Automatic Epileptic Seizure Onset Detection Using Matching Pursuit

A case Study

Sorensen, Thomas Lynggaard; Olsen, Ulrich L.; Conradsen, Isa; Duun-Henriksen, Jonas; Kjaer, Troels W.; Thomsen, Carsten E.; Sørensen, Helge Bjarup Dissing

Published in:
Engineering in Medicine and Biology Conference

Link to article, DOI:
10.1109/IEMBS.2010.5627265

Publication date:
2010

Document Version
Publisher's PDF, also known as Version of record

32nd Annual International Conference of the IEEE EMBS
Buenos Aires, Argentina, August 31 - September 4, 2010

Automatic Epileptic Seizure Onset Detection Using Matching Pursuit: A Case Study

Thomas L. Sorensen†, Ulrich L. Olsen†, Isa Conradsen†, Jonas Henriksen†,
Troels W. Kjaer*, Carsten E. Thomsen‡ and Helge B. D. Sorensen†

Abstract— **An automatic alarm system for detecting epileptic seizure onsets could be of great assistance to patients and medical staff. A novel approach is proposed using the Matching Pursuit algorithm as a feature extractor combined with the Support Vector Machine (SVM) as a classifier for this purpose. The combination of Matching Pursuit and SVM for automatic seizure detection has never been tested before, making this a pilot study. Data from red different patients with 6 to 49 seizures are used to test our model. Three patients are recorded with scalp electroencephalography (sEEG) and three with intracranial electroencephalography (iEEG). A sensitivity of 78-100% and a detection latency of 5-18s has been achieved, while holding the false detection at 0.16-5.31/h. Our results show the potential of Matching Pursuit as a feature extractor for detection of epileptic seizures.**

I. INTRODUCTION

About 1 % of the world's population suffers from epilepsy [1][2], making it one of the most frequent neurological disorders only outnumbered by stroke and headache [3]. About 75 % of epilepsy patients can be seizure free on antiepileptic drugs, and some of the remaining 25 % can be treated with other procedures, like surgical resection of the epileptic focus, a vagus nerve stimulator or a ketogenic diet [4].

The goal of this study is to build an automatic onset detection for epileptic seizures. Such an alarm would give patients suffering from epilepsy an opportunity to leave their homes knowing that family or medical personnel can come to their rescue if they encounter a seizure. Furthermore it is important to register the number of seizures the patient encounter in a given time frame. This can give medical doctors insight on how well a treatment is working. It can also be important to know when a patient has a seizure, in case of acute treatment, or if a tracer drug has to be administered for an ictal SPECT-scan. An automatic trigger for the vagus nerve stimulator is another possibility, since it has the greatest effect if it is activated early in the seizure [5]. An automated seizure detection system would also assist in detecting seizures in large encephalography (EEG) data sets, that often include recordings from several days.

Automatic seizure detection is not a new idea. Through the past couple of decades many attempts have been made, to find the optimal algorithm for classification, primarily using intracranial EEG (iEEG) or scalp EEG (sEEG) [6]. More recently other approaches have been attempted such as accelerometers, electromyography (EMG) and angular velocity recordings [1][4].

We have applied the Matching Pursuit algorithm on both iEEG and sEEG data providing features, which will be used with the Support Vector Machine (SVM) classifier. The algorithm was first used to study ictal EEGs by Jouny et al. in 2003 [7]. However this is the first time SVM has been combined with Matching Pursuit for seizure onset detection.

II. METHOD

A. Clinical data

We have included six patients (pt.) with a total of 133 seizures in 305 hours of recordings (rec.) in this study. To investigate if the robustness of the algorithm depends on whether data is collected intracranially or extracranially, two of the patients are recorded with sEEG and two are recorded with iEEG. The EEG-data is recorded at the Epilepsy Monitoring Unit (EMU) at Rigshospitalet University Hospital, Copenhagen. The sEEG-data are recorded at a sampling frequency of 200 Hz from patients admitted for diagnostic workup, using Stellate™ Harmonie with 21-25 EEG channels, placed using the 10-20 system.

TABLE I
PATIENT INFORMATION

Pt.	Sex	Age	Rec.	Modality	Type	# of Seizures
P1	M	6	49 h	sEEG	pGTCS	10
P2	M	63	8 h	sEEG	CPS	49
P3	F	33	44 h	sEEG	SPS	35
P4	M	45	95 h	iEEG	CPS	20
P5	F	28	66 h	iEEG	SPS/CPS	13
P6	M	45	43 h	iEEG	SPS/CPS	6
Sum			305 h			133

- pGTCS = primary Generalized Tonic Clonic Seizures
- CPS = Complex Partial Seizures
- SPS = Simple Partial Seizures

†Department of Electrical Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark
*Department of Clinical Neurophysiology, Rigshospitalet University Hospital, Copenhagen, Denmark
‡Department of Odontology, University of Copenhagen, Denmark

T.L. Sorensen: thomas.lynggaard@gmail.com
U.L. Olsen: ulrich.olsen@gmail.com
H.B.D. Sorensen: hbs@elektro.dtu.dk

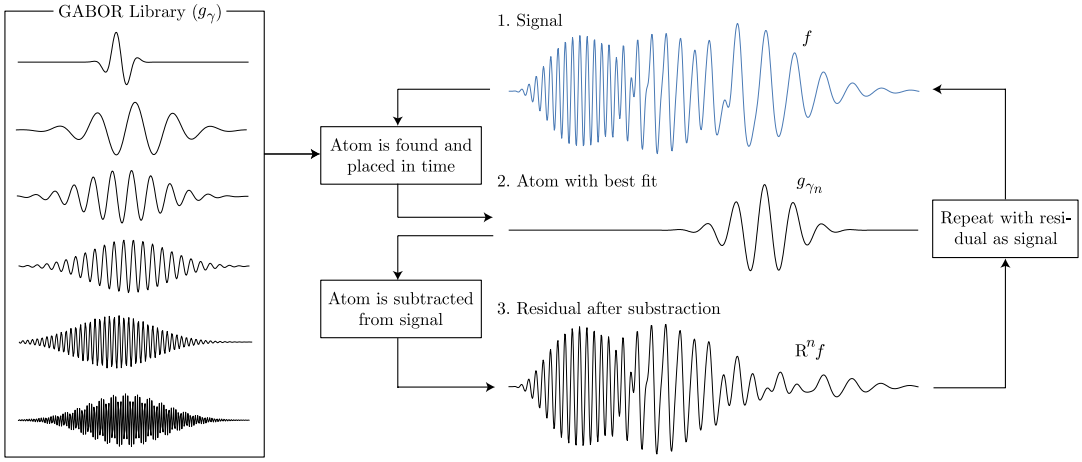

Fig. 1. A flowchart of the Matching Pursuit algorithm. First the inner product between all Gabor functions g_γ in the library and the signal f is found. The Gabor function g_{γ_n} which leads to the highest inner product is chosen and subtracted from the signal f, using the inner product value as the Gabor function amplitude. The process is then repeated with the residual $R^n f$ as the signal.

The iEEG-data are recorded from patients monitored prior to epilepsy surgery. It is recorded at a sampling frequency of 200 Hz using grid and strip electrodes. The seizures are outlined by a specialist in clinical neurophysiology (Troels W. Kjaer, MD, Ph.D.), to have a frame of reference for training the detection algorithm, and for the calculation of latency between the true EEG onset and the onset estimated by the algorithm. Furthermore three channels that clearly show epileptic activity are chosen by the specialist, which greatly minimizes the amount of data. Table I shows information for the four patients and their seizure types.

B. Matching Pursuit

The Matching Pursuit algorithm was developed in the early 1990's by Mallat and Zhang [9]. The fundamental concept is to approximate a signal by the sum of functions found in a dictionary. The functions found in the dictionary, are referred to as atoms. Often Gabor functions are used, which are sinusoids multiplied by a gaussian function [9]. This study uses the Gabor function for the decomposition. The software used to compute the Matching Pursuit algorithm, is the original software by Mallat and Zhang, rewritten by Dr. C. Jouny and Dr. P. Franaszczuk from the Epilepsy Reseach Laboratory at Johns Hopkins Medical Institutions. It can be downloaded at: http://erl.neuro.jhmi.edu/mpsoft/. Dr. Supratim Ray from Harvard Medical School has adapted the code to run from within MATLAB.

To find the Gabor function that best describes the signal, the inner product between the Gabor function and the signal is calculated. The Gabor function that leads to the highest inner product is used, see Fig. 1. The inner product is then used as the amplitude for the Gabor function and the Gabor function is subtracted from the signal. An iterative process is run until the energy is below a specified threshold. The Matching Pursuit algorithm is described by equation (1) where a finite number of Gabor functions, m, are used to decompose the signal [9].

$$f = \sum_{n=0}^{m-1} \langle R^n f, g_{\gamma_n} \rangle \, g_{\gamma_n} + R^m f \tag{1}$$

$R^n f$ is the residual of the n'th iteration and g_{γ_n} is the n'th Gabor function.

Because the Gabor functions are sinusoids limited in time by a gaussian function, a non-complex Gabor function is given by:

$$g_\gamma(t) = K(\gamma) e^{-\pi\left(\frac{t-u}{s}\right)^2} \cos\left(\omega(t-u) + \phi\right), \tag{2}$$

where $\gamma = \{u, s, \omega, \phi\}$. u shifts the atom in time, s defines the width of the atom, ω is the frequency of the atom in rad/s and ϕ is the phase of the atom. $K(\gamma)$ is a scaling factor making $||g_\gamma(t)|| = 1$. Matching Pursuit is an iterative process and therefore a stop criteria is needed. Jouny et. al [7] introduced a stop criteria that is simply an energy threshold of the last atom. The decomposition will continue to run until the energy of the last atom is lower than the specified energy threshold.

C. Features

The Matching Pursuit algorithm returns the number of Gabor functions needed to reconstruct the signal. In [7], focus is on the number of atoms m needed to describe the signal. They state that the complexity of ictal EEG is higher than the complexity of interictal EEG, thus more atoms are needed to describe ictal EEG signals than is needed for interictal EEG. Therefore we apply the density of Gabor functions normalized with the window length, *Gabor Atom Density* (GAD) [7][8], as a feature for seizure detection. The equation for GAD is:

$$GAD = \frac{m}{\Delta t \cdot \Delta f} = \frac{2 \cdot m}{N}, \tag{3}$$

where Δt is the window size in seconds and Δf is the size of the frequency scale; making N the window size in samples.

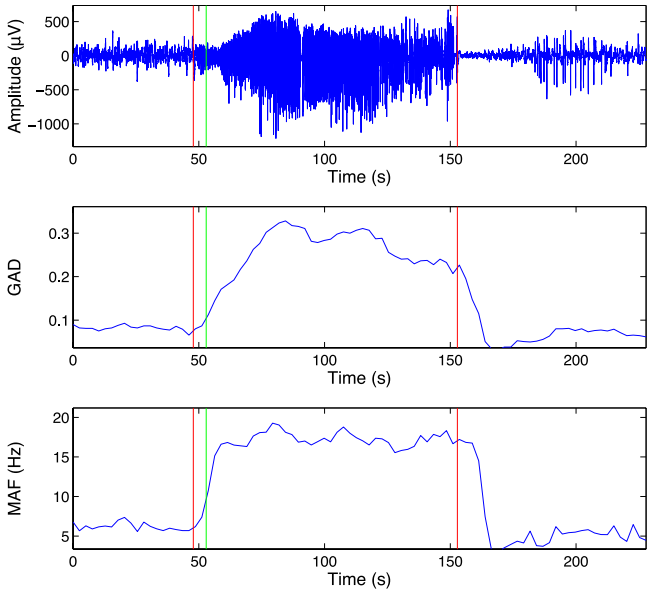

Fig. 2. Top: The raw EEG signal of Patient P4. Middle: The GAD curve calculated from (3). Bottom: The MAF curve calculated from (4) The red line marks the seizure start and end placed by the specialist. The green line marks the seizure onset detection by the SVM classifier.

We found that the frequencies increases significantly in the EEG during a seizure. This makes the mean of all the gabor functions within a window, *Mean Atom Frequency* (MAF), another possible feature enabling the algorithm to accurately distinguish between seizure and non-seizure. MAF can be described by:

$$MAF = \frac{1}{m}\sum_{n=1}^{m}\omega_{g_{\gamma_n}}, \tag{4}$$

where $\omega_{g_{\gamma_n}}$ is the frequency in rad/s of the atom g_γ.

It has been proved that working on data from multiple channels increases the accuracy, why three focal channels are used for this study. Taking the GAD and MAF for each of the three channels, gives a total of six features. The features are fed as input to the SVM classifier.

For the scalp EEG the features are extracted from the signal by applying a 512 sample window to a 200 Hz signal, hence 2.56 seconds. For iEEG the window size is 2048 samples, also sampled at 200 Hz. The reason for the different window sizes is to optimize the difference between interictal and ictal data, and is a result of testing the performance with different sizes. An overlap of 75% is used on both modalities resulting in a temporal resolution of 0.64 and 2.56 seconds respectively. For the stop criteria an energy threshold of 200 μV^2 was used for sEEG, as was found in [7] to be the optimum. For the iEEG data an energy threshold of 500 μV^2 was used as it results in the largest difference between the interictal and ictal data. On Fig. 2 the two features calculated for a 230 second period is shown. During the seizure, a clear rise in both GAD and MAF can be observed.

D. Support Vector Machine

The SVM algorithm was developed by C. Cortes and V. Vapnik in 1995 [12][13]. In this context it is applied for classification of the features. The MATLAB implementation of the SVM classifier is used for this study.

The SVM is trained using the first 30% of the seizures for causality, combined with three times the amount of background data randomly selected from the first part of the EEG recordings. The remaining seizures were used for testing along with as much background data, from after the training period, as possible.

E. Post processing

To eliminate artifacts of short periods, a temporal constraint was applied to the SVM output. The constraint only allows detection of a seizure after the SVM has classified a number of consecutive epochs containing seizure activity. This post processing greatly reduces the number of false positives, but also leads to an increase in detection latency. The number of consecutive epochs depends on the use of the algorithm. One should seek to specify the number that best fulfill the purpose of the algorithm. The constraint has only been applied to the sEEG data, as the false detection rate in the iEEG data was significantly smaller, and the constraint has therefore been omitted. For sEEG five consecutive epochs were required to be classified, by the SVM as a seizure, before the algorithm accepted it as a seizure. The optimal number of five consecutive epochs (for this purpose) was found by empirical evaluation.

III. Results

Fig. 3 shows the two features, GAD and MAF, for one channel recorded using iEEG. It can be seen that the seizure epochs are well separated from the background data and confined to a relatively small area. This leads to a low *False Detection Rate* (FDR), which is particularly true for the iEEG data, as can be seen in table II.

The sEEG data have higher FDRs and by looking at the feature plots for these data (not depicted) it can be seen that, though they are separated from the background data, they are not as clearly separated as the iEEG data are.

For some patients it was found that both GAD and MAF peaked early in the seizure and then decayed steadily to the end of the seizure. This indicates that these features are best suited for seizure onset detection, and might not be able to

TABLE II
RESULTS SUMMARY

Patient	# of Seizures	Sensitivity (%)	FDR (1/h)	Latency (s)
P1	10	100	0.59	18.3
P2	49	91.4	5.31	9.1
P3	35	-	-	-
P4	20	95.0	0.20	6.06
P5	13	77.8	0.16	5.67
P6	6	100	1.8	7.44

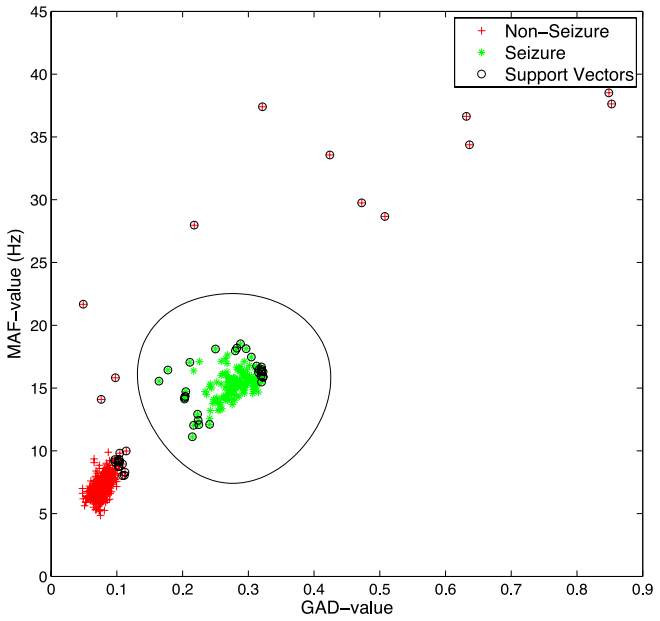

Fig. 3. The SVM training from patient P4 - channel 1. The decision boundary is based on the training seizure- and non-seizure epochs from the two features. The circled epochs are the support vectors used for defining the decision boundary, which is developed by using a radial basis kernel as mapping function, with $\sigma = 1$.

detect seizure duration. This has not been investigated in this study, but is a subject for our future work.

The algorithm is patient-specific. This means that the SVM is trained for each patient, and then tested on the rest of the seizures and background activity for the same patient.

From Table II it can be seen that the algorithm in this initial case study is capable of detecting 78-100% of all seizures, while keeping the FDR between 0.16-5.31/h. The detection latency was found to be 5-8s for iEEG and 9-18s for sEEG. We were unable to detect any seizures of P3 with the features used in this study.

IV. Discussion & Conclusion

The reason for the high FDR in patient P2, is due to a large amount of artifacts. The algorithm is sensitive to the artifacts found in this particular patient. A solution to avoid the artifact could be to pre-filter the data before feeding it to the algorithm. The seizure dynamics for P3 are different from the other patients. This results in a poor detection and a high FDR. For P3 different features are needed for effective seizure detection.

Another issue is the number of electrodes. We used three focal electrodes for the detection. In patient P4 and P5 the epileptic focus is mainly in one hemisphere, but the patients do encounter some seizures that have focus in the opposite hemisphere. These seizures are not detected, since no electrodes from the focus area was used in the algorithm. A better way would be to include all the available electrodes. This could also introduce the idea of focus localization.

This is a case study, so the full potential of the method is not fully investigated. The goal is to improve performance, especially on the onset detection latency which does seem plausible. With an artifact removal algorithm, some of the artifacts introduced when the medical personnel are interacting with the patient etc. could be avoided. This would remove the need for the epoch constraint, and thereby result in a better latency performance. Furthermore a larger patient database would show the robustness of the method better, and perhaps introduce the idea of a non-patient specific approach based on different seizure types. Also a larger patient database could show if there is a difference in modality, i.e. iEEG or sEEG, and if there is significant differences between the different types of seizures. Results show a tendency for iEEG to be the more robust modality for this algorithm, but further work will have to be conducted to confirm these findings.

The idea of using Matching Pursuit for seizure detection seems very promising, and we will investigate it further to achieve even better results than presented in this paper.

V. Acknowledgments

The authors would like to thank neurophysiology assistant Lennart Derm from Rigshospitalet University Hospital, Copenhagen, for invaluable help collecting, preprocessing and reformatting the EEG-data. Also a special thanks to Dr. Supratim Ray from Harvard Medical School for his help with implementing the Matching Pursuit algorithm into MATLAB.

References

[1] I. Conradsen, S. Beniczky, P. Wolf, D. Terney, T. Sams and HBD. Sorensen, Multi-modal Intelligent Seizure Acquisition (MISA) system - A new approach towards seizure detection based on full body motion measures. *Conf Proc IEEE Eng Med Biol Soc*, vol. 1, 2009, pp. 2591-2595.

[2] A.T. Tzallas, M.G. Tsipouras and D.I. Fotiadis, Epileptic Seizure Detection in EEGs Using Time-Frequency Analysis. *IEEE transactions on information technology in biomedicine*, vol. 13, No. 5, 2009, pp. 703-710.

[3] T.J Nowak and A.G. Handford, Pathophysiology: Concepts and Applications for Health Care Professionals. *McGraw-Hill Higher Education*, 2004, p. 633.

[4] T.M.E. Nijsen, P.J.M. Cluitmans, J.B.A.M. Arends, P.A.M. Griep, Detection of Subtle Nocturnal Motor Activity From 3-D Accelerometry Recordings in Epilepsy Patients. *Conf Proc IEEE Eng Med Biol Soc.*, vol. 54 no. 11, 2007, pp. 2073-2081.

[5] E.J. Hammond, B.M. Uthman, S.A. Reid and B.J. Wilder, Electrophysiological Studies of Cervical Vagus Nerve Stimulation in Humans: I. EEG Effects, *Epilepsia*, vol. 33, issue 6, 1992, pp. 1013-1020.

[6] J. Gotman, Automatic recognition of epileptic seizures in the EEG. *Electroencephalography Clin. Neurophys.*, vol. 54, 1982, pp. 530-540.

[7] C.C. Jouny, P.J. Franaszczuk and G.K. Bergey, Characterization of epileptic seizure dynamics using Gabor atom density *Clinical Neurophysiology*, vol. 114, 2003, pp. 426-437.

[8] C.C. Jouny, P.J. Franaszczuk and G.K. Bergey, Signal complexity and synchrony of epileptic seizures: Is there an identifiable preictal period? *Clin. Neurophys.*, vol. 116, 2005, pp. 552-558.

[9] S.G. Mallat and Z. Zhang, Matching Pursuits With Time-Freqency Dictionaries, *IEEE Transactions on signal processing*, vol. 41 No. 12, 1993, pp. 3397-3415.

[10] C.C. Jouny, G.K. Bergey and P.J. Franaszczuk, Partial seizures are associated with early increase in signal complexity, *Clinical Neurophysiology* vol. 121, 2010, pp. 7-13.

[11] P. Durka. Matching Pursuit and Unification in EEG Analysis. *Artech House*, 1st edition, 2007, ISBN 13: 978-1-58053-304-1.

[12] C. Cortes and V. Vapnik, Support-vector networks *Machine Learning*, vol. 20, issue.3, 1995, pp. 273-297.

[13] N. Christianini and J. Shawe-Taylor, An introduction to Support Vector Machines and other kernel-based learning methods. *Cambridge University Press*, 2007. ISBN 0 521 78019 5.