Automatic absence seizure detection evaluating matching pursuit features of EEG signals

Katerina Giannakaki
School of Electrical and Computer Engineering
Technical University of Crete
Chania, Greece
katerina_gian@yahoo.gr

Giorgos Giannakakis
Institute of Computer Science
Foundation for Research and Technology - Hellas
Heraklion, Greece
ggian@ics.forth.gr

Pelagia Vorgia
Institute of Computer Science
Foundation for Research and Technology - Hellas
Heraklion, Greece
vorgia.p@uoc.gr

Manousos A. Klados
Dept. of Biomedical Engineering
School of Life and Health Sciences
Aston University
Birmingham, United Kingdom
m.klados@aston.ac.uk

Michalis Zervakis
School of Electrical and Computer Engineering
Technical University of Crete
Chania, Greece
michalis@display.tuc.gr

***Abstract*—This paper evaluates the usage of matching pursuit (MP) features of electroencephalographic (EEG) signals and classification techniques on automatic absence seizure detection. Absence epileptic seizures are neurological disorders which are manifested as abnormal EEG patterns. Matching pursuit algorithm is able to decompose a signal into components with specific time-frequency characteristics. It is a robust technique especially when there is complex, multicomponent signal. In the present study, a clinical dataset containing 40 annotated absence seizures in long-term EEG recordings from pediatric epileptic patients (with age 6.0±2.9 years) was analyzed. The extracted MP features fed an automatic classification schema which achieved a time window based discrimination accuracy of 98.5%. As indicated by the study's results, the proposed features and analysis methods can be a promising addition to the area of automatic absence seizures detection.**

Keywords—seizure, EEG, matching pursuit, seizure detection, classification

I. Introduction

Epilepsy is a neurological disorder which constitutes the most common brain disorder in childhood. According to the World Health Organization, it is estimated that in 2019, epilepsy affects around 50 million people worldwide [1]. In 2017, the prevalence and incidence of epilepsy are estimated to be 6.38 and 0.67 per 1000 persons respectively [2].

Although there are various diagnostic techniques such as magnetoencephalography (MEG), electrocardiography (ECG), magnetic resonance imaging/functional magnetic resonance imaging (MRI/fMRI), positron-emission tomography (PET), computed tomography (CT) etc, however electroencephalography (EEG) remains the most widely adopted clinical technique for seizure diagnosis, detection, and anticipation [3]. The advantages of EEG are, among others, that it is a relative non-invasive and accurate method, thus being ideal for clinical practice. In daily home monitoring through wearable devices ECG gains ground, however EEG retains its superiority in relation to ECG in terms of predictive evidence, localization ability, and temporal resolution [4].

Seizure detection is a fundamental step in research and confrontation of epilepsy. An expert neurologist performs visual seizure detection combining the patient's clinical picture, the EEG (main diagnostic tool) and other modalities as synchronized video, ECG, fMRI, etc. This procedure, especially during long-term monitoring, may be time-consuming and tedious, limiting treatment's effectiveness.

Various methods have been developed based on EEG signal analysis supporting automatic seizure detection [3, 5-7]. The absence seizures, which are investigated in this study, are generalized seizures that present specific patterns during ictal period. Advanced signal processing techniques enable the signal decomposition on its time/frequency components utilizing basis functions approach [8]. Especially, in complex and multicomponent signals, these methods have the advantage of eliminating noisy time-frequency cross terms. Matching Pursuit (MP) algorithm [9] provides an efficient signal decomposition using time-frequency components (atoms). It has been used in EEG analysis [10, 11] and in few studies related to seizure detection [12-14]. A challenging issue is to use decomposition methods in order to extract robust, representative features for seizure discrimination.

Machine learning techniques have been widely used in the analysis of EEG signal, offering positive results in various research areas such as discrimination between emotional states [15], enhancement of brain-computer interfaces [16], motor imagery, and epileptic seizure detection [17].

In this study, we use MP algorithm for automatic seizure detection, proposing two features by combining selected MP atoms that provide characteristic information. These features are, in our view, more representative in relation to features in mentioned literature. Then, feature ranking and classification methods are used for discriminating interictal/ictal periods.

II. Clinical Protocol and Data Acquisition

A. Inclusion criteria and ethics

Subjects participating in this study are patients diagnosed with absent seizures. They had presented at least one seizure event in the last month, which made them eligible for inclusion in the study. The study's protocol has been approved by the appropriate scientific board of the University Hospital of Heraklion. Informed consent was obtained from all patients following a detailed explanation of the study objectives and protocol to each patient and/or caregiver. All caregivers/patients provided written informed consent.

B. Procedure

Patients that met the inclusion criteria as evaluated by two expert neuropediatricians, were admitted to the hospital. Their

2471-7819/19/$31.00 ©2019 IEEE
DOI 10.1109/BIBE.2019.00165

TABLE I. METHODS AND PARAMETERS USED IN THE ANALYSIS

Patient Code	# seizures	Age	Gender
PAT_01	4	6	Male
PAT_02	3	9	Female
PAT_04	27	7	Male
PAT_05	6	2	Male
Total seizures	**40**		

medical health record was created including clinical data about demographics, medical history, family history, medication, epilepsy classification, etc. An EEG cap with 10/20 electrode system was placed in the head of the patient, a camera was placed opposite the patient's bed and additional sensors for recording the breath rate and SpO2 were utilized. Video and surface EEG were recorded simultaneously for each patient during routine long-term hospital monitoring. The EEG signals were recorded at 21 scalp loci of the international 10-20 system (channels Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, A1, A2, O1, O2), with all electrodes referenced to the earlobe. An electrode placed in the middle of the distance between Fp1 and Fp2 on the subject's forehead served as ground. EEG data were sampled at 256Hz.

C. Dataset

This study's population consisted of 4 patients diagnosed with active absence epilepsy, 3 males and 1 female (with age 2-9 years) with 40 annotated absence seizures. The EEG recordings were independently evaluated and annotated for epileptic seizures and pathological findings by two expert neuropediatricians. All the epileptic seizures identified, were classified as absence like generalized seizures according to the criteria of the International League Against Epilepsy (ILAE) [18]. Table I presents patients' demographic data as well as selected clinical data.

III. METHODS

A. Matching pursuit algorithm

The matching pursuit algorithm [9] is an iterative procedure which provides a mathematical formulation of approximation of a signal using a set of functions (atoms). The redundant set of time-frequency atoms is called dictionary D. The dictionary can be comprised of any arbitrary function, however, in this study we construct the dictionary D from Gabor functions which are considered effective in approximating EEG signals [19]

$$g_\gamma(t) = K(\gamma) e^{-\pi\left(\frac{t-u}{\sigma}\right)^2} \sin\big(2\pi f(t-u)\big) \tag{1}$$

where K(γ) is a coefficient such that $||g_\gamma||=1$, $\gamma = [u, f, \sigma]$ denotes dictionary's function parameters, u is the translation in time, f is the frequency, σ is the Gaussian spread.

The algorithm looks for the Gabor function that best matches to an inner pattern of the original signal x over a redundant set of atoms selected from the dictionary. This is done by successive approximations of x with orthogonal projections on elements of the dictionary, i.e. the inner product between the Gabor function and the signal. This inner product is then subtracted from the signal and the next iteration take place. Let $R^0 x = x$ .We suppose that we have computed the n^{th} order residue $R^n s$ for $n \geq 0$. We choose an element $g_{\gamma_n} \in D$ from the dictionary D which best matches the signal $R^n x$ (the residual left after subtracting the results of previous iterations). The residue $R^n x$ can be also decomposed into:

$$\begin{cases} R^n x = \left\langle R^n x, g_{\gamma_n} \right\rangle g_{\gamma_n} + R^{n+1} x \\ g_{\gamma_n} = \arg\max\limits_{g_{\gamma i} \in D} \left| \left\langle R^n x, g_{\gamma i} \right\rangle \right| \end{cases} \tag{2}$$

where $argmax_{g\gamma i} \in_D$ means the $g_{\gamma i}$ giving the largest value of the product $<R^n x, g_{\gamma i}>$. The iterative procedure of decomposition stops either when the energy of residual signal is below a preset cut-off level ε or, alternatively after a predetermined number of iterations M. After M iterations, a matching pursuit decomposes a signal x into:

$$x = \sum_{n=0}^{M-1} \left\langle R^n x, g_{\gamma_n} \right\rangle g_{\gamma_n} + R^m x \tag{3}$$

where $R^m s$ is the residual vector after m iterations and $\left\langle x, g \right\rangle = \int_{-\infty}^{\infty} x(t)\bar{g}(t)dt$ denotes inner product of functions s and g. This inner product $a_n = \left\langle R^n x, g_{\gamma_n} \right\rangle$ represents also the magnitude of the selected atom. Because the orthogonality of $R^{n+1} x$ and to $g_{\gamma n}$ is valid in each step of the procedure, the form of energy conservation law becomes:

$$\|x\|^2 = \sum_{n=o}^{M} \left| \left\langle R^n x, g_{\gamma_n} \right\rangle \right|^2 + \left\| R^m x \right\|^2 \tag{4}$$

When the iterative procedure terminates, the selection of Gabor atoms from dictionary is completed.

B. MP features extraction

The MP algorithm estimates a number of Gabor atoms that best describes the signal in decreasing order in terms of signal's energy variation described. Each atom has specific time-frequency properties which can be described by its corresponding frequency f_n, and specific energy which can be described by its corresponding magnitude α_n. In order to provide specific features for a selected sliding window, 3 features were estimated, the first being the mean amplitude (*MA*) value of the selected atoms

$$MA = \frac{1}{M} \sum_{i=1}^{M} a_i$$

We also introduced two new metrics, the weighted frequency mean (*WFM*)

$$WFM = \frac{\sum_{i=1}^{M} a_i f_i}{\sum_{i=1}^{M} a_i}$$

and the mean-product frequency (*MPF*)

$$MPF = \sum_{i=1}^{M} a_i f_i$$

which both describe weighted metrics of frequency on the selected sliding time window. In our point of view, the *WFM* and *MPF* are considered to be better metrics in relation to the mean frequency described in relevant study [13], because they take into account the contribution of each atom's spectral content through its amplitude α_i. The selection of these specific features was founded on the observation that the frequency and the EEG signal envelope (amplitude) were increased during seizures.

C. Feature classification

In this study, the discrimination between the two states under investigation (interictal, ictal) was performed by evaluating MP features via classification schemes comparison. The classification schemes used are summarized and presented in TABLE II. The Trivial classifier classifies everything in the most frequent class, and is used as a reference point for the performance of the other classifiers, since it is considered to represent random classification. In order to assess the performance of each classification scheme the classification accuracy, sensitivity and specificity metrics were used, which are given by the equations

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

$$Sensitivity = \frac{TP}{TP + FN}\ ,\ Specificity = \frac{TN}{FP + TN}$$

where TP is the true positive, TN the true negative, FP is the false positive and FN the false negative cases. As used in this study, sensitivity is the proportion of cases predicted as belonging to preictal period, and actually belonged to preictal period, while specificity is the proportion of cases predicted as belonging to ictal period, that actually belonged to ictal period.

The classification schemes (classifier and its parameters) were cross-validated in order to evaluate their performance and select the best combination. A standard 10-fold cross validation method was utilized for each classification scheme for testing the system's performance.

IV. Results

A. EEG preprocessing and MP dictionary construction

The EEG recordings were digitized in sampling frequency f_s=256 Hz. Artifacts related with subject's activity (body movements, eye blinks, spikes, head movements, chewing, general discharges) contaminate EEG recordings with unwanted noise components. The suppression of artifacts and spikes was performed using Independent Component Analysis (ICA) [20]

A sliding temporal window of Δt=2sec and a step of 0.5sec was used for the signal analysis. This approach was chosen in order to track EEG temporal dynamics in interictal phase as well as the transition from interictal to ictal phase. It was checked that the increase of time window does not affect the system performance significantly. As the dataset of seizure detection is highly non-balanced, i.e. data from ictal periods are much fewer than interictal periods, a balanced dataset should be ensured for formulating a proper model. Thus, all samples from ictal and their double samples from interictal periods were selected for the subsequent analysis.

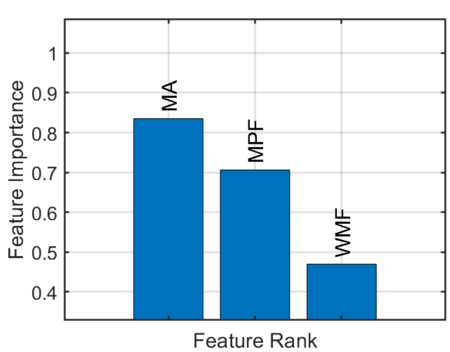

Fig. 2: Features ranking according to the Fisher discrimination ratio for the discrimination problem (interictal, ictal states)

TABLE II. Classification Accuracies (10-fold cross validation) for the Selected feature set used in this study

Classifiers	Accuracy (%)	Sensitivity (%)	Specificity (%)
Trivial Classifier	66.7	-	-
KNN (k=5)	98.1	98.2	90.6
GLM	97.8	99.8	93.8
NVB	97.4	98.3	96.7
LDA	97.3	99.7	99.6
SVM	**98.5**	**99.8**	**96.8**

The parameters used for the construction of MP dictionary was N_{dict}=512 samples, σ_i=[2:2:256] samples, f_i=[1:0.5:30] Hz and u=[-N_{dict}/2:2:N_{dict}/2-1] leading to a dictionary of 15.360 different atoms.

B. MP features

The features WMF, MPF and MA were calculated for each sliding temporal window. A typical temporal evolution of the EEG signal and its corresponding MA timeseries is presented in Fig. 1. These features form the feature matrix then fed the classification scheme. The features' importance and relevance for the investigated discrimination problem were assessed using Fisher discrimination ratio [21] and their ranking is presented in Fig. 2. One can observe that MA achieves the highest ranking, thus it is the most relevant feature for the discrimination problem.

Then, the features set was evaluated in terms of its discrimination ability between interictal and ictal state. A 10-fold cross-validation technique on sample basis was used with the trivial classifier and Naïve Bayes (NVB), K-Nearest Neighbors (KNN), Generalized Linear Model (GLM), Linear Discriminant Analysis (LDA), Support Vector Machines (SVM) classifiers. The classification accuracy results are shown in TABLE II.

It can be observed from TABLE II that SVM classifier outperforms all other classification schemes with a classification accuracy of 98.5%. The distribution of the 2 top-ranked features (MA, MPF) is presented in Fig. 3 as a classification plot with the decision boundaries of SVM for the 2 features along with samples of a testing fold (blue:

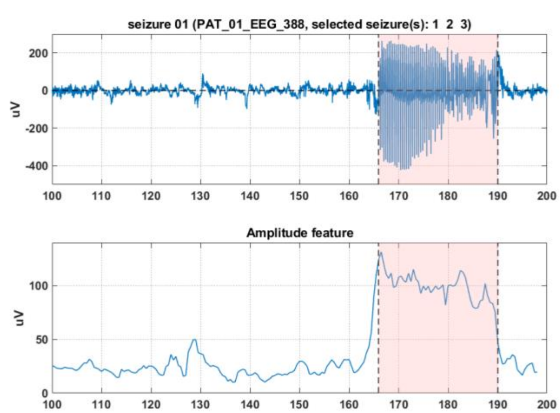

Fig. 1: EEG signal of electrode Fp2 (upper graph) and its corresponding MA temporal evolution. The vertical dashed lines represent the onset and ending of the seizure.

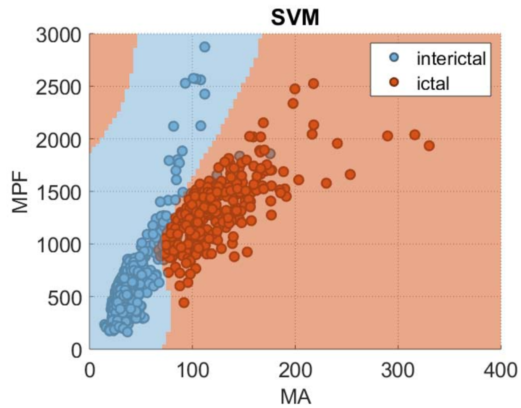

Fig. 3: Distribution visualization of the 2 top-ranked features MA and MPF of one classification fold distributions of interictal (blue) and ictal period (red)

interictal, red: ictal). This figure presents the data separability achieved using this study's features and the discrimination efficiency of the proposed methodology.

V. Discussion

The aim of this study was to train a model that could effectively discriminate between interictal and ictal periods. Matching Pursuit algorithm based features were extracted using sliding time window, followed by various classifiers.

The MP algorithm is a decomposition method that extracts best matching Gabor atoms to the inner signal patterns following the principle of maximal energy. The fact that these atoms represent different levels of signal's energy make the extraction of one specific feature for a selected time-window a difficult task. Relevant literature use metrics such as mean of frequencies which is not representative of the resultant dominant frequency. We propose two features that, in our view, represent the resultant frequency in the selected time window using weighted frequency metrics in the same way MP operates, i.e. using the weighted coefficients of MP.

These MP features relate to the strength and frequency of Gabor atoms that compose the EEG signal and provide good separability between interictal and ictal states. Machine learning techniques led to a best achieved time-window classification of 98.5% using 10-fold cross validation. These results indicate that these MP features are features that can be used effectively in seizure detection procedure.

In the future, seizure based classification will be estimated in combination with other EEG or ECG features using personalized information which is closer to the modern clinical approach. Furthermore, the predictive ability of the feature set will be evaluated on extensive datasets.

Acknowledgment

The authors would like to thank Spyridon Voutoufianakis, neuropediatrician, for the seizure annotation and Giorgos Livas for his valuable technical support.

References

[1] "Epilepsy: a public health imperative: summary," World Health Organization report, 2019.

[2] K. M. Fiest, K. M. Sauro, S. Wiebe, S. B. Patten, C.-S. Kwon, J. Dykeman, T. Pringsheim, *et al.*, "Prevalence and incidence of epilepsy A systematic review and meta-analysis of international studies," *Neurology,* vol. 88, pp. 296-303, 2017.

[3] G. Giannakakis, V. Sakkalis, M. Pediaditis, and M. Tsiknakis, "Methods for seizure detection and prediction: an overview," in *Modern Electroencephalographic Assessment Techniques, Neuromethods*. vol. 91, ed: Humana Press, New York,, 2015, pp. 131-157.

[4] G. Giannakakis, M. Tsiknakis, and P. Vorgia, "Focal Epileptic Seizures Anticipation based on Patterns of Heart Rate Variability Parameters," *Computer Methods and Programs in Biomedicine,* vol. 178, pp. 123-133, 2019.

[5] G. Giannakakis, V. Sakkalis, M. Pediaditis, C. Farmaki, P. Vorgia, and M. Tsiknakis, "An approach to absence epileptic seizures detection using Approximate Entropy," in *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2013, pp. 413-416.

[6] V. Sakkalis, G. Giannakakis, C. Farmaki, A. Mousas, M. Pediaditis, P. Vorgia, and M. Tsiknakis, "Absence seizure epilepsy detection using linear and nonlinear EEG analysis methods," in *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2013, pp. 6333-6336.

[7] G. Giannakakis, P. Vorgia, S. Voutoufianakis, M. Pediaditis, D. Champsas, H. Dimitriou, and M. Tsiknakis, "Low Frequency Preictal Eeg Waves Preceding Epileptic Seizures Originating From Right Frontal Lobe," *Epilepsia,* vol. 56, pp. 94-95, 2015.

[8] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM review,* vol. 43, pp. 129-159, 2001.

[9] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *Ieee Transactions on Signal Processing,* vol. 41, pp. 3397-3415, 1993.

[10] P. Durka and K. Blinowska, "Analysis of EEG transients by means of matching pursuit," *Ann Biomed Eng,* vol. 23, 1995.

[11] G. Giannakakis, M. Vavatsioula, C. Papageorgiou, and K. Nikita, "Statistical Time-Frequency Differentiations of ERP in Dyslexia using Matching Pursuit Algorithm," in *6th European Symposium on Biomedical Engineering (ESBME 2008), Chania, Greece*, 2008.

[12] P. J. Franaszczuk, G. K. Bergey, P. J. Durka, and H. M. Eisenberg, "Time–frequency analysis using the matching pursuit algorithm applied to seizures originating from the mesial temporal lobe," *Electroencephalography and Clinical Neurophysiology,* vol. 106, pp. 513-521, 1998.

[13] Z. Emigdio, L. Trujillo, A. Sotelo, P. Legrand, and L. N. Coria, "Regularity and matching pursuit feature extraction for the detection of epileptic seizures," *J Neurosci Methods,* vol. 266, pp. 107-125, 2016.

[14] T. L. Sorensen, U. L. Olsen, I. Conradsen, J. Henriksen, T. W. Kjaer, C. E. Thomsen, and H. B. Sorensen, "Automatic epileptic seizure onset detection using matching pursuit: a case study," in *Annual International Conference of the IEEE Engineering in Medicine and Biology*, 2010.

[15] K. Giannakaki, G. Giannakakis, C. Farmaki, and V. Sakkalis, "Emotional State Recognition Using Advanced Machine Learning Techniques on EEG Data," in *2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS)*, 2017, pp. 337-342.

[16] Y. Zhang, Y. Wang, G. Zhou, J. Jin, B. Wang, X. Wang, and A. Cichocki, "Multi-kernel extreme learning machine for EEG classification in brain-computer interfaces," *Expert Systems with Applications,* vol. 96, pp. 302-310, 2018.

[17] A. H. Shoeb and J. V. Guttag, "Application of machine learning to epileptic seizure detection," in *Proceedings of the 27th International Conference on Machine Learning (ICML-10)*, 2010, pp. 975-982.

[18] A. T. Berg, S. F. Berkovic, M. J. Brodie, J. Buchhalter, J. H. Cross, W. van Emde Boas, J. Engel*, et al.*, "Revised terminology and concepts for organization of seizures and epilepsies: report of the ILAE Commission on Classification and Terminology, 2005–2009," *Epilepsia,* vol. 51, pp. 676-685, 2010.

[19] J. Sinkkonen, H. Tiitinen, and R. Näätänen, "Gabor filters: an informative way for analysing event-related brain activity," *J Neurosci Methods,* vol. 56, pp. 99-104, 1995.

[20] A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural Networks,* vol. 13, pp. 411-430, 2000.

[21] B. Seijo-Pardo, I. Porto-Díaz, V. Bolón-Canedo, and A. Alonso-Betanzos, "Ensemble feature selection: homogeneous and heterogeneous approaches," *Knowledge-Based Systems,* vol. 118, pp. 124-139, 2017.