Scaring or Pleasing: Exploit Emotional Impact of An Image

Bing Li
National Laboratory of pattern Recognition,Institute of Automation, Chinese Academy of Sciences
bli@nlpr.ia.ac.cn

Songhe Feng *
School of Computer & Information Technology, Beijing Jiaotong University, Beijing China
shfeng@bjtu.edu.cn

Weihua Xiong
National Laboratory of pattern Recognition,Institute of Automation, Chinese Academy of Sciences
wallace.xiong@gmail.com

Weiming Hu
National Laboratory of pattern Recognition,Institute of Automation, Chinese Academy of Sciences
wmhu@nlpr.ia.ac.cn

ABSTRACT

Automatic image emotion analysis has emerged as a hot topic due to its potential application on high-level image understanding. Considering the fact that the emotion evoked by an image is not only from its global appearance but also interplays among local regions, we propose a novel affective image classification system based on bilayer sparse representation (BSR). The BSR model contains two layers: The global sparse representation (GSR) is to define global similarities between a test image and all the training images; and the local sparse representation (LSR) is to define similarities of local regions' appearances and their co-occurrence between a test image and all the training images. The experiments on real data sets demonstrate that our system is effective on image emotion recognition.

Categories and Subject Descriptors

I.5.2 [**Pattern Recognition**]: Design Methodology-Classifier design and evaluation—*complexity measures, performance measures*; H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing

Keywords

affective image classification, bilayer sparse Representation

1. INTRODUCTION

An image is worth a thousand words and does not just contain the physical words. It has the ability to convey emotional information that makes us feel happy, excited, disgusted or curious, etc. Therefore, it is interesting to design and develop an automatic image classification system based on its affective content [8].

Due to the well-known issue of semantic gap in image understanding, very limited attention has been focused on developing techniques related to affective image classification. According to whether a learning procedure is involved or not, existing affective images classification methods can be roughly divided into two kinds of strategies: machine learning based strategies and domain knowledge based strategies [5]. Methods belonging to the first category aim to train a mapping function between low-level features and high-level emotional semantics in a manner of "black box" [6]. On the other hand, the methods from the latter category try to build hierarchal model using image itself [1]. Both kinds of existing affective image classification algorithms mentioned above share a big drawback. Only global-level visual features are employed for the affective images classification, while some important local and contextual information is ignored, which makes the classification performance unsatisfactory. However, the affective feeling of an image has been proved to be affected by not only its global view or local regions' appearances, but also contextual interrelations among regions as well as reciprocity between local regions and global background[3].

*indicates the equal contribution to the first author

MM'12, October 29–November 2, 2012, Nara, Japan.
ACM 978-1-4503-1089-5/12/10.

2. SYSTEM OVERVIEW

In this paper, we propose an affective image classification system based on bilayer sparse representation model (BSR) that contains two layers: global sparse representation (GSR) and local sparse representation (LSR). Figure 1 gives out an overview of the proposed system. The GSR is to reconstruct the test image using training images from global viewpoints. Principal Component Analysis (PCA) is applied on the global feature to reduce the dimension before it is fed into BSR module. An input image is also segmented into regions. And these regions' appearances and interplay will be used to define local features' similarity (LSR) between test image and training ones.Therefore, the contextual reciprocity between global features and local regions are considered through mutual constraints in our system. The reconstruction coefficient obtained from BSR is used to get the reconstruction residual of each class to get the final emotional class label of the test image.

3. FEATURE EXTRACTION

In the GSR layer of the proposed system, the HSV color histogram feature, Gabor texture feature, and bag of words

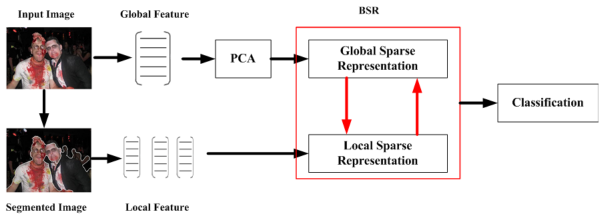

Figure 1: Framework of the proposed system.

on SIFT descriptor three kinds of global features are used. After merging these feature vectors, we can obtain a 300D global feature vector. Then the Principal component analysis is used to reduce the dimension to 150D. In the scenario of local feature extraction, we firstly apply the well-known JSEG algorithm [2] to segment each image into homogeneous regions. For each region, the HSV color feature, color emotion feature [7] and Gabor texture feature is also extracted.

4. BILAYER SPARSE REPRESENTATION FOR EMOTION CLASSIFICATION

Generally speaking, an image can be viewed as a bag of certain regions. To a certain extent, the regions' appearances and their co-occurrence can reflect the scene content and corresponding affective feeling. In order to consider both appearances and co-occurrence property of regions in the same image, the multi-task joint sparse representation [10]is introduced to reconstruct the regions in a test image using those regions from the same training image as possible as we can. On the other hand, besides the interplay among regions of any image, the global visual character is also another important factor related with an image's affective expression. Therefore, the sparse coding technique is used to reconstruct the global visual appearance of a test image using the training images in the global layer. Then the mutual constraints of the two layers is used to make the BSR model simultaneously take the local context and global-local context into account. The details of the proposed BSR model can be found in [4].

In the BSR model, the key issue is to find the reconstruction coefficient matrix and coefficient vector in both layer. After getting the coefficients vector and coefficient matrix , the final class label that is assigned to the test image is the one that gives the smallest reconstruction residual in both global and local layer.

5. EVALUATION

The emotion categories used this paper are the word list defined in a psychological study on affective images that has been widely used in affective images classification [9][6][5]. The emotional output categories $\mathbf{T}$ ={Amusement, Awe, Contentment, Excitement, Anger, Disgust, Fear, Sadness}. The proposed system is evaluated on two image sets. The first one is the International Affective Picture System (IAPS) set[1] [9] . The IAPS set is a standard emotion-evoking image set in psychology. It consists of 716 natural scene images taken by professional photographers and depicts complex scenes containing objects, people, and landscapes. The images used in our experiment are from a subset with 394 images from the IAPS set that has been used in many affective image classification researches [9][6][5]. The second image set is the artistic photographs (AP) set[2] that are collected by Machajdik et al [6] from an art sharing site[3]. The AP set contains 806 images obtained using the emotion categories as search terms in the art sharing site.

[1]http://csea.phhp.ufl.edu/media.html

6. CONCLUSION

Automatic affective images classification, whose final goal is to understand images from emotional semantic, has attracted more and more people's attention. Considering the fact that the emotion evoked by an image is not only from its global features but also interplays among local regions, we propose a novel context-aware classification system based on bilayer sparse representation (BSR). Our method takes into account global images' similarities and local regions similarities as well as the regions' co-occurrence property for affective images classification. The power of our proposed system is demonstrated by experimental results on two real image data sets.

7. ACKNOWLEDGMENTS

This work is partly supported by the National Nature Science Foundation of China (No. 61005030, 60935002, 60825204 and 61100142) and Chinese National Programs for High Technology Research and Development (863 Program) (No. 2012AA012503 and No. 2012AA012504).

8. REFERENCES

[1] C. Colombo, A. D. Bimbo, and P. Pala. Semantics in visual information retrieval. *IEEE Multimedia*, 6(3):38–53, 1999.

[2] Y. Deng and B. S. Manjunath. Unsupervised segmentation of color-texture regions in images and video. *IEEE TPAMI*, 23(8):800–810, 2001.

[3] B. Li, W. Xiong, and W. Hu. Web horror image recognition based on context-aware multi-instance learning. In *Proc. of ICDM*, pages 1158–1163, 2011.

[4] B. Li, W. Xiong, W. Hu, and X. Ding. Context-aware affective images classification based on bilayer sparse representation. In *Proc. of ACM MM*, 2012.

[5] S. Y. Liu, D. Xu, and S. H. Feng. Emotion categorization using affective-plsa model. *Optical Engineering*, 49(12):127201, 2010.

[6] J. Machajdik and A. Hanbury. Affective image classification using features inspired by psychology and art theory. In *Proc. of ACMMM*, pages 83–92, 2010.

[7] L. Ou, M. Luo, A. Woodcock, and A. Wright. A study of colour emotion and colour preference. part 1: Colour emotions for single colours. *Color Res. & App.*, 29(3):232–240, 2004.

[8] W. Wang and Q. He. A survey on emotional semantic image retrieval. In *Proc. of ICIP*, pages 117–120, 2008.

[9] V. Yanulevskaya, J. C. van Gemert, and K. Roth. Emotional valence categorization using holistic image features. In *Proc. of ICIP*, pages 101–104, 2008.

[10] X. Yuan and S. Yan. Visual classification with multi-task joint sparse representation. In *Proc. of CVPR*, pages 3493–3500, 2010.

[2]http://www.imageemotion.org/
[3]http://www.deviantart.com