ELSEVIER

Medical Engineering & Physics 21 (1999) 431–438

Medical Engineering & Physics

www.elsevier.com/locate/medengphy

Classification of the myoelectric signal using time-frequency based representations

K. Englehart [a,*], B. Hudgins [a], P.A. Parker [b], M. Stevenson [b]

[a] *Institute of Biomedical Engineering, University of New Brunswick, Fredericton, N.B., Canada*
[b] *Department of Electrical and Computer Engineering, University of New Brunswick, Fredericton, N.B., Canada*

Received 26 February 1999; accepted 20 July 1999

Abstract

An accurate and computationally efficient means of classifying surface myoelectric signal patterns has been the subject of considerable research effort in recent years. Effective feature extraction is crucial to reliable classification and, in the quest to improve the accuracy of transient myoelectric signal pattern classification, an ensemble of time-frequency based representations are proposed. It is shown that feature sets based upon the short-time Fourier transform, the wavelet transform, and the wavelet packet transform provide an effective representation for classification, provided that they are subject to an appropriate form of dimensionality reduction.

Keywords: Myoelectric signal; EMG; Dimensionality reduction; Principal components analysis; Wavelet; Wavelet packet; Time-frequency representation; Neural networks; Pattern recognition; Classification

1. Introduction

The myoelectric signal (MES), collected at the skin surface, has become an important tool in rehabilitation due to the ease with which it may be acquired. The MES provides information about the neuromuscular activity from which it originates, and this has been fundamental to its use in clinical diagnosis, and as a source of control for assistive devices and schemes of functional electrical stimulation. The signal is utterly complex however, as it is influenced by many factors due to the electrophysiology and the recording environment.

It is the complexity of the MES that has presented the greatest challenge in its application to the control of powered prosthetic limbs. Whether or not an artificial limb is an acceptable replacement for a human limb depends upon the expectations of the affected individual, their motivation to incorporate this device into their lifestyle, and the functionality of the device. Clearly, these issues are not mutually exclusive: a potential user will be more motivated to learn how to use a highly functional device. If the system is to replace a single function, such as the opening/closing of a hand, there is relatively little control information required. For this reason, fittings of this type have been very successful. The extension to controlling multiple functions however, is a much more difficult problem. Unfortunately, these are the requirements of those with high-level amputations, and these are the individuals who could stand to benefit the most from a functional replacement of their absent limbs.

The greatest success in multifunction myoelectric control has been realized by pattern recognition of the MES. This work proposes a means of improving the accuracy of these methods in the form of a time-frequency based signal representation. These methods are easily extensible to other applications requiring quantitative or decision-based analysis of the MES.

1.1. Background

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

* Corresponding author. Tel.: +1-506-453-4966; fax: +1-506-453-4827.

E-mail address: kengleha@unb.ca (K. Englehart)

1350-4533/99/$ - see front matter
PII: S1350-4533(99)00066-1

based on an estimate of the amplitude [1] or the rate of change [2] of the MES. This information is used to specify the function to be performed: the *state* of the device. Once the state is selected, it may be driven at a constant speed, or its speed may be controlled in a manner proportional to the myoelectric activity [3]. Although these systems have been very successful, amplitude and rate coded schemes do not provide sufficient information to reliably control more than one device [4].

There have been many attempts to increase the number of states available from the surface MES. Some schemes have used many channels (or sites) of amplitude coded information [5–7] or other statistical measures [8]. A vector of features may then be subject to some form of pattern recognition to assign the state. The requirement of several electrode sites, however, introduces severe problems in locating and maintaining the integrity of patterns of MES activity. Others have attempted to increase the capacity of information from one or two channels of MES activity by using time-series models [9,10]. The results were promising, but the method was sensitive to changes in signal amplitude.

In each of these cases, the *steady state* MES (that produced during constant effort) was used. The steady-state MES however, has very little temporal structure due to the active modification of recruitment and firing patterns needed to sustain a contraction [11]. This is due to the establishment of feedback paths, both intrinsic (the afferent neuromuscular pathways) and extrinsic (the visual system). In a departure from conventional steady-state analysis, Hudgins [12,13] investigated the information content in the *transient* burst of myoelectric activity accompanying the onset of sudden muscular effort, It was found that significant temporal structure exists in these transient MES bursts and that this temporal structure encodes information important for pattern discrimination. Hudgins devised a control system for powered upper-limb prostheses using time-domain features (zero crossings, mean absolute value, mean absolute value slope and trace length) and a simple multilayer perceptron artificial neural network as a classifier. This controller identified four types of muscular contraction using signals measured from the biceps and triceps. This classifier performs well, but improved classification performance would benefit the functionality and, ultimately, the acceptance of artificial limbs controlled by the MES.

In the quest to improve classification accuracy, one has the choice of improving the classifier or the means of signal representation (the feature set). Although some classifiers demonstrate obvious advantages over others, it is the signal representation that most dramatically affects the classification performance, and this is the focus here. Given that transient MES patterns have structure in both time and frequency, it is suggested that the signal energy which would discriminate amongst contraction types would be best concentrated in a dual representation. Consequently, this work explores the efficacy of feature sets derived from time-frequency representations.

2. Methodology

2.1. Problem definition

Although the focus of this investigation is upon signal representation, the classification task is a multi-stage process, and some of the stages are closely coupled. This is illustrated in Fig. 1.

The *feature extraction* stage is defined to be the initial transformation chosen to represent the measured signals. This investigation will compare the performance of Hudgins' time domain (TD) features, and those derived from the short-time Fourier transform (STFT), the wavelet transform (WT), and the wavelet packet transform (WPT.)

The role of *dimensionality reduction* is to retain information that is important for class discrimination and discard that which is irrelevant. A classifier with fewer inputs has fewer adaptive parameters to be determined, leading to a classifier with better generalization properties. Dimensionality reduction strategies may be categorized according to their objective function:

1. *Feature selection* methods attempt to determine the best subset of the original feature set. Feature selection will be performed here using an Euclidean distance class separability (CS) criterion [14]. Feature selection using CS may be regarded as a *supervised* method, since the features are ranked using class membership information.
2. *Feature projection* methods attempt to determine the best combination of the original features. Feature projection was performed using principal components analysis (PCA) which produces an uncorrelated feature set by projecting the data onto the eigenvectors of the covariance matrix [15]. PCA provides a means of *unsupervised* dimensionality reduction, as no class membership qualifies the data when specifying the eigenvectors of maximum variance.

Although the emphasis here is not upon the classifier, the performance of the feature extraction and the dimensionality reduction is dependent upon the capabilities of the *classifier*. The performance of each form of signal representation will be evaluated in the context of a linear discriminant analysis (LDA) classifier [14] and a multilayer perceptron (MLP) c1assitier [22]. The LDA and the MLP are easily implemented and well understood representatives of statistical and neural classifiers, respectively.

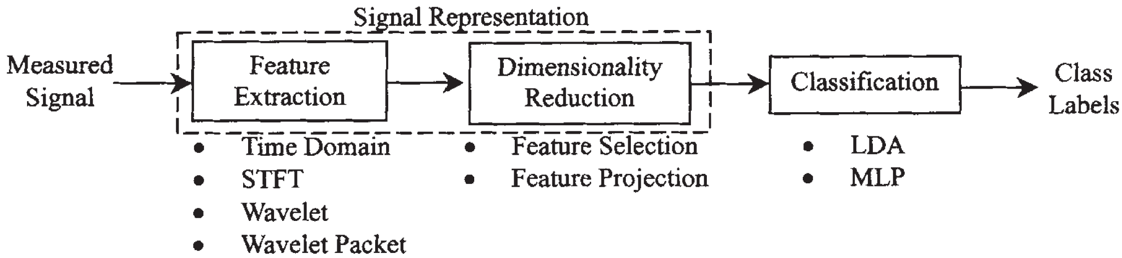

Fig. 1. A breakdown of the classification problem, and the methods subject to investigation.

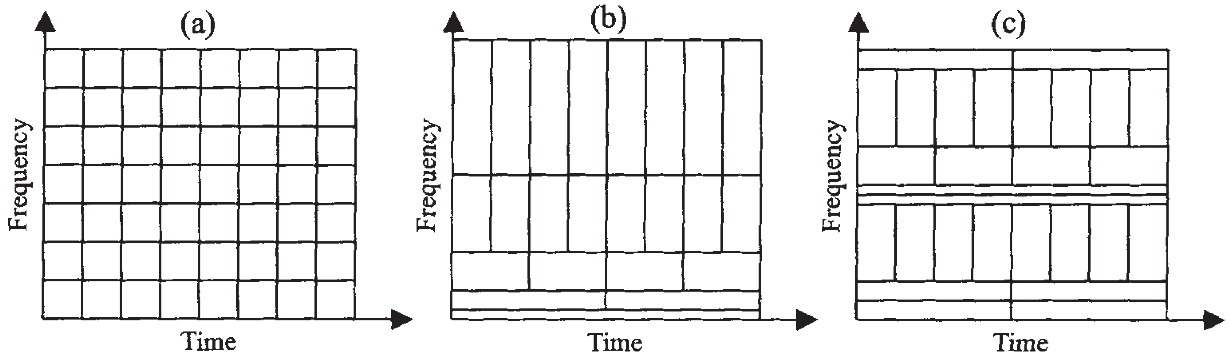

Fig. 2. The time-frequency tiling of (a) the STFT, (b) the WT, and (c) the WPT. Whereas the tiling of the STFT and the WT is fixed, the tiling of the WPT may be adapted to suit a particular application. The tiling in (c) is therefore an arbitrary example.

2.2. Time-frequency representations

Time-frequency representations (TFRs) have received considerable attention in such diverse fields as speech recognition, and the classification of radar, underwater acoustic and geoacoustic signals. In many applications, especially those involving signal classification, it is essential that efficient transforms accompany these representations. For this reason discrete, linear TFRs (the STFT, the WT, and the WPT) are preferable to quadratic TFRs and the continuous wavelet transform in real-time applications [16]. This is certainly the case here, such that a perceptible delay in state selection is undesirable in prosthetic control.

The fundamental difference between linear TFRs 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 wavelet transform is variable — the aspect ratio of the cells varies such that the frequency resolution is proportional to the centre 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. An illustration of the time-frequency tiling subtended by the STFT, the WT, and the WPT is shown in Fig. 2.

The most common approach to specifying the WPT tiling is by selecting that which minimizes the reconstruction error, using an entropy cost function [17]. 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 [18]. This is the method that will be used here.

2.3. The transient MES data

A roster of 16 healthy subjects participated in this study. Four classes of myoelectric signal patterns were collected from the biceps and triceps, corresponding to flexion and extension of the elbow, and pronation and supination of the forearm. Each pattern consists of two channels of n=256 points, sampled at 1000 Hz. See Hudgins et al. [13] for details on the experimental protocol. The data were divided into a training set (100 patterns), a test set (150 patterns), and a validation set (150 patterns)[1]. The validation set provides an estimate of the classification performance of the test set. Consequently, the validation set was used to specify the dimen-

[1] The cardinality of the training, validation, and test sets was specified to provide sufficient data to train the classifiers, estimate the behaviour of the test set, and estimate the classification performance, respectively. The variance of the test set and validation set classification performance could be further reduced by cross-validation techniques, but due to considerable computational expense, these methods have not been considered here.

sionality of the reduced feature set when using CS and PCA, by prescribing the dimension at which the classification error was minimized.

3. Results

For each feature set, there are two factors that affect its efficacy when applied to transient MES classification. These are:

1. **Transform parameters**. Each feature set has parameters that influence the nature of the transform.
2. **Dimensionality reduction**. The performance of some feature sets depends critically upon an appropriate form of dimensionality reduction.

3.1. Transform parameters

Optimization of transform parameters was done empirically, by selecting the methods that subtend the best generalization performance, on average, across all subjects. Table 1 delineates the transform parameters that minimize the test set classification error.

The Hamming STFT window was preferred to Rectangular, Hanning, Blackman, Bartlett, and Kaiser windows [19]. The selection of mother wavelet type was made amongst all possible orders of the following families: Daubechies, Coiflet, Symmlet, Vaidyanathan, Beylkin and biorthogonal spline [20]. The WPT cost function was selected amongst several criteria, including both compression and class separability based objective functions.

3.2. Dimensionality reduction

Using the transform parameters specified above, the performance of each feature set was determined in the context of feature selection (CS) and feature projection (PCA) based dimensionality reduction. Both CS and PCA specify a subset of the $K \leq N$ most informative features, where N is the dimension of the original feature set, and K is the dimension of the reduced feature set. For each scheme the dimension, K, which subtends the best generalization performance must be determined. Fig. 3 shows the effect of feature set dimension upon the validation set classification error, averaged across all subjects.

Table 1
The tansfrom parameters which yield the lowest classification error on the transient MES dataset

STFT	WT	WPT
Window width: 64 ms Window overlap: 50% Window type: Hamming	Mother wavelet: Coiflet-4	Mother wavelet: Symmlet-5 Cost function: Class seperablility (Euclidean distance)

The most striking result is that, when using the TFR based feature sets, PCA easily outperforms CS dimensionality reduction. This is especially dramatic when using a LDA classifier, which cannot accommodate high-dimensional input data as well as a MLP. Indeed, the TD feature set easily outperforms the TFR methods when using CS. When using PCA however, the TFR based methods subtend lower classification error than the TD feature set, especially when using the LDA. It is evident that the STFT, WT and WPT require a *linear projection*, rather than a *subset selection* to accommodate their high-dimensional representation of transient MES patterns.

3.3. The relative performance of the feature sets

Although Fig. 3 reveals some information as to the relative performance amongst the feature sets under consideration here, the most meaningful measure is that of the test set error. For each subject, the best dimension was selected as that which minimizes the validation set error, and the test set error was evaluated at this dimension. Fig. 4 depicts the test set classification error, averaged across all subjects.

From Fig. 4(a) it is evident that CS is ineffective when using the TFR based feature sets. The transient MES has a large degree of within-class variance and correspondingly, the energy of the transient MES is liberally dispersed in the time-frequency domain. This renders any subset of the coefficients of a high-dimensional TFR inadequate for discrimination. The time domain features are more effective, as they occupy a lower dimension and "smooth" the within-class variance. When using the TD features, the application of CS dimensionality reduction slightly improves the generalization performance. For all feature sets, the MLP outperforms the LDA. This is due to its ability to handle higher dimensional input vectors and its capacity to construct nonlinear class decision boundaries.

In Fig. 4(b), when using PCA, the TFR feature sets demonstrate an obvious advantage over the TD sets. When using a LDA classifier, classification performance improves as one progresses from the unreduced TD to PCA-reduced TD features, and from STFT to WT to WPT based features. A similar improvement from TD to STFT is apparent when using a MLP classifier, but the wavelet and wavelet packet methods do not match the performance of the STFT. Overall, the best performance (6.25% error or 93.75% accuracy) is achieved when using a LDA to classify a PCA-reduced WPT feature set.

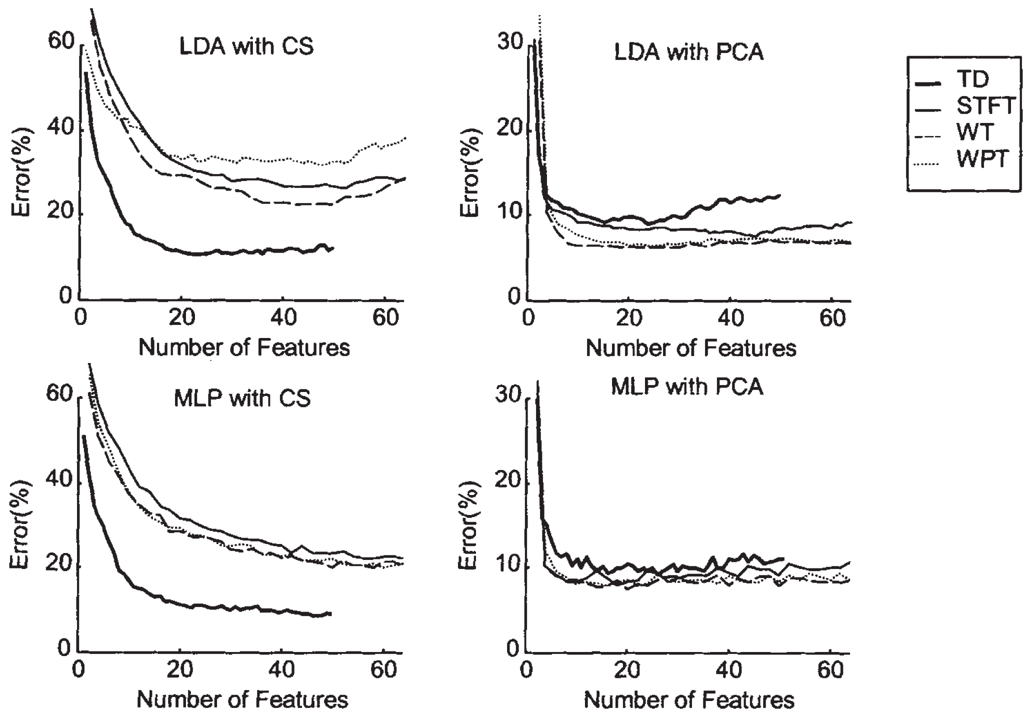

Fig. 3. The effect of feature set dimension upon the validation set error, when using CS and PCA dimensionality reduction. The response, averaged across all subjects, is shown for Hudgins' time domain features (TD), the STFT, the WT and the WPT. Results are given for both LDA and MLP classifiers.

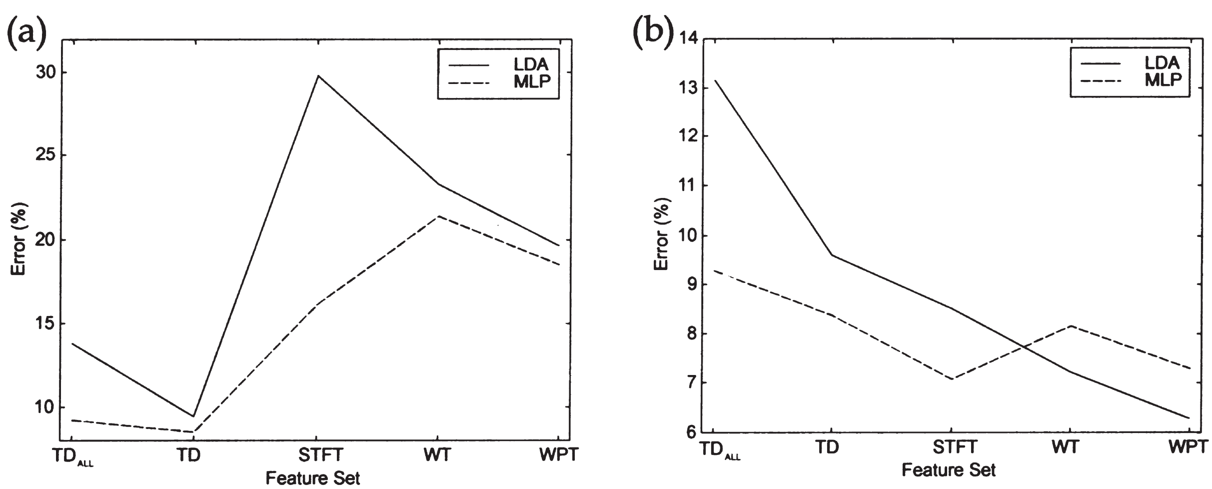

Fig. 4. The test set classification error, averaged across all subjects. The results are shown for each of the feature sets: (a) when using CS and (b) when using PCA. The performance when using the full time domain set (TD_{ALL}) has been included as well.

Although Fig. 4(b) suggests an improvement when using TFR based feature sets, it implies no statistical validation. A substantial amount of variance in the test set classification error is due to inter-subject variability, which obscures the relative performance between feature sets. Consider, for each subject i, a normalization of the error for each reduced feature set:

$$\overline{\mathrm{TD}}_{\mathrm{ALL}i}=\frac{\mathrm{TD}_{\mathrm{ALL}i}}{\Sigma_i},\ \overline{\mathrm{TD}}_i=\frac{\mathrm{TD}_i}{\Sigma_i},\ \overline{\mathrm{STFT}}_i=\frac{\mathrm{STFT}_i}{\Sigma_i},\ \overline{\mathrm{WT}}_i=\frac{\mathrm{WT}_i}{\Sigma_i},\ \overline{\mathrm{WPT}}_i=\frac{\mathrm{WPT}_i}{\Sigma_i}, \qquad (1)$$

where

$$\Sigma_i=\mathrm{TD}_{\mathrm{ALL}i}+\mathrm{TD}_i+\mathrm{STFT}_i+\mathrm{WT}_i+\mathrm{WPT}_i \qquad (2)$$

provides a measure of the overall "ability" of subject i. This normalizes an individual's response and conveys only the relative performance of each feature set, allowing the response from each subject to be interpreted on the same scale. Fig. 5 depicts the mean response of the PCA-reduced feature sets, normalized by the "ability" index. Superimposed on these plots is the scatterplot of the response of each subject. Only the PCA results are shown, since they represent the superior method.

The relative performance of these normalized data show the same trend as the absolute error of Fig. 4, and the individual responses are well-clustered about the mean for each feature set. These data were then subject to a one-way ANOVA with *Feature Set* as the factor. Analyzing the LDA results, one can easily reject the null hypothesis of equal means (p=0.000) and state that there is a significant effect due to *Feature Set*. Table 2 presents the results of a Sheffe *post hoc* test [21], which reveals the significance levels amongst pairs of feature sets.

These results imply that

1. A significant improvement is gained by using PCA reduction on the TD features: TD outperforms TD_{ALL}.
2. All PCA-reduced TFR based feature sets are significantly superior to TD_{ALL}.
3. When using a PCA/LDA combination, there is no significant improvement between adjacent feature sets: TD→STFT, STFT→WT, and WT→WPT, but a significant difference amongst all other combinations, This indicates a distinct trend toward improvement in the progression TD_{ALL}→TD→STFT→WT→ WPT.

The same analysis was performed upon the results using the MLP. The results within each feature set are not clustered as tightly as those of the LDA. Also, there is no distinct improvement from TD_{ALL}→TD→STFT→WT→WPT. Applying a one-way ANOVA to the normalized scores yields p=0.029, allowing one to reject the null hypothesis at α=0.05 across all feature sets. A Sheffe *post hoc* test however, indicates no significant differences amongst any of the individual feature sets.

Table 2
The results of a Scheffe test upon the test set error, with a factor of feature set. The bold entries represent a mean difference that is significant at the α=0.05 level

Feature set 1	Feature set 2	Significance, p
$\overline{\text{TD}}_{\text{ALL}}$	$\overline{\text{TD}}$	**0.001**
	$\overline{\text{STFT}}$	**0.000**
	$\overline{\text{WT}}$	**0.000**
	$\overline{\text{WPT}}$	**0.000**
$\overline{\text{TD}}$	$\overline{\text{TD}}_{\text{ALL}}$	**0.001**
	$\overline{\text{STFT}}$	0.372
	$\overline{\text{WT}}$	**0.012**
	$\overline{\text{WPT}}$	**0.000**
$\overline{\text{STFT}}$	$\overline{\text{TD}}_{\text{ALL}}$	**0.000**
	$\overline{\text{TD}}$	0.372
	$\overline{\text{WT}}$	0.608
	$\overline{\text{WPT}}$	**0.044**
$\overline{\text{WT}}$	$\overline{\text{TD}}_{\text{ALL}}$	**0.000**
	$\overline{\text{TD}}$	**0.012**
	$\overline{\text{STFT}}$	0.608
	$\overline{\text{WPT}}$	0.655
$\overline{\text{WPT}}$	$\overline{\text{TD}}_{\text{ALL}}$	**0.000**
	$\overline{\text{TD}}$	**0.000**
	$\overline{\text{STFT}}$	**0.044**
	$\overline{\text{WT}}$	0.655

4. Discussion

The results of the previous section imply that

1. PCA is vastly superior to CS based dimensionality reduction,
2. a distinct improvement attributable to feature set is evident in the progression TD_{ALL}→TD→STFT→WT→WPT when using a LDA, and
3. a LDA classifier often matches or exceeds the per-

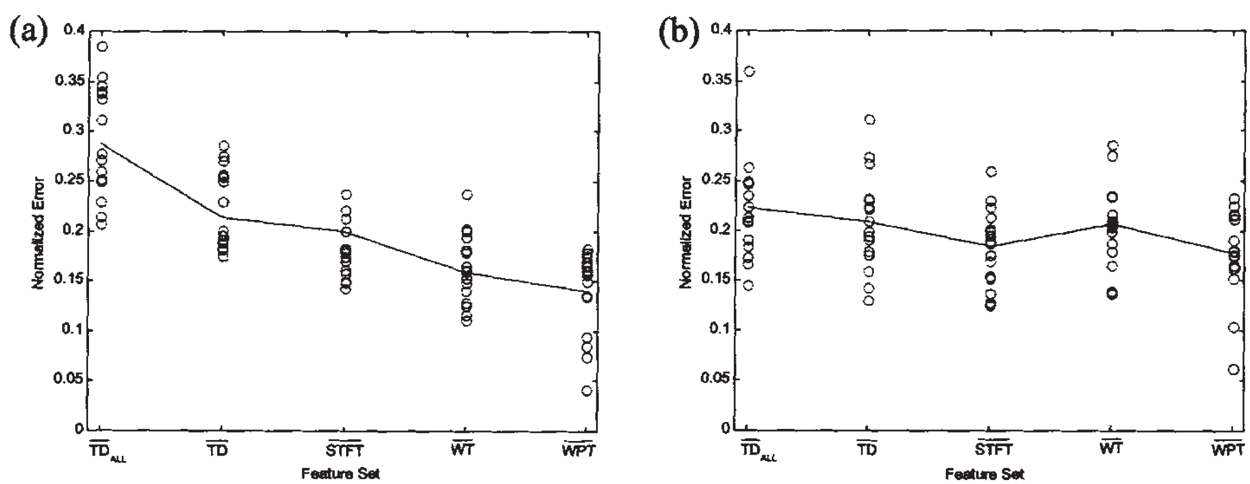

Fig. 5. The normalized classification error of the test set. The results are shown for (a) a LDA classifier, and (b) a MLP classifier.

formance of the MLP when using a PCA reduced feature set.

The reasons for the superiority of PCA to CS when classifying the transient MES are twofold:

1. *The projected features are mutually uncorrelated*. By projecting the data onto the orthonormal axes of maximum variance, the covariance structure is removed. If there are significant linear dependencies in the original feature space, then it may be possible to discard most of the lesser principal components with little loss of information. In the situation where information is liberally dispersed amongst the original feature set, a PCA will consolidate this information much more effectively than feature selection.
2. *The method is unsupervised*. Although it may seem counterintuitive, the knowledge of class membership may actually deteriorate the efficacy of a dimensionality reduction technique. This is because embedding class membership information into the method will bias the representation to the training data in the same manner that a classifier may be biased, hampering the generalization performance [14]. Class separability feature selection methods rely upon class membership in their feature evaluation criteria. PCA, on the other hand, uses no prior knowledge of class membership, and does not experience bias toward the training set. If the variance in the data can be explained by the signal (rather than the noise), then the leading principal axes will tend to pick projections with good separations.

In the high-dimensional space of the STFT, the WT and the WPT, there is almost certainly a significant degree of linear dependency amongst coefficients. The loose structure of the transient MES subtends a substantial degree of within-class dispersion in the time-frequency domain. PCA appears to effectively accommodate these effects. Conversely, feature selection requires so many features to provide adequate discrimination that the resulting dimensionality subtends poor generalization. The improvement that PCA offers to TD features is not as pronounced as that given to the TFR sets, as the original dimensionality is relatively low.

Of particular interest as well is that, when using TFR feature sets subject to PCA, the LDA classifier occasionally exhibits better generalization performance than the MLP classifier. This is despite the fact that the MLP enjoys the advantage over the LDA of being capable of prescribing nonlinear class boundaries, so as to encompass the capabilities of the LDA. To explain the LDA's performance, consider an arbitrary low-dimensional signal representation in which the class boundaries are indeed nonlinear. In this situation, a MLP will most certainly outperform a LDA. Now consider partitioning the feature space (in time, frequency, or some other domain) such that a larger feature set is formed. As the feature set dimensionality grows, the degree of nonlinearity between class boundaries must diminish. In the high dimensional feature space of a TFR, it is unlikely that highly nonlinear bounds exist between the classes. If a significant degree of linear dependency exists as well, a PCA will project the TFR coefficients onto a relatively low dimensional space, while preserving the linearity that exists between classes in the higher dimensional space. The fact that the PCA-projected TFR features have reasonably linear class boundaries and that they have relatively low dimension diminishes the advantage that a MLP may have over a LDA.

A MLP that is appropriately trained and that has an appropriate number of hidden layer nodes will always match, if not exceed, the performance of a LDA. Due to the need to automate MLP training over a large number of trials however, the hidden layer size was fixed at eight and the stopping criterion was a fixed number of iterations (200). These generalizations were determined by empirical analysis of the validation set data. For a given subject however, the size of the MLP may be inappropriate, or the network may be overtrained or undertrained. Both of these factors will inhibit the generalization performance of the MLP. The LDA does not require heuristic specification of its architecture or training algorithm, yet it consistently performs very well.

The advantage that PCA does offer to MLP classifiers is with respect to training time. A speedup of the back-propagation algorithm may result from the application of PCA since the Hessian matrix of the cost function is more diagonalized than usual. This generates an appropriate scaling of the learning rate along each weight axis independently [22].

5. Conclusion

It has been shown that, when using TFR based feature sets, PCA provides a far more effective means of dimensionality reduction than feature selection by CS. Moreover, by preprocessing the feature set with PCA prior to classification, a LDA — a classifier that is easier to implement and to train than a MLP — may be used without degrading performance. It has also been demonstrated that when using a PCA/LDA combination, there is a significant improvement in performance in the progression $TD_{ALL} \rightarrow TD \rightarrow STFT \rightarrow WT \rightarrow WPT$. The best performance is exhibited when using a WPT/PCA/LDA combination, yielding an average classification error of 6.25%. This represents a significant improvement over Hudgins' method which, for these data, subtend an average error of 9.25%.

Hudgins' pattern recognition based myoelectric control system has been implemented as an embedded controller, based upon a digital signal processing microprocessor [23]. This controller has shown early success in clinical trials. The signal representation proposed by this work can be easily implemented into this controller's software; it is anticipated that the resulting improvement in accuracy will enhance the functionality of prosthetic control. The WPT has a complexity on the order of $n \log n$, and the stages of PCA and LDA, once trained, require a simple matrix multiplication in the feedforward path. Algorithm complexity estimates indicate that the feedforward computation required of the WPT/PCA/LD combination will easily meet real-time constraints in this controller, which has relatively modest signal processing capacity.

A PCA-reduced time-frequency signal representation has been shown to effectively accommodate the loosely structured waveforms of the transient MES. This suggests that the same success may be found upon analyzing other biological signals, many of which exhibit similar characteristics. Moreover, the results of this work can be easily generalized to other procedures that require quantitative and decision-based analysis.

Acknowledgements

The authors acknowledge the assistance of the Natural Sciences and Engineering Research Council of Canada and the Whitaker Foundation.

References

[1] Dorcas D, Scott RN. A three state myoelectric control. Medical and Biological Engineering 1966;4:367–72.
[2] Childress DA. A myoelectric three state controller using rate sensitivity, in Proc. 8th ICMBE, Chicago, IL, 1969:S4–5.
[3] Parker PA, Scott RN. Myoelectric control of prostheses. CRC Critical Reviews in Biomedical Engineering 1986;13(4):283–310.
[4] Vodovnik L, Kreifeldt J, Caldwell R, Green L, Silgalis E, Craig P. Some topics on myoelectric control of orthotic/prosthetic systems, Rep. EDC 4-67-17, Case Western Reserve University, Cleveland, OH, 1967.
[5] Schmeidl H. The I.N.A.I.L. experience fitting upper-limb dysmelia patients with myoelectric control. Bulletin of Prosthetics Research 1977;10(27):17–42.
[6] Wirta RW, Taylor DR, Findley FR. Pattern recognition arm prosthesis: A historical perspective — Final Report. Bulletin of Prosthetics Research 1978;10(30):8–35.
[7] Almstrom C, Herberts P, Korner L. Experience with Swedish multifunction prosthetic hands controlled by pattern recognition of multiple myoelectric signals. Int Orthopaed 1981;5:15–21.
[8] Saridis GN, Goatee TP. EMG pattern recognition for a prosthetic arm. IEEE Trans Biomed Eng 1982;29:403–12.
[9] Graupe D, Salahi J, Kohn KH. Multifunction prosthesis and orthosis control via microcomputer identification of temporal patter differences in single-site myoelectric signals. J Biomed Eng 1982;4:17–22.
[10] Doerschuck PC, Guftafson DE, Willsky AS. Upper extremity limb function discrimination using EMG signal analysis. IEEE Trans Biomed Eng 1983;30:18–28.
[11] DeLuca CJ. Physiology and mathematics of myoelectric signals. IEEE Trans Biomed Eng 1979;26:313–25.
[12] Hudgins BS. A New Approach to Multifunction Myoelectric Control, Ph.D. Dissertation, University of New Brunswick, Fredericton, N.B., Canada, 1991.
[13] Hudgins B, Parker PA, Scott RN. A new strategy for multifunction myoelectric control. IEEE Trans Biomedical Engineering 1993;40(1):82–94.
[14] Fukunaga K. Introduction to statistical pattern recognition. 2nd ed. Academic Press: San Diego, CA, 1990.
[15] Bishop CM. Neural networks for pattern recognition. New York: Oxford University Press, 1996.
[16] Cohen L. Time-frequency distributions — A Review. Proc IEEE 1989;77(7):941–81.
[17] Coifman RR, Wickerhauser MV. Entropy-based algorithms for best basis selection. IEEE Trans Information Theory 1992;38(2):713–9.
[18] Saito N, Coifman RR. Local discriminant bases and their applications. J Math Imaging and Vision 1995;5(4):331–58.
[19] Marple SL. Digital spectral analysis with applications. Englewood Cliffs, New Jersey: Prentice-Hall, 1987.
[20] Daubechies I. Ten Lectures on Wavelets, CBMS-NSF Regional Conference Series in Applied Mathematics, SIAM, Philadelphia, 1992:61.
[21] Hicks CR. Fundamental concepts in the design of experiments. New York: Saunders College Publishing, 1993.
[22] Haykin S. Neutral networks: a comprehensive foundation. Don Mills, Ontario: Maxwell MacMillan Canada, Inc, 1994.
[23] Hudgins B, Englehart K, Parker PA, Scott RN. A microprocessor-based multifunction myoelectric control system, 23rd Canadian Medical and Biological Engineering Society Conference, Toronto, May, 1997.