

ビッグデータを用いた観光行動分析手法の開発 ～「写真」に着目して～

大室 ひな¹・神田 佑亮²・助永 雅紀³

¹学生会員 呉工業高等専門学校専攻科 (〒737-8506 広島県呉市阿賀南2-2-11)
E-mail: S18-afxz@kure.kosen-ac.jp

²正会員 呉工業高等専門学校環境都市工学分野 教授 (〒737-8506 広島県呉市阿賀南2-2-11)
E-mail: y-kanda@kure-nct.ac.jp

³非会員 中部電力株式会社 (〒461-8680 愛知県名古屋市中区新町1番地)
E-mail: Sukenaga.Masaki@chuden.co.jp

近年、全国的に地域活性化への取り組みが進められている。その一つとして観光地の活性化が挙げられる。各地域でアンケートなどによって、マーケティングが行われるべきであるが、基礎であるマーケティングが十分行われているとは言い難い。その課題として、マーケティングへの認識不足、アンケート等の労力に対する負担感などが挙げられる。一方で、観光に対し、人々の行動や印象の記述が、SNSにより自ら発信されるようになり、これらのデータがビッグデータとして蓄積されつつある。また、印象を写真で投稿するSNSも登場している。本研究ではこれらのデータを活用した観光マーケティング分析の手法開発や活用可能性について論じる。

Key Words : SNS, bigdata,, travel pattern analysis, picture mining

1. はじめに

地域活性化の一環として、観光振興に取り組む地方が非常に多い。この大きな背景として、訪日外国人旅行者、いわゆるインバウンド観光客の受入環境の整備・充実を総合的に推進することで、訪日外国人旅行者の訪問を促進するとともに、満足度を高め、リピーターの増加を図ること、そしてそれに伴い地域経済の活性化をめざす取り組みを国家的政策として行なっていることが挙げられる。

こうした観光戦略を立案、展開する際に重要なアクションの1つは、継続的かつ的確なマーケティングである。地域を訪れる旅行者を増加させるためには、多くの一般企業が財・サービスを市場に提供し、ニーズやウォンツを満たすための活動を実施する場合と同様に、地域の観光資源を発掘し観光商品に組み込み、市場に対して積極的・戦略的にマーケティング活動を行っていくことが極めて重要である。しかしながら、この重要なマーケティング活動が十分に行われているとは言い難いのが現状である。その原因として、マーケティング活動に関するノウハウや経験の不足、加えて多忙さからマーケティン

グに人材を十分に費やせない、あるいは十分な資金が捻出できない等の課題がある。あるいはマーケティングの一環として観光客向けにアンケート調査を行なったとしても、実施日が限られた数日であったりと、調査手法での課題も抱えている。

一方、最近様々SNSや口コミサイトを通じ、観光客自身が訪問した観光地の印象や評価を「レビュー」し、インターネット上に自発的かつ気軽に書き込むようになって来た。特に最近ではカメラ付き携帯電話、スマートフォンの普及により、写真付きで投稿されるようになって来ている。また、来訪する観光客は、これらのSNSや口コミサイトに掲載されている情報の参照度が極めて高くなってきている。特にこれらのデータは他のビッグデータと同様に、時期を問わずに所定のフォーマットに従い自発的に投稿されたデータがインターネット上に自動的に蓄積される、マーケット分析者側の視点から換言すると、現在のマーケティング調査の主体であるアンケート調査の課題である、調査時期の偏りや回答の受動性といった点では問題のないデータを大量に得られる可能性を秘めている。

ただし、これらのデータを用いるには、いくつか課題が挙げられる。1点目は投稿の偏りである。回答者の属性や投稿時期に偏りがなくどうか、検証を行っておく必要がある。2点目は利用に際し加工が必要である。投稿された内容はマーケティング利用を主としたものではないため、記載された文章や写真等を、マーケティング分析に活用できるように加工を行うとともに、分析手法の確立が必要である。特に、コメントや写真は定量的なデータではないため、これらのデータを定量的に処理する手法が求められるであろう。また社会への実装を考えると、できるだけ簡易な方法であることも求められる。

このような背景から、本研究では観光活性化に資する効果的かつ省力化できるマーケティング手法の確立を目指し、観光SNSや口コミサイト等のビッグデータ、特に「写真」に焦点を当て、観光行動分析を行う手法の開発を目的とする。

2. 関連研究のレビュー

「写真」を表現方法の1つと捉え、撮影された写真を分析する試みは写真投影法と称され、人類学や社会学の領域で着目され始めた¹⁾。都市計画分野では、人々の景観や環境に対する認知を明らかにする、ランドスケープ評価の研究、生活空間や環境の評価や認識の分析に用いられている。また、発達心理学の分野では、個人の内面世界の把握に用いられている。上記のように様々な分野で写真投影法の適用が試みられているが、共通して写真の読み手の主観が結果の解釈に入り込む可能性が指摘され、また写真について定量的に分析していない点もが課題として指摘されている。

近年では、直感的な「パッと見」の印象から解釈を行う「ピクチャマイニング」が提唱され、手法の確立が試みられている。例えば落原ら(2012)²⁾は、高速道路のサービスエリア、パーキングエリアの評価にピクチャマイニングの適用を試みている。インターネットモニターによる写真調査によるデータを用いて、写真からキーワードを抽出しそれを多変量解析を行っているが、分析に必要なサンプル数が課題となっている。

こうした既往研究の課題に対し、観光SNSや口コミサイトの情報を用いることで、ビッグデータの長所であるデータの豊富さや、写真とともに自発的に記述された文章を活用することにより、解釈の偏りに対する妥当性の検証がしやすいなど、従前の写真解析の課題をクリアできる可能性を秘めている。

3. ピクチャマイニングに用いるデータの概要と特性

本研究では、世界最大閲覧数の観光口コミサイト「Trip Advisor」に投稿されたデータを用いる。「Trip Advisor」

に投稿された情報には個人の属性(性別・年代・国籍または移住地)・訪問した観光地の5段階評価・コメント・写真のデータが含まれている(図-1)。



図-1 Trip Advisorの投稿画面

はじめに、観光SNSデータの特性を分析するため、2016年1月～12月に「宮島」で投稿された240件(内訳:日本人87件,外国人153件)のデータについて、月別訪問数(構成比率)を集計し、既存統計(広島県客数の動向2016)と比較した(図-2)。

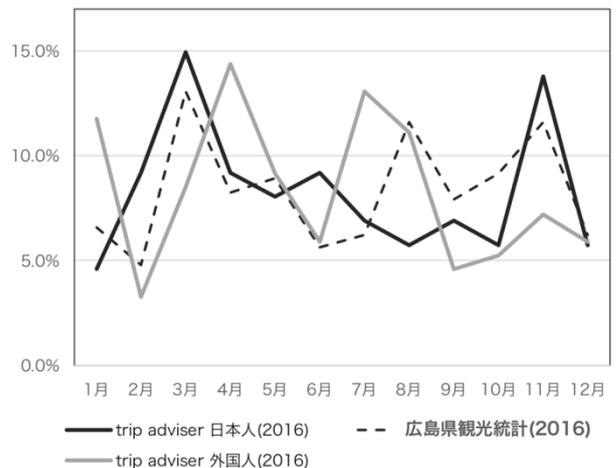


図-2 既存統計の観光入込客数と投稿数の月別比率

既存統計と「Trip Advisor」への日本人の月別投稿比率は概ね同様の傾向となっており、観光SNSデータでの分析は問題ないものといえる。また、外国人の来訪ピークが4月、7-8月であることも読み取れる。

また、「Trip Advisor」では、同一人物の過去の投稿内容も参照できる。すなわち、その人物の観光周遊箇所を把握することが可能となる。ここでは、2017年に「Trip Advisor」に「宮島」で写真つきで投稿した外国人(102人)

を対象に、一連の旅行で日本国内の他地点で投稿した都道府県を集計した (図-3)。

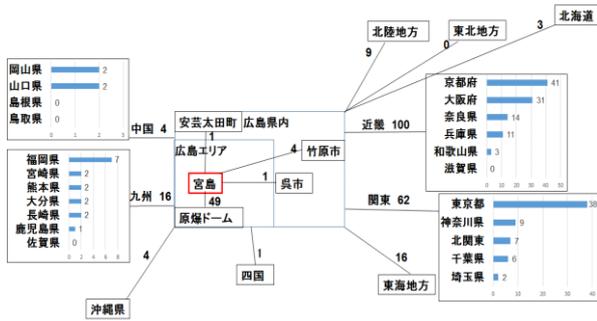


図-3 観光周遊行動分析 (宮島訪問客)

広島県内では、原爆ドームを投稿した人が 49 名 (48%) と非常に多く、一方で他の観光スポットは少なかった。中国地方や九州への回遊も少なかった。逆に多いのは大阪・京都・東京であった。京都は金閣寺など外国人からの関心が高く、大阪はたこ焼きなどの食文化やユニバーサルスタジオジャパンに関心が高かった。東京は日本の中心地として有名であり、東京タワーや東京スカイツリーへの観光が多かった。広島を訪れる外国人は海外日本へのゲートウェイ空港が東京 (羽田・成田) や関西国際空港となり移動していることも推察される。

4. 観光マーケティングのための SNS データを活用したピクチャマイニング手法の構築と検証

(1) 分析手法の構築と妥当性の検証

ピクチャマイニング分析のデータとして「Trip Advisor」から写真を収集し、写真から直感的に思い浮かぶ単語の抽出を行い、名詞・動詞・形容詞の三品詞に分類した。ピクチャマイニングの手順として次の作業を行った。

まず、写真を一枚につき 20 秒程度投影する。そして、その写真を見て想起される単語を 3 品詞別に 20 秒程度で書き出す。

このような方法で、40 名が投稿した写真について、6 名の分析者が抽出した単語の一致性を検証した。

はじめに、6 人で同一の写真を見て単語を抽出し、抽出された単語の 3 品詞ごとの一致率を検証する (図-4)。ここで言う一致率とは、各写真に付き、同じ単語を何人が挙げたかというものである。

このグラフからもわかるように、名詞が一致した単語の比率が一番大きくなっている。6 人の半数以上である 3 人以上の一致率に着目すると特に名詞は一致率が大きく、写真から抽出したときの各個人による認識の違いが少ないといえる。要因と推察されることとして、名詞は分析者により表現はばらつきにくい、形容詞や動詞はほぼ意味が同じであっても表現が多様であるためと考え

られるためである。

次に、抽出された単語について、写真と同時に投稿されたコメント (文章) と一致しているかどうかを検証する。観光口コミデータを用いる利点として、当該観光地の評価がコメントと写真の 2 つの表現方法で記載されているため、写真による印象評価の妥当性を検証できる、すなわち従来の写真投影法で指摘されていた、分析者の主観性によるバイアスに対して検証できることが利点として挙げられる。

検証の方法は、投稿されたコメントの文章を単語レベルに分解する形態素解析を行い、写真で抽出した単語が、投稿されたコメントにもあるかどうかを検証した。なお、分析の対象とした投稿は、「Trip Advisor」(宮島 (厳島神社) で検索した) の写真付きの投稿で、日本人は 2015 年 5 月~2017 年 9 月の 229 件の投稿 (写真 661 枚) で、外国人は 2015 年 5 月~2017 年 10 月の 310 件の投稿 (写真 1229 枚) である。分析結果を図-5 に示す。

写真と文章が一致していた単語は全体の 55% であった。一方、「写真・文章不一致」が 35% を占めているが、これは、写真から抽出された内容とコメントの内容 (文脈) がそもそも一致していないものであった。解釈の違いではなく、投稿者が表現方法の違いにより別の内容を表現したものと思われる。純粋な解釈の間違いは「ミスリード」であり、全体の 5%、「写真・文章不一致」を除外すると、全体の 1 割程度であった。この点から、写真から読み取った要素は、概ね偏った解釈をしていないと言えることができるであろう。

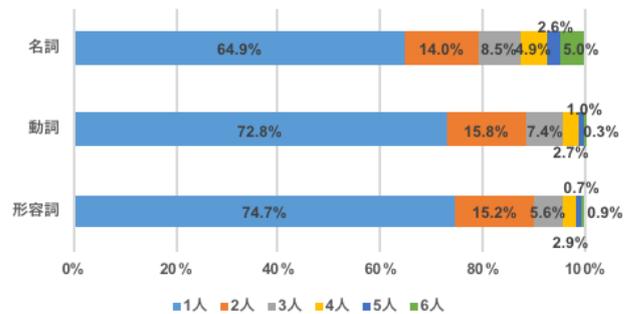


図-4 ピクチャマイニングの一致率の検証

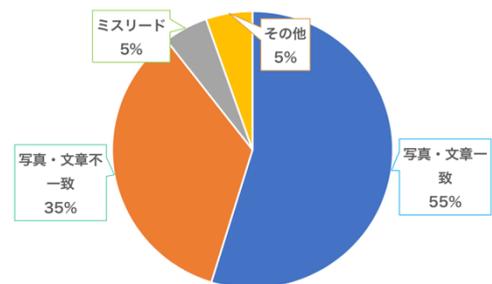


図-5 文字情報との一致率

5. ピクチャマイニング手法を用いた観光資源評価分析のケーススタディ

(1) 要素の出現頻度による観光資源分析

4.で構築した手法を用いて、宮島を訪れる観光客が「Trip Advisor」に投稿した写真から要素抽出(名詞)を行った。分析の対象データは、文字情報との一致性を検証した539件の投稿である。

日本人の投稿の分析結果を図-6に、外国人の結果を図-7に示す。

日本人については、鳥居・海・人などの単語が上位に挙がっている。ピクチャマイニングの日本人の結果からは宮島の魅力や特徴としては、鳥居と海というのは大きなものなのではないかと推察される。

外国人の写真から抽出した単語の結果を見ると、海・鳥居・人など上位に挙がっている単語は日本人の結果とほぼ同じであるが、日本人にはない「ロープウェイ」という単語が上位に出現したりしている。また、外国人を対象とした分析は「石灯籠」や「狛犬」の順位が高い。理由として、外国にはない日本らしい文化を象徴しているものであるためと推察できる。

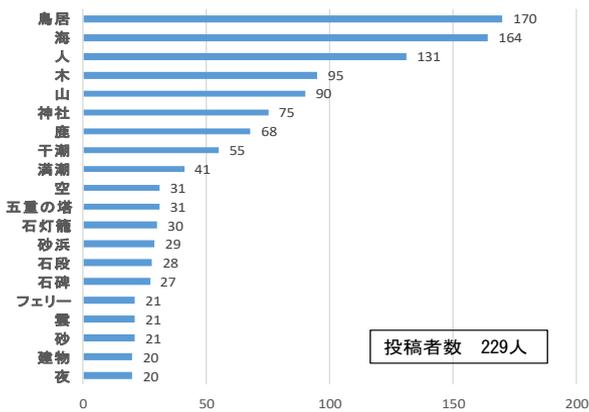


図-6 ピクチャマイニングによる要素の出現頻度 (日本人)

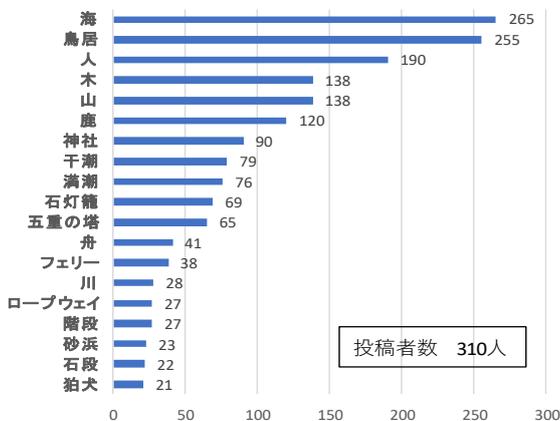


図-7 ピクチャマイニングによる要素の出現頻度 (外国人)

(2) 満足度とその要因に関する定量的分析関分析

「Trip Advisor」上の満足度評価(5段階評価)を目的変数とし、ピクチャマイニングにより抽出された単語と投稿者の個人の属性を説明変数として数量化1類によって分析した。

日本人と外国人とで異なる点として、「鳥居」や「ロープウェイ」・「フェリー」等の乗り物に関する単語が出現すると、日本人は満足度が高くなっていたが、外国人では逆の傾向となり、満足度を低下させる要因となっていた。「石灯籠」・「狛犬」・「石碑」など、日本の文化を感じさせる要素に関しては、外国人は満足度が高くなっていたが、日本人は逆の傾向となった。これらの点から、宮島に関しては日本文化の代表としての側面から、外国人に評価されている可能性が高いと推察できる。

表-1 満足度に関する数量化分析結果

アイテム	日本人		外国人	
	カテゴリー	寄与率	カテゴリースコア	寄与率
海	無	5.4%	-0.0468	6.0%
	有		0.1295	
鳥居	有	1.0%	-0.0164	2.8%
	無		0.0164	
人	有	2.8%	-0.0360	0.7%
	無		0.0551	
木	有	13.1%	-0.0079	16.0%
	無		0.4185	
山	有	6.8%	-0.0597	4.4%
	無		0.1602	
鹿	有	0.8%	0.0105	2.9%
	無		-0.0165	
神社	有	7.7%	-0.1513	3.0%
	無		0.0988	
干潮	有	1.9%	-0.0085	7.5%
	無		0.0516	
満潮	有	8.4%	-0.0336	1.4%
	無		0.2387	
フェリー	有	0.0%	0.0006	5.3%
	無		-0.0007	
川	有	23.0%	0.0138	15.7%
	無		-0.7327	
ロープウェイ	有	12.5%	0.0351	0.2%
	無		-0.3708	
石段	有	8.2%	-0.0033	16.6%
	無		0.2631	
橋	有	4.5%	-0.0146	17.2%
	無		0.1329	
夜	有	3.8%	-0.0092	0.4%
	無		0.1156	
重相関係数			0.4345	0.2223

6. まとめ

本研究の結果から、観光マーケティングで観光地や来訪者の特徴をつかむ際に、観光 SNS に投稿されたビッグデータを集計分析やピクチャマイニング分析により解析することは、観光資源の特徴や訪問者の行動特性をある程度の精度で読み取ることが可能であると示すことができた。しかしながら「写真」が全ての良い点を評価しているわけではないことも示された。このような点から、写真による直感的な解釈と、多変量解析による定量的解釈を組み合わせる行うことが極めて重要であろう。

参考文献

- 1) 林・岡本・藤原(2008)「写真投影法による場所への愛着の測定」、関西学院大学社会学部紀要 (106), pp.15-26,
- 2) 落原・江戸 (2012) 「ピクチャマイニング確立のための基礎的研究～写真調査の類型化と新しい調査方法の課題と方向性～」JMRAアニュアルカンファレンス論文集2012年 p.23-32