11

ACOUSTIC ECHO CANCELLATION USING SUBBAND ADAPTIVE FILTERING

Phillip L. De León II and Delores M. Etter

Department of Electrical and Computer Engineering, University of Colorado, Boulder, CO 80309, USA

1 INTRODUCTION

Adaptive filtering has been successfully used in canceling line echos in the telephone network since the 1960's [179, 443, 387]. In these cases, the adaptive filter is used to model the unknown echo path in a communications channel and then cancel the echo signals from the conversation. An adaptive filter is required because the echo path is not only unknown, but also because it is different for each network path. However, the acoustic echo cancellation problem is not as easily solved with conventional adaptive filtering systems due to the requirements for a much longer adaptive impulse response and for faster convergence since the acoustic echo path is time-varying during the communication transmission. In this chapter, we summarize key results from adaptive filter theory, and discuss the least mean square (LMS) and normalized least mean square (NLMS) algorithms used to adjust an adaptive filter. We then examine adaptive filtering in subbands as one method for acoustic echo cancellation. Simulations (using a measured room response with speech signals) demonstrate that the oversampled subband adaptive system provides improved performance for acoustic echo cancellation applications.

2 LINEAR FIR ADAPTIVE FILTERING

An adaptive system can be characterized by four parts: 1) the adaptive structure being modified, 2) the configuration which contains the adaptive structure, 3) the performance measure used to evaluate the current state of the adaptive structure, and 4) the algorithm used to modify the structure to improve its per-

formance. In this discussion, we focus on the adaptive system used in acoustic echo cancellation.

2.1 The Adaptive Structure

An adaptive FIR filter $\hat{\mathbf{w}}$, shown in Figure 1, is one of the most commonly-used structures for adaptive systems. In an adaptive FIR filter, the filter coefficients (or weights) vary with time, and thus the impulse response and corresponding transfer function of the filter also vary with time.

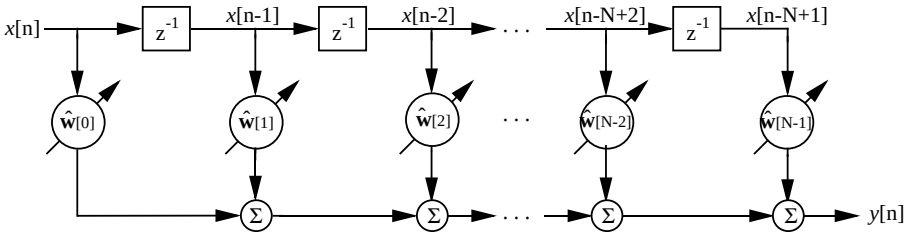

Figure 1 Adaptive FIR filter.

2.2 The Adaptive Configuration

An unknown system $\mathbf{w}$, may be changing over time, may be too complicated to compute directly, or may need to be computed in real-time. In these situations, a system identification or system modeling configuration is used as illustrated in Figure 2. In this configuration, an adaptive filter, $\hat{\mathbf{w}}$ attempts to model the unknown system using the input and the output of the unknown system, $\mathbf{w}$.

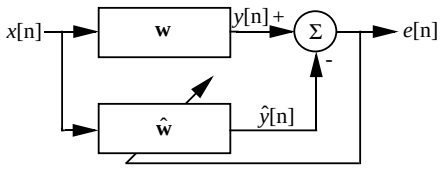

Figure 2 System identification configuration.

To see how the system identification configuration is used in the cancellation of acoustic echos, consider the hands-free teleconferencing system in Figure 3 which shows a speaker and a microphone at each end of the channel. Without echo cancellation systems, a filtered version of the signal from one side of the conversation (for example, originating from A) can pass through the network at B, and is then added to the signal originating from B; this causes an echo to be added to the signal received by A.

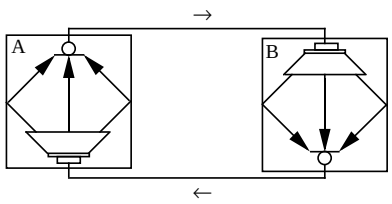

Figure 3 Teleconferencing system.

To eliminate this echo, an adaptive filter is introduced at each end of the teleconferencing system, in parallel to the room echo path, as shown for one end in Figure 4. In this system identification configuration, the adaptive filter, $\hat{\mathbf{w}}$ builds a model of the room echo path, $\mathbf{w}$ and the output from the model is subtracted from the output containing the echos, yielding a signal in the return path that has little or no echo.

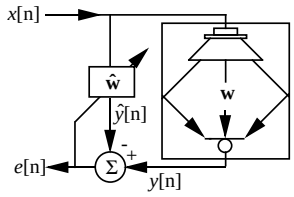

Figure 4 Acoustic echo canceler application.

2.3 The Performance Measure

The objective of the adaptive filter in Figure 2 is to modify an input signal, x such that the resulting signal estimates a desired signal, y. An error signal, e is

computed by subtracting the output of the adaptive filter $\hat{y}$, from y, as shown in Figure 2. If the error signal is large, then $\hat{y}$ is not a good match to y; if the error signal is small, then the output of the filter is a good match to the desired signal. Thus, the error signal provides a *measure* of the ability of the adaptive filter to model the unknown system. The mean-squared error (MSE), ξ, is used as the performance measure for the adaptive filter (instead of just the error itself) because it provides a performance surface that is quadratic with respect to the filter coefficients. This performance surface can be shown to be a function of the correlation of the input signal, x and the cross-correlation of the input and output signals to the unknown system [179, 443]

$$\begin{aligned} \xi[n] &= E[e[n]e^*[n]] \\ &= E\left[(y[n] - \hat{y}[n])(y[n] - \hat{y}[n])^*\right] \\ &= E\left[|y[n]|^2\right] - \hat{\mathbf{w}}^H\mathbf{p} - \mathbf{p}^H\hat{\mathbf{w}} + \hat{\mathbf{w}}^H\mathbf{R}\mathbf{w} \end{aligned} \tag{11.1}$$

where $\hat{\mathbf{w}}$ is a column vector of the adaptive filter coefficients; $\mathbf{R}$ is the input correlation matrix, $E\left[\mathbf{x}[n]\,\mathbf{x}^H[n]\right]$; $\mathbf{p}$ is the cross-correlation vector, $E\left[\mathbf{x}[n]\,y[n]^H\right]$; and ^{H} refers to the conjugate transpose.

It can be shown [443] that the eigenvectors of the input correlation matrix define the principal axes of the error surface and that the corresponding eigenvalues give the second derivatives of the error surface with respect to the principal axes. The effect of these eigenvectors and eigenvalues on the adaptive process will become evident in the next section.

2.4 The Adaptive Algorithm

We now consider algorithms for adjusting the adaptive filter to its optimum, which minimizes the MSE. Since the MSE surface is quadratic, the minimum occurs at the point at which the gradient (derivative of the MSE) with respect to the filter coefficients is equal to zero. Therefore, taking the gradient of the MSE in Equation (11.1), we have

$$\nabla(\xi[n]) = 2\mathbf{R}\hat{\mathbf{w}} - 2\mathbf{p}. \tag{11.2}$$

Assuming that $\mathbf{R}$ is nonsingular, we can then set the gradient to zero and solve for the optimum filter coefficients

$$\hat{\mathbf{w}}_{\text{opt}} = \mathbf{R}^{-1}\mathbf{p} \tag{11.3}$$

where $\mathbf{R}^{-1}$ is the inverse of $\mathbf{R}$. Thus, if $\mathbf{R}$ is known, we can solve directly for the optimum filter coefficients, using Equation (11.3) which is also called the Wiener filter solution.

In many applications, either $\mathbf{R}$ is not known or is time-varying, and thus a direct solution for $\hat{\mathbf{w}}_{\text{opt}}$ may not be feasible or will not be optimal at all times. In these cases, an adaptive algorithm is desired that can start at any initial filter setting, and converge to $\hat{\mathbf{w}}_{\text{opt}}$.

Most adaptive algorithms fall into two categories: gradient-based algorithms (search algorithms that use an estimate of the gradient to determine the direction of the minimum); and recursive least squares (RLS) algorithms (algorithms that use iterative techniques to determine the optimum filter coefficients at any point in time by computing $\mathbf{R}^{-1}[n]$, the inverse of $\mathbf{R}$ at time n). There are many variations of gradient-based algorithms, including steepest descent algorithms and random algorithms. There are also a number of variations of RLS algorithms, including the sequential regression algorithm (SER) and fast RLS algorithms that are able to efficiently estimate $\mathbf{R}^{-1}[n]$.

In the acoustic echo cancellation application, we use a steepest descent algorithm because it is simple and requires fewer computations per iteration than other algorithms. The general form of a steepest descent algorithm is the following

$$\hat{\mathbf{w}}[n+1] = \hat{\mathbf{w}}[n] - \frac{1}{2}\mu\nabla\left(\xi[n]\right) \tag{11.4}$$

where μ is a convergence parameter that controls the size of the step taken at each iteration along the performance surface. The factor $\frac{1}{2}$ is used merely for convenience. If you consider a simple quadratic function of one variable that has only positive function values, a point on the function with a positive gradient (or slope) will be to the right of the minimum, and thus the weight value should be decreased to move toward the optimum weight value. Similarly, a point on this

quadratic function with a negative gradient will be to the left of the minimum, and the corresponding weight value should be increased to move toward the optimum weight value. Thus, in general, the weight increment should be in the direction of the negative gradient, as shown in Equation (11.4).

In real-time processing applications, we usually do not know $\mathbf{R}$ and $\mathbf{p}$, and in addition, they may be changing in time. Thus, we need to use an *estimate* of the gradient. The LMS algorithm [179, 443] uses an instantaneous estimate of the gradient to determine the following algorithm

$$\hat{\mathbf{w}}[n+1] = \hat{\mathbf{w}}[n] + \mu \mathbf{x}[n] e^*[n]. \tag{11.5}$$

Convergence of the LMS algorithm to the optimum weight values depends on the convergence factor, μ. Since the eigenvalues give the steepness of the performance surface, the bound for the step-size (or convergence parameter) is a function of the maximum eigenvalue of the correlation matrix, $\mathbf{R}$ [179, 443]

$$0 < \mu < \frac{2}{\lambda_{max}}. \tag{11.6}$$

Since the eigenvalues are generally not simple to estimate, a practical guideline for the upper bound of the convergence parameter can be developed that uses the fact that the trace of $\mathbf{R}$ is equal to the sum of its eigenvalues. Hence, the maximum eigenvalue is less than or equal to the trace of $\mathbf{R}$ which may be approximated by the filter length times the input signal power. Thus, an upper bound for μ is

$$0 < \mu < \frac{2}{N \text{ (signal power)}}. \tag{11.7}$$

The NLMS algorithm can be viewed as a modification of the LMS algorithm that gives it a time-varying step-size parameter that accounts for variations in the input signal power. The NLMS algorithm can also be developed as a solution to a constrained optimization problem [179] using the method of Lagrange multipliers. In either case, the final form of the NLMS algorithm is

$$\hat{\mathbf{w}}[n+1] = \hat{\mathbf{w}}[n] + \frac{\tilde{\mu}}{a + \|\mathbf{x}[n]\|^2}\mathbf{x}[n]\, e^*[n] \tag{11.8}$$

where a (used to avoid possible division by zero) is greater than 0, $\tilde{\mu}$ is between 0 and 2, and $\|\mathbf{x}[n]\|^2$ is the norm squared or power of the input vector $\mathbf{x}[n]$.

The next section of this chapter presents a subband adaptive filtering solution for acoustic echo cancellation, and compares it to the full-band solution.

3 THE SUBBAND ADAPTIVE FILTERING SYSTEM

The motivation for adaptive filtering in subbands stems from two well-known problems in full-band adaptive filtering. First, the convergence and tracking of a gradient-based adaptive filter can be very slow if the input correlation matrix is ill-conditioned or equivalently if the input signal has wide spectral dynamic range such as that found in speech [179, 171]. Second, very high-order adaptive filters are computationally expensive. One technique used to overcome these limitations is to decompose the signal into subbands and adaptively filter each subband signal using separate adaptive filters. The filter bank is the primary tool used to perform the subband decomposition. The earliest references to this technique used in acoustic echo cancellation applications were Furukawa in 1984; Itoh, Maruyama, Furuya, and Araseki in 1985; and Kellermann in 1985 [387, 209].

In the subband adaptive filter system for a system modeling configuration (Figure 5), the desired output signal, y, and the input signal, x, are split into M subband signals by analysis filters, $\mathbf{h}_0, \cdots, \mathbf{h}_{M-1}$, and downsampled by a factor of D. A bank of adaptive filters, $\hat{\mathbf{w}}_0, \cdots, \hat{\mathbf{w}}_{M-1}$, each adjust themselves (usually using the NLMS algorithm) so as to minimize their mean squared subband error, e_m, which as mentioned earlier, is taken as the square of the difference between the desired subband signal, y_m and the subband adaptive filter output, $\hat{y}_m$. The subband error signals may be used either directly (as in Figure 5) or indirectly to reconstruct a full-band output. Reconstruction of the full-band signal consists of upsampling by a factor of D and filtering with synthesis filters, $\mathbf{g}_0, \cdots, \mathbf{g}_{M-1}$.

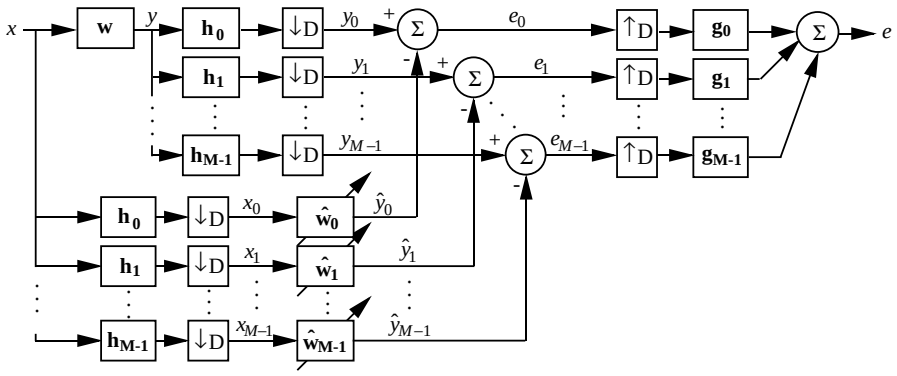

Figure 5 Subband adaptive filtering system (system modeling configuration).

The first benefit of this system results from the fact that the subband adaptive filters are shorter in length than the equivalent full-band adaptive filter (although the total number of adaptive FIR coefficients is usually the same) and operate at a downsampled rate. The lengths of the adaptive filters do not necessarily have to be equal and in fact there may be certain advantages in adjusting the length of each filter to better match the signal characteristics in that band [164]. The second benefit stems from the subband decomposition of the input and desired output signals by the analysis filters. By decomposing the signals, the adaptive filters operate on a smaller bandwidth and can be adjusted to take advantage of this fact. Furthermore, in a critically sampled ($D = M$), subband adaptive filter, the subband decomposition of the input may decrease the spectral dynamic range of the subband input, resulting in faster convergence for gradient-based algorithms. On the downside is the increased design complexity of the technique, the end-to-end delay associated with analysis and synthesis filtering, and the aliasing due to downsampling. In oversampled systems ($D < M$), there is the undesirable potential to increase the spectral dynamic range of the subband input which equivalently increases the eigenvalue disparity ($\lambda_{\max}/\lambda_{\min}$) and in turn, slows the convergence rate of the subband adaptive filter for gradient-based algorithms [279]. All these factors have warranted closer study and analysis in recent years.

4 COMPUTATIONAL COMPLEXITY OF THE SUBBAND ADAPTIVE FILTERING SYSTEM

To gain an understanding of the computational advantage the subband system has over the full-band system, we compute the computational complexity of the M/D oversampled, M-band subband adaptive filtering system (Figure 5), as measured by the number of real multiplications per input sample. This is computed in two parts: the complexity for subband filtering, $C_{\text{subband},1}$ and the complexity for adaptive filtering $C_{\text{subband},2}$.

Assume that x and y are real signals, that analysis and synthesis filtering is implemented with the polyphase uniform DFT filter bank [108], that the prototype analysis/synthesis filter is length L, and that M/D is an integer. Then $C_{\text{subband},1}$ is computed as follows. There are a total of M polyphase filters, each of length L/D operating at a rate of $1/D$ in the filter bank thus requiring $\frac{LM}{D^2}$ real multiplications per input sample. This operation is performed three times: for the analysis filtering of x and y and for the synthesis filtering of $e_0, \cdots, e_{M-1}$. The M-point DFT and IDFT are implemented (assuming M is a power of 2) with a radix-2 FFT which requires approximately $\frac{M}{2}\log_2 M - M$ complex multiplications. For real data, the M-point IDFT can be realized with an $M/2$-point FFT and $M/2$ complex multiplications [104]. This results in $M\log_2\frac{M}{2}$ real multiplications for the analysis filtering of x and y; a similar realization holds for the synthesis filtering of $e_0, \cdots, e_{M-1}$. Thus the total number of real multiplications for subband filtering per input sample is

$$C_{\text{subband},1} = \frac{3LM}{D^2} + 3M\log_2\frac{M}{2}. \tag{11.9}$$

Since the input and desired output signals are real, the DTFT is symmetric. Exploiting this symmetry will require processing of $\frac{M}{2}+1$ of the subbands with subbands $\frac{M}{2}+1, \cdots, M-1$, taken as the respective complex conjugates of subbands $\frac{M}{2}-1, \cdots, 1$. Furthermore, the uniform DFT bank will yield real signals in subbands 0 and $M/2$ and complex signals in the other subbands. Thus two adaptive filters will have real coefficients and $\frac{M}{2}-1$ will have complex coefficients. Assume the length of the impulse response to be modeled is N, each adaptive filter is of length N/D operating at the downsampled rate, and the LMS algorithm (either real or complex) is used for the update. The total number of real multiplications for adaptive filtering per input sample is

$$\begin{aligned} C_{\text{subband},2} &= \frac{2\left(\frac{2N}{D}+1\right)+4\left(\frac{M}{2}-1\right)\left(\frac{2N}{D}+1\right)}{D} \\ &= \left(\frac{2N}{D}+1\right)\left(\frac{2M-2}{D}\right). \end{aligned} \tag{11.10}$$

The complexity for the M/D oversampled, M-band subband adaptive filter system is then taken as the sum of Equations (11.9) and (11.10)

$$\begin{aligned} C_{\text{subband}} &= C_{\text{subband},1} + C_{\text{subband},2} \\ &= \frac{3LM+4MN-4N}{D^2} + \frac{2M-2}{D} + 3M\log_2\frac{M}{2}. \end{aligned} \tag{11.11}$$

Figure 6 contains a plot of the normalized ($C_{\text{subband}}/C_{\text{full-band}}$) subband computational complexity with $L = 129$ and $N = 512$ versus the number of subbands, M, for critically sampled and $2\times$ oversampled systems.

It is clear from the plot (for the parameters used) that critically sampled subband systems with 4, 8, 16, or 32 subbands (in the derivation we assumed M is a power of 2) are computationally more efficient than the equivalent full-band system. Furthermore, $2\times$ oversampled subband systems are more efficient with 16 or 32 subbands.

5 EXPERIMENTAL SETUP

The experimental results presented in this chapter were conducted in software on a general purpose computer and use both full-band and subband adaptive system modeling configurations for comparison. The unknown system, $\mathbf{w}$, is an actual impulse response for a room, sampled at 8 kHz and truncated to 512 samples. In all experiments, $M = 4$ and efficient polyphase, uniform DFT filters are used for analysis and synthesis filtering [108]. The analysis filter prototype was designed using the Parks-McClellan algorithm and is of length 129. The synthesis filters were the same as the analysis filters. Magnitude responses for these filters are illustrated in Figure 7.

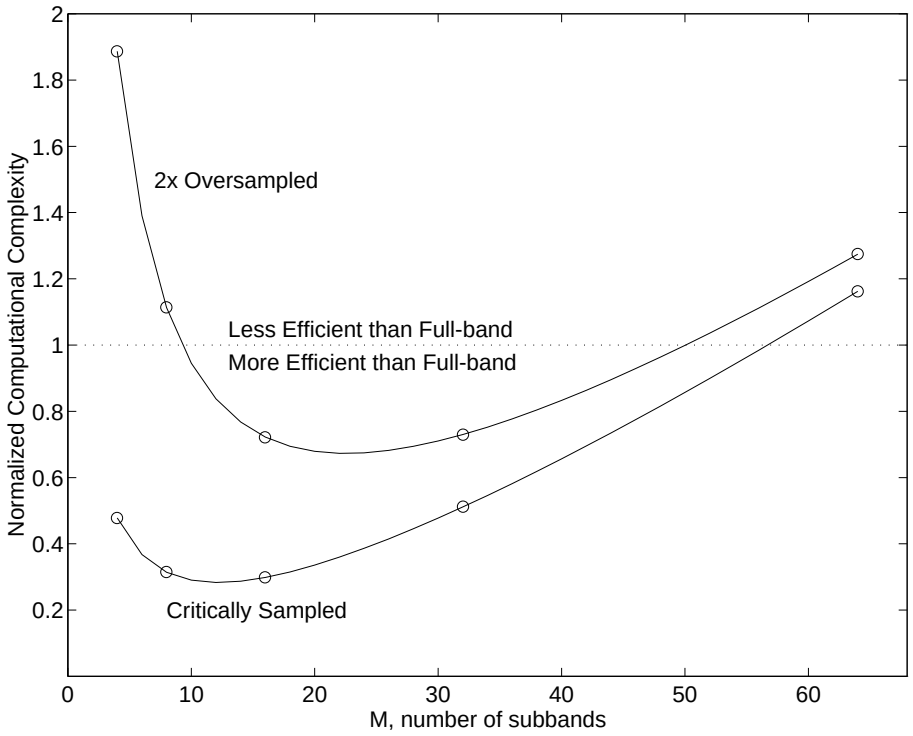

Figure 6 Normalized computational complexity for the subband adaptive filtering system.

The adaptive filters were all of length $\lceil \frac{512}{D} \rceil$ and use the NLMS algorithm for the update with $a = 0.01$ and $\tilde{\mu} = 1/3$. Adaptive filtering was done only in subbands 0, 1, and 2 since $y_3, \hat{y}_3$ were taken as $y_1^*, \hat{y}_1^*$ respectively, due to the complex conjugate symmetry of the uniform DFT filter bank. The adaptive filters in subbands 0 and 2 were real and in subband 1 was complex. Zero-mean, unit-variance white Gaussian noise input was used to measure the convergence of both the full-band and subband systems, and experiments were averaged over 100 simulations. MSE plots were smoothed with a 100 point moving average filter and 2× oversampled subband MSE plots were smoothed with a 50 point moving average filter. Power spectrum plots were obtained from the last 1024 points of the hanning-windowed error signal. In a separate experiment, speech input was also used to measure the performance. The speech signal was composed of the utterance "Zero One Two Three Four Five Six Seven Eight Nine", sampled at 8 kHz, and is shown in Figure 8.

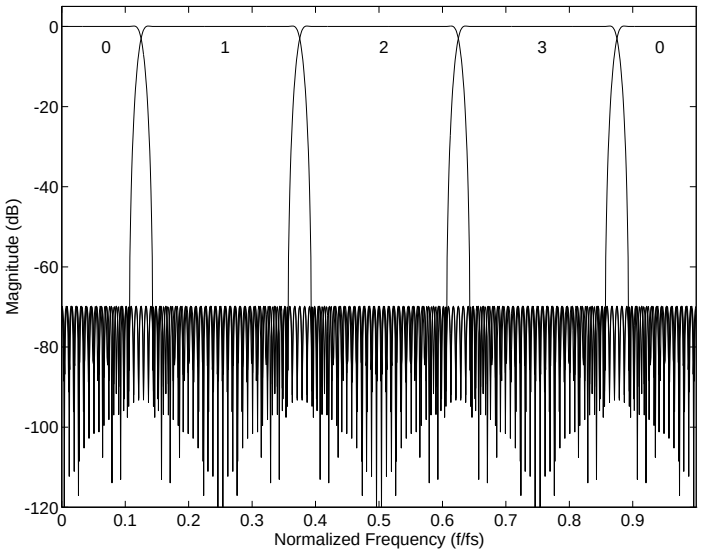

Figure 7 Polyphase, uniform DFT analysis and synthesis filters used in experiments.

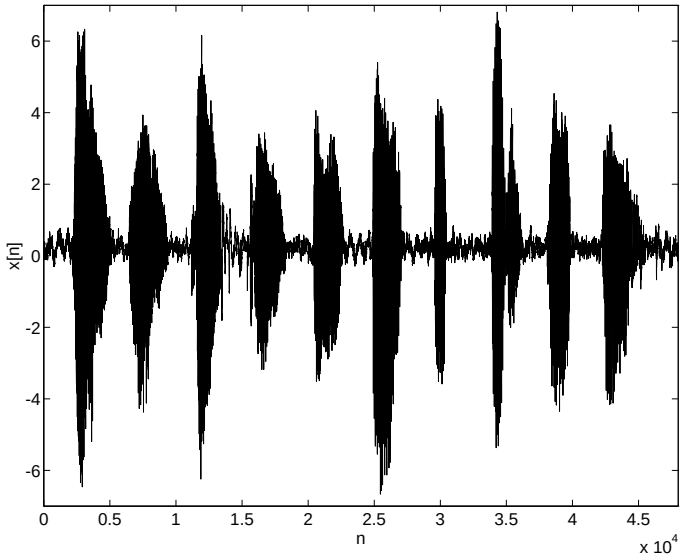

Figure 8 Input speech signal.

6 CRITICALLY SAMPLED SYSTEMS

The first experiments in subband adaptive filtering used critically sampled subbands for maximum efficiency. The primary result of these experiments is the gain in computational efficiency. However, as can be seen in Figure 9, the subband MSEs quickly level off due to the inability to properly model the aliasing, which is present in the critically sampled subbands. The point at which convergence levels off is usually referred to as the asymptotic level.

The high asymptotic levels in the subband MSEs, result in a high asymptotic level for the reconstructed full-band MSE, as shown in Figure 10. Included on this plot for comparison is the MSE for the equivalent full-band system. During the initial convergence the subband system performs equally well as compared to the full-band system. However, as mentioned earlier, the effects of aliasing in the subbands hinders the ability to properly model and the performance tapers off.

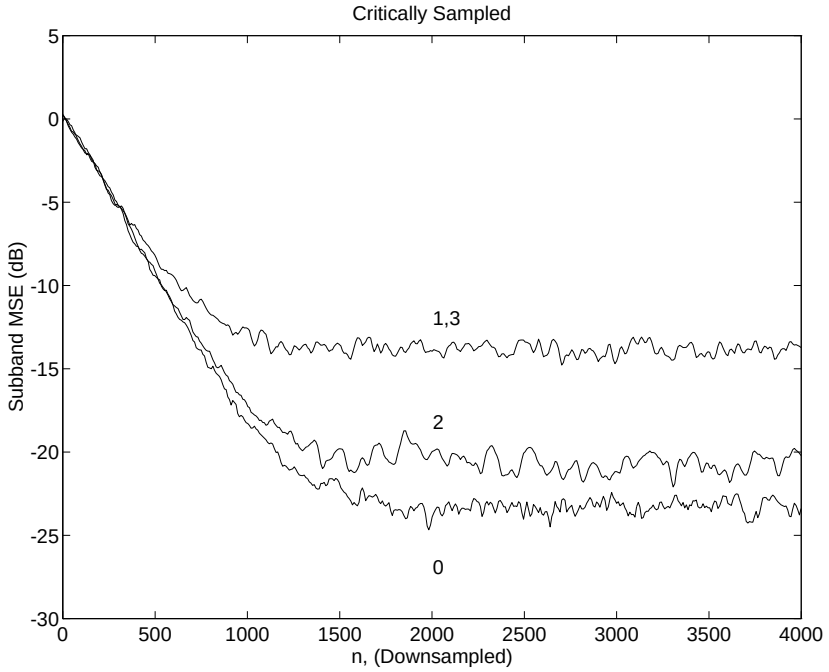

Figure 9 Subband MSEs for the critically sampled subband adaptive filtering system.

Spectral analysis of the reconstructed full-band error signal (Figure 11) shows marked spectral peaks at the crossover frequencies (corresponding exactly to the aliased parts of the spectrum) in the filter bank indicating the inability of the system to properly model in these frequency regions.

We next compare the performance of the critically sampled subband system with the full-band system using speech as the input. As mentioned earlier, the subband system performs a spectral decomposition of the input and for speech, this should be beneficial. Figure 12 illustrates the MSE of both systems and the results show very comparable performance between the two systems.

In an attempt to retain the critical sampling rate and have more accurate modeling (smaller output error), a slightly different subband system was proposed and is illustrated in Figure 13 [165].

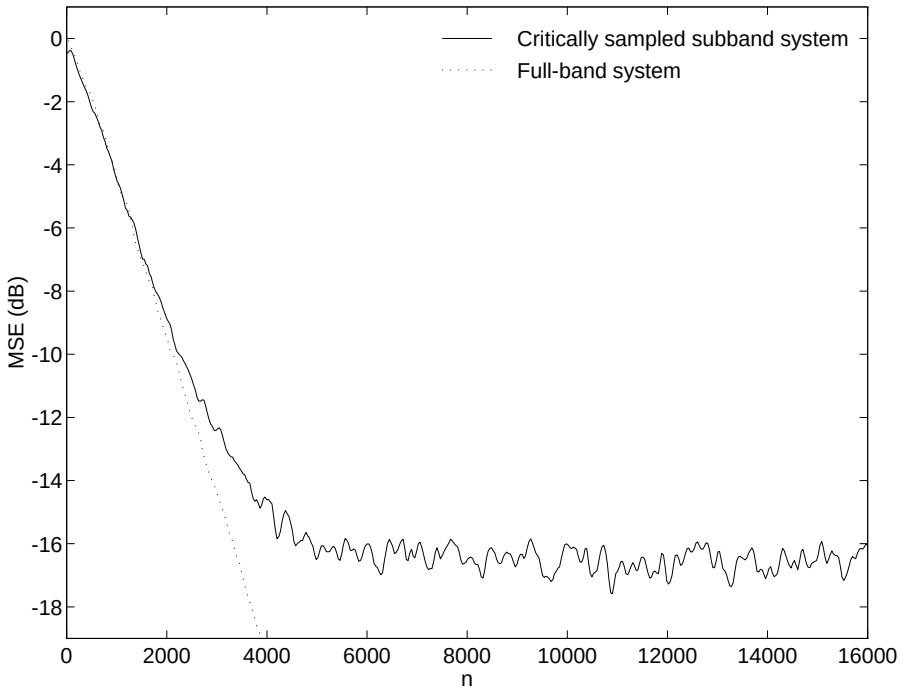

Figure 10 MSE for full-band and critically sampled subband adaptive filtering systems.

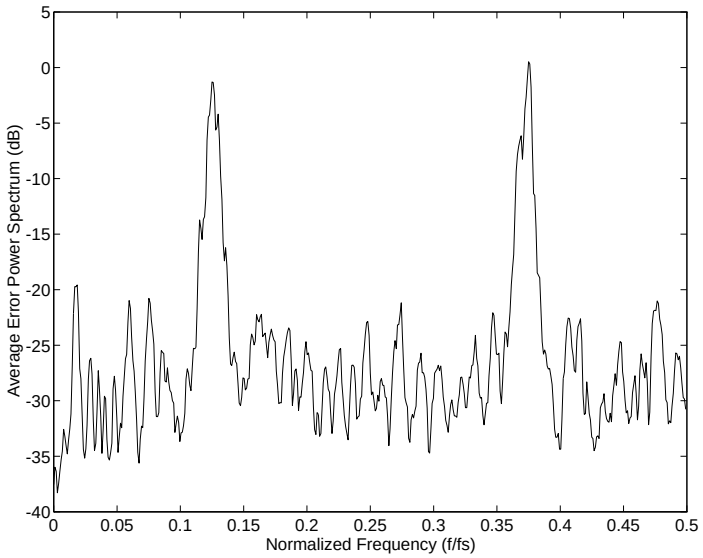

Figure 11 Average power spectrum of the error signal for the critically sampled subband adaptive filtering system.

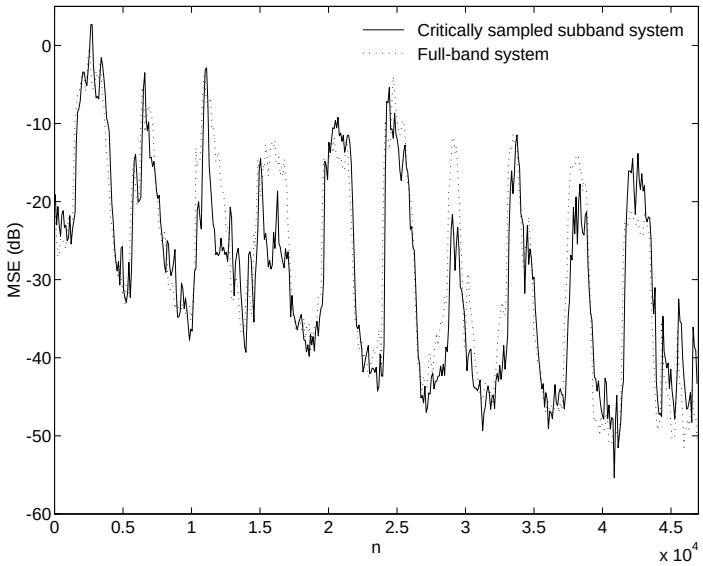

Figure 12 MSE for full-band and critically sampled subband adaptive filtering systems under speech input.

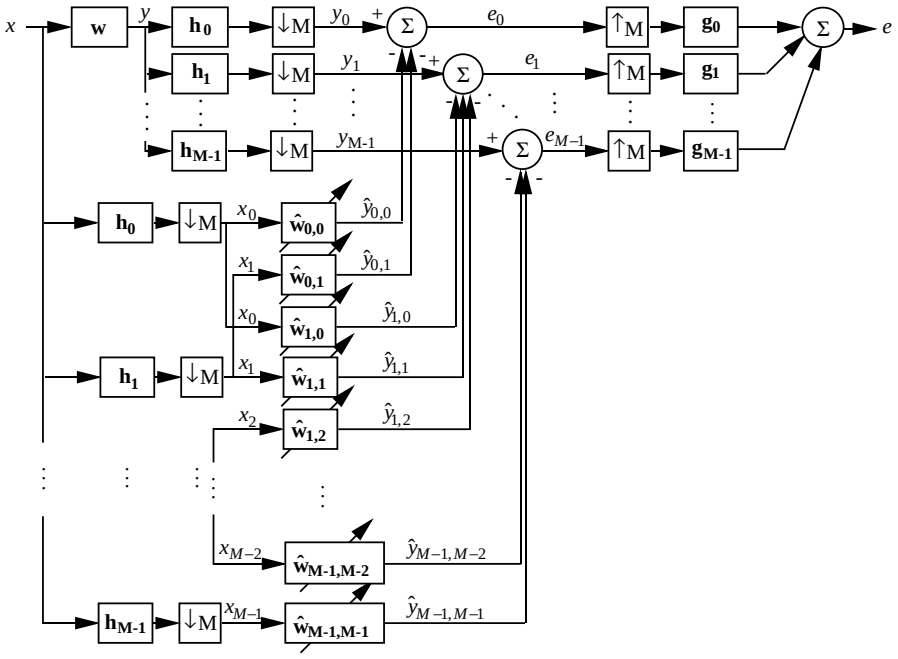

Figure 13 Critically sampled subband adaptive filtering system with adaptive cross filters.

In this system, adaptive cross filters are used in which the input is from one subband and the error signal used for update is from another subband. In experiments with input noise with a speech-like spectrum (USASI noise), it was demonstrated by Gilloire and Vetterli [165] that the critically sampled, eight-band subband system with adaptive cross filters had a lower asymptotic level than the two-band subband system without cross filters. However, there was no significant gain in the convergence rate for either system as compared to the full-band system. In a separate experiment with white noise input, a two-band subband system with adaptive cross filters and perfect reconstruction filter banks achieved perfect modeling, although the convergence rate was slower. It was ultimately concluded, however, that oversampled systems provided the best method of adaptive filtering in subbands, both in terms of convergence and computational efficiency [165]. In the next section, we present results from oversampled subband systems.

7 OVERSAMPLED SYSTEMS

Due to the inability of critically sampled subband systems without cross filters to adequately model in the presence of aliasing, we consider oversampling the subband signals as a means of eliminating the effects of aliasing. This of course, reduces the computational efficiency as compared to the critically sampled system. In these experiments, the subbands are $2\times$ oversampled and from a practical standpoint, this is more oversampling than actually needs to be done. As can be seen in Figure 14, the subband MSEs do not have the characteristic high asymptote that the critically sampled subband MSEs have, but nonetheless exhibit an asymptotic convergence after a fast initial convergence [279].

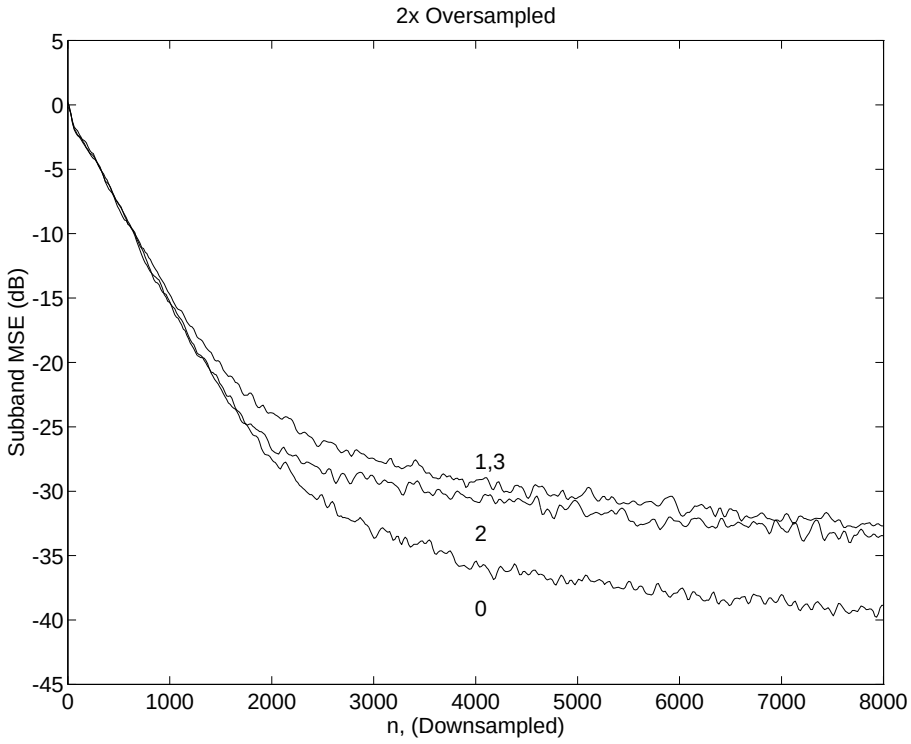

Figure 14 Subband MSEs for $2\times$ oversampled subband adaptive filtering system.

Consequently, the reconstructed full-band MSE for the oversampled subband system (Figure 15) exhibits this slow asymptotic convergence as well. Included on this plot for comparison is the MSE for the full-band system. During the initial convergence the subband system performs better than the full-band system but is less effective afterward. The oversampled subband system does, however, have better convergence as compared to the critically sampled system. In an environment where $\mathbf{w}$ is changing over a period of time shorter than the initial convergence period (as in the case of acoustic echo cancellation), initial convergence will most affect cancellation quality.

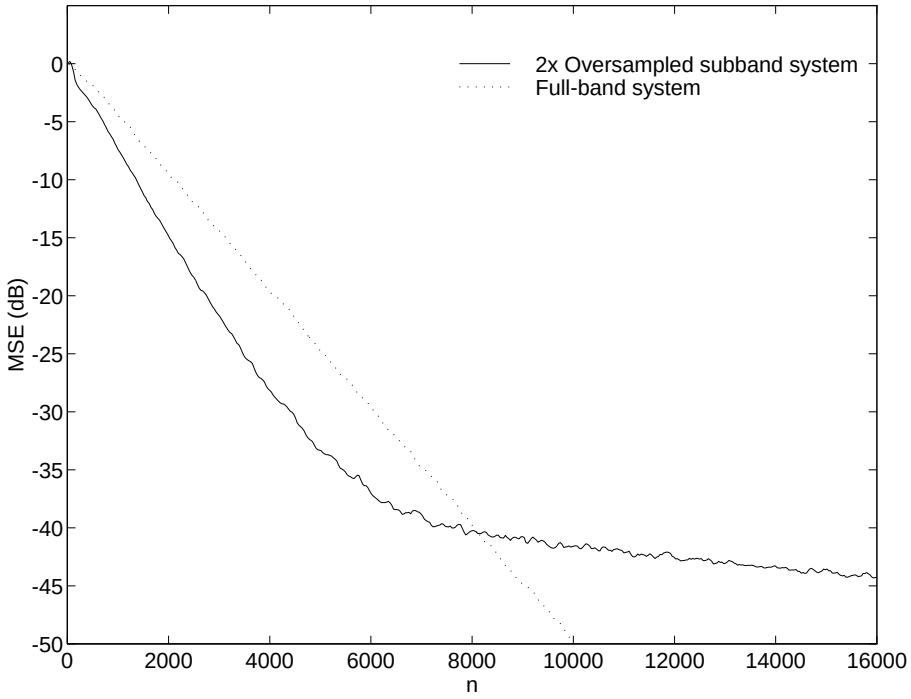

Figure 15 MSE for full-band and 2× oversampled subband adaptive filtering systems.

Spectral analysis of the error signal (Figure 16) also shows marked peaks located in the vicinity of the crossover frequencies. These peaks are not due to aliasing but rather due to the small eigenvalues associated with the analysis filter roll-off in the oversampled case. This characteristic is the basis of the theory of slow asymptotic convergence of LMS acoustic echo cancelers [279, 119].

Figure 17 contains the MSE plot for the full-band and subband systems with speech input. In this plot, the oversampled subband system achieves 0-20 dB increase in convergence over the full-band system as well as the critically sampled subband system but asymptotic convergence is still present.

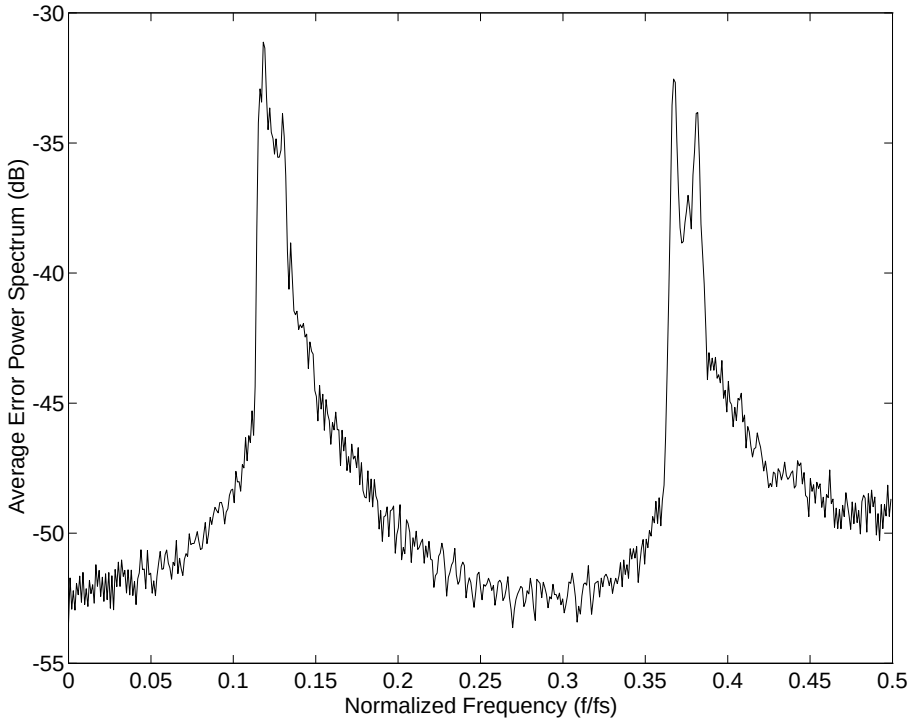

Figure 16 Average power spectrum of the error signal for the $2\times$ oversampled subband adaptive filtering system.

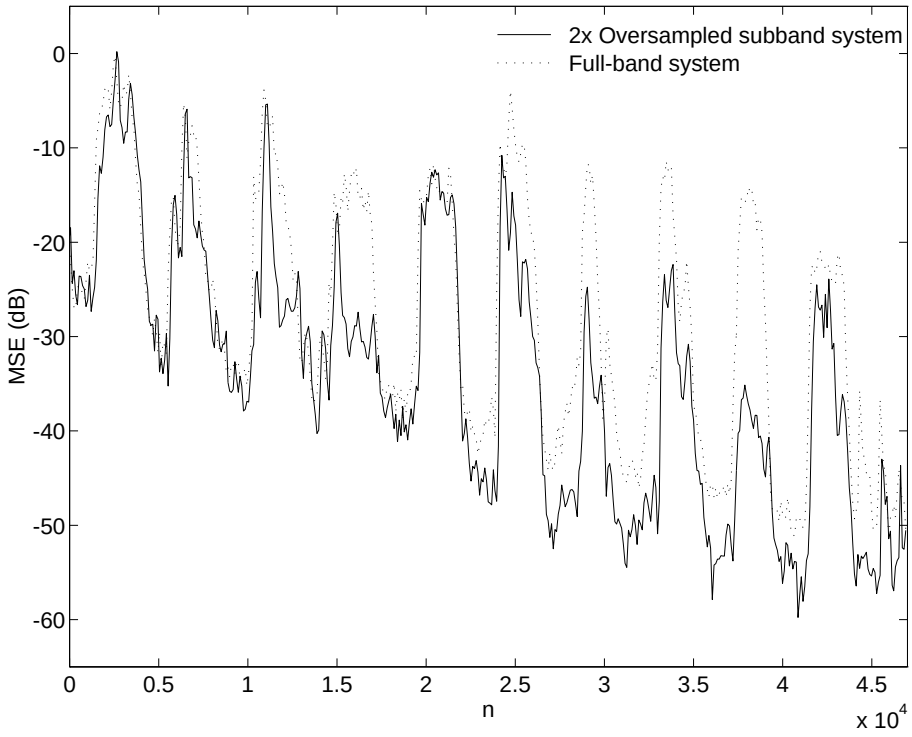

Figure 17 MSE for full-band and 2× oversampled subband adaptive filtering systems under speech input.

8 CONCLUSION

In this chapter we have seen how using multiple adaptive filters in a filter bank can reduce the computational requirements of a very high-order adaptive filter and reduce the effects of disparate eigenvalues. The performance in terms of MSE under white noise input is comparable to the full-band system during the initial convergence but degrades due to aliasing effects in the critically sampled system and slow asymptotic convergence effects in the oversampled system. Under speech input, the critically sampled subband system performs comparably to a full-band system. The 2× oversampled subband system with speech input, however, offers as much as a 20 dB reduction in MSE (in our experiments) when compared to the full-band system or critically sampled subband system.

FURTHER READING

J. Chen, H. Bes, J. Vandewalle, and P. Janssens, "A New Structure for Sub-band Acoustic Echo Canceler," in *Proc. IEEE ICASSP '88* (New York, NY) pp. 2574-2577.

A. Gilloire and M. Vetterli, "Adaptive Filtering in Sub-bands," in *Proc. IEEE ICASSP '88* (New York, NY) pp. 1572-1575.

W. Kellermann, "Analysis and Design of Multirate Systems for Cancellation of Acoustical Echoes," in *Proc. IEEE ICASSP '88* (New York, NY) pp. 2570-2573.

S. Mitra, M. Petraglia, and A. Mahalanobis, "Structural Subband Implementation of Adaptive Filters," in *Proc. 24th Asilomar Conference* (Pacific Grove, CA), Nov. 1990, pp. 232-236.

J. Shynk,"Frequency-domain and Multirate Adaptive Filtering," IEEE ASSP Magazine, vol. 9, no. 1, pp. 14-37, Jan. 1992.

H. Yasukawa, S. Shimada, and I. Furukawa, "Acoustic Echo Canceller with High Speech Quality," in *Proc. IEEE ICASSP '87* (Dallas, TX) pp. 2125-2128.