FA6-10:15

ADAPTIVE FREQUENCY SAMPLING FILTERS

Robert R. Bitmead
Department of Electrical & Electronic Engineering
Post Office, James Cook University, Qld 4811
Australia

Brian D.O. Anderson
Department of Electrical & Computer Engineering
The University of Newcastle, NSW 2308
Australia

SUMMARY

We present two new structures for adaptive filters based on the idea of frequency sampling filters and gradient based estimation algorithms. These filters have a finite impulse response (FIR) and can be thought of as attempting to approximate a desired frequency response at given points on the unit circle. The filters operate in real time with no batch processing of signals as is the case when using the discrete Fourier transform. They result in a marked reduction in dimension of the time-domain problem of fitting an Nth order FIR transversal filter to a collection of length 2 transversal filters and further to a collection of N scalar filters. The advantages of this are then discussed.

INTRODUCTION

Adaptive filters find application in many areas of signal processing including automatic equalization, adaptive antenna arrays, doppler radar systems, adaptive line enhancers, noise and echo cancellers, adaptive control, pattern recognition etc. One particularly familiar and useful approach to the construction of adaptive filters is to choose a suitably parametrised model structure to approximate the dynamic system under investigation and then to select an algorithm to tune or adapt the variable parameters of this model to improve the approximation, possibly in a time-varying environment, as measured by some given criterion. The thrust of the work of this paper will be to present a modelling structure together with appropriate adaption algorithms which for certain classes of problems results in better, more predictable adaptive performance than current schemes.

Perhaps the most familiar structure for an adaptive filter is the tapped delay line or transversal filter, and we shall naturally refer to it as the yardstick for comparison with our structure to be presented. We shall briefly describe the adaptive tapped delay line filter.

Suppose we are given two random sequences $\{x_i\}$ and $\{d_i\}$ and we wish to approximate $\{d_i\}$ as closely as possible by an Nth-order moving average of the x_i. Then, denoting $X_i^T = (x_i, x_{i-1}, \ldots, x_{i-N+1})$ and $W^T = (w_1, w_2, \ldots, w_N)$, we seek a value of W to minimise the criterion $E(d_i - W^T X_i)^2$. This is a familiar Wiener filtering or linear regression problem for which solution procedures have long been studied in certain cases, eg. when the two sequences are wide sense stationary and covariances are known. In the slowly time-varying situation with unknown distributions different procedures to estimate W are used as described below.

There are two immediate points to note about this filter. Firstly, it is FIR (finite impulse response) and indeed the optimum solution for W can be regarded as a best finite impulse response approximation of length N to the system linking $\{x_i\}$ and $\{d_i\}$. Since it is FIR there are no problems of filter stability. Secondly, the performance criterion above is quadratic in W so that simple gradient-based methods for estimating the minimising W can be used. One common gradient-based estimation algorithms is LMS [1].

$$W_{i+1} = W_i + \mu X_i (d_i - X_i^T W_i) \tag{1}$$

where μ is a fixed gain which determines the convergence rate, and hence also the ability to track time variations.

Transversal adaptive filters, with the above structure and algorithm, may be viewed as attempting to find the best FIR approximation to a desired response by directly estimating the values $\{W_i\}$ of the impulse response. These filters then have an obvious "time domain character" about them and this is entirely suited to many applications such as echo cancellers where the resulting impulse response has heuristic meaning. There are other applications however, such as discussed by Griffiths [2], where the impulse response is not of primary interest but rather the frequency response is desired. For this class of problems, where frequency domain information is desired, it may be more sensible to attempt the adaption of the filter in a frequency domain setting. It is this kind of approach to adaptive filtering which we present.

Rather than trying to fit an FIR system to that desired, we approach the problem by estimating the value of the desired frequency response at specified points around the unit circle. In doing this we are freely admitting that we are making an approximation to the ideal system. The course taken is to use fixed frequency sampling filters (FSF) with adaptive gains.

These frequency sampling filters are familiar from FIR digital filter design and will be more fully discussed in section 2. They effectively act as a bank of comb filters, each passing a single narrow band. We process the complex-valued gain of each band independently and, when tuned, the gains approximate the value of the transfer function across each of the bands. This has the effect of breaking down the adaptation problem from a possibly very large dimensional quadratic minimisation to a collection of two-variable quadratic minimisations - the two variables being the real and imaginary (or in-phase and quadrature) parts of the frequency response at the given frequency. We then show how this 2-dimension problem can be decoupled into two scalar minimisations by proper choice of adaptation algorithm. This yields a potentially very fast and predictable convergence rate.

While the idea of fitting wide sense stationary time series in the frequency domain is certainly not new and the use of frequency domain adaptive filters has been proposed and examined by others, see eg. [3-7], one novelty of our approach is the use of frequency sampling filters (FSF) for the separation of frequency components. The advantage of this is that the frequency sampling filters are <u>isochronous</u> with the signal sampling frequency, that is, the FSF produces a real-time output sequence as a digital filter with clock frequency identical to and in step with the input sampling frequency. This is to be contrasted with the FFT which requires batch processing of N time samples of the signal at a time to produce spectrum estimates at discrete frequencies spaced by N^{-1} times the sampling frequency. This real time isochronous operation of the FSFs has led to their output being described as a 'sliding spectrum' [8]. We will show that the use of a sliding spectrum allows us to implement a simple real time adaptation scheme which attempts to minimise a collection of independent scalar unconstrained quadratic problems, as opposed to a constrained vector minimisation as in Waltzman and Schwartz [4]. (While [5,6,7] do not initially require a constrained minimisation, they do if a real valued adaptive filter is required. We note also that appending sequences of zeros to the input to the FFT is necessary to prevent circular convolution. This is noted in [3,6] but is not mentioned in [5,7]). These advantages of our method will be discussed more fully as the presentation progresses and will be drawn together in the conclusion.

FREQUENCY SAMPLING FILTERS (FSF)

We now briefly describe frequency sampling filter and then give some possible implementations of them which require real signals only. FSF are well known in the area of FIR digital filter design [8,9].

Background and Description.

The underlying idea of FSF is that an FIR filter of impulse response length N can be completely described by giving the value of its frequency response at N points equally spaced around the unit circle - this corresponds to specifying the DFT of the filter's impulse response.

Having chosen the N points we then build filters which have zeros at exactly N-1 of these points. Such a filter is

$$H_k(z) = \frac{1-z^{-N}}{1-W_N^k z^{-1}} = \sum_{i=0}^{N-1} W_N^{ki} z^{-i} \qquad (2)$$

where $W_N^k = \exp(j\frac{2\pi k}{N})$, which has zeros at W_N^i for $i \neq k$ and may be considered as having a frequency response which passes only those frequencies centred at and close to W_N^k and excluding others outside this band. As N is made large the frequency response $H_k(e^{j\omega})$ assumes the form of a sinc- or sampling-function in ω - hence the name frequency sampling filter. Banks of these filters centred on different frequencies and with different gains may then be connected in parallel to approximate a desired response. This is further discussed in [6,7]. It should be noted that the transfer function values are exactly matched at the N points.

One immediate benefit of these FSF can be seen in that the output of each elemental filter $H_k(z)$ is approximately the component of the input sequence at frequency $z = \exp(j\frac{2\pi k}{N})$. A bank of these filters then produces a collection of spectral components which approximate those of the input. These components do overlap slightly in frequency but have zero contribution from other filters at the centre frequency. Furthermore the output spectral components are isochronous with the input, ie. they arise in real time as a sequence at the same sampling frequency as the input, and hence have been termed a 'sliding spectrum' as opposed to batch processed FFTs. This is similar in concept to the idea of comb filters in radar systems [10].

Equation (2) gives an indication of two possible methods for the implementation of $\bar{H}_k(z)$. The first method is to construct $\bar{H}_k(z)$ as an (N-1)-length FIR filter via, say, a tapped delay line. For a bank of these elemental filters $\{\bar{H}_k(z)\}$ this requires a tapped delay line for each filter. The second approach is to pass the entire input signal through the common 'comb' filter $1-z^{-N}\alpha^N$ and then in each of the parallel branches to implement a single pole oscillator filter $(1-\alpha W_N^k z^{-1})^{-1}$, where α is a number marginally less than one to guarantee the stability of the oscillators, usually $\alpha = 1-2^{-26}$. Each of these approaches has its obvious advantages.

Implementations for real transfer functions and signals

For transfer function approximation it is necessary to follow each elemental filter with a complex gain equal to the desired value of the frequency response at the particular frequency point on the unit circle. For implementation of these filters for real systems and signals it is desirable to use real operations on real sequences only, and we now show two ways of achieving this. Both methods arise through the collecting together of complex conjugate pairs of elemental filters.

Clearly the FSF which has a real centre frequency, $H_0(z)$, given by

$$H(z) = \frac{1-z^{-N}}{1-z^{-1}} \qquad (3)$$

and which passes the DC component, requires only a real gain to specify the system DC performance. Thus the zeroth elemental filter/gain combination has the form $A_0H_0(z)$ where A_0 is the real gain.

The k-th elemental filter with its associated complex gain has the form

$$(A_k+jB_k)\quad \frac{(1-z^{-N})}{1-W_N^k z^{-1}}$$

where A_k and B_k represent the real and imaginary parts of the transfer function gain at $z = W_N^k$. The (N-k)th elemental filter has centre frequency W_N^{-k}, and the real character of the impulse response implies that the associated complex gain must be A_k-jB_k. Grouping together the conjugate pair we have

$$\frac{1-z^{-N}}{1-W_N^k z^{-1}}(A+jB) + \frac{1-z^{-N}}{1-W_N^{-k} z^{-1}}(A-jB) = \frac{(1-z^{-N})(C+Dz^{-1})}{1-2\cos(\frac{2\pi k}{N})z^{-1}+z^{-2}} \qquad (4)$$

where $C = 2A$ and $D = -2A\cos(\frac{2\pi k}{N})-2B\sin(\frac{2\pi k}{N})$, so that the filter may be constructed by a comb filter $1-z^{-N}$ followed by a second-order (damped) oscillator followed by a two-tap tapped delay line with real weights C and D.

The second method appears more complicated, involving discrete Hilbert Transformers, but will be shown to possess certain advantages. We recognise that the frequency response of an ideal Hilbert Transformer is given by

$$\mathcal{H}(e^{j\omega}) = \begin{cases} -j & 0 \leq \omega < \pi \\ j & \pi \leq \omega < 2\pi \end{cases}$$

so that the left hand side of (4) may be written

$$\left(\frac{1-z^{-N}}{1-W_N^k z^{-1}} + \frac{}{1-W_N^{-k} z^{-1}}\right)(A-\mathcal{H}(z)B) \qquad (5)$$

$$= \frac{(1-z^{-N})(1-\cos(\frac{2\pi k}{N})z^{-1})2}{1-2\cos(\frac{2\pi k}{N})z^{-1}+z^{-2}}(A-\mathcal{H}(z)B)$$

The ideal Hilbert Transformer acts as a 90° phase shifter at all frequencies and is unrealizable as such because its impulse response is non causal and stretches from $-\infty$ to ∞ with odd symmetry. Approximate Hilbert Transformers must therefore be used, the impulse responses of which must approximate the impulse response of the ideal case, left-truncated and right-shifted to become causal. The right shift, in effect the same as a delay, necessitates the incorporation of delays in paths parallel to the transformer. The design of simple wideband bandpass FIR digital Hilbert Transformers has been discussed by Rabiner and Schafer [11] and they tabulate many minimax design data with specifications of passband (at frequencies close to the real axis Hilbert Transformation corresponds essentially to producing the quadrature signal from the in-phase component, which is not a well-posed problem), transfer function approximation error and impulse response length/delay (the delay is half the impulse response length). In simulations these designs with *impulse response* length as low as seven were entirely satisfactory and easily implemented.

ADAPTIVE FREQUENCY SAMPLING FILTERS

Estimation structures and adaptive algorithms

Having specified our chosen modelling structure of a bank of FSFs with pairs of variable real gains, the next step in formulating the adaptive filter is to describe the general estimation structure and devise suitable algorithms for the adaption of the variable gains in the model.

The general estimation structure is illustrated in Figure 1 and we shall refer to this for the definition of symbols, signals etc. Here we have $H_k(z)$ as the function multiplying $(C+Dz^{-1})$ in (4), $(A-\mathcal{H}(z)B)$ in (5) and $H_0(z)$ as above for our collection of real valued FSFs, the alternative depending on whether or not Hilbert Transformers are used. It should be noted that the hypothetical 'Ideal System' $G(z)$ in Figure 1 with input sequence $\{x_i\}$ and output sequence $\{y_i\}$ need not be a tangible physical system on which we are conducting our experiment but could simply be a mental construct representing an abstract linking between the two sequences. Obviously an interpretation of the adaptive filter as an approximate model then needs to be decided.

The aim of the exercise is to choose the variable gains A_k, B_k so that the error between narrow band components $y_{k,i}-\hat{y}_{k,i}$ is small in a certain sense. We treat this as a collection of independent (decoupled) estimation problems, for each of the 2-vectors $(A_k,B_k)^T$ and the scalar A_0, and concentrate without loss of generality on one real elemental filter from the bank of FSF - we shall discuss the DC filter $H_0(z)$ later. The rationale behind this decoupling is that the outputs of the FSF for different k have an approximate orthogonality property by virtue of their being in different narrow frequency bands.*

We shall consider firstly the adaptive implementation of the FSFs using the filters of the first kind above. We call this the <u>direct implementation</u>. We then examine the set up with filters involving the Hilbert Transformer and refer to this as the <u>transformed implementation</u>. As we are studying a single representative component, we shall drop the subscript k from the sequences.

For the direct implementation we recall (4) and write $\{s_i\}$ for the output of the filter

$$(1-z^{-N})(1-2\cos\frac{2\pi k}{N}z^{-1}+z^{-2})^{-1}$$

driven by $\{x_i\}$. Thus we seek to have

$$\hat{y}_i = C\,s_i + D\,s_{i-1}$$

which is equivalent to a simple 2-tap delay line. We may use LMS, (1), to update estimates $\hat{C}_i$, $\hat{D}_i$ of the optima C and D by using $W_i = (\hat{C}_i, \hat{D}_i)^T$, $d_i = y_i$ and $X_i = (s_i, s_{i-1})^T$. This implementation is shown in Figure 2. The most striking comparison between this adaptive filter and the time domain transversal filter is that they both utilise the same update algorithm but the time domain N-vector adaptation is reduced to a collection of $\lfloor N/2 \rfloor$ 2-vector adaptations plus one scalar adaptation. This and its consequences will be discussed more fully after the following presentation of the transformed implementation.

*In case $\{x_i\}$ is wide-sense stationary, the outputs are mutually uncorrelated, other than for effects induced by the overlap of passbands of the nonideal elemental filters. If $\{x_i\}$ is nonstationary, but approximately stationary, the outputs will be only approximately uncorrelated even discounting the overlap of passbands.

Denoting by $\{u_i\}$ the output sequence of the FSF $(1-z^{-N})(2-2\cos\frac{2\pi k}{N}z^{-1})(1-2\cos\frac{2\pi k}{N}z^{-1}+z^{-2})^{-1}$ driven by $\{x_i\}$ and writing $\{v_i\}$ as the Hilbert Transform of the sequence $\{u_i\}$ we have from (5) that ideally

$$\hat{y}_i = Au_i - Bv_i \tag{6}$$

where A and B are respectively the real and imaginary parts of the gain at the particular frequency. This again conforms to a 2-vector identification problem amenable to the LMS algorithm with

$$W_i = (\hat{A}_i,\hat{B}_i)^T,\quad d_i = y_i \text{ and } X_i = (u_i,-v_i)^T \text{ in (1).}$$

As remarked earlier an ideal Hilbert Transformer is an unrealizable filter and approximate transformers require the introduction of a delay (usually half the length of the impulse response for FIR approximations) to try to compensate for the noncausality of the ideal filter. Denoting this delay by ℓ we see that the implementation requires the use of $\hat{y}_{i-\ell} = Au_{i-\ell} - Bv_i$. The transformed implementation is shown in Figure 3.

Performance and Comparison of Structures

We now turn to consider the performance of and comparisons between the direct and transformed FSF approaches and the time domain approach. As stated above the obvious point of comparison between these approaches is the reduction in the dimension of the X vector in the LMS algorithm. The primary effects of this are to alter the convergence rate of the adaptive filter and to allow us to predict this rate more accurately.

When the sequences $\{X_i\}$ and $\{d_i\}$ are random it is straightforward to show that, subject to independence [1] or mild forms of asymptotic independence [12], the LMS algorithm converges exponentially fast to a neighbourhood of the best solution at a rate $(1-\beta)^j$, where β depends roughly linearly on μ (the adaption gain) and the minimum eigenvalue of the covariance

$$R = E[X_iX_i^T]$$

or the time average of this quantity, and the degree of dependence of neighbouring X_i. This property arises basically because these methods are a form of steepest descent procedure to minimize a quadratic form with average weighting R with step size proportional to μ (which in turn must be less than the maximum eigenvalue of R for LMS). These steepest descent or gradient methods typically exhibit slow convergence when the condition number of the quadratic form is large [13]. This dependence of convergence rate upon condition number of steepest descent methods is alleviated in the use of Newton-Raphson algorithms by premultiplying the gradient step by R^{-1}. This latter technique is exemplified in the more rapidly converging Recursive Least Squares algorithms.

When the dimension of X_i is large it is often difficult to estimate both the minimum and maximum eigenvalues of R and so a slow convergence rate often occurs because one naturally chooses a conservative value of μ. Quite apart from these problems of ignorance, ill-conditioning problems become increasingly prevalent with increasing dimension. As the dimension is decreased the severity of these problems decreases until, with a *scalar quadratic minimisation*, the steepest descent procedure is equivalent to Newton-Raphson. The possible advantages of a collection of small dimension minimizations over one large dimension problem then become apparent and we will demonstrate how the proposed implementations allow effective reduction to a collection of N scalar adaptations.

These sorts of arguments are also advanced in [3-7] in support of the use of FFT-based adaption algorithms. However, as [3] makes clear, one may well be facing a constrained optimization problem, a fact which can negate much of the advantage.

The DC FSF requires only a single real gain so that its associated adaptive algorithm is a scalar minimization for both FSF implementations. Consequently the steepest descent methods converge as fast as Newton-Raphson. Furthermore the convergence rate is easily related to the magnitude of the scalar $\{X_i\}$ sequence. We may therefore easily choose an appropriate value for μ. This will be discussed again later.

We now estimate the convergence rate and performance of the 2-vector adaptations of the two adaptive FSF implementations. We use the following notation: C and D, and A and B are the optimum values of the coefficients (assumed stationary for the moment) for the two implementations with a minimum mean square error criterion; $\{n_i\}$ is the minimum mean square error $y_i - Cs_i - Ds_{i-1}$ and $y_i - Au_i + Bv_i$ (delays assumed to be already accounted for); $\tilde{\omega}_i$ is the coefficient estimate error $(\hat{C}_i-C, \hat{D}_i-D)^T$ or $(\hat{A}_i-A, \hat{B}_i-B)^T$.

The adaptation algorithms for the direct implementation are described by the equations

$$\tilde{\omega}_{i+1} = \begin{bmatrix} 1-\mu s_i^2 & -\mu s_i s_{i-1} \\ -\mu s_i s_{i-1} & 1-s_{i-1}^2\mu \end{bmatrix} \tilde{\omega}_i + \mu n_i \begin{bmatrix} s_i \\ s_{i-1} \end{bmatrix} \qquad (7)$$

The associated homogeneous equation for sufficiently small μ is known to be exponentially convergent under various mild conditions, see eg. [12]. Its time constants determine for (7) the time constants associated with convergence of the moments of $\tilde{\omega}_i$. We shall now examine the homogeneous version of (7) to gain insight into what these time constants are.

Our arguments lack full rigour, but are certainly suggestive. Because s_i is a narrow band process, we represent it as

$$s_i = r_i \cos \frac{2\pi ki}{N} + \phi_i \qquad (8)$$

Here, r_i and ϕ_i are slowly-varying processes. The coefficient matrix in (7) can be written as

$$\begin{bmatrix} 1-\mu s_i^2 & -\mu s_i s_{i-1} \\ -\mu s_i s_{i-1} & 1-\mu s_{i-1}^2 \end{bmatrix} = \tfrac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \times$$

$$\begin{bmatrix} 1-\mu(s_i+s_{i-1})^2 & -\mu(s_i^2-s_{i-1}^2) \\ -\mu(s_i^2-s_{i-1}^2) & 1-\mu(s_i-s_{i-1})^2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

where

$$(s_i+s_{i-1})^2 = 4r_i^2 \cos^2\left(\frac{2\pi ki}{N} + \frac{\pi k}{N} + \frac{\phi_i+\phi_{i-1}}{2}\right) \times \cos^2\left(\frac{\pi k}{N} + \frac{\phi_i-\phi_{i-1}}{2}\right)$$

$$(s_i-s_{i-1})^2 = 4r_i^2 \sin^2\left(\frac{2\pi ki}{N} + \frac{\pi k}{N} + \frac{\phi_i+\phi_{i-1}}{2}\right) \times \sin^2\left(\frac{\pi k}{N} + \frac{\phi_i-\phi_{i-1}}{2}\right)$$

$$s_i^2-s_{i-1}^2 = -r_i^2 \sin 2\left(\frac{2\pi ki}{N} + \frac{\pi k}{N} + \frac{\phi_i+\phi_{i-1}}{2}\right) \sin\left(\frac{2\pi k}{N} + \phi_i-\phi_{i-1}\right)$$

If μ is small enough in (7), the fast frequency $(\frac{4\pi k}{N})$ variations in the s_i^2, s_{i-1}^2, $s_i s_{i-1}$ terms are inconsequential, ie. the coefficient matrix can be replaced to a good approximation by its slow-varying part:

$$\begin{bmatrix} 1-\mu s_i^2 & -\mu s_i s_{i-1} \\ -\mu s_i s_{i-1} & 1-\mu s_{i-1}^2 \end{bmatrix}$$

$$\approx \tfrac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1-2\mu r_i^2 \cos^2(\frac{\pi k}{N}) & 0 \\ 0 & 1-2\mu r_i^2 \sin^2(\frac{\pi k}{N}) \end{bmatrix} \times \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

This shows that the time constants associated with (7) are those associated with the scalar equations

$$\lambda_{i+1} = [1-2\mu r_i^2 \cos^2 (\tfrac{\pi k}{N})]\lambda_i \qquad (9a)$$

and

$$\xi_{i+1} = [1-2\mu r_i^2 \sin^2 (\tfrac{\pi k}{N})]\xi_i \qquad (9b)$$

Particularly for small k, ie. frequencies close to DC, the second equation will have a much longer time constant than the first.

We shall now analyse the transformed algorithm, and show that again two equations are obtained; the time constants however are the same and independent of the frequency band.

The transformed algorithm has the following description:

$$\tilde{\omega}_{i+1} = \begin{bmatrix} 1-\mu u_i^2 & \mu v_i u_i \\ \mu v_i u_i & 1-\mu v_i^2 \end{bmatrix} \tilde{\omega}_i + \mu n_i \begin{bmatrix} -u_i \\ v_i \end{bmatrix} \qquad (10)$$

We replace u_i, v_i by $\alpha_i \cos \frac{2\pi ki}{N}$ and $\beta_i \sin \frac{2\pi ki}{N}$, where, if u_i, v_i are gaussian and exact Hilbert transforms, α_i and β_i are also gaussian and band limited to $2\pi/N$. The coefficient matrix in (10) becomes

$$\begin{bmatrix} 1-\mu\alpha_i^2 \cos^2 \frac{2\pi ki}{N} & \tfrac{1}{2}\mu\alpha_i\beta_i \sin \frac{4\pi ki}{N} \\ \tfrac{1}{2}\mu\alpha_i\beta_i \sin \frac{4\pi ki}{N} & 1-\mu\beta_i^2 \sin^2 \frac{2\pi ki}{N} \end{bmatrix}$$

$$\approx \begin{bmatrix} 1-\tfrac{1}{2}\mu\alpha_i^2 & 0 \\ 0 & 1-\tfrac{1}{2}\mu\beta_i^2 \end{bmatrix}$$

with the approximation involving elimination of fast time-variation. The time constants associated with (10) are then those associated with

$$\lambda_{i+1} = (1-\tfrac{1}{2}\mu\alpha_i^2)\lambda_i \qquad (11a)$$

$$\xi_{i+1} = (1-\tfrac{1}{2}\mu\beta_i^2)\xi_i \qquad (11b)$$

and these are the same, since α_i, β_i are identically distributed.

We may use the results of [14] to estimate the time constants of (9a), (9b), (11a) and (11b). These results state that eg. (11a) converges with time constant at least as fast as the deterministic equation

$$\lambda_{i+1} = [1-\tfrac{1}{2}\mu E(\alpha_i^2)]\lambda_i$$

when α_i is ergodic. Generalization to the nonstationary case is possible. The differing time constants of equations (9) and (11) become evident.

The overall effect of the FSF adaptive filters compared to the time domain is to reduce a single N-vector problem to a collection of $\lfloor N/2 \rfloor$ 2-vector problems plus one scalar problem, with each 2-vector problem being equivalent to a pair of 1-vector problems.

Whilst pointing out the advantages of these adaptive FSFs, firstly of the direct implementation over time domain implementation, and then of the transformed implementation over the direct, it is pertinent to reiterate that the dominant dictator of which filter to use is the applications context. For example, if a measure of transient performance is desired by approximating an impulse response then the use of FSF would clearly be unwarranted over the time domain approach. Similarly, the use of the transformed implementation instead of the direct implementation is not necessarily better. Indeed the convergence rate of the direct algorithm for the estimate of C+D may be faster than that of the transformed method - C+D is related to the transfer function magnitude - so that the direct method is to be preferred in some instances.

Simulations of both types of adaptive frequency sampling filter were conducted for a variety of different systems including both those FIR systems which could be exactly modelled by a collection of weighted FSFs and systems (FIR and IIR) which could not. The performance of both adaptive filter implementations were compared for both types of systems above, with and without noise, and with differing gains.

The Hilbert transformers for the transformed implementation were approximated by FIR Minimax approximations of [11]. It was found that approximate transformers of impulse response length 7 were workable for frequency separations of $2\pi/20$ radians and that those with length 15 were entirely adequate.

The gain μ used in both implementations was appropriately scaled to the signal power, as is common, to allow comparison between the convergence rates and it was found that, as predicted, the convergence rate of the transformed implementation was indeed faster than that of the direct implementation especially close to the real axis. The difference between rates decreased as the frequency approached $\pi/2$ radians. Except for systems with transfer functions such as $1-z^{-1}$, which are perfectly describable by the direct implementation, the transformed algorithm appeared to perform more accurately.

There are several effects limiting the accuracy of the transformed adaptive FSF. Firstly, there is model mismatch due to the inability of describing exactly the plant system by a bank of weighted FSFs. This manifests itself as an extra forcing function to the homogeneous parameter update equations resulting in a time-variation of the parameter about the optimum value. The smaller the value μ is, the smaller is the amplitude of this time variation. Allied with this mismatch and contributing time variations possibly with D.C. offset is the inaccuracy of the Hilbert transformer. In the simulations performed this latter effect was negligible, however. Associated with model mismatch is the overlapping of neighbouring frequency sampling filters due to their non-ideal bandpass nature - this too may introduce a time variation and offset in the parameter estimate.

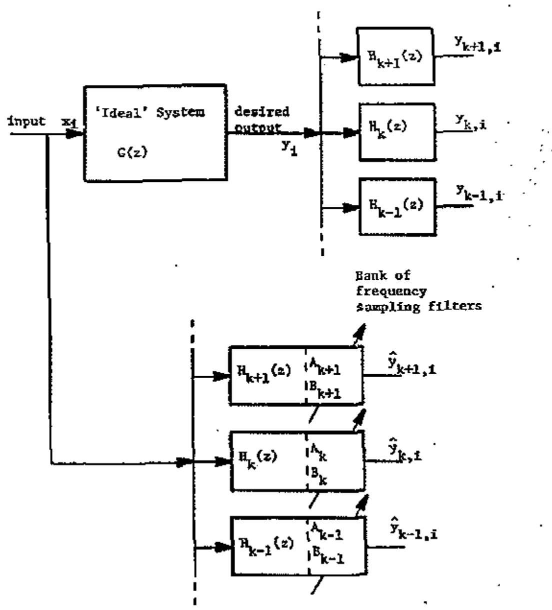
CONCLUSIONS

We have presented tow new adaptive filters based on frequency sampling/sliding spectrum ideas and LMS-type estimation algorithms. The first of these filters - the direct implementation - involves a bank of filters each followed by a separate 2-tap delay line which incorporates the adaptive gains. This was shown to reduce the N-vector transversal filter problem of the time domain to a collection of $\lfloor N/2 \rfloor$ 2-vector transversal filter problems plus a single scalar problem. The second filter - the transformed implementation - involves a similar bank of filters and the use of a Hilbert Transformer at the output of each together with the two adaptive gains. While this latter scheme is more complicated in concept, it allows us to further decouple the 2-vector adaptations so that the single N-vector problem reduces to N scalar ones.

The advantages of this reduction in dimension of the adaptation were shown to be related to the convergence rate, to our ability to predict this rate and performance, to have confidence in our design, and to add more flexibility to our design by choosing different values of parameters for different frequency bands. The convergence rate is related to the eigenvalue properties of a certain matrix having the same dimension as the transversal filter involved. By lowering this dimension through using FSF methods we reduce the likelihood of condition number problems and increase the predictability of the adaptive performance.

Finally it should again be remarked that we do not intend that adaptive FSF's should be used to the exclusion of familiar time domain filters, especially tapped delay line filters, but rather that they should be thought of as another valid alternative possessing certain advantages in many situations.

The choise of adaptive filter should be dictated by the ultimate application.



FSF with adjustable gains

Figure 1.

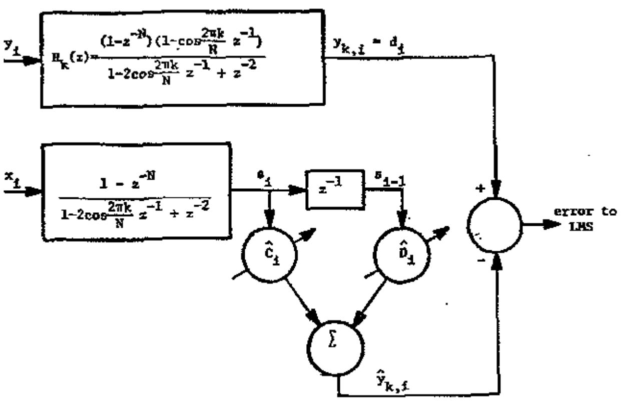


Figure 2. Direct Implementation.

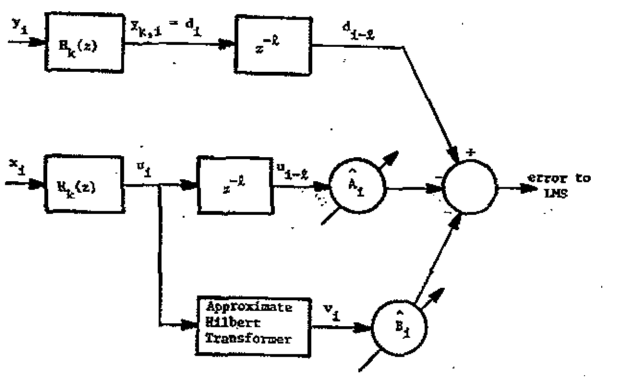


Figure 3. Transformed Implementation

REFERENCES

[1] B. Widrow, J.M. McCool, M.G. Larimore and C.R. Johnson, Jr., 'Stationary and Nonstationary Learning Characteristics of the LMS Adaptive Filter', Proc. IEEE, Vol. 64, No. 8, Aug. 1976, pp. 1151-1162.

[2] L.J. Griffiths, 'Rapid Measurement of Digital Instantaneous Frequency', IEEE Trans. Acoust. Speech and Sig. Proc., Vol. ASSP-23, No. 2, April 1975, pp. 207-222.

[3] T. Waltzman and M. Schwartz, 'Automatic Equalisation using the Discrete Frequency Domain', IEEE Trans. Info. Theory, Vol. IT-19, No. 1, Jan. 1973, pp. 59-68.

[4] W.S. Hodgkiss, 'Adaptive Array Processing: Time vs Frequency Domain', Int. Conf. on Acoust. Speech and Sign. Proc. 1979, Washington DC, April 1979, pp 282-285.

[5] M. Dentino, J. McCool and B. Widrow, 'Adaptive Filtering in the Frequency Domain', Proc. IEEE, Vol. 66, No. 12, December 1978, pp. 1658-1659.

[6] D. Mansour, 'Frequency Domain Adaptive Filter by Overlap-Save Method', Tech. Report, Department of Electrical and Computer Engineering, University of California at Santa Barbara, 1979.

[7] N.J. Bershad and P.L. Feintuch, 'Analysis of the Frequency Domain Adaptive Filter', Proc. IEEE, Vol. 67, No. 12, December 1979, pp. 1658-1659.

[8] B. Gold and C.M. Rader, Digital Processing of Signals, McGraw-Hill Book Co., New York, 1969.

[9] L.R. Rabiner and B. Gold, Theory and Application of Digital Signal Processing, Prentice-Hall Inc., Englewood Cliffs, N.J., 1975.

[10] M.I. Skolnik, Introduction to Radar Systems, McGraw-Hill Book Co. New York, 1962.

[11] L.R. Rabiner and R.W. Schafer, 'On the Behaviour of Minimax FIR Digital Hilbert Transformers', Bell System Tech. Journal, Vol. 53, No. 2, Feb. 1974, pp. 363-390.

[12] R.R. Bitmead and B.D.O. Anderson, 'Performance of adaptive estimation algorithms in dependent random environments', IEEE Trans. Auto. Control, August 1980.

[13] D.G. Luenberger, Introduction to Linear and Non-Linear Programming, Addison-Wesley Publishing Co. Reading MA. 1973.

[14] B.D.O. Anderson and R.R. Bitmead, 'Exponential Stability of a Random Difference Equation', Tech. Report No. 8001, Department of Electrical Engineering, University of Newcastle.

ACKNOWLEDGEMENT

This work was supported by the Australian Research Grants Committee.