

写真共有サイトに投稿された写真群を活用したテーマ別観光マップの作成

真田 風[†] 倉田陽平[†] 相 尚寿[†]首都大学東京 大学院都市環境科学研究科 観光科学域[†]

1. はじめに

近年観光に対するニーズは多様化している。しかし、雑誌などで提供されている観光情報は、多数の旅行者が望む一般的な観光情報であり、多様化する観光ニーズに対応できているとは言い難い。そこで、CGM (Consumer Generated Media) におけるデータを用いることで、効果的な観光情報が獲得できると考えられる[1]。中でも、Flickr やフォト蔵をはじめとする写真共有サイトの利用拡大や、GPS センサ搭載携帯デバイスの普及により、位置情報の付いた写真が Web 上に大量にアップロードされており、それらの写真データを活用した観光情報の研究が盛んに行われている。例えば、Web 上の大量の写真を画像分類して提示する観光マップの提案[2]や、写真撮影の位置軌跡を利用した旅程推薦システムの開発[3]、写真群の位置情報を元に人気撮影箇所をヒートマップとして可視化する観光ポテンシャルマップの作成[4]等がある。しかしいずれの研究も、大量の写真データを用いて、大勢の人が注目した場所を見出し、観光情報として提供することを目的としているために、近年の観光ニーズへの対応は難しい。

一方で、一連の観光ポテンシャルマップ研究の中で「写真のタイトルやタグの中身に踏み込み写真を選別する、テーマ別のマップ作成の可能性」が指摘されている。今までは、あくまで「旅行中に撮影された写真」全てを用いることで人気写真撮影箇所を表現していた。これに対し、使用する写真の抽出の際にキーワード(タイトルやタグの情報)検索を行うなどすることで、特定の観光資源に関する観光マップが作成できるようになることが期待される。図1は、「猫」をキーワードに写真を抽出して作成を行った、猫をテーマとした東京周辺の観光マップの例である。

そこで本研究では、観光ポテンシャルマップ研究を土台に、テーマ別観光マップの作成に取

り組む。そして、低コストで有用なCGMデータを用いた、テーマ別観光マップ作成の方法論を導き出すことを目的とする。

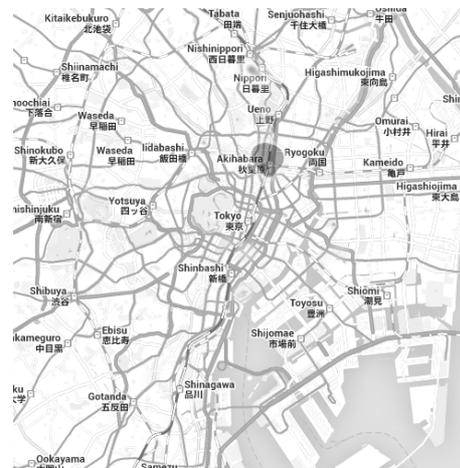


図1. 猫マップの例

(秋葉原周辺に猫の写真が多いことを示している)

2. 方法

テーマ別観光マップを作成するにあたり、まずはどのようなテーマが実際に地図化に適しているのかを評価する必要がある。さらに、抽出データ数の多いテーマに対しては、先行研究同様、テーマに適した有効な写真を自動的に選別するためのルールの構築が必要である。これは、枚数が多すぎると人力での判別が追いつかず、今後の他地域や他期間への応用が難しいからである。以上をふまえて、調査は2段階で行った。

3. 調査1 データの抽出

調査1では、地図化が期待できそうなテーマの候補を選定し、対応するキーワードによってサンプルとなる写真を抽出し、さらに目視によって有効写真と不要写真の判別を行った。対象地域は東京駅を中心に東西10km、南北10kmの範囲とし、2013年の1年間に撮影されFlickrに投稿された写真を用いた。地図化するテーマとして選定したのは、夜、雨、猫、桜、ラーメン、コスプレ、東京スカイツリー、などをはじめ

めとした計 23 テーマである。各テーマについて、有効写真枚数、有効写真率、有効写真の投稿者数と投稿写真枚数の偏り（ジニ係数）、等のデータを元に、テーマとしての地図化適性を考察した。表 1 に、各テーマにおけるそれぞれのデータを示す。

この結果として、例えば①雨のように目視での判別が難しいテーマは有効写真率が低くなるということや、②富士山のような有名な観光地名は、その地を行先として含む旅行の写真全体にタグとして付けられ、まとめて投稿される場合があることから不要写真を多く抽出してしまうということ、③猫やコスプレなど、同一被写体が様々なポーズや表情を見せるようなテーマは、一部の熱狂的なファンにより同じような写真が複数枚投稿されるケースが見られ、投稿者ごとに投稿枚数に偏りが出る、ということがわかった。

4. 調査 2 自動選別ルール構築

調査 2 では、調査 1 において有効写真枚数の多かった、夜、猫、桜、東京スカイツリーの 4 テーマについて、決定木学習法を用いて、有効写真を自動選別するためのルールを構築した。説明変数として、Flickr に記録されている写真のメタデータ（写真のタイトル、タグ、撮影日時等）や個々の写真に記録されている Exif データ（撮影機材のメーカー、シャッタースピード等）を用いた。構築したルールは、他地域で抽出されたデータに適用することで、その有効性を検証した。

この結果として、①夜の有効写真を自動で選別するためには、シャッタースピードという、そのテーマの写真の特性に合った説明変数が有効で、他地域への応用も利きやすいということや、②一部の人による大量写真投稿がある猫などのテーマは、投稿者の情報に依存しない説明変数を用意することが、他地域にも応用できるルールを構築するためには必要である、ということなどがわかった。

5. おわりに

本研究では地図化を行いたいテーマごとに、抽出されたデータを分析し考察することで地図化適性に影響する要因を整理し、データ数が多いテーマに対しては自動選別ルールの構築を行い、有効なルール構築に必要な説明変数の性質を明らかにすることで、テーマ別観光マップの作成手法の一部を示した。

表 1. 各テーマの有効写真枚数，有効写真率，ジニ係数

テーマ	有効 写真枚数	有効 写真率	ジニ係数
夜①	3873	60.8%	0.662574
夜②	5075	80.4%	0.697701
雨①	199	30.5%	0.409319
雨②	306	46.9%	0.433643
雪①	214	26.8%	0.540330
雪②	409	51.3%	0.606895
虹	10	1.8%	0.088888
夕焼け	811	91.1%	0.602395
朝焼け	43	48.9%	0.275862
猫	1359	67.4%	0.899872
蝶	33	58.9%	0.363636
ホテル	0	0%	-
桜	3154	86.0%	0.711455
梅	191	70.0%	0.531219
紫陽花	130	98.5%	0.532692
銀杏	174	76.3%	0.681321
楓	87	70.2%	0.492017
そば	239	41.4%	0.720977
寿司	413	56.9%	0.646604
ラーメン	581	67.2%	0.544562
着物	72	62.6%	0.349483
コスプレ	1338	92.3%	0.898355
神輿	224	65.3%	0.688928
クリスマスツリー	110	50.2%	0.610774
富士山	51	33.8%	0.280451
東京スカイツリー	1408	48.8%	0.657485

参考文献

- [1]長尾光悦 (2012) CGM からの観光情報の獲得. 情報処理学会第 74 回全国大会講演論文集 2012(1), 505-507, 2012-03-06
- [2]王佳な・野田雅文・高橋友和・出口大輔・井手一郎・村瀬洋 (2011) Web 上の大量の写真画像分類して提示する観光マップの提案. 電子情報通信学会技術研究報告. IE, 画像工学 110(456), 153-158, 2011-02-28
- [3]奥山幸也・柳井啓司 (2011) 写真撮影の位置軌跡を利用した旅行支援システム. DEIM Forum, F7-6
- [4]倉田陽平 (2013) 観光ポテンシャルマップの信頼性向上に向けて—ソースとなる投稿写真データの自動選別ルールの構築—. 第 22 回地理情報システム学会学術大会, 東京, 2013 年 10 月, CD-ROM