MYOELECTRIC SIGNAL CLASSIFICATION USING A FINITE IMPULSE RESPONSE NEURAL NETWORK

K.B. Englehart, B.S. Hudgins, M. Stevenson and P.A. Parker
The University of New Brunswick, Dept. of Electrical Engineering
and the Institute of Biomedical Engineering

ABSTRACT

Recent work by Hudgins [1] has proposed a neural network-based approach to classifying themyoelectric signal (MES) elicited at the onset of movement of the upper limb. A standardfeedforward artificial network was trained (using the backpropagation algorithm) to discriminateamongst four classes of upper-limb movements from the MES acquired from the biceps and triceps muscles. The approach has demonstrated a powerful means of classifying limb function intent from the MESduring natural muscular contraction, but the static nature of the network architecture fails to fullycharacterize the dynamic structure inherent in the MES. It has been demonstrated [2] that a finite-impulseresponse (FIR) network has the ability to incorporate the temporal structure of a signal, representing therelationships between events in time and providing translation invariance of the relevant feature set The application of this network architecture to limb function discrimination from the MES is described here.

Keywords — Neural Networks, Myoelectric Signal, Prosthetic Control

INTRODUCTION

A significant body of research has focused upon extracting sufficient information from the myoelectric signal (MES) to describe limb function intent This information may then be used toactuate a powered prosthetic device or manage a scheme of functional electrical stimulation. Basic feature extraction has subtended a useful, but crude control signal: robust, dextrous, andnatural control of external devices remains an elusive goal.

Classification schemes proposed in most previous work have attempted discrimination using features extracted from the *steady-state* MES, elicited by constant contractions. The following are some notable contributions[1].

(i) Amplitude-coded Systems
- one/two channel [2] (Dorcas et al.[4]
- many channel (Jacobsen et al.[5])

---

[1] This outline is far from complete - for a more detailed overview, refer to [3].

[2] The term *channel* implies an independent monopolar or bipolar recording from a distinct site.

(ii) Pattern Recognition-Based Systems
- one/two channel (Graupe *et al.*[6], Saridis *et al.*[7]
- many channel (Wirta *et al.*[8], Almstrom *et al.*[9], Freedy *et al.*[10])

There are fundamental limitations to these schemes. *Many channel* systems are clinically inconvenient— electrode contact integrity is difficult to establish and maintain. It is difficult (sometimes impossible) to train an individual to effectively control more than a few movement types if the required contractions are contrived. Most control schemes do not offer a *natural* interface in which limb functionintent is discerned by contractions resembling normal muscular effort. The need to limitthe number of channels and to maximize the information extracted from each channel has prompted investigation into further detail of the MES.

Hudgins [1] has demonstrated that the transient burst patterns of the MES acquired at the onset of upper-limb movement possess a significant degree of temporal structure[3]. The MES burst patterns typical of four classes of limb movement are shown in Figure 1.
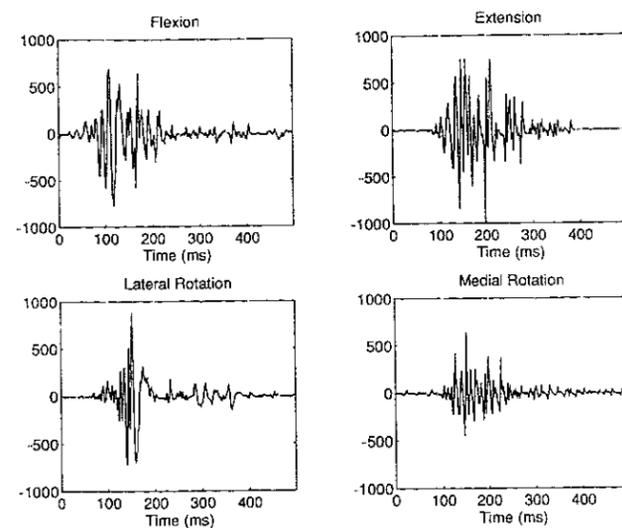

Figure 1. The myoelectric signal acquired at the onset of four types of upper limb movement: elbow flexion/extension and lateral/medial rotation of the wrist.

Further, Hudgins proposed a means of extracting control information from these transient patterns. A feedforward artificial neural network was trained (using the backpropagation algorithm) to discriminate amongst four classes of upper-limb movements. The work was significant in that it demonstrated a means of classifying limbfunction intent from the MES acquired during *natural muscular contractions.*

---

[3] Hudgins used a bipolar recording arrangement, one active lead on the biceps and one on the triceps.

NEURAL NETWORK CLASSIFICATION

The network was used to classify bursts of the MES accompanying the onset of movement, as described above. Hudgins investigated the effects of segmenting the 240*ms* window into a number of frames to capture the temporal excursion of the feature set. The features extracted from each frame were chosen to capture distinct details of the burst patterns:
- zero-crossings (ZC),
- trace length (TL),
- mean absolute value (MAV), and
- differential mean absolute value (dMAV).

The inherent tradeoff in reducing frame size to obtain more frames within the 240*ms* data record was that, although the temporal resolution increases, the variance of the features also increases and the dynamic range of the features decreases. Moreover, since all time-delayed frames in the input pattern must be presented to the neural network at once, more frames means more inputs, which means more network weights. Network training complexity grows quadratically and networkgeneralization suffers with additional weights.

For the static feedforward classifier used in his work, Hudgins determined that six 40*ms* time frames offered an acceptable compromise amongst the factors of this tradeoff. The static network that was used is shown in Figure 2.
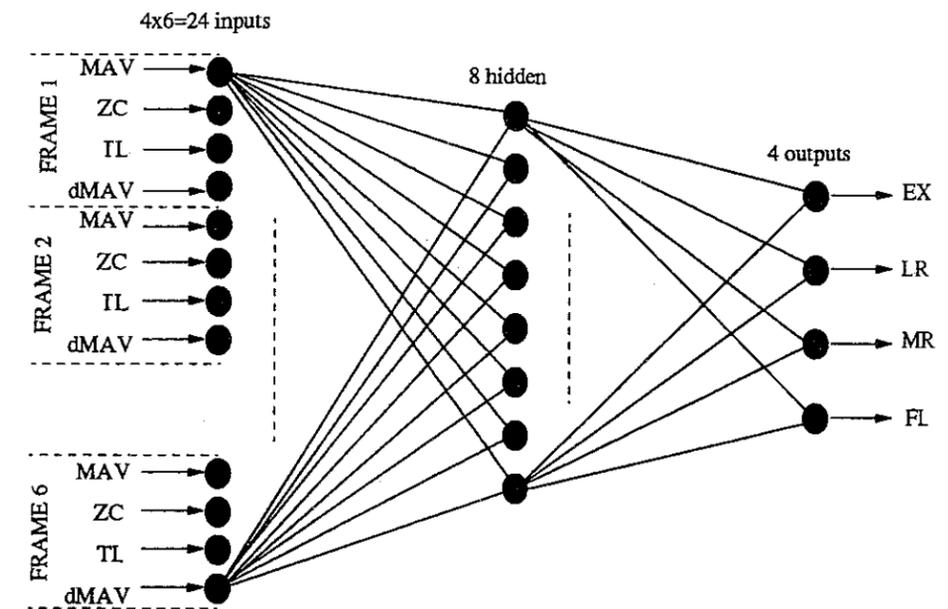

Figure 2. The static feedforward network used by Hudgins.

Each pattern presentation consists of sets of four features from six successive time frames — the number of neurons in the input layer is 4 x 6 = 24. Hudgins determined — *via* cross-validation classification rates — that eight neurons in the hidden layer offered the best generalization applied to a set of test patterns.

Classification performance was impressive. Most normally-limbed subjects were capable of correctly selecting one of four motions more than 90 percent of the time with very little training. The success of limb-deficient subjects was more variable, depending on the nature of their deficiency, but most individuals were capable of approaching the classification rate of the normally-limbed subjects. The static neural network classifier, however, imposes some limitations:

- A static network must have all features, at all relevant time instances, presented simultaneously to the network during training and classification. The input layer and thus, the network can become very large as the number of time frames increases.
- The network is sensitive to misalignment of the observation window. Accurate triggering and framing of the data window is imperative.
- The network has no *a priori* knowledge of the temporal nature of the burst pattern feature set. A network constrained to model temporal signals can be smaller and may be trained more quickly than a naive static network.

A novel neural network classifier is introduced here to address these limitations intrinsic to a static network implementation.

THE FIR NEURAL NETWORK CLASSIFIER

In his doctoral dissertation [2], Eric Wan has described a modification to standard feedforward networks to incorporate temporal structure in the architecture. Each connection in the network is no longer a simple scalar, but an FIR *filter*. Consider an FIR structure of order M with input $x(k) = [x(k), x(k-1),..., x(k-M)]^T$ and weights $w = [w(0), w(1),..., w(M)]^T$. The linear output of the FIR is

$$y(k) = \sum_{n=0}^{M} w(n)x(k-n) = w \cdot x(k). \tag{1}$$

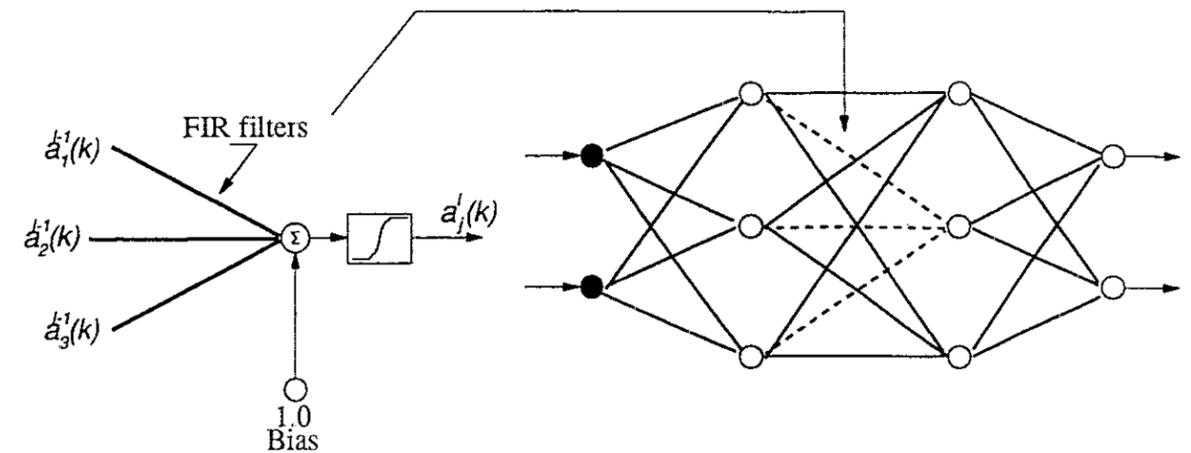Consider now a network of FIR neurons, depicted in Figure 3.

Figure 3. The FIR neuron and network.

Within the network of neurons, the coefficients for the synaptic filter connecting neuron i to neuron j in layer l are specified by the vector

$$w_{ij}^l = [w_{ij}^l(0), w_{ij}^l(1),..., w_{ij}^l(M^{(l)})]^T \tag{2}$$

where $M^{(l)}$ is the order of the filters in layer l. The scalar bias weight for neuron j in layer l is notated as w_{bj}^l. The summing junction for neuron j in layer l at the kth weight update forms the spatio-temporal integration:

$$s_j^l(k) = \sum_i \sum_{n=0}^{M^{(l)}} w_{ij}^l(k)a_i^{l-1}(k-n) = \sum_i w_{ij}^l \cdot a_i^{l-1}(k)) + w_{bj}^l \tag{3}$$

where

$$a_i^{l-1}(k) = [a_i^{l-1}(k), a_i^{l-1}(k-1),..., a_i^{l-1}(k-M^{(l)})]^T \tag{4}$$

is the vector of delayed activations from the previous layer[4].

The inputs to the network are $x_i(k) \stackrel{\Delta}{=} a_i^0(k) \; \forall \; i \in \{1..N\}$, and the outputs are $y_i(k) \stackrel{\Delta}{=} a_i^L(k) \; \forall \; i \in \{1..M\}$. The number of neurons in layer l is denoted by $N^{(l)}$. The overall architecture of an L-layer network may be described as

---

[4] In general $a_i^l(k) = f(s_i^l(k))$; the activation function $f()$ is often *tanh*(), but many possibilities exist.

$$N^{(1)} \times N^{(2)} \times \ldots \times N^{(L)} \text{ neurons}$$

$$M^{(2)}: \ldots: M^{(L)} \text{ FIR tap}[5]$$

The focus of this work was to investigate the ability of the FIR network to perform the MES classification task, and to identify the strengths it may have as an alternative to a static network classifier. The first approach was to use the same feature-based signal representation as that employed by Hudgins: this allowed a direct comparison of classification and generalization performance. Next, to exploit the temporal architecture of the FIR network, alternative signal representationsoffering greater temporal resolution of the input pattern were considered.

The FIR structure of nearest complexity to that of Hudgins' classifier is a three-layer network with

$$4 \times 8 \times 4 \text{ neurons}$$

$$5:0 \text{ taps}$$

The number of neurons in the input and the output layer were constrained to match the dimension of the input pattern (four features) and the output (four classes). To compare the performance of the FIR network to a static network, the network complexities should be, in some sense, comparable. If the two networks are specified to have the same number of weights, then the static network will necessarily have more neurons than the FIR network. If the number of neurons in the two networks are equal, the FIR network have fewer weights. Since the weights of a static network that is equivalent to a given FIR network are necessarily constrained, a direct comparison is not straightforward. A comparison based on an equal number of neurons in the hidden layer is a more appropriate approach.

The normalized mean squared error (nMSE) trajectories of the FIR network for two data sets are shown in Figure 4. The error trajectories for both the training set data and the test set data are given [6]. As would be expected, the network error is always greater applied to the test set data. Both test sets exhibit a point of minimum nMSE, with further training adversely affecting network generalization. Figure 5 shows the trajectory of the classification rate upon the test set during training. Both sets exhibit a point during training of maximum classification; further training adversely affects the network generalization. Curiously, the classification rate does not begin to degrade past the point of maximum network generalization, as implied by the test set nMSE of Figure 4b. The discrete nature of classification apparently obscures the loss of generalization implied by the continuous measure of the nMSE cost function. The choice of eight hidden layer nodes appears to have been a good one for the FIR network. A network with fewer hidden layer nodes demonstrated

---

[5] The orders are given for FIR synaptic structures of layers 2 to L. Since $x_i(k) \stackrel{\Delta}{=} a_i^0(k)$, by definition, $M^{(1)} = 0$ and the scalar weight is unity.

[6] The largest learning factor that would guarantee convergence was used. This was determined heuristically: for this network, $\mu=0.1$ (in each layer) seemed to be a stable choice; training with $\mu=0.5$ exhibited intermittent instability.

inferior training set discrimination, and consequently, poor generalization. Additional hidden layer nodes caused the network to "memorize" the training set too much — generalization upon the test
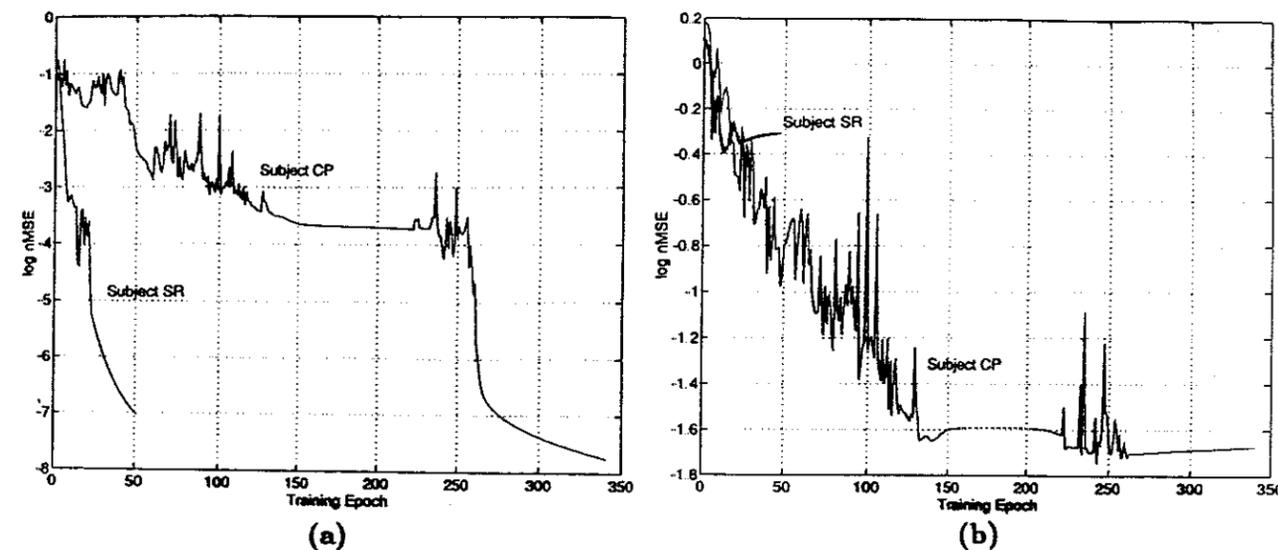

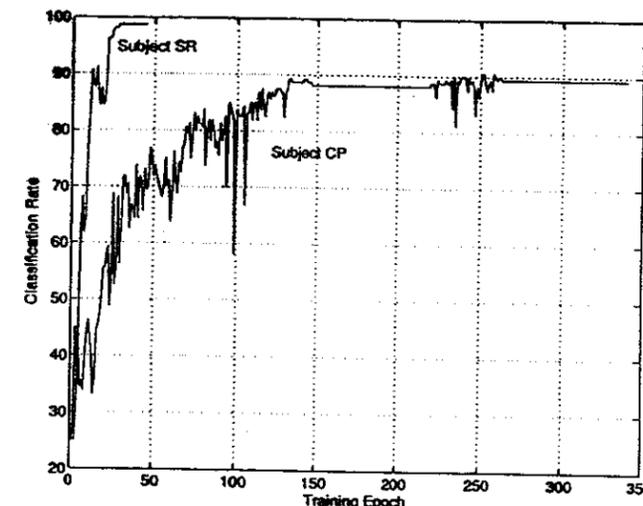**Figure 4.** The nMSE during training: (a) Training Set, (b) Test Set

Figure 5. The classification rate upon the test set during training.

The best classification performance of the FIR network and a standard backpropagation network upon the data from the same two subjects are shown in Table 1.

Subject	Time Frames	Static Network (correct/trials)	FIR Network (correct/trials)
SR	6	79/80	79/80
CP	6	64/80	64/80

Table 1. Classification performance.

For both the "good" and the "bad" data set, theclassification performance of both networks is essentially the same. The strengths offered by the FIR architecture are the ability to capture the temporal excursion of the feature set, and to compensate for misalignment of the patterns presented to the network. The small number of frames (*six*) as defined by Hudgin's temporal segmentation scheme limits the range of freedom to specify FIR orders to capture hierarchies of subfeatures, and to accommodate temporal skew in the feature set. There are simply too few frames to play with. For this reason, alternatives to a simple (5:0) FIR tap structure did not offer improved classification performance.

Consider signal segmentation yielding a greater number of time frames. The strength of the FIR network is in capturing the temporal structure of a waveform, suggesting a signal representation that is as temporally resolvedas possible. To capture the temporal dimension of a waveform pattern, a static network would need to contain many time-delayed inputs and thus, many more weights. Network training complexity grows quadratically and network generalization suffers with additional weights. The FIR network implements the time delays longitudinally; additional time delays (and thus, data points per raw pattern presentation) add to training complexity linearly.

The "bad" dataset is the more interesting one for analysis, so it is the one to be further scrutinized here. Reducing the frame size to 12 frames of 20 *ms* gave slightly better performance for both the static and FIR network; the added temporal resolution appears to have supplied more information. To generate a representation with much greater temporalresolution, each burst was segmented into forty 40 *ms* overlapping frames, the boundsof successive frames incremented by 5*ms*[7]. The trajectories of the classification rate during training for the FIR networks corresponding to 6, 12 and 40-frame representations are shown in Figure 6. In each case the classification rate is the percentage of correct classifications upon a set of test data of patternsfrom each of four movement types. It is apparent that the classification rate improves withincreased temporal dimension,

---

[7] At some level of overlap the redundancy between frames negated the advantage of increased temporal dimension; the 7/8 overlap here is an arbitrary initial choice.

but not markedly so. A feature set with greater information capacitymay be what is needed to improve classification performance. It is entirely possible that this data setcannot be further re-solved — some of the patterns may simply be too far from the bounds of their correctfeature space.
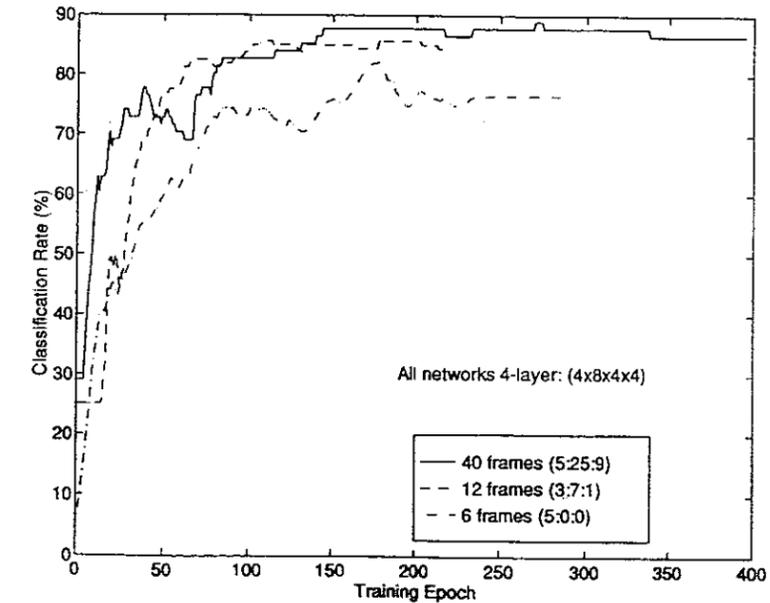

Figure 6. The classification rate upon the test set during training.

CONCLUSIONS

The advantages offered by the FIR network classifier are:

• The temporal dimension of the feature set is accommodated *longitudinally* through the network, significantly reducing the size and training complexity of the network.

• The network can compensate for temporal misalignment of characteristic features in each pattern presentation.

The focus of current work assigns the FIR network to the task of enhancing the robustness of classifier performance. For signals of finite support, as are the MES burst patterns here, awaveform classifier offers significant advantages:

1. **It may act as a continuous classifier.** The network can be trained on a continuous stream of input data, with an associatedcontinuous desired response, identifying the regions of as-serted input patterns. A network trained in this manner would becapable of acting as a *on-line* classifier, identifying a valid class (limb movement intent) from a continuous waveform input.

2. **It may be trained to classify using a subset of the entire pattern.** Instead of training the classifier to approach a desired response only at the *instant* the entire waveform fills the network (and zero everywhere else), a weighted *range* of nonzero desired response is given. Specifying the desired response to build from zero (when valid class data is first presented) to a maximum (when the data has filled the network) assigns a *confidence range* on the output of the network. The network now acts as a *sequential classifier*: if the output exceeds a confidence threshold, possibly before all data has been presented, a classification result is available. The availability of a result upon a subset of the entire input pattern would decrease the response time of the classifier; sequential classification has been the goal of previous work using amplitude-based schemes [11]. Classification with partial data is not possible using a static network; all data must be presented simultaneously for the result to be meaningful.

The representation of the MES is a fundamental issue as well. An optimal representation of the burst patterns is under scrutiny: among candidates are wavelet basis and principal component coefficients.

ACKNOWLEDGEMENTS

REFERENCES

1. Hudgins, B.S., P.A. Parker, and R.N. Scott, "*A New Strategy for Multifunction Myoelectric Control*", IEEE Trans. Biomed. Eng., vol. 40. No. 1, Jan., 1993.

2. Wan, E., "*Finite Impulse Response Neural Networks with Applications in Time-Series Prediction*," Ph.D. Dissertation, Stanford University, Stanford CA, November, 1993.

3. Parker, P.A. and R.N. Scott, "*Myoelectric Control of Prostheses*," CRC Critical Reviews in Biomedical Engineering, vol. 13, issue 4, pp. 283-310, 1986.

4. Dorcas, D.S. and R.N. Scott, "*A Three-State Myoelectric Control*," Medical and Biological Engineering, 4, p. 367, 1966.

5. Jerard, R.B. and Jacobsen, S., "*Laboratory Evaluation of a Unified Theory for Simultaneous Multiple Axis Artificial Arm Control*," Journal of Biomedical Engineering, vol. 102, pp. 199-207, 1980.

6. Graupe, D. and W.K. Cline, "*Functional Separation of EMG Signal via ARMA Identification Methods for Prosthetic Control Purposes*," IEEE Trans. on Systems, Man and Cybernetics, SMC-5, No. 2, pp. 252-259, 1975.

7. Saridis, G., "*EMG Trained, Prosthetic Control*," Joint Coordinating Forum for the International Advanced Robotics Programme, Ottawa, Canada, pp. 15.1-15.9, 1988.

8. Wirta, R.W., Taylor, D.R. and F.R. Finlay, "*Pattern Recognition Arm Prosthesis: A Historical Perspective — Final Report*," Bulletin of Prosthetics Research, vol. BPR 10-30, pp. 8-35, Fall, 1978.

9. Almstrom, C., Herberts, P. and L. Korner, "*Experience with Swedish Multifunction Prosthetic Hands Controlled by Pattern Recognition of Multiple Myoelectric Signals*" Int. Orthopaed, vol. 5, pp. 15-21, 1981.

10. Freedy, A., Hull, F.C., Lucaccini, L.F. and J. Lyman, "*A Computer-Based Learning System for Remote Manipulator Control*," IEEE Transactions on Systems, Man and Cybernetics, SMC-1, No. 4, pp. 356-363, October, 1971.

11. Bérube, J.L., P.A. Parker, R.E. Gander, and V.A. Dunfield, "*Digital Myoelectric Signal Processor with Adaptive Signal Boundaries*", Medical & Biological Engineering and Computing, pp. 349—352, July, 1984.