Feature Parameter Optimization for Seizure Detection/Prediction

R. Esteller* #, J. Echauz #, M. D'Alessandro‡, G. Vachtsevanos‡ and B. Litt†,.

IntelliMedix, Atlanta, USA
* Universidad Simón Bolívar, Caracas, VENEZUELA.
‡ Georgia Institute of Technology, Atlanta, USA
† University of Pennsylvania, Philadelphia, USA

Abstract-**When dealing with seizure detection/prediction problems, there are three main performance metrics that must be optimized: false positive rate, false negative rate, detection delay or, if the problem is seizure prediction, it is desirable to obtain the greatest prediction time achievable. Tuning specific extracted features to individual patients can lead to improved results. The processing window length is also an important parameter whose optimization may significantly affect performance. In this study we propose an approach for selecting the window length for the particular detection/prediction problem. This approach is applicable to other feature parameters suitable for tuning or optimization**.

I. INTRODUCTION

Even though there are differences between seizure detection and prediction, there are also some similarities in the methodology used to approach each one as well as some common issues, such as feature extraction and two-state classification. While the two classes in seizure detection are seizure onset and non-seizure onset, in seizure prediction the two classes are preseizure (preictal) and non-preseizure (not preictal).

Some of the first attempts to detect seizures were accomplished during the seventies by Prior et al. [1] and Ives et al. [2]. These investigators intended to identify tonic-clonic and other conspicuous seizures, respectively. Both aimed to detect the seizures at any time during their evolution; without regard to detection delay. In addition, no attempt was made to tune prediction parameters to individual patients. In a similar fashion, Murro et al. [3] and Harding [4] performed similar work, but generalized to all patients. In the late nineties Qu and Gotman [5] proposed a seizure-onset detector introducing the idea of tuning quantitative features to individual subjects.

In this study a methodology for tuning the window length or any other feature parameter is proposed, and analyzed for the particular problem of seizure onset detection. Section II describes the problem and general background, Section III explains an optimization methodology, and presents the results, and Section IV provides the discussion and conclusions of this study.

II. GENERAL BACKGROUND

For most detection/prediction problems, the running/sliding window method is the technique used to extract features from continuous data. Feature extraction is performed through a running window, as sketched in Figure 1. The shaded area is the sliding observation window, which moves through the data as features are computed. The data points inside this sliding window are used for the feature generation as the window moves through the data.

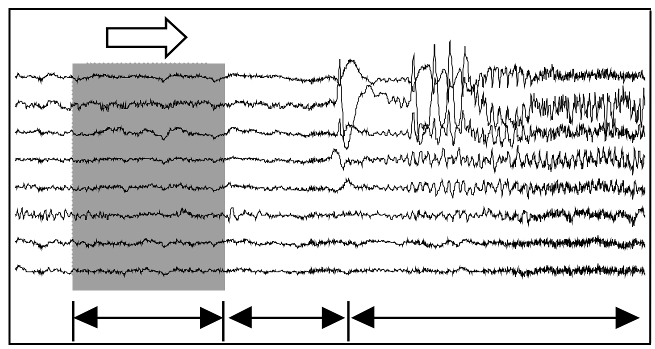

Figure 1: Running window technique

Therefore, this observation window is continually collapsed into a feature vector by means of formulas and algorithms that take preprocessed EEG epochs as inputs and produce scalar quantities as outputs, which then become the components of the feature vector. Two levels of features can be defined: *instantaneous features* and *historical features*, which are sketched in Figure 2.

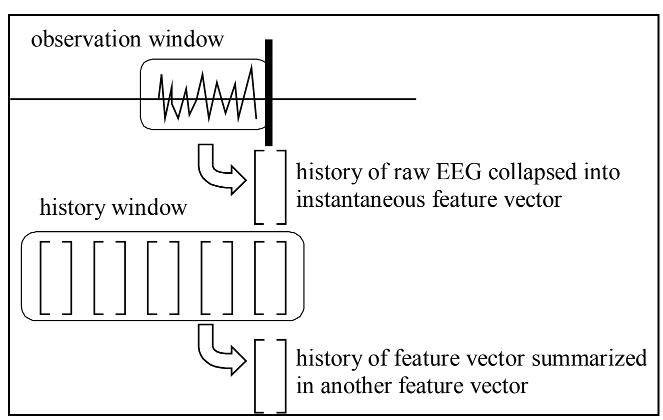

Figure 2: Types of Features

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
25 Oct 2001	N/A	-

Title and Subtitle Feature Parameter Optimization for Seizure Detection/Prediction	**Contract Number**
	Grant Number
	Program Element Number
Author(s)	**Project Number**
	Task Number
	Work Unit Number
Performing Organization Name(s) and Address(es) Universidad Simon Bolivar Caracas, Venezuela	**Performing Organization Report Number**
Sponsoring/Monitoring Agency Name(s) and Address(es) US Army Research, Development & Standardization Group (UK) PSC 802 Box 15 FPO AE 09499-1500	**Sponsor/Monitor's Acronym(s)**
	Sponsor/Monitor's Report Number(s)
Distribution/Availability Statement Approved for public release, distribution unlimited	
Supplementary Notes Papers from 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference cd-rom., The original document contains color images.	
Abstract	
Subject Terms	
Report Classification unclassified	**Classification of this page** unclassified
Classification of Abstract unclassified	**Limitation of Abstract** UU
Number of Pages 4	

Instantaneous features are computed directly from the original signal (IEEG data) through a running observation window. *Historical features* are "features of features" that require a second, third or higher level of feature extraction, which entails the evolution of the *history* of features through time. Over a large set of instantaneous and historical features extracted (candidate features), feature parameter optimization takes place.

Several factors are taken into account when determining the window length to be used in the analysis. Among them, data stationarity, data length required to compute the features, sampling frequency, maximizing the distinguishability between epochs containing seizures and those that do not and between epochs that are preictal and those that are not preictal, if seizure prediction is the problem, and minimizing detection delay. A compromise has to be achieved between the requirement that a data window be sufficiently long to compute specific IEEG (intracranial EEG) features and that it be short enough to assume data stationarity. An IEEG segment of tens of seconds can be considered quasi-stationary, depending on the patient's behavioral state [6]-[7].

III. Window Optimization Methodology and Results

An original methodology for selecting processing window size is proposed in this study. This methodology arises as an answer to the issues of how to effectively select the window size to compute specific features, and how to create the feature vector when the features extracted require data sets of different lengths. These questions emerged during the development of the feature extraction stage of a broader problem of seizure onset detection and prediction [8]-[9]. This optimization methodology can be accomplished in two ways. In one case, when the classifier to be used in the detection/prediction system is known a priori, the objective function used in the optimization can be any combination of false positives (FPs), false negatives (FNs), and detection delays or prediction times obtained after the classifier output. In the other case, when the classifier has not been determined yet, an objective function aimed at maximizing the class separability is used. In this study, the second optimization option was ultilized, therefore, the goal of the optimization was to maximize the distinguishability between the seizure-onset and no seizure-onset classes, or preictal and no-preictal classes for the prediction case. The scheme of Fig. 3 summarizes the procedure. In this scheme, each of the selected features is computed for different sliding window sizes.

Specifically, in the present analysis 90 different window sizes were selected within the range of 50 points (0.25 seconds) to 9000 points (45 seconds). This window range was selected to include the maximum window size to satisfy quasi-stationarity of the data segments [7] [11] and the minimum window size required to compute the feature [6]. All these windows were shifted 90 points (0.45 seconds) along the IEEG sequence, while the running window method described earlier was used to generate the features. These 90-point shifts fix the maximum delay in the onset detection (d in Figure 5) to 0.45 seconds, assuming features capable of detecting the seizure onset as soon as one sample of the ictal IEEG is within the sliding window. There is also a trade-off between this maximum detection delay for features capable of detecting the onset as soon as one ictal sample goes into the sliding window, and storage capacity of the system. The shorter this detection delay or the smaller the window shifting, the greater the memory space required.

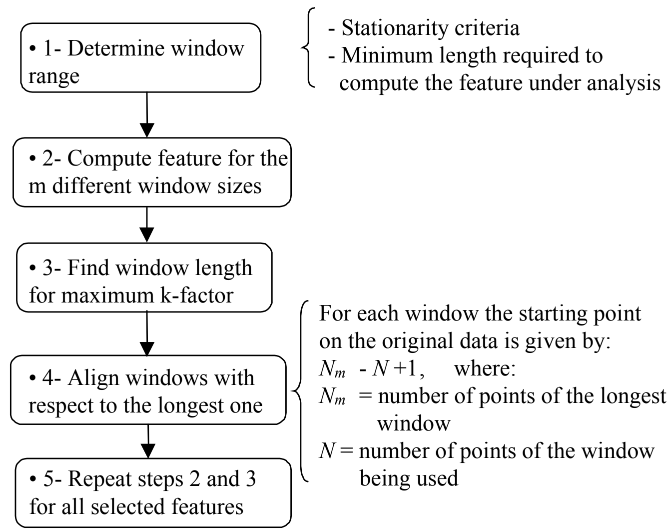

Figure 3: Window Size Selection for Maximum Distinguishability between Classes

After each feature is computed for different windows, the *k*-factor given in (1) is computed as a measure of effectiveness of each feature.

$$K = \frac{|\mu_1 - \mu_2|}{\sqrt{(\sigma_1^2 + \sigma_2^2)/2}}, \tag{1}$$

where
K is the *k*-factor (measure of effectiveness of the feature),
μ_i is the mean of feature for class *i*,
σ_i^2 is the variance of feature for class *i*.

For each seizure record, the window size corresponding to the maximum *k*-factor was chosen to preceed the analysis. Then, a visual verification followed to confirm that the window lengths that maximize the *k*-factor in each record clustered around some value. This mean value was chosen as the window length for the feature under consideration. Figure 4 illustrates the variation of the *k*-factor for the fractal dimension feature, as the window size is changed for four different seizure records. Note the so-called "optimal" window length within approximately 1000 and 1500 points.

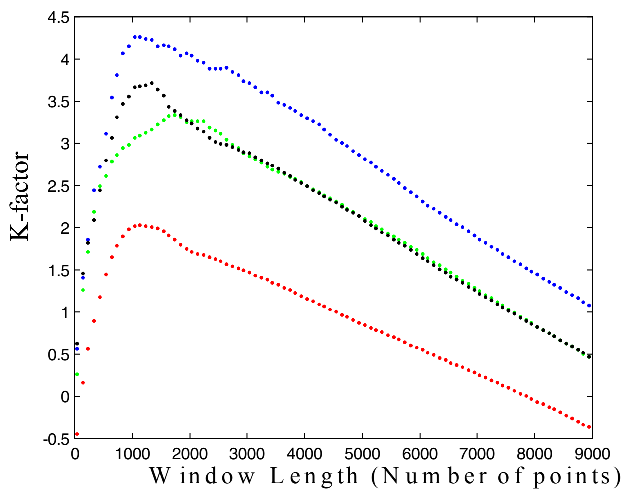

Figure 4: *K*-Factor from the Fractal Dimension for Different Window Sizes

Typically, the window sizes that maximized the *k*-factor were different for each feature. Therefore, a strategy was required to allow the creation of feature vectors from features extracted with different sliding window sizes, which implies that the features do not coincide in time and have different time spans between consecutive values. To have a perfect time alignment and identical time span across features, two conditions must be satisfied. The first condition guarantees the same time span for consecutive values on all the features. This was achieved by making the observation window displacement equal for all the window sizes on all the features. The second condition requires the alignment of all the observation windows with respect to the right border of the longest window, as shown in Figure 5. The effect of applying equal displacement of the observation window even for features with different window sizes is that the number of overlapping points on each observation window will change from feature to feature, while the shifting points remain constant.

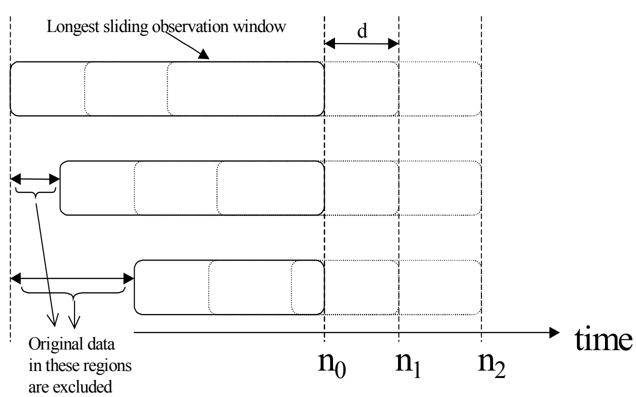

Figure 5: Time Alignment and Time Span for Different Window Sizes

Using this approach, historical and instantaneous features can be combined by extracting historical features from the instantaneous feature utilizing a shift of one-feature-sample for the observation window and a time alignment of the historical features. Intuitively, this type of approach could outperform those that rely only on instantaneous features.

The window length of the best three features of 6 patients were selected following the procedure described. The results for the other features always presented a global maximum like in the feature on Fig. 4, and the window lengths ranged from 700 points for the to 6000 points for the in patient .

IV. Discussion and Conclusion

A data driven methodology for window length selection and/or feature parameter selection in detection/prediction classification problems has been proposed and studied. The results obtained showed that in all features and patients studied there exists a global maximum for the *k*-factor corresponding to an "optimal" window length that maximizes class separability. Class separability was determined as a measure proportional to the distance between the mean of each class and inversely proportional to the average variance of both classes (k-factor), however any other class separability measure or objective function that suits the particular problem goals can be used. The methodology applied to the specific problem of seizure-onset detection can be used as well for seizure prediction problems, and/or for any other detection/prediction problem. In addition, this methodology can be extended to "optimize" other feature parameters related to each particular feature, such as a scale factor, the type of window used (Hanning, Hamming, Barlett, rectangular, etc.), etc.

Further research is required to study this methodology in the case when the classifier is used as part of the feature parameter evaluation and the objective function is computed directly from the classifier output, rather than from the feature values directly. It is also important, to analyze the behavior of the method with other objective functions and when dealing with other feature parameters.

Acknowledgments

This work was supported by the Epilepsy Foundation, The American Epilepsy Society, The Whitaker Foundation, and The National Institutes of Health grant #MH-62298RO1.

References

[1] P. F. Prior, R. S. M. Virden, and D. E. Maynard, "An EEG device for monitoring seizure discharges," *Epilepsia*, vol. 14, pp. 367-372, 1973.

[2] J. R. Ives, "The on-line computer detection and recording of spontaneous temporal lobe epileptic seizures from patients with implanted depth electrodes via a radio

telemetry link," *Electroencephalography and Clinical Neurophysiology*, vol. 37, pp. 205, 1974.

[3] A. M. Murro, et al. "Computerized seizure detection of complex partial seizures," *Electroencephalography and Clinical Neurophysiology*, vol. 19, pp. 330-333, 1991.

[4] G. W. Harding, "An automated seizure monitoring system for patients with indwelling recordings electrodes," *Electroencephagraphy and Clinical Neurophysiology*, vol. 86, no. 6, pp. 428-437, 1993.

[5] 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 Trans. on Biomed. Eng.*, vol. 44, no. February No. 2, pp. 115-122, 1997.

[6] K. Lehnertz and C. E. Elger, "Spatio-temporal dynamics of the primary epileptogenic area in temporal lobe epilepsy characterized by neural complexity loss," *Electroencephalography and Clinical Neurophysiology*, vol. 95, pp. 108-117, 1995.

[7] F. H. Lopes da Silva, "EEG analysis: theory and practice," in *Electroencephalography, Basic Principles, Clinical Applications and Related Fields*. Baltimore, MD: Urban ans Schwarzenberg, 1987, pp. 871-897.

[8] Esteller R., Ph.D dissertation "Detection of seizure onset in epileptic patients from intracranial EEG signals", Georgia Institute of Technology, April 2000.

[9] B. Litt, G. Vachtsevanos, J. Echauz, and R. Esteller, "Method and apparatus for predicting the onset of seizures based on features derived from brain activity signals," *provisional U.S. patent application, attorney docket no. 05010.0067*, August 1998.

[10] R. Esteller, B. Litt, J. Echauz, G. Vachtsevanos, "Adaptive method and apparatus for forecasting and controlling neurological disturbances under a multi-level control," *provisional U.S. patent application, attorney docket no. I080 1020*, November 2000.

[11] D. Qin, "A comparison of techniques for the prediction of epileptic seizures," in *Proc. of 8th IEEE Symposium on Computer-Based Medical Systems*. Lubbock, Texas, 1995, pp. 151-157.