

# スポット画像特徴量を用いた 観光スポットマイニング手法 「ジェネリック観光地」抽出技術

クロステック開発部 かつみ ひさお やまだ わたる  
**勝見 久央 山田 渉**  
おちあい けいいち  
**落合 桂一**

AI技術を用いて各地域から新たな観光スポットをマイニングすることができれば、観光客の一極集中や、難航する地方都市への観光客の誘致などの社会的課題に対する新たな解決策となり得る。本稿では有名な観光スポットに類似し、これらの代替となり得るスポットを「ジェネリック観光地」と呼称し、AI技術によってWeb上のデータからジェネリック観光地をマイニングするアイデアとその実装案を紹介する。また、実際にWeb上のデータからジェネリック観光地をマイニングし、その結果を評価した内容について解説する。本稿で紹介する取組みが今後AI技術を活用した観光産業への新たな一歩となることが期待される。

## 1. まえがき

移動手段やメディアの急速な発達により、我々はさまざまな場所を旅行することが可能になった。一方で、特定の観光地に観光客が一極集中するオーバーツーリズムが問題視されている [1]。オーバーツーリズムは、旅行者と当該観光地の住民の双方に悪影響を及ぼす深刻な問題である。これに対する有効な取組みとして、①既存の観光スポットの観光客受入れ容量を増やす、②既存の観光スポット間で観光客を時間的、空間的に分散させる、③新たに観光

スポットを創出し、観光客をそちらに誘導するという3つが挙げられる [2]。

これらのうち①では、観光施設の駐車場を整備するなどのさまざまな取組みが進んでいる。②の取組みの代表例としては、リアルタイムの混雑状況を観光スポットへの推薦に反映して混雑回避を図る手法 [3] や、Web上の画像投稿サイトにおいて、観光スポット画像ごとの投稿ユーザの属性などを分析し、観光客よりも地元住人に好まれているような穴場スポットをマイニングする手法 [4] などがある。また、③のように、これまで観光スポットとみなされ

©2022 NTT DOCOMO, INC.

本誌掲載記事の無断転載を禁じます。

本誌に掲載されている社名、製品およびソフトウェア、サービスなどの名称は、各社の商標または登録商標。

ていなかったようなスポットを観光資源とみなし、都市内の観光スポットの上限数を増やすような取組みも重要である。

ドコモは、③の新たな観光スポットを創り出す取組みとして、すでに有名なスポットに見た目が似ていることが話題となり、さらに人気になったスポットが存在することに着目した。例えば、日本国内で有名な「竹田城跡」では、ペルーのマチュ・ピチュに見た目が似ていることがメディアなどで話題となり、その結果、観光客が増加した。このように、これまで観光スポットとみなされていなかったスポットに対して、有名スポットに類似しているという観点で観光価値を見いだすことができれば、新たな観光スポットをマイニングすることができる。ドコモは、こうした既存の有名なスポットに類似し、これらの代替となり得るスポットを「ジェネリック観光地」と名付けた。

本稿では、画像処理AI技術などによって、既存の有名観光スポットに見た目が類似したスポットを

ジェネリック観光地として提示する手法を解説する。さらに、Web上のスポット画像を用いて提案手法の評価実験を行った内容について解説する。

## 2. ジェネリック観光地抽出技術

提案手法の概要を図1に示す。本手法では、ジェネリック観光地のマイニング対象となる候補スポットおよび既存の有名観光スポットの画像をWeb上から収集し、これらすべての組合せについて画像類似度を算出する。そして、画像類似度が上位の組合せに含まれる候補スポットをジェネリック観光地として出力する。

本手法とそのアイデアは、スポット画像のみを用いて新たな観光スポットをマイニングするものであり、一般的な観光スポットマイニング技術と比較して、事前のスポットの口コミや投稿画像などの収集量に依存せずに新たな観光スポットをマイニングできる可能性がある。さらに、地元住民すらもこれま

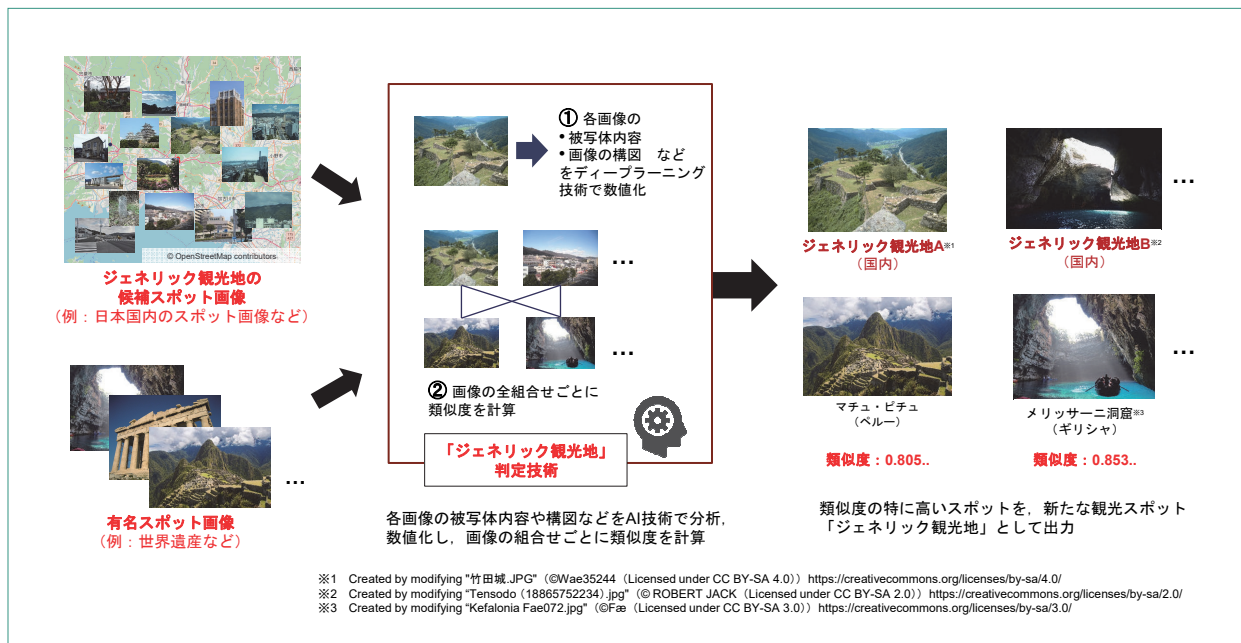


図1 提案手法の概要

で観光スポットと認識していなかったようなスポットに対しても新たな観光価値を見いだすことができる。

### 3. ジェネリック観光地抽出技術

本手法は主に4つの処理から構成される（図1）。

手順①：Web上から、任意の範囲内で撮影された候補スポット画像を収集する。

手順②：Web上から、名称を基に有名スポット画像を収集する。

手順③：収集された各候補スポット画像、有名スポット画像から、すべての組合せの画像類似度を算出する。

手順④：画像類似度の上位 $N$ 件の組合せに含まれる候補スポットをジェネリック観光地として出力する。

手順①、②では、Web上から得られたさまざまなスポット画像を利用することが想定される。また、本手法の評価実験では、写真投稿サイトであるFlickr<sup>\*1</sup>上に投稿されたスポット画像を利用した。

手順③では、すべての候補スポット画像と有名スポット画像の組合せの画像類似度を算出した。画像類似度の算出にあたっては、各候補スポット、有名スポットの画像を特徴量ベクトル<sup>\*2</sup>に変換し、それらのコサイン類似度<sup>\*3</sup>を各スポット画像同士の類似度とした。スポット画像の特徴量ベクトルとしては、各画像の内容などの意味的特徴量や、構図などの構造的特徴量を抽出する手法が有用であると考えられる。前者としては、各画像に山や河川などの365種類の風景カテゴリの成分がどの程度含まれているかを数値化したVisual Concept [5]がある。後者としては、各画像の構図や明暗箇所などの特徴を抽出するGIST特徴量 [6] などがある。

最後に、手順④ですべての組合せのうち、類似度

上位の組合せを基にジェネリック観光地を出力する。本手法では、類似度上位の $N$ 件までに含まれる候補スポットをジェネリック観光地と判定して出力した。

### 4. 評価実験

提案手法の妥当性を検証するため、評価用データを作成した。さらに、提案手法を用いて評価用データから実際にジェネリック観光地をマイニングし、その結果が正解例とどの程度一致しているかを誤差として算出し、評価した。

#### 4.1 評価用データの作成

評価用データの作成方法の概要を図2に示す。まず、すでに海外の有名スポットに似ていることで知られている日本国内の観光スポット5カ所（図2の左の表）を、ジェネリック観光地の正解例とした。

次に、これらの正解例を用いて、ジェネリック観光地の正解例を含む候補スポット画像をFlickr上から収集した。具体的には、5つの正解例が属する標準地域メッシュ<sup>\*4</sup>内で撮影された画像を、候補スポット画像として収集した。標準地域メッシュとして、日本国内を約10km四方のメッシュに分割する2次元メッシュを用いた。図2右側は、該当の2次元メッシュ内に評価データとして利用したジェネリック観光地の正解例を★印、候補スポット画像の撮影地点を●印でプロットしたものである。このような方法で2,353枚の候補スポット画像が得られた。

また、5つの正解例に対応する有名スポット画像については、正解例に含まれる5つの有名スポット名の英名をそれぞれクエリ<sup>\*5</sup>としてFlickr上から得られた画像5枚ずつ、合計25枚を利用した。ただし、実際に有名スポット名の検索結果として得られる画像には、スポットそのものを撮影した画像だけでなく、人物や料理などを撮影したようなスポット画像以外の画像がノイズとして含まれる。評価用データ

\*1 Flickr：米国Oath Inc. の商標または登録商標。

\*2 特徴量ベクトル：データのパターンや特徴を数値の列であるベクトルで表現し、プログラムなどで扱えるようにしたもの。

\*3 コサイン類似度：2つのベクトルの向きがどの程度近いかを数値化したもの。

\*4 標準地域メッシュ：日本の各地域データを集計する用途などのために、日本国内を緯度経度に従ってメッシュ状に分割したもの。

\*5 クエリ：データベースに対する問合せ（処理要求）。

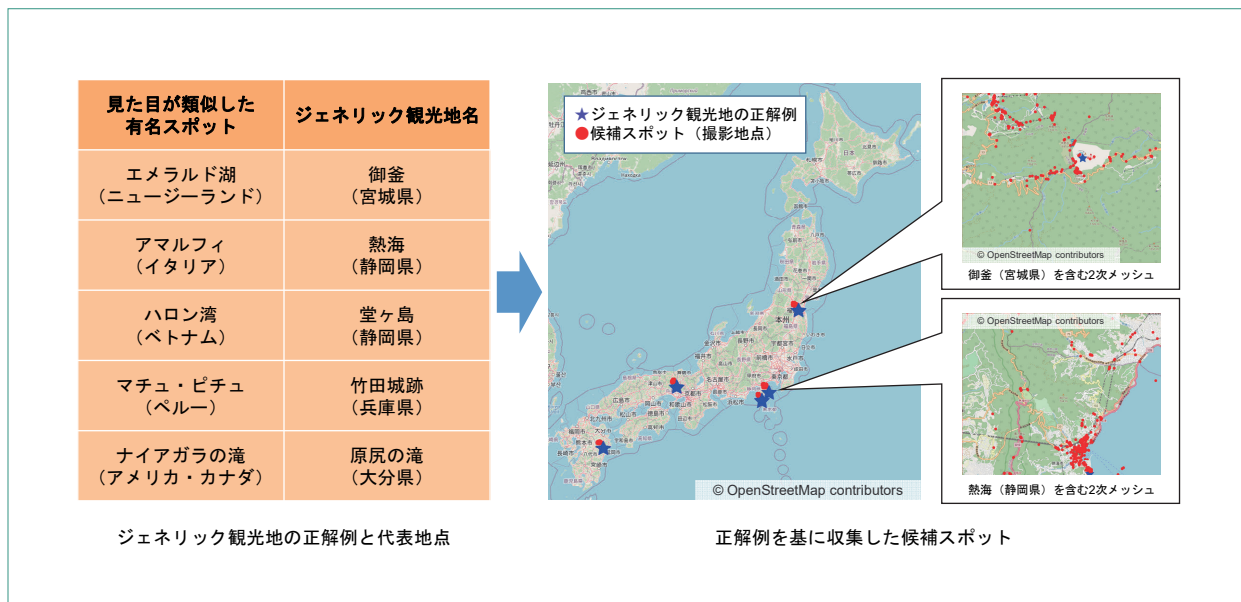


図2 評価用データの作成

からはこのようなノイズとなる画像を人手で除去した。

なお、上記の収集画像は、評価用データの作成段階で投稿者がFlickr上で設定していたクリエイティブ・コモンズ・ライセンス<sup>\*6</sup>表記を基に、研究用途での利用が可能な画像のみを対象として得られたものである。

#### 4.2 スポット画像の特徴量ベクトル化

本稿で提案するジェネリック観光地のマイニング手法では、各スポット画像を特徴量ベクトル化してからスポット画像間の類似度を算出する。評価時には特徴量ベクトル化の方法として次の3種類の方法を用いて、それぞれ得られた出力結果を比較した。

##### 手法①：Visual Concept

この手法では、各スポット画像が山、河川、海など365種類の風景カテゴリのそれぞれにどの程度当てはまるかをニューラルネットワーク<sup>\*7</sup>で判定し、スポット画像をそれぞれ365次元の特徴量ベクトルに変換した。なお、ニューラルネットワークは風景画像データセットである

Places365を用いて学習されたResNet18を用いた [7]。

##### 手法②：GIST特徴量

各スポット画像をGISTにより特徴量ベクトル化した。GIST特徴量は、各スポット画像内の大まかな明暗の分布などの構造的な特徴を抽出したものである。

##### 手法③：埋込みベクトル<sup>\*8</sup>

ニューラルネットワークが段階的に特徴量の抽出を行う性質を考慮し、手法①で用いたニューラルネットワークが最終的にVisual Conceptとして365次元の特徴量ベクトルを算出する前段階の512次元のベクトルを特徴量ベクトルとして利用した。このようにして得られる特徴量ベクトルには、ニューラルネットワークが手法①で最終的にVisual Conceptとして365種類の風景カテゴリを推論する直前の、より抽象度の高い特徴が埋め込まれており、365種類のカテゴリ上では異なるカテゴリに分類されてしまうような構造物などに対しても類似し

\*6 クリエイティブ・コモンズ・ライセンス：クリエイティブ・コモンズが定めるルールに従って、著作物の利用ルールなどについて意思表示したもの。

\*7 ニューラルネットワーク：生物の脳の働きを数理的モデルによって模倣し、数値的なパターンの認識や推論ができるようなモデル。

\*8 埋込みベクトル：高次元の特徴量から必要な情報を低次元に絞り込んでベクトル化したもの。



ているものとして扱い易くなる。

### 4.3 評価結果と考察

上記の方法で作成した評価用データから、提案手法を用いて $N$ カ所 ( $N=10, 20, 30$ ) のジェネリック観光地をマイニングした。手法③を用いて $N=30$

としてマイニングされたジェネリック観光地のスポット画像を図3に、それらの位置を図4に示す。マイニングされたジェネリック観光地と正解例の誤差については、図4右側に示すように、各メッシュ内でジェネリック観光地としてマイニングされた合計 $N$ カ所のスポットと正解例の位置との平均距離

ジェネリック観光地と判定されたスポット	竹田城跡 (兵庫県) "Takeda-jo ruins" © cotaro70s (CC BY-ND 2.0)	熱海周辺 (静岡県) "Atami Port in Atami City, Shizuoka, Japan" © Izu navi (CC BY 2.0)	原尻の滝 (大分県) "原尻の滝 1" © wa_me (CC BY-NC-ND 2.0)	堂ヶ島 (静岡県) "APC_0064_iPhone.jpg" © Misako Kuniya (CC BY-NC 2.0)	御釜 (宮城県) "御釜" © NISSAN MOTOR CO., LTD. (CC BY-NC-ND 2.0)
類似している有名スポット	マチュ・ピチュ (ペルー) "Machu Picchu" © F Delventhal (CC BY 2.0)	アマルフィ (イタリア) "Amalfi cartolina" © Antonio Salsano (Salsano2016 - SIAE) (CC BY-NC-ND 2.0)	ナイアガラの滝 (アメリカ・カナダ) "Niagara Falls" © Peter Miller (CC BY-NC-ND 2.0)	ハロン湾 (ベトナム) "Quang Ninh VN - Ha Long Bay 03" © Daniel Mennerich (CC BY-NC-ND 2.0)	エメラルド湖 (ニュージーランド) "Emerald Lake" © Andrea Schaffer (CC BY 2.0)

※画像は各画像の下部に示す作品名および著作者の作品から複製または変更したものを、その下部に示すCreative Commonsライセンスに記載された条件に従って使用している。

図3 手法③ ( $N=30$ ) によってジェネリック観光地としてマイニングされたスポット例

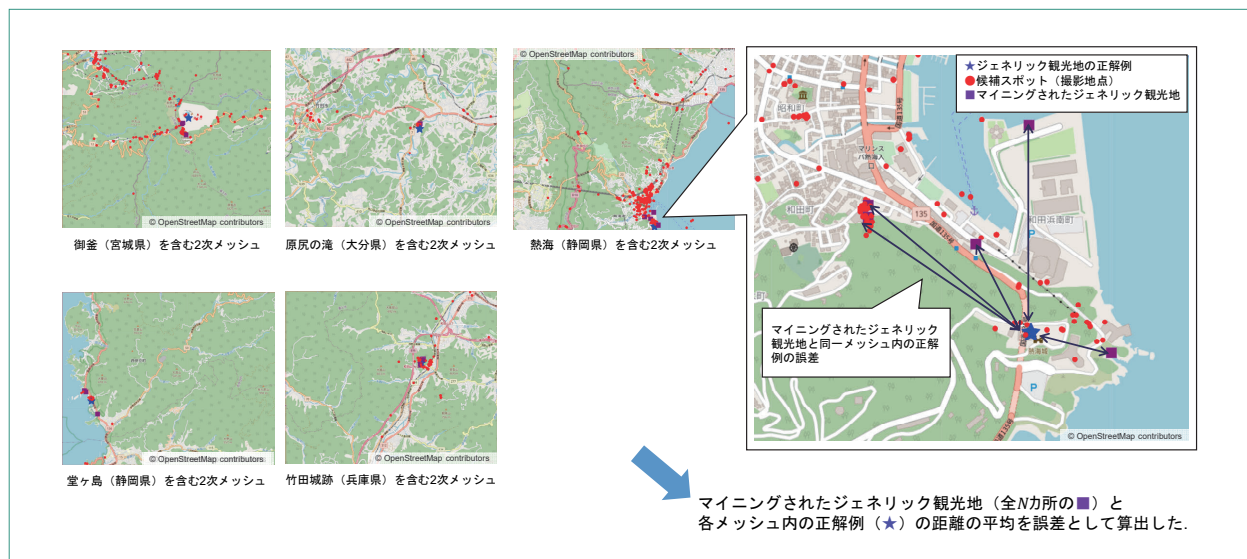


図4 手法③ ( $N=30$ ) によってジェネリック観光地としてマイニングされたスポットの位置と誤差の算出方法

(km)として算出した。ジェネリック観光地として、より正解例と近い候補スポットをマイニングできているほど、この誤差は小さくなり、正解例と位置が完全に一致するスポットをマイニングすると誤差は0となる。ただし、Web上から収集されるスポット画像には同じ対象を撮影したものでも撮影地点が異なるものが含まれるため、正解例を正しくジェネリック観光地としてマイニングしていても誤差は必ずしも0になるとは限らない。例えば図4右側に示すように、熱海周辺でのマイニング結果では、正解例に含まれるアマルフィに類似した熱海のスポット画像を撮影することができる複数カ所の撮影スポットがジェネリック観光地としてマイニングされていた。最終的に得られた各特徴量化手法と $N$ の組合せごとの誤差の評価結果を表1に示す。

図3、4および表1の結果から、おおむね正解例の近傍で有名スポットに類似した見た目の候補スポットをジェネリック観光地としてマイニングできていると考えられる。また、表1に示すように、すべての場合で手法③がその他の手法を上回る結果となった。さらに、手法①と手法②を比較すると、すべての場合で前者が後者を上回る結果となった。これらの結果から、スポット画像の被写体の内容などの意味的な特徴が、構図などの構造的な特徴よりジェネリック観光地のマイニングに有用であると考えられる。

本手法は、候補スポット画像に対してクラスタリング<sup>\*9</sup>などの処理を行わないため、入力される候補スポット画像や有名スポット画像の内容や枚数に

関わらず手法を適用することができると考えられる。今後、評価データの規模を拡大した場合も同様の結果が得られるかどうかを評価する予定である。

## 5. あとがき

本稿では、既存の有名スポットに類似したスポットをジェネリック観光地と呼称し、新たな観光スポットとしてマイニングするアイデアについて解説した。また、ニューラルネットワークなどを用いたマイニング手法を提案し、Web上のスポット画像を用いて評価を行った内容について述べた。ただし、本稿で解説した提案手法はWeb上から収集した5カ所の正解例を基に評価を行ったものであるため、今後、提案手法が大量の候補スポット画像や有名スポット画像の中からも正しくジェネリック観光地をマイニングできるかどうかや、マイニングされたジェネリック観光地を観光客の一極集中の緩和や観光客の誘致に活用できるかどうかを検証する予定である。

本稿で解説したアイデアや提案手法が、AI技術を活用してこれまで観光スポットと見なされていなかったスポットから新たな観光価値を見だし、オーバーツーリズムだけでなく地方創生などへも活用されるような取組みの第一歩となることを期待している [8]。また、昨今の新型コロナウイルス感染防止を意識した新たなライフスタイルの受容とも相まって、海外や国内の遠く離れた場所への旅行だけでなく、これまで観光価値に気付かなかった近場の

表1 各特徴量化手法と $N$ の組合せごとの誤差

	上位10スポットの平均誤差 (km)	上位20スポットの平均誤差 (km)	上位30スポットの平均誤差 (km)
手法①Visual Concept	2.18	2.20	2.22
手法②GIST	2.56	2.35	2.45
手法③埋込みベクトル	0.33	0.27	0.34

\*9 クラスタリング：大量のデータを類似した特徴をもつデータ同士でグループにして分割すること。

さまざまなスポットを新たな視点で捉え、その隠れた魅力を再認識するようなニューノーマル\*10な観光スタイル確立への一歩となれば幸いである。

### 文 献

- [1] M. Duignan : “Overtourism? Understanding and Managing Urban Tourism Growth beyond Perceptions : Case Studies,” United Nations World Tourism Organisation (UNWTO), pp.34-39, Mar. 2019.
- [2] T. Mainil, E. Eijgelaar, J. Klijs, J. Nawijn, and P. Peeters : “Research for TRAN Committee-Health tourism in the EU : a general investigation,” European Parliament, Directorate General for Internal Policies, 2017.
- [3] M. Hidaka, Y. Kanaya, S. Kawanaka, Y. Matsuda, Y. Nakamura, H. Suwa, M. Fujimoto, Y. Arakawa and K. Yasumoto : “On-site Trip Planning Support System Based on Dynamic Information on Tourism Spots,” Smart Cities, Vol.3, No.2, pp.212-231, Apr. 2020.
- [4] C. Zhuang, Q. Ma, X. Liang and M. Yoshikawa : “Anaba : An obscure sightseeing spots discovering system,” 2014 IEEE International Conference on Multimedia and Expo (ICME), pp.1-6, Sep. 2014.
- [5] C. Peters, T. Deselaers, N. Ferro, J. Gonzalo, G. J. F. Jones, M. Kurimo, T. Mandl, A. Penas and V. Petras : “Evaluating Systems for Multilingual and Multimodal Information Access,” 9th Workshop of the Cross-Language Evaluation Forum, CLEF 2008, Aarhus, Denmark, Revised Selected Papers, p.527, 2008.
- [6] A. Oliva and A. Torralba : “Modeling the Shape of the Scene : A Holistic Representation of the Spatial Envelope,” International journal of computer vision, Vol.42, No.3, pp.145-175, May 2001.
- [7] B. Zhou, A. Lapedriza, A. Khosla, A. Oliva and A. Torralba : “Places : A 10 Million Image Database for Scene Recognition,” IEEE Transactionson on Pattern Analysis and Machine Intelligence, Vol.40, Issue 6, pp.1452-1464, Jul. 2017.
- [8] NTTドコモ報道発表資料 : “日本初！海外の有名スポットにそっくりな四国の景観を発見するフォトコンテストを開催 ～遠出が難しい今、海外旅行気分を味わえる場所を発見！風景画像の類似度をAIで判定～,” Jul. 2021.

---

\*10 ニューノーマル：社会の環境や情勢の変化に伴い、不可逆的に新たな常識が定着した状態。