

Flickr と Google Cloud Vision API によりテーマ別観光マップを作る試み

倉田 陽平[†] 真田 風[‡] 鈴木 祥平[†] 石川 博

[†] 首都大学東京 都市環境学部 〒192-0397 東京都八王子市南大沢 1-1

[‡] 福島県立テクノアカデミー会津 〒969-3527 福島県喜多方市塩川町御殿場 4-16

・首都大学東京 システムデザイン学部 〒191-0065 東京都日野市旭が丘 6-6

E-mail: [†] ykurata@tmu.ac.jp

あらまし 筆者らは写真共有サイト Flickr 上の位置情報付き写真を用いて、見所があると思しき箇所を地図上に可視化した「観光ポテンシャルマップ」を作成し、さらに写真の種類を限定することで「桜」「夜景」「猫」などの「テーマ別観光ポテンシャルマップ」の作成も試みてきた。先の研究では、写真の精選を自動化するため、写真の Exif データを用いた決定木ルールを利用した。一方、Google 社が Cloud Vision API を一般公開したことにより、画像認識による写真の内容推定がきわめて容易になった。そこで本研究では、テーマごとの写真の精選に際し Cloud Vision API を適用し、どのようなテーマであればテーマ別観光ポテンシャルマップの作成が精確かつ効率的に行えるのかを検討した。

キーワード 写真共有サイト, Flickr, 位置情報付き写真, Google Cloud Vision API, 観光ポテンシャルマップ

1. はじめに

旅行者は、旅行中、何かに心動かされたとき、その対象について写真撮影を行う傾向にある。さらに最近では、そのような写真は SNS 上へ投稿されることが少なくない。したがって、SNS に投稿された写真群、とりわけ位置情報がついたものを活用することで、旅行者に強い印象を与えるような場所を抽出できるようになることが期待できる。実際に位置情報付き SNS データを活用し、観光客向けの新たな地図を構築しようとする研究はいくつも見られる（たとえば[1-3]）。筆者らも写真共有サイト Flickr (<https://www.flickr.com/>) 上の位置情報付き写真を利用し、旅行者と思しき投稿者の写真[4]あるいは観光・レジャー中に撮影されたと思しき写真[5]を選別し、地図上にヒートマップとしてプロットすることで、観光資源の存在可能性を示す地図（観光ポテンシャルマップ/Potential-of-Interest Maps）を作成した。さらに筆者らは、Flickr API のキーワード検索を活用することで、様々な「テーマ別の観光資源地図」を描くアイデアを提起した[5]。たとえば、検索キーワード「猫」で得られた膨大な写真の中には、屋外の猫の写真もあれば、室内猫や招き猫の写真もある。そこで前者のみを選別し、その撮影位置の分布を地図化すれば、猫の写真が撮影されがちな（おそらく猫と遭遇しやすい）場所のわかる地図が得られるはずである。個人旅行者が増加し、旅行目的が多様化している今日、このようなソーシャルデータを活用した観光資源地図は、今までに無い有用な観光情報となり得ると期待される。

しかし、ここで課題となるのは、いかにテーマに即した写真を効率よく抽出するかである。Flickr のキーワード検索から得られた写真群について、一枚一枚、

本当にテーマに該当するものか否かを人力でチェックするには限界がある。そこで真田[6]は決定木学習を用いて、5 テーマ（「夜景」「猫」「桜」「ラーメン」「東京スカイツリー」）に対する写真の自動精選方法の構築を試みた。しかしこの方法では写真のメタデータ（Exif データ）しか用いられておらず、肝心の「何が写っているか」という被写体情報までは考慮されていなかった。一方、2016 年に一般公開された画像認識クラウドサービス Google Cloud Vision API (<https://cloud.google.com/vision/>) は、写真の被写体を機械判定し、ラベリングする機能を有しており、これを用いれば被写体情報を得ることが極めて容易になった。そこで本研究では、この Google Cloud Vision API を写真の精選に用いることで、テーマ別の観光資源地図の効率的な作成を試みる。とくに様々なテーマについて試行を行うことで、どのようなテーマが観光資源地図の自動作成に適しているのかを検討する。

2. 方法

Flickr API (<https://www.flickr.com/services/api/>) を用いて、テーマごとに、テーマに即した語句（日本語・英語一語ずつ）を検索キーワードとして、2014 年 1 月 1 日から 2015 年 12 月 31 日の間に東京都心（東京駅を中心とする半径 15km 以内）で撮影・投稿された位置情報付き写真を検索する。次に、得られたテーマ別の投稿写真群について、単一の評定者の目視によりテーマに該当する写真か否かを判別し、正解データを構築する。また一方で、Google Cloud Vision API を用いて個々の写真の被写体情報（ラベル）を取得する。なお各写真のラベルは、Google Cloud Vision API の仕様により最大で 5 つ得られる（撮影内容が不明瞭のとき

は、1つも得られないこともある)。そしてテーマごとに、出現したすべてのラベルの出現回数を数え、出現頻度上位5つのラベルを抽出する。たとえば、「着物」というテーマであれば costume, clothing, pedestrian, kimono, art が頻出ラベルとして抽出される。そして、これらの頻出ラベルを用いて写真の精選を行った際の有効性を評価する。たとえば検索キーワード「着物」で得た写真群の中に含まれる「costume」というラベルが与えられた写真の一群が、正解にどれくらい一致するのかを評価する。この評価を各頻出ラベルについて行い、テーマごとに最良のラベルを決定する。そして最後に、精選ラベルを有する写真群の撮影位置分布を地図化し、正解写真の撮影位置分布と乖離していないかを検証する。

3. テーマの選定

本研究では観光の様々なシーンを想定し、以下の13テーマを地図化対象として検討した。

- ・ 街中で偶然に遭遇するようなもので、観光対象になりうるもの
 - …着物 (着物を着た人), 猫, コスプレイヤー
- ・ 季節性のある自然資源で、観光対象になるもの
 - …梅, 桜, 銀杏
- ・ イベント関連資源
 - …クリスマスツリー, ハロウィン (ハロウィンの飾り付け, ならびに仮装した人々)
- ・ 遠望対象となる観光資源
 - …富士山, 東京スカイツリー
- ・ 時間帯や気象条件により鑑賞価値が変動する景観
 - …夕焼け, 夜景
- ・ 観光客 (とくに訪日外国人) の注目が高い飲食物
 - …ラーメン

上記は基本的には [6]で挙げたテーマを継承しているが、[6]では「朝焼け」「虹」「蝶」「ホテル」「紫陽花」「楓」の各テーマについては投稿事例があまりなかったと報告されていたため、今回は対象にしなかった。一方、「着物」「クリスマスツリー」「富士山」についても投稿事例が少ないと報告されているが、観光における重要度を考え、本研究の対象とした。

4. テーマ別の検索結果とテーマ適合状況

テーマごとに使用した検索キーワードと抽出状況を表1にまとめた。検索の結果、得られた写真枚数では「桜」が最も多く (11194 枚)、ついで「コスプレイヤー」(6866 枚)、「東京スカイツリー」(6466 枚) となった。また、検索結果中の適合写真枚数についても、「桜」「コスプレイヤー」「東京スカイツリー」の順となった。一方、「着物」「クリスマスツリー」「銀杏」「富

士山」については、検索結果枚数・適合写真枚数ともに低い値となった。

適合率に注目すると、最高値は順に「コスプレ」(95%)、「夕焼け」(90%)、「梅」(88%)であった。これらのテーマは、被写体情報を用いた精選を行わなくても、キーワード検索の結果だけでかなり正確な観光資源地図が描けそうである。一方、適合率が低いのは、「富士山」(41%)、「猫」(48%)、「東京スカイツリー」(60%)、「着物」(60%)、「クリスマスツリー」(61%)、「ハロウィン」(67%)、「ラーメン」(67%)の順であった。「富士山」と「東京スカイツリー」の適合率が低い主原因は、どの写真にも同じタグ群 (旅行中に訪れた観光スポットをすべて列挙したもの) を付与するような「面倒臭がり」な投稿者が見られたためである。一方、「猫」については猫カフェや原宿キャットストリートの写真、「着物」については着物店や和テイストのコスプレ写真、「クリスマスツリー」に関しては一般の電飾やライトアップの写真、「ハロウィン」についてはハロウィン関連のポスターや同時期の東京ディズニーランドの写真、「ラーメン」についてはラーメン店の外観や自炊ラーメンの写真の投稿が多く混じっていた。

表1の一番右に示した「投稿者ジニ係数」は投稿者の投稿枚数の不平等の指標であり、1に近いほど一部の投稿者による大量投稿で占められていることを示す。最高値を示したのは「コスプレイヤー」(0.92)であり、このテーマには熱烈な写真投稿者がいることが伺える。次いで「東京スカイツリー」(0.83)、「梅」(0.82)、「猫」(0.80)も高い値となった。一方、「着物」(0.46)、「富士山」(0.50)は比較的低い値となったが、これは様々な旅行者が概ね少しずつ撮るようなテーマであるためと考えられる。

テーマ	検索キーワード	検索結果枚数	適合写真枚数	適合率	投稿者ジニ係数
着物	着物, kimono	367	221	60.2%	0.46
猫	猫, cat	3179	1514	47.6%	0.80
コスプレイヤー	コスプレ, cosplay	6866	6543	95.3%	0.92
梅	梅, cherry blossom	1041	918	88.2%	0.82
桜	桜, cherry blossoms	11194	9345	83.5%	0.77
銀杏	銀杏, ginkgo	529	408	77.1%	0.68
クリスマスツリー	クリスマスツリー, Christmas tree	400	243	60.8%	0.64
ハロウィン	ハロウィン, halloween	934	630	67.5%	0.77
富士山	富士山, Mt. Fuji	753	305	40.5%	0.50
東京スカイツリー	東京スカイツリー, Tokyo Skytree	6466	3908	60.4%	0.83
夕焼け	夕焼け, sunset	2405	2158	89.7%	0.59
夜景	夜景, night view	3046	2417	79.3%	0.78
ラーメン	ラーメン, ramen	2131	1420	66.6%	0.67

表1. テーマ別の検索キーワードと検索結果

5. テーマ対応ラベルの評価と選定

テーマごとの頻出ラベル（Google Cloud Vision API が各写真に与えたラベルのうち、頻出上位 5 位までのもの）と、それぞれのラベルの有無を写真の精選基準に使ったときの各評価値（適合率等）を表 2 にまとめた。

テーマ	頻出ラベル	選別後のうち適合		適合率	再現率	F尺度
		写真枚数	写真枚数			
着物	costume	88	63	71.6%	28.5%	0.41
	clothing	82	57	69.5%	25.8%	0.38
	pedestrian	39	21	53.8%	9.5%	0.16
	kimono	29	28	96.6%	12.7%	0.22
	art	26	4	15.4%	1.8%	0.03
猫	mammal	1829	987	54.0%	65.2%	0.59
	cat	1450	815	56.2%	53.8%	0.55
	vertebrate	1425	758	53.2%	50.1%	0.52
	cat like mammal	698	390	55.9%	25.8%	0.35
	small to medium sized cats	634	335	52.8%	22.1%	0.31
コスプレイヤー	costume	2586	2565	99.2%	39.2%	0.56
	clothing	2505	2481	99.0%	37.9%	0.55
	cosplay	1839	1830	99.5%	28.0%	0.44
	anime	1683	1677	99.6%	25.6%	0.41
	festival	250	250	100.0%	3.8%	0.07
梅	flower	803	734	91.4%	80.0%	0.85
	plant	784	716	91.3%	78.0%	0.84
	branch	582	528	90.7%	57.5%	0.70
	blossom	471	428	90.9%	46.6%	0.62
	flora	316	289	91.5%	31.5%	0.47
桜	flower	9577	8458	88.3%	90.5%	0.89
	plant	9013	7920	87.9%	84.8%	0.86
	blossom	6863	6216	90.6%	66.5%	0.77
	cherry blossom	5974	5530	92.6%	59.2%	0.72
	branch	4913	4353	88.6%	46.6%	0.61
銀杏	tree	363	325	89.5%	79.7%	0.84
	plant	244	217	88.9%	53.2%	0.67
	season	182	173	95.1%	42.4%	0.59
	woody plant	169	159	94.1%	39.0%	0.55
	autumn	138	132	95.7%	32.4%	0.48
クリスマスツリー	christmas decoration	227	200	88.1%	82.3%	0.85
	christmas tree	182	169	92.9%	69.5%	0.80
	christmas lights	131	111	84.7%	45.7%	0.59
	decor	70	69	98.6%	28.4%	0.44
	tree	60	53	88.3%	21.8%	0.35
ハロウィン	festival	199	190	95.5%	30.2%	0.46
	costume	192	186	96.9%	29.5%	0.45
	clothing	160	156	97.5%	24.8%	0.39
	advertising	130	8	6.2%	1.3%	0.02
	banner	69	3	4.3%	0.5%	0.01
富士山	horizon	182	129	70.9%	42.3%	0.53
	skyline	181	117	64.6%	38.4%	0.48
	city	128	80	62.5%	26.2%	0.37
	sky	91	57	62.6%	18.7%	0.29
	atmospheric phenomenon	81	47	58.0%	15.4%	0.24
東京スカイツリー	skyscraper	2815	2542	90.3%	65.0%	0.76
	skyline	2665	2239	84.0%	57.3%	0.68
	tower	2072	1997	96.4%	51.1%	0.67
	landmark	1955	1918	98.1%	49.1%	0.65
	city	1744	1146	65.7%	29.3%	0.41
夕焼け	horizon	1164	1095	94.1%	50.7%	0.66
	skyline	950	881	92.7%	40.8%	0.57
	sky	734	689	93.9%	31.9%	0.48
	city	620	569	91.8%	26.4%	0.41
	afterglow	606	576	95.0%	26.7%	0.42
夜景	night	785	770	98.1%	31.9%	0.48
	skyline	761	660	86.7%	27.3%	0.42
	lighting	725	673	92.8%	27.8%	0.43
	city	689	623	90.4%	25.8%	0.40
	cityscape	562	523	93.1%	21.6%	0.35
ラーメン	food	1607	1262	78.5%	88.9%	0.83
	dish	1554	1239	79.7%	87.3%	0.83
	cuisine	1504	1210	80.5%	85.2%	0.83
	noodle	857	723	84.4%	50.9%	0.64
	soup	690	559	81.0%	39.4%	0.53

表 2. 抽出されたテーマ対応ラベルとその評価値

まずラベルの内容に注目すると、たとえばテーマ「着物」に対しラベル「kimono」が 4 位、テーマ「猫」に対しラベル「cat」が 2 位になっているように、各テ

ーマに直接対応するようなラベルが必ずしも最頻出とは限らない結果となった。一方、最頻出のラベルは、たとえば「猫」に対して「mammal」、「梅」「桜」に対して「flower」のように、包括的な概念を指す名詞であった。

次に選別前後の適合率の変化を比較すると、ほとんどのラベルについて適合率の向上が見られた。とりわけ上昇度合が著しいのは、「着物」に対する「kimono」（60.2%→96.6%）、「銀杏」に対する「autumn」（77.1%→95.7%）、「ハロウィン」に対する「clothing」（67.5%→97.5%）、「富士山」に対する「horizon」（41%→70.9%）、「東京スカイツリー」に対する「landmark」（60.4%→98.1%）である。また、元々高い適合率がさらに向上した例としては、「コスプレイヤー」に対する「anime」（95.3%→99.6%）、「夕焼け」に対する「afterglow」（89.7%→95%）、「桜」に対する「cherry blossom」（83.5%→92.6%）などがみられた。これらの結果は、機械判定による被写体情報を考慮することで、テーマに即した画像の精選が一般にうまくいくことを示唆している。ただし、例外的に「猫」に対しては、どのラベルを用いても適合率の改善があまり見られなかった。これは「猫」画像と判定されても、それがテーマに即した「屋内にいる猫」なのか、それとも「室内にいる猫」はわからないためである。

一般的に情報検索の評価では、適合率だけでなく、再現率や F 尺度（適合率と再現率の調和平均）も用いられる。そこで、これらの計算も行ったところ（表 2）、いずれのテーマも適合率と再現率のトレードオフが確認された。適合率最大となるラベルと F 尺度最大となるラベルとが一致したテーマは「夜景」（ラベル：night）と「富士山」（ラベル：horizon）の 2 例のみだった。

観光客向けの地図を作成することを考えると、誤った情報による誤誘導は可能な限り避けるべきである。そのため、もし精選の結果それなりの適合写真が残存するのであれば、たとえ再現率が低くても、より適合率が高いほうが好ましい地図が得られると考えられる。他方で、精選後に残る写真枚数が少なすぎると、地図（ヒートマップ）としての見栄えが不十分になりかねない。そこで、様々なケースについて観察を行い、「東京都心を対象に 300 枚の撮影位置データがあれば鑑賞に堪える地図ができる」という経験則を見出し、これをふまえて「テーマごとの最頻出ラベル 5 つのうち、精選後の適合写真枚数 300 枚以上となるもので、適合率が最大になる頻出ラベルを用いて写真精選を行い、その結果得られた写真の撮影位置分布を地図化する」という方針を設けた。この方針のもとで選ばれたラベルは以下の通りである。

- ・ 着物：該当なし

- ・ 夕焼け：afterglow
- ・ 梅：flower
- ・ 富士山：該当なし
- ・ コスプレ：cosplay
- ・ 桜：cherry blossom
- ・ 銀杏：tree
- ・ 猫:cat
- ・ スカイツリー：landmark
- ・ ハロウィン：該当なし
- ・ ラーメン：noodle
- ・ クリスマスツリー：該当なし
- ・ 夜景：night

6. 地図化

各テーマについて、全適合写真（正解データ）の撮影位置分布と、5章で選んだ頻出ラベルの有無を基準に精選した写真の撮影位置分布とをそれぞれ地図化し、比較を行った。なお、「着物」「クリスマスツリー」「ハロウィン」「富士山」については、どの頻出ラベルを用いても適合画像数が300枚残らないため、地図化は行わなかった。

まずテーマ「猫」の適合写真の空間分布（図1）については、①秋葉原駅周辺、②中野駅北側、③代々木八幡、④井の頭公園、⑤登戸駅北側、⑥生田緑地にて強いピークが、また⑦谷中、⑧池袋駅東側、⑨代々木上原でもそれに準じたピークが生じた。⑦は地域猫が観光資源化している地域であり、③④⑥も野良猫が愛されている地域である。一方、①②⑧については、野良猫が多いとは思えない都市地域だが、猫カフェが多いという共通特徴があるため、猫好きが屋外で偶発的に出会った野良猫を写真投稿する機会が必然的に多くなるのではないかと考えられる。

次に、「猫」または「cat」で検索した Flickr の写真群に対し、Google Cloud Vision API がラベル「cat」を与えた写真の撮影位置分布（図2）を見ると、大塚駅周辺、新宿駅東口、原宿駅周辺にもピークが見られる。この地域で撮影された写真を見てみると、猫カフェ店内の写真が数多く見られた。この原因は、先述の通り、今回の手法では屋内の猫は必然的に除外できないためであると考えられる。

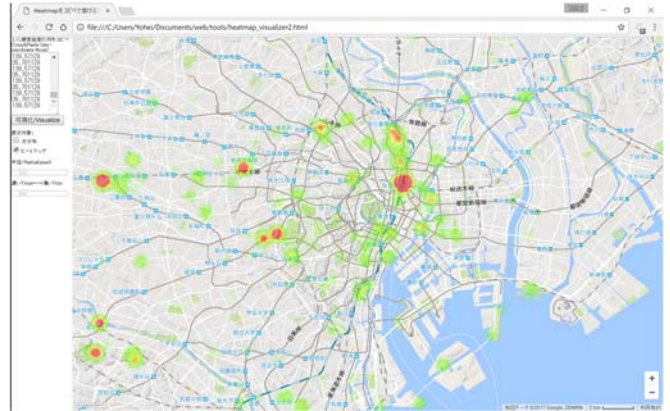


図1. 「猫」写真の撮影位置分布

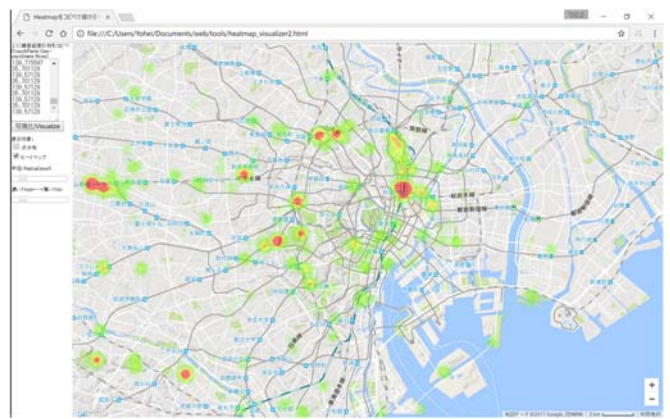


図2. ラベル「cat」で精選された猫写真の撮影位置分布

次に「コスプレイヤー」の適合写真の撮影位置分布については、大変強いピークが東京ビッグサイト周辺に、それ次ぐピークが秋葉原駅周辺に見られた（図3）。ラベル「cosplay」による精選結果も、ほぼ相違ない空間分布が得られた（図は省略）。なお、レインボーブリッジ北側の会場にも写真の投稿が見られるが、これは位置情報の誤りによるノイズであった。

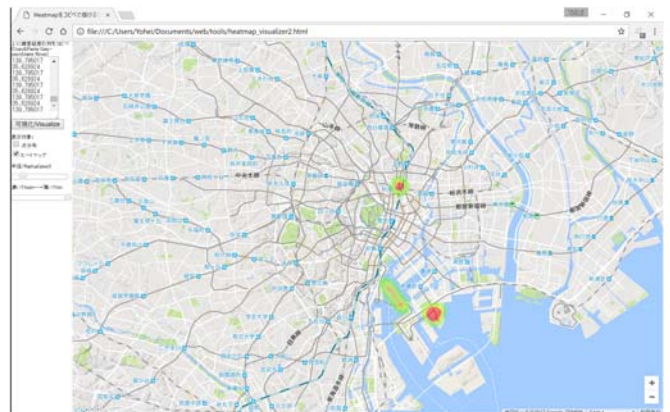


図3. 「コスプレイヤー」写真の撮影位置分布

「梅」の適合写真の撮影位置分布については、北東から順に①北綾瀬駅東側、②押上駅と亀戸駅の間地点、③湯島駅近辺、④後楽園駅近辺、⑤梅ヶ丘駅、⑥西馬込駅周辺駅でピークが見られた(図4)。一方、ラベル「blossom」による精選結果も、ほぼ相違ない空間分布が得られた(図は省略)。②③④⑥はいずれも梅の名所として観光客を集める亀戸天神、湯島天神、後楽園、池上梅園に対応している。また①は大谷田公園、⑤は羽根木公園であり、調べてみると、いずれも梅の木を多く有する公園であることがわかった。

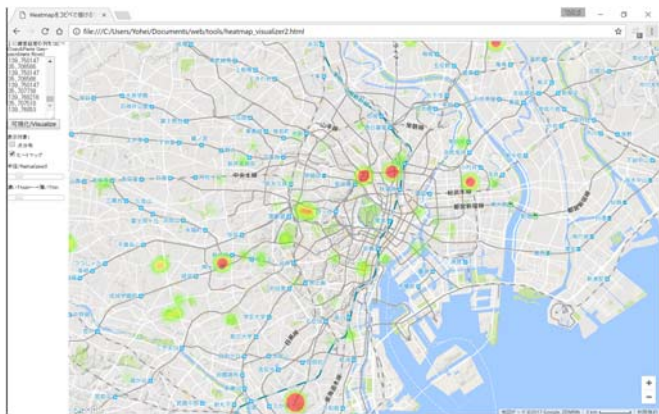


図4. 「梅」写真の撮影位置分布

「桜」の適合写真の撮影位置分布については、北から順に①駒込駅周辺、②上野駅西側、③皇居の北西側、④新宿駅南東、⑤原宿駅の西方、⑥中目黒駅周辺にピークが現れた(図5)。一方、ラベル「cherry blossom」による精選結果についても、ほぼ相違ない空間分布が得られた(図は省略)。①~⑥はいずれも桜の名所として知られる六義園、上野公園、千鳥ヶ淵、新宿御苑、代々木公園、目黒川と対応している。

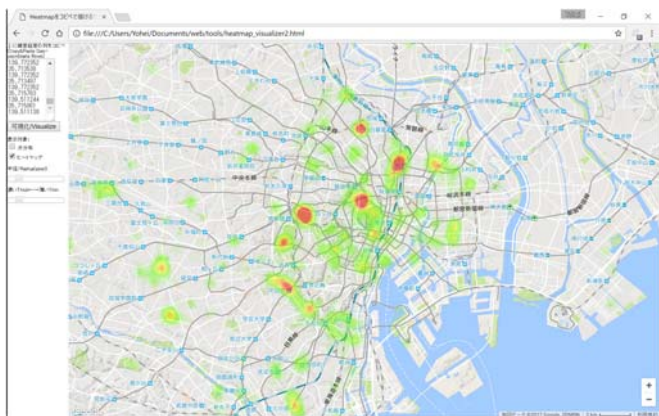


図5. 「桜」写真の撮影位置分布

「銀杏」の適合写真の撮影位置分布については、北東から順に①上野駅西側、②本郷、③皇居の北、④皇

居の南南西、⑤新宿駅の南東、⑥外苑前駅周辺、の各地区にピークが現れた(図6)。これらは銀杏の名所として知られる上野公園、東京大学本郷キャンパス、靖国神社、国会議事堂前、新宿御苑、神宮外苑に対応している。一方、ラベル「tree」による精選結果については、小石川植物園近辺の色がやや濃くなるという違いが得られた(図7)。そこでこの地域の写真を見てみると、紅葉していない銀杏の写真が含まれていた。今回は「銀杏」または「ginkgo」で検索した写真群に対し、画像判別で「tree」というラベルを与えられた写真を抽出したため、紅葉時期以外の銀杏は必然的に除外できない。この問題が植物園という場所で顕在化したと考えられる。

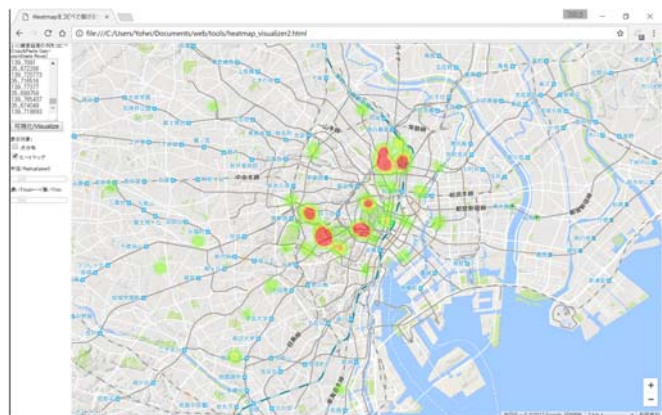


図6. 「銀杏」写真の撮影位置分布

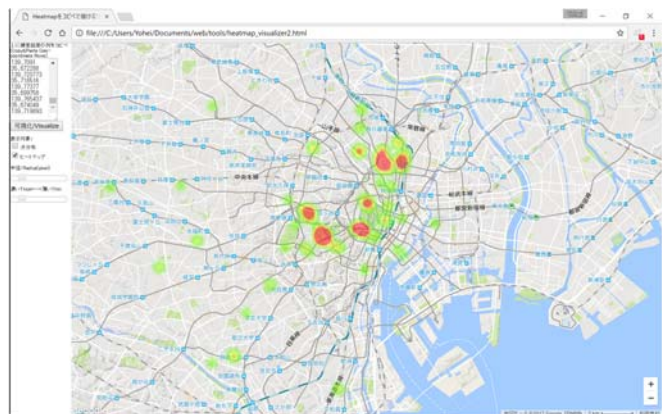


図7. ラベル「tree」で精選された銀杏写真の撮影位置分布

「東京スカイツリー」の適合写真の撮影位置分布については、①スカイツリーが立地する押上から、その西の浅草にかけて、特に隅田川沿いと横十間川沿いに連続した集積が見られたほか、②少し離れた荒川沿いでもピークが見られた(図8)。一方、ラベル「tower」による精選結果についても、ほぼ相違ない空間分布が得られた(図は省略)。東京スカイツリーは、巨大建造

物でありながら、建物密集地帯に存在するため、写真撮影箇所は川沿いに偏る結果になったと考えられる。

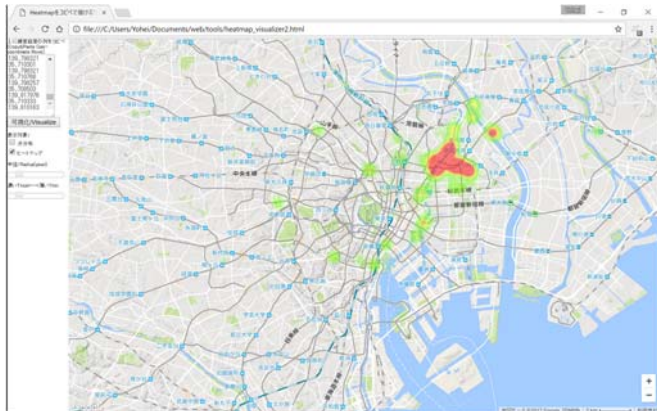


図 8. 「東京スカイツリー」写真の撮影位置分布

「夕焼け」の適合写真の撮影位置分布については、①東京都庁、②文京シビックセンター、③浜松町貿易センタービル、④東京スカイツリー、⑤お台場海浜公園付近にピークが見られた(図 9)。①～④は展望が有名な各高層建築物群であり、⑤もレインボブリッジ越しに眺める都市景観で知られる。一方、ラベル「afterglow」による精選結果では、①～④が消え、代わりに東京西部の郊外の三カ所に強いピークが生じた(図 10)。この原因は、高層ビル群を含むような夕焼けの写真に対し「afterglow」というラベルがあまり付与されなかったためではないかと考えられる。

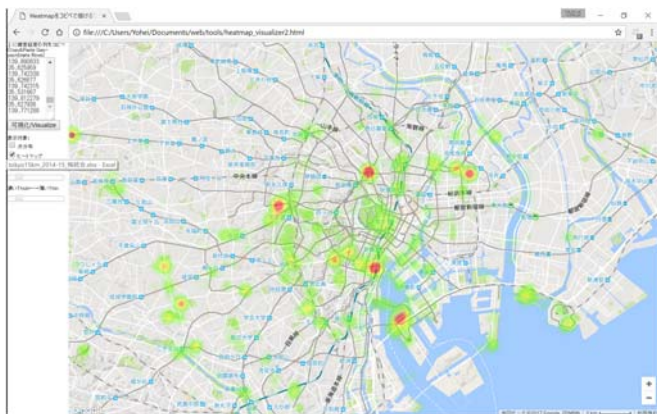


図 9. 「夕焼け」写真の撮影位置分布

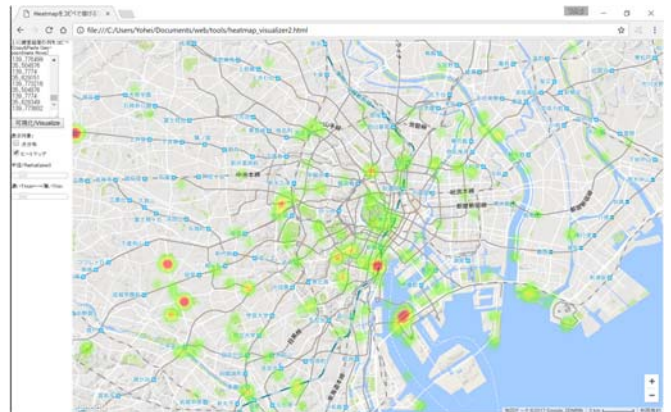


図 10. ラベル「afterglow」で精選された夕焼け写真の撮影位置分布

「夜景」の適合写真の撮影位置分布については、①新宿副都心、②渋谷、③六本木、④丸の内、⑤東京スカイツリーに強いピークが現れ、ついで⑥東京タワー、⑦浜松町、⑧浅草、⑨お台場にもそれに準じたピークが現れた(図 11)。「夕焼け」の結果と比較すると、渋谷、六本木、丸の内という繁華街が現れるのが特徴である。この差は、「夕焼け」は西側の展望を前提にしているが、「夜景」は展望がなくても、都市的な景観があれば成立するためだと考えられる。一方、ラベル「night」による精選結果は様相がやや異なり、浅草やお台場のピークが強まる一方で、他地域の集積が相対的に低下した(図 12)。これはおそらく浅草やお台場が夜になると相対的に暗く「night」らしさが生じるのに対し、他地域は相対的に明るい繁華街で、写真に「night」というラベルが付与されにくいためだと考えられる。

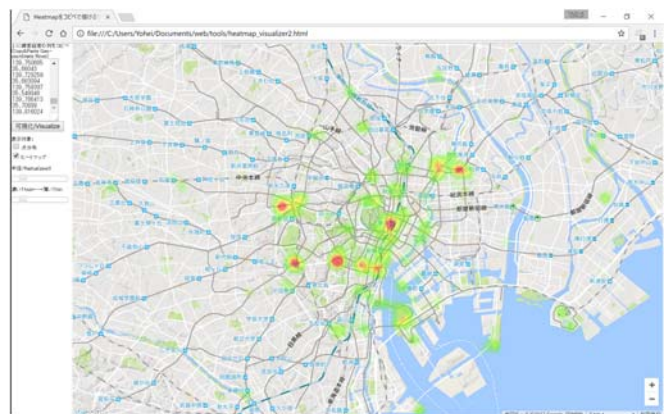


図 11. 「夜景」写真の撮影位置分布

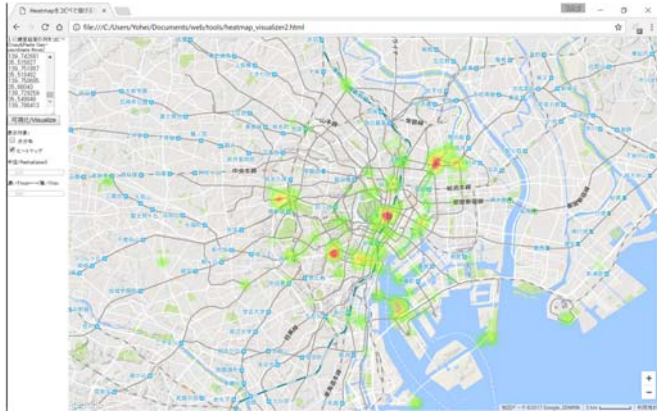


図 12. ラベル「night」で精選された夜景写真の撮影位置分布

最後に「ラーメン」の写真撮影箇所については、多極的な分布がみられた。特徴すべきは、新宿、渋谷、池袋といった繁華街よりも、桜台、仙川、新代田といった郊外の駅周辺に強いピークが存在している点である。一方、ラベル「noodle」による精選結果についても、相違ない空間分布が得られた（図は省略）。これらの郊外の駅周辺にはいずれも「ラーメン次郎」の支店があり、当該地域の投稿写真を見てもこの店と思しき写真が多い印象であった。ただし他の支店のある地域では、さほど強いピークは生じていなかった。

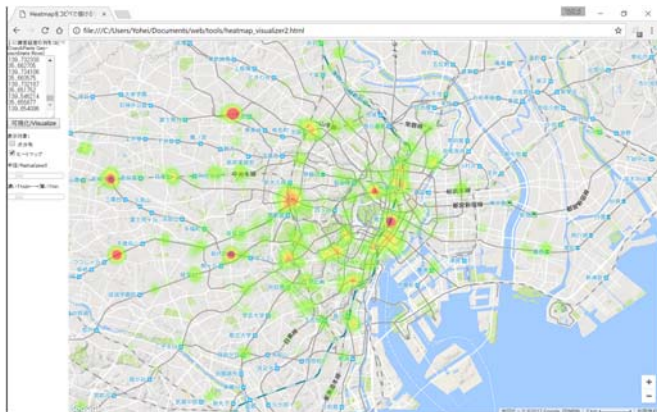


図 13. 「ラーメン」写真の撮影位置分布

7. まとめ

本研究では、テーマ別の観光資源地図を効率よく作成するために、まず Flickr でテーマに即したキーワード検索を行い、そこで得られた写真を Google Cloud Vision で機械判定し、特定のラベルが付与されたか否かをもって精選する手法を提案し、13 種類の様々なテーマについて検証を行った。この結果得られた示唆をまとめると、以下の通りである。

- ・ Flickr 上でのキーワード検索の際に観光スポット名称を用いると、訪問先をタグに羅列する投稿者の投

稿に攪乱され、適合率が低くなる傾向にある（例：「東京スカイツリー」「富士山」）

- ・ 一部の人々が熱狂的に写真撮影を行うようなテーマ（今回の例では「コスプレイヤー」「梅」「スカイツリー」、他に考えられる例としては「鉄道」「飛行機」「工場夜景」「廃墟」「野鳥」など）は、少数投稿者の大量投稿によって空間分布に偏りが生じる恐れがあるが、その特殊性ゆえ、キーワード検索だけで適合度が高くなることも期待できる
- ・ 地図化にはパブリック空間（屋外をはじめ観光客が入ることのできる空間）で撮影された投稿写真の理由が前提となるため、プライベート空間にも多く見られるような事物をテーマにすると、適合率が低くなりやすい（今回の例では「猫」、他に考えられる例では、自炊が一般的に行われる料理など）

以上を踏まえると、「夕焼け」「夜景」「桜」「銀杏」など、観光客が一般に撮影を行うような普遍的テーマで、かつプライベート空間では撮影されないようなものが、地図化のテーマとして向いていると考えられる。さらに、精選に用いるラベルは、適合度の高さだけを理由に選ぶと、そのニュアンスの微妙なずれから、得られる分布図が多少偏る危険性がある（例：「夕焼け」に対する「afterglow」、 「夜景」に対する「night」）。このため、意味や再現性に考慮しながら、慎重にラベルを選択することが望ましいと考えられる。

なお本研究では、写真の精選にあたって単一のラベルの利用のみを考えたが、複数のラベルを複合的に用いることも考えられる。たとえば、「猫」写真の精選にあたっては、「cat」のようなキーワードとともに、屋内・屋外を特徴付けるキーワード（たとえば「room」など）を併用すれば、より写真精選の精度が挙がるのが期待できる。このような複数ラベルの組み合わせによる写真精選は、今後の研究課題である。

謝辞：本研究の遂行にあたっては、総務省 SCOPE の支援を利用した（研究課題名「訪日外国人旅行者を対象とした地域情報マイニング技術の研究開発」代表：難波英嗣）。

参考文献

- [1] Shirai, M., Hirota, M., Yokoyama, S., Fukuta, N., and Ishikawa, H.: Discovering Multiple Hot Spots using Geo-tagged Photographs. ACM SIGSPATIAL, 2012.
- [2] 長尾 光悦: CGM をベースとした観光情報提供方法に関する考察. 観光情報学会第 9 回全国大会発表概要集, 26-27 (2012).
- [3] 斎藤一・横川祥司: 環状誤字所と位置情報付きツイート分析に基づいたアプリケーション「EmoNavi」の観光利用の検討. 北海道情報大学紀要 28(1), 103-110, 2016.
- [4] 倉田陽平: 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, CD-ROM, 2013.

- [5] Yohei Kurata : Potential-of-Interest Maps for Mobile Tourist Information Services. ENTER 2012, pp. 239-248, 2012.
- [6] 真田風 : 写真共有サイトの位置情報付き写真データを利用したテーマ別観光マップの作成. 首都大学東京大学院都市環境科学研究科観光科学域修士論文 2015.