Methods to Evaluate the Performance of Fetal Electrocardiogram Extraction Algorithms

M Martínez, J Calpe, E Soria, JF Guerrero, G Camps, L Gómez

Grupo de Procesado Digital de Señales, Dpto de Ingeniería Electrónica, Universitat de València, Valencia, Spain

Abstract

One of the main problems for the evaluation of fetal electrocardiogram (FECG) extraction algorithms is the difficulty in giving objective measurements of the quality of the recovered signals, and therefore of the proposed extraction method. This is the reason why most of the papers determine the behavior of a certain algorithm by visual inspection criteria. We propose the use of simulated registers and the definition of indexes that determine the quality of an extraction algorithm, without requiring visual inspection. This way we will be able to have objective measurements to determine the quality of an algorithm or we will have a criterion to fix the operation parameters of the method that show better performance

1. Introduction

Defining objective measures for establishing the quality of a given procedure is one of the main problems found in the inverse problem in electrocardiography; i.e., estimating the source signals from the registers acquired with surface electrodes. This is the reason why, most papers on fetal ECG extraction, propose visual inspection of the waveforms as the only technique to evaluate the performance of the proposed procedures, making it difficult to compare different approaches to the problem. An additional difficulty is the lack of a specific ECG database to assure a fair comparison among the results.

Another implication of this lack of numerical evaluation criteria is the difficulty to apply extraction methods that contain parameters that must be adjusted, since visual inspection is both slow and subjective. Our proposal is to use simulated registers and define objective indexes in order to measure the quality of the FECG signal recovery and so, fix the optimal parameters of the extraction methods and compare among different procedures. The aforementioned indexes determine the similarity between the original signals given by the simulator and those recovered by the proposed method.

In order to validate the usefulness of the approach, we have to show if these parameters are suitable to determine the quality of the signals and that they show an homogenous behavior for the complete set of signals generated by our simulator.

2. Material and methods

2.1. Indexes definition

The measurement of the quality of the FECG extraction procedures is based on the estimation of the similarity of the recovered signals and those original records used to generate the synthetic inputs to the FECG simulator. The measure indexes proposed are those commonly used in data compression literature [1]. There, the two main parameters to compare the similarity between the original and recovered signal are the Per Cent Root-Mean Square Difference (PRD) and the Correlation Coefficient (COEF) whose definitions are shown in (1) and (2).

$$PRD = \frac{\sum_{i=1}^{N}\{x_{ori}(i) - x_{rec}(i)\}^2}{\sum_{i=1}^{N}\{x_{ori}(i)\}^2}\cdot 100 \qquad (1)$$

$$\rho = \frac{\sum_{i=1}^{N}\{x_{ori}(i) - \overline{x_{ori}}\}\{x_{rec}(i) - \overline{x_{rec}}\}}{\sqrt{\sum_{i=1}^{N}\{x_{ori}(i) - \overline{x_{ori}}\}\sum_{i=1}^{N}\{x_{rec} - \overline{x_{rec}}\}}} \qquad (2)$$

where sub-index *ori* denotes the original signals and *rec* the signal recovered by the algorithm.

The correlation coefficient takes values in the range $-1 \leq \rho \leq 1$ and is an indicator of the morphologic similarity between two signals; i.e. the amplitude of each signal does not affect the result.

The PRD is more a troublesome parameter to work with. The closer this value is to zero the similar are both signals. But the lack of an upper limit makes it difficult to compare figures from different pairs of signal, as it

0276-6547/01 $17.00 © 2001 IEEE

depends from the amplitude of the signals. In order to avoid this, and enable the comparison among pairs of signals, these must be normalized[1] previous to obtain their PRD.

A third parameter to evaluate the similarity between signals is the signal to noise ratio (SNR), whose expression is (3),

$$SNR = 10 * \log_{10}\left(\frac{S}{N}\right) \quad (dB) \tag{3}$$

In the general FECG inverse problem, this is not an operative definition for the SNR, since it requires knowing the contribution of the fetal signal and the noise. However, in our case, this information is available in synthetic registers obtained from the simulator.

In generating the simulated records, we have made a distinction among different types of noise sources to determine the contribution due to each of them. Nevertheless, the SNR figure quantifies the relation between the fetal signal and the rest of the unwanted components (maternal contribution, Gaussian noise, etc.).

The application of any parameter to evaluate the goodness of any extraction method is not simple when real registers are used. Indexes such as COEF and PRD are not applicable since the source signals are unknown; in fact, their obtaining is the aim of the problem. The only estimator used in the bibliography is the SNR of the recovered signal, but this is an indicator of the quality of the signal rather than the goodness of the extraction method. Therefore, the aforementioned definitions of the parameters are not applicable in a direct way. Two alternative procedures are found in [2], and [3], and both consider that the extracted register whose SNR is being estimated, only contains fetal components and uncorrelated noise. If this hypothesis is accepted, the signal is segmented in pulses, which are arranged as columns of a matrix. Then, its singular value decomposition is performed, and the SNR figure is obtained as the quotient between the first singular value (fetus signal energy) and the sum of the rest singular values (noise energy), as it is shown in expression (4)

$$SNRsvd = \sqrt{\frac{\sigma_1^2}{\sum_{i=2}^{N} \sigma_i^2}} \tag{4}$$

where, σ_i are the singular values of the matrix. The main drawback of this estimation is that it tends to overestimate the SNR, especially when the fetal component is very low and the maternal contribution has not been completely cancelled.

[1] This normalization implies that the mean value is forced to be zero and the standard deviation one.

If the rows of the previous matrix are normalized, and we assume that the repetitive component of the signal and the noise are not correlated, the SNR between two pulses $\mathbf{x}(i)$, $\mathbf{x}(j)$ can be defined by (5).

$$SNRcor = \sqrt{\frac{\mathbf{x}(i)^T \mathbf{x}(j)}{1 - \mathbf{x}(i)^T \mathbf{x}(j)}} \quad i \neq j \tag{5}$$

If all available pulses (N) are averaged, the following expression is obtained:

$$SNRcor \approx \sqrt{\frac{S}{1-S}} \quad S = \frac{2}{N(N-1)} \sum_{i=0}^{N-2} \sum_{j=i+1}^{N-1} \mathbf{x}(i)^T \mathbf{x}(j) \tag{6}$$

2.2. Records

The records used in this study contain one thoracic and one abdominal derivation sampled at 1KHz. They have been obtained with a simulator that enables the setting of parameters such as fetal to mother signal to noise ratio (SNRfm), fetal to Gaussian noise signal to noise ratio (SNRfr), fetal to electromyogram signal to noise ratio (SNRfe), presence of base line oscillations and mains interference. In order to generate the registers, real pulses obtained from non-pregnant females were repeated. The SNRfm spans from –4 dB to –30 dB, and the SNRfr from 0dB to 30dB. No electromyogram contribution was considered. Previous to use any of the registers, they were pre-processed to reduce base line wander by using the algorithm described in [6].

2.3. Indexes estimation

The algorithm that has been used to cancel maternal contribution is the one described in [4]. The algorithm has an adjust parameter whose range is]0,1[. The trials have been done for 20 different values. Using the extracted registers obtained by the algorithm and the original noiseless signal provided by the simulator, six different indexes are calculated:

a. Total correlation coefficient (coef).
b. Average correlation coefficient (coef_m).
c. Total PRD (prd).
d. Average PRD (prd_m).
e. SNR quotient calculated according to (4) (SNRcoc1).
f. SNR quotient calculated according to (6) (SNRcoc2).

Parameters e and f are the quotient of SNR parameters before and after maternal cancellation as shown in expression (6)

$$SNRcoc = \frac{SNRtotal\_rec}{SNRtotal\_ori} \tag{6}$$

One-minute registers have been generated, but, in order to avoid the effects of algorithm's settling transient states, the parameters have been estimated considering the segment between 30 and 50 seconds (around 30 beats).

An additional comment must be made about the aforementioned parameters. Even though, in all cases the recovered and original fetal signals are compared, two procedures have been followed to estimate PRD and COEF. The first considered the complete 20-second tract, and the second just took into account a time window around all the fetal complexes, as shown in figure 1. The reason for this second approach is that this region is more heavily affected by a faulty operation of the cancellation algorithm.

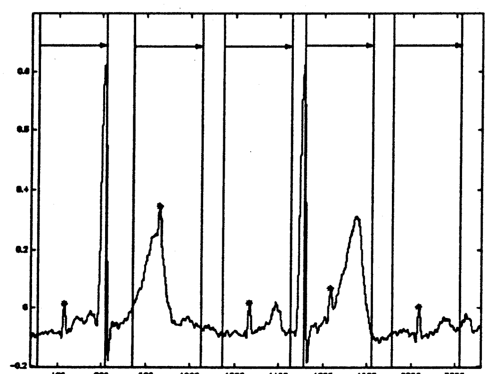

Fig 1. Intervals where parameters in pulse-to-pulse measurements are evaluated.

3. Results

Firstly, COEF and PRD parameters obtained pulse-by-pulse and with the complete 20-second tract have been compared. In order to make a sound comparison, different situations must be considered, therefore, in a first step, SNRfr was fixed and made a sweep on SNRfm, afterwards, SNRfm was fixed and simulations were carried out changing the contribution of the Gaussian noise. The results are similar for all cases as show in figures 2 to 4.

X-axis contains the 20 values given to the parameter that defines the behavior of the extraction algorithm considered. Higher values for the correlation coefficient indicate a better performance of the algorithm.

A first conclusion suggested by these figures is that both strategies to obtain the parameters; i.e. pulse-by-pulse and complete registers, show the same trend. So we conclude that both ways are equally significant, even though the pulse-by-pulse option yields higher indexes.

Another conclusion shown by figures 3 and 4 is that an increase in COEF entails a decrease in the PRD values.

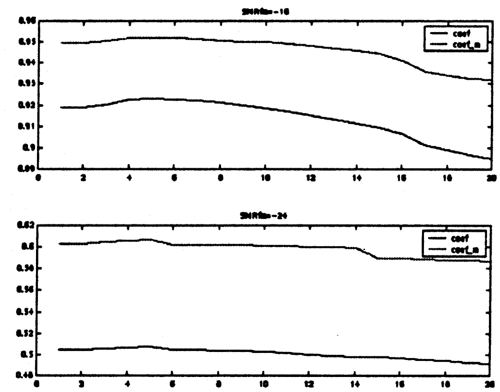

Fig 2. Comparison between the average and total correlation coefficient as SNRfm varies.

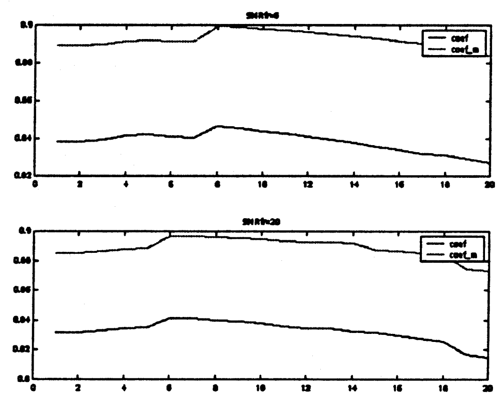

Fig 3. Comparison between the average and total correlation coefficient as SNRfm varies.

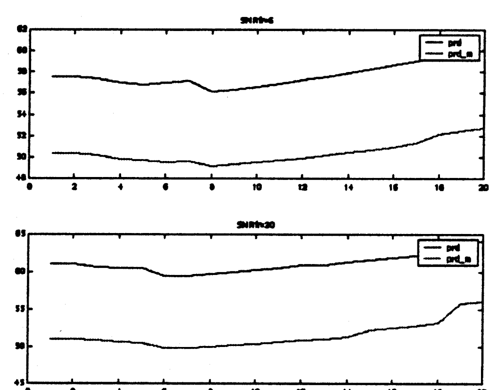

Fig 4. Comparison between the average and total PRD as SNRfr varies.

SNRcoc figures are always higher than unity. Since the desired signal (fetal component), has lower amplitude than the maternal, the application of a cancellation algorithm must improve the original SNR.

In order to check whether both SNR estimation procedures give the same results, figure 5 shows the outcomes obtained for SNRcoc1 on the X-axis and SNRcoc2 on the Y-axis.

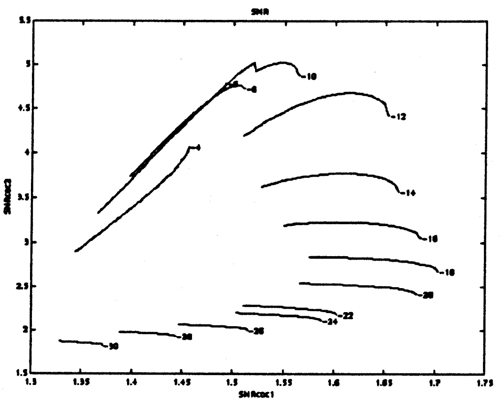

Fig.5. SNRcoc obtained with simulated registers varying SNRfr

The expected result for two equivalent procedures should be a straight line. As this is not the case, we conclude that these parameters highly depend on the characteristics of the signals so they are not robust enough to be considered.

4. Conclusions

This paper proposes several parameters to quantify the performance of the algorithms proposed to extract fetal ECG registers. The fittest parameter seems to be the correlation coefficient calculated over the whole register. It gives the same information than the correlation coefficient calculated on a pulse-by-pulse basis, that requires the segmentation of signal in pulses and applying time-windows. Contrary to the PRD, its range is defined between [-1, 1]. SNR parameters are interesting because their estimation do not require knowing the source fetal signals, and thus they can applied to real registers. Despite this, they did not show acceptable performance because of the difficulty to define robust and confident bearing estimates.

The proposed method has been applied to fix the operating parameters of the maternal cancellation algorithms described in [4] and [5]. In the latter, an artificial neural network with a high number of weights must be trained so, establishing a numerical evaluation index is extremely useful. The use of this figure of merit and simulated registers with tunable parameters are proposed as a tool to compare different approaches or algorithms to the FECG extraction problem

References

[1] Jalaleddine S.M.S., Hutchens C., Strattan R., Coberly W. "ECG data compression techniques. A unified approach", IEEE Trans on Biomed Eng, 1990; Vol. 37, pp. 329-343.

[2] Kanjilal P., Palit S., Saha G. "Fetal ECG extraction from single channel maternal ECG using singular value decomposition", IEEE Trans on Biomed Eng, 1997; Vol. 44, pp. 51-59.

[3] Outran N. J. "Intelligent Pattern Analysis of the Foetal Electrocardiogram", PhD Thesis. University of Plymouth. 1997

[4] Martínez M. ,Soria E., Calpe J., Guerrero J., Magdalena J.R. "Application of the Adaptive Impulse Correlated Filter algorithm to recover Fetal Electrocardiogram", Computers in Cardiology, 1997;Vol. 24, pp. 9-12.

[5] Camps G., Martínez M., Soria E. "Fetal ECG Extraction using an FIR Neural Network", Computers in Cardiology, 2001; Vol. 28.

[6] Martínez M., Martín J.D., Calpe J., Guerrero J., Soria E., Muñoz J. "Removing Base Line Wander and Power Line interference in non-invasive FECG registers.", Chicago 2000 World Congress on Medical Physics and Biomedical Engineering. 2000

Address for correspondence.

Marcelino Martínez Sober
Dpto. Ingeniería Electrónica. Facultad de Física
Dr. Moliner, 40. Burjassot
46100 Valencia (SPAIN)
marcelino.martinez@uv.es