

# Flickr データに基づいた時空間イベント検出

谷 生<sup>†, a</sup>

熊野 雅仁<sup>†‡, b</sup>

小堀 聰<sup>†‡, c</sup>

木村 昌弘<sup>†‡, d</sup>

† 龍谷大学大学院理工学研究科電子情報学専攻 ‡ 龍谷大学理工学部電子情報学科

a) *t18m014@mail.ryukoku.ac.jp* b) *kumano@rins.ryukoku.ac.jp* c) *kobori@rins.ryukoku.ac.jp*  
d) *kimura@rins.ryukoku.ac.jp*

**概要** 近年、Twitterなどのソーシャルメディアにおいて、ユーザから投稿された発信時間およびGeo-tagなどの時空間情報付き tweet (文書情報) がビッグデータとして成長するに従い、一般の人々の実世界の行動を捉え、地域の分析や開発に利活用することが期待されつつある。Shin らは、地図システムで普及しているタイルマップとして区画化された地域群に対して、時空間情報付きの文書情報を用いてアノマリーイベントを検出する STExNMF を提案し、単語群により時空間イベントの説明を試みている。近年、ユーザが文書以外に写真を投稿する傾向が高まっているため、写真情報を有効活用できれば、文書情報では限界がある各地域についてのユーザの嗜好や意見をより深く捉える可能性が期待される。本研究では、Shin らの STExNMF を時空間情報付きの写真における画像情報を用いる方法として拡張し、時空間イベントを検出する枠組みを与える。そして、京都市を対象として収集した Flickr からの大量の写真データを用い、各地域の特色分析の観点から、時空間イベント検出の結果について画像群を用いて説明を試みる。

**キーワード** ソーシャルメディアマイニング、時空間イベント、アノマリー検出、観光都市開発

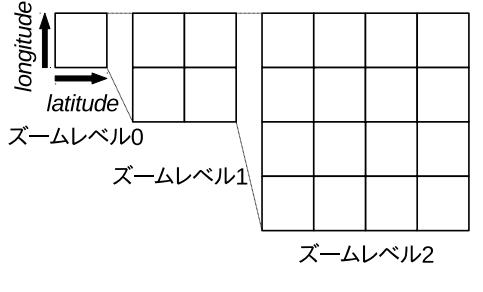
## 1 はじめに

Twitterなどのソーシャルメディアが普及し、各ユーザから投稿された発信時間およびGeo-tag が付随する時空間情報付き tweet がビッグデータとして成長するに従い、ある時ある場所で起きている実世界の時空間イベントを検出する研究 [1, 2] や、アノマリーイベント検出を促すための可視化分析に関する研究 [3, 4]、アノマリーイベント検出を行う研究 [6, 5, 7] など、文書情報と時空間情報を含んだビッグデータに基づく時空間イベント検出に関する研究が活発に行われている。

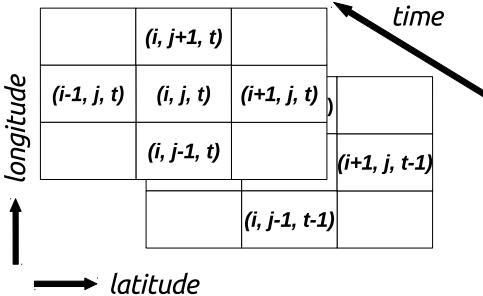
ところで、時空間情報を含んだビッグデータに関する研究の新たな活用先として、地域開発へ応用 [8] することが期待されているが、その一つのアプローチとして時空間イベント検出を応用する方法が考えられる。時空間情報付 tweet は各地域に存在する事物やある時期に存在するイベントに対する人々の嗜好や意見を反映している可能性がある。また、各地域の魅力を捉える場合、少なくとも周辺の地域には無い特色や、ある時期に特有のイベントを捉えることが望ましいと考える。すなわち、一般の人々によって投稿された時空間情報付の嗜好や意見に基づき、各地域のある期間に特有の対象を検出することで各地域ならではの特色を客観的に捉え直すなど、地域開発の支援に応用することが考えられる。Shin ら [7] は、地図システムで普及しているマップタイルで区画化された各地域に着目し、隣接するタイルおよび時間的に隣接する過去の期間とは異なる傾向を示す時空間アノマリーイベントを検出するため、NMF [12] を拡張した

STExNMF を提案し、大量の時空間情報付 tweet に基づいて検出されたアノマリーイベントの代表的な単語群から実世界上の時空間イベントの説明を試みている。

一方、近年、ソーシャルメディアのユーザは、文書情報の発信よりも、写真を撮影して投稿する機会が増えつつある。tweet では、時空間情報が付随していてもその地域と関係のない話題をつぶやく場合があるものの、geo-tag 付写真は撮影者が魅力を感じた対象に遭遇したときに撮影される傾向があり、その画像は撮影地点付近の対象を捉えるだけでなく、文書による表現が難しい対象の魅力やユーザの嗜好を内包している可能性がある。そのため、時空間情報付写真データから地域の魅力を検出し、文書情報だけでは限界があった人々の嗜好や意見を捉えることで地域開発への貢献が期待される。Geo-tag 付写真の画像データを用いて実世界上の人の行動イベントを捉える研究として、個人が撮影した写真群を画像の特徴から分類したりアノテーションを行う研究 [10, 11] がある。しかし、多くの人々が撮影した写真群の画像データから時空間イベントを検出し、各地域に特有の客観的な特色を抽出する取り組みは挑戦的な課題である。本研究では、一般の人々が撮影した大量の時空間情報付き写真データから各地域の特色を捉える研究の第一歩として、文書情報と時空間情報を活用した Shin ら [7] の STExNMF を写真の画像情報と時空間情報を用いる手法として拡張し、与えられた都市に対してマップタイルに基づいて区画化された各地域の時空間イベントを検出する枠組みを与える。そして、京都市を対象として写真共有サイト Flickr から収集した大量の写真画像データを用い、各地域の特色分析の観点から、時空



(a) マップタイルのズームレベル



(b) 時空間マップタイルとタイル座標

図 1 マップタイルによる時空間の区画化

間イベントの検出結果について画像群を用いて説明を試みる。

## 2 時空間イベント検出

与えられた都市に対して、Google map などの Web ベース地図システムで広く用いられているマップタイルによりその都市を区画化し、Flickr から得られる写真データを用いて、区画化された各地域における時空間アノマリーイベントの検出を考える。

### 2.1 特徴量の抽出

マップタイルは、ズームレベル（図 1(a) 参照）と呼ばれる地図の表示倍率を指定することによって構築される。我々は時空間イベントを論じるので、対象とする都市に対して、時間解像度を指定し通常の空間的マップタイルを時間軸方向に拡張した時空間マップタイルを考え、そのタイル座標を導入する（図 1(b) 参照）。ここに、座標  $(i, j, t)$  のタイルは、通常のマップタイルでの区画された地域  $(i, j)$  の時間ステップ  $t$  における状況を表し、タイル  $(i-1, j, t)$ ,  $(i, j-1, t)$ ,  $(i, j, t-1)$  などは、その近傍タイルを表す。本研究では、入力データは Flickr から得られる写真データであるので、タイル  $(i, j, t)$  には地域  $(i, j)$  で時間ステップ  $t$  に撮影された写真群が対応している。ここに、そのような写真の総数を  $n_{i,j,t}$  とする。我々は、各タイルにおけるアノマリーイベントの検出を考える。

入力写真データからなる画像集合に対して、画像認識

などによく行われるように、SIFT [13] や SURF [14] などの特徴量に基づいて画像を表現し、クラスタリング技術を援用することで、各画像  $v$  を語彙数  $m$  の Bag-of-Visual-Words(BoVW)  $v = (v_1, \dots, v_m)$  で表現する。ただし、 $v_1 + \dots + v_m = 1$  と正規化している。各タイル  $(i, j, t)$  に対して、それに対応する写真画像を BoVW 表現し、 $m$  次元ベクトルを  $m \times 1$  行列に変換する。そして、それら  $n_{i,j,t}$  個の  $m \times 1$  行列を列方向に並べて、 $m \times n_{i,j,t}$  行列  $A_{i,j,t}$  を構築する。ここに、行列  $A_{i,j,t}$  はタイル  $(i, j, t)$  の特徴量であり、その各列は対応する写真画像を表していることに注意する。

### 2.2 NMF

非負値行列分解 (NMF) 法を基本として、各タイル  $(i, j, t)$  のアノマリーイベントを検出することを考える。まず、潜在トピック数を  $k$  として、標準的な NMF [12] を適用し、

$$A_{i,j,t} \approx W_{i,j,t} H_{i,j,t}$$

となる  $m \times k$  非負行列  $W_{i,j,t}$  と  $k \times n_{i,j,t}$  非負行列  $H_{i,j,t}$  を求める。そして、 $W_{i,j,t}$  の  $k$  個の列ベクトルに対応して、タイル  $(i, j, t)$  の  $k$  個のトピックを抽出する。

### 2.3 STExNMF

各タイル  $(i, j, t)$  のアノマリーイベントを検出するために、さらに STExNMF 法 [7] を援用することを考える。STExNMF 法では、時空間マップタイルの近傍情報を利用して、NMF 法が抽出するトピックを時空間的に特徴あるものとすることを試みる。

まず、空間的および時間的にどのくらい離れたタイルまでを近傍タイルとするかを定義し、それに基づいてタイル  $(i, j, t)$  の近傍タイル集合  $\mathcal{U}_{i,j,t}$  を決定する。

すべての近傍タイル  $g \in \mathcal{U}_{i,j,t}$  に対して、標準的な NMF 法を適用し  $m \times k$  非負行列  $W_g$  を求め、それを写真数で重みづけした行列  $n_g W_g$  を考える。そして、それら  $|\mathcal{U}_{i,j,t}|$  個の  $m \times k$  行列を列方向に並べて、 $m \times k |\mathcal{U}_{i,j,t}|$  行列  $\tilde{W}_{i,j,t}$  を構築する。ここに、行列  $\tilde{W}_{i,j,t}$  はタイル  $(i, j, t)$  の近傍タイル集合  $\mathcal{U}_{i,j,t}$  のトピック全体を表していると考えられる。 $\tilde{W}_{i,j,t}$  に標準的な NMF 法を適用し、

$$\tilde{W}_{i,j,t} \approx W'_{i,j,t} \tilde{H}_{i,j,t}$$

となる  $m \times k$  非負行列  $W'_{i,j,t}$  と  $k \times k |\mathcal{U}_{i,j,t}|$  非負行列  $\tilde{H}_{i,j,t}$  を求める。ここに、行列  $W'_{i,j,t}$  はタイル  $(i, j, t)$  の近傍タイル集合  $\mathcal{U}_{i,j,t}$  の共通トピック群を表していると考えられる。

次に、 $0 < \alpha < 1$  とし、タイル  $(i, j, t)$  の情報  $A_{i,j,t}$  を削減した  $\alpha A_{i,j,t}$  に対して、それを近傍タイル集合の共通トピック群  $W'_{i,j,t}$  で説明するために、行列  $W'_{i,j,t}$  を固

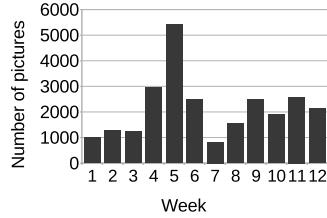


図 2 京都データの写真数変遷

定して  $\alpha A_{i,j,t}$  に標準的な NMF 法を適用し,

$$\alpha A_{i,j,t} \approx W'_{i,j,t} H'_{i,j,t}$$

となる  $k \times n_{i,j,t}$  非負行列  $H'_{i,j,t}$  を求める. そして, 行列  $A_{i,j,t} - W'_{i,j,t} H'_{i,j,t}$  を閾値処理して得られる  $m \times n_{i,j,t}$  非負行列  $R_{i,j,t}$  に標準的な NMF 法を適用し,

$$R_{i,j,t} \approx W''_{i,j,t} H''_{i,j,t}$$

となる  $m \times k$  非負行列  $W''_{i,j,t}$  と  $k \times n_{i,j,t}$  非負行列  $H''_{i,j,t}$  を求める.  $W''_{i,j,t}$  の  $k$  個の列ベクトルに対応して, タイル  $(i, j, t)$  の  $k$  個のトピックを抽出する.

#### 2.4 アノマリーイベントの検出

各タイル  $(i, j, t)$  に対して, 求めたトピック行列  $W''_{i,j,t}$  の第  $\ell$  列ベクトルを  $w_{i,j,t}(\ell)$  とする. ここに,  $\ell = 1, \dots, k$  である. 我々は, Shin ら [7] に従い, 各  $m$  次元ベクトル  $w_{i,j,t}(\ell)$  がタイル  $(i, j, t)$  の一つのアノマリーイベントを表しているとみなす. 本研究では, 第一步として,  $w_{i,j,t}(\ell)$  とコサイン類似度が高い画像を抽出し, それら全体によって検出したアノマリーイベントを解釈する.

### 3 実験

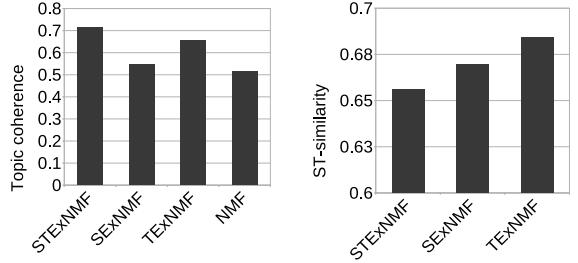
#### 3.1 実験データ

時空間アノマリーイベント検出法の性能評価を行うため, 実データを用いた実験を行った. 日本でも有数の観光地である京都市を対象とし, 写真共有サイト Flickr から Geo-tag が付与された写真データを収集した. このデータセットをここでは京都データと呼ぶ.

#### 3.2 実験設定

実験では, Flickr における 2013 年 3 月 4 日から 2013 年 5 月 26 日までの 12 周間分のデータを対象とした. 写真の総数は 25,847 枚であった. 京都データにおける週ごとの写真数の変遷を図 2 に示す. 本研究では, マップタイルのズームレベルを 13 に設定した.

Shin ら [7] が, 抽出する潜在トピック数  $k = 2$  として文書を対象とした STExNMF 適用した際, 一方のトピックに時空間アノマリーイベントが含まれる傾向があった. また, Shin らはパラメータ  $\alpha$  に関して,  $\alpha$  を変動させて STExNMF の性能比較を行い,  $\alpha = 0.9$  のとき優れた



(a) Topic coherence

(b) ST-similarity

図 3 性能評価

性能を発揮することを示した. 本研究における, 画像を BoVW で表現して STExNMF を適用した手法を用いても同様の効果が得られるかを確かめるため, Shin らと同様に抽出する潜在トピック数  $k = 2$ , パラメータ  $\alpha = 0.9$  とした. そして, タイル  $(i, j, t)$  に対する近傍タイルについて, 空間的な近傍タイルを  $(i \pm 1, j \pm 1, t)$  の 8 タイルと定義し, 時間的な近傍タイルを  $(i, j, t-1), \dots, (i, j, t-4)$  の 4 タイルと定義した. これらを合わせた 12 タイルをタイル  $(i, j, t)$  に対する近傍タイルの集合とした.

STExNMF の効果を調べるうえで, 比較法として, NMF (標準的な NMF を適用した場合), SExNMF (隣接するタイルに対する注目するタイルの空間的アノマリーイベントを捉える場合), TExNMF (過去のタイルに対する注目するタイルの時間的アノマリーイベント場合) と比較を行う.

#### 3.3 性能評価

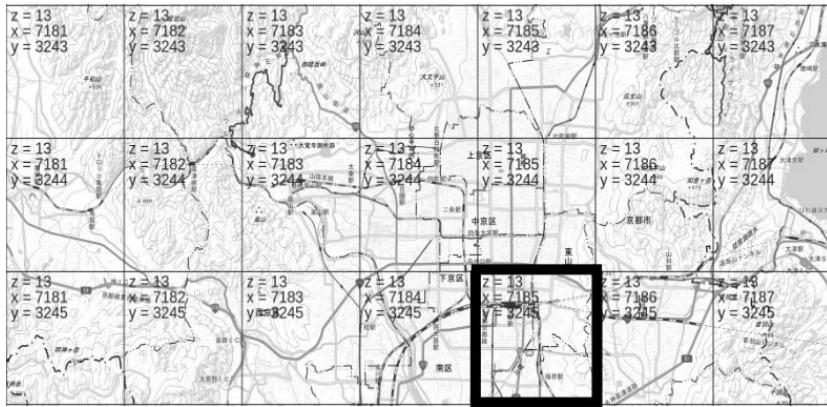
STExNMF および三つの比較法に対して以下に示す 2 つの指標に基づいて時空間アノマリーイベント検出法の性能評価を行った.

##### 3.3.1 トピックコヒーレンス

自己相互情報量 (PMI) に基づいたトピックコヒーレンスにより, 抽出した個々のトピックの性質を評価する. ただし, トピックコヒーレンスは以下のように定義される. まず, 各タイルに対して, 抽出されたトピックベクトルから, 最も代表的な Visual Word を 10 個抽出する. そして, それらの任意の 2 つの Visual Word  $i$  と  $j$  に対して, その PMI を次のように計算する.

$$\text{PMI}(i, j) = \log \frac{P(i, j)}{P(i)P(j)}$$

ここで,  $P(i, j)$  はそのタイル内の同一写真画像における  $i$  と  $j$  の共起確率を表し,  $P(i)$  (または  $P(j)$ ) は  $i$  (または  $j$ ) のそのタイル内の写真画像における発生確率を表す. 次に, それらの平均値を計算することで, 各タイルの PMI スコアとする. 各手法のトピックコヒーレンスは, PMI スコアを全てのタイルで平均することにより計算される. トピックコヒーレンスに関する性能評価の結



(a) 京都市周辺のマップタイル（ズームレベル 13）



(b) 京都駅から稻荷駅を含むマップタイル

図 4 マップタイルの例



(a) トピック 1



(b) トピック 2

図 5 STExNMF で検出されたトピック

果を図 3(a) に示す。STExNMF が最も性能が高かった。

### 3.3.2 時空間類似スコア

時空間類似スコア (ST-sim score) は、タイル  $(i, j, t)$  から抽出したトピックの近傍タイルに対する独自性を測定する。タイル  $(i, j, t)$  の ST-sim score を

$$\text{sim}_{ST}((i, j, t)) = \frac{1}{|\mathcal{U}_{i,j,t}|} \sum_{g \in \mathcal{U}_{i,j,t}} \|(W''_{i,j,t})^T W_g\|_{1,1}$$

と定義する。ST-sim score はそのスコアが低いほど独自性の高いトピックに対応する。時空間類似スコアに関する性能評価の結果を図 3(b) に示す。この指標においても、STExNMF が最も性能が高いことが分かる。

## 3.4 分析結果

京都市のズームレベル 13 におけるマップタイルの例を図 4 に示す。図 4(a) は京都市周辺のマップタイルで

あり、図 4(b) は特に写真数が多かった図 4(a) の黒枠のタイルを拡大したものである。図 4(b) の左上に京都駅があり、右下に伏見稻荷大社から近い稻荷駅がある。また、このタイルの中央には京都市で有名な鴨川も含む区域となっている。本研究では、図 4(b) のタイルに焦点を当てた分析結果を報告する。

抽出されたトピックの説明を行ううえで、各トピックにおいて尤度の高い上位 20 の写真画像のうち、比較する 4 つの方法の結果どうしで異なる特徴を持つ代表写真を抽出し、それらの写真群に基づいてトピックの説明を試みる。

図 5 は、STExNMF を用いた場合の結果である。図 5(a) のトピック 1 では、特長的な絵馬の写真が目立っていた。この絵馬の写真は、伏見稻荷大社で有名なキツネと鳥居の絵馬であり、比較法のいずれのトピックでもこれらの



(a) トピック 1



(b) トピック 2

図 6 NMF で検出されたトピック



(a) トピック 1



(b) トピック 2

図 7 SExNMF で検出されたトピック



(a) トピック 1



(b) トピック 2

図 8 TExNMF で検出されたトピック

写真が上位 20 の写真群に含まれてはいなかった。また、この SExNMF のトピック 1 では、昼間の桜や夜桜の撮影が多く、注目するタイルの隣接空間や過去に対して異なる撮影動向を捉えている可能性がある。次に、図 5(b) のトピック 2 では、京都駅の前にある昼間や夜の京都タワーや、京都駅、鴨川、伏見稻荷大社の千本鳥居の撮影が多く、日常的に撮影され得る写真が多く含まれていた。これらの結果から、写真画像を用いた SExNMF

は、文書を用いたものと同様に、一方のトピックで時空間アノマリーイベントを捉え得る傾向があることが示唆された。

一方、図 6 は、標準的な NMF を用いた結果である。図 6(a) のトピック 1 では、京都駅前の京都タワーや京都駅を捉える写真が多くかった。図 6(b) のトピック 2 では、伏見稻荷大社の千本鳥居を撮影する写真が多くかった。このように、標準的な NMF では、京都駅を撮影するト

ピックと伏見稻荷大社の特に千本鳥居を撮影するいずれも日常的なトピックに大きく分かれる結果となった。また、図 7 は、SExNMF の結果である。図 7(a) のトピック 1 では、標準的な NMF 法による図 6(a) のトピック 1 と同様に京都タワーや京都駅に関する撮影が多かったが、鴨川を撮影する写真も目立っていた。これは、注目したタイルで日常的に川（鴨川）が注目されているというその地域の特色を捉えている可能性がある。次に、図 7(b) のトピック 2 では、桜や京都駅周辺が撮影される写真が含まれていた。これも、隣接するタイルに対して、この時季に桜の写真を撮るという地域の特色を捉えている可能性がある。最後に、図 8 は、TExNMF の結果である。図 8(a) のトピック 1 では、同様に、京都タワーや京都駅に関する撮影が多かった。また、図 8(b) のトピック 2 ではライトアップされた桜が多い傾向があった。図 7(b) と異なり、図 8(b) のトピック 2 では撮影された桜がライトアップされたもののみであった。すなわち、このタイルでは、空間的に隣接するタイルや過去のタイルに対してライトアップされた桜が時空間アノマリーイベントとして捉えられている可能性がある。

以上の比較法の結果を踏まえ、STExNMF の結果と比較した場合、ライトアップされた桜が顕著に時空間アノマリーイベントとして検出される傾向を示すことが予想され、その点において、STExNMF のトピック 1 (図 5(a)) では、その傾向を示しているものの、比較法では捉えていなかった、注目した区域の特色として伏見稻荷大社の独特な絵馬が人々によって注目されたことを捉えている可能性がある。これは、隣接するタイルと過去すなわち空間と時間という二つの観点を同時に捉えることで得られた STExNMF 法の効果であることが期待されるが、その要因については詳細に調べる必要がある。これらの結果は、写真画像を用いた時空間イベントの検出において、地域特有の特徴を分析することができる可能性を示唆しているものと考えられる。

#### 4 まとめ

本研究では、地図システムで普及しているタイルマップとして区画化された地域群に対して、時空間情報付きの文書情報を用いてアノマリーイベントを検出する STExNMF を時空間情報付きの写真における画像情報を用いる方法として拡張した。そして、京都市を対象として収集した Flickr からの大量の写真データを用い、抽出されたトピック群に関する時空間イベント検出結果についてそれらのトピックの代表的な画像群を用いて各地域の特色を捉え得るかについて説明を試み、写真データを用いた時空間イベントの検出を行うアプローチにおける第一歩として、その可能性を示した。写真データを用

いた時空間イベントの検出は地域開発につながると期待している。

#### 参考文献

- [1] Schubert, E., Weiler, M., and Kriegel, H., "SPOTHOT: Scalable Detection of Geo-spatial Events in Large Textual Streams," in *Proceedings of SSDBM'16*, 2016, pp.8:1–8:12.
- [2] Huang, Y., Li, Y., and Shan, J., "Spatial-Temporal Event Detection from Geo-Tagged Tweets," *ISPRS International Journal of Geo-Information*, 2018, 7(4), 150; pp. 1–21.
- [3] Thom, D., Bosch, H., Koch, S., Worner, M., and Thomas, Ertl., "Spatiotemporal Anomaly Detection through Visual Analysis of Geolocated Twitter Messages," in *Proceedings of PACIFICVIS'12*, 2012, pp. 41–48.
- [4] Chae, J., Thom, D., Bosch, H., Jang, Y., and Maciejewski, R., "Spatiotemporal Social Media Analytics for Abnormal Event Detection and Examination using Seasonal-Trend Decomposition," in *Proceedings of VAST'12*, 2012, pp. 143–152.
- [5] Cheng, T. and Wicks, T., "Event Detection using Twitter: A Spatio-Temporal Approach," *PROLSE ONE*, 2014, 9(6): e97807.
- [6] Witayangkum, A., Horanont, T., Sekimoto, Y., and Shibasaki, R., "Anomalous Event Detection on Large Scale GPS Data from Mobile Phones Using Hidden Markov Model and Cloud Platform," in *Proceedings of UbiComp'13*, Workshop on Pervasive Urban Applications, 2013, pp. 1219–1227.
- [7] Shin, S., Choi, M., Choi, J., Langevin, S., Bethune, C., Horne, P., Kronenfeld, N., Kannan, R., Barry, D., Park, H., and Choo, J., "STExNMF: Spatio-Temporally Exclusive Topic Discovery for Anomalous Event Detection," in *Proceedings of ICDM'17*, 2017, pp. 435–444.
- [8] 秋山 祐樹, “ビッグデータで描く地域の現在と未来の姿 (特集:位置情報から未来社会を構想する)”, 日本地域開発センター, 地域開発 620, 17-21, 2017-06.
- [9] Chen, S., Yuan, X., Wang, Z., Guo, C., Liang, J., Wang, Z., Zhang, X.L., and Zhang, J., "Interactive visual discovering of movement patterns from sparsely sampled geo-tagged social media data," *IEEE Transactions on Visualization and Computer Graphics*, 2016, vol. 22, no. 1, pp. 270–279.
- [10] Yuan, J., Luo, J., Kautz, H., and Wu, Y., "Mining GPS Traces and Visual Words for Event Classification," in *Proceedings of MIR'08*, 2008, pp. 2–9.
- [11] Bossard, L., Guillaumin, M., and Gool, L. V., "Event Recognition in Photo Collections with a Stopwatch HMM," in *Proceedings of ICCV'13*, 2013, pp. 1193–1200 .
- [12] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, 1999, vol. 401, no. 6755, pp. 788?–791.
- [13] Lowe, D., "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, 2004, vol. 60, no. 2, pp. 91–110.
- [14] Bay, H., Tuytelaars, T., and Van Gool, L., "SURF: Speeded up robust features", in *Proceedings of ECCV'06*, 2006, pp. 404–417.