Combining Image Databases for Affective Image Classification

Hye-Rin Kim
Dept. of Computer Science Yonsei University
Seoul, Republic of Korea
Email: hrkim@cs.yonsei.ac.kr

In-Kwon Lee
Dept. of Computer Science Yonsei University
Seoul, Republic of Korea
Email: iklee@yonsei.ac.kr

***Abstract*—Affective image classification has attracted much attention in recent years. However, the production of more exact classifiers depends on the quality of the sample database. In this study, we analyzed various existing databases used for affective image classification and we tried to improve the quality of the learning data by combining existing databases in several different ways. We found that existing image databases cannot cover the overall range of the arousal-valence plane. Thus, to obtain a wider distribution of emotion labels from images, we conducted a crowd-sourcing-based user study with Amazon Mechanical Turk. We aimed to construct several different versions of affective image classifiers by using different combinations of existing databases, instead of using one. We used low-level features in our classification experiments to explore the discriminatory properties of emotion categories. We report the results of intermediate comparisons using different combinations of databases to evaluate the performance of this approach.**

Keywords–image emotion; emotion-based classification

I. EMOTION-BASED CLASSIFICATION

A. *Image collections*

Recently, many researchers have reported studies of emotion extraction from images. Several key issues influence the affective classification of images. In particular, it is necessary to obtain ground-truth emotion labels for images. However, obtaining high quality emotion-based images is not easy because of human subjectivity and there are no standard models of emotions. In general, researchers have conducted large-scale user studies to obtain emotion information with two types of emotion models: categorical and continuous models. Categorical models give a discrete value to an emotion using a word, such as happy, sad, or gloomy. By contrast, continuous models represent specific emotions as coordinates in a multidimensional space (a two-dimensional plane is usually preferred, which is called the arousal-valence plane) and we used this type of model in our experiments.

International Affective Picture System (IAPS) is a database of pictures that are used to elicit a range of emotions, which Lang et al. [1] employed in experimental studies of affective image classification. Mikels et al. [2] introduced a subset of the IAPS database for the categorization of images, which we used in our research to obtain the arousal and valence values of the pictures.

Geneva Affective PicturE Database (GAPED) contains 730 images with emotional values [3]. GAPED has four specific types of negative contents, including spiders, snakes, and negative scenes. The positive pictures mainly comprise images of human and animal babies, and nature scenes. The pictures are rated according to their arousal, valence, and congruence values.

The Nencki Affective Picture System (NAPS) [4], is another affective image database, which comprises 1,356 realistic, high-quality photographs with five subject categories (people, faces, animals, objects, and landscapes). The images were given affective arousal and valence ratings by 204 participants, who were mostly European.

Obtaining emotion information using crowd-sourcing Machajdik et al. [5] obtained emotion information based on categorical labels. Furthermore, the range of arousal-valence values is highly limited in other databases, as shown in Figure 1(a). Therefore, we collected arousal and valence values for the images in Machajdik et al.'s database based on a large-scale user survey. A total of 199 subjects were recruited to participate in the survey using Amazon Mechanical Turk and the subjects provided 6787 responses. We collected at least six responses for each image and each subject provided an average of 33 responses. Figure 1(b) shows the distribution of the emotion labels obtained in the survey, which demonstrates that the combined database was more evenly distributed in the arousal-valence plane compared with the original database.

B. *Image Features*

In this study, we applied most of the features used in previous studies, which are mainly related to color and texture. In addition, we used a new feature called color harmony (f31, f32 in TableI), which is based on color perception theory. Recently, several statistical studies have proposed methods for computing the harmony between colors. We employed one of these methods [9] to compute the harmony between

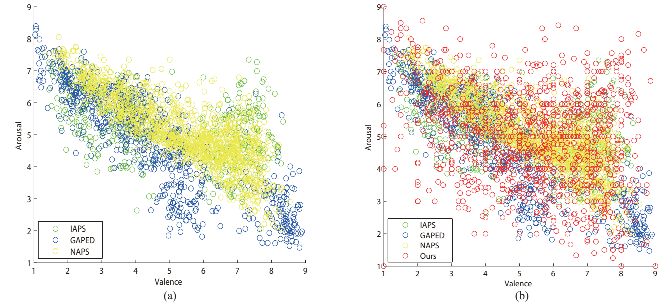

Figure 1. (a) Arousal-Valence distribution of images using three existing databases. (b) our user-study results are added (red dots)

 ISBN: 978-1-61208-382-7

TABLE I. OVERVIEW OF FEATURES IN OUR METHOD.

Feature	Description	character	Feature	Description	character
f1, f2, f3	The histogram of hue, saturation and value of image	color	f21, f22, f23	Average saturation for the first, second and third largest segment	color
f4, f5, f6	Average of hue, saturation and value of image	color	f24, f25, f26	Average value for the first, second and third largest segment	color
f7	The hue section that used in image over threshold	color	f27	Color descriptor in [6]	color
f8	The number of hue sections that used in image over threshold	color	f28	Color consistancy in [7]	color
f9, f10, f11	Activity, Weight and Heat of image [8]	color	f29	The existance of basic color	color
f12, f13	Mean and standard deviation of the magnitude of Gabor filtered image	texture	f30	The number of used colors for each basic colors	color
f14, f15, f16, f17	Energy, Entropy, Contrast, Homogeneity of gray scale image	texture	f31	Average color harmony of the most used ten colors	color
f18, f19, f20	Average hue for the first, second and third largest segment	color	f32	Color harmony between two colors among the ten representative colors	color

TABLE II. CLASSIFICATION PERFORMANCE USING VARIOUS COMBINATIONS OF DATABASES.

Database	5 fold cross validation	No. of images
GAPED	0.80	730
GAPED+NAPS	0.68	2086
GAPED+IAPS	0.64	1119
NAPS	0.60	1356
NAPS+IAPS	0.59	1745
IAPS + Machajdik + GAPED+ NAPS	0.54	3561
NAPS + Machajdik	0.54	2044
GAPED + Machajdik	0.54	1816

representative colors in image. For each image, we extracted 10 representative colors using k-means clustering and we then computed the harmony among all of the colors. The features used in this study are listed in Table I.

II. Classification

Given a set of features, we aimed to construct an appropriate classifier to estimate the emotion in a given image. We used the public library A Library for Support Vector Machines [10] to compute the nonlinear hyperplanes for class separation. To evaluate the classification performance, we divided the emotion space into four classes where the point (5, 5) was at the center of the arousal and valence axes. Based on the ratings in the database, all of the images were labeled according to one of the four classes for training. We performed a 5 fold cross-validation because we lacked a ground-truth database. The classifier was trained using various combinations of databases. Table II shows the classification performance based on 4 four categories in for each combination. The results show that the GAPED database recorded the best performance in with our scheme so far.

III. Conclusion

In this study, we compared the affective classification performance of different combinations of existing image databases, where we included the results of a user study to compensate for the lack of data. The main contributions of our study can be summarized as follows: 1) We performed a crowd-sourcing-based user survey to collect emotion information for a large set of images; 2) We evaluated emotion-based image databases using various combinations of categories. There is no research for affecitve classification using the combination of various databases. Therefore, we tried to find a research using GAPED database which recorded the best performance in our scheme, but couldn't find it. Statistically, the accuracy for catogorical affective classification is less than 80%. We leave the exact comparison with other methods for future work. We will also construct a more appropriate regression-based model to estimate the arousal and valence coordinates for images. In addition to low-level features, we may consider the use of high-level semantics to obtain better performance, which are employed widely in aesthetics as new features.

Acknowledgment

This research is supported by Ministry of Culture, Sports and Tourism (MCST) and Korea Creative Content Agency (KOCCA) in the Culture Technology (CT) Research & Development Program 2014.

References

[1] P. Lang, M. Bradley, and B. Cuthbert, "International Affective Picture System (IAPS): Affective Ratings of Pictures and Instruction Manual. Technical Report A-6," The Center for Research in Psychophysiology, University of Florida, 2005.

[2] J. A. Mikels, B. L. Fredrickson, G. R. Larkin, C. M. Lindberg, S. J. Maglio, and P. A. Reuter-Lorenz, "Emotional category data on images from the International Affective Picture System," Behav Res Methods., vol. 37, 2005, pp. 626–630.

[3] E. S. Dan-Glauser and K. R. Scherer, "The Geneva affective picture database (GAPED): a new 730-picture database focusing on valence and normative significance," Behav Res Methods., vol. 43, 2011, pp. 268–277.

[4] A. Marchewka, L. Zurawski, K. Jednorog, and A. Grabowska, "The Nencki Affective Picture System (NAPS). Introduction to a novel standardized wide range high quality realistic pictures database," Behav Res Methods., vol. 46, 2014, pp. 596–610.

[5] J. Machajdik and A. Hanbury, "Affective Image Classification Using Features Inspired by Psychology and Art Theory," in Proceedings of the International Conference on Multimedia, ser. MM '10. ACM, 2010, pp. 83–92.

[6] J. van de Weijer and C. Schmid, "Coloring Local Feature Extraction," vol. 3952, 2006, pp. 334–348.

[7] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-Based Color Constancy," Image Processing, IEEE Transactions on, vol. 16, no. 9, 2007, pp. 2207–2214.

[8] M. Solli and R. Lenz, "Color Based Bags-of-Emotions," in Computer Analysis of Images and Patterns, vol. 5702. Springer Berlin Heidelberg, 2009, pp. 573–580.

[9] L.-C. Ou and M. R. Luo, "A Colour Harmony Model for Two-Colour Combinations," Color Research and Applications, vol. 31, no. 3, 2006, pp. 191–204.

[10] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," vol. 2, 2011, pp. 27:1–27:27, software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm [retrieved: December, 2014].