

Multichannel ECG Data Compression Method Based on a New Modeling Method

A Prieto*, C Mailhes**

UAM Iztapalapa, Mexico City, Mexico*, TeSA-ENSEEIH, Toulouse, France**

Abstract

This work addresses the problem of multichannel ECG data compression. The proposed method is based on the identification of a FIR system. The idea of this method is to consider different independent leads (I, II, V1 to V6) as the outputs of linear filters in which the excitation signal is one of these leads suitably chosen.

For purposes of compression, the parameters of the FIR system are quantized and coded.

Compression results are presented in terms of the Signal-to-Noise Ratio (SNR) and Compression Ratio (CR). Results obtained shown a good reconstruction level with a high compression ratio.

1. Introduction

At present, the incessant data generated in medical centers remains a major problem to be resolved. This problem is reinforced with the patients monitoring and the emergent domain of telemedicine. A possibility to solve this problem involves data compression.

One of the most important biomedical signals is the ElectroCardioGram (ECG). Many studies of ECG data compression have been developed, especially about one lead data compression [1]. However, the multilead ECG data compression case has been treated in reduced scale. The reasons are maybe due to the problem difficulty and the complexity of the obtained algorithms. Among these works: subband coding [2], the AZTEC multilead [3], parameter modeling [4] and the use of the Karhunen-Loève Transform [5].

In next sections, a new multilead ECG data compression based on a FIR system identification is presented. Section 2 contains the main idea of the proposed method. A general description of the system identification is presented in section 3. In section 4, the general compression method is outlined and described. The results of this method are presented in section 5. The last section is devoted to the work conclusion.

2. Method basis

The proposed method basis is to give an answer to the question: Considering the fact that a multichannel ECG represents a measure of the same phenomenon (heart electrical activity) captured with different electrodes, is it possible to get a relative simple expression linking one lead to the others?

The answer to this question is positive: a linear relationship between leads II and III with lead I is known. Other relationships are not evident.

The method basis consists of a system where several leads are the outputs of linear filters driven by a lead suitably chosen. In figure 1, lead I is shown as an example of input signal.

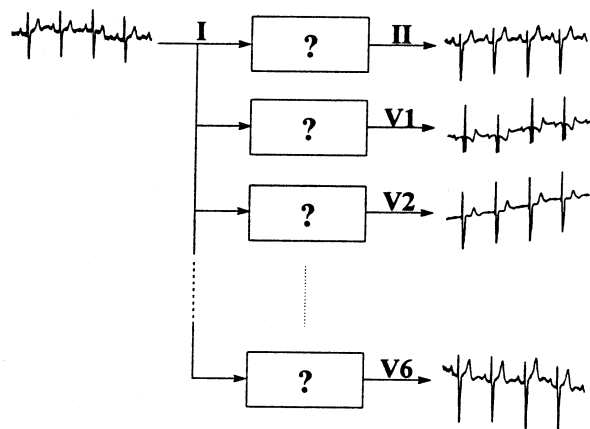


Figure 1. Method basis using system identification concept.

This idea seems attractive from a compression point of view. Indeed, once the filters have been identified is just necessary to code the filters parameters and the input signal to retrieve another leads. A better performance can be obtained considering the coding of the model error.

3. System identification

3.1. Model choice

The choice of a suitable model to give an answer to the proposed question, is not an easy task. Indeed, many models, linear or non-linear, can be considered. In this work, a linear model based on a FIR filter is used. This choice was retained for two reasons: first, it presents a simple expression linking input and output and, second, it has no problem of stability (important situation for compression purposes).

3.2. Identification based on FIR filters

The expression between the input signal $x(n)$ and the output signal $y(n)$ of FIR filters to be identified, is given by:

$$\hat{y}(n) = \sum_{i=0}^{p-1} a_i x(n-i)$$

where a_i represents the FIR filter coefficients and p the filter order. These coefficients correspond to the filter impulse response ($h_i = a_i$).

The problem is to estimate the impulse response of FIR filters from input and output signals data minimizing the error between $y(n)$ and $\hat{y}(n)$. To estimate these coefficients, analysis windows of $x(n)$ and $y(n)$ of length N , with $N \gg p$, are considered. Estimated parameters are given by:

$$\theta = (X'X)^{-1} X'Y$$

where

$$\theta = \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ M \\ a_{p-1} \end{bmatrix} \quad Y = \begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ M \\ y(p-1) \\ \vdots \\ M \\ y(N-1) \end{bmatrix}$$

$$X = \begin{bmatrix} x(0) & 0 & \Lambda & 0 \\ x(1) & x(0) & \Lambda & 0 \\ \vdots & \vdots & \Lambda & \vdots \\ x(p-1) & x(p-2) & \Lambda & x(0) \\ \vdots & \vdots & \Lambda & \vdots \\ x(N-1) & x(N-2) & \Lambda & x(N-p) \end{bmatrix}$$

3.3. Signal input choice

Another important choice is the lead driven filters. That is not a simple task. Indeed, among the independent leads considered (I, II, V1 to V6) just one must be chosen as input signal. In this work, an empirical criterion based on signals energy is used to determine the input signal.

To obtain a better performance, leads were classified in two big classes: first class involves standard leads I and II, second class involves leads V1 to V6. This division implicate that two input signal, one for every class, must be selected. The selected input signal in every class is the lead with the most important energy.

3.4. FIR filter order

It is well known that a better performance of a FIR filter is obtained when the order increases. However, a higher order represents a considerable computation in the inversion of the matrix $X'X$ and importance reduction from a compression point of view. Figure 2 shows the model error $e(n) = \hat{y}(n) - y(n)$ in terms of the FIR filter order p . These results are obtained considering the whole multilead ECG database. Based on these results and considering the compromise: low error power with a low order, the order of all FIR filters was fixed to 100.

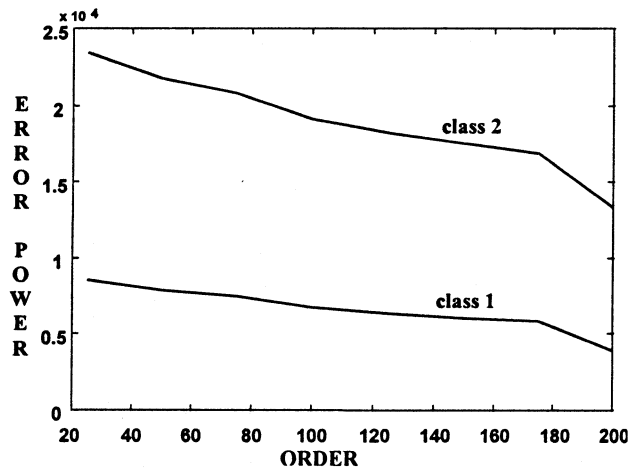


Figure 2. Estimate of the error power in terms of the filter order.

4. General compression scheme

Figures 3 and 4 show the general coder and decoder for the proposed method. In these figures, the input signal of every FIR filter to be identified is not the original signal but a reconstructed signal obtained after a compression process. Thus, the same signal is known simultaneously by coder and decoder with no introduction of additional reconstruction errors.

Given the importance of the input signal, a high reconstruction level must be imposed. Therefore, a Discrete Cosine Transform (DCT) was chosen as compression method of the input signal with a good reconstruction level in average (SNR=27 dB) and a low Compression Ratio (CR=4).

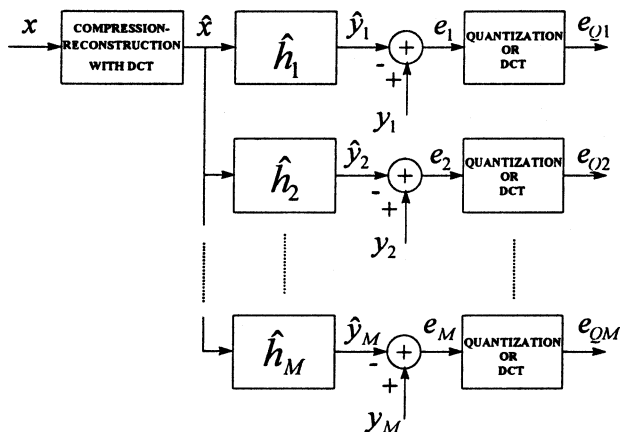


Figure 3. General scheme for the coder.

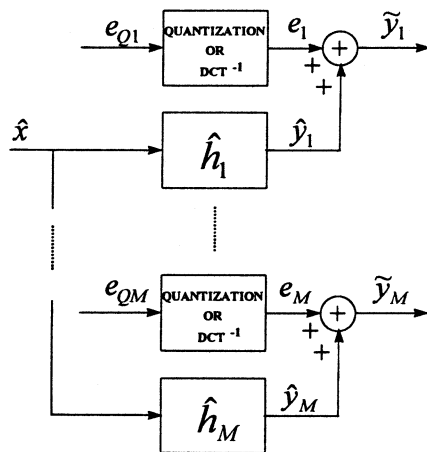


Figure 4. General scheme for the decoder.

In figure 3, the model error is processed in two ways. First way considers the direct quantization of error samples. Second way uses an orthogonal transform, in this case the DCT, applied to the model error. In this last case, transformation coefficients are quantized and coded using a energy criterion. The inverse processes are applied in the decoder (figure 4). These two procedures were implemented based on a study of the model error distribution. Results of this study showed that to obtain a better performance of the proposed method, a model error coding was necessary.

5. Results

The above-mentioned compression method is evaluated using two criteria: the Compression Ratio (CR) and the Signal-to-Reconstruction Noise Ratio (SNR). The measure of the compression ratio is defined as the ratio of bits used to represent the signal before and after compression:

$$CR = \frac{\text{original sample bit number}}{\text{reconstructed sample bit number}}$$

The signal-to-reconstruction noise ratio is given by:

$$SNR = 20 \log \left(\frac{\sigma_x}{\sigma_{x-\hat{x}}} \right) \quad (dB)$$

where $x(n)$ is the original signal, $\hat{x}(n)$ the reconstructed signal after compression and σ a standard deviation estimator.

The proposed method was implemented and evaluated on the CSE multilead data base [6]. As mentioned, only independent leads are concerned: I, II, V1 to V6. The results obtained in terms of compression ratio and SNR are presented in figure 5. In this figure, two cases are shown: first case considers the direct quantization of the model error and second case considers the application of DCT to the model error and next the quantization of these coefficients.

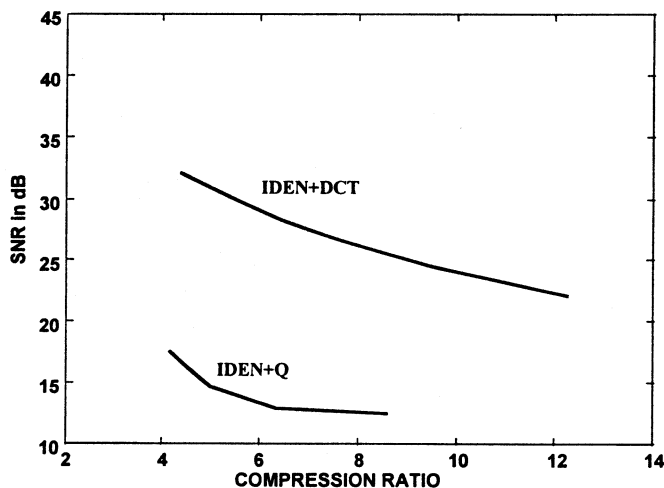


Figure 5. Multichannel ECG compression results.

In figure 5, two remarks can be done: First, the application of DCT to the model error gives better results in relation with direct quantization of the model error. Second, excellent results with DCT are obtained. Indeed, in average, SNR=12 dB with a CR=14, are obtained.

These results with DCT applied to model error, are better than those obtained with direct quantization because the DCT exploits in a better way the redundancy contained still in the model error.

6. Conclusion

In this work a new multichannel ECG data compression is presented. This method is based on the identification of a FIR system where inputs and outputs are the independent leads I, II, V1 to V6 suitably chosen. For the compression purposes, parameters of the FIR system and input signals are quantized and coded.

Results given are encouraging: good levels reconstruction with high compression ratios are obtained. Better results were obtained with the DCT applied to the model error in relation with the direct quantization of this.

Acknowledgements

We want to thank to the Consejo Nacional de Ciencia y Tecnología (CONACyT) and Universidad Autónoma Metropolitana (UAM) in México, for the support to this work.

References

- [1] Jalaliddine SMS, Hutchens CG, Strattan RD, Coberly WA. ECG data compression – A unified approach. *IEEE Trans. on Biomed. Eng.* 1990; 37:329-343.
- [2] Çetin AE, Köymen H, Aydin MC. Multichannel ECG data compression by multirate signal processing and transform domain coding techniques. *IEEE Trans. on Biomed. Eng.* 1993; 40:495-499.
- [3] Mammen CP, Ramamurthi B. Vector quantization for compression of multichannel ECG. *IEEE Trans. on Biomed. Eng.* 1990; 37:821-825.
- [4] Cohen A, Zigel Y. Compression of multichannel ECG through multichannel long-term prediction. *IEEE Eng. In Med. and Biol.* 1998; 109-115.
- [5] Olmos S, Echevarría D, Jané R, Laguna P. Compresión de señales ECG multiderivacionales mediante expresiones ortogonales óptimas. *Comm. del XV Congreso Anual de la Soc. Esp. de Ing. Biomed.* 1997.
- [6] Commission of the European Communities: Medical and Public Health Research. Common standards for quantitative electrocardiography, CSE Multilead Atlas. 1988, Leuven.

Address for correspondence.

Alfonso Prieto-Guerrero.
Universidad Autónoma Metropolitana, Unidad Iztapalapa
Departamento de Ingeniería Eléctrica
Av. San Rafael Atlixco 186
Col Vicentina, C.P. 09340
México, D.F., México.
apg@xanum.uam.mx