COMPEL - The international journal for computation and mathematics in electrical and electronic engineering

Joint analysis of simultaneous EEG and eye tracking data for video images

Dominik Szajerman, Piotr Napieralski, Jean-Philippe Lecointe,

Joint analysis of simultaneous EEG and eye tracking data for video images

Dominik Szajerman and Piotr Napieralski
Institute of Information Technology, Lodz University of Technology, Lodz, Poland, and
Jean-Philippe Lecointe
Laboratory of Electrotechnical Systems and Environment LSEE, Artois University, Béthune, France

Received 14 July 2018
Revised 14 July 2018
Accepted 18 July 2018

Abstract

Purpose – Technological innovation has made it possible to review how a film cues particular reactions on the part of the viewers. The purpose of this paper is to capture and interpret visual perception and attention by the simultaneous use of eye tracking and electroencephalography (EEG) technologies.

Design/methodology/approach – The authors have developed a method for joint analysis of EEG and eye tracking. To achieve this goal, an algorithm was implemented to capture and interpret visual perception and attention by the simultaneous use of eye tracking and EEG technologies. All parameters have been measured as a function of the relationship between the tested signals, which, in turn, allowed for a more accurate validation of hypotheses by appropriately selected calculations.

Findings – The results of this study revealed a coherence between EEG and eye tracking that are of particular relevance for human perception.

Practical implications – This paper endeavors both to capture and interpret visual perception and attention by the simultaneous use of eye tracking and EEG technologies. Eye tracking provides a powerful real-time measure of viewer region of interest. EEG technologies provides data regarding the viewer's emotional states while watching the movie.

Originality/value – The approach in this paper is distinct from similar studies because it takes into account the integration of the eye tracking and EEG technologies. This paper provides a method for determining a fully functional video introspection system.

Keywords Control systems, Sensors, Optimal design, Numerical analysis, Applied electromagnetism, Condition monitoring

Paper type Research paper

1. Introduction

Electroencephalography (EEG) is an electrophysiological monitoring method used to record the electrical activity of the brain. It can be used for estimating electric neuronal activity distribution. Eye tracking is a powerful tool to measure viewer region of interest objectively. An eye tracker is a device for measuring either the point of the gaze or the motion of an eye relative to the head. EEG and eye tracking technologies have been investigated in human–computer interaction (HCI) contexts since the 1980s. The initial studies in this field focused primarily on reading comprehension models based on the allocation of eye fixations (Just and Carpenter, 1980). EEG-based brain–computer interfaces (BCIs) detect events of the users' attention for HCI without manual intervention (Donchin *et al.*, 2000).

Human affective states may be used to recognize users' emotions or provide full proportional control over the assistive devices for disabled or injured persons

COMPEL - The international journal for computation and mathematics in electrical and electronic engineering

0332-1649
DOI 10.1108/COMPEL-07-2018-0281

(McMurrough *et al.*, 2013; Półrola and Wojciechowski, 2012). The combination of recording eye tracking and EEG is useful for many purposes. These include:

- fixation control;
- microsaccades signal distortions;
- ocular artifact correction;
- measuring saccadic reaction times;
- stimuli gaze-contingently; and
- improving BCIs.

In the film industry, it is important to use such signals to analyze viewer reaction. In the literature, we can find non-physiological and physiological signals for emotion recognition. Non-physiological signals include facial expression (Pantic and Rothkrantz, 2000; Staniucha and Wojciechowski, 2016), speech or gesture (Maksymiv *et al.*, 2016; Pantic and Rothkrantz, 2000). Physiological signals such as EEG, pupillary diameter (PD), electromyogram (EMG) and electrocardiogram (ECG), seem to be more effective and accurate. EEG which records brain activities provides responses to emotional states. In this paper, we offer a novel approach for a physiological-based analysis of EEG signals with eye activity, for the assessment of particular viewer reactions (Szajerman and Napieralski, 2017).

2. Related work

Modulation of the EEG activity can be used for discrete or continuous real-time control of the "Computer-User Interface" (van Gerven and Jensen, 2009; Treder *et al.*, 2011; Szajerman *et al.*, 2016). Visual BCIs often use intense visual stimulation by eye-trackers for simultaneously capturing the subjects' focus of attention by detecting covert shifts of visual attention. Trachel *et al.* (2013) propose a future interface that could enhance HCI by displaying visual information at the focus of attention. Putze *et al.* (2013, 2016) recorded gaze movements (via eye tracking) and brain activity (via electroencephalography) for implicit selection in a graphical user interface. This solution works without any manual intervention of the user. This method allows the localization processes to investigate how the modalities interact.

Jangraw and Sajda (2013) developed a similar hybrid classifier capable of combining signals into a fully functional mobile HCI system. Affective states are psycho-physiological constructs which vary along three principal dimensions (Harmon-Jones *et al.*, 2013):

- valence;
- arousal; and
- motivational intensity.

They found that the intensity broadens the cognitive range, affecting high motivational intensity of the narrow cognitive range.

Initial work in this field focused primarily on developing methods for affective state recognition. The first approach to recognizing human emotions and their affective power was proposed by Richard A. McFarland (McFarland, 1985), who investigated the relationship between skin temperature changes and emotions accompanying music. A similar approach can be found in the work of (Kim *et al.*, 2004; Jonghwa and André, 2008) who proposed using the heart rate and skin conductance to recognize an audience's affective states. A proposed user-independent system was based on physiological signal databases obtained from an electrocardiogram, skin temperature variation and electrodermal activity. They find the influence of emotion on the autonomic nervous system. Hou *et al.* (2016) proposed a combination of feature extraction, feature selection and SVM classification for feature like/dislike analysis. Some related work has been done by the researchers to classify the emotions of a movie audience by SVM classification (Xu *et al.*, 2008; Souvannavong *et al.*, 2004; Hee and Loong-Fah, 2006; Hanjalic and Xu, 2005; Salway and Graham, 2003). The recognition by video of viewers' emotions can be performed by the analysis of EEG signals and facial expressions (Soleymani *et al.*, 2009; Shiliang *et al.*, 2008; Cowie *et al.*, 2001). Facial emotions analysis is an indispensable element of emotion analyses. Mouth features' extraction became one of the most representative face regions in the context of emotions retrieval.

The electroencephalogram method combined with eye tracking algorithms can be used to recognize human affective states Liu *et al.* (2016). Human perception with a functional brain response can bring much information by joint data mining of EEG and eye tracking signals. This approach was applied by Scholler *et al.* (2012) to the measurement of human perception of video quality change. This change gives rise to specific components in the EEG that can be detected on a single-trial basis.

Bodala *et al.* (2015) investigated the utility of a specific type of eye movement, referred to as microsaccades, as a marker for the analysis of ERPs during a workload task for an online estimation of 28 points, such as mental workload and fatigue. They show a clear correlation of ERP activations to both latency and activation areas.

Zheng *et al.* (2014) presented a new emotion recognition method which combined EEG and pupil diameter to recognize emotions. They used two fusion strategies (feature-level fusion and decision-level fusion) to build emotion recognition models. The level of feature-level fusion and decision-level fusion combining EEG signals and eye tracking data can improve the performance of the emotion recognition model. For reduction of space dimensionality, block scalar quantization, or additive white noise can be used with the KLT by maximizing the sum of fourth-order moments Puchala (2013).

All presented solutions, both those related to HCI and those related to affective state recognition, identify features from EEG data, synchronized and analyzed with respect to the events obtained from eye tracking. This enables a more accurate acquisition of data from these two devices for the purpose of brain states classification. The present paper explores the integration of the aforementioned features into a fully functional video introspection system.

3. Equipment

The research involved two types of data sources: eye tracking and electroencephalography. The former was provided by Tobii EyeX Controller eye-tracker (Plate 1), the latter by NeuroSky Mindwave EEG/BCI device (Plate 2). To integrate them, both devices were connected to a computer which acquired data and displayed visual stimuli at the same time. The whole operation was conducted using software designed by the authors of the present study.

The Tobii EyeX Controller tracks the user's eyes at a distance of 50-90 cm, acquiring the gaze position with a 70-Hz frequency and about 1 cm accuracy on a typical 17-inch laptop screen. To some extent, it also tracks the user's head, so the head needs to be fixed. EyeX uses USB3.0 to communicate with the system.

The NeuroSky Mindwave measures and outputs the EEG power spectrums (alpha, beta waves, etc.) and brain wave metrics (attention and meditation), among others. It consists of a headset, an ear-clip and an arm with one EEG sensor designed to rest on the forehead above

Plate 1.
Tobii EyeX controller

Plate 2.
NeuroSky Mindwave EEG/BCI device

the eyes. After preprocessing, the data are typically provided at a frequency of 1 Hz. Mindwave uses a USB-Bluetooth dongle to communicate with the system.

4. System design

Both devices presented in the previous chapter need a proper SDK to communicate with the designed application, as they not only have different requirements, but offer unlike programming philosophies and different programming interfaces.

Our system's task is to support two phases of the study. The first is a test of a group of subjects (recipients of the movie). This means showing them the film and recording EEG signals and eye tracking data that were created during the viewing. The second is the support in subsequent analysis (individual and/or group) of recorded data. To build a system that allows for joint analysis of EEG and eye tracking data, the proper architecture should be proposed. The following requirements were expected to be met by the system:

- The system should be composed of two modules: the recording, and the player, which allows for simultaneous analysis. Such a distinction means that the recording of EEG signals and eye tracking information during movie playback can be performed by a technician who has only the necessary program options available (the recorder). Conversely, in the analytical part (the player), the researcher has access to extensive tools, but without the risk of accidentally running the recording nor the loss or distortion of the data. Further analysis resulted in the use of limited player functions in the recording phase for the presentation of the tested movie.
- Because of the large diversity and variability of access to the equipment, it is necessary to architecturally adapt the application for easy expansion to support new devices.
- Because of the separation of system functionality into two modules, it is necessary to develop a method of communication between them. This is necessary both at the level of the recorder's and the player's cooperation during recording, as well as during the subsequent exchange of data recorded offline.
- For exchanging and storing the recorded data, use the human-readable format. It should be independent of the movie format (2D, 3D), its resolution (SD, HD) and frame rate. At the same time, it should allow for expansion in the event that new equipment will offer more types of signals useful for recording.
- The recorder should allow for efficient and effective creation of data sets. Which means that for a given movie, we can test many viewers in the shortest possible time. The recorder–user interface should take into account the single selection of the film and global parameters and then allow for it to be reduced to the minimum necessary settings to examine each next subject.
- The movie player should support popular video formats. The best solution is to integrate it with operating system codecs so that it automatically adapts to the user's environment and his/her preferences or needs.
- The tool, in addition to traditional data analysis, should be prepared for visual presentation and study of the data. This is required by the target users who are film-makers accustomed to working with (motion) pictures.
- The aforementioned visual presentation should be easily expandable with new indicators and elements to be able to evolve with the needs of users and the capabilities of the recording equipment.

The above requirements, led to the design of a scalable architecture providing the division of the system into two computer applications. Their main components are presented in Figure 1. They are described in the following subsections.

4.1 The recorder

The Recorder application project had to meet requirements (2-5). The first of them was fulfilled by using abstraction, generalizing support for two groups of devices:

(1) Eye-trackers, which mainly read the position of the user's gaze (viewing) point, can also provide additional data, such as the degree of eye closing (or blink) and for some devices the diameter of the pupil. These parameters have been selected as an abstraction of eye tracking data entering our system. The Eye Tribe device used in

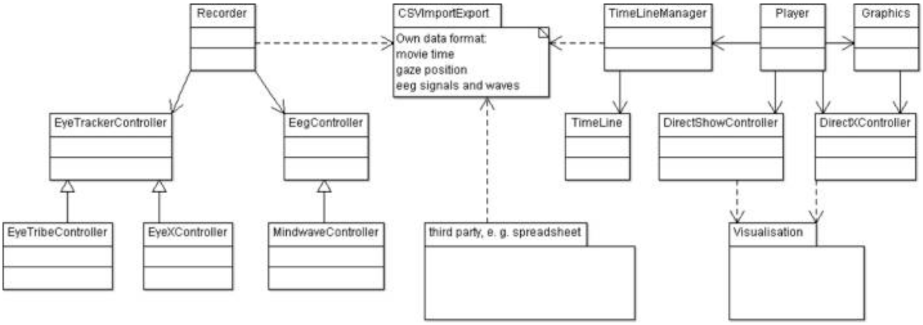


Figure 1.
System architecture

the initial phase of research was abandoned by its development team. Therefore, it quickly turned out that the scalability of the solution with subsequent types of devices is very useful and has been practically proven. The second supported eye-tracker type during the presented research was the Tobii EyeX.

(2) EEG reading devices are much more diverse because of their earlier use in medicine. Such solutions allow the use of multiple electrodes and transmit data in the form of raw waveforms. Therefore, they require modules for additional signal aggregation into forms more suited to the requirements of our research. Conversely, non-medical solutions for interactive applications and computer games have recently appeared. Software embedded in their drivers and the SDKs contain interpreting modules that transmit ready-made power spectra of brain waves or even more aggregated parameters such as attention, meditation, excitement or frustration for further processing. Such parameters have been selected as an abstraction of EEG signals entering the system. However, this does not prevent the addition of modules for handling medical grade devices.

The requirement (3) is for communication between recorder and player. It works by various means. The first of these is starting the player via the recorder. Such a launch can be made explicitly, by pressing the “Run player” button – then the full version of the player starts and the user has access to all analytical options. The second way to start the player by the recorder takes place automatically when the technician starts the test – then the player works in full screen mode and does not allow any action, so the subject (the tested recipient of the film) can only watch the movie, and hihe/sher EEG and eye tracking parameters are recorded by the recorder. The third method of communication occurs between both applications. Player displaying the movie has knowledge about the passage of time (current movie playback time), which is passed to the recorder. Recorder needs this parameter to properly record time-stamped EEG and eye tracking signals. This is done using the operating system’s message system. The fourth way of communication between these modules is the exchange of recorded data itself. It takes place in the manner described in the next section.

The requirement (5) applies to the user interface. Figure 2 shows two recorder configuration screens. The first one allows us to configure input devices and to calibrate those that require it. The second screen globally configures the movie data for the study and a place for recorded data. As we can see, the only parameter set for each examined person is hihe/sher ID (e.g. name and/or surname). After this we can start the test by “Start examination” button.

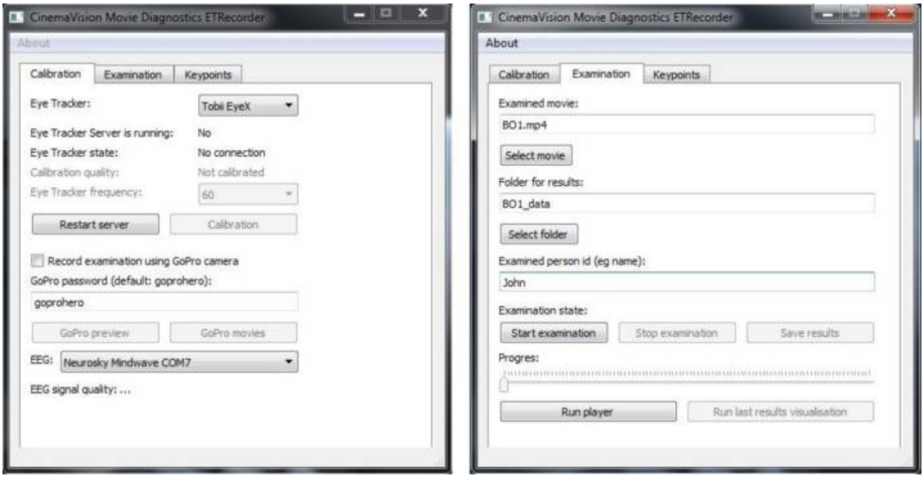


Figure 2.
View of two tabs in the main window of the recorder

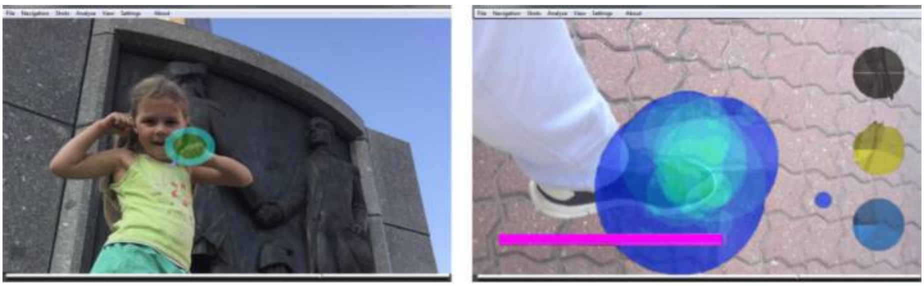
Figure 3 shows the main components of the recorder application. We can see the abstractions of the controllers of the eye-tracker (*EyeTrackerController*) and EEG (*Eeg-Controller*), which are specific to real devices: *EyeTribeController*, *EyeX-Controler* and *MindwaveController*.

4.2 Data exchange file

Requirement (4) has been completed by the use of a human-readable, comma separated file format. Its nature allows for the easy definition of the component parts of the signal record. At the same time, by simply naming and adding columns, there is the possibility of flexibly scaling the solution.

The listing below shows the part of the eye tracking record for the first four measurements, which took place successively in the 1st, 22st, 24st and 44st milliseconds of the film's duration.

In accordance with the definitions contained in the first line of the file, subsequent columns contain two coordinates (x, y) of the subject's gaze location, the eye closing mark (a "1" means open) and the pupil diameter (the recording device did not support this option).



Notes: Left: GDI drawing; Right: DirectDraw drawing

Figure 3.
Two screens of the main window of the player

In the case of the gaze point coordinates, the decimal point/comma problem was solved by rescaling – multiplying values in the range $<0, 1>$ by 10,00,000.

```
time_ms;pos_x_1m_normalized;pos_y_1m_normalized;blink_bool;
  pupil_diameter_100nm
#Examination date: 2017_05_06_15_58_31
#Examination id: John
1;338999;634601;1;0
22;334998;640965;1;0
24;342813;632962;1;0
44;336522;623247;1;0
```

The listing below presents a fragment of the EEG signal recording. In the same way as in the previous case, the first line contains the definition of the contents of individual columns. These are: time [ms], signal quality (the device declares the reliability of the measurements made at the current moment; "0" means flawless), the level of attention, the level of meditation and subsequent raw power spectra of individual brain waves.

```
time_ms;eeg_quality;eeg_attention;eeg_meditation;
  mindwave_delta;mindwave_theta;mindwave_alpha1;mindwave_
alpha2;mindwave_beta1;mindwave_beta2;
  mindwave_gamma1;mindwave_gamma2
#Examination date: 2017_04_12_13_39_17
#Examination id: John
866;0;78;67;692000;128848;13295;28788;19549;10332;9876;4655
1840;0;57;75;929405;37902;7062;16494;18813;2647;3713;100328
24;0;29;77;952900;185131;98966;29804;23724;14030;5513;16913
876;0;21;70;709127;92409;12380;5617;10297;6681;2237;1690
```

As we can see, the measurement frequency closely depends on the device and its software. In the case shown, the EEG device is several dozen times slower than the eye-tracker. This results in the later need to match both recordings, which is the subject of the solution presented in this paper. Common support for CSV files is shown in Figure 1 as the CSVImportExport module.

4.3 The player

The player application project had to meet requirements (5-7). It is invoked both during the movie testing and during visual analysis of the gathered data.

Requirement (6) applies to video formats. We decided to use DirectShow, which is the application programming interface (API) used in media streaming in Windows OS. With it, it is possible to both record and play media. In our project, we used the two options offered:

(1) Playing movie clips. DirectShow covered the demand for the support of many file formats, and the codecs registration system handled the file formats which the user worked on. These include, but are not limited to: ASF, WMV, MPEG, WAV, MP3 etc.

(2) Applying graphic objects to the movie content being played. This was done using the interface *IDirect3DSurface*, which let us pass a surface on which GDI or Di-rectDraw draws to DirectShow rendering pipeline. This resulted in the overlapping (also with transparency) of the rendered graphic objects on the displayed movie image.

The second point above leads to requirement (8). The use of GDI or DirectDraw allowed for the rendering of any shape whose task was to represent selected recorded eye tracking and EEG parameters.

Figure 3 shows the Player action. First, we tested GDI. It allows us to easily draw predefined geometric figures (circles), but at the same time it limits the use of transparency levels to one preset for the whole image. This allowed drawing precise images, which, however, could not be complicated. It was useful for presenting and analyzing results for a single subject. On the left side of the Figure 3, a user can see the location of the gaze point of the subject marked with a circle, the green filled circle means the eyes were open, and the thickness of the cyan ring means the length of time looking at this point. This makes it easy to visually assess the interest of the subject at a given point in the image. In turn, on the right side of Figure 3 we see the effect of the DirectDraw-based drawing algorithm. It was necessary to create our own drawing procedures (circles, bars, pie charts), but at the same time it allowed the use of many colors and degrees of transparency. As a result of this, we made a heatmap showing the position of the gaze points of many subjects at the same time (various sizes of circles, here also the size shows the length of time of viewing), mixing colors allows to create a heatmap showing the number of subjects looking at a given area of the image. In addition, on the screen the magenta bar indicates the average pupil size of subjects and on the right side the values of three recorded EEG parameters: signal quality, attention and meditation. They are presented as pie charts.

The above-mentioned pie charts were a suggestion of filmmakers – users and system testers who helped us find optimal data presentation methods for analysis. As a result, our system met the requirement 8 for the visualization module to be extendable with new possibilities.

5. Dataset

Movie audience research with the use of eye-tracker and EEG equipment in the context of the image quality of the film is not yet very common. Therefore, this field has not yet received large databases for research, such as for example, computer vision. Hence, to conduct out joint analysis research, we decided to build our own data collection.

The experiment was conducted on a group of 23 volunteers – students of Lodz University of Technology. They were asked to wear the EEG headset and sit at 70 cm from the 17-inch computer monitor with the eye-tracker attached. The participants were shown a 1-min and 12-s video at 25 FPS which gave 1,800 presented frames. During the presentation, the eye tracking and EEG data were recorded. All waveforms: gaze position from eye-tracker and attention level from EEG, were stored in separate files for each participant.

The recording speed of the eye-tracker was around 30-60 Hz. As it was not stable, the minimum number of recorded samples for a user was 2,290 and the maximum was 4,605. At the same time, the numbers of samples acquired from the EEG fit in the range 72-74, which gave a recording speed of 1 Hz. The instability of the eye tracking recording frequency, the inconsistency of both recording frequencies and the lack of possibility of synchronization caused by differences between the devices' APIs made it necessary to create a joint analysis method.

6. Experiments and methods

Visual analysis allowed for intuitive discovery and anticipation of the relationship between the tested signals, which, in turn, allowed for a more accurate validation of hypotheses by appropriately selected calculations.

The joint analysis method consists of two stages:

- the preprocessing; and
- the collation of measures.

The former stage consisted of sampling recorded signal waveforms with a constant step corresponding to 15 ms (Figure 4). Each sample was calculated by interpolation of its neighbors from the recorded data. This procedure enabled the collation of all waveforms, which was followed by the next stage aimed at finding the relationships between eye-tracker-recorded and EEG-recorded data.

The following measures were used during the analysis of the eye tracking data. Each value was calculated at a specific moment. Of these three measures, the first two are based on the average of all gaze points (the centroid). Figure 5 shows the centroid value calculated for an example frame from the test video.

The first measure – *spread of gaze points* was calculated as in equation (1) (Figure 6) which could be treated as the equality of variance for a set of two-dimensional points:

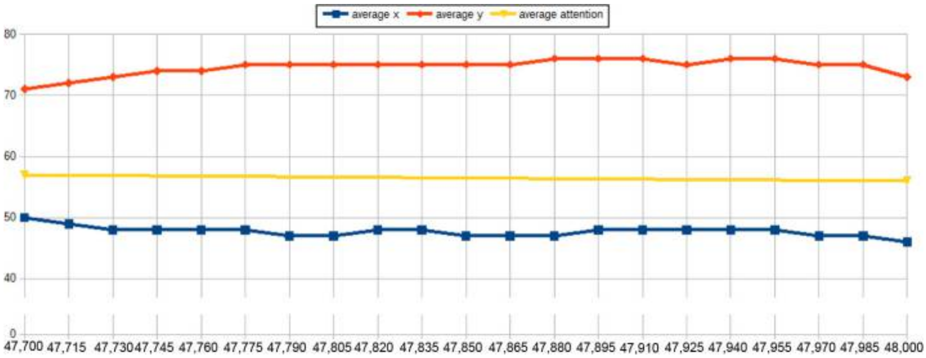


Figure 4.
The sampling process

Figure 5.
Measured values of gaze points (circles) and their centroid (white dot); source: the designed software

$$s = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \bar{x})^2 + (y_i - \bar{y})^2} \tag{1}$$

where: N is the number of gaze points ($N = 23$ in our experiment), (x_i, y_i) are the coordinates of a subsequent gaze point ($\bar{x}, \bar{y}$) is the average of all gaze points' coordinates (the centroid).

As the user's gaze points are more inconsistent, the spread s is bigger. Figure 6 shows two example frames from the test video with two different gaze point distributions and the s value they gave.

The second measure [equation (2)] is the *collective speed* (Figure 7) which represents the speed of the centroid of the gaze points:

$$c = \sqrt{\left(\overline{x^t} - \overline{x^{t-1}}\right)^2 + \left(\overline{y^t} - \overline{y^{t-1}}\right)^2} \tag{2}$$

where: ($\overline{x^t}$, $\overline{y^t}$) is the centroid of all the gaze points in the current time frame, and ($\overline{x^{t-1}}$, $\overline{y^{t-1}}$) is the centroid in the previous time frame.

The time step is known (15 ms); thus, c could be interpreted as the speed. The third measure is the *average of the individual speeds* [equation (2), Figure 7], which could be better than the collective speed of detected eyesight wandering of individual participants:

$$a = \frac{1}{N}\sum_{i=1}^{N}\sqrt{(x^t - x^{t-1})^2 + (y^t - y^{t-1})^2} \tag{3}$$

where: (x^t, y^t), (x^{t-1}, y^{t-1}) are the individual gaze position coordinates in the current and the previous frame, respectively.

During the analysis, with the second source of data – electroencephalographic one, only the attention signal was considered. Because the device's software and SDK provides a ready-to-use, normalized value, it was sufficient to calculate its average for all users (Figure 7).

The signal from eye-tracking is very variable. Because of the nature of the construction and operation of the eye, the recording, in addition to the intuitively perceived direction of looking, also contains a well-visible component of tremor. While observing the recorded results, it is very tiring for the person analyzing them, and at the same time it adds nothing to the conclusions. For this reason, the presentation module in the player contains the option of filtering the position of the recorded gaze points. This is done with moving average understood as unweighted mean of the previous samples g_t^{filtered} in the time t():

Source: The designed software

Figure 6.
The spread equals 0.298 for the left frame and 0.191 for the right one; overlapping circles create a heat map

$$g_t^{\text{filtered}} = \frac{1}{n}\sum_{i=0}^{n-1} g_{t-i} \tag{4}$$

where n is the quantity of the previous samples taken into account, and $g_{t\text{-}i}$ is the unfiltered, raw gaze position. The default n = 8, but it is fully configurable in the system.

Another visual aspect of facilitating visual analysis is to show the time of the gaze point focus duration on the current position.

Again, the tremor of this point must be taken into account in the subsequent moments. However, this time a certain tolerance was assumed expressed in per cent of the size of the image being examined. The focus time τ_t in the time t is the number of previous consecutive picture frames, during which the pointe gaze was within the tolerance range d relative to the current position:

$$\tau_t = t - max\{i \in (0,\, t) \wedge \| \mathrm{g}_t - g_{t-i} \| \ > d\} \tag{5}$$

where t and i are integers and g_t, g_{t-i} are positions of gaze points in specified time frames. The τ_t value is then used to calculate the size of the circles included in the heatmap, if the "Consider focus time" option is enabled in the system (e.g. Figure 3, right side). If this option is disabled, the circles have the same size (e.g. Figure 5 and 6).

7. Results

All the measures mentioned in the previous section were collated. The result is presented in Figure 8.

After collating the data, the following observations were made:

- The decrease in the spread of gaze points precedes the increase in the average attention level. This may indicate that the observer's sight was attracted to a specific element in the image, which then resulted in generating an interest signal in the brain.
- The decrease in the average attention level precedes the increase in the speeds (especially the average individual speed) of gaze points, which may indicate a loss of

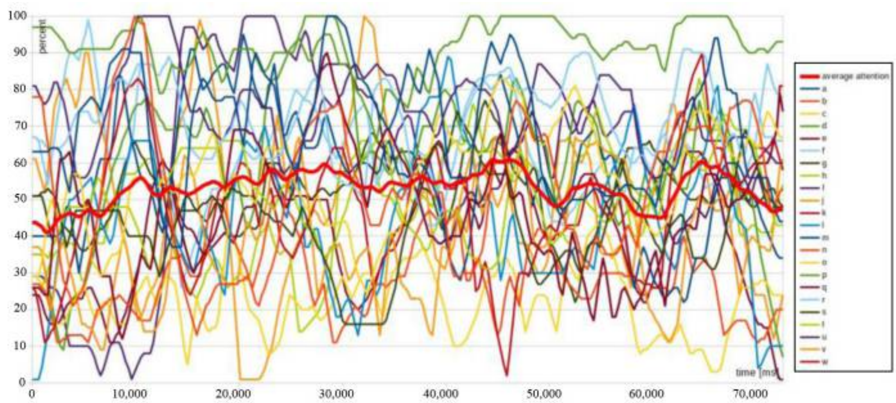


Figure 7.
The individual values of attention measurements (a-w) and average value versus time

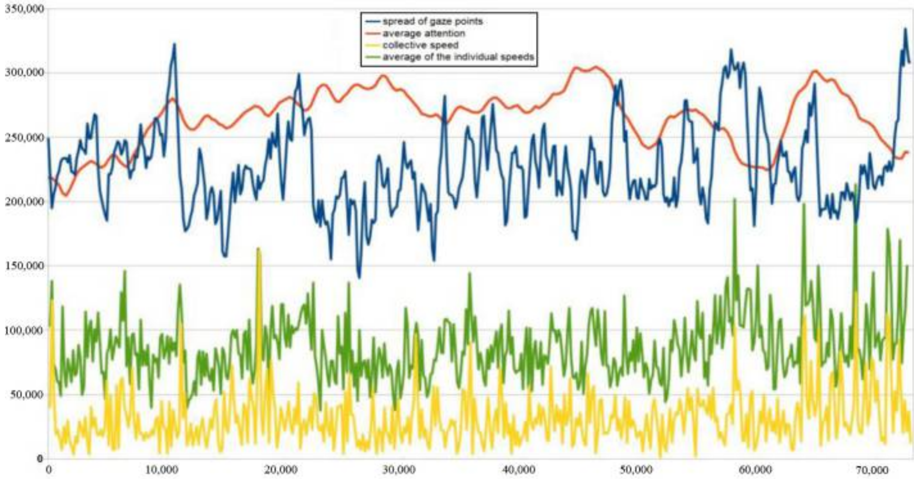


Figure 8.
Calculated measures versus time

interest in a particular element in the image, followed by further image exploration in the search for other interesting items.

8. Conclusions and future work

The proposed method of joint analysis could be successfully used for the analysis of data collected using technologies as different as eye tracking and EEG. The designed system for analysis and visualization of BCI and eye tracking data are as yet at an early stage of development. As the presented data analysis evolves, new methods will be added for jointly developing brain and visual perception models with a view to understand how the human brain reacts to movies.

References

Bodala, I.P., Kukreja, S., Li, J., Thakor, N.V. and Al-Nashash, H. (2015), "Eye tracking and EEG synchronization to analyze microsaccades during a workload task", in 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), *IEEE*.

Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W. and Taylor, J.G. (2001), "Emotion recognition in human-computer interaction", *IEEE Signal Processing Magazine*, Vol. 18 No. 1, pp. 32-80, doi: 10.1109/79.911197.

Donchin, E., Spencer, K.M. and Wijesinghe, R. (2000), "The mental prosthesis: assessing the speed of a p300-based brain-computer interface", *IEEE Transactions on Rehabilitation Engineering*, Vol. 8 No. 2, pp. 174-179, doi: 10.1109/86.847808.

Hanjalic, A. and Xu, L. (2005), "Affective video content representation and modeling", *IEEE Transactions on Multimedia*, Vol. 7 No. 1, pp. 143-154, doi: 10.1109/tmm.2004.840618.

Harmon-Jones, E., Gable, P.A. and Price, T.F. (2013), "Does negative affect always narrow and positive affect always broaden the mind? Considering the influence of motivational intensity on cognitive scope", *Current Directions in Psychological Science*, Vol. 22 No. 4, pp. 301-307, doi: 10.1177/0963721413481353.

Hee, L.W. and Loong-Fah, C. (2006), "Affective understanding in film", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 16 No. 6, pp. 689-704, doi: 10.1109/tcsvt.2006.873781.

Hou, Y., Xiao, T., Zhang, S., Jiang, X., Li, X., Hu, X., Han, J., Guo, L., Miller, L.S., Neupert, R. and Liu, T. (2016), "Predicting movie trailer viewer's 'like/dislike' via learned shot editing patterns", *IEEE Transactions on Affective Computing*, Vol. 7 No. 1, pp. 29-44, doi: 10.1109/taffc.2015.2444371.

Jangraw, D.C. and Sajda, P. (2013), "Feature selection for gaze, pupillary, and EEG signals evoked in a 3D environment", in Proceedings of the 6th workshop on Eye gaze in intelligent human machine interaction: gaze in multimodal interaction - GazeIn '13, *ACM Press*.

Jonghwa, K. and André, E. (2008), "Emotion recognition based on physiological changes in music listening", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30 No. 12, pp. 2067-2083, doi: 10.1109/tpami.2008.26.

Just, A.M. and Carpenter, A.P. (1980), "A theory of reading: from eye fixations to com- prehension", *Psychological Review*, Vol. 87 No. 4, pp. 329-354, doi: 10.1037/0033-295x.87.4.329.

Kim, K.H., Bang, S.W. and Kim, S.R. (2004), "Emotion recognition system using short-term monitoring of physiological signals", *Medical and Biological Engineering and Computing*, Vol. 42 No. 3, pp. 419-427, doi: 10.1007/bf02344719.

Liu, S., Lv, J., Hou, Y., Shoemaker, T., Dong, Q., Li, K. and Liu, T. (2016), "What makes a good movie trailer?", in Proceedings of the 2016 ACM on Multimedia Conference - MM '16, *ACM Press*.

McFarland, A.R. (1985), "Relationship of skin temperature changes to the emotions accompanying music", *Biofeedback and Self-Regulation*, Vol. 10 No. 3, pp. 255-267, doi: 10.1007/bf00999346.

McMurrough, C., Ranatunga, I., Papangelis, A., Popa, O.D. and Makedon, F. (2013), "A development and evaluation platform for non-tactile power wheelchair controls", in *Proceedings of the 6th International Conference on PErvasive Technologies Related to Assistive Environments - PETRA '13, ACM Press*.

Maksymiv, O., Rak, T. and Menshikova, O. (2016), "Deep convolutional network for detecting probable emergency situations", IEEE First International Conference on Data Stream Mining and Processing (DSMP), pp. 199-202.

Pantic, M. and Rothkrantz, L.J.M. (2000), "Automatic analysis of facial expressions: the state of the art", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 22 No. 12, pp. 1424-1445, doi: 10.1109/34.895976.

Półrola, M. and Wojciechowski, A. (2012), "Computer science in novel applications, human-centered interdisciplinary applications".

Puchala, D. (2013), "Approximating the KLT by maximizing the sum of fourth-order moments", *IEEE Signal Processing Letters*, Vol. 20 No. 3, pp. 193-196, doi: 10.1109/LSP.2013.2237764.

Putze, F., Hild, J., Ka«rgel, R., Herff, C., Redmann, A., Beyerer, J. and Schultz, T. (2013), "Locating user attention using eye tracking and EEG for spatio-temporal event selection", in *Proceedings of the 2013 international conference on Intelligent user interfaces - IUI '13, ACM Press*.

Putze, F., Popp, J., Hild, J., Beyerer, J. and Schultz, T. (2016), "Intervention-free selection using EEG and eye tracking", in *Proceedings of the 18th ACM International Conference on Multimodal Interaction - ICMI 2016, ACM Press*.

Salway, A. and Graham, M. (2003), "Extracting information about emotions in films", in *Proceedings of the eleventh ACM international conference on Multimedia -MULTI- MEDIA '03, ACM Press*.

Scholler, S., Bosse, S., Treder, M.S., Blankertz, B., Curio, G., Muller, K.R. and Wiegand, T. (2012), "Toward a direct measure of video quality perception using EEG", *IEEE Transactions on Image Processing*, Vol. 21 No. 5, pp. 2619-2629, doi: 10.1109/tip.2012.2187672.

Shiliang, Z., Qi, T., Shuqiang, J., Qingming, H. and Wen, G. (2008), "Affective MTV analysis based on arousal and valence features", in 2008 IEEE International Conference on Multimedia and Expo, *IEEE*.

Soleymani, M., Kierkels, J.M., Chanel, G. and Pun, T. (2009), "A bayesian framework for video affective representation", in 2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, *IEEE*.

Souvannavong, F., Merialdo, B. and Huet, B. (2004), "Latent semantic analysis for an effective region-based video shot retrieval system", in *Proceedings of the 6th ACM SIGMM international workshop on Multimedia information retrieval - MIR '04, ACM Press*.

Staniucha, R. and Wojciechowski, A. (2016), "Mouth features extraction for emotion classification", in 2016 Federated Conference on Computer Science and Information Systems (FedCSIS), pp. 1685-1692.

Szajerman, D. and Napieralski, P. (2017), "Joint analysis of simultaneous EEG and eye tracking data for video picture", in 2017 18th International Symposium on Electro-magnetic Fields in Mechatronics, Electrical and Electronic Engineering (ISEF) Book of Abstracts, *IEEE*.

Szajerman, D. Warycha, M. Antonik, A. and Wojciechowski, A. (2016), "Popular brain computer interfaces for game mechanics control", doi: 10.1007/978-3-319-43982-211.

Trachel, R., Brochier, T. and Clerc, M. (2013), "Enhancing visuospatial attention performance with brain-computer interfaces", in *CHI '13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA '13*, ACM Press.

Treder, M.S., Bahramisharif, A., Schmidt, M.N., van Gerven, M. and Blankertz, B. (2011), "Brain-computer interfacing using modulations of alpha activity induced by covert shifts of attention", *Journal of NeuroEngineering and Rehabilitation*, Vol. 8 No. 1, pp. 24, doi: 10.1186/1743-0003-8-24.

van Gerven, M. and Jensen, O. (2009), "Attention modulations of posterior alpha as a control signal for two-dimensional braininterfaces", *Journal of Neuroscience Methods*, Vol. 179 No. 1, pp. 78-84, doi: 10.1016/j.jneumeth.2009.01.016.

Xu, M., Jin, J.S., Luo, S. and Duan, L. (2008), "Hierarchical movie affective content analysis based on arousal and valence features", in *Proceeding of the 16th ACM international conference on Multimedia - MM '08, ACM Press*.

Zheng, W.L., Dong, B.N. and Lu, B.L. (2014), "Multimodal emotion recognition using EEG and eye tracking data", in 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, *IEEE*.

Further reading

NeuroSky technologies (2015), "Specifications brainwave sensing headset".

Tobii, A.B. (2017), "Specifications for the tobii eye tracker 4C".

Corresponding author

Piotr Napieralski can be contacted at: piotr.napieralski@p.lodz.pl