A Wavelet-Based Continuous Classification Scheme for Multifunction Myoelectric Control

Kevin Englehart*, *Member, IEEE*, Bernard Hudgins, *Member, IEEE*, and Philip A. Parker, *Senior Member, IEEE*

***Abstract*—This work represents an ongoing investigation of dexterous and natural control of powered upper limbs using the myoelectric signal. When approached as a pattern recognition problem, the success of a myoelectric control scheme depends largely on the classification accuracy. A novel approach is described that demonstrates greater accuracy than in previous work. Fundamental to the success of this method is the use of a wavelet-based feature set, reduced in dimension by principal components analysis. Further, it is shown that four channels of myoelectric data greatly improve the classification accuracy, as compared to one or two channels. It is demonstrated that exceptionally accurate performance is possible using the steady-state myoelectric signal. Exploiting these successes, a robust online classifier is constructed, which produces class decisions on a continuous stream of data. Although in its preliminary stages of development, this scheme promises a more natural and efficient means of myoelectric control than one based on discrete, transient bursts of activity.**

***Index Terms*—EMG, myoelectric, pattern recognition, principal components analysis, prosthesis, Wavelet, wavelet packet.**

I. Introduction

THE myoelectric signal (MES), recorded at the surface of the skin, has been used for many diverse applications, including clinical diagnosis, and as a source of control of assistive devices and schemes of functional electrical stimulation. This work seeks to improve the functionality and ease of control of powered upper-limb prostheses using the myoelectric signal.

Many myoelectric control systems are currently available that are capable of controlling a single device in a prosthetic limb, such as a hand, an elbow, or a wrist. These systems extract control information from the MES based on an estimate of the amplitude [1] or the rate of change [2] of the MES. Although these systems have been very successful, they do not provide sufficient information to reliably control more than one function (or device) [3]; the extension to controlling multiple functions, is a much more difficult problem. Unfortunately, these are the requirements of those with high-level (above the elbow) limb deficiencies, and these are the individuals who could stand to benefit most from a functional replacement of their absent limbs.

Manuscript received March 28, 2000; revised November 7, 2000. This work was supported by the Natural Sciences and Engineering Research Council of Canada and the Whitaker Foundation. *Asterisk indicates corresponding author.*

*K. Englehart is with the Department of Electrical and Computer Engineering and the Institute of Biomedical Engineering, University of New Brunswick, P.O. Box 4400, Fredericton, NB, E3B5A3, Canada, (e-mail: kengleha@unb.ca).

B. Hudgins and P. A. Parker are with the Department of Electrical and Computer Engineering and the Institute of Biomedical Engineering, University of New Brunswick, Fredericton, NB, E3B5A3, Canada

Publisher Item Identifier S 0018-9294(01)01570-1.

II. Background

A. Pattern Recognition-Based Control

In an attempt to increase the information extracted from the MES, investigators have proposed a variety of feature sets, and have utilized pattern recognition methods to discriminate amongst desired classes of limb activation. Most work in MES classification has considered the steady-state MES: that collected during a maintained (usually constant-force) contraction. Hudgins *et al*. [4] were the first to consider the information content of the transient bursts of myoelectric activity that accompany the onset of contraction. These data were acquired in a single MES channel, using a widely spaced bipolar electrode pair placed on the biceps and triceps. The data were acquired by triggering on an amplitude threshold of a moving average of the absolute value of the transient waveforms. The structure inherent in the early portion of these transient bursts (roughly the first 100 ms) suggested a promising means of MES classification. Hudgins developed a control scheme based upon a set of simple time domain statistics and a multilayer perceptron artificial neural network classifier, capable of classifying four types of upper limb motion from the MES acquired from the biceps and triceps. This control scheme demonstrated greater discriminant ability than any other at the time, and allowed a user to evoke control using muscular contractions that resemble those normally used to produce motion in an intact limb. This system has been implemented as an embedded controller [5], and is currently undergoing clinical trials.

Although the accuracy of Hudgins' controller is good, (roughly 10% error, averaged over a set of ten subjects) there is an obvious motivation to reduce the error as much as possible. This would enhance the usability of the system as perceived by the user, and allow greater dexterity of control. A number of approaches have appeared in the literature that have used the transient signal as prescribed by Hudgins, seeking to improve the accuracy of the approach using dynamic artificial neural networks [6], genetic algorithms [7], fuzzy logic classifiers [8], and self-organizing neural networks [9]. Absent from this work however, was a direct comparison with Hudgins' method, and none has suggested a clearly superior method.

B. Signal Representation

Instead of focusing upon the classifier, the authors have demonstrated in previous work that the classification performance is more profoundly affected by the choice of feature set [10]. Specifically, a wavelet-based approach is described that, in direct comparison to Hudgins' time domain approach, exhibits superior performance. The performance of Hudgins' time domain feature set (TD), and those based upon the short-time Fourier transform (STFT), the wavelet transform (WT), and the wavelet packet transform (WPT) were compared using a new data set. This work is briefly described here to provide the context of the current investigation.

A roster of 16 healthy subjects participated in the study. Four classes of myoelectric signal patterns were collected, corresponding to flexion/extension of the elbow, and pronation/supination of the forearm. The data were acquired from two channels, located at the biceps and triceps, each pattern consisting of two channels of 256 points, sampled at 1000 Hz. The data were divided into a training set (100 patterns) and a test set (150 patterns).

The fundamental difference between the STFT, the WT, and the WPT is in the manner in which they partition the time-frequency plane. The STFT has a *fixed* tiling; once specified, each cell has an identical aspect ratio. The tiling of the WT is *variable*—the aspect ratio of the cells varies such that the frequency resolution is proportional to the center frequency. This tiling has been shown to be more appropriate for many physical signals, but the partition is nonetheless still fixed. The WPT provides an *adaptive* tiling—an overcomplete set of tilings are provided as alternatives, and the best for a given application is selected.

Each of the STFT, WT, and WPT implementations were empirically optimized to yield the best possible classification performance from the ensemble of 16 normally limbed subjects. For the STFT, it was determined that (from a number of taper windows) a Hamming window of length 64 points with an overlap of 50% gave the best performance. When using the WT, a Coiflet mother wavelet (of order four) yielded better accuracy than a host of other wavelet families of varying order [11]. The WPT experienced the best performance when using a Symmlet mother wavelet (of order five). A number of methods were considered as candidates to determine the best tiling of the WPT. The most common approach to specifying the WPT tiling is by selecting that which minimizes the reconstruction error, using an entropy cost function [12]. This may be considered optimal for signal compression, but may be inappropriate for signal classification. A modified form of this algorithm has been proposed that seeks to maximize the discriminant ability of the WPT by using a class separability cost function [13]. It is established in [10] that this discriminant cost function does indeed produce the best classification performance. A detailed description of the signal representation is given in an Appendix to this paper.

From each subject, the TD, STFT, WT, and WPT feature sets were computed. Subsequently, each feature set was subject to dimensionality reduction using principal components analysis (PCA), so as not to overwhelm the classifier with high-dimensional data. It is shown in [10] that the application of PCA is critical to the success of the time-frequency-based feature sets, and that PCA is clearly superior to other forms of dimensionality reduction. Although the classification performance is not sensitive to the dimensionality of the PCA-reduced feature set, it was demonstrated that at least five PCA features are needed, and more than thirty unnecessarily burdens the classifier. Twenty PCA coefficients are used in the analyses described here.

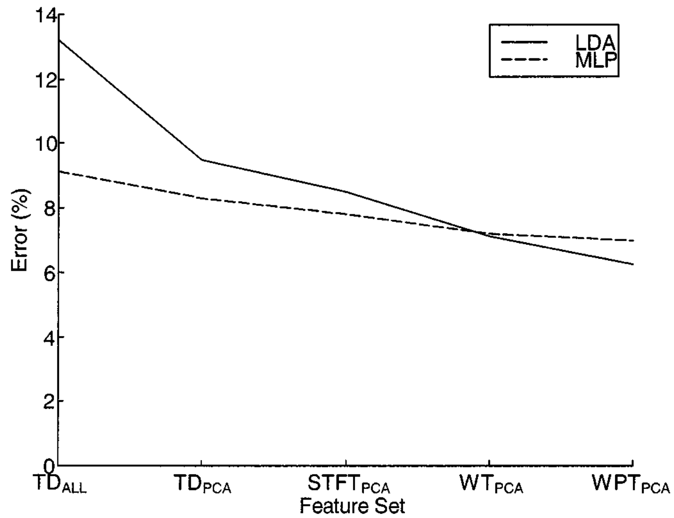


Fig. 1. The test set classification accuracy of the four-class problem demonstrated in [10]. The results of two classifiers are shown: a linear discriminant analysis (LDA) and a multiplayer perceptron (MLP).

Fig. 1 depicts the test set classification error, averaged over the ensemble of 16 subjects.

In this figure, the subscript (ALL) indicates that the entire TD feature set was used (as done by Hudgins). The subscript PCA indicates that PCA was used to reduce the feature set. The figure indicates that the performance improves in the progression $\mathrm{TD} \rightarrow \mathrm{STFT} \rightarrow \mathrm{WT} \rightarrow \mathrm{WPT}$, indicating the relative efficacy of the feature sets. Another important observation is that the LDA classifier performs as well as or better than that MLP classifier for the time-frequency-based features sets. This, presumably, is due to the fact that the PCA dimensionality reduction has the effect of linearizing the discrimination task of the classifier.

These results are encouraging, in that a more powerful feature set has been realized in the form of the PCA-reduced WPT. This investigation seeks to extend this promising approach in the following ways: to consider the benefit gained by using more channels of MES activity, to compare the performance of transient versus steady-state data, and to demonstrate the feasibility of continuous myoelectric control of a multifunction prosthesis.

III. Methodology

Two experiments are described in this work to elucidate these factors. The first experiment compares the performance of a two-channel configuration to that of a four-channel system, and compares the performance when using transient and steady-state data. It is shown that exceptionally good performance is achieved using four-channel steady-state data with as many as six classes. The second experiment, bolstered by these results, demonstrates the capabilities of the four-channel system acting on a continuous stream of data.

In the first experiment, data were acquired[1] from 11 normally limbed individuals, recording four channels of MES from elec-

[1]These data were acquired during the course of an M.Sc.E. thesis by Sentiono Leowinata.

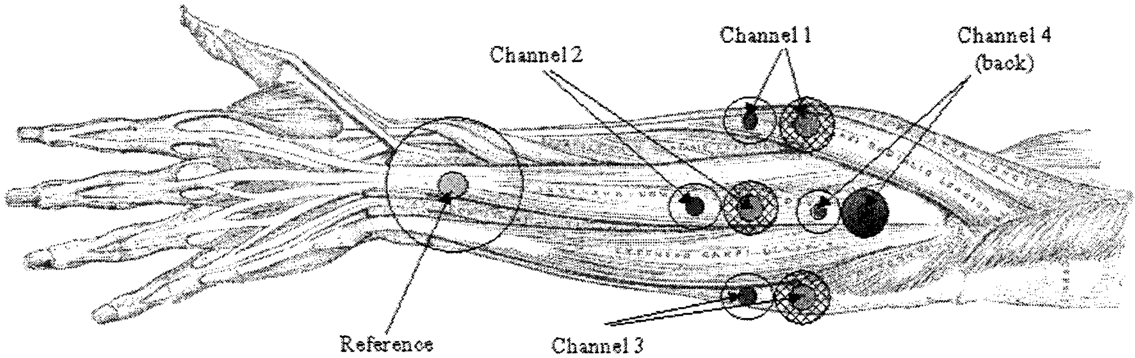


Fig. 2. The electrode placement used in the four-channel MES acquisition. Four bipolar electrode pairs (Red Dot—3M Corp.) were used with a reference at the wrist. Although difficult to show on the figure, the top and bottom electrode pairs are at the same level of the forearm.

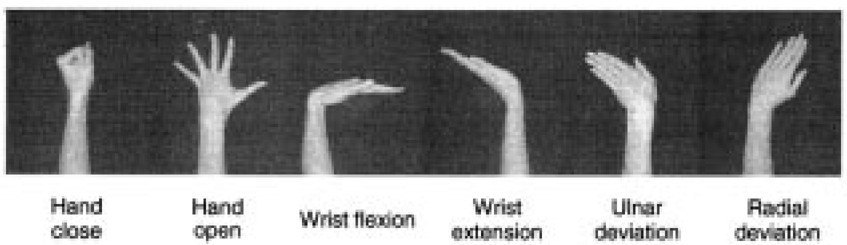


Fig. 3. The six classes of motion used in the four channel experiments.

trodes placed on the medial side, top, lateral side and bottom of the forearm, as depicted in Fig. 2.

Each subject generated six different classes of motion: hand close/open, flexion/extension of the wrist, and ulnar/radial deviation of the wrist,[2] as shown in Fig. 3.

Each subject produced two sets of data: one comprising transient bursts, and another consisting of steady-state signals. Because this is a prosthetic control problem, the contraction levels are arbitrary as long as they are reasonably consistent, and comfortable enough for the subject to reproduce in daily use without fatigue. Each bipolar channel was acquired used Ag-AgCl electrodes spaced at 2 cm. Each record was 256 ms in duration (256 points, sampled at 1000 Hz, prefiltered between 10 and 500 Hz). In each dataset, 80 patterns were generated in each class, resulting in a total of 480 patterns. These data were evenly divided into training and test sets of 240 patterns, and then subject to classification by the time-frequency methods described in the previous section.

In the second experiment, steady-state MES data were acquired using the same four-channel configuration as described in Experiment 1. The subject was asked to produce constant-force contractions from each of the six classes for 5-s intervals, and then repeat the pattern, generating 60 s of data. The first 30-s set of six classes was used as a training dataset, and the second 30-s set was used as a test dataset. These data were divided into discrete 256-sample records, and presented to the system in the same manner as in Experiment 1 to train the system. After training, the continuous stream of steady-state data was subject to classification using a sliding window that progressed across the entire record.

[2]The rationale for recording from the forearm is that the underlying musculature directly contributes to each of the six types pf contraction; it would be difficult to contrive six distinct classes of motion directly actuated by the biceps/triceps pair.

IV. Experiment 1—Results

Each feature set was again used in the analysis, as well as another proposed by Leowinata, which comprises the normalized auto-correlation and cross-correlation functions of the channels [14]. Here, this feature set will be denoted AC. The following results depict the classification performance of the test set of data, averaged across the 11 subjects. Each feature set has been subject to PCA dimensionality reduction, and classified using the LDA (the MLP, having shown no advantage in Fig. 1, has been omitted in this and subsequent analyses).

Consider first a four-class problem (using wrist flexion/extension and hand open/close). Fig. 4 shows the results for each feature set when using transient data for two channels (top and bottom electrodes) and for four channels. It is clear that four MES channels offers improved accuracy, as compared to two channels. It is also evident that the WT and WPT feature sets offer the best performance, corroborating the results in the previous experiment. Note that the AC feature performs poorly, relative to the others. This is consistent throughout each comparative analysis.

Now consider the same set of four classes elicited as steady-state contractions. Fig. 5 shows the same improvement when progressing from two to four channels, and the relative efficacy of the WPT feature set.

Having concluded that four channels offer a distinct advantage to two channels, we will use the four-channel configuration for the remainder of the analysis. A direct comparison of the accuracy when using four-channel transient and steady-state data is shown in Fig. 6. The results indicate that the steady-state data

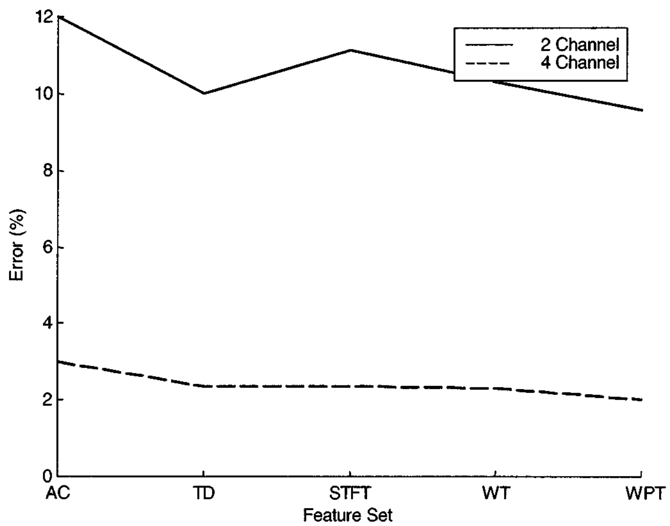


Fig. 4. The classification accuracy using four classes of transient MES data. All feature sets have been subject to PCA and hence, the subscript PCA has been dropped.

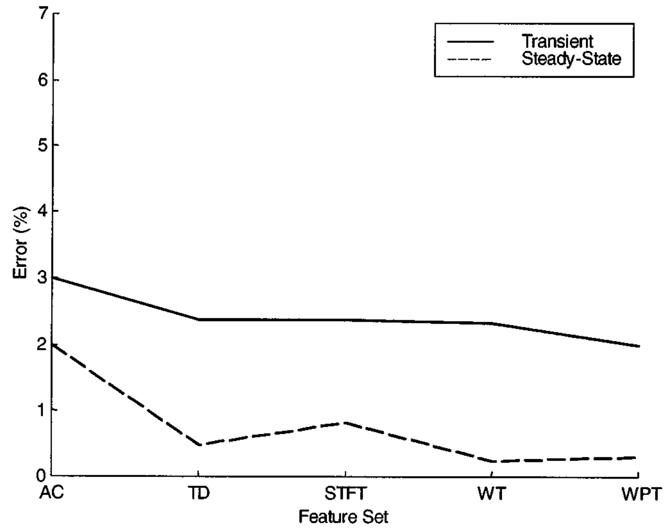


Fig. 6. Four classes, four channels: the relative performance of transient versus steady-state data.

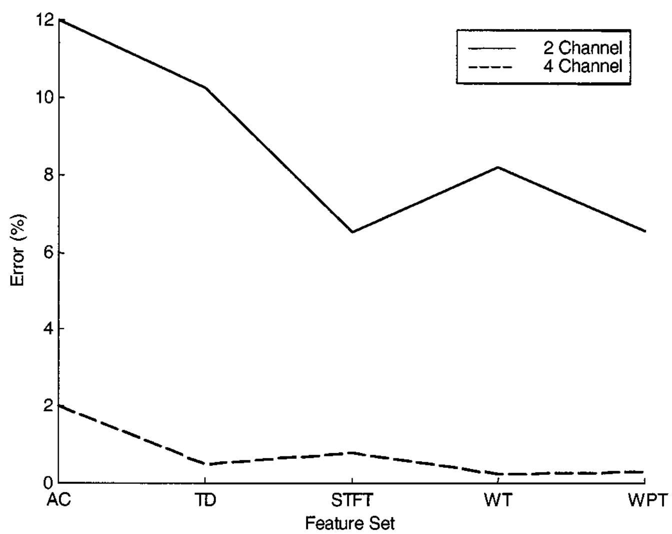


Fig. 5. The classification accuracy using four classes of steady-state MES data.

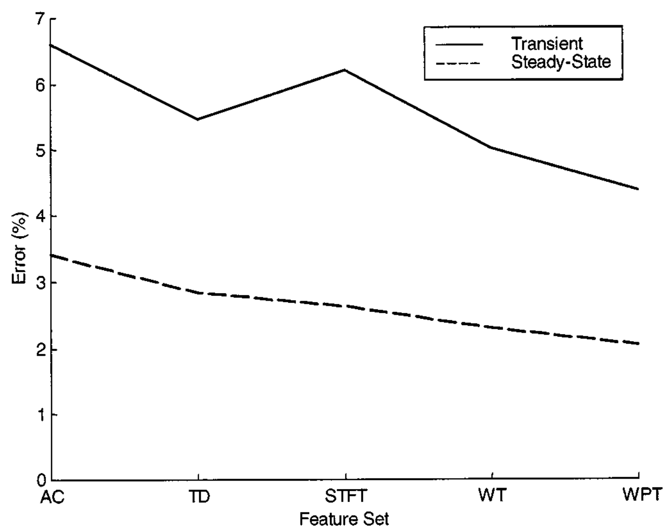


Fig. 7. Six classes, four channels: the relative performance of transient versus steady-state data.

contains greater discriminating information than the transient data. A justification for these results is described in the next section. One can observe that the error rate is approaching zero for the WT and WPT feature sets for the steady-state data; indeed, a majority of the subjects achieve an error rate of zero. This degree of accuracy is unprecedented for this problem.

The exceptional accuracy with four channels of steady-state data suggests one might make the problem a bit more difficult. It may be desired to classify a greater number of classes, to provide more functionality to a prosthetic system. Consider now a six-class problem, including all wrist and hand motions. The classification results for the transient and steady-state data are shown in Fig. 7.

Again, the steady-state data exhibits distinct superiority to the transient data, and the WPT feature set demonstrates the best performance. Although the performance is not as good as in the four-class problem, the WPT feature set yields an error rate of two percent, which is still exceptional given the difficulty of the problem.

Another advantage that the steady-state data has over the transient data pertains to the effect of record length. The analyses described previously used a record length of 256 points (256 ms). This may be considered the maximum record length, dictated by the allowable response time of the classifier (it is generally agreed that 300 ms is the longest acceptable delay in a prosthetic control system). The classification performance degrades rapidly as the record length of the transient data is decreased from 256 to 128 to 64 to 32 samples, as shown in Fig. 8. The WPT feature set has been chosen to demonstrate the effect of record length; the other feature sets experience essentially the same effect with record length.

When using the steady-state data however, this degradation is not as profound, as shown the figure. Indeed, the performance suffers very little when reducing to 128 ms. This suggests that

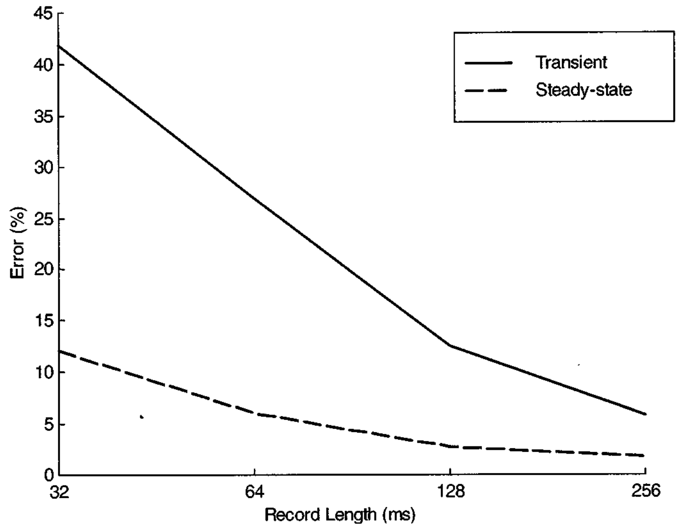


Fig. 8. Six classes, four channels: the effect of record length upon classification performance, using a WPT feature set upon transient and steady-state MES data.

shorter records of steady-state data may be used, if a faster system response is desired.[3]

V. Experiment I—Discussion

The following observations can be made from the preceding analyses.

- Four channels of MES are clearly preferable to two. This suggests an investigation into the benefits that may be gained by further increasing the number of channels. Whereas the number of channels was practically limited in the past by the sheer bulk of the instrumentation and the difficulty in maintaining good contact, advances in electrode array miniaturization, fabrication, and interfacing techniques have made multichannel systems more feasible.
- For the same set of subjects, the steady-state data was classified more accurately than the transient data. As well, the steady-state data suffers less degradation with shorter record lengths.
- The wavelet and wavelet packet-based feature sets outperform the others in every scenario.
- A four-channel steady-state system, using a WPT feature set performs exceptionally well, yielding 0.5% error when discriminating four classes, and 2% error with six classes.

The basis for the improvement when adding channels is obvious: the MES activity recorded from the side electrodes (channels 1 and 3) contribute additional information about the articulated contraction. The superiority of the steady-state data to the transient data however, was somewhat unexpected. In [4], Hudgins *et al.* demonstrated that a single MES channel (with widely spaced electrodes placed on the biceps and triceps) exhibits significant structure in the first 100 ms preceding initiation of a contraction. It is uncertain as to whether this structure is due to electrophysiological determinism, or due other phenomena such as skin stretch potentials, or the motion of the electrodes relative to the underlying musculature. Regardless, this structure was presumed to contribute significantly to discrimination amongst contraction types. Kuruganti *et al.* [15] verified that the accuracy of Hudgins' method could be improved by using two bipolar channels with a localized pickup region over each of the biceps and triceps, instead of a single channel with a large pickup region. In a multiple channel configuration, it appears that this fine structure plays a lesser role to the gross activity of the constituent muscle activity. The more localized bipolar electrodes in a multichannel configuration appear to degrade the incidence of fine structure in the waveforms. These localized channels do, however, communicate more information about the relative activity of spatially separated muscle groups, which is conveyed by the gross activity of the steady-state signal.

At face value, these analyses describe a method that discriminates the surface MES with greater accuracy than any previous work. Another, perhaps more important implication of these results, is that one may abandon the need to detect and frame transient bursts of MES activity. Instead of requiring an individual to elicit a contraction from rest (a rather awkward imposition when performing a sequence of tasks), classification may be performed on a continuous stream of steady-state data, as one switches from one contraction type to another. This *continuous classifier* could produce classification results as often as the feature extraction processing delay would allow (a factor of the feature set, the record length, and the processing power).

VI. Experiment II—Results

A second experiment was carried out to provide a demonstration of the capabilities of such a classifier. The system was trained using disjoint 256-sample frames of data extracted from the continuous training data. For 256-sample records, each 5-s interval (for each class) yields 19 patterns, resulting in $19 \times 6 = 114$ patterns in the training dataset. A set of test patterns was generated in the same way. These discrete records were classified in the same manner as in Experiment 1, to provide an indication as to what type of performance might be expected of the continuous classifier. Classification results when using these data were very typical of those obtained in Experiment 1: the test set error for the six-class problem using WT and WPT features sets was 4.3% and 3.5%, respectively.

Having determined the PCA, LDA and WPT tiling coefficients from the training session (storing them for later reference), a continuous classifier could then be constructed. Instead of acting on an ensemble of disjoint records of MES data, the continuous classifier produces a series of decisions using a sliding window of MES activity. This illustrated in Fig. 9, which depicts one (of the four) MES channels, and the sliding window of activity used to generate classification decisions.

In addition to the type of feature set to be used, the parameters which affect the classifier's performance include the record length (which will be denoted N), and the window increment (denoted M). We have seen from the results of Experiment 1 that N significantly affects the classification rate, so it is likely

[3]Of course, the system is also subject to the response time and damping factor of the prosthesis. A faster control system response would offer no benefit to a slower prosthetic system.

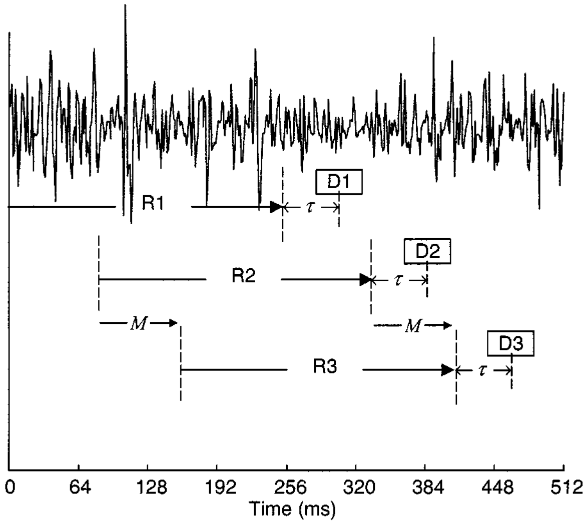


Fig. 9. The sliding window used in continuous classification. Here, a sequence of 256 ms records (R1, R2, and R3) are used to make decisions at times D1, D2, and D3. In this example, the window index increments by M; the processing delay τ must be less than M.

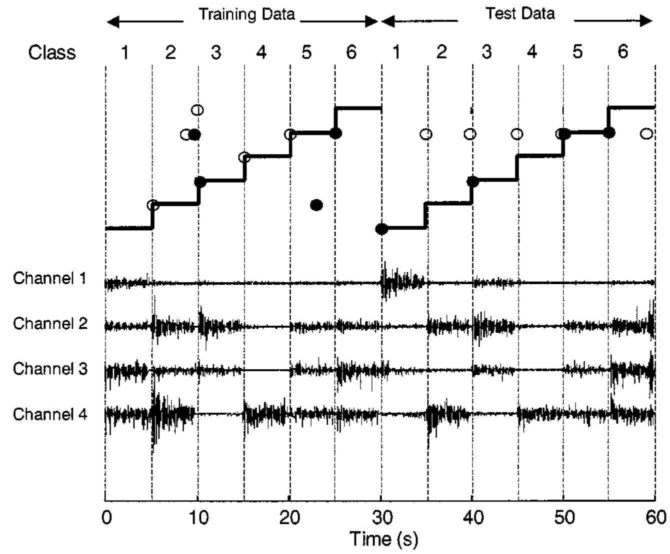


Fig. 10. A session of continuous classification using a WT feature set. In this analysis, $N = 256$ and $M = 128$. One minute of four-channel MES data is shown, with the class targets (the staircase in the upper portion) and the classification errors (the open circles superimposed on the staircase). The errors made using a majority vote of decisions made in the past 500 ms are shown as filled circles.

that $N = 256$ samples is preferable here. It is not clear, however, what is the best choice for M. The window increment affects the rate at which class decisions will be made, and consequently, it determines the real-time constraint of the system (the processor must compute the feature set and generate a decision before the next batch of data arrives). A small value of M produces a dense stream of class decisions with respect to time, which may improve the response time of the classifier and, by utilizing this redundant information, improve the classification accuracy.

A sample session of continuous classification is shown in Fig. 10. Here, a WT feature set[4] was used with $N = 256$ and $M = 128$, yielding 466 decisions over the 60-s interval of data. The first 30 s of data represent the training data, and the latter 30 s, the test data. The class targets are represented by the staircase, with the level indicating the class number (and labeled at the top of the graph). The open circles superimposed on the targets represent errors made by the classifier, plotted at the level of the predicted class. Clearly, most of the errors are clustered about the zones of transition between classes, which are to be expected, since the muscle activity is in a state of flux.

In this analysis, the test set error is 7.8%. Recall that when training the system (using disjoint 256-sample records), the test set error when using this WT feature set was 4.3%. This is a favorable comparison, with the observation that most of the errors appear to occur at the transition regions in Fig. 10. When using the WPT feature set, the error is 6.8% as compared to 3.5% in the batch training session.

The dependency of the accuracy at different window increments was investigated. The classification accuracy at values of $M = 32$, 64, 128, and 256 samples was found to be essentially the same. This is true for both the WT and the WPT feature set. Although at smaller values of M more errors were made, the ratio of error to decisions remains fairly constant.

[4]The WT was chosen for this example due to its computational efficiency. Although the WPT exhibits greater accuracy in Experiment 1, its advantage over the WT is marginal in most cases.

With a denser stream of class decisions however, one may combine adjacent decisions in an attempt to improve classification accuracy. A simple approach to post-processing the sequence of class decisions is to take a majority vote of recent decisions. A majority vote was performed using the current decision and each decision made in the previous 500 ms. In Fig. 10, the positions where the majority vote scheme produces errors are indicated by filled circles. These errors are almost exclusively restricted to the transition regions of the activity, essentially eliminating errors in the midst of an interval of steady-state contraction. The error rate using a majority vote scheme is roughly half that of the unprocessed stream of class decisions, regardless of the feature set (WT or WPT) or window increment.

The improvement due to using a majority vote scheme may reduce to an academic exercise however, since the inertia of a prosthetic device will serve to smooth the stream of class decisions, and forgive spurious errors. The important observation to be made, however, is that the continuous classifier produces a very reliable decision stream for this six-class problem.

The remaining issue to be discussed is whether the classification scheme can meet the real-time constraints of the problem. The system must perform feature extraction (either WT or WPT), PCA, and LDA in the window increment time, M. Fortunately, in the feedforward path, each of these operations are computationally efficient. The complexity of the WT is on the order N (for the Coiflet mother wavelet) [11] and the complexity of the WPT is $N \log N$ [12]. The PCA is the product of a $N \times 1$ by an $N \times N$ matrix, and the LDA is the product of a $K \times 1$ by $K \times P$ matrix (where $K = 20$ features and $P =$ six classes).

The processing delays were empirically evaluated using a 450-MHz Pentium III-based workstation. The computation was performed in Matlab (The Mathworks, Natick, MA); the matrix multiplications were built-in functions, and the WT and WPT

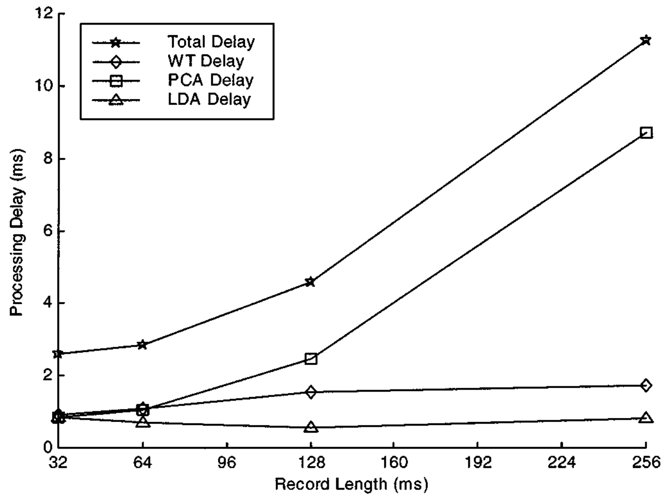


Fig. 11. The processing delays associated with the WT, PCA, and LDA stages of the system at various record lengths.

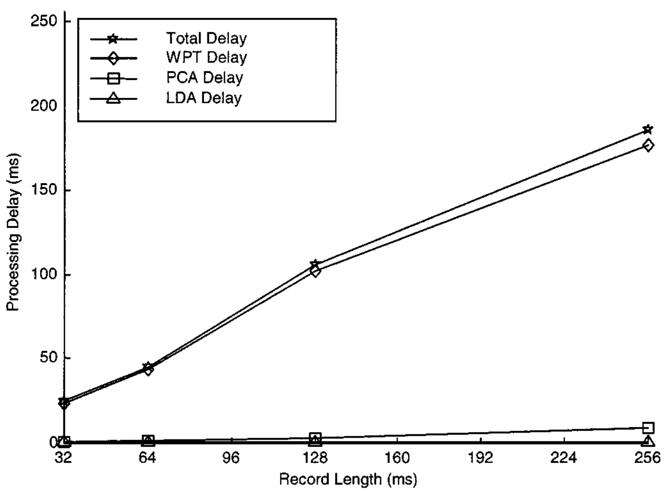


Fig. 12. The processing delays associated with the WPT PCA, and LDA stages of the system at various record lengths.

routines were compiled C code. Fig. 11 shows the processing delays using various record lengths, N, when using the WT feature set.

As the record length N increases, the bulk of the processing delay is associated with the PCA stage. The WT is extremely computationally efficient, increasing linearly with N and never exceeding 2 ms of processing time. To meet the real-time constraints imposed on the system, the delay must be less than the window increment; it is clear that this constraint is easily met for window increments greater than 12 ms. Fig. 12 shows the same delays when using a WPT feature set. The WPT demands substantially greater computation (it represents the bulk of the delay); with a record length of 256 ms, the real-time constraint will be met only if the window increment is greater than 200 ms. These delays are relative to the capacity of the chosen computing platform; presumably, an embedded system with dedicated signal processing hardware could meet or exceed these processing requirements.

VII. Experiment II—Discussion

This experiment has demonstrated the viability of a continuous classification scheme that is impressively accurate. Almost all misclassifications seem to be clustered about the regions of transition between classes. These errors may be forgiven in actual use as long as the region in which errors are registered is short compared to the desired dynamic response time of the system. That is, if one wishes to manipulate a prosthetic device and be capable of establishing control of a device (or switching control between devices) within 300 ms of the intent actuated by a contraction, the errors produced between states must be limited to this interval. This is achieved at each transition in Fig. 10, where $M = 128$ and no more than one error is encountered at any transition (even two errors would fall within 256 ms). It is likely that the inertia associated with actuation of the prosthetic device will forgive any spurious misclassifications upon transition, until the steady-state activity has been established.

Although some insight as to the capabilities of such a system has been illustrated here, there are some issues yet to be resolved that are currently under investigation.

1) The system must know when to actuate the prosthetic devices, and when to suppress actuation. With a constant stream of decisions being produced, the actuation must be gated by some means. This might be accomplished by including an additional "inactive" class in the training session, by imposing a lower threshold of MES activity, or a combination of both. The development of this strategy is as important as classification accuracy in terms of usability.
2) The steady-state data in this investigation comprises contractions that are of roughly the same intensity (although subjects were not instructed to maintain a consistent level of effort). The performance of the system when using contractions of varying intensity must be investigated. This is essential if one is to implement proportional control of the devices (where the velocity is proportional to the intensity of the muscular effort). Otherwise, only ON/OFF control (one speed) may be used.
3) The system has been shown to be very accurate in discriminating six classes of motion. Is it possible that combined motions (for example, hand close/wrist flexion) might be classified? This would enable simultaneous control of devices, which would enhance the anthropomorphism of control, offering benefits of functionality and dynamic cosmesis.
4) To what extent will additional channels of myoelectric activity improve the classification performance? Will a many-channel grid of electrodes offer the discrimination needed to resolve combined/simultaneous activities?

It should be noted that, with the motivation of developing a continuous classifier, the time-frequency analysis tools (the STFT, WT, and WPT) developed for the transient signals were directly transferred to the analysis of the steady-state signals. It is clear from the results presented that the classification performance is superior when using steady-state data, although one would expect that these nonstationary signal analysis tools

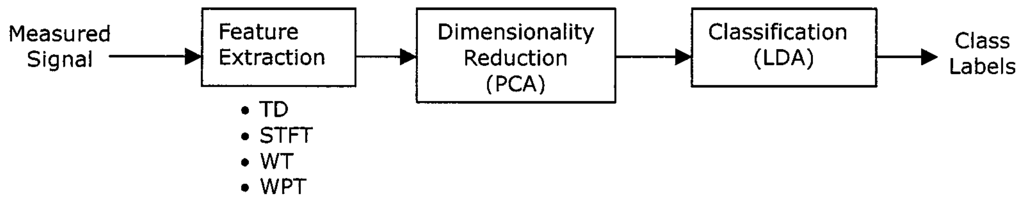


Fig. 13. The stages comprising the classification problem, and the methods subject to investigation.

are not required for these data, which are essentially stationary. Indeed, reducing the segmented STFT to a single-window FFT produces essentially the same classification results as the STFT with 64 sample windows, overlapping by 50%. One cannot perform a similar comparison with the WT and WPT, as they are inherently time-coherent. It still remains, however, that the WT and WPT outperform the other feature sets when using the steady-state data. This is not due to their ability to capture temporal information, which is absent in the steady-state data, but to their ability to model the elemental basis of the myoelectric signal. The wavelet functions themselves closely resemble the motor unit action potentials that constitute the gross myoelectric signal. To this end, the optimum wavelet and wavelet packet parameters for classifying steady-state signals are currently under investigation.

VIII. Conclusion

A wavelet-based approach to MES classification has been described that exhibits very good accuracy when using two channels of MES activity, and even better accuracy when using four channels. Steady-state data has been shown to outperform transient data for the same ensemble of 11 subjects. The WPT/PCA feature set, with four channels of steady-state data allows four classes of motion to be classified with a average of 0.5% error, and six classes with 2% error.

Given the efficacy of the WT and WPT-based feature sets, a continuous classification scheme has been described. The continuous classifier represents a promising new mode of controlling prosthetic devices. It represents a more natural and efficient means of myoelectric control than one based on discrete, transient bursts of activity, promising to reduce the mental burden of a user, and the dexterity of control.

Appendix
A Detailed Discussion of Signal Representation

The problem of signal classification may be thought of as a multistage process as shown in Fig. 13.

The measured signal (each channel of myoelectric signal) is subject to feature extraction, in this case producing a feature set consisting of TD, STFT, WT, or WPT coefficients. This feature set is then subject to dimensionality reduction using PCA, yielding a reduced feature set of five coefficients from each channel. The PCA coefficients from each channels are then combined to produce an aggregate feature set; for six channels, the LDA classifier would be presented with 30 features. A more detailed discussion of these stages follows.

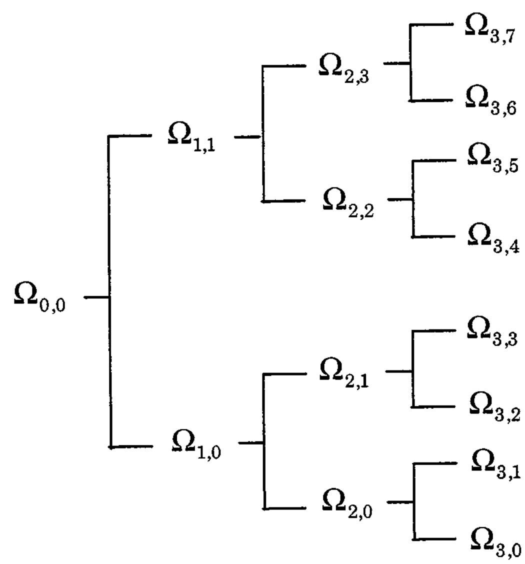


Fig. 14. A decomposition of $\Omega_{0,0}$ into binary tree-structured subspaces using the WPT (with $J = 3$).

A. Feature Set Specification

The reader is referred to [4] for a detailed description of the TD feature set. The parameters of the STFT were empirically optimized to yield the best classification rate on an ensemble of transient MES data [16]. These parameters include the taper window, the size of the taper window, and the degree of overlap. It was determined that, for $N = 256$ sample records of transient MES, a 64 sample Hamming window with an overlap of 50% gave the best performance. The sole parameter of the WT is the choice of mother wavelet. Various orders of the following wavelet families were considered: Symmlet, Coiflet, Daubechies, Haar, Vaidyanathan, Beylkin, and Biorthogonal.

For the WT, the Coiflet family of order four gave the lowest classification error. For the WPT, Symmlet mother wavelet of order five was best.

The WPT may be thought of as a tree of subspaces. The root node of the tree (the original signal space) is $\Omega_{0,0}$. A subspace $\Omega_{j,k}$ is decomposed into two orthogonal subspaces $\Omega_{j,k} \rightarrow \Omega_{j+1,2k}$ and $\Omega_{j,k} \rightarrow \Omega_{j+1,2k+1}$. Here j denotes scale, as before, and k indicates the subband index within the scale.[5] Each subspace $\Omega_{j,k}$ is spanned by 2^{n_0-j} basis vectors $\{w_{j,k,n}\}_{n=0}^{2^{n_0-j}-1}$, where $n_0 = \log_2 N$. These basis vectors are the wavelet packet basis functions. A decomposition to scale $J = 3$ is shown in Fig. 14.

[5]The WT has only two subbands per scale, high and low, with $k = 0, 1$.

The set of subspaces in the WPT binary tree is a redundant set, that is, the transform yields a binary tree of coefficients comprising 2^N possible orthonormal bases, where N is the record length. The power of the WPT is that a "best basis" can be chosen for a specific task, if it can be properly identified from the ensemble of possible candidates. To determine the *best basis*, it is necessary to evaluate and compare the efficacy of many bases. To this end, a *cost function* must be chosen to represent the goal of the application. The *best-basis selection algorithm* has its origins in signal compression [12], [17], and the cost functions associated with compression all entail some form of entropy measure. This may be considered optimal for signal compression, but may be inappropriate for signal classification.

A modified form of this algorithm was proposed by N. Saito in his Ph.D. dissertation [13]. He termed the algorithm the *local discriminant basis* (LDB) algorithm, implying that an orthonormal basis is selected from the binary wavelet packet tree which most discriminates data from a given set of classes. The measure of class separability is conveyed by the discriminant measure D. An n-feature discriminant measure can be defined as $D(\mathbf{p}, \mathbf{q})$, where $\mathbf{p} = \{p_i\}_{i=1}^n$, $\mathbf{q} = \{q_i\}_{i=1}^n$ are measures used to represent the n features from two different classes. Of several discriminant measures investigated in [16], that which was found to give the best performance for MES classification was symmetric relative entropy [18]

$$D(\mathbf{p}, \mathbf{q}) \doteq \sum_{i=1}^{n} p_i \log \frac{p_i}{q_i} + \sum_{i=1}^{n} q_i \log \frac{q_i}{p_i}. \tag{1}$$

In order to optimize the representation with respect to the time-frequency localization characteristics of the wavelet packet basis, the input parameters to D are the *time-frequency energy maps* of each class.

Definition: Let $\{\mathbf{x}_i^{(c)}\}_{i=1}^{N_c}$ be a set of training signals belonging to class c, where N_c is the number of patterns in class c. The *time-frequency energy map* of class c is a table of positive real values indexed by (j, k, n)

$$\Gamma_c(j, k, n) \doteq \frac{\sum_{i=1}^{N_c} \left(\mathbf{w}_{j,k,n}^T \mathbf{x}_i^{(c)}\right)^2}{\sum_{i=1}^{N_c} \left\|\mathbf{x}_i^{(c)}\right\|^2} \tag{2}$$

for $j = 0, \ldots, J$, $k = 0, \ldots, 2^j - 1$, $n = 0, \ldots, 2^{n_0 - j} - 1$. That is, Γ_c is computed by accumulating the squares of the transform coefficients for each entry in the binary packet tree (j, k, n), and normalizing by the total energy of the signal belonging to class c.

Since the algorithm must choose the best set of subspaces from the binary packet tree, the response from individual temporal locations from within a subspace must be summed. For K classes, the overall discriminant measure for the subspace $\Omega_{j,k}$ is thus

$$D\left(\{\Gamma_c(j, k, \bullet)\}_{c=1}^K\right) \doteq \sum_{n=0}^{2^{n_0 - j} - 1} D(\Gamma_1(j, k, n), \ldots, \Gamma_K(j, k, n)) \tag{3}$$

We are now ready to develop a specification of the LDB algorithm. Let $B_{j,k}$ denote a set of basis vectors belonging to the subspace $\Omega_{j,k}$, arranged in matrix form

$$B_{j,k} = \lfloor \mathbf{w}_{j,k,0}, \mathbf{w}_{j,k,1}, \ldots, \mathbf{w}_{j,k,2^{n_0 - j} - 1} \rfloor^T. \tag{4}$$

Let $A_{j,k}$ represent the LDB for the training set restricted to the span of $B_{j,k}$, and let $\Delta_{j,k}$ be a work array containing the discriminant measure of the node (j, k).

The LDB Algorithm [13]: Given a training dataset consisting of K classes of signals $\{\{\mathbf{x}_i^{(c)}\}_{i=1}^{N_c}\}_{c=1}^K$,

- Step 0) Choose a time-frequency decomposition method. That is, specify a WPT, the depth of decomposition J, and the discriminant measure D.
- Step 1) Construct the time-frequency energy maps Γ_c for $c = 1, \ldots, K$.
- Step 2) Begin at level J: set $A_{j,k} = B_{j,k}$ and $\Delta_{j,k} = D(\{\Gamma_c(J, k, \bullet)\}_{c=1}^K)$ for $k = 0, \ldots, 2^J - 1$.
- Step 3) Determine the best subspace $A_{j,k}$ for $j = J - 1, \ldots, 0$, $k = 0, \ldots, 2^j - 1$ by the following rule:

  Set $\Delta_{j,k} = D\left(\{\Gamma_c(j, k, \bullet)\}_{c=1}^K\right)$
  if $\Delta_{j,k} \geq \Delta_{j+1,2k} + \Delta_{j+1,2k+1}$,
  then $A_{j,k} = B_{j,k}$,
  else $A_{j,k} = A_{j+1,2k} + A_{j+1,2k+1}$ and
  set $\Delta_{j,k} = \Delta_{j+1,2k} + \Delta_{j+1,2k+1}$.

- Step 4) Order the N basis functions in the LDB by their power of discrimination.
- Step 5) Use the L $(\ll N)$ most discriminating basis functions in the LDB for classifier features.

When Step 3 has been completed, we are left with $A_{0,0}$, which is the LDB restricted to the span of $B_{0,0} \equiv \Re^N$: a complete orthogonal basis. The chosen LDB consists of a set of disjoint subspaces, which form a cover of the time-frequency plane. Each subspace $\Omega_{j,k}$ contains $2^{n_0 - j}$ basis vectors. The total number of basis functions is always N, where $N = 2^{n_0}$ is the length of each signal $\mathbf{x}_i^{(c)}$. The pruning algorithm is fast $(O(N))$ since the measure D has been chosen to be additive.

Saito's LDB algorithm, in steps 4 and 5, ranks the features and chooses a subset $L \ll N$ of these determined to be most discriminant. This form of dimensionality reduction can be considered a form of *feature selection*. This approach was found to perform poorly with MES data, due to the high variance of the signal and consequently, the WPT features. This algorithm was modified in [16] by replacing steps 4 and 5 with PCA dimensionality reduction.

PCA involves projecting the features onto their eigenvectors and retaining those which correspond to the largest eigenvalues. PCA is not designed for class discrimination, rather, it is optimized for signal compression. The inherent assumption in its use for dimensionality reduction in the context of classification is that the signal variance accounts for a significant portion of the discriminant information amongst classes. It is well known that, for many physical signals, this is true, and certainly seems to be the case for the MES. Admittedly, higher-order discriminant information may be lost when using PCA, but for high-dimensional feature sets (such as those considered here), this higher-order information seems to contribute little. Other projection-based methods of dimensionality reduction are available that are tailored for classification, such as projection pursuit [19], but these methods have proven unsatisfactory because they are exploratory methods, tend to be sensitive to outliers, and are computationally intense.

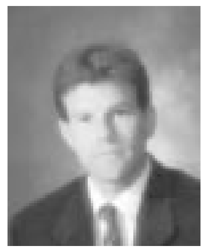
The efficacy of the method described here is the result of the ability of PCA to gather the essential discriminant information from the highly stochastic, high-dimensional feature sets generated by the STFT, WT, and WPT applied to the MES.

ACKNOWLEDGMENT

The authors would like to acknowledge the contribution of S. Leowinata for his efforts in compiling the four-channel MES database.

REFERENCES

[1] D. Dorcas and R. N. Scott, "A three state myoelectric control," *Med. Biol. Eng.*, vol. 4, pp. 367–372, 1966.
[2] D. A. Childress, "A myoelectric three state controller using rate sensitivity," in *Proc. 8th ICMBE Conf.*, vol. S4-5, Chicago, IL, 1969.
[3] L. Vodovnik, J. Kreifeldt, R. Caldwell, L. Green, E. Silgalis, and P. Craig, "Some Topics on Myoelectric Control of Orthotic/Prosthetic Systems," Case Western Reserve Univ., Cleveland, OH, Rep. EDC 4-67-17, 1967.
[4] B. Hudgins, P. A. Parker, and R. N. Scott, "A new strategy for multifunction myoelectric control," *IEEE Trans. Biomed. Eng.*, vol. 40, pp. 82–94, Jan. 1993.
[5] B. Hudgins, K. Englehart, P. A. Parker, and R. N. Scott, "A microprocessor-based multifunction myoelectric control system," presented at the 23rd Canadian Medical and Biological Engineering Society Conf., Toronto, ON, May 1997.
[6] K. Farry, I. D. Walker, and R. G. Baraniuk, "Myoelectric teleoperation of a complex robotic hand," *IEEE Trans. Robot. Automat.*, vol. 12, pp. 775–788, Oct. 1996.
[7] K. Englehart, B. Hudgins, M. Stevenson, and P. A. Parker, "Classification of transient myoelectric signals using a dynamic feedforward neural network," presented at the World Congress Neural Networks, Washington, DC, July 1995.
[8] S. Leowinata, B. Hudgins, and P. A. Parker, "A multifunction myoelectric control strategy using an array of electrodes," presented at the 16th Annu. Congress International Society Electrophysiology and Kinesiology, Montreal, P.Q., Canada, 1998.
[9] P. J. Gallant, "An Approach to Myoelectric Control Using a Self-Organizing Neural Network for Feature Extraction," master's thesis, Queens Univ., Kingston, ON, Canada, 1993.
[10] K. Englehart, B. Hudgins, P. A. Parker, and M. Stevenson, "Classification of the myoelectric signal using time-frequency-based representations," *Medical Engineering and Physics (Special Issue: Intelligent Data Analysis in Electromyography and Electroneurography)*, vol. 21, pp. 431–438, 1999.
[11] I. Daubechies, "Ten lectures on wavelets," in *CBMS-NSF Regional Conference Series in Applied Mathematics*. Philadelphia, PA: SIAM, 1992, vol. 61.
[12] R. R. Coifman and M. V. Wickerhauser, "Entropy-based algorithms for best basis selection," *IEEE Trans. Inform. Theory*, vol. 38, pp. 713–719, Feb. 1992.
[13] N. Saito and R. R. Coifman, "Local discriminant bases and their applications," *J. Math. Imag. Vis.*, vol. 5, no. 4, pp. 337–358, 1995.
[14] S. Leowinata, B. Hudgins, and P. A. Parker, "A multifunction myoelectric control strategy using an array of electrodes," presented at the 16th Annu. Congress International Society Electrophysiology and Kinesiology, Montreal, PQ, Canada, 1998.
[15] U. Kuruganti, B. Hudgins, and R. N. Scott, "Two-channel enhancement of a multifunction control scheme," *IEEE Trans. Biomed. Eng.*, vol. 42, Jan. 1995.
[16] K. Englehart, "Signal Representation for Classification of the Transient Myoelectric Signal," Ph.D. dissertation, Univ. New Brunswick, Fredericton, NB, Canada, 1998.
[17] M. V. Wickerhauser, *Adapted Wavelet Analysis from Theory to Software*. Wellesley, MA: AK Peters, Ltd., 1994.
[18] S. Kullback and R. A. Leibler, "On information and sufficiency," *Ann. Math. Stat.*, vol. 22, pp. 79–86, 1951.
[19] P. J. Huber, "Projection pursuit (with discussion)," *Ann. Stat.*, vol. 13, no. 2, pp. 435–475, 1985.



Kevin Englehart (S'90–M'99) received the B.Sc. degree in electrical engineering and the M.Sc. and Ph.D. degrees from the University of New Brunswick (UNB), Fredericton, NB, Canada, in 1989, 1992, and 1998, respectively.

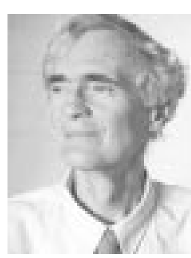
He is currently an Assistant Professor of Electrical and Computer Engineering, and a Research Consultant to the Institute of Biomedical Engineering at UNB. His research interests include neuromuscular modeling and biological signal processing using adaptive systems, pattern recognition, and time-frequency analysis.

Dr. Englehart is a Registered Professional Engineer, and a member of the IEEE Engineering in Medicine and Biology Society, the IEEE Computer Society, and the Canadian Medical and Biological Engineering Society.

Bernard Hudgins (M'97) received the Ph.D. degree from the University of New Brunswick (UNB), Fredericton, NB, Canada, in 1991.

He is currently a Senior Research Associate with the Institute of Biomedical Engineering at UNB. His Primary research interests are in the area of myoelectric signal processing for the control of artificial limbs.

Dr. Hudgins was the recipient of a Whitaker Foundation Investigator Award and recently spent two years on a NATO workgroup assessing alternative technologies for cockpit applications. He has been the Region 7 (Canada) representative on the IEEE EMBS Advisory Committee for the past two years.



Philip A. Parker (S'70–M'73–SM'86) received the B.Sc. degree in electrical engineering from the University of New Brunswick (UNB), Fredericton, NB, Canada, in 1964, the M.Sc. degree from the University of St. Andrews, St. Andrews, U.K., in 1966, and the Ph.D. degree from the UNB in 1975.

In 1996, he joined the National Research Council of Canada as a Communications Officer and the following year he joined the Institute of Biomedical Engineering, UNB, as a Research Associate. In 1976, he was appointed to the Department of Electrical Engineering, UNB, and currently holds the rank of Professor in that department. He is also a Research Consultant to the Institute of Biomedical Engineering, UNB. His research interests are primarily in the area of biological signal processing.