The Time-Sequenced Adaptive Filter

EARL R. FERRARA, JR., MEMBER, IEEE, AND BERNARD WIDROW, FELLOW, IEEE

Abstract—A new form of adaptive filter is proposed which is especially suited for the estimation of a class of *nonstationary* signals. This new filter, called the time-sequenced adaptive filter, is an extension of the least mean-square error (LMS) adaptive filter. Both the LMS and time-sequenced adaptive filters are digital filters composed of a tapped delay line and adjustable weights, whose impulse response is controlled by an adaptive algorithm. For stationary stochastic inputs the mean-square error, which is the expected value of the squared difference between the filter output and an externally supplied "desired response," is a quadratic function of the weights—a paraboloid with a single fixed minimum point which can be sought by gradient techniques, such as the LMS algorithm. For nonstationary inputs however the minimum point, curvature, and orientation of the error surface could be changing over time. The time-sequenced adaptive filter is applicable to the estimation of that subset of nonstationary signals having a recurring (but not necessarily periodic) statistical character, e.g., recurring pulses in noise. In this case there are a finite number of different paraboloidal error surfaces, also recurring in time.

The time-sequenced adaptive filter uses multiple sets of adjustable weights. At each point in time, one and only one set of weights is selected to form the filter output and to be adapted using the LMS algorithm. The index of the set of weights chosen is synchronized with the recurring statistical character of the filter input so that each set of weights is associated with a single error surface. After many adaptations of each set of weights, the minimum point of each error surface is reached resulting in an optimal *time-varying* filter. For this procedure, some *a priori* knowledge of the filter input is required to synchronize the selection of the set of weights with the recurring statistics of the filter input. For pulse-type signals, this *a priori* knowledge could be the location of the pulses in time; for signals with periodic statistics, knowledge of the period is sufficient.

Possible applications of the time-sequenced adaptive filter include electrocardiogram enhancement and electric load prediction.

I. INTRODUCTION

A SIGNAL-enhancing technique for statistically stationary signals based on conventional least mean-square (LMS) adaptive filtering has been proposed [1]. This technique yields a substantial reduction in background noise but often at the expense of considerable signal distortion at moderately low signal-to-noise ratios (SNR's). This paper demonstrates that by modeling certain signals as nonstationary stochastic processes, an optimally time-varying adaptive filter may be derived which can significantly outperform the comparable LMS adaptive filter. Applications of this technique include electrocardiogram enhancement and electric load prediction.

Manuscript received March 25, 1980; revised December 10, 1980. This work was supported by the National Science Foundation under Grant ECS-7808526, by the Office of Naval Research under Contract N00014-76-C-0929, and by the Naval Air Systems Command under Contract N00019-80-C-0483.

E. R. Ferrara, Jr., was with Stanford University, Stanford, CA. He is now with ESL Inc., Sunnyvale, CA 94086.

B. Widrow is with Stanford University, Stanford, CA 94305.

An adaptive transversal filter consists of a tapped delay line connected to an adaptive linear combiner that adjusts the gain of (or "weights") the signals derived from the taps of the delay line and combines them to form an output signal. A minor extension of the adaptive transversal filter includes a separate input which is fixed. This input is multiplied by a "bias weight" and then summed with the other weighted signals to form the output. A bias weight is often used when the filter input and desired response have nonzero means.

The input signal vector X_j of the adaptive linear combiner is defined as

$$X_j^T = \begin{bmatrix} 1 & x_j & x_{j-1} \cdots x_{j-(n-1)} \end{bmatrix}^T. \tag{1.1}$$

The input signal components are assumed to appear simultaneously on all input lines at discrete times indexed by the subscript j. The weighting coefficients or multiplying factors $w_0, w_1, \cdots, w_n$ are adjustable. The weight vector W is

$$W^T = [w_0 \quad w_1 \quad w_2 \cdots w_n]^T. \tag{1.2}$$

The output y_j is equal to the inner product of X_j and W:

$$y_j = X_j^T W = W^T X_j. \tag{1.3}$$

The error ϵ_j is defined as the difference between the desired response d_j (an externally supplied input sometimes called the "training signal") and the actual response y_j:

$$\epsilon_j \triangleq d_j - X_j^T W = d_j - W^T X_j. \tag{1.4}$$

In adaptive filtering applications the desired response is usually composed of some underlying signal to be estimated plus additive noise uncorrelated with both the signal and the filter input.

Assume that the sequence of pairs $\{(d_j, X_j)\}_{j=1}^{\infty}$ is a stochastic process which need not be stationary. The expectations in this paper will be taken over the ensemble described by this stochastic process. The correlation matrix at time j, defined by

$$R_j = E\left[X_j X_j^T\right] \tag{1.5}$$

is assumed to be positive definite. The cross-correlation vector is defined by

$$P_j = E\left[d_j X_j\right]. \tag{1.6}$$

We will be interested in the mean-square error at time j,

0096-3518/81/0600-0679$00.75 © 1981 IEEE

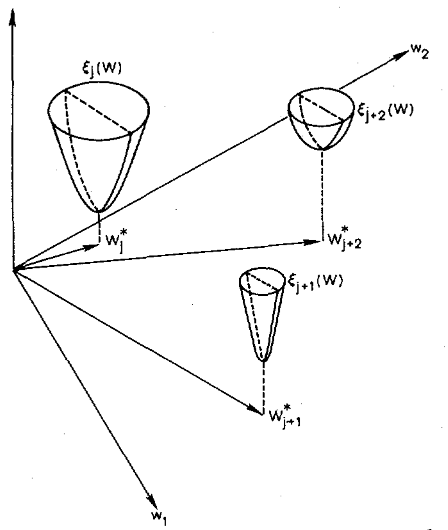


Fig. 1. Time-varying error surface for a nonstationary process.

given by

$$\xi_j = E\left[\left(d_j - W_j^T X_j\right)^2\right]$$
$$= E\left[d_j^2\right] - 2W_j^T P_j + W_j^T R_j W_j. \tag{1.7}$$

It can be shown [3] that the optimal weight vector W_j^* which minimizes the mean-square error is given by

$$W_j^* = R_j^{-1} P_j. \tag{1.8}$$

This vector W_j^* will be called the "Wiener weight vector" at time j. If R_j and P_j were known, performance could be optimized by choosing W_j according to (1.8). More often, however, these statistics are not known and an adaptive approach must be used for approximating W_j^*. Note that the error surface (1.7) is a quadratic function of the weight vector at any particular time and can be pictured as a concave hyperparaboloidal surface, a function that never goes negative. Choosing $W_j = W_j^*$ corresponds to operation at the minimum of the error surface at time j. With nonstationary inputs, the minimum point, orientation, and curvature of the error surface could be changing over time, as shown in Fig. 1. If, however, the desired signal and input signal vectors are jointly stationary then the statistics R_j and P_j are constant, and in accord with (1.7) only a single error surface need be considered. In this case gradient search techniques may be used to find the minimum. One method that has proven to be very useful is the Widrow–Hoff LMS algorithm [3]–[5], based on the method of steepest descent [6], [7]. According to this method, the "next" weight vector W_{j+1} is equal to the "present" weight vector W_j plus a change proportional to an estimate of the negative gradient of the error surface:

$$W_{j+1} = W_j + 2\mu\epsilon_j X_j. \tag{1.9}$$

Signals composed of recurring pulses in noise are highly nonstationary due to their time-varying statistical character. The LMS adaptive filter, being unable to track such rapidly varying nonstationarities, would essentially converge to the best time-invariant filter. For these signals an adaptive filter which could exhibit a rapidly varying impulse response may perform in a vastly superior fashion to the LMS adaptive filter. The utility of time-variable filtering to electrocardiographic signals has been suggested in [8].

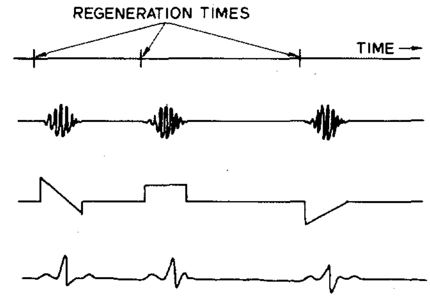


Fig. 2. Statistically recurring signals.

In Section II an extension of the LMS adaptive filter better able to track rapidly varying nonstationarities is proposed. This new filter, called the "time-sequenced adaptive filter," converges to a time-varying solution and is especially suited to the filtering of signals in noise having a recurring (not necessarily periodic) statistical nature. A comparison of optimal time-varying and time-invariant filtering is presented for randomly shaped pulses in noise. Applications are considered in Section III.

II. The Time-Sequenced Adaptive Filter

This section describes an extension of the LMS adaptive filter that allows the weight vector to change freely in time in order to accommodate rapid changes in the statistics of a certain class of nonstationary signals, while allowing slow precise adaptation. The signals to be considered are those whose statistical properties recur at various points in time, called regeneration times. In particular we require that the autocorrelation matrix R_j and cross-correlation vector P_j at any particular time are elements of some finite set and that they occur in identical sequence after each regeneration time. The times between regenerations are allowed to be variable. Thus the *entire* sequence of R matrices and P vectors will not in general be used each cycle, since the occurrence of a regeneration starts the sequence over. Examples of signals which may be modeled as statistically recurring are shown in Fig. 2 and include electrocardiograms, radar signals, and trains of randomly shaped pulses.

Each member of the (R, P) sequence described above possesses a corresponding error surface described by (1.7). Thus there exists a sequence of recurring error surfaces as in Fig. 1. The time-sequence adaptive filter proposed here uses a *multiplicity* of weight vectors—usually one corresponding to each error surface. Since the number of different error surfaces for a statistically recurring process is finite, the number of weight vectors is also finite. These weight vectors will be denoted by $\mathcal{W}_0, \mathcal{W}_1, \mathcal{W}_2, \cdots$. At each station in time, one and only one weight vector is selected, based on the error surface present at that time, and adapted toward the minimum of the error surface by the LMS algorithm. When the minimum point is reached, after many adaptations, the weight vector is identical to the

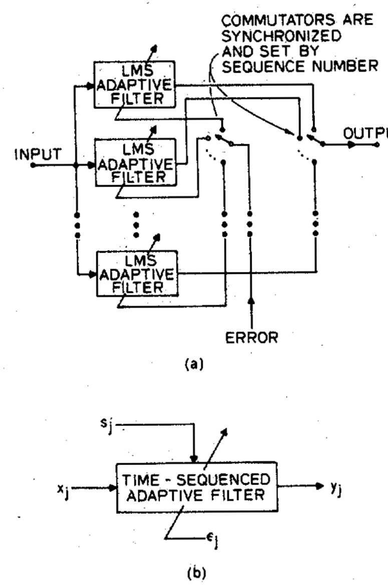


Fig. 3. The time-sequenced adaptive filter. (a) Conceptual realization as a bank of LMS adaptive filters. (b) Symbolic representation.

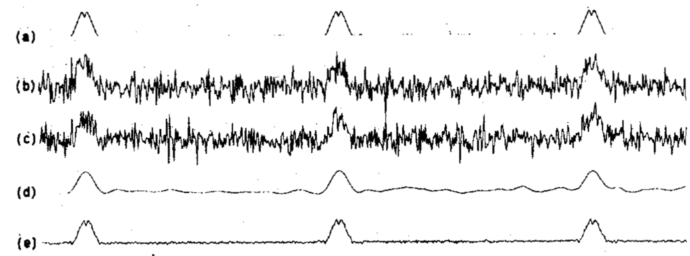


Fig. 4. Comparison of time-sequenced and conventional LMS performance in estimating notched triangle pulses. (a) Underlying signal. (b) Desired response. (c) Filter input. (d) Conventional LMS filtering. (e) Time-sequenced filtering.

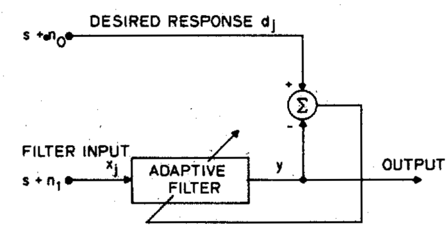


Fig. 5. Block diagram for a signal enhancing experiment. The noises n_0 and n_1 are uncorrelated. The output is an estimate of the signal s.

Wiener weight vector for that error surface, yielding a filter output which is a best least squares match to the desired output at that station in time. Thus each weight vector becomes an expert in filtering a particular portion of the interval between regenerations. For this procedure an external input to the filter, called the sequence number and denoted σ_j, is required to determine the appropriate weight vector to use at time j. Thus when $\sigma_j = i$ the ith error surface is assumed present, so that the ith weight vector is used to form the filter output and then adapted toward the bottom of its error surface. In order to set the sequence number, some *a priori* knowledge of the filter input is required. For pulse-type signals this *a priori* knowledge could be the location of the pulses in time; for signals with periodic statistics (sometimes called cyclostationary [9]), knowledge of the period is sufficient.

Mathematically the time-sequenced adaptive algorithm is

$$y_j = X_j^T \mathcal{W}_{\sigma_j}(j) \tag{2.1}$$

and

$$\mathcal{W}_i(j+1) = \begin{cases} \mathcal{W}_i(j) + 2\mu_i \epsilon_j X_j, & i = \sigma_j \\ \mathcal{W}_i(j), & \text{otherwise} \end{cases} \tag{2.2}$$

where $\mathcal{W}_i(j)$ is the value of the ith weight vector at time j. A different μ is used for each weight vector in order to keep the percent loss in steady-state performance due to the adaptive process, referred to as the "misadjustment," the same for each weight vector. A conceptual block diagram of the time-sequenced adaptive filter is shown in Fig. 3(a), illustrating the process as a bank of LMS adaptive filters. Fig. 3(b) shows the symbolic representation adopted for the time-sequenced adaptive filter.

By using a different weight vector for each error surface, the time-sequenced adaptive filter eliminates the Wiener weight vector tracking error [5] present when the LMS adaptive filter is used with nonstationary signals. However, a different sort of error is introduced in some applications due to uncertainty or jitter in σ_j at any particular time. The performance of the time-sequenced adaptive filter in the face of this uncertainty is analyzed in [2]. If the sequence number can be chosen perfectly, then it can be shown that the time-sequenced adaptive filter converges to the optimal time-varying filter when adaptation is performed slowly enough. This and other properties are discussed in [2].

Although the time-sequenced adaptive filter may seem to be a costly approach to signal processing, the amount of computing involved is essentially the same as that for a single LMS adaptive filter. This is because, in either case, one vector dot product and one weight vector adaptation is required with each input data point. The number of input data samples required for the time-sequenced adaptive filter to converge to its time-varying solution is greater than that required for the LMS adaptive filter to converge to its essentially time-invariant solution, since adaptation is effected for each weight vector only once per regeneration rather than for every data sample. However, the increased performance resulting from the time-varying solution more than compensates for this drawback in many applications. The amount of memory required by the time-sequenced adaptive filter is of course increased due to the multiple weight vectors.

Fig. 4 shows the results of a computer simulation comparing the performance of the time-sequenced and conventional LMS adaptive filters in the signal enhancer of Fig. 5. The concept of the adaptive signal enhancer is described in detail in [1]. All weight vectors in this example contained 75 weights plus a bias weight. The underlying signal, s, to be estimated was the recurring notched triangle pulse of Fig. 4(a). White Gaussian noise was added to this signal and used as the filter input x_j, shown in Fig. 4(c). A similar

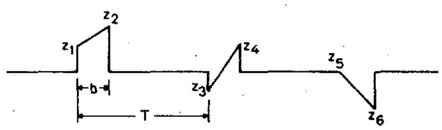


Fig. 6. A signal composed of randomly shaped pulses.

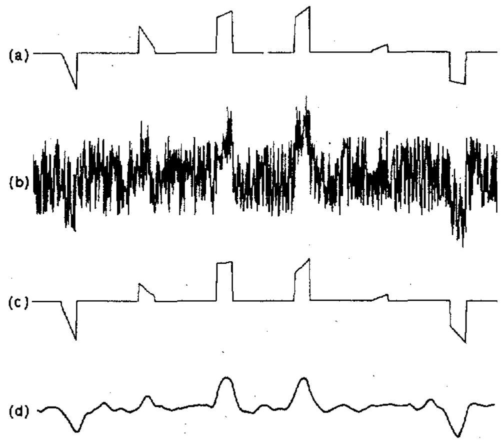


Fig. 7. Simulation comparing time-sequenced and conventional LMS enhancing for a randomly shaped pulse train. Measured improvement of time-sequenced over conventional LMS is 7.2 dB. (a) Underlying waveform. (b) Filter input. (c) Time-sequenced enhancing. (d) Conventional LMS enhancing.

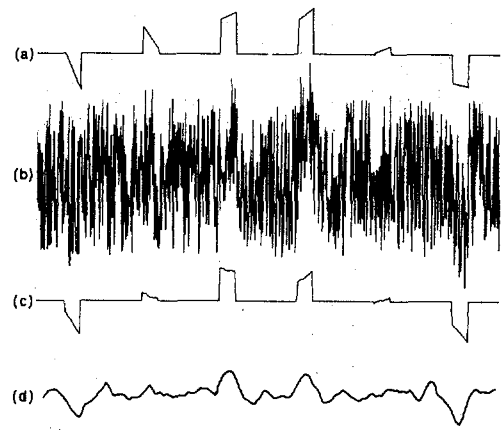


Fig. 8. Simulation comparing time-sequenced and conventional LMS enhancing for a randomly shaped pulse train with increased noise. Measured improvement of time-sequenced over conventional LMS is 3.7 dB. (a) Underlying waveform. (b) Filter input. (c) Time-sequenced enhancing. (d) Conventional LMS enhancing.

waveform having noise independent from the first was used as the desired response, Fig. 4(b). Although the signal components of the desired response and filter input were identical in the experiment, they did not need to be. Assuming perfect knowledge of the regeneration times, the output of the time-sequenced adaptive filter after convergence is shown in Fig. 4(e). Note that the notch is clearly visible. This experiment was repeated using the LMS adaptive filter. The filter output after convergence is shown in Fig. 4(d). Note that the notch has been smoothed beyond recognition. This distortion due to smoothing is typical of the difference between the optimal time-invariant and time-varying filters when pulse-type signals are to be estimated.

Instead of a fixed pulse shape as in the previous example, let the pulses have a randomly varying shape as in Fig. 6. Each pulse is generated by choosing two independent random variables (e.g., z_1 and z_2 in Fig. 6) as the heights of the pulse edges, then connecting these values by a straight line. These random variables are chosen independently from pulse to pulse. Simulations were performed to measure the difference in performance between the optimal time-varying and time-invariant filters in estimating the underlying pulse train in white noise. A time-sequenced adaptive filter was used to learn the optimal time-varying filter. A conventional LMS adaptive filter was used to learn the optimal time-invariant filter. The actual underlying waveform was used as the desired response for both adaptive filters. Although a known desired response would not be generally available, this example provides a fundamental comparison of time-varying and time-invariant approaches. Similar performance can be achieved with a noisy desired response by adapting more slowly. The results of a simulation are shown in Fig. 7. Fig. 7(a) is the underlying waveform to be estimated. Independent white noise was added to this waveform to form the filter input Fig. 7(b). The output of the time-sequenced adaptive filter after convergence is shown in Fig. 7(c) and that of the LMS adaptive filter in Fig. 7(d). The desirability of a time-varying filter is clear. The average square error over 30 cycles was measured for both filter outputs to quantify the increase in performance. The time-sequenced filter reduced the average square error by 7.2 dB as compared to the LMS filter.

Fig. 8 shows the results of another simulation with increased noise power. The measured performance of the time-sequenced filter was 3.7 dB better than that of the LMS filter in this case.

III. Applications of the Time-Sequenced Filter

One application of the time-sequenced adaptive filter is to fetal electrocardiography. The interfering maternal electrocardiogram is adaptively cancelled by synchronizing the adaptive filter with the maternal heart beat. Then, a second adaptive filter is synchronized with the fetal heart beat to enhance it against the remaining background noise. Substantial improvement in performance over conventional LMS adaptive filtering has been demonstrated [2].

The time-sequenced filter can also be used to predict future samples of stochastic processes which have a periodic character. The sequence number input is easy to generate if the period of the process is available, e.g., when the process is known to cycle daily or annually. One possible application is to electric load forecasting [10]. In this application a utility desires to predict power consumption one-half hour to one week in the future based on past values of consumption. Load prediction is necessary to supply electric energy in a secure and economic manner. Although power consumption exhibits a clear daily cycle, the actual demand during a particular hour varies from day to day. The time-sequenced adaptive filter can be used as a predictor for this problem as illustrated in Fig. 9. Samples

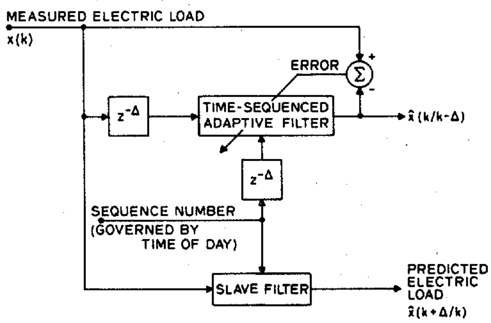


Fig. 9. An adaptive predictor for electric load forecasting.

of the power consumption, $x(k)$, taken every half-hour are used as the desired response of the adaptive filter. These same samples are also used as the adaptive filter input, but first they are delayed by the number Δ of half-hour increments for which prediction is desired. If the adaptive filter has n weights, it will adjust itself to form the best least square predictor of the present load $x(k)$ based on $x(k-\Delta), x(k-\Delta-1),\cdots,x(k-\Delta-n+1)$. Call this estimate of the present given the past $\hat{x}(k/k-\Delta)$. A time-sequenced, rather than conventional LMS, adaptive filter is used since the statistics of power consumption vary approximately on a 24-hour cycle so that a different weight vector should be used for each half-hour of the day. The sequence number of the adaptive filter can be generated by the time of day. To predict the load Δ half-hours in the future a slave filter is required. The slave filter is also a time-varying transversal filter whose weights have been previously calculated and stored by the adaptive filter. The sequence number input to the slave filter determines which of the precalculated weight vectors it should use at that time. Since the filter input and sequence number of the slave filter are advanced Δ time units with respect to the adaptive filter, the slave filter output is the desired estimate for the load Δ units in the future, denoted $\hat{x}(k+\Delta/k)$ in Fig. 9.

IV. Conclusion

The time-sequenced adaptive filter described in this paper is a new form of adaptive filter particularly suited to the optimal estimation of those nonstationary signals having a recurring statistical character. The principle advantages of the method are that it does not require that signal statistics be known *a priori* and that its computational requirements are modest. Possible applications of time-sequenced filtering include fetal electrocardiogram enhancement and electric load prediction.

By conceptualizing the time-sequenced adaptive filter as a bank of LMS adaptive filters, several convergence properties of the time-sequenced filter can be derived [2]. When the sequence number, required to choose the appropriate filter in the bank, is known the time-sequenced filter converges to the optimal time-varying Wiener filter.

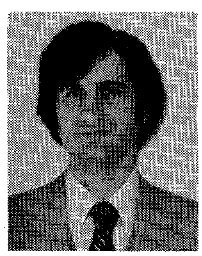
One suggestion for further research is to modify the adaptive algorithm to incorporate the information learned about the optimal solution at one point in time with that learned at other points in time. In this way a faster converging time-sequenced adaptive filter could be realized.

Acknowledgment

The application of the time-sequenced approach to electric load prediction was suggested by C. Richard Johnson, Jr.

References

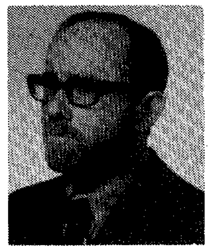
[1] E. R. Ferrara and B. Widrow, "Multichannel adaptive filtering for signal enhancement," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-29, pp. 766–770, this issue.
[2] E. R. Ferrara, "The time-sequenced adaptive filter," Ph.D. dissertation, Stanford University, Stanford, CA, Dec. 1977.
[3] B. Widrow *et al.*, "Adaptive noise cancelling: Principles and applications," *Proc. IEEE*, vol. 63, pp. 1692–1716, Dec. 1975.
[4] B. Widrow, "Adaptive Filters," in *Aspects of Network and System Theory*, R. Kalman and N. DeClaris, Eds. New York: Holt, Rinehart, and Winston, 1971, pp. 563–587.
[5] B. Widrow *et al.*, "Stationary and nonstationary learning characteristics of the LMS adaptive filter," *Proc. IEEE*, vol. 64, pp. 1151–1162, Aug. 1976.
[6] R. V. Southwell, *Relaxation Methods in Engineering Science*. New York: Oxford, 1940.
[7] D. J. Wilde, *Optimum Seeking Methods*. Englewood Cliffs, NJ: Prentice-Hall, 1964.
[8] A. R. Hambley *et al.*, "The use of intrinsic components in an ECG filter," *IEEE Trans. Biomed. Eng.*, vol. BME-21, pp. 469–473, Nov. 1974.
[9] W. A. Gardner and L. E. Franks, "Characterization of cyclostationary random signal processes," *IEEE Trans. Inform. Theory*, vol. IT-21, pp. 4–14, Jan. 1975.
[10] F. D. Galiana *et al.*, "Identification of stochastic electric load models from physical data," *IEEE Trans. Automat. Contr.*, vol. AC-19, pp. 887–893, Dec. 1974.



Earl R. Ferrara, Jr. (S'73–M'78) was born on December 29, 1952. He received the B.S. degree in engineering *magna cum laude* from the University of New Orleans in 1974, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, in 1975 and 1978, respectively. From 1974 to 1977 he was a National Science Foundation graduate fellow.

During the summers of 1973 and 1974 he was associated with Oak Ridge National Laboratory, Oak Ridge, TN, and the Naval Weapons Laboratory, Dahlgren, VA, respectively. Since 1978 he has been on the technical staff of ESL, Inc., Sunnyvale, CA, where he has been engaged in research in adaptive filtering and in the location of multiple cochannel signal sources. During 1979 he served as a Lecturer at Stanford University.

Dr. Ferrara is a member of Phi Kappa Phi.



Bernard Widrow (M'58–SM'75–F'76) was born in Norwich, CT, on December 24, 1929. He received the B.S., M.S., and Sc.D. degrees from the Massachusetts Institute of Technology, Cambridge, in 1951, 1953, and 1956, respectively.

From 1951 to 1956, he was a Staff Member of the M.I.T. Lincoln Laboratory and a Research Assistant in the Department of Electrical Engineering of M.I.T. He joined the M.I.T. faculty in 1956 and taught classes in radar, theory of sampled-data systems, and control theory. In 1959, he joined the faculty of Stanford University, Stanford, CA, and is now Professor of Electrical Engineering. He is presently engaged in research and teaching in systems theory, pattern recognition, adaptive filtering, and adaptive control systems. He is Associate Editor of *Information Sciences* and *Pattern Recognition*.

Professor Widrow is a member of the American Association of University Professors, the Pattern Recognition Society, Sigma Xi, Tau Beta Pi, and is a Fellow of the American Association for the Advancement of Science.