SentiBank: Large-Scale Ontology and Classifiers for Detecting Sentiment and Emotions in Visual Content

Damian Borth[1,2], Tao Chen[1], Rong-Rong Ji[1] and Shih-Fu Chang[1]
[1]Columbia University (USA) and [2]University of Kaiserslautern (Germany)

ABSTRACT

A picture is worth one thousand words, but what words should be used to describe the sentiment and emotions conveyed in the increasingly popular social multimedia? We demonstrate a novel system which combines sound structures from psychology and the folksonomy extracted from social multimedia to develop a large visual sentiment ontology consisting of 1,200 concepts and associated classifiers called *SentiBank*. Each concept, defined as an Adjective Noun Pair (ANP), is made of an adjective strongly indicating emotions and a noun corresponding to objects or scenes that have a reasonable prospect of automatic detection. We believe such large-scale visual classifiers offer a powerful mid-level semantic representation enabling high-level sentiment analysis of social multimedia. We demonstrate novel applications made possible by SentiBank including live sentiment prediction of social media and visualization of visual content in a rich intuitive semantic space.

Categories and Subject Descriptors

H.3.3 [**Information Storage and Retrieval**]: Information Retrieval and Indexing

Keywords

Social Multimedia, Affect, Emotion, Sentiment Analysis, Ontology, Concept Detection

1. INTRODUCTION

A picture is worth one thousand words: visual content offers rich complementary information beyond what the accompanying text reveals. The role of visual content has become even more prominent in increasingly popular social media such as Twitter, where the textual description is limited to very short messages (140 characters). In such cases, extracting information from visual content is critical in understanding the rich emotions, affect, or sentiments conveyed in the multimedia content.

Existing research of sentiment or affect analysis of social multimedia are either restricted to text analysis or direct mapping of low-level visual features to affects [1, 2]. Here we advocate a novel approach focusing on construction of a large concept ontology and associated automatic classifier library that aim at extracting mid-level semantic attributes from visual content. Our effort is inspired by major progresses made in semantic concept discovery and learning in the multimedia and computer vision community in the last decade. But to make high-level affect analysis feasible, we specifically focus on semantic concepts that draw sound foundation from the psychology theory of emotion [3], have close relevance to actual terms used by social media users, and consider practical detectability using current computer vision and multimedia analytics techniques.

We have developed a large-scale visual sentiment ontology which includes 1,200 semantic concepts and corresponding automatic classifiers, called SentiBank [4]. Each concept is defined as an Adjective Noun Pair (ANP), made of an adjective strongly indicating emotions and a noun corresponding to objects or scenes that have a reasonable prospect of automatic detection. We demonstrate several exciting applications enabled by SentiBank in this paper: (1) robust prediction of affects (sentiment or emotion) in visual content, (2) interactive exploration of large image data sets along the high-dimensional sentiment concept space using efficient visualization tools such as emotion wheel and treemap, and finally (3) a multimodal interface (incorporating novel sound and animation effects) for monitoring the sentiment concepts present in live social multimedia streams.

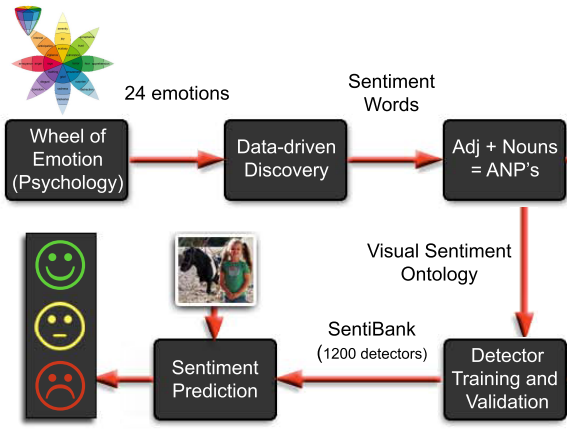

Figure 1: Processes for constructing ontology and learning concept classifiers in SentiBank.

MM'13, October 21–25, 2013, Barcelona, Spain.

http://dx.doi.org/10.1145/2502081.2502268.

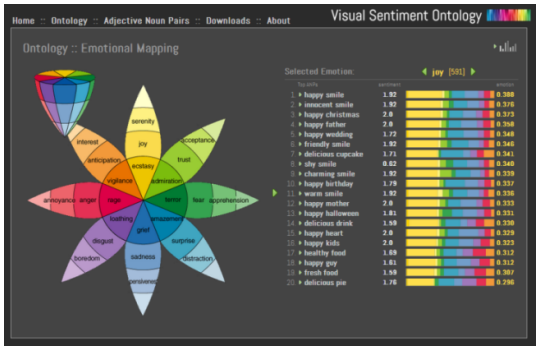

Figure 2: Emotion Sentiment Concept Browser.

Figure 3: TreeMap Sentiment Concept Browser.

2. SENTIBANK ONTOLOGY AND CLASSIFIERS

Fig. 1 summarizes the processes for constructing the visual sentiment ontology and large concept classifiers in SentiBank. The process starts with a well-established psychological model of emotions called *Plutchik's Wheel of Emotions* [3]. First, a data-driven mining step is performed for each of the 24 emotions in Plutchik's theory to find related tags. Images and videos from Flickr and YouTube are retrieved with emotions as query words. Then, tags frequently co-occurring with each emotion are found (e.g., "joy" leads to "happy" and "beautiful"). We then applied text parsing and lexical based sentiment analysis tools to find "polarized" (either positive or negative) adjectives, which are then paired with nouns frequently used together as phrases in the retrieved image set to form more than 3,000 Adjective Noun Pairs (ANP). We then trained a machine learning classifier for each ANP by using the corresponding training images crawled from Flickr. The final set of classifiers, filtered by detection accuracy, includes 1,200 ANP concept detectors, covering 178 adjectives and 313 nouns. The majority of the classifiers has accuracy, measured in F-score, higher than 0.6 (with all concept detectors having AP@20>0), verifying reasonable detectability of classifiers. The full list of concepts and details of the training process can be found in [4].

3. APPLICATIONS AND DEMOS

To demonstrate the power of the proposed sentiment ontology and SentiBank concept classifiers, in the following we describe a few novel applications for large image exploration and live media monitoring. The interactive demos can also be viewed online at http://visual-sentiment-ontology.appspot.com/.

Demo A: Emotion Sentiment Concept Browser

Taking advantage of the psychology principled structure in the SentiBank ontology, we have developed an efficient browser for exploring big image sets (currently more than 500,000 Flickr images) in the rich emotion space (Fig. 2). The wheel provides an intuitive map for arranging emotions based on their relations and intensities (8 emotion groups and 3 intensity levels). Upon selecting a specific emotion, users can view the related ANP concepts (shown on the right) together with information about the sentiment value and distribution of associated emotions of each ANP. This gives users a quick overview of the entire emotion landscape, related ANP concepts, and their corresponding sentiment values. Users can then zoom in to get detailed information such as example images and detector accuracy of each specific ANP concept.

Demo B: Treemap Sentiment Concept Browser

The hierarchical structure of SentiBank ontology allows development of a treemap style browsing system for navigating through the ontology and the mapped image set (Fig. 3). The four-level structure of the TreeMap includes the ontology root, emotion groups, adjectives, and finally specific ANPs. The size of each block in the map is proportional to the number of children nodes at the top levels and the number of Flickr images at the leaf node. The color varying from green to red illustrates the sentiment value. The user can browse the ontology by zooming in and out following the edges of the tree. The leaf node will also show the average image and link to the sample images of the corresponding ANP. Compared to the emotion wheel described above, the treemap offers a quick summary of nodes at each level of the ontology, augmented by color and size cues to indicate the additional attributes such as sentiment and data volume.

Demo C: Live Monitoring of Visual Sentiment and Emotions

We have also extended the treemap browser to a new visualization system for monitoring the sentiment and ANP concepts detected in images in live social media streams such as those from Twitter. To help users comprehend rich information in a large number of dimensions, we utilize multimodal cues in a new way by mapping the detected sentiment value of an incoming image to the pitch of a music note, i.e., positive (negative) images are displayed together with playing of a high (low) pitch music note. In addition, the specific ANP concepts detected will be highlighted in the corresponding areas in the treemap in a synchronized animation manner. We found such multimodal visualization tools enhanced with sound and animation effects can significantly improve the usability and attractiveness of the system.

4. REFERENCES

[1] J. Machajdik and A. Hanbury. Affective Image Classification using Features inspired by Psychology and Art Theory. *ACM MM*, 2010.

[2] J. Jia et al. Can we Understand van Gogh's Mood? Learning to infer Affects from Images in Social Networks. *ACM MM*, 2012.

[3] R. Plutchik. *Emotion: A Psychoevolutionary Synthesis.* Harper & Row New York, 1980.

[4] D. Borth, R. Ji, T. Chen, T. Breuel, and S-F Chang Large-scale Visual Sentiment Ontology and Detectors Using Adjective Noun Pairs *ACM MM*, 2013.