Understanding the Emotional Impact of Images

Xiaohui Wang[†‡♯], Jia Jia[†‡♯], Peiyun Hu[♯], Sen Wu[♯], Jie Tang[♯], Lianhong Cai[†‡♯]
[†]Key Laboratory of Pervasive Computing, Ministry of Education
[‡]Tsinghua National Laboratory for Information Science and Technology (TNList)
[♯]Department of Computer Science and Technology, Tsinghua University
wangxh09@mails.tsinghua.edu.cn, jjia@tsinghua.edu.cn, tsinghua.hp@gmail.com, ronaldosen@gmail.com,
jietang@tsinghua.edu.cn, clh-dcs@tsinghua.edu.cn

Categories and Subject Descriptors: H.3.3 [Information Search and Retrieval]: Retrieval models

General Terms: Algorithms

Keywords: Emotional impact, social networks, prediction model

1. INTRODUCTION

Images often embed users' emotional impact. Understanding the emotional impact can benefit many applications such as image search and context-aware recommendation. However, fulfilling the task is not a trivial issue, and the biggest challenge is to model the images and to capture the intrinsic relationships between the image features and the emotional impact. Traditional methods treat the relation modeling as a simple regression problem, and use SVM, Random Forest and Naive Bayes directly on various low-level visual features. However, these methods usually ignore the essential emotion related characteristics, which narrow the deep and high level understanding of emotions. Another, challenge usually ignored is that most images are generated in social networks, where users have complex and subtle influence with the emotional impact of each other.

We study the problem of understanding the emotional impact of images and its applications. We introduce a novel notion of *aesthetic effects* as the intermediate layer to model high level semantics of images. We develop a *semi-supervised factor graph* model which incorporates both the single image features and the image correlations to better predict the emotional impact. The proposed method is general and can be applied to various applications. In particular, we consider two applications: Affective image adjustment and emotion prediction from images.

2. SOLUTIONS

We propose a novel framework of understanding the emotional impact of images, which is shown in Figure 1.

Problem formulation. Modeling emotional impact for massive Internet dataset. Given a partially labeled network $G = (V^L, V^U, E, \mathbf{X})$, where V^L and V^U are the set of images with labeled/unlabeled emotions, $E \subset V \times V$ is the edge set which represents the correlations between images, and $\mathbf{X}$ is a $|V| \times d$ feature matrix with each element x_{ij} denoting the value of the j^{th} feature of image v_i. Our goal is to learn an emotional category prediction function $f : G = (V^L, V^U, E, \mathbf{X}) \to Y$, where $Y = \{y_1, \ldots, y_{|V|}\}$, and y_i indicates the category of image v_i.

Aesthetic effects. Here we use the aesthetic effects to denote quantitative feature descriptions inspired by art theories. The Aesthetic effects are a few low dimensional features, denoted as $D = \{D_1, D_2\}$, where D_1 is a combined feature space related to the color theme of an image, and $D_2 = (wc, hs)$ is the *image-scale* space [1] which is a two-dimension space (*warm-cool* and *hard-soft*) and is well recognized in art design. Psychological experiments have shown that color themes are crucial for the recognition of emotions. Here we adopt the 5-color themes as their representations, which is compatible with industrial practices. In [1], 490 5-color themes and 180 keywords are labeled with the image-scale space based on psychophysical investigations. With these data, we learn a continuous mapping from the feature space D_1 to the image-scale space D_2 by the LASSO regression framework.

Prediction model. To better utilize the Internet images and their connections, and to overcome the *noise*, we adopt a partially-labeled factor graph model (PFG) for learning and predicting the emotional impact of images. As noted in *Problem formulation*, we denote the labels of image nodes as $Y = \{y_1, \ldots, y_{|V|}\}$, where y_i is a hidden variable associated with v_i. For each image node v_i, we define the features from the aesthetic effects as a vector $\mathbf{x}_i$. Then we can construct a factor graph model by defining the following two factors: (1) feature factor: $f(y_i, \mathbf{x}_i)$ represents the posterior probability of the emotional category y_i given the feature vector of image v_i; and (2) correlation factor: $g(y_i, N(y_i))$ denotes the correlation between the emotional categories, where $N(y_i)$ is the set of correlated categories to y_i. Finally, we define the joint distribution over Y as $P(Y|G) = \prod_i f(y_i, \mathbf{x}_i) g(y_i, N(y_i))$. The final labels are calculated by maximizing the posterior probability P [2].

Applications. Benefited from the generality of the aesthetic effects layer and prediction model, the proposed framework has various applications. Specifically, we consider two applications: (1) Affective image adjustment. The system supports automatically changing the emotional impact of an image driven by a single word. (2) Emotion prediction from images. Based on the factor graph model, we design a system to predict the emotion from images uploaded to the largest photo sharing website, Flickr. Besides, we also try to use this model to predict the emotional impact of artworks, such as Van Gogh's paintings.

Acknowledgements. This work is supported by the National Basic Research Program (973 Program) of China(2011CB302201), and Science Foundation of China (61003094), and National Significant Science and Technology Projects of China under Grant (No.2011ZX01042-001-002). Jie Tang is supported by Natural Science Foundation of China (61073073), Chinese National Key Foundation Research (No. 60933013).

3. REFERENCES

[1] S. Kobayashi. *Art of Color Combinations*. Kosdansha International, 1995.

[2] W. Tang, H. Zhuang, and J. Tang. Learning to infer social ties in large networks. In *ECML/PKDD'11*, pages 381–397, 2011.

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

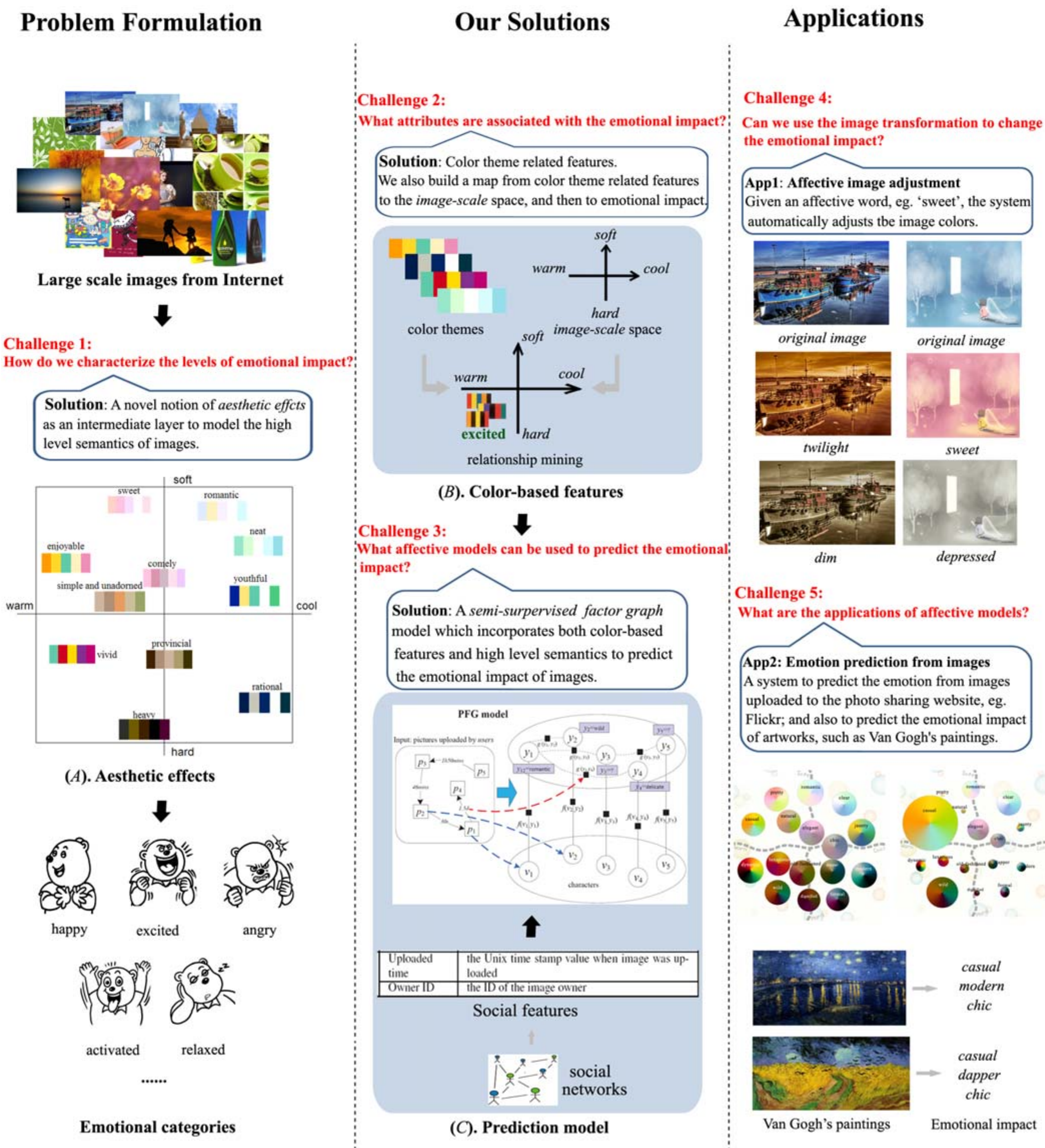

Figure 1: Framework. We propose a novel framework of emotional impact understanding from images. The aesthetic effects (*A*) are closely related to the aesthetic expressions of images (e.g. elegant, classic, romantic). We use the aesthetic effects as the intermediate layer to reduce the gap between the visual features and the high-level semantics. A factor graph based semi-supervised prediction model is presented by incorporating the aesthetic effects and various connections in social networks (*C*). The method is applied in two applications: affective image adjustment and emotion prediction from images.