

ジオソーシャル写真を用いた観光スポットの検知と景観評価

木上 智貴[†] 馬 強[†]

[†] 京都大学大学院情報学研究科社会情報学専攻 〒606-8501 京都市左京区吉田本町

E-mail: [†]kinoue@db.soc.i.kyoto-u.ac.jp, ^{††}qiang@i.kyoto-u.ac.jp

あらまし 近年、観光ビッグデータや ICT を活用し、訪日外国人旅行者の動向や関心事項を把握することは、観光立国の実現に向けた観光資源の調査やプロモーションの実施に結びつくという期待から、注目を集めている。その中で、SNS 上に公開されている位置情報が付与された画像を利用して、景観を自動分析することは、ユーザ視点の観光資源の再発見において重要である。従来手法の多くは、ヒューリスティックに画像の特徴量を選択して景観評価を行っているが、景観地を表現する特徴量には客観的なものだけでなく主観的なものも考慮されるべきである。本研究では、テキストを基に、日本の景観への評価を自動で行うことを目指す。本稿では、観光スポットを検出するクラスタリングとそのパラメータ調整手法について述べた後、テキストデータから二種類の特徴量を抽出し、景観評価に利用する手法を提案する。

キーワード 観光, 景観評価, 特徴抽出, DBSCAN

1 はじめに

観光は日本経済を支える重要な成長分野であり、観光立国に向けた取り組みは国を挙げて行われている。2017 年の訪日外国人人数は 11 月の時点で累計 2616 万 9000 人に達しており、2016 年の年間累計を既に超えた結果となっている。日本政府は観光立国推進基本計画の中で訪日外国人人数 4000 万人など、様々な目標を掲げている。

しかし、現状では、急増する観光客は、東京・大阪・京都といったいわゆるゴールデンルートに集中する傾向にあり、社会インフラなどの負担や住民・観光客へのストレスが大きい一方で、地方への経済効果の恩恵は限定的である。

観光ビッグデータや情報通信技術を活用し、訪日外国人旅行者の行動に基づいた調査を行うことによって、観光資源の掘り起こしや磨き上げ・インフラの整備・プロモーション等を実施し、リピーターを含む訪日外国人旅行者数の増加を図ると同時に、観光の個人化や地域・時期の分散化を狙う試みが為されている。

多岐に渡る観光ニーズに応え、観光の個人化を実現するためには、多様な観光資源の発見とその分析が重要である。従来までの観光資源の価値分析は、現地調査に頼ってきており、コストの掛かるものである。Twitter や Flickr 等の SNS 上に公開されている位置情報やタグがメタデータとして付与された画像を対象に、景観地を直接自動分析することは、ユーザ視点の観光資源の再発見において重要な取り組みである [1]。そこで、本研究では、SNS にアップした写真とそのメタデータを分析して、観光客視点からの観光スポット発見とその評価手法を提案する。

本研究では、まず、SNS にアップした写真の位置情報を利用して、ユーザ視点からのスポットを検出する。観光業界や行政が公開されている、既存のスポットとの差異を明らかにしていく。スポットの検知は、位置情報付けの写真データのクラスタ

リングに基づいて行うが、既知の有名スポットの範囲情報を用いて、クラスタリング手法に必要なパラメータを自動調整する手法を提案している。

観光スポットで撮影された写真の分析においては、写真に写されている景観物の色・配置・種類などを考慮する際に、環境心理学の知見が応用されてきた。Shen [2, 3] は、Tveit らが提案した 9 個の尺度 [4] の内、“coherence”、“visual scale”、“imageability” に焦点を絞り、画像からそれぞれの尺度について計算を行い景観評価を行う手法を提案している。しかしながら、Shen の手法は、多様性のある景観の評価として一般性に欠けている。特に、これらの海外の景観から整理された特徴量は日本の風景の特徴をうまく表現できない場合がある。

Hsu は、日本の景観に関する研究文献から 21 個の形容詞を対象に、日本の景観評価の尺度を分類し、景観の選好に影響を及ぼす因子の構造を明らかにしている [5]。本研究では、SNS などにアップしたスポットで撮影した写真の記述を分析して、記述から景観を評価するための主観と客観的な特徴量の埋め込み表現を学習して、その埋め込み表現を用いてスポットの景観を評価することを試みる。

景観地と呼ばれるスポットやエリアには、その場所を景観地足らしめている様々な要因が存在している。景観地を強く印象付けるシンボルや、景観の色合いやパターン、そして数多く存在するランドマークは、景観地の特徴として大きな影響力を持っている。しかしながら、景観地を特徴付ける要因は、必ずしも視覚的に容易に判断できるもののみとは限らないことがしばしばある。歴史的な建築物は有名な観光スポットであることが多いが、多くはその歴史や雰囲気強い影響力を持つ。

例えば、京都は日本の中でも最も主要な観光地の一つであるが、一方で日本全国には小京都と呼ばれる観光地が存在している。小京都とはその景観や町並みや風情が京都と似ているという特徴があるスポットを指し、魅力的な観光地として推薦されている。ただ単にお寺や紅葉や歴史的建造物が存在するだけで

は、小京都と呼ばれることはない。この点で、小京都と呼ばれるスポットは、お寺や紅葉といった客観的に確認できる要因に加え、「京都らしさ」という主観的な特徴を併せ持った場所であると言える。

客観的な特徴だけでなく主観的な特徴量を景観評価において考慮することで、多角的な景観地の分析や、景観地推薦のパーソナル化が可能となる。

本稿の構成は次の通りである。第2節では、関連研究を紹介する。第3節では、観光資源の発見のために行うクラスタリングにおいて、最適なパラメータを自動調整する手法を説明する。第4節では、品詞に基づき2種の特徴量を抽出し、景観の評価を行う手法を説明する。第5節では、実験について述べる。第6節では、本稿をまとめると共に今後の課題を示す。

2 関連研究

2.1 景観評価

2.1.1 環境心理学における景観評価の尺度・分類に関する研究

環境心理学において景観評価とは、環境や景観に対して与える意味や価値付けを把握することにより、人間と環境との間を結ぶ関係性の解明を課題とするものである。景観の良し悪しは、人間が環境をいかに認識しているかの表れとして考えられており、多くの研究が為されている [6-9]。Kaplan ら [9] は、同質の自然景観の中にも、景観の選好に差を生み出す決定的な要因が含まれているという仮定のもとで、景観の好ましさに関する4つの評価因子を提唱した。また、Hsu は景観選好における評価因子の構造を解明する上で、Kaplan らの提唱する評価因子の有効性を確かめた [5]。

Hsu は、日本の研究成果から抽出した21個の形容詞を対称に、『好む意識に与える要因』という設定を築くにあたり、4段階の因果順序を想定し、選好に影響を及ぼす因子の構造を明らかにした [5]。[5] では、「好きな」という結果に強い影響を与える語彙として、「自然な」・「良い」・「美しい」の3つがあると提唱している。

2.1.2 観光情報学における景観評価

観光資源の開拓や、増加する観光需要へ対応すべく、POIs(points of interests) で撮影された画像を分析することで、観光価値を推定する取り組みが行われている。

Shen は、環境心理学での景観評価尺度の定義を解釈し、それらの尺度が自然景観の画像内でどのように表現されるかを人間の知識をもとに再定義を行った [2,3]。人間の知識による定義に基づいて、画像を構成している色やオブジェクトを対象に画像の特徴を抽出する手法を提案している。Ge は、特定の都市の文化や特性といった側面は、都市で撮影された画像内で頻繁に出現するオブジェクトによって表現されるという考えのもと、教師なしランキング手法を提案している [10]。Zhuang らは、写真の撮影者と閲覧者の行動を分析して、景観を間接評価する手法を提案している [11]。

これらの手法には、画像処理で扱いづらい尺度は、評価が困

難であるという問題点がある。

現に、Shen [2,3] は Tveit ら [4] が提唱する9個の評価尺度の内、3個のみを採用して評価を行っている。

2.2 特徴抽出

画像や文章などのメディアオブジェクトを表現する特徴を、定量的な変数とみなすことで、より低次元のベクトルに圧縮して表現し、オリジナルのメディアオブジェクトを特徴量として端的に表すことができる。データを表す最適な特徴量は、人間の試行錯誤によって探索されていたが、ニューラルネットワークを用いて特徴抽出もプロセスも自動的に行う研究が、様々な種類のメディアを対象に盛んに行われている。

情報処理において、任意の単語は文字列のインデックスに過ぎず、意味の含まれていない one-hot ベクトルとして処理される。Mikolov らが提案した word2vec [12] は、ニューラルネットワークを利用し、大量のテキストデータを学習することで、それぞれの単語を数百次元のベクトルで表す分散表現を獲得する手法である。これにより、'king' - 'man' + 'woman' = 'queen' のように、単語を意味に基いた形式で演算処理することが可能となる。また、近い意味を持つ単語同士が、ベクトル空間上で近い位置に存在している意味ネットワークが作り出される。word2vec は、大量のコーパスデータから単語間の共起関係を得るものであり、本研究では、景観地を表現する単語の分散表現を取得するにあたり、word2vec と同様の学習モデルを利用する。

3 観光スポットの検知

観光の個人化に向けて多様な観光資源を発見すべく、本研究では、SNS にアップした写真データの位置情報を用いて、観光スポットの検知とその評価を行う手法を提案する。実際に観光客が訪れているスポットを検知できるだけでなく、サービスプロバイダーが想定している観光スポットとの差を検討することで、穴場となるスポットや潜在的な観光価値を計ることが可能である。本節では、ユーザーが撮影した位置情報付き写真をクラスタリングすることで行われる観光スポットの検出と、その際の最適なパラメータの調整手法について述べる。図1は、提案の概要を示したものである。

3.1 クラスタリング

観光スポットのクラスタリングにおいては、パラメータとしてクラスタ数を指定する必要がないというメリットを持つ DBSCAN が適していると考えられる。その他にも DBSCAN のメリットには、外れ値に対する頑健性の高さや、クラスタの形状に制約が無いことが挙げられる。DBSCAN クラスタリングは、2種類のパラメータを必要とする。1つ目はデータ点から探索を行う半径を表す ϵ であり、2つ目はクラスタを形成するために必要となるデータ点の最小数を表す $minPts$ である。DBSCAN には、この2種のパラメータの決定が困難であるというデメリットが存在する。それぞれの点は以下の3種類に分類される。

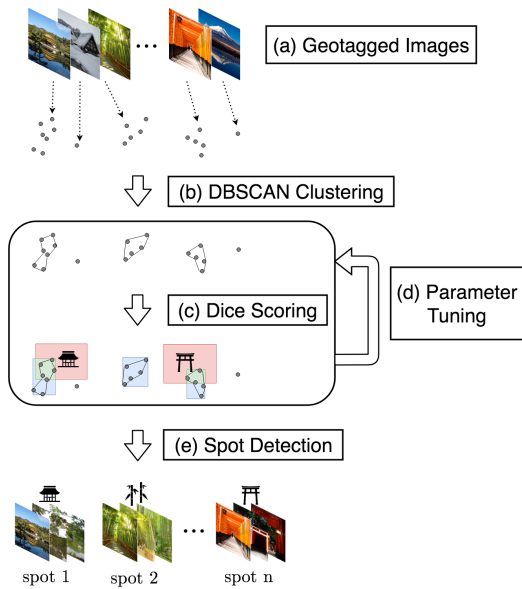


図 1 観光スポット検知の概要：SNS 上に投稿された位置情報付きの画像を対象とする (a). DBSCAN クラスタリングを用いてデータ点を密度に基づきクラスタリングする (b). 生成されたクラスターと、既知の主要スポットとの被覆度に基づきクラスタリングの結果におけるスコアを計算する (c). (c) のクラスタリングと (d) のスコア計算を繰り返し、最適なパラメータの自動調整を行う (d). (d) で得られたパラメータを用いて、最終的なスポットの検知を行う (e).

- コア点

ある点 p から距離 ϵ 以内に、 p 自身を含め $minPts$ 以上の点が存在する場合、点 p はコア点である。

- 境界点

ある点 q から距離 ϵ 以内に、 q 自身を含め $minPts$ 以上の点が存在しないが、他のコア点から到達できる場合、点 q は境界点である。

- 外れ値点

コア点でも境界点でもない点は、外れ値点である。DBSCAN クラスタリングでは、コア点の集合がクラスターを形成し、境界点はいずれかのクラスターに割り当てられる。外れ値点はいずれのクラスターにも割り当てられることなく、外れ値として扱われる。

3.2 パラメータチューニング

DBSCAN クラスタリングでは 2 種類のパラメータに対してそれぞれ最適な値を決定する必要がある。本研究では、各クラスターや各主要スポットを被覆する最小の矩形 MBR (Minimum Bounding Rectangle) を導入し、それらの被覆度により 2 種のパラメータを自動調整する手法を提案する。

主要な観光スポットは当然訪れるユーザーも多く、多数の写真が撮影されるため、これらのスポットは DBSCAN クラスタリングで検知されるべきであると考えられる。そこで、検知されるべきスポットとして、主要なスポットとその領域矩形を予

め定めておく。これらのスポットを検出するクラスターと、予め定めた矩形との類似度が高ければ高いほど、優れたクラスタリングであると仮定し、両者の領域の重複する面積を基準にパラメータチューニングを行う。具体的には以下の手順で算出する。図 2 に提案手法を示す。

(1) m 個主要スポットを対象に、予め矩形領域 $s = \{s_1, s_2, \dots, s_m\}$ を定める。矩形 $s_i (1 \leq i \leq m)$ の面積を a_{s_i} と表す。

(2) DBSCAN クラスタリングによって得られた n 個のクラスター $c = \{c_1, c_2, \dots, c_n\}$ それぞれについて、 c を被覆する最小の矩形 $s = \{t_1, t_2, \dots, t_n\}$ を定義する。なお、矩形の向かい合う対辺は、緯線及び経線と平行になるようにする。矩形 $t_j (1 \leq j \leq n)$ の面積を a_{t_j} と表す。

(3) $s_i (1 \leq i \leq m)$ と $t_j (1 \leq j \leq n)$ の重複領域を $r_{i,j}$ とし、その面積 $a_{r_{i,j}}$ と定義する。

(4) $\forall i, \forall j (1 \leq i \leq m, 1 \leq j \leq n)$ において、Dice 係数 $d(s_i, t_j) = \frac{2a_{r_{i,j}}}{a_{s_i} + a_{t_j}}$ を算出する。

(5) 主要スポット s_i の Dice 係数は、 $D(s_i) = \max\{d(s_i, t_j) | 1 \leq j \leq n\}$ で表す。すなわち、1 つの主要スポットと重複領域を持つ矩形が複数存在する場合は、それぞれの Dice 係数が最も高いものを、その主要スポットの Dice 係数の値として使用する。

(6) $DSC(\epsilon', minPts') = \sum_{1 \leq i \leq m, 1 \leq j \leq n} \frac{D(s_i)}{m}$ を、 $\epsilon = \epsilon'$ と $minPts = minPts'$ のときの精度スコアとし、この値が最も高くなる ϵ' と $minPts'$ を最適なパラメータとして決定する。

4 観光スポットの特徴量

4.1 主観的及び客観的特徴量

従来手法では、スポットの特徴量として考えられていた要素は、スポットを構成する物体や、その色やパターンなど、視覚的に判断が容易であるものを対象にしたものが主であった。確かに印象的な物体やランドマークといった情報は、景観地を特徴として高い影響力を持つが、一方で風情や雰囲気といった感覚的に認識される要素も景観の特徴として非常に強い影響力を持っていることは事実である。

景観地の特徴量を考えるにあたり、客観的な特徴量だけでなく主観的な特徴量を考えることは、景観分析の多様性の向上をもたらす。

景観の特徴量として、本研究では客観的特徴量と主観的特徴量の二種類を仮定し、以下のように定義する。

- 客観的特徴量

客観的特徴量とは、対象物に対して視覚的に確認される要素を表す。本研究での景観における客観的特徴量としては、「犬」・「バス」・「金閣寺」といった対象そのものの名前・種類・固有名詞を指すものの分散表現とする。

- 主観的特徴量

客観的特徴量とは、対象物に対する感じ方や受け取り方を表す。

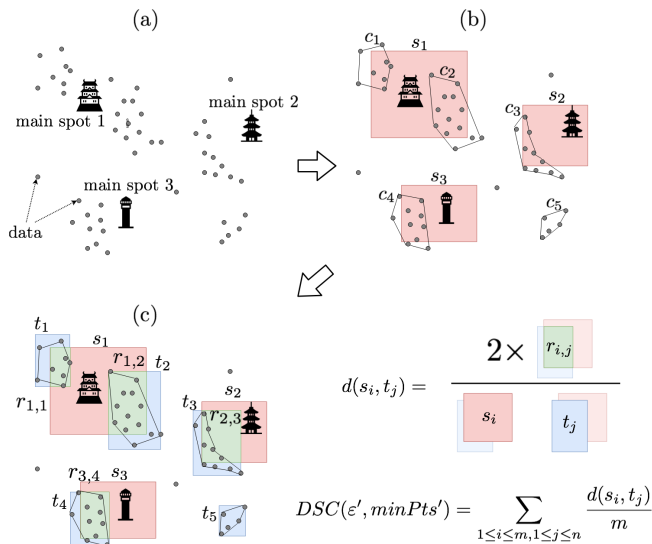


図2 Minimum Bounding Rectangle: DBSCANによりクラスタリングされるべきデータ点と、検出されるべき主要スポットが存在する (a). m 個の各主要スポットに対して、予めその矩形領域 s_i を定めておく. また, DBSCAN クラスタリングにより, n 個のクラスタ c_j が検出される (b). 各クラスタ c_j を完全に被覆する最小の矩形領域 t_j を定義することができる. その際, s_i と t_j が重複する領域を $r_{i,j}$ とする (c). これらの各矩形の組み合わせに対して, Dice 係数 $d(s_i, t_j)$ を計算し, その平均値を $\varepsilon = \varepsilon'$ と $\minPts = \minPts'$ の際の DBSCAN クラスタリングのスコア $DSC(\varepsilon', \minPts')$ とする.

本研究での景観における客観的特徴量としては、「美しい」・「伝統的」・「自然な」など、対象の様子を形容したり、或いは説明となる概念を指すものの分散表現とする.

4.2 概観

スポットの評価を行う提案手法の概観を図3に示す.

(a) クラスタリング

SNS 上に投稿される位置情報付き写真を対象に、景観の特徴表現を抽出することが提案手法における目的である. 3章で提案した手法を用いて、テキストデータが付与されている画像をクラスタリングし、分析対象とすべきスポットと写真集合を明らかにする.

(b) テキストを1つのコーパスに統合

景観地を表現する文章を分析することにより、使用されている単語を意味を考慮した関係性を得ることが狙いである. 文書集合として、全ての画像に付与されている文章を一つにマージし、景観評価における学習コーパスとする.

(c) スポット毎にテキストを統合

同一クラスタ内の写真に付与されているテキストデータは、同一スポットに対する描写であると考えられる. そこで、スポットにおける説明文書集合として、各クラスタ内の文章の一つにマージし、それぞれのスポットの文書として扱う.

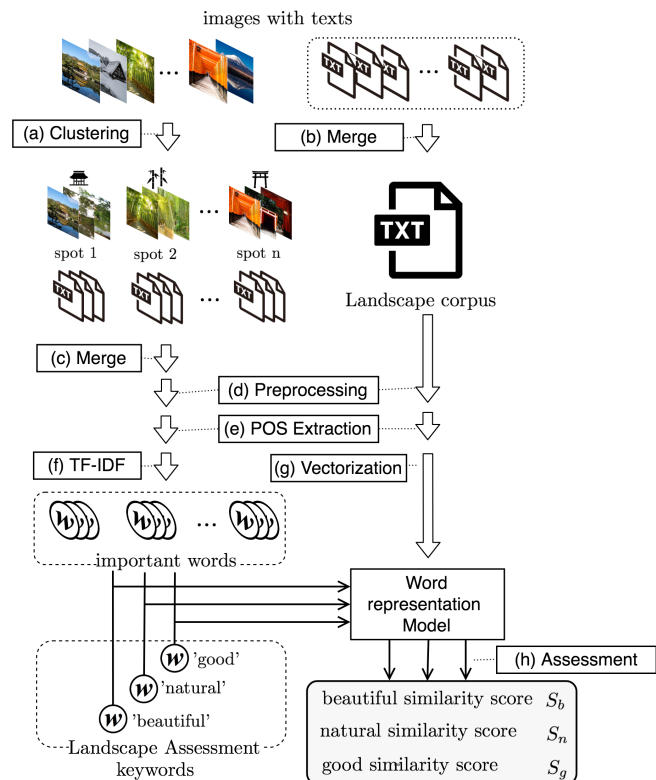


図3 スポットの評価手法の概観

(d) 前処理

(b) あるいは (c) で得られたテキスト文書集合に対して、前処理を行う. SNS 上に投稿される画像という性質上、URL や HTML のタグといった、被写体とは無関係の文字列も含まれており、これらのノイズは事前に取り除くことでコーパスの品質を高める.

(e) 品詞に基づく単語抽出

入力テキストデータに対し、形態素解析に基づいて単語への分割処理と品詞解析を行う. また、主観的および客観的な特徴量を表す表現として、特定の品詞のみを抽出する. スポットを表す特徴的な語を抽出し、コーパスから、主観的な特徴表現と客観的な特徴表現との関係性を捉えることが目的である.

(f) TF-IDF による重要語の発見

各スポット毎に、重要語を発見すべく、TF-IDF の計算を行う. あるスポットの文章内で数多く出現する一方で、他のスポットの文章内には出現していない語ほど、そのスポットを特徴付ける重要なワードであると考え、単語毎に重み付けをおこなう.

(g) 単語分散表現のモデルの獲得

コーパス内に含まれている単語をベクトルに変換する処理を行う. 分散表現を使用することで、大量の単語を低次元に圧縮して表現する.

(h) 類似度に基づく景観評価

(f) で取得された各スポットの重要ワードと、(g) で獲得した単語空間を用いることで、各スポットの景観評価を行う. 環境心理学で利用されている分類基準を利用し、3つの指標に対して類似度計算を行うことでスコアを算出する.

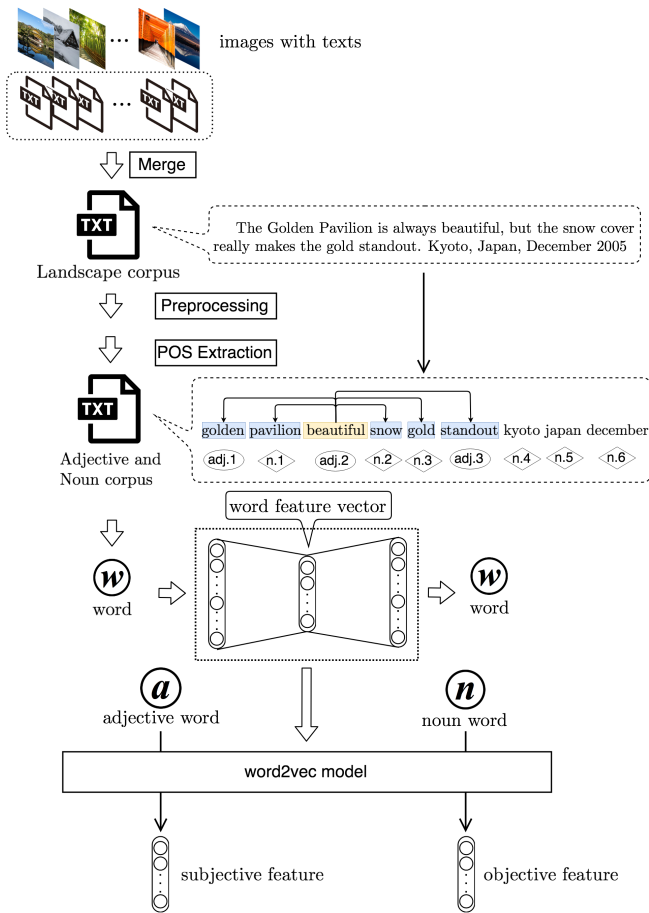


図4 テキストデータからの特徴抽出プロセス

4.3 テキストデータからの特徴抽出と、景観の評価

本研究における主観的特徴量および客観的特徴量を表す語として、テキストデータのうち形容詞と名詞を扱う。形容詞と名詞のみに対象を限定し、単語空間を獲得するモデルを構築することで、単語間の関係性を得ることを目的とする。図4はテキストデータから特徴抽出を行うプロセスを示したものである。

本手法は景観の特徴を、景観画像に付与されているテキストデータを利用することによって抽出を試みるものである。また、その際に、景観の主観的及び客観的な特徴量を表現する要素として、名詞および形容詞に焦点を当て、その関係性を深層学習の応用により獲得しようとするものである。本手法ではword2vecでの手法を景観地のコーパスデータに応用する。

本研究では、SNS上に投稿する際に併せて付与されている説明文をマージしたものを学習コーパスとして利用する。言語処理を行う上で前処理は必要不可欠である。とりわけ、SNS上に投稿されている画像には、URLやHTMLタグなど、景観とは無関係の内容も数多く含まれているため、このようなノイズとなるテキストを取り除いておくことで、分析の精度の向上に結びつく。

通常のword2vecでは、膨大な種類が存在するコーパス内の単語を、数百次元にまで次元削減することにより、単語の意味情報を保った分散表現を獲得することができる。本プロセスによって達成されるべき目標は、客観的特徴量を表す名詞と、主

観的特徴量を表す形容詞との間に存在する相関関係を、分散表現の形式で獲得することである。本手法では、前処理を行ったテキストデータをそのままword2vecモデルの学習コーパスとして用いるのではなく、着目すべき品詞を絞り、より効果的な表現を獲得することを目指す。そこで、形態素解析により得られた品詞情報のもと、集約したコーパスから名詞と形容詞のみに絞って単語を抽出し、登場した順に並べたものを学習すべきコーパスとして用意する。

5 実験

5.1 YFCC100M データセット

本項では、使用するデータセットYFCC100M [13]について説明する。YFCC100Mは、2004年から2014年にFlickrにアップロードされた約9920万枚の画像と80万本の動画から構成されており、多種多様なメタデータが添付されている。メタデータには、Flickr上でのID・ユーザ情報・撮影機材に関する情報・撮影時刻・投稿時刻・緯度経度による位置情報・タイトル・説明文・タグなどの情報が含まれている。

本研究では、京都府・大阪府・滋賀県・奈良県で撮影された231245件を対象に、画像とメタデータの収集を行った。

収集した画像データに付与されているテキストデータのうち、説明文(Description)を対象に統計情報分析を行った。有効な説明文のテキストが付与されているデータは、全231245件中84955件であった。この84955件の説明文を形態素解析にかけ、含有するワードに対する調査を行った。形態素解析にはPythonのライブラリの一つであるnlTK¹を使用した。表1は、名詞(単数形)・形容詞それぞれについて、含まれるワードの例と件数を示したものである。

表1 名詞と形容詞のそれぞれのワードの例とその件数

noun(singular)	count	adjective	count
temple	4287	Japanese	3099
castle	2589	main	1972
shrine	2322	famous	1650
area	2050	original	1323
photo	1751	great	1084
garden	1749	old	785
bei	1524	good	761
picture	1739	small	749
city	1410	important	674
building	1284	great	1084
culture	1144	popular	587
station	1137	traditional	585
hall	1071	beautiful	561
way	991	large	542
Country	957	wooden	527

表1から読み取れるように、“temple”, “castle”, “shrine”といった客観的特徴量を表現する名詞が含まれている。また、形

1: <https://www.nltk.org/>

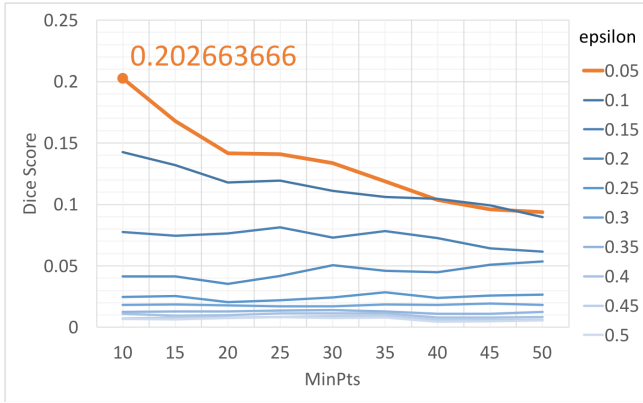


図5 大域的にパラメータを設定した際の結果

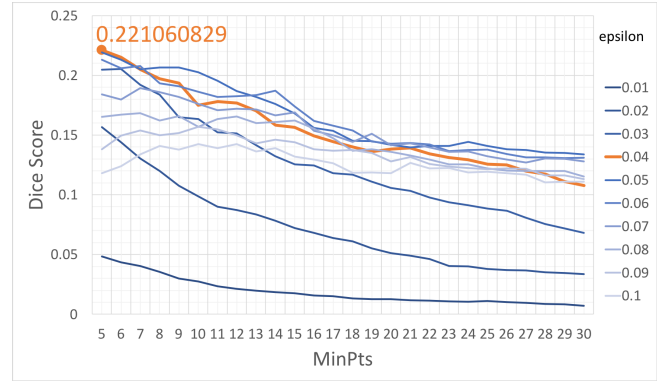


図6 ϵ の値を小さく設定した際の結果

容詞の中には“great”, “old”, “small”, “traditional”, “beautiful” など、主観的特徴量を表現していると考えられるワードも存在している。一方, “photo”, “original” など、画像の内容とは無関係のワードも見られる。これは、説明文やタグに撮影機材・パーソナル情報などを記載する Flickr サービスの風習に影響を受けていると考えられる。

5.2 クラスタリングによるスポット検知

3節で述べた手法を用いて、DBSCAN クラスタリングと、そのパラメータチューニングを行った。収集したデータの内、京都市内で撮影された 120321 件のデータ点をクラスタリングの対象とした。また、予め指定しておく検出されるべき主要スポットとして、京都観光旅館連盟が発行している「京都の観光便利帖」に掲載されている 177 の主要寺社仏閣を定めた。

まず始めに、パラメータの範囲を、

$$\epsilon = \{0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5\}$$

$$\text{minPts} = \{5, 10, 15, 20, 25, 30, 35, 40, 45, 50\}$$

と大域的に設定してチューニングを行った。図5は、出力結果を示したものである。

その結果、 $\epsilon = 0.05, \text{minPts} = 5$ のときに、最も高いスコアを得た。また、スコアの分布からも見られるように、半径 ϵ の値が小さいほど、スコアは高くなる傾向が見られた。一方、 minPts の値はスコアに大きな影響を与えている傾向は確認できなかった。

そこで、パラメータの範囲を、

$$\epsilon = 0.01i \quad (1 \leq i \leq 10)$$

$$\text{minPts} = j \quad (5 \leq j \leq 30)$$

と、 minPts の範囲は大きく変更せず、 ϵ の値を比較的小さく設定し、さらに細かい値の範囲でチューニングを行った。図6は、出力結果を示したものである。

その結果、 $\epsilon = 0.04, \text{minPts} = 5$ のときに、最も高いスコアを得た。図5,6を見ると、 ϵ が大きくなるにつれて、スコアは小さくなる傾向があることが分かる。しかし、 $\epsilon = 0.01$ の時など、 ϵ の値が極めて小さいときは、大小関係が逆転している。また、 minPts に関しては、 ϵ が小さい場合は、 minPts が大

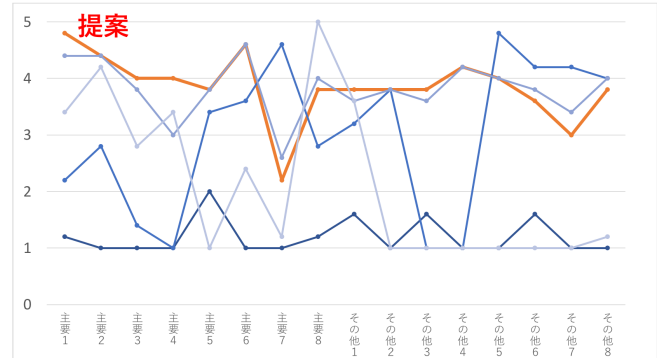


図7 クラスタリング結果に対するユーザ評価平均値の比較

きくなるにつれてスコアは小さくなる傾向にあるが、それ以外のケースにおいては、スコアに及ぼす影響はあまり見られない。

次に、得られたパラメータの妥当性を評価すべく、ユーザ実験を行った。獲得された $\epsilon = 0.04, \text{minPts} = 5$ のペアを含む、5通りのパラメータの組合せで、16箇所のスポットに対してクラスタリングの妥当性を5段階で評価を行った。16箇所のスポットには、主要なスポット8箇所と、それ以外のスポット8箇所を対象にした。図7はユーザ評価値の平均を比較した結果を示したものである。我々の提案手法で得られたパラメータは、最も高い妥当性を示した。これ以降では、 $\epsilon = 0.04, \text{minPts} = 5$ でのクラスタリング結果を利用して実験を進める。

5.3 テキストデータからの特徴抽出と、景観の評価

今回の実験では、クラスタリングにより検知された各スポットに対する分析として、各スポットがどれほど「美しい」スポットか、あるいはどれほど「自然な」スポットか、あるいはどれほど「良い」スポットであるかを求めたいものとする。これらの基準は、Hsu らが環境心理学における景観選好の基準として有効であるとした因子構造に基いて定めたものである [5]。

3つの指標に対するスコアは、TF-IDF により導出した各スポットの重要語と、'beautiful', 'natural', 'good' の各語との類似度により求める。類似度計算の際に使用するモデルの作成方法を変化させることにより、分類結果にどのような差が生じるかを確認する。図8は、類似度計算によるスポットの評価分類プロセスを示したものである。

収集したデータセットから有効な説明文の文章テキストを

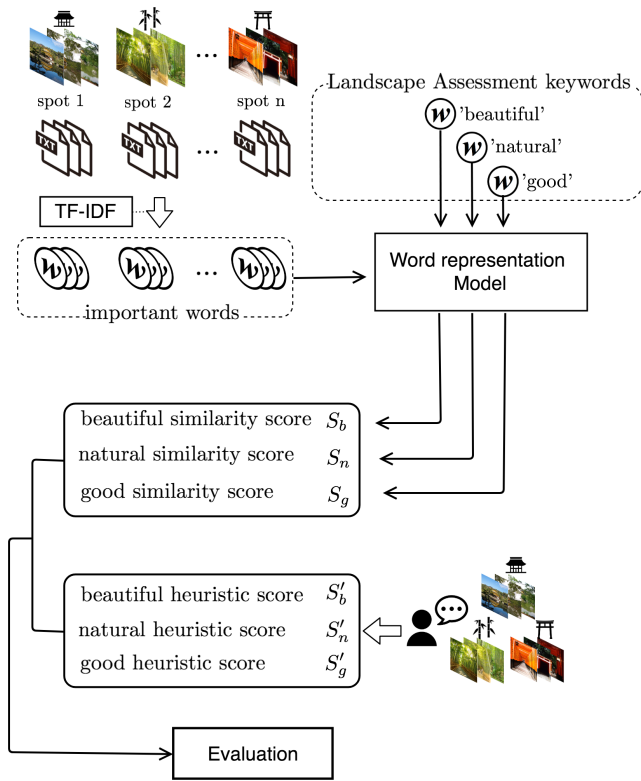


図 8 類似度計算実験のプロセス

マージし、品詞による単語抽出を行わずにそのまま入力コーパスとして扱ったものをベースラインとする。また、学習時に 1 ワードに対して考慮される語数の大小が結果にもたらす影響を確かめるために、ウィンドウサイズの値を変化させた際の結果についても比較を行った。モデルの学習方法には skip-gram を用いた。

提案手法で算出したスポットの評価と、ユーザ視点でのスポットの評価との相関を確かめるべく、ユーザ評価実験を行った。クラスタリングの実験と同じく、16 箇所のスポットに対して評価を下した。各スポットの景観画像に対して、「美しい」・「自然だ」・「良い」の 3 つの指標をそれぞれ 5 段階で点数を付けた。ユーザ評価の平均値と、提案手法により得られた類似度との相関値を算出し、ウィンドウサイズ毎に提案手法とベースラインとの比較を行った。図 9,10,11 は実験結果を示したものである。

結果から見て取れるように、我々の提案手法は、ベースラインと比較すると、ユーザ視点での景観評価と高い相関を示すことが確認できた。しかしながら、指標毎に結果にばらつきがあることや、相関値の絶対値が高いものでも 0.4 以下に留まっていることから判断できるように、モデルやコーパスの作成方法には取り組むべき課題が残されていると言える。これは、各スポット毎に付与されているテキスト量に差があることが原因の一つとして挙げられる。主要スポットは投稿されるデータ数も多いため潤沢なテキスト量が確保できる一方で、本来発見すべき潜在的観光資源には十分なテキストが付与されていないケースも多く、スポットに対する適切な特徴量抽出に至っていないと考えられる。

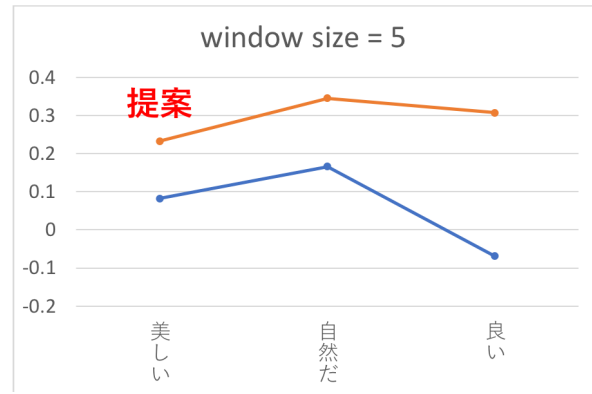


図 9 window size = 5 の時の結果

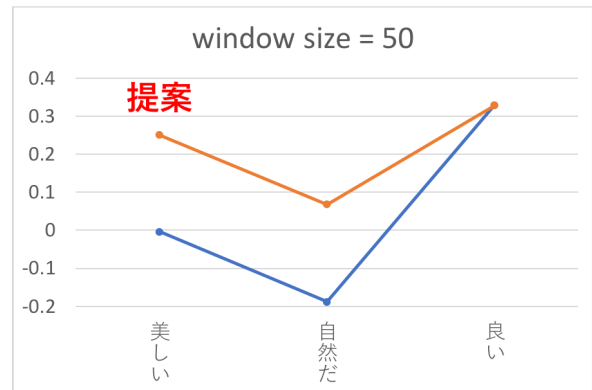


図 10 window size = 50 の時の結果

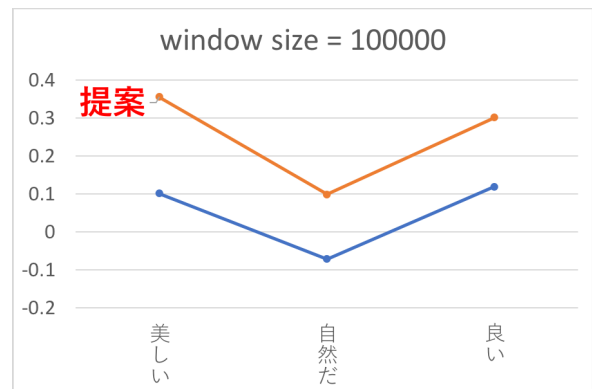


図 11 window size = 100000 の時の結果

6 まとめと今後の課題

観光産業は日本において重要な成長分野であり、政府は観光の促進に力を入れる一方で、観光地の局所集中が大きな問題となっている。本研究では、ユーザー視点で多様な観光資源を発見するために、ユーザが SNS を上にアップロードした画像を分析することで、観光スポットを検知し、そしてそのスポットが持ちうる潜在的な価値を評価する手法を研究する。写真に付与された地理情報を利用する DBSCAN クラスタリングを行うことで、地図上にスポットを表すクラスタを検知することは可能である。DBSCAN クラスタリングは、クラスタ数を事前に決める必要がなく、様々な形や大きさのクラスタを発見す

ることが可能なため、地理情報付きデータ点のクラスタリングに適している。しかし、DBSCAN クラスタリングを行うには2種類のパラメータが必要であり、これらのパラメータの適切な値を求めるのは困難な課題である。そこで我々はまず最初に、DBSCAN アルゴリズムにおける二種類のパラメータを自動調整する手法を提案した。クラスタリングによって生成されたクラスタを完全に被覆する最小の領域矩形 MBR (Minimum Bounding Rectangle) と、既知の有名スポットの範囲情報を用い、それらの重複領域から計算されるスコアによって、クラスタリングに必要なパラメータを自動調整する手法を提案した。実験では、提案手法により得られたパラメータを含む、複数のパラメータの組合せに対してユーザ評価実験を行い、提案手法が最も高い妥当性を示すパラメータを自動で調整していることを確かめた。また、我々は、景観の評価についても研究した。従来までの景観評価の手法は、人間の経験や知識に基いた評価尺度に依存しており、多様性のある景観の評価としては一般性に書いている問題点がある。また、従来研究において整理されてきた特徴量は、海外の景観をベースに提案されてきたものが多く、日本の景観の特徴をうまく表現できない場合がある。そこで我々は、2種類の特徴量を定義し、景観値についてのデータからこれら2種の特徴量を深層学習を応用して抽出する手法を提案した。まず特徴量として、客観的な特徴量と主観的な特徴量を定義した。客観的な特徴量とは、スポットを構成する物体や、その色やパターンなど、視覚的に判断が容易なものであり、この客観的な特徴量は従来の景観評価手法においても利用されてきた。次に、主観的な特徴量とは、景観の佇まいや雰囲気など、具体的には認識できないが、スポットの風情や魅力に大きな影響力を与えている要因である。我々の手法では、景観スポットの画像に付与されたテキストデータから、主観的な特徴量と客観的な特徴量の2つを抽出し、環境心理学の分野で定義された尺度を利用して評価を行う枠組みを執っている。テキストデータにおいて、名詞が客観的な特徴量を表現し、形容詞が主観的な特徴量を表現するという仮定した。スポット毎に抽出した重要語と、word2vec の学習モデルを利用して得られた分散表現を利用し、単語間の類似度に基づいて計算を行った。実験結果によると、ベースラインと比較して、提案手法はユーザ視点での景観評価と高い相関を示すことが確認できた。しかし、我々の手法は2種の特徴量の関係性を完全には抽出できていない。コーパスの処理や学習モデルのパラメータ調整は、今後の研究において精査される必要がある。また、テキストデータだけでなく画像データからも客観的な特徴量と主観的な特徴量を抽出する深層学習モデルについて研究を進めることで、景観の画像からの直接的な景観評価が可能となる。また、テキストデータから得られた特徴量と、画像データから得られた特徴量を共通空間上に写像することで、マルチモーダルな埋め込みベクトル空間を構築し、景観評価手法に応用していく研究も今後と課題として考えている。

謝辞

本研究の一部は総務省 SCOPE(172307001) による。

本研究のデータの提供にあたり、桂田佳代子氏（京な株株式会社）に御協力を戴いた。ここに感謝の意を表す。

文 献

- [1] 馬強 et al. 観光情報学の最前線-観光の分散化と個人化を促進する集合知活用情報技術. *情報処理*, 58(3):220–226, 2017.
- [2] Yizhu Shen, Min Ge, Chenyi Zhuang, and Qiang Ma. Sightseeing value estimation by analyzing geosocial images. In *Multimedia Big Data (BigMM), 2016 IEEE Second International Conference on*, pages 117–124, 2016.
- [3] Yizhu Shen, Chenyi Zhuang, and Qiang Ma. Element-oriented method of assessing landscape of sightseeing spots by using social images. In *Asia-Pacific Web (APWeb) and Web-Age Information Management (WAIM) Joint Conference on Web and Big Data*, pages 66–73, 2017.
- [4] Mari Tveit, Åsa Ode, and Gary Fry. Key concepts in a framework for analysing visual landscape character. *Landscape research*, 31(3):229–255, 2006.
- [5] HSU CHUNG-PENG. 景観選好における評価因子の構造に関する研究. PhD thesis, 東京大学, 2014.
- [6] Nirmala Kalidindi, A Le, J Picone, L Zheng, H Yaqin, and V Rudis. Scenic beauty estimation of forestry images. In *Southeastcon'97. Engineering new New Century., Proceedings.*, pages 337–339, 1997.
- [7] 田中平八. 環境景観に対する好悪感の規定因-認知モデルにもとづいた予測的評価法. *総合都市研究 特別号*, 56:37–46, 1995.
- [8] 芝田征司. 自然環境の心理学. *環境心理学研究*, 1(1):38–45, 2013.
- [9] Kaplan. R.and Kaplan. S.and Brown. T. Environmental preference: A comparison of four domains of predictors. *environment and behavior*, 21(5), 509–530.
- [10] Min Ge, Chenyi Zhuang, and Qiang Ma. A ranking based approach for robust object discovery from images of mixed classes. pages 71–83. AIRS 2017, 2017.
- [11] Chenyi Zhuang, Qiang Ma, Xuefeng Liang, and Masatoshi Yoshikawa. Discovering obscure sightseeing spots by analysis of geo-tagged social images. In *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*, pages 590–595, 2015.
- [12] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems*, pages 3111–3119, 2013.
- [13] Bart Thomee, David A Shamma, Gerald Friedland, Benjamin Elizalde, Karl Ni, Douglas Poland, Damian Borth, and Li-Jia Li. Yfcc100m: the new data in multimedia research. *Communications of the ACM*, 59(2):64–73, 2016.