Comparative Study of Nonlinear Metrics to Discriminate Atrial Fibrillation Events from the Surface ECG

M Julián[1], R Alcaraz[2], JJ Rieta[1]

[1]Biomedical Synergy, Electronic Engineering Dept., Universidad Politécnica de Valencia, Spain
[2]Innovation in Bioengineering Research Group. University of Castilla-La Mancha, Cuenca, Spain

Abstract

The atrial activity (AA) of surface ECGs in atrial fibrillation (AF) has been studied to evaluate and compare the ability of several nonlinear metrics in the discrimination of AF events by estimating their regularity. The AA signals were extracted by applying an adaptive QRST cancelation method. Next, the study of surrogate data was performed and revealed the nonlinear behavior of AA. The following nonlinear metrics were studied: Sample Entropy (SampEn), Fuzzy Entropy, Spectral Entropy, Lempel-Ziv Complexity, Hurst Exponents, Generalized Hurst Exponents and AA average power. The Dominant Atrial Frequency (DAF) was added to the study as a reference. To reduce noise and ventricular residues, the Main Atrial Wave (MAW) was obtained applying a selective filtering to the AA signal centered on the DAF. Next, the MAW regularity was estimated using the aforementioned parameters. As conclusion, the study of MAW regularity improved results obtained from direct AA regularity estimation in all the analyzed metrics. Furthermore, some of them achieved better results than SampEn, a widely used metric in the classification of organization-related events in AF. Therefore, they could be considered as promising tools in the quantification of AF events from the surface ECG.

1. Introduction

Non-linear metrics allow to obtain information related to the underlying mechanisms of physiological processes [1]. Some metrics like Sample Entropy [2] and Fuzzy Entropy [3] are related to the existence of similar patterns in the time series; the Lempel-Ziv Complexity [4] is a measurement of the arise of different patterns in a symbol sequence obtained by encoding the signal using a finite symbol set. Other metrics like the Hurst exponents [5] and the Generalized Hurst exponents [6] measure the existence of long-term self-dependencies in the signal. Finally, Spectral Entropy [7] is related to the spectral complexity of the signal.

In the present work the usefulness of different non-linear regularity metrics applied to AF signals with organization-dependent events has been assessed. There are no previous studies comparing different regularity indices using the same data set. With the aim of performing an unbiased evaluation, a freely available database was analyzed.

2. Materials

The AF Termination Database available in Physionet [8], which contains 80 recordings of paroxysmal AF, was analyzed. The signals are labeled as non-terminating (N) which lasted at least 1 hour in AF after the end of the recording, immediately-terminating (T), whose termination happened 1 second after the end of the recording and soon-terminating (S), which terminated 1 minute after the end of the recording [9].

For the analysis lead V1 was chosen and the signals upsampled to 1 kHz to obtain better alignment in QRST subtraction. All signals were preprocessed to improve the analysis. First, baseline wander was removed making use of bidirectional high pass filtering with 0.5 Hz cut-off frequency. Then, high frequency noise was reduced using an eight order IRR Chebyshev low pass filter with 70 Hz cut-off frequency. Finally, Powerline interference was removed with an adaptive notch filter, which preserves the ECG spectral information [10]. Then, the Atrial Activity (AA) was extracted by applying an adaptive QRST cancelation method described in [11].

3. Methods

3.1. Previously analyzed metrics

The metrics analyzed in previous works include the Dominant Atrial Frequency, the *f* wave average power and Sample Entropy. The Dominant Atrial Frequency (DAF) is defined as the frequency corresponding to the highest Power Spectral Density (PSD) amplitude in the 3–9 Hz band and it is considered as an indirect non-invasive AF organization estimator [12]. To obtain the DAF the PSD was computed using the Welch Periodogram with a 4096-points Hamming window, 50% overlapping between adjacent sections and 8192-points Fast Fourier Transform.

The average power of the f waves (fWP), which is defined as the root-mean-squared value of the normalized signal amplitude, represents the energy carried by the f waves within the time interval under analysis. This average power has demonstrated its ability to behave as a robust indicator of the signal amplitude in previous studies [13]. The signal normalization, referred to the R-peak amplitude, avoids all the effects that could affect the amplitude of the ECG during recording like, for example, the skin conductivity or the electrode impedance.

Sample Entropy (SampEn) is a widely used AF organization metric [14]. SampEn estimates the regularity of a time series, understood as the probability that two sequences similar for m points remain similar for $m + 1$ points [2]. Larger values of SampEn are related to more complex sequences. A mathematical description of SampEn can be found in [14]. Two input parameters must be specified for SampEn: a run length m and a tolerance r. In this work, the chosen values were $m = 2$ and $r = 0.35$ times the standard deviation of the signal [15].

Previous studies have shown noise and ventricular residues in the AA signal degrades the performance of non-linear regularity metrics [16]. To reduce these nuisance signals the Main Atrial Wave (MAW) was extracted by linear phase FIR selective filtering the AA with a 3 Hz bandwith centered on the DAF [16]. SampEn was computed both on the AA and the MAW. Since the MAW allows to reduce the presence of noise and ventricular residues in the signal, the average power of the MAW (MAWP) was also computed in order to evaluate its performance in the classification of events in AF.

3.2. New studied nonlinear metrics

The new studied nonlinear metrics were Spectral Entropy, Fuzzy Entropy, Lempel-Ziv Complexity, Hurst Exponents and Generalized Hurst Exponents. Spectral Entropy (SEn) is a measure of complexity in the spectral domain and is defined as the Shannon's Channel Entropy applied to the normalized PSD of the signal [7]. High values of SEn indicate high spectral complexity. In the present work SEn was computed in the 3–9 Hz band.

Fuzzy Entropy (FuzzEn) estimates the regularity of a time series by computing the similarity degree, defined by a fuzzy function, between two segments of the signal [3]. FuzzEn, and with the remaining features that will be detailed, were computed both on the AA and the MAW. The fuzzy function selected was $D_{ij}^m = \exp(\frac{-d_{ij}}{r})^n$ as proposed by Chen et al. [3]. Since no guidelines exist for optimal selection of the parameters r, m and n to study AF through FuzzEn, the chosen values were $m = 2$, $r = 0.25$ times the standard deviation of the time series as suggested by Pincus [1] and $n = 2$ as suggested by Chen et al. [3].

Lempel-Ziv Complexity (LZC) provides a measure of complexity related to the arise of different substrings along a given sequence [4], larger values of LZC correspond to more complex series. To compute the LZC the signal is codified as a finite symbol sequence. The algorithm applied for calculating LZC is described in [17]. Two different sequence conversion methods were considered: two-symbol and three-symbol sequence conversion.

Hurst Exponents (HE) are interpreted as a measure of the scaling behavior of fractal curves and long-term correlations [5]. HE might provide information about the existence of activation patterns in the atria during AF. The use of HE in the study of AA signals regularity was proposed by Mainardi and Sassi in [18]. In the present study HE was computed over all 1 second segments of the signal and then the average value for each recording was obtained.

Generalized Hurst Exponents (GHE), $H(q)$, are associated with the scaling behavior of the q-order moments of the distribution of increments [6]. In the case of unifractal processes $H(q)$ is independent of q and is equal to the Hurst Exponent for all $q > 0$. When $H(q)$ is not constant the process is called multifractal or multiscaling. $H(q)$ has different interpretations depending on the value of q. In the present study $H(1)$ and $H(2)$ have been computed to study their performance in AA organization assessment.

3.3. Statistical analysis

First, the distributions of the computed metrics were tested by the Kolmogorov-Smirnov and Levene tests. Since the distributions were not normal and homoscedastic, the Kruskal-Wallis test was performed in order to evaluate the differences between groups. A two-tailed of $p < 0.01$ was considered as statistically significant. When a statistically significant difference between groups was found a Mann-Whitney U-test was performed on each pair of groups.

Next, the ROC curves were calculated to determinate the classification thresholds for each of the different cases considered. Finally, leave-one-out cross-validation was applied to the parameters that showed statistically significant differences between groups in order to improve the consistency of the classification results.

In order to check the non-linearity of the AA signals the study of surrogate data was performed. 40 surrogate signals were generated from each original AA by the iterative amplitude adjusted Fourier Transform algorithm described in [19]. The analyzed statistics were SampEn, HE and LZC. The Wilcoxon T test was performed in order to compare the distributions of the indices calculated from the original and surrogate signals.

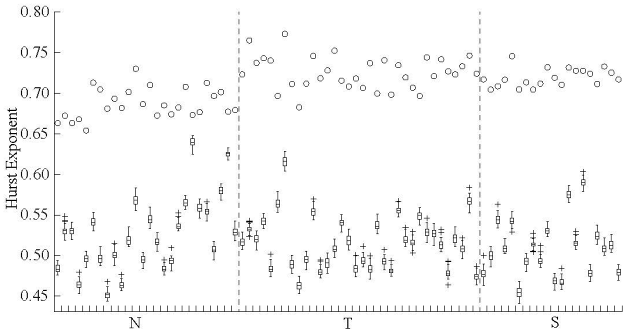

Figure 1. Hurst exponent values for original signals and its corresponding surrogate data sets (boxplot)

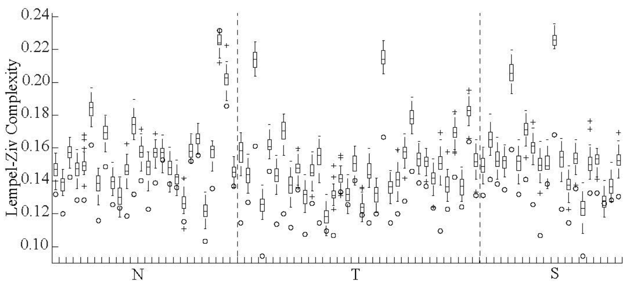

Figure 2. Lempel-Ziv Complexity values for original signals and its corresponding surrogate data sets (boxplot)

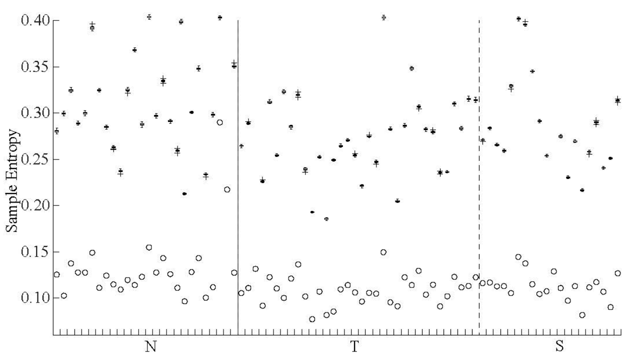

Figure 3. Sample Entropy values for original signals and its corresponding surrogate data sets (boxplot)

4. Results

The analysis of surrogate data showed the nonlinear behavior of the signals. Figures 1, 2 and 3 show the statistical values for the original and surrogate data. Regarding the regularity estimation, all the non-linear metrics studied show that the signals in the groups S and T, which correspond to terminating AF events, are more regular than those in the group N. However, none of the studied parameters provided statistically significant differences between the groups S and T. Tables 1 and 2 show the classification accuracy values of the metrics applied to the AA and MAW respectively. On the other hand eave-one-out cross-validation results are shown in Table 3.

Table 1. Results of classification accuracy for the metrics computed on the atrial activity

	N vs. T	N vs. S	T vs. S	N vs. (T & S)
DAF	88.3%	84.7%	66.7%	86.3%
fWP	73.4%	71.8%	63.0%	76.3%
SampEn	76.7%	74.0%	64.8%	75.0%
FuzzEn	71.7%	69.6%	66.7%	71.3%
SEn	86.7%	89.1%	63.0%	72.2%
LZC (2 sym)	73.3%	76.1%	63.0%	72.2%
LZC (3 sym)	70.0%	74.0%	66.7%	72.5%
HE	75.0%	67.4%	66.7%	75.0%
GHE (q=1)	71.7%	69.6%	64.8%	77.5%
GHE (q=2)	71.7%	71.7%	63.0%	77.5%

Table 2. Results of classification accuracy for the metrics computed on the main atrial wave

	N vs. T	N vs. S	T vs. S	N vs. (T & S)
MAWP	78.3%	76.1%	63.0%	78.8%
SampEn	88.3%	87.0%	65.0%	91.3%
FuzzEn	90.0%	91.3%	66.7%	92.5%
LZC (2 sym)	93.3%	91.3%	70.4%	92.5%
LZC (3 sym)	90.0%	91.3%	66.7%	90.0%
HE	83.3%	87.0%	63.0%	87.5%
GHE (q=1)	95.0%	93.5%	66.7%	93.8%
GHE (q=2)	93.3%	93.5%	66.7%	93.8%

Table 3. Leave-one-out cross-validation results

	N vs. T	N vs. S	T vs. S	N vs. (T & S)
DAF	85.0%	72.3%	57.4%	82.5%
fWP	63.3%	69.6%	53.7%	61.3%
SampEn	73.3%	65.2%	57.4%	73.8%
SEn	76.7%	87.0%	48.2%	75.0%
MAWP	65.0%	69.6%	53.7%	62.5%
MAW-SampEn	85.0%	89.1%	54.0%	86.3%
MAW-FuzzEn	88.3%	87.0%	53.7%	90.0%
MAW-LZC (2s)	91.7%	91.3%	55.6%	90.0%
MAW-LZC (3s)	90.0%	84.8%	59.3%	88.8%
MAW-HE	81.7%	87.0%	57.4%	83.8%
MAW-GHE(1)	90.0%	87.0%	58.0%	91.3%
MAW-GHE(2)	90.0%	89.1%	57.4%	91.3%

Statistical significance in the discrimination between terminating and non-terminating AF episodes was found for DAF, fWP, SEn and SampEn computed on the AA and for all the MAW parameters. Moreover, the metrics FuzzEn, LZC and GHE achieved better classification results than DAF and SampEn. Furthermore, the MAW regularity evaluation improved the classification results for all the metrics computed both on the AA and the MAW.

5. Discussion

The study of MAW regularity improved results obtained from direct AA regularity evaluation for all the analyzed non-linear metrics, which is consistent with previous works reported in [14, 16].

In this work, the use of FuzzEn applied to the MAW improved slightly classification results obtained with SampEn, however its direct application to the AA did not improve the results attained with SampEn. Therefore, further research is needed to assess the usefulness of FuzzEn in the AF organization estimation.

The classification results obtained in the present work suggest that LZC could be useful in the evaluation of AF organization. Regarding the number of symbols used in the signal conversion, the best accuracy values were obtained with the binary conversion, but the results obtained with the three-symbol conversion were similar; however, further studies are needed to decide the optimum symbol set in the context of AF organization estimation.

The Generalized Hurst Exponent GHE($q = 2$) is related to the autocorrelation function, and like LZC is related to the PSD of the signal; the results presented in this work suggest that both parameters are significantly correlated. The signals corresponding to terminating AF episodes show higher values of GHE than those in the non-terminating group. Further research should be performed to establish the optimal use of GHE in AA signals.

6. Conclusions

In the present study it has been demonstrated that several non-linear metrics allow to classify events in AF from the surface ECG. Moreover, the MAW regularity study improved results obtained from direct AA regularity estimation in all the analyzed nonlinear metrics. Furthermore, the analyzed metrics FuzzEn, LZC and GHE applied to the MAW improved the classification results obtained with DAF and SampEn in the present study; therefore, they could be considered as promising tools in the quantification of AF events from the surface ECG. Nevertheless, the robustness of these results should be validated in further studies with wider databases.

Acknowledgments

Work funded by TEC2010–20633 from the Spanish Ministry of Science and Innovation and PPII11–0194–8121 from Junta de Comunidades de Castilla-La Mancha.

References

[1] Pincus SM. Assessing serial irregularity and its implications for health. Ann N Y Acad Sci 2001;954:245–267.

[2] Richmann JS, Moorman JR. Physiological time-series analysis using approximate entropy and sample entropy. Am J Physiol Heart Circ Physiol 2000;278 (69:H2039–H2049.

[3] Chen W, Zhuang J, Yu W, Wang Z. Measuring complexity using FuzzyEn, ApEn, and SampEn. Medical Engineering Physics 2009;31:61–68.

[4] Lempel A, Ziv J. On the complexity of finite sequences. IEEE Transactions on Information Theory 1976;Vol. 22, NO 1:75–81.

[5] Mandelbrot BB, Ness JWV. Fractional brownian motions, fractional noises and applications. SIAM Review 1968; 10:422–437.

[6] Barabási AL, Vicsek T. Multifractality of self-affine fractals. Physical Review A 1991;44(4):2730–2733.

[7] Rezek IA, Roberts SJ. Stochastic complexity measures for physiological signal analysis. IEEE Transactions on Biomedical Engineering 1998;45 (9):1186–1191.

[8] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. Circulation 2000;101(23):E215–E220.

[9] Moody GB. Spontaneous termination of atrial fibrillation: A challenge from physionet and computers in cardiology 2004. Computers in Cardiology 2004;31:101–104.

[10] Sörnmo L, Laguna P. Biomedical Signal Processing in Cardiac and Neurological Applications. Elsevier Academic Press, 2005.

[11] Alcaraz R, Rieta JJ. Adaptive singular value cancelation of ventricular activity in single-lead atrial fibrillation electrocardiograms. Physiological Measurement 2008;29:1351–1369.

[12] Holm M, Pehrson S, Ingemansson M, Sörnmo L, Johansson R, Sandhall L, Sunemark M, Smideberg B, Olsson C, Olsson SB. Non-invasive assessment of the atrial cycle length during atrial fibrillation in man: introducing, validating and illustrating a new ECG method. Cardiovasc Res 1998;38:69–81.

[13] Alcaraz R, Rieta JJ. Time and frequency recurrence analysis of persistent atrial fibrillation after electrical cardioversion. Physiological Measurement 2009;30:479–489.

[14] Alcaraz R, Rieta JJ. A review on sample entropy applications for the non-invasive analysis of atrial fibrillation electrocardiograms. Biomedical Signal Processing and Control 2010;5 (1):1–14.

[15] Alcaraz R, Abásolo D, Hornero R, Rieta JJ. Optimal parameters study for sample entropy-based atrial fibrillation organization analysis. Computer Methods and Programs in Biomedicine 2010;99:124–132.

[16] Alcaraz R, Rieta JJ. Sample entropy of the main atrial wave predicts spontaneous termination of paroxysmal atrial fibrillation. Medical Engineering Physics 2009;31:917–922.

[17] Aboy M, Hornero R, Abásolo D, Álvarez D. Interpretation of the lempel-ziv complexity measure in the context of biomedical signal analysis. IEEE Transactions on Biomedical Engineering 2006;Vol. 53, No. 11:2282–2288.

[18] Mainardi LT, Sassi R. Analysis of scaling behaviour of ECG signals during atrial fibrillation. Computers in Cardiology 2005;32:627–630.

[19] Schreiber T, Schmitz A. Surrogate time series. Physica D 2000;142:346–382.

Address for correspondence:

Matilde Julián Seguí
Biomedical Synergy. Electronic Eng. Dept. 7F
Universidad Politécnica de Valencia
Camino de Vera s/n. 46022 Valencia (Spain)
e–mail: majuse@upv.es