A surface electromyography classification system

A Surface Electromyography Classification System

P. A. Kaplanis[1,2], C. S. Pattichis[1,3], C. I. Christodoulou[3],
L. J. Hadjileontiadis[4] , V.C. Roberts[2], T. Kyriakides[1]

[1]The Cyprus Institute of Neurology and Genetics, P.O. Box 23462, Nicosia, Cyprus
[2]Kings College School of Medicine and Dentistry, University of London, UK
[3]Dept. of Computer Science, University of Cyprus, P.O. Box 20537, Nicosia, Cyprus
[4]Dept. of Electrical & Computer Engineering, Aristotle University of Thessaloniki, Greece

E-mail: p.a.kaplanis@cytanet.com.cy,

Abstract: The objective of this study was to investigate whether surface electromyography (SEMG) signals may be used to discriminate between normal (NOR), myopathy (MYO) and neuropathy (NEUR) subjects. Ninety-one normal subjects participated and 20 patients (11 MYO and 9 NEUR). Recorded signals were obtained for 5 seconds at 10, 30, 50 70 and 100% of Maximum Voluntary Contraction (MVC). Five parameters were obtained, two from the time domain (turns and zero crossings per second), two from the frequency domain (median frequency and total power per second) and one from the bi-frequency domain (bispectrum peak amplitude). The KNN classifier with the leave-one-out method was used for classification whereas the Self Organising Feature Maps (SOFM) was used for cluster analysis and visualization of the group separation. Results have shown that separation of normal subjects from neuromuscular diseased patients is possible with a success rate of the order of 83%, whereas separation of subjects from the two types of patients (myopathy and neuropathy) is again high and of the order of 77%.

Index terms: Surface electromyography, myopathy neuropathy, SOFM, KNN.

Introduction

To be able to identify whether or not a subject is suffering from a neuromuscular disorder or even more to identify the correct disorder type a needle electromyography (EMG) examination necessitating constant expert supervisor and all the drawbacks associated with an invasive examination. Analysis of signals recorded with surface electrodes from various research groups aiming to discriminate normal from abnormal subjects have produced poor results and have hence been generally neglected or approached with suspicion.

In this study, we present only selected results obtained from a large group of subjects (91 normal subjects and 20 patients), where fourteen parameters were examined [1].

In particular the influence of transition from a high to a low force level (100%, 70%, 50%, 30% and 10%) of the maximum voluntary contraction (MVC) was examined. Five parameters from the three domains (time, frequency and bi-frequency) were used for the analysis and comparison of the EMG signals acquired from the controls and patients. In addition Self Organising Feature Maps (SOFM) and the Kohenen Neural network (KNN) leave one out method parameter classification technique were adopted to classify the subjects into the following classes; normal (NOR) subjects (non-diseased), myopathy (MYO), or neuropathy (NEUR) patients.

Materials and Methods

Data Capture. SEMG were recorded from the biceps brachii (BB) muscle of all subjects. Eighteen non-diseased subjects, from a total of 91 were selected with the method of age/gender, in order to be tested against the patients separately.

Recording was done using a four-bar EMG active probe with an interelectrode distance of 10 mm and a bar width of 1 mm. The electrode block was placed on the BB, in such a way so that the second electrode was at a distance equal with 1/3 of the BB length towards the shoulder

A bar configuration has been preferred from the well-accepted circular configuration, since the former intersects with more fibers than the latter. By intersecting more fibers greater amplitude will be recorded [2]. From the four bars of the electrode, the second was used as allocation index. Its predefined placement ensures that all four electrodes lay between the innervation zone of the motor unit and the tendon. The differential recordings were recorded simultaneously one from each pair of the electrode bars. Only the highest value though of the pair was used for this analysis. Recordings were performed for 5 seconds at 10%, 30%, 50%, 70% and 100% of the maximum voluntary contraction (MVC). An antialiazing band pass filer [20÷500Hz] was initially applied on the recorded signals, which were then sampled with a sampling frequency of 1 kHz at a 12-bit digitization resolution.

Time domain analysis. Two parameters were estimated from the analysis of SEMGs in the time domain. The number of turns per second was defined as the number of slope reversals separated from the previous and the following turn by an amplitude difference greater than 20 μV in one second. The number of zero crossings per second was defined as the number of sign reversals exceeding a threshold of 20 μV in a period of one second.

Power spectral analysis. The following measures were estimated from the power spectrum curve of the analyzed SEMGs: median frequency and total power per second $P_{t/s}$ The signal was segmented at 512 points with 25% overlap.

Bispectral analysis. From the Bispectrum domain, the Bispectrum peak amplitude was calculated.

For a zero-mean, stationary process {*X(k)*}, the third-order cumulant (TOC) is defined as the expected value of the triple product

$$R(m,n) = E\{X(k)X(k+m)X(k+n)\}, \qquad (1)$$

and the bispectrum is defined as the Fourier transform of the TOC sequence [3].

$$B(\omega_1,\omega_2) = \sum_{m=-\infty}^{m=\infty}\sum_{n=-\infty}^{n=\infty} R(m,n)\, e^{-j(w_1 m + w_2 n)} \qquad (2)$$

The signal was segmented into overlapped records, for reducing the variance of the estimated BS, using a window of 512 points in length with a 25% overlap.

Classifications. The KNN classifier was used in this study for classification of subjects, between the three groups of subjects, as well as between two groups (normals and patients). In the KNN algorithm, in order to classify a new input pattern, its *k* nearest neighbours from the training set are identified. The new pattern is classified to the most frequent class among its neighbours based on a similarity measure that is usually the Euclidean distance as defined below [4].

$$Ed_{xy} = \|x-y\| = \sqrt{\sum_{i=1}^{N}(x_i - y_i)^2}$$

In this study, the KNN subject classification system was implemented for different values of k=1,3,5 and 7, and for 25 features (5 force level x five parameters), and it was tested for the following:

1. 18 NOR, 9 MYO and 9 NEUR
2. 18 NOR and 18 patients as one group
3. 91 NOR, 11 MYO and 9 NEUR
4. 91 NOR and 20 patients as one group

In order to evaluate the performance of a cluster analysis system the following parameters need to be first defined: (i) True positive (TP), decision occurs when the positive diagnosis of the system coincides with a positive diagnosis according to the Neurophysiologist (ii) False positive (FP) decision occurs when the system made a positive diagnosis that is not in agreement with the Neurophysiologist (iii) False negative, occurs when the system made a negative diagnosis that is not in agreement with the Neurophysiologist and (iv) True negative (TN) both the physician and the system indicate the absence of a positive diagnosis.

The following measures may then be defined [5]:

Percentage of correct classifications (%CC)

$$\%CC = 100 \times (TP + TN)/n \qquad (4)$$

n is the number of subjects

Percentage of false positives (%FP)

$$\%FP = 100 \times FP/(TN + FP)$$

(5)

Percentage of false negatives (%FN)

$$\%FN = 100 \times FN/(TP + FN)$$

(6)

Sensitivity

$$SE = 100 \times TP/(TP + FN) \qquad (7)$$

Sensitivity is the likelihood that an event will be detected given that it is present

Specificity

$$SP = 100 \times TN/(TN + FP) \qquad (8)$$

Specificity is the likelihood that the absence of an event will be detected given that it is absent

Recall

$$RE = 100 \times TP/(N_{NEUR} + N_{MYO}) \qquad (9)$$

where N_{NEUR} is the number of NEUR patients and N_{MYO} is the number of MYO patients

Recall is the number of positive diagnoses correctly made by the system, divided by the total number of positive diagnoses made by the physician.

Precision

$$PR = 100 \times TP/(TP + FP) \qquad (10)$$

Results and Discussion

The average values for the five selected parameters at the corresponding force levels for the three groups of subjects, normals (NOR), myopathy patients (MYO) and neuropathy patients (NEUR) are listed in Table 1 for all 91 normal subjects (top) and for the selection of 18 subjects (below). The values listed in Table 1 have been used to produce Tables 2 and 3 and the plots in Figure 1. These selected parameters present significant differences, ($p<0.05$, Wilcoxon signed rank test), between corresponding force levels and between one force level and another [1].

Table 2, presents the clustering results, including all measures listed earlier for different values of k value for two classes, (Normal and Abnormal). Table 3 is a similar table but for three classes (Normal, Myopathy and Neuropathy).

From Table 2, the highest percentage of correct classifications (%CC), between Normal and Abnormal (patients), is given for k=1 with 91 normal subjects compared with 20 patients (83%), which gives also the highest value in specificity (92). The success rate if compared with a similar one obtained (87.5%) [6], may seem low, however one needs to remember that this present, technique utilizes surface electromyography and voluntary stimulation alone. Furthermore the patients here were suffering from different type of diseases and were classified only as myopathy and neuropathy patients and irrespective of their disease severity. From Table 3 the highest value of correct classification and specificity is given for the same combination i.e. when k=1, and 91 normal subjects are compared with 11 myopathy patients and 9 neuropathy patients. However this should be seen with caution since the unequal number of subjects per class creates a bias in favor of the larger classes. The more reliable results are when equal number of subjects per class with the best result being 69%.

Figure 1 plots the Self Organising Feature Maps (SOFM), showing the distribution of all 91 normal subjects and 18 patients and of a selection of 18 subjects and 18 patients.

Table 1. Average values for the five parameters for the three groups of subjects at 10, 30, 50, 70 and 100% MVC, with the standard deviation in parenthesis for all 91 normal subjects (top table) and for the selected 18 normal subjects (below).

	PERCENTAGE OF MAXIMUM VOLUNTARY CONTRACTION														
	10%			30%			50%			70%			100%		
	NOR	MYO	NEUR	NOR	MYO	NEUR	NOR	MYO	NEUR	NOR	MYO	NEUR	NOR	MYO	NEUR
t/s	6	23	17	12	29	25	17	35	33	22	38	37	28	47	46
zc/s	3	17	10	7	23	18	12	29	26	26	32	31	20	36	38
Pfmed (Hz)	109	103	102	104	103	118	102	102	121	100	97	117	96	102	117
Pt/s V^2	0.5	1.2	0.7	1.1	1.5	1.1	1.5	1.8	1.4	1.8	1.9	1.6	2.2	2.4	1.9
Bmax V^3sec	2.0	2.2	1.9	2.2	2.3	2.0	2.3	2.4	1.8	2.4	2.7	1.8	2.5	2.7	2.0

	PERCENTAGE OF MAXIMUM VOLUNTARY CONTRACTION														
	10%			30%			50%			70%			100%		
	NOR	MYO	NEUR	NOR	MYO	NEUR	NOR	MYO	NEUR	NOR	MYO	NEUR	NOR	MYO	NEUR
t/s	6	23	17	11	29	25	18	35	33	25	38	37	32	47	46
zc/s	3	17	10	7	23	18	11	29	26	14	32	31	23	36	38
Pfmed (Hz)	124	103	102	121	103	118	115	102	121	113	97	117	107	102	117
Pt/s V^2	0.5	1.2	0.7	1.0	1.5	1.1	1,4	1.8	1.4	1.7	1.9	1.6	2.1	2.4	1.9
Bmax V^3sec	1.8	2.2	1.9	1.7	2.3	2.0	1.9	2.4	1.8	2.0	2.7	1.8	2.2	2.7	2.0

Table 2. KNN Classifier performance results for two classes (Normal and Abnormal)

K	NORMAL	ABNORMAL	%CC	%FP	%FN	SE	SP	RE	PR
1	91	20	83	8	60	40	92	44	53
1	18	18	61	28	50	50	72	50	64
3	91	20	77	24	20	80	76	89	42
3	18	18	58	33	50	50	67	50	60
5	91	20	72	31	15	85	69	94	38
5	18	18	69	22	39	61	78	61	73
7	91	20	81	14	40	60	86	67	48
7	18	18	67	28	39	61	72	61	69

Table 3. KNN Classifier performance results for three classes (Normal [NOR], Myopathy [MYO] and Neuropathy [NEUR])

K	NOR	MYO	NEUR	%CC	%FP	%FN	SE	SP	RE	PR
1	91	11	9	77	8	86	14	92	11	22
1	18	9	9	61	28	50	50	72	50	64
3	91	11	9	68	34	36	64	76	39	24
3	18	9	9	61	25	29	71	75	56	71
5	91	11	9	65	31	25	75	69	50	24
5	18	9	9	47	35	54	46	65	33	50
7	91	11	9	60	37	17	83	62	56	24
7	18	9	9	56	25	33	67	75	44	67

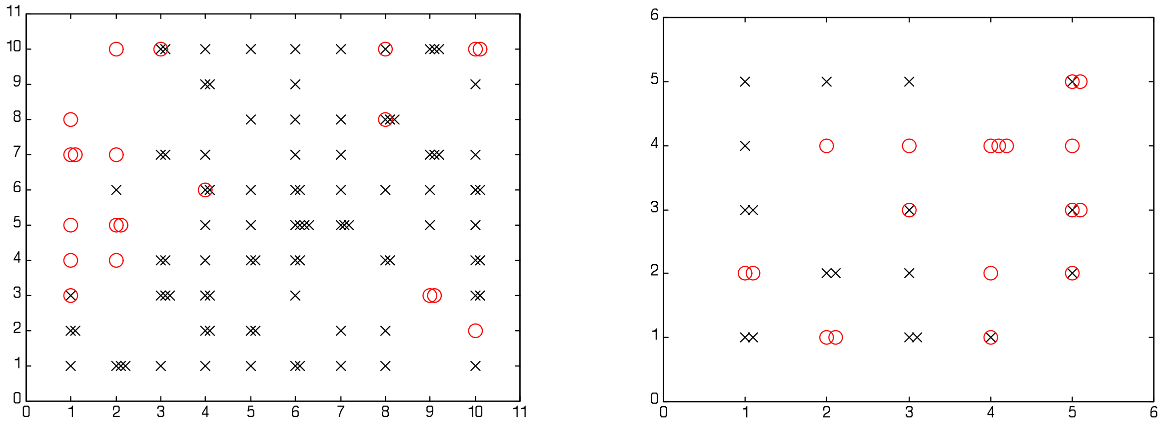

Figure 1. Left plot presents the SOFM classification between 91 normal subjects (x) and 18 patients (o) and right plot distinguishes between selected 18 normal subjects (x) and 18 patients (o).

Conclusions

Results have shown that surface electromyography signals when applied into our system may be used with a high degree of accuracy to discriminate between normal subjects and neuromuscular diseased patients. Furthermore patients may be discriminated from normal subjects into groups (normal, myopathy or neuropathy) with a slightly less degree of correct classification. However it is believed that even better results may be obtained when more parameters are applied into the system. Still, even these results are quite promising and may be used without causing any discomfort to the patients, to help the neurophysiologist cross -reference their diagnosis as to if and what type of disease a patient is suffering from.

Acknowledgements

This work was made possible through the kind support of the Cyprus Institute of Neurology and Genetics.

References

[1] Roberts V.C, Kaplanis P. A, Pattichis C. S., (2003): SEMG analysis in normal and pathological subjects, Proc. of the World Congress of Medical Physics and Biomedical Engineering, Sydney, Australia.

[2] De Luca C J, (1995): "Surface electromyography. Detection and recording," *A publication by the Neuromuscular Research Center*, Boston University

[3] Hinich M.J. (1982): "Testing for Gaussianity and Linearity of a stationary time series," *J. Time and Series analysis*, Vol. 3, pp. 169-176.

[4] Christodoulou C.I., Pattichis C.S., Pantziaris M, Nicolaides A, (2003): "Texture based classification of Artherosclerotic Carotid Plaques", IEEE Transactions on Medical Imaging, Vol., 22, No. 7, July.

[5] Pattichis C S, (1992): Thesis submitted for the degree of the Doctor of Philosophy in the faculty of Engineering, Queen Mary and Westfiled College, University of London, February.

[6] Christodoulou C.I., Pattichis C.S., Fincham W.F., (1998): "A Modular Neural Network Decision Support System in EMG Diagnosis", *Journal of Intelligent Systems,* Special issue on Computational Intelligent Diagnostic Systems in Medicine, ed. by C.N. Schizas, Volume 8, Nos. 1-2, pp. 99-143.