ON THE SELECTION OF TRENDING IMAGE FROM THE WEB

Dongfei Yu, Xinmei Tian[*]

University of Science and Technology of China
ydf2010@mail.ustc.edu.cn, xinmei@ustc.edu.cn

Tao Mei, Yong Rui

Microsoft Research
{tmei, yongrui}@microsoft.com

ABSTRACT

The recommendation of trending images has become a popular feature used by commercial search engines to attract public attention. By browsing through trending images, search engine users can discover trending events at a glance. However, the selection of trending images is very challenging and remains an open issue. Most existing work is highly dependent on editorial efforts, though some preliminarily identify a few plain features for trending images. In this paper, we investigate a set of perceptual factors that can distinguish trending images from common ones. We propose a set of trending-aware features based on several common criteria, which reflect the characteristics of trending images. We further construct a manually labeled dataset based on a commercial search engine's query log over a two-week timespan. We evaluate our proposed method on this dataset and the results demonstrate its effectiveness.

***Index Terms*—** Trending image selection, image search, trending-aware features

1. INTRODUCTION

Nowadays, massive amounts of data regarding user searches, clicks, browsing activities from portal sites is being logged on back-end servers in real time. How to mine social attention, interest and trends from user logs are hot research topics. Trend recommendation is an effective and efficient way to satisfy one's curiosity about public focus. A trending image, the representative image for the trending event, is usually used for various purposes: acting as proof of a news event, catching readers' attention, filling up pages, or adding value by bringing a new dimension (aesthetic or informational) to text [1]. In a word, the usage of trending images provides a more convenient approach for capturing trending events.

Many commercial search engines, like Yahoo!, Bing, and Google, recommend trending content to users, as shown in Figure 1. On the Google Trends[1] site, the term "trending" is defined as, "search topics which have the largest increase in search or click volume since the previous period". To our best knowledge, most work on trending events detection from search engines is based on diverse statistics of search volume. Additionally, most of these services provide trending images for catching a user's eye quickly amidst noisy information, as well as enhancing users' vivid experience of trending events.

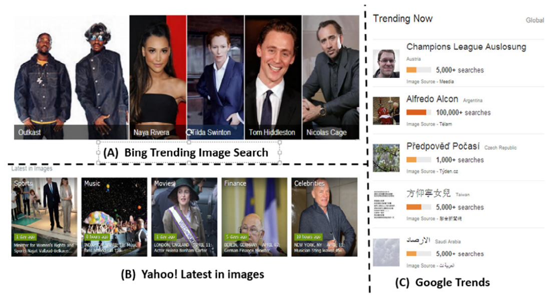

Fig. 1. Trending image recommendation from three commercial search engines. The screen shots were captured on 4/12/2014.

In trending image recommendation systems, there are two components: trending event detection and trending image selection. Automatic trending event detection has been studied for years and has already achieved great success, so our work focuses on the second component, *i.e.* trending image selection. According to our survey, however, trending images are usually selected manually in practical systems and are highly dependent on editorial efforts. Markkula *et al.* conducted a field study on news photo selection for journalists [2]. Stottinger *et al.* tried to translate journalists' requirements into low-level visual features for image search in a journalistic work context [3]. The recent work done by Wu *et al.* has the same objective on trending image selection [4]. Their framework focused on suggesting personalized trending search queries. Nevertheless, when coming to image selection for these suggested search queries, they used a very naive single-view method, and performed a simple user study to evaluate their trending image selection results.

Based on extensive observations and motivated by user study in [2], we conduct a comprehensive study on trending image suggestion and propose a novel method for automatic trending image selection. We first investigate several general criteria that trending images would follow, such as temporal modality on the click count of images, contextual cues on

[*]This work is supported by the NSFC under the contract No.61201413, the Specialized Research Fund for the Doctoral Program of Higher Education No. WJ2100060003, the Fundamental Research Funds for the Central Universities No. WK2100060007, No. WK2100060011, WK2100100021.

[1]http://www.google.com/trends/

web pages, *etc*. Based on these criteria, we propose a set of features to describe the characteristics of trending images, including popularity, burstness, representativeness, freshness, and aesthetic quality. Then, a supervised learning model is trained to combine all the aforementioned features. To evaluate the performance of our method, we conduct extensive experiments on a manually labeled web image dataset. The result demonstrates the effectiveness of our proposed method.

2. RELATED WORK

Trending image selection is a practical problem, which is novel, yet not well-defined in research communities. However, some similar problems, such as canonical image selection and event visualization, have been studied thoroughly across multimedia and computer vision fields. Therefore, we will also review those works here.

Trending Image Selection Significant work has been done to conduct field studies on journalists' and archivists' practices in newspaper editorial offices for pictorial Information Retrieval, thus providing a systematic guideline on news photo selection for journalists [1, 2]. Common selection criteria, such as topicality, technical and biographical criteria, impression to be conveyed, passport photos/formal portraits, and aesthetic criteria have been proposed by user studies [2]. Contextual factors (nature of articles, publishing section, space reserved for the image, and layout of the page) formed a selection frame for suitable images [1]. Following this work, Stottinger *et al*. tried to translate journalists' requirements into low-level visual features, such as *Color/B&W, Shooting Distance, Color and Composition*, and offered these features as image filters in the search system used by the journalists [3]. Wu *et al*. compared the distinctive power of two features, *i.e.* visual consistency, and burstness, to select the trending-aware representative images [4]. This work is closely related to ours; however, their method is too simplistic since only few separate features are leveraged to select trending images.

Canonical Image Selection Although prior work rarely contributed to automatic trending image selection, a lot of attention was indeed paid to canonical image selection for visual objects, such as landmarks [5], and products [6, 7]. Kennedy *et al*. used location and other metadata, as well as tags and visual features, to cluster images with the similar views of landmarks, and then leveraged some specific rules to rank clusters and representative images [5]. Berg *et al*. defined *an iconic image* for an object category as an image with a large object clearly separated from the background. They performed the object/background segmentation and extracted shape descriptors to analyze the object's similarity [8, 9]. Jing *et al*. constructed a similarity graph based on matching local visual features, in order to find single canonical images from a collection of commonly searched-for products [6]. Wang *et al*. defined a simple yet effective score function to measure the representativeness of an image, which was based on the similarity with nearest neighbors and textual relevance with a category [10]. The VisualRank algorithm was also applied to image re-ranking for visual consistency and authority, which was motivated by PageRank algorithm [7, 11]. It provides an effective approach to measure the representativeness of an image in an image collection. However, canonical image selection only measures visual similarity among image collections, and completely ignores contextual factors.

Event Visualization Event-based media analysis has recently drawn much attention due to the prevalence of social multimedia. The representation of events with multimedia data is one of the most attractive applications. Sahuguet *et al*. extracted time-sensitive events from Google Trends results using a simple statistical approach for a specific topic, then these events were illustrated on a timeline with videos mined from social media sharing platforms. They clustered the video sets based on textual features and then took the centroid of each cluster as representative data and ranked the items by mean similarity of each cluster in [12, 13]. Liu et at. generated the photo collage through the Google image search engine with the event title as a parameter, while filtering out these photos for which the cosine distance between its textual metadata and the event title were below a given threshold [14]. They retrieved images from Flickr with a geotag or title, and filtered the results using a time interval of five days. Then the irrelevant media was removed through visual analysis [15]. However, these work never considers the trending aspect of media, especially temporal modality. Moreover, powerful features in social media platform, such as time filters, geotags, and event titles are not available or directly accessible for most web images.

3. SELECTION OF TRENDING IMAGES

The framework of our proposed method for trending image suggestion is shown in Figure 2. We first identify the trending queries/events from a query log. Then for each detected trending query, we harvest a set of candidate trending images from the web. A set of trending-aware features are extracted for each candidate image in the image pool. Finally, a trending image classification model is trained and used for trending image selection.

3.1. Trending Query Detection

We detect trending queries from a query log using a method similar to [16]. We calculate the Buzz Score of a query as:

$$BuzzScore(q_j|\mathcal{Q}_d) = \sum_{s=d-1}^{d-n} \frac{1}{d-s}(P(q_j|\mathcal{Q}_d) - P(q_j|\mathcal{Q}_s)) \tag{1}$$

where n is the time window length and $\mathcal{Q}_d$ is the query set at day d. $P(q_j|\mathcal{Q}_d)$ is the likelihood (count) of the query q_j given the query set $\mathcal{Q}_d$ at day d [16].

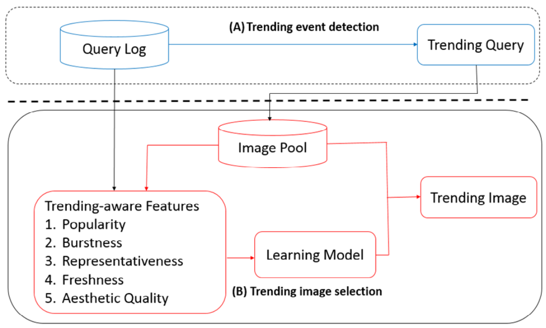

Fig. 2. The framework of our approach. The system consists of two components: (A) Trending event detection, (B) Trending image selection. The system can be shown in the front-page of an image search engine and updated automatically.

In order to cluster similar queries together, [16] proposed the generalized count to boost those queries with fewer terms. In [16], the modified Buzz Score is defined as:

$$BuzzScore'(q_j|\mathcal{Q}_d) = BuzzScore(q_j|\mathcal{Q}_d) \times log(1 + v(q_j, d) + v^*(q_j, d)) \quad (2)$$

where $v(q_j, d)$ is the count of q_j at day d. $v^*(q_j, d)$ is the generalized count of the query q_j at day d, which is defined as the total count of queries which fully contain query q_j at day d. Therefore, a trending query with short terms is easier to detect, and it represents a cluster of queries in the same topic.

All queries in day d are sorted according to their $BuzzScore'$ in descending order. The top-ranked queries are treated as trending queries/events in day d.

3.2. Criteria for Trending Image Selection

We refer to the representative and high-quality images of trending events which attract public attention on the web as **trending images**. Before we explain how to derive a set of distinctive features to describe the properties of a trending image, we first investigate what common criteria trending images should follow.

A. Query-dependent relevance

According to our definition above, a trending image must be relevant to a trending event. Yet, image understanding is still an open issue. It is difficult to mine semantic events from a single image directly. Fortunately, images are not isolated in web pages in most cases; they are surrounded by various textual cues, such as titles, tags, and descriptions. Therefore, these contextual cues can be used to find images associated with trending events.

In this work, we submit trending queries into an image search engine and collect the top-ranked images as the candidate trending image set. Thus, relevance can be guaranteed for topicality as proposed in [2].

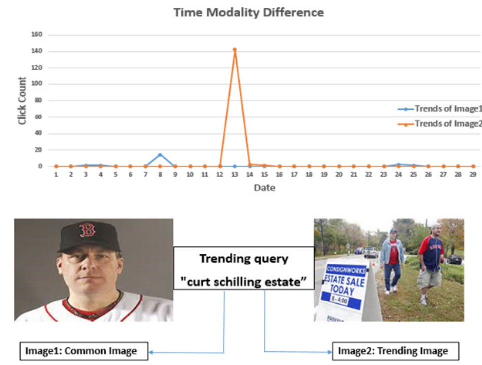

Fig. 3. Temporal modality comparison between trending and common images. The trending query "curt schilling estate" was detected due to the breaking news "Former Boston red sox pitcher Curt Schilling holds estate sale at former Massachusetts home on Oct 12th".

B. Time-dependent modality

Specific temporal modality is an important signature of trending images. There are many ways to characterize the temporal modality of an image, for example, how an image's click-count varies with time. One example is illustrated in Figure 3. In order to distinguish trending images from common ones by click-through data, we analyze the temporal modality of the images in the candidate set. In this paper, three significant indicators are discussed as follows:

Popularity Obviously, breaking events usually attract more attention than common events. Therefore, there will be larger search and click volume for trending images during the same period.

Burstness According to News Communication Theory, a trending event usually evolves in four stages: formation, break-out, distribution, and fading out. When a trending event breaks, search engines will capture significant bursts in search and click volume about the event. Burstness is an important signal to recognize a trending image.

Recency Few people can predict when a trending event will happen, so breaking news is usually time sensitive. Editors intensively edit, illustrate, post, and reproduce their news reports to trace the trending event. Therefore, the time stamp of an image also plays an important role in trending image selection.

C. Content-dependent quality

Web images have large variations in quality since they may be taken by professional and amateur photographers from different points of view using different devices. Trending images are selected for the purpose of filling up the landing page, so authority and high-quality are essential attributes.

Visual authority When a trending event breaks, various news media outlets will follow the trend from different points of view (see Figure 4). Consequently, many

Fig. 4. The visual authority of trending images (with red borders) among the image set. The trending query "Conan O'Brien MTV" is detected because he will host the 2014 MTV Movie Awards. The images with red borders were captured during the live show.

near-duplicate news photos on the trending event appear on the web. [7] has proved that an image is more authoritative when many similar photos are available.

Aesthetic quality According to [2], aesthetic criteria plays an important role in the final stage of selecting news photos for journalists. At this time, a group of candidate photos has been retrieved, so an image with higher quality will be more likely to be selected among them.

3.3. Trending-Aware Features

We design a set of trending-aware features to quantitatively represent the criteria discussed above.

3.3.1. Popularity

To describe the popularity of an image I_k associated with the trending query q_j detected at day d, we count <trending query, image URL> pairs within a time window (day d and the past n days). The larger the click volume, the more popular the image tends to be. The popularity $P_p(I_k|q_j)$ is determined by the likelihood of I_k given q_j at day d as follows:

$$P_p(I_k|q_j, d) = \frac{\sum_{t=d}^{d-n} c_{jk}(t)}{\sum_{I_k} \sum_{t=d}^{d-n} c_{jk}(t)} \tag{3}$$

where $c_{jk}(t)$ is click count of I_k given the query q_j at day t. $\sum_{I_k} \sum_{t=d}^{d-n} c_{jk}(t)$ is the total click count of all images during days $(d-n, d)$ to query q_j.

3.3.2. Burstness

Buzz Score has been used to detect trending queries in a query log previously. This concept can also be adopted to measure the burstness of an image by calculating the increasing rate of an image's click volume over time. URL Buzz Score of an image I_k is defined as:

$$\begin{aligned} &BuzzScore(I_k|q_j) \\ &= \sum_{s=d-1}^{d-n} \frac{1}{d-s}(P_p(I_k|q_j, d) - P_p(I_k|q_j, s)) \end{aligned} \tag{4}$$

where $P_p(I_k|q_j, d)$ is the likelihood of image I_k given query q_j at day d.

3.3.3. Freshness

Trending images associated with a trending event are usually uploaded right after the event breaks out. An old photo that has existed on the Internet for a long time is rarely related to a current event. However, the image search engine we used in our work doesn't provide available API for image filtering with customized time. To overcome this problem, we measure the freshness of an image by calculating the time gap between the moment an image first appeared on the web and the one the trending event began. We define the freshness of an image as:

$$P_f(I_k|q_j, d) = d - T(I_k) \tag{5}$$

where the trending query q_j is detected at day d, and $T(I_k)$ is the time stamp when the image I_k is retrieved by the image search engine.

3.3.4. Representativeness

Representativeness is proposed to estimate visual authority among the candidate image set. The Bag-of-Visual-Words model with the Scale Invariant Feature Transformation (SIFT) is applied to represent the visual feature of an image [17]. Then, we use a cosine metric to measure the visual similarity of any pair of images. Finally, a random walk process is performed on a query-dependent graph [7], whose nodes represent images associated with the trending query and edges are weighted by the similarity of two images. The edge weight $a_{k,l}$ is defined as:

$$a_{k,l} = \frac{sim(I_k, I_l)}{\sum_{I_s \in KNN(I_k)} sim(I_k, I_s)} \tag{6}$$

Therefore, the visual authority (representativeness) $P_r(I_k|q_j)$ of image I_k can be formulated as:

$$\mathbf{P_r} = \alpha A^T \mathbf{P_r} + (1-\alpha)\frac{1}{m} \tag{7}$$

where $\mathbf{P_r}$ is the visual authority vector of the image set associated with the trending query q_j, A is the transition matrix consisting of $a_{k,l}$, and m is the size of the vector $\mathbf{P_r}$.

3.3.5. Aesthetic quality

Aesthetic quality assessment is a popular research topic in computer vision. As a semantic attribute of images, aesthetic quality measures how visually appealing an image is from the view of human beings. Great efforts have been devoted to researching various aesthetic features and learning models for automatic photo quality assessment [18, 19, 20]. Most work refers to the experience and skills of professional photographers, such as the rule of thirds and visual weight balance, and proposes some common criteria: simplicity, realism, and basic photographic technique. In this paper, we follow the recent work [21] to predict an aesthetic quality score for each image.

3.4. Learning Model

In this paper, we formulate the trending image selection into a binary classification problem. Each image in the candidate trending image set can be represented by the features introduced in Section 3.3. To combine all these features and enhance the predictive power, we train various learning models including the Linear-kernel SVM, RBF-kernel SVM, Random Forest, Boosting, and Neural Networks. Finally, we predict a decision value for each image and rank the image set associated with the same trending event.

4. EXPERIMENTS

4.1. Data Set

There is little work on automatic trending image selection, and there is no publicly available dataset. Therefore, we collected such a dataset to evaluate our method.

Query log We collected about 0.4 billion query logs during two weeks (2013/10/01∼2013/10/14) from a commercial image search engine. Each log consists of five items $<$User ID, Market, Query, Request Time, URLs$>$. In this paper, only en-US market logs were kept. After removing robot agents and low frequency queries, there are about 24M unique users, 5M unique queries, and 21M unique URLs in the dataset.

Candidate trending images The trending queries are detected based on Eqn. (2). We set the time window length $n = 3$. For the remaining 11 days, we collect the top 10 queries with the highest *BuzzScore'* as trending queries each day, which yields 110 trending queries in total. For each trending query, we gather all search queries that contain the same terms as the trending query in the query log to form a trending topic. Next, we submit all queries in the same trending topic into the image search engine by its API and collect the top 150 returned images for each query. In order to reduce the redundancy, we only keep one image for repeated images with the same URL. Finally, we select at most top 100 click-volume images as our candidate trending image set for the trending query.

Trending image labelling In order to get ground truth of trending images, eight volunteers were invited to label the candidate image set, seven males and one female. All of them were familiar with the search engine and could comprehend English articles quickly. For each trending query, the volunteers searched related trending news appearing during the same time the event was detected. Since they were aware of the trending event, the volunteers labeled each image as a trending image (positive instance) or a common image (negative instance), just as photo editors in a newsroom might do.

4.2. Experimental Setting

We divide our dataset into two parts. The first part consists of 40 trending queries from 2013/10/04∼2013/10/07, which are used for training. The second part includes the remaining 70 trending queries from 2013/10/08∼2013/10/14, which are used for testing. We use the famous open source library OpenCV version 2.4.9 with a machine learning module to train our trending image selection model. To compare the predictive performance of different learning models, we try several common ones including the Linear-kernel SVM, RBF-kernel SVM, Random Forest, Boosting, and Neural Networks. For all models, parameters are set via cross validation on training set. We calculate the trending score for each test image, which represents the probability of the image to be trending. For each trending query from the test set, its candidate images are ranked according to their trending scores. The performance is measured by the widely used non-interpolated Average Precision (AP). We average the AP values over all the 70 testing queries to get the Mean Average Precision (MAP) to measure the overall performance.

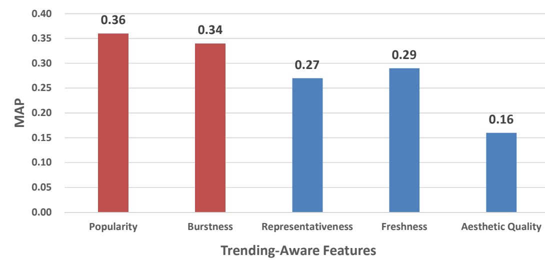

Fig. 5. The distinctive power of individual trending-aware features for trending image selection.

4.3. Experimental Results

As introduced in Section 3.3, we propose five trending-aware features. Here we first investigate the effectiveness of each individual feature on the trending metric. We rank images based on each individual feature, and calculate the MAP value to measure relevance between each feature and the trending label. The results are shown in Figure 5. We can see that *Popularity* and *Burstness* are the two most discriminative features in characterizing the trending images. *Representativeness* and *aesthetic quality* only consider the visual content, while *freshness* only considers the upload time, resulting in poor performance due to insufficient information. Therefore combining these different, yet complementary features should be a more effective method to measure the trending degree of an image.

To combine those features together, we directly concatenate them into a long feature vector and try different learning models to conduct the trending image selection. The experimental results are given in Figure 6. We can see that Linear-SVM and RBF-SVM outperform all the other models and also achieve better performances than all individual features in Figure 5. Therefore, we select the Linear-kernel SVM model to combine all the extracted features in our method.

As aforementioned, there is little work on automatic trending image selection. Some similar work concerns canon-

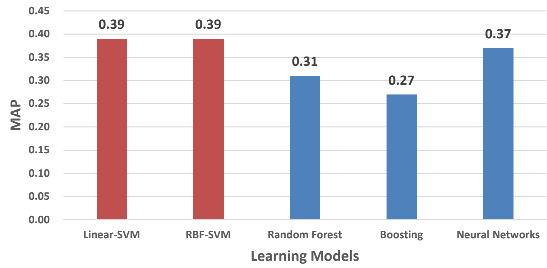

Fig. 6. The performance of different learning models with all features combined.

Table 1. Comparison between our method and related works.

Method	MAP	Improvement(%)
[7]	0.27	44.4
[4]	0.34	14.7
Ours	**0.39**	-

ical image selection, where the *VisualRank* algorithm [7] is commonly used, and the recent work focuses on personalized trending image search suggestion [4], which assumes that the burstness of click volume over time should be the most powerful predictor for trending images. We compare our method with these two methods. The experimental results are summarized in Table 1. We achieve 44.4% improvement when compared to the *VisualRank* algorithm, and a boost of 14.7% when compared to recent solutions in [4]. This is because [7] only considers visual content without taking other contextual factors into account and [4] leverages one single aspect with a naive ranking method.

5. CONCLUSION

In this paper, we present a learning-based trending image selection method. First, we detect trending queries from a query log based on the topic Buzz Score. Then, we automatically select trending images from an image set collected by a commercial search engine. We carefully select a set of trending-aware features to depict different views of trending images, including popularity, burst, representativeness, freshness, and aesthetic quality. In order to make full use of the information from these different features, we combine the complementary features by employing a learning model. Comparison experiments on a large-scale dataset show that our method outperforms other methods.

6. REFERENCES

[1] Stina Westman and Pirkko Oittinen, "Image retrieval by end-users and intermediaries in a journalistic work context," in *IIiX*. ACM, 2006, pp. 102–110.

[2] Marjo Markkula and Eero Sormunen, "End-user searching challenges indexing practices in the digital newspaper photo archive," *Information Retrival*, vol. 1, pp. 259–285, 2000.

[3] Julian Stottinger, Jana Banova, Thomas Ponitz, Nicu Sebe, and Allan Hanbury, "Translating journalists' requirements into features for image search," in *VSMM*. IEEE, 2009, pp. 149–153.

[4] Chun-Che Wu, Tao Mei, Winston H. Hsu, and Yong Rui, "Learning to personalize trending image search suggestion," in *SIGIR*. ACM, 2014, pp. 727–736.

[5] Lyndon S. Kennedy and Mor Naaman, "Generating diverse and representative image search results for landmarks," in *WWW*. ACM, 2008, pp. 297–306.

[6] Yushi Jing, Shumeet Baluja, and Henry Rowley, "Canonical image selection from the web," in *CIVR*, 2007, pp. 280–287.

[7] Yushi Jing and Shumeet Baluja, "Pagerank for product image search," in *WWW*, 2008, pp. 307–316.

[8] Tamara L Berg and D. A. Forsyth, "Automatic ranking of iconic images," Tech. Rep., University of California, Berkeley, 2007.

[9] Tamara L Berg and Alexander C Berg, "Finding iconic images," in *CVPR Workshops*. IEEE, 2009, pp. 1–8.

[10] Xin-Jing Wang, Zheng Xu, Lei Zhang, Ce Liu, and Yong Rui, "Towards indexing representative images on the web," in *MM*. ACM, 2012, pp. 1–8.

[11] Winston H. Hsu, Lyndon S. Kennedy, and Shih-Fu Chang, "Video search reranking through random walk over document-level context graph," in *MM*. ACM, 2007, pp. 971–980.

[12] Mathilde Sahuguet and Benoit Huet, "Socially motivated multimedia topic timeline summarization," in *SAM*. ACM, 2013, pp. 19–24.

[13] Mathilde Sahuguet and Benoit Huet, "Mining the web for multimedia-based enriching," in *MMM*. Springer International Publishing, 2014, pp. 263–274.

[14] Xueliang Liu and Benoit Huet, "Event representation and visualization from social media," in *14th Pacific-Rim Conference on Multimedia*. Springer International Publishing, 2013, pp. 740–749.

[15] Xueliang Liu, Raphael Troncy, and Benoit Huet, "Finding media illustrating event," in *ICMR*. ACM, 2011, pp. 58:1–58:8.

[16] Ziad Al Bawab, George H. Mills, and Jean-Francois Crespo, "Finding trending local topics in search queries for personalization of a recommendation system," in *SIGKDD*. ACM, 2012, pp. 397–405.

[17] Csurka G., Bray C., Dance C., and Fan L., "Visual categorization with bags of keypoints," in *ECCV*, 2004, pp. 1–22.

[18] Yan Ke, Xianou Tang, and Feng Jing, "The design of high-level features for photo quality assessment," in *CVPR*. IEEE, 2006, pp. 419–426.

[19] Subhabrata Bhattacharya, Rahul Sukthankar, and Mubarak Shah, "A framework for photo-quality assessment and enhancement based on visual aesthetics," in *MM*. ACM, 2010, pp. 271–280.

[20] Bo Geng, LinJun Yang, Chao Xu, Xian-Sheng Hua, and Shipeng Li, "The role of attractiveness in web image search," in *MM*. ACM, 2011, pp. 63–72.

[21] Zhe Dong and Xinmei Tian, "Effective and efficient photo quality assessment," in *SMC*. IEEE, 2014.