

# 卒業論文概要書

## Summary of Bachelor's Thesis

学科名 Department	情報通信	氏名 Name	木村凌太	指導員 Advisor	渡辺 裕 印 Seal
研究指導 Research guidance	オーディオビジュアル情報処理研究	学籍番号 Student ID number	1W143035-1 CD		
研究題目 Title	写真共有サービスを用いた旅行ルート推薦の基礎検討 A study on travel route recommendation with photo sharing service				

### 1. まえがき

近年、技術の進歩により、ユーザは容易に情報を発信することができ、他者が公開している様々な情報を簡単に得ることができる。また、位置情報を用いた新たなサービスも進歩している。しかし、この変化により、ユーザが処理可能な量を遥かに上回る情報が流通するようになり、真に必要な情報の選別を困難にしている。この影響は旅行にも見られ、旅行先を選ぶ際、有名な観光名所や特定の場所のみが選択の対象とされている。そのため、世間では知られていないような価値ある観光地の情報を得ることが困難になりつつある。そこで本研究では居住者と旅行者の地理的条件の違い[1]を利用し、世間的に馴染みのない価値ある観光地を穴場スポットとして検出する。また、その穴場スポットに関して観光地として優れているかを被験者に対して五段階評価のアンケートを行い、本研究の妥当性を検証する。

### 2. 穴場スポットの検出

写真共有サイト Flickr[2]を用いて、対象とする観光地の居住者とそこを訪れる旅行者の写真データを収集する。この時、得られた写真の位置情報を局所外れ値因子法(LOF法)[3]を用いて分類し、穴場スポットの検出する。

### 3. 実験結果評価

実験では、Flickrを用いて得られた写真に対して、局所外れ値因子法を用いることで正しく穴場スポットの写真検出を行えた。また、得られた写真に対してアンケートを行った結果、実験で得られた穴場スポットが旅行者の訪れる観光地と比較しても魅力的であることを確認した。

### 4. まとめ

本研究では、居住者と旅行者の写真の位置情報から局所外れ値因子法を用いて穴場スポットの検出を行い、アンケートにより評価した。アンケートの結果より穴場スポットとして定義した観光地は旅行者が好む有名な観光地と比較しても引けをとらない魅力があることを明らかになった。旅行ルートの推薦に関して、旅行者と居住者の撮影する写真の位置情報を利用し、視覚的に評価を行う本研究の手法の有効性を示した。

### 参考文献

[1]Gigazine, 世界各地の都市の観光客と地元民の行動範囲の違いがわかるマップ(2010年 6月 10日)

[https://gigazine.net/news/20100610\\_locals\\_and\\_tourists/](https://gigazine.net/news/20100610_locals_and_tourists/)

[2]Flickr, <https://www.flickr.com/>

[3] Markus M. Breuning and Hans-Peter Kriegel and Raymond T. Ng and Jörg Sander. LOF: Identifying Density-Based Local Outliers. In ACM SIGMOD 2000, pp. 93–104, 2000.

2017 年度 卒業論文

写真共有サービスを用いた  
旅行ルート推薦の基礎検討

A Study on Travel Route Recommendation  
with Photo Sharing Service

指導教員 渡辺 裕 教授

早稲田大学 基幹理工学部

情報通信学科

1W143035-01

木村凌太

## 目次

第 1 章	序論 .....	3
1.1	研究の背景と目的 .....	3
1.2	関連研究 .....	3
1.3	本論文の構成 .....	3
第 2 章	ルート推薦に関する要素技術 .....	5
2.1	はじめに .....	5
2.2	研究意図 .....	5
2.3	Flickr [7].....	5
2.3.1	Flickr API[8] .....	6
2.3.2	Flickr から大量のデータを取ってくるツール[9].....	6
2.4	Survey monkey[10].....	6
第 3 章	提案手法 .....	7
3.1	はじめに .....	7
3.2	提案手法の概要 .....	7
3.2.1	穴場スポット検出アルゴリズム .....	7
3.3	フローチャート .....	8
3.4	むすび .....	10
第 4 章	提案手法の評価実験と結果, 考察 .....	11
4.1	はじめに .....	11
4.2	実験の概要 .....	11
4.3	実験 1 (穴場スポット検出実験) .....	11
4.3.1	実験 1 の条件 .....	11
4.3.2	実験 1 の結果 .....	11
4.3.3	実験 1 の考察 .....	13
4.4	実験 2 (アンケートを用いた穴場スポットの評価実験) .....	13
4.4.1	実験 2 の条件 .....	13
4.4.2	実験 2 の結果 .....	13
4.4.3	実験 2 の考察 .....	13
4.5	まとめ .....	14
第 5 章	結論と今後の課題 .....	15
5.1	結論 .....	15
5.2	今後の課題 .....	15
謝辞	.....	16

参考文献 .....	17
図一覧 .....	18
表一覧 .....	19

# 第1章 序論

## 1.1 研究の背景と目的

近年、ブログや SNS 等のサービスの発達により、誰もが容易に他者の情報を得ることができるようになった。また、中には写真データに付随する位置情報も多く存在するため Twitter を始めとした位置情報に基づく新たなサービスや研究も行われている。このような新しい技術は多方面に影響を及ぼしており、その一つの例として旅行が挙げられる。一昔前までは旅行情報誌や旅行会社でのツアーが主流であったが、近年では SNS から入手した情報を元に旅行する等、様々な形態で旅行を楽しむ人が増えている。その他にもトリバゴやエクスペディア等の旅行サービスも充実してきている。[1][2]

しかし、選択肢の増加は情報過多を生み出し、人々が自分に有益な情報を効率的に取捨選択することを妨げている。また、充実しているように思えるこれらの旅行サービスだが、得られる情報には知名度の高い観光地の情報が多いという偏りが見られる。そのため、本研究では個人の趣味趣向を考慮した旅行ルートのおすすめに関する基礎検討を行った。

## 1.2 関連研究

Web 上に投稿される大量の写真には GPS 情報を有するものが存在する。王ら[3]は写真共有サイト Panoramio を用いて、大量の写真を風景によって分類し、人々が注目した風景の写真を地域の情報として地図上に表現する方法を検討した。

このようにして作成された観光マップは撮影した風景ごとにカテゴリに分類され、地図上にアイコンとして表示される。また、GPS 情報を有しているため、風景ごとの距離感が地図上で視覚的にわかる。そのため、ユーザにとっては旅行の計画時に実際に現地を訪れた際により具体的なイメージ化が可能になる。また、異なるアプローチ方法として、坂元ら[4]はユーザの気分に応じた旅行ルートのおすすめ手法を提案している。

SNS 上で投稿されるテキストの中から感情に関わる単語と限定した地域の観光地を訪れた際の感想とを比較することで、ユーザの現在の感情に対して最もふさわしい観光地を推薦している。

これらの手法は web 上に投稿される観光地のみを対象としているので世間的に馴染みのない観光地の推薦ができない問題点がある。そこで、本論文では、位置情報を利用することにより、馴染みのない観光地を検出し、推薦できる手法を提案する。LOF 法を利用することにより、提案方式の妥当性を評価する。

## 1.3 本論文の構成

本論文の構成は以下のとおりである。

第 1 章は本章であり、本研究の背景や目的について述べる。

第 2 章では、本論文で用いる要素技術について述べる。

第 3 章では、本論文で提案する手法について述べる。

第 4 章では、実験概要、結果および考察について述べる。

第 5 章では、本論文の結論と今後の課題について述べる。

## 第2章 ルート推薦に関する要素技術

### 2.1 はじめに

本章では、本論文の研究意図と研究で使用するツールの説明し、用語の定義する。

### 2.2 研究意図

技術の進歩により、ユーザは容易に情報を発信することができ、他者が公開している様々な情報を簡単に得ることができる。また、位置情報を用いた新たなサービスも進歩している。しかし、この変化により、ユーザが処理可能な量を遥かに上回る情報が流通するようになり、真に必要な情報の選別を困難にしている。この影響は旅行にも見られ、旅行先を選ぶ際、有名な観光名所や特定の場所のみが選択の対象とされている。そのため、あまり世間では知られていないような価値ある観光地の情報を得ることが困難になりつつある。

そのような状況の中で、観光地の写真に関して、その地に住む居住者とその地を訪れる旅行者の間に差異がみられるというデータがある。[5]これは旅行者がインターネットや SNS を使い、有名な観光地や人気の観光名所を訪れ、そこで写真を撮影する一方で、居住者は有名な観光地の写真だけでなく、ネットでは広まっていない地元民のみが知るマイナーな観光地の写真も撮影する傾向があると思われる。そこで本研究では旅行者と居住者の居住地に関する地理的条件の違いを利用し、世間的に馴染みはないが観光する価値がある観光地の検出を行う。観光する価値があるかどうかを決定するには写真を用いる。

旅行する目的に自然観光、歴史・文化観光等、その地の文化や景観美を重視し、観光する傾向がある。[6]この時、人々は観光地の画像や写真を参考に旅行のイメージを膨らまし、目的地を決定するものと思われる。そのため観光地の画像や写真は旅行の目的地を決定する際、大きな役割を担っていると考えられる。

そこで、本研究においては上記で述べたとおり地理的条件から馴染みのない観光地の検出し、得られた観光地に関して写真を用いて観光地として魅力的かどうかを視覚的に評価することで本研究の有効性を確認する。そこで大量の写真収集が行え、位置情報も同時に得られるという点から写真共有サービス Flickr を用い、観光地に関する写真と位置情報の収集する。

### 2.3 Flickr [7]

Flickr とは、米 YAHOO が運営するオンラインの写真共有サービスである。ユーザは自分で撮影したデジタル画像の写真を Flickr のサーバーへアップロードし、保存・管理することができる。Flickr 上で公開された写真は、誰でも閲覧可能である他、コメントの許可やタグ付け機能を利用しながら、他のユーザと交流することも可能である。また、投稿した写真の位置情報を得ることができ、専用の API が充実しているという点から本研究では Flickr を用いて画像取得する。

### 2.3.1 Flickr API[8]

Flickr API 使用にあたって API Key の取得が必要となる。API Key 取得にあたって Flickr のユーザ登録を行う必要がある。Key の取得後、アプリケーションの商用利用の有無を選択する。商用利用の場合、審査に 2 週間~1 ヶ月を要する。

### 2.3.2 Flickr から大量のデータを取得するツール[9]

本研究においては Flickr に付随する位置情報と投稿された画像を利用する。使用する位置情報として以下の 2 点を用いる。

1. 投稿された場所に関する位置情報
2. 写真投稿者の居住地に関する情報

上記の位置情報は Flickr API を用いて入手することが可能である。しかし、API では投稿された写真と位置情報を同時に表示することができない。本ツールでは上記の二つの位置情報の取得と投稿された写真情報を同時に表示することができる。また、取得する写真に関して、検索対象エリアと対象期間、キーワードを設定することができるため、効率的にデータの取得ができる。

## 2.4 SurveyMonkey[10]

本研究において SurveyMonkey を用いて得られた穴場スポットに関して 5 段階評価のアンケートを作成した。

## 第3章 提案手法

### 3.1 はじめに

本章では、居住者と観光客の行動範囲の違いに関する考察から、穴場スポットを検出するアルゴリズムに関する考え方と手法について述べる。

### 3.2 提案手法の概要

提案手法では、Flickr に投稿される写真データの中で居住者と観光客の行動の違いについて着目する。観光地の写真に関して、その地に住む居住者とその地を訪れる旅行者の間に差異がみられるというデータがある。Gigazine らの報告によれば、旅行者は旅行情報誌やネット等に掲載されている有名な観光名所に多く足を運ぶのに対し、地元民は有名な観光名所だけでなく、情報誌等に掲載されていないようなマイナーなスポットで撮影する傾向がある。そこで本研究では観光地に関して居住者が撮影するスポットのうち旅行者が撮影していない場所を穴場スポットとして定義する。その穴場スポットが世間的には有名ではないが、旅行で訪れる人にとって価値のある観光地である可能性があるとの仮説の下、それらを位置情報から推挙する。推挙結果について、数人の被験者に対してアンケートを行い、観光地としての魅力があるかどうかの5段階評価する。

#### 3.2.1 穴場スポット検出アルゴリズム

穴場スポットの検出に関して、居住者と旅行者の撮影位置の違いを利用する。用いる手法は得られたデータを密度ベースで分類し、他のデータの集合と比較したとき、密度の低い集合として認識された位置情報のデータを穴場スポットの位置情報として検出する。そこで本研究ではデータを密度ベースで分類する代表的な手法である局所外れ値因子法（LOF法）[10]を用いて穴場スポットの検出を行う。

##### 3.2.1.1 局所外れ値因子法（LOF法）

データの近傍の密度を利用することで、そのデータの外れ値の度合いを測り、その値によって外れ値を検出する手法である。

全体の集合を  $D$  とする。LOFにおけるデータ  $x \in D$  における外れ値の度合いを  $LOF(x)$  と表記する。ここで  $D$  はデータ全体の集合である。 $LOF(x)$  を定義するために、いくつかの式を定義する。まず、 $kdist(x)$  は  $x$  に対する  $k$  距離と呼ばれる値で、以下の条件を満たすデータ  $o \in D$  との距離  $d(x, o)$  として定義される。

1. 少なくとも  $k$  個のデータ  $o' \in D \setminus \{x\}$  に対して  $d(x, o') \leq d(x, o)$  が成立する。
2. 高々  $k-1$  個のデータ  $o' \in D \setminus \{x\}$  に対して  $d(x, o') < d(x, o)$  が成立する。

直感的には、上記のデータ  $o$  はデータ  $x$  からの  $k$  番目に近いデータとなる。データ  $x$  から同じ距

離を持つデータが複数存在する場合を考慮した定義と言える.

次に  $kdist(x)$  を利用して,  $N_k(x)$ ,  $rd_k(x, y)$  及び  $lrd_k(x)$  を以下のように定義する.

$$N_k(x) = \{y \in D \setminus \{x\} \mid d(x, y) \leq kdist(x)\} \quad (3.1)$$

$$rd_k(x, y) = \max\{d(x, y), kdist(y)\} \quad (3.2)$$

$$lrd_k(x) = \frac{|N_k(x)|}{\sum_{y \in N_k(x)} rd_k(x, y)} \quad (3.3)$$

これらの式を用いて,  $LOF(x)$  は以下で定義される.

$$LOF(x) = \frac{1}{|N_k(x)|} \sum_{y \in N_k(x)} \frac{lrd_k(y)}{lrd_k(x)} \quad (3.4)$$

### 3.3 フローチャート

提案手法に関するフローチャートを図 3.1 として以下に示す.

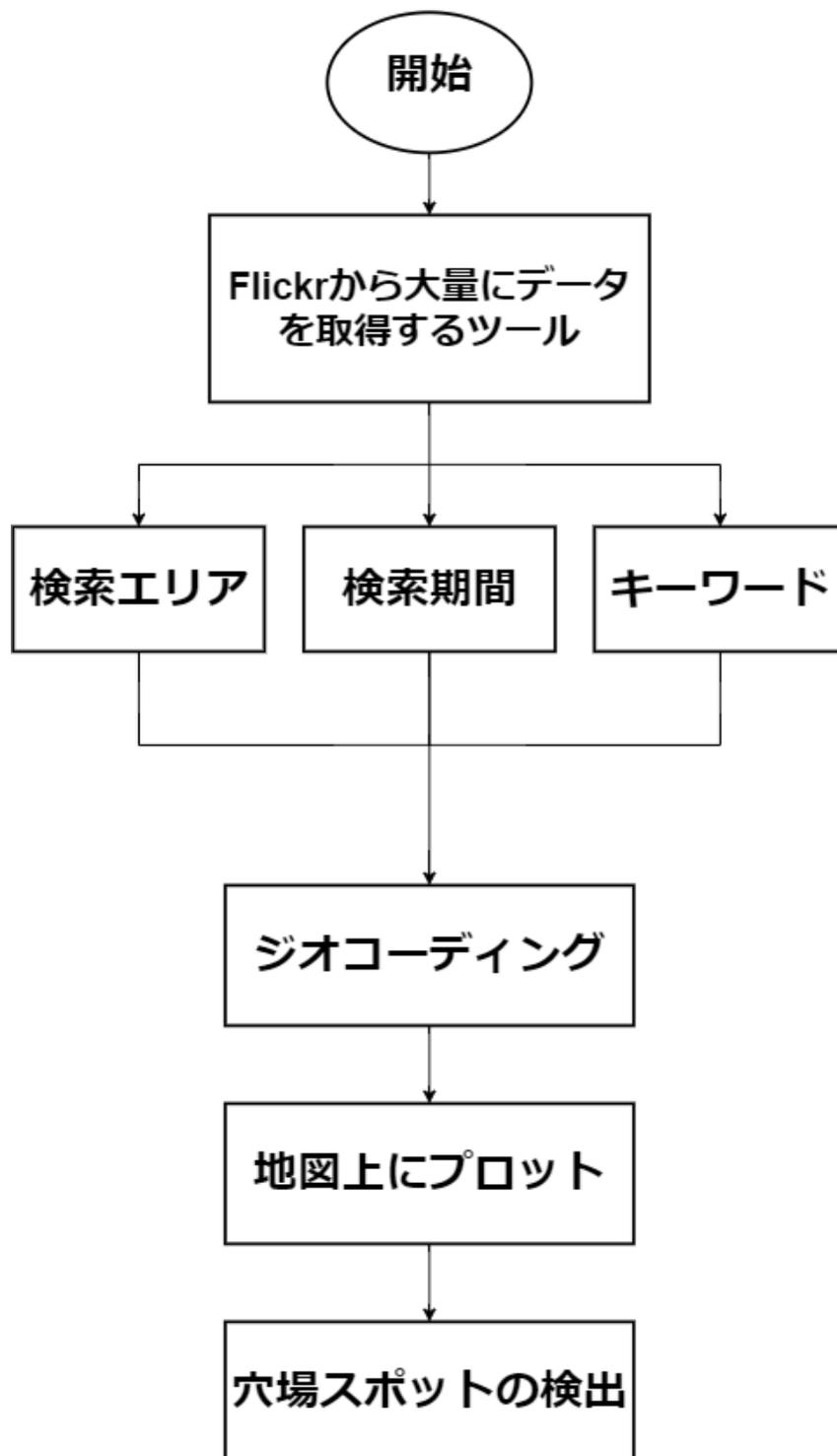


図 3.1 提案手法のフローチャート

### 3.4 むすび

本章では，本論文で提案する居住者と旅行者の撮影位置の違いと LOF 法を用いた穴場スポット検出を組み合わせた手法について述べた．

第 4 章では，本章で説明した提案手法の評価実験を行う．

## 第 4 章 提案手法の評価実験と結果, 考察

### 4.1 はじめに

本章では, Flickr に投稿された画像を用いて第 3 章で説明した提案手法を実験により評価する.

### 4.2 実験の概要

提案手法の評価実験として, 二つの実験を行う. 一つ目の実験では, Flickr に投稿された写真の位置情報と Flickr に登録されている居住者情報を用いて穴場スポット検出を行う. 二つ目の実験では, 検出された穴場スポットに関する写真を被験者に対するアンケートを実施して, 妥当性を評価する.

### 4.3 実験 1 (穴場スポット検出実験)

本実験では, 対象とする地域を設定し, キーワード検索を使って Flickr に投稿された写真を調べる. 表示された写真のうち, 観光地として適切な写真を選び, 写真が撮影された位置情報を入力する. 位置情報は緯度, 経度の情報を利用する. この時, Flickr に設定できる居住者情報から, 対象とする観光地の地域に在住しているユーザを居住者, それ以外の地域に在住しているユーザを旅行者と区別し, 写真を集める. 一定数の写真を収集したら, それを地図上にプロットし, LOF 法を用いて穴場スポットを検出する.

#### 4.3.1 実験 1 の条件

表 4.1 実験条件

検索エリア	京都駅周辺半径 10km 内
検索期間	2016/07/01~2017/07/01
キーワード	kyoto

#### 4.3.2 実験 1 の結果

4.3.1 の条件の下, 旅行者の写真 91 枚, 居住者の写真 71 枚を収集し, その位置情報を地図上 [11]にプロットした図 4.1 を以下のように示す. 赤は旅行者, 青は居住者を示す.

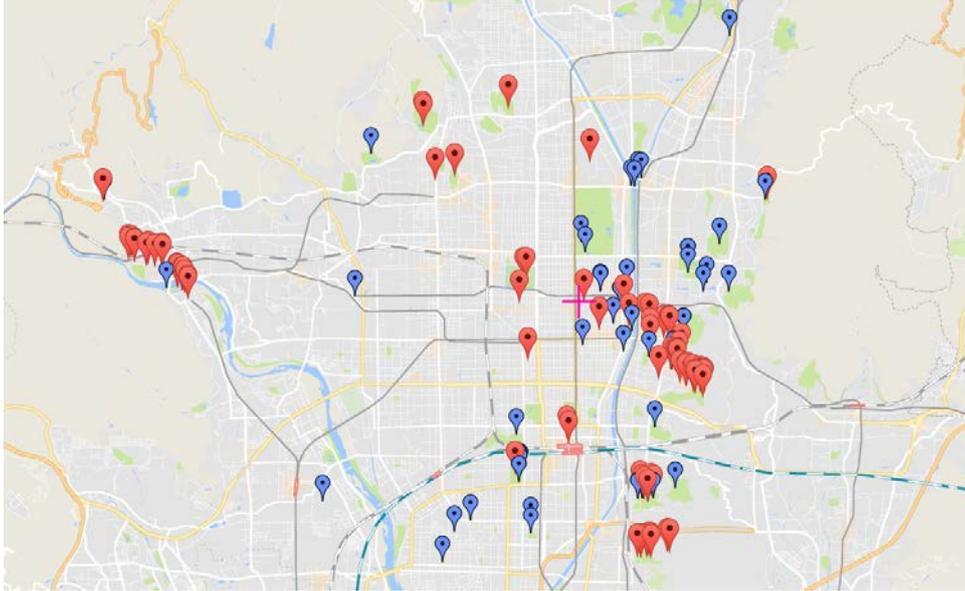


図 4.1 旅行者と居住者の撮影位置の違い

図 4.1 から旅行者の撮影場所は清水寺や伏見稲荷等の有名な観光地に偏っていた。また、得られた位置情報について LOF 法を用い、穴場スポットを検出する。検出した結果を図 4.2 として以下に示す。

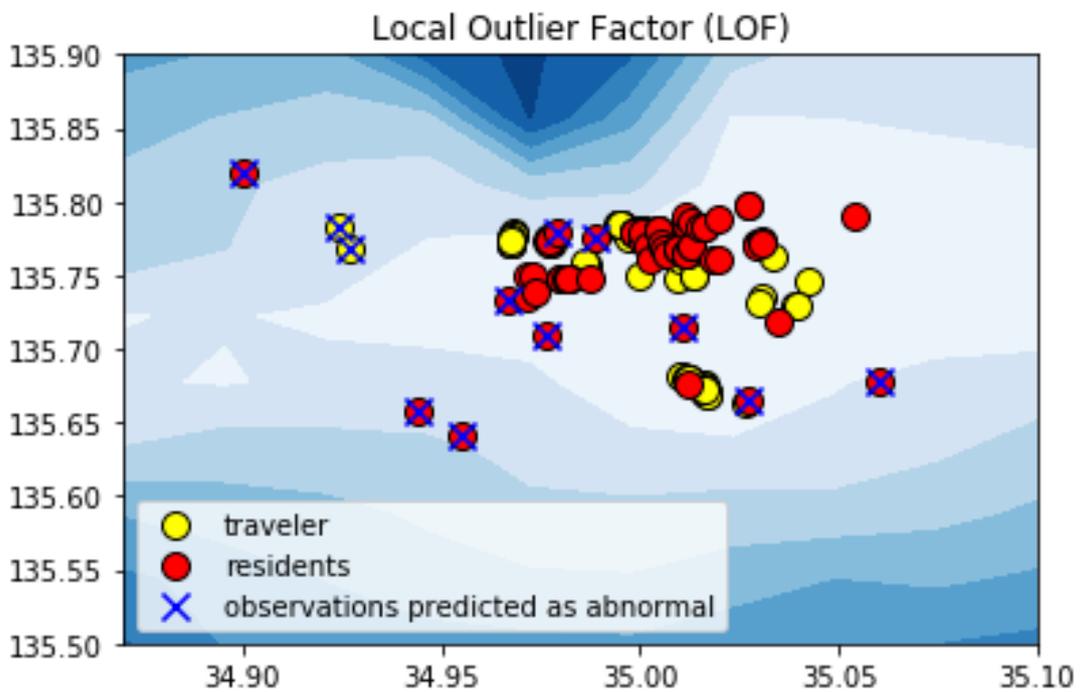


図 4.2 LOF 法を用いて得られた穴場スポットの位置情報

上記の操作から密度の低い集合として検出されたものは

旅行者の写真データ 5 個,

居住者の写真データ 11 個

であった。本研究においては居住者のデータ 11 個を穴場スポットとして考える。

#### 4.3.3 実験 1 の考察

今回は Flickr を利用して写真収集を行い、Flickr に登録してある居住地情報から旅行者と居住者を区別した。しかし、居住地情報を開示していないユーザが多かったため、収集できた写真数が少なかった。また、写真の中には観光地以外の写真も多く見られた。これはユーザが Flickr を写真投稿目的のみに使用している人が多く、他者との交流目的のためではなく自己満足のために使用しているため、不要なプロフィール情報を公開していないユーザが多いと考えられる。

#### 4.4 実験 2 (アンケートを用いた穴場スポットの評価実験)

実験 2 では、実験 1 で得られた穴場スポットに対して被験者数名にアンケートを実施し、検出された穴場スポットが観光地として魅力的かどうかの評価を行う。

##### 4.4.1 実験 2 の条件

- ・アンケートを約 20 人に実施,
- ・旅行者の写真 91 枚, 居住者の写真 71 枚を利用する。
- ・上記の 162 枚の写真を「旅先として魅力的かどうか」を被験者に対して、5 段階 (5:とても思う, 4:思う, 3:普通, 2:思わない, 1:全く思わない) で 162 枚の写真について、被験者が 5 段階評価する。
- ・収集したアンケート結果について、上下 5% ずつのデータを除外し、異常値を取り除く。
- ・データの除外にはエクセルの TRIMMEAN 関数を利用する。

##### 4.4.2 実験 2 の結果

実験条件を以下の表 4.1 にまとめる。

表 4.2 アンケート結果

旅行者	居住者	穴場スポット
3.64	3.51	3.77

##### 4.4.3 実験 2 の考察

結果から旅行者が撮影した写真の方が居住者の撮影した写真より観光地として魅力的に考える人が多いことがわかる。しかし、穴場スポットと旅行者の写真を比較したとき、穴場スポットの

方が魅力的に考える人が多い。これは旅行者にとって馴染みのない場所も有名な観光地以上に魅力的と考える人が一定数いるということを示している。そのため、本研究における仮説が成り立つことを確認した。しかし、実験の条件を変更することで上記の結果は変動すると考えられる。

#### 4.5 まとめ

二つの実験を行った。実験 1 では Flickr から収集した写真を旅行者と居住者の二種類に分け、LOF 法を用いて穴場スポットの検出した。実験 2 においては実験 1 で収集した写真を元に被験者に対して観光地として魅力的かどうかのアンケートを行った。結果から穴場スポットの写真は旅行者の写真と比較しても魅力的であるという結果が得られ、本研究の仮説が成り立つことを確認した。

## 第 5 章 結論と今後の課題

### 5.1 結論

本研究では、居住者と旅行者の写真の位置情報から LOF 法を用いて穴場スポットの検出を行い、五段階評価のアンケートによって評価した。アンケートの結果より穴場スポットとして定義した観光地が旅行先として魅力的であるということが分かった。旅行ルートのおすすめに関して、旅行者と居住者の撮影する写真の位置情報を利用し、視覚的に評価を行う本研究の手法の有効性が証明できた。

### 5.2 今後の課題

Flickr で収集したデータ数が少ない点が課題点として挙げられる。また、本研究では提案手法に関する具体的な推薦方法を示すことができなかった。

## 謝辞

本研究の実験環境を与えてくださり、研究の方向性等の丁寧かつ熱心なご指導を頂いた渡辺裕教授と早稲田大学国際情報通信研究センター招聘研究員の石川孝明様に心から感謝いたします。また、日頃から御意見やアドバイスをくださった研究室の皆様に御礼申し上げます。最後に、私をここまで育ててくださった家族に感謝いたします。

## 参考文献

[1] trivago, <https://www.trivago.jp/>

[2] エクスペディア, <https://www.expedia.co.jp/>

[3] 王佳な, 野田雅文, 高橋友和, 出口大輔, 井手一郎, 村瀬洋, “Web 上の大量の写真に対する画像分類による観光マップの作成”, 情報処理学会論文誌 Vol.52 No.12 3588-3592, December 2011

[4] 坂元陽亮, 高間康史, “RDF データベース構築によるユーザの気分に応じた観光スポット推薦システムの提案”, The 31<sup>st</sup> Annual Conference of Japanese Society for Artificial Intelligence, 2017

[5] Gigazine, 世界各地の都市の観光客と地元民の行動範囲の違いがわかるマップ(2010年6月10日), [https://gigazine.net/news/20100610\\_locals\\_and\\_tourists/](https://gigazine.net/news/20100610_locals_and_tourists/)

[6] 日本交通公社, 2017年度調査結果「旅行年報 2017」第I編日本人の旅行市場 [https://www.jtb.or.jp/wp-content/uploads/2017/10/nenpo2017\\_1-4.pdf](https://www.jtb.or.jp/wp-content/uploads/2017/10/nenpo2017_1-4.pdf), 2017

[7] Flickr, <https://www.flickr.com/>

[8] Flickr API, <https://www.flickr.com/services/api/>

[9] Flickr から大量の画像のデータを持ってくるツール  
[http://www.comp.tmu.ac.jp/kurata/tpm/tpm2h\\_downloader.html](http://www.comp.tmu.ac.jp/kurata/tpm/tpm2h_downloader.html)

[10] SurveyMonkey, <https://jp.surveymonkey.com/>

[11] Google Maps, <https://www.google.co.jp/maps>

**図一覧**

図 3.1	提案手法のフローチャート図	10
図 4.1	旅行者と居住者の撮影位置の違い	12
図 4.2	LOF 法を用いて得られた穴場スポットの位置情報	12

**表一覧**

表 4.1 実験条件-----	11
表 4.2 アンケート結果-----	13