Improved Multiscale Permutation Entropy Measure for Analysis of Brain Waves

Young-Seok Choi
Department of Electronics and Communications Engineering, Kwangwoon University, Seoul, Korea

Abstract
This paper presents a novel multiscale entropy measure for analyzing brain rhythms, called electroencephalograms (EEGs), with an aim of unveiling the underlying dynamics of EEG over multiple time scales. This work is inspired by the known fact that neurological signals such as EEG has distinct dynamics over different temporal scales. To reflect the nonlinear and nonstationary nature of EEGs, the recently developed empirical mode decomposition is incorporated, allowing an EEG to be decomposed into its inherent multiple temporal scale elements, referred to as intrinsic mode functions (IMFs). By computing the permutation entropy of each IMF in a time-dependent manner and averaging them over multiple scales, it yields a data-adaptive multiscale permutation entropy (DAMPE) measure for analyzing the brain waves. Simulation and experimental results show that the proposed DAMPE is efficient to reveal the dynamical changes over multiple scales.

Keywords: Entropy, Electroencephalogram, Multiscale, Empirical mode decomposition

1. Introduction
Electroencephalogram (EEG) is a measurement of the electric activity from the scalp, obtained with the aid of an array of electrodes [1]. After amplification, the signal is usually saved in graphic or digital format. EEG signals have been widely used as a clinical diagnostic and research tool in recent decades [2–4]. The most significant issue of EEG studies is evaluating and quantifying the EEG waves. However, the conventional clinical method of observing EEG signals depends on the technicians’ experience, thus being considered as subjective. Thus, the need for objective measures gives rise to the development of quantitative EEG (qEEG) measure to unveil various neurological states [5].

In recent decades, an entropy based qEEG has been long of interest in characterizing dynamics of EEG recording because EEG signals have nonstationarity [5, 6]. As a primary tool for quantifying EEG signals, an entropy measures a degree of regularity of the underlying EEG signals [7, 8].

One of the widely used entropy measures for analysis of EEG is the permutation entropy (PE). The PE has been developed for quantifying the regularity for nonstationary and noisy time-series [9]. Since the PE computes the probability distribution based on the temporal structure of a time-series, it yields simplicity and robustness as advantages of the PE. Thus the PE is suitable for evaluating the physiological data such as EEG signals.

In addition, the PE is not restricted on the type of a time-series, thus implying its appropri-
ateness for analyzing EEG signals [10][12]. However, since the PE is based on single scale computation, it fails to reflect the dynamical characteristics across multiple temporal scales. However, EEG signals are representative of the interacting mechanisms between numerous neurons across multiple temporal and spatial scales, resulting in dynamical changes in both time and frequency associated with the states of the brain [1]. Thus, the single scale based PE which is calculated using gross EEG signals may not properly reflect the underlying dynamics of EEG over multiple scales.

To address this obstacle, we presents a multiscale based PE measure which utilizes the empirical mode decomposition (EMD) approach [13] for evaluating entropy values. The EMD method, which has been recently introduced as a data-driven and adaptive technique, is known to be appropriate for analyzing nonstationary and nonlinear time-series [13]. It decomposes a time-series into a number of narrowband components, called the intrinsic mode functions (IMFs), by empirically identifying the physical time scales intrinsic to the signal. Thus, due to the potential of EMD, it has been gradually used to analyze neurophysiological recordings such as EEG [14][15]. In addition, since it is known that the EMD behaves as a dyadic filter bank [16], it is well fit for detecting the dynamics of the frequency bands of interest in EEG study [14]. Upon the results of the EMD of EEG, the proposed measure computes the PE using the IMF at each scale, followed by averaging PE values over multiple scales. Thus, the resultant data-adaptive multiscale PE (DAMPE) reflects distinct features over multiple scales.

To demonstrate the performance of the proposed DAMPE, the simulation and experimental studies using a synthetic signal and an animal model during brain injury and recovery after cardiac arrest have been carried out. The performance of the DAMPE was demonstrated by comparing with the conventional single scale based PE in terms of both how well discriminate the degree of uncertainty and how accurately the entropy measure predict neurological outcomes.

The remainder of the paper is organized as follows. Section 2 provides a brief description on the proposed DAMPE measure. Section 3 validates the proposed approach through simulation and experimental studies. Section 4 presents the conclusions.

2. Data-Adaptive Multiscale Permutation Entropy

2.1 Adaptive Decomposition based on Empirical Mode Decomposition

The EMD was developed by Huang et al. to decompose a signal into its IMFs, which contains only one frequency component at each time point as a monocompont/narrowband signal [13]. To be an IMF, it needs to meet two conditions: (1) the number of extrema and the number of zero crossings differ at most by one, and (2) the mean of the local maxima and minima is approximately zero. In an iterative manner, termed a sifting process, EMD extracts the highest frequency oscillation (finest temporal scale) from the underlying time series, being considered as an IMF. The remaining part after the extraction contains lower frequency oscillatory components. The resulting IMFs represent the oscillatory patterns at different scale. This gives rise to the following major feature of EMD. EMD results in basis functions which are derived from the time-series in self-originated way, whereas other conventional methods such as Fourier and wavelet analyses rely on the use of pre-defined basis functions.

Let $x_i$ denote the raw sampled EEG signal. Then, the EMD consists of the following steps:

1. Identify all the local maxima and minima of $x_i$.
2. Interpolate between local maxima and minima respectively, getting an upper envelope $e_i^u$ and a lower envelope $e_i^l$.
3. Compute the mean between $e_i^u$ and $e_i^l$, i.e., $\mu_i = [e_i^u + e_i^l]/2$.
4. Subtract the mean from the original signal $d_i = x_i - \mu_i$.
5. Repeat steps 1–4 until $d(i)$ satisfies the above two criteria to be an IMF. If $d_i$ satisfies conditions, it becomes the first IMF that contains the finest temporal scale in the signal. Also denote as $d_1^i$.
6. Compute the residue $r_1^i = x_i - d_1^i$.
7. Iterate through steps 1–6 with $x_i$ instead of $x_i$ until the residue satisfies some stopping criterion [13].

Through the sifting process, the raw EEG signal $x_i$ is decomposed as follows:

$$x_i = \sum_{k=1}^{K} d_i^k + r_i^K,$$  (1)
where $K$ is the number of all extracted intrinsic mode functions, $d^k_i$ is the $k$th IMF, and $r^K_1$ is the final residue. The last residue $r^K_1$ can be considered as the last IMF, and thus Eq. (1) can be rewritten as $x_i = \sum_{k=1}^{K+1} d^k_i$.

### 2.2 Permutation Entropy

As recently developed by Bandt and Pompe [9], the PE is a measure of dynamic changes of a time-series by comparing neighboring values. The time-series is transformed into a series of ordinal patterns which describes the order relations between the present values and a fixed number of equidistant values at a given past times. Based on the counting of ordinal patterns, called motifs, PE quantifies the relative frequencies of the distinct motifs. Since the computing of PE considers just ordinal patterns not the amplitude of the time-series, thus it leads to robustness against measurement noise.

For a given time series $\{x_i : 1 \leq i \leq N\}$, we can embed the time series in a $m$-dimensional space to derive a reconstruction vector $X_i$ as follows:

$$X_i = \{x_i, x_{i+\tau}, \ldots, x_{i+(m-1)\tau}\}, \quad (2)$$

where $m$ is the embedding dimension and $\tau$ is the time lag. Then, $X_i$ is rearranged in an increasing order:

$$x_{i+(j_1-1)\tau} \leq x_{i+(j_2-1)\tau} \leq \cdots \leq x_{i+(j_m-1)\tau}. \quad (3)$$

In situation where the equality $x_{i+(j_1-1)\tau} = x_{i+(j_2-1)\tau}$ holds, they are ordered according to the values of their corresponding $j$’s, namely, if $j_1 < j_2$, then, $x_{i+(j_1-1)\tau} \leq x_{i+(j_2-1)\tau}$. Now, each vector $X_i$ can be mapped onto a symbol sequence,

$$A(g) = [j_1, j_2, \ldots, j_m], \quad (4)$$

where $g = 1, 2, \ldots, k$, and $k \leq m!$ and $A(g)$ is one of the $m!$ permutation of $m$ distinct symbol sequences $[j_1, j_2, \ldots, j_m]$. The relative frequency of occurrence of permutation pattern $\pi_j$ is denoted as $p(\pi_j)$ and given by

$$p(\pi_j) = \frac{\#\{X_i | X_i \text{ has ordinal pattern } \pi_j\}}{N - (m - 1)\tau}, \quad (5)$$

where $\#$ denotes the number of elements in the set and $j = 1, 2, \ldots, m!$. Thus, a probability distribution $P = \{p(\pi_j), j = 1, 2, \ldots, m!\}$ is defined. Subsequently, for the embedding dimension $m$, and delay time $\tau$, the PE value of a time-series $x = \{x_i : 1 \leq i \leq N\}$ is computed as follows:

$$H(x, m, \tau) = -\sum_{j=1}^{m!} p(\pi_j) \log(p(\pi_j)), \quad (6)$$

where $H(m, \tau)$ is denoted as PE. The maximum value of $H(m)$ is $\log(m!)$ when all permutations of the time series occur with equal probability. The normalized version of $H(m)$ is given by

$$H_n(x, m, \tau) = \frac{H(x, m, \tau)}{\log(m!)}, \quad (7)$$

Hereafter, the PE refers to the normalized $H(x, m, \tau)$, namely, $H_n(x, m, \tau)$. It has been known that the length of time series, $N$, is constrained as $m! \leq N - (m - 1)\tau$ [17]. Moreover, $N > m! + (m - 1)\tau$ is used to avoid undersampling [18]. For practical purpose, $3 \leq m \leq 7$ and delay time $\tau = 1$ have been recommended [9].

### 2.3 Time-Dependent Entropy Computation

Next, the distribution of the time-varying individual oscillatory components obtained in Eq. (1), i.e., $d_k(i)$, are utilized to evaluate the adaptive subscale entropy. To deal with continuously acquired signals, EEG recording is divided into a number of segments using a sliding temporal window, leading to a time-dependent entropy measure [19]. For a given $\{x_i : 1 \leq i \leq N\}$, a sliding temporal window $w \leq N$ and a sliding interval $\Delta \leq w$ are defined. Then, the $n$th sliding window of the raw EEG signal are defined by

$$w_n(i) = \{x_i; i = 1 + n\Delta, \ldots, w + n\Delta\}, \quad (8)$$

where $n = 0, 1, \ldots, [(N - w + 1)/\Delta]$ and $[x]$ denotes the integer part of $x$.

Then, EMD is incorporated to utilize the underlying time-varying oscillatory components in EEG. Let assume EEG is decomposed by a sifting process, yielding totally $K$ IMFs and one residual which is considered as $(K + 1)$th IMF. A set of IMFs, $EMD[w_n(i)]$, is obtained from EEG in the sliding window $w_n(i)$

$$EMD[w_n(i)] = [d_1^k, d_2^k, \ldots, d^{K+1}_n], \quad (9)$$

where $d^k_n = [d^k_n; i = 1+n\Delta, \ldots, w+n\Delta]$ for $k = 1, \ldots, K+1$ denote the $k$th IMF in the $n$th sliding window.

By sliding the window $w$, a data-adaptive PE in the $k$th scale
is defined as

$$PE^k(n) = H_n(d^k_n, m, \tau),$$

(10)

where $k = 1, \ldots, K + 1$. Finally, the data-adaptive PE values in multiple scales are averaged over all scales, yielding the DAMPE

$$DAMPE(n) = \frac{1}{K+1} \sum_{k=1}^{K+1} PE^k(n).$$

(11)

3. Simulation and Experimental Studies

3.1 Simulation Study

To compare the multiscale Renyi entropy with the single scale based one, i.e., Renyi entropy of gross EEG, computer simulation was carried out. A synthesized signal consisting of Gaussian distribution and multiple sinusoidal components was used, which is shown in Figure 1(a). The sampling frequency for the synthetic signal was 256 Hz. For the first 4 s, the synthetic signal has Gaussian distribution. Following periods of the synthetic signal has different number of sinusoids in time-dependent manner as follows: From 4 s to 6 s, it consists of 4 sinusoids whose frequencies at 1, 5, 10, and 20 Hz. From 6 s to 8 s, it is composed of 2 sinusoids with 1 and 5 Hz. During last 4 s, the random permutation surrogate of the period between 4 s and 8 s was included. Figure 1(b) depicts the results of the conventional PE and the proposed DAMPE, respectively. As can be seen, two entropy measures show similar levels for Gaussian distribution. From 4 s to 8 s, the proposed DAMPE shows more distinct than the PE in accordance with the decrease of sinusoids. During last 4 s, the DAMPE increases to be comparable with the PE.

3.2 Experimental Study

EEG signals were recorded from rats during experiments in rodents subjected to controlled periods of normal circulation and asphyxial cardiac arrest with the goal of assessing brain dynamics following such an injury. Two channels of EEG using epidural screw electrodes (Plastics One, Roanoke, VA, USA) were recorded continuously in the right and left parietal areas of the rats. The experimental model of brain injury by cardiac arrest has been approved by Animal Care and Use Committee of the Johns Hopkins Medical Institutions. This rat model has been previously validated to study multiple aspects of calibrated brain injury after asphyxial cardiac arrest, including the physiologic parameters, short-term and long-term neurobehavioral outcomes, EEG recovery, and histology [20].

Nine adult male Wistar rats (300±25 g) were used. Anesthesia was induced with 4% halothane in 50% N₂:50% O₂. A 10 min of baseline trend was recorded including 5 min washout period to ensure that halothane did not influence the EEG. Subsequently, 7 min asphyxia was induced by stopping and disconnecting the ventilator and clamping the tracheal tube. The duration of cardiac arrest was determined by the mean arterial blood pressure being below 10 mmHg. Cardio pulmonary resuscitation was carried out by chest compression until return of spontaneous circulation which was decided a spontaneous the mean arterial blood pressure greater than 60 mmHg. EEG

Figure 1. Time evolution of the PE and DAMPE for synthetic signal with time varying frequency components. (a) Synthetic signal in time domain. (b) Comparison of the PE and DAMPE.
signals were recorded using two channels from the right and left parietal regions of rat’s brain using subdermal needle electrodes (Plastics One).

The signals were digitalized using CODAS, a data acquisition package (DATAQ Instruments Inc., Akron, OH, USA). A sampling rate of 250 Hz and a 12 bit resolution of A/D converter were used for digitization of the data. All rats were resuscitated and neurological outcome was evaluated by neurological deficit score (ranging from 0 = worst to 80 = best) consisting of level of arousal, cranial nerves and sensory motor assessments, reflexes, and occurrence of clinically appreciable seizures [20]. The neurological deficit score (NDS) was calculated by an independent observer 72 h after asphyxial cardiac arrest injury. Figure 2 shows the time trend of the experiment, with Phase I (the control period), Phase II (the global ischemic brain injury), and Phase III (the recovery period).

Figure 4(a)–4(b) show the time evolutions of the PE and DAMPE measures for three rats which have different NDS values, which implies distinct neurological outcome of rats. In these plots, after washout around 15 min, entropies of three rats dramatically fall to approximately zero, followed by rapid increase from 35-40 min. In Figure 4(a), the value of the PE during recovery are not clearly separable for different animals with good (79) to bad (50) NDS values. On the other hand, the values of the DAMPE in Figure 4(b) are consistently separable for the three animals with different NDSs. This result indicates that the higher neurological score, the higher entropy value at the end of the four hour recovery period.

To assess with a larger sample, the values of the PE and DAMPE of nine rats including the previous three rats were
calculated as shown in Table 1. To demonstrate the entire trend, entropies for each rat were averaged over selected intervals and the average of recovery phase (30-240 min from the start of experiment). To analyze the capability of entropies as a predictor of neuronal outcome, Pearson correlation coefficient and hypothesis testing using \( p \)-value were evaluated between NDS and entropies over the selected intervals and the whole recovery period recorded (30-240 min). From Table 1, Pearson correlation coefficients between the DAMPE and NDS were more significant than the PE over all given time slots. Additional hypothesis testings using a Student-\( t \) distribution (\( n=9 \)) reveals that the DAMPE is more closely correlated to NDS than its counterpart, the conventional PE.

4. Conclusions

This work presented a new multiscalar quantifier of brain waves, e.g., EEG signals which takes into account the dynamic changes over multiple time scales. By utilizing a data-adaptive decomposition tool, i.e., EMD, the proposed measure extracts locally changing features from fine to coarse scales of EEGs. Following the PE computation for IMFs over multiple scales, it leads a data-adaptive multiscale entropy measure for revealing dynamic changes in EEGs. Through the simulational and experimental studies, the proposed DAMPE has shown its effectiveness over the conventional PE in terms of the distinctiveness between different conditions.

Conflict of Interest

No potential conflict of interest relevant to this article was reported.

Acknowledgements

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2016R1D1A1B03934014).

References

[1] G. Buzsáki, Rhythms of the Brain, Oxford: Oxford University Press, 2006.
[2] R. Ferenets, T. Lipping, A. Anier, V. Jäntti, S. Melto, and S. Hovilehto, “Comparison of entropy and complexity measures for the assessment of depth of sedation,” IEEE Transactions on Biomedical Engineering, vol. 53, no. 6, pp. 1067–1077, 2006. https://doi.org/10.1109/tbme.2006.873543
[3] C. Stam, B. Jelles, H. Achtereekte, S. Rombouts, J. Slaets, and R. Keunen, “Investigation of EEG non-linearity in dementia and Parkinson’s disease. Electroencephalography and Clinical Neurophysiology, vol. 95, no. 5, pp. 309-317, 1995. https://doi.org/10.1016/0013-4694(95)00147-q
[4] D. E. Kim, J. H. Yu, and K. B. Shim, “EEG feature classification based on grip strength for BCI applications,” International Journal of Fuzzy Logic and Intelligent Systems, vol. 15, no. 4, pp. 277-282, 2015. https://doi.org/10.5391/ijfis.2015.15.4.277
Table 1. Statistical Results of PE and DAMPE. PE results enclosed in parentheses. (r: Correlation coefficient, p: p-value, NDS: Neurological Deficit Score)

| Animal | #  | 30-60 min | 60-120 min | 120-240 min | Ave | NDS |
|--------|----|-----------|------------|-------------|-----|-----|
| 1      | 1  | 0.40 (0.64) | 0.53 (0.66) | 0.65 (0.65) | 0.58 (0.65) | 46  |
| 2      | 2  | 0.17 (0.35) | 0.35 (0.44) | 0.55 (0.54) | 0.43 (0.48) | 50  |
| 3      | 3  | 0.39 (0.43) | 0.56 (0.54) | 0.73 (0.69) | 0.63 (0.61) | 59  |
| 4      | 4  | 0.73 (0.81) | 0.87 (0.91) | 0.84 (0.86) | 0.83 (0.87) | 74  |
| 5      | 5  | 0.48 (0.44) | 0.72 (0.67) | 0.85 (0.79) | 0.75 (0.70) | 74  |
| 6      | 6  | 0.59 (0.65) | 0.70 (0.69) | 0.82 (0.78) | 0.75 (0.74) | 74  |
| 7      | 7  | 0.63 (0.63) | 0.72 (0.85) | 0.72 (0.66) | 0.70 (0.71) | 75  |
| 8      | 8  | 0.52 (0.60) | 0.75 (0.77) | 0.83 (0.87) | 0.76 (0.79) | 78  |
| 9      | 9  | 0.50 (0.57) | 0.72 (0.75) | 0.70 (0.70) | 0.67 (0.69) | 80  |
| r      | r  | 0.72 (0.40) | 0.84 (0.70) | 0.73 (0.68) | 0.81 (0.71) |     |
| p      | p  | 0.03 (0.30) | 0.004 (0.03) | 0.02 (0.04) | 0.008 (0.03) |     |

[5] N. V. Thakor and S. Tong, “Advances in quantitative electroencephalogram analysis methods,” Annual Review of Biomedical Engineering, vol. 6, pp. 453-495, 2004. [doi.org/10.1146/annurev.bioeng.5.040202.121601]

[6] X. Yu, S. M. Park, K. E. Ko, and K. B. Sim, “Discriminative power feature selection method for motor imagery EEG classification in brain computer interface systems,” International Journal of Fuzzy Logic and Intelligent Systems, vol. 13, no. 1, pp. 11-18, 2013. [doi.org/10.5391/IJFIS.2013.13.1.11]

[7] S. M. Pincus, “Approximate entropy as a measure of system complexity,” Proceedings of the National Academy of Sciences, vol. 88, no. 6, pp. 2297-2301, 1991. [doi.org/10.1073/pnas.88.6.2297]

[8] J. S. Richman and J. R. Moorman, “Physiological time-series analysis using approximate entropy and sample entropy,” American Journal of Physiology-Heart and Circulatory Physiology, vol. 278, no. 6, pp. 2039-2049, 2000. [doi.org/10.12691/bse-2-1-3]

[9] C. Bandt and B. Pompe, “Permutation entropy: a natural complexity measure for time-series,” Physical Review Letters, vol. 88, no. 17, article ID. 174102, 2002. [doi.org/10.1103/physrevlett.88.174102]

[10] N. Nicolau and J. Georgiou, “The use of permutation entropy to characterize sleep electroencephalograms,” Clinical EEG and Neuroscience, vol. 42, no. 1, pp. 24-28, 2011. [doi.org/10.1177/155005941104200107]

[11] X. Li, G. Ouyang, and D. A. Richards, “Predictability analysis of absence seizures with permutation entropy,” Epilepsy Research, vol. 77, no. 1, pp. 70-74, 2007. [doi.org/10.1016/j.eplepsys.2007.08.002]

[12] E. Olofsen, J. Sleigh, and A. Dahan, “Permutation entropy of the electroencephalogram: a measure of anaesthetic drug effect,” British Journal of Anaesthesia, vol. 101, no. 6, pp. 810-821, 2008. [doi.org/10.1093/bja/aen290]

[13] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, and Q. Zheng, “The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis,” Proceedings: Mathematical, Physical and Engineering Sciences, vol. 454, pp. 903-995, 1998. [doi.org/10.1098/rspa.1998.0193]

[14] C. M. Sweeney-Reed and S. J. Nasuto, “A novel approach to the detection of synchronisation in EEG based on empirical mode decomposition,” Journal of Computational Neuroscience, vol. 23, no. 1, pp. 79-111, 2007. [doi.org/10.1007/s10827-007-0020-3]

[15] H. Amoud, H. Snoussi, D. Hewson, M. Doussot, and J. Duchee, “Intrinsic mode entropy for nonlinear discriminant analysis,” IEEE Signal Processing Letters, vol. 14,
[16] P. Flandrin and P. Goncalves, “Empirical mode decompositions as data-driven wavelet-like expansions,” International Journal of Wavelets Multiresolution and Information Processing, vol. 2, pp. 477-496, 2004. [https://doi.org/10.1142/s0219691304000561]

[17] J. M. Amig, S. Zambrano, and M. A. Sanjun, “True and false forbidden patterns in deterministic and random dynamics,” Europhysics Letters, vol. 79, no. 5, article ID. 50001, 2007. [https://doi.org/10.1209/0295-5075/79/50001]

[18] G. Ouyang, C. Dang, D. A. Richards, and X. Li, “Ordinal pattern based similarity analysis for EEG recordings,” Clinical Neurophysiology, vol. 121, no. 5, pp. 694-703, 2010. [https://doi.org/10.1016/j.clinph.2009.12.030]

[19] Y. S. Choi, M. Koenig, X. Jia, and N. Thakor, “Quantifying time-varying multiunit neural activity using entropy-based measures,” IEEE Transactions on Biomedical Engineering, vol. 57, no. 11, pp. 2771-2777, 2010. [https://doi.org/10.1109/tbme.2010.2049266]

[20] L. Katz, U. Ebmeyer, P. Safar, A. Radovsky, and R. Nenmar, “Outcome model of asphyxial cardiac arrest in rats,” Journal of Cerebral Blood Flow & Metabolism, vol. 15, no. 6, pp. 1032-1039, 1995. [https://doi.org/10.1097/00003246-199401000-00250]

Young-Seok Choi is an associate professor in the Department of Electronics and Communications Engineering at Kwangwoon University. He received the B.S. degree in electronic and electrical engineering from Hanyang University, Seoul, Korea, and the Ph.D. degree in electrical and computer engineering from Pohang University of Science and Technology (POSTECH), Pohang, Korea in 2000 and 2007, respectively. His current research interests include neural signal processing, brain—machine interface, neural engineering and machine learning.

E-mail: yschoi@kw.ac.kr