ALGORITHM FOR COMPRESSION OF EMG SIGNALS

P. A. Berger[1], F. A. O. Nascimento[1], J. C. Carmo[1], A. F. Rocha[1], Icaro dos Santos[2]

[1]Grupo de Processamento Digital de Sinais - Department of Electrical Engineering, University of Brasília – UnB

[2]Department of Electrical and Computer Engineering, University of Texas at Austin, USA

Abstract— **This paper presents an algorithm for the compression of electromyography (EMG) signals. Its main application is in the efficient construction banks of EMG signals. The developed algorithm uses an orthogonal transform followed by a thresholding process for choosing significant coefficients. Many orthogonal transforms were studied and the Daubechies-4 Wavelet Transform has shown a better performance for data compression. The proposed algorithm was implemented, and the experimental results have shown that a compression gain greater than 10:1 can be achieved without perceptible errors in the decoded signal.**

Keywords—**Compression, EMG, wavelets**

I. INTRODUCTION

The technology for treatment of digitized signals from the physical world has been opening new fields of study and it has been propitiating the creation of algorithms and specialist technological tools. In this context, the analog signal should be digitized in agreement with Nyquist criterion and processed or stored in mass memory. The biological signals are included in this class. The electromyography (EMG) signal has been object of study by researchers of several areas of the human knowledge. The construction of data banks with EMG signals constitutes an interesting possibility, as it would provide access to a significant number of physiologic phenomena, facilitating the study of these phenomena, its mathematical modeling and, also, the development of new techniques for the treatment of the signal. One inconvenience of the EMG data bank is the relatively high amount of memory necessary for its storage. The acquisition of multiple signals for a relatively long time (a few minutes) can make the construction of a database with a high amount of signals unfeasible. In order to build a database with compressed signals, it is important to study the characteristics of these signals. The compression of the EMG signals leads to an efficient representation of the information, facilitating its transmission by physical means and its storage for subsequent analysis [1].

This paper presents a data compression algorithm for EMG signal. It is based on the discrete wavelet transform. In order to evaluate the quality of decoded signals we chose a compression scheme based on the discrete cosine transform which was used as a reference transform. The proposed compression scheme gives more priority to the fidelity of the recovered EMG wave than for the compression rate.

II. METHODOLOGY

A. Orthogonal Transforms

Many studies have been published on transform coding methods, which were applied to ECG and EEG signals. However, we only find a few studies specifically on EMG signals. Different methods for error-free compression of EMG signals were compared to the other methods based on transforms [3]. These methods based on transform coding showed good results.

The Karhunen-Loève orthogonal transform is considered optimal in the statistical sense [4]. However, this transform does not have a fast algorithm, its basis functions depend on the signal statistics and it requires a high computational effort, which restricts its use. The Discrete Cosine Transform (DCT) is considered sub-optimal but it has a fast algorithm which requires only O(N.log2N) operations. Given a discrete signal x[n], with n=0, 1,…, N-1, [4]. Equations (1) and (2) define the Discrete Cosine Transform (DCT).

$$C[k] = w[k]\sqrt{\frac{2}{N}}\sum_{n=0}^{N-1} x[n]\cos\left[\left(n+\frac{1}{2}\right)\frac{k\pi}{N}\right], 0 \le k \le N-1 \qquad (1)$$

$$x[k] = w[k]\sqrt{\frac{2}{N}}\sum_{n=0}^{N-1} C[k]\cos\left[\left(n+\frac{1}{2}\right)\frac{k\pi}{N}\right], 0 \le n \le N-1 \qquad (2)$$

Where,

$$w[k] = \begin{cases} \frac{1}{\sqrt{2}}, \text{ if } k = 0; \\ 1, \text{ if } k = 1,2,..., N-1 \end{cases} \qquad (3)$$

On the other hand, the wavelet transform is a representation in time-scale of the analysed signal, where the basis functions are translations, dilations and compressions of a function. This function is called mother wavelet. The temporal analysis is carried out by using compressed versions (or versions with high frequency) of the mother wavelet. The analysis in frequency is carried out by dilated versions (or versions with low frequency) of the mother wavelet. The choice of a Wavelet Transform is essentially the choice of the filter bank. The Daubechies bases, studied in this work, corresponds to a sub-band filter bank outlined using FIR filters and they adapt efficiently to the solution of the proposed problem.

The present work was accomplished with the support of CNPq, an entity of the Brazilian Government to the scientific and technological development.

In order to obtain orthogonal Daubechies wavelet basis functions, one departs from appropriate coefficients h_n, and then investigates if they correspond to an orthogonal wavelet basis. If a finite number of h_n is different from zero, the function (denominated mother wavelet) is reduced to a finite combination of linear functions with compact support and, therefore, automatically the function possesses compact support. The research results in a collection of coefficients Mh_n, where M=2,3,4,..., and $0 \le n \le 2M-1$. For instance, for four coefficients we have:

$$(h_0, h_1, h_2, h_3) = \left(\frac{1+\sqrt{3}}{4\sqrt{2}}, \frac{3+\sqrt{3}}{4\sqrt{2}}, \frac{1+\sqrt{3}}{4\sqrt{2}}, \frac{3-\sqrt{3}}{4\sqrt{2}}, \frac{1-\sqrt{3}}{4\sqrt{2}} \right) \quad (4)$$

Beginning from these coefficients one can build the function called father wavelet by solving the dilation equation:

$$\phi(x) = \sqrt{2} \sum_{k=0}^{3} h_k \phi(2x - k) \quad (5)$$

Once $g_n=(-1)^n[h_{-n+1}]$ we have that

$$(g_0, g_1, g_2, g_3) = \left(\frac{1-\sqrt{3}}{4\sqrt{2}}, \frac{-3+\sqrt{3}}{4\sqrt{2}}, \frac{3+\sqrt{3}}{4\sqrt{2}}, \frac{-1-\sqrt{3}}{4\sqrt{2}} \right) \quad (6)$$

and the basis function is:

$$\psi(x) = \sqrt{2} \sum_{k=0}^{3} g_k \phi(2x - k) \quad (7)$$

B. Emg pre-processing: time-invariant de-noising

When the EMG signal is digitized, it has various kinds of noise due to physiological and instrumentation causes. In order to get a better performance of the compression technique we decided to pre-process the EMG signal by applying a de-noising algorithm. Time invariant wavelet de-noising was used in this case and, as a second result of this investigation, it was expected to evaluate the wavelet transform that has the best performance in the sense of space domain decorrelation of signal.

In order to choose the orthogonal wavelet basis which better represents the EMG signal, an algorithm was developed to verify the data correlation in the transform domain. This same process was adopted to evaluate the DCT-based algorithm.

The performance in the sense of noise reduction was measured by the mean squared error (MSE). Fig. 1 shows the MSE versus the number of discarded coefficients in a noisy EMG signal of some of the wavelet functions of Daubechies' family. The percentage of small discarded energy coefficients is shown in the horizontal axis and the mean square error (MSE) is shown in the vertical axis. It can be verified that Daubechies 4 presents the maximum noise reduction, eliminating about 93% of its coefficients.

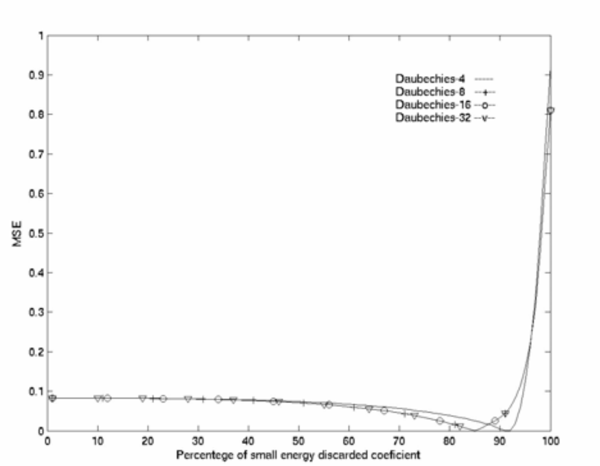

Fig. 1. Simulation with an example of the original EMG signal and the correspondent signal after noise reduction.

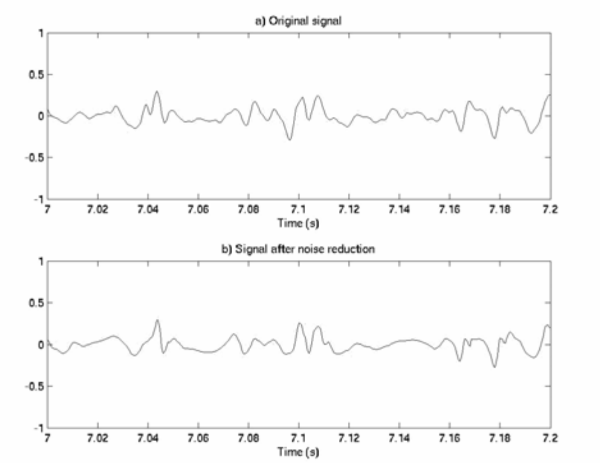

Fig. 2. Example: original signal (above), signal after noise reduction (below).

The objective to identify the wavelet transform which clusters more signal energy in fewer coefficients was achieved. Thus, it is possible to use this result to create an efficient compressed representation of the original signal.

C. Transform compression performance for EMG signals

The performance of the compression scheme was measured objectively according to two criteria: the compression ratio and the signal to noise ratio (SNR) of the reconstruction. The compression ratio is defined as the ratio between the number of bits used to represent this signal before and after compression. The estimate of the compression rate includes the amount of bits used to represent the lateral information necessary for the decodification. This information comprises a vector, which contains the information of the position of the relevant coefficients and the amount of bits used to quantify these coefficients. The signal to noise ratio of the reconstruction is defined as:

$$\mathrm{SNR} = 10\log\left(\frac{E_x}{E_e}\right) \tag{8}$$

where E_x is the signal energy and E_e the reconstruction error energy.

The Daubechies´ family of wavelets was chosen for coding EMG signals because they showed a better compression performance. The compression was done by discarding the smaller M transformed coefficients. When we increase the amount of eliminated coefficients, the compression rate also increases. However, the fidelity of reconstructed signal gets worse.

It was observed that the Daubechies-4 function takes better advantage of the statistical characteristics of the EMG signals and, therefore, presents a better performance when used for the compression of these signals. As expected, when the resolution scale is increased in the wavelet transform, the codification performance is better, since the transform can better extract the signal statistics. However, for the scale resolution above L=8, no improvement in the performance is observed, while the computational effort is larger. The following figure shows the comparison between the compression performance between the DCT and DWT. It is observed, for the same compression rate, that the DWT presents a better signal to noise ratio.

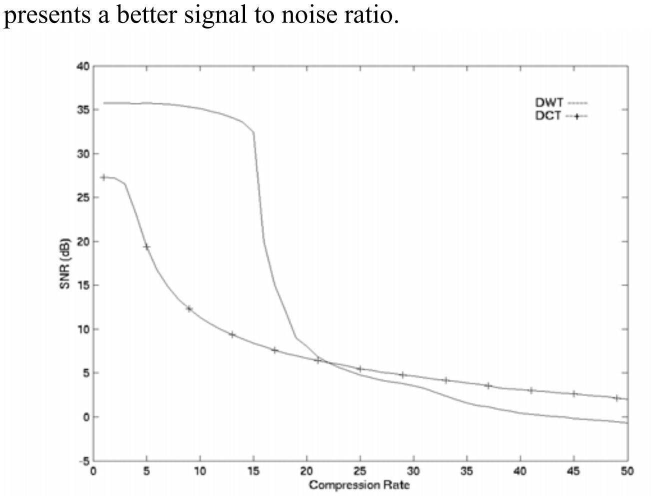

Fig. 3. Comparison between DWT and DCT.

D. Compression algorithm overview

As the original EMG is a band limited signal [5, 6], there is a resolution scale above which there is little or no energy. The basic idea in the transform coding is to use an orthogonal transform to compact the signal energy in a small number of transformed coefficients and to eliminate those coefficients that are considered to be non-relevant, according to a criterion of reconstruction error control as presented in the following section.

The block diagram of the compression algorithm is shown in Fig. 4.

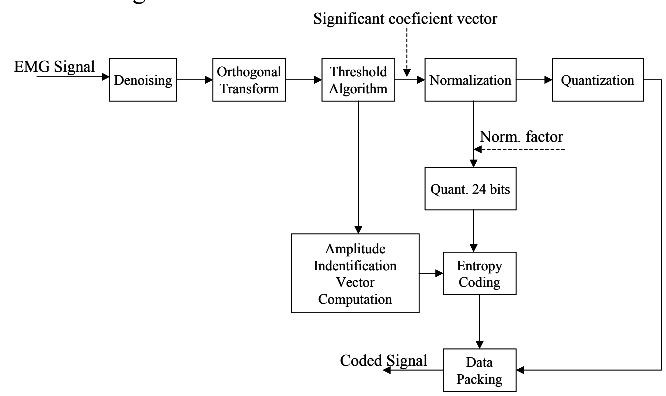

Fig. 4. Compression algorithm block diagram.

The coding algorithm is composed by the de-noising step using the time-invariant wavelet transform, the orthogonal transform, the selection of the relevant coefficients, the normalization, the quantization of those coefficients, and computation and entropy coding of the side information necessary for the decoding process.

The process of coefficient selection is developed and implemented taking into account the nature of the EMG signals. It uses a threshold algorithm. The coefficients of the transforms that are below the threshold are set to zero. This results in:

$$F[k] = \begin{cases} F[k]; if: k \geq \lambda \\ 0; if: k < \lambda \end{cases}, \lambda = \delta \max\{|F[k]|\}_{k=0,1,\ldots,N-1} \tag{9}$$

where *F[k]* is the kth coefficient of the Wavelet transform, and δ is the factor which defines the threshold. Computer simulations have shown that δ must be equal or below 0.01 for high accuracy signal reconstruction. A vector of significant amplitude positions is then constructed using one bit as an identifier, as shown if Fig. 5. If a coefficient is bigger than the threshold, the index position is set to 1 (one) and, otherwise, set to 0 (zero).

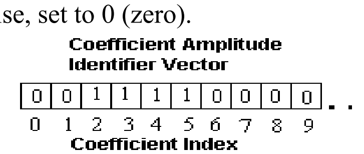

Fig. 5. Coefficient amplitude identifier.

The coefficient amplitude identifier vector is an information overhead that needs to be transmitted or stored. More efficiency is achieved if this information is coded in order to increase the compression ratio. The *significant coefficient vector* needs to be quantized, and the *amplitude identification vector*, which identifies the significant wavelet coefficients is compressed by an entropy coding. Finally, the compressed EMG signal is packed and stored in memory.

The decoding process is quite simple; it implies the unpacking of the coded data, the de-quantization and de-normalization of the significant coefficient vector, the decoding the overhead information and use of this information to restore the position of the significant coefficient vector and, finally, the inverse transform, which results in the reconstructed signal. Figure 6 shows a block diagram illustrating the decoding algorithm.

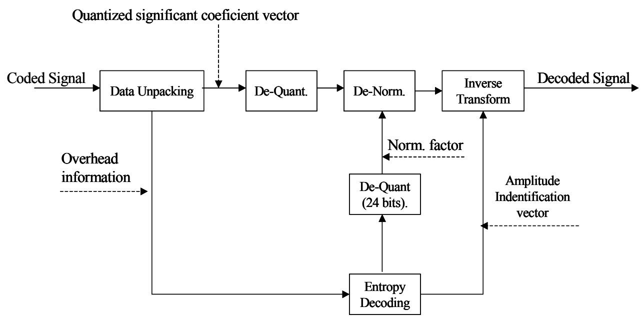

Fig. 6. Block diagram of the decoding algorithm.

III. RESULTS

Experimental results were obtained by processing EMG signals digitized at 2000 samples per second represented with 16 bits per sample. All signals were pre-processed by the de-noising algorithm before applying the coding process. The algorithm was carried out by using the Daubechies-4 functions and a scale resolution of L=8. Fig. 7 and Fig. 8 show simulated examples with the original EMG signal and the reconstructed signal after the decoding of the compressed data.

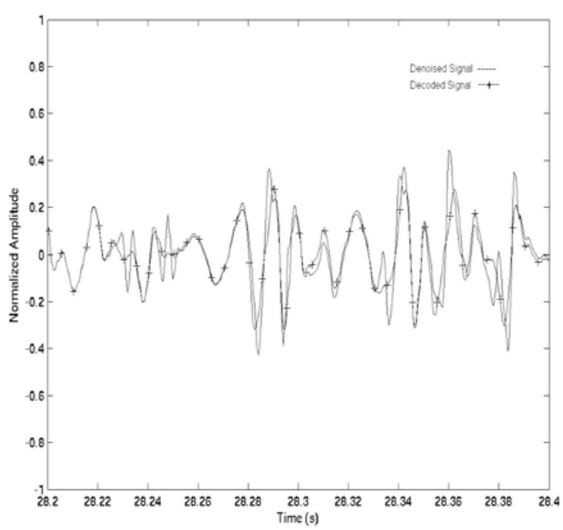

Fig 7. Comparison between the original and the decoded EMG signal. The Compression rate is 16 times and RSR = 17.95 dB.

IV. CONCLUSION

In this paper, we studied the performance of compression of EMG signals by using Wavelet Transform when we vary the parameters such as the mother wavelet function and the scale resolution. It was observed that, in the wavelet functions of Daubechies' family, the function, which presents a better performance is the Daubechies-4. It was also verified that when the resolution scale is increased, the transform performance is improved until L = 8. Scale resolution greater than 8 leads to a large computational effort with no significant improvement in the performance. Performance comparison between DCT and DWT has shown a better performance for the DWT.

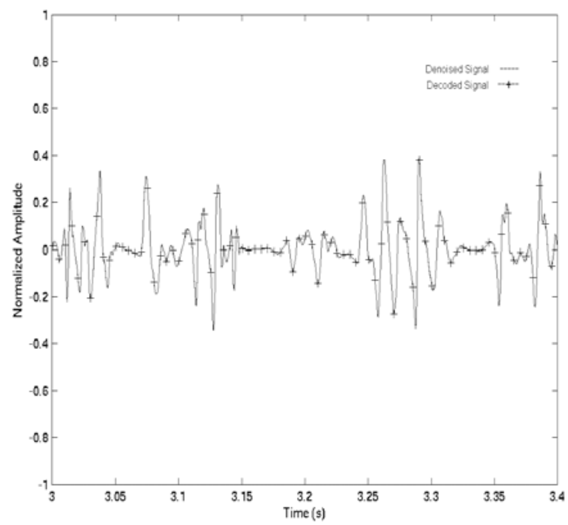

Figure 8. Almost perfect reconstruction. The compression rate is =15 times and RSR = 32.68 dB.

REFERENCES

[1] G. L. Antoniol, P. Tonella, "EEG dates compression technique", *IEEE Trans. On Biomed. Eng.*, vol. 44 (2), pp. 105-114, 1997.

[2] S. M. Jalaleddine, C. G. Hutchens, R. D. Satrattan, W. A. Coberly, "ECG dates compression techniques - the unified approach.," *IEEE Trans. On Biomed. Eng*, vol. 37, no. 4, pp. 329-343, 1990.

[3] A. Guerrero, C. Maihes, "On the choice of an electromyogram data compression method", *Proc. 19th annual International Conference of the IEEE Engineering in Medicine Biology Society*, 1997, pp. 1558-1561.

[4] A. Gersho, R. M. Gray, *Vector quantization and signal compression*, Norwell, NJ: Kluwer Academic Publishers, 1992, pp. 225-252.

[5] C. J. De Lucas, "Physiology and mathematics of myoelectric signal," *IEEE Trans. Biomed Eng.*, vol. 11, no. 4, pp. 251-279, 1979.

[6] P. A. Berger, F. A. O. Nascimento, A. F. Rocha, J. C. Costa, J. C. Carmo, Model for Simulation of Isometric EMG Signals" (in Portuguese), XVIII Brazilian Congress of Biomedical Engineering, São José dos Campos, Publication in magnetic media.