

大量の位置情報付き Twitter 画像データからの 視覚的イベント検出

金子 昂夢¹ 柳井 啓司¹

1. はじめに

近年、スマートフォンのような位置情報を扱えるデバイスの普及に伴い、撮影した画像に位置情報を付加することが容易となった。また、ツイートと呼ばれる短文を投稿できるサービス Twitter の普及により、自分が今何をしているのか、どこにいるのかといったことを投稿する人が増加してきた。ツイートには本文の他に画像や位置情報を添付でき、スマートフォンを利用して外出先で撮影した位置情報付き画像を即座に投稿することができる。Twitter は Web マイニング対象として広く研究されており、イベントの検出を試みた研究も多くある。

榊ら [3] は、Twitter のユーザを実世界のイベントを監視するソーシャルセンサと捉えたモデルを作り、イベントの検出とその位置の推定を行うシステムを作成した。タイムラインを監視することでリアルタイムな検出を行い、高精度かつ高速な検出が行われた。

中地ら [1] は、あらかじめ「台風」、「正月」といったキーワードや 2011 年 3 月 11 日から 2011 年 3 月 12 日といった期間を設けて Twitter から位置情報付き画像ツイートを収集し、クエリに対する代表的な画像を選出することで画像付きのイベント検出を試みた。

本研究では、Twitter に投稿された位置情報付き画像ツイートから視覚的なイベントの検出を試みる。イベントを検出し、その情報を画像と共に地図上に表示することで、今どこで何が起きているのかを視覚的に捉えることを目標とする。

2. 提案手法

2.1 キーワードの検出

2.1.1 局所的なバースト検出

出現頻度が急上昇（バースト）したキーワードを検出するために、まず、ツイートの本文に対して形態素解析を行い、品詞が名詞である単語をタグとして抽出した。

検出には、ユーザ人口の違いを考慮し、対象を緯度・経度 1 度ずつのグリッドに分割し、以下の式 1 により各地域の重みを算出した。

$$W_{i,j} = \frac{M + s}{N_{i,j} + s} \quad (1)$$

式 1 において、 i, j はそれぞれ緯度・経度によるグリッドを表し、 M は最も出現するユーザ数が多い地域のユーザ数、 $N_{i,j}$ はグリッド i, j に出現するユーザ数、 s は標準偏差である。

そして、タグの各単語について、以下の式 2 によるスコアが閾値以上となるものを抽出した。

$$S_{k,d,i,j} = (N_{k,d,i,j} - N_{k,d-1,i,j})W_{i,j} \quad (2)$$

式 2 において、 k は各単語、 d は注目している日、 N はそれぞれの条件で単語 k をタグに含むツイートを投稿したユーザ数である。

2.1.2 キーワードの統合・補完

抽出されたキーワードのうち、キーワードを含むツイートが 50% 以上同じである同日・同地域のキーワードは同じイベントに関するキーワードとして統合し、さらに、キーワードを含むツイート内で前後の文字が 80% 以上同じならその文字でキーワードを再帰的に補完した。

2.2 代表画像抽出

2.2.1 画像特徴量

画像特徴量は SIFT による Fisher Vector と RGB カラーヒストグラムを用いた。Fisher Vector [2] は、BoF における量子化誤差を低減したものであり、確率密度関数に GMM を仮定し、局所特徴群のモデル化を行った。各画像から dense sampling した SIFT を PCA により 64 次元とし、GMM の混合数を 64 として 8192 次元の Fisher Vector で表現した。また、画像を 2×2 に 4 分割し、それぞれの部分画像から RGB 空間における 64 次元の色頻度を抽出し、256 次元の RGB カラーヒストグラムを生成した。

2.3 クラスタリング

まず、イベントを表す画像群を抽出するため、2 つの特徴量を用いてクラスタリングを行った。クラスタリングには Ward 法を用い、エラーの増加量に対して閾値を設け、任意のクラスタ数に分類されるようにした。

ただし、今回は 2 つの画像特徴量を使用するため、Ward 法に用いるクラスタのエラーを以下の式 3 のように定義した。

$$E(C) = \sum_{x \in C} ((x_{FV} - \overline{x_{FV}})^2 w_{FV} + (x_{RGB} - \overline{x_{RGB}})^2 w_{RGB}) \quad (3)$$

¹ 電気通信大学 大学院情報理工学専攻 総合情報学専攻 〒182-8585 東京都調布市調布ヶ丘 1-5-1

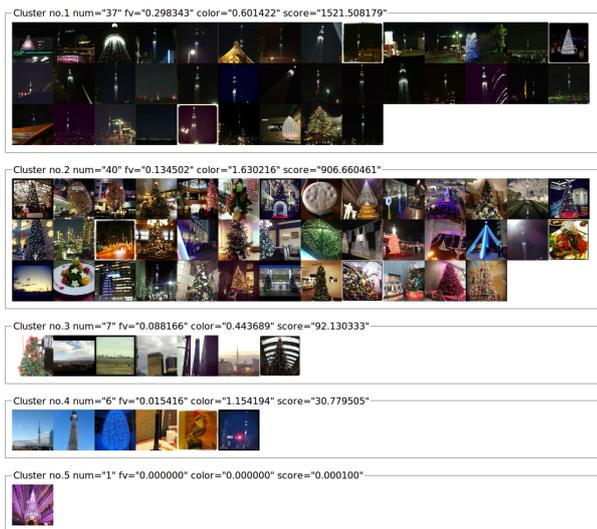


図 1 「ツリー」のクラスタリング結果

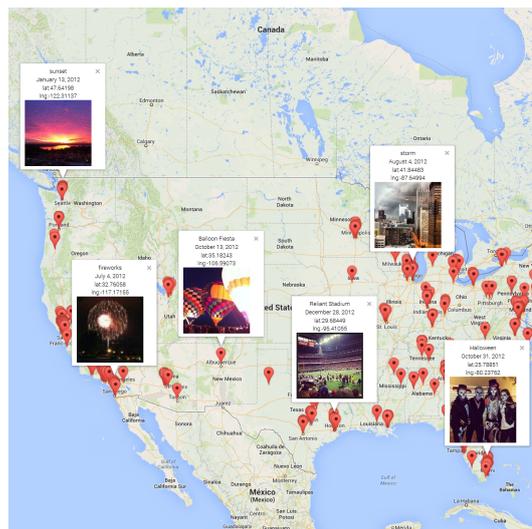


図 2 マーカーをクリックした例

式 3 において, x は画像の各特徴ベクトル, \bar{x} はクラスタ C における各特徴量の中心, w は各特徴量の重みである.

次, クラスタリング結果から, 各クラスタについて以下の式 4 によりスコアを付け, ランキングした.

$$V_C = \frac{n_C^2}{E(C)} W_{i,j} \quad (4)$$

式 4 において, n_C はクラスタ C に分類された画像枚数, $W_{i,j}$ は式 1 によって得られた, キーワードが検出された地域の重みである.

そして最上位のクラスタの内, 最も中心に近い画像を代表画像として選出し, 他の情報と共に地図上にマッピングした.

3. 実験

3.1 データセット

今回の実験を行うにあたって, Twitter に投稿された位置情報付き画像ツイートから 2 つのデータセットを作成した. 1 つは日本で 2011 年 2 月 10 日から 2012 年 9 月 30 日までに投稿されたものであり, 約 3,000,000 件の位置情報付き画像ツイートである. もう 1 つはアメリカで 2012 年 1 月 1 日から 2012 年 12 月 31 日までに投稿されたものであり, 約 17,000,000 件の位置情報付き画像ツイートである.

3.2 キーワードの検出結果

キーワードの検出を行った結果, 日本では「虹」のような自然現象から, 「花火大会」のような局所的なイベントに関するキーワードが検出され, アメリカでは「Super Bowl」のようなスポーツに関するキーワードが多く検出された.

3.3 クラスタリング結果

クラスタリング結果について, 日本で 2011 年 12 月 23 日に抽出されたキーワード「ツリー」に対する例を図 1 に示す. 各クラスタの右上の値がスコアであり, 上位 2 つのクラスタを見ると上はスカイツリーの画像が, 下はクリスマスツリーの画像が多く分類された結果となった.

3.4 検出結果の表示

検出されたイベントについて, 地図上にマーカーを設置し, いくつかのマーカーをクリックした例を図 2 に示す. Web 上のシステムでは表示された代表画像をクリックすることでクラスタリング結果が表示される.

最終的なイベントの検出結果は表 1 のようになった.

表 1 それぞれのデータセットの検出結果

	日本	アメリカ
イベント数	258	1676
代表画像の適合率 (%)	70.1	75.5

4. まとめ

本研究では, Twitter に投稿された位置情報付き画像ツイートからイベントの検出を行い, 視覚的関連性が高い画像群を抽出することでイベントを視覚的に捉えられるようにした.

今後の課題として, まず, 検出するグリッドの大きさと期間を可変にすることで柔軟にイベントの検出が行えるようにする. また, Twitter のタイムラインを監視することでリアルタイムなイベントの検出を行えるようにする.

参考文献

- [1] Nakaji, Y. and Yanai, K.: Visualization of Real World Events with Geotagged Tweet Photos, *Proc. of IEEE ICME Workshop on Social Media Computing (SMC)* (2012).
- [2] Perronnin, F., Sánchez, J. and Mensink, T.: Improving the Fisher Kernel for Large-scale Image Classification, *Proc. of European Conference on Computer Vision*, pp. 143–156 (2010).
- [3] Sakaki, T., Okazaki, M. and Matsuo, Y.: Earthquake shakes Twitter users: real-time event detection by social sensors, *Proc. of the International World Wide Web Conference*, pp. 851–860 (2010).