

# SNS投稿写真の時空間クラスタリングに基づく類似地域マイニングの検討

陳 ル<sup>†</sup> 川西 康友<sup>†</sup> 井手 一郎<sup>†</sup> 平山 高嗣<sup>††</sup> 道満 恵介<sup>†††</sup>

出口 大輔<sup>††††</sup> 村瀬 洋<sup>†</sup>

<sup>†</sup> 名古屋大学 大学院情報学研究科 〒464-8601 愛知県名古屋市千種区不老町

<sup>††</sup> 名古屋大学 未来社会創造機構 〒464-8601 愛知県名古屋市千種区不老町

<sup>†††</sup> 中京大学 工学部 〒470-0393 愛知県豊田市貝津町床立 101

<sup>††††</sup> 名古屋大学 情報連携統括本部 〒464-8601 愛知県名古屋市千種区不老町

あらまし 知らない土地に旅行する際、旅行を円滑に計画するためには、行き先の雰囲気や予め把握することが重要である。ある地域が、既に知っている地域のある時期の雰囲気と似ていることが分かれば、旅行する時期のその地域の雰囲気を直感的に把握できる。これを実現するため、我々は時空間クラスタリングにより SNS に投稿された写真を時期情報付き地域に分割し、SNS 投稿写真の画像内容に基づいて雰囲気が類似する地域の組を検出する手法を検討する。  
キーワード 旅行, SNS 写真, 時空間クラスタリング

## A Study on Similar Geo-region Mining based on Spatio-temporal Clustering of Social Media Photos

Lu CHEN<sup>†</sup>, Yasutomo KAWANISHI<sup>†</sup>, Ichiro IDE<sup>†</sup>, Takatsugu HIRAYAMA<sup>††</sup>, Keisuke

DOMAN<sup>†††</sup>, Daisuke DEGUCHI<sup>††††</sup>, and Hiroshi MURASE<sup>†</sup>

<sup>†</sup> Graduate School of Informatics, Nagoya University

<sup>††</sup> Institute of Innovation for Future Society, Nagoya University

Furo-cho, Chikusa-ku, Nagoya-shi, Aichi, 464-8601 Japan

<sup>†††</sup> School of Engineering, Chukyo University, 101 Tokodachi, Kaizu-cho, Toyota-shi, Aichi, 470-0393, Japan

<sup>††††</sup> Information Strategy Office, Nagoya University

Furo-cho, Chikusa-ku, Nagoya-shi, Aichi, 464-8601 Japan

**Abstract** When travelling to a place that we had never been, it is important to grasp the atmosphere there beforehand to smoothly plan the travel. If a place's atmosphere is similar to somewhere we have already known at a certain season, we can intuitively grasp the atmosphere of that place when we travel. To realize this, we propose a method that divides SNS photos into spatio-temporal region using clustering and detects pairs of similar spatio-temporal regions based on the contents of social media photos.

**Key words** Travel, social media photos, spatio-temporal clustering

### 1. はじめに

休暇の過ごし方として、旅行が人気を博している。近年、旅行中の自由度が高いため、団体旅行より個人旅行を選択する旅行者数が年々増加している。このような個人旅行では、旅行者自身が旅行計画を策定する必要がある。

それを支援するため、観光地を推薦する様々なシステムがある。その多くは、キーワードからジャンルを絞り、ランドマークを中心に観光地を推薦する。ランドマークは明示的に指定す

ることができるが、それとは異なり、言語で明示的に表現しにくい興味対象を調べたい場合、満足な結果を得ることは難しい。このような場合は、旅行者が知っている別の地域を例に、その地域の雰囲気と似ている地域を推薦する方法が考えられる。

滝本ら [1] は年間を通して SNS に投稿された写真に基づく地域間の類似度算出により雰囲気が似ている場所をマイニングする手法を提案した。しかし、時期によって雰囲気が大きく変わる地域も多い。この研究では、写真が撮影された時期の情報が考慮されていないため、特定の時期しか雰囲気が似ていないこ

とを表現することができない。そのため、一般に特定の短期間しか滞在しない旅行者への支援としては不十分である。

このような背景から、本研究は、時期情報を含めた旅行支援の実現をするために、時期を考慮して類似地域をマイニングすることを目的とする。本報告では、滝本ら [1] の手法における地域のクラスタリング手法を拡張し、写真が撮影された時期まで考慮したクラスタリング手法を提案し、その結果から得られた類似地域マイニングの結果について報告する。

## 2. 時空間クラスタリングに基づく類似地域マイニング

### 2.1 時間情報付き地域の獲得

本研究では、旅行者視点の旅行支援を実現するため、大勢の人の興味を引く対象が存在する地理的・時期的に連続した範囲を時期情報付き地域と定義し、時空間情報が付与された SNS 投稿写真を時空間クラスタリングすることで得られるクラスタに属する写真が撮影された場所を含む矩形の領域として求める。

クラスタリング時には、DBSCAN で得られたクラスタのうち、時期方向に長いクラスタに対して、分割型クラスタリングを適用し、写真をより小さいクラスタに分割する。また、地理的範囲が重なり、撮影した時期が近いクラスタには、同じような注目対象が存在する可能性が高いと考えられるため、凝集型クラスタリングを適用し、クラスタ間距離が近い・クラスタ内の画像内容が似ているクラスタを統合する。ただし、少数の人にしか興味を持たれていない（少数のユーザの投稿からなる）クラスタは、対象外として除去する。

### 2.2 時期情報付き地域の特徴記述及び類似地域マイニング

本研究では、滝本ら [1] の手法と同様に、画像の視覚的内容を表す Visual Concept に基づいて地域の特徴を表現する。また、時期情報付き地域間の類似度算出法も滝本ら [1] の手法に倣い、正規化相互相関を用いて地域間の類似度を計算する。

## 3. 類似度算出実験

本節では、SNS 投稿写真の時空間クラスタリングに基づいた提案手法の有効性を確かめるため、実際の SNS 投稿写真を用いて類似地域をマイニングする実験を行なった。

### 3.1 実験条件及び結果

実験では、Flickr からのデータセット YFCC100M [2] のうち日本で撮影された時空間情報付き写真約 91 万枚を使用した。

DBSCAN のパラメータとして  $Eps = 0.01$  と  $MinPts = 100$  (地理的距離およそ 1km, 時間間隔およそ 4 日に相当) を初期設定とし、時間方向に長いクラスタに対して  $MinPts$  を 100 ずつ増やすことにした。

また、地理的範囲が重なり、時間間隔が 1 ヶ月以内のクラスタを対象に、クラスタ間距離として画像内容の余弦距離を用いて凝集型クラスタリングを適用した。ただし、地理的範囲が重ならないときには、クラスタ間距離を  $\infty$  とした。

類似地域のマイニングでは、類似地域判定に用いる正規化相互相関のしきい値を 0.7 とし、それを満たす時期情報付き地域の組を類似地域として抽出した。

表 1: 時期情報付き類似地域マイニング結果の一部

地域 A	時期	地域 B	時期	類似度
岐阜県可児市	5 月下旬	千葉県八千代市	6 月上旬	0.9958
栃木県芳賀郡	9 月下旬	三重県鈴鹿市	10 月上旬	0.9924
福島県いわき市	7 月下旬	大阪府大阪市	8 月中旬	0.9733
...	...	...	...	...



(a) 岐阜県可児市花フェスタ記念公園



(b) 千葉県八千代市大和田新田

図 1: 検出された類似地域の組の例

上記の条件で得られた時期情報付き地域の総数は 633 箇所、それらの総組み合わせ 200,028 組のうち、38,129 組が類似地域として検出された。結果の一部を表 1 に示す。

### 3.2 考察

得られた類似地域の組の例を図 1 に示す。このような場所では、花に注目したイベントが行われるため、その開催時期に自然風景に関する写真が多く、類似地域として検出された。また、実際に調べたところ、図 (a) の花フェスタ記念公園は世界最大級のバラ園であることが分かった。

そのほか、スポーツの開催地など、試合の時期に人が集まり、戦況の報告や参加の記念のために写真を撮影して共有するため、類似地域として検出された。

## 4. まとめ

本報告では、従来手法では表現できない特定の時期だけ似ていることに対し、撮影した時期情報を加えてクラスタリングを行い、ある時期だけ似ている地域の組をマイニングする手法を検討した。実際の投稿写真を用いた実験により、提案手法で花祭りがある地域の組やスポーツ開催地の組など特定の時期だけ類似する地域を検出できることを確認した。

### 文 献

- [1] 滝本広樹, 川西康友, 井手一郎, 平山高嗣, 道満恵介, 出口大輔, 村瀬洋, SNS 投稿写真の視覚的内容に基づく類似地域マイニング, 信学技報 MVE2017-100, Mar. 2018.
- [2] B. Thomee, D.A. Shamma, G. Friedland, B. Elizalde, K. Ni, D. Poland, D. Borth and L. J. Li, YFCC100M: The new data in multimedia research, Comm. ACM, Vol.59, No.2, pp.64-73, Feb. 2016.