



HAL
open science

Emotional interactive movie: adjusting the scenario according to the emotional response of the viewer

Romain Cohendet, Matthieu Perreira da Silva, Anne-Laure Gilet, Patrick Le Callet

► **To cite this version:**

Romain Cohendet, Matthieu Perreira da Silva, Anne-Laure Gilet, Patrick Le Callet. Emotional interactive movie: adjusting the scenario according to the emotional response of the viewer. *EAI Endorsed Transactions on Creative Technologies*, 2017, 4 (10), <10.4108/eai.4-9-2017.153053>. <hal-01675135>

HAL Id: hal-01675135

<https://hal.science/hal-01675135v1>

Submitted on 7 May 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

Emotional interactive movie: adjusting the scenario according to the emotional response of the viewer

Romain Cohendet^{1,*}, Matthieu Perreira Da Silva¹, Anne-Laure Gilet², and Patrick Le Callet¹

¹IRCCyN, UMR 6597, University of Nantes, France

²LPPL, EA 4638, University of Nantes, France

Abstract

Emotional interactive movie is a kind of film unfolding in different ways according to the emotion the viewer experiences. The movie is made of several sequences; their combination determines the particular scenario experienced. In this paper, we describe the system and its implementation by providing combination selection criteria. We measured neurophysiological and ocular activities of 60 individuals viewing an experimental interactive short-movie composed of 12 different scenarios. For this purpose, we combined an electroencephalography headset which directly provides emotional data with an eye-tracker in order to simultaneously investigate the position of viewer's gaze. From the collected data analysis, we propose a functional version of the emotional interactive movie, which was used in a so called "emotional cinema" during public exhibitions.

Keywords: Interactive movie, Emotiv Epoc headset, Arousal, Eye tracking

Received on 23 November 2015, accepted on 11 February 2016, published on 04 January 2017

Copyright © 2017 Romain Cohendet *et al.*, licensed to EAI. This is an open access article distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/3.0/>), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi: 10.4108/eai.4-9-2017.153053

1. Introduction

Art has always thrived on technological developments. New tools provide artists with new ways to express their creativity. Consequently, uses change and original forms of interaction emerge between artistic contents and individuals who experience it. Interactive film is one such example. It concerns not only the diffusion of audiovisual content created by the artist but also the way it is displayed, which depends on the viewer behavioral responses. Such interaction proposes a constantly renewed user experience, and probably increases the quality of the user's experience and the feeling of presence [1]. Additionally, it provides creators with the opportunity to improve their control over the user experience with the possibility to define the interaction criteria, allowing them to better convey their artistic intention.

'Interactive film' is a generic term that can refer in the literature to different concepts: interactive movie, a kind of

highly cinematic video game [e.g. 1]; interactive video, a system that allows proactive access to video content [e.g. 2]; interactive cinema, that give the audience a role in the way the movie is displayed [3]; interactive film as a form of interactive art, that aims to transform viewers into observers [4]. This paper focuses on the latter, with the following definition: interactive film is a film that considers viewers' reactions to provide them with a specific content. The film is cut into several sequences; their adjustment determines the particular coherent scenario which will be experienced. The sequences articulation – and consequently the resulting specific scenario – is built based on interactivity with the viewer, depending on its reactions. At certain moments of the film, called junctions, a sequence is selected from among several possible to assign the film a specific structure.

2. 'Emotional' interactive film

A film induces number of reactions (physical, attentional, emotional...) which can be measured by different devices

*Corresponding author. Email: romain.cohendet@univ-nantes.fr

(e.g. camera, eye tracking, motion detectors). Today there is a growing interest in human emotional reaction, which results in a number of studies constantly increasing. With the explosion of human-computer data exchanges, emotions are indeed recently considered as a solution to improve human-technology interactions [5]. To meet the need of researchers in emotional audiovisual material a database of 9800 emotional video clips has recently become available, showing the interest on this issue [6]. The emotional interactive film falls within this context. The viewer-film interaction is based on the viewers’ emotional reactions. Beyond a new form of digital entertainment, this film is an invitation to explore a new way of digital sharing and consumption.

Many current studies in relevant fields are based on a dimensional view of emotion, which regard emotions as having bipolar basic dimensions [7]. Most frequently studied dimensions are arousal and valence. Valence refers to attraction or aversion to a specific object or event and is part of a pleasant-unpleasant continuum. Arousal corresponds to a level of vigilance or activation, and range from calm to excited. Relative to valence data, arousal data can be quite easily collected by suitable devices (e.g. galvanic skin response sensor [8], heart rate monitors [9]). Recent technical possibilities for collecting arousal data have contributed to highlight new perspectives in computer sciences where emotion was not traditionally considered. The emerging field of affective computing identified several applications in which a real-time measure of arousal could be of great interest: human-machine interaction, assisted learning, perceptual information retrieval, human health, preventive medicine, creative arts and entertainment... [10]. The emotional interactive film presented in this paper, which is based on a real-time measurement of arousal induced in viewers, falls within the latter category.

3. System

The work presented in this paper is a preliminary work to Marie-Laure Cazin’s “emotional cinema” project [11] (see Figure 1). We describe in this section the system setup used for this preliminary work.

3.1. EEG device

We used an electroencephalographic (EEG) headset to measure the arousal activity of viewers watching the movie. Electroencephalography is a cerebral exploration method that measures real-time neurophysiological activity through electrodes placed on the scalp. The analysis of the recorded signals informs us about brain processes at work [12]. Scientific community interest for automatic recognition of emotion from EEG signals has grown up with the exponential development of new applications in the brain-computer interface domain and development of new forms of human-centered interactions with digital media [13]. EEG devices intended for a broad audience, handy to use and relatively affordable, have recently been released. One of

them is the Emotiv Epoc headset we used in our experiment, which directly provides emotional data on certain dimensions including arousal (called excitement) [14]. The validity and reliability of the device were verified regarding the raw data it provides [15], and it has been show that raw data can be used for emotion recognition [16].



Figure 1. The ‘emotional cinema’ project, in which our work is used. The spectator is equipped with an Epoc EEG headset. A secondary screen give him a feedback of his own emotional state.

3.2. The film used

The film we used, “Mademoiselle Paradis”, is a multi-scenario movie written and directed by Marie-Laure Cazin [11]. It lasts about 20 minutes (depending on the triggered scenario), and includes two junctions, with four different choices for the first junction and three for the second one (for a total of 12 different scenarios) (see Figure 2). Aside from the selectable sequences, the rest of the film is the same for all viewers.

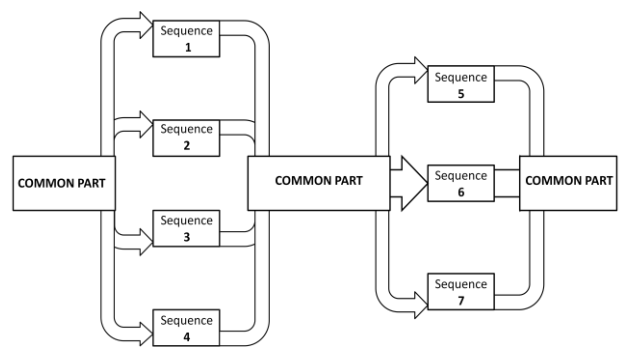


Figure 2. Repetitive and variable structures of *Mademoiselle Paradis*.

4. Experiment

In order to have a fully functional system for the first public exhibition, a number of issues had to be solved. This section

describes the different challenges that we had to address and the experiment we have conducted to solve them.

4.1. System requirements

A first need was to define a criterion to select a sequence at a junction. For this first prototype, we arbitrarily decided that the sequence that conveys the emotion the closest to the viewer's feeling should be selected. The idea was to follow the viewer's way of feeling to extend the experience he lived. The emotional measurement was the arousal data provided by the Emotiv EEG system. So the first goal of this experiment was to calculate an 'arousal score' for each selectable sequence, which would be compared with the emotional activity of the viewer while watching the film.

Secondly, we had to determine at which points in the film the viewers' arousal state would be recorded for the choice of the selectable sequences. These *recording* sequences had to meet several criteria: (1) be close in time to the junctions, so that the selected sequence is related to the current viewer excitement, (2) be non-neutral but induce a significant emotional arousal and (3) elicit sufficiently varied reactions in the audience (on the basis of which the viewers can be differentiated to adapt the scenario according to what they experience). With regard to point 2, we used the emotional data provided by the Emotiv software associated with the EEG headset to characterize the film sequences according to the arousal they induce. To meet the third criterion, we used an eye tracker to characterize the film sequences depending on the viewers' ocular behaviours. The commercial EEG device we used (as it was a will of Marie-Laure Cazin, for practical reasons in its using in Emotional cinema), provides emotional data from recordings carried out via the electrodes by using a black box model: we do not know what corresponds to these data. In particular, maybe the variation in the arousal is due to something else than the content displayed (e.g. noise due to body movements). An eye-tracker records eye movements of a viewer, from which can be deduced locations on the screen where he gazes [17]. By using it, we ensured that the audience' reactions were in relation with the film displayed. Eye tracking opens a window on brain processes: on cognitive processes (e.g. decision-making [18], memory [19], metacognitive monitoring [20]), and on emotional processes because the way of looking at a picture is related to the emotions it provokes (e.g. [21]). As ocular behaviour reflects underlying cerebral processes, we supposed the more varied it would be in the audience, the more different would be the processing performed by the different viewers. By characterizing the parts of the film as a function of their discriminative power derived from audience' ocular behaviours, and of their emotional intensity with the arousal measurement, we expected to find key sequences for emotional recording, that is to say emotional and which elicit varied reactions.

4.2. Participants

We conducted the experiment in the IRCCyN laboratory facilities (Nantes, France). A total of 60 people (33 males and 27 females) compensated for their participation took part in the experiment. Their age ranged from 19 to 61 years, with an average age of 24,1 years. All participants had normal or corrected – with glasses or contact lenses – visual acuity.

4.3. Material

We worked on four fixed versions of the film *Mademoiselle Paradis*: A, composed of sequences 1 (for junction 1) and 5 (for junction 2), B (sequences 2 and 6), C (sequences 3 and 7) and D (sequences 4 and 5). So, all selectable sequences were seen by at least one group of 15 viewers. The movie lasts about 20 minutes (depending on the version), with the first junction at 07:32 and the second at about 11:00. We worked only on the 11 minutes preceding the second junction (where the last junction occurs).

EEG data was acquired using Emotiv headset at 128 Hz, from its 14 electrodes [14]. The arousal was computed by means of the Emotiv intelligent framework that interprets the signals from each electrode to offer a real-time summary of the user's feeling. An SMI RED eye tracker was simultaneously used to collect gaze data. The device records 50 gaze points/sec for each eye.

4.4. Procedure

The experiment took place in an experiment room specifically designed and equipped for audiovisual quality assessment (Rec. ITU-R BT.509). Upon their arrival, the participants were divided into four groups corresponding to the four presented versions of the film. They were installed on a seat at a distance of 150cm (three times the height of the screen) from the 40-inch monitor (screen resolution was 1920×1080), and paired with the EEG headset. A 5-points calibration of the eye tracking system was done. The experimenter then started the movie.

4.5. Results

We conducted an analysis of the collected data in order to 1/ highlight parts of the film that resulted in both a significant arousal and a great inter-observer variability in the ocular behavior and 2/ calculate an 'arousal score' for each selectable sequence. Analysis was performed on 59 subjects for the eye tracker and 52 subjects for EEG (due to data recording problems) and focused on the description of the film shots. We chose to manually cut the movie (preceding the second junction) into shots because they are the shortest unit used in a movie to further express emotion, ideas and movement. To compare the two – arousal and gaze – measures, presented data are normalized (standard score).

4.5.1. Eye tracking data analysis

From the raw data collected by the eye tracker, we extracted the gaze points of each participant for each movie frame (see Figure 3). We then calculated the gaze dispersion within each frame (DF) as:

where $d(p, c)_i$ is the euclidian distance between the gaze points i and the centroid c and n the number of gaze points in the frame.

The mean gaze points dispersion per shot (DS) was calculated as:

with n the number of frames in the shot.

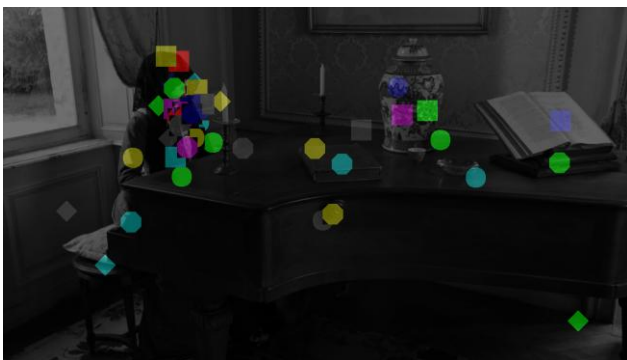


Figure 3. An image from *Mademoiselle Paradis*. Each colored geometric shape represents one viewer’s gaze point.

4.5.2. Arousal data analysis

From the arousal data provided by the Emotiv system, we defined each shot in terms of mean arousal recorded on participants. The goal was twofold: to describe each selectable sequence with a simple score of arousal intensity, and to determine “emotional shots” where the viewers’ reactions would be recorded.

The arousal score of a selectable sequence A_s viewed by n participants and made up of m arousal samples was calculated as:

with the sample j for the participant i .

The obtained values are presented in Table 1.

We performed the same calculation to obtain an arousal value for each shot.

Table 1. Mean arousal scores for the selectable sequences

Sequence number	Junction 1				Junction 2		
	1	2	3	4	5	6	7
Arousal score (A_s)	-0,69	-0,27	1,61	0,49	-1,31	1,05	0,02
SD	0,27	0,15	0,24	0,21	0,38	0,26	0,26

4.5.3. Discriminative film parts

To determine on which parts of the movie the spectator’s emotional activity would be recorded (i.e. the recording sequences we wanted all at once emotional and eliciting varied reactions in the audience), we plotted together arousal values and gaze points dispersions of the shots (see Figure 4). We chose the shots with both high arousal intensity and important gaze points dispersion to record emotional activity. For junction 1 (which occurs between shots 40 and 41), selected shots are 35 to 37 for a total length of 19 seconds. For junction 2 (which occurs between shots 58 and 59), selected shots are 47 to 49 for a total length of 17 seconds. Note that the shots 47 to 49 are time-shifted for the other versions of the film (here they apply to version A) and therefore must be shifted accordingly.

We found a positive linear correlation between arousal and GP dispersion per shot (DS) ($r=.30, p<.05$), suggesting first, that the arousal the Emotiv EEG provides is not independent of the displayed content, and second that the way a person looks at a scene is related to the emotion he feels. However, this suggestion should be taken with caution, because – as we aforementioned – we do not know what corresponds to the arousal the Emotiv software provides; maybe, for example, the movements of the eyes are captured by the electrodes on the scalp and taken into account in the arousal opaque calculation (eye movements are a well-known artefact in EEG signals).

In addition, it is interesting to note that the mean arousal values of the shots were positively correlated ($r=.65, p<.001$) with the arousal dispersions of the shots (measured as the SD of the arousal values collected for each shot), suggesting that the most emotional sequences are also the sequences where the emotional reactions are the most varied.

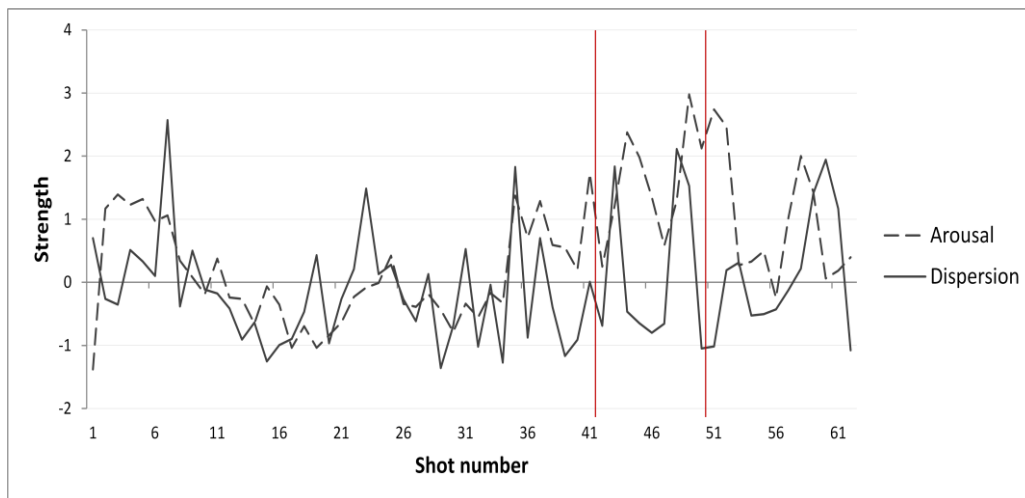


Figure 4. Arousal and gaze points dispersion scores per shot (version A). Data are normalized. Vertical red lines correspond to junctions 1 and 2.

4.5.4. Criterion to select a sequence

We defined the criterion to select the sequence as: “the sequence for which arousal score is the closest to the arousal score of the particular viewer watching the movie (calculated in real-time as a mean on the film parts defined in the precedent section) is selected”. So, the selected sequence (SS) at a junction is the one that having the lowest distance d :

where A_v is the arousal score for the particular viewer, A the arousal score of the selectable sequence previously calculated, j is the junction number (1 or 2) and k is the selectable sequence number.

5. Discussion

Our paper presents a pilot study which aimed to propose a first functional adaptation of a movie that changed content depending on the emotions felt by the viewer. We have selected shots where it is relevant to record the viewer’s emotional reactions, and proposed a criterion to adapt the film content according to them. This work has been used in the Marie-Laure Cazin’s emotional cinema, of which the first public exhibition took place at Stereolux in Nantes (France), on 29 and 30 April 2014.

From its first exhibition, it appeared that some selectable sequences tended to be a little more often selected than others. This inequality was not important enough, however, that we decided to correct it for subsequent public exhibitions, by adjusting the selection of sequences according to their probability of occurrence.

In the situation where more than one viewer interacts with the film, as was the case for emotional cinema, where two spectators were equipped with an EEG device, the selection of a sequence at a junction depended of only one spectator, which changed at the subsequent junction. We can think of other ways; for example, in a family viewing, we could imagine taking into account only the viewer who feels the more intense emotion – this could increase the viewers’ immersion in the film.

Part of our study aimed at characterizing the film shots according to the arousal they induced and ocular behaviors they generated on a set of viewers, using data provided by the Emotiv EEG device and the eye tracker. The way we processed had required an experiment with participants; to avoid this part, other methods could be adopted. Different approaches are proposed in the literature to computationally infer emotional information from a film. The approach proposed by [22], based on convolutional neural networks, enables to calculate dimensional affective scores for 1-second video portions, on the arousal and valence dimensions (the model is described in details in [23]). Another approach was proposed by Soleymani et al. [24], based on a Bayesian classification framework, to classify film scenes into three affective categories: ‘calm’, ‘positive excited’, or ‘negative excited’. Malandrakis et al. [25] have also proposed a system, based on hidden Markov models, to predict intended emotion video frames convey. By using one of these methods, one could automatically label film sequences according to the emotional information they contain (that is, without going through a preliminary experiment requiring several participants). In the same way, computer vision algorithms enable to extract visual saliency from a frame (e.g. the GBVS model [26]). We could use this predicted visual saliency as an indicator to characterize the film parts according to the probability

they would generate more or less different ocular behaviours in the audience (due to the more or less important number of regions of interest).

We chose an interaction with the film based on emotional arousal: other criteria could have been chosen. In conjunction with the development of brain computer interfaces, methods have been proposed to infer the emotional valence from EEG signals (e.g. [27, 28]). If the real-time measurement of emotional valence is less simple and (maybe) less reliable than arousal measurement, the incorporation of this aspect of the user experience could enrich the interaction, and allow the artist to better enforce his artistic intention. We must add that it is possible to match discrete affective states (such as sadness, relaxation, joy) with conjunctions of values on the arousal and valence bipolar axes [29]; therefore, adding the valence information could allow the system to interact with the viewers through words that make more sense than ‘arousal’ or ‘valence’ for him.

Also, as aforementioned, other measurement techniques could be used to capture emotion in real time. Emotion recognition is a very active field, since emotional information is highly valuable in number of applications, and development of new tools to collect and analyze emotional data accompany fundamental advances. The industry surely will make fruitful of such tools, in particular through two topics related to emotional movie we are very interested in: the personalization of shared digital contents by adding data relative to user experience in delivering systems, and the autonomous creation of digital artworks from contents collected in databases and structured based on users’ feedback. The latter is probably the next big step for digital art.

6. Acknowledgments

We thank Marie-Laure Cazin, director of the film *Mademoiselle Paradis*, and Stereolux, a cultural association located in Nantes (France), which was also part of this project.

References

- [1] Hu, J., Janse, M., & Kong, H. J. (2005). User experience evaluation of a distributed interactive movie. In *HCI International*.
- [2] Zhang, D., Zhou, L., Briggs, R. O., & Nunamaker, J. F. (2006). Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information & management*, 43(1), 15-27.
- [3] Hales, C. (2015). 3 Interactive Cinema in the Digital Age. *Interactive Digital Narrative: History, Theory and Practice*, 36.
- [4] Edmonds, E., Turner, G., & Candy, L. (2004, June). Approaches to interactive art systems. In *Proceedings of the 2nd international conference on Computer graphics and interactive techniques in Australasia and South East Asia* (pp. 113-117). ACM.
- [5] Picard, R. W. (2010). Affective computing: from laughter to IEEE. *Affective Computing*, *IEEE Transactions on*, 1(1), 11-17.
- [6] Baveye, Y., Dellandrea, E., Chamaret, C., & Chen, L. (2015). LIRIS-ACCEDE: A Video Database for Affective Content Analysis. *Affective Computing*, *IEEE Transactions on*, 6(1), 43-55.
- [7] Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *Journal of behavior therapy and experimental psychiatry*, 25(1), 49-59.
- [8] Shi, Y., Ruiz, N., Taib, R., Choi, E., & Chen, F. (2007, April). Galvanic skin response (GSR) as an index of cognitive load. In *CHI'07 extended abstracts on Human factors in computing systems* (pp. 2651-2656). ACM.
- [9] Wulfert, E., Roland, B. D., Hartley, J., Wang, N., & Franco, C. (2005). Heart rate arousal and excitement in gambling: Winners versus losers. *Psychology of Addictive Behaviors*, 19(3), 311.
- [10] Picard, R. W., & Picard, R. (1997). *Affective computing* (Vol. 252). Cambridge: MIT press.
- [11] www.marielaurecazin.net/; www.facebook.com/lecinemaemotif
- [12] Sanei, S., & Chambers, J. A. (2008). *EEG signal processing*. John Wiley & Sons.
- [13] Liu, Y., Sourina, O., & Nguyen, M. K. (2010, October). Real-time EEG-based human emotion recognition and visualization. In *Cyberworlds (CW), 2010 International Conference on* (pp. 262-269). IEEE.
- [14] Emotiv, E. E. G. (2012). *System Electroencephalography*. URL <http://www.emotiv.com/>.
- [15] Debener, S., Minow, F., Emkes, R., Gandras, K., & Vos, M. (2012). How about taking a low-cost, small, and wireless EEG for a walk?. *Psychophysiology*, 49(11), 1617-1621.
- [16] Pham, T. D., & Tran, D. (2012, January). Emotion recognition using the emotiv epoc device. In *Neural Information Processing* (pp. 394-399). Springer Berlin Heidelberg.
- [17] Duchowski, A. (2007). *Eyetracking methodology: Theory and practice* (Vol. 373). Springer.
- [18] Glöckner, A., & Herbold, A. K. (2011). An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making*, 24(1), 71-98.
- [19] Mancas, M., & Le Meur, O. (2013, September). Memorability of natural scenes: The role of attention. In *Image Processing (ICIP), 2013 20th IEEE International Conference on* (pp. 196-200). IEEE.
- [20] Roderer, T., & Roebers, C. M. (2014). Can you see me thinking (about my answers)? Using eye-tracking to illuminate developmental differences in monitoring and control skills and their relation to performance. *Metacognition and learning*, 9(1), 1-23.
- [21] Schupp, H. T., Stockburger, J., Codispoti, M., Junghöfer, M., Weike, A. I., & Hamm, A. O. (2007). Selective visual attention to emotion. *The Journal of Neuroscience*, 27(5), 1082-1089.
- [22] Baveye, Y., Dellandrea, E., Chamaret, C., & Chen, L. (2015). Deep learning vs. kernel methods: Performance for emotion prediction in videos. In *Affective Computing and Intelligent Interaction (ACII), 2015 International Conference on* (pp. 77-83). IEEE.

- [23] Baveye, Y. (2015). Automatic prediction of emotions induced by movies (Doctoral dissertation, Ecole Centrale de Lyon).
- [24] Soleymani, M., Kierkels, J. J., Chanel, G., & Pun, T. (2009). A bayesian framework for video affective representation. In *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on* (pp. 1-7). IEEE.
- [25] Malandrakis, N., Potamianos, A., Evangelopoulos, G., & Zlatintsi, A. (2011). A supervised approach to movie emotion tracking. In *Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on* (pp. 2376-2379). IEEE.
- [26] Harel, J., Koch, C., & Perona, P. (2006). Graph-based visual saliency. In *Advances in neural information processing systems* (pp. 545-552).
- [27] Schaaff, K. (2008). EEG-based emotion recognition. Universitat Karlsruhe (TH).
- [28] Horlings, R., Datcu, D., & Rothkrantz, L. J. (2008). Emotion recognition using brain activity. In *Proceedings of the 9th international conference on computer systems and technologies and workshop for PhD students in computing* (p. 6). ACM.
- [29] Valenza, G., Lanata, A., & Scilingo, E. P. (2012). The role of nonlinear dynamics in affective valence and arousal recognition. *Affective Computing, IEEE Transactions On*, 3(2), 237-249.