Myoelectric Signal Processing: Optimal Estimation Applied to Electromyography—Part I: Derivation of the Optimal Myoprocessor

NEVILLE HOGAN AND ROBERT W. MANN, FELLOW, IEEE

Abstract—This paper (Part I of II) describes the development of a novel technique for processing the electrical activity of muscle which uses multiple channels of myoelectric activity. A phenomenological mathematical model of myoelectric activity is formulated. From this model, a mathematical statement of the optimal myoelectric signal processor is derived, and some of its properties are investigated. This mathematical statement encompasses and places in perspective almost all single-channel myoprocessor developments to date, as well as specifying the optimal multiple-channel myoprocessor. An experimental demonstration of the efficacy of this processor is presented in a subsequent paper (Part II).

NOMENCLATURE

$p(x/y)$	Probability density function for x given y.
$E\{\ \}$	Expectation.
$\vert\ \vert$	Absolute value of a scalar; determinant of a matrix.
$\Gamma(\ )$	Gamma function.
$\zeta(\ )$	Characteristic function.
$\Phi(\ )$	State transition matrix.
$H(\ )$	Weighting function, transfer function or frequency response function in time, Laplace or frequency domains, respectively.
$S(\ )$	Power spectral density.
$R(\ )$	Autocorrelation function.
ρ	Correlation coefficient.
C or Σ	Covariance matrix.
σ, σ^2	Standard deviation, variance.
Λ	Eigenvalue matrix.
λ	Eigenvalue.
Φ	Eigenvector matrix.
I	Identity matrix.
M	Myoelectric activity.
F	Muscle force.
$\hat{F}$	Muscle force estimate.
u	Transformed myoelectric activity.
W	Prewhitened myoelectric activity.

Manuscript received August 29, 1979; revised February 8, 1980. This work was supported in part by the Rehabilitation Services Administration, U.S. Department of Health, Education, and Welfare, under Grant 23-P-55854/1 and the Whitaker Professorship of Biomedical Engineering at the Massachusetts Institute of Technology, Cambridge, MA.

The authors are with the Department of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139.

INTRODUCTION

BECAUSE of its potential for providing an easily accessible quantitative estimate of the active state of muscle, myoelectric activity has seen wide application in diverse fields, e.g., clinical diagnosis, experimental medicine, bioengineering, physical therapy, psychotherapy, rehabilitation, prostheses/orthoses, etc. Despite this wide attention, the fidelity achieved using common myoprocessing techniques is poor. Attempts to provide a quantitative myoelectric estimate of muscle activity have been plagued by the presence of an apparent noise component of very large amplitude and low frequency which is superimposed on the desired signal [1]. Over the past decade, many attempts have been made to improve myoprocessor performance, but most of these improvements have evolved serendipitously, without any comprehensive theoretical justification. This is primarily because when viewed from the physiological standpoint, the problem appears to be intractably complex, and when viewed from the signal-processing standpoint, the problem is seen to be fundamentally nonlinear. Part I of this paper reports on a project whose outcome was: 1) the introduction of a novel myoprocessing technique, and 2) the mathematical derivation of the optimal myoprocessor. Part II reports on the experimental demonstration of the performance of this myoprocessor, which is almost an order of magnitude better than the common myoprocessor.

This project was part of our ongoing effort to develop a myoelectrically controlled, externally powered prosthesis—one of the most demanding applications of myoelectric activity. Although myoelectrically controlled, externally powered prostheses are currently in use, these devices are of limited capability, and the opinion has grown that there are fundamental problems associated with myoelectric control. It has been commonly observed that the operation of myoelectrically controlled prostheses requires an inordinately high degree of concentration and is heavily dependent upon direct feedback of prosthesis position, particularly via vision [2]-[6]. This is manifestly unacceptable, and the primary focus of recent research on externally powered assistive devices has been the provision of substitute sensory feedback paths (i.e., cutaneous or neural stimulation) to replace the lost proprioceptive senses [6]-[11]. The underlying presumption is that difficulties with myoelectric control are due to sensory lacunae, and that

0018-9294/80/0700-0382$00.75

in the absence of position and movement sensors, such as the skin receptors or joint afferents, and the subsequent integration of this sensory information in the brain, the operator is largely ignorant of what his muscles are doing and requires external sensory feedback to inform him of his actual movements. It is important to note that this argument is based on the predominant neurophysiological opinion of a decade ago (contemporaneous with the research presaging current prostheses) that proprioceptive feedback was all-important in the execution of normal movements [12]-[14]. Neurophysiological research has recently provided new insights, but the philosophy underlying research in myoelectrically controlled devices has yet to reflect this new knowledge. Strong experimental evidence now exists that the major component of the movement control which was hitherto attributed to sensory feedback loops can be accomplished in the absence of any feedback [15]-[19]. This new neurophysiological knowledge means that provision in a myoelectrically controlled device of a high-quality, surrogate sensory feedback pathway alone will not be sufficient to guarantee successful control. Instead, the most important requirement is a forward path which will ensure accurate and timely communication to the machine of the motor intent of the operator. Given a forward path of sufficient quality, the performance of at least the ballistic portion of movements may be achieved without need of external feedback. Note that we do not suggest that the amputee will be able to dispense with position feedback entirely, but that the need for observation during movement will be reduced. On the other hand, inaccuracies in the forward path will prohibit fine control, even in the presence of external feedback. In effect, the fidelity of the forward path of the system determines the limits of the obtainable performance. For these reasons, we are of the opinion that a major impediment to successful myoelectric control is the low fidelity achieved using common myoprocessors. There is a clear need for improvement, and this provided the motivation for the work reported in this paper.

Multichannel Myoprocessing: A New Approach

The basis of the new processing technique and the justification for the mathematical analysis are provided by considering the physiology and electrophysiology of muscle. The mechanical output of the muscle is a combination of the outputs of all active motor units in the muscle. Each of these, in turn, is a combination of the outputs of the individual contractile fibers which, in turn, are a combination of the outputs of the individual sacromeres which, in turn, are a combination of individual actin/myosin interactions. Similarly, the electrical activity of the muscle is a combination of individual electrical events in the muscle. In both cases, the macroscopic activity of the muscle is a weighted sum across both space and time of a large number of individual events, but the weighting in the electrical case is radically different from the mechanical case. This is due to the electrode and tissue attenuation which limits the pickup region of the electrodes. In effect, the electrodes "look at" only a small sample of the total population of active muscle fibers. As a result, a spatiotemporal sampling artifact ensues which is manifested as large-amplitude, low-frequency "noise" contaminating the myoprocessor output. Elimination of a sampling artifact requires an increased sample size. To accomplish this without compromising the selectivity of the detector, the new technique of combining multiple channels of myoelectric activity into a single myoprocessor output was proposed [1].

To apply this new technique, two questions have to be answered: what is the best way to process the activity of each individual electrode, and what is the best way to combine the individual electrode activities?

Functional Modeling of Myoelectric Activity

To answer these questions, a model of myoelectric activity was developed. Because the relation between the contraction of a muscle and its myoelectric activity is extremely complex, a detailed, structural description of even the firing of a single muscle is inordinately involved. As our objective was to design a practical myoprocessor, the approach taken was phenomenological. The model developed provides a functional description of the observed relation between muscle contraction and myoelectric activity, but does not describe the underlying structural processes. This point is important when interpreting our mathematical results.

Shown schematically in Fig. 1, the model describes myoelectric activity as a random process whose instantaneous amplitude distribution for a fixed level of muscle activity is Gaussian with zero mean:

$$p\,\{M(t)\} = \frac{1}{\sqrt{2\pi}\,\sigma}\exp\left[-\frac{1}{2}\cdot\frac{M(t)^2}{\sigma(F)^2}\right]; \qquad E\,\{M(t)\} = 0. \tag{1}$$

As myoelectric activity is a sum of conditionally independent events, the Law of Large Numbers justifies this assumption, and experiments verify it (at least for surface myoelectric activity) [20]-[24]. The assumption of zero mean is justified on the basis of experimental observations.

The other parameter of the amplitude distribution is its variance, and this is assumed to be modulated multiplicatively by a static or memoryless function of muscle force which is, in general, nonlinear:

$$\sigma^2 = \sigma\,[F(t)]^2. \tag{2}$$

The time dependence between successive values of myoelectric activity is completely characterized by the autocorrelation function or, equivalently, the power spectral density. In general, myoelectric activity is a nonstationary random process, as muscle force may vary, but as the frequency components of muscle force are almost an order of magnitude lower than those of myoelectric activity, the process may be described as "secularly varying," that is, stationary over the time available for processing.

Because we have assumed multiplicative modulation, the shape of the power spectral density of myoelectric activity will be invariant, independent of force. This assumption is justified by experimental data [25], [26]. We further assume the power-spectral density to be a rational function of frequency. This assumption is made primarily for mathematical con-

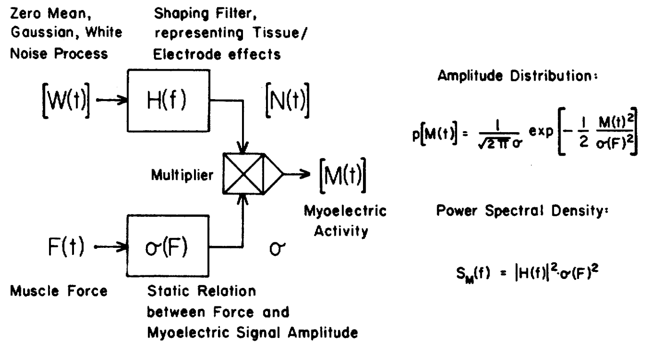


Fig. 1. A functional mathematical model representing myoelectric activity as a band-limited, zero-mean, Gauss-Markov process $[N(t)]$ which is amplitude-modulated by a static function of muscle force.

venience, but it is also a reasonable description of experimental observations. (See Part II of this paper.) Thus (see Appendix II),

$$S_M(f) = Q \cdot |H(f)|^2 \cdot \sigma(F)^2 \tag{3}$$

where Q = constant and $|H(f)|^2$ = ratio of polynomials in f^2.

This set of assumptions permits myoelectric activity to be modeled as a zero-mean, Gaussian white noise process passed through a linear constant-coefficient filter (representing a combination of the frequency content of subcutaneous myoelectric activity and the filtering effect of transmission through tissue and detection by electrodes) which is subsequently multiplied by a static, nonlinear function of muscle force. A physical situation corresponding to this model is that of a muscle contracting under isometric, nonfatiguing conditions. In this case, the parameters describing the myoelectric activity are a function only of the muscle force. This situation obtains when the relevant, residual, but now dysfunctional muscles in an amputee's stump are used to command the movements of an externally powered prosthesis.

Optimization Procedure

In myoelectric control, the basic function of the myoprocessor is to permit the human operator to communicate his intent to the machine. We assume that the operator's intent with respect to the machine can be determined unequivocally from an observation of the appropriate set of muscle forces. Fundamentally, this assumption is based on Newton's laws which guarantee that every movement is causally determined by a force, but more specifically, it has been shown that given the appropriate set of muscle forces and a knowledge of the kinematic state of the prosthesis and the relevant musculoskeletal anatomy, the required prosthesis actuator torques can be determined [27], [28].

The task now is to measure muscle force using myoelectric activity. The model of Fig. 1 describes myoelectric activity as a Gauss-Markov process and suggests the use of the well-developed theory of linear Kalman filters. However, the frequency components of the desired signal, muscle force, differ from those of the available information, myoelectric activity, and, in addition, the "signal" and "noise" are multiplied, not added. Thus, the problem is fundamentally nonlinear, and Kalman-filter theory cannot be applied, even as an approximation. The approach taken was to treat the problem as that of estimating a statistical parameter of a random process rather than that of extracting a signal from noise. Using a combination of state-space methods and statistical decision theory, the optimum myoprocessor was defined as the maximum likelihood estimator of muscle force. Given some very general and unrestrictive conditions, the classical maximum likelihood estimator is known to be consistent and efficient, although not necessarily unbiased.

The use of the maximum likelihood estimator and state-space methods make the optimization procedure particularly simple. Also, in the interest of simplicity, a discrete formulation was used. From the probability density function for an individual sample of myoelectric activity $M(t)$ and the autocorrelation function or power spectral density relating successive samples of myoelectric activity, we can derive the probability density function (or likelihood function) for a finite set of samples of myoelectric activity, given a value of muscle force:

$$p(M/F) = p(M_1 \cdots M_N/F)$$

where $M = [M(t_1), \cdots, M(t_N)]^T$ is the sequence of values of myoelectric activity and F is the muscle force.

The maximum likelihood estimator of muscle force $\hat{F}$, given a set of observed values of myoelectric activity M_0, is defined by

$$\left.\frac{d}{dF} p(M/F)\right|_{\substack{F=\hat{F} \\ M=M_0}} = 0$$

where $\hat{F}$ is the optimal estimator of muscle force and M_0 are the observed values of myoelectric activity.

Case I: Single Channel, Uncorrelated Samples

To illustrate the properties of the resulting optimum myoprocessor, we will first suppose myoelectric activity to be sampled at equally spaced intervals Δt seconds apart where Δt is such that successive samples are uncorrelated, i.e.,

$$R_M(\tau) = 0; \quad \tau > \Delta t.$$

Because the samples are taken from a Gaussian process, uncorrelatedness implies independence. Because only a finite time T is available in which to produce the estimate of muscle force F, we specify that there are only a finite number N of samples where $N = T/\Delta t$. Thus, we have N independent samples of myoelectric activity, each distributed according to

$$p(M_i/F) = \frac{1}{\sqrt{2\pi}\,\sigma(F)} \cdot \exp\left[-\frac{1}{2}\frac{M_i^2}{\sigma(F)^2}\right]; \quad i = 1 \cdots N.$$

Hence, the likelihood of this sequence occurring is given by

$$p(M/F) = \prod_{i=1}^{N} p(M_i/F).$$

The best estimate of force on the base of this set of samples is that for which this likelihood is a maximum. The resulting expression defining the best estimator F is

$$\hat{F} = \sigma^{-1}\left\{\left[\frac{1}{N}\prod_{i=1}^{N} M_i^2\right]^{1/2}\right\}. \tag{5}$$

In this expression, σ^{-1} denotes the inverse functional relationship between muscle force F and standard deviation σ. Details of this derivation are given in Appendix I.

What this derivation has yielded is a reiteration of the well-known fact that a Gaussian process is completely described by its mean and variance. In this case, the mean is zero, by assumption; thus, the variance carries all the information. The optimal estimate of force is, therefore, obtained by first estimating the variance by squaring and averaging and then using the known relation between force and variance to obtain the force estimate. An operational block diagram of the process is shown in Fig. 2.

Properties of the Optimal Myoprocessor

As with all optimization procedures, the nature of the results depends heavily upon the precise meaning of "optimal" or "best." To provide some insight into the properties of the maximum likelihood estimate, we will examine some of the characteristics of the optimal processor. First, we need a precise form for the relation $\sigma = \sigma(F)$. This can be obtained from data relating muscle force to mean rectified myoelectric activity because (see Appendix I)

$$\sigma = \sqrt{\frac{\pi}{2}} \cdot E\{|M|\}.$$

This relation has been found to be both linear [29]-[31] and nonlinear [32]-[36], with the consensus favoring a nonlinear relationship. We will use the form

$$\sigma = k \cdot F^a. \tag{6}$$

This relation fits the data of Vredenbregt and Rau [35] with a correlation coefficient of 0.98 (see Appendix I). Using (6) in (5), we obtain

$$\hat{F} = \left\{\frac{1}{k}\left[\frac{1}{N}\sum_{i=1}^{N} M_i^2\right]^{1/2}\right\}^{1/a}.$$

It is shown in Appendix I that the force estimate thus defined is distributed according to an inverse Gamma-n distribution [37]. The mean value of $\hat{F}$ is given by

$$E\{\hat{F}\} = \frac{\Gamma(N/2 + 1/2a)}{\Gamma(N/2)} \cdot \left[\frac{2}{N}\right]^{1/2a} \cdot F.$$

Thus, except for the case when $a = 0.5$, $\hat{F}$ is a biased estimator of muscle force. N is the number of degrees of freedom in the estimate of the total signal variance and is given by $N = 2B_sT$ where B_s is the statistical bandwidth of the signal and T is the averaging time. Typical values for the parameters are $a = 1.74$, $B_s = 100$, and $T = 0.25$ [38]. These yield an expected value for the estimate of

$$E\{\hat{F}\} = 0.996F.$$

Thus, the effect of the bias is small—less than 0.4 percent—and is easily corrected. This bias is characteristic of the maximum likelihood method. Application of least-squares methods would probably correct this minor fault.

Under constant force conditions, we can consider the mean value of $\hat{F}$ as the desired signal and the variations about this mean as "noise." We can define the signal-to-noise ratio as

$$\text{SNR} \triangleq \left[\frac{E\{\hat{F}\}^2}{E\{(\hat{F} - E\{\hat{F}\})^2\}}\right]^{1/2}$$

which can be shown to be (see Appendix I)

$$\text{SNR} = \left[\frac{\Gamma(N/2 + 1/a) \cdot \Gamma(N/2)}{\Gamma(N/2 + 1/2a)^2} - 1\right]^{-1/2} = \text{constant}.$$

Thus, the variability of the estimate increases with the mean value of the estimate. The variability of the actual force also increases with the mean level of muscle force; thus, the properties of the optimal estimator derived above are quite reasonable.

Case II: Single Channel, Dependent Samples

Case I above assumes samples spaced far enough apart that successive samples were independent of one another. At closer spacings, these samples become interdependent. This dependence is given by the autocorrelation function of the process which is a property of the linear constant coefficient filter in the model of Fig. 1. For simplicity, we assumed this filter to be a first-order, low-pass filter (generalization to a higher order filter is straightforward although tedious). This and the earlier modeling assumptions permit us to assume the stationary first-order Markov property:

$$p(M_i/M_{i-1}, M_{i-2}, \cdots, M_1, F) = p(M_i/M_{i-1}, F) \tag{7}$$

where the sequence $M_1 \cdots M_N$ is ordered such that M_N is the most recent sample, i.e.,

$$M_i = M(t_i); \quad t_1 < \cdots t_i < \cdots t_N.$$

The Markov property means that given an initial likelihood function, subsequent likelihood functions can be determined from a knowledge of the transition likelihood function, the right-hand side of (7). To assume the Markov property is to assume that the process is state determined in the probabilistic sense. The transition likelihood function $p(M_i/M_{i-1}, F)$ is derived from the autocorrelation function $R_M(\tau)$ via the state transition matrix $\Phi(\tau)$ of the linear, time-invariant, shaping filter, as shown in Appendix II.

Given the transition likelihood function, we use Bayes' rule and the Markov property to express the likelihood function for the sequence $M_1 \cdots M_N$ as

$$p(M/F) = \left[\prod_{i=2}^{N} p(M_i/M_{i-1}, F)\right] \cdot p(M_1/F).$$

From this point, the derivation proceeds much as in the previous case and is presented in Appendix II. Because we have accounted for the interdependence of successive samples, there is no restriction on the sample spacing. The continuous-

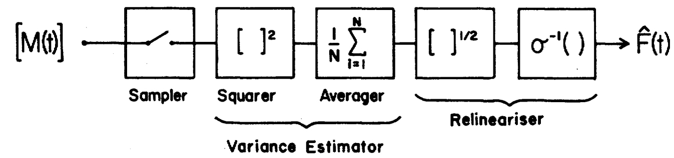


Fig. 2. Block diagram of the discrete-time version of the optimal single-channel estimator of muscle force. Samples are assumed uncorrelated.

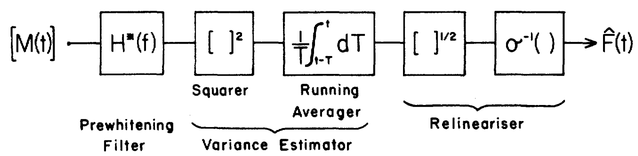


Fig. 3. Block diagram of the continuous-time version of the optimal single-channel estimator of muscle force.

time version of the optimum myoprocessor in this case will be as shown in Fig. 3. The only addition to the processor of Case I is the inclusion of a complementary or "prewhitening" filter $H^*(f)$. In theory, the prewhitening filter acts to remove the correlation between successive samples of myoelectric activity which had been introduced by the shaping filter $H(f)$.

The optimal processor of Fig. 3 must be interpreted carefully. The fact that the mathematics indicate a complementary filter to remove the correlation introduced by the shaping filter is a direct consequence of our initial assumption of a white noise (or purely random) process. The complementary filter is an attempt to restore the spectrum of the signal to its initial "white" condition–hence, the name "prewhitening" filter. Clearly, in the real physical situation, we do not have a white process–a white process is merely a convenient mathematical fiction. The relationship between the prewhitening filter $H^*(f)$ and the shaping filter $H(f)$ for the continuous case is shown in Fig. 4. This clearly indicates that the prewhitening filter is unrealizable as it requires infinite gain at some frequencies. However, as indicated in Fig. 4, an approximation to the prewhitening filter is realizable.

Putting aside the considerations of realizability, the blind application of this technique of prewhitening is unwise. At high frequencies, the signal amplitude is not simply reduced by the filtering effects of the skin–there is no signal at high frequencies. Any attempts to "restore" the signal at these frequencies will merely serve to amplify background noise–a feature of the real world which was not included in the model. Nonetheless, the technique can be applied over those frequency ranges for which the signal exists and is not dominated by background noise. Certainly no frequency range containing useful signal information should be attenuated.

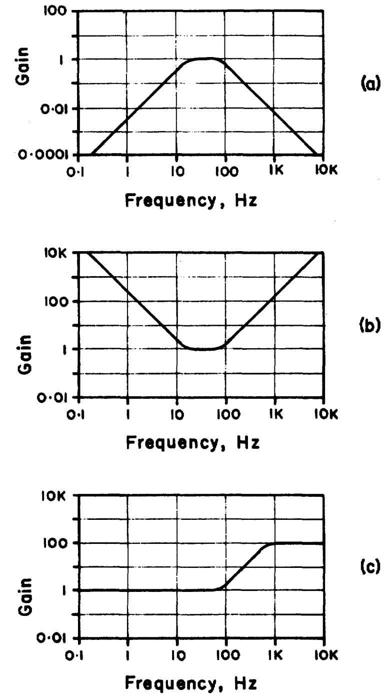


Fig. 4. Diagram of the relationship between the shaping filter $H(f)$ shown in (a) and the prewhitening or complementary filter $H^*(f)$ shown in (b). Because $H(f)$ goes to zero at high and low frequencies, the relation $H(f) \cdot H^*(f) = 1$ requires $H^*(f)$ to become infinite at these frequencies. However, practical approximations to $H^*(f)$ are possible, one of which is shown in (c).

Prewhitening via Electrode Configuration

An interesting approach to prewhitening is to modify the filtering function due to the muscle tissue, skin, and electrodes [40]. From simple analytical considerations, it can be shown that the differential electrode configuration which is used to eliminate common mode interference acts to filter the myoelectric activity through a "frequency comb" as sketched in Fig. 5. The effect is much like a diffraction effect–the difference results in destructive interference of certain frequency components. The location of the dips in the spectrum is determined by the velocity of propagation of the motor-unit

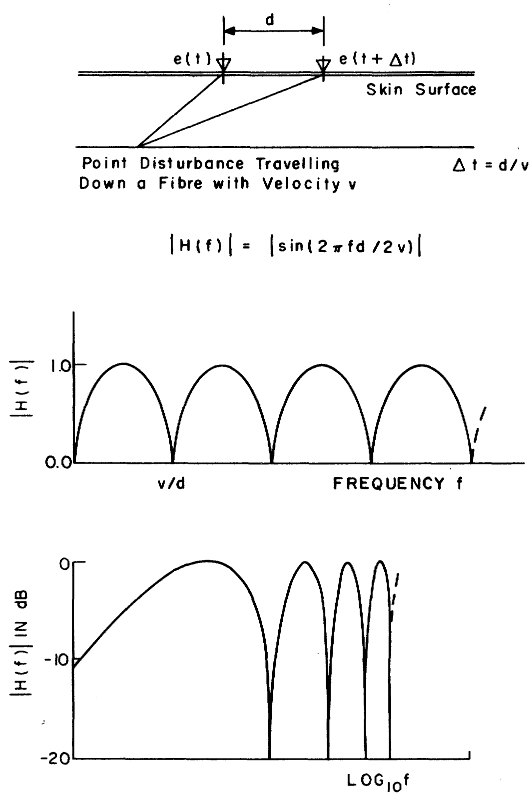


Fig. 5. Diagram of the filtering effects of the spacing between the two electrodes of a differential pair.

action potential along a muscle fiber and the spacing of the differential electrodes. Typical values for these parameters place the first dip at a frequency of 100 Hz–right in the middle of the bandwidth of surface myoelectric activity. By reducing the electrode spacing, this dip can be moved to a higher frequency, thus increasing the effective bandwidth of the myoelectric activity or, alternatively, acting to make the frequency spectrum of myoelectric activity look more white. The result will be an improvement in myoprocessor performance.

Case III: Multiple Channels, Uncorrelated Samples

The preceding analysis yielded the optimal myoprocessor for the single-channel case. The same approach is readily adapted to the multiple-channel case (see Appendix III). A multidimensional Gaussian amplitude distribution is assumed whose parameters are modulated by a static, nonlinear function of muscle force. As before, the mean is assumed zero; thus, in this case, all of the information is carried in the covariance matrix C. This covariance matrix can be written as

$$C = S^T \cdot R \cdot S = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_m \end{bmatrix} \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1m} \\ \rho_{21} & 1 & & \\ \vdots & & \ddots & \\ \rho_{m1} & & & 1 \end{bmatrix} \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_m \end{bmatrix}$$

where S is a matrix of the standard deviations of individual channels and R is a matrix of the correlations between pairs of channels. In general, all σ_i and ρ_{ij} are functions of F. However, considerable algebraic simplification is obtained by performing an orthonormal transformation on the vector of myoelectric activities $M(t)$:

$$u(t) = \Phi^T \cdot M(t).$$

Φ is a matrix of unit eigenvectors of $C(F)$ and is defined by

$$C(F) \cdot \Phi = \Phi \cdot \Lambda(F); \qquad \Phi^T \cdot \Phi = I$$

where I is the identity matrix and $\Lambda(F)$ is the diagonal matrix of eigenvalues:

$$\Lambda(F) = \begin{bmatrix} \lambda_1(F) & & 0 \\ & \ddots & \\ & & \lambda_m(F) \end{bmatrix}.$$

The advantage of this transformation is seen in the covariance matrix of the vector of transformed myoelectric activities u:

$$E\{u \cdot u^T\} = \Phi^T E\{M \cdot M^T\} \Phi = \Phi^T \cdot C(F) \cdot \Phi = \Lambda(F).$$

Thus, the covariance matrix is a diagonal matrix, and the individual transformed myoelectric activities u_j are uncorrelated. The transformation is linear; therefore, the vector of transformed myoelectric activities u is Gaussian distributed. Consequently, the individual transformed myoelectric activities are independent.

A geometric interpretation will clarify the effect of this transformation. Contours of constant likelihood are defined by

$$p(M/F) = \text{constant} = (2\pi)^{-m/2} |C(F)|^{-1/2} \cdot \exp\left[-\tfrac{1}{2} \cdot M^T \cdot C^{-1} \cdot M\right].$$

Taking the natural logarithm yields

$$M^T \cdot C^{-1} \cdot M = 2 \log_e \left[\text{constant } (2\pi)^{-m/2} \cdot |C(F)|^{-1/2}\right] = \text{constant}.$$

This equation defines a hyperquadric surface which, in the two-dimensional case, is an ellipse. The effect of the transformation is simply to rotate the coordinate axes; that is, linear combinations of the original myoelectric activities are formed whose coordinates lie along the major and minor axes of the ellipse; see Fig. 6. Because the major axis coincides with one of the new coordinate axes, the two combinations are uncorrelated.

The problem is now to estimate muscle force on the basis of m uncorrelated and independent channels of transformed myoelectric activity u. Knowing that $\lambda_j(F)$ is the variance of u_j, we make an assumption of amplitude modulation just as in the previous cases. We assume that the functions relating each eigenvalue to force are similar in form and differ only by a constant, i.e.,

$$\lambda_j(F) = \lambda_j \cdot \sigma(F)^2 \qquad \text{all } j$$

or

$$\Lambda(F) = \Lambda \cdot \sigma(F)^2$$

where $\sigma(F)$ is a static nonlinear function of force. This assumption says that the shape of the hyperquadric surface of constant likelihood is independent of force–only its size or amplitude depends on force. In terms of the original covariance matrix $C(F)$, this assumption is equivalent to assuming a constant correlation matrix R, and assuming that the functions relating the variances of the original myoelectric activities to force are similar, differing only by a constant, i.e.,

$$\sigma_j(F)^2 = k_j \sigma_1(F)^2 \qquad \text{for all } j.$$

Although different researchers investigating different muscles and electrode locations have found different forms for these functions, this latter assumption seems quite reasonable. The former assumption of a constant correlation matrix is somewhat more dubious. It is possible, for example, that as force increases, a number of active motor units within range of the electrodes are recruited in a pattern such that the correlation matrix changes with force; however, this is likely to be a weak effect (see Part II of this paper). In any case, these assumptions afford considerable mathematical simplification, so with the above comments in mind, we proceed without further apology. Note that the assumption of constant correlation is the spatial analogy of the previous assumption that the shape

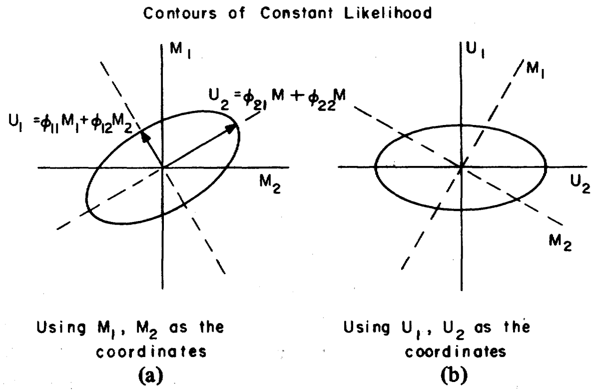


Fig. 6. (a), (b) Geometric interpretation of the orthonormal transformation showing its effect to be a rotation of the coordinate axes.

of the frequency spectrum is constant. This analogy will prove useful.

Now, as in case I, we assume a sequence of N serially uncorrelated samples of the orthonormally transformed myoelectric activity $u_1 \cdots u_N$ where

$$u_i = u(t_i) = [u_1(t_i) \cdots u_j(t_i) \cdots u_m(t_i)]^T.$$

From here, the derivation proceeds as in the previous cases (see Appendix III) and yields as the optimal myoprocessor

$$\hat{F} = \sigma^{-1} \left\{ \left[\frac{1}{N} \sum_{i=1}^{N} \frac{1}{m} \sum_{j=1}^{m} \frac{u_{ij}^2}{\lambda_j} \right]^{1/2} \right\}. \tag{8}$$

To get a clearer picture of what this equation means, we define the transformation:

$$W = \Lambda^{-1/2} \cdot u = \Lambda^{-1/2} \cdot \Phi^T \cdot M \tag{9}$$

where

$$\Lambda^{-1/2} = \begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\sqrt{\lambda_m}} \end{bmatrix}.$$

Thus,

$$\sum_{j=1}^{m} \frac{u_{ij}^2}{\lambda_j} = u^T \cdot \Lambda^{-1} \cdot u = W^T \cdot W = \sum_{j=1}^{m} W_j^2.$$

Substituting into (8), the expression for the optimal myoprocessor becomes

$$\hat{F} = \sigma^{-1} \left\{ \left[\frac{1}{N} \sum_{i=1}^{N} \frac{1}{m} \sum_{j=1}^{m} W_{ij}^2 \right]^{1/2} \right\}.$$

Thus, the required process is: perform the transformation from M to W, then square each channel, and average across space (index j) and across time (index i). Finally, use the nonlinear relation between force and variance to obtain the force estimate. An operational block diagram is shown in Fig. 7.

The covariance of W is

$$E\{W \cdot W^T\} = \Lambda^{-1/2} \cdot E\{u \cdot u^T\} \cdot \Lambda^{-1/2} = \Lambda^{-1/2} \cdot \Lambda(F) \cdot \Lambda^{-1/2} = I \cdot \sigma(F)^2,$$

that is, each of the W_i has the same variance $\sigma^2(F)$. For this reason, the transformation defined by (9) is known as a prewhitening transformation. This is because the orthogonal components of M are transformed into components of W which are uncorrelated and of equal variance. The transformation is illustrated in Fig. 8. Spatial prewhitening is directly analogous to frequency domain prewhitening. A frequency spectrum can be regarded as being composed of a number of orthogonal harmonic components, each of which carries a portion of the total variance of the signal. Frequency domain prewhitening acts to equalize the contribution of each component to the total variance. In direct analogy, spatial prewhitening acts to equalize the contribution of each channel of myoelectric activity to the total variance.

Pursuing the analogy between spatial prewhitening and frequency domain prewhitening, it can be shown that similar caveats apply. If one attempts to combine two channels of myoelectric activity which are perfectly correlated, the prewhitening transformation is degenerate and would require infinite gains. The reason for this failure of the prewhitening is because if the two channels of myoelectric activity are perfectly correlated, the second channel adds no new information. Now, suppose that instead of a correlation of one, we have a correlation of slightly less than one. In this case, the second channel does add some new information, and we do not have a degenerate case as above. However, a very large gain is necessary. Although it was not included in the model, any real situation will have additive noise, and the effect of prewhitening will be to amplify this noise. Thus, although the second channel of myoelectric activity may add a small amount of new information, it is not worthwhile to try to use it. As in the frequency domain case, the message is clear: proceed with caution.

Case IV: Multiple Channels, Dependent Samples

The next logical step would be to extend the analysis to account for the time dependence between samples in the multichannel case. Brief reflection on this problem reveals that, in this case, we would have to deal with an m-input, m-output transfer function representing the effects of the transmission of myoelectric activity to the skin electrode interface. The dimensionality of this problem is extremely cumbersome, and the benefits of the pursuit are not enticing. Inductive reasoning indicates that the analysis would most likely require the optimal myoprocessor to prewhiten in the spatial domain, prewhiten in the frequency domain, and then proceed as before. At this stage, we felt it more important to seek experimental verification of the efficacy of the optimal myoprocessor than to pursue the analysis.

Comparison of Analysis with Previously Published Results

The optimal single-channel myoprocessor derived above encompasses and places in perspective almost all of the improvements of myoprocessing reported in recent literature.

The averaging circuit has been shown to be superior to the common first-order, low-pass filter by Kreifeldt [41] and Gottlieb and Agarwal [42]. Note that under the stated assumptions of our analysis, not only is it superior, it is also the best.

Variations of the nonlinear demodulation with and without

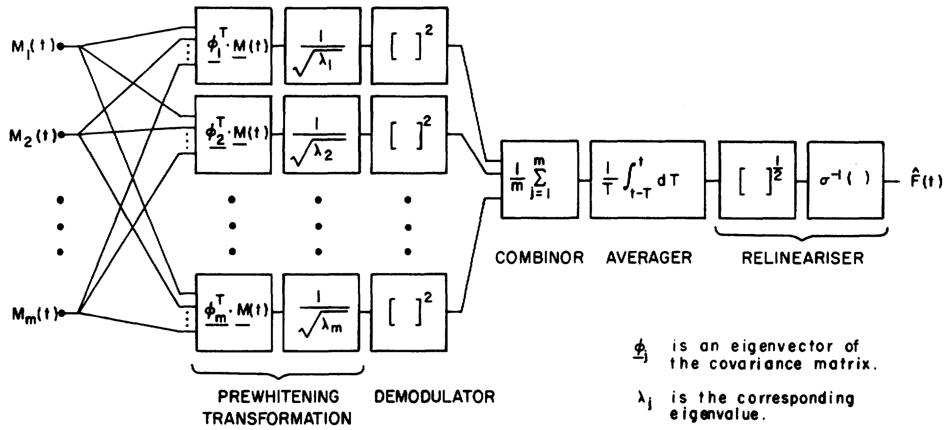


Fig. 7. Block diagram of the optimal multichannel estimator of muscle force. Samples are assumed uncorrelated.

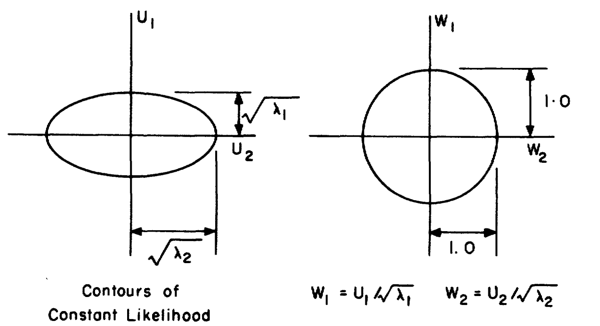


Fig. 8. Geometric interpretation of the effect of the prewhitening transformation showing the rescaling of the orthogonal components to equal variance.

some form of subsequent relinearization were examined by Kreifeldt and Yao [43]. Under the assumption made by Kreifeldt and Yao that mean rectified myoelectric activity varies linearly with force, our analysis predicts that of the processors with relinearization, the square law processor will perform best. This is borne out by their experimental data (except at the 5 percent contraction level). At the 25 percent contraction level, the superiority is marked. It is difficult to put the remainder of their results in perspective because they violate the condition $E\{\hat{F}\} = F$.

The nonlinear transfer characteristic of a root-law demodulator such as Kreifeldt and Yao recommend has been criticized by Kadefors [44] on the grounds that excessive nonlinearity "reduces the number of separable signal levels." In fact, a nonlinear transfer characteristic is necessary in order to meet the requirements that $E\{\hat{F}\} = F$. The overall transfer characteristic of the common myoprocessor using a simple rectifier is "linear" in that the output varies linearly with the mean amplitude of myoelectric activity; however, because of the nonlinear relation between muscle force and myoelectric activity, a nonlinear transfer characteristic is required in order for the processor output to vary linearly with muscle force.

Finally, both the idea of electronic prewhitening and the idea of prewhitening by reducing electrode spacing were suggested by Kaiser *et al.* [20]. Detailed analysis of the situation was performed by Lindstrom [45]. Experimental work was performed by Monster [3] which confirmed bandwidth increase. The effect of reduced electrode spacing on the signal-to-noise ratio of the myoprocessor output is demonstrated experimentally in Part II of this paper.

As can be seen from the above, the analysis agrees well with data reported in the literature. The particular value of the analysis lies not so much in the fact that the reported improvements are derived, but in that it serves to put them in perspective and provides a coherent basis for comparison and evaluation of each improvement. For example, on the one hand, it elucidates why prewhitening must be approached with caution, and on the other hand, it shows how some of the objections to relinearization can be removed.

Discussion

The mathematical techniques developed to solve the myoprocessing problem—a combination of state-space methods and statistical decision theory—are applicable to a broad class of nonlinear estimation problems, particularly those in which the contaminating noise is multiplicative and the desired variable includes zero in its range. This situation is ill-represented in the literature, with the most attention being focused on the problem of estimation in the presence of additive noise. This emphasis on linear filtering of additive noise has led to the often-inappropriate application of textbook results. As pointed out earlier, the myoelectric signal-processing problem is fundamentally nonlinear, and the use of linear techniques is, at best, a poor approximation. Because of the techniques we used, we were able to avoid the assumption made by Parker *et al.* [24] that the myoelectric signal takes on a series of discrete amplitudes, and instead solve for a fully proportional myoprocessor.

The techniques presented here will have wide application in processing other bioelectric signals. For example, recent research has shown the feasibility of detecting the activity of a nerve bundle directly using a chronically implanted nerve electrode [46]. The signal from a nerve bundle is a spatiotemporal combination of the activities of a large number of individual nerve fibers; hence, it will probably be best described as a statistical process, and the methods presented in this paper can be used to decode it.

Although we do not take the analysis any further in this paper, our method shows considerable promise, and we recommend that it be pursued in future work. We derived the optimal myoelectric estimate of muscle force from a simple functional model of myoelectric activity which assumed a Gaussian amplitude distribution. This is adequate for surface myoelectric activity, but for some applications, e.g., intramuscular electrodes, a more detailed structural model such as that developed by De Luca [47] will be necessary. For surface electrodes, De Luca's model reduces to the model we used.

We assumed that surface myoelectric activity depends solely on muscle force. In fact, myoelectric activity is related to muscle force F, muscle length L, velocity of shortening V, and a host of other physiological variables which reflect such things as state of fatigue, temperature, etc. Denoting all of these latter variables by ϕ, we can write this in noncausal form as

$$g(F, M, L, V, \phi) = 0.$$

In causal form, this equation should be written

$$F = g(M/L, V, \phi),$$

that is, the kinematic variables L and V and the other physiological variables ϕ act as parameters of the relation between the output force and the input activity. This is borne out by the experiments of Vredenbregt and Rau [35]. For most purposes, the physiological variables ϕ may be considered constant (i.e., no fatigue) and the kinematic variables may be assumed constant (as in the case of an amputee's stump muscles) or measured directly by means of goniometers. Thus, with minor modifications, the optimal myoprocessor derived above can still be applied.

We assumed that the relation between muscle force and myoelectric activity is static or memoryless. This is not the case. In addition to a pure time delay between myoelectric activity and muscle force [38], [25], the amplitude of myoelectric activity appears to depend on the rate of change of muscle force as well as on its instantaneous value [38], [48]. If this behavior can be incorporated into our mathematical analysis, we anticipate a significant improvement in myoprocessor performance, particularly during the dynamic phase of muscle activity. Finally, our method should be modified to include in the model such effects as additive noise. An approach to both of the above would be to use Bayesian methods to allow the current likelihood function for a sequence of samples to reflect prior information from past values of muscle force. There is some difficulty in specifying the form of the prior distribution, but this should not be insurmountable.

Appendix I
Optimal Single-Channel Myoprocessor, Uncorrelated Samples

The maximum likelihood estimator of muscle force is defined by

$$\frac{d}{dF} p(M/F)\bigg|_{\substack{F=\hat{F}\\ M=M_0}} = 0. \tag{A.1}$$

The location of the extremum is unaffected by any monotonic transformation; hence, (A.1) is equivalent to

$$\frac{d}{dF} [\ln p(M/F)]\bigg|_{\substack{F=\hat{F}\\ M=M_0}} = 0. \tag{A.2}$$

Individual samples of myoelectric activity are Gaussian distributed:

$$p(M_i/F) = \frac{1}{\sqrt{2\pi}\,\sigma(F)} \cdot \exp\left[-\frac{1}{2}\cdot\frac{M_i^2}{\sigma(F)^2}\right]. \tag{A.3}$$

The likelihood function for a finite set of N serially uncorrelated (and by the Gaussian assumption, independent) samples of myoelectric activity is

$$p(M/F) = p(M_1 \cdots M_N/F) = \prod_{i=1}^{N} p(M_i/F).$$

Applying (A.2) and (A.3),

$$\begin{aligned}\ln P(M/F) &= \ln \prod_{i=1}^{N} P(M_i/F) = \sum_{i=1}^{N} \ln P(M_i/F)\\ &= -N \ln\sqrt{2\pi} - N\ln\sigma(F) - \frac{1}{2}\\ &\quad \cdot \sigma(F)^{-2} \cdot \sum_{i=1}^{N} M_i^2\end{aligned}$$

$$\frac{d}{dF}[\ln p(M/F)] = \sigma(F)^{-1} \cdot \frac{d}{dF}[\sigma(F)] \cdot \left[\sigma(F)^{-2}\sum_{i=1}^{N} M_i^2 - N\right] = 0.$$

Because $\sigma(F)$ is a monotonic function of force,

$$\sigma(F)^{-1}\frac{d}{dF}[\sigma(F)]\bigg|_{F=\hat{F}} \neq 0$$

$$\therefore \frac{d}{dF}[\ln p(M/F)]\bigg|_{\substack{F=\hat{F}\\ M=M_0}} = 0 \Rightarrow \sigma(\hat{F})^{-2}\sum_{i=1}^{N} M_i^2 - N = 0$$

$$\therefore \hat{F} = \sigma^{-1}\left\{\left[\frac{1}{N}\sum_{i=1}^{N} M_i^2\right]^{1/2}\right\}. \tag{A.4}$$

The properties of this estimator are demonstrated by deriving expressions for its mean value $E\{\hat{F}\}$ and its signal-to-noise ratio, defined by

$$\mathrm{SNR} \triangleq \left[\frac{E\{F\}^2}{E\{(\hat{F} - E\{\hat{F}\})^2\}}\right]^{1/2}. \tag{A.5}$$

We need the likelihood function for $\hat{F}$. With the change of variables $Y = M^2$, (A.3) becomes

$$p(Y/F) = \text{constant} \cdot Y^{-1/2} \cdot \exp\left[-\frac{1}{2}\cdot\frac{Y}{\sigma(F)^2}\right].$$

This is a Gamma density function:

$$p_\gamma(Y/\alpha, \beta) = \begin{cases}\dfrac{\beta^\alpha}{\Gamma(\alpha)} \cdot Y^{\alpha-1} \cdot e^{-\beta Y}, & Y > 0;\ \alpha, \beta > 0\\ 0, & Y < 0\end{cases} \tag{A.6}$$

with $\alpha = \frac{1}{2}, \beta = 1/2\sigma(F)^2$.

The likelihood function for the variable

$$z = \sum_{i=1}^{N} Y_i = \sum_{i=1}^{N} M_i^2$$

is found by convolving $p(Y/F)$ with itself N times, which is most easily performed via the characteristic function $\zeta(t)$, the inverse Fourier transform of the likelihood function:

$$\zeta_Y(t) = E\{e^{itY}\} = \int_{-\infty}^{\infty} e^{itY} p(Y/F)\, dY.$$

Using (A.6) yields

$$\zeta_Y(t) = \left[\frac{\beta}{\beta - it}\right]^{\alpha}.$$

Under the inverse Fourier transform, convolution becomes multiplication; thus,

$$\zeta_Z(t) = \prod_{i=1}^{N} \zeta_{Y_i}(t) = \left[\frac{\beta}{\beta - it}\right]^{N\alpha}.$$

Hence, Z is Gamma distributed (A.6) with

$$\alpha_Z = N\alpha = \frac{N}{2}; \quad \beta_Z = \beta = \frac{1}{2\sigma(F)^2}.$$

From (A.4),

$$\hat{F} = \sigma^{-1}\left\{\left[\frac{Z}{N}\right]^{1/2}\right\}. \tag{A.7}$$

To proceed, we need an explicit form for the relation $\sigma = \sigma(F)$ and its inverse. This is obtained from the data of Vredenbregt and Rau [35] relating muscle force and mean rectified myoelectric activity $E\{|M|\}$ if we note that

$$E\{|M|\} = \int_{-\infty}^{\infty} |M| \cdot p(M)\, dM$$

$$= 2 \cdot \int_{0}^{\infty} M \cdot \frac{1}{\sqrt{2\pi}\,\sigma} \cdot \exp\left[-\frac{1}{2} \cdot \frac{M^2}{\sigma^2}\right] \cdot dM.$$

Substitute $V = \frac{1}{2} \cdot (M^2/\sigma^2)$; $E\{|M|\} = \sigma\sqrt{2/\pi} \cdot \int_0^\infty e^{-v}\, dv = \sigma\sqrt{2/\pi}$. Thus, $E\{|M|\}$ is proportional to σ. For $\sigma = \sigma(F)$, we assumed a relation of the form $\sigma = k \cdot F^a$; transform this to a linear equation by taking the logarithm

$$\ln \sigma = \ln k + a \ln F$$

and fit Vredenbregt and Rau's data to this equation using least-squares linear regression techniques. The value obtained for a was 1.74, with a correlation coefficient of 0.98, indicating an adequate fit. k is an arbitrary scaling constant.

From (A.7), we now have

$$\hat{F} = \left[\frac{Z}{k^2 N}\right]^{1/2a}$$

$$E\{\hat{F}\} = \int_{-\infty}^{\infty} \left[\frac{Z}{k^2 N}\right]^{1/2a} \cdot p(Z/F)\, dZ$$

$$= \frac{\Gamma(N/2 + 1/2a)}{\Gamma(N/2)} \cdot \left[\frac{2}{N}\right]^{1/2a} \cdot F.$$

Thus, the mean value of the estimate $\hat{F}$ is proportional to actual muscle force F.

$$E\{\hat{F}^2\} = \int_{-\infty}^{\infty} \left[\frac{Z}{k^2 N}\right]^{2/2a} \cdot p(Z/F)\, dZ$$

$$= \frac{\Gamma(N/2 + 1/a)}{\Gamma(N/2)} \cdot \left[\frac{2}{N}\right]^{1/a} \cdot F^2$$

$$E\{(\hat{F} - E\{\hat{F}\})^2 = E\{F^2\} - E\{\hat{F}\}^2$$

$$= \left[\frac{\Gamma(N/2 + 1/a)\,\Gamma(N/2)}{\Gamma(N/2 + 1/2a)^2} - 1\right] \cdot E\{\hat{F}\}^2. \tag{A.8}$$

Using (A.8) in (A.5),

$$\mathrm{SNR} = \left[\frac{\Gamma(N/2 + 1/a) \cdot \Gamma(N/2)}{\Gamma(N/2 + 1/2a)^2} - 1\right]^{-1/2} = \text{constant}.$$

Thus, the standard deviation about the mean is proportional to the mean.

Appendix II
Optimal Single-Channel Myoprocessor, Dependent Samples

The differential equation describing the linear time-invariant filter of Fig. 1 is written in state-space notation as

$$\frac{d}{dt} x(t) = A \cdot x(t) + b \cdot w(t); \quad N(t) = c^T \cdot x(t). \tag{A.9}$$

Solution of this equation yields the matrix superposition integral:

$$x(t) = \Phi(t - t_0) \cdot x(t_0) + \int_{t_0}^{t} \Phi(t - \tau) \cdot b \cdot w(\tau) \cdot d\tau \tag{A.10}$$

or

$$N(t) = \int_{-\infty}^{t} c^T \cdot \Phi(t - \tau) \cdot b \cdot w(\tau) \cdot d\tau$$

$$= \int_{-\infty}^{t} H(t - \tau) \cdot w(\tau)\, d\tau. \tag{A.11}$$

$\Phi(t - \tau) = \exp\,[A(t - \tau)]$ is the state transition matrix.

$H(t - \tau) = c^T \cdot \Phi(t - \tau) \cdot b$ is the system weighting function; its Laplace transform is the system transfer function $H(s)$:

$$H(s) = \frac{N(s)}{w(s)} = c^T \cdot (sI - A)^{-1} \cdot b.$$

Substituting $s = i2\pi f$ yields the frequency response function $H(f)$:

$$H(f) = c^T \cdot (i2\pi f I - A)^{-1} \cdot b.$$

This is a ratio of complex-valued polynomials in f.

The autocorrelation function of $N(t)$ is found as follows:

$$R_N(t_1, t_2) \triangleq E\{N(t_1) \cdot N(t_2)^T\}. \tag{A.12}$$

Using (A.9),

$$R_N(t_1, t_2) = c^T \cdot E\{x(t_1)\, x(t_2)^T\} \cdot c = c^T \cdot R_x(t_1, t_2) \cdot c.$$

Assume that $t_1 \geqslant t_2$:

$$E\{x(t_1) \cdot x(t_2)^T\} = \Phi(t_1 - t_2)\, E\{x(t_2)\, x(t_2)^T\} + \int_{t_2}^{t_1} \Phi(t_1 - \tau) \cdot b \cdot E\{w(\tau) \cdot x(t_2)^T\}\, d\tau.$$

Because the current state $x(t_2)$ is independent of future inputs $w(\tau)$,

$$E\{w(\tau) \cdot x(t_2)^T\} = 0; \qquad \tau \geqslant t_2.$$

Therefore,

$$E\{x(t_1) \cdot x(t_2)^T\} = \Phi(t_1 - t_2) \cdot E\{x(t_2) \cdot x(t_2)^T\}$$

$$R_N(t_1, t_2) = c^T \cdot \Phi(t_1 - t_2) \cdot R_x(T_2, t_2) \cdot c.$$

In the stationary case, $R_N(t_1, t_2) = R_N(t_1 - t_2) = R_N(\tau)$:

$$R_N(\tau) = c^T \cdot \Phi(\tau) \cdot R_x(0) \cdot c.$$

For algebraic simplicity, assume a first-order, low-pass filter:

$$R_N(\tau) = \Phi(\tau) \cdot R_x(0) \cdot c^2; \qquad R_N(0) = R_x(0) \cdot c^2$$

$$R_N(\tau) = \Phi(\tau) \cdot R_N(0); \qquad \tau \leqslant 0.$$

By symmetry, a similar result holds for $\tau \geqslant 0$; thus,

$$R_N(\tau) = \Phi(|\tau|) \cdot R_N(0).$$

The autocorrelation function of myoelectric activity is

$$R_M(\tau) = \Phi(|\tau|) \cdot R_N(0) \cdot \sigma(F)^2.$$

For convenience, we assume that $R_N(0) = 1$. For a first-order, low-pass filter with time constant τ_c, $\Phi(\tau) = e^{-\tau/\tau_c}$.

Thus,

$$R_M(\tau) = e^{-|\tau|/\tau_c} \cdot \sigma(F)^2$$

or

$$R_M(\tau) = \rho \cdot \sigma(F)^2 \qquad \text{where} \quad \rho = e^{-|\tau|/\tau_c}. \tag{A.13}$$

Thus, the correlation coefficient ρ is a function of the time separation of two samples τ and the "correlation time" of the process τ_c.

The power spectral density of $M(t)$ is defined by

$$S_M(f) \triangleq \int_{-\infty}^{\infty} R_M(\tau)\, e^{-i2\pi f t}\, d\tau. \tag{A.14}$$

Using (A.11) and (A.12) in (A.14) yields

$$S_M(f) = H(-f) \cdot H(f) \cdot S_w(f) \cdot \sigma(F)^2$$

where $S_w(f)$ is the power spectral density of the input white noise process, constant with respect to frequency, i.e., $S_w(f) = Q$. Because $H(-f)$ is the complex conjugate of $H(f)$,

$$S_M(f) = |H(f)|^2 \cdot Q \cdot \sigma(F)^2.$$

Thus, the power spectrum of myoelectric activity is a ratio of polynomials in f^2.

Assume N equally spaced samples of myoelectric activity $M_1 \cdots M_N$; $M_i = M(t_i)$; $t_1 < \cdots < t_i < \cdots t_N$. Each of these samples is Gaussian distributed (A.3) and the samples are serially dependent. The likelihood function for the N samples is found by repeated application of Bayes' rule:

$$p(M_1 \cdots M_N/F) = p(M_N/M_{N-1}, \cdots, M_1, F) \cdot p(M_{N-1}/M_{N-2}, \cdots, M_1, F) \cdots p(M_1/F)$$

and application of the first-order Markov property:

$$p(M_1 \cdots M_N/F) = p(M_N/M_{N-1}, F) \cdot p(M_{N-1}/M_{N-2}, F) \cdots p(M_1/F) = \left[\prod_{i=2}^{N} p(M_i/M_{i-1}, F)\right] \cdot p(M_1/F). \tag{A.15}$$

To find the transition likelihood function $p(M_i/M_{i-1}, F)$, we apply Bayes' rule and use (A.3):

$$p(M_i/M_{i-1}, F) = \frac{p(M_i, M_{i-1}/F)}{p(M_{i-1}/F)}.$$

Write

$$M = \begin{bmatrix} M_i \\ M_{i-1} \end{bmatrix}; \qquad E\{M\} = \mathbf{0};$$

$$E\{M \cdot M^T\} = \begin{bmatrix} R_M(0) & R_M(\Delta t) \\ R_M(-\Delta t) & R_M(0) \end{bmatrix} = \sigma(F)^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} = \Sigma(F);$$

Δt = sample spacing.

$$p(M_i, M_{i-1}/F) = p(M/F) = \frac{1}{2\pi \cdot |\Sigma(F)|^{1/2}} \cdot \exp\left[-\frac{1}{2} M^T \cdot \Sigma(F)^{-1} \cdot M\right]$$

$$|\Sigma(F)|^{1/2} = \sigma(F)^2 \sqrt{1 - \rho^2}$$

$$M^T \cdot \Sigma(F)^{-1} \cdot M = \frac{(M_i^2 - 2\rho M_i M_{i-1} + M_{i-1}^2)}{\sigma(F)^2\,(1 - \rho^2)}$$

$$p(M_{i-1}/F) = \frac{1}{\sqrt{2\pi}\,\sigma(F)} \cdot \exp\left[-\frac{1}{2}\frac{M_{i-1}^2}{\sigma(F)^2}\right]$$

$$p(M_i/M_{i-1}, F) = \frac{1}{\sqrt{2\pi}\,\sigma(F)\sqrt{1-\rho^2}} \cdot \exp\left[-\frac{1}{2}\cdot\frac{(M_i - \rho M_{i-1})^2}{\sigma(F)^2(1-\rho^2)}\right]. \quad \text{(A.16)}$$

Using (A.3), (A.15), and (A.16) in (A.2),

$$\ln p(M_1 \cdots M_N F) = -N \ln\sqrt{2\pi} - N \ln \sigma(F) - (N-1)\ln\sqrt{1-\rho^2} - \frac{1}{2\sigma(F)^2} \cdot \left[\sum_{i=2}^{N} \frac{(M_i - \rho M_{i-1})^2}{(1-\rho^2)} + M_1^2\right]$$

$$\frac{d}{dF}\ln p(M_1 \cdots M_N/F) = \sigma(F)^{-1}\frac{d}{dF}[\sigma(F)] \cdot \left\{\sigma(F)^{-2}\left[\sum_{i=2}^{N}\frac{(M_i - M_{i-1})^2}{(1-\rho^2)} + M_1^2\right] - N\right\}.$$

As before,

$$\sigma(F)^{-1}\cdot\frac{d}{dF}[\sigma(F)]\bigg|_{F=\hat{F}} \neq 0$$

$$\therefore \frac{d}{dF}[\ln p(M_1 \cdots M_N/F)]\bigg|_{\substack{F=\hat{F}\\ M=M_0}} = 0$$

$$\Rightarrow \sigma(\hat{F})^{-2}\left[\sum_{i=2}^{N}\frac{(M_i - \rho M_{i-1})^2}{(1-\rho^2)} + M_1^2\right] - N = 0$$

$$\sigma(\hat{F})^2 = \frac{1}{N}\left[\sum_{i=2}^{N}\frac{(M_i - \rho M_{i-1})^2}{(1-\rho^2)} + M_1^2\right] \quad \text{(A.17)}$$

$$\hat{F} = \sigma^{-1}\left[\left\{\frac{1}{N}\left[\sum_{i=2}^{N}\frac{(M_i - \rho M_{i-1})^2}{(1-\rho^2)} + M_1^2\right]\right\}^{1/2}\right]. \quad \text{(A.18)}$$

To understand (A.18), recall that M_i was modeled as linearly filtered white noise. The discrete time form of the filter is obtained by assuming the white noise process to be sampled at intervals of Δt, yielding a purely random sequence of steps which are then input to the system of (A.9) and (A.10):

$$N(t+\Delta t) = c^T \cdot \Phi(\Delta t) \cdot \dot{x}(t) + \int_t^{t+\Delta t} c^T \cdot \Phi(t-\tau) \cdot b \cdot d\tau \cdot w(t).$$

For the first-order low-pass filter assumed above,

$$\frac{dx(t)}{dt} = -\lambda \cdot x(t) + \lambda \cdot w(t); \qquad N(t) = x(t); \qquad \lambda = 1/\tau_c$$

$$N(t+\Delta t) = e^{-\lambda\Delta t}\cdot N(t) + \frac{1}{-\lambda}(e^{-\lambda\Delta t} - 1)\cdot\lambda\cdot w(t).$$

Using $\rho = e^{-\lambda\Delta t}$ (A.13) and $M(t) = \sigma(F)\cdot N(t)$, we get

$$M_i = \rho M_{i-1} + \sigma(F)(1-\rho)\cdot w_{i-1}. \quad \text{(A.19)}$$

Let

$$u_{i-1} = \frac{M_i - \rho M_{i-1}}{\sqrt{1-\rho^2}}. \quad \text{(A.20)}$$

Thus,

$$\hat{F} = \sigma^{-1}\left[\left\{\frac{1}{N}\left[\sum_{i=2}^{N} u_{i-1}^2 + M_1^2\right]\right\}^{1/2}\right].$$

Using (A.19), we see that

$$u_{i-1}^2 = \frac{(M_i - \rho M_{i-1})^2}{(1-\rho^2)} = \frac{(1-\rho)^2}{(1-\rho^2)}\cdot\sigma(F)^2\cdot w_{i-1}^2.$$

Thus, (A.20) is a recursive formula defining a filter to restore the sequence of samples of myoelectric activity to a purely random or "white" sequence. The term involving ρ is a scaling factor to correct for the variance of the purely random input sequence, as can be seen by taking the expectation of (A.17):

$$E\{\sigma(\hat{F})^2\} = \frac{1}{N}\left[\sum_{i=2}^{N}\frac{E\{M_i^2 - 2\rho M_i M_{i-1} + \rho^2 M_{i-1}^2\}}{1-\rho^2} + E\{M_1^2\}\right]$$

$$= \frac{1}{N}\left[\sum_{i=2}^{N}\frac{\sigma(F)^2 - 2\rho\cdot\rho\sigma(F)^2 + \rho^2\sigma(F)^2}{1-\rho^2} + \sigma(F)^2\right]$$

$$= \frac{1}{N}\left[\sum_{i=2}^{N}\sigma(F)^2\frac{(1-\rho^2)}{(1-\rho^2)} + \sigma(F)^2\right] = \sigma(F)^2.$$

As no restriction was placed on Δt, to obtain the continuous-time form of the estimator, we take Δt to zero in the limit, with $N = T/\Delta t$. However, this limit is indeterminate and requires the use of mean-square calculus [39] for its evaluation. Instead, we argue by induction that as the effect of (A.20) is to convert M_i to a "white" sequence, the continuous-time estimator requires a prewhitening or complementary filter, i.e., one whose frequency-response function is such that

$$H^*(f)\cdot H(f) = 1$$

in order to convert the power spectrum of $M(t)$ to that of a white noise process.

Appendix III
Optimal Multiple-Channel Myoprocessor, Uncorrelated Samples

The activities of m channels of myoelectric activity are denoted by

$$M(t) = [M_1(t) \cdots M_m(t)]^T.$$

The vector $M(t)$ is Gaussian distributed with zero mean:

$$p\{M(t)/F\} = \frac{1}{(2\pi)^{m/2}|C(F)|^{1/2}} \cdot \exp\left[-\frac{1}{2}\cdot M(t)^T\cdot C(F)^{-1}\right]\cdot M(t) \quad \text{(A-21)}$$

where $C(F)$ is the covariance matrix defined by

$$C(F) \triangleq E\{M(t) \cdot M(t)^T\}; \quad E\{M(t)\} = \mathbf{0}.$$

An orthonormal transformation is performed on the vector $M(t)$:

$$u(t) = \Phi^T \cdot M(t) \tag{A.22}$$

where $C(F) \cdot \Phi = \Phi \cdot \Lambda(F)$; $\Phi^T \cdot \Phi = I$. Φ is a matrix of unit eigenvectors. $\Lambda(F)$ is a diagonal matrix of eigenvalues. The transformation of (A.22) is linear; thus, the likelihood function for the vector u is Gaussian:

$$p(u/F) = \frac{1}{(2\pi)^{m/2} |\Lambda(F)|^{1/2}} \cdot \exp\left[-\frac{1}{2} \cdot u^T \cdot \Lambda(F)^{-1} \cdot u\right]. \tag{A.23}$$

The amplitude-modulation assumption is generalized by assuming that

$$\Lambda(F) = \Lambda \cdot \sigma^2(F); \quad \Phi = \text{constant}.$$

Now as in case I, assume N vector samples of myoelectric activity spaced in time such that their serial correlation is approximately zero. Orthonormal transformation yields N vector samples of uncorrelated activity, $u_1 \cdots u_N$. By the Gaussian assumption, uncorrelatedness implies independence; hence, the likelihood function for the N samples is

$$p(u_1 \cdots u_N/F) = \prod_{i=1}^{N} p(u_i/F). \tag{A.24}$$

Using (A.23) and (A.24) in (A.2),

$$\ln p(u_1 \cdots u_N/F) = \sum_{i=1}^{N} \left[-\ln(2\pi)^{m/2} - \ln|\Lambda(F)|^{1/2} - \frac{1}{2} u_i^T \cdot \Lambda(F)^{-1} \cdot u_i\right].$$

Because $\Lambda(F)$ is diagonal, its determinant and inverse have simple forms:

$$|\Lambda(F)| = |\Lambda| \, \sigma(F)^{2m} = \left[\prod_{j=1}^{m} \lambda_j\right] \cdot \sigma(F)^{2m}$$

$$\Lambda(F)^{-1} = \Lambda^{-1} \cdot \sigma(F)^{-2} = \begin{bmatrix} \frac{1}{\lambda_1} & & \\ & \ddots & \\ 0 & & \frac{1}{\lambda_m} \end{bmatrix} \cdot \frac{1}{\sigma(F)^2}$$

$$u_i^T \cdot \Lambda(F)^{-1} \cdot u_i = u_i^T \cdot \Lambda^{-1} \cdot u_i \cdot \sigma(F)^{-2} = \frac{1}{\sigma(F)^2} \sum_{j=1}^{m} \frac{u_{ij}^2}{\lambda_{ij}}$$

$$\therefore \ln p(u_i \cdots u_N/F) = -N \cdot \ln(2\pi)^{m/2} - N \cdot \ln|\Lambda|^{1/2} - N \cdot m \cdot \ln \sigma(F) - \frac{1}{2\sigma(F)^2} \sum_{i=1}^{N} \sum_{j=1}^{m} \frac{u_{ij}^2}{\lambda_j}$$

$$\frac{d}{dF}[\ln p(u_1 \cdots u_N/F)] = \sigma(F)^{-1} \frac{d}{dF}[\sigma(F)] \cdot \left[\sigma(F)^{-2} \sum_{i=1}^{N} \sum_{j=1}^{m} \frac{u_{ij}^2}{\lambda_j} - N \cdot m\right].$$

As before,

$$\sigma(F)^{-1} \frac{d}{dF}[\sigma(F)]\bigg|_{F=\hat{F}} \neq 0.$$

$$\therefore \frac{d}{dF}[\ln p(u_1 \cdots u_N/F)]\bigg|_{\substack{F=\hat{F} \\ M=M_0}}$$

$$\Rightarrow \sigma(F)^{-2} \sum_{i=1}^{N} \sum_{j=1}^{m} \frac{u_{ij}^2}{\lambda_j} - N \cdot m = 0$$

$$\hat{F} = \sigma^{-1}\left\{\left[\frac{1}{N} \sum_{i=1}^{N} \frac{1}{m} \sum_{j=1}^{m} \frac{u_{ij}^2}{\lambda_j}\right]^{1/2}\right\}.$$

Acknowledgment

The authors wish to acknowledge the valuable counsel of Prof. T. B. Sheridan and Prof. E. Bizzi of the Massachusetts Institute of Technology and Dr. C. J. De Luca of the Children's Hospital Medical Center.

References

[1] N. Hogan, "A review of the methods of processing EMG for use as a proportional control signal," *Biomed. Eng.*, vol. 11, pp. 81-86, Mar. 1976.

[2] R. Kadefors, "Control components in rehabilitation engineering," Res. Lab. Med. Electron., Goteborg, Sweden, Final Rep. 9:73, Dec. 1973.

[3] A. W. Monster, "Basic aspects of the communication problem in the clinical use of assistive devices," Ph.D. dissertation, School Eng., Case Western Reserve Univ., Cleveland, OH, Jan. 1970.

[4] D. Radonjic and C. Long, "Why myoelectric control is so difficult," in *Proc. 3rd Int. Symp. External Contr. of Human Extremities*, Dubrovnik, Yugoslavia, Aug. 1970, pp. 59-67.

[5] L. Vodovnik and S. Rebersek, "Myoelectric and myomechanical prehension systems using functional electrical stimulation," in *The Control of Upper Extremity Prosthesis and Orthoses*, P. Herberts *et al.*, Eds. Springfield, IL: Thomas, 1974.

[6] R. W. Mann and S. D. Reimers, "Kinesthetic sensing for the EMG-controlled 'Boston arm,'" *IEEE Trans. Man-Machine Syst.*, vol. MMS-11, pp. 110-115, Mar. 1970.

[7] D. A. Alles, "Information transmission by phantom sensations," *IEEE Trans. Man-Machine Syst.*, vol. MMS-11, pp. 85-91, Mar. 1970.

[8] R. K. Sundstrom, "An investigation of cutaneous phantom sensations produced via electrical stimulation," S.M. thesis, Dep. Mech. Eng., M.I.T., Cambridge, Aug. 1974.

[9] G. B. Rollman, "Electrocutaneous stimulation," in *Proc. Conf. Cutaneous Commun. Syst. and Devices*, F. A. Geldard, Ed., Monterey, CA, 1973, p. 42.

[10] R. E. Prior, J. Lyman, P. A. Case, and C. M. Scott, "Supplemental sensory feedback for the VA/NU myoelectric hand–Background and preliminary designs," *Bull. Prosthetics Res.*, vol. BPR 10-26, pp. 170-191, Fall 1976.

[11] T. A. Rohland, "Sensory feedback in upper-limb prosthetic systems," *Inter-Clinic Inform. Bull.*, vol. 23, pp. 1-8, June 1974.

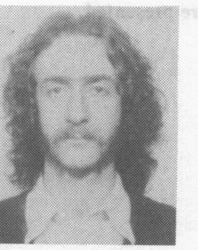
[12] P. A. Merton, "Speculations on the servo-control of movement," in *The Spinal Cord*, C. E. W. Wolstenholme, Ed. London, England: Churchill, 1953, pp. 247-255.

[13] R. Granit, *Receptors and Sensory Perception.* New Haven, CT: Yale Univ. Press, 1955.

[14] P. H. Hammond, "An experimental study of servo action in human muscular control," in *Proc. Int. Conf. Med. Electron.*, Inst. Elec. Eng., London, England, 1960, pp. 190-199.

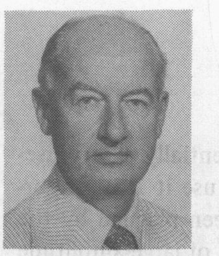
[15] A. Polit and E. Bizzi, "Processes controlling arm movements in monkeys," *Science*, vol. 201, pp. 1235-1237, Sept. 1978.

[16] E. Bizzi, P. Dev, P. Morasso, and A. Polit, "Effect of load disturbances during centrally initiated movements," *J. Neurophysiol.*, vol. 41, pp. 542-556, May 1978.
[17] E. Bizzi, A. Polit, and P. Morasso, "Mechanisms underlying achievement of final head position," *J. Neurophysiol.*, vol. 39, pp. 435-444, Mar. 1976.
[18] J. H. J. Allum, "Responses to load disturbances in human shoulder muscles: The hypothesis that one component is a pulse test information signal," *Exp. Brain Res.*, vol. 22, pp. 307-326, 1975.
[19] J. Dichgans, E. Bizzi, P. Morasso, and V. Tagliasco, "Mechanisms underlying recovery of eye-head co-ordination following bilateral labyrinthectomy in monkeys," *Exp. Brain Res.*, vol. 18, pp. 548-562, 1973.
[20] E. Kaiser, R. Kadefors, R. Magnusson, and J. Petersen, "Myoelectric signals for prosthesis control," *Medicinsk Teknik/Medico Teknik*, no. 1, 1968.
[21] L. Dhareshwar, "Crosstalk in myoelectric control systems," M.Sc. thesis, Univ. New Brunswick, Fredericton, N.B., Canada, 1967.
[22] E. Kwatny, D. H. Thomas, and H. G. Kwatny, "An application of signal processing techniques to the study of myoelectric signals," *IEEE Trans. Bio-Med. Eng.*, vol. BME-17, pp. 303-313, 1970.
[23] H. Roesler, "Statistical analysis and evaluation of myoelectric signals," in *The Control of Upper Extremity Prostheses and Orthoses*, P. Herberts *et al.*, Eds. Springfield, IL: Thomas, 1974.
[24] P. A. Parker, J. A. Stuller, and R. N. Scott, "Signal processing for the multistate myoelectric channel," *Proc. IEEE*, vol. 65, pp. 662-674, May 1977.
[25] G. Weltman, H. Groth, and J. Lyman, "An analysis of bioelectric prosthesis control," Biotechnol. Lab., Univ. California, Los Angeles, Tech. Rep. 59-49, July 1959.
[26] A. Nightingale, "The analysis of muscle potentials by means of a Muirhead-Pametrada wave analyser," *Muirhead Tech.*, vol. 11, no. 4, pp. 27-28, 1957.
[27] S. C. Jacobsen and R. W. Mann, "Control systems for artificial arms," in *Proc. IEEE Conf. Syst., Man, Cybern.*, Boston, MA, 1973.
[28] R. B. Jerard, "Application of a unified theory for simultaneous multiple axis artificial arm control," Ph.D. dissertation, Dep. Mech. Eng., Univ. Utah, Salt Lake City, Dec. 1976.
[29] O. C. J. Lippold, "The relation between integrated action potentials in a human muscle and its isometric tension," *J. Physiol.*, vol. 117, pp. 492-499, 1952.
[30] G. C. Knowlton *et al.*, "Relation between electromyographic voltage and load," *J. Appl. Physiol.*, vol. 9, pp. 473-476, 1956.
[31] V. F. Inman, H. J. Ralston, C. M. Saunders, B. Feinstein, and E. Wright, "Relation of human electromyogram to muscle tension," *Electroencephalogr. Clin. Neurophys.*, vol. 4, pp. 107-194, 1952.
[32] A. H. Bottomley, "The control of muscles," in *Progress in Biocybernetics*, vol. 1, N. Wiener and J. P. Schade, Eds. Amsterdam: Elsevier, 1964, pp. 124-131.
[33] E. Kuroda, V. Klissouras, and H. J. Milsum, "Electrical and metabolic activity and fatigue in human isometric contraction," *J. Appl. Physiol.*, vol. 29, pp. 358-367, 1970.
[34] A. Nightingale, "The graphic representation of movement II. Relationship between muscle force and the electromyogram in the stand-at-ease position," *Appl. Phys. Med.*, vol. 5, pp. 187-191, 1960.
[35] J. Vredenbregt and G. Rau, "Surface electromyography in relation to force, muscle length and endurance," in *New Developments in EMG and Clinical Neurophysiology*, J. E. Desmedt, Ed. New York: Karger, 1973.
[36] R. Alter, "Bioelectric control of prostheses," Sc.D. dissertation, Dep. Elec. Eng., M.I.T., Cambridge, 1965.
[37] H. Raiffa and R. Schlaiffer, *Applied Statistical Decision Theory*. Cambridge, MA: M.I.T. Press, 1961.
[38] N. J. Hogan, "Myoelectric prosthesis control: Optimal estimation applied to EMG and the cybernetic considerations for its use in a man-machine interface," Ph.D. dissertation, Dep. Mech. Eng., M.I.T., Cambridge, Aug. 1976.
[39] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York: Academic, 1970.
[40] R. Kadefors, A. W. Monster, and I. Petersen, "A new aspect on electrode design in myo-electric control systems," in *Proc. 8th Int. Conf. Med. Biol. Eng.*, Chicago, IL, 1969.
[41] J. G. Kreifeldt, "Signal versus noise characteristics of filtered EMG used as a control source," *IEEE Trans. Biomed. Eng.*, vol. BME-18, Jan. 1971.
[42] G. L. Gottlieb and G. C. Agarwal, "Filtering of electromyographic signals," *Amer. J. Phys. Med.*, vol. 49, 1970.
[43] J. G. Kreifeldt and S. Yao, "A signal-to-noise investigation of nonlinear electromyographic processors," *IEEE Trans. Biomed. Eng.*, vol. BME-21, July 1974.
[44] R. Kadefors, "Control components in rehabilitation engineering," Dep. Appl. Electron., Chalmers Univ. Technol., Goteborg, Sweden, Final Rep., 1973.
[45] L. H. Lindstrom and R. I. Magnusson, "Interpretation of myoelectric power spectra: A model and its applications," *Proc. IEEE*, vol. 65, pp. 653-662, May 1977.
[46] C. J. De Luca and L. D. Gilmore, "Voluntary nerve signals from severed mammalian nerves: Long-term recordings," *Science*, vol. 191, pp. 193-195, Jan. 1976.
[47] C. J. De Luca, "Physiological and mathematical basis of myoelectric signals," in *Muscles Alive*, 4th ed., J. V. Basmajian, Ed. Baltimore, MD: Williams & Wilkins, 1978, ch. 3.
[48] G. L. Gottlieb and G. C. Agarwal, "Dynamic relationship between isometric muscle tension and the electromyogram in man," *J. Appl. Physiol.*, vol. 30, pp. 345-351, 1971.



Neville Hogan was born in Dublin, Ireland. He received the Diploma degree in engineering (with distinction) from Dublin College of Technology, Dublin, Ireland, in 1970, and the S.M. Mechanical Engineering, Mechanical Engineer, and Ph.D. Mechanical Engineering degrees from the Massachusetts Institute of Technology, Cambridge, in 1973, 1976, and 1977, respectively.

He spent a year in Irish industry as a Product Development and Design Engineer followed by a year as a Research Associate and Lecturer in the Department of Mechanical Engineering and the Department of Psychology at the Massachusetts Institute of Technology. He presently holds an appointment as Assistant Professor of Mechanical Engineering at M.I.T. Professional interests are in design, system dynamics, and control, with research activities in the control of movement and the development of manipulatory assistive devices.



Robert W. Mann (SM'71-F'79) received his vocational high school training in Brooklyn, NY. He received the S.B. degree, S.M. degree in mechanical engineering, and the Sc.D. degree from the Massachusetts Institute of Technology, Cambridge, in 1950, 1951, and 1957, respectively.

His career revolves around engineering design. He worked at Bell Labs as a Draftsman before serving in World War II in teletypewriter and cryptographic maintenance. Postwar employment included Bell Labs where he worked on microwave apparatus. After completing the S.M. degree he accepted professional employment as a Design Engineer in the Dynamic Analysis and Control Laboratory, Massachusetts Institute of Technology, in 1951, and was soon head of its Design Division where he supervised the design, development, and testing of missile components. He was appointed Assistant Professor in Mechanical Engineering in 1953. In 1957 he assumed responsibility for the Engineering Design Division of the academic department. His research and teaching are oriented towards the rehabilitation of the blind and amputees. Computer translation into braille and braille typewriters, ultrasonic mobility aids for the blind traveller, and the "Boston Arm" prosthesis for the upper extremity amputee are among his projects which have reached practical application. His current work ranges across human rehabilitation, degenerative arthritis and skeletal joints, and biomechanics.