Associative Memories and Diagnostic Classification of EMG Signals

C. Shirota[1], M. Y. Barretto[1], C. Itiki[1]

[1]Biomedical Engineering Laboratory, Escola Politécnica da Universidade de São Paulo, Brazil
E-mail: {camila, mateus, cinthia}@leb.usp.br

Abstract

In this work, associative memories are used for diagnostic classification of needle EMG signals. Vectors containing 44 autoregressive coefficients represent each signal and are presented as stimuli to associative memories. As the number of training stimuli increases, the method recursively updates associative memories. The obtained classification results are equivalent to the ones provided by the traditional Fisher's discriminant, indicating the feasibility of the proposed method.

1 Introduction

Adaptive techniques have been applied to the study of electromyographic (EMG) signals. Artificial neural networks were used to control functional electrical stimulation in paraplegics, by adapting to changing environments and allowing patient's interaction with the network's operation [1]. An adaptive feature extraction algorithm was used for the classification of motion commands used in prosthetic arms [2]. Adaptive classification of motor unit action potentials was designed to deal with shape changes due to muscular fatigue [3].

EMG signals have also been used for diagnostic purposes [4]. Clinicians diagnose patients, based on the knowledge or experience that has been acquired over the years. An automatic classification technique should mimic a physician, by storing previous classification results and improving its diagnostic ability. However, most papers that deal with automatic EMG classification for diagnostic purposes present non-adaptive techniques [5], [6], [7].

In this work, we propose an adaptive EMG classification technique for diagnostic purposes and study its feasibility. EMG signals are represented by autoregressive models, since encouraging classification results have been presented in the literature [6], [7]. Associative memories are used as classifiers, for they provide reliable results [8] and they can be computed by a recursive algorithm. In order to validate the obtained classification results, they are compared to the ones obtained by the traditional Fisher's discriminant.

2 Methodology

2.1 Signal Acquisition and Autoregressive Modeling

A data bank with needle EMG signals was used in this work. These signals were recorded at 50% MVC, from the *biceps brachii* muscle, using a 10kHz low-pass filter, at the Duke University Medical Center.

Signals were acquired at the Biomedical Engineering Laboratory of the University of São Paulo with a 12-bit A/D converter and a sampling rate of 25kHz. We selected 800-ms EMG signals, recorded from 6 normal patients (72 signals), 6 myopathic patients (56 signals) and 5 neuropathic patients (43 signals). All signals were classified as stationary, both in mean and variance, by the run test [9], for a 5% significance level and 20 segments. EMG signals were modeled as outputs of autoregressive models. Burg's method [10] was used to estimate the autoregressive coefficients. Order 44 was used in this work, since previous studies [8] showed that this order provided the best classification results for the same data bank.

2.2 Classification

The use of a two-step classification procedure is reported in the literature [5], [6]. We used the two-step classification, since it provided better results than a single-step classification, in a previous research using the same data bank [11]. The first classification step separated EMG signals into normal and pathological classes. The second step separated the signals previously classified as pathological into neuropathic and myopathic classes.

The classification procedure was implemented by two linear classifiers—associative memories and Fisher's discriminant—, which are described in the following subsections.

2.2.1 Linear Associative Memories

According to psychologists, an associative memory (AM) often results from learning the relationship between a stimulus and a response (operant conditioning) [12]. The same terminology is used in the mathematical formulation of associative memories

[13]. A response r is associated to a stimulus s, through an associative memory M, according to equation 1:

$$r = M \cdot s \,. \tag{1}$$

The conditioning process is called training. Among all the signals from the data bank, a training set of signals is created. Autoregressive coefficients of each training signal form a column of the training stimuli matrix. The known classification of the signal is represented in a column of the training response vector. For each signal, an integer number represents the classification. In the first classification step, '1' stands for normal and '2' for pathological signal. In the second step, '1' stands for neuropathic, while '2' stands for myopathic. An associative memory vector M_{j+1} is estimated, by using the training stimuli matrix S_j and the training response vector R_j [13]

$$M_{j+1} = R_j \cdot S_j^+ \,, \tag{2}$$

where S^+_j indicates the Moore-Penrose inverse of the training stimuli matrix S_j [14].

As the number j of stimuli in the training matrix S_j increases, the computation of the generalized inverse becomes more burdensome. The stimuli matrix at the j-th iteration can be expressed as:

$$S_j = \begin{bmatrix} S_{j-1} & s_j \end{bmatrix} , \tag{3}$$

where S_{j-1} is the previous stimuli matrix and s_j is the newest stimulus vector to be included in the training group. As one can observe, the number of columns in the stimuli matrix increases with the number of training signals. As a consequence, the direct computation of the generalized inverse could be a hindrance to the use of associative memories as classifiers. However, Greville's recursion [14] enables us to update the generalized inverse by calculating:

$$S_j^+ = \begin{bmatrix} S_{j-1}^+ - S_{j-1}^+ \, s_j \, b_j^+ \\ b_j^+ \end{bmatrix} , \tag{4}$$

where b^+_j is the generalized inverse of vector b_j, which is defined by

$$b_j = s_j - S_{j-1} \, S_{j-1}^+ \, s_j \,. \tag{5}$$

If b_j given by equation 5 is equal to zero, then we must use another formula to compute b_j

$$b_j = \frac{1 + s_j^t \left(S_{j-1} S_{j-1}^t\right)^+ s_j}{s_j^t \left(S_{j-1} S_{j-1}^t\right)^+ \left(S_{j-1} S_{j-1}^t\right)^+ s_j} \left(S_{j-1} S_{j-1}^t\right)^+ s_j \,, \tag{6}$$

where the superscript 't' indicates the transpose matrix. In our Matlab® implementation, we calculated the norm of vector b_j given by equation 5. Whenever this norm was smaller than 10^{-7}, we substituted b_j by the value provided by equation 6. This procedure avoided numerical errors.

The generalized inverse of matrix S_j is obtained by computing the generalized inverse of column vector b_j, which has length 44 (the number of autoregressive coefficients). Vector b_j itself is either a function of the previously computed S^+_{j-1} or a function of the generalized inverse of the product $(S_{j-1} \, S^t_{j-1})$, which is square matrix of size (44 x 44). In this way, we can compute S^+_j by inverting smaller matrices or vectors that do not increase in size, as the number of training signals increases.

We also provide the recursion to compute the associative memory vector:

$$M_{j+1} = M_j - M_j \, s_j \, b_j^+ + r_j \, b_j^+ , \tag{7}$$

where r_j is the response corresponding to the newest training stimulus vector s_j, and M_j is the previous memory vector.

In order to evaluate the performance of these adaptive associative memories, we must use test stimuli s_k of known classification responses. The response r_k associated to each test stimulus is estimated by

$$\hat{r}_k = M_j \cdot s_k \,, \tag{8}$$

for j varying from 1 to the number of training stimuli and for k varying from 1 to the number of test stimuli. These responses are real values and do not correspond exactly to the integer values associated to the classes. As a consequence, we must use a criterion to separate the obtained responses into the classes. In the first classification step, responses r_j below 1.5 were classified as normal signals (class 1), while responses r_j greater or equal to 1.5 were classified as pathological signals (class 2). This same value was used in the second classification step, in order to separate neuropathic (class 1) from myopathic signals (class 2).

2.2.2 Fisher's Linear Discriminant

Fisher's linear discriminant (FLD) classifies a signal associated to a stimulus vector s_j, by estimating the response [15]

$$\hat{r}_j = v^T s_j \,. \tag{9}$$

In order to compute vector v, training stimuli must be separated into classes 1 and 2. For class 1, the mean stimulus vector u_1 and the covariance matrix Σ_1 are calculated. Similarly, for class 2, the mean stimulus vector u_2 and the covariance matrix Σ_2 are computed. The optimal projection direction that separates classes 1 and 2 is given by vector v:

$$v = \left(\frac{(n_1 - 1)\Sigma_1 + (n_2 - 1)\Sigma_2}{n_1 + n_2 - 2} \right)^{-1} \cdot (u_1 - u_2) \,, \tag{10}$$

where n_1 and n_2 are the number of stimuli in classes 1 and 2 respectively. The stimulus s_j is separated into

class 1, if the response r_j is above a threshold value. Otherwise, it is separated into class 2.
In this work, both AM and FLD were implemented in Matlab®.

2.2.3 Training group

In a real system, the diagnosis of a patient is based on previous knowledge. Similarly, an initial set of subjects was used to compose the initial training group. Three subjects were chosen to form the initial training group, providing 11 normal, 10 neuropathic and 10 myopathic signals.
Eight additional subjects were later added to the training group, at a random order. This random order of subjects was used because physicians cannot choose the order in which patients arrive at the clinic. All the signals of a given subject were inserted one-by-one into the training process. The following order of insertion was used: 12 normal, 15 normal, 7 myopathic, 9 neuropathic, 9 myopathic, 6 neuropathic, 16 normal and 12 myopathic signals.
At each training step, one signal was classified using all the information stored in the previous training group. Then, autoregressive coefficients of this signal were included in the next training stimuli group. The response associated to each stimulus varied according to the type of training: unsupervised or supervised [16]. Unsupervised training was based on the results given by the classifier. The estimated class for each new signal was inserted as the next training response. Supervised training was based on the known signal classification, which was provided by a teacher. For supervised training, the known classification of the subject was included as response, regardless of the estimated class. The known classification inserted in supervised training could represent the diagnosis provided by a physician, based on clinical evidence and exam results, other than EMG.

2.2.4 Test Group

In order to test the variation of classification rate with the increasing number of training stimuli, we needed signals that were not used previously in training. All six of the remaining subjects had already been separated for the test group, providing 18 signals of each type. All test signals were classified several times, as each new training signal was included in the training group.

3 Results

Classification rates obtained by AM are shown in Figure 1, while results obtained by FLD are illustrated in Figure 2. The horizontal axis presents the number of training signals used at each iteration and ranges from 31 (initial training group) to 117 (final training group). The vertical axis represents the correct classification rate in percentage. It is equivalent to the ratio between the number of correct classifications and the total number of test signals (54). Dashed lines represent the classification rate obtained for unsupervised training, while continuous lines represent supervised training.

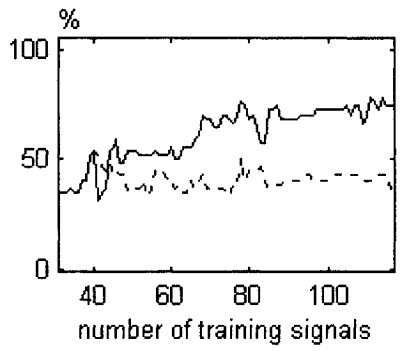

Fig. 1. Classification rates obtained by AM, for supervised (continuous line) and unsupervised training (dashed line).

According to figures 1 and 2, supervised training (continuous line) clearly showed better results than unsupervised training (dashed line), for both classifiers (AM and FLD). This result was expected, since the correct training is fundamental for the classifier's performance.

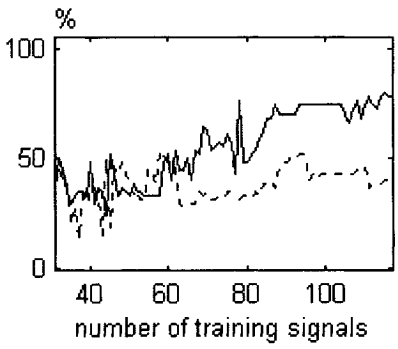

Fig. 2. Classification rates obtained by FLD, for supervised (continuous line) and unsupervised training (dashed line).

Unsupervised training seemed to stabilize at values close to 40%, which is similar to the classification rate obtained for the initial training group. This result shows that both classifiers did not 'learn' from

unsupervised training. On the other hand, supervised training presented increasing classification rates that reached values above 75%. The performance of both supervised classifiers improved as the size of the training group increased. This result shows that the supervised classifiers were able not only to store but also to use previous training information, in order to improve classification rates.

Comparing FLD and AM results, one can observe that FLD provides a more erratic behavior than AM, whenever less than 60 signals are used in training.

The method presented in this work uses simple update equations, whose computation does not increase in complexity, as additional training signals are included. So, the method can be applied to a larger group of signals and subjects. If we had more subjects in our data bank, we could have provided the classification rates in percentage of subjects. However, since we had only six test subjects, we provided our final classification rates (above 75%) in percentage of signals. The classification rates presented in the literature for FLD (60% to 87.5%) [7], [6] and neural networks (47.5% to 90%) [17], [7] are given in percentage of subjects and were obtained for the same parameters of our study—autoregressive coefficients of EMG signals.

4 Conclusion

This work presented a recursive method to update associative memories. These memories 'adapted' as new signals and their known classification were included in the training process. The method was applied to 117 training signals and 54 test signals. The obtained 75% classification rate showed the method's feasibility, since it is within the range of 47.5% to 90% presented by other methods in the literature [6], [7], [17].

Training signals were included one-by-one in the training group and associative memories were updated by recursive equations. As a consequence, the computational load did not increase as new information was added to the training process. This is one of the advantages of associative memories in comparison to other methods.

These results indicate that recursive associative memories could be applied, in the near future, to a large number of signals in a clinical setting, providing high classification rates.

5 Acknowledgements

This work was partially supported by FAPESP and CNPq. We thank Prof. Sanders from the Duke University Medical Center, who kindly provided the EMG data bank used in this study.

References

[1] Graupe, D., Kordylewski, H. (1995) Artificial Neural Network Control of FES in Paraplegics for Patient Responsive Ambulation. IEEE Trans. Biom. Eng. 42(7): 699-707.

[2] Lee, S., Kim, J., Park, S. (1996) An Enhanced Feature Extraction Algorithm for EMG Pattern Classification. IEEE Trans. on Rehabilitation Engineering 4(4): 439-443.

[3] Gazzoni, M., Farina, D., Merletti, R. (2004) A new method for the extraction and classification of single motor unit action potentials from surface EMG signals. Journal of Neuroscience Methods 136(2): 165-177.

[4] Desmedt, J. E. (1989) Computer-Aided Electromyography and Expert Systems. Elsevier-Science, Amsterdam.

[5] Berzuini, C., Maranzana-Figini, M., Bernardinelli, L. (1982) Effective use of EMG parameters in the assessment of neuromuscular diseases. International Journal of Bio-Medical Computing 13: 481-499.

[6] Inbar, G. F., Noujaim, A. E. (1984) On Surface EMG spectral characterization and its application to diagnostic classification. IEEE Trans. on Biomedical Engineering 31(9): 597-604.

[7] Abou-Chadi, F. E., Nashar, A., Saad, M. (2001) Automatic analysis and classification of surface electromyography. Frontiers of Medical and Biological Engineering 11: 13-29.

[8] Shirota, C., Barretto, M. Y., Itiki, C. (2004) Classificação de sinais eletromiográficos de agulha por memórias associativas e modelagem auto-regressiva. Proceedings of the International Federation for Medical and Biological Engineering 5(1): 959-962.

[9] Bendat, J. S., Piersol, A. G. (1986) Random Data - Analysis and Measurement Procedures, 2nd ed. John Wiley & Sons, New York.

[10] Marple, S. L. (1987) Digital Spectral Analysis: with Applications. Prentice-Hall, Englewood Cliffs.

[11] Barretto, M. Y., Kohn, A. F., Itiki, C. (2004) Modelagem auto-regressiva e discriminante de Fisher na classificação de sinais eletromiográficos de agulha. Proceedings of the International Federation for Medical and Biological Engineering 5(1): 939-942.

[12] Kandel, E. R., Schwartz, J. H., Jessell, T. M. (1991) Principles of Neural Science, 3rd ed. Elsevier Science, New York.

[13] Kohonen, T. (1984) Self-Organization and Associative Memory. Springer-Verlag, Berlin.

[14] Graybill, F. A. (1983) Matrices with Applications in Statistics. Wadsworth, Pacific Grove.

[15] Nadler, M., Smith, E. P. (1993) Pattern Recognition Engineering. John Wiley, New York.

[16] Haykin, S. (1999) Neural Networks: a Comprehensive Foundation, 2nd ed. Prentice Hall, Upper Saddle River.

[17] Pattichis, C., Elia, A. G. (1999) Autoregressive and cepstral analyses of motor unit action potentials. Med. Eng. Phys. 21: 405-419.