journal homepage: www.intl.elsevierhealth.com/journals/cmpb

Automatic segmentation of long-term ECG signals corrupted with broadband noise based on sample entropy

Pau Micó[a], *Margarita Mora*[b], *David Cuesta-Frau*[a,*], *Mateo Aboy*[c]

[a] *Technological Institute of Informatics, Polytechnic University of Valencia, Alcoi Campus, Plaza Ferrándiz y Carbonell 2, 03801 Alcoi, Spain*
[b] *Department of Applied Mathematics, Polytechnic University of Valencia, Alcoi Campus, Spain*
[c] *Department of Electrical Engineering, Oregon Institute of Technology, Oregon, USA*

ARTICLE INFO

Article history:
Received 28 April 2009
Received in revised form
15 August 2009
Accepted 24 August 2009

Keywords:
Signal segmentation
Sample entropy
ECG records
Signal noise

ABSTRACT

Biomedical signals are nonstationary in nature, namely, their statistical properties are time-dependent. Such changes in the underlying statistical properties of the signal and the effects of external noise often affect the performance and applicability of automatic signal processing methods that require stationarity. A number of methods have been proposed to address the problem of finding stationary signal segments within larger nonstationary signals. In this framework, processing and analysis are applied to each resulting locally stationary segment separately.

The method proposed in this paper addresses the problem of finding locally quasi-stationary signal segments. Particularly, our proposed algorithm is designed to solve the specific problem of segmenting semiperiodic biomedical signals corrupted with broadband noise according to the various degrees of external noise power. It is based on the sample entropy and the relative sensitivity of this signal regularity metric to changes in the underlying signal properties and broadband noise levels.

The assessment of the method was carried out by means of experiments on ECG signals drawn from the MIT-BIH arrhythmia database. The results were measured in terms of false alarms based on the changepoint detection bias. In summary, the results achieved were a sensitivity of 97%, and an error of 16% for records corrupted with muscle artifacts.

1. Introduction

Biomedical signals are intrinsically nonstationary because their underlying statistical properties change with time. This source of nonstationarity is intrinsic in the sense that the origins are physiological in nature [1]. In addition to this intrinsic nonstationarity, an additional source of nonstationarity present in biomedical signals obtained in practical settings is the external noise and the corresponding changes in noise characteristics such as noise power and noise bandwidth. Such changes often decrease the performance of automatic signal processing methods, especially when the algorithm parameters are selected for the entire time series but these parameters might be locally unsuitable. In certain application areas a nonstationary time series may be considered as a concatenation of stationary segments where properties can be assumed homogeneous [2].

Decomposition of signals into stationary or quasi-stationary intervals is a well-known problem often referred to as time series segmentation [3]. The exact segmentation of a nonstationary time series is a computationally intensive problem that cannot be easily solved, especially when dealing with long signals or when the statistical properties

* *Corresponding author*. Tel.: +34 651028565.
E-mail address: dcuesta@disca.upv.es (D. Cuesta-Frau).
0169-2607/$ – see front matter © 2009 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.cmpb.2009.08.010

of the signal or noise change very slowly and continuously. The methods available typically assume a piecewise stationary signal model or noise model and detect the instants of change. Signal segmentation due to changes in noise characteristics such as noise power and bandwidth is typically an easier problem than segmentation due to changes in underlying signal characteristics. Physiologic signal parameters often change slowly, gradually, and continually which make the segmentation significantly more difficult to accomplish.

Time series segmentation is very important in many fields and applications: speech processing [4] (coding, synthesis, speaker identification, estimation of speech parameters), biomedical signal processing [5] (detection of onsets of spikes in electroencephalograms, P-waves in electrocardiograms, denoising), quality control [6], Internet traffic fluctuations [7], among many others. A number of approaches have been used to solve the time series segmentation problem:

- Dynamic programming based: This approach guarantees the global optimality of the segmentation when it can be quantified in terms of a cost function [8–11].
- Top-down methods: These methods start with an unsegmented time series and add one point at a time. At the ith step, they add a new boundary point by splitting one of the segments. This is repeated until a stopping condition is met [12,13,3].
- Bottom-up methods: In these methods each point is seen as a segment and then consecutive segments are merged till a predefined number of segments is reached or the error exceeds a threshold [14–16].
- Random: The algorithms start with an arbitrary segmentation. Boundary points are randomly selected, taken away, and then a search is performed to find the best position to put them back [17].
- Sliding window methods: These methods fix the left boundary and try to place the right boundary as far as possible. When a parameter of the current segment exceeds a threshold, fix the current segment and proceed with the next one. Repeat until the sequence ends [18,19].
- Model based. A segmentation model assumes a predefined number of segment-types or a prior distribution of the parameters. The determination of the actual number of segments is achieved by training and comparing several separate models [20–22].

Most of these methods are complex, computationally intensive, and difficult to implement [20]. Additionally, some of them require knowledge of a number of parameters in advance such as the number of segments [23], the stopping rules [18], thresholds [24], number of models or states [25], model complexity or degree [21], or need training data [26]. The complexity of the segmentation methods greatly depends on the specific segmentation problem they are aimed at solving. While complex algorithms are required to solve the general segmentation problem, simpler algorithms can be developed to address more specific segmentation problems encountered in practical applications.

In this paper we describe a new method to automatically segment long term biomedical signals that overcomes some of the problems stated above for a particular case of the segmentation problem, namely, the segmentation of signal segments corrupted by different degrees of broadband noise. The main strengths of the method proposed are:

- Simple and efficient underlying algorithm: The proposed method is based on the well known SampEn algorithm [27] which is very simple to implement, and computationally efficient. Additionally, windowed SampEn can be computed incrementally.
- Non-supervised: Changes of SampEn related to noise are of very high amplitude in comparison to SampEn changes related to signal. Thus, thresholds generalize very well and no customization is necessary.
- On-line application: The input data series can be processed sequentially, in contrast to top-down or bottom-up based methods. There is a small computation delay due to the signal window employed.
- A-Priori signal model: The method does not require and does not assume a statistical model of the signal.

Our proposed algorithm is designed to solve the specific problem of segmenting semiperiodic biomedical signals corrupted with broadband noise according to the various degrees of signal-to-noise-ratio (SNR) while not creating segments due to statistical changes in the underlying biomedical signals properties such as changes in frequency, amplitude, and other signal properties within physiological normality. Consequently, the resulting segments are still nonstationary in the sense that the underlying statistical properties of the signal change within the segment but within each of the segments the corrupting broadband noise is stationary.

The problem of signal segmentation according to the broadband noise power is significant in several biomedical applications. For instance, changes in broadband in biomedical signals is often an indicator of a change in the level of physical activity [28], and decreases the performance of other algorithms such as thresholding for wavelet-based denoising [29], fiducial points detection [30] and interval measurements. Broadband noise power changes are one of the most typical signal changes since they can be caused by patient activity, electrode-skin contact degradation, external source interference, and multiple other causes. The method proposed is based on a recently described property of sample entropy (SampEn): high sensitivity to noise changes in semiperiodic signals [31]. Fig. 1 graphically illustrates this sensitivity.

We chose to illustrate the segmentation algorithm by applying it to electrocardiogram (ECG) signals because of their widespread use, although the method can also be applied to other semiperiodic biomedical signals such as arterial blood pressure, intracranial pressure, plethysmogram, or respiration signals.

The rest of the paper is structured as follows. In Section 2.1, we describe the SampEn metric used to detect noise level changes. Section 2.2 is devoted to describe the relationship between signal noise and SampEn, and demonstrate the suitability of this metric for the purpose of the paper. The complete algorithm is introduced in Section 2.3. Next, a set of comprehensive experimental studies is covered in Section 3, including a description of the data set in Section 3.1, the experiments in Section 3.2, and the assessment parameters in Section 3.3.

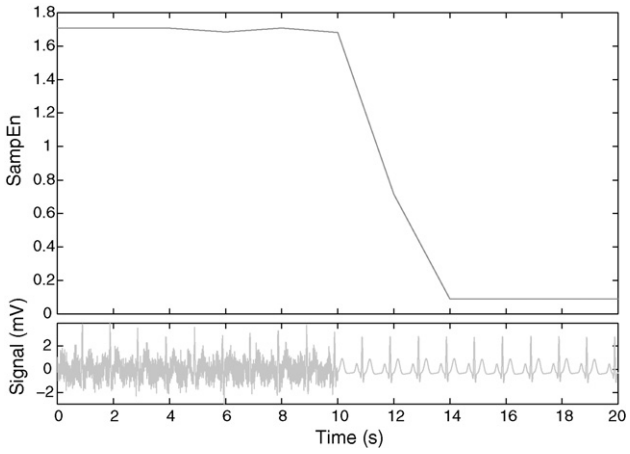

Fig. 1 – SampEn and noise power change in an ECG. The signal exhibits a 5 dB SNR from 0 to 10 s, and it is noiseless for the remaining 10 s. Such a change in the noise power level elicits a significant change in SampEn. This example shows a straightforward case for illustrative purposes.

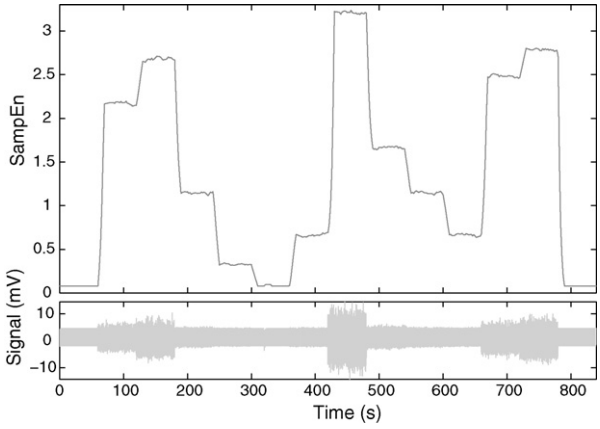

Fig. 2 – Illustrative example showing a synthetic ECG signal corrupted by noise with different SNRs (14 segments with SNRs: 100, 0, −5, 10, 20, 100, 15, −10, 5, 10, 15, −3, −6, 100 dB). Noise power changes trigger SampEn significant variations.

Results are discussed in Section 4. Finally, Section 5 includes a number of concluding remarks.

2. Algorithm description

The proposed algorithm described here is aimed at automatically segmenting an ECG into homogeneous epochs of different lengths, according to changes in broadband noise power. It is based on the relationship between SampEn and broadband noise. All the elements involved in this method are described in the following sections.

2.1. *Underlying regularity metric: sample entropy*

The underlying regularity metric used in our algorithm is sample entropy [27]. SampEn has been proposed as a regularity metric to overcome some of the limitations associated with approximate entropy (ApEn) [31]. SampEn is a nonlinear metric that estimates the regularity in time series. It takes the time series and looks for similar patterns with the same length. The more frequent and likely these patterns are, the smaller is the entropy level of the processed series. The sample entropy of a time series $\langle x(n)\rangle$ of length N, SampEn(m, r, N) is computed as follows:

1. Take m vectors $X_m(1), X_m(2), \ldots, X_m(N-m+1)$, defined as $X_m(i) = [x(i), x(i+1), \ldots, x(i+m-1)]$, for $1 \le i \le N-m+1$. These vectors are m consecutive values of x, commencing at the ith sample.
2. The distance between vectors $X_m(i)$ and $X_m(j)$, $d[X_m(i), X_m(j)]$ is defined as:

$$d[X_m(i), X_m(j)] = \max_{k=1,\ldots,m}(|x(i+k) - x(j+k)|) \quad (1)$$

For a given $X_m(i)$, count the number of $j(1 \le j \le N-m, j \ne i)$, such that $d[X_m(i), X_m(j)] \le r$. This number is denoted as $B_i(r)$. For $1 \le i \le N-m$, two new values are defined and computed, $B_i^m(r) = (1/(N-m-1))B_i(r)$ and $B^m(r) = (1/(N-m))\sum_{i=1}^{N-m} B_i^m(r)$, where $B^m(r)$ is the probability that two sequences coincide for m points, and $A^m(r)$ is the probability that coincide for $m+1$ points.
3. Length is increased to $m = m+1$, and previous steps are repeated to obtain the counterpart of B with this new value of m, $A_i^m(r) = (1/(N-m-1))A_i(r)$ and $A^m(r) = (1/(N-m))\sum_{i=1}^{N-m} A_i^m(r)$.
4. Finally, compute SampEn as SampEn$(m, r) = \lim_{N\to\infty}\{-\log[A^m(r)/B^m(r)]\}$. Since the time series length is finite, SampEn is estimated as SampEn$(m, r, N) = -\log[A^m(r)/B^m(r)]$.

2.2. *Relationship between sample entropy and broadband noise power*

The relationship between broadband noise and SampEn in biomedical signals was first reported in the SampEn characterization study described in [31]. The results indicated that, for quasi-periodic signals, the SampEn increases as the SNR decreases (defined as SNR $= 10\log(\sum_1^N x(n)^2/\sum_1^N z(n)^2)$, being $z(n)$ the noise), in other words, SampEn is positively correlated with the noise power, as shown in Figs. 1 and 2. Fig. 2 shows an illustrative example of an ECG signal corrupted by noise with different SNRs. The figure shows how noise power changes cause clear dips and spikes as a consequence of the close relationship between SampEn and broadband noise power. Additionally, SampEn is nearly independent of the underlying signal power (Fig. 3). Practically, this result indicates that when broadband noise corrupts a quasi-periodic biomedical signal with a line-spectra (e.g. arterial blood pressure, intracranial pressure, and ECG) the SampEn increases as a function of the noise power. On the other hand, in broadband signals, increasing the power of broadband noise does not result in significant increases in SampEn. Consequently, in biomedical signals that are broadband in nature and that cannot be accurately modeled as quasi-periodic signals, the SampEn is nearly independent of additive noise power, and therefore this

method cannot be applied. In these cases, segmentation is still possible using SampEn but based on the noise bandwidth instead of the noise power.

Our method takes advantage of this relationship between broadband noise and SampEn to detect noise changes between consecutive intervals. The method is simple to implement and computationally inexpensive. Furthermore, it can be applied in real time. This method is also robust since SampEn is not influenced by the underlying semiperiodic signal power. The swing of the SampEn due to noise power changes is large.

2.3. *Segmentation algorithm description*

The segmentation algorithm must split the signal into homogeneous segments in a non-supervised way. The beginning of

Table 1 – SampEn segmentation method results for records corrupted with synthetic white gaussian broadband noise. The SNRs for the noise intervals were: 9, 6, 3, 0, −3, −6, −9, −6, −3, 0, 3, 6, 9 **and 12 dB. The actual changepoints were located at 129, 258, 387, 516, 644, 773, 902, 1031, 1160, 1289, 1418, 1547 and 1675 s. Results are expressed in terms of detection time offset (in seconds). False negatives (FNs) are represented by the symbol '–'. The number of false positives (FPs) is shown on the right column.**

mitdb	129	258	387	516	644	773	902	1031	1160	1289	1418	1547	1675	FPs
SampEn segmentation method results for the mitdb records plus white gaussian broadband noise														
100	0	0	0	0	1	1	1	1	1	0	1	0	2	0
101	0	–	–	1	1	1	1	1	1	1	0	1	2	0
102	1	0	0	0	1	1	1	0	1	1	0	0	2	0
103	0	1	0	0	1	1	1	–	0	1	1	1	2	1
104	1	1	1	1	1	1	1	0	1	0	1	1	1	3
105	0	0	0	0	1	1	1	3	1	1	0	–	–	0
106	0	0	0	6	1	1	0	2	1	0	1	0	2	2
107	1	0	1	0	2	2	1	2	2	0	1	0	2	0
108	1	0	1	0	1	4	5	0	1	0	1	1	–	1
109	0	0	0	0	1	1	1	1	1	1	1	1	3	0
111	2	0	0	0	2	1	–	0	–	0	1	0	2	0
112	0	0	–	0	1	–	1	0	0	1	1	0	2	0
113	1	2	2	1	1	1	1	1	1	1	1	0	2	0
114	0	0	1	1	2	1	1	0	4	1	1	2	1	2
115	1	0	1	0	2	1	1	1	1	–	0	0	2	1
116	0	1	0	0	1	1	–	0	1	0	0	1	2	1
117	1	0	1	0	0	3	1	1	1	0	1	0	1	0
118	1	0	1	–	2	1	1	1	1	1	1	1	5	2
119	1	1	–	–	1	5	1	1	1	1	1	1	1	0
121	0	0	5	–	2	1	1	0	1	1	1	0	1	0
122	1	0	1	0	2	1	1	1	1	1	0	1	1	1
123	9	2	0	0	1	1	1	1	0	0	1	1	1	1
124	1	0	2	–	–	2	1	1	0	1	0	1	2	0
200	0	9	1	2	1	3	–	1	0	1	0	1	2	2
201	1	1	1	0	1	1	1	1	0	3	0	0	1	1
202	1	0	0	0	1	–	3	2	2	0	1	1	2	0
203	–	0	0	0	–	–	–	4	1	1	1	0	2	1
205	0	0	–	0	3	2	1	0	1	1	–	0	1	1
207	–	0	–	–	–	4	0	0	0	1	0	1	2	2
208	4	0	1	0	1	1	1	1	9	1	1	1	4	1
209	0	0	–	0	2	1	1	1	1	1	1	4	2	0
210	0	2	1	0	0	1	–	0	–	1	0	0	2	3
212	5	0	7	1	1	1	1	2	1	1	1	1	3	1
213	0	0	0	0	2	0	1	9	0	6	0	1	1	1
214	1	1	1	0	1	1	1	1	1	1	1	1	2	0
215	0	1	0	1	1	1	1	0	1	1	1	1	2	0
217	1	1	0	1	1	1	0	7	0	1	3	0	2	1
219	0	0	0	0	1	1	1	0	1	1	0	1	2	0
220	2	0	0	2	1	1	1	1	1	1	1	1	2	0
221	0	0	0	0	1	1	1	0	0	1	1	0	1	0
222	2	0	1	0	–	2	–	1	1	–	1	1	–	0
223	0	0	0	0	1	2	1	1	1	0	–	1	2	0
228	0	1	–	–	1	2	1	2	0	2	1	1	1	0
230	1	1	–	1	1	1	1	1	1	0	1	0	2	1
231	1	1	1	0	1	1	0	1	1	0	–	1	1	0
232	1	1	0	–	1	1	1	3	–	1	1	1	2	0
233	1	1	0	1	1	0	1	0	1	1	1	1	1	0
234	1	1	0	–	2	1	1	1	1	–	0	1	3	2

each stationary interval will be marked by an index q_j. The lengths of the segments will be adapted to the local properties of the analyzed signal. In this case, the homogeneity of a segment is defined in terms of the homogeneity of the SNR.

The algorithm proceeds as follows. Given a signal of length L and a window length N, the SampEn of all the windows centered at every sample $x(i)$ is computed sequentially ($(N/2) < i < L - (N/2)$ to avoid border effects). Each computed value of SampEn, s_i is compared with two thresholds, termed upper threshold t_h and lower threshold t_l. These thresholds are calculated from two parameters of a window of previous SampEn measurements (since last changepoint), their mean s_{mean} and their standard deviation s_{std}. If s_i falls out of the interval defined by the thresholds, a changepoint q_j is set at i.

Table 2 – SampEn segmentation method results for records corrupted with synthetic pink colored noise ($1/f$). The SNRs for the noise intervals were: 9, 6, 3, 0, −3, −6, −9, −6, −3, 0, 3, 6, 9 and 12 dB. The actual changepoints were located at 129, 258, 387, 516, 644, 773, 902, 1031, 1160, 1289, 1418, 1547 and 1675 s. Results are expressed in terms of detection time offset (in seconds). False negatives (FNs) are represented by the symbol '–'. The number of false positives (FPs) is shown on the right column.

mitdb	129	258	387	516	644	773	902	1031	1160	1289	1418	1547	1675	FPs
SampEn segmentation method results for the mitdb records plus pink colored broadband noise														
100	0	0	0	0	1	1	1	0	0	0	1	0	1	1
101	0	–	1	0	1	1	0	1	1	1	1	1	2	2
102	1	0	0	0	1	1	1	1	1	0	0	3	2	1
103	0	1	0	0	1	1	1	–	0	1	0	1	3	1
104	1	1	–	1	1	1	1	0	1	0	1	1	1	5
105	0	0	0	0	1	0	0	1	0	0	0	–	–	1
106	2	0	0	2	1	1	1	1	0	0	2	0	2	2
107	1	0	0	0	1	1	1	1	1	1	0	0	2	0
108	0	0	0	0	1	1	1	0	1	0	1	1	–	3
109	1	0	0	0	1	1	1	1	0	1	0	2	2	3
111	1	0	0	0	–	1	–	0	–	2	0	2	2	1
112	0	0	–	0	0	–	2	1	0	1	0	1	2	0
113	0	0	0	0	1	1	1	0	1	1	0	1	1	0
114	0	0	0	0	1	0	0	1	1	0	0	1	4	3
115	0	1	0	0	1	0	1	0	1	–	0	1	2	2
116	0	0	0	0	1	1	–	0	1	0	0	2	4	1
117	0	0	1	0	1	–	1	1	1	0	3	0	–	1
118	1	0	0	0	1	1	1	1	0	1	1	0	5	1
119	0	0	–	0	1	1	1	1	1	1	1	1	1	1
121	0	0	–	1	0	0	1	1	0	1	1	0	1	1
122	1	0	0	0	1	1	1	0	0	0	0	1	2	2
123	9	0	0	0	1	1	1	1	1	0	1	0	2	1
124	0	0	0	0	–	0	1	1	0	0	1	0	1	0
200	0	–	0	0	1	1	4	0	0	1	0	0	2	3
201	0	0	0	1	1	1	1	6	0	1	1	5	2	1
202	0	0	0	1	1	2	0	1	9	0	1	1	1	1
203	–	0	0	0	–	6	3	0	0	1	3	1	2	2
205	0	0	–	0	1	0	1	0	1	1	1	0	–	1
207	–	0	–	–	–	5	0	1	0	0	0	1	–	2
208	0	0	0	0	1	1	1	1	9	0	0	2	2	2
209	0	0	–	0	1	1	1	1	1	0	0	1	2	1
210	0	–	0	1	1	1	–	0	–	1	0	0	2	1
212	5	0	0	0	1	1	1	1	0	1	0	1	3	4
213	0	0	0	0	1	0	1	3	0	1	0	1	2	1
214	1	0	0	0	1	1	1	0	0	0	1	0	2	1
215	0	–	1	1	1	1	1	0	1	0	1	0	2	2
217	1	1	0	1	1	1	0	–	0	1	3	0	2	1
219	1	0	0	0	1	1	0	1	0	1	0	1	2	1
220	1	0	1	0	1	1	1	0	1	0	0	0	2	0
221	0	0	0	0	1	1	1	0	0	0	1	0	1	0
222	–	0	0	0	–	1	–	1	0	1	0	1	4	2
223	0	0	0	0	1	1	1	0	0	1	6	0	2	1
228	0	0	1	1	1	2	0	1	0	1	0	1	1	3
230	0	0	0	0	1	1	0	1	0	0	1	1	1	2
231	0	0	0	0	1	1	0	–	1	0	–	1	2	0
232	0	0	0	–	1	1	1	1	–	0	1	1	2	2
233	0	1	1	0	1	1	0	0	0	1	0	0	2	2
234	0	0	0	–	2	1	1	1	0	–	0	1	2	2

The detailed steps of the algorithm are:

1. **Input:** SampEn s_i calculation window length N.
2. **Input:** Parameter α, used to compute $t_h = s_{mean} + \alpha s_{std}$, and $t_l = s_{mean} - \alpha s_{std}$.
3. **Input:** Signal $\langle x(n) \rangle$ of length $L \gg N$.
4. **Initialization:** Compute s_1 as the SampEn at $x(N/2)$ for the first signal window $\{x(1), x(2), \ldots, x(N)\}$.
5. **Initialization:** Initialize parameters s_{mean}, s_{std}, t_h and t_l.
6. **Initialization:** $j = 1$, $q = \langle \varnothing \rangle$.
7. **FOR** $i = (N/2) + 1$ **TO** $i = L - (N/2)$ **DO**
8. **BEGIN**
9. $s_i = \text{SampEn}(m, r, N)$ for $\{x(i), x(i+1), \ldots, x(i+N)\}$
10. **IF** $s_i < t_l$ **OR** $s_i > t_h$ **THEN** {Set changepoint}
11. **BEGIN**
12. $q_j = i$

Table 3 – Summary of SampEn segmentation results for records corrupted with synthetic broadband noise (white gaussian and pink colored) in terms of quantitative measures.

mitdb	Gaussian white noise					Pink colored noise				
	TP	FP	FN	S (%)	Err (%)	TP	FP	FN	S (%)	Err (%)
Quality estimators for the mitdb records plus broadband noise										
100	13	0	0	100	0	13	1	0	100	8
101	11	0	2	85	15	12	2	1	92	23
102	13	0	0	100	0	13	1	0	100	8
103	12	1	1	92	15	12	1	1	92	15
104	13	3	0	100	23	12	5	1	92	46
105	11	0	2	85	15	11	1	2	85	23
106	13	2	0	100	15	13	2	0	100	15
107	13	0	0	100	0	13	0	0	100	0
108	12	1	1	92	15	12	3	1	92	31
109	13	0	0	100	0	13	3	0	100	23
111	11	0	2	85	15	10	1	3	77	31
112	11	0	2	85	15	11	0	2	85	15
113	13	0	0	100	0	13	0	0	100	0
114	13	2	0	100	15	13	3	0	100	23
115	12	1	1	92	15	12	2	1	92	23
116	12	1	1	92	15	12	1	1	92	15
117	13	0	0	100	0	11	1	2	85	23
118	12	2	1	92	23	13	1	0	100	8
119	11	0	2	85	15	12	1	1	92	15
121	12	0	1	92	8	12	1	1	92	15
122	13	1	0	100	8	13	2	0	100	15
123	13	1	0	100	8	13	1	0	100	8
124	11	0	2	85	15	12	0	1	92	8
200	12	2	1	92	23	12	3	1	92	31
201	13	1	0	100	8	13	1	0	100	8
202	12	0	1	92	8	13	1	0	100	8
203	9	1	4	69	38	11	2	2	85	31
205	11	1	2	85	23	11	1	2	85	23
207	9	2	4	69	46	8	2	5	62	54
208	13	1	0	100	8	13	2	0	100	15
209	12	0	1	92	8	12	1	1	92	15
210	11	3	2	85	38	10	1	3	77	31
212	13	1	0	100	8	13	4	0	100	31
213	13	1	0	100	8	13	1	0	100	8
214	13	0	0	100	0	13	1	0	100	8
215	13	0	0	100	0	12	2	1	92	23
217	13	1	0	100	8	12	1	1	92	15
219	13	0	0	100	0	13	1	0	100	8
220	13	0	0	100	0	13	0	0	100	0
221	13	0	0	100	0	13	0	0	100	0
222	9	0	4	69	31	11	2	2	85	31
223	12	0	1	92	8	13	1	0	100	8
228	11	0	2	85	15	13	3	0	100	23
230	12	1	1	92	15	13	2	0	100	15
231	12	0	1	92	8	11	0	2	85	15
232	11	0	2	85	15	12	2	1	92	23
233	13	0	0	100	0	13	2	0	100	15
234	11	2	2	85	31	11	2	2	85	31
Globals	578	32	46	93	13	583	71	41	93	18

13. $j = j + 1$
14. Reset s_{mean} and s_{std}
15. **ENDIF**
16. Update parameters s_{mean}, s_{std}, t_h and t_l.
17. **ENDFOR**
18. **Output:** Changepoints $(q_1, q_2, \ldots, q_m)$ at which the SampEn variation thresholds between two consecutive windows are exceeded.

The selection of the parameter N depends on the abruptness of the change to be detected. When small windows are used to calculate SampEn, the measures have a large variance. Alternatively, large windows may have a large bias [31]. The exact values of these parameters are not crucial, values of a few seconds generalize very well. Regarding parameter α, as the SampEn variation is great even for relative small changes in noise level (Fig. 2),

Table 4 – SampEn segmentation method results for records corrupted with real muscle artifacts. The SNRs for the noise intervals were: 6, 100, 6, 100, 6, 100, 6, 100, 6, 100, 6, 100, 6 and 100 dB. The actual changepoints were located at 129, 258, 387, 516, 644, 773, 902, 1031, 1160, 1289, 1418, 1547 and 1675 s. Results are expressed in terms of detection time offset (in seconds). False negatives (FNs) are represented by the symbol '–'. The number of false positives (FPs) is shown on the right column.

mitdb	129	258	387	516	644	773	902	1031	1160	1289	1418	1547	1675	FPs
SampEn segmentation method results for the mitdb records corrupted with real muscle artifacts														
100	12	0	12	4	12	2	12	2	10	2	10	1	0	0
101	15	4	2	3	5	8	2	9	2	7	3	7	7	3
102	2	5	4	5	5	8	5	6	3	8	5	6	6	0
103	6	7	4	7	7	6	5	6	1	10	3	6	6	0
104	–	–	6	1	7	6	5	6	3	7	3	6	6	2
105	4	3	4	7	5	4	3	–	–	15	1	8	12	2
106	8	5	4	9	5	14	–	6	3	6	3	10	6	0
107	4	7	4	7	3	12	5	8	1	8	5	8	2	1
108	11	7	2	9	5	6	5	4	10	4	1	8	10	1
109	2	7	10	7	3	12	5	6	3	12	5	8	6	1
111	4	7	4	3	5	7	6	8	5	6	1	6	4	2
112	4	5	6	7	7	0	3	10	5	10	7	6	8	0
113	2	5	2	7	5	12	7	8	3	8	3	6	4	1
114	2	7	4	5	3	7	3	6	3	6	3	6	4	1
115	4	7	8	1	8	4	3	10	9	7	3	4	8	2
116	6	5	6	5	3	8	1	5	5	3	1	2	5	1
117	2	7	0	9	3	9	3	6	1	6	3	6	4	1
118	2	9	–	–	5	0	5	8	1	12	3	6	–	4
119	2	7	2	7	3	9	3	8	1	11	3	0	2	1
121	4	5	6	9	3	8	5	4	3	8	3	6	2	0
122	2	9	4	3	1	7	3	8	1	8	1	8	6	5
123	4	5	4	7	3	6	3	8	1	8	3	6	4	0
124	4	7	2	9	3	12	5	8	3	10	5	10	10	1
200	2	9	8	5	3	12	5	10	5	12	5	6	10	4
201	2	7	2	5	3	8	3	8	3	4	5	6	6	1
202	4	7	4	7	5	6	7	12	5	6	3	6	6	0
203	2	7	8	11	9	–	12	6	9	10	15	10	6	1
205	2	7	4	5	3	8	3	8	1	8	3	6	4	0
207	2	3	6	9	5	8	3	8	3	6	5	4	10	3
208	0	9	6	5	3	9	5	6	1	2	3	0	4	2
209	2	1	6	9	5	2	5	6	1	10	3	4	4	1
210	4	–	10	3	3	8	5	9	1	8	3	6	4	2
212	2	5	–	–	–	2	7	10	1	8	3	8	6	4
213	11	1	6	5	3	6	3	10	3	10	5	6	6	1
214	6	5	14	5	5	8	3	6	5	15	9	10	4	3
215	–	11	12	9	5	10	3	10	3	12	3	8	6	0
217	4	7	6	5	3	8	1	8	1	11	3	6	2	2
219	0	9	2	5	3	8	1	8	1	12	11	6	4	4
220	4	7	0	7	7	6	5	10	3	8	5	6	4	0
221	6	3	2	9	3	6	1	–	1	6	5	6	6	3
222	2	7	4	5	5	8	5	8	3	4	5	6	6	0
223	2	9	12	5	7	12	–	9	1	6	5	6	6	2
228	4	7	4	5	3	10	3	8	1	10	3	12	4	3
230	4	7	12	1	5	8	9	2	1	12	1	8	8	3
231	12	3	4	5	3	8	5	6	7	6	9	2	6	0
232	2	7	8	7	3	12	3	8	1	10	1	8	4	1
233	2	7	6	1	1	10	1	12	1	–	1	6	4	10
234	4	7	4	7	3	8	3	8	5	6	1	10	4	2

Table 5 – SampEn segmentation method results for the records in the Noise Stress Test Database (nstdb) corrupted with real electrode motion artifacts. The ECGs are created alternating noiseless with noisy segments. The SNRs for all the noise intervals in each record were: 6, 12 and 18 dB. The end of the record's name is related to this SNR value. The actual changepoints were located at 300, 420, 540, 660, 780, 900, 1020, 1140, 1260, 1380, 1500, 1620 and 1740 s. Results are expressed in terms of detection time offset (in seconds). False negatives (FNs) are represented by the symbol '–'. The number of false positives (FPs) is shown on the right column.

mitdb	300	420	540	660	780	900	1020	1140	1260	1380	1500	1620	1740	FPs
SampEn segmentation method results for the nstdb records corrupted with real electrode motion artifacts														
118e06	3	7	9	5	15	9	3	7	5	7	5	13	5	2
118e12	5	9	15	7	–	7	7	9	7	7	5	–	9	2
118e18	9	11	14	11	–	7	11	7	–	13	5	9	–	2
119e06	3	7	3	5	3	9	3	9	7	7	5	7	3	0
119e12	7	11	5	9	5	11	7	13	9	9	5	11	9	0
119e18	9	9	5	7	5	7	9	13	–	11	5	5	11	1

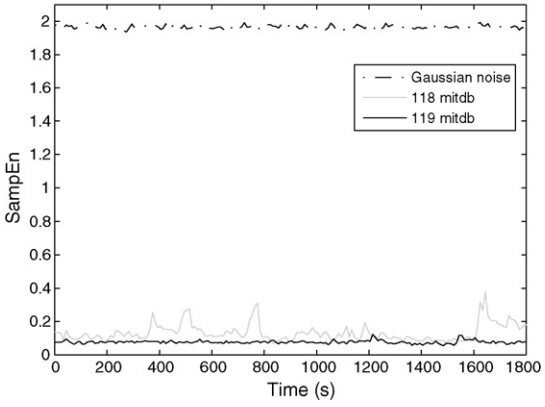

Fig. 3 – Comparison of SampEn measurements for real noiseless ECG signals (MIT database records 118 and 119) and white noise. As can be observed, the SampEn for gaussian white noise is several orders of magnitude greater than that of the signal.

the selection of the exact value for it is also very flexible.

3. Experimental studies

3.1. *Dataset*

The experimental database was composed of a set of ECG records drawn from the MIT-BIH arrhythmia database ($f_s = 360$ Hz) [32]. The noisy ECG recordings were created using calibrated amounts of white noise, pink noise, artifact muscle noise from MIT record 'ma' and electrode motion noise from MIT record 'em' [33]. The noise power level was changed at points as described in Section 4.

3.2. *Experiments*

By combining real signals with synthetic and real noise from the MIT database, we configured the following two sets of experiments:

- Real ECGs plus white gaussian and pink noise. The noiseless real recordings were corrupted with synthetic noise in several intervals. All the signals from the arrhythmia database were used.
- Real ECGs plus real noise. The noiseless recordings were corrupted with noise record 'ma' or 'em' in several intervals.

3.3. *Quantitative performance assessment*

The assessment of the experimental results was carried out in terms of the following quantitative measures:

- TP: True positive. We termed TP the detection of a noise power change at a signal location within ±15 s of an actual change.
- FP: False positive. A FP corresponds to a false alarm, it takes place when a nonexistent noise power change is detected.
- FN: False negative. False negatives account for undetected actual changes
- Sensitivity: The ratio of correctly detected changes, $S = TP/(TP + FN)$.
- E: Error. Defined as $E = (FP + FN)/(\text{real number of transitions})$.

4. Results and discussion

The results for the experiments described in Section 3.2 are shown in Tables 1–5. Tables 1, 2, 4 and 5 depict the changepoint detection bias in seconds. Table 3 shows the quantitative performance assessment in case of synthetic noise for each register. The specific parameters used were $m = 2, r = 0.25$ ([31,34,27]), $N = 14$ s (window overlapping 13 s) and $\alpha = 2.5$. Tables 6 and 7 illustrate this parameter flexibility by means of ROC values for a subset of experiments. For illustrative purposes, Fig. 4a depicts s_i for a favorable segmentation case, where the SampEn differences between consecutive intervals are clear, whereas the plot in Fig. 4b corresponds to a more difficult segmentation case due to high SNR and therefore a low SampEn. In order to assess the influence of the abruptness of the noise power changes, some experiments were repeated using a linear interpolation of the noise power level between intervals. If the size of the temporal window used to compute SampEn was larger than the duration of the noise power interval, the same results were obtained. An example of the noise power changes on ECG segmentation is shown in Fig. 5.

We also carried out an additional experiment to enable a comparative assessment of the performance of the method

Table 6 – ROC values for the *N* parameter in the SampEn segmentation algorithm. The SampEn threshold α has been set to 3.0 and one entropy sample per second is calculated. Three real records from the mitdb have been used (100, 107 and 122). The records have been corrupted with synthetic white gaussian broadband noise. The SNRs for the noise intervals were: 9, 6, 3, 0, −3, −6, −9, −6, −3, 0, 3, 6, 9 and 12 dB. The actual changepoints were located at 129, 258, 387, 516, 644, 773, 902, 1031, 1160, 1289, 1418, 1547 and 1675 s. The false positive rate (FPR) and true positive rate (TPR) mean values for the three experiments are presented. As it can be seen in the table, the detection results do not depend on *N* in the analyzed range.

N	FPR	TPR
ROC values for the *N* parameter in the SampEn segmentation algorithm		
2	0	1
4	0	1
6	0	1
8	0	1
10	0	1
12	0	1
14	0	1

proposed. We chose a method based on a time-scale approach to perform a detection of abrupt spectral changes in a non stationary signal. A stationarity index is obtained from the time-scale representation, in our case an estimation of the noise power level based on the first detail power of the wavelet transform. For each time value, two intervals are considered on both sides of this central point. If the distance between the intervals sharply peaks, there is a change at the studied instant. The selection of this specific method was based on its simplicity, ease of implementation, similar application framework, and good reported results. Its details can be found in the paper [35]. The comparative results of this experiment are shown in Tables 8 and 9 for real noise 'ma' and 'em', respectively.

Table 7 – ROC values for the α parameter in the SampEn segmentation algorithm. The window length *N* has been set to 2 s. Three real records from the mitdb have been used (100, 107 and 122). The records have been corrupted with synthetic white gaussian broadband noise. The SNRs for the noise intervals were: 9, 6, 3, 0, −3, −6, −9, −6, −3, 0, 3, 6, 9 and 12 dB. The actual changepoints were located at 129, 258, 387, 516, 644, 773, 902, 1031, 1160, 1289, 1418, 1547 and 1675 s. The false positive rate (FPR) and true positive rate (TPR) mean values for the three experiments are presented. The values for α that optimize the ROC curve are between 2 and 4.

α	FPR	TPR
ROC values for the α parameter in the SampEn segmentation algorithm		
1.0	0.47	0.72
1.5	0.20	0.63
2.0	0.04	0.94
2.5	0.01	0.99
3.0	0	1
3.5	0	1
4.0	0	1
4.5	0	0.79
5.0	0	0.47
5.5	0	0.52
6.0	0	0.23

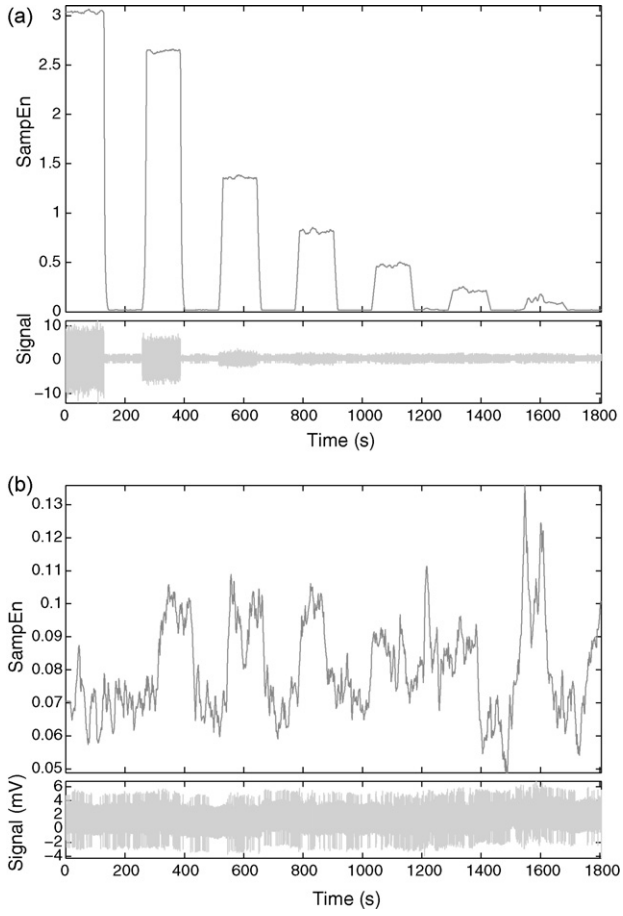

Fig. 4 – (a) Example of a favorable segmentation problem because of the clear differences between the broadband noise power in consecutive intervals (MIT record 119 with 14 different SNR levels: −20, 100, −15, 100, −5, 100, 0, 100, 5, 100, 10, 100, 15, 100). (b) Example of a difficult segmentation problem because of the high SNR that reduces the differences between the underlying signal and the noise, with regard to SampEn (MIT record 119e24 with 14 SNR intervals). Odd ones correspond to 24 dB em noise, and even ones are noiseless.

The experiments were aimed at finding the estimated changepoints and study their bias in relation to the actual ones. The results for different conditions of SNR and changepoints location showed a small detection bias in general, they fell within the ±15 s interval in most cases. Additionally, no transitions were detected at points where no change took place, in other words, the number of FP was 0 in many cases. In a few experiments, the performance of the method was poor because the signals were already noisy. This is more apparent for signals corrupted with muscle noise.

Noise power level change minimum detectable step was 3 dB. The noise influence is too small to elicit a significant SampEn change for smaller variations, specially within inter-

Table 8 – Summary of segmentation results for records corrupted with muscle artifacts in terms of quantitative measures. Comparative study.

mitdb	Time-scale based					Proposed				
	TP	FP	FN	S (%)	Err (%)	TP	FP	FN	S (%)	Err (%)
Comparison between the method proposed and a time-scale based method										
100	12	5	1	92	46	13	0	0	100	0
101	12	5	1	92	46	13	3	0	100	23
102	8	7	5	62	92	13	0	0	100	0
103	12	3	1	92	31	13	0	0	100	0
104	5	2	8	38	77	11	2	2	85	31
105	13	2	0	100	15	11	2	2	85	31
106	10	3	3	77	46	12	0	1	92	8
107	13	2	0	100	15	13	1	0	100	8
108	11	5	2	85	54	13	1	0	100	8
109	13	2	0	100	15	13	1	0	100	8
111	12	3	1	92	31	13	2	0	100	15
112	11	2	2	85	31	13	0	0	100	0
113	12	4	1	92	38	13	1	0	100	8
114	11	6	2	85	62	13	1	0	100	8
115	11	4	2	85	46	13	2	0	100	15
116	11	2	2	85	31	13	1	0	100	8
117	11	5	2	85	54	13	1	0	100	8
118	12	3	1	92	31	10	4	3	77	54
119	13	3	0	100	23	13	1	0	100	8
121	11	3	2	85	38	13	0	0	100	0
122	12	3	1	92	31	13	5	0	100	38
123	11	2	2	85	31	13	0	0	100	0
124	13	2	0	100	15	13	1	0	100	8
200	11	4	2	85	46	13	4	0	100	31
201	11	2	2	85	31	13	1	0	100	8
202	11	4	2	85	46	13	0	0	100	0
203	8	5	5	62	77	12	1	1	92	15
205	12	3	1	92	31	13	0	0	100	0
207	12	1	1	92	15	13	3	0	100	23
208	12	4	1	92	38	13	2	0	100	15
209	13	3	0	100	23	13	1	0	100	8
210	10	5	3	77	62	12	2	1	92	23
212	11	4	2	85	46	10	4	3	77	54
213	13	3	0	100	23	13	1	0	100	8
214	12	1	1	92	15	13	3	0	100	23
215	8	6	5	62	85	12	0	1	92	8
217	13	2	0	100	15	13	2	0	100	15
219	13	3	0	100	23	13	4	0	100	31
220	9	5	4	69	69	13	0	0	100	0
221	9	3	4	69	54	12	3	1	92	31
222	11	4	2	85	46	13	0	0	100	0
223	13	2	0	100	15	12	2	1	92	23
228	8	3	5	62	62	13	3	0	16	23
230	13	2	0	100	15	13	3	0	100	23
231	12	4	1	92	38	13	0	0	100	0
232	12	4	1	92	38	13	1	0	100	8
233	13	2	0	100	15	12	10	1	92	85
234	13	3	0	100	23	13	2	0	100	15
Globals	543	160	81	87	39	607	81	17	97	16

vals with high SNR. Therefore, some changepoints cannot be detected in these cases. From the experiments, we concluded that at least SNR changes should be equal or greater than 3 dB when SNR is 6 dB or higher. However, since small SNR changes in an otherwise almost clean signal do not imply noticeable changes in signal processing methods performance, there is actually no need to segment signals in these cases.

Furthermore, if the SNR is relatively high (above 20 dB), regardless of the SNR difference between consecutive segments, the method is unable to detect some changepoints either. This was somehow expected, since these cases correspond to very low power noise and therefore it is not necessary to carry out any segmentation.

With regard to the result differences between the two sets of experiments, it is apparent that the algorithm performance with 'ma' noise is lower than with synthetic white gaussian noise. The performance of the proposed algorithm improves with the bandwidth of the noise process since higher band-

Table 9 – Summary of segmentation results for records corrupted with electrode motion artifacts in terms of quantitative measures. Comparative study.

mitdb	Time-scale based					Proposed				
	TP	FP	FN	S (%)	Err (%)	TP	FP	FN	S (%)	Err (%)
Comparison between the method proposed and a time-scale based method										
118e06	13	2	0	100	15	13	2	0	100	15
118e12	13	2	0	100	15	11	2	2	85	31
118e18	10	4	3	77	54	10	2	3	77	38
119e06	13	2	0	100	0	13	0	0	100	0
119e12	13	3	0	100	0	13	0	0	100	0
119e18	11	2	2	85	31	12	1	1	92	15
Globals	73	15	5	93	19	72	7	6	92	16

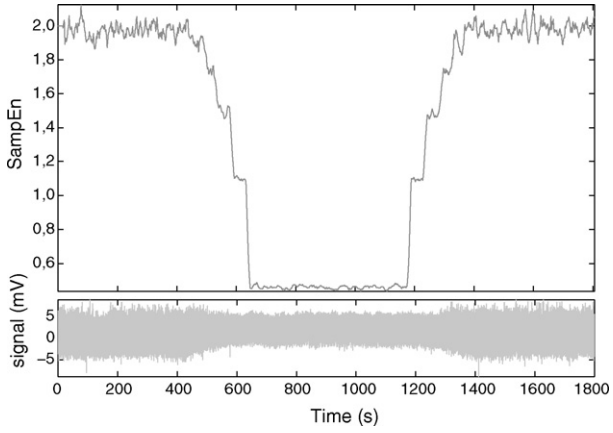

Fig. 5 – Nonabrupt noise power change. Changepoints (only two are shown for the sake of clarity) are correctly detected if the size of the windows is larger than the duration of the progressive change.

width results in higher SampEn values. As the bandwidth increases, the algorithm performs better and there is a performance degradation for colored noise as the bandwidth decreases. Practically, this algorithm is designed to be used with broadband noise and it can be used for colored noise with a continuous spectrum as long as the underlying biomedical signal is quasi-periodic (i.e. with line spectra).

In summary, the proposed method works well without parameter optimization for ECG signals corrupted by noise provided three conditions are met: (1) the noise must be broadband, (2) the SNR should be below 25 dB, otherwise the signal is almost clean and there is no need to segment it, and (3) the SNR difference between consecutive segments should be at least 3 dB, else the signal processing methods are not significantly influenced by the change. If the noise power changes are gentle instead of abrupt, the parameter N should be updated accordingly. If the resulting ramp fits into the windows, the detection is performed as in the previous cases.

5. Conclusion

We presented a novel and simple method to segment semiperiodic biomedical signals based on SampEn variations. This method can be implemented on real-time applications. It requires no filtering or signal domain transform. No prior knowledge about the number of segments, signal features or convergence conditions are necessary either. Additionally, this method can open a new field of applications of regularity metrics that further improve the results achieved in this work.

The segmentation parameters were kept constant for all the experiments since this method is intended to be unsupervised. However, some parameter adjustment can increase the accuracy of the results in case even more accuracy is required, specially for relatively clean signals, small SNR changes, or signals with smooth noise level changes instead of abrupt changes.

The algorithm performance decreases when the SNR difference between consecutive segments is smaller than 3 dB, the two SNRs are very high, or the noise is not broadband. This is compatible with the objectives of the method since small noise power changes do not cause changes in other signal processing methods parameters, there is no need to segment almost noiseless signal intervals, and other types of noise require different methods.

It is important to emphasize that our proposed algorithm is designed to solve the specific problem of segmenting semiperiodic biomedical signals corrupted with broadband noise according to the various degrees of SNR while not creating segments due to statistical changes in the underlying biomedical signals properties such as changes in frequency, amplitude, and other signal properties within physiological normality. Because of this, the resulting segments are still technically nonstationary in the sense that the underlying statistical properties of the signal change within the segment but within each of the segments the corrupting broadband noise is stationary. The problem of signal segmentation according to the broadband noise power is significant in several biomedical applications (for example, threshold selection for denoising, fiducial point detection and heartbeat classification) and our results indicate that can be solved elegantly using a simple and computationally efficient algorithm based on the SampEn. This metric may also prove useful to solve other segmentation problems.

Conflict of interest

None declared.

REFERENCES

[1] S. Tong, Z. Li, Y. Zhu, N.V. Thakor, Describing the nonstationarity level of neurological signals based on quantifications of time-frequency representation, IEEE Transactions on Biomedical Engineering 54 (10) (2007) 1780–1785.

[2] P. Prandoni, M. Vetterli, Goodwin, Optimal time segmentation for signal modeling and compression, in: IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 3, 1997, pp. 2029–2032.

[3] M. Lavielle, Optimal segmentation of random processes, IEEE Transactions on Signal Processing 46 (5) (1998) 1365–1373.

[4] S. Tahir, A. Shaameri, S. Salleh, Time-varying autoregressive modeling approach for speech segmentation, in: Sixth International, Symposium on Signal Processing and its Applications, vol. 2, 2001, pp. 715–718.

[5] A. Pedrycz, W. Gacek, A genetic segmentation of ecg signals, IEEE Transactions on Biomedical Engineering 50 (10) (2003) 1203–1208.

[6] B. Feil, J. Abonyi, O. Nemeth, P. Arva, Monitoring process transitions by kalman filtering and time-series segmentation, Computers and Chemical Engineering 29 (2005) 1419–1427.

[7] S. Jana, R. Dey, Change detection in teletraffic models, IEEE Transactions on Signal Processing 48 (3) (2000) 846–853.

[8] R. Bajwa, R. Owens, T. Kelliher, Simultaneous speech segmentation and phoneme recognition using dynamic programming, in: IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-96), vol. 6, 1996, pp. 3213–3216, doi:10.1109/ICASSP.1996.550560.

[9] A. Kehagias, E. Nidelkou, V. Petridis, A dynamic programming segmentation procedure for hydrological and environmental time series, Stochastic Environmental Research and Risk Assessment 20 (2006) 77–94.

[10] M. Goodwin, J. Laroche, A dynamic programming approach to audio segmentation and speech/music discrimination, in: IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP-04), vol. 4, 2004, pp. 309–312.

[11] D. Lemire, A better alternative to piecewise linear time series segmentation, in: SIAM Data Mining 2007, 2007.

[12] P. Bernaola-Galvan, R. Roman-Roldan, J.L. Oliver, Segmentation and long-range fractal correlations in DNA sequences, Physical Review E 53 (5) (1996) 5181–5189.

[13] T. Truonga, C. Lina, S. Chen, Segmentation of specific speech signals from multi-dialog environment using svm and wavelet, Pattern Recognition Letters 28 (11) (2007) 1307–1313.

[14] E. Keogh, P. Smyth, A probabilistic approach to fast pattern matching in time series databases, in: Proceedings of the 3rd International Conference of Knowledge Discovery and Data Mining, AAAI Press, 1997, pp. 24–30.

[15] E. Keogh, M. Pazzani, An enhanced representation of time series which allows fast and accurate classification, clustering and relevance feedback, in: Proceedings of the 4th International Conference of Knowledge Discovery and Data Mining, AAAI Press, 1998, pp. 239–241.

[16] N.C. Maddage, M.S. Kankanhalli, H. Li, Effectiveness of signal segmentation for music content representation, in: 14th International Multimedia Modeling Conference (MMM-08), vol. 4903 of Lecture Notes in Computer Science, Springer, 2008, pp. 477–486, doi:10.1007/978-3-540-77409-9-45.

[17] J. Himberg, K. Korpiaho, H. Mannila, J. Tikanmaki, H.T.T. Toivonen, Time series segmentation for context recognition in mobile devices, in: IEEE International Conference on Data Mining, 2001, pp. 203–210.

[18] Y.N. Rao, J.C. Príncipe, Time series segmentation using a novel adaptive eigendecomposition algorithm, Journal of VLSI Signal Processing 32 (2002) 7–17.

[19] D.G. Galati, M.A. Simaan, Automatic decomposition of time series into step, ramp, and impulse primitives, Pattern Recognition 39 (11) (2006) 2166–2174, doi:10.1016/j.patcog.2005.10.010.

[20] E. Punskaya, C. Andrieu, A. Doucet, W.J. Fitzgerald, Bayesian curve fitting using mcmc with applications to signal segmentation, IEEE Transactions on Signal Processing 50 (2002) 747–758.

[21] A. Kehagias, V. Fortin, Time series segmentation with shifting means hidden markov models, Nonlinear Processes in Geophysics 13 (2006) 339–352.

[22] P. Fearnhead, Exact Bayesian curve fitting and signal segmentation, IEEE Transactions on Signal Processing 53 (6) (2005) 2160–2166, doi:10.1109/TSP.2005.847844.

[23] A. Moghaddamjoo, Automatic segmentation and classification of ionic-channel signals, IEEE Transactions on Biomedical Engineering 38 (2) (1991) 149–155.

[24] A. Ukil, R. Zivanovic, Automatic signal segmentation based on abrupt change detection for power systems applications, in: Second International Symposium on Communications, Control and Signal Processing, 2006.

[25] J. Fwu, P. Djuric, Automatic segmentation of piecewise constant signal by hidden markov models, Signal Processing Workshop on Statistical Signal and Array Processing, IEEE 0 (1996) 283, doi:http://doi.ieeecomputersociety.org/10.1109/SSAP.1996.534872.

[26] A. Raskinis, G. Raskinis, Application of symbolic machine learning to audio signal segmentatin, Nonlinear Speech Modeling and Applications 3445 (2005) 397–403.

[27] J.S. Richman, J.R. Moorman, Physiological time-series analysis using approximate entropy and sample entropy, American Journal of Physiology-Heart and Circulatory Physiology 278 (6) (2000) H2039–H2049.

[28] G. Clifford, L. Zapanta, B. Janz, J. Mietus, C. Youn, R. Mark, Segmentation of 24-hour cardiovascular activity using ecg-based sleep/sedation and noise metrics, in: Computers in Cardiology, 2005, pp. 595–598, doi:10.1109/CIC.2005.1588171.

[29] S. Mahmoodi, B.S. Sharif, Signal segmentation and denoising algorithm based on energy optimisation, Signal Processing 85 (9) (2005) 1845–1851, doi:10.1016/j.sigpro.2005.03.016.

[30] R.O. Morales, M.A.P. Sanchez, J.V.L. Ginori, R.G. Abalo, R.R. Ramirez, Evaluation of qrs morphological classifiers in the presence of noise, Computers and Biomedical Research 30 (3) (1997) 200–210, doi:10.1006/cbmr.1997.1437.

[31] M. Aboy, D. Cuesta-Frau, D. Austin, P. Mico-Tormos, Characterization of sample entropy in the context of biomedical signal analysis, in: IEEE 29th Annual International Engineering in Medicine and Biology Society Conference (EMBC-07), 2007, pp. 5942–5945.

[32] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.-K. Peng, H.E. Stanley, PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals, Circulation 101 (23) (2000) e215–e220.

[33] G. Moody, W. Muldrow, R. Mark, A noise stress test for arrhythmia detectors, pp. 381–384, in: Computers in Cardiology, vol. 11, 1984.

[34] D. Lake, J.S. Richman, M.P. Griffin, J.R. Moorman, Sample entropy analysis of neonatal heart rate variability, American Journal of Physiology. Regulatory, Integrative and Comparative Physiology 283 (3) (2002) 789–797.

[35] H. Laurent, C. Doncarli, Stationarity index for abrupt changes detection in the time-frequency plane, IEEE Signal Processing Letters 5 (2) (1998) 43–45.