Fine-grained permutation entropy as a measure of natural complexity for time series

Article *in* Chinese Physics B · July 2009

DOI: 10.1088/1674-1056/18/7/011

Xiaofeng Liu
Hohai University

Vol 18 No 7, July 2009
1674-1056/2009/18(07)/2690-06

Chinese Physics B

Fine-grained permutation entropy as a measure of natural complexity for time series*

Liu Xiao-Feng(刘小峰)[a)b)†] and Wang Yue(王 越)[a)]

[a)] The Key Laboratory of Robot and Intelligent Technology of Shandong Province, and College of Information and Electrical Engineering, Shandong University of Science and Technology, Qingdao 266510, China

[b)] Institute of Artificial Intelligence and Robot, Xi'an Jiaotong University, Xi'an 710049, China

(Received 10 November 2008; revised manuscript received 11 January 2009)

In a recent paper [2002 *Phys. Rev. Lett.* **88** 174102], Bandt and Pompe propose permutation entropy (PE) as a natural complexity measure for arbitrary time series which may be stationary or nonstationary, deterministic or stochastic. Their method is based on a comparison of neighbouring values. This paper further develops PE, and proposes the concept of fine-grained PE (FGPE) defined by the order pattern and magnitude of the difference between neighbouring values. This measure excludes the case where vectors with a distinct appearance are mistakenly mapped onto the same permutation type, and consequently FGPE becomes more sensitive to the dynamical change of time series than does PE, according to our simulation and experimental results.

Keywords: complexity, entropy, dynamical change, fine-grained symbolization
PACC: 0547, 0254, 8700

1. Introduction

Complexity is widely used to examine intuitive properties of physical systems and to compare time series. Recently Bandt and his colleagues proposed the concept of permutation entropy (PE) as a natural complexity measure for time series, the advantages of which are its conceptual simplicity and high computational efficiency.[1,2] Bandt and Pompe[1] provide a variety of interesting simulation results which show that PE is similar to Lyapunov exponents in measuring the complexity of some chaotic dynamical systems. In experimental results they also show that PE, in comparison with the zero-cross rate, is robust to the choice of window length and observation noise for the detection of voiced speech segments. More importantly PE is invariant to nonlinear monotonous transformations.[1] There is a large and growing literature concerned with the measure of complexity. Included in this literature is the concept of approximate entropy (ApEn) proposed by Pincus,[3] which is a measure of the regularity of chaotic and non-stationary time series, and indicates the rate of generation of new information. From the examination of time series for similar epochs, it is shown that the more frequent and more similar epochs in the time series, the lower the corresponding ApEn. This implies that a lower value of ApEn reflects a higher degree of regularity. The ApEn is defined as follows: m is fixed to be a positive integer, and r is fixed to be a positive real number. A sequence of vectors $\boldsymbol{X}(1)$, $\boldsymbol{X}(2), \ldots, \boldsymbol{X}(N-m+1)$ is then extracted from the time series $\{x(i)\,|1 \le i \le N\}$. These vectors are constructed as $\boldsymbol{X}(i) = \{x(i), x(i+1), \ldots, x(i+m-1)\}$, of course, for each i, $1 \le i \le N-m+1$. Define $d[\boldsymbol{X}(i), \boldsymbol{X}(j)] = \max(|x(i+k-1) - x(j+k-1)|)$, with $k = 1, 2, \ldots, m$ for vector $\boldsymbol{X}(i)$ and $\boldsymbol{X}(j)$, $C_i^m(r) =$ (number of j such that $d[\boldsymbol{X}(i), \boldsymbol{X}(j)] \le r)/(N-m+1)$ and $\Phi^m(r) = (N-m+1)^{-1} \sum_{i=1}^{N-m+1} \ln[C_i^m(r)]$, then $\mathrm{ApEn}(u, m, r, N) = \Phi^m(r) - \Phi^{m+1}(r)$. Richman and Moorman propose the notion of sample entropy (SampEn),[4] which is conceptually similar to ApEn but overcomes the disadvantages of ApEn by excluding self-matches in the definition of $C_i^m(r)$, $C_i^m(r) =$ (number of j such that $d[\boldsymbol{X}(i), \boldsymbol{X}(j)] \le r, j \ne i)/(N-m+1)$. The SampEn has been successfully and widely applied in clinical cardiovascular studies.

In contrast to all known complexity measures, PE is robust to observation or dynamical noise, and moreover it is conceptually simple and computationally ex-

*Project supported by the National High Technology Research and Development Program of China (Grant No 2007AA04Z238) and the Qingdao Foundation for Development of Science and Technology, China (Grant No 06-2-2-10-JCH).
†Corresponding author. E-mail: xfliu@sdust.edu.cn

tremely fast, which enables it to be applied for huge data sets without any preprocessing and fine-tuning of parameters. Bandt and Pompe[1] argue that it should be carried out naturally to map a time series into a symbol sequence, so they ingeniously take partitions by comparing neighbouring values, thereby contributing almost all features of PE. On the other hand, only the order pattern is involved, which means that coarse-grained symbolization omits important features characterizing the time series, being the magnitude of the difference between neighbouring values. It causes the failure of PE in capturing the dynamical change in some time series. This motivates us to work on improvements. In the present paper we take into account the magnitude of the difference between neighbouring values when symbolizing the time series. The proposed measure not only retains all the advantages and merits of PE, such as easy and fast computation, robustness, and simplicity, but more importantly it improves the performance for detecting the dynamical change of time series and approximates more closely to the Lyapunov exponent for the chaotic time series. We demonstrate its excellent features in different examples which include some of the most studied families of chaotic systems, as well as experimental data.

The paper is organized as follows. In Section 2 we review the main results and notion of PE and then provide our definition of fine-grained PE (FGPE). We also demonstrate the performance of FGPE for detecting dynamical changes in time series. In Section 3 we show that our FGPE is more like the Lyapunov exponent for the logistic map than is PE. Section 4 contains the main results on epileptic seizure detection. We finally close with concluding remarks in Section 5.

2. Permutation entropy and fine-grained permutation entropy

Permutation entropy: given a scalar time series $\langle x(i)\rangle = \{x(i), i = 1, 2, \ldots, N\}$, where N is the total number of data points. The parameter n, the embedding dimension of the vector to be formed, must be specified in advance, and then an embedding procedure forms $N-n+1$ vectors $\boldsymbol{X}(1), \boldsymbol{X}(2), \ldots, \boldsymbol{X}(N-n+1)$, $\boldsymbol{X}(i) = [x(i), x(i+1), \ldots, x(i+n-1)]$, $i = 1, \ldots, N-n+1$. Then the n real values of vector $\boldsymbol{X}(i)$ are permutated in increasing order: $\{x(i+k_1-1) \le x(i+k_2-1) \le \ldots \le x(i+k_n-1),\ 1 \le k_1, k_2, \ldots, k_n \le n\}$. As a result, any vector $\boldsymbol{X}(i)$ is mapped onto a vector $[k_1, k_2, \ldots, k_n]$, which is one of the $n!$ permutations of n distinct symbols $\{0, 1, \ldots, n-1\}$. Let $Q(\pi)$ denote the amount of any permutation pattern appearing in the time series, and its relative frequency is

$$p(\pi) = \frac{Q(\pi)}{N-n-1}. \tag{1}$$

The PE of order $n \ge 2$ is defined as

$$H(n) = -\sum_{i=1}^{n!} p(\pi_i) \ln p(\pi_i). \tag{2}$$

Obviously $H(n)$ is bounded in $[0, \ln(n!)]$ where the lower bound is reached for the time series with increased or decreased sequence values and the upper bound for a completely random system. In general, $H(n)$ is normalized by $(n-1)$,[1]

$$h_n = H(n)/(n-1). \tag{3}$$

The fault of PE: given vectors of arbitrary time series, [0.95, 0.93, 0.30, 0.08] and [0.97, 0.11, 0.09, 0.07], they are intuitively distinct from each other as shown in Fig.1(a), but both of them are symbolized as the same ordinal pattern 3210. Furthermore, let us consider a time series consisting of two or more data slices with distinct appearance but the same permutation type. Here, we construct a time series with a data slice $\{0.15,\ 0.2,\ 0.05,\ 0.3\}$ recurring $2M$ times and other two data slices $\{0.17,\ 0.28,\ 0.005,\ 0.9\}$ and $\{0.18,\ 0.31,\ 0.001,\ 1.7\}$, all of which have the same permutation type 1203. To satisfy the condition that there are no or only very few tied ranks, a random noise is added to the time series. It should be pointed out that the noise is small enough to hold the permutation type of the slices unchanged, as shown in Fig.1(b). To compute PE, we choose the length of the sliding window as $W = 1000$, step size=1, and embedding dimension $n = 4$. We expect that the permutation entropy is capable of detecting bursts in the time series. However PE remains constant for all these slices, as shown in Fig.1(c). The sample entropy with $W = 1000$, step size=1, $m = 2$ and $r = 0.2 \cdot \mathrm{SD}$, in contrast, sharply falls corresponding to bursts in the time series, as shown in Fig.1(d). The causes of PE's failure to capture the dynamical changes in amplitude for the time series, such as bursts appearing in the time series, is that the coarse-grained symbolization of PE omits the more dynamically changing information in the time series. However, it is well known that the detection of dynamical changes in complex systems plays a key role in areas such as weather forecasting, pathological diagnosis, earthquake prediction, stock prediction, and market forecasting.

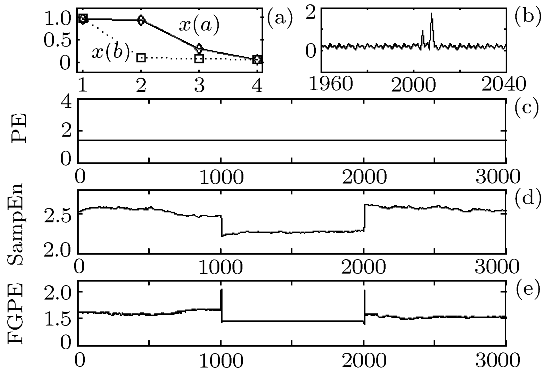

Fig.1. Comparison of complexity measures among PE, SampEn, and FGPE: (a) $X(a) = [0.95, 0.93, 0.30, 0.08]$ and $X(b) = [0.97, 0.11, 0.09, 0.07]$; (b) part of the constructed time series $\langle x(i)\rangle$; (c) PE of $\langle x(i)\rangle$ (length of sliding window $W = 1000$; step size=1, order $n = 4$); (d) SampEn of $\langle x(i)\rangle$ (length of sliding window $W = 1000$, step size=1, $m = 2$ and r equals to 0.2 times standard deviation, i.e. $r = 0.2\cdot\mathrm{SD}$); (e) FGPE of $\langle x(i)\rangle$ (length of sliding window $W = 1000$, step size=1, order $n = 4$ and $\alpha = 1$).

PE focuses only on the order of the elements in the time series. The values of the elements, as comparatively key information, is however not taken into consideration, which suggests that it is not concerned about the degree to which the neighbouring elements differ from each other. To overcome this fault we proposed a concept of FGPE in which a factor q is introduced to quantify the difference between the neighbouring elements in the vector $\boldsymbol{X}(i)$, and is added in permutation type as an additional element

$$q = \left[\frac{\max\left(D\left(i\right)\right)}{\mathrm{SD}\left\langle d\left(i\right)\right\rangle \times \alpha}\right], \tag{4}$$

where [.] returns the largest integer value that is equal to or less than a number, $\langle d(i)\rangle$ is a difference series, defined by $\langle d(i)\rangle = \{|x(i+1) - x(i)|, i = 1, \ldots, N-1\}$, $\max(D(i))$ returns the maximum value of differences between the neighbouring elements in vector $\boldsymbol{X}(i)$, $\mathrm{SD}\langle d(i)\rangle$ returns the standard deviation of difference series $\langle d(i)\rangle$, α is a precision regulation factor. According to Eq.(4), if $\alpha > \dfrac{\max(D(i))}{\mathrm{SD}\langle d(i)\rangle}$, the only possible value of q is zero. So this additional element does not affect the result of permutation, which means that the entropy value will be exactly the same as PE. If α falls in the interval $\left(0, \dfrac{\max(D(i))}{\mathrm{SD}\langle d(i)\rangle}\right]$, then q has more possible values, which results in more permutation type π. The closer the α to zero, the more permutation types it will produce. In the present paper, we apply $\alpha = 1$.

In the calculation of FGPE, let $Q'(\pi)$ denote the amount of any permutation pattern acquired by this newly defined method. Then, the relative frequency is

$$p'(\pi) = \frac{Q'(\pi)}{N-n-1}. \tag{5}$$

The FGPE of order $n \geq 2$ is defined as

$$\mathrm{FGPE} = -\sum_{i=1}^{n!} p'(\pi_i) \ln p'(\pi_i). \tag{6}$$

Baisically, q carries the information of the data's actual magnitude into the calculation of FGPE. Thus FGPE considers both the permutation and magnitude of the data. In contrast, PE only takes the rank and location of data into account.

Let us reconsider the time series $\langle x(i)\rangle$ by FGPE. Fixing $\alpha = 1$, we find that permutation types of $\{0.15,\ 0.2,\ 0.05,\ 0.3\}$ and $\{0.17,\ 0.28,\ 0.005,\ 0.9\}$ are 1203 2 and 1203 9, respectively, which enable FGPE to distinguish the two parts of the series. We find that FGPE is sensitive to bursts in time series as shown in Fig.1(e), which is consistent with SampEn. An indicator of the amplitude information q as additional element is appended to the permutation symbols. This not only renders the result of FGPE analogous to the one of SampEn, but also excludes the case where vectors with a distinct appearance are mistakenly mapped onto the same permutation type. More importantly, the time consumption of FGPE is slightly more than PE, but far less than SampEn.

3. Results of time series from logistic maps

As shown in Figs.2(b) and 2(c), we can clearly see that both PE and FGPE have a very similar appearance over the whole chaotic regime, and are very similar to Lyapunov exponents. A parameter δ is introduced to accurately compare PE and FGPE in terms of the approximation to the Lyapunov exponent, which is defined by

$$\delta = \mathrm{SD}\left(\left|\lambda\left(i\right) - h_n\left(i\right)\right|\right), \qquad i = 1, \ldots, N-n+1, \tag{7}$$

where $\lambda(i)$ is the element of the Lyapunov exponent series, $h_n(i)$ is the element of FGPE or PE series.

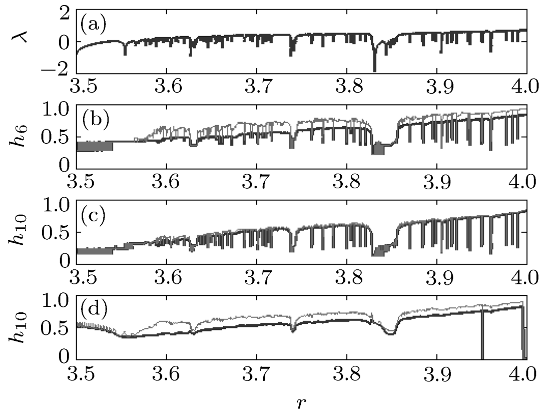

Fig.2. Logistic equations for varying control parameter r (step $\Delta r = 10^{-4}$): (a) Lyapunov exponent λ; (b) permutation entropy h_6 (lower line) and FGPE h_6 (upper line); (c) h_{10} of PE (lower) and FGPE (upper); (d) h_{10} of PE (lower) and FGPE (upper) with additive Gaussian dynamical noise at standard deviation SD = 0.00025 ($T = 10^6$ data for each r value).

We have made a comparison of parameter δ between PE and FGPE, by fixing the step at $\Delta r = 10^{-4}$ and data length $T = 10^6$, and varying the embedding dimension n from 6 to 10. We find that the FGPE approximates more closely to the Lyapunov exponent than does PE in terms of parameter δ (see Table 1) and, moreover, the approximation between FGPE and the Lyapunov exponent can be regulated by factor α in the FGPE, by making δ smaller and thus the approximation better.

Table 1. Degree of approximation of PE and FGPE to the Lyapunov exponent.

n	FGPE		PE
	$\alpha = 1$	$\alpha = 0.5$	
6	0.14571	0.13101	0.16763
7	0.14071	0.13056	0.15347
8	0.13523	0.12844	0.14375
9	0.13221	0.12765	0.13768
10	0.1302	0.12569	0.13347

We have investigated the relation between the data length and the embedding dimension, and taken 1000 time series of size $T = 10^k$ with $k = 3, \ldots, 6$. In Fig.3(a), we find that PE increases up to $n = 7, 8$ and then decreases, which replicates the result reported by Bandt and Pompe.[1] Bandt and Pompe[1] point out that the limit of PE for $n \to \infty$ exists and is 1. In contrast to PE, we see in Fig.3(b) that FGPE monotonically decreases up to $n = 10$.

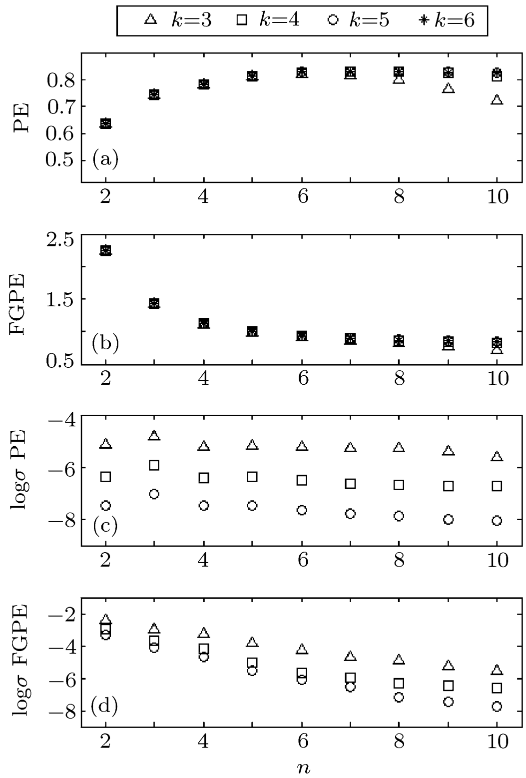

Fig.3. Logistic map with $r = 4$: (a) mean PE of order n for varying length $T = 10^k$ of time series; (b) mean FGPE of order n for varying length $T = 10^k$ of time series; (c) corresponding standard deviation σ of PE; (d) corresponding standard deviation σ of FGPE.

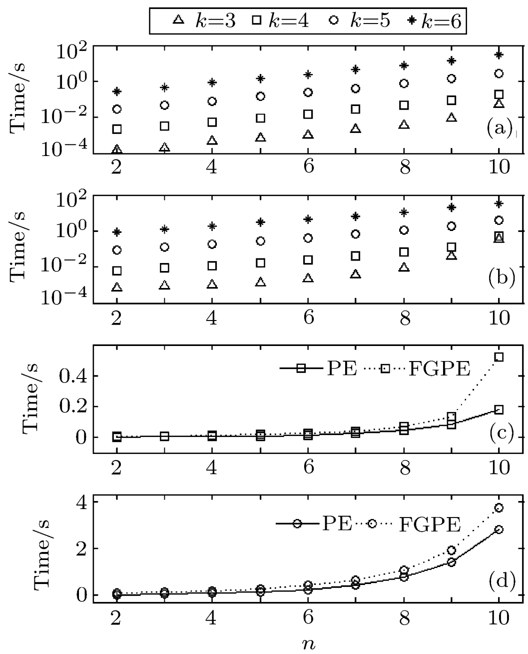

Fig.4. Time consumption (logarithm value) for computing PE and FGPE: (a) time consumption of PE of order n for varying length $T = 10^k$ of time series; (b) time consumption of FGPE of order n for varying length $T = 10^k$ of time series; (c) time consumption of PE and FGPE of order n for $T = 10^4$; (d) time consumption of PE and FGPE of order n for $T = 10^5$.

In view of the fact that fast calculation is an important feature of PE, we also compare the time consumption between FGPE and PE for computing the aforementioned time series. The time consumption for computing PE monotonically increases with the dimension order n for a time series with constant length, and likewise increases with the data length for the constant order (see Fig.4(a)). This is the same case for computing FGPE (see Fig.4(b)). As shown in Figs.4(c) and 4(d) we find that the time consumption for FGPE and PE is of the same order of magnitude and that the time consumption for FGPE increases evidently in comparison with that for PE only for the case $r > 9$. As a result, FGPE still possesses the feature of fast computation.

4. Results on the epileptic seizure detection

PE is well suited for the exploration of long and complex time series, and so the application of PE to the analysis of epileptic activity has been widely explored by researchers in related fields. In a recent study by Cao *et al*,[5] PE is utilized as an effective tool for seizure detection. Cao and his colleagues analyse intracranial electroencephalogrophy (EEG) signals from different patients, and find that PE has a sharp drop followed by a gradually increase slightly after the seizure. The paramount feature of PE, high computational speed, makes it promising for clinical seizure prediction. Li and his colleagues[6] have explored the predictability analysis of absence seizures with PE. They demonstrate that the values of PE are at a higher level during normal state and gradually decrease prior to the seizure, and moreover, they show that PE performs better for seizure prediction than that done by sample entropy. These findings are consistent with previous studies which reported the low signal complexity of EEGs during epileptic seizures.[5,6] In the present paper, we also analyse epileptic EEGs with PE and FGPE, respectively. The EEG data consists of five sets, each containing 100 single-channel EEG signals of 23.6 s with sample rate 173.61 Hz. These data are available from the EEG database in the Bonn University (for a detailed description of the data the reader is referred to Ref.[7]). It is found that the values of both PE and FGPE evidently decrease with the seizure state in comparison with the normal state which is highly suggestive that both of them are capable of distinguishing the seizure state from the normal state as shown in Fig.5. Furthermore, the values of FGPE are much higher than those of PE regardless of the state, and the discrepancy between the normal state and the seizure state is far more salient for FGPE in contrast to that of PE.

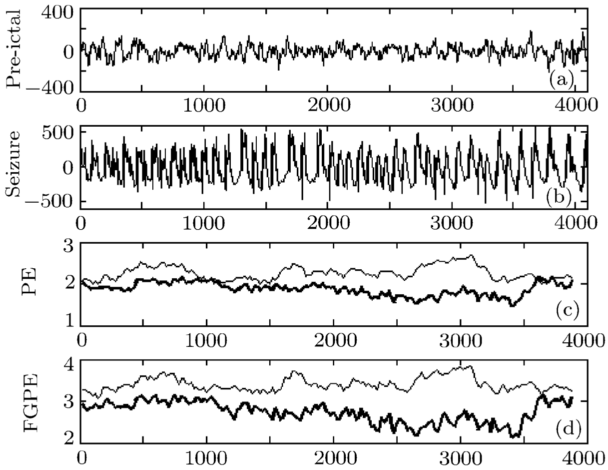

Fig.5. The complexity analysis of epileptic time series from the EEG database in the Bonn University: (a) preictal EEG signals from N040 of set C, measured in seizure-free intervals in the epileptogenic zone; (b) seizure EEG signals from S040 of set E with seizure activity; (c) PE of EEG in normal state (upper line) and seizure state (lower line) (length of sliding window $W = 200$ samples, order $n = 4$); (d) FGPE of EEG in normal state (upper line) and seizure state (lower line) (length of sliding window $W = 200$ samples, order $n = 4$).

5. Conclusion

As stated in the abstract, the main objective of the present paper is to improve PE. The magnitude of the difference between neighbouring values is taken into account when symbolizing the time series. Similar to PE, FGPE can also be applied in arbitrary real-world time series. FGPE is still computationally fast, although the time consumption of FGPE is slightly larger in comparison with PE. Most importantly, it improves the performance for detecting the dynamical change of time series and approximates more closely to the Lyapunov exponent for the chaotic time series from the logistic map in comparison with that of PE. In view of the fact that FGPE holds almost all the merits of PE, such as easy and fast computation, robustness, and simplicity, FGPE is a promising tool for complexity measurement of huge data sets with dynamical changes.

Acknowledgement

We are grateful to Professor Mike Collier and Dr Bin Yan for assistance with respect to the English.

References

[1] Bandt C and Pompe B 2002 *Phys. Rev. Lett.* **88** 174102
[2] Bandt C, Keller G and Pompe B 2002 *Nonlinearity* **15** 1595
[3] Pincus S M 1991 *Proc. Natl. Acad. Sci.* USA **88** 2297
[4] Richman J S and Moorman J R 2000 *Am. J. Physiol Heart Circ. Physiol.* **278** H2039
[5] Cao Y, Tung W W, Gao J B, Protopopescu V A and Hively L M 2004 *Phys. Rev.* E **70** 046217
[6] Li X, Ouyang G and Richards D A 2007 *Epilepsy. Res.* **77** 70
[7] Andrzejak R G, Lehnertz K, Mormann F, Rieke C, David P and Elger C E 2001 *Phys. Rev.* E **64** 0619