Volume 3, Issue 10, October 2013 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Survey in Classification of Natural Scene Images Based on Emotion

Sharmishtha D. Ronge
Department of Computer Network
Sinhgad College of Enginerring, Pune, India

Abstract— Emotion modeling evoked by natural scenes is demanding issue. In this paper, we propose a novel scheme for investigating the emotion reflected by a natural scene, considering the human emotional state. Emotion in natural scene images performances a significant role in the way humans see an images. founded on the emotion (happiness, unhappiness, fear, anger etc.) of any human being the images that are examined by that individual can have a important influence in a sense that if the person is for demonstration in happy mood and views an image that is satisfying then would have a better sense of addition in the direction of that image and would not accept an image that depicts unhappiness as an emotion. Whereas distinct persons may interpret the same images in different ways, we still can build a universal classification for different emotions. An important interaction was found between facial signs and the emotional content of the scenes, displaying a response benefit for facial signs accompanied by congruent scenes. This benefit was robust against increasing task load. Taken together, the result show that the surrounding scene is a significant factor in identifying facial expressions, Although, However, we rarely encounter an isolated facial expression; generally, we perceive a person's facial reaction as part of the surrounding context Recognition of facial expressions has traditionally been investigated by giving facial expressions without any information. In three trials, participants were required to categorize facial expressions (disgust, fear, happiness) that were shown against backgrounds of natural scenes with either a congruent or an incongruent emotional significance.

Keywords— Natural Scene, image classification, facial expressions, CBIR, image Description

I. INTRODUCTION

Emotion in natural scene images plays an important role in the way humans perceive an image. Based on the emotion (happiness, sadness, fear, anger etc.) of any human being the images that are viewed by that person can have a significant impact in a sense that if the person is for example in happy mood and views an image that is pleasing then would have a better sense of attachment towards that image and would not accept an image that depicts sadness as an emotion. Although different people may interpret the same image in different ways, we still can build a universal classification for different emotions. A substantial allowance of study has established that the visual scheme is highly efficient at seeing both faces and facial expression Specifically, the trials here use a set of six basic emotion happiness, sadness, anger, surprise, fear and disgust. These are the six emotions. Two datasets are utilized in the experiments, extracted from reside periodical. After, the instance space is reduced by applying feature assortment techniques.

The majority of functional magnetic resonance imaging Studies of emotion have used visual stimuli, most often employing pictures of expressive human faces or natural scenes to elicit emotional re-activity. The use of emotional faces and scene images has yielded a considerable amount of data, and here we seek to identify the similarities and distinctions in blood oxygen level dependent signal that is associated with these classes of eliciting stimuli, using a coordinate-based meta-analysis technique, through which a comparison can be made between studies using faces and those using scenes. The studies of emotional processing typically contain a neutral control condition (inexpressive face, or non-emotional scene) that theoretically serves to remove simple effects of visual perception, thus revealing the emotional process. Here we seek to identify potential differences that may remain in comparisons of emotional face and emotional scene perception, to aid in data interpretation across published reports. Perhaps more importantly, we seek to identify the brain regions consistently activated during visual perception of emotional stimuli, whether via the expressive face or the evocative scene.

Many experiments made schematic faces as materials to study emotion. Although the schematic faces were impoverished as compared with photographs of real human faces, they appeared to be potent effective facial stimuli. Schematic faces appeared to communicate emotional meaning effectively and showed disruptions in perception when inverted similar to those found with photographed faces. Furthermore, it has recently been shown that schematic faces elicit event-related potentials that are similar to those elicited by photographs of faces. Therefore, given that schematic faces contain fewer feature confounds than photographs and appear to be effective affective face stimuli, it may actually be preferable to use schematic faces in studies in which the perception of facial expression is studied. Detecting negative faces quickly and responding to them as efficiently as possible has an evolutionary benefit that is argued to have formed the basis for selection [10]. In addition, visual system can anticipate the appearance of new visual information. It would be adaptive to be able to selectively prioritize relevant objects, even if those faces have yet to emerge. That has clouded interpretation of the results of studies comparing the effectiveness of different emotional expressions in guiding focal attention is that it has proven difficult to determine whether the observed differences in the speed with which faces expressing different emotions are detected reflects a difference in the emotions expressed by the faces or a difference in the component parts or features that distinguish the faces. By definition, faces expressing different emotions, such as anger and happiness, consist of different composites of features. Given these differences, any evidence showing differential guidance of focal attention by unattended faces expressing different emotions can often be accounted for in terms of the different features, rather than in terms of the different emotions expressed by the faces. Clouded interpretation of previous studies revolves around the question of what constitutes satisfactory evidence that unattended information guides focal attention. In a number of studies, the underlying assumption has been that the only satisfactory evidence that a face guides focal attention is a pattern of findings showing that the speed with which a face is detected is relatively unaffected by the number of distracters faces.

Content-Based Image Retrieval (CBIR) has been discussed in the technical literature as a method that may develop into an efficient image search and retrieval technique. Prior work on medical image retrieval has mainly focused on extracting low-level visual features (e.g., color, texture, shape, spatial layout) and then using them directly to compute image similarity. Extensive experiments have shown, however, that low-level image features cannot always capture the biomedical semantic concepts in the image.1 This poses a serious shortcoming in applying CBIR to routine clinical use, where image content is defined in terms of biomedical concepts. In general, it is challenging to link high-level semantic concepts and automatically-extracted, low-level image features. Therefore, to support query by semantic concept, there is a compelling need for CBIR systems to provide maximum support towards bridging the semantic gap between the low-level visual features and the semantics in biomedical concepts.

II. Literature Survey

A. Emotions in Images:

In this study, we characterize and discuss about key facets of the problem of computational inference of aesthetics and emotion from images. We start with a background consideration on beliefs, photography, paintings, visual arts, and psychology. This is followed by introduction of a set of key computational difficulties that the research community has been striving to explain and the computational framework required for solving them. We furthermore recount information groups available for performing evaluation and summarize some real-world submissions where research in this domain can be engaged. An important number of papers that have tried to solve difficulties in aesthetics and emotion inference are reviewed in this tutorial. We also discuss future main headings that researchers can follow and make a strong case for gravely trying to solve problems in this research domain.

B. Fuzzy GIST Emotion Detection from Natural Scene Image:

Emotion modeling evoked by natural scenes is demanding issue. In this paper, we propose a innovative scheme for analyzing the emotion reflected by a natural scene, considering the human emotional rank. Based on the notion of initial GIST, we developed the fuzzy-GIST to construct the emotional characteristic space. According to the connection between emotional factors and the individual features of image, L*C*H* hue and orientation data are chosen to study the relationship between human's reduced level emotions and images characteristics. And it is realized that we need to investigate the visual characteristics at semantic grade, so we incorporate the fuzzy concept to extract characteristics with semantic meanings. Furthermore, we treat emotional electroencephalography (EEG) utilizing the fuzzy reasoning based on possibility theory rather than broadly utilized accepted likelihood idea to generate the semantic characteristic of the human emotions. Fuzzy-GIST comprises of both semantic visual data and linguistic EEG feature, it is used to represent emotional gist of a natural view in a semantic grade. The emotion evoked by images is predicted from fuzzy- GIST by utilizing a support vector appliance, and the mean attitude tally (MOS) is utilized for performance evaluation for the suggested design. The trials results display that positive and contradictory emotion can be identified with high correctness for a given dataset.

C. K-Nearest Neighbor classifier

The k-Nearest-Neighbor (kNN) classification is one of the most fundamental and simple nonparametric classification methods. For k nearest neighbors, the predicted class of test sample x is set equal to the most frequent true class among k nearest training samples. In our work we used the mat lab implementation of kNN classifier. We tested several values of k. Best results were obtained by k = 10.

D. Support Vector Machine classifier

Support Vector appliances (SVM) are based on the concept of decision hyper plane. The SVM finds a linear dividing hyper plane with a maximal margin in the higher dimensional space. For our experiments, the LIBSVM bundle with the radial cornerstone function (RBF) kernel was engaged. LIBSVM implements the “one-against-one” approach for multi-class classification. For n = 8 categories there are n(n 1) 2 =28 single classifiers and each one teaches data from two different classes. Each binary classification is considered to be a voting.

III Methods

A. Stimuli and Apparatus

Stimuli were randomly plotted in the cell of a 10×10 virtual matrix. The overall matrix dimensions were 20cm wide×20cm high. The visual search items consisted of schematic faces of positive emotion (happy faces, see Fig. 1. the left one), neutral emotion (neutral faces, see Fig. 1. the middle one) and negative emotion (sad faces, see Fig. 1. the right one) presented on a black background. The individual stimuli faces were 18mm wide×18mm high. In the all-element and gap conditions, the display size was 6, 8 or 10 items. An equal number of old items and new items were always present in the display. The target was present on 50% of the trials.

Fig.1 The Example of the stimulus displays used in Experiment

B. Design and Procedure

The participants were instructed to find the unique image in each display as quickly as possible while maintaining high accuracy. Importantly, the participants were not told which target they would be searching for on any given trial. At the beginning of each trial, every condition had six search types. The six search types as search a happy face from neutral faces or sad faces, search a happy faces or neutral faces, and search a neutral face from happy faces or sad faces. The search types of all-element condition were the same as the half-element condition except that another kind of faces were also presented in the display, such as searching a happy face from negative faces and neutral faces, and so on; thus the display size was double that of the half-element condition. For the gap condition, following the fixation cross, some faces were first displayed for 1000 ms, after which the other kinds of faces were added to the display. For example, some happy faces appeared for 1000 ms, and then a negative face and some neutral faces were displayed. Each trial started with a white fixation cross in the center of the screen for 1000ms. Following this, the search display was presented and remained until the participants responded to either the presence or absence of the face that searched. After the participants responded or did not respond after 3000ms, the display disappeared and a new trial began. All participants completed 18 blocks of trials lasting approximately 90min, with block order counterbalanced across the participants. Each block had 60 trials, with an equal number of present and absent trials within a block was presented in random order. Before each block of trials, the participants received a short practice block. If the participants had a low correct percent, they would have to engage the practice again until they passed.

IV Results

A. Reaction time (RT)

The data analyze error trials were first identified and then removed. In the data analyze, at first, we analyzed whether the preview benefit existed when we searched the schematic faces and then analyzed the difference of the three kinds of schematic faces in search in detail. Before the correct reaction times (RTs) were analyzed, the outliers in each cell were removed, using a recursive procedure (see Van Selst & Jolicoeur, 1994). The highest and lowest RTs were removed, one at a time, and the mean and the standard deviation (SD) of the resulting distribution were calculated. If the extreme RT was more than four SDs from the mean, it was considered an outlier and was removed. This procedure was repeated until no outliers remained. A total of 2.02% of the trials were removed in this manner. The remaining RT data were then evaluated by a 2 3 4 analysis of variance (ANOVA) to assess target type (positive vs. Negative face) and set size (7, 11, 15, and 19). All of three main effects were significant: Condition, $F(2, 9) = 393.92$, $p < 0.001$; Target, $F(1, 10) = 398.62$, $p < 0.001$; and Display size, $F(2, 9) = 193.11$, $p < 0.001$. Present trials were faster than absent trials, RTs increased as the display size increased, and RTs were fastest in the half-element condition and slowest in the all element condition (see Fig. 1.). The two separate two-way within subject ANOVAs showed the comparison of the gap condition and each of the two other conditions. RTs in the gap condition and half element condition had no significant difference, $F(1, 10) = 1.437$, $p = 0.258$, which showed the performance in the gap condition was as efficient as the half-element condition. RTs were faster in the gap condition than in the all-element condition, $F(1, 10) = 593.63$, $p < 0.001$, which showed the performance in the gap condition, was more efficient than the all-element condition.

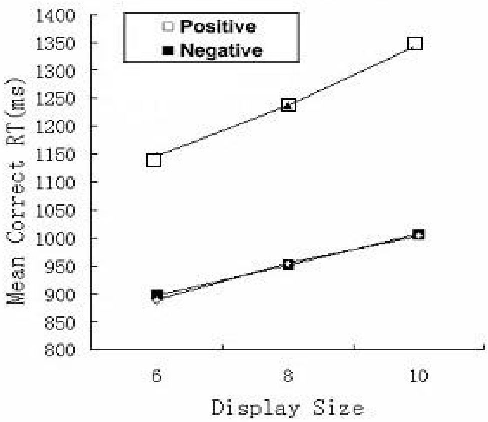

Fig. 2 Mean correct reaction times (RTs) as a function of condition and display size

B. The Difference among Happy Faces, Sad Faces and Neutral Faces

Search a happy face, a sad face and a neutral face respectively. All of the three main effects were significant: Face type, $F(2, 9) = 128.64$, $p < 0.001$; Target, $F(1, 10) = 381.90$, $p < 0.001$; and Display Size, $F(2, 9) = 182.46$, $p < 0.001$. RTs were fastest in the search of a neutral face and slowest in the search of a happy face, present trials were faster than absent trials, and RTs increased as the display size increased. There was also a significant Face Type×Display Size interaction, $F(4, 7) = 9.51$, $p < 0.01$, and Target×Display Size interaction, $F(2, 9) = 459.26$, $p < 0.001$. No other interactions were significant. Search a happy face in sad faces versus search a sad face in happy faces. The difference of the two kinds of search was significant: RTs in the search of a sad face in happy faces were faster than in the search of a happy face in sad faces, $F(1, 10) = 8.69$, $p < 0.05$. The results showed the search was asymmetry. Search a happy face in neutral faces versus search a sad face in neutral faces. The difference of the two kinds of search was significant: RTs in the search of a sad face in some neutral faces were faster than in the search of a happy face, $F(1, 10) = 6.09$, $p < 0.05$.

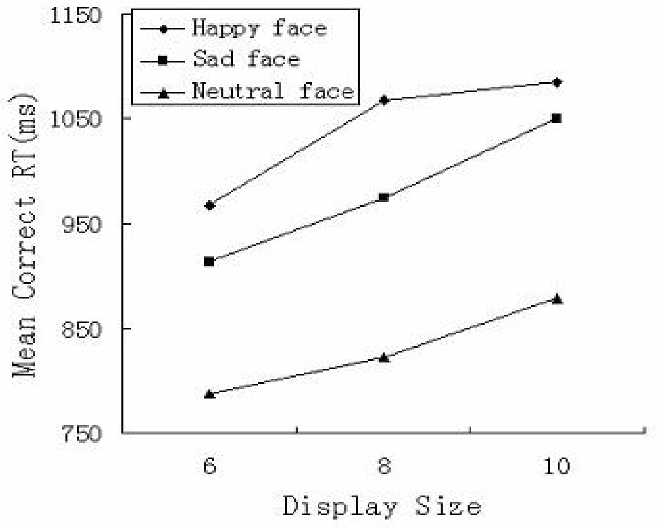

Fig. 3 Mean correct reaction times (RTs) as a function of three kinds of faces

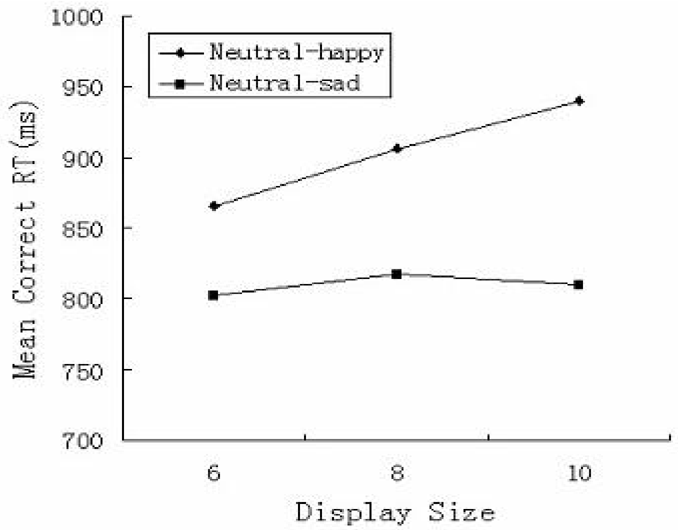

Fig. 4 Mean correct reaction times (RTs) as a function of three kinds of faces

V Discussion

The results showed that it was faster when searching a sad face than searching a happy face. However, in half element condition, when sad faces and happy faces were both to be used as distracted, it was faster when searching a neutral face in sad faces than in happy faces. Mack and Rock (1998) found the participants detect the presence of a positive face more easily than detect a negative face. By this view, it seemed suspect able to draw a conclusion that negative faces capture attention more effectively than happy faces. One possible explanation is that the results are different in different search condition and different tasks [4]. The results showed that it was usually faster to search a neutral face than to search a happy face or a unhappy face. This difference might be induced by the local feature of the mouth in the faces (see Fig. 1.). However, Eastwood et al. (2001) approved that the differential guidance of focal attention was due to holistic face perception instead of part differences. Although the schematic face images used in the present study were impoverished compared with photographs of real human faces, they have appeared to function as potent effective facial stimuli. This is consistent with a number of findings in the literature. For example, schematic faces appear to communicate emotional meaning effectively (Aronoff, Barclay, & Stevenson, 1988; Cuceloglu, 1970; McKelvie, 1973). Furthermore, it has recently been shown that schematic faces elicit event related potentials that are similar to those elicited by photographs of faces (Sagiv & Bentin, in press). Therefore, given that schematic faces contain fewer feature confounds than do photographed faces and yet appear to function as effective face stimuli it may actually be preferable to use schematic faces in studies in which the perception of facial expression is investigated Our conclusion that attention can be guided by unattended faces expressing emotion is consistent with other findings showing that visual search is influenced by the global representations of faces that are formed by composites of parts. Compared visual search for a unique feature (e.g., down arc) when it was located in 6, 12, or 18 triplets of features (e.g., up arcs) arranged to form either faces or meaningless patterns. They found that visual search for the unique feature was less efficient, as was indicated by the slopes of the search functions, when the features were arranged to form faces. On the basis of these results, Suzuki and Cavanagh concluded that a global representation of a face can have priority during visual search and that the processing of the global representation can even pre-empt access to such features as the curvature of an arc. The present findings extend these conclusions by suggesting not only that faces are perceived outside of the focus of attention, but also that emotional expressions are perceived outside of the focus of attention and play a functional role in guiding focal attention.

VI Conclusion

To study emotion, although the schematic faces were impoverished as compared with photographs of real human faces, they appeared to be potent affective facial stimuli. Schematic faces appeared to communicate emotional meaning effectively and showed disruptions in perception when inverted similar to those found with photographed faces. The results showed that it was usually faster to search a neutral face than to search a happy face or a unhappy face. This difference might be induced by the local feature of the mouth in the faces. Emotion Detection from Natural Scene Images is a new concept in an innovative field of Image Processing domain. Image processing domain has always proven to be a challenging Criterion in field of research and development. This survey demands a thorough study of every concept related to Image retrieval, emotion detection and CBIR (Content Based image Retrieval) technique.

References

[1] Aronoff, J., Barclay, A. M., & Stevenson, L. A. (1988). “The recognition of threatening facial stimuli”. *Journal of Personality & Social Psychology*, 54, 647-655.

[2] Van Selst, M., & Jolicoeur, P. (1994). “A solution to the effect of sample size on outlier elimination”. *Quarterly Journal of Experimental Psychology*, 47A, 631-650.

[3] Hansen, C. H., & Hansen, R. D. (1994). “Automatic emotion: Attention and facial efference*”. In P. M. Niedenthal & S.Kitayama (Eds.),* “The heart‟s eye: Emotional influences in perception and attention “(pp. 217- 243). *San Diego: Academic Press.*

[4] Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). “What causes the face inversion effect?” *Journal of Experimental Psychology: Human Perception* & Performance, 21, 628-634.

[5] John D. Eastwood, Daniel Smilek, and Philip M. “Merikle Differential attention guidance by unattended faces expressing positive and negative emotion” *University of Waterloo, Waterloo, Ontario, Canada, Perception & Psychophysics 2001, 63 (6), 1004- 1013*

[6] Fang Hao1,2, Hang Zhang1,2, and Xiaolan Fu1, “ Modulation of Attention by Faces Expressing Emotion: Evidence from Visual Marking”, *Graduate School of the Chinese Academy of Sciences*, Beijing 100049, China

[7] Farah, M. J., Wilson, K. D., Drain, M., Tanaka, J. N.: “What is “special” about face perception?” Psychological *Review*. 105 (1998) 482-498

[8] Fenske, M. J., Eastwood, J. D.: *Modulation of focused attention by faces expressing emotion: evidence from flanker tasks*. Emotion. 3 (2003) 327

[9] Smilek, D., Eastwood, J. D., Merikle, P.: “Does unattended information facilitate change detection?” *Journal of Experimental Psychology: Human Perception & Performance*. 26 (2000) 480-487

[10] McKelvie, S. J. (1973). “The meaningfulness and meaning of schematic faces”. *Perception & Psychophysics*, 14, 343-348.

[11] Eastwood, J., Smilek, D., Merikle, P. M.: *Differential attentional guidance by unattended faces expressing positive and negative emotion*. Perception & shchophysics. 63 (2001) 1004-1013

[12] Morris, J. S., Öhman, A., & Dolan, R. J. (1998). “Conscious and unconscious emotional learning in the human amygdale”. *Nature*, 393, 467-470.

[13] Sagiv, N., & Bentin, S. (in press). *Structural encoding of human and schematic faces: Holistic and part-based processes*. Journal of Cognitive Neuroscience.

[14] Suzuki, S., & Cavanagh, P. (1995). *Facial organization blocks access to low-level features: An object inferiority effect*. Journal of Experimental Psychology: Human Perception & Performance, 21, 901-913.

