

# 景観評価のための特徴自動抽出に向けたデータセットの作成について

木上 智貴<sup>†</sup> 馬 強<sup>†</sup>

<sup>†</sup> 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: <sup>†</sup>kinoue@db.soc.i.kyoto-u.ac.jp, <sup>††</sup>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、本研究では、深層学習の応用による特徴量の自動抽出を目指す。本稿では、特徴自動抽出のフレームワークとそれに用いる学習データセットを作成する手法について述べる。

キーワード 景観評価, 観光, データセット, 深層学習

## 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「世界が訪れたい日本」<sup>(注1)</sup>を目指し、観光立国推進基本計画の中で訪日外国人人数4000万人という目標を打ち立てたが、現状の増加率が続けばその達成も可能である。

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 (Information and Communication Technology, 以下、ICTと呼ぶ) 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行し、リピーターを含む訪日外国人旅行者数の増加を測ると同時に、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。

多様な観光ニーズに応え、観光の個人化を実現するため、多様な観光資源の発見と開発が重要である。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<sup>(注2)</sup> や Flickr<sup>(注3)</sup> 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]。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、画像から各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 [2,3]。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。

そこで、我々は景観画像の特徴づける特徴量を、人間のヒューリスティックによって決定するのではなく、深層学習を応用することにより自動的に抽出する手法について研究を行っている。それにより、多種多様な景観地に依拠して人間の知識から評価尺度を選択する必要なく、景観地のタイプに応じた評価を行うことが期待できる。

深層学習においては学習に用いるデータセットの質が結果を左右する大きな要因となる。そこで、本稿では、環境心理学における景観評価の尺度に対応したキーワードに注目し、画像に付与されているメタデータを対象に画像を分類することによって、特徴量の自動抽出に向けたデータセットを作成する手法を提案する。

本論文の構成は次の通りである。まず、第2節では環境心理学の領域における景観評価の分類と、それらを用いた景観評価についての関連研究について述べる。第3節では、景観評価のための特徴自動抽出のフレームワークについて述べる。第4節では、特徴抽出に必要な学習データセットの作成について説明する。第5節では成果のまとめと今後の課題について述べる。

## 2. 関連研究

### 2.1 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。景観のよし悪しは、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 [5-8]。Kaplan ら [8] 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。これらの評価因子は環境心理学において最も参照される標準的位置付けを得ている [9]。ま

(注1) : [http://www.mlit.go.jp/kankochou/news02\\_000307.html](http://www.mlit.go.jp/kankochou/news02_000307.html)

(注2) : <https://twitter.com/>

(注3) : <https://www.flickr.com/>

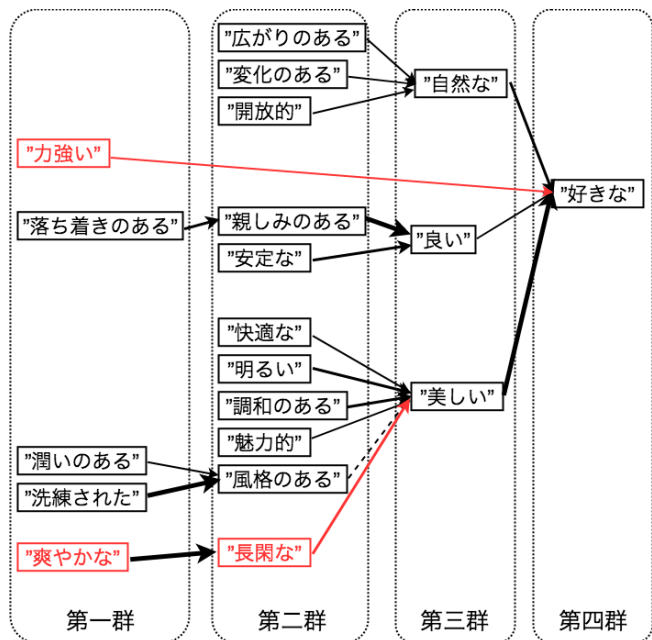


図1 選好に影響を及ぼす因子の構造

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 [9].

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、「好きな」という結果に至る変数として、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 [9]. 図 1 は、Hsu が整理した因果構造である。赤字は因果性メカニズムを説明しにくい 3 因子を表す。線の太さは因果関係の強さを表す。実線は有意、点線はやや有意であることを表す。黒線は正、赤線は負のパス係数を表す。

## 2.2 ヒューリスティックな景観評価

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 [2,3] 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。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 [10] 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。

Zhuang [11] 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。

これらの手法によって対応する尺度を数値化しスコアを算出することはできるが、画像処理で扱いづらい評価尺度は、数値化するメソッドの定義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。

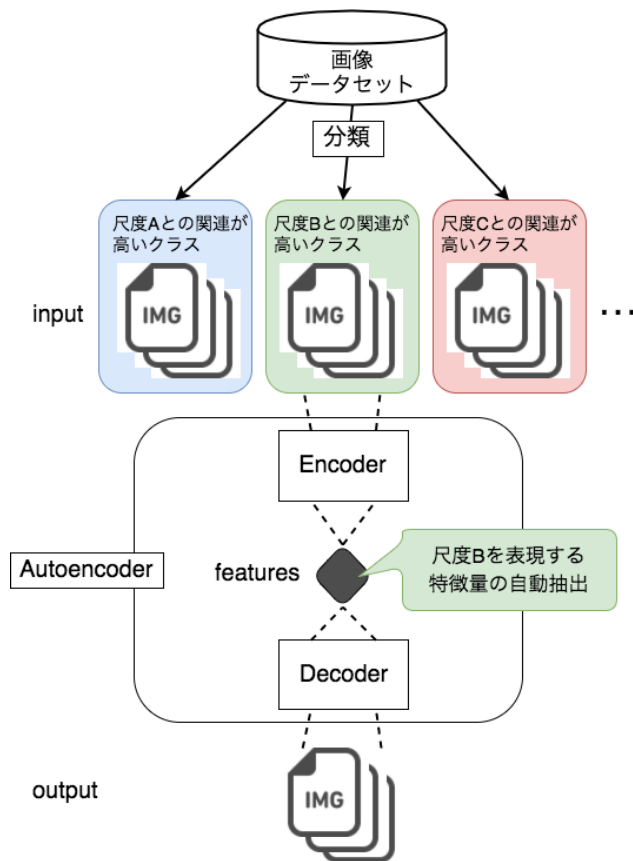


図2 特徴自動抽出のフレームワーク

## 3. 特徴自動抽出のフレームワーク

本研究では深層学習を応用することで画像から直接特徴量を自動抽出する手法を提案する。特徴自動抽出のフレームワークを図 2 に示す。特徴量の抽出にはオートエンコーダを用いる。ニューラルネットワークによって入力画像の次元削減を行い、得られる特徴表現によって入力と同じ画像が出力されるようにニューラルネットワークを学習する。このオートエンコーダによって入力画像群が有する特徴を人間の知識に基くことなく抽出することができる。

深層学習においては学習データセットが重要である。観光地には様々なタイプが存在しており、画像データセットをそのまま学習データセットとして用いたとしても、有効な特徴量を容易に得ることができない。本研究では全ての画像データセットに対してオートエンコーダで特徴抽出を行うのではなく、環境心理学の評価尺度それぞれに対して特徴抽出を行う。そのためには、使用する評価尺度をクラスとし、画像データセットを各クラスに分類する作業が必要である。次節では、評価尺度に対応するデータセットの作成について説明する。

## 4. データセットの作成

### 4.1 概要

本項では、オートエンコーダを利用した特徴自動抽出のフレームワークを実現するためのデータセット作成について説明する。データセット作成の概要を図 3 に示す。本稿ではデータ

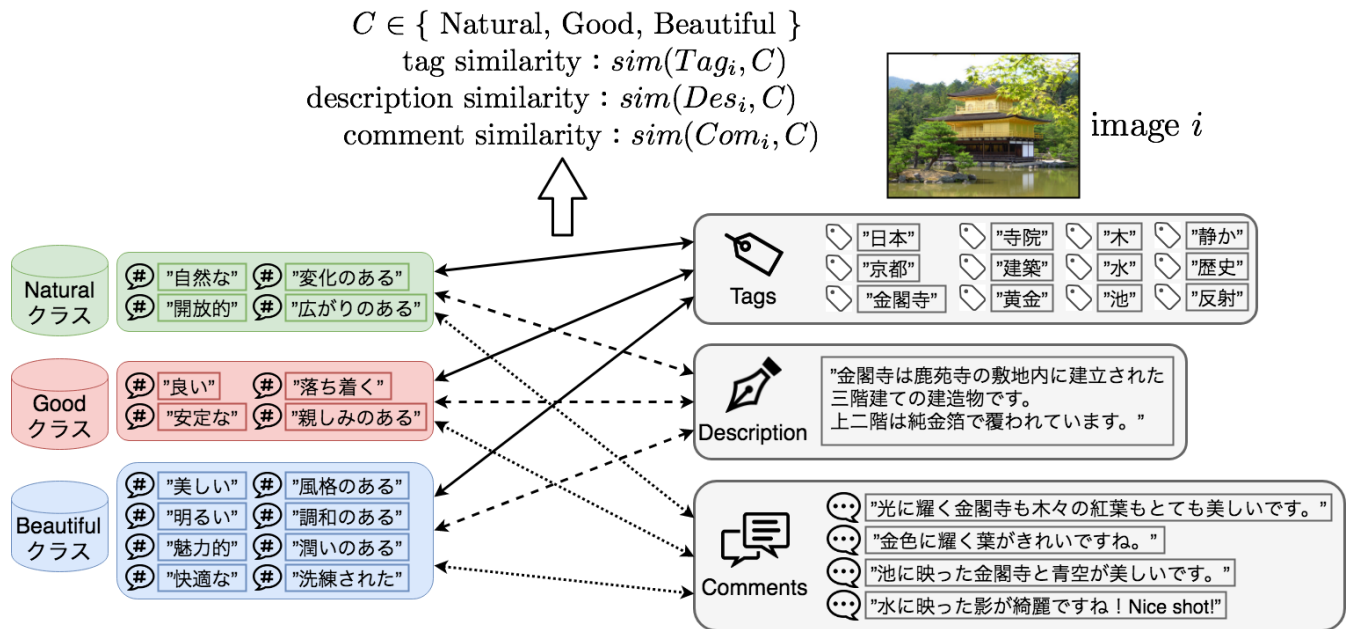


図 3 データセット作成のフレームワーク

セットの画像群に付与されたメタデータを利用することで、画像を Natural クラス・Good クラス・Beautiful クラスの 3 クラスへ分類する手法を提案する。この分類は、Hsu が明らかにした因子構造において、第四群の変数である「好きな」に関わる有効な指標として、第三群の「自然な」・「良い」・「美しい」の 3 変数が挙げられていることを参考にしている。付与されたテキスト形式のメタデータと、因子相関に基づいたキーワードとの類似度を基準に、画像の分類を行う。

#### 4.2 YFCC100M

分類する画像データセットには、YFCC100M [12] を利用する。YFCC100M は、2004 年から 2014 年に Flickr にアップロードされたクリエイティブ・コモンズのメディアオブジェクトから成るデータセットであり、約 9920 万枚の画像と 80 万本の動画から構成されている。メタデータには、Flickr の識別子・撮影者・撮影したカメラ情報・撮影時刻・アップロード時刻・位置情報・タイトル・説明文・タグが含まれている。コメント・お気に入り・フォロワー等の情報は含まれていないが、これらの情報は Flickr が公開している API から取得することができる。タグデータは人・動物・場所・旅行など豊富で多様なエンティティで構成されており、ユーザが付与するタグと、写真の内容から自動的に付与されるタグの 2 種類が存在する。表 1 は、画像検知により求められたメディアのタイプの上位を示すものである。

#### 4.3 キーワードの分散表現による分類

本項では、単語の分散表現を利用した画像分類について説明する(図 3)。flickr の画像に付与されているテキスト形式のメタデータには、タグ(Tag)・説明文(Description)・コメント(Comment)が存在する。Hsu [9] により構造が明らかにされた形容詞の因子(図 1)を、各クラスのキーワードと定義する。単語ベクトルから成る分散表現の獲得には、テキストコーパスに対する以下のような適切な前処理が不可欠となる。

表 1 The top of visually detected concepts.

Concept	Photos	Videos
outdoor	44,290,738	266,441
indoor	14,013,888	127,387
people	11,326,711	56,664
nature	9,905,587	47,703
architecture	6,062,789	11,289
landscape	5,121,604	28,222
monochrome	4,477,368	18,243
sport	44,354,325	25,129
building	4,174,579	7,693
vehicle	3,869,095	13,737
plant	3,591,128	11,815
blackandwhite	2,585,474	10,351
animal	2,317,462	9,236
groupshot	2,271,390	4,392
sky	2,232,121	11,488
water	2,089,110	15,426
text	2,074,831	5,623
road	1,796,742	12,808
blue	1,658,929	10,273
tree	1,641,696	6,808
hill	1,448,925	6,075
shore	1,439,950	8,602
car	1,441,876	4,067
head	1,386,667	8,984
art	1,391,386	2,248

##### 4.3.1 クリーニング

テキストコーパス内に含まれるノイズを除去する。除去するノイズは以下の通りである。

- " () ". " [ ] ". " [ ] " などの括弧
- "@ " から始まるメンションワード
- 全角の空白

- URL
- html タグ

### 4.3.2 形態素解析

日本語テキストデータをコーパスとしているため、形態素解析により単語の分割処理を行う。本研究では形態素解析器に、NEologd<sup>(注4)</sup> 辞書を追加した Janome<sup>(注5)</sup> を用いる。数値表現の重要性は低いため、数字を全て 0 に置換し、語彙数を削減する。取得するワードは、名詞・形容詞・形容動詞の原型のみに絞った。

### 4.3.3 ストップワードの除去

日本語ストップワード辞書の Slothlib<sup>(注6)</sup> から、景観に関連が深いと思われるワードを除き、コーパスから除去を行った。本研究において、Slothlib に含まれるが、除去を行わなかった単語は以下の通り。

”春” ”夏” ”秋” ”冬” ”村”

### 4.3.4 分散表現

Word2Vec [13] により、画像に付与されている各メタデータ及び、各クラスのキーワードの分散表現を扱うことができる。Word2Vec により、作成したコーパスからモデルを構築することで、大量の単語を数百次元のベクトルに圧縮することが可能となる。それぞれの単語について単語ベクトルが表現できるため、コサイン類似度などの手法で単語間の類似度を測ったり、あるいは単語同士での演算を行うことができる。本手法においては、画像に付与されたテキストデータ中に含まれている単語と、各クラスにおけるキーワードの 2 語において、それぞれモデルから分散表現を得ることができるため、単語間あるい単語と文書の類似度を計算することが可能である。これらの各類似度を、画像が景観選好に影響を及ぼす因子の度合いとみなすことで、画像のクラス分類への適用を提案する。

こうして各クラス別に分類された画像群を学習データセットとして、3 節で述べた深層学習のフレームワークが実現される。また、景観評価を行うにあたり、得られたデータセットは位置情報でクラスタリングすることにより、観光地の景観画像データとして扱うことが可能であると考えられる。

## 4.4 類似度計算

画像  $i$  のタグ集合  $Tag_i$  と、クラス  $C \in \{\text{Natural, Good, Beautiful}\}$  の類似度  $sim(Tag_i, C)$  は以下のように求められる。

$$sim(Tag_i, C) = \sum_{l=1}^M \sum_{j=1}^N \frac{f(t_l, k_{C_j}) \cdot w_{k_{C_j}}}{M \cdot N} \quad (1)$$

ここで、 $N$  はクラス  $C$  に含まれるキーワードの総数、 $k_C$  はクラス  $C$  に含まれるキーワードの集合、 $w_{k_{C_j}}$  は、クラス  $C$  に含まれるキーワード  $k_{C_j}$  の重みとなるパス係数、 $M$  は画像  $i$  に付与されているタグでコーパスに含まれているものの総数、 $t_l$  は画像  $i$  に付与されている  $l$  番目のタグ、 $f(a, b)$  は単語  $a$  と単語  $b$  の類似度を求める関数である。

画像  $i$  の任意のタグ  $t_l (1 \leq l \leq M)$  について、クラス  $C$  の各キーワード  $k_{C_j} (1 \leq j \leq N)$  との類似度は、関数  $f$  によって計算される。その値にパス係数による重み  $w_{k_{C_j}}$  を掛け合わせた値の平均値を、タグ  $t_l$  とクラス  $C$  の類似度  $sim(t, C)$  と定義する。画像  $i$  の全てのタグに対する類似度  $sim(t, C)$  の平均値を、画像  $i$  とクラス  $C$  の類似度  $sim(Tag_i, C)$  と定義している。

また、画像  $i$  の説明文に処理を行い抽出されたワードの集合  $Des_i$  と、クラス  $C \in \{\text{Natural, Good, Beautiful}\}$  の類似度  $sim(Des_i, C)$  は、式 1 において

$$Tag_i = Des_i$$

と代入することで求まる。

同様に、画像  $i$  のコメントに処理を行い抽出されたワードの集合  $Com_i$  と、クラス  $C \in \{\text{Natural, Good, Beautiful}\}$  の類似度  $sim(Com_i, C)$  は、式 1 において

$$Tag_i = Com_i$$

と代入することで求まる。

$sim(Tag_i, C)$  と  $sim(Des_i, C)$  と  $sim(Com_i, C)$  の平均値を、画像  $i$  とクラス  $C$  との総合的な類似度と定義し、分類の指標とする。

## 5. 実 験

本実験では、Flickr に付与されているテキストデータの内、タグデータのみを対象にし、各クラスのキーワードとの類似度に基づいて、画像の分類を行った。

### 5.1 コーパスデータ

景観選好に影響を及ぼすキーワードに対応する単語の分散表現を得るためのテキストデータソースとして、以下の 2 種類のデータベースを用いた。

#### 5.1.1 ALAGIN 京都観光ブログの評価情報付与データ<sup>(注7)</sup>

京都観光ブログは、日本語ブログ記事のデータベースである。京都観光を中心とした内容で、執筆者は 47 名、合計 1041 記事 (平均 480 字) から構成されている。データには、記事 ID・記事タイトル・執筆者名 (ハンドルネーム)・ブログ本文・執筆日が付与されている。本実験では、全 1041 記事のブログ本文を利用し、各記事を一つの文書としてコーパスに加えた。

#### 5.1.2 CD-毎日新聞 2007 データ集<sup>(注8)</sup>

毎日新聞の東京・大阪本社の朝夕刊最終版を対象とした、1994 年度から 2007 年度までのタグ付全文記事データ集である。データには、記事 ID・キーワード・掲載日付・記事本文などのタグが付与されている。本実験では、全 96279 件の記事の内、「国際」・「経済」・「スポーツ」のカテゴリを除いた、59667 件の記事本文を利用し、各記事を一つの文書としてコーパスに加えた。

### 5.2 画像データ

#### 5.2.1 対 象

flickr にアップロードされている画像から、「金閣寺」・「富士

(注4) : <https://github.com/neologd/mecab-ipadic-neologd>

(注5) : <http://mocobeta.github.io/janome/>

(注6) : <http://svn.sourceforge.jp/svnroot/slothlib>

(注7) : <https://alaginrc.nict.go.jp/resources/nict-resource/li-info>

(注8) : <http://www.nichigai.co.jp/sales/mainichi/mainichi-data.html>



図 4 使用した画像例

山」・「白川郷」・「伏見稲荷大社」・「嵐山」のワードがタグ付けされている画像を、それぞれ約 10 枚ずつランダムに選出した。

### 5.2.2 タグ

選出した画像に付与されているタグは、以下のタイプに大別される。

#### (1) 名称タグ

撮影されている対象の名称を記載したタグである。(例) “金閣寺”・“Kinkaku-ji”・“kinkakuji”

#### (2) エリアタグ

撮影されているスポットの地域を記したタグである。(例) “京都”・“kyoto”・“日本”・“Japan”・“geo:lat=36.2631033”

#### (3) 機材タグ

撮影に仕様した機材に関するタグである。(例) “nikon24120f4”・“Canon EOS 60D”・“Mirrorless”

#### (4) 特徴タグ

撮影されているスポットの特徴・要素・説明を記したタグである。(例) “湖”・“黄金”・“反射”・“世界遺産”・“symmetry”・“sky”・“mountain”・“serene”・“outdoor”

#### (5) その他のタグ

その他の、写真とは直接関係のないタグである。(例) “観光”・“gettyimagesJapan13Q1”・“DiamondClassPhotographer”

本実験では、これらのタグの内、特徴タグのみを抽出した。また、日本語データをコーパスとして扱っているため、外国語のタグは日本語に翻訳を行い利用する。

本実験で利用した画像例を、図 4 に示す。また、本実験で利用した画像のタグの例を、表 2 に示す。

表 2 使用した画像に付与されたタグの例

金閣寺	富士山	白川郷	伏見稲荷大社	嵐山
Asia	Fuji	Snow	Fushimi	嵯峨野
Japan	Lake	White	Inari	京都
Nikon	cloud	House	Shrine	嵐山
Travel	sky	Old house	伏見稲荷大社	新緑
nikon24120f4	grass	合掌造り	Shinto	Kyoto
Kinkakuji	water	白川郷	Kyoto	日本
金閣寺	mountain	Shirakawago	Japan	Japan
Golden	landscape	Sky	2014	landscape
Pavilion	serene	Blue	geometric	forest
tourism	tree	Pond	alley	outdoor
kyoto		outdoor		road
garden		snowbank		serene

### 5.3 実装

本実験では、画像に付与されているテキストデータの内、タグのみを対象に類似度計算を行った。すなわち、画像  $i$  とクラス  $C$  の類似度は、画像  $i$  のタグ集合  $Tag_i$  とクラス  $C$  の類似

度  $sim(Tag_i, C)$  で表されるという条件のもとで実験を行った。

各クラスのキーワードとして図 1 の内、第二群と第三群のワードを用いた。また、類似度に掛け合わせる重みとして、Hsu [9] が因子分析で求めたパス係数を利用した。

### 5.4 結果と考察

#### 5.4.1 結果

表 3, 表 4, 表 6, 表 5, 表 7 は、各タグ付き画像に対する、各クラスとの類似度と、最も類似度が高いクラスを示している。

表 3 金閣寺の画像の分類結果

$i$	class	$sim(i, N)$	$sim(i, G)$	$sim(i, B)$
1	Good	0.1707	0.1709	0.1679
2	Natural	0.1856	0.1774	0.1691
3	Good	0.1689	0.1737	0.1343
4	Natural	0.1470	0.1396	0.1405
5	Good	0.1427	.01484	0.1408
6	Natural	0.1588	0.1472	0.1429
7	Natural	0.1471	0.1396	0.1405
8	Natural	0.1465	0.1434	0.1328
9	Natural	0.1530	0.1373	0.1391
10	Natural	0.1719	0.1558	0.1626

表 4 富士山の画像の分類結果

$i$	class	$sim(i, N)$	$sim(i, G)$	$sim(i, B)$
1	Natural	0.1692	0.1433	0.1523
2	Natural	0.1697	0.1446	0.1209
3	Natural	0.1771	0.1652	0.1572
4	Good	0.1800	0.1830	0.1662
5	Good	0.1525	0.1597	0.1505
6	Natural	0.1621	0.1397	0.1493
7	Natural	0.1559	0.1474	0.1523
8	Natural	0.1710	0.1607	0.1678
9	Good	0.1429	0.1468	0.1326
10	Natural	0.1590	0.1452	0.1521

表 5 白川郷の画像の分類結果

$i$	class	$sim(i, N)$	$sim(i, G)$	$sim(i, B)$
1	Beautiful	0.1624	0.1633	0.1732
2	Good	0.1684	0.1779	0.1657
3	Good	0.1619	0.1746	0.1492
4	Beautiful	0.1471	0.1482	0.1486
5	Natural	0.1504	0.1279	0.1351
6	Natural	0.1591	0.1549	0.1492
7	Natural	0.1599	0.1580	0.1485
8	Good	0.1598	0.1727	0.1645
9	Natural	0.1495	0.1143	0.1416
10	Natural	0.1593	0.1536	0.1523

表 6 伏見稲荷大社の画像の分類結果

$i$	class	$sim(i, N)$	$sim(i, G)$	$sim(i, B)$
1	Good	0.1574	0.2115	0.1796
2	Good	0.1423	0.1531	0.1434
3	Good	0.1465	0.1637	0.1554
4	Good	0.1495	0.1942	0.1761
5	Good	0.1441	0.1847	0.1451
6	Good	0.1427	0.1592	0.1519
7	Good	0.1398	0.1595	0.1539
8	Beautiful	0.1131	0.1191	0.1305
9	Beautiful	0.1551	0.1594	0.1615
10	Beautiful	0.1225	0.1317	0.1353
11	Good	0.1601	0.1731	0.1567

表 7 嵐山の画像の分類結果

$i$	class	$sim(i, N)$	$sim(i, G)$	$sim(i, B)$
1	Natural	0.1455	0.1419	0.1393
2	Beautiful	0.1478	0.1189	0.1652
3	Beautiful	0.1600	0.1538	0.1655
4	Good	0.1802	0.1836	0.1561
5	Natural	0.1740	0.1262	0.1499
6	Good	0.1694	0.1780	0.1656
7	Natural	0.1918	0.1849	0.1676
8	Good	0.1681	0.1720	0.1530
9	Natural	0.1558	0.1260	0.1468
10	Good	0.1849	0.1878	0.1701
11	Natural	0.1637	0.1559	0.1556

#### 5.4.2 評価基準

soft classification の形式を執る今回の分類結果を、ケーススタディとして評価するため、正解のクラスを定め評価を行った。Hsu は、対応分析により評価対象と評価言語の関係を分析した結果、「山や池」と「変化のある」、「住宅や路地」と「落ち着いたもの」、「寺院や古民家」と「魅力的」の間に近い布置関係があることを明らかにした [9]。この布置関係と図 1 の因子構造から、本実験での正解クラスを以下のように定める。

金閣寺 : Beautiful クラス  
 富士山 : Natural クラス  
 白川郷 : Good クラス  
 伏見稲荷大社 : Beautiful クラス  
 嵐山 : Natural クラス

#### 5.4.3 評価

- 正解率 *Accuracy*

テスト画像データ数は 52 個、分類結果が正解だった画像数は 18 個。よって、正解率は 0.3462 であった。

- 適合率 *Precision*

実験結果に対する各クラス  $C$  の分割表は以下の通り。

表 8  $C_N$  の分割表

	$C_N$ に属する	$C_N$ に属さない
$C_N$ に分類	12	12
$C_N$ 以外に分類	不明	不明

表 9  $C_G$  の分割表

	$C_G$ に属する	$C_G$ に属さない
$C_G$ に分類	3	18
$C_G$ 以外に分類	不明	不明

表 10  $C_B$  の分割表

	$C_B$ に属する	$C_B$ に属さない
$C_B$ に分類	3	4
$C_B$ 以外に分類	不明	不明

全画像データセットの枚数とそのクラスの内訳が不明であるが、各クラス  $C$  に対する適合率は計算する事ができる。

表 8 より、Nature クラスの適合率は 0.5 であった。表 9 より、Good クラスの適合率は 0.1429 であった。表 10 より、Beautiful クラスの適合率は 0.4286 であった。

#### 5.4.4 考察

正解率が低くなった背景には以下の要因の影響が推測される。

- 景観評価に関するコーパスから単語毎の分散表現を得ようと試みたが、コーパスに占める毎日新聞データの割合が非常に大きいため、トピックに十分沿った特徴空間が得られなかった。

- 本実験では Flickr に付与されているタグデータのみを扱ったため、類似度の算出においてはデータが不十分である可能性がある。

- 今回のモデルにおいて、各クラスを特徴づけるキーワードの選定及びその重みの適正が不明瞭である。

- コーパスからは日本語の単語特徴空間を得たが、Flickr に付与されているテキストデータは外国語の割合が多く、単純な翻訳では正しい単語ベクトルを選択できていない。

また、Good クラスの適合率が低くなった背景には以下の要因の影響が挙げられる。

- Good クラスはキーワードの数が少ないため、類似度の高いワードの有無が大きく影響する。

- “良い”というワードが非常に幅の広い単語であるため、意図通りの分散表現が得られていない。

## 6. まとめと今後の課題

本稿では、景観評価を行うために深層学習を応用して特徴自動抽出を行うフレームワークの実現に向けた、データセットの作成手法について述べた。具体的には、環境心理学の領域で提唱されている因子相関に基づき、対応するキーワード群と、画像に付与されているテキストデータとの類似度を基準に分類を行う手法について述べた。

以下の事項を今後の課題と考えている。

- 画像に付与されているタグデータだけでなく、説明文およびコメントのデータも利用し、類似度計算を行う。

- 景観評価というトピックにより沿った分散表現を得るべく、観光レビューサイトや観光ブログをクローリングすることにより適切かつ十分量のコーパスを獲得する。

- 獲得した日本語の単語ベクトルと外国語との適切なマッ

ピング手法を得る。或いは、日本語の分散表現だけでなく、外国語の分散表現の利用を検討する。

- 現在は関連研究で利用されたパス係数を類似度の重みとして用いているが、単語の頻度情報などを利用した新たな重みの利用により、結果の差異を確認する。

## 謝 辞

本研究の一部は科研費(16K12532)と総務省SCOPE(172307001)による。

## 文 献

- [1] 馬強ほか. 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, Vol. 58, No. 3, pp. 220–226, 2017.
- [2] Yizhu Shen, Min Ge, Chenyi Zhuang, and Qiang Ma. Sight-seeing value estimation by analyzing geosocial images. In *Multimedia Big Data (BigMM), 2016 IEEE Second International Conference on*, pp. 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*, pp. 66–73, 2017.
- [4] Mari Tveit, Åsa Ode, and Gary Fry. Key concepts in a framework for analysing visual landscape character. *Landscape research*, Vol. 31, No. 3, pp. 229–255, 2006.
- [5] 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.*, pp. 337–339, 1997.
- [6] 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, Vol. 56, pp. 37–46, 1995.
- [7] 芝田征司. 自然環境の心理学. 環境心理学研究, Vol. 1, No. 1, pp. 38–45, 2013.
- [8] Kaplan. R.and Kaplan. S.and Brown. T. Environmental preference: A comparison of four domains of predictors. *environment and behavior*, 21(5), 509-530.
- [9] 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.
- [10] Min Ge, Chenyi Zhuang, and Qiang Ma. A ranking based approach for robust object discovery from images of mixed classes. pp. 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*, pp. 590–595, 2015.
- [12] 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*, Vol. 59, No. 2, pp. 64–73, 2016.
- [13] 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*, pp. 3111–3119, 2013.