Automatic detector of High Frequency Oscillations for human recordings with macroelectrodes

R. Zelman [Student Member, IEEE],

Department of Biomedical Engineering at McGill University and the Montreal Neurological Institute, Montreal, QC, H3A 2B4, Canada, (514-398-6644 x00445; fax: 514-398-8106)

F. Mari,

Montreal Neurological Institute, Montreal, QC, H3A 2B4, Canada

J. Jacobs,

Montreal Neurological Institute, Montreal, QC, H3A 2B4, Canada

M. Zijlmans,

Montreal Neurological Institute, Montreal, QC, H3A 2B4, Canada

R. Chander, and

Montreal Neurological Institute, Montreal, QC, H3A 2B4, Canada

J. Gotman

McGill University and the Montreal Neurological Institute, Montreal, QC, H3A 2B4, Canada

Abstract

High Frequency Oscillations (HFOs) in the EEG are a promising biomarker of epileptogenic tissue. Given that the visual marking of HFOs is highly time-consuming and subjective, automatic detectors are necessary. In this study, we present a novel automatic detector that detects HFOs by incorporating information of previously detected baselines. The detector was trained on 72 channels and tested on 278, achieving a mean sensitivity of 96.8% with a mean false positive rate of 4.86%. This low rate is reasonable since only visually marked baseline segments were considered as the true negatives. This detector could be useful for the systematic study of HFOs and for their eventual clinical application.

I. Introduction

Epileptic conditions affect approximately 0.5% of the population [1]. About 20% of the patients are refractory to medication, so in some cases with presumed focal epilepsy the resection of the epileptic focus can be considered. When non-invasive techniques (scalp EEG and imaging) give only a rough estimate of the epileptogenic region but cannot ensure the precise location of the focus, invasive recording are sometimes required.

Traditionally, only frequencies up to 70Hz were considered of clinical relevance. Recent findings in rodents and humans showed the presence of higher frequency activity (up to

500Hz) and a possible relation of High Frequency Oscillations (HFOs) with epileptogenesis [2–7].

HFOs are spontaneous EEG patterns in the range 80–500Hz, consisting of at least 4 oscillations that can be distinguished from background. The lack of a quantitative definition makes the detection of HFOs difficult and subjective. Visual marking of HFOs can be performed (e.g.: [6, 7]), but it is highly time consuming (it takes about 10 hours to visually mark HFOs in a 10-channel 10-minute recording) and subjectivity is inevitable. Thus, the development of automatic HFO detectors is crucial for the systematic study of HFOs and for their eventual utilization in clinical settings.

Only a handful automatic detectors based on different energy functions exist [4, 8, 9]. In all these detectors the energy of the signal is compared with a threshold that is computed based on the segment of EEG under consideration (including the events), under the assumption that HFOs are rare events.

In this study, we propose a different approach in which we first detect baseline segments and then compute the energy threshold with respect to those baselines. By considering the detected baselines locally, the characteristic of the background surrounding the events is considered for detection. This is particularly important in channels where the background is not perfectly flat and in channels where a large number of interictal epileptiform discharges (spikes) are present.

II. Methods

A. Patient information

Between September 2004 and April 2008, 45 patients with medically intractable epilepsy underwent depth macro-electrode (surface area 0.8mm^2) implantation at the Montreal Neurological Hospital. The depth EEG (SEEG), recorded with the Harmonie system (Stellate, Montreal, Canada), was low-pass filtered at 500Hz and sampled at 2000Hz, allowing for the identification of HFOs. Twenty patients were randomly selected, but one had to be excluded due to continuous artefacts. All patients gave informed consent in agreement with the Research Ethics Board of the Montreal Neurological Institute and Hospital (MNI).

B. Channels and Events Selection

HFO events were identified independently by two experienced reviewers in all functioning channels during one minute of slow wave sleep with the method described in [7]. We considered as our gold standard those events jointly marked by the two reviewers. The duration of the gold standard (visual) HFO events was the intersection between their markings. In addition, baseline segments (where it was clear that no oscillation was present) were visually marked.

Channels with nearly continuous high frequency activity or less than one visually identified HFO or baseline were excluded from this study. The database therefore consisted of 19 one-minute sections, each with 6 to 36 channels. Given that the variability within channels of a

patient is as large as the variability across patients, all channels were considered independently. Twenty percent of the channels were randomly selected and used for training the detector, resulting in 278 channels for testing the performance of the detector. Of these channels, 97 were temporal and 181 were neocortical (NC); 73 were in the seizure onset zone SOZ of the patient. These channels included 5238 visually identified HFO events and 51076 visually identified baselines, with mean length of 237ms (range: 70 to 398ms) that were used as the gold standard events.

C. Automatic Detection of HFOs

Given that HFO are short oscillatory events that “stand out” from background activity, a logical approach to identify these events is an energy based detector. Existing automatic HFO detectors are based on the comparison of the energy of the signal with the EEG epoch that includes the events. The main difference among published detectors is the type of energy function that is computed on the filtered signal, either the root mean square (RMS) amplitude [4], the short-time line length [9], or the Hilbert envelope [8]. In contrast, this detector tries to look at the problem from a perspective more similar to the way that human reviewers mark: by comparing the EEG with the surrounding previously identified baseline.

D. The MNI detector

The proposed detector (referred to as *MNI detector*) consists of a baseline detector block and an energy based event detector block (that incorporates the baseline information). The MNI detector aims to detect as many events as possible, even if a large number of false detections take place.

The first step is the detection of baselines. We define as baselines, segments of EEG where there is no oscillatory activity of any kind. The baseline detector [10] is based on the wavelet entropy (WE) which measures the degree of randomness (vs. oscillatory behavior) in the signal [11]. To enhance the oscillatory characteristics of the signal, the WE was applied to the autocorrelation of the filtered (80–450Hz) signal [10]. Thus, the wavelet power is computed as

$$W_n(a, b) = \left| \frac{1}{\sqrt{a}} \int_{t=0}^L r_n(t) \psi^* \left(\frac{t-b}{a} \right) dt \right|^2 \quad (1)$$

where $\psi(a, b)$ is chosen as the complex Morlet wavelet and $r_n(t)$ is the autocorrelation of the n^{th} EEG segment of length $L=75\text{ms}$. The normalized wavelet power is

$$P_n(a, b) = \frac{W_n(a, b)}{\sum_a W_n(a, b)} \quad (2)$$

and the WE_n is obtained as

$$WE_n(b) = - \sum_a P_n(a, b) \log_{10}[P_n(a, b)] \quad (3)$$

For a white noise signal, there will be similar contributions at all scales and the maximum WE is obtained as $WE_{\max} = \log_{10}(1/\#a)$ where $\#a$ is the number of possible scales. Thus, a segment is considered as baseline (ie: without oscillatory behaviour) when the minimum WE_n for segment n is larger than the threshold obtained from training ($0.56 WE_{\max}$).

The next block detects possible HFOs based on the energy defined as the moving average of the root mean square amplitude of the filtered signal. Each channel is band-pass filtered (80–450Hz) and segments with energy above threshold during more than 15ms were considered HFOs. The energy threshold is obtained by computing the empirical cumulative distribution function (CDF) of each 2-sec baseline segment and returning the value at the 95 percentile as the threshold. Thus, as indicated in [9], no assumption regarding the normality of the energy distribution is necessary. This block of the detector is similar to [4] in the energy function and to [9] in the threshold computation, but unlike all the other detectors, the MNI detector considers the threshold with respect to the detected baselines. A band-pass FIR equiripple filter was used ($f_{\text{Stop1}}=70\text{Hz}$; $f_{\text{Pass1}}=80\text{Hz}$; $f_{\text{Pass2}}=450\text{Hz}$; $f_{\text{Stop2}}=460\text{Hz}$; stopband attenuation = -60dB). The signal was filtered forward and backwards to obtain zero-phase. All processing was implemented based on MATLAB. All values were obtained from the optimization described below.

Therefore, instead of considering the energy with respect to the complete EEG (as in [4, 8, 9]), the EEG is divided in segments where 2 seconds of baseline are found (e.g.: if in the first 12 seconds of data, there are 2 seconds of detected baselines, this 12 seconds are considered as an epoch and the energy threshold of the epoch is unique). In this way, the local characteristics of the baseline around the point of interest are considered in a similar way than it is considered when visually marking HFOs.

E. Performance Metrics

Receiver Operating Characteristic (ROC) curves measure the performance of the detector when varying the energy threshold, representing different levels of sensitivity/specificity. The visually marked HFOs were considered as the true positive (P) events and the visually marked baselines as the true negative (N) events. Segments of EEG where nothing was visually marked by the reviewers were not taken into account, under the assumption that there might be HFOs in those segments, which were not visible to the human reviewer. The average ROC was computed by computing the ROC curve for each channel and averaging each threshold value across channels [12].

Thus, a TP was an HFO identified by the automatic detector that was visually marked; a FP was an automatically detected event that actually corresponded to a visually marked baseline; a FN was a visually marked HFO that was missed by the automatic detector; and a TN was a visually marked baseline where no HFOs were identified.

In addition, Cohen's Kappa coefficient [13] was computed to compare the automatic detector with the gold standard in each channel. This coefficient measures the degree of agreement between two reviewers (in this case one human and one automatic), taking into account the agreements made by chance. $Kappa < 0$ reflects an agreement due purely to chance and $Kappa = 1$ indicates complete agreement.

Not only are the percentages of detected HFOs important, but also whether the order of the channels with respect to the relative number of events is preserved, since channels with a high rate of HFOs have been associated with the SOZ [7] and their removal with good surgical outcome [14]. The Ranking Distance [15] measures the cost of obtaining a different order of channels (ordered from more to fewer events) when detecting events automatically compared to the order obtained when visually identified events. For instance, if channel A has 3 visually detected events, channel B 25 and channel C 5, we expect the automatic detector to detect more events in channel B than in the other two. RKD score ranges from 0 to 1, with 0 representing exactly the same ranking, and 1 corresponding to a ranking completely upside down. An RKD score smaller than 0.2 is usually considered as a measure of good agreement. Thus, even though we accept a large number of false positives under the assumption that there might be events not seen by the reviewer, the relative number of false positives to true positives must be equivalent in all the channels. For this measure, the automatically detected events identified in the whole EEG are considered.

F. Parameter Optimization

The parameters of the baseline detector were optimized on 36 channels with the following criteria: maximize the duration of detected baseline that overlaps with the reference, constrained by global (median across channels) $FPR = 0.05$ ($FPR = 1 - \text{Specificity}$). Each channel "voted" their best parameters (i.e.: the parameters to obtain the largest baseline overlap), and the set of parameters with most votes was selected.

For optimizing the parameters of the second block, the events detector, a similar approach was taken with another 36 channels. In this case, each channel "votes" with its best parameters to obtain maximum Kappa, constrained by: global $FPR = 0.1$; FPR per channel = 0.2; Sensitivity per channel = 0.8. In both cases, each channel could vote for more than one parameter set (i.e.: more than one set of parameters could yield the maximum sensitivity).

III. RESULTS

The optimized parameters of the MNI detector were: duration of each baseline segment = 75ms; wavelet entropy threshold = 0.56 of WE_{max} ; minimum duration of HFO event = 15ms; energy threshold = 95 percentile of CDF of baseline; short RMS window = 10ms; baseline length = 2s. Twenty of the 36 channels voted for the selected baseline parameters, while 15 of the 36 channels voted for the selected parameters of the HFOs detector block (and 11 channels voted for a very similar set). This set of parameters is used for the following results, unless specified.

HFO events were detected when their energy was larger than a threshold computed from automatically detected baselines. To illustrate this concept, Fig. 1 shows an example of a correctly detected baseline and HFO.

The performance of the detector (with a threshold of 95%) was: Sensitivity 96.8 $\pm$ 11.41% (median: 100%) and FPR: 4.86 $\pm$ 8.91% (median: 2.19%), averaged across channels (Fig. 2). In the automatic baseline detector a mean baseline length of 29.6sec per channel (range: 1.5 to 55.8sec) was detected.

Fig. 3 shows the average ROC curve (in black) across channels for different thresholds for the test dataset. Even though the operating point (0.95) was obtained from training data, in the test data it is situated in the right most corner. The area under the curve (AUC) was 0.99.

The average Kappa was 0.63 $\pm$ 0.28, which corresponds to a moderate agreement. The mean RKD per patient was 0.42 (range: 0.01 to 0.78). In some patients the RKD score was low, indicating good agreement with the reviewers ranking (4 patients with RKD<0.2). However, for most patients the RKD score was high, indicating poor correspondence with reference rankings.

To further ensure that the parameter selection was representative of all channels, in particular in the event detector block where less than half the channels voted for the selected parameters as their first choice, the distribution of parameters with respect to each channel's vote (which parameter/s maximized the channel's sensitivity) was analysed. Fig. 4 shows that the location of the 7 parameters with more votes in parameter space is very similar for 6. For 90% of the channels these 6 parameters provided the best kappa. Thus, for almost all channels their best parameter was located within a dense cloud, suggesting that the detector could be tuned to a range of parameters, which is a desirable property.

As a corollary to the above, when slightly changing the parameters (duration of HFO event=20ms; and short RMS window = 15ms), the performance remained almost the same (Sensitivity: 97.28 $\pm$ 11.53%; FPR: 5.41 $\pm$ 8.95%). Finally, when changing the energy function to the envelope computed with the Hilbert transform, the results are also very similar, achieving a sensitivity: 96.79 $\pm$ 12.25% (median: 100%) and FPR: 5.12 $\pm$ 9.05% (median: 2.15%). Thus, the MNI detector is robust to variability in the parameters and even in the way the energy of the signal is computed.

IV. DISCUSSION and FUTURE WORK

In this study, a new approach for HFO detection was proposed. The MNI detector first detects baseline segments, where no oscillatory activity is present, and then compares the energy of the EEG signal with that of the detected baselines. In this way, the local characteristics of the background are taken into consideration. The addition of baseline information is particularly important in those channels in which the background is not flat everywhere, such as channels where a large number of spikes are present.

The detector is suitable for all brain locations, since it detected HFOs in all patients, in temporal and NC areas, and regardless of whether the channel was in the SOZ. Moreover,

changes in parameters and energy function did not modify the performance. This is suggestive that it could generalize appropriately (with the corresponding training) to different situations, such as change in sleep stage, different contact sizes or distance to seizures. However, this remains to be tested. In agreement with other detectors, the MNI detector aims to detect as many events as possible. A second step should ensure the elimination of FP, particularly those corresponding to artefacts and non epileptogenic events. A GUI for human validation, a classification scheme or a combination of both could be incorporated to the detector.

Furthermore, HFOs were detected independently of their characteristics. A formal classification step in which HFOs could be divided into different categories and separated from sharp spikes is desirable. An ultimate goal would be the identification of pathological HFOs distinctly from physiological HFOs. However, given that a formal definition of what is pathologic does not exist [16], this is a particularly challenging task.

Only events that corresponded to visually detected HFOs or baselines were considered in this study. This strict definition of the true negative events can explain the extremely low FPR. Therefore, even though automatic detections occurred, a good proportion of the EEG data was only analyzed in relation to the RKD score, but not in relation to possible FP. Furthermore, there was a great variability in RKD across subjects probably due to FP detections in these segments of EEG. Analysis of events in these EEG segments could identify events not visible to the human reviewer or on the contrary indicate FP due to artefacts. Further investigation in this regard is required.

Given that visually marking of HFOs is highly time consuming, in order for this new biomarker of epilepsy to be clinically useful, HFOs need to be detected automatically. This study focused on the evaluation of the performance of the MNI detector with respect to clear reference events, but in order for the detector to be used as a clinical tool a (semi- or automatic) post-processing step is required to identify and remove FP detections. We hypothesize that most events detected in areas not marked by the experts would be rejected in this second step, but this remains to be tested.

Channels containing continuous high frequency activity were excluded from this study because it was not possible to identify baseline segments. However, most of these channels contain high rates of HFOs and are related to epileptogenic regions. Thus, an analysis of different types of backgrounds is necessary to establish clearer definitions for automatic HFO detection in those channels.

Finally, a systematic comparison of the performance of the MNI detector with other detectors [4, 8, 9] should be performed. Each detector was developed for a different purpose (such as detection of HFOs above 200Hz [8]), location (only temporal [4]) or electrode size (micro-electrodes in [4], micro- and macroelectrodes in [17]). Therefore, a systematic study comparing the detectors on the same data set is required.

V. CONCLUSION

In conclusion, a new approach to HFO detection was presented in which HFO events were detected when their energy was larger than a threshold computed from automatically detected baselines. The high sensitivity of the detector and robustness to variability makes it a promising tool for the study of HFOs in humans.

Acknowledgments

This work was supported by grant MOP-10189 from the Canadian Institutes of Health Research. R. Z. was supported by National Science and Engineering Research Council (NSERC) Postgraduate Scholarship (PGSD).

References

1. Niedermeyer, E., Lopes da Silva, FH. *Electroencephalography : basic principles, clinical applications, and related fields*. 5. Philadelphia: Lippincott Williams & Wilkins; 2005.
2. Bragin A, Engel J Jr, Wilson CL, Fried I, Buzsaki G. High-frequency oscillations in human brain. *Hippocampus*. 1999; 9:137–42. [PubMed: 10226774]
3. Bragin A, Engel J Jr, Wilson CL, Fried I, Mathern GW. Hippocampal and entorhinal cortex high-frequency oscillations (100–500 Hz) in human epileptic brain and in kainic acid-treated rats with chronic seizures. *Epilepsia*. Feb.1999 40:127–37. [PubMed: 9952257]
4. Staba RJ, Wilson CL, Bragin A, Fried I, Engel J Jr. Quantitative analysis of high-frequency oscillations (80–500 Hz) recorded in human epileptic hippocampus and entorhinal cortex. *J Neurophysiol*. Oct.2002 88:1743–52. [PubMed: 12364503]
5. Jirsch JD, Urrestarazu E, LeVan P, Olivier A, Dubeau F, Gotman J. High-frequency oscillations during human focal seizures. *Brain*. Jun.2006 129:1593–608. [PubMed: 16632553]
6. Urrestarazu E, Chander R, Dubeau F, Gotman J. Interictal high-frequency oscillations (100–500 Hz) in the intracerebral EEG of epileptic patients. *Brain*. Sep.2007 130:2354–66. [PubMed: 17626037]
7. Jacobs J, LeVan P, Chander R, Hall J, Dubeau F, Gotman J. Interictal high-frequency oscillations (80–500 Hz) are an indicator of seizure onset areas independent of spikes in the human epileptic brain. *Epilepsia*. Nov.2008 49:1893–907. [PubMed: 18479382]
8. Crepon B, Navarro V, Hasboun D, Clemenceau S, Martinerie J, Baulac M, Adam C, Le Van Quyen M. Mapping interictal oscillations greater than 200 Hz recorded with intracranial macroelectrodes in human epilepsy. *Brain*. Jan.133:33–45.
9. Gardner AB, Worrell GA, Marsh E, Dlugos D, Litt B. Human and automated detection of high-frequency oscillations in clinical intracranial EEG recordings. *Clin Neurophysiol*. May.2007 118:1134–43. [PubMed: 17382583]
10. Chander, R. M Eng Thesis. 2007. Algorithms to detect high frequency oscillations in human intracerebral EEG; p. viip. 123
11. Rosso OA, Blanco S, Yordanova J, Kolev V, Figliola A, Schurmann M, Basar E. Wavelet entropy: a new tool for analysis of short duration brain electrical signals. *J Neurosci Methods*. Jan 30.2001 105:65–75. [PubMed: 11166367]
12. Fawcett T. An introduction to ROC analysis. *Pattern Recognition Letters*. Jun.2006 27
13. Cohen J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*. 1960; 20:37–46.
14. Jacobs J, Zijlmans M, Zelmann R, Chatillon CE, Hall J, Olivier A, Dubeau F, Gotman J. High-frequency electroencephalographic oscillations correlate with outcome of epilepsy surgery. *Ann Neurol*. Sep 4.2009 67:209–220.
15. Zelmann R, Zijlmans M, Jacobs J, Chatillon CE, Gotman J. Improving the identification of High Frequency Oscillations. *Clin Neurophysiol*. Aug.2009 120:1457–64. [PubMed: 19576848]
16. Engel J Jr, Bragin A, Staba R, Mody I. High-frequency oscillations: what is normal and what is not? *Epilepsia*. Apr.2009 50:598–604. [PubMed: 19055491]

17. Worrell GA, Gardner AB, Stead SM, Hu S, Goerss S, Cascino GJ, Meyer FB, Marsh R, Litt B. High-frequency oscillations in human temporal lobe: simultaneous microwire and clinical macroelectrode recordings. *Brain*. Apr.2008 131:928–37. [PubMed: 18263625]

Fig. 1.

Example of detected HFO event surrounded by detected baselines. Top: unfiltered EEG segment; bottom: filtered (80–450Hz) EEG. The concordance between automatically detected (AUTO) HFOs and visually identified (reference) HFOs is excellent.

Fig. 2.

Boxplot of Sensitivity and FPR (equivalent to 1-Specificity) when automatically detecting HFOs in 278 channels. Black line: median of the distribution; boxes: 0.25 and 0.75 quantiles; crosses: outliers.

Fig. 3. ROC showing the average performance of the detector. Black points: analyzed thresholds (only some threshold values are shown); Gray lines: 5 and 95 quantiles of ROC of individual channels.

Fig. 4.

Scatter plot of set of most voted parameters in parameter space. Size of the bubble corresponds to the number of channels that voted for this parameter. Most parameters are within a clustered cloud. Boxes show the corresponding set of parameters for the two parameters with most votes. One dimension of parameter space (baseline length) is not shown since all parameters had a value of 2sec.