DETECTING EVENTS IN VIDEO SEQUENCE OF VIDEO-EEG MONITORING

D. Murashov [1, *], Yu. Obukhov [2], I. Kershner [2], M. Sinkin [3]

[1] Federal Research Center "Computer Science and Control" of RAS, 119333, Moscow, Russia - d_murashov@mail.ru
[2] Kotel'nikov Institute of Radio Engineering and Electronics of RAS, 125009, Moscow, Russia – (yuvobukhov, ivan_kershner)@mail.ru
[3] N.V. Sklifosovsky Research Institute for Emergency Medicine of Moscow Healthcare Department, 129090, Moscow, Russia - msinkin@mail.ru

Commission II, WG II/5

KEY WORDS: Video-Electroencephalographic Monitoring, Video Sequence, Diagnostic Event, Optical Flow, Epileptic Seizure

ABSTRACT:

In this paper, an algorithm for automated detecting diagnostic events in video channel of video and electroencephalographic (EEG) monitoring data is presented. The analysis of video sequences is focused on identifying a group of frames with high or very low (depending on the type of seizure) dynamics of informative areas according to a criterion calculated during processing of the optical flow. The preliminary results of the analysis of real clinical data are given and compared with data obtained from the synchronous EEG. The results showed the possibility in principle of reliable diagnosing epileptic seizures and distinguishing them from non-epileptic events.

1. INTRODUCTION

This paper deals with the problem for automated detecting diagnostic events in video channel of video and electroencephalographic monitoring data is presented. Video-electroencephalographic (Video-EEG) monitoring is a method of long - term synchronous registration of electroencephalography (EEG) and video image. Simultaneous video recording of the clinical condition of the patient and the bioelectric activity of the brain (i.e. EEG) allows one to diagnose epileptic seizures reliably and differentiate them with events of non-epileptic nature.

The duration of EEG monitoring is usually 24 hours or more, and if used in intensive care units it can last for weeks. Visual analysis of large amounts of data obtained during long-term video–EEG monitoring requires huge labour costs and special training of clinical neurophysiologists. This determines the urgency of development new methods for detection, quantitative analysis and metric classification of diagnostic objects and removal of artifacts in long-term video-EEG monitoring of patients.

The methodology of EEG analysis is traditionally based on the visual analysis of curves, for the selection of non-artifact record fragments and the analysis of its background structure, single (epileptiform) graphical elements and their specific patterns with specificity in relation to various clinical states (Hirsch, 2010). In most cases, the algorithmic capabilities of the software for video EEG instruments are limited to preprocessing multi-channel EEG signals, indicating the likelihood of record artifacts, calculating inter-channel coherence and sources of electrical activity. To simplify the assessment of large volumes of visual information, a mathematical analysis of the oscillations is used with a graphical presentation of the results — a quantitative EEG (qEEG) (Duffy, 1994). However, the method of visual presentation of quantitative EEG in the form of trends and histograms does not take into account many artifacts, in particular, chewing and movement of the patient's head. When detecting high-amplitude plots on the qEEG histogram, physician needs to revise the in video record a fragment of interest for visual assessment and differentiation of an epileptic and artifact event. For this purpose, they use not only the electrographic pattern, but also the analysis of video.

Currently, several methods of automatic detection of seizures according to EEG data have been proposed (Hirsch, 2010; Tzallas, 2007; Antsiperov., 2017; Obukhov, 2017). However, all of them use only the native EEG without taking into account the video image, and their accuracy is insufficient for widespread clinical practice. The analysis of publications in periodicals and monographs in the subject domain showed the lack of publications on methods for automatic recognition of epileptic seizures in video sequences obtained during the video-EEG monitoring. Therefore, it is necessary to develop methods and algorithms for automated detection of diagnostic events in long-term video-EEG recordings, which will improve the reliability of their classification and significantly reduce the time for analyzing large amounts of video EEG data and increase their diagnostic significance.

In this paper, we propose an algorithm for detecting two types of diagnostic events in the video channel of the video-EEG monitoring data. The first type of events is observed for patients in coma. The second type of events is fixed for active patients and can be observed in the form of activity freezing for a few seconds.

---

* Corresponding author

2. ALGORITHM FOR DETECTING DIAGNOSTIC EVENTS

When creating methods and algorithms for automated detection of diagnostic events in video-EEG recordings, it is assumed that the decision on event detection is made when an event is detected simultaneously in two channels - EEG and video.

2.1 Detecting Events for Patients in Coma

In this case, it is necessary to analyze informative areas in patient images with visible particular muscle contractions. Informative areas are usually associated with the details of a person's face (eyes, nose, and mouth). Analysis of video sequences, which recorded seizures, showed a variety of appearance of these seizures. For example, for a non-convulsive attack, only rather weak muscular contractions are observed in the region of the patient's mouth. At the same time the rest of the facial muscles remain motionless. In another case, there are more intense contractions of the muscles of the mouth and periodic movements of the head with immobile muscles in the eye area. In a number of patients, intense contractions of the muscles are observed all over the face, and contractions of the neck muscles and head movements are also possible. In the absence of seizures, the frames of the video data are quite static for the studied group of patients.

One of the possible ways of detecting events on a video channel can be associated with an analysis of the dynamics of small areas associated with the details of a person's face (eyes, nose, mouth). The literature presents a wide range of methods for localizing such areas (Yow, 1997; Viola, 2004; Zhang, 2016; Singh, 2017; Shi, 2018; Baroffio, 2018; Araujo, 2015). It should be noted that the images of video sequences obtained during video-EEG monitoring have the following features. First, arbitrary angle of video recording the patient's face. This feature eliminates the use of methods based on the property of face symmetry, and methods that require full frontal image of the face. Second, presence of medical equipment, partially covering the details of the face. This circumstance also complicates the task of localizing diagnostically important areas. Third, Informative areas may not be associated with characteristic points of the face (eyes, corners of the mouth, etc.). Such areas, for example, may be neck areas. Therefore, traditional methods of localization of characteristic points and details of the face may not be applicable.

In this paper, we propose to detect diagnostic events by the magnitude of the criterion characterizing the degree of activity of the region of interest. The area of interest will be the portion of the frame that includes patient's face, head, and neck areas (see Figure 1). As a criterion of the activity of the region of interest, the total optical flow calculated for each frame of the video sequence will be used:

$$J(i) = \sum_{x=0}^{W-1}\sum_{y=0}^{H-1}\sqrt{V_x^2(x,y,i)+V_y^2(x,y,i)}+\delta(i),\ i=1,...,N \qquad (1)$$

where $J(i)$ is a criterion value calculated in the frame number i; W,H are the frame width and height; $V_x(x,y,i)$ and $V_y(x,y,i)$ are the optical flow values in axial directions X and Y in the frame number i at a pixel with coordinates (x,y); $\delta(i)$ is a noise.

Since the noise component is present in the model (1), the smoothed value of the activity criterion $\hat{J}(i)$ should be used to detect events. The smoothed $\hat{J}(i)$ value is obtained using a discrete version of the Kalman-Bucy filtering algorithm (Kalman, 1961). We apply the Kalman-Bucy algorithm, since it provides the optimal estimate in the sense of minimum error variance. The decision to fix a diagnostic event will be made according to the threshold rule. To avoid false alarms of the detector due to short-term spikes, the decision about the occurrence of an event will be made if the value of $\hat{J}(i)$ exceeds a predetermined threshold in a sequence of frames of length not less than M. Thus, the decision rule is formulated as follows:

$$Event = \begin{cases} 1,\ \text{если}\ \hat{J}(i) \geq T\ u\ i - i_0 \geq M; \\ 0,\ \text{если}\ \hat{J}(i) < T\ \text{или}\ i - i_0 < M, \end{cases} \qquad (2)$$

where $Event$ is an event indicator, T is a threshold value, i_0 is a frame number from which the inequality $\hat{J}(i) \geq T$ is taking place, M is the length of the sequence of frames required to make a decision about the appearance of a diagnostic event. The threshold value is defined as:

$$T = \hat{J}_0 + k\sigma, \qquad (3)$$

where $\hat{J}_0$ is computed as a mean value of $\hat{J}(i)$ in a fragment of video sequence with low dynamics of the scene, σ is a standard deviation of $\hat{J}(i)$, k is a coefficient.

Figure 1. A frame from Video-EEG record. Patient is in coma

Thus, the algorithm for detecting diagnostic events in the video channel of video-EEG monitoring data consists of the following steps.
1. Reading frame number 1 of a video sequence.
2. Computing the total optical flow in the frame of the video sequence according to the formula (1).
3. Computing of the smoothed value of the activity index $\hat{J}(i)$.
4. Checking conditions (2) and (3). If the condition $\hat{J}(i) \geq T$ is satisfied, the current frame number $i_0 = i$ is stored. If the condition is not satisfied, go to step 1.

5. Repeating steps 1–3. If the conditions $\hat{J}(i) \geq T$ and $i - i_0 \geq M$ are satisfied, it is decided that an event is detected. In the opposite case, go to step 1.

2.2 Detecting Events for Active Patients

Diagnostic events of a different type manifest as freezing for a few seconds of active patients (see Figure 2). As in the previous case, we propose to fix events by the value of the criterion (1). In contrast to the case considered above, the occurrence of an event corresponds to a minimum of the activity index. The decision rule is as follows:

$$Event = \begin{cases} 1, \; если \; \hat{J}(i) \leq T \; u \; i - i_0 \geq M; \\ 0, \; если \; \hat{J}(i) > T \; или \; i - i_0 < M, \end{cases} \quad (4)$$

and the threshold value is calculated by the formula:

$$T = \hat{J}_0 - k_1 \sigma, \quad (5)$$

where k_1 is the coefficient. Algorithm for event detecting is the same as presented in subsection 2.1, but the decision rule is (2), (3) is replaced with rule (4), (5).

In the next section, a computational experiment is conducted to evaluate the effectiveness of the proposed algorithms.

Figure 2. A frame from Video-EEG record. Seizure appears as freezing

3. EXPERIMENT

The developed algorithm is implemented in the MatLab software environment. To compute $V_x(x,y,i)$ and $V_y(x,y,i)$ in (1), Lucas–Kanade algorithm (Lucas, 1981) is applied. This algorithm for computing the optical flow is chosen from the condition of the highest performance in comparison with other methods. The magnitude of the smoothed mobility index is calculated using a discrete version of the Kalman-Bucy filtering algorithm (Kalman, 1961). The values of the filtering algorithm parameters are selected when processing test video sequences, and based on the best ratio of the error and speed values.

The experiment includes two stages. At the first stage, the detection algorithm was applied to five videos of patients in

coma. In three records epileptic seizures are detected, including non-convulsive one. For each of the five video sequences, the parameters of the decision rule (2), (3) were determined from fragments with low scene dynamics. The number of frames corresponding to the shortest event duration in condition (2) was chosen equal to 75, which corresponds to a time interval of 2.5 seconds. The value of the coefficient in (3) was selected as $k = 1.1$. Examples of graphs of the criterion $J(t)$ and its smoothed value $\hat{J}(t)$, as well as the event indicator $Event$ for two fragments of the video-EEG monitoring of the same patient are shown in Figures 3 and 6. Here and below, instead of the variable i designating the frame number, we use the variable $t = i / FrameRate$, where t is the time, and $FrameRate$ is the frame rate of the video. Figure 3 demonstrates detection of diagnostic events. The graphs in Figure 6 correspond to video record without seizures. For the videos, corresponding to the graphs shown in Figures 3 and 6, the parameters in expression (3) for determining the threshold value T were found to be equal to $\hat{J}_0 = 1604$ and $\sigma = 127.9$.
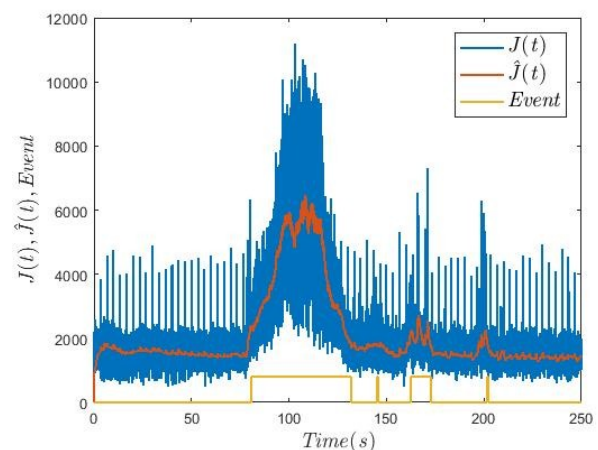

Figure 3. Graphs of the criterion $J(t)$, smoothed estimate $\hat{J}(t)$, and event indicator $Event$. Events are detected between 75 and 130 seconds, 144 and 147 seconds, 165 and 175 seconds, and between 202 and 207 seconds

Figures 4 and 5 show the graphs of the processed EEG signals from the synchronous recording of video-EEG monitoring. Figure 4 shows the projection of the ridge of the wavelet spectrogram on the power spectral density (PSD) and time axes. Using adaptive threshold, the ridge points (maximum values of the power spectral density at each time point), which lie above this threshold, were calculated. Close points of the ridge, lying above the threshold, were combined into clusters - fragments of the ridge, marked in black on the graph. These ridge fragments are interpreted as episodes of suspicious activity, similar to an epileptic seizure.

Figure 5 shows the EEG signal in one of the channels, filtered by a Butterworth filter of the 8th order with a passband from 5 to 22 Hz and notch filters at frequencies multiple of 50 Hz. The sampling frequency of the signal is equal to 500 Hz. The black color indicates the suspicious intervals obtained by analyzing the ridges of the wavelet spectrograms. From Figures 3-5 one can see the intervals of synchronous muscle and brain activity, which may correspond to epileptic seizures. These intervals are

found between 75 and 130 seconds, 144 and 147 seconds, 165 and 175 seconds, and between 202 and 207 seconds.
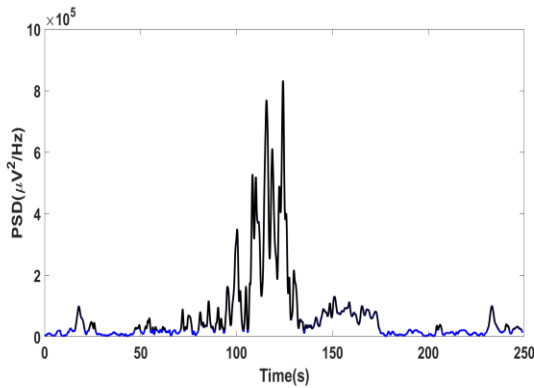

Figure 4. projection of the ridge of the wavelet spectrogram on the power spectral density (PSD) and time axes
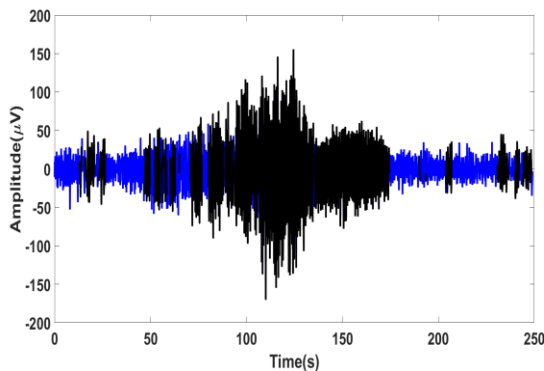

Figure 5. EEG signal in one of the channels, filtered by a Butterworth filter of the 8th order and notch filters
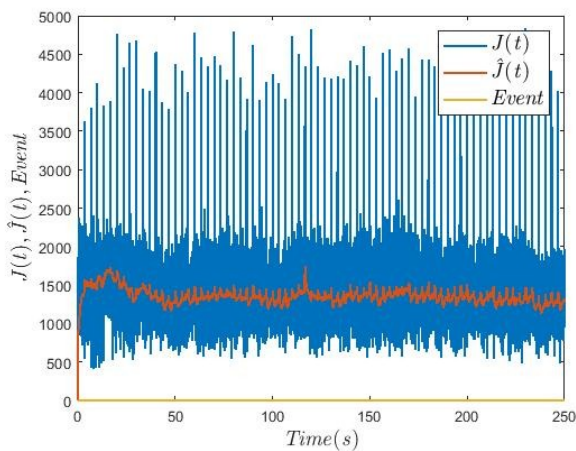

Figure 6. Graphs of the criterion $J(t)$, smoothed estimate $\hat{J}(t)$, and event indicator $Event$. The analyzed video does not contain seizures

At the second stage of the experiment, the data of video-EEG monitoring of active patient were analyzed. In this case, an epileptic seizure is observed as the patient's freezing for a few seconds. The parameters of the decision rule (4), (5) were determined from the video record fragments. For the video, the

analysis results of which are shown in Figure 7, the parameters of the expression (5) were found as $\hat{J}_0 = 4483$ and $\sigma = 797.2$. The value of the number of frames corresponding to the shortest duration of the detected event in condition (4), as at the first stage of the experiment, was chosen equal to 75, which corresponds to a time interval of 2.5 seconds. The value of the coefficient in (5) was selected as $k = 2$. The graphs of the criterion $J(t)$, its smoothed value $\hat{J}(t)$, and the event indicator $Event$ for a fragment of the video sequence of the video-EEG monitoring are shown in Figure 7. In Figure 7 it can be seen that in the interval between 30 and 35 seconds, corresponding to the freezing of the patient in the frame, a seizure was detected.
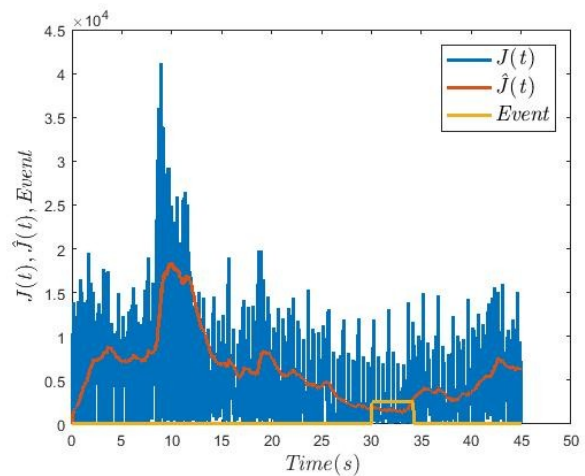

Figure 7. Graphs of the criterion $J(t)$, its estimated value $\hat{J}(t)$, and event indicator $Event$ obtained from active patient video record. Event is detected between 30 and 35 seconds

Figures 8 and 9 show graphs of the processed EEG signals from synchronous recording of video EEG monitoring. Figure 8 shows the projection of the ridge of the wavelet spectrogram on the axes PSD, time.
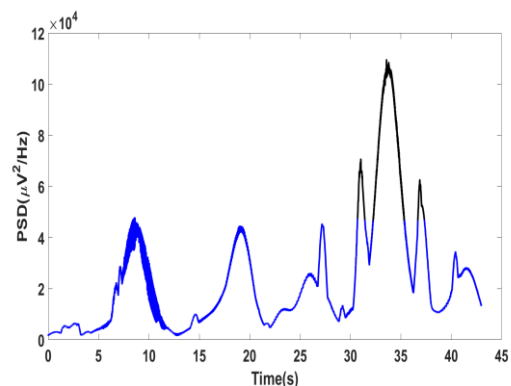

Figure 8. Projection of the ridge of the wavelet spectrogram on the power spectral density (PSD) and time axes

Figure 9 shows the filtered EEG signal in one of the channels. In the Figure 8 and 9, a fragment of the ridge and the corresponding interval of the EEG signal are marked in black.

This fragment is interpreted as an episode of suspicious activity. From Figures 7-9, you can see the interval of synchronous low muscle and high brain activity between, which may correspond to an epileptic seizure.
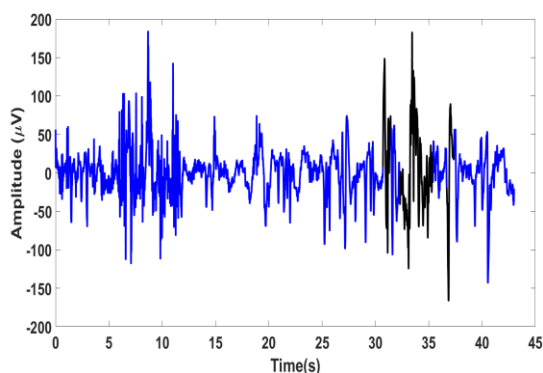

Figure 9. EEG signal in one of the channels, filtered by a Butterworth filter of the 8th order and notch filters

4. CONCLUSIONS

As part of the development of the technology for detecting epileptic seizures from video-EEG monitoring data, an algorithm for automatic detection of diagnostic events based on the analysis of the quantitative characteristics of the patient's activity in video records is proposed. Analysis of video sequence is focused on identifying a group of frames with high or low scene dynamics using a criterion calculated from the optical flow magnitude. The preliminary results of the analysis of real clinical data for patients in coma and active patients are presented. The results of the analysis showed the efficiency of the proposed algorithm at different angles of shooting and partial occlusion of the patient's face with the details of medical equipment. A comparison of the results of diagnostic event detection from the video record with data obtained from the synchronous EEG was made. The comparison showed the possibility in principle of reliable diagnosing epileptic seizures and distinguishing them from non-epileptic events. Further research will be aimed at developing a method for detecting epileptic seizures based on a joint analysis of synchronous EEG and video channels.

ACKNOWLEDGEMENTS

REFERENCES

Antsiperov, V. E., Obukhov, Y. V., Komol'tsev, I. G., & Gulyaeva, N. V., 2017. Segmentation of quasiperiodic patterns in EEG recordings for analysis of post-traumatic paroxysmal activity in rat brains. *Pattern Recognition and Image Analysis*, *27*(4), pp. 789-803.

Araujo, G.M., 2015. Detection and Tracking of Facial Landmarks in High Definition Video Sequences. Ph.D. Thesis. COPPE/UFRJ. http://www02.smt.ufrj.br/~eduardo/teses/gabriel_matos-doutorado.pdf. (02 April 2019).

Baroffio, L., Cesana, M., Redondi, A., Tagliasacchi, M., Tubaro, S., 2016. Fast keypoint detection in video sequences. In 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1342-1346.

Duffy, F. H., Hughes, J. R., Miranda, F., Bernad, P., Cook, P., 1994. Status of quantitative EEG (QEEG) in clinical practice, *Clinical Electroencephalography*, *25*(4), pp. vi-xxii.

Hirsch, L., Brenner, R. (Eds.), 2011. *Atlas of EEG in critical care*. John Wiley & Sons.

Lucas, B. D., Kanade, T., 1981. An iterative image registration technique with an application to stereo vision. In: Proceedings of Imaging Understanding Workshop, pp. 121-130.

Kalman, R. E., Bucy, R. S., 1961. New results in linear filtering and prediction theory. *Journal of basic engineering*, *83*(1), pp. 95-108.

Obukhov, K., Kershner, I., Komol'tsev, I., Maluta, I., Obukhov, Y., Manolova, A., Gulyaeva, N., 2017. An approach for EEG of post traumatic sleep spindles and epilepsy seizures detection and classification in rats. *Pattern Recognition and Image Analysis*, 27(1), pp. 114-121.

Shi S., 2017. Facial Keypoints Detection, arXiv:1710.05279v1 [stat.ML] . https://arxiv.org/pdf/1710.05279. (02 April 2019).

Singh, A., Patil, D., Reddy, M., Omkar, S. N., 2017. Disguised face identification (DFI) with facial keypoints using spatial fusion convolutional network. In Proceedings of the IEEE International Conference on Computer Vision, pp. 1648-1655.

Tzallas, A. T., Tsipouras, M. G., & Fotiadis, D. I., 2007. Automatic seizure detection based on time-frequency analysis and artificial neural networks. *Computational Intelligence and Neuroscience*, *2007*. doi.org/10.1155/2007/80510.

Viola, P., Jones, M. J., 2004. Robust real-time face detection. *International journal of computer vision*, *57*(2), pp. 137-154.

Yow, K. C., Cipolla, R., 1997. Feature-based human face detection. *Image and vision computing*, *15*(9), pp. 713-735.

Zhang, S., Meng, C., 2016. Facial keypoints detection using neural network. Stanford Report. http://cs231n.stanford.edu/reports/2016/pdfs/007_Report.pdf (02 April 2019).