

Article

Comparing Performance of Nonlinear Complexity Metrics for Assessing Quality of ECGs Collected in Real Time

Yatao Zhang ^{1,2,*}, Hai Liu ², Shoushui Wei ¹ and Chengyu Liu ³

¹ School of Control Science and Engineering, Shandong University, Jinan, China
² School of Mechanical, Electrical & Information Engineering, Shandong University, Weihai, China
³ School of Instrument Science and Engineering, Southeast University, Nanjing, China
* Correspondence: zytboy@sdu.edu.cn; Tel.: +86-187-6919-7261

Abstract: We compared performance of a novel encoding Lempel-Ziv complexity (ELZC) with approximate entropy (ApEn), sample entropy (SmpEn) and permutation entropy (PerEn) as nonlinear metric to assess ECG quality. Firstly to compare performance of discerning randomness and inherent nonlinear properties within time series, this study calculated the aforementioned four nonlinear complexity values on several typical artificial time series i.e., Gauss noise, two kinds of noisy time series, two kinds of Logistic series and periodic series, respectively. Then for analyzing sensitivity of the aforementioned four complexity methods to content level of different types noise within ECG recordings, we investigated variation trend of ELZC, ApEn, SmpEn and PerEn in several synthetic ECG recordings containing different types noise (i.e., baseline wander, muscle artefacts, electrode motion, power line and mixed noise) and different signal noise ratios (i.e., 15, 10, 5, 0, -5 and -10 dB). Finally, the four complexity methods were employed to classify the quality of real ECG recordings from the PhysioNet/Computing in Cardiology Challenge 2011 (CINC 2011) of the MIT databases, then receiver operating characteristic curves (ROC) and their corresponding area under curve (AUC) were yielded. The results showed ELZC could not only distinguish randomness and chaotic within time series but also reflect content level of noise within time series, and the highest AUC of PerEn, ELZC, SmpEn and ApEn were 0.850, 0.695, 0.474 and 0.461, respectively. The results demonstrated PerEn and ELZC were more effectively than ApEn and SmpEn for assessing ECG quality.

Keywords: ECG quality assessment; complexity; entropy; ROC

1. Introduction

In practice, considerable amount of ECG recordings collected via wearable device and smart phone cannot be used in clinic because their quality is so poor that their main waveform cannot be identified, so ECG recordings in real time have to be classified their quality before diagnosis [1-3].

At present, most methods of assessing ECG quality employed time and frequency domain features extracted from ECG recordings. Langley et al. [4] just relied on time domain features i.e., flat lines, saturation, baseline drift, high and low amplitude and steep slope, to assess quality of ECG recordings, and the corresponding assessment accuracy was 85.7% on test data. Similarly Kužílek et al. [5] summarized six waveform rules to assess ECG quality and obtain 83.6% on test data. In most cases, waveform features are easy to be identified and calculated, however an assessing quality method has poor generation ability when it just relies on waveform features. Frequency domain features i.e., power spectrum density (PSD) can reflect energy distribution within time series, so it helps to extract some detailed inherent information from physiological series. So PSD was usually used for assessing quality of ECG recordings [6-8]. In [6], Zaunseder et al. proposed an assessment method based on thirty five frequency domain features generated from PSD, and assessment result of the method was 90.4% and higher than that of the other methods just based on time domain features. In fact some ECG recordings can be classified as unacceptable signals by some obvious

waveform features i.e., flat line and high amplitude. So some assessment methods combining waveform features with frequency features not only improve assessment accuracy but also help for calculation efficiency. Clifford et al. [7] and Zhang et al. [8] proposed an assessing method based on combination of time and frequency features, respectively. Comparing with the previous methods, the combination methods of multi features can provide relatively more comprehensive information within the ECG recordings. However, the methods are difficult to further improve their classification accuracy, and their generation ability also needs to be improved. So it is necessary to find new features that can reflect the inherent nature of physiological series. Actually nature of ECG signal is nonlinear, so nonlinear complexity approaches can be used for extracting the inherent nonlinear properties of ECG recordings to assess quality of ECG signals [8, 9]. Some nonlinear complexity approaches i.e., classical Lempel-Ziv complexity (CLZC), approximate entropy (ApEn), sample entropy (SmpEn) and permutation entropy (PerEn) have been extensively used for biomedical signal analysis [10-13]. Zhang et al. [14] evaluated performance of CLZC on ECG quality assessment and concluded that the CLZC was only sensitive to high frequency noise. However in practice ECG signals are usually contaminated by the mixed noise including high frequency, low frequency and power line noise, so performance of CLZC is not satisfactory when it is used as quality assessment metric. In addition, the ApEn, SmpEn and PerEn has been not yet reported to be used for assessing ECG quality. Zhang et al. [15] proposed an encoding Lempel-Ziv complexity (ELZC) to analyze the nonlinear complexity of physiological signal, and the approach can better discern random and chaotic character within time series than other LZ complexity approaches. In this study, we compared the ELZC with ApEn, SmpEn and PerEn as quality metric for assessing ECG quality.

In this study, we firstly calculate ELZC, ApEn, SmpEn and PerEn of six typical artificial time series i.e., Gauss noise, two kinds of noisy series contained different level of noise, two kinds of Logistic series (different μ) and periodic series so as to comparing performance of the four complexity methods for discerning randomness and the inherent nonlinear properties within time series. Then we calculated the four aforementioned complexity values on artificial synthetic ECG signals with different types noise (i.e., baseline wander (BW), muscle artefacts (MA), electrode motion (EM), power line (PL) and mixed noise) and Signal to Noise Ratio (SNR) so as to not only analyze the four complexity methods whether can reflect different types of noise, but also evaluate sensitivity of the four methods to noise content level of the artificial synthetic ECG signals. Finally we used the ELZC, ApEn, SmpEn and PerEn on real ECG recordings for assessing quality by receiver operating characteristic curve (ROC) and the corresponding area under curve (AUC). The study is organized as follows: Section 2 describes the four complexity methods and dataset including artificial time series and real ECG signals. The experimental results and discusses are described in Section 3. Section 4 concludes the current study.

2. Materials and Methods

2.1. Artificial synthetic and Real ECG signals

For presenting comprehensive study, we used three types of datasets including six typical artificial time series, the artificial synthetic ECG recordings and the real ECG recordings from the MIT database of the PhysioNet/Computing in Cardiology Challenge 2011 (CinC 2011).

First, six typical artificial time series were employed in this study. Gau noise represented a pure noise, MIX(0.4) means that 40 percent of a period series of length N were randomly chosen and replaced with independent identically distributed random noise, and it was used as a mixed noise series, MIX(0.2) was similar. Logistic (Logi) mapping series represent nonlinear series and are defined as

$$x_{n+1} = \mu \times x_n \times (1 - x_n), \quad (1)$$

where $1 < \mu \leq 4$. In this study, Logi(4.0) represented the Logistic series that μ were set as 4.0, and Logi(3.8) was similar. Logi(3.5) was defined as the same way and represented the periodic series.

Second, in this study, we used 48 real ECG recordings from the MIT arrhythmia database and the real common noises within ECG signals from the MIT Noise Stress Test Database (NSTDB) to construct the artificial synthetic noisy ECG recordings [16-18]. The 48 arrhythmia ECG recordings collected from the real clinical environment, each with duration of 30 mins and a sample rate of 360 Hz. Baseline correction was performed for removing the main BW noise because the noise contained in the raw signals affected the content level of BW noise within ECG recordings leading to inaccurate results in further study. The NSTDB provides three common noises that can be typically found in real clinical application, i.e., MA, EM and BW. In addition, another common noise existed in several clinic environments i.e., PL was also used to combine the artificial synthetic ECG recordings.

SNR reflects the content level of noise within signal, and SNR is defined as

$$SNR = 10 \times \log_{10} \left(\frac{P_{signal}}{P_{noise}} \right), \quad (2)$$

where P_{signal} and P_{noise} denote the power of the clean ECG and the power of the noise, respectively.

Five kinds of synthetic noisy ECG signal, i.e., the real ECG plus BW, EM, MA, PL and hybrid noise, respectively, were employed for calculating sensitivity of the ELZC on different SNR in this study. In practice, ECG signals usually cannot be used for clinical purpose when SNR of the signals lower than -10 dB because the main waveform of ECG signals cannot be recognized due to a lot of noise. So the range of SNR of the four synthetic ECG signals is from -10 dB to 15 dB. For each SNR level, 48 repeats were yielded and the mean value and standard deviation of each type signal were calculated in this study. In fact, ECG signals are usually contaminated by a mixed noise instead of one of the aforementioned noises i.e., BW, EM, MA and PL noise. Thus we built the hybrid noise and added to the real ECG recordings, with the same proportion (25%) for each noise type. Figure 1 shows the aforementioned five artificial noisy ECG signals with SNR from -10 dB to 15 dB with steps of 5 dB.

Third, the real data from CinC 2011 [2] consisted of training set in this study. Training set includes 1000 12-lead ECG recordings, and it is recordings open to all researchers and given quality label as 'acceptable' or 'unacceptable'. In training set, each recording lasts for 10s. The ECG recordings with lead-fall is no need to calculated their complexity because waveforms of the recordings look like a straight line so that the recordings are easy to be identified, and assessment accuracy is also effected by the recordings with lead-fall. So this study first removed the ECG recordings with lead-fall in training set, and then the remained ECG signals included 767 acceptable and 95 unacceptable ECG signals.

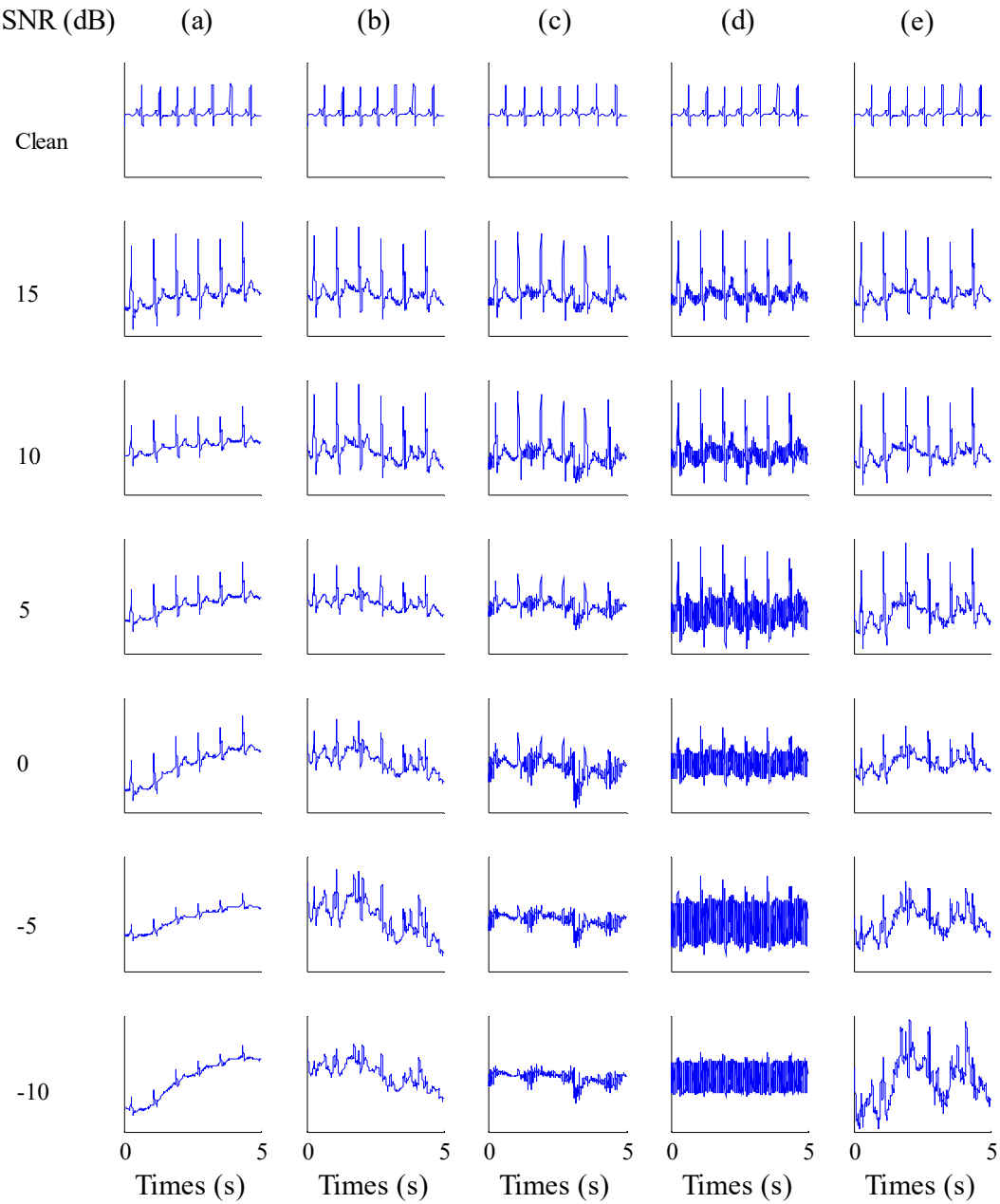


Figure 1. The real ECG recordings and the artificial synthetic noisy ECG recordings with SNR from -10 to 15 dB: (a) The real and the clean plus BW noise; (b) The real and the clean plus EM; (c) The real and the clean plus MA; (d) The real and the clean plus PL and (e) The real and the clean plus mixed noise.

2.2. Encoding Lempel-Ziv complexity (ELZC)

The classical LZ complexity consists of two steps. Firstly, an original time series is transformed into a new binary symbolic sequence by comparing with the mean or median of the original series, and then the LZ value from the binary sequence is calculated. In this study, the original series was transformed into an 8-state symbolic (3-bit binary) sequence by an encoding way.

Each x_i within the original signal $X=x_1, x_2, \dots, x_n$ is transformed into a 3-bit binary symbol $b_1(i)b_2(i)b_3(i)$, and the detailed process consists of three steps and is described as follows [15]:

Step 1, the $b_1(i)$ is determined by comparing x_i with the mean of signal X , and $b_1(i)$ is set 0 when the x_i is less than the mean, otherwise the $b_1(i)$ is 1.

Step 2, the $b_2(i)$ is 0 when the difference between x_i and x_{i-1} is less than 0, otherwise the $b_2(i)$ is set to 1. Initially, $b_2(1)$ is set to 0.

Step 3, calculated process of the third digit $b_3(i)$ is relatively complex, a variable Flag is first denoted as follows:

$$Flag(i) = \begin{cases} 0 & \text{if } |x_i - x_{i-1}| < dm \\ 1 & \text{if } |x_i - x_{i-1}| \geq dm \end{cases}, i = 2, 3, \dots, n, \quad (3)$$

where dm is the mean distance between adjacent points within signal X . Subsequently, $b_3(i)$ is calculated as follows:

$$b_3(i) = NOT(b_2(i) XOR Flag(i)), i = 2, 3, \dots, n, \quad (4)$$

where $b_3(i)$ is 0.

After the symbolic process, the LZ value of the new symbolic sequence will be calculated, and the detailed calculation process is detail in [15]. In fact, the LZ complexity counter $c(n)$ is closed related to the new subsequence of consecutive characters within the symbolic sequence. For calculating the counter $c(n)$, the symbolic sequence B is scanned from left to right and the $c(n)$ is increased one unit when a new subsequence is encountered. First, S and Q are represented two subsequences of B respectively, and SQ is the concatenation of S and Q . The subsequence $SQ\pi$ yielded from SQ after its last character is deleted (π is the operation deleting the last character in a sequence). $v(SQ\pi)$ is the vocabulary of all subsequences of $SQ\pi$. Initially, $c(n)=1$, $S=s(1)$ and $Q=s(2)$, then $SQ\pi=s(1)$. Generally, $S=s(1), s(2), s(3), \dots, s(r)$ and $Q=s(r+1)$, and so $SQ\pi=S$. Q is a subsequence of $SQ\pi$ instead of a new sequence when it belongs to $v(SQ\pi)$. Then Q is replaced with $s(r+1), s(r+2)$ and used to judge if it belongs to $v(SQ\pi)$ or not. The aforementioned processes are repeat until $Q=s(r+1), s(r+2), \dots, s(r+i)$ and it is a new sequence instead of a subsequence of $SQ\pi$, then $c(n)=c(n)+1$. Thereafter $S=s(1), s(2), \dots, s(r+i)$ and $Q=s(r+i+1)$. This above procedure is repeated until Q is the last character of B . in fact the LZ complexity counter $c(n)$ may be normalized as the $C(n)$

$$C(n) = c(n) \frac{\log_{\alpha}(n)}{n}, \quad (5)$$

where n is the length of signal X , α is the number of possible symbols contained in the new sequence. Generally, the normalized complexity $C(n)$ is consider instead of $c(n)$ in practice.

2.3. ApEn, SmpEn and PerEn

• ApEn and SmpEn

Pincus proposed ApEn as a metric to quantified regularity of a time series [19], and it means the probability of new pattern within time series when the dimension increases from m to $m+1$. The process of calculating ApEn is described as follows: Let S be a time series of length N and $S = s_1, \dots, s_N$, and reconstruct a vector \mathbf{x}_i of the embedded dimension m and $\mathbf{x}_i = s_i, s_{i+1}, s_{i+2}, \dots, s_{i+m-1}$ for $1 \leq i \leq N-m+1$ where m indicates the embedding dimension. The distance d_{ij} between the two vectors \mathbf{x}_i and \mathbf{x}_j is calculated where $1 \leq i, j \leq N-m+1$.

$$d_{ij} = \max |s(i+k) - s(j+k)|, k = 0, 1, \dots, m-1, \quad (6)$$

For each vector \mathbf{x}_i , the number of the distance d_{ij} within $r \times SD$ is found where SD is the standard deviation of the time series S , and the ratio of the number to the total number of vector $N-m+1$ is calculated as $C_i^m(r)$.

$$C_i^m = \frac{1}{N-m+1} \text{num}\{d_{ij} < r\}, i = 1, 2, \dots, N-m+1, \quad (7)$$

174 Then the average degree of similarity for all of i is defined as

$$\Phi^m(r) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln C_i^m(r), \quad (8)$$

175 Similarly, when the embedded dimension is $m+1$, the corresponding $C_i^{m+1}(r)$ and $\varphi^{m+1}(r)$ can be
176 obtained.

$$C_i^{m+1} = \frac{1}{N-m} \text{num}\{d_{ij} < r\}, i = 1, 2, \dots, N-m, \quad (9)$$

$$\Phi^{m+1}(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} \ln C_i^{m+1}(r), \quad (10)$$

177 Then, the ApEn is described as follows:

$$ApEn = \lim_{N \rightarrow \infty} [\Phi^m(r) - \Phi^{m+1}(r)], \quad (11)$$

178 In fact, the length N is not infinite, so the ApEn can be calculated by Eq. 12 when N is a finite
179 number.

$$ApEn = \Phi^m(r) - \Phi^{m+1}(r), \quad (12)$$

180 In fact, the SmpEn is a modified complexity method based on ApEn. Comparing with ApEn,
181 the SmpEn does not include self-similar patterns, and it also does not depend on data size. Finally
182 the SmpEn can be given as

$$SmpEn = -\ln[B^{m+1}(r) / B^m(r)], \quad (13)$$

183 where $B^m(r)$ is the probability that two m -dimension vectors will match, and similarly $B^{m+1}(r)$ is the
184 probability that two $m+1$ dimension vectors will match, and m is the embedding dimension[20].

185 In this study, both of ApEn and SmpEn were calculated with r as $0.15 \times SD$ where SD was the
186 standard deviation of the data series S , and m as 2.

187 • PerEn

188 The PerEn algorithm was detailed in [21-24]. A time series $\{s_i, i = 1, 2, \dots, n\}$ firstly is
189 reconstructed to generate the following matrix in phase space

$$\begin{bmatrix} S_1 \\ \vdots \\ S_i \\ \vdots \\ S_k \end{bmatrix} = \begin{bmatrix} s_1 & s_{1+\tau} & \cdots & s_{1+(m-1)\tau} \\ \vdots & \vdots & & \vdots \\ s_i & s_{i+\tau} & \cdots & s_{i+(m-1)\tau} \\ \vdots & \vdots & & \vdots \\ s_k & s_{k+\tau} & \cdots & s_{k+(m-1)\tau} \end{bmatrix}, \quad (14)$$

190 where m is the embedding dimension and τ is delay time, and $k = n - (m-1) \times \tau$. Each a row of the matrix
191 represents a reconstructed component, so there are a total number of k reconstruction components
192 in the matrix. The i th reconstructed component S_i contained m number of real values can be
193 arranged in an increasing order as

$$s_{i+(j_1-1)\tau} \leq s_{i+(j_2-1)\tau} \cdots \leq s_{i+(j_m-1)\tau}. \quad (15)$$

194 If there exists two or more reconstructed components are equal, e.g. $x_{i+(j_1-1)\tau} = x_{i+(j_m-1)\tau}$, they can be
195 sorted according to the values of j_1 and j_2 , that is, $x_{i+(j_1-1)\tau} < x_{i+(j_2-1)\tau}$ when $j_1 < j_2$. Accordingly, each of
196 reconstructed components can be transformed into a symbol series as

$$S_l = j_1 j_2 \cdots j_m, \quad (16)$$

197 where $l = 1, 2, \dots, k$ and $k \ll m!$, and $m!$ is the largest number of distinct symbols. The symbol sequence
198 S_l is one kind of arrangement that is mapped onto the m number symbols (j_1, j_2, \dots, j_m) . If the

probability of the occurrence of each symbol sequence is p_1, p_2, \dots, p_k , respectively, the PE of k kinds of different symbol sequences of time series s_i in terms of Shannon entropy can be defined as [24]:

$$H_p(m) = -\sum_l^k p_l \ln p_l, \quad (17)$$

When all the symbol sequences have the same probability namely $p_l=1/m!$, $H_p(m)$ can generate the maximum value $\ln(m!)$. Finally $H_p(m)$ can be normalized as [24]:

$$0 \leq H_p = H_p / \ln(m!) \leq 1, \quad (18)$$

The value of H_p represents the randomness degree of the time series s_i , and it described local order structure of the time series. The smaller the value of H_p , the more regular and inerratic the time series is. The change of H_p can reflect and magnify the minute details of the time series [21].

2.4. Testing performance of ELZC, ApEn, SmpEn and PerEn for discerning randomness and nonlinear properties within time series

In fact, a good quality metric should be able to discern randomness and the inherent nonlinear within physiological signals instead of confusing because physiological signal are nonlinear time series instead of random series. ELZC has a satisfied performance on discerning randomness and nonlinear properties within time series [15]. For comparing performance of the four aforementioned complexity methods, this test was designed to calculate ELZC, ApEn, SmpEn and PerEn on six typical artificial time series mentioned in Section 2.1. In this test, 20 samples for each type of series were employed, and length of each sample was 100, 500 and 2000 points.

2.4. Analyzing sensitivity of ELZC, ApEn, SmpEn and PerEn to different types noise and different SNR

In ECG quality assessment, a satisfied quality metric should be able to reflect different types and content level of noise within physiological series, meaning the metric is sensitive to different SNR of signal. Aiming to evaluate sensitivity of ELZC to noise levels of ECG signals, we compared ELZC, ApEn, SmpEn and PerEn values of the aforementioned synthetic noisy ECG signals in Section 2.1 i.e., the real ECG signals plus BW, EM, MA, PL and hybrid noise, respectively. In this test, 50 samples for each type of synthetic ECG signals were used.

2.5. Using Receiver Operating Characteristic (ROC) Curve to verify classification performance of ELZC, ApEn, SmpEn and PerEn

In this test, we calculated sensitivity (Se) and specificity (Sp) of the ApEn, SmpEn, PerEn and ELZC methods on all leads of the real ECG recordings for generating ROC curve to compare classification performance of the aforementioned nonlinear methods. For getting Se and Sp , we calculated ELZC of each lead of the real 12-lead ECG recordings from the training set A of the CinC 2011, and the thresholds were selected from 0 to 1 in steps of 0.005. An ECG recordings was marked as unacceptable when its ELZC value was higher than these threshes, otherwise it was recognized as an acceptable recording. For comparing assesement performance with ELZC, we also calculated the Sp and Se of ApEn, SmpEn and PerEn using the same method in this test. The two indexes Se and Sp are usually used for evaluating classification results, and the two concepts are described as follows:

$$\begin{aligned} Se &= \text{true positive} / \text{total number of unacceptable ECG recordings} \\ Sp &= \text{true negative} / \text{total number of acceptable ECG recordings} \end{aligned} \quad (19)$$

where true positive is number of identified unacceptable recordings among the unacceptable ECG recordings, and ture negative is number of identified acceptable recordings among the acceptable ECG recordings. Actually in this study normalization of data was necessary because the range of values obtained by the aforementioned methods (i.e., ApEn, SmpEn, PerEn and ELZC) was

different, so the Min-Max normalization approach was employed to normalize the data in this study. The approach is described as follows:

$$x^* = \frac{x - \min(x)}{\max(x) - \min(x)}, \tag{20}$$

where x and x^* are an original time series and the new series after normalizing respectively.

3. Results and Discussions

3.1. Results of discerning randomness and nonlinear properties within time series

Figure 2 shows ApEn, SmpEn, PerEn and ELZC values of six aforementioned artificial time series including Gau, MIX(0.4), MIX(0.2), Logi(4.0), Logi(3.8) and periodic series on three time lengths 100, 500 and 2000, respectively. The PerEn and ELZC values exhibit monotonically decrease in order of Gau, MIX(0.4), MIX(0.2), Logi(4.0), Logi(3.8) and periodic series on all type time lengths. The ApEn values exhibit fluctuations on MIX(0.2) and Logi(4.0) when time length is 100, and monotonically decrease in aforementioned order on 500 and 2000 length. Figure 2 shows the SmpEn values monotonically decrease in aforementioned order on 100 of time length, and exhibits fluctuations on MIX(0.2) and Logi(4.0) when time lengths are 500 and 2000 points respectively. The results indicated the ELZC and PerEn approaches had better performance for discerning randomness and nonlinear properties within time series on all type time lengths than ApEn and SmpEn. In fact Both of ELZC and PerEn do not be performed on the original time series to calculate complexity values. The ELZC approach employs firstly the encoding coarse grain method to transform the original series into a new symbolic sequence then to calculate its complexity values. Similarly PerEn calculates complexity values of a symbolic sequence generated from a reconstructed matrix instead of the original series. For the ELZC and PerEn approaches, the symbolic sequence can reflect properly nonlinear properties within time series instead of randomness because the symbolic process preserves the inherent properties within the original series, losing random component i.e., noise and unexpected components. Random components within the original series affect calculation accuracy of ApEn and SmpEn approaches when the two approaches are performed on the original time series.

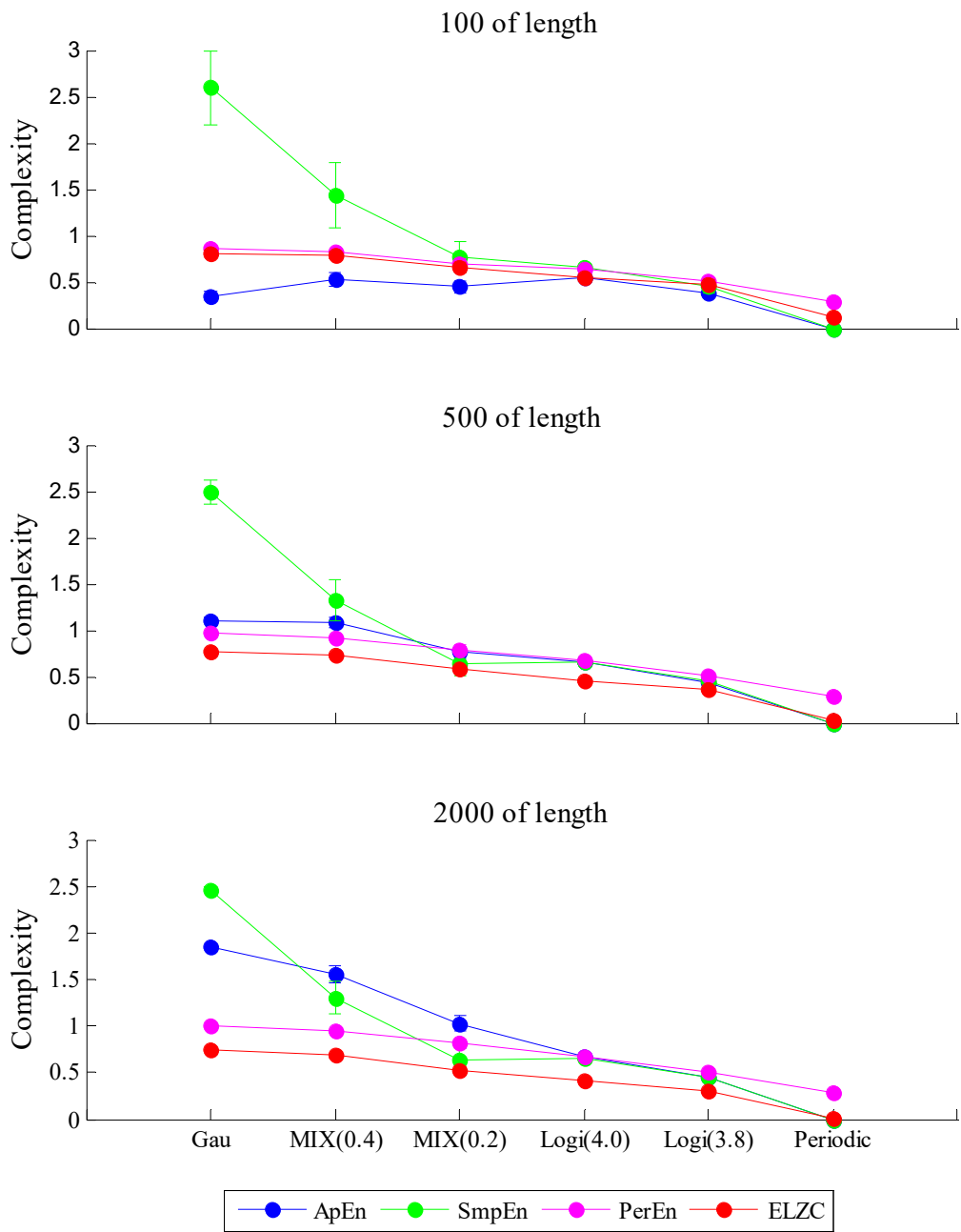


Figure 2. The ApEn, SmpEn, PerEn and ELZC Values of six typical artificial series on 100, 500 and 2000 time length, respectively

3.2. Results of sensitivity analysis

Table 1 shows the ELZC, ApEn, SmpEn and PerEn values for five artificial signals i.e., the clean ECG plus BW, the clean plus EM, the clean plus MA, the clean plus PL and the clean plus hybrid noise on several SNR from -10 to 15 dB, in steps of 5 dB. Table 1 shows ELZC values of all synthetic ECG except that of the clean ECG plus PL noise exhibit monotonically decreasing with increase of SNR. ApEn values of the clean ECG plus MA noise yield monotonically decreasing with increase of SNR but those of other synthetic ECG signals are not. SmpEn and PerEn values of the clean ECG plus BW noise exhibit monotonically decreasing with increase of SNR but those of other synthetic ECG signals are not. In fact the clean ECG should be nonlinear signal instead of random signal, so randomness within the clean ECG mixed several types noise i.e., BW, EM, MA and the

hybrid noise should increase with decreasing in SNR. The change trend of ELZC values, namely decreasing with increasing in SNR, because nonlinear properties within ECG signals are more obviously with SNR increasing instead of randomness. Table 1 shows the ELZC values of the clean plus PL is 0.0919 when SNR is -10 dB, and the ELZC values increase with increasing of SNR. Complexity of pure PL noise is nearly zero because PL is periodic signal, so complexity of the clean ECG recordings plus PL noise should be higher than that of periodic signal. Furthermore in contrast to complexity of the periodic signal, that of the clean plus PL noise should increase with the increase of SNR because nonlinear properties within the signal increases.

Table 1. The ELZC values for clean ECG plus BW, EM, MA and mixed noise on several SNR from -10 to 15 dB with steps of 5 dB.

Methods	SNR	Clean plus BW noise	Clean plus EM noise	Clean plus MA noise	Clean plus PL noise	Clean plus hybrid noise
ApEn	-10	0.2136±0.0322	0.3415±0.0212	0.8251±0.0077	0.4507±0.1044	0.3928±0.0192
	-5	0.2479±0.0401	0.3611±0.0249	0.8188±0.0147	0.5868±0.1222	0.4065±0.0224
	0	0.2766±0.0554	0.3667±0.0391	0.7862±0.0360	0.7223±0.1315	0.4019±0.0361
	5	0.2988±0.0806	0.3570±0.0662	0.7119±0.0681	0.8277±0.1323	0.3792±0.0641
	10	0.3097±0.1013	0.3401±0.0917	0.6081±0.0940	0.8634±0.1104	0.3509±0.0902
	15	0.3128±0.1126	0.3252±0.1073	0.4965±0.1074	0.7834±0.1073	0.3299±0.1064
SmpEn	-10	0.1713±0.0166	0.2544±0.0118	0.5005±0.0134	0.3525±0.0833	0.3059±0.0102
	-5	0.1852±0.0253	0.2625±0.0181	0.5069±0.0200	0.4624±0.1131	0.3085±0.0155
	0	0.1996±0.0428	0.2649±0.0349	0.4999±0.0324	0.5876±0.1340	0.3010±0.0316
	5	0.2123±0.0700	0.2594±0.0620	0.4698±0.0558	0.6978±0.1409	0.2819±0.0589
	10	0.2204±0.0924	0.2466±0.0862	0.4184±0.0800	0.7535±0.1259	0.2576±0.0842
	15	0.2211±0.1047	0.2336±0.1007	0.3546±0.0958	0.7147±0.1086	0.2377±0.0999
PerEn	-10	0.8246±0.0160	0.6990±0.0181	0.7326±0.0061	0.5597±0.0025	0.7283±0.0109
	-5	0.8152±0.0256	0.7066±0.0295	0.7285±0.0085	0.5651±0.0041	0.7287±0.0202
	0	0.8040±0.0382	0.7218±0.0425	0.7261±0.0142	0.5733±0.0064	0.7353±0.0327
	5	0.7936±0.0504	0.7399±0.0537	0.7266±0.0238	0.5840±0.0079	0.7479±0.0457
	10	0.7860±0.0574	0.7560±0.0600	0.7308±0.0356	0.5948±0.0105	0.7694±0.0558
	15	0.7803±0.0620	0.7665±0.0622	0.7405±0.0461	0.6094±0.0158	0.7681±0.0609
ELZC	-10	0.6285±0.0178	0.5672±0.0144	0.5622±0.0037	0.0919±0.0211	0.6098±0.0109
	-5	0.6033±0.0300	0.5604±0.0263	0.5581±0.0051	0.1280±0.0302	0.6000±0.0190
	0	0.5687±0.0455	0.5514±0.0388	0.5544±0.0125	0.1730±0.0408	0.5839±0.0324
	5	0.5349±0.0600	0.5368±0.0514	0.5500±0.0228	0.2324±0.0524	0.5611±0.0468
	10	0.5107±0.0747	0.5199±0.0645	0.5472±0.0369	0.3010±0.0655	0.5337±0.0615
	15	0.4985±0.0092	0.5053±0.0726	0.5408±0.0489	0.3852±0.0775	0.5109±0.0706

3.3. ROC of the ELZC on real data

Figure 3 shows ROC curves and its corresponding area under the ROC curve on each lead of the 12-lead ECG recordings. Figure 3 shows the ROC curves of ApEn and SmpEn on all leads of 12-lead ECG recordings are relatively overlap and lower than the ROC curves of PerEn and ELZC, so the corresponding AUCs of ApEn and SmpEn are also relatively smaller than that of PerEn and ELZC. Figure 3 shows that the PerEn generates the highest ROC curves, so it obtains the largest AUCs on leads I, II, III, aVR, aVF, aVL, V1 and V2 that are 0.824, 0.744, 0.675, 0.850, 0.647, 0.736, 0.647 and 0.662, respectively. The ELZC generates the second highest ROC curves on the aforementioned 8 leads, and Figure 3 also shows it obtains the second largest AUCs on

aforementioned leads. In fact the ELZC can also get the largest AUCs of 0.609, 0.336, 0.695 and 0.532 on leads V3, V4, V5 and V6, respectively.

PerEn and ELZC have relatively larger AUCs for classifying physiological series quality, especially PerEn, because the PerEn and ELZC have a satisfied performance to distinguish randomness and the inherent nonlinear properties within time series. Conversely, the ApEn and SmpEn keep smaller AUCs so that the two complexity methods cannot effectively classify quality of physiological series because they cannot discern accurately randomness and chaotic within time series. In practice, many other factors cause yield poor quality of physiological series except the lower SNR. For example poor electrode contact can cause poor quality of physiological time series so that the main waveform of the series cannot be recognized. However these physiological series are not contaminated by noise, so the series keep higher SNR. Additional, some unacceptable time series have lower SNR because noise, however the time series keep a quasi-periodic property, so that complexity of the series keeps a relatively lower level. This reason cause that classification performance of ELZC is relatively lower than that of PerEn.

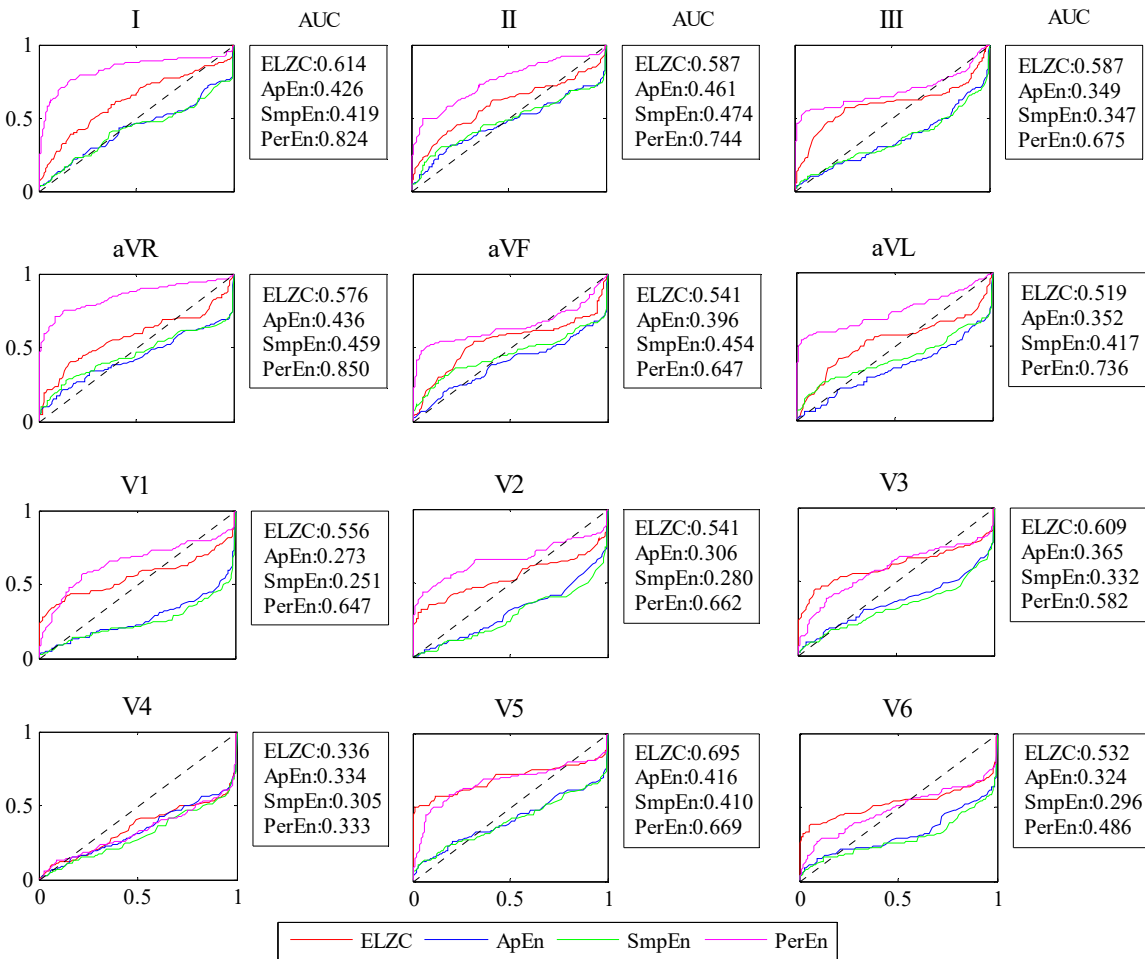


Figure 3. ROC curve of the ApEn, SmpEn, PerEn and ELZC on all leads i.e., I, II, III, aVR, aVF, aVL, V1, V2, V3, V4, V5 and V6, respectively, and its corresponding AUC.

4. Conclusions

In this study, we compared a novel ELZC complexity method with ApEn, SmpEn and PerEn for ECG quality assessment. The experiment results indicate that ELZC and PerEn have satisfied performance for distinguishing randomness and inherent nonlinear properties within time series, and ELZC can efficiently reflect types of noise and content level of noise contained in physiological signal because the ELZC values of the synthetic noisy signals decrease monotonically with increase

of SNR except that of the clean ECG plus PL noise. The experiment results on real ECG recordings indicate that the ELZC and PerEn are a satisfied quality metric for assessing quality of physiological series. In practice, the ELZC and PerEn metrics have to be combined with the other quality metrics i.e., waveform and frequency metrics for physiological signal quality assessment.

Acknowledgments: Project supported by the China Postdoctoral Science Foundation (No. 2017M612280), the National Natural Science Foundation of China (No. 61473174), and the Natural Science Foundation of Shandong Province, China (No. ZR2015AM015). We would like to thank the MIT PhysioNet for providing the open source code of PerEn and Cardiology Challenge database.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Clifford, G.D.; Moody, G.B. Signal quality in cardiorespiratory monitoring. *Physiological Measurement* **2012**, *33*, E01.
2. Moody, G.B. Physionet/computing in cardiology challenge 2011, July 2011. URL: <http://physionet.org/challenge/2011>.
3. Li, Q.; Clifford, G.D. Signal quality and data fusion for false alarm reduction in the intensive care unit. *Journal of Electrocardiology* **2012**, *45*, 596-603.
4. Langley, P.; Di Marco, L.Y.; King, S.; Duncan, D.; Di Maria, C.; Duan, W.; Bojarnejad, M.; Zheng, D.; Allen, J.; Murray, A. An algorithm for assessment of quality of ECGs acquired via mobile telephones, Proc. of the 38th Computing in Cardiology, Hangzhou, China, **2011**, 281-284.
5. Kužílek, J.; Huptych, M.; Chudáček, V.; Spilka, J.; Lhotská, L. Data driven approach to ECG signal quality assessment using multistep SVM classification, Proc. of the 38th Computing in Cardiology, Hangzhou, China, **2011**, 453-455.
6. Zaunseider, S.; Huhle, R.; and Malberg, H. Assessing the usability of ECG by ensemble decision trees. Proc. of the 38th Computing in Cardiology, Hangzhou, China, **2011**, 277-280.
7. Clifford, G.D.; Behar, J.; Li, Q.; Rezek, I. Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms. *Physiological Measurement* **2012**, *33*, 1419-1434.
8. Zhang, Y.T.; Liu, C.Y.; Wei, S.S.; Wei, C.Z.; Liu, F.F. ECG quality assessment using a kernel support vector machine and genetic algorithm with a feature matrix. *Journal of Zhejiang University-SCIENCE C (Compute & Electron)* **2014**, *15*, 564-573.
9. Arfan Jaffar, M.; Hussain A.; Mirza, Anwar M. Fuzzy Entropy and Morphology based fully automated segmentation of lungs from CT scan images. *International Journal of Innovative Computing, Information and Control* **2009**, *5*, 4993-5002.
10. Lempel, A.; Ziv, J. On the complexity of finite sequences. *IEEE Trans Inf Theory* **1976**, *22*, 75-81.
11. Balasubramanian, K.; Nair, S.S.; Nagaraj, N. Classification of periodic, chaotic and random sequences using approximate entropy and Lempel-Ziv complexity measures. *Pramana* **2015**, *84*, 365-372.
12. Yentes, J.M.; Hunt, N.; Schmid, K.K.; Kaipust, J.P.; McGrath, D.; Stergiou, N. The appropriate use of approximate entropy and sample entropy with short data sets. *Annals of biomedical engineering* **2013**, *41*, 349-365.
13. Ravelo-García, A. G.; Navarro-Mesa, J. L.; Casanova-Blancas, U.; Martin-Gonzalez, S.; Quintana-Morales, P.; Guerra-Moreno, I.; Canino-Rodríguez, José M.; Hernández-Pérez, E. Application of the permutation entropy over the heart rate variability for the improvement of electrocardiogram-based sleep breathing pause detection. *Entropy* **2015**, *17*, 914-927.
14. Zhang, Y.T.; Wei, S.S.; Di Maria, C.; Liu, C.Y. Using Lempel-Ziv Complexity to Assess ECG Signal Quality. *Journal of Medical and Biological Engineering* **2016**, *36*, 625-634.
15. Zhang, Y.T.; Wei, S.S.; Liu, H.; Zhao, L.N.; Liu, C.Y. A novel encoding Lempel-Ziv complexity algorithm for quantifying the irregularity of physiological time series. *Computer methods and programs in biomedicine* **2016**, *133*, 7-15.
16. Moody, G. B.; Mark, R. G. The impact of the MIT-BIH Arrhythmia Database. *IEEE Engineering in Medicine and Biology Magazine* **2001**, *20*, 45-50.
17. Goldberger, A.L.; Amaral, L.A.N.; Glass, L.; Hausdorff, J.M.; Ivanov, P.C.; Mark, R.G. et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* **2000**, *101*, E215-E220.

371
372
373
374
375
376
377
378
379
380
381
382
383
384

18. Moody, G.B.; Muldrow, W.E.; Mark, R.G. Noise stress test for arrhythmia detectors. *Computers in Cardiology* **1984**, 11. 381-384.

19. Pincus, S.M. Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences*, **1991**, 88, 2297-2301.

20. Chen, Y.; Oyama-Higa, M.; Pham, T.D. Identification of Mental Disorders by Hidden Markov Modeling of Photoplethysmograms. *Biomed. Inform. Technol* **2014**, 404, 29-39.

21. Bandt, C.; Pompe, B. Permutation entropy: a natural complexity measure for time series, *Phys. Rev. Lett.* **2002**, 88, 174102-1–174102-4.

22. Yan, R.; Liu, Y, Gao R X. Permutation entropy: a nonlinear statistical measure for status characterization of rotary machines. *Mechanical Systems and Signal Processing*, **2012**, 29: 474-484.

23. Li, J.; Yan, J.; Liu, X.; Ouyang, G. Using permutation entropy to measure the changes in EEG signals during absence seizures. *Entropy* **2014**, 16, 3049-3061.

24. Cao, Y.; Tung, W.; Gao, J.B.; Protopopescu, V.A.; Hively, L.M. Detecting dynamical changes in time series using the permutation entropy, *Phys. Rev. E* **2004**, 70, 46217.