ELSEVIER

Journal of Neuroscience Methods 156 (2006) 360–367

JOURNAL OF NEUROSCIENCE METHODS

www.elsevier.com/locate/jneumeth

Classification of EMG signals using wavelet neural network

Abdulhamit Subasi [a,*], Mustafa Yilmaz [b], Hasan Riza Ozcalik [a]

[a] *Kahramanmaras Sutcu Imam University, Department of Electrical and Electronics Engineering, 46500 Kahramanmaraş, Turkey*
[b] *University of Gaziantep, Faculty of Medicine, Department of Neurology, 27310 Gaziantep, Turkey*

Received 5 January 2006; received in revised form 9 February 2006; accepted 3 March 2006

Abstract

An accurate and computationally efficient means of classifying electromyographic (EMG) signal patterns has been the subject of considerable research effort in recent years. Quantitative analysis of EMG signals provides an important source of information for the diagnosis of neuromuscular disorders. Following the recent development of computer-aided EMG equipment, different methodologies in the time domain and frequency domain have been followed for quantitative analysis. In this study, feedforward error backpropagation artificial neural networks (FEBANN) and wavelet neural networks (WNN) based classifiers were developed and compared in relation to their accuracy in classification of EMG signals. In these methods, we used an autoregressive (AR) model of EMG signals as an input to classification system. A total of 1200 MUPs obtained from 7 normal subjects, 7 subjects suffering from myopathy and 13 subjects suffering from neurogenic disease were analyzed. The success rate for the WNN technique was 90.7% and for the FEBANN technique 88%. The comparisons between the developed classifiers were primarily based on a number of scalar performance measures pertaining to the classification. The WNN-based classifier outperformed the FEBANN counterpart. The proposed WNN classification may support expert decisions and add weight to EMG differential diagnosis.

Keywords: Electromyography (EMG); Motor unit potential (MUP); Autoregressive method (AR); Wavelet neural networks (WNNs)

1. Introduction

More than 100 neuromuscular disorders that influence the spinal cord, nerves or muscles are present. Early finding and diagnosis of these diseases by clinical examination and laboratory tests is crucial for their management as well as their anticipation through prenatal diagnosis and genetic counselling. Such information is also valuable in research, which may lead to the understanding of the nature and eventual treatment of these diseases (Christodoulou and Pattichis, 1999). Motor unit morphology can be studied by recording its electrical activity, known as electromyography (EMG). In clinical EMG motor unit potentials (MUPs) are recorded using a needle electrode at slight voluntary contraction. The MUP reflects the electrical activity of a single anatomical motor unit. It represents the compound action potential of those muscle fibres within the recording range of the electrode. Features of MUPs extracted in the time domain such as duration, amplitude and phases proved to be very helpful in differentiating between muscle and nerve diseases with the duration measure being the key parameter used in clinical practice (Pattichis and Pattichis, 1999). With increasing muscle force, the EMG signal shows an increase in the number of activated MUPs recruited at increasing firing rates, making it difficult for the neurophysiologist to distinguish the individual MUP waveforms. EMG signal decomposition and MUP classification into groups of similar shapes give significant information for the assessment of neuromuscular pathology (Christodoulou and Pattichis, 1999).

Nevertheless, the measurement of the duration parameter is a complicated task depending on the neurophysiologist and/or the computer-aided method used. The description of an extensively accepted criterion that will allocate the computer-aided measurement of this parameter is still absent (Stalberg et al., 1986). On the other hand, frequency domain features of MUPs like the mean or median frequency, bandwidth and quality factor give supplementary information for the assessment of neuromuscular disorders and it has recently been shown that the discriminative power of the MUP mean or median frequency is comparable to the duration measure (Pattichis and Elia, 1999) or the spike duration measure (Pfeiffer and Kunze, 1993). Recent advances in computer technology have made automated EMG analysis feasible. Although a number of computer-based quantitative EMG

* Corresponding author. Tel.: +90 344 219 12 53; fax: +90 344 219 10 52.
E-mail address: asubasi@ksu.edu.tr (A. Subasi).

0165-0270/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.jneumeth.2006.03.004

analysis algorithms have been developed, some of them are commercially available, practically none of them have gained broad acceptance for widespread routine clinical use. Pattichis and Elia (1999) used autoregressive and cepstral analyses combined with time domain analysis in classification of EMG signals. Also De Michele et al. (2003) described how the proper use of the wavelet cross-correlation analysis on surface signals of the above two different muscles allows a more comprehensive classification of subjects and, at the same time, a reliable temporal evolution analysis of Parkinson's disease.

Pattichis et al. (1995) used MUP parameters as input to a sequential parametric pattern recognition classifier. Loudon et al. (1992) used eight MUP features as input to a statistical pattern recognition technique for classification. The decomposition of superimposed waveforms used a combination of procedural and knowledge-based methods. Finally Hassoun et al. (1994a,b) proposed a system called neural network extraction of repetitive vectors for electromyography (NNERVE), and they used the time domain waveform as input to a three-layer artificial neural network (ANN) with a "pseudo unsupervised" learning algorithm for classification. Christodoulou and Pattichis (1999) used two different pattern recognition techniques for the classification of MUPs. They used an artificial neural network (ANN) technique based on unsupervised learning, using a modified version of the self-organizing feature maps (SOFM) algorithm and learning vector quantization (LVQ) and a statistical pattern recognition technique based on the Euclidean distance. In addition, Schizas and Pattichis (1997) used genetics-based machine learning as pattern classifiers in EMG. There are numerous limitations in the existing quantitative EMG analysis methods, which limit their wider applicability in clinical practice.

The theory of wavelets can be exploited in understanding the universal approximation properties of wavelet neural networks (WNNs), and in providing initialization heuristics for fast training. WNNs offer a good compromise between robust implementations resulting from the redundancy characteristic of non-orthogonal wavelets and neural systems, and efficient functional representations that build on the time–frequency localization property of wavelets (Daubechies, 1992; Sureshbabu and Farrell, 1999; Xu and Ho, 2002). Much research has been done on applications of WNNs, which combine the capability of artificial neural networks in learning from processes and the capability of wavelet decomposition (Zhang and Benveniste, 1992; Pati and Krishnaparasad, 1993), for identification and control of dynamic systems (Sureshbabu and Farrell, 1999; Zhang et al., 1995; Wong and Leung, 1998). Zhang and Benveniste (1992), proposed a new notation of wavelet network as an alternative to feedforward neural networks for approximating any square-integrable non-linear functions based on the wavelet transform theory, and a backpropagation algorithm is adopted for wavelet network training. Zhang et al. (1995) described a wavelet-based neural network for function learning and estimation, and the structure of this network is similar to that of the radial basis function network except that the radial functions are replaced by orthonormal scaling functions. In addition, Zhang (1997) presented wavelet network construction algorithms for the purpose of non-parametric regression estimation.

This paper presents a comprehensive investigation of the practicality of using an AR model and WNN to extract classifiable features from EMG. Here, AR power spectral density (PSD) was used to define EMG signal representations. A variety of features based on this model were classified with wavelet neural network. The system is intended to decompose EMG signals at low to moderate force levels. The proposed techniques were successfully applied in the classification and decomposition of EMG signals recorded from normal (NOR) subjects and subjects suffering from myopathy (MYO) and neurogenic (NEU) disorder. By using AR PSD and WNN resulted in the best classification percentages than FEBANN method.

2. Materials and method

2.1. Subjects and data acquisition

All the measurements from patients and control group were done in Neurology Department of University of Gaziantep. Diagnostic criteria for the subjects selected were based on clinical findings; on the other hand, if it is required, muscle biopsy was performed. Normal, myopathic and neurogenic subjects were evaluated by expert doctors. All the EMG data, collected from 27 subjects have been analyzed. Data were recorded from 7 healthy subjects (three males and four females) with ages ranging from 10 to 43 years (mean age ± standard deviation (S.D.): 30.2 ± 10.8 years), 7 myopathic subjects (four males and three females) with ages ranging from 7 to 46 years (mean age ± standard deviation (S.D.): 21.5 ± 13.3 years) and 13 neurogenic subjects (eight males, five females) with ages ranging from 7 to 55 years (mean age ± standard deviation (S.D.): 25.1 ± 17.2 years).

An EMG system (*Keypoint*; *Medtronic Functional Diagnostics*, *Skovlunde*, *Denmark*) with standard settings was used. The EMG signal was acquired from the biceps brachii muscle using a concentric needle electrode (0.45 mm diameter with a recording surface area 0.07 mm^2; impedance at 20 Hz below 200 kΩ). At least 20 different MUPs were obtained from each muscle via five to seven muscle insertion. Between two sites, the needles were withdrawn for at least 5 mm. The position of the needle near to active muscle fibers was guided by acoustic and visual control of the EMG signal. The EMG signal was recorded at force levels approximately 30% of maximum voluntary contraction (MVC) under isometric conditions. The signal was acquired for 5 s, bandpass filtered at 5–10 kHz, and sampled at 20 kHz with 12-bit A/D resolution. The EMG signal was then low-pass filtered at 2 kHz (Fig. 1).

2.2. Autoregressive modelling of EMG signals

The AR method consists of modelling the EMG signal as the output of a linear filter driven by a white noise. This filter, referred to as AR, is a linear combination of the previous samples (regressive) of the output itself (auto). The equation of a classical

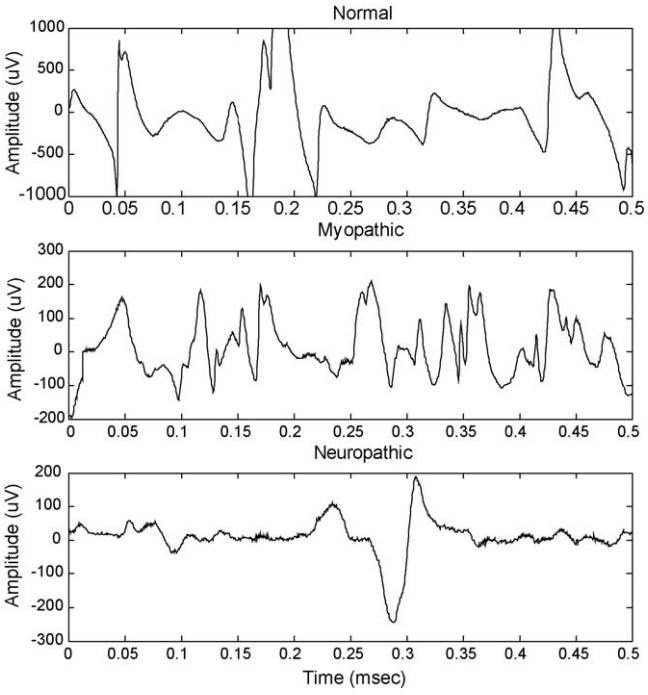

Fig. 1. Normal, myopathic and neurogenic EMG signals.

AR complex process, in a non-stationary context, is given by:

$$x(n) = -\sum_{i=1}^{p} a_i x(n-i) + e(n) \quad (1)$$

where $a_i(n)$ are AR complex parameters, p the AR model order, $e(n)$ is a white complex noise and n is the sample time. The main steps involved in the spectrum estimation procedure using AR modelling are as follows: (a) optimal model order determination; (b) AR parameters estimation; (c) estimating the power spectral density using the above parameters. Determination of the optimal AR model order is an important part of the whole procedure since too low a model order tends to smooth the actual spectrum and too high order tends to introduce spurious peaks in the power spectrum. One of the commonly used methods for determining model order is called the Akaike information criterion (AIC) (Akaike, 1974). It is based on minimizing an information theoretic criterion (Proakis and Manolakis, 1996; Guler and Ubeyli, 2003; Subasi et al., 2006). One of the most popular methods for estimating the AR parameters of a sequence of N data points is using Burg's algorithm (Tseng et al., 1995; Proakis and Manolakis, 1996; Pardey et al., 1996). In this, the AR parameters are estimated using a constrained least squares minimization procedure. This is a statistically accurate model when the amplitudes of the signal have a zero mean Gaussian distribution. However, it can be used to a good approximation even when the signals have a different distribution. It assumes that the data sequence is stationary i.e. the first and second order statistics of the sequence $x(n)$ do not change with time. The spectrum estimation method based on autoregressive (AR) modelling yields better resolution without the problem of spectral 'leakage' (Muthuswamy and Thakor, 1998; Guler et al., 2001). A very good tutorial on the theoretical foundations of the different

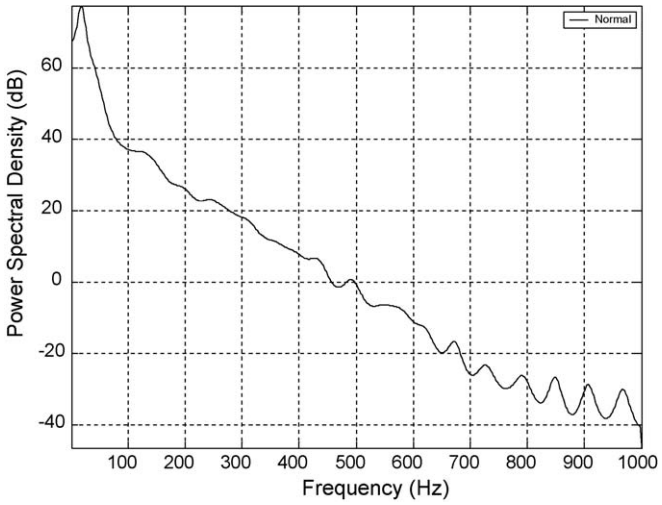

Fig. 2. AR Burg spectrum of normal EMG signal.

methods of spectrum estimation can be found in Kay and Marple (1981).

The estimation problem of these parameters covers the formation and resolution of the set of equations, which can easily be calculated. AR spectral estimation of data's estimation from the order p was given with the following equation (Isaksson et al., 1981).

$$P(f) = \frac{\sigma p^2 \, \Delta t}{\left|1 + \sum_{i=1}^{p} a_{p_i} \, \mathrm{e}^{-j2\pi f i \, \Delta t}\right|^2} \quad (2)$$

where $a_{p_0} = 1$. Thus, for estimation of spectral power density of EMG signal only existence of p number a_{p_i} parameters, and σ_p^2 parameters, which are the variances of white noise, are sufficient AR coefficients, which identify the amplitude rates. Fig. 2 shows an AR Burg spectrum of normal EMG signal, Fig. 3 shows an AR Burg spectrum of myopatic EMG signal and Fig. 4 shows an AR Burg spectrum of neurogenic EMG signal.

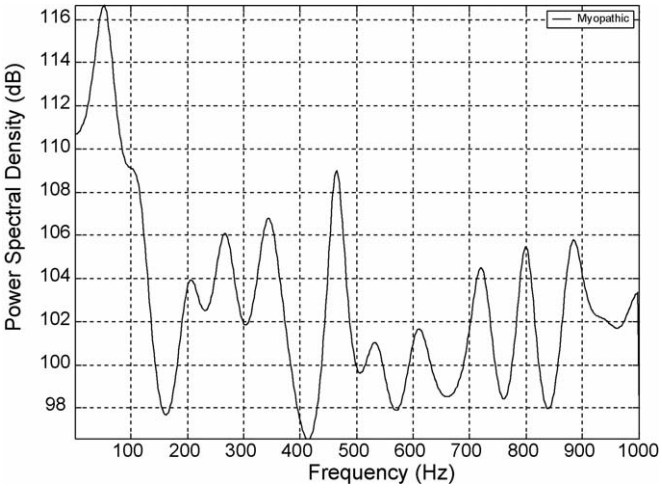

Fig. 3. AR Burg spectrum of myopathic EMG signal.

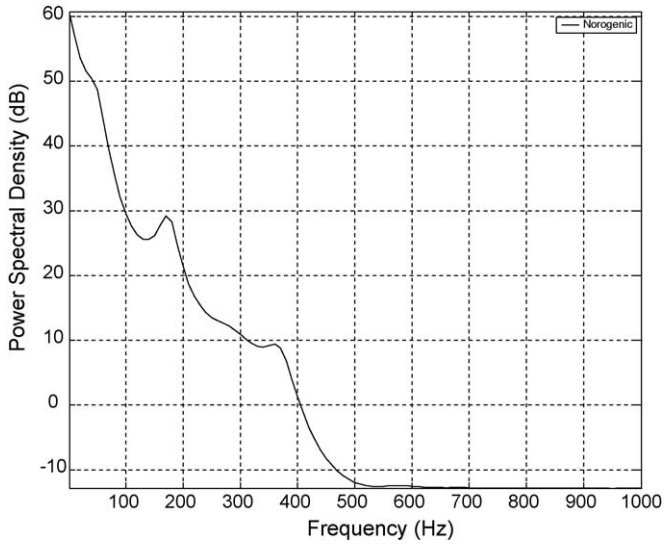

Fig. 4. AR Burg spectrum of neurogenic EMG signal.

2.3. Classification using artificial neural networks

Artificial neural networks (ANNs) are formed of cells simulating the low-level functions of biological neurons. In ANN, knowledge about the problem is distributed in neurons and connections weights of links between neurons. The neural network must be trained to adjust the connection weights and biases in order to produce the desired mapping. At the training stage, the feature vectors are applied as input to the network and the network adjusts its variable parameters, the weights and biases, to capture the relationship between the input patterns and outputs. ANNs are particularly useful for complex pattern recognition and classification tasks. The capability of learning from examples, the ability to reproduce arbitrary non-linear functions of input, and the highly parallel and regular structure of ANN make them especially suitable for pattern classification tasks (Fausett, 1994; Haykin, 1994; Basheer and Hajmeer, 2000).

ANNs are widely used in the biomedical field for modelling, data analysis and diagnostic classification (Alkan et al., 2005; Subasi, 2005; Subasi and Ercelebi, 2005). The most frequently used training algorithm in classification problems is the back-propagation (BP) algorithm, which is used in this work also. There are many different types and architectures of neural networks varying fundamentally in the way they learn; the details of which are well documented in the literature (Dreiseitl and Ohno-Machado, 2002). In this paper, two neural networks relevant to the application being considered, i.e., classification of EMG data will be employed for designing classifiers; namely the FEBANN and WNN.

2.3.1. Wavelet neural networks

Wavelet neural networks (WNN) is a new network based on wavelet transform (Pati and Krishnaparasad, 1993), in which discrete wavelet function is used as the node activation function. Because the wavelet space is used as characteristic space of pattern recognition, the characteristic extraction of signal is realized by weighted sum of inner product of wavelet base and signal vector. Furthermore, because it combines the function of time–frequency localization by wavelet transform and self-studying by neural network, the network possesses doughty capacity of approximate and robust. In this paper, a WNN was designed with mono-hidden-layer forward neural network with its node activation function based on dyadic discrete Morlet wavelet basic function (Subasi et al., 2005). The Morlet transform was able to unambiguously locate three classes. Significantly, we were able to limit the set of interrogated scales to exclude those that correspond to structural features of the EMG. Note that the pattern of background noise in the Morlet was similar across all EMG signals.

Wavelet transforms have emerged as a means of representing a function in a manner that readily reveals properties of the function in localized regions of the joint time–frequency space. The applications of WNN are usually limited to problems of small input dimension. The main reason is that they are composed of regularly dilated and translated wavelets. The number of wavelets in the WNNs drastically increases with the dimension (Zhang, 1997). Some work has been done on reducing the size of the WNN by removing the redundant candidates (Wong and Leung, 1998; Xu and Ho, 2002). Galvao et al. (2004) worked recently on the use of wavelet neural networks as a non-linear regression structure using high dimensional data.

A wavelet $\psi_j(x)$ is derived from its mother wavelet $\psi(z)$ by the relation

$$\psi_j(x) = \psi\left(\frac{x - \boldsymbol{m}_j}{\boldsymbol{d}_j}\right) = \psi(z_j) \tag{3}$$

where the translation factor $\boldsymbol{m}_j$ and the dilation factor $\boldsymbol{d}_j$ are real numbers in $\Re$ and $\Re^*$, respectively. The family of functions generated by ψ can be described

$$\Omega_c = \left\{ \frac{1}{\sqrt{\boldsymbol{d}_j}}\psi\left(\frac{x - \boldsymbol{m}_j}{\boldsymbol{d}_j}\right), \quad \boldsymbol{m}_j \in \Re \quad \text{and} \quad \boldsymbol{d}_j \in \Re^* \right\} \tag{4}$$

A family Ω_c is said to be a frame of $L^2(\Re)$ if there exists two constants $c > 0$ and $c < +\infty$ such that for any square integrable function f, the following inequalities hold:

$$c||f||^2 \leq \sum_{\substack{j \\ \psi_j \in \Omega_c}} |\langle \psi_j, f \rangle|^2 \leq C||f||^2 \tag{5}$$

where $||f||$ denotes the norm of function f and $\langle f, g \rangle$ the inner product of functions f and g. Families of wavelet frames of $L^2(\Re)$ are universal approximators.

For the modelling of multivariable processes, multidimensional wavelets must be defined. In the present work, we use multidimensional wavelets constructed as the product of N_i scalar wavelets (N_i being the number of variables)

$$\Psi_j(x) = \prod_{k=1}^{N_i} \psi(z_{jk}) \quad \text{with} \quad z_{jk} = \frac{x - \boldsymbol{m}_{jk}}{\boldsymbol{d}_{jk}}, \tag{6}$$

where $\boldsymbol{m_j}$ and $\boldsymbol{d_j}$ are the translation and dilation vectors, respectively. Families of multidimensional wavelets generated according to this scheme have been shown to be frames of $L^2(\Re^{N_i})$

(Oussar et al., 1998). Wavelet networks were presented in the framework of static modelling architecture, where the network output y is computed as

$$y = y(x) = \sum_{j=1}^{N_w} c_j \Psi_j(x) + \sum_{k=0}^{N_j} b_k x_k. \quad (7)$$

It can be viewed as a network with an input vector of N_i components, a layer of N_w weighted multidimensional wavelets and a linear output neuron. The coefficients of the linear part of the networks would be called direct connections.

Wavelet network training consists in minimizing the usual least-squares cost function

$$J(\boldsymbol{\theta}) = \frac{1}{2}\sum_{n=1}^{N}(y_p^n - y^n)^2, \quad (8)$$

where vector $\boldsymbol{\theta}$ includes all network parameters to be estimated: translations, dilations, weights of the connections between wavelets and output and weights of the direct connections; N is the number of elements of the training set, y_p^n is the output of the process for example n and y^n is the corresponding network output.

In the framework of the discrete wavelet transform, a family of wavelets can be defined as

$$\Omega_d = \{\alpha^{m/2}\psi(\alpha^m x - n\beta), (m,n) \in Z^2\}, \quad (9)$$

where α and β are constants that fully determine, together with the mother wavelet ψ, the family Ω_d. Actually, relation (9) can be considered as a special case of relation (4), where

$$\begin{matrix} \boldsymbol{m}_j = n\alpha^{-m}\beta \\ \boldsymbol{d}_j = \alpha^{-m} \end{matrix}, \quad (10)$$

These relations show that, unlike the continuous approach, wavelet parameters cannot be varied continuously; therefore, gradient-based techniques cannot be used to adjust them. Generally, training wavelet networks stemming from the discrete transform (Zhang and Benveniste, 1992; Zhang, 1997) is performed using the Gram–Schmidt selection method. This approach usually generates large networks, which are less parsimonious than those trained by gradient-based techniques. This may be a drawback for many applications (Oussar et al., 1998).

2.3.2. Wavelet neural network classifier

Wavelets offer many attractive features for the analysis of physical signals, including universal approximation properties, robustness against coefficient errors (Daubechies, 1992), and joint input-space/frequency localization. Since EMG signals possess a combination of slow variations over long periods, with sharp, transient variations over short periods, WNNs seem to be a more natural choice than other mainstream neural networks for EMG analysis. A multidimensional wavelet $\Psi(z_{jk})$ can be induced from a scalar wavelet $\Psi(z)$ via an affine vector–matrix transformation of the input x (Zhang and Benveniste, 1992). We have introduced a variation to this idea for obtaining multidimensional wavelets that are radially symmetric with respect to

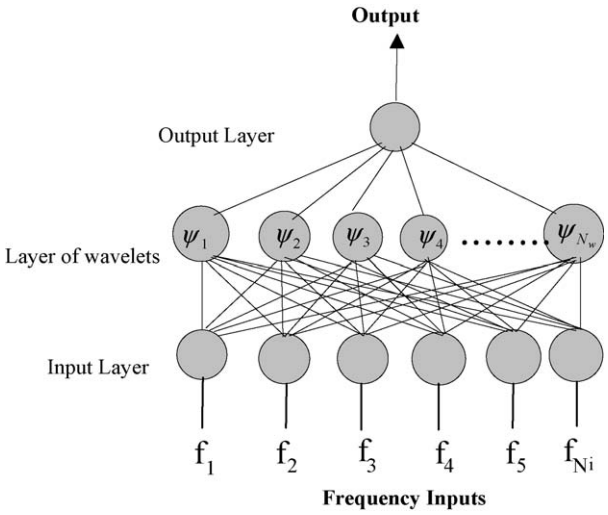

Fig. 5. A wavelet neural network model.

n-dimensional translation vectors $\boldsymbol{k}$, by dilating the Euclidean distance between input and translation vectors:

$$\psi(z_{jk}) = \sqrt{j^n}\psi(j||x - \boldsymbol{k}||), \quad j \le 0 \quad (11)$$

We are interested in the non-linearities introduced by the WNN only to the extent that they can improve a function approximation. Therefore, the WNN is equipped with a linear discriminant portion that can quickly account for a linear trend in the input/output data. The wavelet nodes are specifically trained to approximate only the "wave-like" components in the function. The resulting WNN classifier is shown in Fig. 5. For a detailed presentation of this method see Oussar and Dreyfus (2000) and Subasi et al. (2005).

2.3.3. Selection of network parameters

For solving the pattern classification problem, we used ANN employing back-propagation training algorithm. The advantages of this type of neural network are effective training algorithm and better understanding of the system behaviour. Selection of network input parameters and performance of neural network are important to distinguish between normal (NOR), myopathic (MYO) and neurogenic (NEU) subjects.

When using a neural network, decisions must be taken on how to divide data into a training set and a test set. In this study, 18 of 27 subjects were used for training and the rest of them were used for testing. In order to obtain a better network generalization five training subjects were used as cross-validation set.

The classification scheme of 1-of-C coding has been used for classifying the signal into one of the output categories. For each type of EMG signals, a corresponding output class is associated. The feature vector set, x represents the ANN inputs, and the corresponding class, once coded, constitutes the ANN outputs. In order to make the neural network training more efficient, the input feature vectors were normalized so that they fall in the range [0,1.0]. Since the number of output classes is 3, the ANN has three outputs, which produce a code for each class. The outputs are represented by basis vectors:

[0.9 0.1 0.1] = normal (NOR);
[0.1 0.9 0.1] = myopathy (MYO);
[0.1 0.1 0.9] = neurogenic (NEU).

Each dummy variable is given the value 0.1 except for the one corresponding to the correct category, which is given the value 0.9. Using target values of 0.1 and 0.9 instead of the common practice of 0 and 1 prevents the outputs of the network from being directly interpretable as posterior probabilities (Kandaswamy et al., 2004). The output vector associated to the modified input vector $\boldsymbol{x}_k$, $k = 1, 2, \ldots, K$ is noted $\boldsymbol{y}_k$, with K the number of EMG signals (Subasi, 2005).

2.3.4. Cross validation

Cross validation (CV) (Basheer and Hajmeer, 2000; Haselsteiner and Pfurtscheller, 2000) is often used for comparing two or more learning ANN models to estimate which model will perform the best for the problem at hand. With n-fold CV, the available data is partitioned into n disjoint subsets, the union of which is equal to the original set. Each learning model is trained on $n - 1$ of the available subsets, and then tested on the one subset, which was not used during training. This process is repeated n times, each time using a different test set chosen from the n available partitions of the training data, until all possible choices for the test set have been exhausted. The n test set scores for each learning model are then averaged, and the model with the highest average test set score is chosen as the one most likely to perform well on unseen data (Sureshbabu and Farrell, 1999).

2.3.5. Measuring error

Given a random set of initial weights, the outputs of the network will be very different from the desired classifications. As the network is trained, the weights of the system are continually adjusted to reduce the difference between the output of the system and the desired response. The difference is referred to as the error and can be measured in several ways. The most common measurement is SSE and MSE. SSE is the average of the squares of the difference between each output and the desired output (Fausett, 1994; Haykin, 1994; Basheer and Hajmeer, 2000; Haselsteiner and Pfurtscheller, 2000). In this study, mean squared error (MSE) was used for measuring performance of the neural network.

2.4. Development of ANN model

The purpose of the modelling phase in this application was to build up classifiers that are able to identify any input combination as belonging to either one of the three classes: normal, myopathic or neurogenic. For developing the neural network classifiers, 600 examples were randomly taken from the 900 examples and used for training the neural networks. The remaining 300 examples were kept aside and used for testing the validity of the developed models. The class distribution of the samples in the training and testing data set is summarized in Table 1.

The FEBANN was designed with AR spectrum of EMG signal in the input layer; and the output layer consisted of three nodes representing normal, myopathic or neurogenic disorder.

Table 1
Class distribution of the samples in the training and test data sets

Class	Training set	Test set	Total
Normal	300	100	400
Myopathic	300	100	400
Neurogenic	300	100	400
Total	900	300	1200

The beginning architecture of the network was examined using one and two hidden layers with a variable number of hidden nodes in each. It was found that one hidden layer is adequate for the problem at hand. Hence, the required network will contain three layers of nodes. The training procedure started with one hidden node in the hidden layer, followed by training on the training data (600 data sets), and then by testing on the validation data (300 data sets) to examine the network's prediction performance on cases never used in its development. Then, the same procedure was run repetitively each time the network was expanded by adding one more node to the hidden layer, until the best architecture and set of connection weights were obtained. Using the modified error–backpropagation algorithm for training, a training rate of 0.001 and momentum coefficient of 0.95 was found optimum for training the network with various topologies. The selection of the optimal network was based on monitoring the variation of error and some accuracy parameters as the network was expanded in the hidden layer size and for each training cycle. The mean of squares of error representing the mean of square of deviations of ANN solution (output) from the true (target) values for both the training and test sets was used for selecting the optimal network. A computer program that we have written for the training algorithm based on backpropagation of error was used to develop the FEBANNs.

According to the theory, the number of nodes in the hidden layer of the network is equal to that of wavelet base. If the number is too small, WNN may not reflect the complex function relationship between input data and output value. On the contrary, a large number may create such a complex network that might lead to a very large output error caused by over-fitting of the training sample. It was noticed that the best performance was obtained for the training set, validation test set, and separate test set with those models whose hidden layer had 50 neurons or more. Thus the optimum number of neurons required in the hidden layer is 50.

3. Results and discussion

The decomposition of real EMG signals into their constituent MUPs and their classification into groups of similar shapes is a typical supervised learning pattern classification problem. The number of MUP classes composing the EMG signal, the number of MUPs per class, and the shape of the MUP waveforms are unknown. The problem gets even more complex because of MUP waveform inconsistency, jitter of single fiber potentials and MUP superpositions. Any computerized method for EMG analysis should necessitate no operator involvement; should be fast, robust and reliable; and achieve high success rate in order

to be of clinical use. EMG data collected from 27 subjects were analyzed using the FEBANN and WNN. Data were recorded from 7 normal (NOR) subjects, 7 subjects suffering from myopathy (MYO) and 13 subjects suffering from neurogenic (NEU) disorder. Diagnostic criteria were based on clinical opinion, biochemical data or muscle biopsy. Only subjects with no history or signs of neuromuscular disorders were considered as normal.

3.1. Results of classification experiments

The computed AR spectrums were used as the inputs of the FEBANNs employed in the architecture of WNN. For each EMG signal frame (1024 samples) AR power spectral density were computed. In this application, there were three classes: healthy, myopathy and neurogenic. Classification results of the WNN were displayed by a confusion matrix. The confusion matrix showing the classification results of the WNN is given below.

Confusion matrix

Output/desired	Result (normal)	Result (myopathic)	Result (neurogenic)
Result (normal)	94	3	3
Result (myopathic)	8	92	0
Result (neurogenic)	11	3	86

According to the confusion matrix, 6 healthy subject was classified incorrectly by the WNN as a subject suffering from myopathy and neurogenic disorder, 14 neurogenic subject was classified as a normal or subject suffering from myopathy and 8 subject suffering from myopathy was classified as a normal.

The test performance of the WNN was determined by the computation of the following parameters:

Specificity: number of correct classified healthy subjects/number of total healthy subjects.
Sensitivity (*myopathy*): number of correct classified subjects suffering from myopathy/number of total subjects suffering from myopathy.
Sensitivity (*neurogenic*)*:* number of correct classified subjects suffering from neurogenic disorder/number of total subjects suffering from neurogenic disorder.
Total classification accuracy: number of correct classified subjects/number of total subjects.

The values of these statistical parameters are given in Table 2. As it is seen from Table 2, the WNN classified healthy subjects, subjects suffering from myopathy and subjects suffering from neurogenic disorder with the accuracy of 94%, 92%, 86%, respectively. The healthy subjects, subjects suffering from myopathy, and subjects suffering from neurogenic disorder were classified with the accuracy of 90.7%. The correct classification rates of the FEBANN were 91% for healthy subjects, 89% for subjects having myopathy and 84% for subjects having neurogenic disorder. Thus, the accuracy rates of the WNN presented for this application were found to be higher than that of the FEBANN.

Table 2
Comparison of FEBANN and WNN models for EMG signal classification

Statistical parameters	FEBANN (%)	WNN (%)
Specificity	91	94
Sensitivity (myopathic)	89	92
Sensitivity (neurogenic)	84	86
Total classification accuracy	88	90

Table 2 shows the classification success rate on 300 MUPs, obtained from EMG recordings. The classification success rate was defined as the percentage ratio of the correctly identified MUP classes by the classifiers and the number of true MUP classes present in the signal as identified by an experienced neurophysiologist. The average success rate for the WNN was 90.7% and for the FEBANN technique 88%. Examining the classification success rate for each class, the highest success rate was obtained for the NOR group and the lowest for the NEU group. This was the case for two ANN. The lowest success rate for the NEU group is attributed to the more complex and variable waveform shapes. In addition, as shown in Table 2, the WNN improved significantly the success rate for the NEU group compared to the FEBANN. In general, where two classification methods unsuccessful to categorize a MUP class, it was due to waveform variability. In some rare cases, MUP classes with very similar shapes were grouped together.

The testing performance of the neural network diagnostic system is found to be satisfactory and we think that this system can be used in clinical studies in the future after it is developed. This application brings objectivity to the evaluation of EMG signals and its automated nature makes it easy to be used in clinical practice. Besides the feasibility of a real-time implementation of the expert diagnosis system, diagnosis may be made more accurately by increasing the variety and the number of parameters. A “black box” device that may be developed as a result of this study may provide feedback to the neurophysiologists for classification of the EMG signals quickly and accurately by examining the EMG signals with real-time implementation.

4. Conclusions

The presented work for decomposing and classifying EMG signals is capable of extracting useful clinical information from simultaneously acquired EMG signals. Based on a set of EMG signals used for evaluation it was demonstrated that WNN classify the subjects with sufficient accuracy and speed to provide clinically useful parameter values relating to detailed aspects of the structure and function of the motor units of a muscle. An artificial neural network that classifies patients as having or not having neuromuscular disorders provides a valuable diagnostic decision support tool for physicians. We developed a wavelet neural network for classifying neuromuscular disorders of full spectrum EMG recordings. This novel method uses AR spectrum of EMG as the input to a wavelet neural network with three discrete outputs: normal, myopathic and neurogenic.

In conclusion, the pattern classification techniques as described in this work make possible the development of a fully automated EMG signal analysis system which is accurate,

simple, fast and reliable enough to be used in routine clinical environment. Future work will evaluate the algorithms developed in this study may be integrated into a hybrid diagnostic system for neuromuscular diseases based on ANN where EMG, muscle biopsy, biochemical and molecular genetic findings, and clinical data may be combined to provide a diagnosis.

References

Alkan A, Koklukaya E, Subasi A. Automatic seizure detection in EEG using logistic regression and artificial neural network. J Neurosci Meth 2005;148:167–76.

Akaike H. A new look at the statistical model identification. IEEE Trans Autom Control 1974;AC-19:716–23.

Basheer IA, Hajmeer M. Artificial neural networks: fundamentals, computing, design, and application. J Microbiol Meth 2000;43:3–31.

Christodoulou CI, Pattichis CS. Unsupervided pattern recognition for the classification of EMG signals. IEEE Trans Biomed Eng 1999;46(2):169–78.

Daubechies I. Ten lectures on wavelets. CBMS-NSF regional series in applied mathematics, vol. 61. Philadelphia: SIAM; 1992.

De Michele G, Sello S, Carboncini MC, Rossi B, Strambi S. Cross-correlation time–frequency analysis for multiple EMG signals in Parkinson's disease: a wavelet approach. Med Eng Phys 2003;25:361–9.

Dreiseitl S, Ohno-Machado L. Logistic regression and artificial neural network classification models: a methodology review. J Biomed Inform 2002;35:352–9.

Fausett L. Fundamentals of neural networks architectures, algorithms, and applications. Englewood Cliffs, NJ: Prentice Hall; 1994.

Galvao RKH, Becerra VM, Calado JMF, Silva PM. Linear-wavelet networks. Int J Appl Math Comput Sci 2004;14(2):221–32.

Guler I, Kiymik MK, Akin M, Alkan A. AR spectral analysis of EEG signals by using maximum likelihood estimation. Comput Biol Med 2001;31:441–50.

Guler I, Ubeyli ED. Application of classical and model-based spectral methods to ophthalmic arterial Doppler signals with uveitis disease. Comput Biol Med 2003;33:455–71.

Haselsteiner E, Pfurtscheller G. Using time-dependent neural networks for EEG classification. IEEE Trans Rehab Eng 2000;8:457–63.

Hassoun MH, Wang C, Spitzer AR. NNERVE: neural network extraction of repetitive vectors for electromyography—Part I: algorithm. IEEE Trans Biomed Eng 1994a;41:1039–52.

Hassoun MH, Wang C, Spitzer AR. NNERVE: neural network extraction of repetitive vectors for electromyography—Part II: performance analysis. IEEE Trans Biomed Eng 1994b;41:1053–61.

Haykin S. Neural networks: a comprehensive foundation. New York: Macmillan; 1994.

Isaksson A, Wennberg A, Zetterberg LH. Computer analysis of EEG signals with 4 parametric models. Proc IEEE 1981;69(4):451–61.

Kandaswamy A, Kumar CS, Ramanathan RP, Jayaraman S, Malmurugan N. Neural classification of lung sounds using wavelet coefficients. Comput Biol Med 2004;34(6):523–37.

Kay SM, Marple SL. Spectrum analysis—a modern perspective. Proc IEEE 1981;69:1380–419.

Loudon GH, Jones NB, Sehmi AS. New signal processing techniques for the decomposition of EMG signals. Med Biol Eng Comput 1992;Nov:591–9.

Muthuswamy J, Thakor NV. Spectral analysis methods for neurological signals. J Neurosci Meth 1998;83:1–14.

Oussar Y, Dreyfus G. Initialization by selection for wavelet network training. Neurocomputing 2000;34:131–43.

Oussar Y, Rivals I, Personnaz L, Dreyfus G. Training wavelet networks for nonlinear dynamic input–output modelling. Neurocomputing 1998;20:173–88.

Pati YC, Krishnaparasad PS. Analysis and synthesis of feedforward neural networks using discrete affine wavelet transformations. IEEE Trans Neural Networks 1993;4(1):73–85.

Pattichis CS, Pattichis MS. Time-scale analysis of motor unit action potentials. IEEE Trans Biomed Eng 1999;46(11):1320–9.

Pattichis CS, Elia AG. Autoregressive and cepstral analyses of motor unit action potentials. Med Eng Phys 1999;21:405–19.

Pattichis CS, Schizas CN, Mittleton LT. Neural network models in EMG diagnosis. IEEE Trans Biomed Eng 1995;42:486–96.

Pardey J, Roberts S, Tarassenko L. A review of parametric modelling techniques for EEG analysis. Med Eng Phys 1996;18(1):2–11.

Pfeiffer G, Kunze K. Frequency analysis and duration of motor unit potentials: reliability and diagnostic usefulness. Elect Clin Neurophrol 1993;89:365–74.

Proakis J, Manolakis D. Digital signal processing principles, algorithms, and applications. NJ: Prentice Hall; 1996.

Schizas CN, Pattichis CS. Learning systems in biosignal analysis. BioSystems 1997;41:105–25.

Stalberg E, Andreassen S, Falck B, Lang H, Rosenfalck A, Trojaborg W. Quantitative analysis of individual motor unit potentials: a proposition for standardized terminology and criteria for measurement. J Clin Neurophysiol 1986;3(4):313–48.

Subasi A, Erçelebi E, Alkan A, Koklukaya E. Comparison of subspace-based methods with AR parametric methods in epileptic seizure detection. Comput Biol Med 2006;36:195–208.

Subasi A, Alkan A, Koklukaya E, Kiymik MK. Wavelet neural network classification of EEG signals by using AR model with MLE preprocessing. Neural Networks 2005;18(7):985–97.

Subasi A, Ercelebi E. Classification of EEG signals using neural network and logistic regression. Comput Meth Prog Biomed 2005;78(2):87–99.

Subasi A. Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients. Exp Syst Appl 2005;28:701–11.

Sureshbabu A, Farrell JA. Wavelet-based system identification for nonlinear control. IEEE Trans Autom Control 1999;44(2):412–7.

Tseng S-Y, Chen R-C, Chong F-C, Kuo T-S. Evaluation of parametric methods in EEG signal analysis. Med Eng Phys 1995;17:71–8.

Wong K, Leung A. On-line successive synthesis of wavelet networks. Neural Process Lett 1998;7:91–100.

Xu J, Ho DWC. A basis selection algorithm for wavelet neural Networks. Neurocomputing 2002;48:681–9.

Zhang Q. Using wavelet network in nonparametric estimation. IEEE Trans Neural Networks 1997;8(2):227–36.

Zhang Q, Benveniste A. Wavelet networks. IEEE Trans Neural Networks 1992;3(6):889–98.

Zhang J, Walter GG, Lee W. Wavelet neural networks for function learning. IEEE Trans Signal Process 1995;43(6):1485–97.