LEMPEL-ZIV COMPLEXITY ANALYSIS FOR THE EVALUATION OF ATRIAL FIBRILLATION ORGANIZATION

Daniel Abásolo1, Raúl Alcaraz2, José Joaquín Rieta3, Roberto Hornero4
1Centre for Biomedical Engineering, Division of Mechanical, Medical and Aerospace Engineering, Faculty of Engineering and Physical Sciences, University of Surrey, United Kingdom
d.abasolo@surrey.ac.uk
2Innovation in Bioengineering Research Group, University of Castilla-La Mancha, Spain
3Biomedical Synergy, Electronic Engineering Department, Universidad Politécnica de Valencia, Spain
4Biomedical Engineering Group, University of Valladolid, Spain

ABSTRACT
The Lempel-Ziv (LZ) complexity is a non-linear time series analysis metric that reflects the arising rate of new patterns along with the sequence. Thus, it captures its temporal sequence and, quite conveniently, it can be computed with short data segments. In the present work, a detailed analysis on LZ complexity is presented within the context of atrial fibrillation (AF) organization estimation. As the analysed time series depend on the original sampling rate (fs), we evaluated the relationship between LZ complexity and fs. Furthermore, different implementations of LZ complexity were tested. Our results show the usefulness of LZ complexity to estimate AF organization and suggest that the signals from a terminating paroxysmal AF group are more organized (i.e. less complex) than those from the non-terminating paroxysmal AF group. However, the diagnostic accuracy was not as high as that obtained with sample entropy (SampEn), another non-linear metric, with the same database in a previous study (92% vs. 96%). Nevertheless, the LZ complexity analysis of AF organization with sampling frequencies higher than 2048 Hz, or even its combination with SampEn or other non-linear metrics, might improve the prediction of spontaneous AF termination.

KEY WORDS
Biomedical signal processing, Electrocardiogram, Atrial fibrillation, Lempel-Ziv complexity.

1. Introduction
Non-linear analysis of physiological signals has proven to be a valuable tool to obtain “hidden information” related to underlying mechanisms [1], [2]. A non-linear analysis technique that has been used with that aim is Lempel-Ziv (LZ) complexity [3]. LZ complexity measure reflects the arising rate of new patterns in a time series [4]. Hence, it could provide a way of estimating a time series structure or organization.

The organization estimation of atrial fibrillation (AF) is a relevant aspect in order to better understand its mechanisms. Although AF affects up to 1% of the general population [5], [6], its mechanisms are still unexplained [7]. Recently, the application of other non-linear metrics, like sample entropy (SampEn), has revealed information that could be used for a better treatment of the arrhythmia [8–10].

In this pilot study we analysed AF organization with LZ complexity. As the time series depends on the original electrocardiogram (ECG) sampling rate (fs), we evaluated the changes in LZ complexity with different fs values. Furthermore, given that one needs to transform the original time series into a finite symbol sequence before LZ complexity calculation, the impact that different transformations have on the results was also studied. Given the fact that no method to generate AF signals with a priori controlled organization exists, the use of simulated signals is not possible. As a consequence, real AF signals with organization-dependent events were selected.

2. Material and Methods
2.1 Paroxysmal AF database
The analysed database consisted of fifty Holter recordings of 30 s from Physionet [11]. Two leads (II and V1) were available. The database included non-terminating AF episodes, which were observed to continue in AF for, at least, one hour following the end of the excerpt. We will refer to them as N group. Moreover, it also included AF episodes terminating immediately after the end of the extracted segment. These signals will be denoted as T group. In order to digitise the signals, a sampling rate of 128 Hz and 16-bit resolution were used. Nevertheless, the signals were also up-sampled to 1024 Hz to achieve a better alignment for QRST complex subtraction, as previously suggested [12]. This procedure is necessary if one wants to extract the atrial activity (AA) from surface ECGs. A cubic splines interpolation method was used due to its good performance [13]. After the AA extraction, the residual signal was down-sampled back to 128 Hz.
V1 was chosen as it has been shown that AA is prevalent in this lead [14]. Furthermore, signals were pre-processed to improve their analysis. Baseline wander was removed with bidirectional high pass filtering (0.5 Hz cut-off frequency) [15]. Then, high frequency noise was reduced with an eighth order bidirectional IIR Chebyshev low pass filtering (cut-off frequency of 70 Hz) [16]. Finally, mains interference was removed with a notch filter, which preserves the ECG spectral information [17].

2.2 Lempel-Ziv complexity

Lempel and Ziv introduced the so called Lempel-Ziv (LZ) complexity measure for sequences of finite length in 1976 [3]. This metric provides a complexity measure in a one-dimensional signal that can be calculated even for short data segments [18]. In addition, LZ complexity reflects the number of distinct substrings and the rate of their recurrence along the given sequence [4]. Larger LZ complexity values correspond to more complex time series.

In the biomedical signal processing field, this complexity metric has been used in the detection of ventricular tachycardia and fibrillation [18], in the study of the brain function [19], to characterize brain information transmission [20] or to quantify the depth of anaesthesia [21], [22], to mention just a few examples of its widespread use.

Before calculating LZ complexity, the signal must be transformed into a finite symbol sequence \( P \). In this pilot study we have considered two different sequence conversion methods:

a) 0-1 sequence conversion. The signal data were converted into a binary sequence (i.e. a 0-1 sequence) by comparison with a threshold. We called that binary sequence \( P = s(1), s(2), \ldots, s(n) \), with \( s(i) \) defined by [21]:

\[
s(i) = \begin{cases} 
0 & \text{if } x(i) < T_d \\
1 & \text{if } x(i) \geq T_d
\end{cases}
\]

(1)

We considered two different thresholds, the median value and the mean. Although partitioning about the median is robust to outliers [23], for the sake of comparison we have also considered the mean.

b) 0-1-2 sequence conversion. In this conversion method we need two different thresholds to transform the original signal into a 0-1-2 sequence. In order to estimate them, we calculated the mean \( x_{\text{mean}} \), the median \( x_{\text{median}} \), maximum \( x_{\text{max}} \) and minimum \( x_{\text{min}} \) for each time series. The thresholds were defined as \( T_{d1} = x_{\text{median}} - \frac{|x_{\text{max}}|}{16}, \) \( T_{d2} = x_{\text{mean}} + \frac{|x_{\text{max}}|}{16} \) when using the mean [21], and as \( T_{d1} = x_{\text{median}} - \frac{|x_{\text{mid}}|}{16} \) and \( T_{d2} = x_{\text{median}} + \frac{|x_{\text{max}}|}{16} \) when using the median. The data were then converted into a 3 symbol sequence \( P = s(1), s(2), \ldots, s(n) \), with \( s(i) \) defined by [21]:

\[
s(i) = \begin{cases} 
0 & \text{if } x(i) \leq T_{d1} \\
1 & \text{if } T_{d1} < x(i) < T_{d2} \\
2 & \text{if } x(i) \geq T_{d2}
\end{cases}
\]

(2)

Once the original signal has been transformed into a finite symbol sequence \( P \), this is scanned from left to right for both conversion methods and a complexity counter \( c(n) \) is increased by one unit every time a new subsequence of consecutive characters is encountered. A detailed algorithm explaining how is \( c(n) \) computed can be found in [18], [21], [22].

However, the \( c(n) \) value might vary with the sequence length. Thus, in order to obtain a complexity measure independent of the sequence length, \( c(n) \) should be normalized. It has been proven [3] that the upper bound of \( c(n) \) is given by:

\[
c(n) < \frac{n}{(1 - e_n) \log_{\alpha} (n)}
\]

(3)

where \( n \) is the length of the sequence, \( \alpha \) is the number of different symbols in the symbol set and \( e_n \) is a small quantity and \( e_n \to 0 \) \((n \to \infty)\). In general,

\[
\lim_{n \to \infty} c(n) = b(n) = \frac{n}{\log_{\alpha} (n)}
\]

(4)

and \( c(n) \) can be then normalized via \( b(n) \):

\[
C(n) = \frac{c(n)}{b(n)}
\]

(5)

\( C(n) \) is the normalized LZ complexity. As this metric reflects the arising rate of new patterns within the sequence, it captures its temporal structure.

2.3 Strategy to estimate AF organization

In order to adequately estimate LZ complexity in surface ECG in AF, firstly the ventricular activity – which is of much greater amplitude than AA – has to be removed with a cancellation technique. With that aim, a QRST cancellation method was used, since only one lead was considered. The highest variance eigenvector of all the ECG beats was considered as the ventricular template for the cancellation. This QRST template provided a higher quality AA extraction in short AF recordings than the one obtained by averaging all the beats [24].

Since QRST morphology is affected by respiration, patient movement, etc., QRST residua and noise are often present in the extracted AA signal [14]. Thus, the main atrial wave (MAW) has to be extracted from the AA [25].
It can be considered as the fundamental waveform associated to the AA. Its wavelength is the inverse of the AA dominant frequency (Figure 1) [26].

To extract the MAW, one has to first compute the AA power spectral density. To that aim, Welch’s periodogram (with a Hamming window of 4096 points in length, a 50% overlapping between adjacent windowed sections and a 8192-points Fast Fourier Transform) was used [27]. The dominant atrial frequency (DAF) was chosen as the highest amplitude frequency within the 3–9 Hz range [26] and the MAW was obtained by selective filtering of the AA centred on the DAF. A linear phase FIR filter was used to prevent distortion [28]. It was designed with the Chebyshev approximation, 3 Hz bandwidth and 768 coefficients [8], [29].

AF organization was estimated via the computation of the MAW $LZ$ complexity. However, as the MAW length depends on the ECG sampling rate, for high $fs$, a low MAW time interval would be analysed, while the opposite also holds. Thus, the selection of a sampling rate as the one defined by the Nyquist criterion does not guarantee accurate organization estimation in AF. Hence, the MAW was down-sampled to 64 Hz or up-sampled to 256, 512, 1024 and 2048 Hz and $LZ$ complexity of the resulting time series was computed. Since the original sampling frequency was 128 Hz, a cubic splines method was used when interpolation was needed.

Differences between the $LZ$ complexity values of the MAW from terminating and non-terminating paroxysmal AF episodes were evaluated with Kruskal-Wallis test for all $fs$ values. Differences were considered statistically significant when $p$-value < 0.01. Additionally, the discriminative abilities of paroxysmal AF termination for the different $LZ$ complexity implementations were calculated with receiver operating characteristic (ROC) curves. A ROC curve is a graphical representation of the trade-offs between sensitivity (here the proportion of non-terminating paroxysmal AF episodes correctly detected) and specificity (percentage of terminating episodes properly identified for the spontaneous AF termination prediction). Accuracy was the number of terminating and non-terminating AF episodes correctly classified.

3. Results

Tables 1, 2, 3 and 4 show the $LZ$ complexity results with different implementations for non-terminating and terminating paroxysmal AF and the different values of $fs$. It can be noticed that the non-terminating paroxysmal AF group is characterized by more complex signals than the terminating one, i.e. the $LZ$ complexity values are higher. Statistically significant differences between the $LZ$ complexity values of both groups for sampling frequencies over 256 Hz were found in all cases but the two symbol conversion with the median as threshold. Furthermore, for larger sampling frequencies, $LZ$ complexity is lower, something that suggests a more regular or organized signal.

Tables 5, 6, 7 and 8 summarise the classification results, namely specificity, sensitivity and accuracy. Moreover, the $LZ$ complexity value that maximizes accuracy is also included. The discriminative differences between subsets increase with the sampling frequency. Accuracy was over 70% in all cases where statistically significant differences were found. It is noteworthy that it reached a maximum value of 92% when the sampling frequency was 2048 Hz and $LZ$ complexity was computed with a three symbol conversion and the median as threshold. On the contrary, for $fs \leq 128$ Hz, a very limited predictive ability was obtained for the different $LZ$ complexity implementations.

4. Discussion

About 18% of paroxysmal AF degenerates into persistent AF in less than 4 years [30] and there are quite serious risks associated to the latter, such as thrombus formation predisposition within the atria, which can cause stroke or any other thromboembolic events [31]. As appropriate interventions may terminate the arrhythmia and prevent chronification of AF, early prediction of paroxysmal AF maintenance is of vital importance. On the other hand, the prediction of spontaneous AF termination could avoid unnecessary therapy, reduce the associated clinical costs, and improve the patient’s quality of life.

In a recent study the prediction of spontaneous AF termination with $SampEn$ was achieved with a diagnostic accuracy of 96% [32]. In this pilot study we wanted to evaluate if that prediction could be improved with a different non-linear analysis technique, namely $LZ$ complexity, whilst at the same time testing the different parameters that play a role in the computation of this metric.
This also happened when the same database was analysed between the terminating and non-terminating patients. Complexity does not reflect significant differences included. Significant differences (p-value < 0.01) are marked with an asterisk.

### Table 2. Mean and SD of LZ complexity with a two symbol conversion and the mean as threshold for non-terminating and terminating paroxysmal AF. The p-value is also included. Significant differences (p-value < 0.01) are marked with an asterisk.

<table>
<thead>
<tr>
<th>fs (Hz)</th>
<th>N group</th>
<th>T group</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>0.3550 ± 0.0218</td>
<td>0.3318 ± 0.0550</td>
<td>0.0658</td>
</tr>
<tr>
<td>128</td>
<td>0.2463 ± 0.0149</td>
<td>0.2333 ± 0.0332</td>
<td>0.2463</td>
</tr>
<tr>
<td>256*</td>
<td>0.1697 ± 0.0089</td>
<td>0.1559 ± 0.0214</td>
<td>0.0067</td>
</tr>
<tr>
<td>512*</td>
<td>0.1116 ± 0.0052</td>
<td>0.1002 ± 0.0110</td>
<td>1.58*10^-5</td>
</tr>
<tr>
<td>1024*</td>
<td>0.0719 ± 0.0037</td>
<td>0.0619 ± 0.0065</td>
<td>7.25*10^-7</td>
</tr>
<tr>
<td>2048*</td>
<td>0.0440 ± 0.0030</td>
<td>0.0361 ± 0.0038</td>
<td>8.70*10^-8</td>
</tr>
</tbody>
</table>

### Table 3. Mean and SD of LZ complexity with a three symbol conversion and the median as threshold for non-terminating and terminating paroxysmal AF. The p-value is also included. Significant differences (p-value < 0.01) are marked with an asterisk.

<table>
<thead>
<tr>
<th>fs (Hz)</th>
<th>N group</th>
<th>T group</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>0.3624 ± 0.0313</td>
<td>0.3345 ± 0.0602</td>
<td>0.1451</td>
</tr>
<tr>
<td>128</td>
<td>0.2689 ± 0.0249</td>
<td>0.2431 ± 0.0440</td>
<td>0.0310</td>
</tr>
<tr>
<td>256*</td>
<td>0.1952 ± 0.0172</td>
<td>0.1720 ± 0.0260</td>
<td>8.62*10^-7</td>
</tr>
<tr>
<td>512*</td>
<td>0.1363 ± 0.0101</td>
<td>0.1143 ± 0.0155</td>
<td>2.59*10^-6</td>
</tr>
<tr>
<td>1024*</td>
<td>0.0883 ± 0.0068</td>
<td>0.0714 ± 0.0090</td>
<td>1.03*10^-7</td>
</tr>
<tr>
<td>2048*</td>
<td>0.0532 ± 0.0042</td>
<td>0.0422 ± 0.0054</td>
<td>6.00*10^-8</td>
</tr>
</tbody>
</table>

### Table 4. Mean and SD of LZ complexity with a three symbol conversion and the mean as threshold for non-terminating and terminating paroxysmal AF. The p-value is also included. Significant differences (p-value < 0.01) are marked with an asterisk.

<table>
<thead>
<tr>
<th>fs (Hz)</th>
<th>N group</th>
<th>T group</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>64</td>
<td>0.3616 ± 0.0340</td>
<td>0.3335 ± 0.0604</td>
<td>0.1369</td>
</tr>
<tr>
<td>128</td>
<td>0.2692 ± 0.0265</td>
<td>0.2436 ± 0.0437</td>
<td>0.0236</td>
</tr>
<tr>
<td>256*</td>
<td>0.1956 ± 0.0167</td>
<td>0.1726 ± 0.0262</td>
<td>7.49*10^-4</td>
</tr>
<tr>
<td>512*</td>
<td>0.1363 ± 0.0101</td>
<td>0.1142 ± 0.0159</td>
<td>2.04*10^-6</td>
</tr>
<tr>
<td>1024*</td>
<td>0.0881 ± 0.0063</td>
<td>0.0711 ± 0.0090</td>
<td>8.73*10^-8</td>
</tr>
<tr>
<td>2048*</td>
<td>0.0534 ± 0.0045</td>
<td>0.0419 ± 0.0053</td>
<td>7.05*10^-8</td>
</tr>
</tbody>
</table>

For sampling frequencies lower than 128 Hz, LZ complexity does not reflect significant differences between the terminating and non-terminating patients. This also happened when the same database was analysed with SampEn [32] and suggests that the sampling frequency should be at least 256 Hz in order to accurately estimate AF organization. Nevertheless, an important difference between this study and [32] is that the diagnostic accuracy of LZ complexity increases with the sampling frequency, irrespective of the use of a 2 or 3 symbol conversion and the mean or median as threshold. The best results were obtained with 2048 Hz. This could imply that more accurate information about MAW evolution over time could be obtained with a higher sampling frequency, something that could also be related...
with the fact that LZ complexity shows that time series are more organized when $s$ is higher.

With respect to the number of symbols used in LZ complexity evaluation, the best diagnostic accuracy was obtained with 3 (92%). Previous studies claimed that a two symbol sequence conversion was enough to study the dynamic complexity of a signal [4], [19], [20]. However, our results suggest that a three symbol conversion might give more detailed insight into the differences between the terminating and non-terminating paroxysmal AF groups. A possible explanation could be that more information from the MAW is kept during the coarse-graining process when a 0-1-2 conversion is used. However, results were not conclusive, as the second best diagnostic accuracy, only slightly lower than the previously mentioned one (90%), was obtained with a 2 symbol conversion. Results were not conclusive either regarding the use of the mean or the median as threshold in the symbol conversion procedure. Although the median is robust to outliers [23], the use of the mean as threshold could be justified when they are not present. Further studies with a larger database are needed to elucidate which symbol conversion choice is optimal in order to compute LZ complexity in the context of AF organization and the prediction of spontaneous AF termination.

Some limitations of our study should be mentioned. First of all, the use of paroxysmal AF recordings sampled at a fixed rate of 128 Hz could be a problem. The up-sampling process requires interpolation and this could provoke an underestimation of the analysed time series complexity with LZ complexity. Ideally, it would be better to digitise the recordings at different sampling rates and then calculate its LZ complexity. Nonetheless, it has been shown that the cubic splines interpolation method of AA did not have a significant impact on results obtained with SampEn [32]. Secondly, the accuracy values were estimated with the whole subsets, something that could cause problems like over-fitting and bias. In the future, leave-one-out cross validation could be used to prevent these problems.

5. Conclusion

We have shown the usefulness of LZ complexity to estimate AF organization. Our results suggest that the terminating paroxysmal AF group signals are more organized (i.e. less complex) than those from the non-terminating paroxysmal AF group. These results agree with invasive studies [33], [34]. However, the diagnostic accuracy was not as high as in a previous study, where SampEn was used to analyse the same database (92% vs. 96%). Nevertheless, the LZ complexity analysis of AF organization with sampling frequencies higher than 2048 Hz, or even its combination with SampEn or other non-linear metrics, might improve the prediction of spontaneous AF termination. Further studies are needed to address these questions.

Acknowledgements

This work was supported by the projects TEC2007-64884, TEC2008–02241 and TEC2010-20633 from the Spanish Ministry of Science and Innovation and PI1209-0224-5983 and PI11C09-0036-3237 from Junta de Comunidades de Castilla La Mancha.

References

atrial fibrillation, Medical and Biological Engineering and Computing, 46(7), 2008, 625–635.


