

第IV部門

ジオタグ付き写真を用いた関西地方の観光者の選好と行動の分析

京都大学工学部

学生会員 ○江端 哲矢

京都大学大学院工学研究科

正会員 木村 優介

1. はじめに

近年、日本を訪れる外国人観光客が急増し、観光地の混雑が問題となっている。混雑の対策として、来訪者は少ないものの、有名な観光地と似たような魅力がある別の観光地や観光経路に観光客を誘導することが考えられる。現在、「歴史・文化に興味あり／なし」「大人向け／子供向け」などの選択肢の一対比較に基づいて観光経路を推薦するサービスがある¹⁾。しかし観光者自身が適切な重みを選択するのは難しい。観光者が過去に撮影した写真をもとに観光経路を決めることのできるような仕組みがあれば、自身の選好の理解と、その結果に応じた観光経路の提案が可能と考えられる。

本研究の目的は、複数の観光地を周遊する観光者を写真により把握される選好に応じて分類し、クラスタごとに特徴的な観光行動を把握することである。そのために、まず関西地方を周遊する観光者が撮影したジオタグ付きの（位置情報を有する）写真を写真共有サービスから収集し、それらを被写体に応じて分類する。次に写真の分類をもとに、観光者の選好を把握・分類し、GIS（地理情報システム）により分析する。関西地方には多くの外国人観光客が訪れており、京都・大阪・奈良などの観光地が分散して立地している。本研究では北は天橋立、南は潮岬、西は姫路、東は長浜を含んだ地域を対象とする。

2. 方法

本研究では、画像共有サービスの一つである Flickr を用いて観光者の選好を写真の被写体から分析した。まず撮影者を観光者と在住者に分類するにあたり、16ヶ所の代表的な観光地（東山、南禅寺・銀閣寺、河原町、二条城、金閣寺、嵐山、伏見稲荷、宇治、奈良、大阪城、USJ、難波、関西空港、高野山、神戸、姫路）で写真を収集した。各撮影者について写真が最初に撮影された日時と最後に撮影された日時に着目し、それが 21 日を超える撮影者を在住者とみなし削除するとともに、観光地のうち一ヶ所ではしか写真を撮影していない撮影者を削除した。残った撮影者を観光者とみな

し、関西地方全体で 572 人の撮影者による 45897 枚の写真を対象とした。

これらの写真を Google Cloud Platform の Vision API のラベル検出機能を用いて分析した。各写真について最大 20 個のラベルとその信頼度（検出されたラベルの確からしさ、0 から 1 の実数）が付加された。検出されたラベル 5689 種類のうち、出現回数が 20 回以下であるラベルを取り除いた。残った 1350 種類のラベルについて写真 1 枚ごとにラベルの出現データを作成し、変数の数を縮約するために主成分分析を行った。

続いて、主成分分析の結果をもとに k-means 法を用いてクラスタ分析を行った。Gap 統計量により最適なクラスタ数は 21 となった。実際に 21 前後のクラスタ数で試してみたところ、クラスタ数を 22 にすると「像」のクラスタが現れたが、より多くしても既存の要素が分割されるだけで、新たな要素は確認できなかった。よって 22 個のクラスタで分類を行った。これらのクラスタを確認し、内容が類似しているクラスタを統合して、15 種類のクラスタに写真を分類した。最後に、写真のクラスタの構成比をもとに撮影者を k-means 法を用いて 8 種類のクラスタに分類した。

3. 写真の分類と観光者の選好の推定の結果

主成分分析の結果について、第 1 主成分：建造物／自然、第 2 主成分：近代的／歴史的、第 3 主成分：観光／日常、第 4 主成分：花・植物／水・空、第 5 主成分：近景／遠景の軸が得られた。これ以降の軸に関しては明確な軸が得られなかったが、第 5 主成分までの累積寄与率は 28.9%と低いため、累積寄与率が 70%に達する第 71 主成分までを使用した。写真の分類について、クラスタ分析の結果、表 1 のように分類された。建造物や歴史的建造物、植物などを示す 15 のクラスタに分類された。

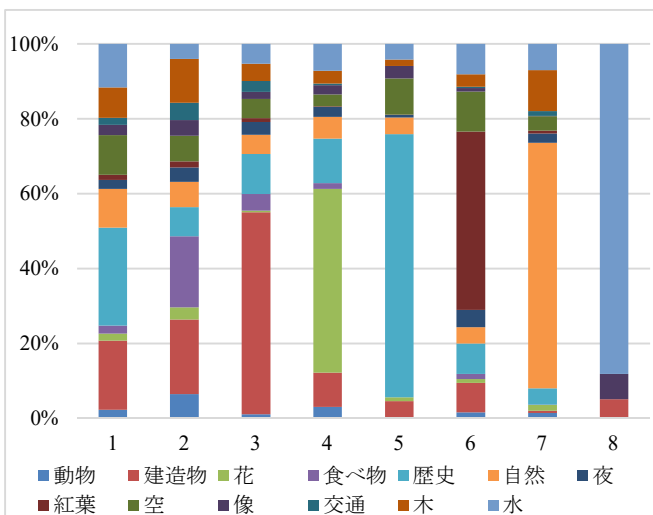
観光者の分類については、表 2 のように分類された。桜や紅葉など、ある特定のクラスタの写真を多く撮影する観光者もいる（クラスタ 3 から 8）一方、各クラスタの写真をバランスよく撮影している観光者（クラ

表 1：写真の分類の結果

クラス	写真数	頻出ラベル
建造物	9033	Architecture, Building, Urban area
歴史	5788	Temple, Architecture, Chinese architecture
木	2833	Tree, Plant, Botany
空	2572	Architecture, Tree, Sky
自然	2397	Tree, Sky, Vegetation
人物	2309	Tourism, People, Tree
水	2256	Water, Nature, Water resources
食べ物	2055	Food, Cuisine, Dish
紅葉	1686	Tree, Leaf, Autumn
花	1598	Flower, Spring, Tree
夜	985	Night, Light, Lighting
交通	922	Transport, Mode of transport, Train
動物	857	Wildlife, Mammal, Vertebrate
像	730	Statue, Sculpture, Art
その他	9876	Vehicle, Building, Font

表 2：観光者の分類の結果

クラス	説明	観光者合計	写真合計
1	Balanced 1	177	12456
2	Balanced 2	162	8366
3	Building	124	6492
4	Flower	38	1208
5	History	31	217
6	Red leaf	24	2048
7	Nature	10	62
8	Water	6	13

図 1：各観光者クラスターの写真分類の構成比
スタ 1, 2) が多い。

4. クラスターごとの観光行動の把握と考察

撮影場所と観光者の選好について、GIS 上でヒートマップを作成して考察する。まず関西地方全体での撮影場所の分布をみると、以下の点が把握できる。桜や紅葉を多く撮影している人はそれらで知られている観光地での撮影が多い。／建造物の写真が多いクラスターは大阪での写真が多い。／様々な種類の写真をバランスよく撮影しているクラスターについては主要な観光地で撮影しているものが多く、明らかな分布の違いは見られない。

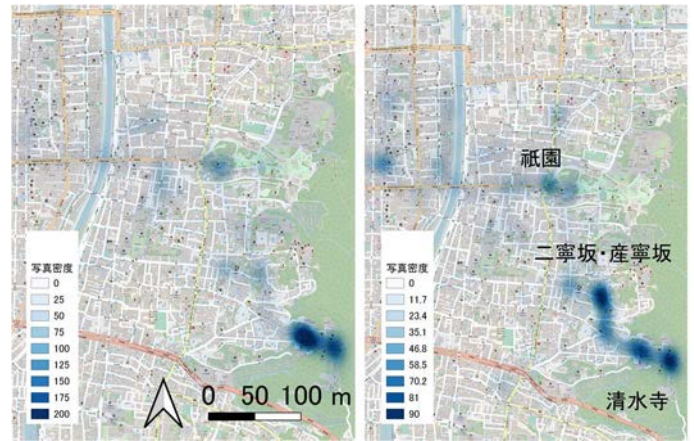


図 2：京都市東山地域のヒートマップ（左：クラスター 1、右：クラスター 2）

次に、観光地周辺での分布に着目する。クラスター 1 と 2 について、東山地域に着目して撮影場所の分布をみると（図 2）、クラスター 1 は清水寺で多く撮影しているのに対して、クラスター 2 は清水寺だけではなく二寧坂や産寧坂、祇園などでも多く撮影している。写真の分類の分布について観察してみたが、大きな違いは見られなかった（図 1）。またほかの地域に関してもクラスター 1 はクラスター 2 に比べてピークがあまり多く観測されなかった。このことからクラスター 2 の方はより広い物事に対し興味を示すクラスターであり、観光行動に違いが生じていると考えられる。

5. 結論

本研究では、関西地方で撮影されるジオタグ付き写真を写真共有サービスから取得して解析することで、観光者がどのように行動しているかを分析した。いずれのクラスターにおいても、同じ観光地では写真の内容の違いはあまり見られなかった一方、クラスターごとに写真が多く撮影されている場所が異なっていることが明らかになった。写真の分類の結果、観光者の行動や、各観光地の性質を示すことができ、観光者の分類の結果、観光経路を推薦するためのパラメータを推測できる可能性が示された。

参考文献

- 1) 京都市観光協会：京都観光 Navi, AI 自動モデルコース& ルート案内, <https://ja.kyoto.travel/concierge/>, アクセス日：2020 年 2 月 23 日。
- 2) 石川眞梨子, 木村優介：写真共有サービスを用いた来街者の選好と観光行動の時空間的特性に関する研究, 景観・デザイン研究講演集, No.15, pp.372-379, 2019.
- 3) 倉田陽平, 真田風, 鈴木祥平, 石川博：Flickr と Google Cloud Vision API によりテーマ別観光マップを作る試み, 第 9 回データ工学と情報マネジメントに関するフォーラム, p7-4, 2017.