Epileptic EEG Signal Analysis and Identification Based On Nonlinear Features

Yuedong Song, Jon Crowcroft
Computer Laboratory
University of Cambridge
Cambridge, United Kingdom
ys340@cam.ac.uk

Jiaxiang Zhang
MRC Cognition and Brain Sciencs Unit
Cambridge, United Kingdom
jiaxiang.zhang@cbu.cam.ac.uk

***Abstract*— In this paper, two non-linear complexity measures, namely approximate entropy and sample entropy are investigated as feature extraction methods for evaluating the regularity of the epileptic EEG signals. Furthermore, in order to obtain more efficient feature extraction for EEG signals, an optimized algorithm for sample entropy measure (O-SampEn) is proposed which removes the calculation redundancy and optimizes the computation procedure for sample entropy measure. Clinical EEG data was obtained from 20 intracranial electrodes placed within the epileptogenic zone in five epilepsy patients during both interictal and ictal periods. In terms of the experimental results, both sample entropy and approximate entropy analysis show lower values during epileptic seizures, which mean an increase of EEG signal regularity during ictal state. Compared with approximate entropy, the feature extraction based on sample entropy measure is more sensitive to EEG signal variety caused by epileptic seizures, approximately 10.14%~20.02% higher than the results using approximate entropy. In addition, the proposed optimized algorithm for sample entropy can run 9.52~36.16 times faster than the original sample entropy algorithm according to the simulation. High discrimination ability and fast computation speed of the proposed optimized sample entropy algorithm demonstrate its huge potential as a novel feature extraction method for real-time epileptic seizure detection.**

Keywords*— *Approximate entropy, Epileptic EEG, Feature extraction, Optimized sample entropy.

I. INTRODUCTION

Epilepsy is a sudden and recurrent brain malfunction and is a disease revealing an enormous and hypersynchronous activity of the nerve cells within the brain. Approximately one in every 100 individuals worldwide suffers from epilepsy [1]. An epileptic seizure is characterized by paroxysmal occurrence of synchronous oscillations. Electroencephalogray (EEG) is a non-invasive, low-cost and effective tool for diagnosing brain diseases in the clinical setting by showing the electrical activity of the brain [2]. This opens a window for physicians and researchers to study the brain's functional activities. EEG has great value for epilepsy diagnosis: it not only helps physicians to know the type of epileptic seizure, but also supplies information for medical treatment. Abnormal epileptic waveforms can be found in EEGs if there is abnormal discharge during ictal [3]. When epilepsy is diagnosed accurately, around three-fourths of epilepsy patients are able to be treated by means of medication and surgery. Epilepsy patients who need surgery are asked to come to epilepsy divisions in hospitals where these patients go through long-term presurgical assessments for localizing the epileptogenic foci in the brain [4]. During the evaluation period, a large number of EEG recordings from several electrodes are obtained from epilepsy patients, and then these EEG recordings are visually inspected and marked by neurologists for identifying epileptic seizure information. This information has great impact on localizing the epileptogenic foci in the brain and determining which region in the brain needs to be resected during neural surgery [5]. The visual marking of such long EEG recordings by human experts is obviously a very tedious, time-consuming and high-cost task, especially considering the large number of epilepsy patients in hospitals and the long-term EEG recordings. In addition, visual analysis of EEG signals is not a highly objective process. For a same EEG segment, different experts can propose different judging results, and even for the same expert, his detection result for one EEG segment can be different on different evaluations; a recent investigation of four trained experts showed that only 92% inter-expert sensitivity was obtained [6]. Therefore, automated detection of epilepsy looks very appealing if it can be made more reliable than this.

With the rapid development of non-linearity theory, complexity analysis is becoming a popular field for studying nonlinear dynamics of EEG time series. The largest advantage of complexity analysis superior to other non-linear methods such as Lyapunov Exponent [7] and Correlation Dimension [8] is that it needs a much lower quantity of data when analysing time series data like EEGs. Complexity can reflect the regularity of dynamic systems. The behaviour of various systems is different, and thus the regularity of the behaviour from these systems is also different. Complexity is capable of describing these differences and then further discriminating these systems. In information theory, 'entropy' represents the irregularity of systems, and many complexity concepts are related to entropy. Entropy is a concept handling predictability and randomness, with higher values of entropy always related to less system order and more randomness [9]. Since physiological time-series signals like EEGs are considered chaotic [8], entropy can supply recognizable variation for normal and abnormal physiological signals. Recently several entropy based measures have been applied in EEG studies. Li, Ouyang and Richards [9] investigated Permutation Entropy (PE) as a feature extracted to predict the absence seizures of genetic absence epilepsy rats by analysing EEG recordings. Permutation entropy was also utilized for analysing human

978-1-4673-2560-8/12/$31.00 ©2012 IEEE

epileptic EEG signals [10]-[12]; it is shown that the EEG during epileptic seizures is characterized by a lower value of Permutation Entropy than the normal EEG, which helps discriminate the two kinds of EEG signals successfully. Wang, Miao and Xie [13] have shown some initial investigations on Best Basis-based Wavelet Packet Entropy as a feature extraction approach for epileptic seizure detection. The general trend in current study of automatic epileptic seizure detection has focused on high accuracy but has not considered the time taken for feature extraction, which should be an important factor of developing an EEG-based detection device for epileptic seizures because fast feature extraction speed for the device not only means shorter development cycle of detection systems and faster updating release, which hugely lowers the cost of the system development, but also means the detection system can be efficiently used in a real-time setting. Therefore, some identification models with high identification accuracy may not be satisfactory when considering the trade-off between the identification accuracy and the execution time.

In this paper, two non-linear complexity measures, namely approximate entropy and sample entropy are investigated as feature extraction methods for evaluating the regularity of the epileptic EEG signals. Furthermore, in order to obtain more efficient feature extraction for EEG signals, an optimized algorithm for sample entropy measure (O-SampEn) is proposed which removes the calculation redundancy and optimizes the computation procedure for sample entropy measure.

II. MATERIALS AND METHODS

A. EEG Database

The EEG time series data employed in this study come from Department of Epileptology, Bonn University, Germany. The data has been depicted by Andrzejak, et al., [14] and used in epilepsy diagnosis and epileptic seizure detection research widely. The whole EEG data were taken from five healthy subjects and five epileptic patients. A total of 100 single-channel EEG data of 23.6s duration for each of the three categories (normal, interictal and ictal) are recorded. Set Z was taken from surface EEG signals recorded using extracranial electrodes of five healthy volunteers with eyes open. Major artifacts, e.g., due to eye or hand movements, have been manually removed by the authors of the dataset. Sets F and S originated from the EEG archive recorded using intracranial electrodes of presurgical diagnosis of five epilepsy patients. Signals in Set F were recorded from within the epileptogenic zone and it contains only brain activity measured during seizure free intervals, Set S contains only seizure activity. All EEG signals were recorded with the same 128-channel amplifier. The data were digitized at 173.6 samples per second at 12-bit resolution. Band pass filter was set to 0.53-40Hz. In this study, Set F (interictal EEG) and Set S (ictal EEG) are chosen for the study of epileptic seizure detection. Both of the datasets come from the epileptogenic zone. Each dataset has 4096 sampling points.

B. Approximate Entropy and Sample Entropy

Approximate Entropy (ApEn) proposed by Pincus [15] is a measure used to quantify the regularity or predictability of a time series. It is defined as the logarithmic likelihood that sequences of templates of certain length which are close to each other will remain close on the next incremental comparison. The more complex the time series is, the larger the value of ApEn. The ApEn algorithm counts each vector as matching itself in order to prevent the happening of ln(0) in the computations, which results in the bias of ApEn. In [16], a new family of statistics called Sample Entropy (SampEn) was introduced and characterized to avoid the bias of ApEn. The SampEn is also less sensitive to noise and can be applied to short-length time series data. Additionally, it is resistant to short strong transient interferences (outliers) such as spikes. These characteristics make Sample Entropy an appealing tool for non-linear analysis of physiological signals.

For calculating the SampEn, the embedding dimension (m) and vector comparison threshold (r) must be specified. It is common to set the embedding dimension parameter m to be m=1, 2 or 3 and to set the vector comparison threshold r to be some percentage of the standard deviation of the time series so as not to depend on the absolute amplitude of the signal [16]. SampEn(m,r,N) is the negative logarithm of the conditional probability that two sequences similar for m points remain similar at the next point, where self-matches are not included in calculating the probability. Thus, a larger value often corresponds to more irregularity or complexity in the time series data. The value of the SampEn is determined as shown in the following steps:

1) Given N data points from a time series {x(n)}= x(1), x(2),...,x(N), take m vectors $X_m(1),...,X_m(N-m+1)$ defined as $X_m(i)=[x(i),x(i+1),...,x(i+m-1)]$, for $1\le i\le N-m+1$. These vectors stand for m consecutive x values, starting at the ith sample.

2) Let r denote the noise filter level which is defined as

$$r = g \times Std \quad for \;\; g = 0.1, 0.2, ..., 0.9 \tag{1}$$

Where Std represents the standard deviation of the data sequence X.

3) The distance between vectors $X_m(i)$ and $X_m(j)$, $d[X_m(i),X_m(j)]$, is defined as the maximum absolute difference between their scalar components:

$$d[X_m(i),X_m(j)] = \max_{k=0,...,m-1}(|x(i+k)-x(j+k)|). \tag{2}$$

4) 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 represented as B_i. Then, for $1\le i\le N-m$,

$$B_i^m(r) = \frac{1}{N-m-1}B_i. \tag{3}$$

Here, note that only the first $N-m$ vectors of length m are considered in order to ensure that for $1 \le i \le N-m$, the vector $X_{m+1}(i)$ is also defined.

5) Define $B^m(r)$ as

$$B^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} B_i^m(r). \quad (4)$$

6) We increment the dimension to m+1 and compute A_i as the number of $X_{m+1}(i)$ within r of $X_{m+1}(j)$, where j ranges from 1 to $N-m$ $(j \neq i)$. We then define $A_i^m(r)$ as

$$A_i^m(r) = \frac{1}{N-m-1} A_i. \quad (5)$$

7) We define $A^m(r)$ as

$$A^m(r) = \frac{1}{N-m} \sum_{i=1}^{N-m} A_i^m(r). \quad (6)$$

Thus, $B^m(r)$ represents the probability that two sequences will match m points, whereas $A^m(r)$ represents the probability that two sequences will match for m+1 points. Since the time series length is finite, SampEn is estimated as

$$SampEn(m,r) = \ln[\frac{B^m(r)}{A^m(r)}]. \quad (7)$$

C. Optimization Algorithm for Sample Entropy

In terms of the definition of sample entropy, the sample entropy algorithm can be optimized from the following three aspects:

Firstly, there exist a lot of repetitive calculations within two loops for B^m and A^m which can be simply avoided by incorporating these two loops into one loop. Here, it is worth noting that the length m of vector comparison for calculating A^m is actually equal to m+1 although the representing form of comparisons within these two loops is the same (both are represented as N-m-1 and N-m).

Secondly, evaluation of both vector pairs $\{X_m(i), X_m(j)\}$ and $\{X_m(j), X_m(i)\}$ for similarity during each loop operation is not necessary, because we know:

$$\max_{k=0,\ldots,m-1}(|x(i+k)-x(j+k)|) \le r \Leftrightarrow \max_{k=0,\ldots,m-1}(|x(j+k)-x(i+k)|) \le r$$

Thirdly, we can take advantage of the fact that:

$$\max_{k=0,\ldots,m-1}(|x(i+k)-x(j+k)|) \le r, |x(i+m)-x(j+m)| \le r \Leftrightarrow$$

$$\max_{k=0,\ldots,m-1}(|x(i+k)-x(j+k)|) \le r, m = m+1$$

In the light of the above-mentioned three aspects, the optimized sample entropy algorithm (O-SampEn) can be depicted as follows:

Optimized sample entropy algorithm (O-SampEn)

(1): *for* $i=1$ *to* $N-m$
(2): $B[i]=0$
(3): *end*
(4): *for* $i=1$ *to* $N-m-1$
(5): $A[i]=0$
(6): *end*
(7): *for* $i=1$ *to* $N-m-1$
(8): *for* $j=i+1$ *to* $N-m-1$
(9): *if* $\max_{k=0,\ldots,m-1}(|x(i+k)-x(j+k)|) \le r$
(10): $B[i]=B[i]+1$
(11): $B[j]=B[j]+1$
(12): *if* $|x(i+m)-x(j+m)| \le r$
(13): $A[i]=A[i]+1$
(14): $A[j]=A[j]+1$
(15): *end*
(16): *end*
(17): $i=N-m$
(18): *for* $j=1$ *to* $N-m-1$
(19): *if* $\max_{k=0,\ldots,m-1}(|x(i+k)-x(j+k)|) \le r$
(20): $B[i]=B[i]+1$
(21): $B[j]=B[j]+1$
(22): *end*
(23): $B^m=0$
(24): *for* $i=1$ *to* $N-m$
(25): $B^m=B^m+B[i]/(N-m-1)$
(26): *end*
(27): $B^m=B^m/(N-m)$
(28): $m=m+1$
(29): $A^m=0$
(30): *for* $i=1$ *to* $N-m$
(31): $A^m=A^m+A[i]/(N-m-1)$
(32): *end*
(33): $A^m=A^m/(N-m)$
(34): $O-SampEn=-\ln(A^m/B^m)$

Although the computational complexity of the O-SampEn algorithm is still $O(N^2)$, however it runs much faster compared to the original edition of SampEn computation. The algorithm is explained as follows:

Firstly, two counters, B and A, represented as arrays are created. These respectively hold the values of B_i in equation (10) for vectors with size m and the values of A_i in equation (12) for vectors with size m+1. The counters are initialised to zero (steps 1-6), because self-matches are excluded in the definition of sample entropy, so it is not necessary to check each vector with itself. All vectors with size m are generated, and points $x(i)$ and $x(j)$ in these vectors are compared correspondingly (step 7). In order to remove repetitive comparisons between pairs of vectors $\{X_m(i), X_m(j)\}$ and $\{X_m(j), X_m(i)\}$, j begins from $i+1$ (step 8), which enables the value of j to be larger than the value of i all the time. The examination for similarity between two vectors with size m is executed at step 9. The ith and jth components of the counter array B are increased by one when two vectors are examined to be similar (steps 10-11). The pairs of vectors

$\{X_m(i), X_m(j)\}$ with size m+1 is also examined for similarity (step 12) by means of examining the last components of vectors $X_m(i)$ and $X_m(j)$, because the first m components have already been examined at step 9. The counter array A is increased correspondingly when similar vectors are discovered (steps 13-14). From the above-mentioned calculation, we can find that the vector $X_m(N-m)$ is not included in the examination for vector similarity, because the loop stops at the $(N-m-1)^{th}$ point in the signal and not the $(N-m)^{th}$ point. The examinations for similarities among the vector $X_m(N-m)$ and other vectors are conducted at steps 17-22, and the value of counter array B is increased correspondingly. Finally, steps 23-27 calculate B^m, steps 29-33 calculate A^m and the sample entropy is calculated at step 34.

III. EXPERIMENTS AND DISCUSSIONS

Although the pattern length parameter *m* and the threshold *r* of the EEG time series data play an important role in determining the results of ApEn and SampEn, there exist no guidelines to set the values of these parameters [16], [17]. In essence, the accuracy and confidence of the entropy estimate improves when the number of matches of length *m* and *m*+1 increases. The number of matches can be increased by means of selecting small *m* and large *r*. However, if *r* is too large, some fluctuations of the signal are not detected, and if *r* is too small, inferior conditional probability estimate is often obtained [18]. The values of m and r that are employed in the experiments are described as follows:

1) m = 1, 2, 3;
2) r = 10% - 90% of standard deviation of the EEG data sequence in increases of 10%;

We selected these values for m and r since values for m>3 lead to some EEG segments with SampEn (m>3, r) = log(0). Values of ApEn and SampEn are calculated for all interictal (seizure free epileptogenic zone segments) and ictal (epileptic seizure segments) EEG signals. Using rectangular-window with fixed sizes, data frames, each with 4096 sampling points, are formed and the values of both ApEn and SampEn are computed for each data frame. In order to find the optimal parameter combinations of SampEn and ApEn calculation for the discrimination between interictal EEG time series and ictal EEG time series, Student's t-test was utilized for evaluating the statistical differences between the estimated SampEn and ApEn values with different parameter combinations of *m* and *r* mentioned above for interictal EEG time series and ictal EEG time series. The t-test is a statistical analysis method for assessing the probability of discrimination between two groups of data. The mean and variance of each group were used for computing the discrimination probability (p-value). We calculated means and variances of SampEn and ApEn values with different parameter combinations for interictal EEG signals and ictal EEG signals, respectively. In terms of these values, we tabulate p-values of two sided t-test for investigating the potential of discrimination between groups with extracted SampEn and ApEn for different combinations of *m* and *r*, and the results are revealed in TABLE I and TABLE II.

TABLE I. STATISTICAL ANALYSIS RESULTS FOR DIFFERENT PARAMETER COMBINATIONS OF SAMPEN CALCULATION BETWEEN INTERICTAL EEG SIGNALS AND ICTAL EEG SIGNALS

	r=0.1*SD	r=0.2*SD	r=0.3*SD	r=0.4*SD	r=0.5*SD	r=0.6*SD	r=0.7*SD	r=0.8*SD	r=0.9*SD
m=1	0.3504	0.2704	0.0042	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001
m=2	<0.0001	<0.0001	<0.0001	0.0897	0.9777	0.0958	0.0019	<0.0001	<0.0001
m=3	<0.0001	<0.0001	<0.0001	<0.0001	0.0077	0.1209	0.7531	0.3125	0.0192

TABLE II. STATISTICAL ANALYSIS RESULTS FOR DIFFERENT PARAMETER COMBINATIONS OF APEN CALCULATION BETWEEN INTERICTAL EEG SIGNALS AND ICTAL EEG SIGNALS

	r=0.1*SD	r=0.2*SD	r=0.3*SD	r=0.4*SD	r=0.5*SD	r=0.6*SD	r=0.7*SD	r=0.8*SD	r=0.9*SD
m=1	0.4511	0.1370	0.0949	0.0025	<0.0001	<0.0001	<0.0001	0.0038	0.0170
m=2	<0.0001	<0.0001	<0.0001	0.1320	0.4817	0.9751	0.2252	0.0016	<0.0001
m=3	<0.0001	<0.0001	0.0023	0.0095	0.4516	0.1937	0.2217	<0.0001	0.1741

As we see in TABLE I, most groups for SampEn parameter combinations show significant difference with over 95% confidence. For m=1, high discrimination rate with over 99.99% confidence can be found when r is in the range of 0.4~0.9*SD. The group with m=1, r=0.3*SD has discrimination rate of 99.58%, which is also very high although it is not as good as the above-mentioned groups. Groups with small r values (0.1~0.2*SD) get poor discrimination rates of 64.96% and 72.96%, respectively. Hence, for short-length embedding dimensions such as m=1, the separation rate between interictal EEG signals and ictal EEG signals arises with increase of the threshold value r according to the results listed above. However, high discrimination rates with over 99.99% confidence are achieved with small threshold value r (0.1~0.3*SD) when the embedding dimension m=2. Groups with high r values such as 0.7~0.9*SD still preserve high discrimination degree of over 99.99% and 99.81%, but the separation rates of the groups with parameter combinations of m=2, r=0.4~0.6*SD drops greatly and each group has no difference according to statistical analysis (p>5%). For m=3, results of the discrimination rates are contrary to that with m=1. High discrimination rate with over 99.99% confidence can be seen when threshold value r is in the range of 0.1~0.4*SD, and then it decreases sharply with the increase of value r.

From the p-value listed in TABLE II, we found the result generated using ApEn parameters for each group is not as good as those using SampEn parameters. For ApEn, three sets of high discrimination rates of over 99.99% appears when m=1, r=0.5~0.7*SD, and there are four groups with over 99.99% significant differences when m=2, r=0.1~0.3*SD and r=0.9*SD, respectively. As for m=3, high separation rates with over 99.99% confidence appear when the threshold value r is set to be 0.1*SD, 0.2*SD and 0.8*SD, respectively.

TABLE III. COMPARISON OF THE PERFORMANCE OF SAMPEN AND APEN

	SampEn	ApEn
Number of pairs with significant difference	<0.0001 for 15 groups	<0.0001 for 10 groups

Compared with ApEn-based feature extraction method which has 10 groups of parameter combinations with over 99.99% separation capability for interictal EEG and ictal EEG, SampEn-based feature extraction method generates more groups (15 groups) with the same high separation rate as shown in TABLE III. In order to further test the capability of SampEn and ApEn for discriminating EEG signals during seizure-free state and EEG signals during seizure. A variety estimation criterion δ is defined as follows:

$$\delta = \frac{n_a - n_b}{n_a} \times 100\% \tag{8}$$

where n_a and n_b represent the values of SampEn or ApEn extracted from EEG signals during interictal state and during ictal state, respectively. If we calculate the δ of sample entropy, then that δ value reflects the varying range of the values of sample entropy from interictal state to ictal state. For different feature extraction methods, a larger value of δ usually means better discrimination capability for two groups of epileptic EEG datasets. We calculated the SampEn and ApEn for interictal EEG signals and ictal EEG signals with nine parameter combinations of m and r which have the highest discrimination rate of over 99.99% for both SampEn and ApEn

TABLE IV. THE δ VALUES OF SAMPEN AND APEN WITH OPTIMAL PARAMETER COMBINATIONS EXTRACTED FROM INTERICTAL EEG SIGNALS AND ICTAL EEG SIGNALS

	m=1 r=0.5*SD	m=1 r=0.6*SD	m=1 r=0.7*SD	m=2 r=0.1*SD	m=2 r=0.2*SD	m=2 r=0.3*SD	m=2 r=0.9*SD	m=3 r=0.1*SD	m=3 r=0.2*SD	Average value	p-Value
ApEn	33.75%	31.26%	40.16%	32.11%	21.74%	32.84%	34.53%	37.93%	42.34%	34.07% ± 5.9789	4.6507e-5<0.0001
SampEn	43.89%	42.45%	51.60%	45.35%	41.76%	49.54%	52.28%	51.17%	53.80%	47.98% ± 4.6232	

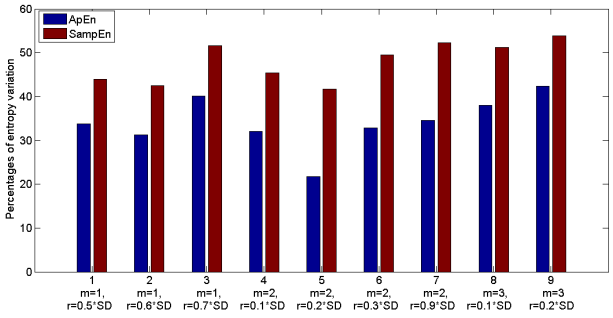

Figure 1. Comparison of SampEn and ApEn variation with optimal parameter combinations extracted from interictal EEG signals and ictal EEG signals.

calculation. As we can see from TABLE IV and Fig. 1, the δ value of SampEn is higher than that of ApEn in all 9 sets of results, which means the varying range of the values of sample entropy from interictal state to ictal state is larger than that of approximate entropy from interictal state to ictal state. The results were compared statistically by Student's t-test. The varying range of SampEn increases by 10.14%~20.02% compared to the varying range of ApEn, and there is an increase of 13.91% of the varying range on average with over 99.99% significance by using SampEn calculation. In terms of these results, we can see that SampEn performs much better than ApEn in discrimination capability. Next, we compared the computational speed of the proposed O-SampEn measure with the original SampEn measure by varying the length of EEG signals. Generally speaking, although the running time for all algorithms can be faster by means of a more advanced processor, another programming language, a different operation platform or a compiler rather than an interpreter, the chief conclusions for the algorithm efficiency itself remain unchanged. Five rectangular windows are formed using 4096, 2048, 1024, 512 and 256 sampling points respectively, such that each EEG signal is divided into 1, 2, 4, 8, or 16 small segments. For each signal length condition, the values of SampEn and O-SampEn for interictal and ictal EEG signals are calculated on the basis of each set of parameter combinations respectively, and their corresponding calculation time is recorded simultaneously. The average calculation time of SampEn algorithm and O-SampEn algorithm on all 15 set of parameter combinations is shown in TABLE V.

TABLE V. COMPARISON OF CALCULATION TIME BETWEEN O-SAMPEN AND SAMPEN

Signal length (sampling points)	Calculation time of O-SampEn (s)	Calculation time of SampEn (s)	Time ratio of two algorithms
256	0.0029±0.0022	0.0276±0.0069	9.52
512	0.0042±0.0027	0.0733±0.0074	17.45
1024	0.0086±0.0051	0.2190±0.0087	25.47
2048	0.0232±0.0079	0.7078±0.0736	30.51
4096	0.0697±0.0181	2.5204±0.0387	36.16

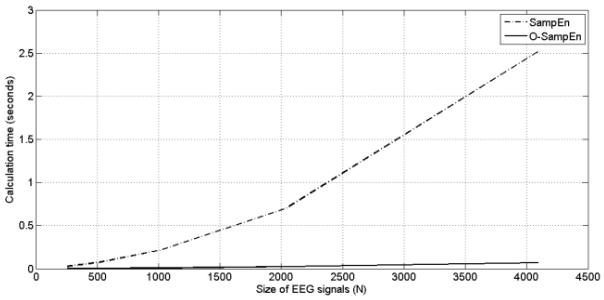

Figure 2. CALCULATION TIME WITH RESPECT TO SIGNAL LENGTH FOR EEG SIGNALS.

From TABLE V, we can see that the calculation time of both algorithms increases with the increase of signal length, and for each signal length, the O-SampEn algorithm all runs much faster than the original SampEn algorithm. We also find that O-SampEn runs 9.52 times faster than SampEn when the data length is 256 sampling points, however, the time ratio between two algorithms raises when the signal length is varied from 256 to 4096 sampling points. O-SampEn can run about 36 times faster than SampEn when the size of EEG signal reaches 4096

sampling points. That is because O-SampEn algorithm optimizes the original algorithm by the following two aspects: Firstly, the vector pairs $\{X_m(i), X_m(j)\}$ in the loop is compared only once, i.e., the vector pairs are compared forward instead of being compared backward, which removes repetitive comparisons in the SampEn algorithm. When the signal length increases, the number of removed repetitive comparisons also increase correspondingly, thus saving more and more time. Secondly, according to the definition of O-SampEn, two loops for B^m and A^m are incorporated into one loop and the similar calculation procedure between B^m and A^m is simplified by only examining the next point of the vector $X_m(i)$ when the vector dimension is increased by one. Hence, more examining time for vector similarity is saved with the increase of signal length. Fig. 2 shows the correlation between the calculation time and the signal length which is represented as N here. As reflected by the slopes of the two curves representing O-SampEn and SampEn respectively in the figure, the optimized sample entropy algorithm runs faster than the original sample entropy algorithm for larger size of EEG signal. Therefore, O-SampEn can demonstrate its advantage in computation efficiency with respect to SampEn more obviously for long time series data, which shows that O-SampEn is more suitable than SampEn for clinical applications.

IV. CONCLUSIONS

The randomness of non-linear time series data can be measured as entropy, which exhibits recognizable variations for normal and abnormal physiological signals. In this paper, we investigated the feasibility of two entropy measures, namely sample entropy and approximate entropy, as feature extraction methods to discriminate between interictal and ictal EEG signals. Furthermore, in order to obtain more efficient feature extraction for EEG signals, we proposed an optimized algorithm for sample entropy measure (O-SampEn) which removes the calculation redundancy and optimizes the computation procedure for sample entropy measure. The experimental results reveal that the values of both sample entropy and approximate entropy fall suddenly during epileptic seizure, which proves that EEG time series is more regular during epileptic seizure than during seizure-free intervals. The feature extraction based on SampEn calculation is more sensitive to EEG signal variety caused by epileptic seizures, approximately 10.14%~20.02% higher than the results using approximate entropy. In addition, the proposed optimized algorithm for sample entropy can run 9.52~36.16 times faster than the original sample entropy algorithm according to the simulation. High discrimination ability and fast computation speed of the proposed optimized sample entropy algorithm demonstrate its huge potential as a novel feature extraction method for real-time epileptic seizure detection.

REFERENCES

[1] L. D. Iasemidis, D. S. Shiau, W. Chaovalitwongse, J. C. Sackellares, P. M. Pardalos, J. C. Principe, P. R. Carney, A. Prasad, B. Veeramani, and K. Tsakalis, "Adaptive epileptic seizure prediction system," IEEE Transactions on Biomedical Engineering, vol. 50, no. 5, 2003, pp. 616-627.

[2] C. J. Stam, J. P. Pijn, P. Suffczynski, and F. H. Lopez da Silva, "Dynamics of the human alpha rhythm: evidence for non-linearity," Clinical Neurophysiology, vol. 110, no. 10, 1999, pp. 1801-1813.

[3] N. Pradhan, P. K. Sadasivan, and G. R. Arunodaya, "Detection of seizure activity in EEG by an artificial neural network: a preliminary study," Computers and Biomedical Research, vol. 29, 1996, pp. 303-313.

[4] H. Ocak, "Optimal classification of epileptic seizures in EEG using wavelet analysis and genetic algorithm," Signal Processing, vol. 88, 2008, pp. 1858-1867.

[5] H. Ocak, "Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy," Expert Systems with Applications, vol. 36, 2009, pp. 2027-2036.

[6] K. Wilson, M. L. Scheuer, C. Y. Plummer, B. Young, and S. Pacia, "Seizure detection: correlation of human experts," Clinical Neurophysiology, vol. 114, 2003, pp. 2156-2164.

[7] E. D. Übeyli, "Automatic detection of electroencephalographic changes using adaptive neuro-fuzzy inference system employing lyapunov exponents," Expert Systems with Application, vol. 36, 2009, pp. 9031-9038.

[8] S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, "Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection," IEEE Transactions on Biomedical Engineering, vol. 54, 2007, pp. 1545-1551.

[9] X. L. Li, G. X. Ouyang, and D. A. Richards, "Predictability analysis of absence seizures with permutation entropy," Epilepsy Research, vol. 77, 2007, pp. 70-74.

[10] A. A. Bruzzo, B. Gesierich, M. Santi, C. A. Tassinari, N. Birbaumer, G. Rubboli, "Permutation entropy to detect vigilance changes and preictal states from scalp EEG in epileptic patients: a preliminary study," Neurological Science, vol. 29, 2008, pp. 3-9.

[11] Y. Cao,W. Tung, J. B. Gao, V. A. Protopopescu, and L. M. Hively, "Detecting dynamical changes in time series using the permutation entropy," Physical Review E, vol. 70, 2004, pp. 046217.

[12] N. Nicolaou, and J. Georgiou, "Detection of epileptic electroencephalogram based on permutation entropy and support vector machines," Expert Systems with Applications, vol. 39, 2012, pp. 202-209.

[13] D. Wang, D. Miao, and C. Xie, "Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection," Expert Systems with Applications, vol. 38, no. 11, 2011, pp. 14314-14320.

[14] R. G. Andrzejak, K. Lehnertz, C. Rieke, F. Mormann, P. David, and C. E. Elger, "Indications of nonlinear deterministic and finite dimensional structures in time series of brain electrical activity: dependence on recording region and brain state," Physical Review E, vol. 64, 2001, pp. 061907.

[15] S. M. Pincus, "Approximate entropy as a measure of system complexity," Proceedings of the National Academy of Sciences of the United States of America, vol. 88, no. 6, 1991, pp. 297-301.

[16] J. S. Richman, and J. R. Moorman, "Physiological time series analysis using approximate entropy and sample entropy," American Journal of Physiology, vol. 278, no. 6, 2000, pp. 2039-2049.

[17] X. N. Chen, I. C. Solomon, K. H. Chon, "Comparison of the use of approximate entropy and sample entropy: applications of neural respiratory signal," Proceedings of the 27th Annual International Conference of the IEEE Engineering in Medicine and Biology, 2005, pp. 4212-4215.

[18] A. Molinaro, R. Simon, and R. Pfeiffer, "Prediction error estimation: a comparison of resampling methods," Bioinformatics, vol. 21, no. 15, 2005, pp. 3301-3307.