"BioVid Emo DB": A Multimodal Database for Emotion Analyses validated by Subjective Ratings

Lin Zhang[1], Steffen Walter[1], Xueyao Ma[1], Philipp Werner[2], Ayoub Al-Hamadi[2], Harald C. Traue[1], and Sascha Gruss[1]

[1] Departement of Medical Psychology, University of Ulm, Germany
{lin.zhang, sascha.gruss}@uni-ulm.de
[2] Institute for Information Technology and Communications, University of Magdeburg, Germany

***Abstract*—The precondition of productive data mining is having an efficient database to work on. The BioVid Emo DB is a multimodal database recorded for the purpose of analyzing human affective states and data mining related to emotion. Psychophysiological signals such as Skin Conductance Level, Electrocardiogram, Trapezius Electromyogram and also 4 video signals were recorded. 5 discrete emotions (amusement, sadness, anger, disgust and fear) were elicited by 15 standardized film clips. 94 participants watched them, rated them in terms of the experienced emotional level and selected the film clips that evoked the strongest emotion. A preliminary analysis of the subjective ratings made during the experiment is presented. The dataset is available for other researchers.**

Keywords—database; data mining; emotion recognition; film clips; discrete emotions

I. Introduction

Human emotions are complicated and interesting and our scientific knowledge about them is still very limited. A series of changes in psychophysiological activity, brain activity and facial expressions can be detected when a person experiences stress or emotions [1], [2]. Affective computing provides a possibility to identify emotional states of humans, even to predict a users´ needs and wishes through various machine learning algorithms in human-computer interactions [3].

The field of emotion classification has widely attracted the focus of computer scientists, psychologists and cognitive scientists since 1995 [3]. Diverse feature selection and classification algorithms have been utilized during this period [4]. But to the best of our knowledge, available open databases on emotions are still underrepresented. Moreover, in most of those databases, the elicited emotions were based on the dimensional rather than on the discrete emotional model [5], [6], [7]. Therefore, building a discrete emotion database was our primary intention for analyzing and classifying discrete emotions.

II. BioVid Emo DB

The BioVid Emo DB contains high-quality data of induced discrete basic emotions and is now accessible to other research groups for, e.g. testing new classification techniques or further data mining in terms of emotions, see Table I. For each emotion, three film clips have been presented during the experiment. For each participant, only the data of the video is provided for which they felt the strongest emotion, e.g. if a participant felt most disgusted by the first 'disgust' clip, only the video signals and psychophysiological parameters of this particular clip is provided in the database. Two different theoretical frameworks, a discrete and a dimensional model of emotions, are widely used in present-day studies during emotion induction due to their perspective advantages [8], [9]. The reason the first model is used in the present paper is that emotions based on the discrete model are more straightforward and specifiable compared to the dimensional model. This is necessary when addressing some aspects regarding emotional states, e.g. anger and disgust can be described explicitly in the discrete model but both are characterized by low valence and high arousal in the dimensional model [10]. Further, discrete emotions can be easily projected into the dimensional affective space which is the valence-arousal-dominance space (VAD space), e.g. through self-reports.

There are many kinds of standardized methods to induce emotions in the laboratory. External stimuli include pictures [11], [12], film clips [13], [14] and music [15], while internal stimuli refer to self-paced imagination methods [16] or autobiographical memory tasks [17]. Among them, emotion elicitation using film clips has obvious merits. For instance, it is a dynamic and multichannel method. Furthermore, according to meta-analyses, film clips seem to be one of the most effective ways to elicit emotions [18].

III. Experimental Design

A. Participants

A total of 94 subjects participated in the experiment, which had been recruited from three different age groups: 18-35 years (35 subjects, consisting of 16 men and 19 women), 36-50 years (31 subjects, consisting of 13 men and 18

TABLE I. BioVid Emo DB

BioVid Emo DB	
Method of emotion elicitation	***Data availability***
Film clips 3×5 conditions	Video signal Psychophysiological parameters Data as .txt files

Sponsors: Transregional Collaborative Research Center SFB/TRR 62 Companion Technology for Cognitive Technical Systems of the German Research Foundation (DFG) and China Scholarship Council (CSC)

women) and 51-65 years (28 subjects, consisting of 15 men and 13 women).

None of them had affective or related disorders. Recruitment was carried out through notices posted at the university for the 18-35 age group and through the press for the other two age groups.

Due to missing and/or corrupted recordings, only the data of 86 subjects are available in the database.

B. *Film Clips for Emotion Elicitation*

Referring to previous studies [19], [27], 15 standardized film clips were selected in total for inducing the following five different basic emotions: amusement, sadness, anger, disgust and fear (three clips for each emotion). In terms of denoting these emotions, we stick to [19]. The lengths of the clips vary from 32 to 245 seconds (average length 68 s). A description of the film clips used is given in Table II.

C. *Setup*

The experiment was performed in a quiet laboratory environment with controlled illumination. The indoor lighting consisted of LED panels which imitated evenly daylight and it was consistent for all subjects. Both the door and the curtains were kept closed in order to keep the room quiet and block the natural light during the experiment. All film clips were displayed on a 19-inch monitor with a resolution of 1440 × 900 pixels. All participants sat roughly 1 meter from the screen and were asked whether the volume of film clips was pleasant before the actual start of the experiment.

For recording psychophysiological data and trigger information, Biotrace software and a Nexus-32 amplifier (www.mindmedia.nl) with a sampling rate of 512 Hz were used. For video recording, three AVT Pike F145C cameras and a Kinect Sensor were utilized. One of the Pike cameras was fixed in front of the participant while the other two were placed at the left and right side for capturing a frontal face in case the participant turned its head 45° to the left or right respectively.

All three cameras were triggered synchronously at a frame rate of 25 Hz and recorded at a resolution of 1388 × 1038 colored pixels. The Kinect Sensor was set above the frontal Pike camera to capture depth maps (640 × 480 pixels, ca. 30 Hz) and color images (1280 × 1024 pixels, ca. 10 Hz) [23].

The synchronization of the Kinect and Pike video streams was done by some sort of clapperboard action in both streams. In order to synchronize all video streams with the psychophysiological data, the Nexus-32 device was utilized to record a frequency divided version of the camera trigger signal along with the biosignals [23]. The whole experimental setting is shown in Fig. 1.

TABLE II. DESCRIPTION OF THE FLIM CLIPS

Emotion	Sequence	Film title	Length (s)	Clip Description
Amusement	1	On Golden Pond	32	A woman named Ethel is walking through the forest while she meets her daughter Chelsea.
	2	When Harry met Sally	149	Harry and Sally are old friends. They are discussing whether Harry can notice it if a woman fakes an orgasm.
	3	An Officer and a Gentleman	118	A man kisses a woman and carries her out of a factory.
Sadness	1	The Champ	245	A boxer is seriously injured and dying while his son enters.
	2	The Killing Fields	83	Pran has a weeping farewell with his friends.
	3	An Officer and a Gentleman	101	A friend of the couple had already died before they find him.
Anger	1	Witness	91	A group of Amish is insulted and harassed by teenagers.
	2	Gandhi	128	Gandhi is treated unfairly and beaten by a policeman.
	3	Cry Freedom	167	A group of Blacks is attracked and some teenagers and children are shot by soldiers.
Disgust	1	The Godfather	66	The head of a dead and blooding horse is found in Jack´s bed when he wakes up in the morning.
	2	Maria´s Lovers	68	Ivan is attacked by a rat while sleeping.
	3	Pink Flamingos	34	A woman eats a dog´s feces.
Fear	1	Silence of the Lambs	205	A young woman follow a dangerous and anomalous killer into a dark basement.
	2	Halloween	208	A woman enters a strange house in the midnight, which is full of dead.
	3	Marathon Man	125	A man is tortured by some people.

D. Procedure

At the beginning of the experiment, an instruction was displayed to the participant: *By clicking the 'next' button several film clips will be shown to you. Please express your facial expression freely and spontaneously. Don´t try to suppress your emotions.*

For the elicitation of each of the five basic emotions, three film clips were presented sequentially without a pause. After each 3-clip basic-emotion-presentation the participants had to fill in self-assessments questionnaires. They had to select the film with the strongest emotional content and give a rating for valence (ranging from unpleasant to pleasant), arousal (ranging from calm to excited/activated), amusement, sadness, anger, disgust and fear on nine points´ scales. Following that, the subjects were given a pause of 2 minutes while displaying a light-yellowish screen with a small black fixation cross in order to neutralize any affective state (see Fig. 2).

E. Measured Parameter

1) Psychophysiological signals

a) Skin Conductance Level (SCL): Two electrodes of the sensor were positioned on the distal phalanges of the index and ring fingers. The electrodermal activity is considered as a good indicator of an affective state of a person because it is under strict control of the sympathetic nervous system. When a person is experiencing states such as stress, excitement or surprise, the skin conductance level shows fluctuations after one to three seconds due to eccrine sweat gland activity [20].

b) Electrocardiogram (ECG): Average action potential of the heart was measured using two Ag/AgCl electrodes on the skin. One electrode was positioned on the upper right and one on the lower left of the body. Commonly derived features of the ECG are heart rate (HR), interbeat interval and heart rate variability (HRV). ECG, which reflects the variation of the cardiac electrical potential over time, can predict affective states to some extent. For example, HRV has been studied in association with anxiety and hostility and HR has been used to distinguish between different emotions [21].

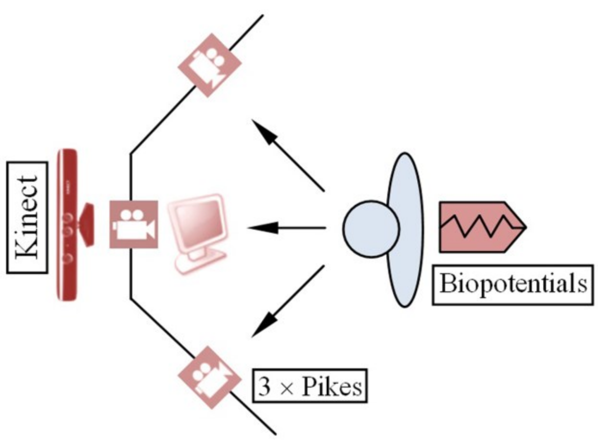

Fig. 1. Experimental Setting

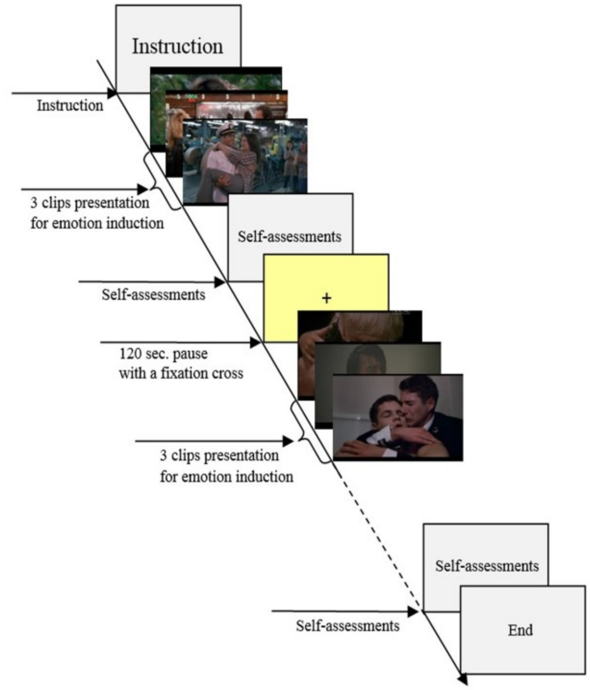

Fig. 2. Experimental Design

c) Trapezius Electromyogram (tEMG): Two Ag/AgCl electrodes were place on the descending portion of the upper trapezius muscles. An increase of muscle tone relates to an increasing activity of the sympathetic nervous system, while a decrease is in relation to parasympathetic arousal. tEMG has been proved to correlate with stress level and can reflect possible head movements during the experiment [22]. A third Ag/AgCl electrode placed on the subject's neck served as a reference for both the EMG and ECG.

During the experiment, none of the participants complained about the attached biopotential sensors. Therefore, we assumed that the sensors did not affect their emotional states.

2) *Video signals:*

The setup for video recording is also shown in Fig. 1. The participants were allowed to freely move their heads as the cameras could capture the face even if it was in out-of-plane rotations. All videos were encoded using the HuffYUV codec and transcoded to H.264 afterwards. The depth map streams are encoded in a lossless format.

IV. Analysis of Subjective Ratings

In this section, the effectiveness of film clips in eliciting target emotions based on subjective report will be evaluated firstly. Then, the projection of discrete emotions onto dimensional emotions will be considered.

Fig. 3 provides descriptive statistics of the recorded ratings of experienced emotions elicited by different film clips: All target emotions were successfully induced. The means of subjective ratings on target emotions have reached around seven points and other non-target emotions were averagely rated five points or less. But there is an exception in anger clips: sadness was also experienced by participants when they were watching film clips which intended to elicit anger. The mean of the subjective ratings on sad hereby is 6.45 points while anger scores a similar mean of 7.65 points.

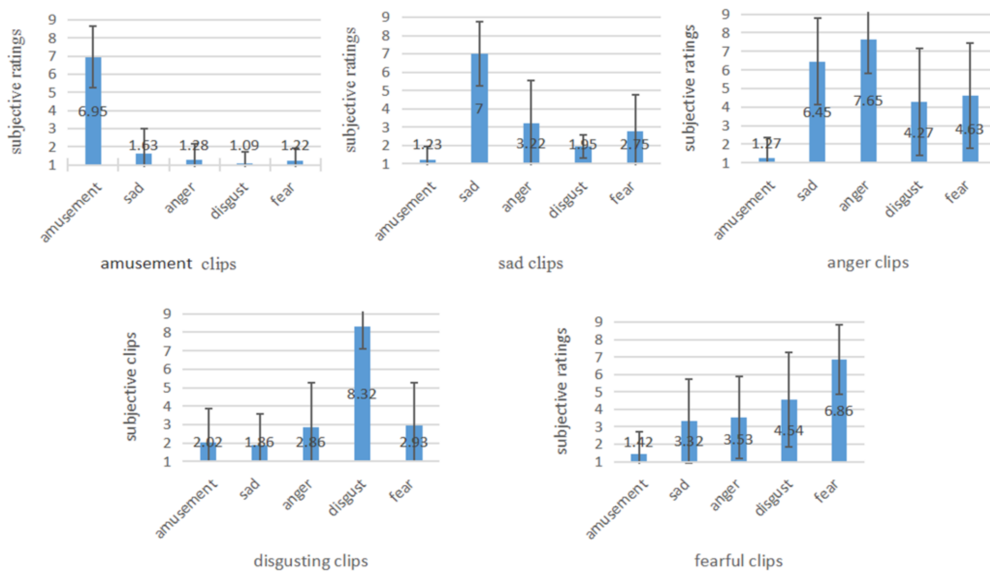

Fig. 3. Discrete Emotion Ratings for Film Clips

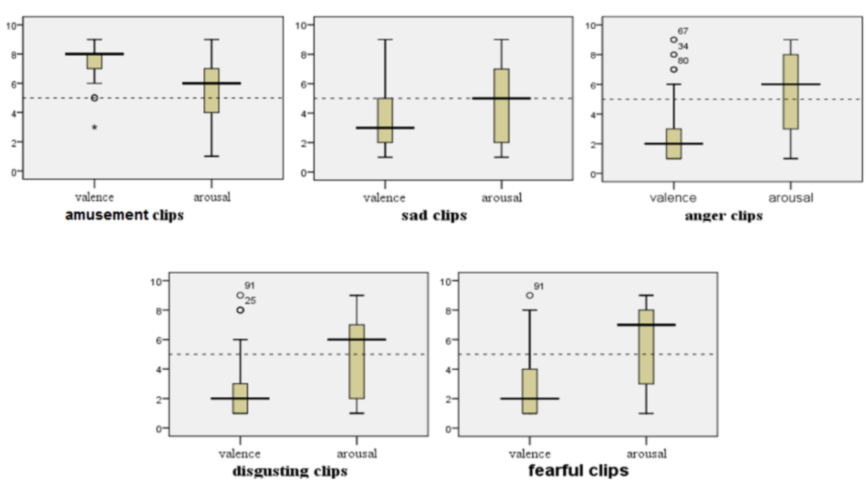

Fig. 4. Valence - Arousal Ratings for Film Clips

TABLE III. COMPARISON BETWEEN THE BIOVID EMO DB AND OTHER EXISTING DATABASES

Database	Number of Participants	Method of Elicitation	Induced Emotions	Video	Peripheral Biosignals	Method of Validation
DEAP [5]	32	music clips (one minute)	valence-arousal-dominance	yes	yes	Numerical rating (Valence, Arousal, Dominance)
MAHNOB-HCI [6]	27	film clips (34.9-117s)	valence-arousal-dominance	yes	yes	SAM
OPEN_EmoRec_ II [7]	30	standardized pictures presentation (20s) and human-computer interaction	valence-arousal-dominance	yes	yes	SAM
BioVid Emo DB	**86**	**film clips (32-245s)**	**discrete emotions (amusement sadness anger disgust fear)**	**yes**	**yes**	**Numerical rating (Valence, Arousal, Amusement, Sadness, Anger, Disgust, Fear)**

Valence and arousal levels were also evaluated for all film clips (see Fig. 4.). At the valence level, the mean of subjective ratings on amusement films have reached around eight points while the means of the other films are below four. Participants felt pleasant when watching happy clips and unpleasant when watching sad, anger, disgusting and fearful clips. At the arousal level, subjective ratings of fearful clips achieved the highest scores while sad clips got the lowest. Taking that into account, participants were activated strongly when watching fearful clips and felt medium aroused or calm when watching sad clips.

V. DISCUSSION

The BioVid Emo DB combines psychophysiological signals with video signals for discrete basic emotions that were effectively elicited by film clips. Compared to existing emotion databases, the presented database has some differences and advantages, for example, more participants were included and a discrete emotional model was considered (see Table III). This database can be of interest to researchers in various fields, from affective computing to multimedia.

Subjective ratings were analyzed preliminarily in this paper. Both sadness and anger were reported almost equally in anger clips. A reason for that might be the similarity of the subjective feelings of these two emotions [24]. This should be considered in future researches when it comes to emotion recognition e.g. in companion systems. Therefore, the finding of psychophysiological and/or video features to distinguish these two emotions could be of great relevance. Furthermore, the connection between discrete and dimensional emotions was verified in this study, which is consistent with other researches [25]. The large variance of the valence-arousal ratings of the different emotions could result from individual differences. Although we have already analyzed the effect of age on movie choice and subjective ratings in another paper [26], it is far from sufficient. Further analyses on individual differences such as gender and personality should be made and evaluated.

In the presented study the lengths of the film clips vary from 32 to 245 seconds. This might be a problem when it comes to individual-related classification tasks. Therefore we recommend that researchers who are interested in machine learning should select the length of the shortest clip, use it as a ‘cutting’ reference length for the other emotional clips and finally use sliding windows (e.g. 5 s with 4 s overlap) for feature extraction.

Comparing to naturalistic databases, higher emotional intensities were induced because of the controlled laboratory environment. But for real life applications, a natural context will be considered in future works.

VI. Conclusion

The next step will be the extraction of features or feature patterns which are most relevant for a robust emotion recognition or classification, respectively. Therefore different feature selection and machine learning techniques will be tested and evaluated.

In conclusion, a new multimodal emotion database has been presented in this paper. The good performance of emotion elicitation validated by subjective ratings and the quality of the recorded data makes the BioVid Emo DB a valuable contribution to the field of emotion analyses and affective data mining.

ACKNOWLEDGMENT

This research was supported by grants of the Transregional Collaborative Research Center SFB/TRR 62 Companion Technology for Cognitive Technical Systems funded by the

German Research Foundation (DFG) and doctoral scholarships funded by the China Scholarship Council (CSC) for Lin Zhang and Xueyao Ma.

AVAILABILITY

The BioVid Emo DB is available for all research groups but for non-commercial research only. Feel free to visit www.emotion-lab.org (download section) to get your own copy. All you or your advisor (a permanent position in an academic institute is required) have to do is to fill in and sign an agreement form. For comments or questions please contact {sascha.gruss, steffen.walter or lin.zhang}@uni-ulm.de.

REFERENCES

[1] D. Rösner, et al. Is There a Biological Basis for Success in Human Companion Interaction?. In: International Conference on Human-Computer Interaction. Springer International Publishing, 2016. S. 77-88.

[2] H. C. Traue, et al. A framework for emotions and dispositions in man-companion interaction. Coverbal Synchrony in Human-Machine Interaction. New Hampshire, USA: Science Publishers, 2013, S. 99-140.

[3] R. W. Picard, R. Picard. Affective computing. Cambridge: MIT press, 1997.

[4] L. Zhang, et al. Classification analysis for the emotion recognition from psychobiological data. International Symposium on Companion Technology, Biundo-Stephan, S & Wendemuth, A. (Eds) Proceedings of the International Symposium on Companion Technology, 2015, S. 149-154

[5] S. Koelstra, et al. Deap: A database for emotion analysis; using physiological signals. IEEE Transactions on Affective Computing, 2012, 3. Jg., Nr. 1, S. 18-31.

[6] M. Soleymani, J. Lichtenauer, T. Pun & M. Pantic. A multimodal database for affect recognition and implicit tagging. IEEE Transactions on Affective Computing, 2012, 3. Jg., Nr. 1, S. 42-55.

[7] S. Rukavina, S. Gruss, S. Walter, H. Hoffmann, & H. C. Traue. OPEN_EmoRec_II-A Multimodal Corpus of Human-Computer Interaction. International Journal of Computer, Electrical, Automation, Control and Information Engineering, 2015, 9. Jg., Nr. 5, S. 977-983.

[8] S. Rukavina, et al. Affective Computing and the Impact of Gender and Age. PloS one, 2016, 11. Jg., Nr. 3, S. e0150584.

[9] J. A..Hinojosa, , et al. Affective norms of 875 Spanish words for five discrete emotional categories and two emotional dimensions. Behavior research methods, 2016, 48. Jg., Nr. 1, S. 272-284.

[10] I. C. Christie, B. H. Friedman, Autonomic specificity of discrete emotion and dimensions of affective space: a multivariate approach. International journal of psychophysiology, 2004, 51. Jg., Nr. 2, S. 143-153.

[11] P. Lang, Bradley, M. Margaret. The International Affective Picture System (IAPS) in the study of emotion and attention. Handbook of emotion elicitation and assessment, 2007, 29. Jg.

[12] C. A. Frantzidis, et al. On the classification of emotional biosignals evoked while viewing affective pictures: an integrated data-mining-based approach for healthcare applications. IEEE Transactions on Information Technology in Biomedicine, 2010, 14. Jg., Nr. 2, S. 309-318.

[13] J. J. Gross, R. W. Levenson, Emotion elicitation using films. Cognition & emotion, 1995, 9. Jg., Nr. 1, S. 87-108.

[14] S. D. Kreibig, et al. Cardiovascular, electrodermal, and respiratory response patterns to fear and sadness inducing films. Psychophysiology, 2007, 44. Jg., Nr. 5, S. 787-806.

[15] J. Kim, E. André, Emotion recognition based on physiological changes in music listening. IEEE transactions on pattern analysis and machine intelligence, 2008, 30. Jg., Nr. 12, S. 2067-2083.

[16] C. A. Kothe, S. Makeig, J. A. Onton. Emotion recognition from EEG during self-paced emotional imagery. In: Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on. IEEE, 2013. S. 855-858.

[17] C. Mills, S. D'mello. On the validity of the autobiographical emotional memory task for emotion induction. PloS one, 2014, 9. Jg., Nr. 4, S. e95837.

[18] H. C. Lench, S. A. Flores, S. W. Bench. Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: a meta-analysis of experimental emotion elicitations. Psychological bulletin, 2011, 137. Jg., Nr. 5, S. 834.

[19] J. Hewig, D. Hagemann, J. Seifert, M. Gollwitzer, E. Naumann & D. Bartussek. Brief report. Cognition & Emotion, 2005, 19. Jg., Nr. 7, S. 1095-1109.

[20] Khalfa, S., Isabelle, P., Jean-Pierre, B., & Manon, R. Event-related skin conductance responses to musical emotions in humans. Neuroscience letters, 2002, 328. Jg., Nr. 2, S. 145-149.

[21] F. Agrafioti, D. Hatzinakos, A. K. Anderson. ECG pattern analysis for emotion detection. IEEE Transactions on Affective Computing, 2012, 3. Jg., Nr. 1, S. 102-115.

[22] J. Wijsman, B. Grundlehner, J. Penders & H. Hermens. Trapezius muscle EMG as predictor of mental stress. In: Wireless Health 2010. ACM, 2010. S. 155-163.

[23] S. Walter, et al. The BioVid heat pain database data for the advancement and systematic validation of an automated pain recognition system. In: Cybernetics (CYBCONF), 2013 IEEE International Conference on. IEEE, 2013. S. 128-131.

[24] D. Keltner, P. C. Ellsworth, K. Edwards. Beyond simple pessimism: effects of sadness and anger on social perception. Journal of personality and social psychology, 1993, 64. Jg., Nr. 5, S. 740.

[25] T. Eerola,; J. K. Vuoskoski. A comparison of the discrete and dimensional models of emotion in music. Psychology of Music, 2010.

[26] D. Hazer, et al. Emotion Elicitation Using Film Clips: Effect of Age Groups on Movie Choice and Emotion Rating. In: International Conference on Human-Computer Interaction. Springer International Publishing, 2015. S. 110-116.

[27] S. Andrea, S. D. Kreibig "Eliciting positive, negative and mixed emotional states: A film library for affective scientists." Cognition and Emotion 30.5 (2016): 827-856.