

感情的な発言への共感を利用した観光情報の提示手法

菅野 裕基¹ 高田 秀志²

概要: 人は観光を行う前に、あらかじめ観光地についての情報を SNS から収集することがある。従来の研究では、観光地に訪れた人がどのような感情を抱くかについて、発言内容に含まれる感情による分類を行っていた。しかし、分類によって得られた情報自体には他のユーザからの評価が考慮されていないため、それが単なる個人の意見なのか、それとも、多数の人に共通する意見なのかを判別することは難しい。そこで、情報に対する他のユーザからの評価として、発言に対する共感を考慮して情報を提示することで、観光地に対する理解を支援することを本研究の目的とする。本研究では、Twitter を情報源として用い、感情を表現する単語を集めた辞書を構築して観光地についての発言を分類した上で、共感の指標となる「いいね」の数が多い順に並び替えて提示する。評価実験を行った結果、観光地を評価するうえで役立つことがわかった。一方で、観光地とは関係のない感情を含む情報が提示されてしまう場合があるため、観光情報に適した辞書を構築する必要があることが分かった。

A Presentation Method of Tourism Information Using Empathy for Emotional Comments

YUKI SUGANO¹ HIDEYUKI TAKADA²

1. はじめに

人は観光を行う前に、あらかじめ観光地についての情報を調べておくことがある。従来は観光情報誌などから情報収集することがよく行われていたが、最近ではパソコンやスマートフォンなどの情報端末全般の普及によって、インターネットから情報収集することが多くなった。インターネット環境の中でも、SNS には多くのユーザが情報を発信していることから、SNS は新しい情報資源の形として利用されている。約 2 千人の若者のユーザを対象にした SNS 利用に関する意識調査 [1] によると、「SNS を利用してよかったこと」という質問に対して、53.9% のユーザが「役に立つ情報が得られる」と答えている。また、36.1% のユーザが「行きたい場所が増えた」と答えており、観光地に関する情報資源として役立っていることも確認できる。

SNS を活用する理由には、投稿に感情が豊かに表現されているという特徴が挙げられる。観光情報専門サイトにおける口コミ情報などは、前提として見ず知らずの人に向け

て発信される情報であることから、かしまった文面が多くなってしまうが、SNS ではそういった堅苦しさがなく、くだけた文面になりやすいことから、素直な感想があらわれる。観光において、訪れた旅行者がどういった感情を抱いていたかを知ることは、観光地に対する想像を形成するために重要であると考えられている。しかし、これまでの研究では、観光地についての感情的な発言を得る手法は提案されているが、観光地に対して大多数の人が抱く一般的な印象を掴みやすくすることを目指していない。すなわち、得られた情報自体には他のユーザからの評価が考慮されていないため、それが単なる個人の意見なのか、それとも、多数の人に共通する意見なのかを判別することは難しい。結果として、ユーザは観光地に対する一般的な印象を評価することができない。

そこで本研究では、観光地に訪れた人がどのような感情を抱くかにおいて、情報に対する他のユーザからの評価を考慮して情報を提示することで、観光地に対する理解を支援することを目的とする。情報の提示においては、SNS 上の観光情報における感情的な発言に着目して、観光地に対する一般的な印象を掴みやすくするために、客観的な評価

¹ 立命館大学院 情報理工学研究所

² 立命館大学 情報理工学部

として、発言に対しての共感を考慮した機能を実現する。情報源としては Twitter を活用し、Twitter に搭載されている「いいね」の数を客観的な指標として用いることで、より多くの人の共感が得られている発言を際立たせて提示する。また、本機能を用いて観光地の評価を行う実験より、観光地を評価するうえでの理解に役立っているかどうかを検証する。

2. 観光における感情的な発言

2.1 関連研究

観光地において、訪れた旅行者の感情を把握することは重要である。杉本らは、観光情報専門サイトにおける口コミに含まれている感情的な発言から有用な情報を導けると考えて、感情を表現する単語を取得して口コミを感情ごとに分類するためのシステムを開発した [2]。一方で、武富らは、SNS の 1 つである Twitter は、ツイート（投稿）できる文字数に制限があるため、簡潔な文章が多くなり、ユーザは気楽にツイートできるので、感情が素直に表れると主張している [3]。感情的な発言を抽出するという点において、かしくまった文面の多い口コミサイトと比べると、くだけた文面の多い SNS の方がよりユーザに親しみやすい情報を抽出できる可能性があると考えられる。

SNS から感情的な発言の抽出を行っている研究の例として、山本らは、テキスト上で伝えることの難しい、ジェスチャーや表情などの情報を表現する顔文字や、「やったあああ」のような語尾の母音を 3 回以上繰り返す表現など、Twitter 特有の表現にも感情が付加されるとして、感情的な発言の抽出手法を提案している [4]。また、斎藤らは、Twitter 上にある観光地付近の位置情報付きツイートのうち、感情的な発言を取得して、感情ごとにマーカーをつけて地図上に表示することで、旅行者にとって馴染みやすい観光アプリを開発した [5]。アプリによって得られる情報は位置情報付きツイートに限られるが、これらによって提供される情報には、実際に訪れた人の感想が現れている。

これらの研究から、SNS 上の投稿から感情によって分類された情報は、観光に役立つ可能性が高いことがわかる。

2.2 解決すべき課題

これまでの研究では、観光地について SNS 上に発信されている情報のうち、感情的な発言を分類して、得られた情報を旅行者に提供していた。しかし、得られた情報が旅行者にとって役立つのかどうかは旅行者個人の感覚に依存しており、得られた情報によって、観光地に対しての一般的な印象を掴みやすくなるかどうかは考慮されていない。例えば、「冬の貴船神社は寒い」といった個人の感覚に依存する情報ではなく、「冬の貴船神社はライトアップがあり、雪景色と相まって綺麗だ」など、観光地に対してそこへ行くと何があるのか、どんな感情を抱くのかといった一般

的な印象は、観光地を訪れる際の有用な情報である。現状では、個人の意見なのか多数の人に共通する意見なのかの判別が難しく、分類した情報の中で、どのような発言が際立っているのかがわかりづらい。結果として、いろんな人の様々な発言が混在しており、ユーザはその観光地がどのような場所であるかが一目でわかりにくくなっている。そのため、感情的な発言を分類して旅行者に提供するだけでなく、得られた情報をランキングし、観光地に対する一般的な印象を評価しやすい形式で旅行者に提供する必要がある。加えて、情報に対する客観的な支持を考慮していないので、情報の信用度も低いと考えられる。そのため、ランキングにおいては、客観的な指標を活用する必要がある。

以上より本研究では、観光地について SNS 上に発信されている情報から、感情的な発言を抽出して分類した後、他のユーザからの評価を考慮した客観的な指標を用いてランキングし、旅行者に提供することで、観光地に対する理解を支援することを目的とする。

3. 共感を利用した提示手法

3.1 感情的な発言への共感

第 2 章で述べたように、観光地に関する理解を支援するためには、感情的な発言に対する客観的な評価を考慮してランキングすることが重要である。

本研究において、客観的に良い評価が得られている情報とは、他者の同意や賛同といった、共感が得られている情報であると考えられる。実際に観光地へ訪れた人の感想につく共感の声が多くあれば、その旅行者の感想は支持されていると考えられる。共感の多い感想は、その共感の数だけ発言内容の認識が広まっていると考えられ、観光地への訪問時に一般に期待される感想としてみなすことが可能である。また、共感の数だけ旅行者の感想は支持されているので、発言内容に対する信用度も上がっていると考えられる。

以上より、本研究では、客観的な評価を考慮するために、共感を表す指標を利用した提示手法を検討する。

3.2 感情による分類

感情的な発言を取得するために、本研究では感情を表現する単語を集めた辞書を構築して観光地についての発言を分類する。以後、本論文では、感情を表現する単語を感情語と呼ぶ。分類に用いる辞書は、感情表現辞典 [6] を使用する。分類する感情は、ツイートを感情によって分類している研究 [5] を参考に 4 種とする。この研究では、感情を「喜ぶ・好ましい」、「悲しみ・恐れ」、「怒り・嫌」、「驚き」の 4 つに分類している。本研究でもこれに基づいて、「嬉」、「圧」、「哀」、「驚」の 4 つに分類する。分類される感情語は、以下を基準に設定する。

- 「嬉」

感情表現辞典上の「喜」と「好」に対応する肯定的表

現である。

- 「圧」
感情表現辞典上の「怒」と「厭」に対応するストレス表現である。
- 「哀」
感情表現辞典上の「哀」に対応する虚無感のある表現である。
- 「驚」
感情表現辞典上の「驚」に対応する新しい発見のある表現である。

また、発言の取得率の向上を図るため、過去形や丁寧語、また漢字を使わずに用いられることがある単語などを考慮して辞書に登録する。

上で述べた方針に基づいて、1585語の感情語を本研究で使用する辞書へ登録した。内訳は、「嬉」が673語、「圧」が599語、「哀」が184語、「驚」が129語である。登録した単語の例を表1に示す。

表1 感情語の例

感情	登録する単語
嬉	喜ぶ, 嬉しい, ありがたみ, 満たされている, 幸せ, 楽しい, 好き, お気に入り, 面白かった
圧	怒った, 酷い, 気持ちが悪い, 我慢できない, 文句を言った, 嫌い, 憤り, 憎い, 困惑, うんざり
哀	悲しい, 悲痛, 寂しい, 哀愁, 虚無, 泣ける, 不憫, 可哀想, しんみり, 寒い
驚	驚き, びっくり, どきっ, 意外, 衝撃, 動揺, 放心, 思いがけない, 腰を抜かす, 仰天

3.3 共感の利用手法

本研究では、SNS上にある観光地についての感情的な発言を感情語で分類した後、共感となる指標が高い順に発言を並び替えることで、観光地に対する一般的な印象の評価を支援する。情報源にはTwitterを用い、共感となる指標にはTwitterに搭載されている機能である「いいね」の数を用いる。

「いいね」は主に、ツイートに対する興味や関心といった素直な肯定を示し、観光地に対する感想や意見、立場を明らかにするツイートに対してつけられる「いいね」は、同意や共感であると考えられる。したがって、本研究では共感の指標に「いいね」の数を用いて提示する。

3.4 提示手法

感情的な発言を取得するためにツイートを感情語で分類した後、共感の指標である「いいね」の高い順にツイートを並べる。まず、Twitterから観光地名をクエリとして指

定して、ツイートを取得する。次に、得られたツイートのうち、感情語を含んでいるものを抽出する。その後、含んでいる感情語に対応する感情ごとにツイートを分類する。最後に、分類されたツイートを「いいね」の数値が高いものから順に並べ、図1に示すような表形式にまとめる。

本提示手法では、ツイート本文と「いいね」の数を1組にして、それぞれの感情の下に配置する。この表の使用方法として、2通りが考えられる。1つ目は、各感情を比較して、「いいね」の多さによってどの感情が際立っているかを把握する方法である。図1では、「嬉」と「驚」において比較的「いいね」の数が多い。この場合、調査対象の観光地は嬉しさを抱く場所であり、意外性のみられる場所であることがわかる。2つ目は、各感情に属しているツイートの中では、「いいね」の多さによってどんな発言が際立っているかを把握する方法である。図1では、「嬉」という感情においては、「●●最高!!」という発言が際立っていることがわかり、その観光地において「●●」が魅力的な要素であることがわかる。

嬉 😊		圧 😡		哀 😞		驚 😲	
本文	いいね	本文	いいね	本文	いいね	本文	いいね
●●最高!!	54	イラつく	7	●疲れた	13	びっくり	68
	32		2		6		37

図1 ツイート提示のイメージ

4. 検証実験

4.1 実験の方針

第3章で示したような提示手法の有効性を検証するために、「いいね」によるランキングが観光地を把握するのに役立つかどうかを調査することを目的として利用実験を行う。本実験では、20代の情報系の学生10名を対象として、1つの観光地に対して5分から10分程度かかる作業を、1人につき4つの観光地に対して行ってもらう、アンケート調査を行った。

本研究では、観光地に対して一般的な印象を評価できるようにすることを目的として提示手法を定めているため、被験者には指定された観光地に対する評価を行ってもらうことを実験内容とする。想定する状況を北海道旅行として、小樽運河、登別温泉、藻岩山、函館山を評価対象の観光地とする。これらの観光地についての評価を、Excelシート上に提示される感情ごとに色分けされた表形式の情報をもとに、まとめてもらう。

4.2 検証方法

本実験で用いる表は、観光地に対する一般的な印象を評価しやすくなることを狙っている。そのため本実験では、本研究が客観的な指標としている「いいね」によって、取

表 2 アンケート

質問 1	いいねの数が多い感情に属するツイートは、観光地を評価するのに役立ちましたか。
質問 2	4つの感情において、それらの上位のツイートは、観光地を評価するのに役立ちましたか。
質問 3	4つの感情において、それらの下位のツイートが、観光地を評価するのに役立ちましたか。
質問 4	3つ目の質問において、「役立った」「どちらかといえば役立った」と答えた方にお聞きします。参考になったツイートには、どういったものが在りましたか。
質問 5	その他に、気づいた事、発見などあればお願いします。

得た感情的なツイートが観光地に対する一般的な印象を評価しやすくなっているかどうかを検証する必要がある。

本実験で用いる表を検証する観点として、3.4節で述べた使用方法を考慮する必要がある。まず、各感情を比較する方法では、際立っている感情に属しているツイートが実際に観光地を評価するうえで役立ったかどうかを確かめる必要がある。また、各感情の中で目立っている発言を知る方法では、「いいね」の多くついているツイートが実際に観光地を評価するうえで役立ったかどうかを確かめる必要がある。一方で、「いいね」の少ないツイートの中でも役立ったツイートがある場合、そのツイートが情報としてどのような価値を持っていたかを把握する必要もある。

上記を考慮したアンケート項目を表2に示す。アンケート項目の最後に気づきを問う自由記述を設けている。この質問は、質問1から質問4の内容にとらわれずに回答してもらうためのものであり、事前に考慮していた観点以外の知見を知りたいことを意図している。

4.3 実験手順

本節では、実験の流れについて説明する。まず、被験者に4つの観光地について、それぞれの表を配る。その後、表から得られた情報を元に、それぞれの観光地について、その観光地には何があって、どんな場所なのかといった一般的な印象の評価をテキストファイルに記録してもらう。作業が終わり次第、調査にかかった時間をテキストファイルに追記して調査結果を提出してもらう。4つの観光地すべてに対する作業が終了した後、最後に表2のアンケートに回答してもらう。

なお、前節で説明した評価の観点を考慮して、被験者には上位のツイートだけでなく下位のツイートも見るように指示をしている。また、被験者には本実験についての目的を説明したうえで実験に取り組んでもらった。

5. 結果と考察

5.1 用意したデータセット

本節では、実験に用いるために取得した感情語を含むツイートのデータセットについて、取得結果とその考察について述べる。

5.1.1 取得結果

第3章で述べた手法を用いて感情的な発言を取得した結

果、ツイートの総取得数のうち、感情的な発言数の割合は、小樽運河が18%、藻岩山が15%、登別温泉が4%、函館山が9%であった。詳細を表3に示す。また、それぞれの観光地のツイートの中で、各感情の最上位のツイートのいいね数を表4に示す。

表 3 取得したデータ数

	嬉	圧	哀	驚	合計	総取得数
小樽運河	99	7	15	4	125	692
藻岩山	44	5	4	2	55	356
登別温泉	40	6	5	1	52	1065
函館山	12	2	1	3	18	197

表 4 最上位のいいね数

	嬉	圧	哀	驚
小樽運河	163	13	4	4
藻岩山	101	9	42	1
登別温泉	71	1	3	0
函館山	89	0	3	7

5.1.2 考察

まず、ツイート数について述べる。表3より、各ツイートの総取得数に対しての感情的なツイート数の割合が低いと感じられる。原因としては絵文字や顔文字のような、感情語辞書に登録している単語にはない感情表現があったことによるものと考えられる。例えば、風景写真付きのツイートにキラキラした絵文字が添えられていると、この絵文字は「風景がきれいであった」ことを示す感情表現となりうる。観光地についてではなく顔文字についても同様であり、汗をかいているように見える顔文字が添えられていると、この顔文字は「発言者が気まずい状態にあった」ことを示す感情表現となりうる。

本研究では、感情語辞典[6]に登録されている単語を用いて感情的な発言の取得を行ったが、感情表現が言葉以外にもあり、絵文字や顔文字などを用いて感情を表すツイートがみられた。また、感情ごとのツイートのデータ数に偏りがあった。4つの観光地とも「嬉」のツイート数が一番多い。これより、観光地名を含むツイートには、肯定的な発言が多くなると考えられる。あるいは、取得時期によっては、例えば冬場に行くことがオススメの観光地について、夏場に調べると否定的な発言や意外な発言が得られるかもしれないが、本研究では取得時期を考慮していない。

さらに、「いいね」の数については、表4に示されているように、どの観光地においても「嬉」の最上位のツイートにつけられる「いいね」が多くなった。これより、肯定的な発言内容には「いいね」をつけやすいと考えられる。一方で、「嬉」以外の感情では、観光地ごとに「いいね」の数に差が出ているため、感情を4つに分けたことには効果があったと考えられる。

5.2 実験結果

本節では、実験後に収集した観光地の評価結果とアンケート結果について述べる。

5.2.1 観光地の評価結果とアンケート結果

観光地の評価にかかった時間を表5に、評価結果の一例を表6に、アンケートの質問1から質問3の結果を表7に、質問5については表8に示す。質問4については、質問3の回答結果を受けたものであったため、該当者がおらず、未回答であった。

表5 評価にかかった時間

	小樽運河	藻岩山	函館山	登別温泉
A	3分38秒	2分45秒	2分17秒	3分46秒
B	2分21秒	7分28秒	1分21秒	2分04秒
C	6分19秒	7分18秒	2分12秒	7分02秒
D	8分00秒	6分00秒	6分00秒	6分00秒
E	5分19秒	4分05秒	4分16秒	3分32秒
F	5分00秒	5分00秒	5分00秒	5分00秒
G	5分00秒	5分00秒	5分00秒	5分00秒
H	6分45秒	4分29秒	5分09秒	3分22秒
I	5分32秒	6分02秒	5分53秒	5分03秒
J	6分00秒	5分00秒	5分00秒	6分00秒

表6 観光地に対する評価結果の一例

	評価結果
小樽運河	近くにレンガ倉庫がある。 近くにびっくりドンキーがある。 夜景やイルミネーションが楽しめる。
藻岩山	展望台から札幌を一望できる。 展望台から夜景が楽しめる。 冬は特に夜景が綺麗である。
登別温泉	露天風呂がある。 冬景色を楽しめる。 ご飯が美味しい。
函館山	働いている方の温かさが伝わってくる。 100万ドルの夜景が楽しめる。 日本三大夜景の一つである。

アンケートについて、各感情を比較する「いいね」に関する質問1において、すべての被験者が「役立った」あるいは、「どちらかといえば役立った」と答えた。また、各感情における上位のツイートに関する質問2において、すべての被験者が「役立った」、あるいは、「どちらかといえば

表7 質問1から質問3についての回答

	質問1	質問2	質問3
役立った	4	2	0
どちらかといえば役立った	6	8	0
どちらかといえば役立たなかった	0	0	8
役立たなかった	0	0	2

役立った」と答えた。さらに、質問2とは逆に、下位のツイートが役立ったかどうかに関する質問3においては、すべての被験者が「役立たなかった」、「どちらかといえば役立たなかった」と答えた。この質問において肯定的な回答が得られなかったため、続く質問4は未回答である。

5.2.2 考察

まず、評価にかかった時間について考察する。すべての被験者が予定していた10分以内で作業を終了しているので、本提示手法による表を用いて観光地の評価を行うことは容易であったと考えられる。

続いて、観光地に対する評価結果について考察する。1つの観光地についての項目数や文量を指定して評価をまとめてもらっていないので、人によって観光地についての説明量が異なるが、その観光地には何があるかについての把握はできていた。ただし、「嬉」に属するツイートから得られた情報が多く、「嬉」以外のツイートを参照している様子はあまりみられなかった。これは、前節で述べたように、用意したデータセットにおいて、「嬉」の「いいね」の数が、どの観光地でも多かったことに起因していると考えられる。

次に、アンケートの質問1から質問3について述べる。質問1や質問2については肯定的な回答が得られたので、本提示手法の効果があったと考えられる。質問3では否定的な回答が得られ、意図していた通り、下位のツイートが役立つことはなかった。

最後に、質問5について述べる。1つ目の回答を要約すると、最上位のツイートでも「いいね」の数が少ない場合、信用できないという意見である。本研究では、客観的な評価の得られていない発言は信用度が低いとしていたため、想定していた結果と同様の回答が得られた。

2つ目の回答より、観光地と無関係の「自分語り」をするツイートが抽出されていたことがわかった。これは、辞書内に登録していた単語が、観光地とは無関係の話題で使用されていたためであった。例えば、孤独感を表す「ぼっち」という表現により抽出されているツイートがあったのだが、このツイートは観光地についてではなく、発言者自身について発言していた。本研究で構築した感情語辞書は、感情語辞典[6]に登録されている単語を詳細に選定せず、第3章で述べたような基準に沿って登録したので、観光地に対するの感情的な発言のみを取得できるように感情語辞書を調整する必要がある。

3つ目の回答より、同じ内容のツイートを繰り返し行っ

表 8 質問 5 についての回答

回答 1	4 つの感情の中でも種類別に見た時の最上位ツイートとはいっても fav の数が少ない場合もあり, その部分は観光地を評価する情報としてあまり信用するべきではないのかもしれないと思いながら取り組みました.4 種に分別されていたが,fav の数がある程度多いのかそうでないかで自分は観光地を評価しました.
回答 2	不満など意見している人は観光地と直接関係ないことなどを述べていてあまり参考になりませんでした.
回答 3	全く同じ内容のツイートが複数確認されたため、情報に無駄が見られた。

ている bot のようなツイートがランキング上位に来ていたことがわかった. こういったツイートには「いいね」がつきにくく, ランキング下位になると考えていたので, 実装において, 対処をしていなかった. 「いいね」の数によってランキングする前に, 同じ内容のツイートがあるかどうかを判別してデータベースへ格納する必要があると考えられる.

6. おわりに

本研究では, 観光地における感情的な発言の提示手法においての問題点を述べ, 客観的な指標として共感を用いる手法を提案した. 情報源としては, Twitter を用い, 「いいね」の数を共感とみなして, 抽出した感情的な発言を「いいね」の数でランキングして提示する手法を実現した. また, 検証実験により, 提示手法は観光地を評価するのに役立つことがわかった. しかし, 構築した感情語辞書には調整が必要であるうえ, bot などによる重複する内容のツイートの除去を行う必要がある.

今後は, 感情語辞書の調整や重複する内容の除去を行ったうえで, Excel などの既存の表計算ソフトに依存しない独自のソフトウェアを開発し, さらなる実用化を図りたい.

参考文献

- [1] 高校生、大学生、社会人 20 代・30 代の SNS 利用に関する意識調査 (2018.3), 入手先 (https://mmdlabo.jp/investigation/detail_1703.html)
- [2] 杉本祐介, 他 : 口コミを利用したレコメンドに適した感情語の分類方法の検討, 研究報告モバイルコンピューティングとユビキタス通信 (MBL), No. 50(2015).
- [3] 武富厚美, 久野雅樹 : Twitter におけるユーザーの特徴と感情表現, 研究報告自然言語処理 (NL), No. 1(2014).
- [4] 山本湧輝, 熊本忠彦 : Twitter 特有表現を考慮したツイートの多次元感情抽出手法の提案, 2014 年度情報処理学会関西支部 支部大会 講演論文集, Vol. 2014,(2014).
- [5] 斎藤一, 横川祥司 : 感情語辞書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 = Memoirs of Hokkaido Information University, pp. 103–110(2016).
- [6] 中村明 : 感情表現辞典, 東京堂出版 (2016).