Seizure Detection using Wavelet Decomposition of the Prediction Error Signal from a Single Channel of Intra-Cranial EEG

Zisheng Zhang, *Student Member, IEEE*, and Keshab K. Parhi, *Fellow, IEEE*

***Abstract*— This paper presents a novel patient-specific algorithm for detection of seizures in epileptic patients from a single-channel intra-cranial electroencephaolograph (iEEG) recording. Instead of extracting features from the EEG signal, first the EEG signal is filtered by a prediction error filter (PEF) to compute a prediction error signal. A two-level wavelet decomposition of the prediction error signal leads to two detail signals and one approximate signal. Eight features are extracted every one second using a 2-second window with a 50% overlap. These features are input to two different types of classifiers: a linear support vector machine (SVM) classifier and an AdaBoost classifier. The algorithm is tested using the intra-cranial EEG (iEEG) from the Freiburg database. It is shown that the proposed algorithm can achieve a sensitivity of 95.0% and an average false positive rate (FPR) of 0.124 per hour, using the linear SVM classifier. The AdaBoost classifier achieves a sensitivity of 98.75% and an average FPR of 0.075 per hour. These results are obtained with leave-one-out cross-validation. In addition, for 13 out of 18 patients, the AdaBoost classifier requires only one feature, while it requires 4 features for the remaining 5 patients.**

I. Introduction

Approximately 0.7% of the world's population suffers from epileptic seizures. About 50 million people worldwide have epilepsy. Epilepsy is the second most common neurological disorder [1]. Reliable seizure detection, which refers to detecting epileptic seizures based on continuous electroencephalogram (EEG) recordings of epileptic patients, is important for not only improving the lives of epileptic patients, but also in assisting the epileptologists in marking seizures in the (EEG) recordings. A device that can detect seizures can be used in a closed-loop therapy system to deliver an anti-epilepsy drug (AED) or stimulate the brain as needed.

Seizure detection can be viewed as a binary classification problem where one class consists of ictal signals corresponding to an occurrence of the seizure, and the other class consists of normal EEG signals, also referred as interictal signals. It is known that the patterns do vary during ictal and inter-ictal periods in most of the cases [2]. Significant amount of research in seizure detection has been directed towards identifying these discriminating patterns or features [3]–[5].

Power spectral density (PSD) is the most commonly used feature for seizure detection. However, the main drawback of this approach is the high false positive rate (FPR) as PSD increases often in the inter-ictal periods as well. The EEG recordings are highly non-stationary as ictal and interictal patterns vary substantially over different patients. Even for a single patient, ictal and interictal patterns may vary substantially from seizure to seizure and from hour to hour. Most existing seizure detection algorithms suffer from several drawbacks. For example, some algorithms are designed without cross-validation, i.e., the same seizure is used for both training and test. Such algorithms are "overtrained", and may not be able to detect future seizures. Other algorithms are validated using few patients, and are not tested on large datasets containing many patients. In addition, algorithms that work well in shorter recordings fail to work in longer recordings.

Our main objective is to develop an automated algorithm that can reliably detect seizures. The algorithm should also have a low hardware complexity. In the proposed approach, *only a single channel EEG signal* is analyzed for seizure detection. We first filter the EEG signal by a prediction error filter, also known as a whitening filter, to compute an error signal. A 19th-order prediction error filter (PEF) computes the error signal as the difference between the current input sample and the estimate of it. A window based processing is used with a 2-second sliding window with half overlap. The predictor coefficients are recomputed every one second. A two-level wavelet decomposition of the error signal computes the approximate signal and two detail signals. The total energies in a window of the error signal and the three signals from the wavelet decomposition are extracted in two different ways. The features are input to two types of classifiers: a linear support vector machine (SVM) classifier and an AdaBoost classifier. The performance of each classifier is evaluated and compared against the other.

II. Materials and Methods

A. Patient Database

We have trained and tested our algorithm on the Freiburg EEG database [6], which is available to any lab by request. According to [6], this database contains electrocorticogram (ECoG) or iEEG from 21 patients with medically intractable focal epilepsy. We have chosen 18 of the available datasets of 21 patients, who have three or more seizures (the minimum number for cross-validation). Each 2-s-long window of iEEG has been categorized as ictal (containing a seizure), interictal (at least 1 h preceding or postceding a seizure), preictal (in 60 min preceding a seizure onset), or artifact. Half an hour of iEEG recordings preceding preictal and an hour of those postceding seizure offset are excluded from training. The

*This work was carried out at Leanics Corporation.

Z. Zhang and K.K. Parhi are with the Department of Electrical and Computer Engineering, University of Minnesota, MN 55455, USA zhan1116 at umn.edu; parhi at umn.edu

Freiburg database contains six of iEEG recordings from grid, strip, or depth-electrodes, three near the seizure focus (focal) and the other three distal to the focus (afocal). Seizure onset times and artifacts were identified by certified epileptologists. The data were collected at 256 Hz (Patient 12 at 512 Hz) sampling rate with 16 bit analog-to-digital converters.

B. System Architecture

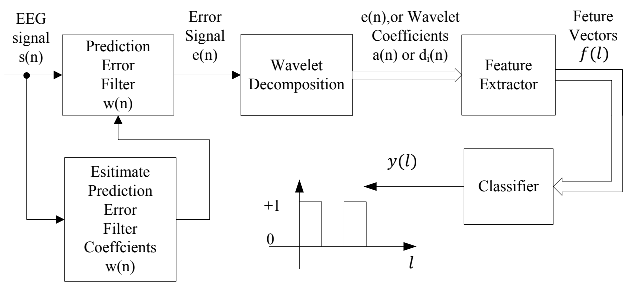

Fig. 1. System architecture for seizure detection

Fig. 1 shows the overall system for seizure detection. Let $s(n)$ denote the single-channel iEEG signal. First the signal $s(n)$ is windowed and filtered by a prediction error filter to compute the error signal $e(n)$. A two-level wavelet decomposition is applied to the error signal to obtain one approximate signal and two detail signals. An 8-dimensional feature vector $f(l) = [f_1(l), f_2(l), ..., f_8(l)]^T$ is extracted by computing the total power for the error signal and the three signals obtained by wavelet decomposition. The feature vectors are then subjected to training and classification. The output of the system $y(l)$ represents the detection signal. Two types of classifiers are considered. These include: the linear SVM and the AdaBoost. The training follows leave-one-out procedure, where the seizure to be tested is not used for training.

C. Feature Extraction

This section describes the method for feature extraction, which includes prediction error filter, a 2-level wavelet decomposition and power computation.

1) Window-based signal processing: In window-based signal processing, the input signal is divided into the input segments (or windows) and the signal is processed segment by segment. Let M denote the length of each segment and L denote the total number windows. Let

$$s_l(n) = s(n + (l-1)M/2)$$
$$n = 0, \dots, M-1, l = 0, \dots, L$$

denote the window signal in the l-th segment. Each segment has a 50% overlap with its neighbour segment.

2) Preprocessing: In the first step, EEG data is preprocessed to remove its mean. The demeaned signal is then filtered by a PEF to remove the predictable component of the EEG signal. Each window is 2 seconds long and has 50% overlap. The PEF is then used to compute the error signal for next one second. Thus, effective feature computation rate is one per second.

Let w_f represent tap-weights vector of an m-tap predictor (or a mth-order PEF). Coefficients of the PEF can be computed by solving the Wiener-Hopf equation: $w_f = R^{-1}r$, where R represents the autocorrelation matrix of the input sample vector of a window, and r represents the cross-correlation vector between the input sample vector and its delayed versions. Levinson-Durbin algorithm is used to solve the above equation [7].

A 19th-order PEF is chosen for this dataset. A singular value decomposition (SVD) of the covariance matrix is performed for patient No. 1 to find the optimal order of the predictor. Fig. 2(a) and Fig. 2(b) show the plots of the percentage of total energy captured by the predictor versus the predictor's order using (a) an hour's inter-ictal data from patient No. 1 while the patient is awake and (b) an hour's inter-ictal data from patient No. 1 while the patient is sleeping, respectively. A 19-tap predictor (equivalently, 19th order or 20-tap PEF) can capture about 95% of the total energy of the signal.

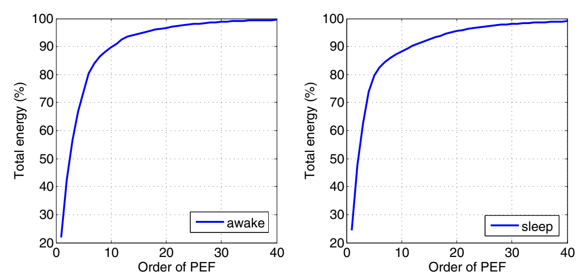

Fig. 2. Percentage of total energy captured by the predictor versus the predictor's order using (a) an hour's inter-ictal data from patient No. 1 while the patient is awake and (b) an hour's inter-ictal data from patient No. 1 while the patient is sleeping.

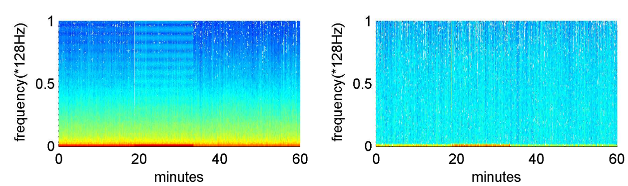

Fig. 3. Spectrograms of the EEG signal (left) and its error signal (right) using interictal recordings for the 16th hour from patient No. 1.

Fig. 3 shows the spectrograms of the EEG signal and its error signal corresponding to the interictal recordings for patient No. 1 in the 16th hour, where undesired harmonics in the interictal period are filtered and the dominance of the low frequencies on the total power is eliminated after prediction error filtering.

3) Discrete wavelet decomposition: A two-level wavelet decomposition is applied to the error signal to compute wavelet coefficients at different levels. The purpose of wavelet decomposition is to decompose the original signal into three disjoint sub-bands [8].

Discrete wavelet transform (DWT) decomposes discrete sequences into discrete wavelet coefficients. The structure of a 2-level wavelet decomposition tree is shown in Fig. 4. The input signal is first passed through a low-pass (LPF) and a high-pass (HPF) filter. Then each filter is followed by a down-sampler with factor of 2. At the next level, the approximation coefficients are further decomposed into approximate and detail coefficients.

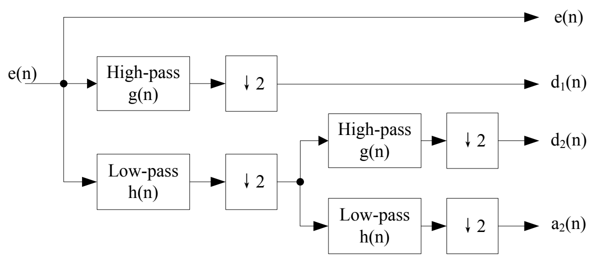

Fig. 4. Structure of a 2-level wavelet decomposition

4) Feature extractor: Two types of features are extracted from the error signal and the wavelet coefficients: one is the total power and the other is the sum of the logarithm of the absolute feature values (also equivalently, logarithm of the product of the absolute feature values). Total power for each segment is obtained by computing the sum of the squared value of the wavelet coefficients (or the error signal). Mathematically, these are computed as:

$$f'(l) = \sum_{n\in I_l} log|e(n)|$$
$$f''(l) = \sum_{n\in I_l} e^2(n)$$

where $I_l = \{(l-1)f_s+1, ..., lf_s\}$ represents the samples of the l-th window. Fig. 5 shows the block diagram of feature extraction, where a total of 8 features ($f_1(l)$ to $f_8(l)$) are extracted from the error signal, $e(n)$, and the wavelet coefficients, $a_2(n)$, $d_2(n)$, and $d_1(n)$; four of these features represent the mean power and the remaining four represent the logarithm of the product of the absolute values. For the AdaBoost classifier, all 8 features are input to the classifier. The classifier always selects between 1 to 4 out of the 8 features.

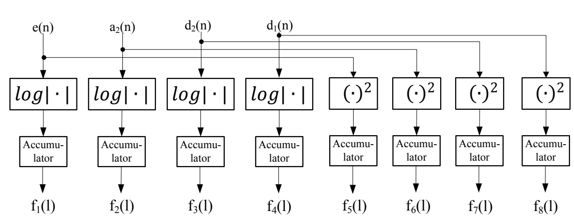

Fig. 5. Feature extraction.

D. Seizure Detection Classification

Two classification methods are used in this paper and their performances are compared. One is classification using a linear Support Vector Machine (SVM) and the other by using AdaBoost.

1) SVM and classification: Detailed descriptions of SVM can be found in [9]. Generally speaking, the SVM seeks to find a hyperplane that achieves the maximum margin between the feature samples correctly classified [10].

The penalty parameter C is usually determined by the cross-validation step [10]. Leave-one-out cross-validation s-trategy, which refers to leaving feature vectors corresponding to a randomly selected seizure out of the training set, is widely used to avoid *overfitting* of the model. After the test data are classified, the hyperplane decision function is smoothed by a moving-average filter in a postprocessing step in the proposed algorithm.

2) AdaBoost: Boosting, formulated by Yoav Freund and Robert Schapire, has been very successful in feature classification [11] due to its adaptivity and strong resistance to overfitting. In prior work, we have applied AdaBoost for predicting seizures in [12]. Furthermore, the computational complexity of the AdaBoost is independent of the dimension of the features; this leads to a classifier with low computational complexity even when large number of features are used. The weak classifiers can be chosen from any commonly used classifier. In our algorithm, the base classifier is defined as a decision stump.

III. EXPERIMENTAL RESULTS

TABLE I
DETECTION PERFORMANCE OF THE SYSTEM USING LINEAR SVM

Patient No.	electrode No.	Total No. of SZ	Sensi-tivity	No. of FP	FP rate
1	1	4	100	1	0.042
3	1	5	100	4	0.167
4	1	5	100	0	0
5	1	5	100	14	0.583
6	2	3	100	13	0.542
7	1	3	100	0	0
9	1	5	100	3	0.125
10	2	5	100	0	0
11	1	4	100	1	0.042
12	1	4	100	0	0
14	1	4	100	0	0
15	1	4	75	0	0
16	3	5	80	8	0.333
17	1	5	100	0	0
18	1	5	100	5	0.208
19	1	4	50	2	0.083
20	1	5	100	1	0.042
21	1	5	100	0	0
Overall		80	95	53	0.124

The parameters for the system are described as follows:

1) For each patient, we apply our algorithms on all electrodes. We select the electrode with best performance.

2) Leave-one-out cross validation is used where one seizure is left out for testing and the classifier is trained using features corresponding to the remaining seizures that constitute the training set. This is repeated with each seizure left out once for testing. The classifier with the best performance over the entire data is selected.

3) A refractory period, which specifies a time period during which the system ignores all the subsequent triggers

TABLE II

DETECTION PERFORMANCE OF THE SYSTEM USING ADABOOST

Patient No.	electrode No.	Total No. of SZ	Sensi-tivity	No. of FP	FP rate
1	1	4	100	0	0
3	1	5	100	0	0
4	2	5	100	0	0
5	2	5	100	5	0.208
6	2	3	100	7	0.292
7	1	3	100	0	0
9	1	5	100	3	0.125
10	2	5	100	0	0
11	1	4	100	0	0
12	1	4	100	0	0
14	3	4	100	1	0.042
15	3	4	75	0	0
16	2	5	100	8	0.333
17	1	5	100	0	0
18	1	5	100	7	0.292
19	1	4	100	1	0.042
20	5	5	100	0	0
21	3	5	100	0	0
Overall		80	98.75	32	0.075

$*$ Features for patient No. 19 are computed as the time difference of the original features.

once it's triggered, is introduced. The refractory period is set to be 10 minutes.

Test Results using linear SVM classifier are shown in Table I. Only the first 4 features $\{f_1(n), .., f_4(n)\}$ are used in the training phase. The average sensitivity is 95% and the average FP rate is 0.124 FP per hour.

Test Results using AdaBoost and all 8 features are shown in Table II. The performance is improved as the sensitivity is increased to 98.75% and the FP rate reduces to 0.075 FP per hour. For patient No. 19, in order to detect all seizures, a new feature was derived by taking the difference of the log features at certain time and at 30s prior to that time point.

IV. DISCUSSION

Many approaches have been presented for detecting seizures in epileptic patients. A seizure detection algorithm that utilizes 3 focus channels was proposed in [4]. In [13], this proposed algorithm was tested on the Freiburg database [6] and achieved a high sensitivity of 96.4% and a false positive rate (FPR) of 0.20 per hour. Another detection algorithm which utilizes 4 bipolar channels and extracts four different types of features was proposed in [14]. This algorithm was tested on the Freiburg database and achieved a high sensitivity of 98.7% and a FPR of 0.27 per hour. Another detection algorithm which uses a single channel signal and 5-level wavelet decomposition was proposed in [15]. This algorithm was also tested on the Freiburg database and achieved a sensitivity of 91.29%. Many other detection algorithms have also been proposed and tested on different databases. A wavelet based automated seizure detection algorithm with four-level wavelet coefficients was proposed in [4] and achieved a sensitivity of 94.2% and a false detection rate of 0.25 per hour. Another algorithm, proposed in [16], achieves a 100% sensitivity and a FP rate of 0.37 per hour. It should also be noted that this algorithm was trained using only the first recorded seizure in each patient and, therefore, has its own limitations.

Table III compares the system performance of the proposed algorithm with prior works. The proposed algorithm for seizure detection has the highest sensitivity (except for the results in [16]) and a significantly lower FP rate than all other prior works when AdaBoost classifier is used. Furthermore, the proposed algorithm uses the least number of features and electrodes. Future work will be directed towards applicability of the proposed method for scalp EEG recordings and long-term recordings.

TABLE III

COMPARISON TO PRIOR WORK

Reference	Sensi-tivity	FPR	No. of electrodes	No. of features
[15]	91.3	-	1	24
[14]	98.7	0.27	4	16
[13]	96.4	0.20	3	24
[4]	94.2	0.25	21	84
[16]	100	0.37	-	6/channel
proposed (SVM)	95.0	0.12	1	4
proposed (AdaBoost)	98.75	0.075	1	$1\sim4$

REFERENCES

[1] M. Leonardi and T. B. Ustun, "The global burden of epilepsy," *Epilepsia*, vol. 43, no. s6, pp. 21–25, 2002.

[2] F. Mormann, T. Kreuz, C. Rieke, R. G. Andrzejak, A. Kraskov, P. David, C. E. Elger, and K. Lehnertz, "On the predictability of epileptic seizures," *Clinical neurophysiology*, vol. 116, no. 3, pp. 569–587, 2005.

[3] Y. Park, L. Luo, K. K. Parhi, and T. Netoff, "Seizure prediction with spectral power of EEG using cost-sensitive support vector machines," *Epilepsia*, vol. 52, no. 10, pp. 1761–1770, 2011.

[4] A. Shoeb, H. Edwards, J. Connolly, B. Bourgeois, S. Ted Treves, and J. Guttag, "Patient-specific seizure onset detection," *Epilepsy & Behavior*, vol. 5, no. 4, pp. 483–498, 2004.

[5] I. Osorio, M. G. Frei, and S. B. Wilkinson, "Real-time automated detection and quantitative analysis of seizures and short-term prediction of clinical onset," *Epilepsia*, vol. 39, no. 6, pp. 615–627, 1998.

[6] https://epilepsy.uni-freiburg.de/freiburg-seizure-predictionproject/eeg database.

[7] S. S. Haykin, *Adaptive Filter Theory, 4th edition*. Prentice Hall, 2002.

[8] S. Mallat, *A wavelet tour of signal processing: the sparse way*. Academic Press, 2008.

[9] S. Theodoridis and K. Koutroumbas, *Pattern Recognition*. Academic Press, 2008.

[10] V. Cherkassky and F. M. Mulier, *Learning from data: concepts, theory, and methods, 2nd edition*. Wiley-IEEE Press, 2007.

[11] Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," in *Proc. of International Conference on Machine Learning (ICML)*, vol. 96, 1996, pp. 148–156.

[12] M. Ayinala and K. K. Parhi, "Low complexity algorithm for seizure prediction using AdaBoost," in *Proc. of IEEE Engineering in Medicine and Biology Society Conference (EMBC)*, 2012, pp. 1061–1064.

[13] J. Henriksen, L. S. Remvig, R. E. Madsen, I. Conradsen, T. W. Kjær, C. E. Thomsen, and H. B. Sorensen, "Automatic seizure detection: going from sEEG to iEEG," in *Proc. of IEEE Engineering in Medicine and Biology Society Conference (EMBC)*, 2010, pp. 2431–2434.

[14] A. Aarabi, R. Fazel-Rezai, and Y. Aghakhani, "A fuzzy rule-based system for epileptic seizure detection in intracranial EEG," *Clinical Neurophysiology*, vol. 120, no. 9, pp. 1648–1657, 2009.

[15] L. M. Patnaik and O. K. Manyam, "Epileptic EEG detection using neural networks and post-classification," *Computer methods and programs in biomedicine*, vol. 91, no. 2, pp. 100–109, 2008.

[16] H. Qu and J. Gotman, "A patient-specific algorithm for the detection of seizure onset in long-term EEG monitoring: possible use as a warning device," *IEEE Transactions on Biomedical Engineering*, vol. 44, no. 2, pp. 115–122, 1997.