

Geo-tag 付き写真群に基づいた観光経路アノマリー検出

岡田 洸一†

熊野 雅仁‡

木村 昌弘‡

† 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

1 はじめに

近年、撮影位置・撮影時間・撮影者情報などメタ情報が付与された大量の写真群から主要撮影スポットや主要経路を抽出し、観光へ応用する研究が注目されている。写真は魅力的な対象に出くわしたときに撮影される傾向があるため、多くの人々の撮影行動をマイニングすることで、様々な魅力を再発見できる可能性がある。Crandallらは大量のGeo-tag付写真の位置情報、撮影時間、撮影者IDから撮影者の移動軌跡が求められることを示し[1]、Luらは撮影者達の移動軌跡を合成して観光経路推薦に応用した[2]。本研究では、年ごとに抽出された主要撮影スポット間を移動する人々の行動に焦点当て、人々が日頃とは逸脱した経路を移動した異常な日を検出することを考える。その異常性の度合を定量化およびランキングすることで、各主要撮影スポットで何が起きたかを異常度の観点から分析する手法へ応用可能な、PCAに基づく観光経路アノマリー検出法を提案する。観光地では桜や紅葉などの旬シーズンに人々の移動経路が変化し得るため、撮影スポットごとに旬シーズンとの関係や、魅力的な観光経路との関係を調べ、行動分析への応用可能性を検証する。

2 観光経路アノマリー

Crandallらは大量のGeo-tag付写真の撮影位置に関する密集度を用いて主に年中人気を得る定番のスポットを抽出できることを示した。しかし、人気スポットを独立に捉えても一度に複数の地域を訪れる定番として存在し得る主要な観光経路は捉えられない。Crandallらは位置情報、撮影時間、撮影者IDから各撮影者の移動軌跡が求められることも示したが、日本の四季のように旬シーズンを持つ地域では、人々の行動が、一年を通じて定番となる主要観光経路以外に四季それぞれにも変化し得る。ただし、四季以外にも、様々な時期に人々の行動が変化し得る可能性がある。本研究では、人々の主要な移動経路とは逸脱した行動を捉える、すなわち観光経路アノマリーを検出することで、観光地の主要な魅力だけでなく、ある時期に特化したさらなる魅力や起きた出来事を異常性の観点から浮き彫りにできる可能性があると考え、その検出法に着目する。

Detecting travel-route anomalies based on a large number of geo-tagged photos

Kouichi OKADA†, Masahito KUMANO‡ and Masahiro KIMURA‡

†Division of Electronics and Informatics, Ryukoku University

‡Department of Electronics and Informatics, Ryukoku University

*http://www.epinions.com **http://konect.uni-koblenz.de

3 提案法

3.1 主要撮影スポット

人気スポット間の移動に着目して観光経路アノマリーを検出する準備として、位置情報、撮影時間、撮影者IDの付随した写真群から撮影位置の密集度に基づき、撮影スポット群を抽出する[3]。次に、主要な撮影スポットに着目するため、全撮影スポットの平均訪問人数 μ_D とその標準偏差 σ_D を求め、 $\mu_D + \sigma_D$ より訪問者数が多い撮影スポット群を主要撮影スポット群 $\{D_k\}_{k=1}^K$ とする。

3.2 観光経路の抽出

本研究では、主要撮影スポット群 $\{D_k\}_{k=1}^K$ における週ごとの人々の移動 $R_k(t) = [R_{k,0}(t), R_{k,1}(t), \dots, R_{k,K}(t)]$ を調べる。ここに、 $R_{k,\ell}(t)$ は第 t 週に D_k から D_ℓ へ移動した人数であり、 $t = 1, \dots, 53$, $\ell = 1, \dots, K$ である。ただし、 $R_{k,0}(t)$ は第 t 週に D_k から $\{D_\ell\}_{\ell=1}^K$ 以外の(主要ではない)撮影スポットに移動した人数であり、 $R_{k,k}(t)$ は、第 t 週にスポット D_k の中では移動したが他の撮影スポットには移動しなかった人数を表す。

3.3 観光経路アノマリーの検出

各主要撮影スポット D_k において $R_k(t)$ が与えられたとき、 t 週目の行動が、定番の行動からどの程度逸脱しているかを表す手法として、PCA (Principal Component Analysis) に基づいた方法を用いる。 $(K+1)$ 次元空間において $\{R_k(t) | t = 1, \dots, 53\}$ を主成分分析し、 $R_k(t)$ の第 j 主成分 $S(D_k, j, t)$ を計算する。 $S(D_k, j, t)$ は、第 t 週の D_k からの人々の移動パターンが平均的な D_k からの移動パターンと、第 j 主成分方向 $e^{(j)} = (e_0^{(j)}, e_1^{(j)}, \dots, e_K^{(j)})$ においてどのくらい離れているかを表していることに注意する。我々は、上位の j に対して、 $S(D_k, j, t)$ を D_k における第 t 週の j -異常度と定義する。ここに、 $S(D_k, j, t)$ は正の値も負の値もつことに注意する。我々は、 j -異常度が高い週として、 $|S(D_k, j, t)|$ の値が大きい週 $t^*(k, j)$ を検出する。このとき、 $|e_\ell^{(j)}|$ の値が大きい $D_{\ell^*(j)}$ に対して、第 $t^*(k, j)$ 週における D_k から $D_{\ell^*(j)}$ への移動人数は、通常の D_k から $D_{\ell^*(j)}$ への移動人数と大幅に異なっていることに注意する。我々は、第 $t^*(k, j)$ 週における D_k から $D_{\ell^*(j)}$ への移動を、 D_k における j -観光経路アノマリーとして検出する。

4 実験

主要な撮影スポット群を抽出して独立に捉える観点や、主要な行動を捉えるだけでは捉えきれない観光行

動を、異常な行動に着目することで浮き彫りにできるかについて検証を行うため、観光名所として、季節に応じて人々の行動が変化し得るとされる日本の京都府と、年中人々の行動が変化しにくい傾向を持つと思われるアメリカのマンハッタンを対象に実験を行った。

4.1 実験設定

写真共有サイト Flickr から、京都とマンハッタン（ニューヨーク）について 2012 年 1 月 1 日から 2012 年 12 月 31 日までの位置情報、撮影時間、撮影者 ID が付随する写真データ群を収集し、実験に使用した。使用データの詳細を表 1 に示す。これらの写真データに対し、主要撮影スポットを抽出したところ、京都データでは $K=59$ 、マンハッタンデータでは $K=46$ となった。次に、 $j = 1, 2, 3$ に対して、各 $S(D_k, j, t)$ を求め、正側で降順に、負側で昇順にランキングを行い、上位 5 位に着目した。注目する上位の週と旬シーズンとの関連性を検証するため、 D_k から旬シーズン抽出を行い、旬の度が高い上位の旬シーズンとの関係を調べた。

表 1: 実験データ

撮影地域	京都	マンハッタン
撮影枚数	52398 枚	287659 枚
ユーザー数	5280 人	55030 人

4.2 旬シーズン抽出

旬シーズンの抽出については、我々が提案した旬シーズン抽出法 [3] を用いる。この旬シーズン抽出法は、より短い期間でより多くの人々が訪問者するほど旬の度合いが高い（バースト性が高い）シーズンとし、各主要撮影スポットから複数の旬シーズンを抽出する手法である。実験では、シーズンを提案法で検出された週の出発スポットと行き先スポットが上位の旬シーズンに属しているかを調べる。

4.3 実験結果

スペースの都合から、京都データに関しては、主要撮影スポット清水寺の結果についてのみ述べる。 $j = 1$ の場合に検出された j -観光経路アノマリーは、4 月 1 日から 4 月 7 日の週における清水寺から清水寺前通り間の移動であった。この週は桜のシーズンであり、この経路は清水寺を来訪する人々の多くが通る道である。一方、 $j = 3$ の場合に検出された j -観光経路アノマリーは、11 月 25 日から 12 月 1 日の週における清水寺と近隣の高台寺間の移動であった。この時期は、清水寺と高台寺が共に紅葉と夜間のライトアップで人気を得ており、同様の魅力を持つ近隣スポット間の移動を異常性の観点から捉えた好例であると思われる。また、旬シーズン抽出法により検出された清水寺の旬シーズンは 10 月 30 日から 12 月 31 日であり、提案法が検出した 3-異常度が高い週を含んでいた。提案法は、清水寺

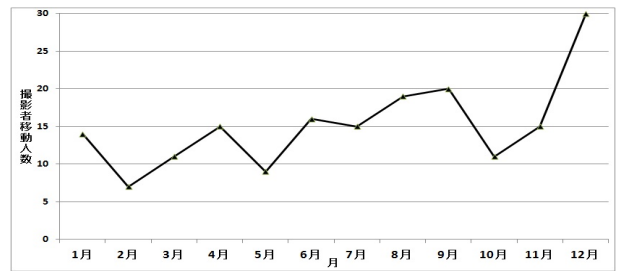


図 1: 移動人数の月別推移

を独立に捉える旬シーズン抽出法では検出しにくい高台寺との関係をも捉えたことに注意する。

次に、マンハッタンデータに関しては、年間を通じて最も撮影数が多かった主要撮影スポットであるワントゥタイムズスクエアの結果についてのみ述べる。 $j = 1$ の場合に検出された j -観光経路アノマリーは、9 月 5 日から 9 月 12 日の週でのワントゥタイムズスクエアから主要撮影スポット以外への移動であった。 $j = 3$ の場合に検出された j -観光経路アノマリーは、12 月 9 日から 12 月 15 日の週でのワントゥタイムズスクエアからロックフェラービル前への移動であった。図 1 に、ワントゥタイムズスクエアからロックフェラービル前への移動人数の年間推移を示す。12 月にその移動人数が急激に増加しており、3-観光経路アノマリーはこの変化を捉えていると考えられる。提案法はこの動きの変化を捉えていると考えられる。また、旬シーズン抽出法により検出されたロックフェラービル前の第 2 位の旬シーズンは 9 月 9 日から 12 月 31 日であり、ワントゥタイムズスクエアにおける 3-異常度が高い週を含んでいたことに注意する。

5 まとめ

本研究では、Geo-tag 付写真を撮影する人々の行動を通じて、日頃から逸脱した行動の異常性、つまり観光経路アノマリーを検出することで、主要撮影スポット単独や、主要な移動経路に着目するだけでは捉えにくい特定の時期に存在し得る観光地の魅力を行動異常性の観点から検出できる可能性を示した。今後は、分析を進め、より検出性能の高い手法を探索する予定である。

参考文献

- [1] D.Crandall, L.Backstrom, D.Huttenlocher, and J.Kleinberg: Mapping the World's Photos, WWW'09.
- [2] X.Lu, C.Wang, J.M.Yang, Y.Pang, and L.Zhang: Photo2trip: Generating travel routes from geo-tagged photos for trip planning. ACM Multimedia'10.
- [3] 熊野, 岩淵, 小関, 小野, 木村: 集合知に基づいたポピュラー撮影スポットに関する旬シーズンの可視化, 芸術科学会論文誌, Vol.13, No.4, pp.218-228, 2014.