Mansoura Engineering Journal

Volume 24 | Issue 2 Article 5

2-10-2021

Automatic Analysis and Classification of Surface EMG.

A. El-Nashar
Communications Engineer in the Delta Electricity Company ,

Fatma Abou-Chadi
Electronics and Communications Engineering Department, Faculty of Engineering, Mansoura University, f-abochadi@ieee.org

M. Saad
Department of Neurology., Faculty of Medicine., El-Mansoura University., Mansoura., Egypt.

Automatic Analysis and Classification of Surface EMG

تحليل و تصنيف إشارات النشاط العضلى الكهربى آليا

A. El-Nashar*, F. E. Z. Abou-Chadi**, and M. Saad***
*Communications Engineer in the Delta Electricity Company
**Department of Electronics and Communications Engineering, Mansoura Univ.
***Department of Neurology, Faculty of Medicine, Mansoura Univ.

الخلاصة- فى هذا البحث تم دمج التمثيل البارامترى الإرجاعى والشبكات العصبية للحصول على منظومة متكاملة لوصف وتصنيف إشارات النشاط العضلى الكهربى المسجلة لمجموعة من الأشخاص الأصحاء والمرضى باستخدام الإلكترود السطحى. وتقوم المنظومة أولا بتعيين عدد من المتغيرات الإرجاعية للإشارة يمكن استخدامها كسمات لتوصيف الأنواع المختلفة من الأمراض. ثم يتم إدخال هذة المتغيرات لمصنفات الشبكات العصبية بغرض تصنيفها آليا. وقد تم استخدام ثلاث نماذج مختلفة من الشبكات العصبية وتمت مقارنة أدائها مع بعضها البعض وأيضا مقارنة أدائها مع خوارزم "فيشر" المعروف وبلغت صحة التصنيف 90% . وقد تبين من النتائج أنه يمكن استخدام إشارات النشاط العضلى المسجلة بالإلكترود السطحى بدلا من الإلكترود الإبرى فى التشخيص.

Abstract- **In this paper, parametric modeling of surface EMG algorithms that facilitates automatic SEMG feature extraction and Artificial Neural Networks (ANN) are combined for providing an integrated system for the diagnosis of myopathic disorders. Three paradigms of ANN were investigated: multilayer backpropagation algorithm, self-organizing feature map algorithm and a probabilistic neural network model. The performance of the three classifiers was compared with that of the old Fisber linear discriminant (FLD) classifiers. The results have shown that the three ANN models give higher performance. The percentage of correct classification reaches 90%. Poorer diagnostic performance was obtained from the FLD classifier. The system presented here indicates that surface EMG, wben properly processed, can replace needle EMG in some clinical applications and can be used to provide the physician with a diagnostic assist device.**

I. INTRODUCTION

ELECTROMYOGRAPHY (EMG) is the study of the electrical activity of muscle, and forms a valuable aid in the diagnosis of neuromuscular disorders. EMG findings are used to detect and describe different disease processes affecting the motor unit, the smallest functional unit of the muscle. With voluntary muscle eontraction, the action potential reflecting the electrical activity of a single anatomical motor unit recorded. It is the compound motor unit action potential (MUAP) of those muscle fibers within the recording range of the needle or surfacc electrodes [1].

The MUAP waveform depends on the motor unit architecture, i.e., on the number of fibers, their sizes, and density, so the analysis of MUAP shape may provide important information about the motor unit structure and its changes. The disease processes which affect the structure and activity of the motor unit are reflected in the changes of MUAP features (particularly those of the durations and amplitudes). These changes may also manifest themselves as polyphasic potentials [2] characterized by an increased number of phases and/or

Accepted June 13, 1999

turns, i.e., in signals of a more complicated shape than the normal MUAP. The changes of the MUAP shape are an important indicator of MU disintegration and compensatory processes [3].

Recently, several authors successfully investigated muscle properties by analyzing the time course of amplitude parameters, muscle fiber conduction velocity, and spectral parameters of the EMG signal during voluntary and electrically elicited isometric contractions [4, 5].

It is also well known and documented that the power spectral density function of the EMG signal undergoes frequency compression during either voluntary or electrically elicited sustained contractions [6], long before the muscle becomes unable to produce the desired force. Such changes are referred to as myoelectric manifestations of localized muscle fatigue.

The spectral content of the EMG signal depends on (a) the number of active motor units whose electrical activity is sensed by the detection probe, (b) their firing rates, (c) the position of the active muscle fibers relative to the detection probe, and (d) the velocity of propagation of depolarization along muscle fibers [7]. During a sustained muscle contraction, the spectral compression is mainly due to a progressive reduction of muscle fiber conduction velocity and to the variation of the spatial distribution of depolarization along the muscle fibers [8]. Therefore, if spectral parameters are studied, it is important to separate their random variations due to estimation errors from those due to physiological events.

Previous approaches for analyzing the time-varying aspects of the EMG signals have used a linear prediction model. Among them, the autoregressive (AR) model has been used to deal with time-varying EMG signals because it emphasizes spectral peaks for time records having a small number of samples [9]. This approach was introduced by Graupe and Cline [10] who attempted to use the surface EMG signal for controlling prostheses. Subsequently, Sherif et al [11] studied the behavior of autoregressive integrated moving average (ARIMA) coefficients of the EMG signal from the deltoid muscle during dynamic contractions. Capponi et al. [12] represented EMG signals, detected from the biceps and triceps muscles, with the time courses of AR coefficients during rapid isometric contractions. Recently, Kiryu et al. [13] investigated the physiological interpretation of AR modeling. They analyzed the time-varying behavior of AR parameters of well-conditioned EMG signals detected during an isometric force-varying ramp contraction. The AR coefficients of the EMG signal could be used as quantitative measures to monitor local muscle fatigue [14, 15].

To further the development of quantitative EMG techniques, the need has emerged for adding automated decision making support to these techniques so that all data is processed in an integrated environment. Towards this goal, Coatrieux et al. [16] applied cluster analysis for the automatic diagnosis of pathology based on MUAP records. Hassoun et al. [17] proposed an automated electromyogram signal decomposition using neural networks. Pattichis et al. [18] utilized artificial neural networks for the automatic classification of EMG features recorded using needle electrodes for normal individuals and patients suffering with neuromuscular diseases. They used seven features derived from the shape of the MUAP waveforms.

The main goal of the present work is threefold: to assure the usage of surface electromyography (SEMG) in clinical diagnosis, to characterize the SEMG signal through the determination of the autoregressive model parameters to be used for comparisons between groups of patients or between an individual record and any population norms that might become available, and to provide an efficient classification for the different pathological cases.

The classification approaches taken here are: the old Fisher linear discriminant algorithm and three models of neural networks: multilayer back-propagation model, self-organizing feature map model and a probabilistic neural network model. A comparison of the performance of the four classifiers is performed for normal individuals and patients suffering with myopathic lesions .

II. METHOD

Twenty eight subjects were used this work: 14 normals and 14 suffering from myopathy. SEMG signal was recorded from the deltoid muscle at 50% Maximum Voluntary

Contraction (MVC) for five seconds using bipolar surface electrodes. The recording points within the muscle are standardized. The Biopac data acquisition system which consists of internal microprocessor MP100 data acquisition card with sixteen-analog input channels and connected to an Apple PC was utilized. The software used in data acquisition is the Acqknowledge Ver.2 software. Simple surface EMG electrodes were used and EMG100 biomedical differential amplifier was used. This system is available at the Medical Electronics Lab., Faculty of Engineering, Mansoura university.

The Sampling Frequency

It has been shown that information exists in the EMG frequency spectrum up to frequencies of 1 kHz, implying that in order to satisfy the Nyquist criterion, a sampling frequency of at least 2 kHz would have to be used. However, for surface myoelectric signals, most of the power in the signal is at low frequencies (below 300 Hz).

Further, consideration of average autocorrelation functions of the EMG recorded from two subjects one for myopathic subject and the other for normal subject from a pair of surface electrodes placed over motor unit of the deltoid muscle, indicates that the difference between these functions is more pronounced at low frequencies. Fig.1. illustrates the average autocorrelation functions of the two records. It can be seen that there is a significant correlation up to the second time lag. This time interval corresponds to a frequency value of about 500Hz. Therefore, it is argued to choose a sampling frequency 1 kHz for diagnosis purposes.

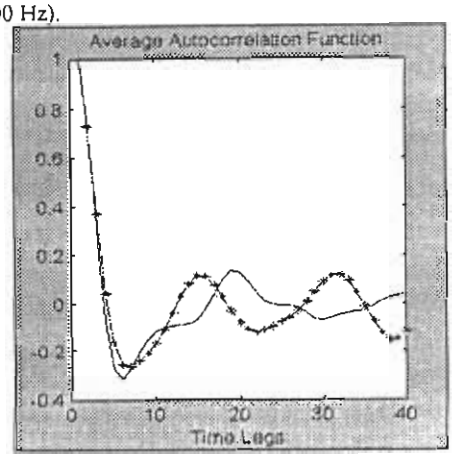

Fig. 1. Typical average autocorrelation functions of the EMG signal recorded during 50 % MVC for normal subject (—) and myopathic subject (•••). The time interval between lags is 1ms.

III. AR MODELING OF SEMG

In AR model each sample *emg(n)* of SEMG is described as a linear combination of previous samples plus an error term $e(n)$ that is independent of past samples, that is:

$$emg(n) = -\sum_{k=1}^{p} a_k \cdot emg(n-k) + e(n) \qquad (1)$$

Where *emg(n)* is the output model signal, a_k are the AR coefficients, *e(n)* is the error sequence and p is the model order. The model represented by (1) can be used in a backward fashion (retrospective regression analysis); the signal at time n is considered as being the linear combination of p future values. The system function is

$$H(z) = \frac{1}{1+\sum_{k=1}^{p} a_k \cdot z^{-k}} \qquad (2)$$

$H(z)$ contains poles only. Thus, the model can work only for signals with well defined peaky spectrum like speech and EEG, and can be fitted also to SEMG. The spectrum of the sequence *emg*(n) can be estimated from the model if we consider $|E(\omega)| = 1$ (white noise sequence), therefore, the spectrum of output signal equal the spectrum of H(z) and can be estimated by substituting $Z \rightarrow e^{-j\omega}$ as follows

$$S(\omega) = |EMG(\omega)|^2 = \frac{1}{\left|1 + \sum_{k=1}^{p} a_k \cdot e^{-j\omega k}\right|^2} \quad (3)$$

The AR coefficients (a_i) are calculated using the covariance method [19] which minimizes the residual energy $\sum_n e^2(n)$.

Spectrum of Surface EMG

The analysis was carried out on consecutive 500 ms segments of EMG signal to ensure the stationarity of the segment where SEMG was found to be stationary on segment length less than 0.64s [8]. Fig.2(a) demonstrates the spectrum of such segment calculated by FFT (fast Fourier transform) routine. A consistent feature of EMG spectra is the many spectral peaks in the region 10-200Hz. The higher frequency regions, above 200Hz, contains minor amplitude information compared to the lower frequency region. Fig.2(b) demonstrates the spectrum of the AR model of the same signal calculated using 20 coefficients. It can be seen here that instead of obtaining six dominant peaks only three are obtained.

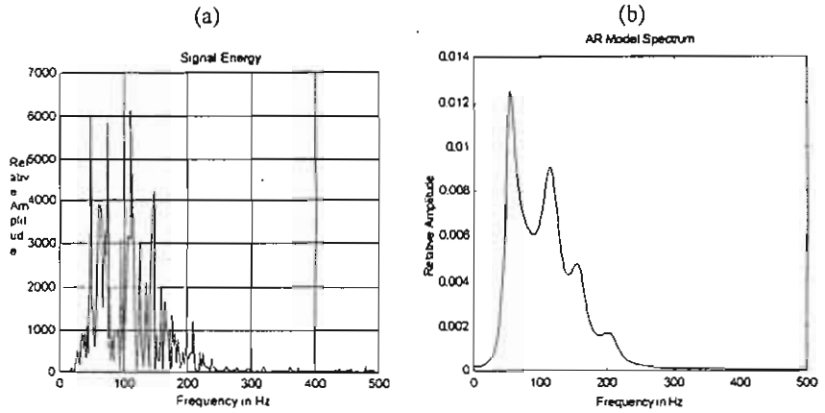

Fig.2. (a) Spectrum of Surface EMG segment Calculated by FFT. (b) AR model spectrum for the same segment with 20 coefficients calculated by covariance method.

The above model, instead of resolving nearby dominant frequencies, resolves frequencies in the high-frequency region that are minor to the dominant ones in the low region.

Fig 3. is an attempt to resolve the problem of dominant frequencies by increasing the model order where the low frequency peak is still not pronounced. The model also tends to resolve the minor frequencies. It became apparent that most of the spectral components lie below 200 Hz as can be seen in Fig.2(a) and Fig. 3.This came in addition to poor spectral envelope matching which resulted from the inclusion of a higher-frequency band. It was decided, therefore, to disregard frequency components above 200 Hz and to reduce the bandwidth of the signal to one-half. The signals were filtered at 230 Hz and resampled at 500 Hz.

Fig.4(a) demonstrates the FFTspectrum of the filtered signal at 230 Hz. Fig. 4(b) illustrates the spectrum of the all-pole 20 coefficients model with the EMG segment. It can be seen how most of the dominant frequencies are resolved and a much better spectral envelope fit is obtained. Therefore, a value of $p = 20$ was chosen to characterize surface EMG segments as a compromise between model size and accuracy of signal representation. After characterizing each segment with 20 coefficients, it was observed that after calculating the spectra of consecutive segments, the signal had a changing spectrum. The number of segments

characterizing each record was calculated as follows. The average of each coefficient was calculated as a function of segment number and its convergence was observed. It was seen that after ten segments the coefficients started to converge. Therefore, ten segments each 500 ms long were chosen to characterize the signal.

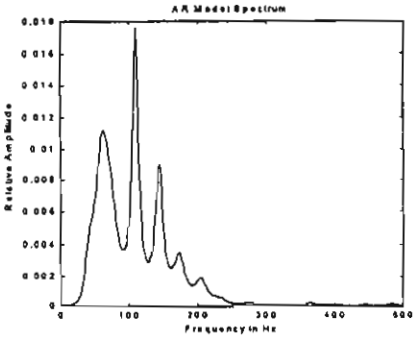

Fig. 3. AR model spectrum of Fig. 2(a) with 30 coefficients

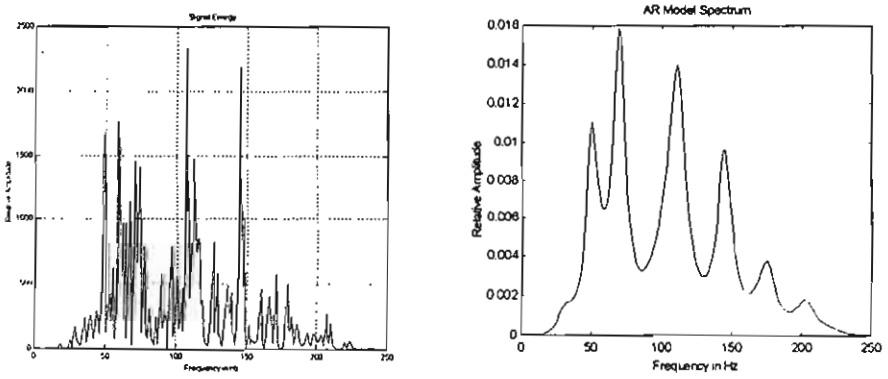

Fig. 4 (a) The FFT spectrum of the EMG signal sampled at 500Hz (b) The AR model spectrum of the same EMG segment with 20 coefficients calculated using the covariance method.

IV. EMG CLASSIFICATION

The classification approaches taken here are: the old Fisher linear discriminant [20] and artificial neural network algorithms. Two paradigms for training the ANN were investigated, supervised and unsupervised. For supervised learning, the well-known back propagation algorithm [21, 22] and for unsupervised learning, the self-organizing feature maps algorithms [23, 24] were implemented. A comparison of the performance of the different classifiers is performed for normal and abnormal individuals.

A. Using Fisher Linear Discriminant Algorithm

The first step in the diagnosis procedure is the construction of the Fisher linear discriminant vector W where the two different classes are introduced as in the following equation [24]:

$$W = C_t^{-1}(\mu_1 - \mu_2) \qquad (4)$$

where μ_1, μ_2 are the means of the two classes $\Im_1, \Im_2$, and C_t is the total *within-class covariance matrix* defined by

$$C_t = \sum_{n \in \Im_1} (X_n - \mu_1) \cdot (X_n - \mu_1)^T + \sum_{n \in \Im_2} (X_n - \mu_2) \cdot (X_n - \mu_2)^T$$

After calculating the vector W, the projection of every subject that participate in the construction of the vector is first calculated and then the projection threshold between the two different classes is determined. After calculating this threshold, the projection of unknown subjects is founded and classified in accordance with the above threshold.

It has been found that the results are dependent on the number of participants in the construction of the Fisher linear discriminant vector FLDV. More participants mean better definition of the threshold between classes, and thus better classification. Using nine normals and nine abnormals to construct of the FLDV, a 60% percent successful classification was obtained for a test a group of five normals and five abnormals.

B. Using Back-Propagation Neural Network

The number of input nodes is 200 using the 20 AR coefficients for 10 SEMG segments, and the number of output is one output node, where the output is corresponding to the two classes, Normal and Abnormal. The number of nodes in the hidden layers is changed (3 - 10 nodes) in order to determine the optimum number of nodes. A scaling method is used to scale the input patterns to give every pattern the same importance.

The results of classification using the back-propagation neural network trained with three different backpropagation algorithms are summarized in Tables 1-3. ANN architectures with three layers (input layer, hidden layer, and output layer with one output node) were used [21, 22]. The ANN architectures are expressed as strings showing the number of inputs, the number of nodes in the hidden layers, and the number of nodes in the output layer. The number of weights, and the training time are tabulated for all models. During the training phase, an error measure of the closeness of weights to a solution can be calculated for each pattern that represents a subject in the training set. This measure is used for determining whether a certain subject has been learned by the system, and is defined by

$$PSS = \sum_{l=1}^{M} (y_l - d_l)^2 \qquad (5)$$

where *PSS*: Pattern sum squares.

M: Number of output nodes (one in this case)

y_l: Calculated output.

d_l: Desired output.

The PSS measure is then summed over all patterns to get the total sum of squares or TSS measure

$$TSS = \sum_{m=1}^{p} \sum_{l=1}^{M} (y_l - d_l)^2 \qquad m = 1, \cdots, p \qquad (6)$$

where p is the number of training patterns (18 in this case; 9 normals and 9 abnormals).

The average error estimated for the output node

$$EE = (TSS / p)^{0.5} \qquad (7)$$

For comparing the results that were obtained by various classification algorithm, common performance metrics have been used [19]. For a given decision suggested by the output neuron, four possible alternatives exist: true positive (TP), false positive (FP), true negative (TN), and false negative (FN). TP decision occurs when the positive diagnosis of the algorithm coincides with a positive diagnosis according to the physician. An FP decision decision occurs when the algorithm made a positive diagnosis that does not coincide the physician. A TN decision occurs when the algorithm and the physician suggest the absence of a positive diagnosis. An

FN decision occurs when the algorithm made a negative diagnosis that does not agree with the physician. From the above measures, correct classification percentage (%CC) has been calculated for the N cases in the evaluation set:

$$\%CC = 100 \times (TP + TN)/N \quad (8)$$

Important issues that characterize the overall performance of the back-propagation algorithms during the training procedure are:

1) The output of all the backpropagation EMG models is limited between one and zero range, so the selection of sigmoidal activation function is preferred. The hidden layer nodes activation functions were also set to sigmoidal function.

Table 1. The results of EMG classification using Neural Network Back-Propagation trained with Variable Learning rate

Model	*Architecture*	*Weights*	*Epochs*	*EE*	*Time (Seconds)*	*Training %CC*	*Evaluation %CC*
1	200-3-1	603	323	10^{-5}	90	100%	70%
2	200-5-1	1005	328	10^{-5}	60	100%	80%
3	200-10-1	2010	285	10^{-5}	60	100%	80%

Table 2. The results of EMG classification using Neural Network Back-Propagation trained with Variable Learning rate (early stopping technique)

Model	*Architecture*	*Weights*	*Epochs*	*EE*	*Time (Seconds)*	*Training %CC*	*Validation %CC*
1	200-3-1	603	188	0.0047	30	90%	90%
2	200-5-1	1005	118	0.0532	30	90%	90%
3	200-10-1	2010	148	0.0067	45	100%	90%

Table 3. Neural Network Back Propagation EMG Models Trained Using Conjugate Gradient Method with Early Stopping Technique to Improve Generalization.

Model	*Architecture*	*Weights*	*Epochs*	*EE*	*Time (seconds)*	*Training %CC*	*Validation %CC*
1	200-3-1	603	28	2.7×10^{-11}	25	100%	80%
2	200-5-1	1005	18	3×10^{-6}	20	100%	90%
3	200-10-1	2010	15	5.4×10^{-5}	20	100%	90%

2) At present, no method other than empirical has been proposed for choosing the architecture of feedforward NN so that for every training algorithm three architectures were created and compared.
3) Training all the NN models is accomplished using batch training.
4) Networks of Table 1 are trained using a variable learning rate algorithm. The first model architecture is insufficient to generalize the network where *Training % CC = 100* but *Evaluation %CC = 70*. Model 2 and 3 are nearly suitable for generalization of network. If we attempt to move towards high architecture network we will over fit the data and the network will be less generalized. The value of TSS will determine the generalization performance of network, high *EE* value will lead to high *Training %CC* but less generalization where the network tend to memorize the training data, small *EE* value will lead to best generalization but may be less *Training %CC*.
5) All Models of Table 2 are trained using variable learning rate algorithm with early stopping technique to determine the optimum value of EE. The available records are subdivided into three sets: training set, and validation set. The problem of determining the suitable network architecture is removed using this technique. The number of epochs is reduced using early stopping technique where the training will continue until validation test failure. In training the networks there is no sufficient records to form an evaluation data set but with the first two sets the basic idea of early stopping technique is clarified.
6) Table 3 is a backpropagation models trained with *Conjugate Gradient Algorithm*. The Conjugate Gradient Training is more fast than old variable learning rate algorithm and is relatively suitable to large size networks than other fast algorithms. Therefore, it was found that it gives the highest performance for the present application.

C- Using Self-Organizing Feature Map

The neural network models in this system were derived using the Kohonen's self-organizing feature maps algorithm [25]. With this algorithm the training process involves the presentation of pattern vectors from the training set one at a time. A wining neuron (node) is selected in a systematic way after all input vectors are presented. A weight adjustment process takes place by using neighborhood concept that shrinks over time and learning coefficient that also decreases with time. After several input vectors are presented, weights will form clusters or vector centers that sample the input space such that the point density function of the vector centers tends to approximate the probability density function of the input vectors [25]. The weights will also be organized such that topologically close output nodes are sensitive to inputs that are physically similar. Thus, the output nodes will be ordered in a natural way.

The results of the self-organizing feature map models that were investigated with no preprocessing of the 200 input feature vector are summarized in Table 4. Models with output grid size 4×4, 6×6, 8×8, and 10×10 were developed. Initial gain factor η was selected to equal 0.9. Training for self-organizing feature map EMG models was carried out for 1000 epochs.

Table. 4 Self-Organizing Feature Map EMG Models

Model	*No. of Inputs*	*No. of Classes*	*Output Grid*	η	*Epochs*	*Time (Seconds)*	*Training %CC*	*Evaluation %CC*
1	200	2	4×4	0.9	1000	75	94.4%	60%
2	200	2	6×6	0.9	1000	140	100%	80%
3	200	2	8×8	0.9	1000	270	100%	80%
4	200	2	10×10	0.9	1000	440	100%	80%

At each training cycle (epoch), the 18 input patterns were presented at random. It was observed that grid sizes below 6×6 were inadequate for producing models with well-separated classes. All models succeed to classify all patterns during training phase except model 1.

The procedure that was followed for assigning normal and abnormal classes to the self-organizing feature maps is presented here.

For every 200 element feature vector

$$x_n = [x_{1,n}, x_{2,n}, \cdots, x_{200,n}]^T \qquad , n = 1, 2, \cdots, N \qquad (9)$$

where N is the number of patterns in the training set (N=18 patterns), there is an output node at the grid for which maximum response R_{max} is caused. This node is assigned the class number of the vector(1 → normal, 0 → abnormal).

Fig. 5(a) shows the nodes where maximum response was caused by the subjects (patterns) in the training set after the completion of the training phase.

0	1	0	1	0	1
0	x	0	1	x	0
x	1	x	x	x	x
1	x	0	1	x	1
x	0	x	x	0	x
x	x	x	1	x	x

0	1	0	1	0	1
0	0	0	1	0	0
1	1	0	1	0	0
1	0	0	1	1	1
0	0	0	1	0	1
0	0	1	1	1	1

Fig.5 Self-Organizing feature maps. (a) Maximum response map after training phase (100%)
(b) Maximum Response with all nodes assigned.

Nodes with "x" values have not been assigned to any class. For this model where the output grid is 6×6 (36 output nodes) with a training set of 18 patterns, at least 18 nodes will not be assigned to any pattern. This means that unknown patterns falling on "x" nodes will not be diagnosed.

During the next phase, the "x" nodes are assigned to one of the classes as follows: the data of each subject in the training set is applied at the input, and the response at a certain "x" node is observed. The class of the subject that cause maximum response at the node is assigned to the node. This procedure is applied for all the "x" nodes until all of them are assigned to a class as shown in the figure 5(b). Fig.6(a) shows the two classes boundaries for all the nodes in the grid. Fig. 6(b) shows the grid after evaluation phase using the test group: five normal and five abnormal subjects which yield 80% correction classification (%CC).

The self-organizing feature maps system compared to the back propagation neural network system has the advantage of the results being presented pictorially. For example, with this system one can relate a certain patient with another patient, find boundary cases, and observe the mapping of a patient over serial examinations. Training effort for self-organizing feature map models was significantly reduced as compared to the back propagation models.

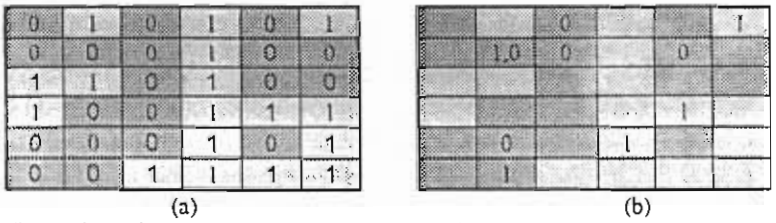

Fig 6 Self-Organizing feature maps. (a) Simplified map showing the two classes boundaries
(b) Maximum response map for the evaluation set (%CC = 80).

D- Using Probabilistic Neural Network

The Probabilistic Neural Network (PNN) is a Baysian classifier put into a neural network architecture [26, 27]. It can be used as a function approximator like the back propagation NN.

The Probabilistic Neural Network should be used only for classification problems where there is a representative training set. It can be trained quickly but has slow recall and is memory intensive. It has solid underlying theory and can produce confidence intervals. This network is simply a Baysian Classifier put into a neural network architecture. The PNN depends on estimation of probability density function for every class using the Gaussian weighting function:

$$g(x) = \frac{1}{n}\sum_{i=1}^{n} e^{-\frac{|x-x_i|^2}{2\sigma^2}} \qquad (10)$$

where n is the number of cases in class, x_i is a specific case in a class, x is the input pattern to be classified, and σ is the width parameter

This formula simply estimates the probability density function (PDF) as an average of separate multivariate normal distributions. This function is used to calculate the probability density function for each class.

The suggested probabilistic neural network consists of three layers. When an input is presented, the first layer computes distances from the input vector to the training input vectors, and produces a vector whose elements indicate how close the input is to a training input. The second layer sums these contributions for each class of inputs to produce as its net output a vector of probabilities. Finally, a compete transfer function on the output of the second layer picks the maximum of these probabilities, and produces a one for that class and a zero for the other classes.

Training the probabilistic neural network with no preprocessing 18 training patterns with different sigma (spread) and ten evaluation patterns set is shown in Table 5.

The effect of the sigma (spread coefficient) on the performance of PNN is sounded in the above table where other feature work may be looking for a training algorithm which train the spread coefficient to the optimum value like the weights of the networks.

Table 5. Results of using the probabilistic neural network classifier

Spread(σ)	0.1	0.5	1	1.5	2
%CC	70%	70%	90%	70%	70%

V. DISCUSSION AND CONCLUSION

An attempt to characterize surface EMG for clinical classification was made. It has been demonstrated, that enough information remains in the recorded surface EMG to allow its usage in clinical classification. Surface EMG was found to have a changing spectra nature with a considerable variance. To overcome the variability nature, a number of segments were chosen to describe one subject. The AR method was selected to characterize the signal since it reduces the dimensionality of spectral characterization. The covariance method was utilized for the parameters estimation process due to its less biased estimation than the autocorrelation method.

Artificial Neural network (ANN) diagnosis models in conjunction with parametric analysis provide an integrated solution to the problem of automated EMG evaluation. This approach is very desirable because it minimizes observer bias, facilitates comparisons of results across individual and different methodologies, and more importantly, provides useful information for helping the physician in reaching a more accurate diagnosis using surface electrodes instead of needle electrodes. The SEMG's used were recorded from normal individuals and neuromuscular disorders (myopathy).

Three paradigms of ANN were investigated: multilayer backpropagation algorithm, self-organizing feature map algorithm and a probabilistic neural network model. The performance of the three classifiers was compared with that of the old Fisher linear discriminant (FLD) classifiers. The results have shown that the three ANN models give higher performance. Poorer diagnostic performance was obtained from the FLD classifier. The percentage of correct classification reaches 90% using the backpropagation multilayer algorithm.

The system presented here indicates that surface EMG, when properly processed, can replace needle EMG in some clinical applications and can be used to provide the physician with a diagnostic assist device. The amount of data collected until now is insufficient for making significant conclusions concerning the accuracy of classification. Yet it is suggested that the methods adopted have the potential of becoming an effective diagnostic device. This is only the first stage of a project aiming at building an expert system for SEMG. Adding new cases and new types of diseases seems to be next and necessary steps.

REFERENCES

[1] De Luca, C. J., "Physiology and Mathematics of Myoelectric Signals," IEEE Trans. Biomed. Eng., vol.BME-26, pp.313-325, 1979.

[2] Richfield, E. K., Cohen, B. A., and Albers, J. W., "Review of quantitative and automated needle electromyographic analyses," IEEE Trans. Biomed. Eng., vol.BME-28, pp.506-514, 1981.

[3] Coatrieux, J.-L., "On-line electromygraphic signal processing system," IEEE Trans. Biomed. Eng., vol.BME-31, pp.199-207, 1984.

[4] McGill, K. C., Cummins, K. L., and Dorfman, L. J., " Automatic decomposition of the clinical electromyogram," IEEE Trans. Biomed. Eng., vol.BME-32, pp.470-477, 1985.

[5] Clancy, E. A., and Hogan, H. " Single site electromygraph amplitude estimation," IEEE Trans. Biomed. Eng., vol.BME-41, pp 159-167, 1994.

[6] D'Alessio, T , Knaflitz, M., Balestra, G., and Paggi, S. "On-line estimation of myoelectric signal spectral parameters and nonstationarities detection," IEEE Trans. Biomed. Eng., vol.BME-40, pp.981-984, 1993.

[7] Tanzi, F. and Taglietti, V., "Spectral analysis of surface motor unit action potentials and surface interference electrommyogram," IEEE Trans. Biomed. Eng., vol.BME-28, pp.318-324, 1981.

[8] Inbar G. F , and Noujaim, A. E., "On Surface EMG Spectral Charactrization and its application to Diagnostic Classification," IEEE Trans. Biomed. Eng., vol.BME-31, pp.597-604, 1984.

[9] Kiryu T., Y. Saitoh, and K. Ishioka, "Investigation on Parametric Analysis of Dynamics EMG Signals by a Muscle-Structured Simulation Model," IEEE Trans. Biomed. Eng., vol.BME-39, pp.280-288, 1992.

[10] Graupe D. and Cline, W.K., "Functional Separation of SEMG Signals via ARMA identification Methods for Prosthesis Control Purpose, IEEE Trans. Syst. Man Cyber., vol.SMC-5, pp.252-259, 1975.

[11] Sherif M. B , Gregor, R. J., and Lyman, J. "Effects of load on Myoelectric Signals: The ARIMA Representation, IEEE Trans. Biomed. Eng., vol.BME-28, pp.411-416, 1981.

[12] Capponi, M., D'Alessio. T., and Laurenti, M., " Sequential estimation of power spectra for nonstationary myoelectric signals, Electromyogr. Clin. Neurophysiol., vol.28, pp.427-432, 1988.

[13] Kiryu T., DE Luca, C. J., and Y. Saitoh, "AR modeling of myoelectric interference signals during a ramp contraction," IEEE Trans. Biomed Eng., vol.BME-41, pp.1031-1038, 1994.

[14] Doerschuk, P.C , Gustafon, D. E., and Willsky, A. S., "Upper extrimity limb function discrimination using EMG signal analysis," IEEE Trans. Biomed. Eng., vol.BME-30, pp.18-29, 1983.

[15] Hefftner, G., Zucchini, W., and Jaros, G.G., " The electromyogram (EMG) as a control signal for functional neuromuscular stimulation-Part II:; practical demonstration of the EMG signature discrimination system," IEEE Trans. Biomed. Eng., vol.BME-35, pp.238-242, 1988.

[16] Coatrieux, J. L., Toulouse, P., Rouvrais, b., and Le Bars, R., "Automatic classification of electromyographic signals," EEG Clin. Neurophysiol., vol.55, pp.333-341, 1983.

[17] Hassoun M. H., Wang, C., and Spitzer, A. R., "Neural Network Extraction of Repetitive Vectors for Electromyography" Part I and Part II, IEEE Trans. Biomed. Eng., vol.BME-41, pp.1039-1061, 1994.

[18] Pattichis, C. S., Schizas, C. N., and Middleton, L. T., " Neural network models in EMG diagnosis," IEEE Trans. Biomed Eng., vol BME-42, pp.468-496, 1995.

[19] Kay, S. M Modern spectral estimation. Prentice Hall, Englewood Cliffs, New Jersey, 1988

[20] Dudu, R. aand Hart, P., Pattern classification and scene analysis. New York, Wiley Interscience, 1975

[21] Haykin S., Neural Networks. A Comprehensive Foundatioo. Prentice-Hall, Inc., 1998.

[22] Beale R. and Jacson, T. Neural Computing: An Introduction. Bristol, Adam Hilger, 1990.

[23] Kohonen T , "The self-organizing map," Proc IEEE, vol.78, no 9, pp 1464-1480, 1990

[24] Schalkoff, R. Pattern recognition: statistical, structural and neural approaches. John Wiley, N. Y , 1992.

[25] Aleksander, I And Morton, H. An introduction to neural computing. Chapman & Hall, London, 1992

[26] Streit Roy, L and Luginbuhl, T. X., "Maximum Likelihood Training of Probabilistic Neural Networks," IEEE Trans. on Neural Networks , September 1994.

[27] Specht, D. F , "Probabilistic neural networks and the polynomial adaline as complementary techniques for classification," IEEE Trans on Neural Networks , vol.NN-1, no.1, pp. 111-121, March 1990.