

旅先での観光地選び支援のための ブログを用いた観光地の印象抽出手法

伊達 賢志^{†1} 北須賀 輝明^{†1}
糸川 剛^{†2} 有次 正義^{†1}

本論文では、ブログから観光地の印象を抽出し、その印象により観光地を提案することを目的とする。ユーザが過去に訪れたことのある観光地や気に入っている観光地と印象の似ている観光地をユーザに推薦する。印象を表す語として、名詞と形容詞、動詞を用いる。まず、観光地に関するブログを集める。そのブログを形態素解析し、それぞれの語を抽出し、TF-IDF法を基にしたTDF-IDF法を用いてそれぞれの語のTDF-IDF値を求める。その値を用いて、ピアソンの積率相関係数により、観光地同士の類似度を算出する。類似度の高い観光地同士を印象の似ている観光地とする。実験では、どの品詞を用いると理想の類似度を算出するのかを確認するために、それぞれの品詞を単独で用いる場合と組み合わせた場合の類似度を比較した。同じカテゴリの観光地の類似度が上位3つに入っているものを成功としてカウントし、成功数の数で評価した。その結果、名詞だけを使った場合の成功数が18中6で最も良くなった。また、名詞、形容詞と動詞それぞれ良い点があることが分かった。

Extracting Impression of Sightseeing Spots from Blogs for Supporting Selection of Spots to Visit in Travel

SATOSHI DATE,^{†1} TERUAKI KITASUKA,^{†1}
TSUYOSHI ITOKAWA^{†2} and MASAYOSHI ARITSUGI^{†1}

In this paper, we try to recommend spots to visit in travel using impression of sightseeing spots from blogs. We recommend spots having same impression as those the user visited and/or likes. We suppose impression can be expressed as nouns, adjectives or verbs in blogs. We first collect blogs concerning sightseeing spots. We then morphological analyze them and extract nouns, adjectives and verbs and assign to them TF-IDF scores calculated by a TDF-IDF based method. Lastly, we find candidate spots to visit that are supposed to be similar to favorite spots of the user by using Pearson product-moment correlation coefficients. For evaluating which part or a combination of parts of speech can

work well for finding good spots, we compare seven combinations of nouns, adjectives and verbs. The results show that nouns themselves could work well in six out of 18 cases in the experiments. Also, they show there are some cases where each combination could work well.

1. はじめに

近年、旅行に行く際、旅行ガイドブックを用いて観光地を探すという従来の方法に加え、インターネットを活用して情報を収集するようになった。ガイドブックは短時間で観光地の情報を得ることができる。ウェブ上には、新たな観光地や穴場などの情報は掲載されている場合がある。ウェブ上の情報のなかでもブログは、情報源として重宝されている。多くの人の書いたブログを閲覧することによって観光地の情報の信頼度や情報量が増す。しかし、膨大な量のブログを読むことは、非常に時間がかかり特に旅先では現実的とは言い難い。

本研究では、このような手間をさけ、観光地の印象を表す印象語として名詞と形容詞、動詞を用いて、観光地の印象を抽出する手法を提案する。自然言語処理の分野では、名詞に注目した研究は、盛んに行われている。本論文で、名詞以外にも形容詞と動詞にも注目した理由として、ブログには、ユーザの行動や感想が書かれており、それらの語が形容詞や動詞として出現するので観光地推薦において有益な情報源になると考えたからである。提案手法を応用することで、これまでにユーザが行った観光地の印象と、旅先の観光地の印象を比較し、ユーザの好みに近い観光地を提案することが可能になると考える。

ユーザの好みの観光地を提案するためには、ユーザの好きな観光地の印象を知り、その印象に適する観光地を推測する必要がある。本研究では、ユーザの好む印象をユーザの好きな観光地から抽出し、その印象と類似の印象がある観光地をユーザに提案する。まず、ユーザが過去に訪れた観光地名や好きな観光地名をクエリとしてブログを収集する。そのブログを形態素解析し印象語を抽出する。この印象語が、観光地の印象である。次に、この印象語に対して、TF-IDF (Term Frequency-Inverse Document Frequency)法を基にした計算によって、観光地ごとに固有の印象語と数値のセットを作成する。事前に全国を網羅した観光地のリストから、同様の手順でユーザの訪れる地域の観光地のセットを作成する。最後

^{†1} 熊本大学大学院自然科学研究科
Graduate School of Science and Technology, Kumamoto University

^{†2} 熊本県立技術短期大学校
Kumamoto Prefectural College of Technology

に、このセット同士の類似度をピアソンの積率相関係数を用いて計り、類似度の高い観光地をユーザに提案する。実験では、ブログ中のどの印象語による観光地提案が有用かを確認するために、名詞、形容詞と動詞を印象語として利用する。また、三つの品詞のうち二つを組み合わせて使用した場合と全てを使用した場合の実験も行う。

本論文の構成は次の通りである。2章で、観光地選びと印象語に関する関連研究を紹介する。3章で、提案手法について説明する。4章で、今回行った実験について説明する。5章で、まとめと今後の課題を述べる。

2. 関連研究

2.1 観光地選びに関する研究

Yuら¹⁾は、GPSデータを使って観光地と旅行プランの推論と提案を行っている。GPSデータとは、GPSポイントが複数集まったものである。GPSポイントとは、GPS機能を使って獲得したときの地点を表し、獲得した時刻と経度、緯度で構成されている。GPSデータは、GPS機能のついている携帯電話などのデバイスに保存されることが多い。Yuらの手法は、GPSデータからユーザの訪れた場所と観光地を抽出し、その観光地に対してスコアを付ける。このスコアはHITS推論モデルに基づいている。HITS推論モデルとはHyperlink-Induced Topic Searchの略で、被参照度(オーソリティスコア)と評価の高いWebページへの参照度(ハブスコア)から、重要性の高いWebページを抽出するアルゴリズムを用いた推論モデルである。本来はWebページを評価するとき用いられるアルゴリズムだが、このアルゴリズムを用いて観光地とユーザの関係を求めて観光地の評価に利用している。このアルゴリズムは、良いWebページには良いリンクページがあり、良いリンクページは良いWebページをリンクしているという考えに基づいている。これと同様に良い観光地は良いユーザが訪れており、良いユーザは良い観光地を訪れていると考えられる。良い観光地とは面白い観光地のことであり、良いユーザとは面白い観光地を知っているユーザである。Yuらは、オーソリティスコアの高い観光地をユーザに提案している。Yuらは、同じ地域にいる全てのユーザに対して同じ観光地を提案している。本論文では、異なる好みのユーザに対しては、異なる観光地を提案している。

丸山ら²⁾は、観光のためのパーソナルナビゲーションシステムを提案している。提案システムは、ユーザが出発地と出発時刻、帰着地と帰着時刻、複数の観光候補地と各地への立ち寄り希望度と時間制約(到着時間帯や滞在時間など)を設定すると、制限時間内で巡回可能かつ最も満足度が高くなるような巡回経路(幾つかの観光地を含む)と各観光地への到着・

出発予定時刻を含むスケジュールを算出しユーザに提示する機能を提供する。このシステムは決定したスケジュールに従い、GPS機能を備えた携帯端末を介し、ユーザにナビゲーション機能を提供する。提案するナビゲーション機能では、現在地を中心とする地図と次の目的地への経路の表示などの空間的な誘導に加え、各目的地での、滞在可能時間の表示や出発時刻の通知などの、時間的な誘導機能を提供する。遺伝的アルゴリズムを用いて準最適なスケジュールを高速に算出するアルゴリズムを設計・開発し、Javaサーブレットとして実装している。PCや携帯端末からウェブインタフェースを介してスケジュールの作成、ナビゲーション機能が利用できる。市販のカーナビゲーションシステム用のデジタル地図を用いた評価実験により、準最適なスケジュールを実用的時間で案内できることなどを確認している。丸山らの提案したシステムの中では、ユーザ自身が訪れたい観光地を入力する仕様となっている。ユーザ満足度は、ユーザの希望通りの観光地が含まれており、指定された時間制約を満たしている場合に増加していくと考えている。丸山らは、ユーザに観光地を指定させているが、本論文では、ユーザが訪れる観光地自体をユーザに提案する。

2.2 印象語に関する研究

Pavelら³⁾は、flickrに投稿された位置情報の付加されている写真のタグを用いて位置情報の付加されていない写真の撮影場所を推論した。flickrとは、個人で撮影した写真をウェブ上で整理、分類、公開できるサービスである。flickrでは、誰かがアップロードした写真に誰でもタグと呼ばれるキーワードを付加して分類することが出来る。flickrにアップロードされた写真に位置情報が付加されている場合は、位置情報の通りの場所に配置される。その写真に対して全てのユーザがタグを付加出来る。提案手法は、まず、位置情報とタグとの関連を機械学習する。その後、位置情報の付加されていない写真のタグから写真の撮影された位置を推測する。Pavelらは、タグと位置情報との関連性を用いて、位置情報のない写真の撮影場所を推論している。本論文では、観光地とブログで用いられる印象語との関連性を用いて印象の似た観光地を提案する。

熊本ら⁴⁾は、漠然とした要求しかなく、ユーザの持つ何らかの判断基準に合致するコンテンツを見つけたいという場合には、印象に基づく検索が有効であると述べている。印象に基づく検索システムへの入力には、通常、複数個の印象表現の中から一つあるいは複数個を選択する、あるいは複数個の印象語に対する多段階評価、複数個の印象語対に対する多段階評価といった形式で行われる。そのため、印象語の選び方によっては、ユーザが見つけたいコンテンツの印象を適切に入力できない可能性がある。そこで、印象語同士の関連性を定義する必要がある。例えば、ユーザが「綺麗な」コンテンツを検索した場合に、「美しい」コンテ

ソツを検索したときと同じ結果になるようにすることが有効である。先行研究では、設計者の知識や経験によって印象語同士の関連性が定義されている。これは、設計者の主観に基づいているという点が課題である。そのようなとき、熊本らは、国語辞書による印象語同士の関連性の定義を提案している。本論文では、印象語同士の関連性を考慮せず、本論文では、全ての印象語が独立しているものとし、熊本らの提案は今後の課題とする。

柴田ら⁵⁾は、街路景観のイメージと物理的特性の対応関係を分析・モデル化し、この対応関係より、自動的に景観画像のイメージを推定し、未知の景観画像を評価する画像検索システムへの応用を目的とする。今までの研究では、色や形などの物理的特性を経験的に決定しており、物理的特性とイメージとの相関関係の分析は余り行われていない。そのため、景観画像のイメージの推定や、未知の景観画像の検索が困難になっている。そこで、人間感情においては、粗い情報処理が行われているという仮説のもとに、物理特性とイメージ語との対応関係を分析する。その結果、明るさを含む色、方向、位置の三つの物理特性とイメージ語とに対応関係があること、そして、イメージ語によって着眼する物理特性が違うことがわかった。柴田らは、イメージ語を「嗜好」や「解放感」といったいくつかのグループに分類する。本論文では、印象語をグループ化しない。印象語をグループ化するに当たって、製作者の主観が含まれ、正しく観光地提案できないと考えたからである。

3. 提案手法

3.1 提案手法の概要

図1に、提案手法の概要を示す。まず、ユーザの好きな観光地をクエリとしてブログを収集する。収集した複数のブログの本文に対して形態素解析を行う。形態素のうち印象語となるものに対して後述の TDF-IDF 値を算出する。次に、事前に全国を網羅した観光地のリストから、同様の手順をユーザの訪れる地域の観光地に対して行う。最後に、これらの観光地の間で、それぞれの印象語と TDF-IDF 値のセットを比較する。比較には、ピアソンの積率相関係数を用いる。ユーザの好きな観光地と類似度の高い観光地をユーザに提案する。

3.2 観光地の印象の抽出

本節では、ユーザの好きな観光地から印象を抽出する手法を説明する。まず、ユーザの過去の旅行経験で訪れた好きな観光地名をクエリとして、ブログを収集する。ユーザの好きな観光地には、ユーザの嗜好を表した観光地への印象が含まれていると考えられる。その観光地への印象となる印象語が、収集したブログに記述されていると考える。観光地について記述されたブログは多く、最新の情報が入手しやすいので有用であると考え。収集した

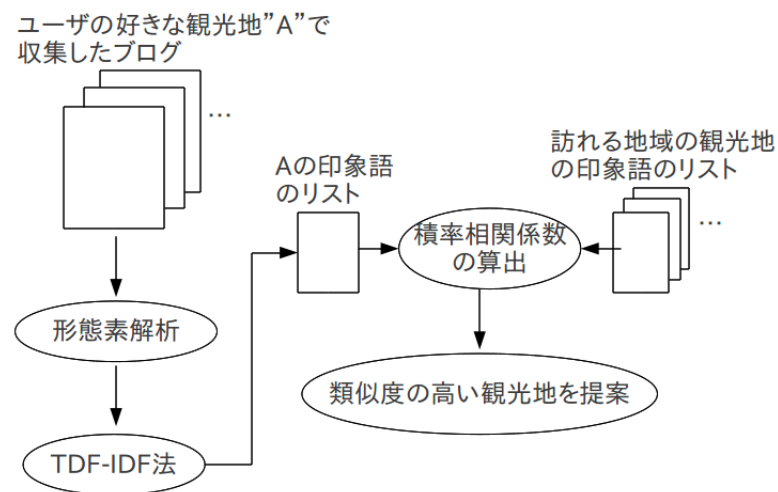


図1 提案手法の概要
Fig. 1 Process flow of proposal

ブログの本文を形態素解析する。一般的なブログは、投稿された日付、ブログのタイトル、URL、本文という情報で構成されている。本提案では印象語として、名詞、形容詞、動詞を用いる。まず、名詞を選択した理由として、ユーザの嗜好が「温泉」や「本」などの具体的な語に現れる可能性を考慮したからである。次に、形容詞を選択した理由として、ユーザの嗜好が「美しい」や「凄い」などの主観的な語に現れる可能性を考慮したからである。最後に、動詞を選択した理由として、ユーザの嗜好が「遊ぶ」や「見る」などの能動的な語に現れる可能性を考慮したからである。ブログの形態素のなかで、印象語にあてはまるものを抽出する。

TF-IDF 法を基にした計算により観光地に関するブログの特長語に数値を付加する。この方法を TDF-IDF (Term's Document Frequency-Inverse Document Frequency) 法と呼ぶ。印象語 i の TDF-IDF 値の計算方法を以下に示す。式 (1) は、印象語 i の TF 値を算出する式である。TF 値は、対象となるドキュメントの中で、特定の単語がどの程度の頻度で出現しているかを表す。つまり、TF 値が大きいほどその単語がドキュメントを特徴づけている単語と判断できるため、TF とはドキュメント中でその単語が持っている重要度といえる。式 (2) は、本論文で用いる印象語 i の TDF 値を算出する式である。式 (2) によって、

ある観光地について収集したブログ中で印象語 i の使われるブログの出現頻度が得られる。右辺の分母は、収集したブログの総数である。ブログの総数とは、ユーザの好きな観光地名で収集したブログとユーザの旅行先の地域にある観光地名で収集したブログの総数である。分子は、印象語 i を含むある観光地のブログ集内でその印象語が使われているブログ数である。TF-IDF 法なら、分子は印象語が一つのドキュメント内で使われている回数だが、一つのブログで 2 度目以降に出現した印象語は、カウントしないことにした。これによって、出現頻度は高いが「ある」や「ない」のような印象語とはなり得ない語の TDF 値を下げられると考えた。

$$\text{印象語 } i \text{ の TF 値} = \frac{n_i}{\sum_k n_k} \quad (1)$$

$$\text{印象語 } i \text{ の TDF 値} = \frac{\text{ある観光地のうち印象語 } i \text{ の出現したブログ数}}{\text{ある観光地のブログの数}} \quad (2)$$

式 (3) は、印象語 i の IDF 値を算出する式である。IDF 値とは、多くのドキュメント中に出現する単語よりも、特定のドキュメント中にしか出現しない単語ほどドキュメントをよく特徴付けていると考えるものである。この値によって、印象語 i の珍しさが得られる。

$$\text{印象語 } i \text{ の IDF 値} = \log \frac{\text{総ブログ数}}{\text{印象語 } i \text{ の出現する総ブログ数}} \quad (3)$$

$$\text{印象語 } i \text{ の TDF-IDF 値} = \text{印象語 } i \text{ の TDF 値} \cdot \text{印象語 } i \text{ の IDF 値} \quad (4)$$

式 (4) は、印象語 i の TDF-IDF 値を算出する式である。TDF-IDF 値とは、TDF 値と IDF 値を掛けたもので特定のブログ集では出現頻度が高いものが、高い数値を示す。これらによって、ある観光地に関するブログから印象語の TDF-IDF 値を算出する。

3.3 印象の似ている観光地の提案

TDF-IDF 値を用いて他の観光地との類似度を計算し、類似度の高いものをユーザに提案する。類似度はピアソンの積率相関係数を用いて表す。ピアソンの積率相関係数 r は式 (5) で表される。ピアソンの積率相関係数は、1 から -1 で表される。1 に近いほど強い正の相関関係があり、-1 に近いほど負の相関関係がある。0 に近いほど相関は弱い。

$$r(X, Y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2} \sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2}} \quad (5)$$

変数 X と変数 Y は、観光地のブログを表しており、2 つの観光地に同じ印象語 i があった個数を n とする。変数 X は、ユーザの好きな観光地であり、変数 Y は、ユーザが旅行する地域にある観光地である。 X_i は、観光地 X のブログ中で出現した印象語 i の TDF-IDF

値である。 Y_i は、観光地 Y のブログ中で出現した印象語 i の TDF-IDF 値である。 \bar{X} は、観光地 X の TDF-IDF 値の相加重平均である。 \bar{Y} は、観光地 Y の TDF-IDF 値の相加重平均である。

4. 実験

4.1 評価対象

本実験では、提案手法の印象語として名詞と形容詞、動詞をそれぞれ用いた場合とそれらを組み合わせて用いた場合の類似度の比較を行う。ユーザの選んだ好きな観光地を表 1 の 18ヶ所とする。ただし、表中の記号は付録での記述に用いたものである。また、これらの観光地は山、温泉、ドーム球場、美術館、城、アミューズメントパークという 6 種類のカテゴリに属するものを 3ヶ所ずつ選んだ。まず、ユーザの好きな観光地それぞれに対して、観光地名をクエリとして 300 件のブログをそれぞれ収集する。収集期間は、2011 年 3 月 24 日から 4 月 30 日までである。収集した 300 件のブログの本文を抽出し、ブログの本文に対して形態素解析を行う。形態素解析を行った後に観光地の印象語の TDF-IDF 値を算出する。印象語は、名詞、形容詞、動詞、名詞と形容詞、名詞と動詞、形容詞と動詞、名詞と形容詞と動詞の 7 パターンを用いる。18ヶ所の観光地のうち 1ヶ所をユーザの好きな観光地とみなし、その他の 17ヶ所の観光地をユーザに提案する観光地候補とみなして、これらの観光地の間で類似度を求める。

表 1 実験で用いる観光地名

Table 1 Sightseeing spots with their categories

観光地のカテゴリ	観光地名
山 (A)	阿蘇山 (I), 富士山 (II), 八ヶ岳 (III)
温泉 (B)	別府 (IV), 指宿 (V), 黒川 (VI)
ドーム球場 (C)	福岡ドーム (VII), 東京ドーム (VIII), 札幌ドーム (IX)
美術館 (D)	熊本現代美術館 (X), 国際国立美術館 (XI), 国際西洋美術館 (XII)
城 (E)	熊本城 (XIII), 名古屋城 (XIV), 大阪城 (XV)
アミューズメントパーク (F)	ディズニーランド (XVI), USJ (XVII), スペースワールド (XVIII)

4.2 評価方法

本実験では、ユーザは好きな観光地と同じカテゴリに属する観光地を好むという想定で評価する。ユーザの好きな観光地と同じカテゴリに属する 2ヶ所の観光地が、ピアソンの積率相関係数の大きさの上位 3 位以内に入っているものを期待通りに観光地が推薦できた観光

地とし、推薦に成功した観光地と呼ぶ。18ヶ所の観光地を6種類のカテゴリに分類する。均等に観光地を選択しているため、それぞれのカテゴリに3ヶ所の観光地が属している。提案手法の有用性を確認するために、ユーザの訪れる地域を指定せず、全てのユーザに対して、提案したい観光地の数が同数になるように観光地を選択した。

4.3 実験結果

実験結果では、印象語として名詞、形容詞、動詞をそれぞれ単独で使用した場合と組み合わせて使用した場合の7パターンの評価結果を比較する。付録の表4、表5、表6、表7、表8、表9、表10は、各印象語を使用した場合の類似度を示す。表2は、それぞれのパターンで推薦に成功した観光地名である。表2に示す通り、名詞だけを印象語として使用した場合は18ヶ所中6ヶ所の観光地で成功し、最も良い結果となった。次に良かったパターンは、名詞と形容詞のパターンと名詞と動詞のパターンで5ヶ所の観光地で成功した。次に、形容詞と動詞のパターンと動詞だけのパターンと全てを使用したパターンで4ヶ所の観光地で成功した。最後に、形容詞だけを使用したパターンで2ヶ所の観光地で成功した。

表2 使用した印象語と推薦に成功した観光地名
Table 2 Results of Impression words and sightseeing spots

印象語	観光地
名詞	黒川, 国際国立美術館, 熊本現代美術館, 国際西洋美術館, 熊本城, 名古屋城
形容詞	別府, 福岡ドーム
動詞	東京ドーム, 札幌ドーム, 熊本現代美術館, 国際西洋美術館
名詞と形容詞	黒川, 熊本現代美術館, 国際国立美術館, 国際西洋美術館, 名古屋城
名詞と動詞	黒川, 熊本現代美術館, 国際国立美術館, 国際西洋美術館, 名古屋城
形容詞と動詞	福岡ドーム, 東京ドーム, 札幌ドーム, 国際西洋美術館
名詞と形容詞と動詞	黒川, 国際国立美術館, 国際西洋美術館, 名古屋城

まず、それぞれの品詞を単独で使用した場合について比較する。表3は、それぞれの観光地名で収集したブログ中で使用された名詞と形容詞、動詞の使用された種類の数である。名詞を単独で使用した場合が、全体でも最も良い結果となった。この理由として、ブログ中で使用される品詞のうち名詞の出現頻度が、形容詞と動詞に比べ高いことが考えられる。出現頻度が多いので多くのデータが集まり、的確にその観光地の印象を表す印象を抽出できる。美術館のカテゴリに属する3ヶ所全ての観光地において成功した。また、城のカテゴリでも3ヶ所中2ヶ所の観光地において成功した。この理由は、これら二つのカテゴリではカテゴリ特有の名詞の TDF-IDF 値が他のカテゴリより高くなったからである。例えば、「美術」,

表3 各観光地ごとの品詞の種類数
Table 3 The number of parts of speech's kinds

観光地	名詞	形容詞	動詞
阿蘇	2192	86	419
富士山	1699	65	328
八ヶ岳	2027	80	390
別府	1744	62	276
指宿	2312	102	492
黒川	2376	85	395
福岡ドーム	2165	78	393
東京ドーム	3594	109	611
札幌ドーム	1960	87	419
熊本現代美術館	2105	87	425
国際国立美術館	2399	89	345
国際西洋美術館	3053	148	601
熊本城	1691	71	292
名古屋城	2015	76	397
大阪城	2706	112	626
ディズニーランド	2357	81	446
USJ	2199	83	409
スペースワールド	2685	69	358

「版画」、「絵画」、「歴史」、「本丸」などのカテゴリ特有の名詞の TDF-IDF 値が高くなった。これらのカテゴリ特有の印象語の TDF-IDF 値が高かったため、同じカテゴリに属する観光地が複数成功した。

形容詞を単独で使用した場合では成功した「別府」と「福岡ドーム」については、名詞を単独で用いた場合では、提案できていない。更に、形容詞を単独で使用した場合に成功している「別府」は、他の6パターンでは1パターンも成功していない。この理由として、「別府」では、主に温泉に関する感想についての形容詞が頻出している。ブログ中での形容詞の出現頻度は少ない。そのため、形容詞の抽出単語数は、名詞や動詞と比べ、非常に少ない。少ない印象語のうち「白い」や「かっこよい」などの観光地の印象を強く表しているものの比率が高かったため、「別府」と「福岡ドーム」は成功したと考えられる。形容詞を単独で使用した場合は、最も成功数が少なかった。この理由として、抽出される形容詞数が、名詞や動詞と比べ少ないからである。少ないデータの中で正しくカテゴリ特有の印象語を抽出できたものが少ないため、このような結果になったと考えられる。

動詞を単独で使用した場合は、4ヶ所の観光地が成功した。ドーム球場と美術館のカテゴリ

りでそれぞれ2ヶ所成功した。この二つのカテゴリに共通することとして、使用される動詞に大きな偏りがあるということがあげられる。ドーム球場の場合は、「打つ」「投げる」「勝つ」など野球に関連した動詞の TDF-IDF 値が非常に高くなった。美術館の場合は、「描く」、「見る」「眺める」などの動詞の TDF-IDF 値が非常に高くなった。逆に、温泉のカテゴリでは「浸かる」や「浴びる」といった入浴に関する動作の TDF-IDF 値が高くなく「書く」や「見る」といった動詞の TDF-IDF 値が高かった。正解しなかったカテゴリでは、直感的には使用頻度が高いと思われる動詞の TDF-IDF 値が低いこと正解しなかった。この理由として、温泉などでは、人それぞれの楽しみ方があり、出現頻度の差が出なかったことが考えられる。

名詞と動詞のパターンと名詞と形容詞のパターンでは、同じ結果を示し、どちらも5ヶ所で成功した。結果が同じになった理由として、名詞の出現頻度の高さが考えられる。名詞の出現頻度が高いため、TDF-IDF 値の上位の大半を名詞が占めていた。そのため、名詞を同時に使用したパターンでは、名詞を単独で使用した場合の実験結果と類似の結果になった。

同様に、形容詞と動詞のパターンでは、動詞の方が形容詞より出現頻度が高いため、動詞を単独で使用した場合の結果とほぼ同じものになった。

名詞と動詞と形容詞を使用した場合は、名詞と動詞を使用した場合の結果と近い結果となった。唯一異なる観光地は「熊本現代美術館」である。異なった理由として、形容詞での類似度が低かったことが考えられる。

4.4 考 察

使用する品詞によって、TDF-IDF 値が高くなる印象語には、特徴があることがわかった。まず、名詞を使用した場合は、地名の TDF-IDF 値が高くなる。また、観光地が属するカテゴリ名が高くなるものが多い。前者の場合、異なる地域の同じカテゴリの観光地より同じ地域にある異なるカテゴリの観光地との類似度が高くなる場合がある。後者の場合、同じカテゴリに属する観光地との類似度が高くなるというメリットがあるが、もし同じカテゴリの観光地がない場合は、正しくユーザに提案できないであろう。形容詞や動詞の場合は、ユーザの好きなカテゴリの観光地が無くても同じ印象語が使われている場所を提案できる。形容詞の場合「ない」や「いい」といった印象語としては不適切だと考えられる語がどの観光地でも抽出される。しかし、どの観光地でも高い値となっているので類似度計算への影響は少ない。動詞の場合「できる」や「やる」といった印象語としては不適切だと考えられる語がどの観光地でも抽出される。しかし、形容詞の場合と同様に、どの観光地でも高い値となっているので類似度計算への影響は少ない。形容詞と動詞でこれらの印象語の

TDF-IDF 値が高くなった理由として、どのブログでも使われているため、TDF 値が非常に高くなり、IDF によって値を下げていないことがあげられる。

ブログから抽出された印象語の中には、その観光地特有のものがある。それらの印象語の TDF-IDF 値が高くなった観光地があった。例えば「別府」の「蒸す」という動詞は、他の温泉のカテゴリでは出現しなかった。普通の温泉地では「茹でる」という動詞の TDF-IDF 値が高くなっている。しかし、別府温泉では、蒸し料理が有名であり他の温泉地との違いとして有名なためこのような結果になった。また、「阿蘇」では「カルデラ」という名詞の TDF-IDF 値が高くなった。観光地の魅力の一つとして、普段は体験できない経験ができることや、その観光地でしか見られない景色などといった希少性がある。このような観光地特有の印象語をユーザに提示することで、より満足度の高い観光地をユーザに提案できると考えられる。

TDF-IDF 値上位の印象語のなかには、ブログの収集期間に開催されたイベントに関連する語が現れた。例えば、美術館のカテゴリでの期間限定の展示物の名前や、ドーム球場のカテゴリでの試合結果などである。このような印象語が多く TDF-IDF 値の上位に現れた場合は、上位の印象語を見るだけで勝敗や活躍した選手名などの試合内容を予測することができる。しかしこれらの印象語は、時と共に変化するので、観光地の印象語としては不適切であると考えられる。

本実験では、山とアミューズメントパークのカテゴリに属する観光地では、1度も成功しなかった。山のカテゴリに属する観光地として「阿蘇山」「富士山」「八ヶ岳」を使用した。「阿蘇山」や「富士山」は、登山以外のブログが多く、その内容にも共通性もあまりなかったため、どのパターンでも成功しなかったと考えられる。また、アミューズメントパークでも、それぞれの観光地で差別化が進んでいるため、「山」のカテゴリと同様に共通性があまりないため、成功しなかったと考えられる。

5. おわりに

本論文では、ブログから観光地の印象を抽出し、その印象による観光地を提案することを目的とした。印象を表す印象語として、名詞と形容詞、動詞を用いた。まず、観光地に関するブログを集める。そのブログを形態素解析し、それぞれの印象語を抽出し、TF-IDF 法を基にした計算によってそれぞれの印象語の TDF-IDF 値を求めた。その値を用いて、ピアソンの積率相関係数により、観光地同士の類似度を算出し、類似度の高い観光地同士を印象の似ている観光地とした。ユーザが過去に訪れたことのある観光地や気に入っている観光地と

類似度の高い観光地をユーザに推薦する．実験では，どの印象語を用いると理想の類似度を算出するのかを確認するために，それぞれの印象語を単独で用いる場合と組み合わせた場合の類似度を比較した．同じカテゴリの観光地の類似度が上位 3 つに入っている場合を推薦に成功したとみなし，成功数の数で評価した．その結果，名詞だけを使った場合の成功数が 18 中 6 で最も良くなった．また，形容詞，動詞それぞれを単独で使用した場合は，名詞だけを使用した場合に成功した観光地以外のものが成功した．このように，名詞と形容詞と動詞でそれぞれ良い点があることが分かった．印象語として名詞を用いた場合，ほかの品詞より文章中での使用頻度が高く，カテゴリの印象をよく表したものが多いため，良い結果が得られた．形容詞の場合，文章中での使用頻度が低く，カテゴリをよく表す形容詞が少ないため，良い結果が得られないことがわかった．動詞の場合，カテゴリである程度使用される動詞が決まる場合に良い結果が得られることがわかった．

今後の課題として，それぞれの品詞のメリットを残し，デメリットを減らす必要がある．提案手法では，名詞の TF-IDF 値が他の品詞より高くなったので，他の品詞と組み合わせて使用する場合に，重み付けし名詞の TDF-IDF 値を抑える方法を検討する．本実験ではユーザの好きな観光地と同じカテゴリの観光地を提案することを成功としたが，異なるカテゴリの観光地を提案する場合の評価方法を考える必要がある．また，印象語同士の関連性を考慮する必要がある．

参 考 文 献

- 1) Yu Zheng, Lizhu Zhang, Xing Xie, Wei-Ying Ma, “Mining Interesting Locations and Travel Sequences from GPS Trajectories”, Proceedings of the 18th international conference on World wide web, pp. 791–800, 2009
- 2) 丸山敦史, 柴田直樹, 村田佳洋, 安本慶一, 伊藤実, “P-Tour: 観光スケジュール作成支援とスケジュールに沿った経路案内を行うパーソナルナビゲーションシステム”, 情報処理学会論文誌, Vol.45, No.12, pp.2678–2687, 2004
- 3) Pavel Serdyukov, Vabessa Murdock, Roelof van Zwol, “Placing Flickr on a map”, Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval, pp.484–491, 2009
- 4) 熊本忠彦, 太田公子, “印象に基づく検索のための印象選定法の提案”, 情報処理学会論文誌, Vol.44, No.7, pp.1808–1811, 2003
- 5) 柴田滝也, 加藤俊一, “街路の景観画像データベースのイメージ語による検索”, 電子情報通信学会論文誌, Vol.J82-D-I, No.1, pp.147–183, 1999

付 録

付録では，印象語として名詞，形容詞，動詞，名詞と形容詞，名詞と動詞，形容詞と動詞，名詞と形容詞と動詞を用いた場合の観光地間のピアソン積率類似度を示す．なお，数値は，1000 倍したものである．表中の記号は，表 1 で用いたものである．

表 4 名詞を用いた場合の類似度 (10^{-3})

Table 4 Degree of similarities using nouns (10^{-3})

カテゴリ	観光地	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII
A	I	1000.0	8.5	145.1	66.5	88.0	258.7	139.9	75.8	83.6	142.6	40.3	47.2	42.9	51.5	78.5	47.9	89.4	191.2
	II	8.5	1000.0	180.5	181.1	181.7	88.3	197.4	111.0	183.8	126.5	23.6	82.9	184.4	111.6	209.4	94.3	163.2	48.3
	III	145.1	180.5	1000.0	168.8	146.6	84.0	176.9	91.6	162.5	77.6	105.6	96.9	165.5	187.6	236.0	219.6	197.8	64.5
B	IV	66.5	181.1	168.8	1000.0	128.3	229.3	179.1	143.4	102.4	185.8	75.9	98.2	92.8	109.1	105.7	138.6	178.3	73.1
	V	88.0	181.7	146.6	128.3	1000.0	208.9	250.7	182.2	279.5	143.8	95.9	104.9	106.0	224.2	193.8	135.4	230.5	93.0
	VI	258.7	88.3	84.0	229.3	208.9	1000.0	99.8	103.9	100.3	96.5	28.0	93.5	78.5	89.6	81.2	82.6	150.6	78.3
C	VII	139.9	197.4	176.9	179.1	250.7	99.8	1000.0	186.4	371.3	213.5	64.4	70.5	34.1	172.5	216.8	65.0	299.7	42.2
	VIII	75.8	111.0	91.6	143.4	182.2	103.9	186.4	1000.0	119.9	148.5	116.5	150.0	53.3	2.6	126.8	218.7	171.7	137.6
	IX	83.6	183.8	162.5	102.4	279.5	100.3	371.3	119.9	1000.0	137.0	131.5	98.1	137.1	141.5	149.3	21.4	213.2	79.7
D	X	142.6	126.5	77.6	185.8	143.8	96.5	213.5	148.5	137.0	1000.0	347.4	243.8	196.0	74.3	206.8	104.5	201.4	144.6
	XI	40.3	23.6	105.6	75.9	95.9	28.0	64.4	116.5	131.5	347.4	1000.0	525.8	104.0	71.8	234.3	116.9	213.0	155.2
	XII	47.2	82.9	96.9	98.2	104.9	93.5	70.5	150.0	98.1	243.8	525.8	1000.0	93.9	-13.7	79.0	122.9	200.9	176.4
E	XIII	42.9	184.4	165.5	92.8	106.0	78.5	34.1	53.3	137.1	196.0	104.0	93.9	1000.0	274.6	197.8	5.0	256.0	-16.7
	XIV	51.5	111.6	187.6	109.1	224.2	89.6	172.5	2.6	141.5	74.3	71.8	-13.7	274.6	1000.0	197.8	36.2	172.3	-9.0
	XV	78.5	209.4	236.0	105.7	193.8	81.2	216.8	126.8	149.3	206.8	234.3	79.0	197.8	197.8	1000.0	137.4	245.0	107.6
F	XVI	47.9	94.3	219.6	138.6	135.4	82.6	65.0	218.7	21.4	104.5	116.9	122.9	5.0	36.2	137.4	1000.0	249.1	126.1
	XVII	89.4	163.2	197.8	178.3	230.5	150.6	299.7	171.7	213.2	201.4	213.0	200.9	256.0	172.3	245.0	249.1	1000.0	147.6
	XVIII	191.2	48.3	64.5	73.1	93.0	78.3	42.2	137.6	79.7	144.6	155.2	176.4	-16.7	-9.0	107.6	126.1	147.6	1000.0

表 5 形容詞を用いた場合の類似度 (10^{-3})
 Table 5 Degree of similarities using adjectives (10^{-3})

カテゴリ	観光地	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII
A	I	1000.0	-91.9	-204.5	303.0	17.3	130.8	173.3	168.4	-74.6	288.8	-61.6	147.2	165.1	2.1	-58.6	160.0	-107.6	206.6
	II	-91.9	1000.0	47.5	-80.5	295.5	322.6	223.0	150.0	359.9	258.3	167.6	241.0	45.4	492.8	125.7	435.1	192.5	-37.0
	III	-204.5	47.5	1000.0	51.9	102.4	364.6	426.5	307.2	434.0	63.5	5.7	23.7	-150.7	196.6	312.2	-138.9	-24.0	49.4
B	IV	303.0	-80.5	51.9	1000.0	318.2	483.4	143.6	212.3	25.4	415.2	182.5	-43.2	174.9	-36.7	122.7	73.1	-51.3	-122.3
	V	17.3	295.5	102.4	318.2	1000.0	307.2	434.6	176.1	116.3	247.7	378.2	242.2	230.0	254.4	477.3	417.8	102.0	-86.9
	VI	130.8	322.6	364.6	483.4	307.2	1000.0	404.1	197.4	161.2	146.8	210.0	72.1	172.6	212.8	139.2	239.9	265.5	-229.1
C	VII	173.3	223.0	426.5	143.6	434.6	404.1	1000.0	512.3	533.3	425.7	420.6	292.7	-88.7	169.3	401.6	166.9	368.1	338.2
	VIII	168.4	150.0	307.2	212.3	176.1	197.4	512.3	1000.0	246.3	406.5	323.9	469.7	104.8	228.9	269.8	396.0	311.9	150.5
	IX	-74.6	359.9	434.0	25.4	116.3	161.2	533.3	246.3	1000.0	336.4	326.6	254.7	-35.1	279.0	494.9	317.9	333.2	165.4
D	X	288.8	258.3	63.5	415.2	247.7	146.8	425.7	406.5	336.4	1000.0	257.6	431.5	378.8	412.1	537.1	220.6	393.5	417.3
	XI	-61.6	167.6	5.7	182.5	378.2	210.0	420.6	323.9	326.6	257.6	1000.0	307.4	228.9	247.1	441.0	302.4	292.2	237.2
	XII	147.2	241.0	23.7	-43.2	242.2	72.1	292.7	469.7	254.7	431.5	307.4	1000.0	204.7	435.7	356.0	190.1	206.1	182.2
E	XIII	165.1	45.4	-150.7	174.9	230.0	172.6	-88.7	104.8	-35.1	378.8	228.9	204.7	1000.0	21.5	202.8	-9.5	60.2	-183.2
	XIV	2.1	492.8	196.6	-36.7	254.4	212.8	169.3	228.9	279.0	412.1	247.1	435.7	21.5	1000.0	222.8	275.2	226.1	6.1
	XV	-58.6	125.7	312.2	122.7	477.3	139.2	401.6	269.8	494.9	537.1	441.0	356.0	202.8	222.8	1000.0	505.2	320.5	115.7
F	XVI	160.0	435.1	-138.9	73.1	417.8	239.9	166.9	396.0	317.9	220.6	302.4	190.1	-9.5	275.2	505.2	1000.0	394.1	81.4
	XVII	-107.6	192.5	-24.0	-51.3	102.0	265.5	368.1	311.9	333.2	393.5	292.2	206.1	60.2	226.1	320.5	394.1	1000.0	192.4
	XVIII	206.6	-37.0	49.4	-122.3	-86.9	-229.1	338.2	150.5	165.4	417.3	237.2	182.2	-183.2	6.1	115.7	81.4	192.4	1000.0

表 6 動詞を用いた場合の類似度 (10^{-3})
 Table 6 Degree of similarities using verbs (10^{-3})

カテゴリ	観光地	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII
A	I	1000.0	306.0	284.6	177.9	333.4	439.9	342.2	182.5	250.9	377.5	350.0	333.3	282.7	220.4	270.4	348.5	284.4	303.8
	II	306.0	1000.0	384.9	197.8	225.5	214.1	307.6	305.7	260.2	141.4	388.6	254.9	271.2	105.7	283.7	428.8	263.6	407.2
	III	284.6	384.9	1000.0	208.5	275.3	435.7	255.9	254.7	329.8	303.4	229.0	180.1	315.8	217.8	312.9	295.1	259.1	371.1
B	IV	177.9	197.8	208.5	1000.0	181.8	284.6	388.8	196.1	311.1	456.1	317.4	307.8	373.0	87.2	249.0	340.5	341.7	198.1
	V	333.4	225.5	275.3	181.8	1000.0	399.1	209.4	237.7	247.6	248.4	159.7	221.2	233.7	183.1	354.9	377.3	279.3	251.8
	VI	439.9	214.1	435.7	284.6	399.1	1000.0	221.9	187.2	139.3	371.7	200.6	244.5	255.3	315.2	235.0	283.2	274.8	229.6
C	VII	342.2	307.6	255.9	388.8	209.4	221.9	1000.0	400.5	442.3	373.4	390.2	406.6	331.1	261.1	325.6	328.7	521.1	197.5
	VIII	182.5	305.7	254.7	196.1	237.7	187.2	400.5	1000.0	464.0	284.5	292.9	387.7	93.2	184.6	340.6	435.0	322.8	368.0
	IX	250.9	260.2	329.8	311.1	247.6	139.3	442.3	464.0	1000.0	343.2	359.0	317.5	338.3	360.5	330.9	343.6	312.9	461.7
D	X	377.5	141.4	303.4	456.1	248.4	371.7	373.4	284.5	343.2	1000.0	393.5	460.0	341.2	188.2	368.4	259.0	384.6	270.7
	XI	350.0	388.6	229.0	317.4	159.7	200.6	390.2	292.9	359.0	393.5	1000.0	595.0	286.4	213.1	396.1	359.5	423.3	341.0
	XII	333.3	254.9	180.1	307.8	221.2	244.5	406.6	387.7	317.5	460.0	595.0	1000.0	171.1	176.2	330.3	468.7	429.0	257.0
E	XIII	282.7	271.2	315.8	373.0	233.7	255.3	331.1	93.2	338.3	341.2	286.4	171.1	1000.0	259.3	302.2	217.9	474.5	205.1
	XIV	220.4	105.7	217.8	87.2	183.1	315.2	261.1	184.6	360.5	188.2	213.1	176.2	259.3	1000.0	265.4	297.2	291.1	222.5
	XV	270.4	283.7	312.9	249.0	354.9	235.0	325.6	340.6	330.9	368.4	396.1	330.3	302.2	265.4	1000.0	309.8	434.3	327.0
F	XVI	348.5	428.8	295.1	340.5	377.3	283.2	328.7	435.0	343.6	259.0	359.5	468.7	217.9	297.2	309.8	1000.0	492.9	366.9
	XVII	284.4	263.6	259.1	341.7	279.3	274.8	521.1	322.8	312.9	384.6	423.3	429.0	474.5	291.1	434.3	492.9	1000.0	397.9
	XVIII	303.8	407.2	371.1	198.1	251.8	229.6	197.5	368.0	461.7	270.7	341.0	257.0	205.1	222.5	327.0	366.9	397.9	1000.0

表 7 名詞と形容詞を組み合わせて用いた場合の類似度 (10^{-3})
 Table 7 Degree of similarities using nouns and adjectives (10^{-3})

カテゴリ	観光地	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII
A	I	1000.0	4.8	132.3	75.3	84.8	257.0	140.5	80.2	76.3	146.2	37.5	51.6	45.8	49.8	70.5	51.3	75.6	191.1
	II	4.8	1000.0	175.4	161.7	191.6	94.5	203.3	110.7	198.4	133.7	30.6	94.5	174.9	124.0	204.6	106.7	169.2	44.0
	III	132.3	175.4	1000.0	164.1	144.3	90.6	187.2	99.9	174.2	76.0	101.0	93.0	148.3	187.8	238.0	201.2	177.4	63.3
B	IV	75.3	161.7	164.1	1000.0	131.7	234.3	175.6	146.5	96.8	189.6	79.2	90.0	92.8	101.6	103.2	135.7	153.7	66.7
	V	84.8	191.6	144.3	131.7	1000.0	210.3	258.5	181.2	270.7	147.7	108.1	112.4	113.3	225.3	207.3	143.9	222.3	87.8
	VI	257.0	94.5	90.6	234.3	210.3	1000.0	104.1	106.6	101.1	96.8	31.7	91.6	79.0	92.3	81.9	85.7	151.4	68.3
C	VII	140.5	203.3	187.2	175.6	258.5	104.1	1000.0	198.7	376.5	224.6	80.3	84.4	27.1	171.3	226.4	69.8	306.9	53.0
	VIII	80.2	110.7	99.9	146.5	181.2	106.6	198.7	1000.0	123.9	159.9	124.3	162.1	53.8	9.3	132.5	222.2	178.5	138.0
	IX	76.3	198.4	174.2	96.8	270.7	101.1	376.5	123.9	1000.0	151.2	141.8	107.2	129.3	149.2	168.9	28.6	224.8	83.2
D	X	146.2	133.7	76.0	189.6	147.7	96.8	224.6	159.9	151.2	1000.0	345.4	249.4	202.2	86.5	221.4	110.4	218.5	154.2
	XI	37.5	30.6	101.0	79.2	108.1	31.7	80.3	124.3	141.8	345.4	1000.0	519.8	110.7	77.7	244.2	124.1	219.2	158.3
	XII	51.6	94.5	93.0	90.0	112.4	91.6	84.4	162.1	107.2	249.4	519.8	1000.0	98.2	1.1	93.0	124.2	202.9	177.9
E	XIII	45.8	174.9	148.3	92.8	113.3	79.0	27.1	53.8	129.3	202.2	110.7	98.2	1000.0	266.1	198.3	3.1	241.2	-21.8
	XIV	49.8	124.0	187.8	101.6	225.3	92.3	171.3	9.3	149.2	86.5	77.7	1.1	266.1	1000.0	199.0	43.8	174.7	-6.2
	XV	70.5	204.6	238.0	103.2	207.3	81.9	226.4	132.5	168.9	221.4	244.2	93.0	198.3	199.0	1000.0	148.1	250.6	108.9
F	XVI	51.3	106.7	201.2	135.7	143.9	85.7	69.8	222.2	28.6	110.4	124.1	124.2	3.1	43.8	148.1	1000.0	250.1	125.1
	XVII	75.6	169.2	177.4	153.7	222.3	151.4	306.9	178.5	224.8	218.5	219.2	202.9	241.2	174.7	250.6	250.1	1000.0	150.5
	XVIII	191.1	44.0	63.3	66.7	87.8	68.3	53.0	138.0	83.2	154.2	158.3	177.9	-21.8	-6.2	108.9	125.1	150.5	1000.0

表 8 名詞と動詞を組み合わせて用いた場合の類似度 (10^{-3})
 Table 8 Degree of similarities using nouns and verbs (10^{-3})

カテゴリ	観光地	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII
A	I	1000.0	43.1	166.9	81.9	123.3	279.4	173.4	89.0	106.9	175.3	85.7	90.5	71.2	73.7	112.8	95.0	125.2	209.0
	II	43.1	1000.0	202.9	183.4	189.2	115.4	229.3	151.6	206.7	132.3	82.3	123.1	209.7	112.2	228.5	170.2	191.9	117.2
	III	166.9	202.9	1000.0	174.8	166.4	141.4	195.0	120.5	202.3	113.9	129.8	112.9	195.7	191.2	251.1	237.9	216.8	110.2
B	IV	81.9	183.4	174.8	1000.0	134.3	236.9	214.5	151.1	138.3	227.9	127.7	143.4	142.7	104.6	131.0	176.3	212.9	93.2
	V	123.3	189.2	166.4	134.3	1000.0	244.7	243.2	191.6	273.8	160.7	107.1	127.5	126.5	216.5	225.9	181.6	243.3	115.8
	VI	279.4	115.4	141.4	236.9	244.7	1000.0	125.0	121.5	110.2	134.2	60.8	120.3	104.2	118.8	116.9	121.8	180.9	101.1
C	VII	173.4	229.3	195.0	214.5	243.2	125.0	1000.0	219.1	380.8	245.1	126.3	142.6	84.8	192.5	242.1	109.1	360.7	67.0
	VIII	89.0	151.6	120.5	151.1	191.6	121.5	219.1	1000.0	164.2	173.9	147.5	193.3	62.5	30.7	170.3	241.7	205.1	164.7
	IX	106.9	206.7	202.3	138.3	273.8	110.2	380.8	164.2	1000.0	191.9	183.2	142.2	181.5	178.8	194.0	71.8	237.7	129.9
D	X	175.3	132.3	113.9	227.9	160.7	134.2	245.1	173.9	191.9	1000.0	352.4	271.5	213.2	87.4	240.5	139.7	248.9	160.1
	XI	85.7	82.3	129.8	127.7	107.1	60.8	126.3	147.5	183.2	352.4	1000.0	534.4	145.6	94.5	270.2	162.9	265.9	178.1
	XII	90.5	123.1	112.9	143.4	127.5	120.3	142.6	193.3	142.2	271.5	534.4	1000.0	108.6	15.0	131.8	185.8	259.9	186.1
E	XIII	71.2	209.7	195.7	142.7	126.5	104.2	84.8	62.5	181.5	213.2	145.6	108.6	1000.0	273.2	212.8	38.4	317.0	13.8
	XIV	73.7	112.2	191.2	104.6	216.5	118.8	192.5	30.7	178.8	87.4	94.5	15.0	273.2	1000.0	207.6	73.7	193.3	18.8
	XV	112.8	228.5	251.1	131.0	225.9	116.9	242.1	170.3	194.0	240.5	270.2	131.8	212.8	207.6	1000.0	170.9	288.8	143.9
F	XVI	95.0	170.2	237.9	176.3	181.6	121.8	109.1	241.7	71.8	139.7	162.9	185.8	38.4	73.7	170.9	1000.0	282.5	153.0
	XVII	125.2	191.9	216.8	212.9	243.3	180.9	360.7	205.1	237.7	248.9	265.9	259.9	317.0	193.3	288.8	282.5	1000.0	182.0
	XVIII	209.0	117.2	110.2	93.2	115.8	101.1	67.0	164.7	129.9	160.1	178.1	186.1	13.8	18.8	143.9	153.0	182.0	1000.0

表 9 形容詞と動詞を組み合わせて用いた場合の類似度 (10^{-3})
 Table 9 Degree of similarities using adjectives and verbs (10^{-3})

カテゴリ	観光地	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII
A	I	1000.0	213.9	197.7	203.0	274.8	396.1	310.8	174.2	175.5	361.7	283.4	291.6	260.8	179.5	207.2	319.6	202.6	283.6
	II	213.9	1000.0	314.5	144.5	244.2	222.4	295.8	269.2	278.5	160.2	341.5	255.6	219.5	151.4	253.1	418.0	245.3	321.5
	III	197.7	314.5	1000.0	185.7	244.8	436.2	282.1	275.1	340.7	259.8	193.6	150.9	223.2	210.7	312.7	218.3	192.1	313.2
B	IV	203.0	144.5	185.7	1000.0	196.8	300.3	348.4	202.2	259.0	443.3	293.5	250.2	326.6	67.1	233.9	309.0	251.8	149.5
	V	274.8	244.2	244.8	196.8	1000.0	383.3	245.5	224.8	224.6	243.4	199.6	224.4	235.4	195.3	374.1	378.0	239.2	202.5
	VI	396.1	222.4	436.2	300.3	383.3	1000.0	232.3	193.6	137.0	343.4	191.9	215.3	236.2	291.4	227.3	283.0	269.3	152.5
C	VII	310.8	295.8	282.1	348.4	245.5	232.3	1000.0	419.5	457.4	382.6	395.4	387.5	237.9	243.9	336.5	302.2	483.4	218.5
	VIII	174.2	269.2	275.1	202.2	224.8	193.6	419.5	1000.0	420.4	303.6	297.2	398.0	87.5	189.1	330.7	429.9	313.0	315.0
	IX	175.5	278.5	340.7	259.0	224.6	137.0	457.4	420.4	1000.0	342.1	353.8	305.3	260.0	344.8	354.7	319.6	317.8	407.9
D	X	361.7	160.2	259.8	443.3	243.4	343.4	382.6	303.6	342.1	1000.0	373.0	453.2	346.8	227.4	393.0	253.1	383.0	292.4
	XI	283.4	341.5	193.6	293.5	199.6	191.9	395.4	297.2	353.8	373.0	1000.0	546.3	263.4	218.2	403.1	353.3	396.3	322.1
	XII	291.6	255.6	150.9	250.2	224.4	215.3	387.5	398.0	305.3	453.2	546.3	1000.0	175.4	218.9	333.4	420.6	378.2	249.8
E	XIII	260.8	219.5	223.2	326.6	235.4	236.2	237.9	87.5	260.0	346.8	263.4	175.4	1000.0	208.1	274.1	177.8	362.1	125.2
	XIV	179.5	151.4	210.7	67.1	195.3	291.4	243.9	189.1	344.8	227.4	218.2	218.9	208.1	1000.0	254.3	294.4	277.3	175.5
	XV	207.2	253.1	312.7	233.9	374.1	227.3	336.5	330.7	354.7	393.0	403.1	333.4	274.1	254.3	1000.0	332.3	404.3	290.7
F	XVI	319.6	418.0	218.3	309.0	378.0	283.0	302.2	429.9	319.6	253.1	353.3	420.6	177.8	294.4	332.3	1000.0	468.4	319.4
	XVII	202.6	245.3	192.1	251.8	239.2	269.3	483.4	313.0	317.8	383.0	396.3	378.2	362.1	277.3	404.3	468.4	1000.0	354.8
	XVIII	283.6	321.5	313.2	149.5	202.5	152.5	218.5	315.0	407.9	292.4	322.1	249.8	125.2	175.5	290.7	319.4	354.8	1000.0

表 10 名詞と形容詞と動詞を組み合わせて用いた場合の類似度 (10^{-3})
 Table 10 Degree of similarities using nouns and adjectives and verbs (10^{-3})

カテゴリ	観光地	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	XVIII
A	I	1000.0	38.7	155.3	89.0	119.4	277.3	172.8	92.1	99.7	177.6	81.9	92.5	73.2	71.5	105.1	96.7	112.6	208.2
	II	38.7	1000.0	197.8	168.5	196.5	119.5	231.7	149.3	216.0	137.2	85.9	130.0	200.1	121.9	223.6	176.7	193.8	110.1
	III	155.3	197.8	1000.0	170.9	163.9	146.2	202.6	127.0	209.8	111.4	125.3	109.1	180.1	191.2	252.4	222.3	199.8	107.9
B	IV	89.0	168.5	170.9	1000.0	137.3	241.0	210.8	153.5	132.6	230.2	128.5	135.2	140.8	98.8	129.0	173.1	192.3	87.2
	V	119.4	196.5	163.9	137.3	1000.0	244.9	249.8	190.4	266.7	163.2	116.8	132.6	131.7	217.7	235.5	187.1	235.5	111.0
	VI	277.3	119.5	146.2	241.0	244.9	1000.0	128.2	123.4	110.4	133.8	62.8	118.0	104.1	120.5	116.8	123.8	180.7	91.8
C	VII	172.8	231.7	202.6	210.8	249.8	128.2	1000.0	228.5	385.0	252.5	136.7	150.3	76.5	190.6	248.5	111.7	361.3	75.3
	VIII	92.1	149.3	127.0	153.5	190.4	123.4	228.5	1000.0	166.3	182.0	153.0	201.8	62.4	35.6	173.4	244.4	208.6	164.0
	IX	99.7	216.0	209.8	132.6	266.7	110.4	385.0	166.3	1000.0	199.6	189.1	147.5	172.0	183.5	206.9	76.1	244.8	131.5
D	X	177.6	137.2	111.4	230.2	163.2	133.8	252.5	182.0	199.6	1000.0	350.5	275.5	218.0	97.6	250.9	143.0	258.2	167.9
	XI	81.9	85.9	125.3	128.5	116.8	62.8	136.7	153.0	189.1	350.5	1000.0	528.9	148.1	98.9	276.7	167.3	267.4	180.2
	XII	92.5	130.0	109.1	135.2	132.6	118.0	150.3	201.8	147.5	275.5	528.9	1000.0	111.6	26.9	140.8	184.7	257.0	187.2
E	XIII	73.2	200.1	180.1	140.8	131.7	104.1	76.5	62.4	172.0	218.0	148.1	111.6	1000.0	265.7	212.3	35.6	299.0	8.2
	XIV	71.5	121.9	191.2	98.8	217.7	120.5	190.6	35.6	183.5	97.6	98.9	26.9	265.7	1000.0	208.2	79.4	194.3	20.3
	XV	105.1	223.6	252.4	129.0	235.5	116.8	248.5	173.4	206.9	250.9	276.7	140.8	212.3	208.2	1000.0	178.8	290.1	143.6
F	XVI	96.7	176.7	222.3	173.1	187.1	123.8	111.7	244.4	76.1	143.0	167.3	184.7	35.6	79.4	178.8	1000.0	282.6	151.4
	XVII	112.6	193.8	199.8	192.3	235.5	180.7	361.3	208.6	244.8	258.2	267.4	257.0	299.0	194.3	290.1	282.6	1000.0	183.1
	XVIII	208.2	110.1	107.9	87.2	111.0	91.8	75.3	164.0	131.5	167.9	180.2	187.2	8.2	20.3	143.6	151.4	183.1	1000.0