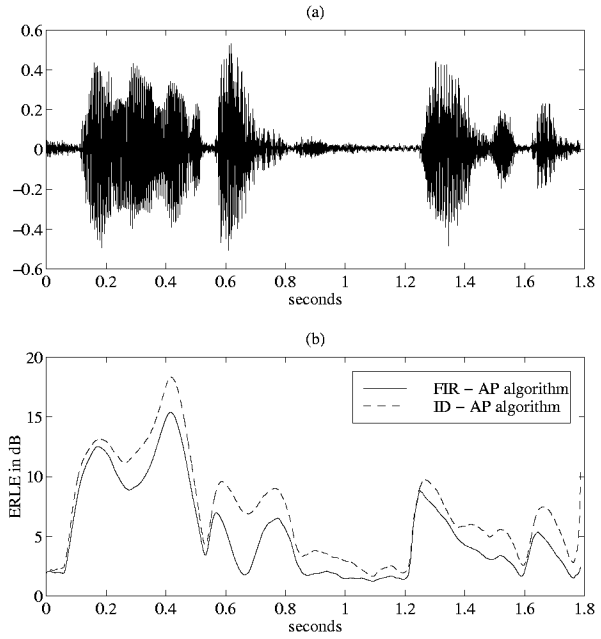

Figure 5: Loudspeaker signal (a) and ERLE (b) for AEC with the AP algorithm in FIR and ID implementation. Both algorithms use the same number of coefficients.

vector of present and past error values (see Fig.2). Note that for $D = 1$ and $\delta = 0$ AP is identical to NLMS.

A measure of performance for AEC is the Echo Return Loss Enhancement (ERLE) which is calculated as

$$\mathrm{ERLE}_n = 10 \log \frac{\sum_{i=-L}^{L} d^2[n-i]}{\sum_{i=-L}^{L} e^2[n-i]} \quad \text{(dB)}.$$

$L = 500$ provides a good average of the current signal energy.

Fig.5a shows the loudspeaker signal which caused the echo. The microphone signal was recorded and compensated with the AP algorithm. In Fig.5b the ERLE of the conventional FIR and ID implementation is compared.

Here, comparison is based on the same number of coefficients ($N = 100$ for FIR and $N_1 = 55, N_2 = 45$ for ID). Again, it can clearly be seen that the ID implementation outperforms the conventional FIR implementation. The reason is the larger effective filter length of the ID system. Note especially that convergence speed does not degrade in the ID implementation as compared to the conventional FIR.

5. DISCUSSION

A new filter structure was proposed which performs implicit decimation and hence, significantly reduces computational requirements or enables the realisation of a larger impulse response length. Because this new filter structure approximates the true impulse response, it can be applied to certain types of systems only, including systems with low-pass characteristics or a smoothed transfer function at higher frequencies.

The performance of the new filter structure was demonstrated in an acoustic echo cancellation application. The room impulse response, which must be modelled in order to compensate its effect, fulfills the above mentioned requirements in most of the cases. The examples show the superiority of the proposed filter compared to conventional FIR implementation.

The advantage of this structure increases further if an algorithm is used with a complexity that is superlinear in the filter length, e.g. RLS. The same is true at higher sampling rates, even for algorithms with a linear complexity. Increasing the sampling frequency by, e.g., the factor of two results in an increase of complexity by the factor of 4 - twice the sampling rate and twice the filter length to realize the same memory. This may be especially interesting in wideband speech transmission at 16 kHz sampling frequency.

6. REFERENCES

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

[2] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, New Jersey, 1986.

[3] K. Ozeki and T. Umeda, "An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties," *Transactions ECEIE*, vol. J67-A, no. 2, pp. 126-132, 1984.