MORPHOLOGICAL CLUSTERING OF INTERICTAL EPILEPTIFORM DISCHARGES IN INTRACRANIAL ELECTROENCEPHALOGRAPHY

P. Vlk[1,2], R. Janca[1], R. Cmejla[1], P. Krsek[3], P. Marusic[3], P. Jiruska[2,3]

[1]Faculty of Electrical Engineering, CTU in Prague
[2]Institute of Physiology, Academy of Science of the Czech Republic
[3]2nd Faculty of Medicine, Charles University in Prague

Abstract

Selected patients with refractory epilepsy can benefit from surgical treatment. The main purpose of presurgical examination is to identify and delineate epileptogenic areas of the brain which should be removed. Epileptogenic areas are determined according to the spatial distribution of seizure onsets, interictal epileptiform discharges or high-frequency oscillations. Specificity of interictal epileptiform discharges to mark epileptogenic tissue is decreased by the fact, that they are also observed outside the epileptogenic areas. To improve the localizing yield of interictal discharges, identification of specific features of the discharges generated only within epileptogenic region is required. The main aim of this project was to develop self-clustering algorithm which will discriminate distinct populations of interictal epileptiform discharges according to the morphology of their waveforms. First step of the developed algorithm extracts nine basic morphological features of each interictal epileptiform discharge detected in band-pass filtered (2-60 Hz) intracranial recordings. Principal component analysis is applied on extracted features to reduce their dimension. Only the first principle components with cumulative variance of 80 % or above are used for clustering. Gaussian Mixture Distribution method is utilized to assign each discharge to appropriate morphological cluster. Results of the clustering algorithm are displayed in the form of cortical maps together with medians of the clustered discharge waveform. Developed algorithm was tested in the model of intracranial EEG signal and on data recorded in patients who underwent intracranial monitoring. Results demonstrate the ability of the algorithm to separate interictal epileptiform discharges according to their morphological features.

1 Introduction

Epilepsy affects approximately 0.5-1 % of population in developed countries and in one third of patients it becomes refractory to antiepileptic drugs. Selected patients with refractory epilepsy can benefit from surgical treatment. Principle of epilepsy surgery is to remove epileptogenic brain areas which are involved in seizure genesis. Currently, we lack parameter which would reliably identify epileptogenic brain tissue. Therefore, resection margins are determined only indirectly based on the information about location of epileptogenic lesion and spatial distribution of epileptiform electrographic phenomena. Interictal epileptiform discharges (IEDs) represent electrographic phenomenon generated in epileptic brain. Their spatial distribution often overlaps with areas of endogenous epileptogenicity. Ability to identify specific features of IEDs generated by epileptogenic brain would improve localization of the brain areas which should be included in the resection to achieve seizure freedom.

The main aim of this project was design and implementation of unsupervised offline algorithm which separates IEDs into clusters according to the morphological features of their waveforms.

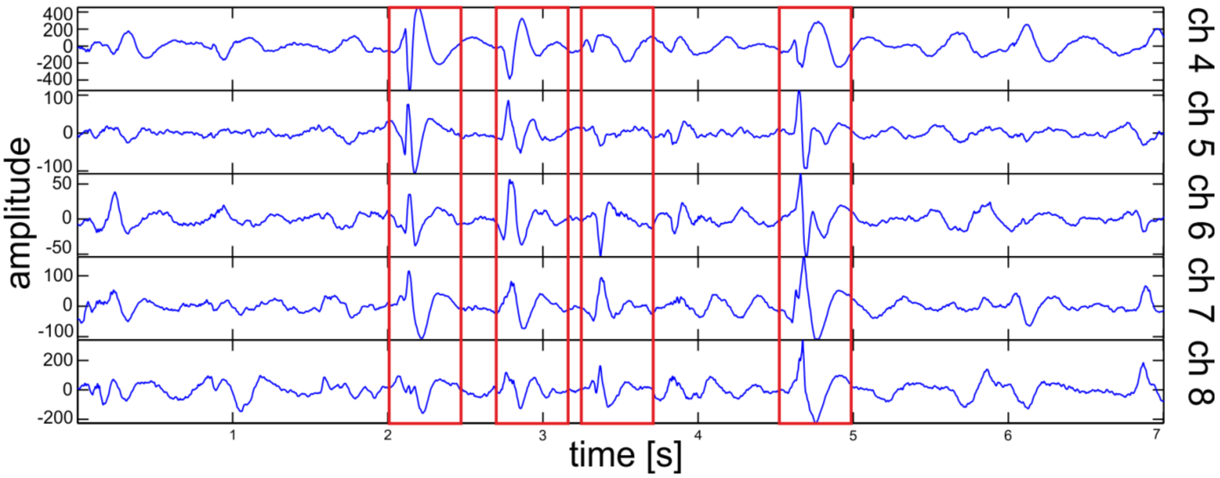

Figure 1: Example of intracranial recordings which contain IEDs. One of the main features of IEDs is their ability to they propagate through the brain. This example shows four IED events (red box) with various spatial pattern of propagation. Note also variation of the morphology of each IED waveform.

2 Methods

2.1 Data acquisition

Data were recorded in patients with refractory epilepsy who underwent invasive exploration at University Hospital Motol in Prague. Signals from subdural and/or depth macroelectrodes were amplified, filtered using aliasing filter at 1/3 of sampling frequency and sampled at frequency 1000 Hz. Data were recorded in reference mode. For the purpose of the clustering they were converted to bipolar mode.

2.2 IEDs detection and signal pre-processing

To automatically detect IEDs we used Hilbert transform detector [1] which was developed in the previous project. Signals were band-pass filtered (2-60 Hz) to preserve signal with spectral information which corresponds to spectral composition of IED. Filtering procedure included following steps: high-pass (2 Hz), biquad notch (50 Hz) and low pass (60 Hz) filtration. High pass filtration (> 2 Hz) involved two-step resampling to achieve sharp cut-off frequency characteristics of the filter. Signal was decimated to 10 Hz using Matlab *resample* function (length of FIR filter was proportional to half of the original sampling frequency). Then Chebyshev Type II low pass filter of 10^{th} order with cutoff frequency of 2 Hz (+ 0,5 Hz crossband with attenuation of 70 dB) was applied. Using this approach, we obtained low-resolution isoline (low frequency interference) of the processed signal. The isoline was then interpolated to default sampling frequency of the original one. During the last step isoline was subtracted from original signal. Biquad filter was used to eliminate 50 Hz additive main hum and its higher harmonic frequencies. Poles of the biquad filter lied on a radius of 0.98 and zeros on a unit circle of Z-plane. Chebyshev Type II low pass filter with cutoff frequency 60 Hz (+ 5 Hz crossband with attenuation of 70 dB with maximum permissible passband loss 5 dB) was used as a low pass filter. Phase delay introduced by application of above mentioned filters was compensated by zero-phase digital filtering.

2.3 Feature extraction

Segments containing detected IED were extracted from the pre-processed signals. Segment size was 600 ms; 150 ms before and 450 ms after the time index of IED detection. Parameters of the segment size were selected so that each segment contains entire IED waveform, i.e. spike and the following wave. IED waveform can be described using myriad of features from time and/or frequency domain. We selected nine basic features and for each IED we determined values of these features.

1) Polarity

Polarity was determined from the peak of the highest absolute value of amplitude in window with dynamic width. Time index of detected event represents midpoint of each dynamic window. If no peak was found, width of the window was increased equally to both sides. For feature extraction we used IEDs with polarity normalized to positive value.

2) Mean value [2]

$$M = \frac{1}{N}\sum_{i=1}^{N}|x_i| \tag{1}$$

Where N is number of samples and x is amplitude of given sample.

3) Curve length [2, 3]

$$CL = \frac{1}{N}\sum_{i=1}^{N-1}|x_{i+1} - x_i| \tag{2}$$

4) Accumulated energy [2]

$$AE = \frac{1}{N}\sum_{i=1}^{N} x_i^2 \tag{3}$$

5) Teager's energy [2, 3]

Average nonlinear energy which represents a measure of energy proportional to both: signal amplitude and frequency.

$$TE = \frac{1}{N}\sum_{i=2}^{N-1}(x_n^2 - x_{i-1}x_{i+1}) \tag{4}$$

6) Entropy of the squared and normalized Teager's Energy [3]

Square of the signal normalized by its sum makes a pseudoprobability mass function (6). Entropy is estimated from this function (7) where n is index of processed sample.

$$TE(n) = x_n^2 - x_{n-1}x_{n+1} \tag{5}$$

$$S(n) = \frac{TE(n)^2}{\sum_{n=1}^{N-1} TE(n)^2} \tag{6}$$

$$ETE = -\sum_{n=1}^{N-1}(S(n)\, log_2\, S(n)) \tag{7}$$

7) Spectral centroid [3]

This parameter estimates frequency which corresponds to the "center of mass" of the spectrum. Where *f* is vector of frequency bands, *S* is vector of computed energy bands of same length and *fs* is sampling frequency.

$$M_1 = \frac{\sum_{f=0}^{fs/2} f \times S(f)}{\sum S} \tag{8}$$

8) Standard deviation of the spectral centroid

$$M_2 = \sqrt{\frac{\sum_{f=0}^{fs/2} f^2 \times S(f)}{\sum S} - M_1^2} \tag{9}$$

9) Global/average-local peak ratio [3]

$$PR = \frac{g^*}{\frac{1}{J-1}\sum_{\{g_i \in G | g_i < g^*\}} g_i} \tag{10}$$

$$\mathbf{G} = \{g_1, \dots, g_J\} \tag{11}$$

Vector **G** describes set of local maxima and $g*$ is global maximum. This parameter is useful to discriminate artifacts.

2.4 Clustering

Principal component analysis (PCA)

Clustering of large number of detected IEDs according to nine parameters (features) represents processing step which is computationally extremely demanding. Thus we have to reduce dimensionality of the parameters' vector space by application of PCA. Result of PCA represents matrix of the same size as the original one (columns corresponds to parameters and rows to segments). Parameters are substituted with principal components which are ranked according to their variance. Only the principle components with cumulative variance of 80 % or above were used for subsequent clustering. Example of eigenvectors in space is shown in Figure 2.

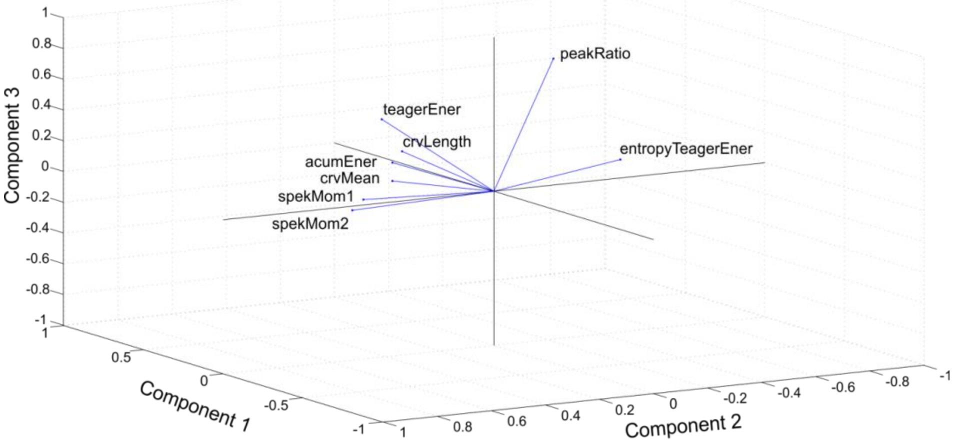

Figure 2: Example of PCA eigenvectors

Gaussian Mixture Distribution (GMD)

GMD is a multivariate distribution that consists of a mixture of one or more multivariate Gaussian distribution components (clusters). This clustering process uses an EM (Expectation Maximization) algorithm to obtain maximum likelihood estimates of each Gaussian mixture distribution component. The number of components for a given GMD is fixed and it must be determined in advance. Unfortunately, there is no established method how to *a priori* estimate number of components. Therefore, for each set of signals we calculated distribution models with the number of components ranging from one to ten [4].

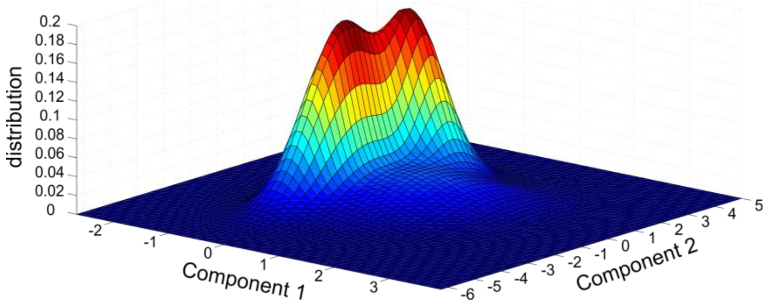

Figure 3: Example of GMD with four clusters of first two principal components

To identify optimal cluster number, we applied on the generated distribution models three different methods of information criterion estimation: methods of Calinski-Harabasz [5], Krzanowski-Lai index [6] and Hartigan's method [7]. Unfortunately, these methods do nto provide consistent results and optimal number of clusters had to be visually verified. Only clusters that contain more than 10 % of IEDs in the largest cluster were considered as significant [8]. These clusters represent output (result) of the entire clustering procedure. Schematics of the clustering algorithm is shown in Figure 4.

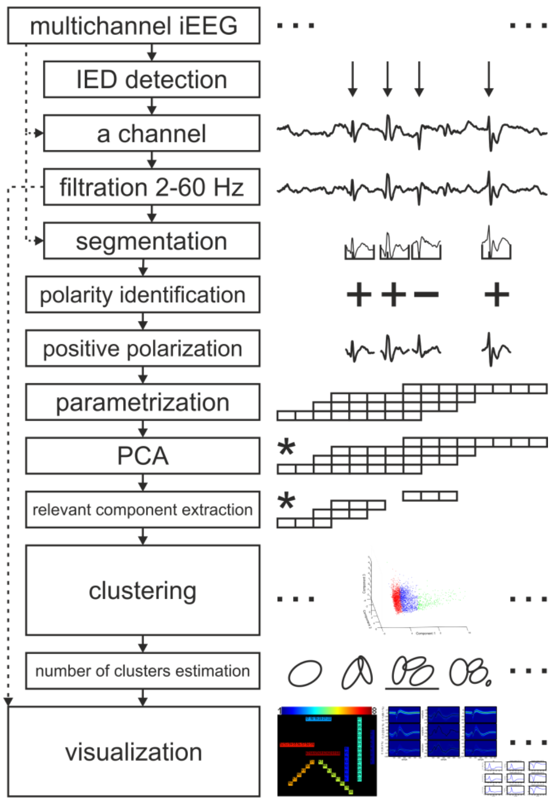

Figure 4: Schematics of proposed clustering algorithm

3 Results

The clustering algorithm was applied to datasets obtained from two patients who underwent invasive exploration as a part of their presurgical examination. Patient A is five years old boy with intractable frontal lobe epilepsy due to tuberous sclerosis complex and patient B is eleven years old boy with multi-focal epilepsy due to meningoencephalitis. Results of the clustering were incorporated into the cortical maps, which contain information about position of implanted electrodes. These diagrams facilitate clinical interpretation of the clustering results together with other clinical data. Example of cortical maps of patient B and its rewritten form is in Figure 5.

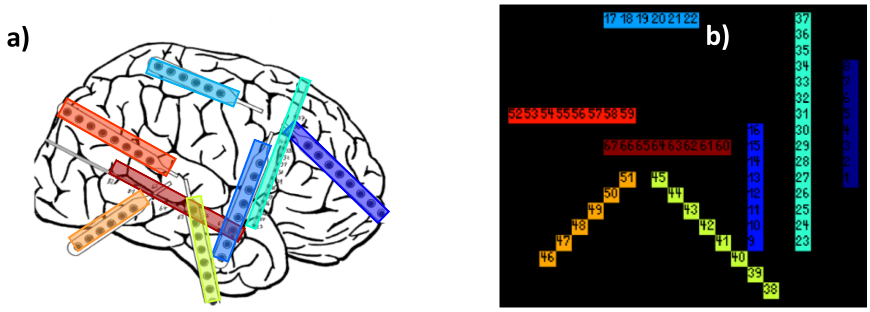

Figure 5: Schematics of electrode placement (a) in patient B and its computerized version (b).

Designed algorithm was able to identify multiple morphological clusters of IEDs in both patients. Three methods of optimal cluster number estimation were tested. Results showed high variability of estimated optimal cluster number Average number of clusters was 5.7±2.6 (median 5). Calinski-Harabasz methods (three clusters in case A, four clusters in case B) provided the best estimates if compared with visual assessment. One cluster in patient A and two clusters in patient B contained artifact and they were visually excluded.

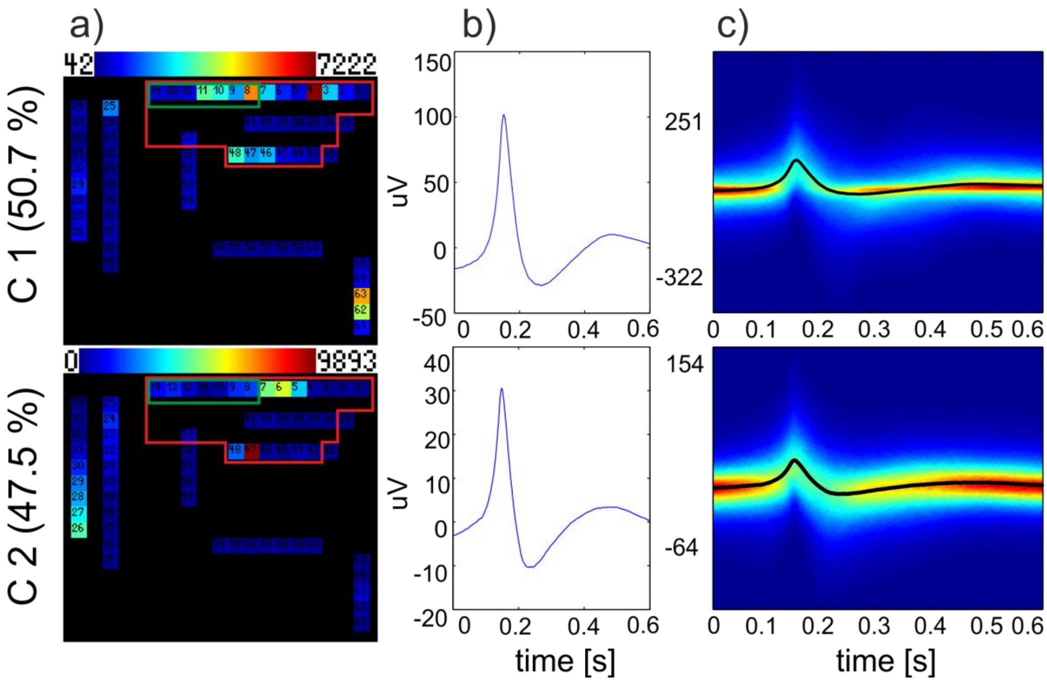

Figure 6: Clustering algorithm applied in patient A. Result demonstrated existence of two relevant clusters C1 and C2 that contain 50.7% and 47.5 % of IEDs respectively. Number of IED in each electrode is visualized in colored cortical maps (a). Red contour represent surgically removed area, green rectangle covers seizure onset zone marked by neurologists. Median waveform of all IEDs from each cluster (b) (64361 realizations of C1 and 60343 of C2). 3D-histograms (c) show variability of clustered IEDs.

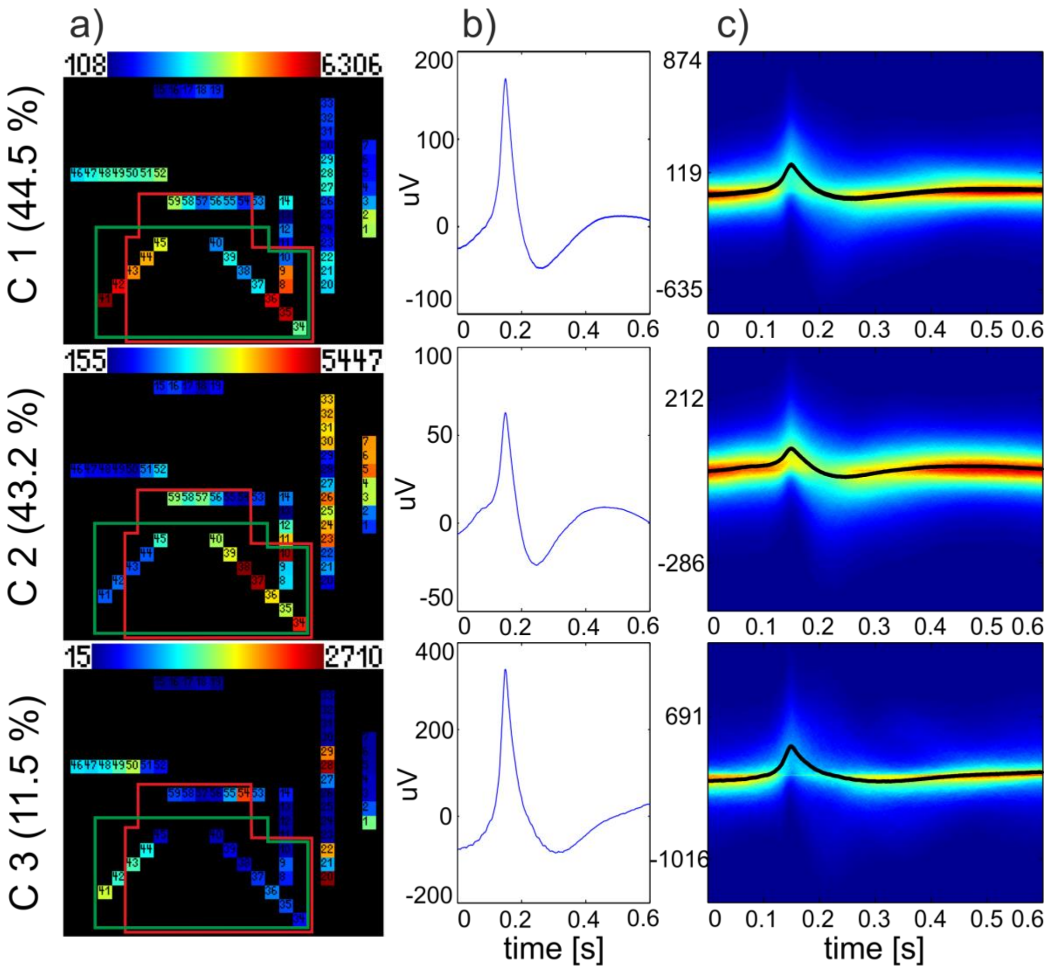

Figure 7: Clustering algorithm applied in patient B. Result demonstrated existence of three relevant clusters C1, C2 and C3 that contain 44.5 %, 43.2 % and 11.5 % of IEDs respectively. Number of IED in each electrode is visualized in colored cortical maps (a). Red contour represent surgically removed area, green rectangle covers seizure onset zone marked by neurologists. Median waveform of all IEDs from each cluster (b) (138430 realizations of C1, 134610 of C2 and 35910 of C3). 3D histograms (c) show variability of clustered IEDs.

4 Discussion

Developed algorithm is able to identify and separate IEDs with different waveforms. Clinical significance of the morphological clustering needs to be determined in future. Our results suggest that morphological clustering can provide additional information about functional organization of the brain areas generating interictal discharges. Previous study showed that IEDs can be reliably clustered according to their spatial propagation profile [1]. This method is able to identify distinct clusters of IEDs and regions of the brain from which they originate. However, spatial clustering technique fails to discriminate clusters with spatial overlap. Morphological clustering can provide tool how to separate spatially overlapping clusters according to the morphology.

Morphological clustering technique assumes that intracerebral propagation of interictal discharges through has linear filter properties (linear transfer function) and does not result in alteration of morphology of propagating IEDs. It has been shown that linear methods can be successfully applied in studies focused on the transmission of the signals through neural networks. However, there is also evidence that propagation pathways have nonlinear features. To determine properties of IEDs propagation and quantification of its transfer function will require experimental verification.

Designed method separates epileptiform activity according to the parameters of their waveforms. It has a potential to be used for unsupervised spike sorting tool. Purpose of spike sorting is to group spikes (action potentials) into clusters based on the similarity of their shapes, when each cluster correspond to action potential firing from single neuron [9].

5 Acknowledgement

This work has been supported by the Ministry of Health of Czech Republic grants (IGA MZ CR NT/11460-4, IGA NT/14489-3, IGA MZ CR NT13357-4), by Karel Janecek Endowment in Research and Science grant (NFKJ 001/2012) and University Hospital Motol grant MH CZ – DRO 00064203 and Student Grant SGS 10/272/OHK4/3T/13.

6 References

[1] JANCA, R., P. JEZDIK, R. CMEJLA, P. KRSEK, J.G.R. JEFFERYS, P. MARUSIC a P. JIRUSKA. Automatic detection and spatial clustering of interictal discharges in invasive recordings. 2013 IEEE International Symposium on Medical Measurements and Applications (MeMeA). IEEE, 2013, s. 219-223. DOI: 10.1109/MeMeA.2013.6549739.

[2] PAN, Yaozhang, Shuzhi Sam GE, Feng Ru TANG a Abdullah AL MAMUN. Detection of Epileptic Spike-Wave Discharges Using SVM. 2007 IEEE International Conference on Control Applications. IEEE, 2007, s. 467-472. DOI: 10.1109/CCA.2007.4389275.

[3] BLANCO, J. A., M. STEAD, A. KRIEGER, J. VIVENTI, W. R. MARSH, K. H. LEE, G. A. WORRELL a B. LITT. Unsupervised Classification of High-Frequency Oscillations in Human Neocortical Epilepsy and Control Patients. Journal of Neurophysiology. 2010-11-02, vol. 104, issue 5, s. 2900-2912. DOI: 10.1152/jn.01082.2009.

[4] TIBSHIRANI, Robert, Guenther WALTHER a Trevor HASTIE. Estimating the number of clusters in a data set via the gap statistic. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 2001, vol. 63, issue 2, s. 411-423. DOI: 10.1111/1467-9868.00293.

[5] CALINSKI, T. a J. HARABASZ. A dendrite method for cluster analysis. Communications in Statistics - Theory and Methods. 1974, vol. 3, issue 1, s. 1-27. DOI: 10.1080/03610927408827101.

[6] KRZANOWSKI, W. J. a Y. T. LAI. A Criterion for Determining the Number of Groups in a Data Set Using Sum-of-Squares Clustering. Biometrics. 1988, vol. 44, issue 1, s. 23-. DOI: 10.2307/2531893.

[7] HARTIGAN, John A. Clustering algorithms. New York: Wiley, 1975, xiii, 351 p. ISBN 04-713-5645-X.

[8] YADAV, R., A. K. SHAH, J. A. LOEB, M. N. S. SWAMY a R. AGARWAL. A novel unsupervised spike sorting algorithm for intracranial EEG. 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2011, s. 7545-7548. DOI: 10.1109/IEMBS.2011.6091860.

[9] FRANKE, Felix, Michal NATORA, Clemens BOUCSEIN, Matthias H. J. MUNK a Klaus OBERMAYER. An online spike detection and spike classification algorithm capable of instantaneous resolution of overlapping spikes. Journal of Computational Neuroscience. 2010, vol. 29, 1-2, s. 127-148. DOI: 10.1007/s10827-009-0163-5.