Ultra-fast Epileptic Seizure Detection Using EMD based on Multichannel Electroencephalogram

Wei Chen, Yan-Yu Lam, Chia-Ping Shen, *Student Member IEEE*, Hsiao-Ya Sung, Jeng-Wei Lin, Ming-Jang Chiu, and Feipei Lai, *Senior Member*, *IEEE*

Abstract **– We present a system to detect seizure and spike in Epilepsy Electroencephalogram (EEG) analysis and characterize different epilepsy EEG types. After extracting features from three EEG types, Normal, Seizure and Spike, with Empirical Mode Decomposition (EMD), we do Analysis of variance (ANOVA) to classify conspicuous features and low-resolution features, and build Gaussian distributions of conspicuous features for probability density function (PDF) to do classification. Using EMD, the recognition rate improved from 70% to 90%. With ANOVA, the recognition rate can reach 99%. The linear model accelerates the system from 2 hours to 90 seconds compare to the previous approach.**

I. Introduction

Epilepsy is a common chronic neurological disorder which afflicts approximately 1% of the population and is characterized by recurrent unprovoked seizures [1]. EEG signals play an important role in the diagnosis of epilepsy. While most traditional work try to classify EEG type based on single channel analysis, how to detect seizure in multichannel becomes highly respected recently [2]. With electrodes at various positions on the scalp of subjects, doctors can observe EEG data real-time on screen and store the instantaneous cerebral electrical activities data in computers.

Several methods have been devised for detection of seizure and spike based on EEG data, including determination of wavelet coefficients [3], eigenvectors [4], time-frequency analysis [5] and principal component analysis [6] in algorithms such as expert systems [7], template methods [8], artificial neural networks [9], wavelet analysis [10], support vector machines [11], Kalman filters [12], independent component analysis [13], and fuzzy c-mean [14]. Our team had proposed some outcome with wavelet transform and approximate entropy using support vector machine [15] [16].

In this research, we present a method to select features. By defining each channel in an EEG signal to a variable and calculate the statistic characteristics of the variables while doing Empirical Mode Decomposition (EMD) [17], integrating all features of one segment to a sample vector, and combining all sample vectors of an EEG type to a sample matrix, we will be able to do Analysis of Variance (ANOVA) to each extracted feature. After selecting the informative feature, we construct Gaussian distributions in hyper-plane for the three EEG types. We then calculate the probability density function (PDF) from one segment to all Gaussian distributions to distinguish the segment to a specific EEG type.

W. Chen, Y. Y. Lam, C. P. Shen, H.Y. Sung and F. P. Lai are with the Graduate Institute of Biomedical Electronics and Bioinformatics, National Taiwan University, Taipei, Taiwan.

J. W. Lin is with the Department of Information Management, Tunghai University, No.181, Sec. 3, Taichung Port Rd., Xitun Dist., Taichung City 40704, Taiwan, R. O. C.

M. J. Chiu is with the Department of Neurology, National Taiwan University Hospital, Taipei, Taiwan.

Correspondence to Wei Chen, 9F., No.187, Sec. 3, Dadu Rd., Taipei, Taiwan, 112; TEL: 886-988153817; E-mail: jpchen9@yahoo.com.tw;

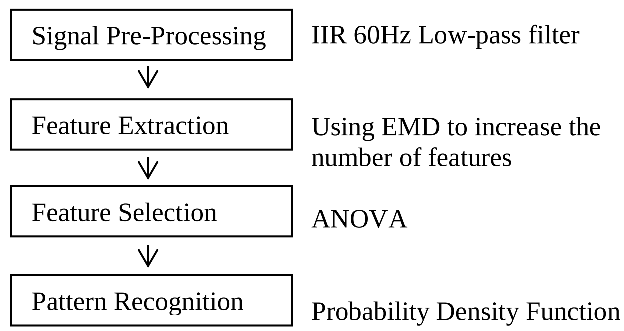

Fig. 1. Classification Process.

II. Material

A. Subjects

Our research uses clinical data from National Taiwan University Hospital (NTUH). The sampling frequency is 200Hz. The EEG data contain 16 channels with the electrodes located according to the ISO 10-20 system [18].

B. Data Recording

All patients were lying down in bed, taking an 8 hours long record including awake and asleep. The proposed work retrieves 600 seconds for each EEG type, normal, seizure, and spike (burst-suppression), from 9 patients.

C. Coding Tool and Environments

In this project, we use MATLAB 2012a as the coding tool. Our operating system is Windows 7. The computer is ASUS, with Intel® Core™ i7-2600 CPU @3.4Gh, 8GB RAM, 64bits operating system.

III. Method

The proposed method used the standard classification process with signal pre-processing, feature extraction, feature selection and pattern recognition as in Figure 1.

A. Signal Pre-Processing

We use IIR as our pre-processing method. Since most EEG artifacts caused by electrical power lines are higher than 60Hz, we can simply filter out all signals exceeding 60Hz [19].

978-1-4799-3163-7/13/$31.00 ©2013 IEEE

B. *Empirical Mode Decomposition*

Empirical Mode Decomposition (EMD), based on Hilbert Huang Transform (HHT), is an algorithm for non-linear and non-stationary signal processing [17].

EMD separates a signal into several Intrinsic Mode Functions (IMFs). The proposed method separates the signal into 4 IMFs before we start extracting the features, increase the input data amount to 5 times the original data size including the original data. The process is shown in Figure 3.

Since IMFs imply bio-continuity of a signal, we expect the extracted features directly link to Epileptic seizure or spike.

C. *Feature Extraction*

For every 400 time points, we slice them into a segment based on the commands from doctors, and construct a 2D raw data matrix ***X*** (1), Figure 2.

$$X = \{X_{i,j}\};\ i_{max} = 400,\ j_{max} = 16, \tag{1}$$

where *i* = time point, *j* = channel number.

We have 300 raw data matrixes for each EEG type. We calculate the characteristics for each segment and combine all segments in one type as one 2D feature matrix ***Y*** (2).

$$Y = \{Y_{m,n}\};\ m_{max} = 100,\ n_{max} = N, \tag{2}$$

where *m* = segment number, *n* = the number of statistic characteristics; ***N*** is the number of statistic feature extracted from ***X***. Note that in this experiment we use 100 segments of each EEG type as training data, the other 600 segments as testing data.

We test several different feature sets in this presented work.

1. Mean of each channel (3), standard deviation of each channel (4), covariance (5) and correlation coefficient between all channels (6). The number of features is 291; ***N*** = 291.

2. Mean of all channels, mean of standard deviation of all channels, mean of covariance among all channels, and sum of all correlation coefficients without self-correlated. The number of features is 4; ***N*** = 4.

3. Mean of all channels, Max, median, and min of standard deviation of each channel, mean of covariance among all channels, and sum of all correlation coefficients without self-correlated, median and min of correlation coefficient. The number of features is 8; ***N*** = 8.

4. Mean and standard deviation of each channel and their 4 stage IMF derived by EMD; Covariance and correlation coefficient among all channels. The number of features is 672; ***N*** = 672.

5. Mean of all channels, Max, median, and min of standard deviation of each channel and their 4 stage IMF derived by EMD. Mean of covariance among all channels, and sum of all correlation coefficients without self-correlated, median and min of correlation coefficient. The number of features is 24; ***N*** = 24.

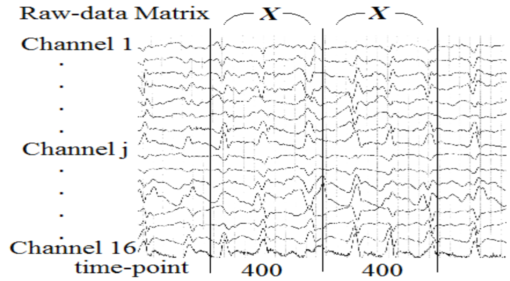

Fig. 2. ***X*** raw data matrix.

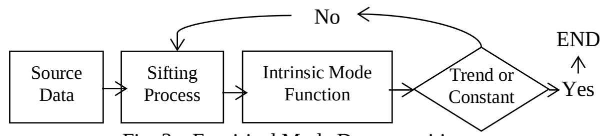

Fig. 3a. Empirical Mode Decomposition

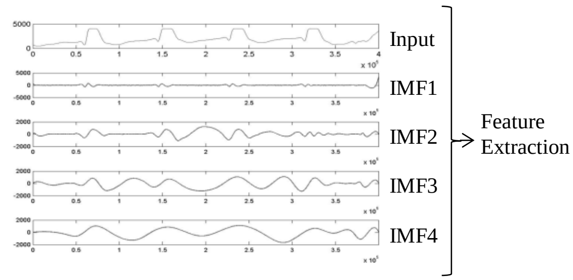

Fig. 3b. Input and IMF1 to IMF4.

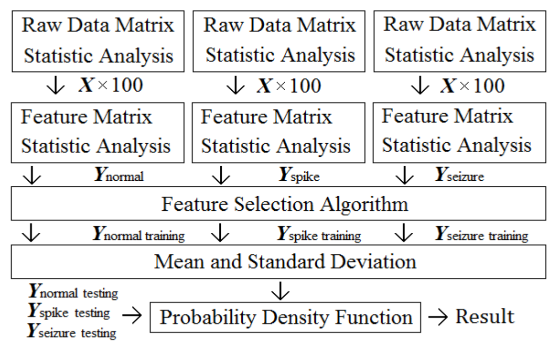

Fig. 4. Block-diagram.

6. 24 features selected by the proposed feature selection algorithm 2 from feature set 1.

7. 24 features selected by the proposed feature selection algorithm 2 from feature set 4.

$$\bar{X}_j = \frac{1}{400}\sum_{i=1}^{400} X_{i,j} \tag{3}$$

$$\sigma_j = \sqrt{\frac{1}{400}\sum_{i=1}^{400}(X_{i,j} - \bar{X}_j)^2} \tag{4}$$

$$\sigma_{jk} = \frac{1}{400}\sum_{i=1}^{400}(X_{i,j} - \bar{X}_j)(X_{i,k} - \bar{X}_k) \tag{5}$$

$$r_{jk} = \frac{\sigma_{jk}}{\sqrt{\sigma_{jj}}\sqrt{\sigma_{kk}}} \tag{6}$$

D. *Feature Selection*

In this experiment, we use ANOVA (8) to do feature selection. Since ANOVA will detect whether a feature is different in different EEG types or not, we use ANOVA to check the ***F*** value for all features and find the features that can tell the difference.

$$\text{Null Hypothesis: } H_0: \mu_1 = \mu_2 \quad (7)$$

$$\boldsymbol{F} = \frac{SS(between)/de(between)}{SS(within)/df(within)} \quad (8)$$

SS(between) = $N \sum_g (\bar{x}_g - \bar{x})^2$

SS(within) = $\sum_m \sum_g (\bar{x}_{mg} - \bar{x}_g)^2$

df(between) = g-1

df(within) = g(m-1), where g is the number of EEG types.

Feature Selection Algorithm: Do ANOVA to two of the EEG types at a time and add the significant features together after doing one after another. The hypothesis is (9).

$$H_0: \mu_1 = \mu_2 \,||\, \mu_1 = \mu_3 \,||\, \mu_2 = \mu_3 \quad (9)$$

Every feature that is good at finding the difference between two EEG types will be chosen. The weightings to separate every two EEG types are equal. For repeated features, choose the top k high $\boldsymbol{F}$ value features in the lowest $\boldsymbol{F}$ value group, where k is the number of repeated features.

E. Pattern Recognition

We calculate the mean $\bar{Y}_n$ and the standard deviation σ_n for $\boldsymbol{Y}$ of each EEG type. Then, we calculate the probability density function (10) for each EEG type.

$$f_{\mu,\sigma^2} = \prod_{m=1}^{100} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{Y_{m,n}-\bar{Y}_n}{\sigma_n})^2} \quad (10)$$

To reduce the complexity of calculation, we apply a log to the PDF and change it to the form (10) and define a new $f'_{\mu,\sigma 2}$.

$$f'_{\mu,\sigma^2} = \sum_{m=1}^{100} \ln \frac{1}{\sigma\sqrt{2\pi}} - \frac{1}{2}(\frac{Y_{m,n}-\bar{Y}_n}{\sigma_n})^2 \quad (11)$$

Since the algorithm compare the probability among each Gaussian distribution of all training EEG record types, and do not need the exact value, we use (11) to do the comparison.

When a new training data join in, the algorithm allows us to re-train the model without running the whole data again. The algorithm is more feasible than SVM in clinical uses.

IV. RESULT

In epilepsy data processing, spike and normal are unlikely distinguishable and the tradeoff between both recognition rates is very important.

Tables I-IV show the results generated by the proposed system, with probability density function.

In the experiment, we can see that EMD does help the system gain more information. From **Table I** to **Table II**, the average accuracy increases from about 70% to about 90%.

TABLE I. FEATURE SETS WITHOUT EMD

Classifying Accuracy	*Feature Set*		
	feature set 1	*features set 2*	*features set 3*
Training data self-predict 300 segments	Normal: 78% Spike: 93% Seizure: 100% Total: 90.3%	Normal: 74% Spike: 85% Seizure: 100% Total: 86.3%	Normal: 85% Spike: 95% Seizure: 100% Total: 93.3%
Testing data prediction 600 segments	Normal: 34.5% Spike: 88% Seizure: 100% Total: 74.2%	Normal: 37.5% Spike: 71% Seizure: 100% Total: 69.5%	Normal: 61.5% Spike: 74% Seizure: 100% Total: 78.5%

TABLE II. FEATURE SETS WITH EMD, WITHOUT FEATURE SELECTION ALGORITHM

Classifying Accuracy	*Feature Set*	
	features set 4	*feature set 5*
Training data self-predict 300 segments	Normal: 90% Spike: 98% Seizure: 100% Total: 96%	Normal: 98% Spike: 100% Seizure: 100% Total: 99.3%
Testing data prediction 600 segments	Normal: 87% Spike: 97% Seizure: 100% Total: 94.7%	Normal: 94.5% Spike: 90% Seizure: 100% Total: 94.8%

a. Feature set 5 is selected based on prior knowledge to the characteristic of each EEG type.

TABLE III. FEATURE SETS WITH FEATURE SELECTION ALGORITHM

Classifying Accuracy	*Feature Set*	
	feature set 6	*feature set 7*
Training data self-predict 300 segments	Normal: 77% Spike: 98% Seizure: 100% Total: 91.7%	Normal: 98% Spike: 100% Seizure: 100% Total: 99.3%
Testing data prediction 600 segments	Normal: 78.5% Spike: 87% Seizure: 100% Total: 88.5%	Normal: 99% Spike: 100% Seizure: 100% Total: 99.7%

Comparing **Table I**, **II** and **III**, the results are better with the proposed feature selection algorithm. Compared to the result of feature set 1, feature set 6 increases its accuracy from 74.2% to 88.5%; Compared to the result of feature set 4, feature set 7 increases its accuracy from 94.7% to 99.7%.

TABLE IV. COMPARE TO OTHER RESEARCHES

Author	*Performance*[a]		
	Sensitivity[b]	*Specificity*[c]	*Precision*[d]
Leistritz et. al. [20]	60%	84%	67%
Sarkela et. al. and Palmu. et. al. [21], [22]	29%	92%	87%
Bhattacharyya et. al. [23], [24]	81%	89%	80%
Shen et al. [16] (same samples in this project)	99.5%	99.8%	99.5%
Proposed Method	100%	99.5%	99.5%

a. Table V only considers spike detection accuracy. b. Sensitivity = TP / (TP + FN).

c. Specificity = TN / (TN + FP). d. Precision = TP / (TP + FP).

True positive (TP) means a spike segment is correctly detected. True negative (TN) means a non-spike segment is correctly detected. False positive (FP) means a non-spike segment is incorrectly detected as spike segment. False negative (FN) means a spike segment is incorrectly detected.

TABLE V. TRAINING AND PREDICTION SPEED COMPARISON

Author	*Time Consumed (100 time tests)*	
	Training	*Prediction*
Shen et al. [16] (same samples in this project)	116$\pm$4 minutes (~6960 seconds)	4$\pm$0.2 seconds
Proposed Method	90$\pm$3 seconds	4$\pm$0.3 seconds

V. EVALUATION

To find more information, we use EMD to extract more informative features. In the comparison of **Tables I** and **II**, we find that EMD is a good feature extraction support mathematic tool. More features with more information that is linked to Epileptic were found with the process of EMD. The result in **Table II** shows a lot better than in **Table I.**

More features do not mean better recognizing rate. In fact, too much useless features make the recognition rate lower instead. As far as concerned, useless features often perform as noise in the classification.

The features selected to distinguish normal EEG data from spike EEG data, to distinguish normal EEG data from seizure EEG data, and to distinguish spike EEG data from seizure EEG data are not the same.

We characterize the properties of each EEG type:

EEG Normal: Signals of this type do not change very often. Their standard deviations in all dimensions are small. They have small covariance and correlation coefficients between each channel. The standard deviations in **IMF1** in EEG normal are smaller than the value in EEG spike.

EEG Spike: Standard deviations in all dimensions are higher than EEG normal. They have smaller correlation coefficients than EEG normal. The standard deviations in **IMF1** in EEG spike are higher than the value in EEG normal.

EEG Seizure: Signals of this type change a lot all the time correspondently. The standard deviations of this type are big compared to spike and normal. Their covariance is big. The best features to distinguish seizure from normal and spike are those standard deviations of raw data.

For **Table IV**, the performance of the proposed method is better than previous work especially in sensitivity. EMD is more sensitive to spike event.

Table V shows the efficiency of linear training algorithms. Fast and adaptive training is more feasible in practical uses.

VI. Conclusion

Currently, EEG interpretation can only be analyzed by neurologists and epileptologists, which often takes a very long time to go through all of the data. This research proposes an ultra-fast classifying system for epilepsy EEG types by using probability density function with EMD and ANOVA feature selection that can help clinical practice.

References

[1] Witte H, Iasemidis L D, and Litt B, “Special issue on epileptic seizure prediction” IEEE Trans. Biomed. Eng. 50 pp. 537–539, 2003.

[2] Xia Y S and Leung H, “Nonlinear spatial-temporal prediction based on optimal fusion” IEEE Trans. Neural Network 17 pp. 975–988, 2006.

[3] Güler I and Ü beyli E D, “Adaptive neuro-fuzzy inference system for classification of EEG signals using wavelet coefficients” J. Neurosci. Methods 148 pp. 113–121, 2005.

[4] Ü beyli E D and Güler I, “Features extracted by eigenvector methods for detecting variability of EEG signals” Pattern Recognit. Lett. 28 pp. 592–603, 2007.

[5] Tzallas A T, Tsipouras M G and Fotiadis D I, “Epileptic seizure detection in EEGs using time frequency analysis” IEEE Trans. Inf. Technol. in Biomed. 13 pp. 703–710, 2009.

[6] Ghosh-Dastidar S and Adeli H, “Principle component analysis-enhanced cosine radial basis function neural network for robust epilepsy and seizure detection” IEEE Trans. Biomed. Eng. 55 pp. 512–518, 2008.

[7] Davey B L, Fright W R, Carroll G J, and Jones R D, “Expert system approach to detection of epileptiform activity in the EEG” Med. Biol. Eng. Comput. 27 pp. 365–370, 1989.

[8] Nonclercq A, Foulon M, Verheulpen D, Cock C D, Buzatu M, Mathys P, and Bogaert P V, “Spike detection algorithm automatically adapted to individual patients applied to spike-and-wave percentage quantification” Neurophysiol. Clin. 39 pp. 123–131, 2009.

[9] Ko C W and Chung H W, “Automatic spike detection via an artificial neural network using raw EEG data: Effects of data preparation and implications in the limitations of online recognition Clin”. Neurophysiol. 111 pp. 477–481, 2000.

[10] Indiradevi K P, Elias E, Sathidevi P S, Nayak S D, and Radhakrishnan K, “A multi-level wavelet approach for automatic detection of epileptic spikes in the electroencephalogram” Comput. Biol. Med. 38 pp. 805–816, 2008.

[11] Acir N and Guzelis C, “Automatic spike detection in EEG by a twostage procedure based on support vector machines” Comput. Biol. Med. 34 pp. 561–575, 2004.

[12] Oikonomou V P, Tzallas A T, and Fotiadis D I, “A Kalman-filter-based methodology for EEG spike enhancement” Comput. Methods Programs Biomed. 85 pp. 101–108, 2007.

[13] Lucia M D, Fritschy J, Dayan P, and Holder D, “A novel method for automated classification of epileptiform activity in the human electroencephalogram-based on independent component analysis” Med. Biol. Eng. Comput. 46 pp. 263–272, 2008.

[14] Inan Z H and Kuntalp M, “A study on fuzzy c-means clustering-based systems in automatic spike detection” Comput. Biol. Med. 37 pp. 1160–1166, 2007.

[15] C.P. Shen, C.C. Chen, S.L. Hsieh, W.H. Chen, J.M. Chan, C.M. Chen, F. Lai and M.J. Chiu, “High-Performance Seizure Detection System Using a Wavelet-Approximate Entropy-fSVM Cascade With Clinical Validation”, Clinical EEG and Neuroscience. DOI: 10.1177/15500594 13483451, 2013.

[16] C.P. Shen, S.T. Liu, W.Z. Zhou, F.S. Lin, Y.Y. Lam, H.Y. Sung, W. Chen, J.W. Lin, M.J. Chiu, M.K. Pan, J.H. Kao, J.M. Wu, F. Lai, “A Physiology-Based Seizure Detection System for Multichannel EEG”, PLoS ONE 8(6): e65862. doi:10.1371/journal.pone.0065862. 2013.

[17] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N. C. Yen, C. C. Tung and H. H. Liu, “The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Non-stationary Time Series Analysis”, Proc. R. Soc. Lond. A, vol. 454, pp. 903- 995, 1998.

[18] Niedermeyer E. and da Silva F.L. “Electroencephalography: Basic Principles, Clinical Applications, and Related Fields.” Lippincot Williams & Wilkins. ISBN 0-7817-5126-8. 2004.

[19] Towle VL, Bolaños J, Suarez D, Tan K, Grzeszczuk R, Levin DN, Cakmur R, Frank SA, Spire JP. "The spatial location of EEG electrodes: locating the best-fitting sphere relative to cortical anatomy". Electroencephalogr Clin Neurophysiol 86 (1): 1–6. doi : 10.1016 / 0013 - 4694 (93) 90061-Y. 1993.

[20] M. Särkelä, S. Mustola, T. Seppänen, M. Koskinen, P. Lepola, K. Suominen, T. Juvonen, H. Tolvanen-Laakso, and V. Jäntti, “Automatic analysis and monitoring of burst suppression in anesthesia,” J. Clin. Monit. Comput., vol. 17, no. 2, pp. 125–34, 2002.

[21] S. Bhattacharyya, J. Mukhopadhyay, A. Majumdar, B. Majumdar, A. Singh, and C. Saha, “Automated burst detection in neonatal EEG,” in Int. Conf. Bio-Inspired Syst. Sig. Process., 2011, pp. 15–21.

[22] L. Leistritz, H. Jäger, C. Schelenz, H. Witte, P. Putsche, M. Specht, and K. Reinhart, “New approaches for the detection and analysis of electroencephalographic burst-suppression patterns in patients under sedation,” J. Clin Monit., vol. 15, no. 6, pp. 357–367, 1999.

[23] B. R. Greene, S. Faul, W. P. Marnane, G. Lightbody, I. Korotchikova, and G. B. Boylan, “A comparison of quantitative EEG features for neonatal seizure detection,” J. Clin. Neurophysiol., vol. 119, no. 6, pp. 1248–1261, 2008.

[24] S. Bhattacharyya, A. Biswas, J. Mukherjee, A. K. Majumdar, B. Majumdar, S. Mukherjee, and A. K. Singh, “Feature Selection for Automatic Burst Detection in Neonatal Electroencephalogram”, IEEE Journal on Emerging and selected topics in Circuits And Systems, vol. 1, no. 4, pp. 469-479, December 2011.