

Chapter 2

Sentiment Analysis Using Social Multimedia

Jianbo Yuan, Quanzeng You and Jiebo Luo

Abstract Sentiment analysis is one of the most active research areas in natural language processing, web/social network mining, and text/multimedia data mining. The growing importance of sentiment analysis coincides with the popularity of social network platforms, such as Facebook, Twitter, and Flickr, which provide a rich repository of people's opinion and sentiment about a vast spectrum of topics. Moreover, the fact that we are exposed to a tremendous amount of data in different forms including text, images, and videos makes sentiment analysis a very challenging task due to its nature of multimodality. In this chapter, in order to provide a big picture of sentiment analysis, we will discuss some of the latest works on topics of sentiment analysis based on visual content and textual content.

2.1 Introduction

Sentiment analysis and opinion mining are fields of study that analyze people's opinions, evaluations, attitudes, and emotions generally from written language [1]. To the best of our knowledge, the term sentiment analysis first appeared in [2]. Because of the explosive communication and information exchanges using social media, researchers are now given the opportunity to access a tremendous amount of texts and images that express people's opinions and sentiments. Therefore, research in sentiment analysis not only has an important impact on Natural Language Processing, but may also have a profound impact on management sciences, political science, economics, and social sciences as they are all affected by opinions.

J. Yuan · Q. You · J. Luo (✉)
Department of Computer Science, University of Rochester,
Computer Studies Bldg, Rochester, NY 14620, USA
e-mail: jluo@cs.rochester.edu

J. Yuan
e-mail: jyuan10@ece.rochester.edu

Q. You
e-mail: qyou@cs.rochester.edu

Sentiment analysis applications have spread to almost every possible domain, from consumer products, services, health care, and financial services to social events and political elections. Many big corporations have also built their own in-house capabilities, e.g., Microsoft, Google, Hewlett-Packard, SAP, and SAS. Such practical applications and industrial interests have provided strong motivations for research in sentiment analysis.

Most recently, social networks such as Twitter and microblogs such as Weibo have become major platforms of information exchange and communication between users, between which the common information carrier is tweets. Social networks such as Twitter and microblogs such as Weibo provide billions of pieces of both textual and visual information, making it possible and imperative to detect sentiment indicated by both textual and visual data, respectively. Multimedia content, such as images, are more likely to express and convey people's subtle feelings compared with text information alone [3]. However, sentiment analysis based on a visual perspective is still in its infancy.

With respect to sentiment analysis, much work has been done on textual information [4–6], as well as online sentiment dictionary [7, 8]. Semantics and concept learning [9–12] based on visual features is another way of sentiment analysis without employing textual information. However, semantics and concept learning approaches are hampered by the limitations of computer vision. The analysis of aesthetics [13, 14], interestingness [15] and affect or emotions [16–19] of images are most related to sentiment analysis based on visual content. Moreover, sentiment analysis based on human activities' content, for example, images of facial expressions, has played a significant role in the fields of psychology and human–computer interaction. Many works have been explored on facial emotion classification [20–22]. Hernandez et al. [23] created a computer vision-based system that automatically encouraged, recognized, and counted smiles on a college campus. Hoque et al. [24] took a step forward by building a novel system that provides ubiquitous access to social skills training based on recognitions of facial expressions, speech, and prosody and responds with verbal and nonverbal behaviors.

There are many existing works on sentiment analysis of social media. In particular, Twitter sentiment analysis is one of the most popular research topics. Most existing methods differ in terms of features and emphasize on different aspects of the problem. Guerra et al. [25] proposed a method to measure the bias of social media users toward a topic. Transfer learning is employed to learn the textual features, in such a way that they can build a more accurate classification model using the user biases as a new feature. However, the identification of users' bias on a particular topic itself may be challenging. In [26], the authors employed label propagation to handle noisy labels and use the network for the propagation of the labels. Their results indicate an improvement of accuracy over existing approaches. In [27], the authors used Twitter as a platform to analyze the language characteristics of mothers during postpartum. Their results indicate that using social media can help discover and understand the health and wellness of women following childbirth. Meanwhile, in [28], a method on

streaming data sentiment analysis is proposed. The core of the solution is a training augmentation procedure. It will automatically incorporate new relevant messages into the training data. In [29], the authors used the social relations extracted from tweets, and applied graph Laplacian to form a sparse formulation. An optimization algorithm is proposed to solve this problem. All of these methods only use textual features for sentiment analysis. Even though noisy labels and network structures are also considered, these approaches did not combine with image features for sentiment analysis, which is another main content feature of tweets.

Meanwhile, other works related to the mining of different aspects of social networks have also been proposed. Kosinski et al. [30] analyzed the likes in Facebook, and discovered that people in social media are more likely to share some common interests with their friends and some particular community. Based on their model, they are able to predict the behavior of the users according to their online social activities. Rao et al. [31] used Bayesian models for latent attribute detection based on topic models. Goel et al. [32] used social media to study the browsing behavior of online users. Wong et al. [33] used online social network data to quantify political leaning using the information extracted from tweets and retweets. Choudhury et al. [34] analyzed the sentiment or mood in social media. They used valence and activation to represent moods. Their work provided validation of conceptualization of human mood.

For social media networks, the network structure itself can also be employed for the analysis of sentiment propagation of different nodes across the network. In [35], the authors used the hyperlinks in the network to analyze the sentiment flow. Their results indicate that a node is significantly influenced by its immediate neighbors. The structure of information propagation graph also illustrates the impact of different sentiment flow patterns. Similarly, users connected in social networks are more likely to have similar opinions. To analyze sentiment, [36] employed network relationship to analyze the sentiment of a group of users over a particular topic. In [29], both the user-content and the user-user relations were exploited for sentiment analysis. More specifically, they proposed a semi-supervised learning framework by using the network relations and formalized the problem in an optimization framework. An empirical study of the proposed framework over two existing Twitter datasets illustrated the improved performance of the algorithm. You and Luo [3] analyzed the sentiment changes of Twitter users using both textual and visual features.

In the remainder of this chapter we will present some latest works on topics of sentiment analysis based on both textual data and visual data. In Sect. 2.2, we introduce Stribute, a novel image sentiment analysis framework based on middle level attributes and eigenface expression detection. In Sect. 2.3, we present a new study aimed at analyzing the sentiment changes of Twitter users due to multimedia data including both visual and textual information. We conclude and look forward to future work in sentiment analysis in Sect. 2.4.

2.2 Sentiment Analysis on Visual Contents

As stated, so far analysis of textual information has been well developed in areas including opinion mining [4, 5], human decision making [5], brand monitoring [37], stock market prediction [38], political voting forecasts [4, 39], and intelligence gathering [40]. Figure 2.1 shows an example of image tweets. In contrast, analysis of visual information covers areas such as image information retrieval [41, 42], aesthetics grading [14] and the progress is relatively behind. On the other hand, a recent study shows that images constitute about 36% of all the shared links on Twitter,¹ which makes visual data mining an interesting and active area to explore. As an old saying has it, an image is worth a thousand words. Much like the textual content-based mining approach, extensive studies have been done regarding aesthetics and emotions in images [13, 15, 43].

Visual content analysis has always been important yet challenging. Thanks to the popularity of social networks, images become a convenient carrier for information diffusion among online users. Aiming to conduct visual content-based sentiment analysis, current approaches include employing low-level features [16, 44, 45], via facial expression detection [46] user intent [47], and understanding images using attribute learning [48, 49]. Sentiment analysis approaches based on low-level features have the limitation of low interpretability, which in turn makes it undesirable for high-level use. Metadata of images is another source of information for high-level feature learning [50]. However, not all images contain such kind of data and researchers are trying to incorporate techniques such as attribute learning and scene understanding before going to final sentiment classification. As for understanding

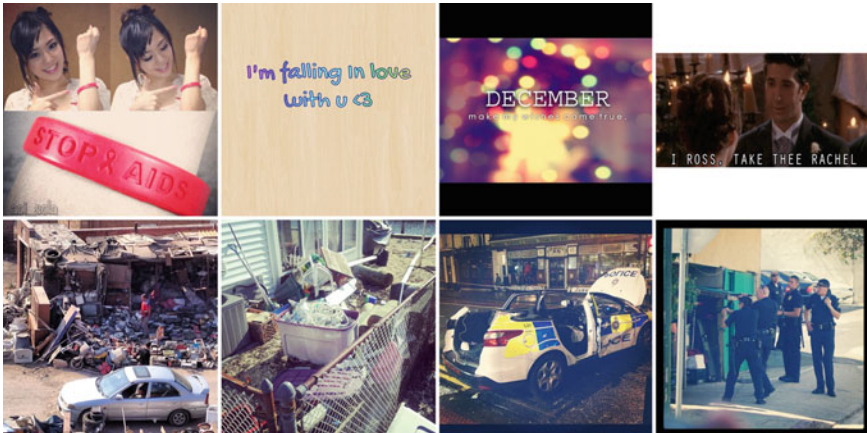


Fig. 2.1 Selected images crawled from Twitter showing (*first row*) positive sentiment and (*second row*) negative sentiments

¹ <http://socialtimes.com/is-the-status-update-dead-36-of-tweets-are-photos-infographic/>.

the visual concepts of an image, [48] established a large-scale Visual Sentiment Ontology (VSO) consisting of more than 3,000 adjective noun pairs and used as detectors for image sentiment. Stribute [49], on the other hand, built 102 middle level attributes and used them as features for sentiment classification.

To understand the diffusion patterns and different aspects of the social images, we need to interpret the images first. Similar to textual content, images also carry different levels of sentiment to their viewers. However, different from text, where sentiment analysis can use easily accessible semantic and context information, how to extract and interpret the sentiment of an image remains quite challenging. In this section, we introduce an image sentiment prediction framework, which leverages the mid-level attributes of an image to predict its sentiment. This makes the sentiment classification results more interpretable than directly using the low-level features of an image. To obtain better performance on images containing faces, we employ eigenface-based facial expression detection as an additional mid-level attribute. An empirical study of the proposed framework shows improved performance in terms of prediction accuracy. More importantly, by inspecting the prediction results, we are able to discover interesting relationships between mid-level attribute and image sentiment.

Compared to the state-of-the-art algorithms, the main contribution of Stribute to this area is two-fold: first, the proposed Stribute, an image-sentiment analysis algorithm based on 102 mid-level attributes, of which results are easier to interpret and ready-to-use for high-level understanding. Second, we introduce eigenface to facial sentiment recognition as a solution for sentiment analysis on images containing people. This is simple but powerful, especially in cases of extreme facial expressions, and contributed an 18 % gain in accuracy over decision making only based on mid-level attributes, and 30 % over the state-of-art methods based on low-level features.

2.2.1 Framework Overview

Figure 2.2 presents the proposed Stribute framework. The idea for this algorithm is as follows: first, we extract scene descriptor low-level features from the SUN Database [47] and use these four features to train the classifiers by Liblinear [16] for generating 102 predefined mid-level attributes, and then use these attributes to predict sentiments. Meanwhile, facial sentiments are predicted using eigenfaces. This method generates really good results, especially in cases of predicting strong positive and negative sentiments, which makes it possible to combine these two predictions and generates a better result for predicting image sentiments with faces. To illustrate how facial sentiment helps refine our prediction based on only mid-level attributes, we present an example in Sect. 2.4, of how to correct the false positive/negative prediction based on facial sentiment recognition.

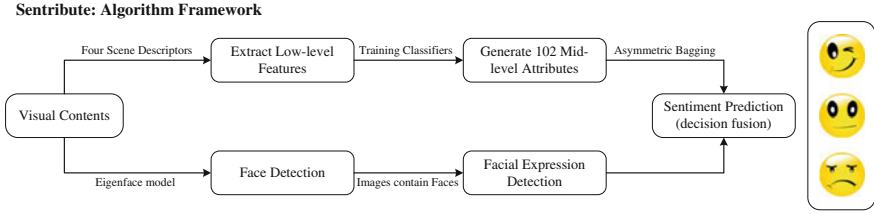


Fig. 2.2 Selected images crawled from Twitter showing **a** positive sentiment and **b** negative sentiments

2.2.2 *Sentribute*

In this section we outline the design and construction of the proposed Sentribute, a novel image sentiment classification method based on mid-level attributes, together with a decision refine mechanism for images containing people. For image sentiment analysis, we conclude the procedure starting from dataset introduction, low-level feature selection, building mid-level attribute classifier, and image sentiment classification. As for facial sentiment recognition, we introduce eigenface to fulfill our intention.

Dataset: Our proposed algorithm mainly contains three steps: first is to generate mid-level attributes labels. For this part, we train our classifier using SUN Database,² the first large-scale scene attribute database, initially designed for high-level scene understanding and fine-grained scene recognition [51]. This database includes more than 800 categories and 14,340 images, as well as discriminative attributes labeled by crowd-sourced human studies. Attributes labels are presented in the form of zero to three votes, of which 0 vote means this image is the least correlated with this attribute, and three votes means the most correlated as shown in Fig. 2.3. Due to this voting mechanism, we have an option of selecting which set of images to be labeled as positive: images with more than one vote, introduced as soft decision (SD), or images with more than two votes, introduced as hard decision (HD). Mid-level attribute classifiers learned based on soft decisions are less likely to be overfitting and less accurate than the classifiers learned based on hard decisions.

The second step of our algorithm is to train sentiment predicting classifiers with images crawled from Twitter together with their textual data covering more than 800 images. Twitter is currently one of the most popular microblog platforms. Sentiment ground truth is obtained from visual sentiment ontology³ with permission of the authors. The dataset includes 1,340 positive, 223 negative and 552 neutral image tweets. For testing, we randomly select 810 images, containing positive (660 image tweets) and negative (150 image tweets). Figure 2.1 shows images chosen from our dataset as well as their sentiment labels.

² <http://groups.csail.mit.edu/vision/SUN/>.

³ <http://www.ee.columbia.edu/ln/dvmm/vso/download/sentibank.html/>.

Camping Attribute	Images given 0 votes	Images given 1 vote	Images given 2 votes	Images given 3 votes
Diving				
Medical Activity				
Cluttered Space				
Fire				

Fig. 2.3 The images in the table above are grouped by the number of positive labels (votes) received from AMT workers. From *left to right* the visual presence of each attribute increases [51]

The final step is facial emotion detection for decision fusion mechanism. We chose to use the Karolinska Directed Emotional Faces dataset [52] mainly because the faces are all well aligned with each other and have consistent lighting, which makes generating good eigenface much easier. The dataset contains 70 men and women over 2 days expressing seven emotions (scared, anger, disgust, happy, neutral, sad, and surprised) in five different poses (front, left prole, right prole, left angle, right angle).

Feature Selection: In this part, we aim to select low-level features for generating mid-level attributes, and we choose four general scene descriptors: GIST descriptor [18], HOG 2x2, self-similarity (SSIM), and geometric context color histogram (GEO-COLOR-HIST) features [53]. These four features were chosen because they are each individually powerful and because they can describe distinct visual phenomena in a scene perspective other than using specific object classifier. These scene descriptor features suffer neither from the inconsistent performance compared to commonly used object detectors for high-level semantics analysis of an image, nor from the difficulty of result interpretation generated based on low-level features.

Generating Mid-level Attribute: Given the selected low-level features, we are then able to train our mid-level attribute classifiers based on SUN Database. We have 14,340 images as training data, and the low-level features of each image add up to more than 170,000 dimensions. For classifier options, Liblinear⁴ outperforms against LibSVM⁵ in terms of training time and maintains similar performance in accuracy in cases where the number of samples are huge and the number of feature dimensions

⁴ <http://www.csie.ntu.edu.tw/~cjlin/liblinear/>.

⁵ <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>.

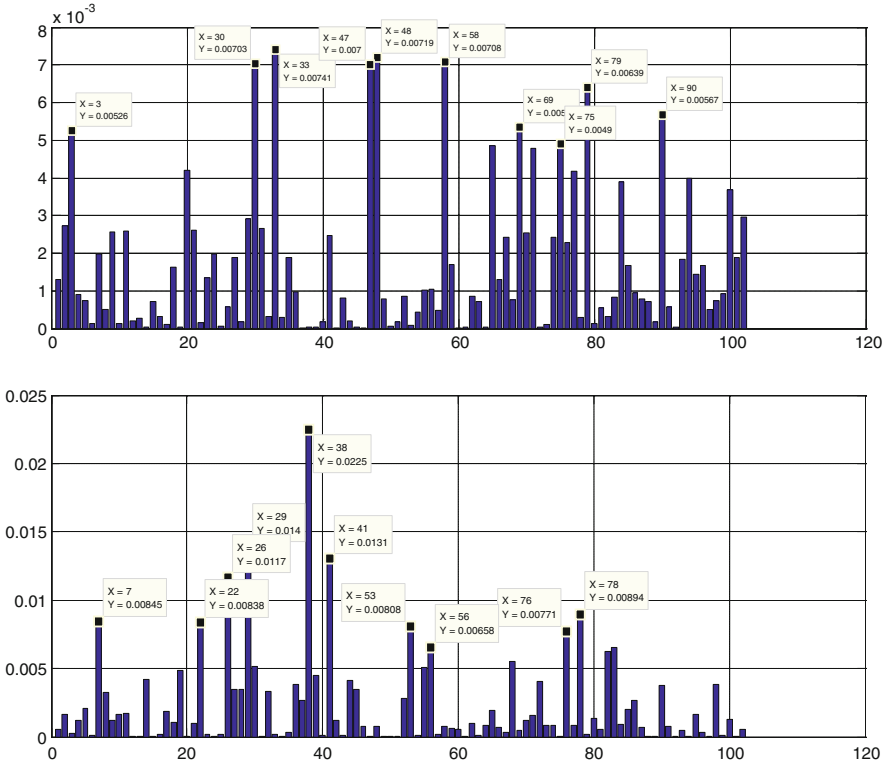


Fig. 2.4 Computing mutual information for each label (*first row is based on SD and second row is based on HD*), where X label indicates the number of each feature and Y label stands for the MI value

is huge. Therefore, we choose Liblinear toolbox to implement SVM algorithm to achieve time saving.

The selection of mid-level attribute also plays an important part in image sentiment analysis. We choose 102 predefined mid-level attributes based on the following criteria: (1) have descent detection accuracy, (2) potentially correlated to one sentiment label, and (3) easy to interpret. We then select four types of mid-level attributes accordingly: (a) Material: such as metal, vegetation; (b) Function: playing, cooking; (c) Surface property: rusty, glossy; and (d) Spatial Envelope [18]: natural, man-made, enclosed.

We conduct mutual information (MI) analysis to discover mid-level attributes that are most correlated with sentiments. Mutual information is a measure of variables' mutual dependence (Fig. 2.4).

For each mid-level attribute, we computed the MI value with respect to both positive and negative sentiment categories (Fig. 2.4). Table 2.1 illustrates the 10 most distinguishable mid-level attributes for predicting both positive and negative labels

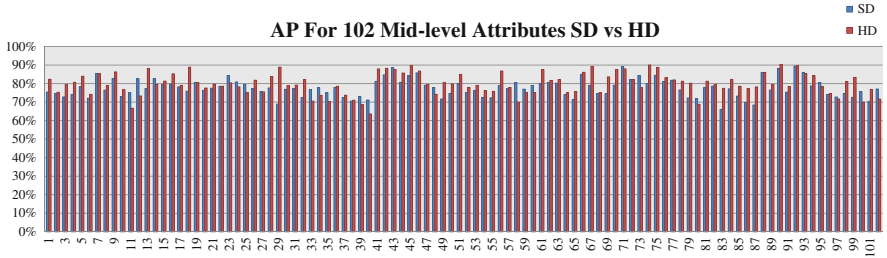


Fig. 2.5 AP of the 102 attributes based on SD and HD

in a descending order based on both SD and HD. Figure 2.5 demonstrates Average Precision (AP) for the 102 attributes we selected, for both SD and HD. It is not surprising to see that attributes of material (flowers, trees, ice, still water), function (hiking, gaming, competing) and spatial envelop (natural light, congregating, aged/worn) all play an important role based on the result of mutual information analysis.

Image Sentiment Classification: In our dataset we have 660 positive samples and 150 negative samples. It is likely to obtain a biased classifier based on these samples alone. Therefore, we introduce asymmetric bagging [54] to deal with biased dataset. Figure 2.6 presents the idea of asymmetric bagging: instead of building one classifier, we now build several classifiers, and train them with the same negative samples together with different sampled positive samples of the same amount. Then we can combine their results and build an overall unbiased classifier.

Facial Sentiment Recognition: Our proposed algorithm, Sentrtribute, contains a final step of decision fusion mechanism by incorporating eigenface-based emotion detection approach. Images containing faces contribute to a great partition of the whole images so that, 382 images from our dataset have faces. Therefore, facial emotion detection is not only useful but important for the overall performance of our algorithm.

In order to recognize emotions from faces we use classes of eigenfaces corresponding to different emotions. Eigenface was one of the earliest successful implementations of facial detection [55]; we modify the algorithm to be suitable for detecting classes of emotions. Although this method is widely appreciated already, we are the first to modify the algorithm to be suitable for detecting classes of emotions, and this method is simple yet surprisingly powerful for detecting facial emotions for front and consistent lightened faces. Note that we are not trying to propose an algorithm that outperforms the state-of-the-art facial emotion detection algorithms. This is beyond the scope of this section.

According to Ekman [56], there are six principal emotions that human’s experience: fear, anger, disgust, happiness, sadness, and surprise. Due to the accuracy of the model and the framework of integrating the results with Sentrtribute, we reduce the set of emotions to positive, neutral, and negative emotions. This is done by classifying the image as one of the seven emotions and then mapping the happy and surprised emotions to positive sentiment, neutral sentiment to itself, and all other emotions to

Table 2.1 Attributes with the top 10 mutual information

TOP 10	Soft decision	Hard decision
1	Congregating	Railing
2	Flowers	Hiking
3	Aged/worn	Gaming
4	Vinyl/linoleum	Competing
5	Still water	Trees
6	Natural light	Metal
7	Glossy	Tiles
8	Open area	Direct sun/sunny
9	Glass	Aged/worn
10	Ice	Constructing

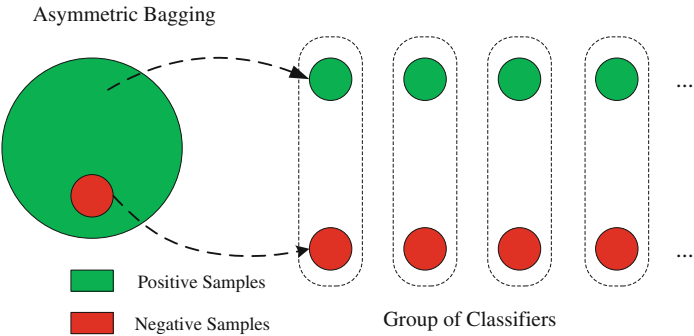


Fig. 2.6 Asymmetric bagging

negative sentiment. At a high level, we are computing the eigenfaces for each class of emotion; we then compare the features of these eigenfaces with the features of the target image projected onto the emotion class space.

The algorithm requires a set of faces to train the classifier (more specifically to find the features of the images). We chose to use the Karolinska Directed Emotional Faces (KDEF) dataset [52] for many reasons, specifically the faces are all well aligned with each other and have consistent lighting, which makes generating good eigenfaces much easier. The dataset contains 70 men and women over 2 days expressing seven emotions (fear, anger, disgust, happy, neutral, sad, and surprised) in five different poses (front, left profile, right profile, left angle, right angle). We use a subset of the KDEF database for our training set, only using the seven frontal emotions from one photographing session.

Training the dataset and extracting the eigenfaces from the images of each emotion class was accomplished by using principal component extraction. We preprocess the

training data by running it through `fdlibmex`,⁶ a fast facial detection algorithm to obtain the position and size of the face. We then extract the face from the general image and scale it to a 64×64 grayscale array; it is then vectored into a 4,096 length vector. We concatenate the individual faces from each class into an $M \times N$ array X , where M is the length of each individual image and N is the number of images in the class. We then are able to find the eigenfaces by using Principal Component Extraction. Principal component extraction converts correlated variables, in our case a set of images, into an uncorrelated variables via an orthogonal transform. We implement principal component analysis by first computing the covariance matrix

$$C = (x - \mu)(x - \mu)^T, \quad (2.1)$$

where μ is the vector of empirical mean of matrix X over each row. The eigenvectors of C (denoted by E^c where c is the emotion class) are then calculated and arranged by decreasing eigenvalues. Only the 20 largest eigenvectors are chosen for each class of facial emotions. The principal eigenfaces are simply the eigenvectors of the system that have the largest eigenvalues. We compute the features F^c of class c as shown below.

$$E^c = PCA(X^c) \quad (2.2)$$

$$F^c = E^c(X^c - \mu^c) \quad (2.3)$$

In order to classify the target image preprocessing is necessary to preprocess the image as we preprocess the training dataset, which we will denote y . The classification of a test face is performed by comparing the distance of the features of the target face (projected onto the emotion subspace) to the features of the eigenfaces of the subspace. We then choose the class that minimizes this function as the predicted class, specifically

$$\arg \min_c \sum_i \| E_i^c(y - \mu) - F_i^c \|, \quad (2.4)$$

where i is each individual feature column vector in the array [55].

Given the distance value we are able to set a threshold value in order to filter out results that are weakly classified. Figure 2.7 shows examples of classified facial emotions.

2.2.3 Experiments

Image Sentiment Classification: As mentioned before, state-of-the-art sentiment analysis approach can be mainly concluded as: (1) textual information-based

⁶ <http://www.mathworks.com/matlabcentral/fileexchange/20976>.



Fig. 2.7 Examples of eigenface-based emotion detection. **a** Classification: Positive. **b** Classification: Negative

sentiment analysis, as well as online sentiment dictionary [7, 8] and (2) sentiment analysis based on low-level features. Therefore, in this section, we set three base-lines: (1) low-level feature-based approach, (2) textual content-based approach [8], and (3) online sentiment dictionary SentiStrength [7].

1. Image Sentiment Classification Performance:

First we demonstrate results of our proposed algorithm, image sentiment classification based on 102 mid-level attributes (SD vs. HD). Both Linear SVM and Logistic Regression algorithms are employed for comparison.

As demonstrated in Table 2.2, performance of precision for both Linear SVM and Logistic Regression outperforms that of recall. Due to the benefits of using asymmetric bagging, we are now able to raise the classification accuracy of negative samples. Smaller number of false positive samples and relatively larger number of detected true positive samples contribute to this unbalanced value of precision and recall performance.

The next thing we are interested in is the comparison against baseline algorithms.

2. Low-level Feature-Based and Textual Content-Based Baselines:

For low-level feature-based algorithm, Ji et al. employed the following visual features: a dimensional color histogram extracted from the RGB color space, a 512-dimensional GIST descriptor [18], a 53-dimensional local binary pattern

Table 2.2 Image sentiment classification performance

		Precision (%)	Recall (%)	Accuracy (%)
Linear SVM	SD	82.6	56.8	55.2
	HD	86.7	59.1	61.4
Logistic regr	SD	84.3	54.7	54.8
	HD	88.1	58.8	61.2

Table 2.3 Accuracy of sentiment classification

(a) Comparison between low-level-based algorithm and mid-level-based algorithm			
	SVM (low)	Logistic regr (low)	SVM (mid)
AC	50 %	53 %	61.4 %

(b) Comparison between mid-level visual content-based algorithm and textual content-based algorithm			
	Contextual polarity	Sentistrength	SVM (mid)
AC	61.7 %	61 %	61.4 %

(LBP), a bag-of-words quantized descriptor using a 1,000 word dictionary with a two-layer spatial pyramid, and a 2,659-dimensional Classemes descriptor. Both linear SVM and logistic regression algorithms are used for classification. For textual content-based algorithm, we choose contextual polarity, a phrase-level sentiment analysis system [6], as well as SentiStrength API.⁷ Table 2.3 shows the results of accuracy based on low-level features, mid-level attributes, and textual contents.

Decision Fusion: The final step of Stribute is decision fusion. By applying eigenface-based emotion detection, we are able to improve the performance of our decision based on mid-level attributes only. We only take into account images with complete face with reasonable lighting condition. Therefore among all the images with faces, we first employ a face detection process and generate a set of 153 images as the testing dataset for facial emotion detection and decision fusion. For each face we detected, we assigned them a label indicating sentiments: 1 for positive, 0 for neutral, and −1 for negative sentiments. We thus computed a sentiment score for each image as a whole. For instance, if we detect three faces from an image, two of them are detected as positive and one of them is detected as neutral, then the overall facial sentiment score of this image is 2. These sentiment scores can be used for decision fusion with the decision made based on mid-level attributes only, i.e., we add up the facial sentiment score and the outputs of the classifiers based on mid-level attributes only returned by our classifiers to implement a decision fusion mechanism. Table 2.4 shows the improvements in accuracy after decision fusion.

⁷ <http://sentistrength.wlv.ac.uk/>.

Table 2.4 Accuracy of Sentribute algorithm

	Accuracy (%)
Mid-level-based prediction	64.71
Facial emotion detection	73.86
Sentribute (after synthesis)	82.35

Figure 2.8 presents examples of true positive (TP), false positive (FP), true negative (TN), and false negative (FN) samples generated by Sentribute. False classified samples show that it is hard to distinguish images only containing texts from both positive and negative labels, and images of big event/celebration (football game or a concert) from those of protest demonstration. They both share similar general scene descriptors, similar lighting condition, and similar color tone. Another interesting false detected sample is the first image shown in false negative samples. Figures make frown expression on their faces, however the sentiment behind this expression is positive since they were meant to be funny. This sample is initially classified as positive based on mid-level attributes only, and then refined as negative because two strong negative facial expressions are detected by our eigenface expression detector. These kinds of images show a better decision fusion metric would be one of our potential improvements.

2.2.4 Conclusion

In this section, we present Sentribute, a novel image sentiment analysis algorithm based on mid-level attributes. Asymmetric bagging approach is employed to deal with unbalanced training data. To enhance the classification performance, eigenface-based emotion detection algorithm is applied, to deal with images containing faces and achieve a significant gain in accuracy over results based on mid-level attributes alone. The proposed algorithm explicitly explores visual content for sentiment analysis by employing mid-level attributes and without using textual content.

2.3 Sentiment Analysis in Multimedia Tweets

Online social networks have attracted the attention of people from both the academia and real-world. In particular, the rich multimedia information accumulated in recent years provides an easy and convenient way for more active communication between people. This offers an opportunity to research people’s behaviors and activities based on those multimedia content that can be considered as social imagematics. One emerging area is driven by the fact that these massive multimedia data contain people’s daily sentiments and opinions. However, existing sentiment analysis typically only pays attention to the textual information regardless of the visual content,

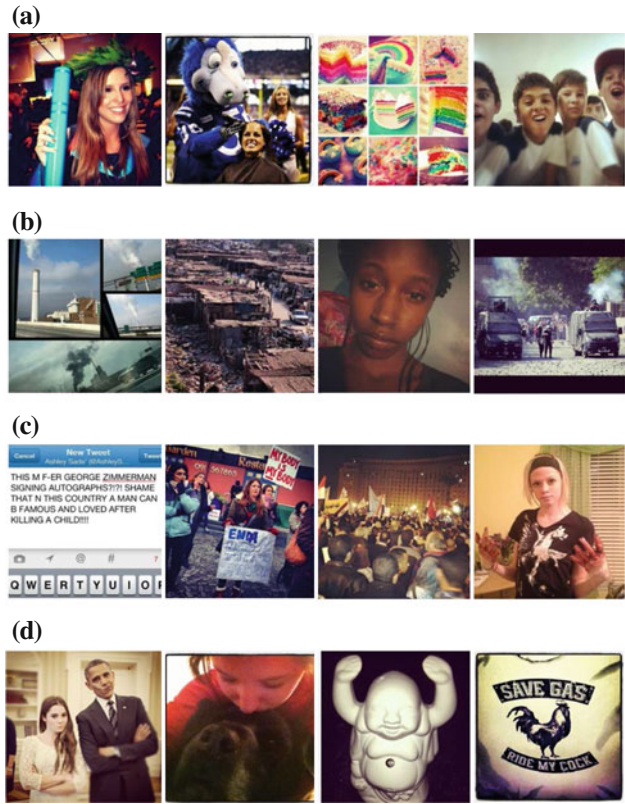


Fig. 2.8 Examples of sentiment detection results by SentiSuite. **a** True positive samples. **b** True negative samples. **c** False positive samples. **d** False negative samples

which may be more informative in expressing people’s sentiments and opinions. In this section, we attempt to analyze the online sentiment changes of social media users using both the textual and visual content. In particular, we analyze the sentiment changes of Twitter users using both textual and visual features. An empirical study of real Twitter datasets indicates that the sentiments expressed in textual content and visual content are correlated. The preliminary results in this section give insight into the important role of visual content in online social media.

Twitter is one of the most influential social networks across the world. Research work of different topics related to Twitter has been published in different conference venues. The large amount of daily generated user content attracted many researchers around the world to analyze potential interesting patterns in social media, including prediction of political election, sentiment analysis, information diffusion, topic trend, etc. However, it should be noted that at the beginning, Twitter as a social platform only allows a maximum of 140 characters to compose users’ messages. However, things changed in 2011, when Twitter allowed online users to post images in their



Fig. 2.9 Example of an image tweet, where the *left* image shows a picture of justin Bieber and the *right* image shows the ejection of Noah during the NBA playoffs

tweets. We denote the tweets that contain images as image tweets. The impacts of image tweets are tremendous. This part will focus on one particular impact of image tweets, namely the impact on sentiment analysis.

Multimedia content, like images, are more likely to express and convey people's subtle feelings compared with text information [3]. With the popularity of smart-phones and convenient social media APPs, more and more people are likely to post image tweets to attract attention from other users in Twitter. Figure 2.9 shows an example of an image tweet, where the big picture conveys more information about the Tweet.

One of the most interesting aspects of Twitter is that people's sentiments in Twitter seem to be related to real social life. For instance, in [57], the authors found that the sentiment changes of Twitter users are closely related to the overall economy situations in the U.S. and the stock market. However, most research on sentiment changes are related to the overall text tweets. Little attention has been paid to the analysis of image tweets. The work described in this part is an attempt toward the analysis of sentiment conveyed in the multimedia content in tweets. We intend to investigate social multimedia analysis, which we refer to as social imagematics. We conduct an empirical study on the sentiments expressed in people's tweets, especially the impact of sentiments in image tweets.

2.3.1 Approaches

As discussed in Sect. 2.1, there are many existing works on sentiment analysis using textual features. In this section, we employ existing algorithms to analyze the sentiment of the textual tweets. For the sentiment analysis of visual features, we build classifiers using low-level and mid-level respectively.

Textual Sentiment Analysis: There are many related works on sentiment analysis of Twitter [26, 29, 33, 36]. Meanwhile, there are also many online services that provide easy access API to evaluate the sentiment of online tweets. Many of these tools⁸ come directly from the academic research. Since we are more concerned with image tweets and the sentiment of images, we directly use existing online service for the sentiment analysis of collected tweets.

In particular, we use the sentiment140⁹ [58]. Sentiment140 is a semi-supervised machine learning approach. It exploits emotions as noisy labels for training data. Moreover, it provides convenient API for the sentiment analysis of different tweets. Typically, one can send the data to the server using HTTP request. The server then returns the sentiment for each line contained in that file. The returned value in this file contains three different values (0, 2, and 4). Here 0 represents the negative sentiment, 4 represents the positive sentiment, and 2 means neutral. In this way, we are able to classify the tweets into different sentiment categories.

Sentiment Changes with the Number of Images: Users in Twitter generally preferred different types of tweets. Some of the users like to post many image tweets, while many other users love to post traditional text tweets. To analyze the sentiments of users with different preferences over image tweets, we conduct an experiment on the relation between the proportion of image tweets and the proportion of positive tweets. We use the textual sentiment analysis in Sect. 2.3.1 to analyze the sentiments of different users. Then, the number of positive tweets over the sum of positive and negative tweets is used to represent the proportion of positive sentiment.

We randomly picked about 300 users and downloaded their tweets using the user timeline API. Figure 2.10 shows that users who like to post many image tweets are more likely to have positive sentiments. On the other hand, for users with fewer proportion of image tweets, the proportion of positive sentiments among these users varies significantly.

Visual Sentiment Classification: Image sentiment analysis is quite challenging. As discussed in [59], the authors used the textual sentiment analysis as the rough labels of the corresponding images. Then, RGB Hist and SIFT features are employed to train a classifier and classify the test images. Their results indicate that the positive and negative sentiments seem to share different interesting image patterns.

In our implementation, we use the image sentiment corpora from visual sentiment ontology¹⁰ with kind permission from the authors. Then according to the dataset, we trained two levels of classifiers. The first classifier only uses the low-level features, which include HOG [60], GIST [18], SSIM [61], and GEO-COLOR-HIST [62]. Different features have different advantages over different tasks [53]. HOG is good for object and human recognition. GIST is another feature designed for scene recognition. On the other hand, SSIM provides measure of invariant scene layout. Meanwhile, geometric color histogram offers a robust histogram feature, which is invariant of scene layout.

⁸ <http://matei.org/ithink/2012/02/08/a-list-of-twitter-sentiment-analysis-tools/>.

⁹ <http://www.sentiment140.com/>.

¹⁰ <http://visual-sentiment-ontology.appspot.com/>.

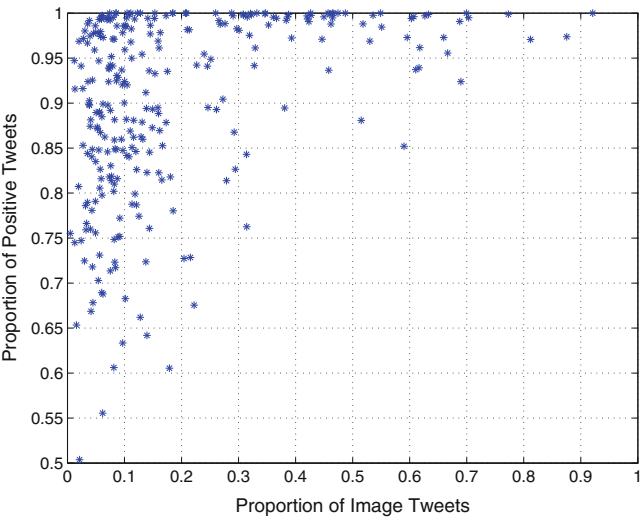


Fig. 2.10 Relationship of proportion of image tweets and the proportion of positive tweets

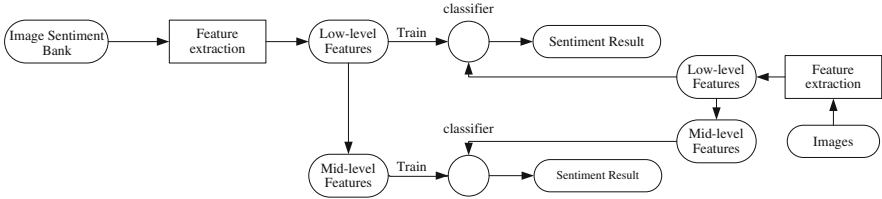


Fig. 2.11 Framework of image sentiment classification using low-level and middle level features respectively

The low-level features can be easily extracted from the given images. Figure 2.11 shows the framework employed for image sentiment classification. The main component in this framework is the low-level and middle level image features. Accordingly, there are two classifiers. In our implementation, we choose liblinear¹¹ as the classifier for both levels due to its scalability in large-scale learning. The first classifier is based on the low-level features discussed above. Based on these low-level features, we also train and learn some middle level features. Middle level features are more interpretable than low-level features. In our implementation, we use the middle level features described in Table 2.5. For each middle level feature, we need to train a classifier, which can determine whether or not the given image contains the corresponding middle level description. By combining all the middle level features, we are able to construct a middle level features description for the given image set. Then, a second-level classifier based on the extracted middle level features is constructed and employed to classify the test images into different sentiment categories.

¹¹ <http://www.csie.ntu.edu.tw/~cjlin/liblinear/>.

Table 2.5 Summary of the middle level features used in this study

Dirt/soil	Matte	Man-made	Rugged scene
Natural light	Dirty	Open area	Cluttered space
Direct sun/sunny	Rusty	Semi-enclosed area	Scary
Electric/indoor lighting	Arm	Enclosed area	Soothing
Aged/orn	Cold	Far-away horizon	Stressful
Glossy	Natural	No horizon	

For all the images contained in image tweets, we then download these images according to the URL contained in the metadata of each image tweet. Then low-level and middle level features are extracted using the same procedure for the training images. In this way, we are able to classify the sentiment of image tweets according to the visual features of the images contained in image tweets.

2.3.2 Experiments

We collect tweets using online Twitter API.¹² Twitter provides different categories of API. We mainly use the Twitter streaming API and Twitter timeline API. In order to choose some relatively active users, we use the streaming API to download over 19 million tweets. Active users simply refer to users who tweet, reply, and retweet more than others over a certain period of time. We chose empirical thresholds (over 100 original tweets in 1 month) to determine the relatively active users. To store such a large amount of data, we use couchdb,¹³ a document database, to store the download tweets. Then, by analyzing the downloaded 19 million tweets, we are able to identify the activity levels of different online users. First we identify over 8,000 users, and we use the timeline API to download the tweets of these users. We collected over 20 million tweets for all the 8,000 users. Next, the tweets of these 8,000 users are further analyzed. Among these 8,000 Twitter users, we further pick out about 300 users who are relatively active in posting both text and image tweets based on the threshold we mentioned because we want to analyze the correlation between sentiments behind text tweets and image tweets. Given these users and the URL contained in their image tweets, we collect all the users' posted images. We got over 90,000 thousand images for these active Twitter users.

In the downloaded 25 million tweets, we analyze the proportion of image tweets. Over the 25 million tweets, about 6 million tweets are image tweets ($5,988,058/25,580,000 = 0.23$). About every 1 in 4 tweets contains images in Twitter. Figure 2.12 shows the distribution of number of retweets. Similar to many other user activities, the distribution is a power law distribution with long tail. Figure 2.12a, b shows that the number of image retweets share a similar distribution,

¹² <https://dev.twitter.com/>.

¹³ <http://couchdb.apache.org/>.

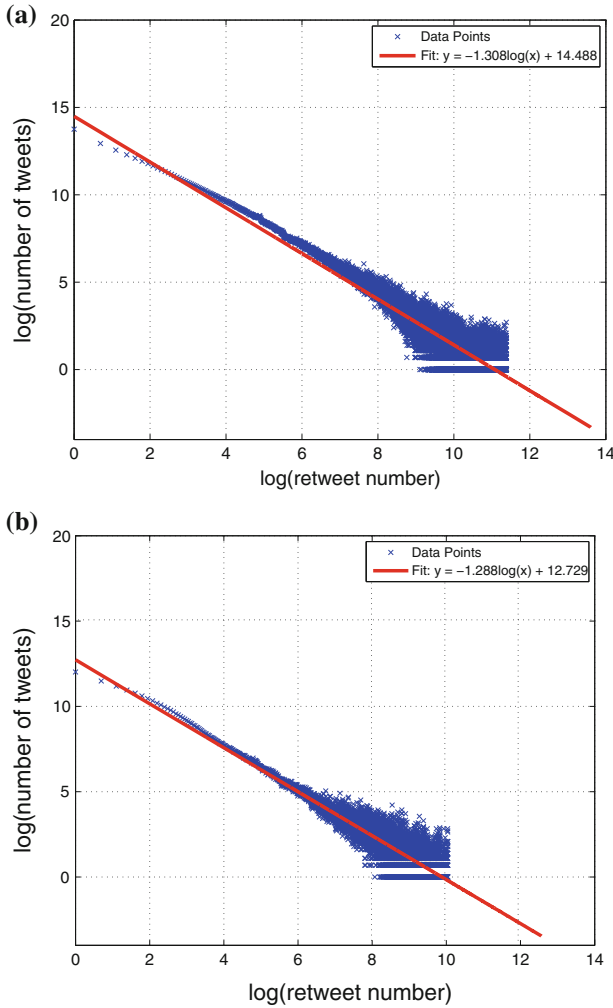


Fig. 2.12 Statistics of retweets number for all tweets and image tweets only. **a** Distribution of number of retweets. **b** Distribution of number of image retweets. **c** Probability distribution of *top* retweets number between 1 and 100. **d** Probability distribution of *top* retweets number between 1 and 1,000

with a slight difference in the slope of the fitted line of the log-log plot of the distribution. If we further look at the cumulative probability distribution of retweets number for all tweets and image tweets only, we can conclude from Fig. 2.12c, d that compared to image tweets, the proportion of tweets that received small number of retweets takes a larger proportion than image tweets. This evidence also verifies the fact that image tweets are more likely to attract online users' attention and are more easily diffused in the social network.

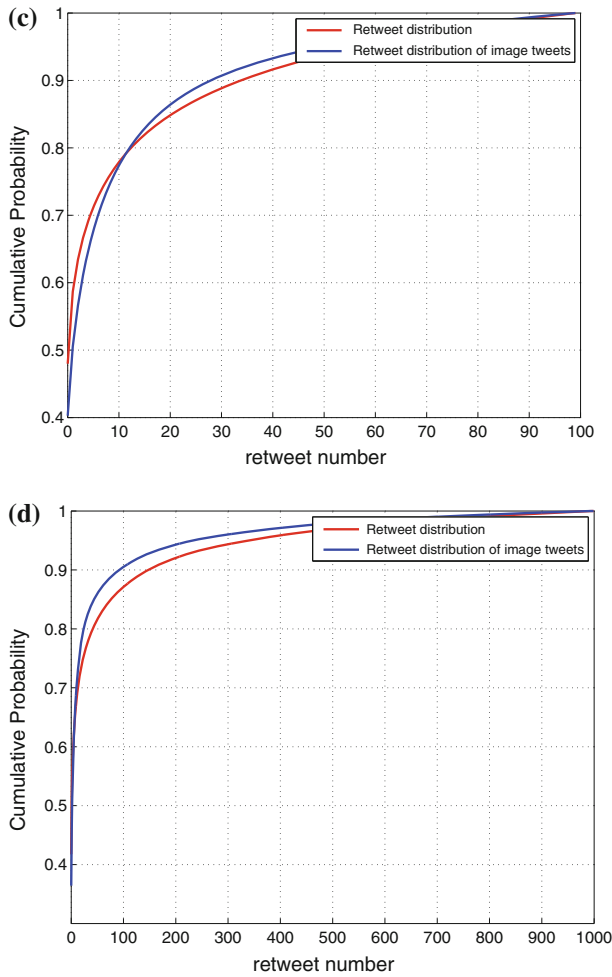


Fig. 2.12 (continued)

Correlation of Sentiment Between Image Tweets and Text Tweets: To illustrate the correlation between text and image tweets, we randomly select 10 users from the 300 users. We employ the methods discussed in Sect. 2.3.1. The sentiment analysis results using text and image features are shown in Figs. 2.13 and 2.14. In both figures, the red line represents the sentiment changes of each user according to the sentiment analysis of using text tweets, while the blue line represents the sentiment changes of each user according to the sentiment analysis of image tweets. The blue lines in the left column give the sentiment analysis using low-level image features, while the blue lines in the right column give the sentiment analysis using middle level image features. In Fig. 2.13, we average the long-term sentiment for each user in terms of

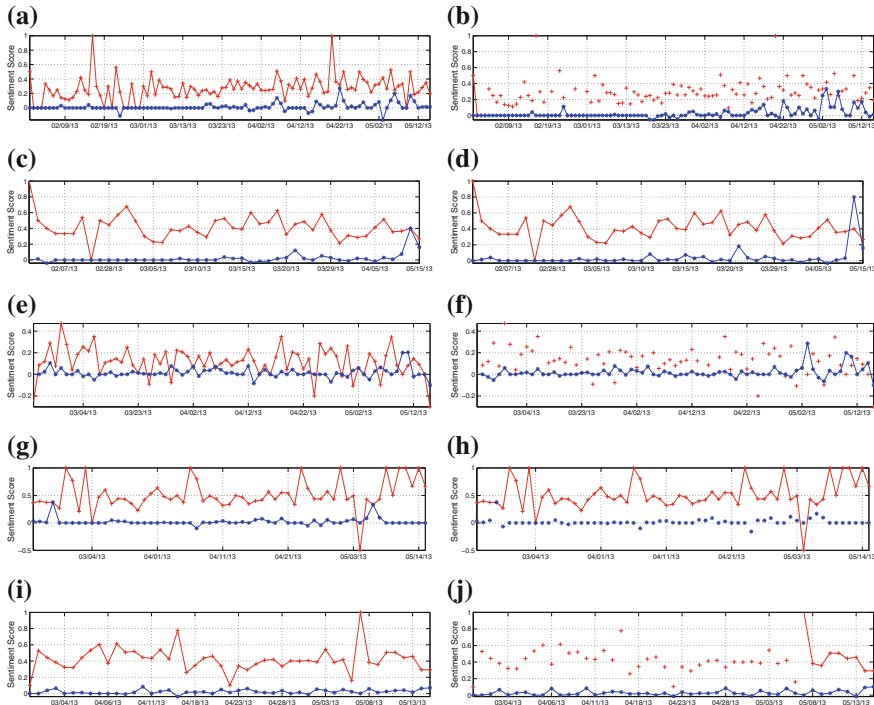


Fig. 2.13 Long-term sentiment changes of tweets and image tweets using low-level and mid-level features. The *red line* represents the sentiment of each user using the textual features and the *blue line* represents the sentiment of each user using the visual features from the image tweets. **a** 110914277. **b** 110914277. **c** 1135866961. **d** 1135866961. **e** 183352499. **f** 183352499. **g** 320657019. **h** 320657019. **i** 341587111. **j** 341587111. **k** 606333611. **l** 606333611. **m** 745235832. **n** 745235832. **o** 910880371. **p** 910880371. **q** 924674300. **r** 924674300. **s** 98005782. **t** 98005782

days, which means that each point represents the average sentiment score for a user. Similarly, in Fig. 2.14, the sentiment is averaged in terms of 1 h.

Table 2.6 shows the correlation coefficients between sentiment of the selected users using text features and image features. Although there is noise in the prediction of user's sentiment, the results indicate that there is still positive correlation between the sentiment expressed in text tweets and image tweets. In particular, for user 606333611, the sentiments are highly correlated. The reasons for this may include two aspects. First, we see this user is a relatively more active user. This can be reflected by the date in the x-axis of the figure. Since Twitter only allows us to download up to 3,200 of a user's most recent statuses, therefore, this user posted many tweets in a relatively short period. Second, there is no negative sentiment predicted by the text tweets. At the same time, for some users, they only have positive sentiment (there is no negative and neutral sentiment), thus the correlation is unavailable.

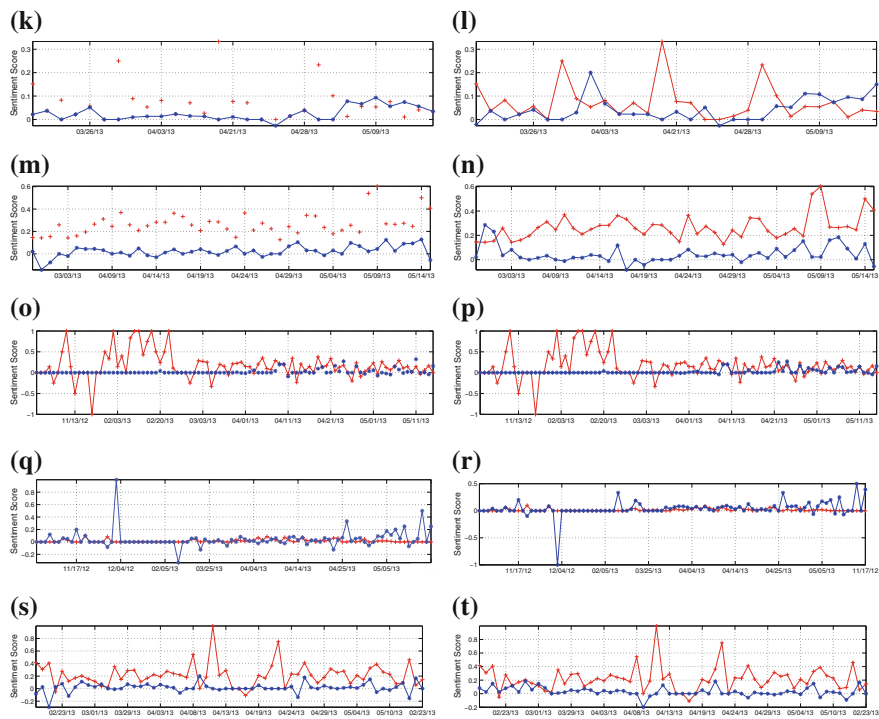


Fig. 2.13 (continued)

However, overall we see that sentiment classification using middle level features seems to be more correlated with the sentiment of using text tweets.

Correlation of Sentiment in a Shorter Period: The above results are averaged in terms of a day. This may not reflect people’s sentiment fluctuation in a particular day. In this section, we average the short-term sentiment of a user in terms of an hour. The results are shown in Fig. 2.14. The results indicate that different users have different sentiment change patterns. Some users are more likely to have emotional fluctuation in terms of both text and image tweets. For some users, their sentiment changes are reflected by text tweets. Meanwhile, some users are more likely to post images to express their sentiment changes. There is a correlation between the sentiment changes for the randomly selected 10 users. Table 2.7 shows the correlation coefficients for the 40 most recent periods. Different from the results in terms of days, in this case some of the correlation coefficients are negative. However, for most users, the correlation coefficients are mostly positive. The results of using low-level visual features and middle level visual features are not consistent all the time. The results on one hand indicate the difficulty in image sentiment analysis. On the other hand, they also illustrate the different patterns of online users in expressing their sentiment.

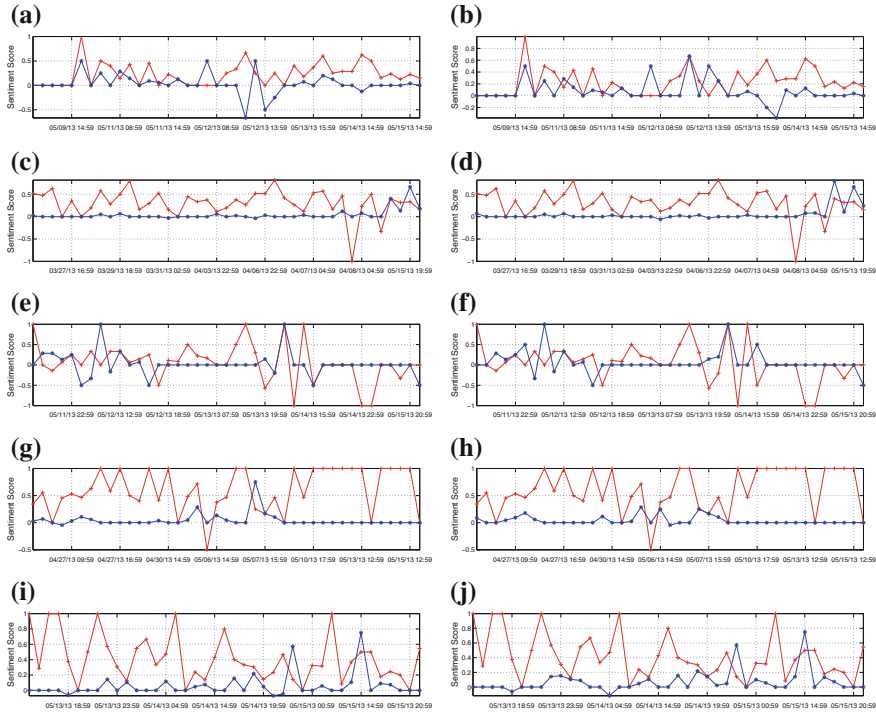


Fig. 2.14 Short-term sentiment of the recent 40 periods. We choose 1 h in which the users posted tweets as one short period. **a** 110914277. **b** 110914277. **c** 1135866961. **d** 1135866961. **e** 183352499. **f** 183352499. **g** 320657019. **h** 320657019. **i** 341587111. **j** 341587111. **k** 606333611. **l** 606333611. **m** 745235832. **n** 745235832. **o** 910880371. **p** 910880371. **q** 924674300. **r** 924674300. **s** 98005782. **t** 98005782

2.3.3 Conclusion

The results in this section are based on preliminary work. Some users are more likely to express their sentiments using image tweets, while others are still more likely to express their sentiment using text tweets. This reveals the challenges in predicting the sentiment of online social network users. The results in this section are encouraging for using the multimedia information for sentiment analysis.

Nevertheless, sentiment analysis is quite challenging for social multimedia. The short text nature of tweets imposes more challenges on this task. The results in this study indicate that both the textual and visual features are informative in determining one's sentiment. We discover the correlation between the sentiment expressed by text tweets and image tweets. At the same time, different users also reveal different behavior patterns in online social networks. Although the results do indicate some correlation between image tweets and textual tweets, to get more robust and more interpretable results, we need more features and more robust data to discover the

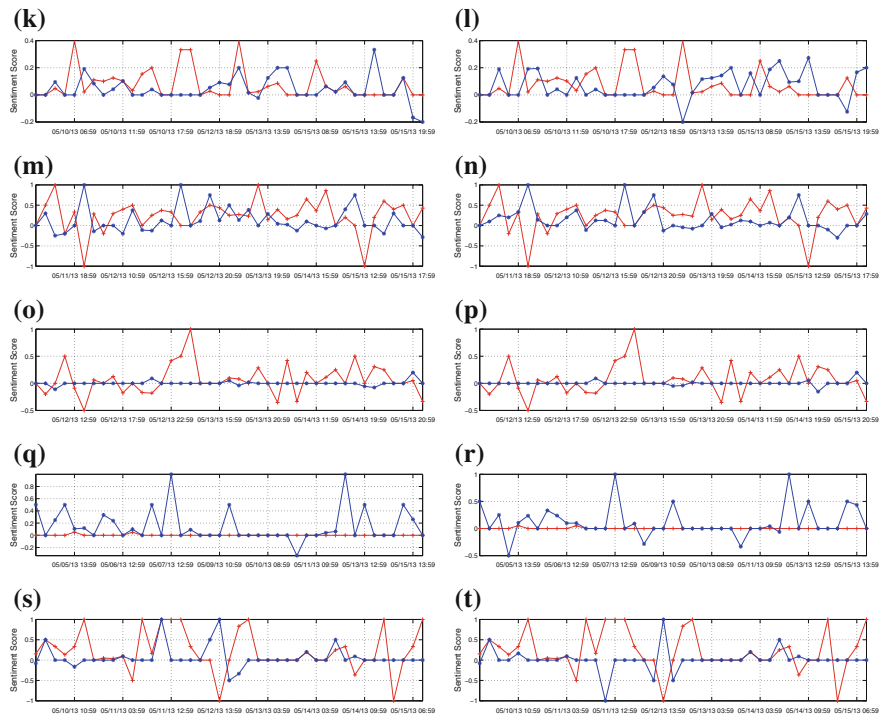


Fig. 2.14 (continued)

Table 2.6 Correlation coefficients of textual sentiment and visual sentiment (NA means not available)

User id	Low-level features	Mid-level features
0110914277	0.132197	0.137298
1135866961	0.059131	0.108657
0183352499	0.038009	0.095219
0320657019	0.105444	0.084368
0341587111	NA	NA
0606333611	0.618337	0.322811
0745235832	NA	NA
0910880371	0.199853	0.023016
0924674300	0.297284	0.317496
0098005782	0.015088	0.166366

influence of multimedia content in the social network. The sentiment analyses of images are still not mature. This, on the other hand, indicates that we have a great opportunity for discovery in this area.

Table 2.7 Correlation coefficients of textual sentiment and visual sentiment for recent 40 periods

User id	Low-level features	Mid-level features
0110914277	0.176150	0.132065
1135866961	0.172788	0.172788
0183352499	0.075004	−0.197358
0320657019	0.226449	0.212064
0341587111	0.150699	0.221518
0606333611	0.398337	0.065079
0745235832	0.089547	0.006048
0910880371	−0.071518	−0.244712
0924674300	0.245525	0.252585
0098005782	−0.127538	−0.027864

2.4 Discussion and Future Work

In this chapter, we have discussed some of the current works in the field of sentiment analysis and presented our new research results on image and multimedia sentiment analysis. We are living in an increasingly open society and individuals are now more and more willing to share feeling with others and listen to others’ opinions at the same time. Due to the enormous growth in social network platforms, sentiment analysis is receiving more attention. Although we now have more data sources at greater scales than ever before, sentiment analysis based on visual and multimodality perspective is still in its infancy. In the computer vision field, the development of attribute learning and deep neural network structures have shown some promising results, which can lead to sentiment analysis approaches such as *Sentribute*. Additionally, from a multimodality perspective, topics on deep multimodal structures are drawing more attention these days. For example, Srivastava showed in [63] that multimodal learning with deep boltzmann machines can improve the classification performance from the joint features extracted from both text and images. These techniques are expected to bring a new chapter to sentiment analysis and opinion mining.

Acknowledgments We would like to thank Prof. Shih-Fu Chang and his group for providing the test dataset and helpful discussions. This work is supported in part by the University of Rochester, and a generous gift from Adobe Systems.

References

1. Liu B (2012) Sentiment analysis and opinion mining. *Synth Lect Hum Lang Technol* 5(1):1–167
2. Nasukawa T Yi J (2003) Sentiment analysis: capturing favorability using natural language processing. In: *Proceedings of the 2nd international conference on knowledge capture*. ACM, pp 70–77

3. You Q, Luo J (2013) Towards social imagematics: sentiment analysis in social multimedia. In: Proceedings of the thirteenth international workshop on multimedia data mining, MDMKDD'13. ACM, New York, pp 3:1–3:8
4. O'onnor B, Balasubramanyan R, Routledge BR, Smith NA (2010) From tweets to polls: linking text sentiment to public opinion time series. In: Proceedings of the international AAAI conference on weblogs and social media, pp 122–129
5. Pang B, Lee L (2008) Opinion mining and sentiment analysis. *Found Trends Inf Retr* 2(1–2):1–135
6. Wilson T, Wiebe J, Hoffmann P (2005) Recognizing contextual polarity in phrase-level sentiment analysis. In: Proceedings of the conference on human language technology and empirical methods in natural language processing. Association for Computational Linguistics, pp 347–354
7. Esuli A, Sebastiani F (2006) Sentiwordnet: a publicly available lexical resource for opinion mining. In: Proceedings of LREC, vol 6, pp 417–422
8. Thelwall M, Buckley K, Paltoglou G, Cai D, Kappas A (2010) Sentiment strength detection in short informal text. *J Am Soc Inf Sci Technol* 61(12):2544–2558
9. Farhadi A, Endres I, Hoiem D, Forsyth D (2009) Describing objects by their attributes. In: IEEE conference on computer vision and pattern recognition (CVPR 2009), pp 1778–1785, IEEE
10. Naphade MR, Lin C-Y, Smith JR, Tseng B, Basu S (2002) Learning to annotate video databases. In: SPIE conference on storage and retrieval on media databases
11. Ordonez V, Kulkarni G, Berg TL (2011) Im2text: describing images using 1 million captioned photographs. In: Neural information processing systems (NIPS)
12. Snoek CG, Worring M (2008) Concept-based video retrieval. *Found Trends Inf Retr* 2(4):215–322
13. Datta R, Joshi D, Li J, Wang JZ (2006) Studying aesthetics in photographic images using a computational approach. In: Computer vision-ECCV 2006. Springer, pp 288–301
14. Marchesotti L, Perronnin F, Larlus D, Csurka G (2011) Assessing the aesthetic quality of photographs using generic image descriptors. In: 2011 IEEE international conference on computer vision (ICCV), pp 1784–1791
15. Isola P, Xiao J, Torralba A, Oliva A (2011) What makes an image memorable? In: 2011 IEEE conference on computer vision and pattern recognition (CVPR). IEEE, pp 145–152
16. Jia J, Wu S, Wang X, Hu P, Cai L, Tang J (2012) Can we understand van Gogh's mood?: learning to infer affects from images in social networks. In: Proceedings of the 20th ACM international conference on multimedia. ACM, pp 857–860
17. Machajdik J, Hanbury A (2010) Affective image classification using features inspired by psychology and art theory. In: Proceedings of the international conference on multimedia. ACM, pp 83–92
18. Oliva A, Torralba A (2001) Modeling the shape of the scene: a holistic representation of the spatial envelope. *Int j comput vis* 42(3):145–175
19. Yanulevskaya V, Uijlings J, Bruni E, Sartori A, Zamboni E, Bacci F, Melcher D, Sebe N (2012) In the eye of the beholder: employing statistical analysis and eye tracking for analyzing abstract paintings. In: Proceedings of the 20th ACM international conference on multimedia, MM'12. ACM, New York, pp 349–358
20. Bartlett MS, Littlewort G, Frank M, Lainscsek C, Fasel I, Movellan J (2005) Recognizing facial expression: machine learning and application to spontaneous behavior. In: IEEE computer society conference on Computer vision and pattern recognition, CVPR 2005, vol 2. IEEE, pp 568–573
21. Fasel B, Luetttin J (2003) Automatic facial expression analysis: a survey. *Pattern Recognit* 36(1):259–275
22. Wan S, Aggarwal J (2014) Spontaneous facial expression recognition: a robust metric learning approach. *Pattern Recognit*. 47(5):1859–1868
23. Hernandez J, Hoque ME, Drevo W, Picard RW (2012) Mood meter: counting smiles in the wild. In: Proceedings of the 2012 ACM conference on ubiquitous computing. ACM, pp 301–310

24. Hoque ME, Courgeon M, Martin J-C, Mutlu B, Picard RW (2013) Mach: my automated conversation coach. In: Proceedings of the 2013 ACM international joint conference on pervasive and ubiquitous computing. ACM, pp 697–706
25. Guerra PHC, Veloso A, Meira W Jr, Almeida V (2011) From bias to opinion: a transfer-learning approach to real-time sentiment analysis. In: Proceedings of the 17th ACM SIGKDD international conference on knowledge discovery and data mining. ACM, pp 150–158
26. Speriosu M, Sudan N, Upadhyay S, Baldrige J (2011) Twitter polarity classification with label propagation over lexical links and the follower graph. In: Proceedings of the first workshop on unsupervised learning in NLP. Association for Computational Linguistics, pp 53–63
27. De Choudhury M, Counts S, Horvitz E (2013) Major life changes and behavioral markers in social media: case of childbirth. In: Proceedings of the 2013 conference on computer supported cooperative work. ACM, pp 1431–1442
28. Silva IS, Gomide J, Veloso A, Meira W Jr, Ferreira R (2011) Effective sentiment stream analysis with self-augmenting training and demand-driven projection. In: Proceedings of the 34th international ACM SIGIR conference on research and development in information retrieval. ACM, pp 475–484
29. Hu X, Tang L, Tang J, Liu H (2013) Exploiting social relations for sentiment analysis in microblogging. In: Proceedings of the sixth ACM international conference on web search and data mining. ACM, pp 537–546
30. Kosinski M, Stillwell D, Graepel T (2013) Private traits and attributes are predictable from digital records of human behavior. In: Proceedings of the national academy of sciences
31. Rao D, Paul M, Fink C, Yarowsky D, Oates T, Coppersmith G (2011) Hierarchical Bayesian models for latent attribute detection in social media. In: Proceedings of the ICWSM, pp 598–601
32. Goel S, Hofman JM, Siner MI (2012) Who does what on the web: a large-scale study of browsing behavior. In: Proceedings of the 6th international AAAI conference on weblogs and social media
33. Wong FMF, Tan CW, Sen S, Chiang M (2013) Quantifying political leaning from tweets and retweets
34. De Choudhury M, Counts S, Gamon M (2012) Not all moods are created equal! Exploring human emotional states in social media. In: Sixth international AAAI conference on weblogs and social media
35. Miller M, Sathi C, Wiesenenthal D, Leskovec J, Potts C (2011) Sentiment flow through hyperlink networks. In: Proceedings of the AAAI CWSM
36. Tan C, Lee L, Tang J, Jiang L, Zhou M, Li P (2011) User-level sentiment analysis incorporating social networks. arXiv preprint [arXiv:1109.6018](https://arxiv.org/abs/1109.6018)
37. Jansen BJ, Zhang M, Sobel K, Chowdury A (2009) Twitter power: tweets as electronic word of mouth. *J Am soc inf sci technol* 60(11):2169–2188
38. Bollen J, Mao H, Zeng X (2011) Twitter mood predicts the stock market. *J Comput Sci* 2(1):1–8
39. Tumasjan A, Sprenger TO, Sandner PG, Welpel IM (2010) Predicting elections with twitter: what 140 characters reveal about political sentiment. In: Proceedings of the fourth international AAAI conference on weblogs and social media, pp 178–185
40. Yang CC, Ng TD (2007) Terrorism and crime related weblog social network: link, content analysis and information visualization. In: Intelligence and security informatics, New Brunswick, pp 55–58
41. Datta R, Joshi D, Li J, Wang JZ (2008) Image retrieval: ideas, influences, and trends of the new age. *ACM Comput Surv (CSUR)* 40(2):5
42. Zhang S, Yang M, Cour T, Yu K, Metaxas DN (2012) Query specific fusion for image retrieval. In: Computer vision-ECCV 2012. Springer, pp 660–673
43. Wang W, He Q (2008) A survey on emotional semantic image retrieval. In: 15th IEEE international conference on image processing (ICIP 2008), pp 117–120
44. Lang PJ, Bradley MM, Cuthbert BN (1999) International affective picture system (Iaps): Technical manual and affective ratings

45. Li B, Feng S, Xiong W, Hu W (2012) Scaring or pleasing: exploit emotional impact of an image. In: Proceedings of the 20th ACM international conference on multimedia. ACM, pp 1365–1366
46. Vonikakis V, Winkler S (2012) Emotion-based sequence of family photos. In: Proceedings of the 20th ACM international conference on multimedia, MM'12. ACM, New York, pp 1371–1372
47. Hanjalic A, Kofler C, Larson M (2012) Intent and its discontents: the user at the wheel of the online video search engine. In: Proceedings of the 20th ACM international conference on multimedia. ACM, pp 1239–1248
48. Borth D, Chen T, Ji R, Chang S-F (2013) Sentibank: large-scale ontology and classifiers for detecting sentiment and emotions in visual content. In: Proceedings of the 21st ACM international conference on multimedia. ACM, pp 459–460
49. Yuan J, McDonough S, You Q, Luo J (2013) Sentribute: image sentiment analysis from a mid-level perspective. In: Proceedings of the second international workshop on issues of sentiment discovery and opinion mining. ACM, p 10
50. Cambria E, Hussain A (2012) Sentic album: content-, concept-, and context-based online personal photo management system. *Cogn Comput* 4(4):477–496
51. Patterson G, Hays J (2012) Sun attribute database: discovering, annotating, and recognizing scene attributes. In: 2012 IEEE conference on computer vision and pattern recognition (CVPR), pp 2751–2758
52. Lundqvist D, Flykt A, Öhman A (1998) The karolinska directed emotional faces-KDEF. CD-ROM from department of clinical neuroscience, psychology section, Karolinska Institutet, Stockholm, Sweden. Technical report, ISBN 91-630-7164-9
53. Xiao J, Hays J, Ehinger KA, Oliva A, Torralba A (2010) Sun database: large-scale scene recognition from abbey to zoo. In: 2010 IEEE conference on computer vision and pattern recognition (CVPR), pp 3485–3492
54. Tao D, Tang X, Li X, Wu X (2006) Asymmetric bagging and random subspace for support vector machines-based relevance feedback in image retrieval. *Pattern Anal Mach Intell IEEE Trans* 28(7):1088–1099
55. Turk MA, Pentland AP (1991) Face recognition using eigenfaces. In: Proceedings of the IEEE computer society conference on computer vision and pattern recognition, CVPR'91. IEEE, pp 586–591
56. Ekman P (1992) An argument for basic emotions. *Cogn Emot* 6(3–4):169–200
57. Gilbert E, Karahalios K (2010) Widespread worry and the stock market. In: Proceedings of the international conference on weblogs and social media, vol 2, pp 229–247
58. Go A, Bhayani R, Huang L (2009) Twitter sentiment classification using distant supervision. Technical report, Stanford, pp 1–12
59. Siersdorfer S, Minack E, Deng F, Hare J (2010) Analyzing and predicting sentiment of images on the social web. In: Proceedings of the international conference on multimedia. ACM, pp 715–718
60. Dalal N, Triggs B (2005) Histograms of oriented gradients for human detection. In: IEEE computer society conference on computer vision and pattern recognition (CVPR 2005), vol 1. IEEE, pp 886–893
61. Shechtman E, Irani M (2007) Matching local self-similarities across images and videos. In: IEEE conference on computer vision and pattern recognition (CVPR'07), pp 1–8
62. Lalonde J-F, Hoiem D, Efros AA, Rother C, Winn J, Criminisi A (2007) Photo clip art. In: *ACM transactions on graphics (TOG)*, vol 26. ACM, p 3
63. Srivastava N, Salakhutdinov R (2012) Multimodal learning with deep Boltzmann machines. In: *NIPS*, pp 2231–2239

Multimedia Data Mining and Analytics

Disruptive Innovation

Baughman, A.; Gao, J.; Pan, J.-Y.; Petrushin, V.A. (Eds.)

2015, XIV, 454 p. 188 illus., 153 illus. in color.,

Hardcover

ISBN: 978-3-319-14997-4