

平成 30 年度 卒業研究論文

題目 観光写真から抽出した
撮影者の好みに基づく
観光スポット推薦に関する研究

指導教員 服部 峻

提出者 室蘭工業大学 情報電子工学系学科

氏 名 中野 広貴

学籍番号 15024124

提出年月日 平成 31 年 2 月 13 日

目次

第 1 章	まえがき	1
第 2 章	関連研究	2
第 3 章	提案手法	3
3.1	観光スポット推薦システム全体の概要	3
3.2	好み抽出フェーズ	4
3.3	観光スポットデータベースの構築	5
3.4	ルート入出力フェーズ	7
第 4 章	k 平均法による観光スポットのクラスタリングにおけるクラスタ数 k に関する実験	9
4.1	画像の特徴量抽出	9
4.2	特徴量の比較と類似度の計算	9
4.3	最適なクラスタ数 k の決定	11
第 5 章	観光写真のクラスタのラベリング	14
5.1	クラスタ内の平均的な画像の検索	14
5.2	クラスタのラベル作成と好みによる検索	14
第 6 章	システムの評価実験	16
6.1	システム評価アンケート調査の実施	16
6.2	評価実験のための各検索手法の実装	16
6.3	システム評価アンケート調査の結果	17
6.4	システムの社会的意義に関する調査	18
第 7 章	今後の課題とまとめ	20
7.1	好み抽出フェーズの改善	20
7.2	観光スポット推薦結果の偏りと改善	21
7.3	潜在的観光スポット発見に用いる写真について	22
7.4	まとめ	24

謝辞 25

参考文献 26

目次

3.1	システムイメージ	3
3.2	好み抽出に使用する写真の例	4
3.3	ルート入出力インターフェイス	7
3.4	ルート入出力インターフェイス (携帯情報端末)	8
4.1	特徴量抽出に使用する写真の例 (夜)	10
4.2	図 4.1 から抽出される特徴量	10
4.3	特徴量抽出に使用する写真の例 (昼)	10
4.4	図 4.3 から抽出される特徴量	10
4.5	観光スポットクラスタの例	11
4.6	クラスタ数 k における室蘭市周辺の全クラスタ内画像類似度 S^{all}	12
4.7	クラスタ数 k における室蘭市周辺から札幌市間の全クラスタ内画像類似度 S^{all}	12
4.8	$k = 100$ の場合に作成されるクラスタの例	13
4.9	$k = 1550$ の場合に作成されるクラスタの例	13
5.1	クラスタラベルの設定と分析の例	15
6.1	写真選択と観光スポット推薦の例	17
6.2	アンケート調査結果	18
6.3	潜在的観光スポットの発見が可能となれば嬉しいか	19
6.4	提案システムを利用したいか	19
7.1	好み抽出フェーズにおける WordNet の上位下位関係の活用	20
7.2	写真の選択枚数と Google Map + 好み情報を用いた観光スポットの推薦評価値の相関関係	21
7.3	写真の選択枚数と 4travel.jp + 好み情報を用いた観光スポットの推薦評価値の相関関係	22
7.4	写真の選択枚数と Flickr + 好み情報を用いた観光スポットの推薦評価値の相関関係	22
7.5	写真の選択枚数と提案 3 手法の評価平均値の相関関係	23

7.6	選択されたタグにおける confidence の平均値	23
-----	---------------------------------------	----

表目次

3.1	図 3.2 の例から抽出されるタグ	4
3.2	ユーザ A の 1 箇所の写真群	5
3.3	ユーザ A の 3 箇所の写真群	5
3.4	ユーザ B の 1 箇所の写真群	5
3.5	ユーザ B の 3 箇所の写真群	5
3.6	各データベースで検索される観光スポットの件数の比較	6

第1章

まえがき

近年、スマートフォンやタブレットなどの情報端末の普及に伴い、ちょっとした旅行時に写真撮影を行い Instagram や Twitter などの SNS に投稿するユーザが増加しているが、同時に、SNS に投稿するために観光地で写真を撮影する機会や枚数も増加しており、個人が撮影する写真は以前と比較して飛躍的に増加していると考えられる。しかし、一部の実際に投稿される写真以外の多くの写真は情報端末に保存されたまま日の目を見ず埋もれていくだけで活用されていない。また、個人の旅行者が未知の場所へ観光に行く際、ガイドブックなどに勧められるがままに観光スポットを巡る場合が多いが、この場合、ガイドブックに掲載されている観光スポットは個々人の「好み」を考慮していないため、旅行者は自身の「好み」から外れた観光スポットを巡ることになってしまう場合も多い。そのため、最終的な観光スポットに対する旅行者の満足度は低くなってしまう。

情報端末に眠る膨大な写真を活用し、それらを観光スポットの推薦に活かすことができれば、新たな旅行機会の創出や、地域活性化に寄与し、観光地に対する満足度向上にも繋がる。また、SNS に投稿された位置情報付きの写真を使用し、世間一般には観光地として認識されていないが、多くの人がそこで写真を撮影している潜在的な観光スポットを発見し推薦することもできると考えられる。

本研究では、ユーザが撮影した写真にはユーザ固有の「好み」が含まれていると仮定し、既存の画像分析技術を用いて写真中の物体の名称をタグとして抽出した上で、それらを分析して抽出したユーザ個々の「好み」タグを用いた新たな手法による観光スポット推薦システムを開発することを目的としている。

第 2 章

関連研究

ユーザ個人々の好みに注目した観光スポット推薦は多く行われている。北村ら [1] は、個人々の撮影した写真に対して一般物体認識を行い、得られたキーワードを利用して写真をグラフ化し旅行の目的を絞り込むユーザインターフェースを提案している。しかし、この研究では最終目標である観光スポット推薦を行う段階まで達しておらず、旅行の目的を絞り込む上でユーザの操作を必要とするため手軽に観光スポットの検索を行うことができるとは言い難い。

また、久保田ら [2] は、SNS に投稿された写真を収集し分析することにより、写真の色から観光スポットの「見ごろ」と「見どころ」を分かりやすく伝える可視化手法を提案している。しかし、この研究ではある特定の観光地の「見ごろ」と「見どころ」をユーザに示すことはできるが、ユーザの知らない観光地を先回りしてユーザに提示し、その魅力を伝えることはできない。

また、青山ら [3] は、SNS に投稿されたジオタグ付き写真を用いて、場所に対する人々の知名度と興味の度合いを算出し、ユーザが移動中に立ち寄ることが可能な知名度の低い寄り道候補を発見し、ランキング形式で提示するシステムを構築している。しかし、このシステムにより提示される寄り道候補は対象のユーザ個人の興味に基づいていないため、必ずしもユーザが満足する寄り道候補が提示されるとは限らない。

これらの研究に対し、本研究はユーザが以前に撮影した観光写真を分析することにより、ユーザの入力を最小限にすると共に、ユーザ自身は知らないが、そのユーザの興味と一致している観光スポットを提示できる。また、同じく SNS に投稿されたジオタグ付き写真を用いるが、広く一般に知られていない潜在的な観光スポットを発見する新しい手法を提案する。

第3章

提案手法

本章では、ユーザが撮影した写真にはユーザ固有の「好み」が含まれていると仮定し、既存の画像分析技術を用いて写真中の物体の名称をタグとして抽出した上で、それらを分析して抽出したユーザ個々の「好み」タグを用いた新たな手法による観光スポット推薦システムについて述べる。

3.1 観光スポット推薦システム全体の概要

本研究における観光スポット推薦システムの処理の流れを図 3.1 に示す。まず、ユーザは自身の観光写真をシステムに入力する。写真は好み抽出フェーズで分析され、ユーザ固有の「好み」タグが抽出される。また、ユーザは次の観光の予定ルートをシステムに入力する。入力されたルートと、好み抽出フェーズで分析された「好み」情報を組み合わせることにより、ユーザに対して「この観光スポットも立ち寄ってみてはいかがでしょうか」という、そのユーザ個人にパーソナライズされた観光スポットの推薦を行う。

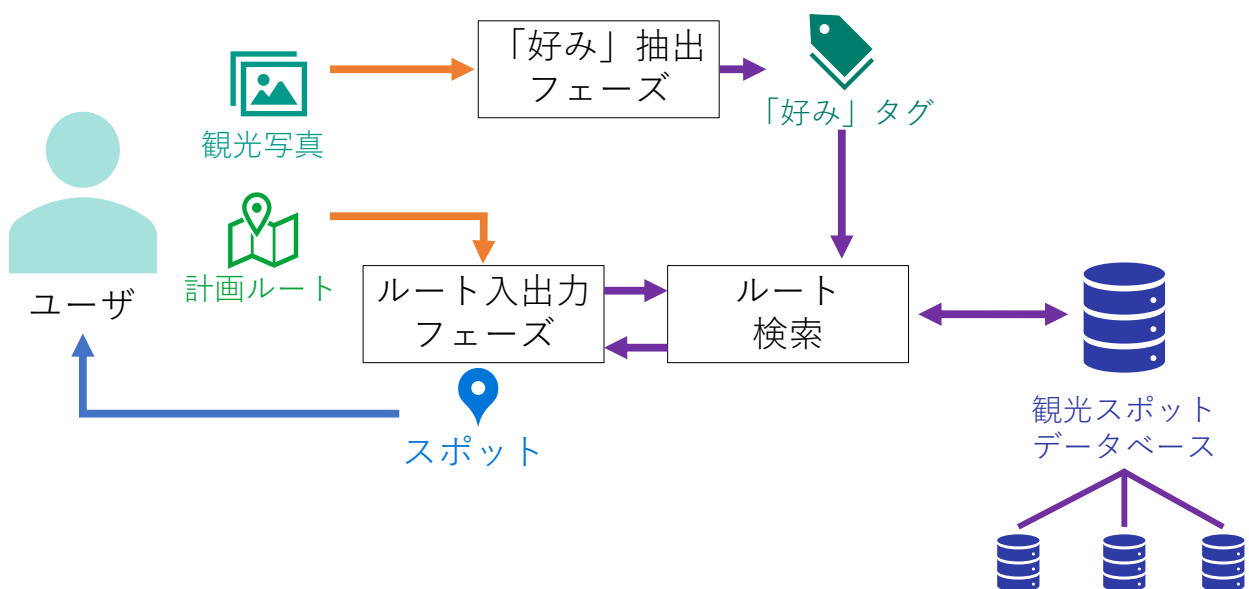


図 3.1 システムイメージ

3.2 好み抽出フェーズ

好み抽出フェーズでは、ユーザから入力された観光写真を画像認識し、ユーザ固有の観光スポットに関する「好み」タグを抽出する。本研究では、既存の画像分析技術として、マイクロソフトのクラウドプラットフォームである Microsoft Azure の Computer Vision API[4] を用いる。この API により、写真に写っている物体の名称とその信頼度スコアがタグとして抽出される。例えば、図 3.2 の写真からは表 3.1 のようなタグを抽出することができる。

表 3.1 の項目の name には入力した写真に写っていると認識された物体名、confidence には各物体名の API による確信度が示されている。複数枚の写真から抽出されたタグ群は 1 つに集約し、confidence の平均値を計算する。本研究では、このタグ群を個人の好みタグセットとして使い、confidence が高いタグから順番に観光スポット推薦に使用する。また、将来的には語の概念辞書である WordNet[5] を用いた上位語検索などの処理を行う。これにより、API で抽出されたタグをそのまま用いるのではなく、タグの概念をより幅広く取り、推薦に活かすことができる。

表 3.2 はユーザ A の函館旅行での観光写真 25 枚から抽出された好みタグの上位 10 件を示している。この表を見ると「電車」「トラム」「トラック」といった特徴的なタグが抽出されていることが分かる。一方、表 3.3 はユーザ A の函館、松前、小樽旅行での観光写真各 25 枚をまとめて抽出した好みタグの上位 10 件を示している。この表を見ると「電車」というタグが取れているが、順位が低く、固有の好みが含まれているようなタグが下位になってしまっている。また、表 3.4 はユーザ B の利尻旅行での観光写真 25 枚から抽出された好みタグの上位 10 件を示している。この表を見ると「猫」「動物」「哺乳類」といった特徴的なタグが抽出されていることが分かる。一方、表 3.5 はユーザ B の利尻、釧路、知床旅行での観光写真各 25 枚をまとめて抽出した好みタグの上位 10 件を示している。この表を見ると、特徴的なタグは順位が低く、固有の好みが含まれているようなタグが下位になってしまっている。この結果から、複数の観光地で撮影された写真をまとめて抽出してしまうと、観光地もしくはユーザ固有の好みタグが取れなくなってしまう危険性が高まることが分かる。



図 3.2 好み抽出に使用する写真の例

表 3.1 図 3.2 の例から抽出されるタグ

name	confidence	name	confidence
table	0.99655	burger	0.31254
food	0.99224	hotdog	0.14531
indoor	0.94887	breakfast	0.13416
sandwich	0.92671	hamburger	0.10358
snack food	0.85508	lunch	0.08399
fries	0.65848		

表 3.2 ユーザ A の 1 箇所の写真群

タグ名	平均 confidence
屋外	0.90908
空	0.55776
電車	0.33536
輸送	0.31030
トラック	0.30355
地面	0.27586
トラム	0.22770
ホワイト	0.22215
道路	0.21858
木	0.11987

表 3.3 ユーザ A の 3 箇所の写真群

タグ名	平均 confidence
屋外	0.74937
空	0.35800
木	0.27902
建物	0.18845
屋内	0.15345
地面	0.14443
道路	0.11224
電車	0.11179
家	0.11037
水	0.11000

表 3.4 ユーザ B の 1 箇所の写真群

タグ名	平均 confidence
屋外	0.70694
猫	0.31896
空	0.27915
山	0.23656
動物	0.22359
哺乳類	0.20679
工場	0.20494
草	0.19599
花	0.17356
自然	0.16539

表 3.5 ユーザ B の 3 箇所の写真群

タグ名	平均 confidence
屋外	0.74810
空	0.30546
草	0.27722
工場	0.21491
自然	0.18200
水	0.17403
動物	0.16644
山	0.15443
木	0.15406
花	0.13029

3.3 観光スポットデータベースの構築

本システムでは、ルート検索において、観光スポット名、位置情報、その他観光スポットの説明テキストから成る観光スポットデータベースを参照する。

本研究では、観光スポットデータベースとして扱う観光スポットのデータセットを複数用意し、どのような観光スポットデータベースが有効かに関する実験を6章で行う。

Google Maps API

旅行者が観光スポットを探したり、旅行ルートを決めたりする際に、一般的に

Google マップが使用されているが、その API である Google Maps API を用いて、抽出された好みタグから観光地の検索を試みる。

観光スポットの口コミの分析

旅行口コミサイト 4travel.jp[6] の口コミを用いて観光地データベースの構築を試みる。データベースの作成は Doc2Vec[7] という機械学習の手法を用いる。Doc2Vec は単語の意味をベクトル表現化する機械学習の手法である Word2Vec[8] を、文や文章といった任意の長さを扱えるように拡張した手法である [9]。

位置情報付きの写真の分析

写真共有コミュニティサイトである Flickr[10] に投稿された位置情報付き写真を用いて、潜在的観光スポットも含む観光スポットデータベース構築を試みる。

Google などの既存の検索サービスを利用する場合、ユーザ自身の知らない観光スポットを検索できる。しかし、得られる結果は検索キーワードに直接関係のあるスポットのみである。一方で口コミから観光スポットのデータベースを構築する場合、データは実際にそのスポットへ訪れたユーザの意見が基となるため、よりそのユーザに合った推薦が可能になると考えられる。また、Doc2Vec を利用して検索を行う場合、キーワードそのものではなくキーワードの意味を用いて検索を行うことができる。つまり、口コミの文章内にキーワードそのものが含まれていなくても、そのキーワードに関係のある観光スポットをユーザに提示できるので、検索キーワードに直接関係の無い観光スポットを得ることができ、Google Maps API の検索サービスと比較して、より「好み」を反映しつつ、ユーザにとっての意外性も考慮した結果となる。

しかし、既存の検索サービスや観光スポットの口コミデータは、既に観光スポットとして広く一般に認識されている場所を示すことはできるが、まだ観光スポットとして認識されていない場所を示すことはできない。そこでさらに、位置情報付き写真を収集し、k 平均法によるクラスタリングを行って観光スポットのデータベースを構築する。この場合、既存のスポット情報を用いないため、一般に観光スポットとして認識されていないが写真が多く撮影されている潜在的観光スポットも推薦できることが期待される。

表 3.6 に室蘭市周辺と札幌市周辺にて各データベースで検索される観光スポットの件数の比較を示す。両地域共に Flickr に投稿された位置情報付き写真を用いた潜在的観光スポットの件数が最も多い。これは、今後新たに観光スポットになり得る潜在的観光スポットの数が、現時点で広く一般に認識されている観光スポットの数よりも多いことを示している。

表 3.6 各データベースで検索される観光スポットの件数の比較

	室蘭市周辺	札幌市周辺
Google Map	97 件	216 件
4travel.jp	55 件	648 件
Flickr	250 件	5750 件

3.4 ルート入出力フェーズ

推薦された観光スポットを表示するためのルート入出力システムを試作した。図 3.3 に試作したシステムのインターフェイスを示す。また、本システムはスマートフォンなどの携帯情報端末で撮影した写真からの観光スポット検索も想定しているため、図 3.4 のように携帯情報端末からシステムへアクセスした場合、その端末に合わせたインターフェイスとなるようレスポンシブデザインを採用している。

ユーザが入力した観光写真から抽出された「好み」情報、次の旅行計画ルートを用いて観光スポットデータベースに問い合わせ、スポット情報を得る。推薦された観光スポット情報は複数表示され、ユーザはそれらを選択することにより、バルーンによってなぜこのスポットが推薦されたのかという情報も確認できる。具体的には「あなたの〇〇という写真から抽出された△△という好みタグに基づいてこの観光スポットが推薦されています」といった情報を確認できる予定である。また、ユーザがこの観光スポットへ行きたいと思えば、それらを経由地として計画ルートに追加することができる。

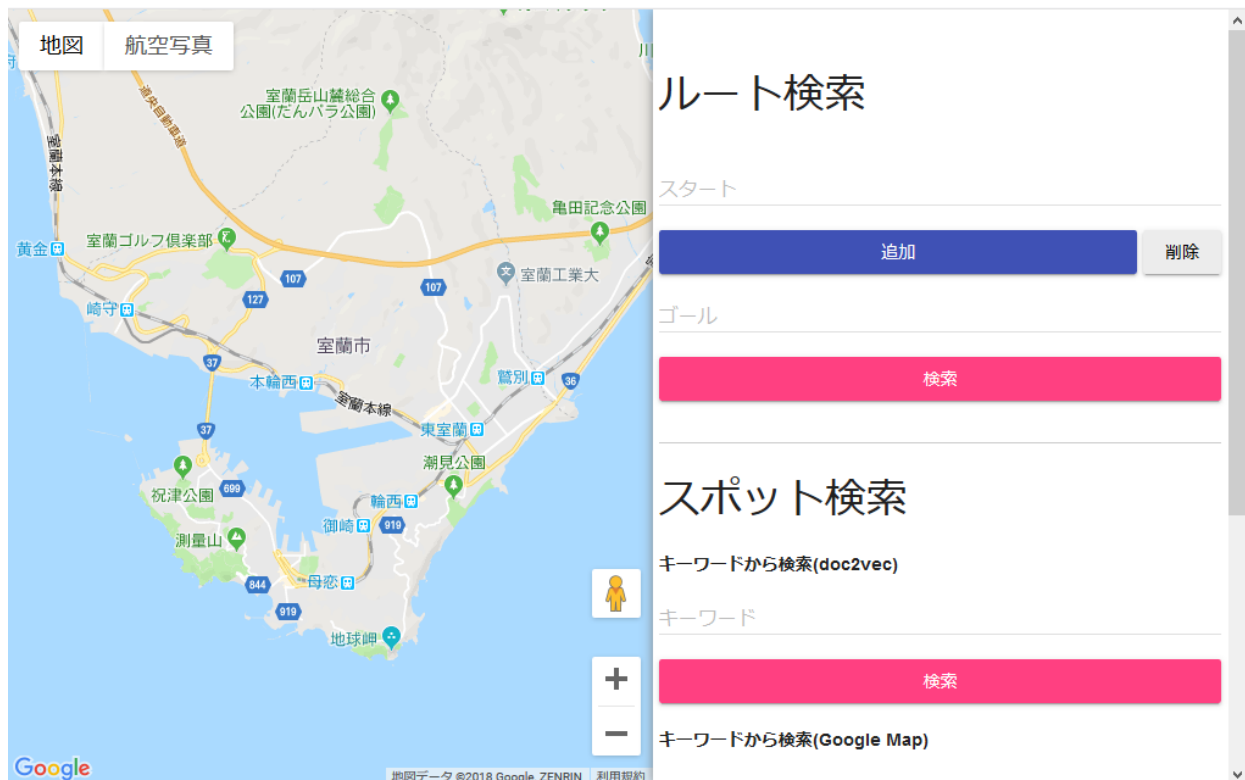


図 3.3 ルート入出力インターフェイス

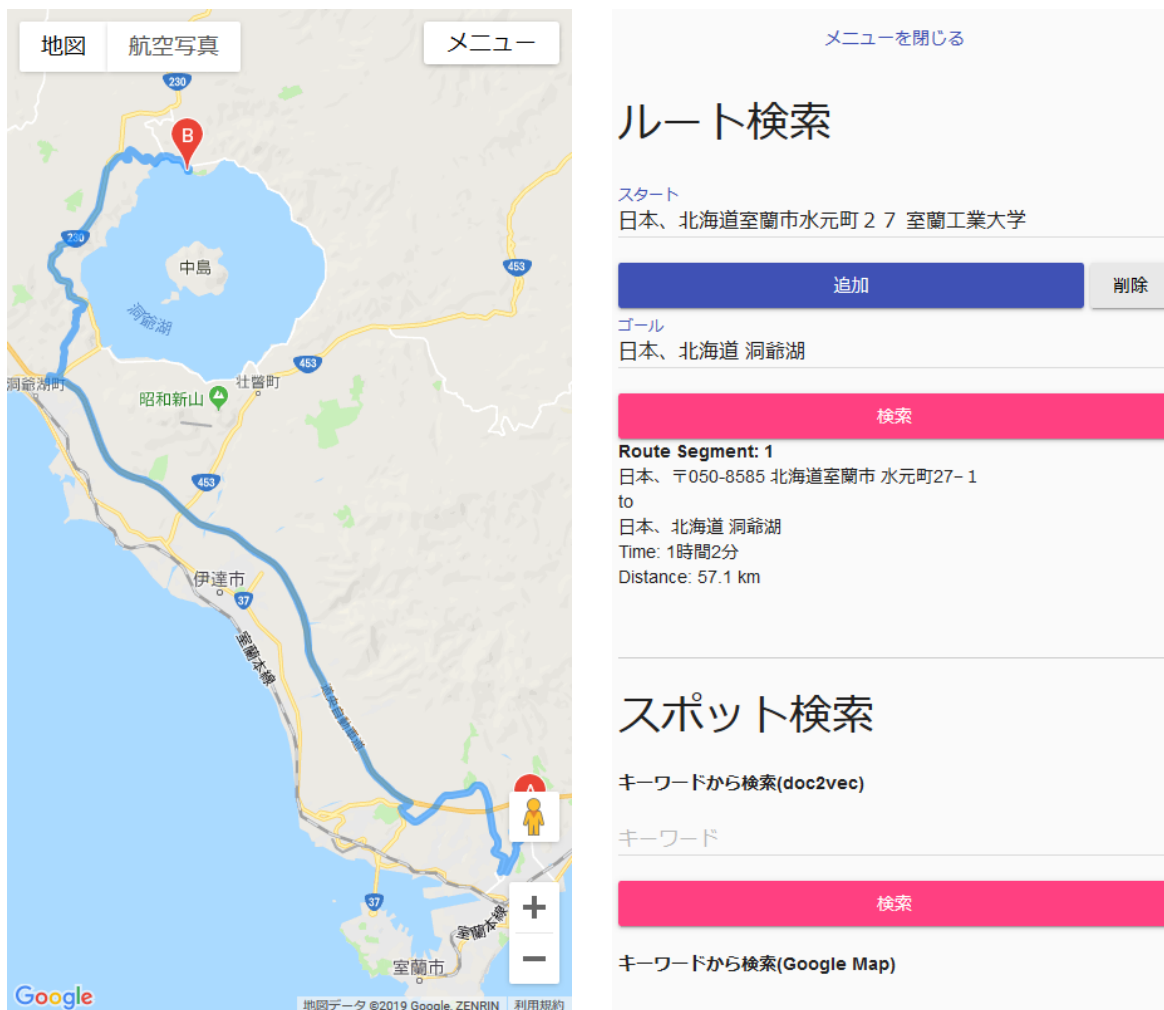


図 3.4 ルート入出カインターフェイス (携帯情報端末)

第 4 章

k 平均法による観光スポットのクラスタリングにおけるクラスタ数 k に関する実験

3.3 節で述べたように，位置情報付きの写真を用いて潜在的観光スポットを含む観光スポットデータベースを構築する．写真に付与された位置情報を基に k 平均法によるクラスタリングを行い，膨大な枚数の写真から位置情報が近い写真をクラスタとして取り出す．これにより各々のクラスタの中心を写真が撮影されたスポットの中心と考えることができる．また，将来的には位置情報だけではなく，内容も考慮したクラスタリングについても検討していく．しかし， k 平均法によるクラスタリングはクラスタ数 k を与える必要がある．本章では最適な k を求めるため，各クラスタ内に含まれる写真の特徴を分析する実験を行う．

4.1 画像の特徴量抽出

クラスタ内の画像の特徴量を比較するために，画像の特徴量を抽出する．本研究では，ImageNet[11] で事前学習した VGG16 を使用する．VGG16 は全 16 層の畳み込みニューラルネットワークであり，画像に対して 1000 クラスの確率推定を出力する [12]．また，VGG16 では出力層の直前に画像の特徴量を持つ 4096 次元のベクトルが計算されており，画像間の類似度を求めるのに用いる．例えば，図 4.1 の写真からは図 4.2 のような特徴量が抽出され，図 4.3 の写真からは図 4.4 のような特徴量が抽出される．これら二つの写真は同じ観光スポットを異なる時間に異なる角度で撮影したものであるが，各特徴量のコサイン類似度を計算すると 0.7303099 となり高い値であることが分かる．

4.2 特徴量の比較と類似度の計算

各クラスタ内の均一性を評価し，最適な k を求めるため，各々の画像特徴量のコサイン類似度を計算する．本研究では同じクラスタに属する画像の全ての組み合わせで類似度を計算し，



図 4.1 特徴量抽出に使用する写真の例 (夜)

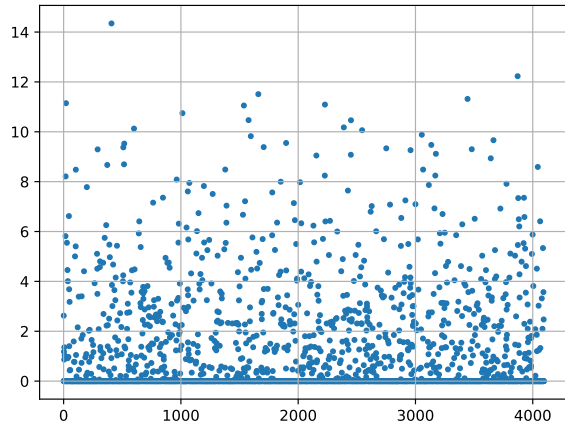


図 4.2 図 4.1 から抽出される特徴量



図 4.3 特徴量抽出に使用する写真の例 (昼)

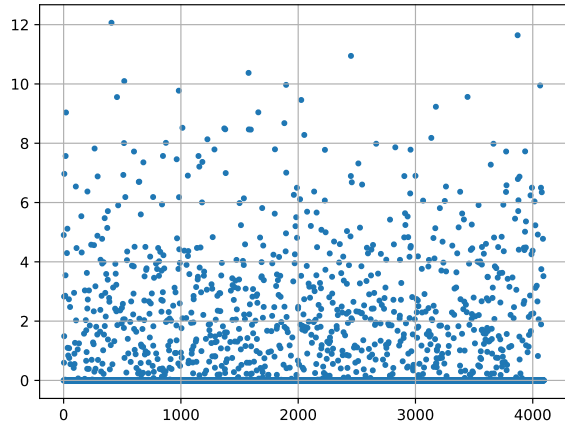


図 4.4 図 4.3 から抽出される特徴量

その平均値をクラスタ内画像類似度 $S_l^{cluster}$ ($l = 1, 2, \dots, M$) とする. ここで M は設定したクラスタ数 k を表している. また, あるクラスタ数 k における全クラスタ内画像類似度の平均値をクラスタ数 k における全クラスタ内画像類似度 S^{all} とし, 各クラスタ内に含まれる写真の枚数を N_l ($l = 1, 2, \dots, M$), 対象とする写真 p_i, p_j ($i, j = 1, 2, \dots, N_l$) の画像特徴量のコサイン類似度を $S(p_i, p_j)$ とした場合,

$$S_l^{cluster} = \begin{cases} \frac{\sum_{1 \leq i < j}^{N_l} S(p_i, p_j)}{N_l C_2} & (N_l \geq 2) \\ 1 & (N_l = 1) \end{cases}$$

$$S^{all} = \frac{\sum_{l=1}^M S_l^{cluster}}{M}$$

となり、クラスタ数 k における全クラスタ内画像類似度 S^{all} が高ければ高いほどクラスタ内に類似した画像が含まれている可能性が高いことを示しており、その k が適当であることを示している。図 4.6 に室蘭市周辺の写真 2509 枚に対するクラスタ数 k における全クラスタ内画像類似度 S^{all} の変化のグラフを示す。また、図 4.7 に室蘭市周辺から札幌市間の写真 15356 枚に対するクラスタ数 k における全クラスタ内画像類似度 S^{all} の変化のグラフを示す。

4.3 最適なクラスタ数 k の決定

前節のクラスタ数 k に関する実験の結果から、室蘭市周辺の画像 2509 枚に対して最適な k を考察する。図 4.6 のグラフからは k が 250 を超えた辺りから類似度 S^{all} の上昇が緩やかになっていることが読み取れる。また、 k を増加させると類似度 S^{all} も増加するが、限られた領域内で写真を細かくクラスタリングし過ぎると本来1つとなるべきスポットが別のスポットとして認識されてしまうため、できる限り k は小さい方が望ましい。よって、この場合の最適な k は 250 程度である。

また、以上の結果から他地域に対しても自動的に k を決定できるようにする。室蘭市周辺の画像 2509 枚に対して $k = 250$ でクラスタリングを行う場合、1 クラスタ当たり平均して 10.036 枚の写真が含まれることになる。この値が最適であると仮定した場合、室蘭市周辺から札幌市間の写真 15356 枚の最適なクラスタ数 k は 1530 であると言える。図 4.7 のグラフから k が 1500 を超えた辺りから類似度 S^{all} の上昇が緩やかになっていることが読み取れるため、前述の仮定はほぼ正しいと考えられる。図 4.8 を見ると、 $k = 100$ の場合、別の場所で撮影された写真が同一クラスタに含まれてしまったが、図 4.9 を見ると、 $k = 1550$ の場合、同じ場所で撮影された写真が同一クラスタに含まれていることを確認できる。また、図 4.5 に今回作成されたクラスタの例を示す。

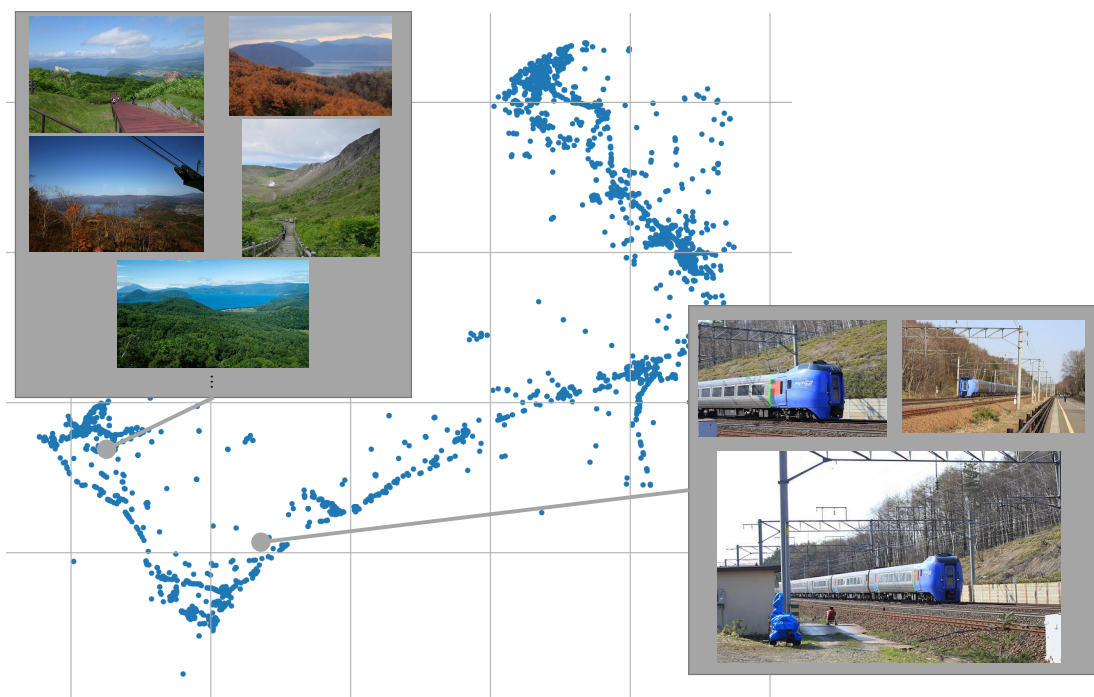


図 4.5 観光スポットクラスタの例

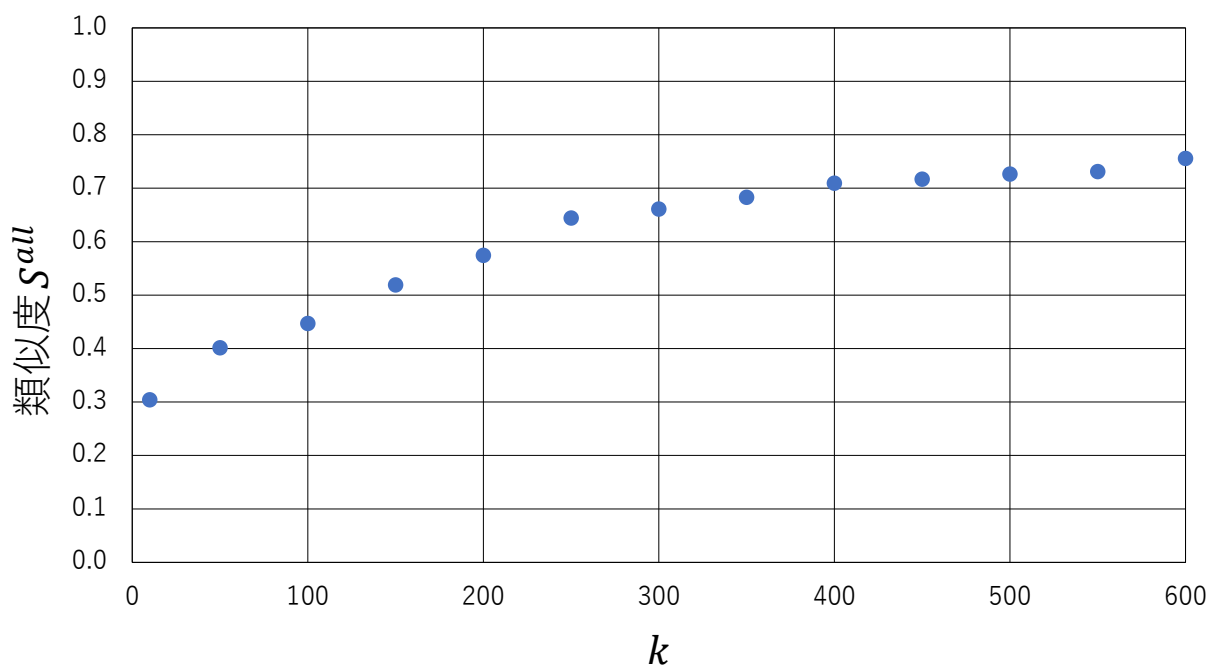


図 4.6 クラスタ数 k における室蘭市周辺の全クラスタ内画像類似度 S^{all}

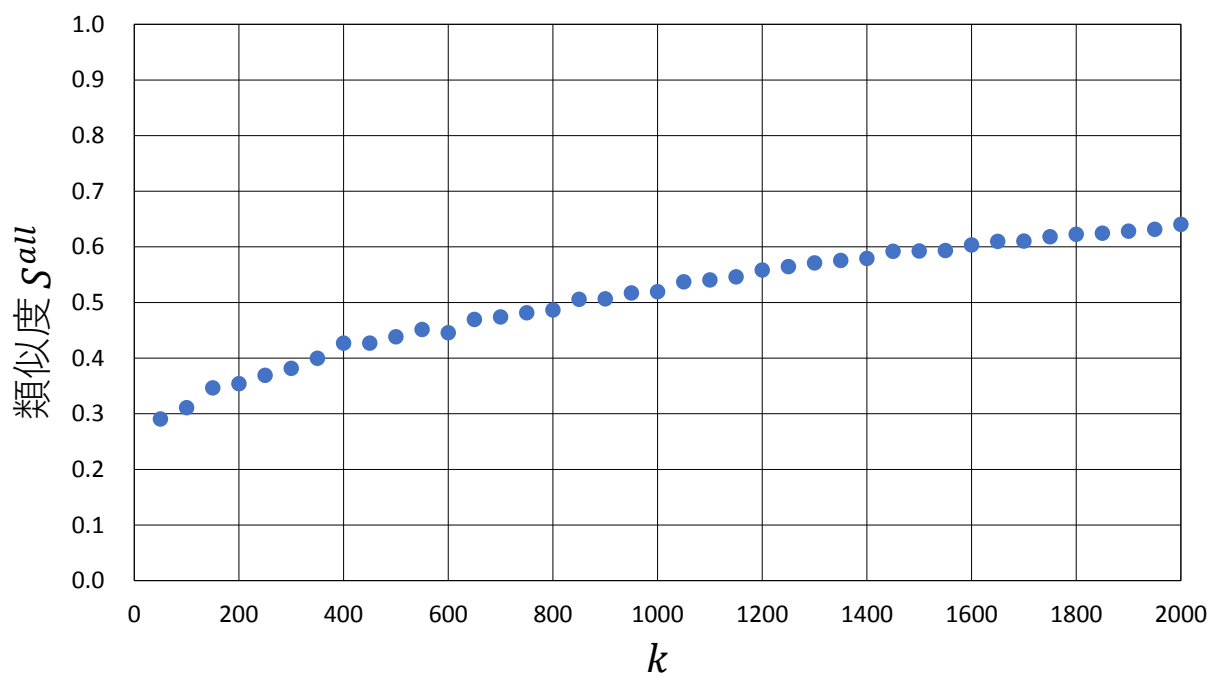


図 4.7 クラスタ数 k における室蘭市周辺から札幌市間の全クラスタ内画像類似度 S^{all}

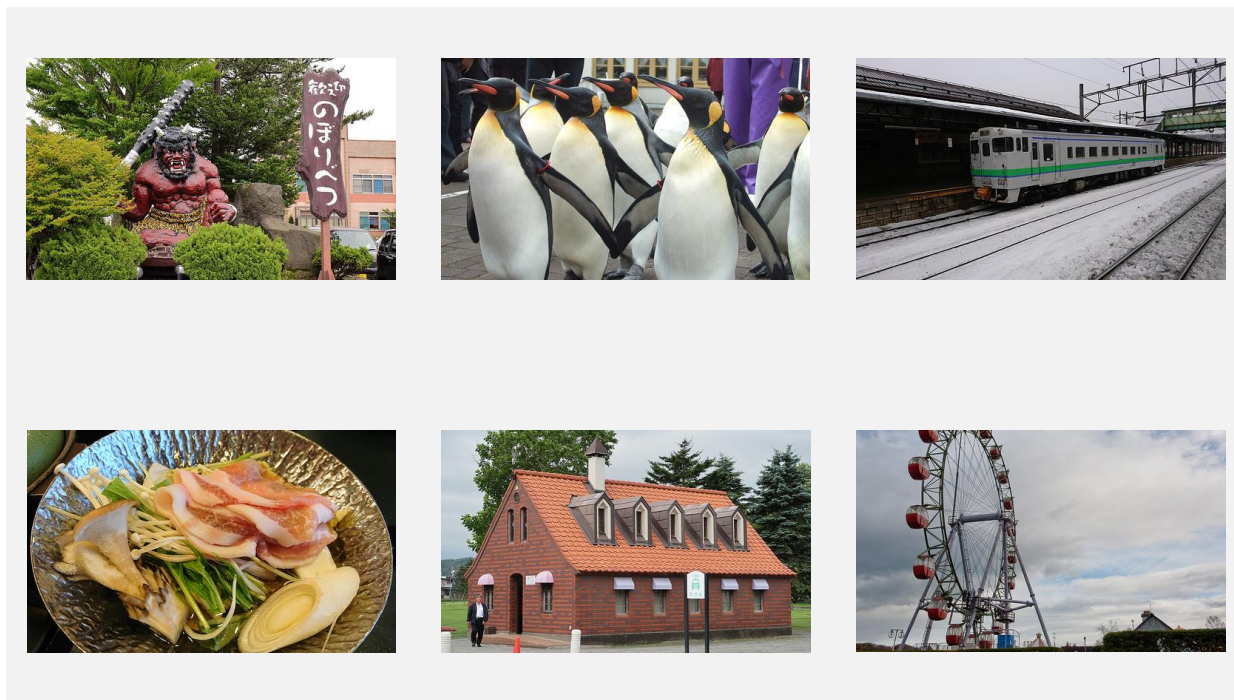


図 4.8 $k = 100$ の場合に作成されるクラスタの例

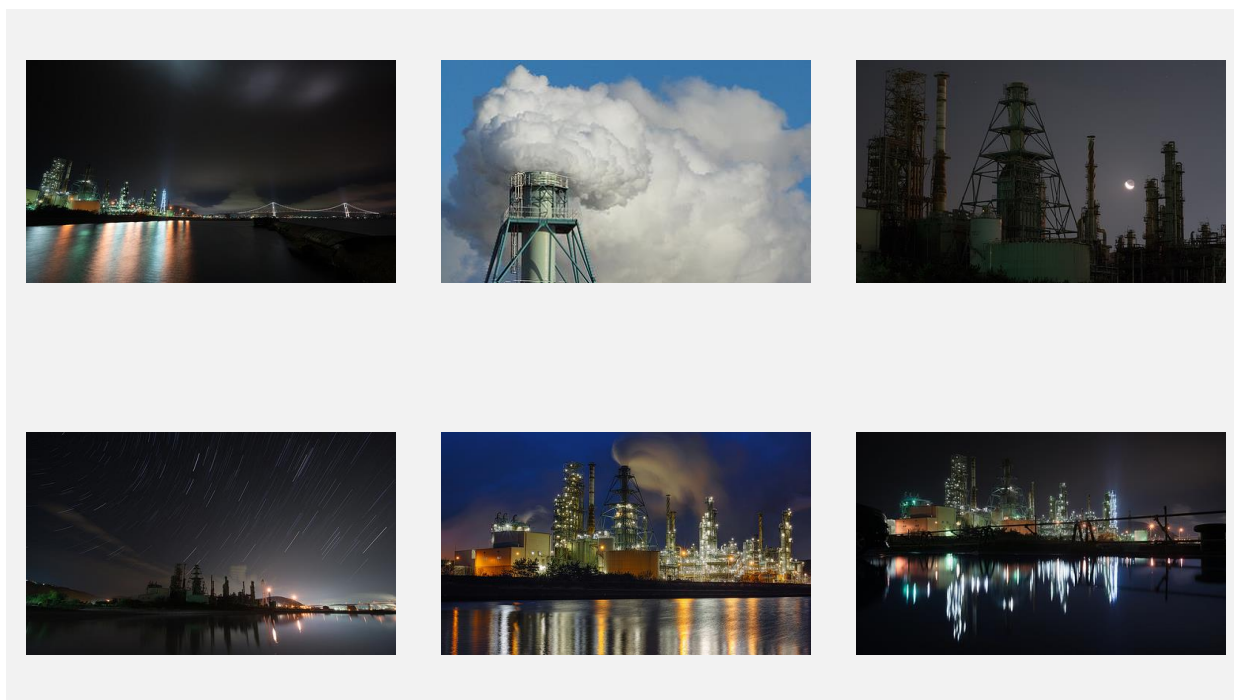


図 4.9 $k = 1550$ の場合に作成されるクラスタの例

第 5 章

観光写真のクラスタのラベリング

前章の実験により，最適と考えられる k の値による観光スポット写真のクラスタリングが完了した．これらのクラスタを観光スポット推薦に活用するため，各々のクラスタに対してラベリングを行う．

5.1 クラスタ内の平均的な画像の検索

クラスタのラベルを作成するため，各クラスタに属する写真からそれぞれ最も典型的な写真として本研究では平均的な写真を選択し，その写真の分析結果をそれぞれのクラスタのラベルとして用いる．ラベルの作成に最も平均的な写真を用いることにより，その地点でいつも撮影できるとは限らない珍しい写真がそのクラスタのラベルとなることを防ぎ，よりその地点で撮影できる一般的な写真をクラスタのラベルとして設定できる．平均的な写真の計算には 4.1 節で用いた VGG16 を使用する．各クラスタ内に属する写真全てを VGG16 で特徴量抽出し，その平均値をクラスタ内の平均特徴量とする．その後，図 5.1 のようにクラスタ内の平均特徴量と最も近い特徴量を持った写真をクラスタのラベル作成に用いる写真として設定する．また，クラスタ内の写真が 1 枚の場合はその写真を，2 枚の場合はランダムで 1 枚を選択しラベル作成に用いる写真として設定する．

5.2 クラスタのラベル作成と好みによる検索

各クラスタのラベルの作成には 3.2 節で用いた Computer Vision API を用いる．この API でクラスタのラベル作成に用いる写真として設定された平均的な写真を分析し，抽出されたタグをそれぞれのクラスタのラベルとして設定する．図 5.1 にクラスタのラベルとして選択された写真の分析の具体例を示す．ここで作成されたクラスタのラベルの confidence ベクトルと，3.2 節で抽出された個人の好みタグの平均 confidence ベクトルを比較し，コサイン類似度が高ければ高いほどそのクラスタに個人の好みと近い写真が含まれていると考えられる．

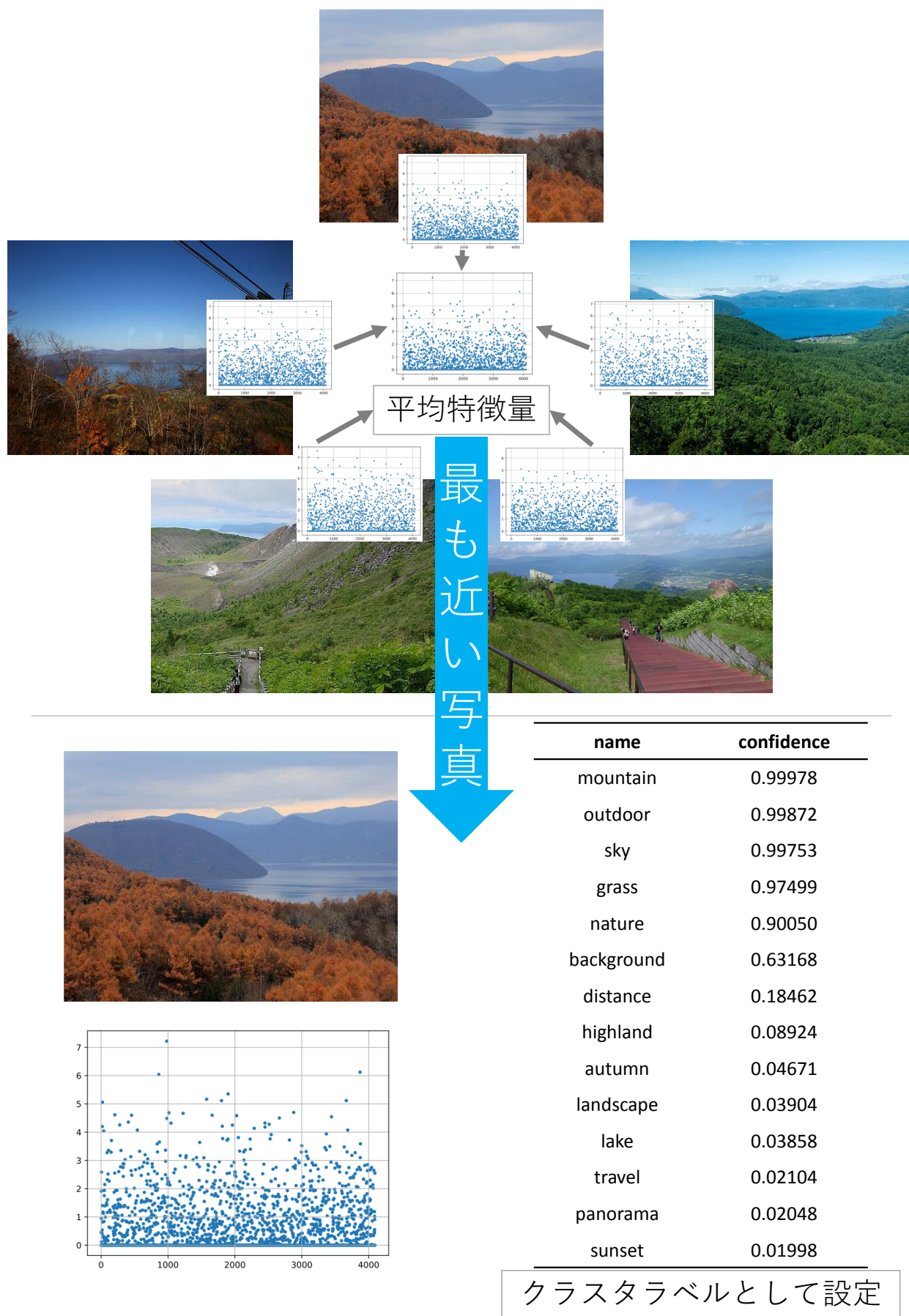


図 5.1 クラスタラベルの設定と分析の例

第 6 章

システムの評価実験

本章では，システムの評価実験として既存の観光スポット推薦手法と比較し，どの程度ユーザの「好み」を反映した観光スポット推薦が行えているか検証を行う．また同時に，潜在的観光スポットと提案システムに関するニーズの調査を行い，本研究の社会的意義を明らかにした．

6.1 システム評価アンケート調査の実施

まず，被験者となる男女 30 名に対して事前調査として観光写真 52 枚を提示し，その中から，もし自分が観光スポットで見かけたら撮影すると思われる写真を複数枚選択してもらい，選択された写真から被験者毎に好み抽出を行う．その後，本調査として北海道の室蘭市周辺から札幌市間を観光ルートとして設定し，その周辺で一般的な手法で検索した観光スポットと，事前調査で抽出された被験者毎の好み情報を基に発見された観光スポット 3 種類の両方（4×5 の計 20 箇所）をランダムに提示する．各被験者はそれらのスポットへ行きたいか，行きたくないかを 5 段階（0 から 4）で評価する．

6.2 評価実験のための各検索手法の実装

一般的な観光スポット検索手法としては Google マップを用いる．Google マップで一般的に観光スポット検索に利用されると思われる『観光地』『観光スポット』『観光』『スポット』の 4 つのキーワードを用いて OR 検索を行う．

好み情報を用いた観光スポット検索としては，一般的な観光スポット検索で用いた手法と，3.2 節で抽出された好み情報とを組み合わせる手法を用いる．好み情報を用いることにより，一般的な観光スポット検索と比較してより個々人の好みを反映した観光スポット推薦が期待できる．例えば，「建物」というタグが好み情報として抽出された場合「建物 観光地」というワードを用いて Google マップで AND 検索を行う．ここでは 5 箇所の観光スポットを検索するため，好みタグの平均 confidence 上位 5 件までを用いて検索を行う．

ロコミデータを用いた観光スポット検索手法では，3.3 節で構築された Doc2Vec のデータ

ベースから抽出された好み情報を用いて観光スポットの検索を行うことにより、他人の口コミデータによる観光スポットの特性を考慮した観光スポット推薦が期待できる。ここでは5箇所の観光スポットを検索するため、好みタグの平均 confidence 上位5件までを用いて検索を行う。

位置情報付き写真を用いた観光スポット検索手法では、構築した Flickr に投稿された位置情報付き写真を用いたデータベースから抽出された好み情報を用いて観光スポットの検索を行う。この手法は上記の手法とは異なり、既存の観光スポット情報を用いないため潜在的観光スポットの発見と推薦が期待できる。具体的には、3.2節で作成した個人の好みタグセットと5章で作成した各クラスタのラベルに含まれる各タグの confidence のコサイン類似度を計算し、値が高ければ高いほど個人の好みと近い写真が含まれているクラスタ（観光スポット）として推薦する。図6.1にあるユーザの写真の選択例及びその写真から推薦される観光スポットの例を示す。



図 6.1 写真選択と観光スポット推薦の例

6.3 システム評価アンケート調査の結果

以上の結果を集計し、各手法による観光スポット推薦結果に対するユーザ評価の平均値を図6.2に示す。Google マップで単純にキーワード検索を行う手法と比較して、本研究における提案手法の3つ何れの結果も有意差が見られた。特に、4travel.jp の口コミを用いた観光スポット推薦は最も良い結果が出ている。これは、4travel.jp に掲載されている情報は元々ある程度一般に認知されている観光スポットであり、観光スポットとしての最低限のクオリティが保証されている為であると考えられる。一方で、Flickr から収集された位置情報付き写真を用いた

潜在的観光スポットを含む観光スポット推薦は他の提案手法と比較して良い結果が出ているとは言い難い。これは、観光スポット検索に用いる個人の好みタグセットと Flickr の位置情報付き写真の両方に問題があると考えられる。詳細は7章で後述する。

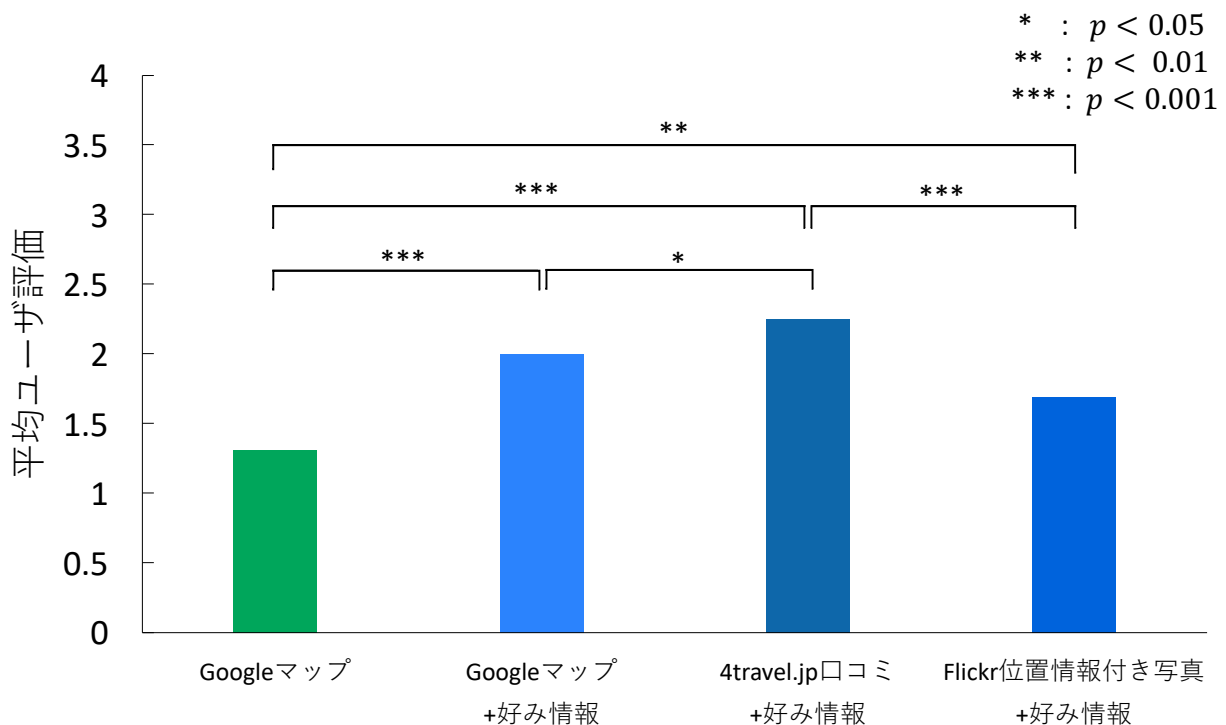


図 6.2 アンケート調査結果

6.4 システムの社会的意義に関する調査

6.1 節のシステム評価アンケートに加えて、被験者に対して潜在的観光スポットの発見が可能となれば嬉しいか、また、提案システムを利用したいか調査を行った。

図 6.3 に『まだ多くの人には認知されていない、潜在的な観光スポットをシステムが推薦してくれたらどう思いますか?』という質問への回答の集計結果を示す。潜在的な観光スポットの発見に対して肯定的な回答を行った被験者が 83% と、否定的な回答を行った被験者 7% に比べて圧倒的に多いことが分かる。このことから、多くの人知らない潜在的観光スポットを発見することができれば嬉しいと考えていることが分かり、社会的な意義も大きいと言える。

図 6.4 に『自分が過去にスマホなどで撮影した写真だけで、自分の好みに合った観光地を自動的に推薦してくれるシステムがあれば使いたいですか?』という質問への回答の集計結果を示す。こちらも従来と比較して少ない労力で自分の好みに合致した観光スポットを提案してくれるシステムに対して肯定的な回答を行った被験者が 70% と、否定的な回答を行った 17% と比較して多いことが分かる。この結果は、従来の個々人の好みを見逃しもしくはユーザに多大な労力を要求する観光スポット検索システムに対する不満を示しており、提案システムに対するニーズも大きいと考えられる。

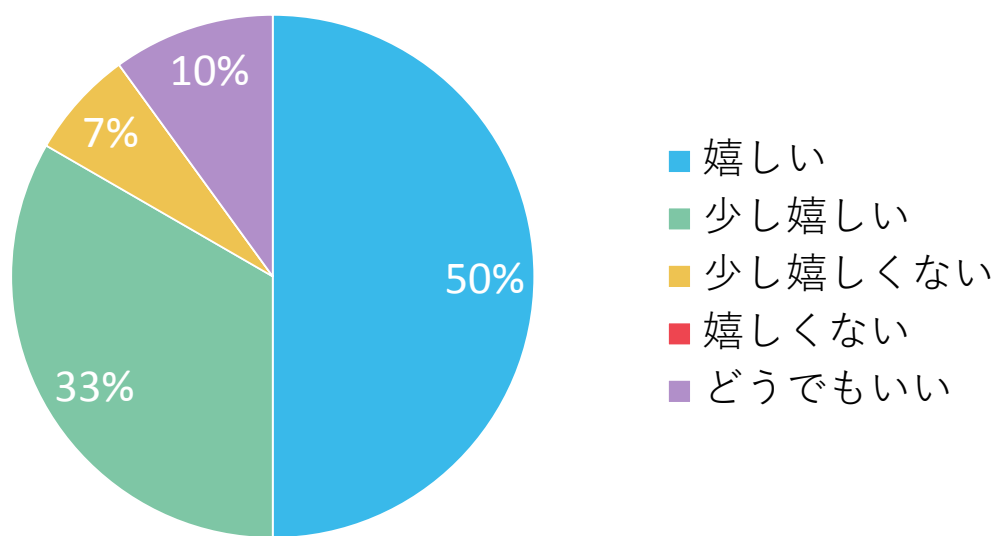


図 6.3 潜在的観光スポットの発見が可能となれば嬉しいか

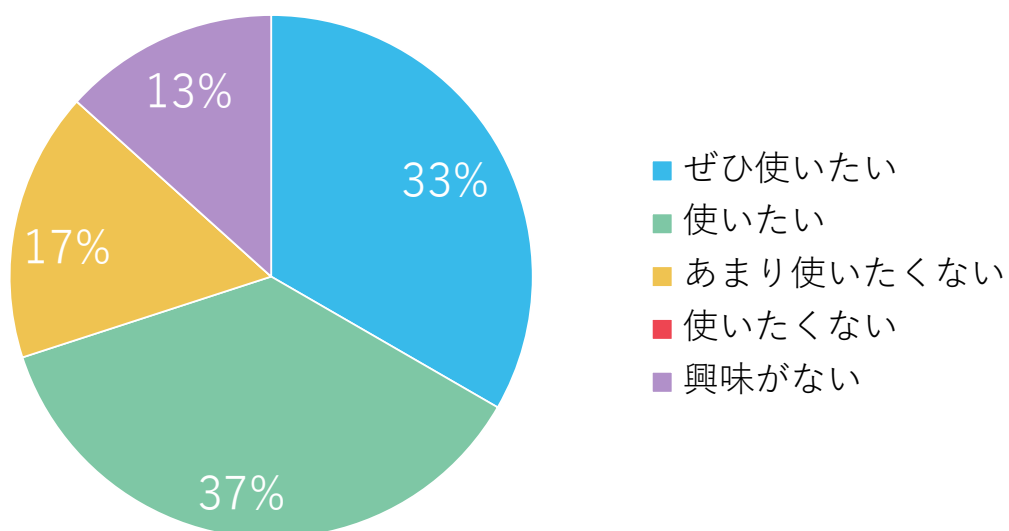


図 6.4 提案システムを利用したいか

第7章

今後の課題とまとめ

本研究では，複数種類の観光スポットデータベースを構築し，それらを用いた新たな観光スポット検索手法を提案した．その結果，いくつかの研究課題が判明した．以降は，提案システムが未だ抱える課題を整理し，それぞれの改善方法について検討する．

7.1 好み抽出フェーズの改善

3.2 節でも述べたが，現段階で完全な好み抽出が出来ているとは言い難い．本研究では写真からの好み抽出を全て Computer Vision API によって行っており，誤検出や関係のないタグの抽出なども起こってしまう可能性がある．また，抽出されるタグが細かくなり過ぎてしまうという問題もある．例えば，あるユーザの写真から sandwich というタグが抽出された場合を考える．もしユーザが sandwich の写真ばかりを撮影しているのであれば問題は無いが，たまたまこの場所で sandwich を撮影しただけで普段は様々な食べ物を撮影しているのかもしれない．この問題に対して，語の概念辞書である WordNet を用いると snack food や dish といった上位のタグを得ることができる．これにより sandwich ばかりではなく potato chip や sukiyaki を撮影するようなユーザに対して，検出された狭い意味の単語に固執せず，より幅広い推薦が可能になる．

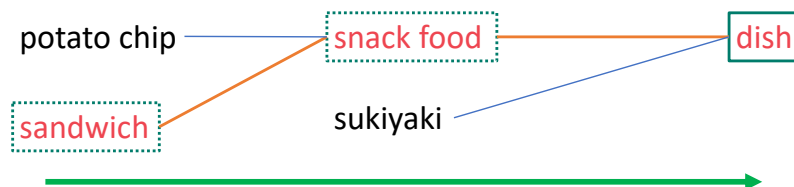


図 7.1 好み抽出フェーズにおける WordNet の上位下位関係の活用

7.2 観光スポット推薦結果の偏りと改善

本研究の実験では被験者に写真を好きなだけ選択させ、それらを分析し個人の好みとして用いている。このことから、3.2節では、ユーザが多くの写真を選択すると一つ一つの写真の特徴が薄まり、ユーザに関係なく似たような推薦結果になってしまうのではないかとこの予想を立てた。そこで、被験者毎の写真の選択枚数と好み情報を用いた観光スポット推薦結果の評価値にどのような関係があるのかという調査を行った。図 7.2 に Google Map と好み情報、図 7.3 に 4travel.jp の口コミと好み情報、図 7.4 に Flickr の写真と好み情報による潜在的観光スポット検索、図 7.5 に提案した 3 手法の平均値による調査結果を示す。これらの調査結果より、予想に反して被験者毎の写真の選択枚数と好み情報を用いた観光スポットの推薦結果とそうでない推薦結果における平均評価値の差に相関関係は無いということが分かった。

しかし一方で、前章のアンケートで抽出された好みタグをランダムで被験者に提示し、『これらの単語はあなたの興味のある単語ですか？興味のある単語にチェックを入れてください。』という質問を行ったところ、興味深い結果が見られた。図 7.6 に各被験者が選択したタグにおける confidence の平均値を示す。この結果から被験者が興味を持ったタグの confidence の平均値は最小値が約 0.007、最大値が約 0.268、平均値が約 0.075 と非常に小さいことが分かる。本研究では推薦に用いるタグを confidence が高い順から選択していたが、おそらく、あるユーザにおける confidence ベクトルに対して、他ユーザ群の confidence の方が高いタグは一般タグとして弱める TF-IDF のようなランキングが必要であると考えられる。

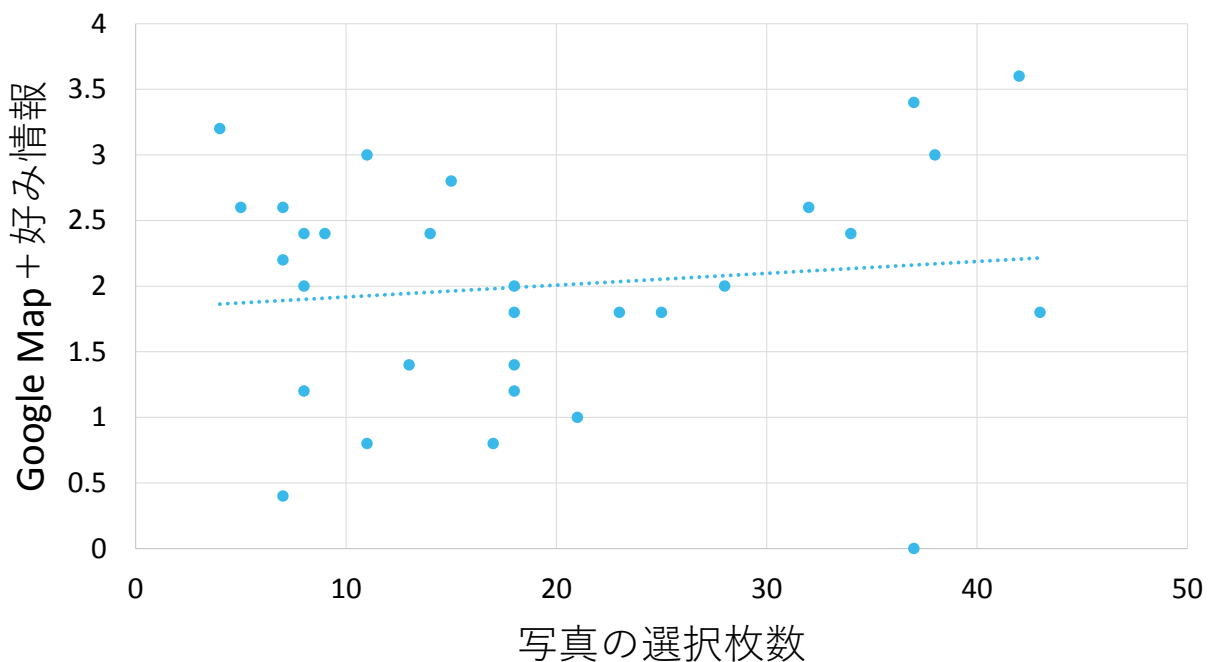


図 7.2 写真の選択枚数と Google Map + 好み情報を用いた観光スポットの推薦評価値の相関関係

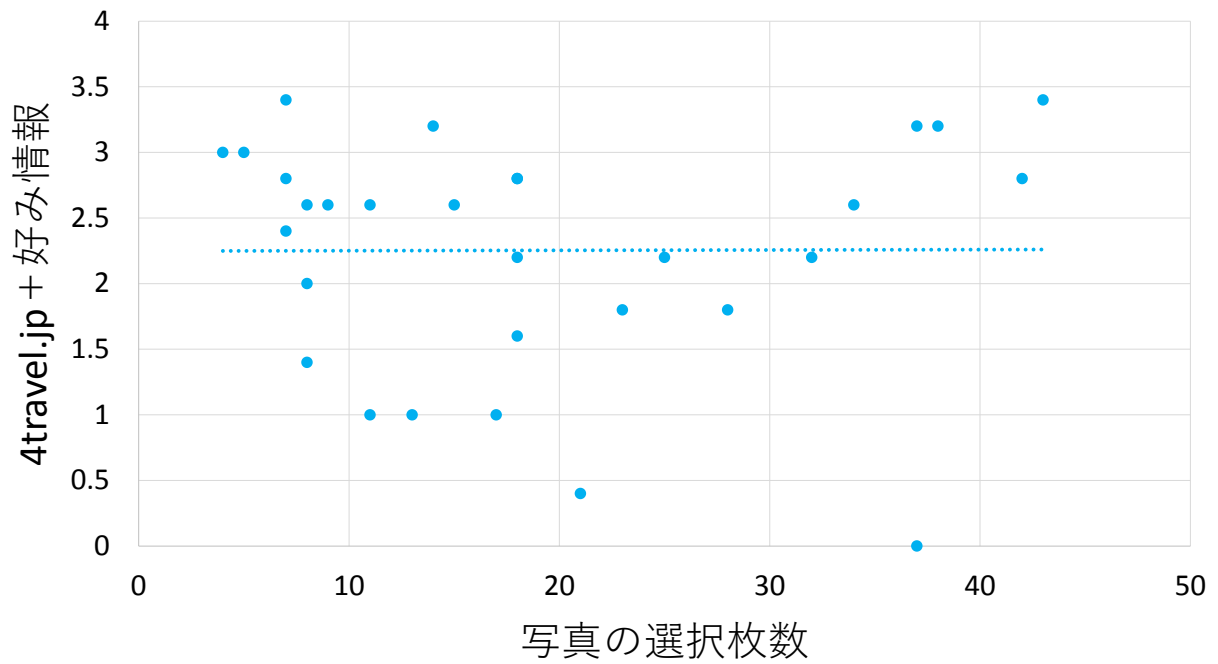


図 7.3 写真の選択枚数と 4travel.jp + 好み情報を用いた観光スポットの推薦評価値の相関関係

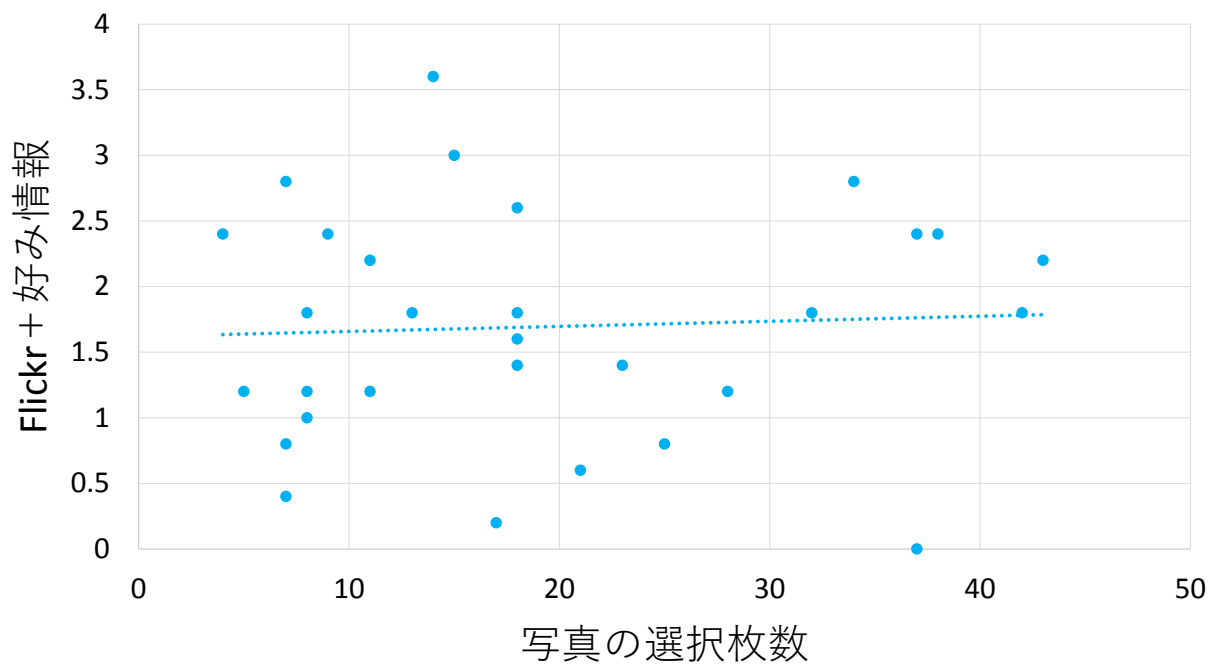


図 7.4 写真の選択枚数と Flickr + 好み情報を用いた観光スポットの推薦評価値の相関関係

7.3 潜在的観光スポット発見に用いる写真について

実験の段階で Flickr にアップロードされている位置情報付きの写真は観光目的で撮影されているものが全てであると仮定していた。しかし、実際にこれらの写真を分析してみると実験

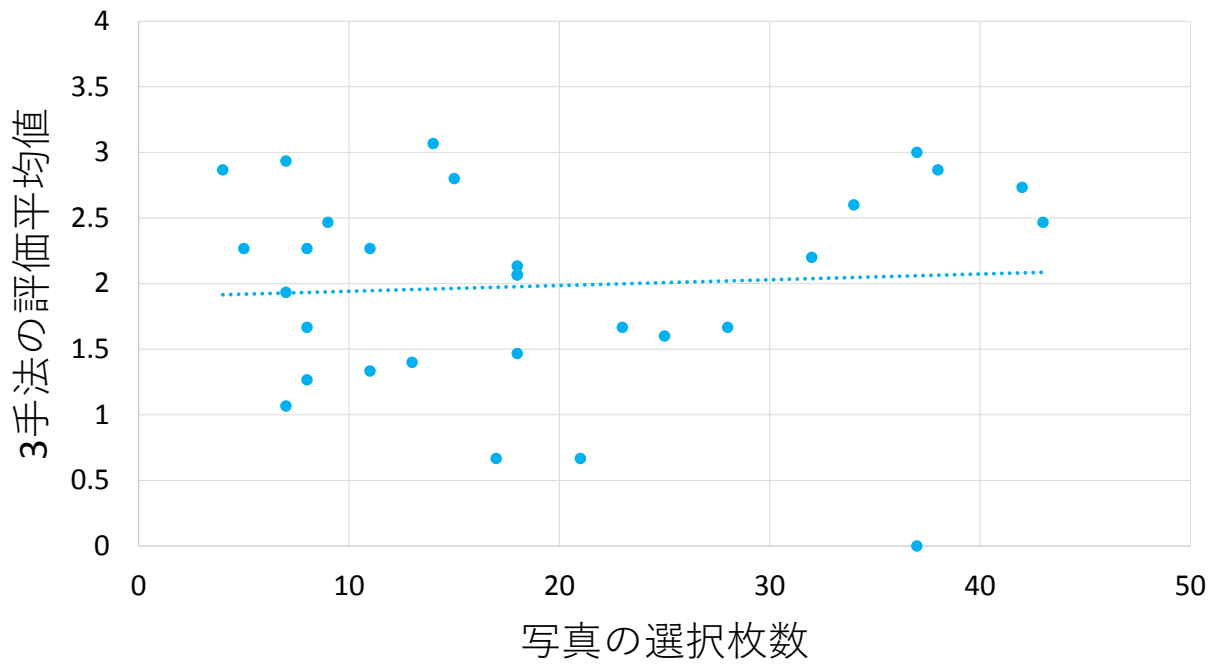


図 7.5 写真の選択枚数と提案 3 手法の評価平均値の相関関係

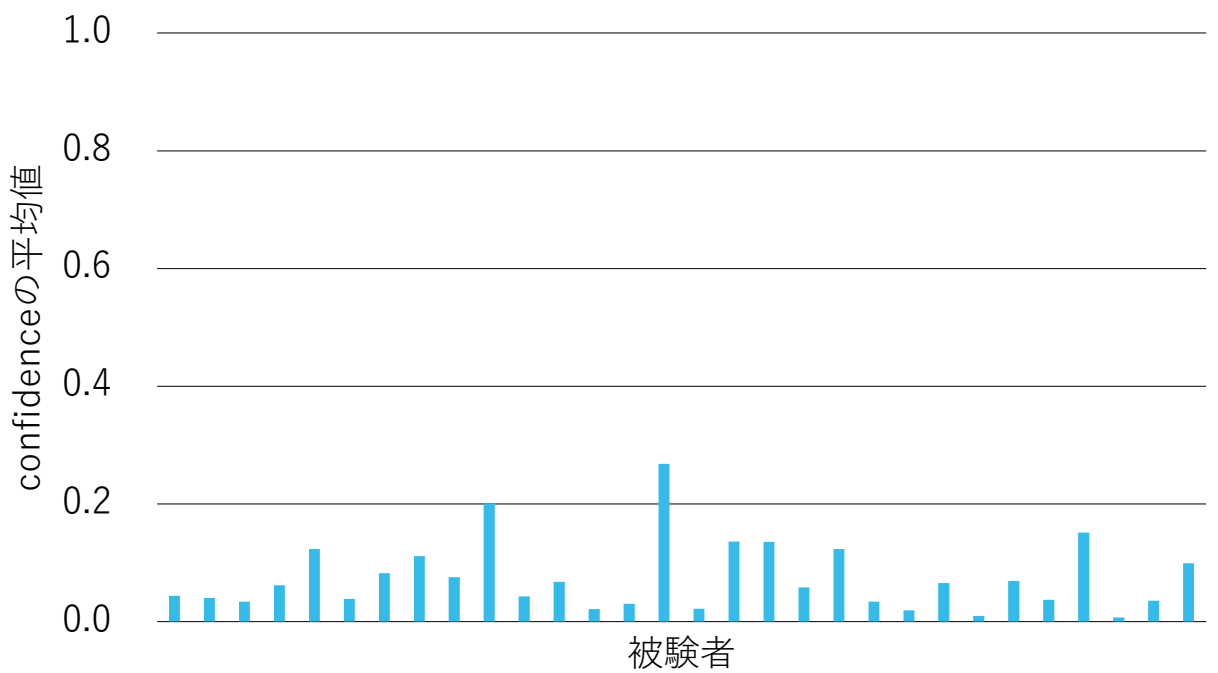


図 7.6 選択されたタグにおける confidence の平均値

には適さないスマートフォンのスクリーンショットや個人の家の中の写真などが稀に含まれていることが分かった。これらの写真は潜在的な観光スポットの発見の際にノイズとして悪影響を及ぼすため、事前に除去されていることが望ましい。今後は観光写真とは言えない写真を取り除くシステム、もしくは位置情報付き観光写真収集に関する新たな手法を検討する必要がある。

7.4 まとめ

本研究により、ユーザが以前に撮影した観光写真を分析することによる、ユーザ自身は知らないが、そのユーザの興味と一致している観光スポットを提示できるようになった。また、SNS に投稿されたジオタグ付き写真を用いた、広く一般に知られていない潜在的な観光スポットを発見する新しい手法を提案した。これらは前述のように、未だ多くの問題を内包しているが、経済社会の発展のために観光立国の実現を目指す我が国の施策にも大きく寄与する可能性を6.3節の結果からも読み取れる。本研究における観光スポット推薦は既存のシステムとは異なり、ユーザによるキーワードの入力を必要としない点が最も特徴的である。これにより、日本語を理解できない外国人に対してもその人それぞれの好みに合った観光スポットを自動的に推薦でき、従来必要であった翻訳などのコストを大幅に削減することができる。また、観光写真を用いた潜在的観光スポット推薦は、現状行政には見逃されていた観光スポットを広く認識させることが可能で、過疎化と人口減少が進行する地方市町村に眠った隠れた観光スポットにスポットライトを当てることができる。

謝辞

本研究に際して，様々なご指導を頂きました服部峻助教に厚く御礼申し上げます。また，日常の議論を通じて多くの知識や示唆を頂いた服部研究室の皆様にも深く感謝の意を表します。

参考文献

- [1] 北村 理紗, 伊藤 貴之, “個人の旅行写真の一般物体認識に基づく観光地推薦,” DEIM Forum 2018, H1-4 (2018).
- [2] 久保田 麻美, 牛尾 剛聡, “SNS による文化と風土の可視化 – Instagram から地域の「見どころ」と「見ごろ」を発見する –,” DEIM Forum 2015, G7-1 (2015).
- [3] 青山 賢, 廣田 雅春, 石川 博, 横山 昌平, “ジオタグ付き写真を用いた知名度が低いにもかかわらず興味の度合いが高い寄り道候補の発見,” DEIM Forum 2015, B5-3 (2015).
- [4] Computer Vision API を利用した画像処理 Microsoft Azure, <<https://azure.microsoft.com/ja-jp/services/cognitive-services/computer-vision/>>.
- [5] 国立研究開発法人情報通信研究機構 (2012) 「日本語 WordNet」, <<http://compling.hss.ntu.edu.sg/wnja/>>.
- [6] 4travel.jp, <<http://4travel.jp/>>
- [7] Jey Han Lau, Timothy Baldwin, “An Empirical Evaluation of doc2vec with Practical Insights into Document Embedding Generation,” Proceedings of the 1st Workshop on Representation Learning for NLP, Berlin, Germany, pp.78–86 (2016).
- [8] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, Jeffrey Dean, “Distributed Representations of Words and Phrases and their Compositionality,” Proceedings of the 26th International Conference on Neural Information Processing Systems, Volume 2, pp.3111–3119 (2013).
- [9] Spot Inc. (2017) 「Doc2Vec の仕組みと gensim を使った文書類似度算出チュートリアル」, <https://deepage.net/machine_learning/2017/01/08/doc2vec.html>.
- [10] Flickr, <<https://www.flickr.com/>>
- [11] Stanford Vision Lab, Stanford University, Princeton University (2016) 「ImageNet」, <<http://www.image-net.org/>>.
- [12] Karen Simonyan, Andrew Zisserman, “Very Deep Convolutional Networks For Large-Scale Image Recognition,” Proceedings of ICLR’15, pp.1–14 (2015).