VISUAL EVENT MINING FROM GEO-TWEET PHOTOS

Takamu Kaneko and Keiji Yanai

Department of Informatics, The University of Electro-Communications, Tokyo
1-5-1 Chofugaoka, Chofu-shi, Tokyo, 182-8585, Japan

ABSTRACT

In this paper, we propose a system to mine events visually from the Twitter stream by making use of "geo-tweet photos" which are tweets including both geotags and photos. Some works on event mining which utilize geotagged tweets have been proposed so far. However, they used no images but only textual analysis of tweet texts. In this work, we detect events using visual information as well as textual information. In the experiments, we show some examples of detected events and their photos such as "rainbow", "fireworks" and "Tokyo firefly festival".

***Index Terms*—** Twitter, Geotagged Images, Event Mining, Geo-Photo Tweet

1. INTRODUCTION

Twitter is a unique microblog, which is different from conventional social media in terms of its quickness and on-the-spotness. Many Twitter's users send messages, which is commonly called "tweets", to Twitter on the spot with mobile phones or smart phones, and some of them send photos and geotags as well as tweets. Most of the photos are sent to Twitter soon after taken, and geotags attached to tweets generally represent the current locations of the Twitter's users.

In fact, we collected about 200 thousand geotagged photos a day on average in December 2012 via Twitter Streaming API. Such geotagged photo tweets are sometimes posted from the places where some events happen such as festival, sport game and earthquake. Then, we can regard Twitter's users as distributed image sensors on events.

We think that Twitter is a promising data source of geotagged of photos, while Flickr has been the most popular data source of geotagged photos in the research community of multimedia so far. Since the characteristic of Twitter is quickness and on-the-spot-ness, the photos on Twitter are different from the photos on Flickr. Flickr has many travel photos, while Twitter has many photos related to everyday life such as food, weather, street scene and some events. Therefore, we think that geo-tweet photos are more useful to understand what happens currently over the world than Flickr geo-photos.

In this paper, we propose a system to mine events visually from the Twitter stream. To do that, we pay attention to the tweets having both geotags and photos. We call such tweets as "geo-tweet photos". So far some works on event mining which utilize geotagged tweets have been proposed. However, they used no images but only textual analysis of tweet texts. On the other hand, in this work, we detect events using visual information as well as textual information. In the experiments, we show some examples of detected events and their photos such as "rainbow", "fireworks" and "Tokyo firefly festival".

2. RELATED WORK

Sakaki et al. [1] regarded Twitter users as social sensors which monitor and report the current status of the places where the users are. They proposed a system which estimates the location of natural events such as typhoons and earthquakes. They used geotagged tweets to estimate event locations.

Lee et al. [2] proposed an event detection system from geotagged tweets. They divided target areas into small subregions, and monitor the number of tweets posted from each sub-region. They regarded the areas where the number of tweets rose suddenly as the event areas where some events happened. In our work, we also examine the daily changes on the number of tweets of each area to detect events.

In these two works, they used textual information extracted from tweets and geo-location information embedded in geotags, and did not use visual information which can be extracted from tweet photos.

As works on geo-tweet photos, Yanai proposed World Seer [3] which can visualize geotagged photo tweets on the online map in the real-time way by monitoring the Twitter stream. Nakaji et al. [4] proposed a system to mine representative photos related to the given keyword or term from a large number of geo-tweet photos. They extracted representative photos related to events such as "typhoon" and "New Year's Day", and successfully compared them in terms of the difference on places and time. However, their system needs to be given event keywords or event term by hand. Then, in this paper, we integrate a method to select representative event photos with automatic detection of event keywords.

3. OVERVIEW

To detect events visually from Twitter stream, we monitor Twitter stream to pick up tweets having both geotags and photos, and store them into a geo-photo tweet database. We apply to this database the proposed visual event mining system which consists of event keyword detection, event photo clustering and representative photo selection. The processing steps of the proposed system are as follows:

(1) Detect event keyword candidates which frequently appear in the tweets posted from specific areas in specific days.

(2) Unify and complement detected event keywords

(3) Select geo-tweet photos corresponding to the event keywords by image clustering

(4) Select a representative photo to each event

(5) Show the detected events with their representative photos on the map

Note that the current system assumes the tweet messages written by Japanese language, since keyword extraction needs to be taken into account of the characteristics of target language. However, it is not so difficult to extend the proposed system to other languages.

4. PROPOSED METHOD

In this section, we explain the detail of each step of the proposed system.

4.1. Event keyword detection

Tweet messages are written in sentences in general. To detect event easier, we extract noun keywords from each tweet message. To do this, we apply the Japanese morphological analyzer, MeCab [1], and extract only noun words as keywords of each tweet after stop-word removal.

To detect events, we search for bursting keywords by examining change of the daily frequency of each keywords within each unit area. The area which is a location unit to detect events are defined in the grids by one degree latitude and one degree longitude as shown in Fig.1. In case that the daily frequency of the specific keyword within one grid area increases greatly, we consider that an event related to the specific keyword happened within the area in that day.

We set up the following equation to decide if an event related to the given keyword happens in the given area. We consider that an event happens if $S_{k,d,i,j}$ is more than the predefined threshold, which was set as 50 in the experiments.

$$S_{k,d,i,j} = (N_{k,d,i,j} - N_{k,d-1,i,j})W_{i,j} \ , \tag{1}$$

where k, d, i, j and $N_{k,d,i,j}$ represent an index of a keyword, an index of date, an index of area grids, and the number users

[1] http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html

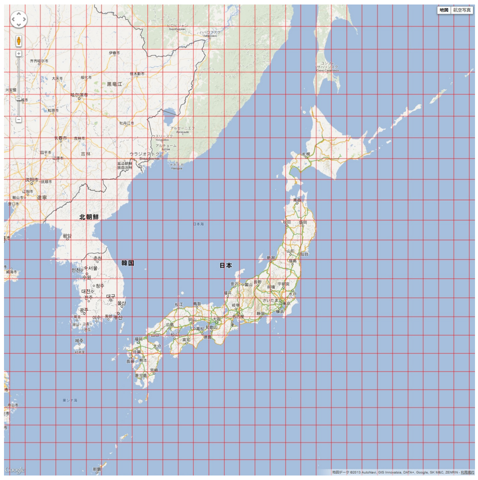

Fig. 1. The grids dividing the Japanese Islands. Each of them is a unit area for event detection.

who posted tweets in the indicated day and area, respectively. $W_{i,j}$ represents a weight to adjust the scale of the number of daily tweet users, which is defined in the following equation:

$$W_{i,j} = \frac{M + s}{N_{i,j} + s} \ , \tag{2}$$

where i, j, N, M and s represents the index of grids, the number of unique users in the given grid, the maximum number of unique users among all the grids (which is equivalent to the number of Tokyo area users), and the standard deviation of user number over all the grids. With this adjusting weight, we can detect events from the areas where tweet users are not so many as well as the areas where so many tweets are always being posted such as Tokyo area.

4.2. Keyword unification and complement

In the previous step, we limited an event keyword to a single noun keyword. However, since some events are represented by compound keywords, the same event are sometimes detected by several keywords independently. In such case, we unify them into a compound keyword related to the same event according to the following heuristics:

(1) In case that more than half of the tweets related to a specific event keyword overlaps the tweets related to another event keyword, the former keywords are integrated and replaced with the latter keywords.

- E.g. “rain” and “typhoon” ⇒ “typhoon”

(2) In case that words just after or before the detected event keyword are the same in more than 80% tweets including the keyword, such words are regarded as being part of a compound event keyword.

- E.g. “Tokyo”, “sky” and “tree” ⇒ “Tokyo Skytree”

4.3. Event photo clustering and representative photo selection

Until the previous step, event keywords and their corresponding tweets have been selected. In this step, we carry out clustering of the photos embedded in the selected event tweets and selecting representative ones from them.

As image features, we use bag-of-features (BoF) with densely-sampled SURF local features and 64-dim RGB color histograms. SURF keypoints are sampled with every 10 pixels in the scale 5, 10 and 15. The size of the codebook for BoF was set as 1000. Both feature vectors are L1-normalized.

For clustering photos, we use the Ward method which is one of agglomerative hierarchical clustering methods. It creates clusters so to minimize the total distance between the center of each cluster and the cluster members. It merges the cluster pairs which bring the minimum total error calculated in the following equation one by one.

$$d(C_1, C_2) = E(C_1 \cup C_2) - E(C_1) - E(C_2) \tag{3}$$

In general, $E(C)$ is defined as the total distance between the center and the members of the cluster. Since we use two kinds of visual features, we defined $E(C)$ to combine them in the following equation.

$$\begin{aligned} E(C) = \sum_{x \in C} ((x_{BoF} - \overline{x}_{BoF})^2 w_{BoF} \\ +(x_{Color} - \overline{x}_{Color})^2 w_{Color}) \ , \end{aligned} \tag{4}$$

where x_{BoF}, x_{Color}, $\overline{x}$ and w represent a BoF feature vector, a RGB color histogram vector, a vector of the center of the cluster, and the weight which is defined as a reciprocal number of the dimension of each feature vector.

We evaluate each of the obtained clusters in terms of visual coherence with the following equation. We designed this equation so that the score of the cluster the member photos of which are similar to each other becomes larger.

$$V_C = \frac{n_C^2}{E(C)} W_{i,j} \ , \tag{5}$$

where n_C represents the number of photos in cluster C, and $W_{i,j}$ is a adjustment weight defined in Eq.(2). When V_C is high, the corresponding cluster is likely to strongly related to the event. On the other hand, in case that V_C is lower, the cluster is expected to be a noise one which is less related to the event. In the experiments, we set the threshold of V_C as 5 which was decided based on the results in the preliminary experiments.

In addition, the cluster having the maximum value of V_C is regarded as a representative cluster, and the photo the visual feature vector of which is the closest to the cluster center is selected as a representative photo for the corresponding event.

5. EXPERIMENTS

In the experiment, we used about three million geo-tweet photos posted from Japan which were collected from the Twitter stream from February 10th, 2011 to September 30th, 2012.

5.1. Experimental results on keyword selection

As results of event keyword extraction for the given dataset, we obtained 306 keywords related to natural phenomena such as "rainbow" and "typhoon" and local events related to "fireworks" and "festival". Part of the results are shown in Table 1. In the table, "Area", "Weight" and "Score" represent the bounds of the grid in terms of latitude and longitude, the value of Eq.(2), and the value of Eq.(1), respectively. Since the area where there are the largest number of unique users who posted geo-photo tweets was Tokyo, the weight value of the Tokyo area become 1.0. Because the other areas have less users than Tokyo, the adjusting weight value become more than 1.0.

As results of keyword unification and complement, the words which originally come from the same compound word such as "fireworks festival" are unified and converted into a compound keyword. Part of the results of unification is shown in Table 2, and part of the results of complement is shonw in Table 3.

5.2. Experimental results on photo clustering

Next, we show some example results of event photo clustering corresponding to five keywords, "fireworks", "tree", "cherry blossoms", "rainbow" and "firefly" (See Table 4 for the detail) in Fig.2, 3, 4 5 and 6. The numbers shown on the right of each photo cluster represent cluster scores. The clusters (with red boxes) having the score which is with more than 5.0 are regarded as event photo clusters, while the rest clusters (with blue boxes) are regarded as non-event clusters unrelated to the corresponding event keyword. Within each cluster, photos are sorted in the ascending order of the distance to the cluster center. From the results, scoring of clusters worked successfully to place more visual clusters in the higher rank.

In Fig.3 ("tree"), the first cluster represents Christmas trees, while the second cluster represents Tokyo Skytree in the night. In Fig.5 ("rainbow"), the score of the second cluster was less than 5, since it contains many of the photos including relatively large and various background regions. In Fig.6 ("firefly"), the first cluster represents illumination event of Tokyo Skytree which was called "Tokyo firefly"

Some detected events are shown on the map with their representative photos in Fig.7. This map is an interactive system based on Google Maps API, and a user can see any event photos by clicking markers.

Finally, 258 events were detected in this experiment. All the 258 detected events can be regarded as being related to some of various kind of actual "events" including weather condition, natural events, festivals and sport games. We evaluated if all the presentative photos of the detected events are relevant to the event or not subjectively by hand, and the precision of the representative photos were 65.5%.

6. CONCLUSIONS

In this paper, we proposed a visual event mining system from the Twitter stream using visual information as well as textual and location information. The system enables us to discover and understand events visually, which is the novel contribution of this work.

For future work, we plan to propose more sophisticated visual event mining methods which integrate visual, textual and location information more closely and more comprehensively. In the current system, the grid size and the term to extract events are fixed to one degree and one day, respectively. We will extend the system so that the grid size and the time unit for detecting events are adjusted automatically depending on events.

In addition, real-time visual event detection is also one of our future works. We also plan to analyze the difference between Tweet photos and Flickr photos in terms of their characteristic.

7. REFERENCES

[1] T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes twitter users: real-time event detection by social sensors," in *Proc. of the International World Wide Web Conference*, 2010, pp. 851–860.

[2] R. Lee and K. Sumiya, "Measuring geographical regularities of crowd behaviors for twitter-based geo-social event detection," in *Proc. of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks*, 2010, pp. 1–10.

[3] K. Yanai, "World seer: a realtime geo-tweet photo mapping system," in *Proc. of ACM International Conference on Multimedia Retrieval*, 2012.

[4] Y. Nakaji and K. Yanai, "Visualization of real world events with geotagged tweet photos," in *Proc. of IEEE ICME Workshop on Social Media Computing (SMC)*, 2012.

Table 1. Part of the list of the extracted keywords

keyword	date	area	weight	score
snow	2011/2/11	34,35,135,136	1.96	135.5
earthquake	2011/3/11	35,36,139,140	1	55
fireworks	2011/8/6	34,35,135,136	1.96	149.2
festival	2011/8/6	34,35,135,136	1.96	68.7
Yodo-river	2011/8/6	34,35,135,136	1.96	72.6
dome	2011/8/10	43,44,141,142	3.96	51.5
rain	2011/8/19	35,36,139,140	1	60
typhoon	2011/9/21	35,36,139,140	1	62
Mt.Fuji	2011/9/24	35,36,138,139	3.35	67
Apple	2011/10/6	35,36,139,140	1	70
Ginza	2011/10/6	35,36,139,140	1	51
Suzuka	2011/10/9	34,35,136,137	3.94	78.8
Circuit	2011/10/9	34,35,136,137	3.94	67
Age	2011/10/23	35,36,135,136	3.46	55.4
eclipse	2011/12/10	34,35,135,136	1.96	84.4
total	2011/12/10	34,35,135,136	1.96	58.9
Christmas	2011/12/24	35,36,136,137	2.9	55.2
New-Year's-Eve	2011/12/31	35,36,139,140	1	68
sunrise	2012/1/1	35,36,139,140	1	84
New-Yesr's-Day	2012/1/1	35,36,139,140	1	69
Meiji	2012/1/1	35,36,139,140	1	50
Eho	2012/2/3	35,36,139,140	1	63
ski	2012/2/11	36,37,138,139	3.69	77.5
Valentine	2012/2/14	35,36,139,140	1	58
Marathon	2012/2/26	35,36,139,140	1	77
Roppongi	2012/3/24	35,36,139,140	1	88
cherry-blossoms	2012/4/28	37,38,140,141	4.18	121.4
super	2012/5/5	35,36,139,140	1	93
moon	2012/5/5	35,36,139,140	1	96
firefly	2012/5/6	35,36,139,140	1	59
mother	2012/5/13	35,36,139,140	1	63
eclise	2012/5/21	35,36,139,140	1	314
annular	2012/5/21	35,36,140,141	3.18	60.5
Tanabata	2012/7/7	34,35,135,136	1.96	56.9
Gion-Festival	2012/7/14	35,36,135,136	3.46	104
Tohoku-Denryoku	2012/7/14	37,38,139,140	4.4	79.2
sky	2012/7/28	35,36,139,140	1	54
tree	2012/7/28	35,36,139,140	1	53
peace	2012/8/6	34,35,132,133	4.08	77.5
Makuhari Messe	2012/8/11	35,36,140,141	3.18	168.9
Awa	2012/8/12	34,35,134,135	3.91	54.8
Okuribi	2012/8/16	35,36,135,136	3.46	104
Daimonji	2012/8/16	35,36,135,136	3.46	83.2
Okinawa	2012/9/8	26,27,127,128	4.17	66.8
wind	2012/9/15	38,39,140,141	3.9	58.5
lakeside	2012/9/15	38,39,140,141	3.9	66.3
Nihon-daira	2012/9/15	34,35,138,139	4.15	78.8

Table 2. Results of keyword unification.

keywords before unification	keyword after unification
fireworks, festival	fireworks
fireworks, festival, Tama-river	fireworks
fireworks, festival, Edo-river	fireworks
fireworks, festival, Lake-Suwa	fireworks
Apple, Ginza	Apple
eclipe, Total	eclipse
Roppongi, Hills	Roppongi
wind, rain	wind
cherry blossoms, beautiful	cherry blossoms
super, moon	super
blue, moon	blue
Sky, tree	sky
Suzuka, circuit	circuit

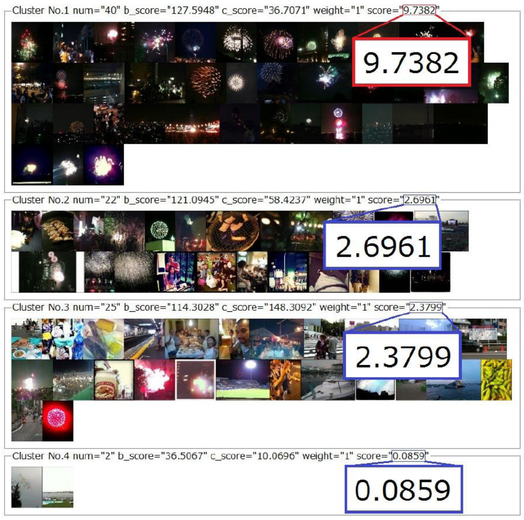

Fig. 2. "Fireworks" photo clusters.

Table 3. Results of keyword completion.

keywords before completion	compound keyword after completion
dome	Sappro Dome
Apple	Apple Store
Meiji	Meiji Shrine
Eho	Eho-maki
ski	ski park
Marathon	Tokyo Marathon
Marathon	Kyoto Marathon
super	super moon
blue	blue moon
firefly	Tokyo firefly
mother	mother's day
gone	Typhoon gone
Tohoku-denryoku	Tohoku-denryoku Big One Studium
sky	sky-tree
circuit	Suzuka Circuit
Peace	Peace Memorial Park
Marine	QVC Marine Filed
Nihon-daira	Outsourcing Studium Nihon-daira

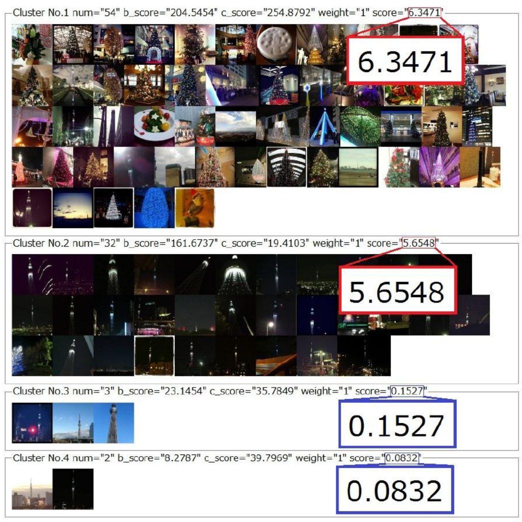

Table 4. Summary for the example results.

event keyword	date	grid (lat,lng)	area	# photos
fireworks	2011/12/23	35,36,139,140	Tokyo	91
tree	2011/12/23	35,36,139,140	Tokyo	91
cherry blossoms	2012/04/21	34,35,135,136	Osaka	57
rainbow	2012/05/04	35,36,139,140	Tokyo	93
firefly	2012/05/06	35,36,139,140	Tokyo	93

Fig. 3. "Tree" photo clusters.

Fig. 4. “Cherry blossoms” photo clusters.

Fig. 5. “Rainbow” photo clusters.

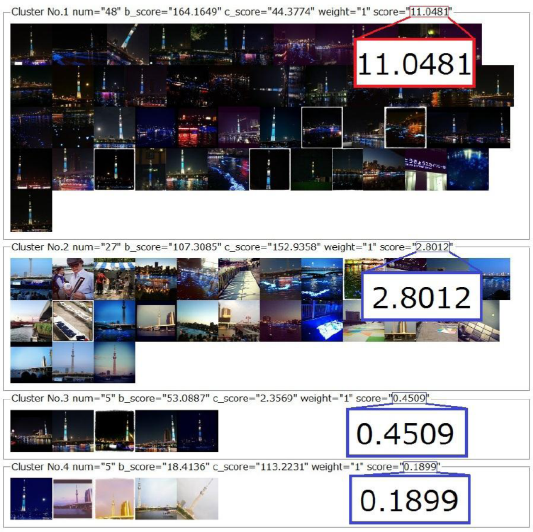

Fig. 6. “Firefly” photo clusters.

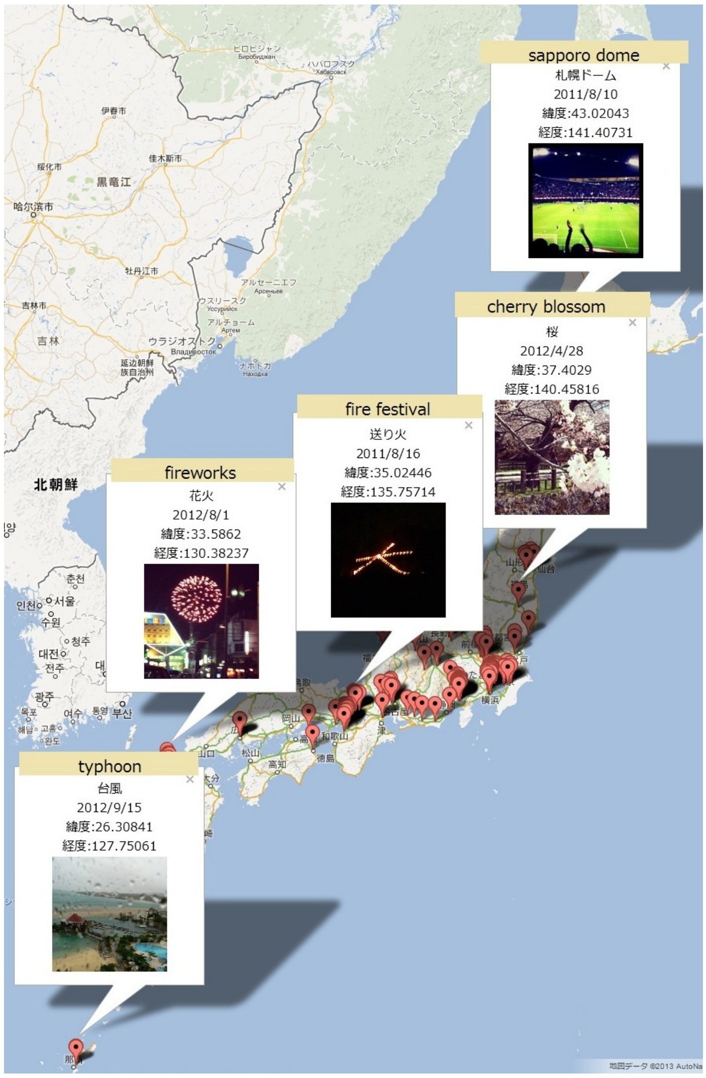

Fig. 7. Some detected events are shown on the map with their representative photos.