34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

Electromyographic Signal Compression Based on Preprocessing Techniques

Wheidima C. Melo[1], Eddie B. L. Filho[2] and Waldir S. S. Júnior[3]

Abstract— **Recently, electromyographic records have been rearranged into two-dimensional arrays and encoded with image compressors, in the same way as image data. However, as a consequence of this reshaping, the correlation among signal segments is generally lost, which reduces the compression efficiency. In the present work, new preprocessing techniques for encoding electromyographic signals as two-dimensional matrices are presented, namely percentage difference sorting and relative complexity sorting, which have the potential to favor the exploitation of the intersegment dependencies. The experiments were carried out with real isometric records acquired in laboratory, that were first preprocessed and then compressed with a JPEG2000 encoder, showing that the proposed framework is effective and outperforms even state-of-the-art schemes present in the literature, in terms of PRD $\times$ Compression Ratio.**

I. INTRODUCTION

Currently, there is a growing interest in biological signal processing [1], mainly due to the myriad of applications that they present, which comprehend areas such as commerce, defense and industry. For instance, there is a great demand for mobile devices capable of monitoring patients [2], databases for behavior comparison and disorder development analysis [3], and tools for assisting physicians in the diagnostic procedure [4].

The most researched biological signals are the Electrocardiogram (ECG) [5], which measures the variation of electrical stimuli applied to atria and ventricles, the Electromyogram (EMG) [6], which records the electrical activity related to the contraction of muscles in human body, and the Electroencephalogram (EEG) [7], that represents the electrical activity of the human brain, that is, oscillatory patterns known as rhythms.

Among these signals, the EMG and the EEG stand out, because in addition to the their importance for pathology analysis, they can also be used in control interfaces. For example, such signals can activate bionic prostheses [8], [9], [10], restoring voluntary movement to accident victims or patients, who suffer from some muscle dysfunction. In particular, the EMG signal directly represents the electrical impulses sent by the nervous system to muscle fibers, and it is present even when a limb is lost [8]. Furthermore, this signal can be easily acquired with surface electrodes (S-EMG), without causing any damage to the patient's skin.

Given that, the need for transmission and/or storage of EMG records tends to increase, which consequently results in the quest for efficient compression methods. However, such schemes, in addition to providing a compact representation, should also be able to preserve the clinical information contained in the signal [11], in such a way that its main features can be extracted, or a medical diagnosis, based on them, can still be carried out.

Although EMG, ECG and EEG records are normally one-dimensional, some authors have used a different approach for their compression, by considering them as images [12], [13], while leaving for the compressor the task to exploit the intra and intersegment redundancies. Besides, in order to increase the two-dimensional signal correlation and consequently the image encoder performance, preprocessing techniques are usually employed [13].

This work proposes two new preprocessing techniques, for compressing EMG signals as images, whose goal is to improve the intersegment dependencies, by increasing correlation among adjacent elements: the percentage difference sorting, which organizes signal segments based on their relative similarity, and the relative complexity sorting, which rearranges signal segments based on their variances and covariances.

Some works [12], [13], [14] used the JPEG2000 [15] for encoding the two-dimensional arrangement, with good results. The same encoder is also employed in the present work.

The remainder of this paper is organized as follows. In Section II, some basic aspects about electromyography are discussed, as well as a justification for the use of two-dimensional compressors. In Section III, some preprocessing techniques available in the literature are presented, and the proposed methodologies are introduced in Section IV. Section V presents the basic compression scheme, and Section VI provides experimental results, with actual EMG signals recorded in laboratory. Finally, Section VII draws the conclusions.

II. THE EMG SIGNAL

The electromyogram is of paramount importance for the biomedical engineering area, because it is generated by the contraction of muscles in human body, when the patient is conscious (similar movements result in similar signal shapes), and can be directly used as a movement-intention indicator [9] or as a diagnosis tool [16].

The basic contributing parts for the electromyogram are the muscle fibers, which are innervated by nerve fibers (extension of neurons), in order to produce an action potential.

*This work was not supported by any organization.

[1]W. C. Melo is with Universidade Federal do Amazonas - UFAM, Av. Gen. Rodrigo Octávio Jordão Ramos, 3000, Manaus - AM, 69077-000, Brazil (e-mail: wheidimawcm@gmail.com)

[2]E. B. L. Filho is with Universidade Federal do Amazonas - UFAM, Av. Gen. Rodrigo Octávio Jordão Ramos, 3000, Manaus - AM, 69077-000, Brazil, and Centro de Ciência, Tecnologia e Inovação do Pólo Industrial de Manaus, Rua Salvador, 391, Adrianópolis, Manaus-AM, 69057-040, Brazil (e-mail: eddie@ctpim.org.br)

[3]W. S. S. Júnior is with Universidade Federal do Amazonas - UFAM, Av. Gen. Rodrigo Octávio Jordão Ramos, 3000, Manaus - AM, 69077-000, Brazil (e-mail: waldirsabino@gmail.com)

A single nerve fiber can innervate multiple muscle fibers, composing a Motor Unit (MU). When the nervous system stimulates a motor neuron, all muscle fibers under its control produce a signal, known as Motor Unit Action Potential (MUAP). Then, the EMG signal is defined as the sum of the MUAPs of all motor units near the electrodes [17].

After being acquired, the EMG signal is amplified and filtered, in order to mitigate noise. The signal amplitude ranges from 0 to 10 mV (peak to peak), with frequencies between 0 and 500 Hz [18].

There are two basic types of EMG: the intramuscular (I-EMG) and the surface (S-EMG), the latter being a very interesting alternative, since it does not cause any damage to the patient's body. However, the S-EMG does not provide a good signal quality, being similar to noise. In summary, it is known that a S-EMG signal captured during constant force and constant angle contractions, with a Maximal Voluntary Compression (MVC) greater than 30%, can be modeled as zero mean, correlation ergodic, stochastic process with Gaussian distribution [19]. Such a result is very important, because it means that adjacent samples already present correlation, that can be exploited by any data compressor.

During the two-dimensional compression of a S-EMG signal, the samples vector is partitioned into segments of length N, which are then organized in a matrix of dimensions $N \times M$. This procedure often results in a signal with low intersegment correlation [12], [13], which may compromise the data compression with image encoders. However, since S-EMG segments can be regarded as independent units, it is possible to increase the compression efficiency by adding another processing stage, immediately prior to the two-dimensional compression [12], [13], [14], in such a way that signal segments are rearranged [12], [13], [14] or even modified, favoring the exploitation of intersegment dependencies.

Therefore, it is worth noticing that the compression of a S-EMG signal as an image may provide good results, as long as a preprocessing stage is added, in order to increase the two-dimensional arrangement correlation and thus the compression efficiency. In addition, improvements in the encoding methodology would concentrate only on preprocessing techniques, given that it is possible to use off-the-shelf images compressors.

III. PREPROCESSING TECHNIQUES IN THE LITERATURE

As mentioned in the previous section, preprocessing techniques are a useful tool for biological signal compression, as they allow an efficient use of encoders that were not specifically developed for this purpose. These techniques are generally based on the characteristics of the signal and the compressor, and can be classified into two groups: techniques with distortion and without distortion.

Normally, techniques with distortion change the signal format, sometimes in a non-invertible way, so that the intersegment dependencies are better exploited. For instance, in [20] the ECG periods are identified and separated, with each image row representing a heartbeat. However, as the patient's condition changes during the exam, even adjacent periods tend to have different lengths. In order to increase the correlation of the arrangement, all ECG periods are normalized to the same length. In [13], the ECG signal also goes through a DC equalization step, in which all periods are clamped at the lowest possible DC level.

On the other hand, techniques without distortion just reorganize the constituent segments. In [13], the authors propose a technique for reordering ECG periods, so that similar signal segments are placed in adjacent rows. The period with the lowest variance is moved to the first row, while the others are sorted in decreasing order of similarity with the first one, according to the mean squared error. Another similarity metric is proposed in [12], in which EMG segments are sorted according to their correlation coefficients.

IV. THE PERCENTAGE DIFFERENCE AND RELATIVE COMPLEXITY SORTING ALGORITHMS

In [12], the authors proposed a column sorting technique based on correlation coefficients. Nonetheless, the correlation coefficient informs whether two variables have a linear relationship or not, which may be direct or inverse. So, it is possible to have two segments with high correlation, but very different amplitudes, which may compromise the exploitation of intersegment redundancies, making the signal more difficult to compress.

The present paper addresses this issue, by presenting a more efficient sorting procedure, called the relative complexity sorting, which reorganizes columns based on their similarities. However, it also takes into account some information about signal complexity, that is, the variance of each segment. The proposed procedure considers the covariance matrix, which has a definition similar to the one presented by the correlation coefficient matrix, but also includes the mean and the variance of the segment to be positioned, in such a way that the rearrangement is jointly given by the linear relationship between segments and the proximity of their amplitudes. The procedure is carried on according to

$$RC(x,y) = C(x,y) + \lambda mean(y) + \gamma var(y), \tag{1}$$

where $RC(x,y)$ is the relative complexity sorting index for segments x and y, $C(x,y)$ is the covariance, $mean(y)$ is the mean value of the segment to be placed, $var(y)$ is the variance of the same segment, and λ and γ are trade-off multipliers.

The two columns that possess the largest relative complexity value are placed in the first two columns. The following columns are then rearranged according to the highest value of relative complexity, based on the last sorted column.

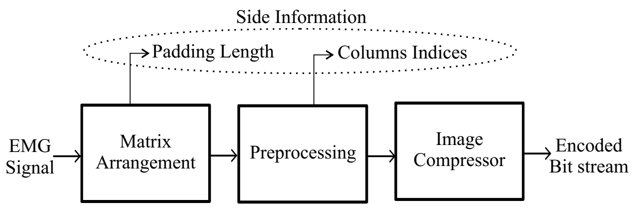
A second sorting procedure is also proposed, which consists in calculating the percentage difference, computed as

$$PD(x,m) = \frac{\displaystyle\sum_{i=0}^{N-1} (x[n] - m[n])^2}{\displaystyle\sum_{i=0}^{N-1} x^2[n]}, \tag{2}$$

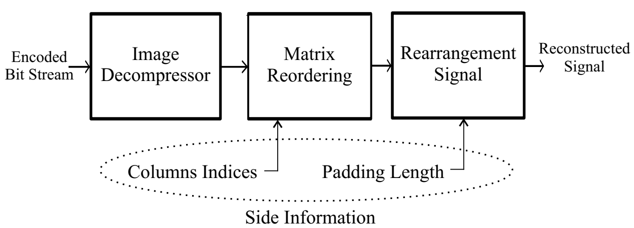
where $PD(x,m)$ is the percentage difference sorting figure for segments x and m, $x[n]$ is the last sorted column, $m[n]$ is the column under analysis, and N is the number of samples in each column. Before calculating the percentage difference, the segment with minimum variance is placed in the first column. The remaining columns are then rearranged, according to the lowest value of percentage difference.

V. THE PROPOSED COMPRESSION FRAMEWORK

The compression scheme used in this work is composed by three steps: signal reshaping, preprocessing and encoding. First, the input EMG record is partitioned, so that the number of samples within each segment is equal to the number of segments, which then occupy the columns of a square matrix. If the last segment is incomplete, its last element is used as padding. Next, the matrix is rearranged according to one of the sorting algorithms, generating a list of column indices that are arithmetically encoded and transmitted, as side information. The resulting two-dimensional signal is then input to a JPEG2000 encoder, which generates the compressed bit stream. At the decoder, all presented step are simply executed in a reverse order. The proposed CODEC is shown in Figure 1.



(a) Block diagram of the encoder system



(b) Block diagram of the decoder system

Fig. 1. Block diagram of the proposed CODEC system.

The EMG signal is compressed by the JPEG2000 as raw data, which means that at least two parameters must be provided: image dimensions and bit depth. However, depending on the bit depth, the quantization step may also be varied, in order to improve signal quality.

VI. SIMULATIONS RESULTS

The proposed scheme was evaluated by running tests with EMG signals collected from the biceps brachii of 13 subjects, during isometric contractions, while they were seated with the upper arm parallel to the torso and sustaining 60% of MVC. The resulting signals were sampled at 2000 Hz and quantized with 12 bits, with a duration ranging from 1.3 to 3.0 minutes. Each input signal was then rearranged into a two-dimensional array and processed by a JPEG2000 encoder (the Kakadu version, available at http://www.kakadusoftware.com). We have used all the default parameters for the JPEG2000 encoder, with CDF (Cohen-Daubechies-Feauveau) 9/7 wavelet kernel (irreversible wavelet transform), and changed the quantization step, which was fixed to 0.000025, and the bit depth, which was set to 16. The quality of the reconstructed signals was evaluated by using the percent root mean square difference (PRD), commonly adopted in the literature, defined as

$$PRD = \sqrt{\frac{\sum_{i=0}^{N-1}(x[i]-\widehat{x}[i])^2}{\sum_{i=0}^{N-1}x^2[i]}} \times 100, \tag{3}$$

where $x[i]$ is the original signal, $\widehat{x}[i]$ is the reconstructed signal, and N is the number of samples. The compression factor is defined as

$$CF = \frac{B_o - B_c}{B_o} \times 100, \tag{4}$$

where the B_o is the total number of bits in the original signal and B_c is the total number of bits in the compressed format, including header information. For each EMG signal used here, $B_o = 12 \times n$, corresponding to the 12 bits resolution and n samples of each signal.

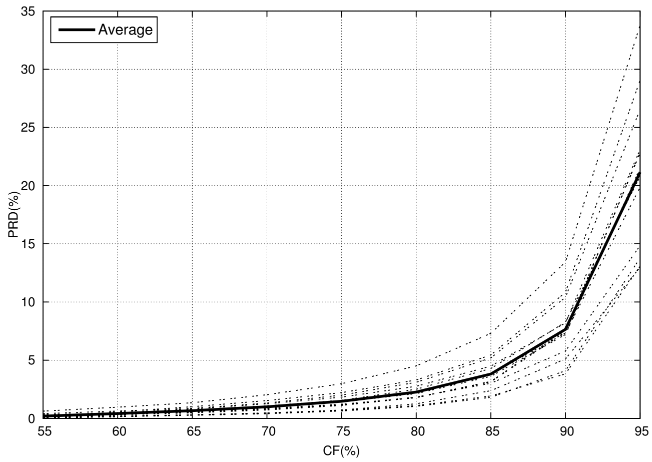
The results for all the 13 isometric EMG signals, with the proposed compression scheme and preprocessing techniques, that is, the percentage difference and the relative complexity sorting procedures, are presented in Fig.2(a) and Fig.2(b), respectively. The percentage difference sorting was employed with λ and γ set to 1. The curves show that their performances are similar, however, the percentage difference sorting tends to present slightly better results. It is also possible to notice that, for CFs below 84%, the proposed scheme keeps a reasonable PRD (< 7%), which is in general enough for not compromising the medical diagnosis [3].

The average results obtained for the proposed algorithms, for all 13 signals, are shown in Table I, along with results for state-of-the-art methods present in the literature. As can be seen, the proposed scheme, with either the percentage difference sorting (*p.d.s.* + JPEG2000) or the relative complexity sorting (*r.c.s.* + JPEG2000), outperformed all other methods for the majority of the CFs, but the one in [3] at a CF of 90%, which uses a spatial domain approach and tries to approximate signal blocks with elements retrived from an adaptive dictionary. The method in [21] is based on wavelets and a dynamic bit-allocation scheme, which uses a Kohonen layer (an artificial neural network), and the one in [22] employs a modification of the EZW algorithm. On the other way, the method in [12] is also based on the JPEG2000 encoder, but uses a different preprocessing technique.

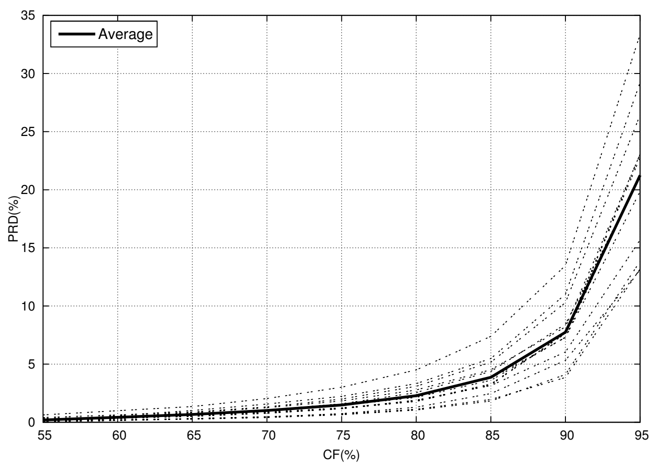
TABLE I
AVERAGE PRD(%) RESULTS FOR ISOMETRIC SIGNALS.

Compression Factor(%)	75	80	85	90
Norris *et al.* [22]	3.8	5	7.8	13
Berger *et al.* [21]	2.5	3.3	6.5	13
Chaffim *et al.* [12]	3.6	4.7	6.8	15.3
Filho *et al.* [3]	1.61	2.51	4.13	7.36
p.d.s. + JPEG2000	1.46	2.26	3.81	7.65
r.c.s. + JPEG2000	1.48	2.28	3.86	7.74

The test signals used in the present work are the same used in [3], which were kindly provided by the authors of [21], and seem to use an acquisition protocol similar to the one in [12].



(a) Performance for the percentage difference sorting



(b) Performance for the relative complexity sorting

Fig. 2. Simulation results: Percent Root Mean Square Difference (PRD) *versus* Compression Factor (CF)

VII. CONCLUSIONS

We presented new preprocessing techniques, that can be directly employed on the compression of EMG signals rearranged as images and are capable of improving the exploitation of the intersegment dependencies. They can be incorporated by any compressor or combined with existing schemes, leading to an enhancement in the quality of the reconstructed signal. In the simulations carried out for validating the proposed methodology, the techniques introduced in Section IV were employed in a JPEG2000-based framework, outperforming state-of-the-art methods present in the literature.

ACKNOWLEDGMENT

The authors wish to thank Professor Dr. Pedro Berger, for providing the test signals used in this work.

REFERENCES

[1] W. A. Sandham and D. J. Hamilton, "Pedagogical Issues for Effective Teaching of Biosignal Processing and Analysis", in proceedings of the *32nd International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 2947-2950, Buenos Aires, Argentina, September 2010.

[2] H. Jin and B. Miao, "Design of Holter ECG System Based on MSP430 and USB Technology", in proceedings of the *1st International Conference on Bioinformatics and Biomedical Engineering*, pp. 976-979, Wuhan, China, July 2007.

[3] E. B. L. Filho, E. A. B. Silva and M. B. Carvalho, "On EMG Signal Compression with Recurrent Patterns", *IEEE Transactions on Biomedical Engineering*, Vol. 55, No. 7, pp. 1920-1923, July 2008.

[4] M. Fani, G. Azemi and B. Boashash, "EEG-based automatic epilepsy diagnosis using the instantaneous frequency with sub-band energies", in proceedings of the *7th International Workshop on Systems, Signal Processing and their Applications*, pp. 187-190, Tipaza, Algeria, May 2011.

[5] L. Sörnmo and P. Laguna, *Electrocardiogram (ECG) Signal Processing*, Wiley Encyclopedia of Biomedical Engineering, Vol. 2, pp. 1298-1313, Wiley, 2006.

[6] R. Merletti and P. A. Parker, *Electromyography: Physiology, Engineering, and Non-Invasive Applications*, Wiley-IEEE Press, 2004.

[7] S. Sanei and J. A. Chambers, *EEG Signal Processing*, Wiley-Interscience, 2007.

[8] M. Gauthaam and S. S. Kumar, "EMG Controlled Bionic Arm", in proceedings of the *National Conference on Innovations in Emerging Technology*, pp. 111-114, Erode, India, February 2011.

[9] T. Hoshino, M. Tomono, R. Furusawa, T. Suzuki, M. Shimojo and K. Mabuchi, "Development of a Motion Support System by Using an Electromyogram," in proceedings of the *IEEE International Conference on Systems, Man and Cybernetics*, Vol. 5, pp. 4432-4437, The Hague, The Netherlands, October 2004.

[10] Z. Zhen and F. Hong-liang, "Prosthetic Controlled System Based on Signal Pattern Recognition of Electroencephalogram", in proceedings of the *International Seminar on Future BioMedical Information Engineering*, pp. 287-290, Wuhan, China, December 2008.

[11] E. B. L. Filho, E. A. B. Silva, M. B. Carvalho, W. S. S. Júnior and J. Koiller, "Electrocardiographic Signal Compression using Multiscale Recurrent Patterns", *IEEE Transactions on Circuits and Systems–I: Regular Papers*, Vol. 52, No. 12, pp. 2739-2753, December 2005.

[12] M. V. C. Costa, J. L. A. Carvalho, P. A. Berger, A. F. Rocha and F. A. O. Nascimento, "Compression of Surface Electromyographic Signals Using Two-Dimensional Techniques", *Recent Advances in Biomedical Engineering*, Vienna: In-Tech, pp. 17-38, October 2009.

[13] E. B. L. Filho, N. M. M. Rodrigues, E. A. B. Silva, S. M. M. Faria, V. M. M. Silva and M. B. Carvalho, "ECG Signal Compression Based on Dc Equalization and Complexity Sorting", *IEEE transactions on Biomedical Engineering*, Vol. 55, No. 7, pp. 1923-1926, July 2008.

[14] A. Bilgin, M. W. Marcellin and M. I. Altbach, "Compression of Electrocardiogram Signals using JPEG2000", *IEEE Transactions on Consumer Electronics*, Vol. 49, No. 4, pp. 833-840, November 2003.

[15] D. S. Taubman and M. W. Marcellin, JPEG2000: Image Compression Fundamentals, Standards and Practice, Boston: Kluwer Academic Publishers, 2002.

[16] R. R. L. Cisi and E. F. C. Júnior, "Human Gait Analyzed by an Artificial Neural Network Model, Proc. of the IV Brazilian Conference on Neural Networks–IV", in proceedings of the *IV Brazilian Conference on Neural Networks*, pp. 148-151, So Jos dos Campos, Brazil, July 1999.

[17] M. Nikolic, "Detailed Analysis of Clinical Electromyography Signals EMG Decomposition: Findings and Firing Pattern Analysis in Controls and Patients with Myopathy and Amytrophic Lateral Sclerosis," *PhD Thesis*, Faculty of Health Science, University of Copenhagen, 2001.

[18] C. J. De Luca, "Surface Electromyography: Detection and Recording", Delsys Incorporated, 2002, retrieved February 23, 2012, from http://www.delsys.com/Attachments_pdf/WP_SEMGintro.pdf.

[19] K. Nazarpour, A. R. Sharafat, and S. M. Firoozabadi, "Negentropy Analysis of Surface Electromyogram Signal," in proceedings of the *13th IEEE Workshop on Statistical Signal Processing*, pp. 974-977, Bordeaux, France, July 2005.

[20] E. B. L. Filho, E. A. B. Silva, Waldir S. S. Júnior and M. B. Carvalho, "ECG Compression using Multiscale Recurrent Patterns with Period Normalization," in proceedings of the *IEEE International Symposium on Circuits and Systems*, Kos, Greece, May 2006.

[21] P. A. Berger, F. A. O. Nascimento, J. C. Carmo, and A. F. Rocha, "Compression of EMG signals with wavelet transform and artificial neural networks," *Institute of Physics Publishing: Physiological Measurement*, Vol. 27, No. 6, pp. 457-465, June 2006.

[22] J. A. Norris, K. Englehart, and D. Lovely,"Steady-state and dynamic myoelectric signal compression using embedded zero-tree wavelets," *Proceedings of the 23rd Annual EMBS International Conference*, pp. 1879-1882, October 2001.