Emotion recognition in computer games and films

IMIE NAZWISKO

Abstract—In last years technology used in game and film creations has formed need to check reaction of users on watched image. Human body react on external stimulus by face microchanges, distortions in electroencephalography, pupil adjustments etc. Those processes can be recorded by specified apparatus thus correct analysis of those characteristics can be automated. Thanks to this authors are able to check reaction of viewers on their creations, or even construct algorithms that can do it automatically.

Index Terms—emotion recognition, pupil reflex, EEG, electroencephalography, emotion clasificcation

1 Introduction

Studies on recognition of the human emotions can be useful at many areas. Starting with psychology studies on behavioural disorder with patient that have problems with expressing emotions, through biology studies on creation of emotions in human body and ending with getting feedback from watched movie. Emotions allows to decide if user like what he see or not. That gives them an opportunity to choose if he wants to end it immediately, or even repeat those emotions again. For artists this informations is very desirable, because they can refine theirs creations based on information gathered under the influence of the viewer's reaction. Thanks to those researches artist will know when user will be more interest in action, and where it will be more dull or touching for them.

In [1] authors are creating theory witch explain generation of emotions in human body. They simplify it to few steps, like in algorithm. First there is a perception of an event then analysis of it based on user's own experience and norms, so finally the event could be classified as certain emotion.

Emotions can be detected by certain characteristics that could be classified to two groups:

- psychological:
  - EEG(electroencephalography),
  - EMG(electromyography),
  - EKG(electrocardiography),
  - pupil diameter.
- non-psychological:
  - text,
  - speech,
  - gestures,
  - facial expressions.

This paper will be focused on group of psychological signals, especially on EEG and pupil diameter. It will be explained how to detect certain emotion based on fusions of the stimuli. There are plenty of researches where authors combine EEG with pupil diameter or even with eye trackers data, and those combined methods are more reliable and with better accuracy then individual ones [2], [3], [6].

- *Lodz University of Technology, Lodz, Poland, EMALJA*

1.1 Subjects and stimuli

Using variety of movie clips, especially selected to those research and shown to participators, the EEG signal and pupil diameter changes was recorded. Key feature of those movie clips is to cover different emotional responses to get results as best and accurate as possible. Psychologists recommended videos from 1 to 10 minutes long for elicitation of single emotion [15].

2 EEG

One of the most popular methods of emotion recognition are based on analysis of electroencephalography signals. Numerous researches [7]–[9] has shown that the brain activity, which EEG collect, is the most reliable source for emotion recognition. Main core of those studies is to find brain regions and frequency bands most related to those emotions. Studies of [11] showed that activation for unpleasant emotions was prominent over the right posterior regions in the alpha band. In [12] authors found that frontal brain electrical activity is closely related to musical emotions, and in [13] authors confirmed theory that gamma band is also related to music emotions.

2.1 Data acquisition

Gathering data of brain activity is done be special EEG cap, where *AgCl* electrodes placed on it are collecting brain activity in certain areas. Most common used layout of electrodes is 10-20 system, shown in Figure 1.

Signals were recorder mostly in 1000–1024 Hz sampling rate. To speed up calculations those characteristics were down sampled to 200–256 Hz. Noises and artefacts reduction were done by applying bypass filter between 0.5 to 70 Hz.

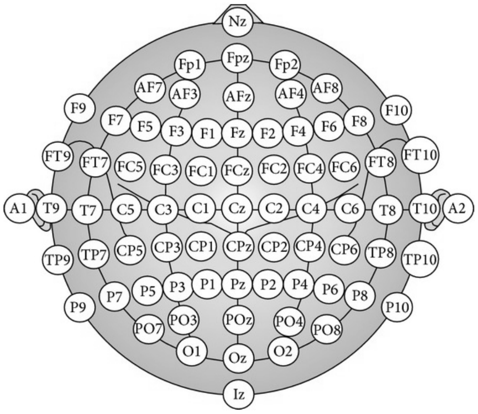

Fig. 1. The EEG cap arrangement for 10-20 system. [14]

2.2 Data extraction

Correlation of certain spectral power of EEG signal and emotions relevant processing was observed [16]. There are multiple methods of extracting power spectral density (PSD) from raw signals. Two of them will be expounded.

First [6] use Fourier transform and Welch algorithm. This method split signal into overlapping segments and the PSD is estimated by averaging the periodograms, in result the power spectrum is smoother. PSD of individual electrode was estimated using 15s long windows with 50 percent overlapping. PSD bands like theta (4 Hz $<$ f $<$ 8 Hz), slow alpha (6 Hz $<$ f $<$ 10 Hz), alpha (8 Hz $<$ f $<$ 12 Hz), beta (12 Hz $<$ f $<$ 30 Hz) and gamma 30Hz $<$ f were extracted from electrodes. In additional 14 symmetrical pairs on the right and left hemisphere was extracted to measure possible asymmetry in brain activity.

Second one use more short-time Fourier transform with non-overlapped Hanning window of 4 s. In addition to PSD the differential entropy (DE), differential asymmetry (DASM) and rational asymmetry (RASM) were extracted and compared. Like in first method five frequency bands was used. Delta (1 Hz $<$ f $<$ 3 Hz), theta (4 Hz $<$ f $<$ 7 Hz), alpha (8 Hz $<$ f $<$ 13 Hz), beta (14 Hz $<$ f $<$ 30 Hz) and gamma (31 Hz $<$ f $<$ 50 Hz). Using equation 1, DE was calculated.

$$h(X) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} exp\frac{(x-\mu)^2}{2\sigma^2} log\frac{1}{\sqrt{2\pi\sigma^2}} exp\frac{(x-\mu)^2}{2\sigma^2} dx = \frac{1}{2} log 2\pi e \sigma^2 \tag{1}$$

where X is Gauss distribution $N(\mu, \sigma^2)$, x is a variable π and e are constants. DASM and RASM are defined as :

$$DASM = h(X_{LEFT}) - h(X_{RIGHT}) \tag{2}$$

$$RASM = h(X_{LEFT})/h(X_{RIGHT}) \tag{3}$$

where X_{LEFT} andX_{RIGHT} are DE features of left and right hemisphere of brain.

2.3 Classification

After the data was collected and extracted the support vector machine (SVM) was used as classifier, in both examples. In [6] they smoothing features using linear dynamic system (LDS).

Feature		Frequency Bands					
		Delta	Theta	Alpha	Beta	Gamma	Total
PSD	Mean	51.60	51.87	54.74	53.23	51.36	**59.04**
	Std	19.56	14.48	16.58	18.06	16.10	20.31
DE	Mean	70.51	47.98	60.18	64.29	68.73	**71.77**
	Std	12.18	15.19	12.94	23.05	20.30	12.03
DASM	Mean	61.08	43.42	49.98	46.96	64.12	**68.37**
	Std	22.45	19.45	15.59	15.21	22.94	23.86
RASM	Mean	61.44	44.90	48.69	48.18	62.71	**66.03**
	Std	22.90	12.14	14.62	15.93	21.11	24.62
ASM	Mean	65.18	44.78	50.29	45.19	63.92	**67.91**
	Std	22.32	13.87	15.91	12.77	22.19	24.45

Fig. 2. The performance of classifiers in % using different kinds of frequency band features. For [6].

Result of classification can be bee seen at Figure 2. ASM feature is concatenation of DASM nad RASM. As we can see, Delta and Gamma frequency bands perform better than Theta and Alpha frequency bands, and total frequency band has a stable and prominent accuracy. Also we can find that, differential entropy features get best accuracies in almost all frequency bands except Theta band (47.98% of DE features is less than 51.87% of PSD features).

In [3] in EEG there was only DE feature. They have used SVM classification with RBF kernel. The linear discrim-

Arousal classification			Valence classification		
Band	Electrode/s	$\sigma^2_{bw}/\sigma^2_{wn}$	Band	Electrode/s	$\sigma^2_{bw}/\sigma^2_{wn}$
Slow α	PO4	0.18	β	T8	0.08
α	PO4	0.17	γ	T8	0.08
θ	PO4	0.16	β	T7	0.07
Slow α	PO3	0.15	γ	T7	0.06
θ	Oz	0.14	γ	P8	0.05
Slow α	O2	0.14	γ	P7	0.05
Slow α	Oz	0.14	θ	Fp1	0.04
θ	O2	0.13	β	CP6	0.04
θ	FC6	0.13	β	P8	0.04
α	PO3	0.13	β	P7	0.04

Fig. 3. The performance of classifiers using different kinds of frequency band features. For [3].

ination criterion was calculated for EEG signals.Dividing between class variance by within class variance for any given feature. For arousal classification, PSD in alpha bands of occipital electrodes was found to be the most discriminant features. In contrast, valence beta and gamma bands of temporal electrodes are more informative. The between class to within class variance ratios are higher for the best arousal EEG features. The higher linear discrimination criterion for best arousal features explains the superior classification rate for arousal dimension.

3 Pupil diameter

Based on human eye observation the theory was forged. Pupil diameter is changing in different emotional states. Disadvantage of this solution is that pupil is highly dependant on the light. First step to gather data from pupil is to create lighting reflex model. Most common and simplest method is principal component analysis (PCA).

3.1 Light reflex model

Assuming that Y is the $M \times N_p$, of pupillary response for the same picture for N_p participant and M samples, Y consists of three components

$$Y = A + B + C \tag{4}$$

where A is the lighting influence, B is the emotional response and C is the noise. Extracting first component from PCA to approximate the pupil response for the lightning changes during the experiments.

3.2 Data acquisition

To gather data the Eye-Tracker devices was used. Those apparatus are able to collect position of the projected eye gaze on the screen, pupil diameter, the moments when the eyes were closed and distance of the participant's eyes to the gaze tracker device. Eye blinking creates gaps in eye gaze and pupil record, thus the linear interpolation was used to replace missing samples.

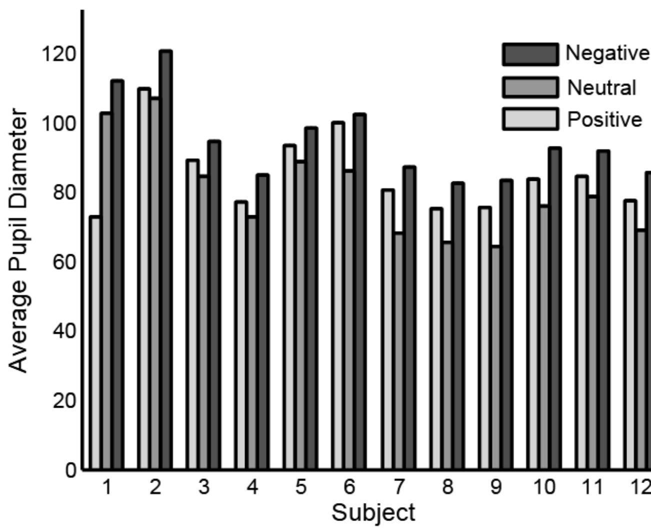

Fig. 4. Average pupil diameter. For [6].

As it can be seen at Figure 4 for 12 participants the smallest average value have neutral emotion, except subject 1.

At Figure 5 the examples of pupillary responses, extracted pupillary lighting reflex, and the residual component after removing the light reflex are given.

3.3 Classification

As it can be seen at Figure 6 the DE feature performs much better than PSD, because DE features have the balance ability of discriminating patterns between lowand high frequency energy.

4 Multimodial Fusion

After the data was gathered the fusion of methods have left. Used in this article examples have implemented the feature level fusion, decision level fusion.

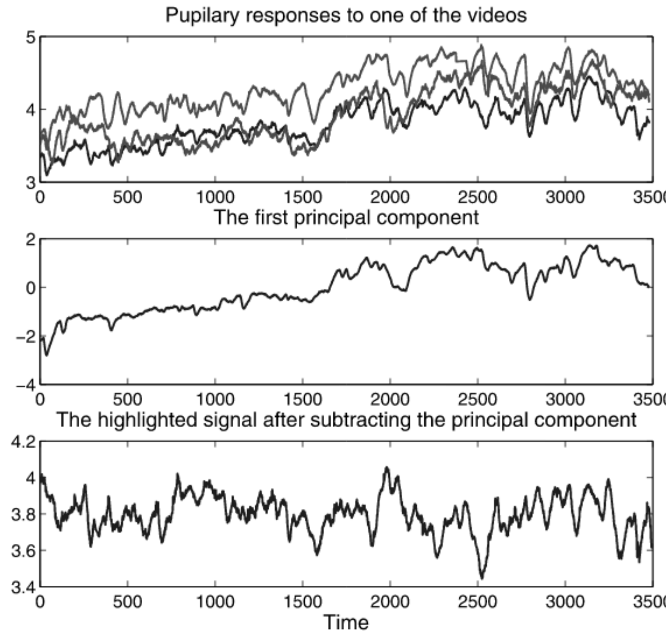

Fig. 5. From top to bottom: In the first plot, there is an example of pupil diameter measures from three different participants in response to one video. The second plot shows the first principal component extracted by PCA from the time series shown in the first plot (the lighting effect). The bottom plot shows the pupil diameter of the blue signal in the first plot after reducing the lighting effect [3]

Exp	Feature	Accuracy	Exp	Feature	Accuracy
1	PSD	65.43	7	PSD	33.95
	DE	**86.42**		DE	**61.73**
2	PSD	56.79	8	PSD	46.91
	DE	**70.37**		DE	**50.62**
3	PSD	54.94	9	PSD	43.83
	DE	**56.79**		DE	**59.88**
4	PSD	60.49	10	PSD	36.42
	DE	**63.58**		DE	**59.88**
5	PSD	37.65	11	PSD	44.44
	DE	**48.77**		DE	**50.62**
6	PSD	33.95	12	PSD	34.57
	DE	**47.53**		DE	**50.62**
Mean	PSD	45.78	Std	PSD	11.03
	DE	**58.90**		DE	**10.25**

Fig. 6. Performence in % of using different features from pupil diameter [6]

5 Conclusion

Emotions are sophisticated mechanism in human body, but knowledge how they work can be helpful in many areas. Those signals can be obtained from many of human impulses, such as pupil reflex or brain signals. Using appropriate techniques and devices those characteristics can be collected and analysed to detect emotions.

References

[1] A. Ortony, G.L. Clore, A. Collins *The Cognitive Structure of Emotions.*, Cambridge University Press, July 1988.

[2] R. A. Calvo, S.D'Mello *Affect detection: An interdisciplinary review of models, methods and their applications*, IEEEE Transactions on Affective Computing, vol 1, no1, pp 18-37, 2010

[3] M. Soleymani, M. Pantic, T. Pun *Multimodal Emotion Recognition in Responce to Videos*, IEEE Transaction on Affective Computing, vol. 3, no. 2, april-June 2012

[4] R. Adolphs, D. Tranel, A.R. Damasio, *Dissociable Neural Systems for Recognizing Emotions*, Brain and Cognition, vol. 52, no. 1, pp. 61-69, June 2003.

[5] A.R. Damasio, T.J. Grabowski, A. Bechara, H. Damasio, L.L.B. Ponto, J. Parvizi, and R.D. Hichwa, *Subcortical and Cortical Brain Activity during the Feeling of Self-Generated Emotions*, Nature Neuroscience, vol. 3, no. 10, pp. 1049-1056, Oct. 2000.

[6] Z. Wei-Long, D. Bo-Nan , L. Bao-Liang *Multimodal Emotion Recognition using EEG and Eye Tracking Data*, IEEE, 2014

[7] Y. P. Lin et al., "EEG-Based Emotion Recognition in Music Listening," in IEEE Transactions on Biomedical Engineering, vol. 57, no. 7, pp. 1798-1806, July 2010. doi: 10.1109/TBME.2010.2048568 http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5458075&isnumber=5484937

[8] Y. Gao, H. J. Lee and R. M. Mehmood, "Deep learninig of EEG signals for emotion recognition," 2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Turin, 2015, pp. 1-5. doi: 10.1109/ICMEW.2015.7169796 http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7169796&isnumber=7169738

[9] Dan Nie, Xiao-Wei Wang, Li-Chen Shi, and Bao-Liang Lu, EEG-based Emotion Recognition during Watching Movies, Proceedings of the 5th International IEEE EMBS Conference on Neural Engineering Cancun, Mexico, April 27 - May 1, 2011 ,https://pdfs.semanticscholar.org/6511/590bc9677922c82747b5d183383f46b50db6.pdf

[10] J. Zhai and A. Barreto, "Stress Detection in Computer Users Based on Digital Signal Processing of Noninvasive Physiological Variables," 2006 International Conference of the IEEE Engineering in Medicine and Biology Society, New York, NY, 2006, pp. 1355-1358. doi: 10.1109/IEMBS.2006.259421 http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4462012&isnumber=4461641

[11] M. Sarlo, G. Buodo, S. Poli, and D. Palomba, Changes in EEG alpha power to different disgust elicitors: the specificity of mutilations, Neuroscience Letters, vol. 382, no.3, pp. 291-296, 2005.

[12] L. A. Schmidt, and L. J. Trainor, Frontal brain electrical activity distinguishes valence and intensity of musical emotions, Cognition and Emotion, vol. 15, no. 4, pp. 487-500, 2001.

[13] M. Li, and B. L. Lu, Emotion classification based on gamma-band EEG, IEEE Int. Conf. Engineering in Medicine and Biology Society, Minneapolis, 2009, pp. 1223-1226.

[14] Jirayucharoensak, Suwicha Pan-ngum, Setha Israsena, Pasin. (2014). EEG-Based Emotion Recognition Using Deep Learning Network with Principal Component Based Covariate Shift Adaptation. TheScientificWorldJournal. 2014. 627892. 10.1155/2014/627892.

[15] Alexandre Schaefer and Frdric Nils and Xavier Sanchez and Pierre Philippot, Assessing the effectiveness of a large database of emotion-eliciting films: A new tool for emotion researchers, 2010 http://dx.doi.org/10.1080/02699930903274322

[16] Aftanas, L.I., Savotina, L.N., Makhnev, V.P. et al. Neurosci Behav Physiol (2005) 35: 951.https://doi.org/10.1007/s11055-005-0151-9