Least–squares Multirate FIR Filters

Roberto Manduchi
Computer Science Department
Stanford University
Stanford, CA 94305
Fax (415)725.1449
manduchi@cs.stanford.edu

Pietro Perona
Dept. of Electrical Engineering
California Institute of Technology
Pasadena, CA 91125
Fax (818)395.2137
perona@vision.caltech.edu

January 30, 1996

Abstract

The authors propose a new least–squares design procedure for multirate FIR filters with any desired shape of the (band–limited) frequency response. The aliasing, inherent in such systems, is implicitly taken into account in the approximation criterion.

1 Introduction

The multirate implementation of FIR filters (see figure 1), introduced by Rabiner and Crochiere [1], makes for reduced computational complexity. In fact, the samples at the output of the FIR filter $g(n)$ that are deleted by the M-fold sampler do not need to be computed, and the null–valued samples introduced by the M-fold interpolator do not contribute to the convolution operated by the FIR filter $h(n)$. Only the case of brick-wall frequency response was considered in [1], and the design technique was inspired by minimax criteria.

We propose a least–squares criterion for the design of multirate FIR filters, to approximate the spectral shape of any desired prototype $d(n)$ (assuming that the necessary band–limiting conditions are met, i.e. that the spectral support of $d(n)$ has length less than $2\pi/M$). The resulting system is linear periodically time–invariant (LPTV [2]), and it is characterized by the M impulse responses $\{t^{(i)}(n+i), 0 \leq i < M\}$, corresponding to the M inputs $\{\delta(n+i), 0 \leq i < M\}$. The fact that the impulse responses differ from each other is usually referred to as *aliasing* effect. The least–squares criterion introduced in this Letter makes for the joint reduction of the approximation error and of the inherent aliasing.

2 Theory

We consider here only the case $M = 2$ (definitions and results are extended straightforwardly to the case of higher M). Define the polyphase components [2] of $g(n)$ as

$$g_0(n) = g(2n)\ ,\ g_1(n) = g(2n+1)$$

One easily shows [2] that

$$t^{(0)}(n) = h * \bar{g}_0(n)\ ,\ t^{(1)}(n+1) = h * \bar{g}_1(n)$$

where $\bar{g}_0(n)$ and $\bar{g}_1(n)$ are obtained by interleaving $g_0(n)$ and $g_1(n)$ with null–valued samples.

We propose the following design criterion: given the kernel $d(n)$ to be approximated, find the filters $g(n)$ and $h(n)$ with given length N_g and N_h respectively, which minimize the approximation error $\mathcal{E}^2$, defined as

$$\mathcal{E}^2 = \frac{\|t^{(0)}(n) - d(n)\|^2 + \|t^{(1)}(n) - d(n)\|^2}{2} \qquad (1)$$

Term $\mathcal{E}^2$ implicitly accounts for both the approximation quality and the aliasing. In fact, if $\mathcal{E}^2$ is small, we may expect both the system's impulse responses to be "close" to $d(n)$, and therefore "close" to each other. More precisely, the following upper bound holds:

$$\|t^{(0)}(n) - t^{(1)}(n)\|^2 \leq 2(\mathcal{E}^2 + \|t^{(0)}(n) - d(n)\|\|t^{(1)}(n) - d(n)\|)$$

No simple closed form solution can be found to the minimization problem, since the error $\mathcal{E}^2$ in (1) is composed of quadratic forms of bilinear expressions in $g(n)$ and $h(n)$. A standard procedure in such cases is based on iterative minimization [3]. Our iterative algorithm is briefly outlined in the remainder. Vectorial notation is used for sequences: a sequence $x(n)$ is represented by a column vector $\mathbf{x}$ whose entries are the samples of $x(n)$. Symbol "^{T}" stands for vector/matrix transposition. We start from an initial guess of $g_0(n)$ and $g_1(n)$, and then iterate through the following two steps:

Optimization of $h(n)$ for fixed $g_0(n), g_1(n)$.

Let $\bar{\mathbf{G}}_0$ and $\bar{\mathbf{G}}_1$ be the Toeplitz matrices representing the filtering with $\bar{g}_0(n)$ and $\bar{g}_1(n)$ respectively. Then

$$\mathcal{E}^2 = \frac{\left(\bar{\mathbf{G}}_0\mathbf{h} - \mathbf{d}\right)^T \left(\bar{\mathbf{G}}_0\mathbf{h} - \mathbf{d}\right) + \left(\bar{\mathbf{G}}_1\mathbf{h} - \mathbf{d}_+\right)^T \left(\bar{\mathbf{G}}_1\mathbf{h} - \mathbf{d}_+\right)}{2}$$

where $\mathbf{d}_+$ is the vector representing $d(n+1)$. Hence, $\mathcal{E}^2$ is minimized for

$$\mathbf{h} = \left(\bar{\mathbf{G}}_0^T\bar{\mathbf{G}}_0 + \bar{\mathbf{G}}_1^T\bar{\mathbf{G}}_1\right)^{-1} \left(\bar{\mathbf{G}}_0^T\mathbf{d} + \bar{\mathbf{G}}_1^T\mathbf{d}_+\right)$$

Optimization of $g_0(n)$ and $g_1(n)$ for fixed $h(n)$.

Let $\mathbf{H}$ be the Toeplitz matrix representing the convolution with $h(n)$, and let $\mathbf{U}$ be a matrix obtained by interleaving the rows of a suitably sized identity matrix with null–valued rows. Then

$$\mathcal{E}^2 = \frac{(\mathbf{HUg}_0 - \mathbf{d})^T (\mathbf{HUg}_0 - \mathbf{d}) + (\mathbf{HUg}_1 - \mathbf{d}_+)^T (\mathbf{HUg}_1 - \mathbf{d}_+)}{2}$$

Error $\mathcal{E}^2$ is minimized for

$$\mathbf{g}_0 = \left(\mathbf{U}^T\mathbf{H}^T\mathbf{HU}\right)^{-1} \mathbf{U}^T\mathbf{H}^T\mathbf{d}\ ,\ \mathbf{g}_1 = \left(\mathbf{U}^T\mathbf{H}^T\mathbf{HU}\right)^{-1} \mathbf{U}^T\mathbf{H}^T\mathbf{d}_+$$

Since the error $\mathcal{E}^2$ does not increase at any iteration and is bounded from below by zero, we are guaranteed to converge to some minimum of $\mathcal{E}^2$. However, the minimum may be just local, and it may be useful to run the algorithm several times with different starting points, choosing the solution that gives the smallest $\mathcal{E}^2$.

3 A design example

We have tested the proposed design technique for a kernel $d(n)$ shaped as the second derivative of a gaussian function, a filter widely used in computer vision (see figure 2). The standard deviation σ was set to 10 and the length of $d(n)$ was 47 samples. The design parameters were: M=4, N_g=N_h=25. The multirate implementation thus requires approximately four times fewer elementary operations per input sample than the direct implementation of $d(n)$. The starting point for the iterative optimization was a constant sequence $g(n)$.

In order to evaluate the multirate system's performance, we may define the *signal to approximation noise ratio*:

$$SNR = \frac{\|d(n)\|^2}{\mathcal{E}^2}$$

and the *signal to aliasing ratio*:

$$SAR = \frac{\|d(n)\|^2}{\max_{i,j}\left\{\|t^{(i)}(n) - t^{(j)}(n)\|^2\right\}}$$

In our case, we obtained SNR=28.3 dB and SAR=25.8 dB.

4 Conclusion

The multirate implementation of band–limited FIR filters makes for the reduction of the computational weight. We have presented a novel least–squares technique to design multirate FIR filters for any shape of the (band–limited) desired frequency response. The technique is based on temporal domain approximation, and the error criterion accounts for both goodness of approximation and aliasing.

References

[1] RABINER, L.R. and CROCHIERE, R.E.: 'A novel implementation for narrow–band FIR filters', *IEEE Trans.*, October 1975, **ASSP-23**, pp.457–464.

[2] VAIDYANATHAN, P.P.: ‘Multirate systems and filter banks’ (Prentice Hall, Englewood Cliffs, NJ, USA, 1993).

[3] GURSKI, G.C., ORCHARD, M.T. and HULL, A.W.: ‘Optimal linear filter for pyramidal decomposition’, *Proc. IEEE ICASSP’92*, 1992, San Francisco, pp. 633–636.

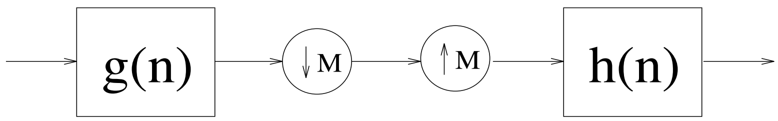
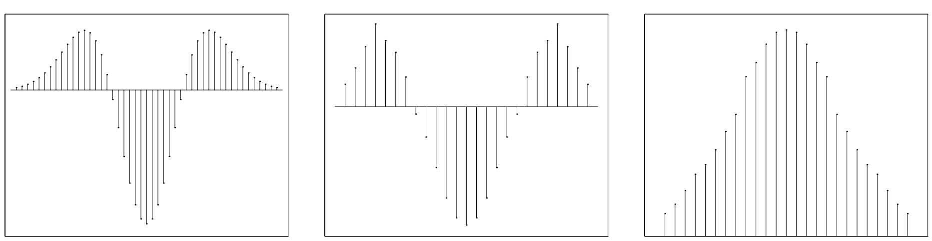


Figure 1: The multirate implementation of a filter.



(a) (b) (c)

Figure 2: The kernel $d(n)$ to be approximated (a), and the filters $g(n)$ (b) and $h(n)$ (c) minimizing $\mathcal{E}^2$.