Combined Analysis of Time and Frequency Series Regularity
Applied to the Study of Atrial Fibrillation

C Vayá¹, JJ Rieta²

¹Innovation in Bioengineering, Universidad de Castilla-La Mancha, Spain
²Biomedical Synergy, Universidad Politécnica de Valencia, Spain

Abstract

Atrial Fibrillation (AF) is the most common arrhythmia encountered in the clinical practice. In this work, a new method based on electrocardiogram (ECG) signal processing is carried out in order to distinguish between AF episodes that will terminate immediately and those that will sustain. This new method is based on a combined analysis of the atrial activity (AA) series regularity in both time and frequency domains. The regularity is measured by using the non-linear regularity estimator Sample Entropy (SampEn). The SampEn of these series of spectral parameters and the SampEn of the AA in the time-domain are studied jointly in a discriminant analysis. In global percentages, forty eight of fifty recordings (96%) were classified correctly by the discriminant analysis, which provided an improvement of 6% with respect to the univariate analysis of SampEn.

1. Introduction

Atrial Fibrillation (AF) is the most common arrhythmia encountered at advanced age. The prevalence of AF remains lower than 1% among the general population, but it increases considerably from sixty years old [1]. The spontaneously terminating AF, frequently refereed as paroxysmal AF (PAF), is often the precursor to sustained AF [2]. Given that sustained AF increases the likelihood of suffering myocardial infarctions and strokes [1], its accurate recognition by means of noninvasive techniques is of great interest to the regular clinical practice. The prediction of PAF maintenance can help to choose the appropriate intervention that may terminate the arrhythmia. Otherwise, the prediction of the spontaneous termination of PAF episodes could avoid unnecessary therapies and their associated clinical costs [3].

In this work, an analysis of AA spectral parameters organization is carried out with the aim to classify between terminating and non-terminating AF episodes. The analysis of these parameters is made in terms of mathematical regularity of their series. The organization is measured by using the entropy estimator Sample Entropy (SampEn) [4, 5]. Entropy estimators have already been used in the characterization of biomedical signals different from ECG [6]. The ECG recordings analysis is completed in five main steps: extraction of the AA, computation of the spectrogram, curve fitting, construction of spectral parameter series, and SampEn computation. The SampEn values are then studied by using univariate and discriminant analyses.

2. Materials

For the present work we have used the database of PhysioNet/Computers in Cardiology Challenge 2004. The database is divided into a learning set and a test set. Each record in the database is a one minute segment of AF that has been extracted from a long term ECG recording. The learning set contains 10 segments of non-terminating AF (group N) and 10 segments of AF that terminates within a second after the end of the record (group T). The test set contains 30 records, approximately half of them belonging to group N, and the rest to group T.

Butterworth filtering of eight order and pass-band from 1 to 50 Hz is applied to each recording. The original sampling rate (fs) of the Holter system was 128 samples per second, but the ECG recordings were interpolated by a factor of 8 so that a fs equal to 1024 resulted. The resultant time-domain higher resolution allowed us to obtain a better cancellation of the QRS complex and a higher length of parameter sequences.

3. Methods

In order to use the ECG as a suitable tool for the analysis of AF, we need to separate the atrial activity from other cardioelectric signals. The extraction of the AA during AF requires nonlinear signal processing since spectra of atrial and ventricular activities (VA) overlap and, in consequence, they cannot be separated by simple linear filtering [7]. Given that the challenge database consisted of
two-lead ECG recordings, we opted to use ABS to separate the AA from the rest of the cardioelectric signal as a previous step to the rest of the analysis [8, 7].

After the obtention of the AA signal, its spectrogram [9] is computed using Hamming windows of 1024 samples in length and 75% overlap. In order to facilitate the spectral parameters extraction, cubic spline fitting is applied to each of the Fourier transforms that constitute the spectrogram. The cubic spline fitting curve from the original data is interpolated so that the resulting frequency increment is 0.01 Hz. Finally, the SampEn of all aforementioned series is obtained as an estimation of their mathematical regularity.

The size of series is around 600 elements, what is large enough since the SampEn is meaningfully applied to more than 100 data points [4]. The SampEn values are evaluated by the t-test and by discriminant analysis, which results are exposed in section 4. The objective pursued by the discriminant analysis is to know if there exists any combination of the aforementioned parameters which improves the decision reliability. The discriminant analysis is carried out in two stages. First, the discriminant function is obtained by considering the learning data set. Then this discriminant line is applied to the test set in order to validate the results.

The first spectral parameter we consider is the main peak frequency (f_p1), which is known to be highly relevant in the characterization of AF [10]. The second parameter is the main peak magnitude, A_1. The third and fourth parameters are the second largest frequency peak (f_p2) and its related peak magnitude, A_2. The fifth parameter is the Spectral Concentration (SC). The SC can be defined as [11]:

\[ SC = \frac{1.17 f_p}{\sum_{f=0.82 f_p}^{1.17 f_p} P_{AA}(f)} \cdot \frac{\sum_{f=0.5 f_p}^{0.5 f_s} P_{AA}(f)}{\sum_{f=0.5 f_p}^{0.5 f_s} P_{AA}(f)} \]

where \( P_{AA} \) is the power spectral density of the AA signal, \( f \) is the frequencies vector, \( f_s \) is the sampling rate (1024 Hz), and \( f_p \) is the main peak frequency of the AA. Other two parameters related to the width of the spectrum main lobe have been used: the 3-dB width of the peak, \( w_{3dB} \), and the power in the 3-dB band, \( P_{3dB} \). This two last parameters have been used in [12] to characterize AF. Two derived parameters, \( \Delta f_p \) and \( A_2 \) are referred to the spectral shape of AA. Similar parameters are used in [13]. The first derived parameter is the normalized distance between \( f_p1 \) and \( f_p2 \), which is expressed as:

\[ \Delta f_p = \frac{(f_p1 - f_p2)}{f_p1} \]

The second derived parameters is the normalized amplitude of the second largest peak, which is defined as:

\[ \bar{A}_2 = \frac{A_2}{A_1} \]

The deviation of the main and second peak magnitudes from their respective mean values are also computed as a dispersion measurement:

\[ d_1 = f_p1 - E(f_p1) \]

\[ d_2 = f_p2 - E(f_p2) \]

where \( E(\cdot) \) represents the average value over the set of periodogram. Finally, the Median Frequency (MF) is obtained as the center of mass of the spectrum:

\[ MF = \frac{\sum_{f=0}^{0.5 f_s} |FT_{AA}(f)| \cdot f}{\sum_{f=0}^{0.5 f_s} f} \]

where \( FT_{AA} \) is the Fourier Transform of the AA. This parameter was previously used in other works to characterize the ventricular fibrillation [14].

4. Results

The results of the t-test applied to the SampEn of the numerical series for the learning set are summarized in Figure 1. These results reveal that it is possible to distinguish between terminating and non-terminating AF in six of the twelve parameters, considering a parameter to be relevant when its bilateral significance is less than 0.05. These six relevant parameters are \( f_p1, f_p2, \Delta f_p, A_1, d_1 \) and \( SC \), which bilateral significances are, respectively, 0.001, 0.005, 0.003, 0.004, 0.015, and 0.001. The mean SampEn in type N recordings is higher than in type T recordings for all these relevant parameters. An optimal decision threshold of 0.1173 has been chosen for \( f_p1 \). By considering this value of threshold, 19 out of 20 learning recordings have been classified correctly. Taking the same threshold for the test set, 26 out of 30 recordings have been classified correctly. This resulted in a percentage of recordings properly classified equal to 95% for the learning set and equal to 86.67% for the test set. The results obtained by this classification are presented in Figure 2 for every recording.

The previous univariate analysis revealed that the SampEn of the spectral parameters \( f_p1, f_p2, \Delta f_p, A_1, d_1 \) and \( SC \) have a bilateral significance lower than 0.05 and, in consequence, all of them are suitable to be used
mean SampEn

(bilateral significance)

Figure 1. Results of the t-test for the SampEn of all the spectral parameters. a) Mean SampEn for groups N and T, b) SampEn bilateral significance between groups.

in discriminant analysis, which is the next step that has been taken. Furthermore, the SampEn of the AA was also computed and a mean difference of 0.25081, with the greatest mean value for the N group, and a bilateral significance equal to 0.004 were figured out by the t-test. This fact suggested to include this parameter in the discriminant analysis along with the spectral parameters. This made possible, in addition, to combine the information obtained from both time and frequency domains.

The obtained discriminant function is a plane given by the equation $x_3 = 0.0355 \cdot x_1 + 1.6 \cdot x_2 + 0.4653$, where $x_1$, $x_2$ and $x_3$ represent the SampEn of $f_{p1}$, $\Delta f_p$ and the AA, respectively. The standardized canonical coefficients of the discriminant function are presented in Table 1. These coefficients are ordered in the table by their importance in the discriminant function. A small value of Wilk’s lambda test [15] significance ($p < 0.001$) was obtained, which indicates the great discriminatory ability of the function. All of the cases used to create the model, i.e. the learning set, were correctly classified. Regarding the test set, 15 out of 16 type N cases and 13 out of 14 type T cases were correctly classified (see Table 2). Expressing this results in percentages, 100 % cases of the learning set and 93.75 % cases of the test set were classified properly.

The discriminant analysis has provided an improvement of the results with respect to the classification by threshold (5% for the learning set and 7.08% for the test set). For that reason, it is worth considering the discriminant analysis in predicting the evolution of AF because this improvement in the classification of AF could be of great importance in the regular clinical practice.

To sum up, a new method based on the mathematical regularity of spectral parameters has been introduced as an original and improved way to predict the evolution of paroxysmal AF episodes. From the results we can deduce that the future evolution of AF affects not only to the values of spectral parameters but also to their variability in time. The SampEn of the spectral parameters is higher for the non-terminating than for the terminating episodes. The spectral parameters mathematical regularity might be associated with the physiological organization of the atrial activation. This good results make this new method a useful tool that can help clinicians in the management of AF.

Table 1. Standardized canonical discriminant function coefficients from the stepwise analysis.

<table>
<thead>
<tr>
<th>Canonical coefficient</th>
<th>$SampEn$ of $f_{p1}$</th>
<th>1.880</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SampEn$ of $f_{p1}$</td>
<td>1.880</td>
<td></td>
</tr>
<tr>
<td>$SampEn$ of AA</td>
<td>1.662</td>
<td></td>
</tr>
<tr>
<td>$SampEn$ of $\Delta f_p$</td>
<td>0.816</td>
<td></td>
</tr>
</tbody>
</table>

5. Conclusions

The discriminant analysis has provided an improvement of the results with respect to the classification by threshold (5% for the learning set and 7.08% for the test set). For that reason, it is worth considering the discriminant analysis in predicting the evolution of AF because this improvement in the classification of AF could be of great importance in the regular clinical practice.

To sum up, a new method based on the mathematical regularity of spectral parameters has been introduced as an original and improved way to predict the evolution of paroxysmal AF episodes. From the results we can deduce that the future evolution of AF affects not only to the values of spectral parameters but also to their variability in time. The SampEn of the spectral parameters is higher for the non-terminating than for the terminating episodes. The spectral parameters mathematical regularity might be associated with the physiological organization of the atrial activation. This good results make this new method a useful tool that can help clinicians in the management of AF.
Table 2. Type N and T correctly classified recordings for both learning and test sets by using the discriminant analysis.

Acknowledgements

This work was partly supported by the projects PAID-04-08 and TEC2007–64884 from the Spanish Ministry of Education and Science.

References


Address for correspondence:
Carlos Vayá
Innovation in Bioengineering.
Universidad de Castilla-La Mancha.
Camino del Pozuelo, s/n, 16071, Cuenca.
carlos.vaya@uclm.es