The LDOS-PerAff-1 Corpus of Face Video Clips with Affective and Personality Metadata

Marko Tkalčič, Jurij Tasič and Andrej Košir

University of Ljubljana Faculty of electrical engineering
Tržaška 25, Ljubljana, Slovenia
marko.tkalcic@fe.uni-lj.si, jurij.tasic@fe.uni-lj.si, andrej.kosir@fe.uni-lj.si

Abstract

This paper presents a corpus of video clips of users responding to emotional stimuli. The corpus is unique for two reasons. First, the emotions are annotated in the valence-arousal-dominance space instead of the usual coarse basic emotions. Second, the subjects are annotated with their personality parameters which offers a new ground for further investigations on personality and emotions. The corpus has been compiled for the needs of our research on recommender system. The paper provides information about the corpus acquisition procedure, corpus basic statistics with few examples and a short description of the research work where the corpus has been used in the past.

1. Introduction

Emotion recognition is an important area of affective computing. Several modalities, or combinations of them, can be used to detect user's emotions in human computer interaction and other areas. Some of these modalities use the video stream of the observed person as source (face expression, posture detection, hand gestures detection). Methods for the detection of emotions are being developed (Zeng et al., 2009). These methods use different corpora of video clips annotated with emotional metadata (Kanade et al., 2000; Pantic et al., 2005). Generally, such corpora are missing contextual information that might be useful in the development of emotion detection algorithms.

We present a corpus of video clips of users responding to emotion elicitation visual stimuli with additional annotations. These annotations include the induced emotive state, end users' personality parameters, demographic data and explicit ratings of the visual stimuli. The corpus presented here has been compiled for the needs of our research work of affective and personality based user modeling in recommender systems (Tkalčič et al., 2009a; Tkalčič et al., 2009b; Tkalčič et al., 2010).

The presented corpus has some unique properties that are not present in other corpora to the best of the authors' knowledge. First, the emotions are annotated in the valence-arousal-dominance (VAD) space which is less coarse then the usual basic emotions space. Second, the subjects are annotated with their personality in the big-five personality space. We believe that the presented corpus can be further exploited. As personality plays an important role in our emotional mechanisms (John and Srivastava, 1999) we believe it is an important contextual information in a corpus used for the development and validation of emotion detection methods. For example, extroverted people are believed to be more expressive than introverted people so the success rate of automatic methods for emotion detection could vary depending on user's personality differences. We encourage the usage of the proposed corpus for research work on emotion detection.

The remaining of the paper includes the acquisition procedure, the description of the corpus with some examples, the details about the distribution of the corpus, experiments where the corpus has been used, the discussion with future work guidelines and few concluding sentences.

2. Acquisition procedure

We used the emotion induction (sometimes referred also as emotion elicitation) experimental approach (Bradley and Lang, 2007). We used a subset of 70 images from the IAPS set of standardized visual stimuli to induce emotive responses in the subjects (Lang et al., 2005; Bradley and Lang, 2007). We assessed the personality of the subjects using the IPIP 50 questionnaire (Goldberg et al., 2006; http://ipip.ori.org/newQform50b5.htm, last accessed February 2010). We had 52 participants involved in the acquisition procedure.

2.1. Requirements for the data corpus

Our research work on recommender systems required a dataset of usage history of users interacting with an image consumption application. We needed the following data fields

- image: represented the item to be consumed as well as the emotion elicitation visual stimulus
- the induced emotive state as a triple in the VAD space
- personality parameters of the big five personality model as described by (Goldberg, 1998)
- explicit ratings on a five level Likert scale
- video clips of the users' responses to the visual stimuli

The general quality measure for a sample of data is that it should be a good representation of the larger set of data it attempts to represent. In our case there were two critical dimensions where the corpus should reflect the wider dataset:

- (i) a wide spectrum of emotive states
- (ii) a wide spectrum of participants' personalities.

We addressed both criteria by carefully selecting the visual stimuli and by analyzing the subjects' personalities.

2.2. Emotion induction procedure

We used a subset of the IAPS database of images for inducing emotive responses. The subset was chosen carefully to cover equally the value-arousal plane (see Fig. 1). The users' goal was to select images for their computer's wallpaper. The users were instructed to watch each image from the subset and give an explicit rating from 1 to 5. During their interaction with the application, that was build in Matlab (see Fig. 2 for a snapshot of the application's GUI), the users were recorded with a web cam. The web cam was positioned on top of the monitor so the participants' gaze was below the camera position. The basic unit of the dataset was thus a video clip of a user responding to a single visual stimulus from the IAPS database.

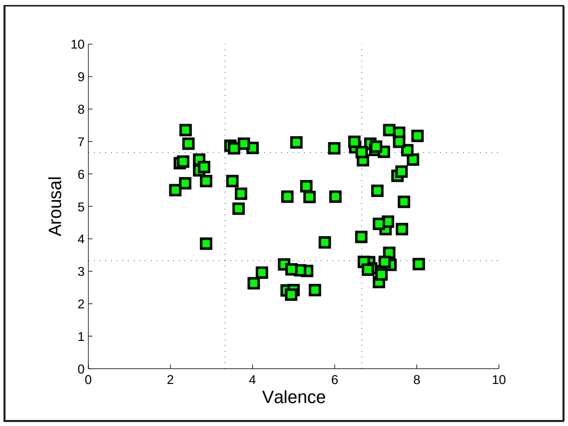

Figure 1: Distribution of the induced emotions of the visual stimuli in the valence-arousal plane.

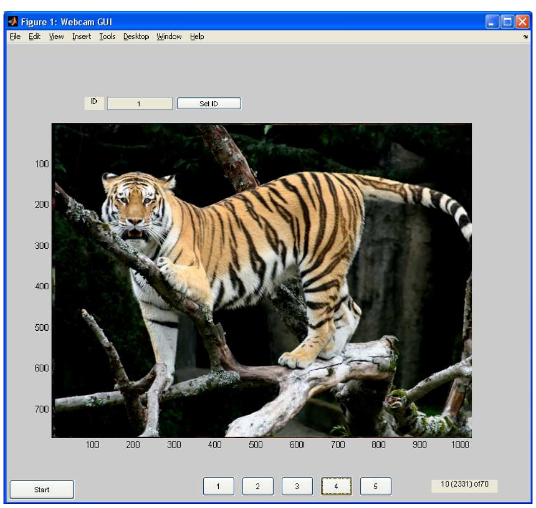

Figure 2: User interface of the application of the emotion induction experiment.

2.3. Personality and its assessment

Personality accounts for the individual differences in their emotional, interpersonal, experiential, attitudinal and motivational styles (John and Srivastava, 1999). The E factor tells the degree of engagement with the external world (in case of high values) or the lack of it (low values). The sub-factors of E are friendliness, gregariousness, assertiveness, activity level, excitement-seeking and cheerfulness. Extrovert people (high score on the E factor) tend to react with enthusiasm and often have positive emotions while introverted people (low score on the E factor) tend to be quiet, low-key and disengaged in social interactions. The N factor refers to the tendency of experiencing negative feelings. People with high N values are emotionally reactive. They tend to respond emotionally to relatively neutral stimuli. They are often in a bad mood which strongly affects their thinking and decision making. Low N scorers are calm, emotionally stable and free from persistent bad mood. The sub-factors are anxiety, anger, depression, self-consciousness, immoderation and vulnerability. The distinction between imaginative, creative people and down-to-earth, conventional people is described by the O factor. High O scorers are typically individualistic, non conforming and are very aware of their feelings. They can easily think in abstraction. People with low O values tend to have common interests. They prefer simple and straightforward thinking over complex, ambiguous and subtle. The sub-factors are imagination, artistic interest, emotionality, adventurousness, intellect and liberalism. The C factor concerns the way in which we control, regulate and direct our impulses. People with high C values tend to be prudent while those with low values tend to be impulsive. The sub-factors are self-efficacy, orderliness, dutifulness, achievement-striving, self-discipline and cautiousness. The sub-domains of the A factor are trust, morality, altruism, cooperation, modesty and sympathy. The A factor reflects individual differences in concern with cooperation and social harmony.

We used the IPIP 50 questionnaire to assess the personality big five factors of the participants. The questionnaire consisted of 50 items, 10 per each big five personality factor.

2.4. Participants

We had 52 students of a secondary school who participated in the experiment. The average age was 18.3 years (SD = 0.56). There were 15 males and 37 females.

3. Corpus

The corpus consists of 3640 video clips of 52 participants responding to 70 different visual stimuli. The video files are segmented by user and by visual stimulus. The annotations are stored in text based files. The participants cover a heterogeneous area in the space of the big five factors.

3.1. File formats

The video files are encoded with the xvid codec and have the resolution of 320 x 240 pixels at the frame rate of 15fps. The filename notation is `USERID_ITEMID.AVI` . For example, the filename `41_1534.avi` represents the video clip of the user with the id 41 responding to the visual stimulus image with the id 1534 from the IAPS database. There is a total of 258 Mbytes of video clips with the duration of 5 hours and 55 minutes.

The annotations are stored in three different formats: excel, semicolon delimited text and ARFF weka format. The filenames are `LDOS-PerAff-1.xls`, `LDOS-PerAff-1.csv` and `LDOS-PerAff-1.arff`.

3.2. Corpus properties

Each video clip is annotated with a line in the annotation file. Tab. 1 shows an extract from the annotations files. The annotations files have the following columns: user_id, image_id, image_tag, genre, watching_time, wt_mean, valence_mean, valence_stdev, arousal_mean, arousal_stdev, dominance_mean, dominance_stdev, big5_1, big5_2, big5_3, big5_4, big5_5 gender, age, explicit_rating, binary_rating. Descriptively, they contain the ID of the observed user, the ID of the observed image used to induce emotion, the recorded watching time, the average watching time of the observed item, the image tag, the image genre, the first two statistical moments of the valence, arousal and dominance values of the induced emotive state of the observed image, the big five parameters of the observed user, the gender, age and explicit rating given by the user to the observed image.
The participants showed heterogeneity in the distribution of the big five personality parameters as can be seen in Fig. 3. Unfortunately it is not possible to assess whether the personality distribution of the participants reflects a wider group of users because such norms do not exist (Goldberg et al., 2006).

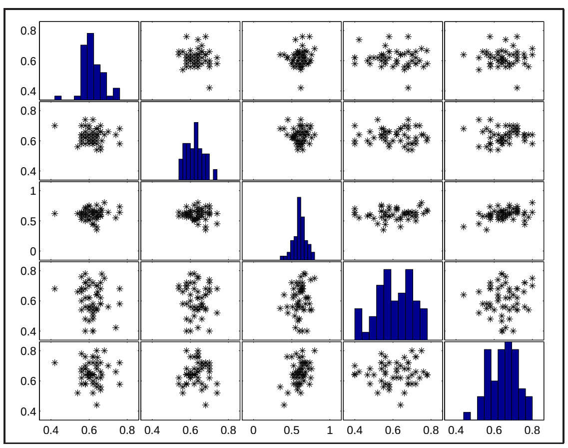

Figure 3: Distribution of the participants' personalities as pairs of big five factors scatter plots and histograms of single factors.

We chose the subset of visual stimuli from the IAPS database randomly with the constraint to cover equally a wide range of induced emotional values in the valence-arousal plane. The distribution of the induced emotions of the visual stimuli used is depicted in Fig. 1. This distribution is more suitable for the development of emotion detection methods that cover the whole valence-arousal plane (all the basic emotions).

3.3. Examples

Tab. 1 shows an extract from the annotation files. Fig. 4 and Fig. 5 shows snapshots from the video files of two subjects. The subject in Fig. 4 offers very little dynamics in face expressivity while the subject in Fig. 5 shows extreme face dynamics.

4. Experimental design and results

In this section we present three experiments that we are conducting on the described corpus. Two of these experiments have been completed (the affective content based recommender and the personality based collaborative recommender) while we are still working on the emotion detection experiment.

4.1. Affective content based recommender system

Part of the presented corpus was used to develop and evaluate an affective user and item modeling approach in a content based recommender (CBR) system. The CBR scenario included users interacting with an application where they could observe images (items) and giving explicit ratings on a five scale Likert scale to each item. The user' goal was to rate images as candidates for the personal computer's wallpaper. We modeled the items with the first two statistical moments of the induced emotive response over a large set of users which was provided by the IAPS database. This represented the affective item model. We also modeled the items with the genre and watching time tags which represented the generic item modeling approach. In the evaluation we compared the performances of both models, the affective and generic, in a CBR scenario. We used several machine learning algorithms to predict the users' ratings. We evaluated the performance of the CBR using confusion matrix based measures precision, recall and F measure. We performed a statistical analysis which showed that the usage of the affective modeling approach yielded significantly better CBR performance than the usage of generic modeling. Parts of the results were published in (Tkalčič et al., 2009a) and (Tkalčič et al., 2010).

4.2. Personality based collaborative recommender system

Another part of the presented corpus was used in an experiment where we evaluated a novel personality based user similarity measure for collabrative filtering. A well known drawback of collaborative filtering methods is the new user problem. It occurs when a new user joins the system and the system has little or no knowledge on the user's preferences. As the user similarity measures rely on explicit ratings given by the users, when there are little ratings available, the algorithm for finding similar users tend to give bad choices. Consecutively, the predicted ratings have low correlation with real ratings. In order to alleviate the new user problem we introduced an initial questionnaire to assess the personality of each user. We chose the IPIP questionnaire with 50 questions (ipi, 2009) which yielded five parameters for each user in the big-five personality model space (Johnson, 2009). We constructed a user similarity measure as an Eclidian distance in the big five personality space. We compared the proposed user similarity measure with a generic rating based user similarity measure in a collaborative recommender system scenario. Again, we calculated the confusion matrix based scalar measures precision,

user_id	image_id	image_tag	genre	watching_time	wt_mean	valence_mean	valence_stdev	arousal_mean	arousal_stdev	dominance_mean	dominance_stdev	big5_1	big5_2	big5_3	big5_4	big5_5	gender	age	explicit_rating
10	6910	Bomber	action	2614	5307	5.31	2.28	5.62	2.46	5.1	2.46	3.2	2.7	2.9	3.5	2.9	0	18	4
10	9331	Assault	action	2240	3214	2.03	1.35	6.04	2.35	0	0	3.2	2.7	2.9	3.5	2.9	0	18	3
10	7052	HairDryer	still	2223	2665	4.93	0.81	2.75	1.8	5.82	1.93	3.2	2.7	2.9	3.5	2.9	0	18	1
10	1280	Rat	animal	1906	3093	3.66	1.75	4.93	2.01	5.05	2.2	3.2	2.7	2.9	3.5	2.9	0	18	1
10	2394	Medicalworker	people	1943	2993	5.76	1.74	3.89	2.26	0	0	3.2	2.7	2.9	3.5	2.9	0	18	3

Table 1: Extract from the annotation table of the corpus.

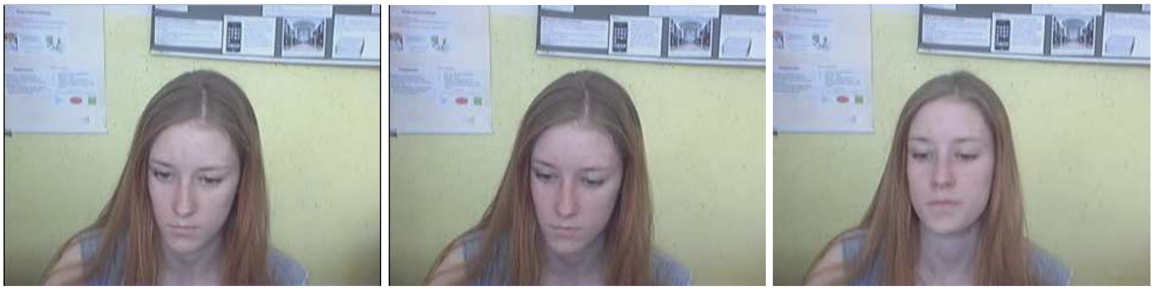

Figure 4: Snapshots of a subject with very low dynamics of face expressivity.

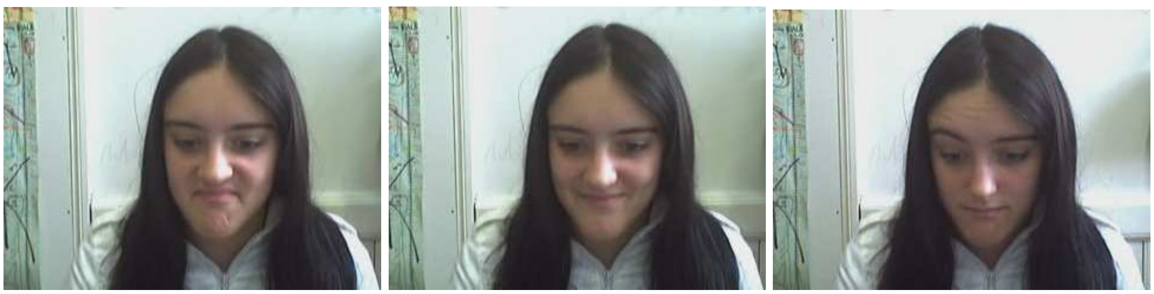

Figure 5: Snapshots of a subject with high dynamics of face expressivity.

recall and F measure. Statistical analysis showed that the proposed personality based user similarity measure yielded significantly better results than the rating based user similarity measure which makes it more suitable not only to alleviate the new user problem but also to use when the new user phase dies away. The results of this research were published in (Tkalčič et al., 2009b).

4.3. Emotion detection from video clips

Part of our ongoing research work is the detection of emotion from face videoclips. Our goal is to develop a method for detecting emotions in users with two novel properties: (i) the inclusion of personality parameters as features and (ii) detection in the valence-arousal-dominance (VAD) space (instead of the coarse space of basic emotions). We intend to use the emotion detection method for automatic tagging of users and items for affective profile building.

The current design of the experiment includes the extraction of the users' faces using the Viola-Jones algorithm (Viola and Jones, 2004) and fine registration using the active appearance model (AAM) tracker developed by (Saragih and Gocke, 2009). We plan to extract low level features using Gabor wavelets and applying the Hidden Markov Model (HMM) to reduce the number of variable features (due to the variable length of video clips) to a fixed number. We will combine personality parameters and low level features in a machine learning (ML) algorithm. We will evaluate several ML algorithms.

5. Copyright and privacy issues

5.1. Distribution of the corpus

Our exclusive interest is the promotion of research. Thus the distribution of the corpus is free for use in academic, not-for-profit research at recognized educational institutions. The researcher who wishes to receive the corpus must fill in and submit the EULA (end user license agreement) form available at `http://slavnik.fe.uni-lj.si/PerAff/` according to the instructions on the form. Within 30 days after receiving your form we will send you a username and password for downloading the corpus. The researchers are expected not to publish or distribute the material in any form.

6. Discussion and future work

The corpus presented in this paper was created to support our work on affective and personality based recommender systems (Tkalčič et al., 2009a; Tkalčič et al., 2009b; Tkalčič et al., 2010). We evaluated the impact that emotive parameters and personality has on user ratings. We proposed a novel approach for modeling users with VAD emotive parameters in the context of an image recommender system. Furthermore we developed a novel user similarity measure based on the big-five personality traits. The personality based user similarity measure yielded statistically equivalent performance of a collaborative recommender system than the usual rating based user similarity measure while withstanding the new user problem.

Beside affective and personality user modeling for recommender systems the corpus can be used for a variety of research work we are unable to foresee right now. Anyway we provide a list of interesting topics where the presented corpus could be used

1. a comparison of efficiency of different expression detection algorithms
2. development of a non intrusive personality detection method based on face video clips of induced emotions
3. relation between expression detection and personality

The latter is, in the authors' opinion, one of the most interesting issues. How can user's personality help emotion detection techniques? For example, knowing that a subject has a certain personality profile could help the emotion detection algorithm to fine tune its internal parameters and thus achieve greater accuracy. The subjects in Fig. 5 and Fig. 4 have different personalities. They differ also in the dynamics of face expressions. It would be interesting to see whether there is a combination of personality parameters that is correlated with face expression dynamics. This would surely be helpful for the emotion detection algorithm to adapt its internal thresholds.
The authors have already undertaken on the emotion detection with the inclusion of personality, as described in Sec. 4.3.. At the time of writing of this paper we have not get so far to be able to include any results.

7. Conclusion

The affective computing community can make good use of the proposed corpus with its personality annotations which makes it unique. It features almost six hours of 52 subjects face video recordings with 70 different induced emotive states. The video sequences are annotated with big-five personality parameters of subjects and metadata related to the items.
In this paper we provide the background for the construction of the corpus. We describe the acquisition procedure and give basic dataset statistics. We also give the description of the recommender systems application where the presented corpus was validated. Instruction for accessing the corpus are given. We suggest a list of topics where the specifics of the LDOS-PerAff-1 corpus could be used.

Acknowledgement

The authors would like to thank the teachers and students from the Gimnazija Poljane school in Ljubljana for their participation. We are also thankful to Matevž Kunaver, Tomaž Požrl and other members of the LDOS group who have helped in the implementation of the acquisition procedure. This work has been partially funded by the European Commission within the 6th framework of the IST under grant number FP6-27312. All statements in this work reflect the personal ideas and opinions of the authors and not necessarily the opinions of the European Commission.

8. References

Margaret M. Bradley and Lang, 2007. *The International Affective Picture System (IAPS) in the Study of Emotion and Attention*, chapter 2. Series in Affective Science. Oxford University Press, 198 Madison Avenue.

Lewis R. Goldberg, John A. Johnson, Herbert W. Eber, Robert Hogan, Michael C. Ashton, C. Robert Cloninger, and Harrison G. Gough. 2006. The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40:84–96.

Lewis R. Goldberg. 1998. What is beyond the big five? *Journal of Personality*, 66(4):495–524.

http://ipip.ori.org/newQform50b5.htm. last accessed: February 2010. Sample ipip big5 questionnaire. web site.

2009. International personality item pool: A scientific collaboratory for the development of advanced measures of personality traits and other individual differences. `http://ipip.ori.org/`, June.

Oliver P. John and Sanjay Srivastava. 1999. The big five trait taxonomy: History, measurement, and theoretical perspectives. In Lawrence A. Pervin and Oliver P. John, editors, *Handbook of Personality: Theory and Research*, pages 102–138. Guilford Press, New York, second edition.

John A. Johnson. 2009. Descriptions used in ipip-neo narrative report. `http://www.personal.psu.edu/faculty/j/5/j5j/IPIPNEOdescriptions.html`, June.

T. Kanade, J.F. Cohn, and Y. Tian. 2000. Comprehensive database for facial expression analysis. In *Proceedings of the fourth IEEE International conference on Automatic Face and Gesture Recognition*, page 46.

P.J. Lang, M.M. Bradley., and B.N. Cuthbert. 2005. International affective picture system (iaps): Affective ratings of pictures and instruction manual. technical report a-6. Technical report, University of Florida, Gainesville, FL.

M. Pantic, M.F. Valstar, R. Rademaker, and L. Maat. 2005. Web-based database for facial expression analysis. In *Proc. IEEE Int'l Conf. Multmedia and Expo (ICME'05), Amsterdam, The Nederlands*, July.

J. Saragih and R. Gocke. 2009. Learning AAM fitting through simulation. *Pattern Recognition*, 42(11):2628–2636.

M. Tkalčič, J. Tasič, and A. Košir. 2009a. Emotive and Personality Parameters in Multimedia Recommender Systems. *ACII 2009 Affective Computing and Intelligent Interaction*, page 33.

Marko Tkalčič, Matevž Kunaver, Jurij Tasič, and Andrej Košir. 2009b. Personality based user similarity measure for a collaborative recommender system. In C. Peter, E. Crane, L. Axelrod, H. Agius, S. Afzal, and M. Balaam, editors, *Proceedings of the 5th Workshop on Emotion in Human-Computer Interaction - Real world challenges*, pages 30–37. Fraunhofer Verlag, September.

Marko Tkalčič, Jurij Tasič, and Andrej Košir. 2010. The need for affective metadata in content-based recommender systems for images. In Mark Maybury, editor, *Multimedia Information Extraction*, chapter 19. MIT Press.

Paul Viola and Michael Jones. 2004. Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154.

Zhihong Zeng, Maja Pantic, Glenn I. Roisman, and Thomas S. Huang. 2009. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(1):39–58, Jan.