

テキストと画像情報を活用した観光地検索支援システムの提案

齊藤 輝 伊藤 恵

本研究では、ユーザの嗜好を考慮した観光地検索システムを提案する。提案するシステムは、テキストと画像情報の両方を解析し、それらから抽出した特徴を基に観光地を推薦することを目的としている。具体的には、ユーザが入力する文章や画像から嗜好を分析し、データベースから特徴に合致する観光地や類似性の高い観光地を推薦する。先行研究では、テキスト解析や画像解析などのアプローチがされていたが、精度の低さが課題とされていた。ゆえに、本研究では、複数の手法を組み合わせることで高い精度を目指す。また、ユーザの観光地に関する嗜好が曖昧な場合でも、システム側から嗜好の候補を提案することで、ユーザの検索負担を軽減し直観的な選択を可能にする。評価方法としては、構築したシステムと既存の地図アプリやブラウザ等と比較し、ユーザの満足度をアンケートで測定する。今後は、システムの開発とその有効性を評価する実験を行う予定である。

1 はじめに

観光産業は急速に発展している。経済産業省の調査 [3] によると、新型コロナウイルスの影響で、2020 年 5 月には旅行業・宿泊業の季節調整済指数 (経済指標などの数値から季節変動の影響を除去した指数。この調査では 2015 年を 100 とし、比較している) は、どちらも 20% を下回ったが、Go To トラベルや全国旅行支援などの政策により、ここ数年で持ち直し、2022 年 10 月には以前よりも高い季節調整済指数となった。そのため、観光産業は今後もさらに発展していくと考えられる。

しかし、観光産業の発展に伴い、課題も見え始めている。観光庁 [2] は、観光客がインターネットを利用して情報を収集する際、求める情報が簡単に見つからず、多様な嗜好を持つ観光客が求める情報と合致していないということが指摘された。そのため、嗜好を把握し、嗜好性に基づいた情報を得られるようにするこ

とが重要である。

そこで本研究では、ユーザの嗜好を考慮した観光地検索システムの提案をする。このシステムでは、まず、文章や画像などユーザが入力した情報を基に、システムが感情や文脈などの特徴を抽出し、ユーザの嗜好を分析する。その結果、用意したデータベースから特徴が当てはまる観光地、また類似性の高い観光地を推薦する。

観光に関する研究は多くされており、様々な手法が提案されている。嗜好を分析する方法は、分析対象のデータ形式で分けると、文章解析と画像解析の二つに分類される。文章解析による研究は、口コミなどのインターネットから収集したデータと、クエリとして入力されるテキストを解析し、類似度の高い観光地を推薦する研究である。画像解析による研究は、Web 等から収集した画像データと、クエリとして入力される画像データの特徴を抽出して類似度し、観光地を推薦する研究である。画像の特徴とは、その観光地のジャンル (自然地形など) や、画像に写る物体の名称 (海、木ばど) を指す。既存の研究では、テキスト解析か画像解析のどちらか一方のみを提案している研究が多い。また、収集しているデータは、口コミ等のテキストか、画像データのどちらか一方に限定されている。

A Proposal of a Tourist Spot Search Support System Using Text and Image Information

Hikaru Saito, 公立はこだて未来大学システム情報科学部 情報アーキテクチャ学科, School of Systems Information Science and Department of Media Architecture, Future University Hakodate.

しかし、テキストから得られる情報と、画像から得られる情報は異なる。Guo ら [1] は、画像とテキストの情報は相補的であり、それぞれが異なる側面から対象を捉えるため、テキストと画像から得られる情報は異なると述べている。そのため、Web などのインターネット上の情報源から、テキストと画像情報の両方を収集することで、より詳細な情報を持つデータベースを構築することが可能だと考えられる。

本研究では、文章解析と画像解析を組み合わせることで、従来より高い精度を目指す。また、検索したいユーザ自身の嗜好が曖昧なものであった場合、その要求を表現することはユーザにとって負担になるだけでなく、システムがユーザの嗜好を汲み取ることが難しくなる。そのため、システム側からユーザ側へ嗜好の候補を提案することで、検索を支援できると考える。さらに、より直観的な選択を可能にするために、観光地の画像を数枚提示し、ユーザが気に入った画像を選択してもらう操作を数回行うことで、嗜好の分析が可能になり、ユーザの情報要求が曖昧であった場合でもユーザが負担に感じることなく検索できる。本研究では、ユーザの検索負担を軽減させるために、システム側から画像の選択肢を提示し、ユーザの情報要求を汲み取ることを目指す。

2 関連研究

2.1 自由記述から観光地を推薦する研究

長尾ら [6] は、自身の要望にあった観光地を見つけることは難しいと考え、多様な要望を自由記述文として入力し最適な観光地を推薦するシステムを提案した。クエリとして入力する自由記述と、各観光地を説明文に対し、形態素解析、ベクトル化、クラスタリングを行うことでベクトル化し、その類似度を計算することで観光地を推薦する。しかし、インターネットから収集したデータ不足による精度の低さや、形態素解析だけでは複雑な文章に対応できず不十分である、という課題が挙げられた。

2.2 画像からユーザ個人の好みを抽出する研究

中野ら [5] は、ユーザ自身が保有する写真には、ユーザ固有の「好み」が含まれていると仮定し、写真を分

析して抽出した「好み」タグを用いた観光スポット推薦システムの開発を行った。写真に写る物体の名称とその信頼度スコアがタグとして抽出することで、ユーザ個人の好みを分析している。結果として、好み抽出が不十分であることや、収集した写真の中に、観光とは関係ない写真がノイズとして影響してしまうという課題が指摘された。

2.3 観光客の嗜好性と観光スポットを数値化して推薦する研究

住友ら [4] は、観光客の嗜好性に合わせた合わせた観光スポットの魅力度を数値化し、観光客の移動時間を考慮した観光経路のアルゴリズムを提案した。観光客の嗜好性を、「観光カテゴリ」に関する嗜好性ベクトルと「感情」に関する嗜好性ベクトルを算出することにより数値化を行っている。作成したアルゴリズムを実装し、募集した 71 名の被験者に利用してもらい、アンケートを実施した。結果として、利用者にとってニーズに合った観光地の提示が可能であると判明した。しかし、すべての被験者に対して嗜好に合った観光地を推薦することができなかったため、観光スポットの魅力度と観光客の満足度の関係について分析し、客観的に評価できる方法を検討する必要がある。

3 提案手法

本研究では、ユーザの嗜好を考慮した、より精度の高い観光地推薦手法の確立を目的としている。1 節で述べたように、テキストから得られる情報、画像から得られる情報は異なる。2.1 節で述べた関連研究はテキストのみ、2.2 節で述べた関連研究は画像情報のみの研究であり、それぞれ片方のデータのみ活用したシステムだったため、高い精度を出すことができなかったと考える。また、2.2 節では関係のない画像がノイズになることが課題であったが、テキストも用いることで除外が可能であると考えられる。そこで、本研究ではインターネットから収集したテキスト・画像情報から特徴を抽出し、ユーザの嗜好を考慮した観光地を推薦するシステムを提案する。インターネットから収集する情報をテキストと画像情報の両方を対象し、抽出した特徴を組み合わせることでより詳細なデー

データベースを構築する。それにより、従来より多い観点で類似度を比較することが可能になり、推薦の精度が向上すると考える。本研究で提案するシステムの概要は以下の通りであり、概要を図1に示す。

1. データベースの構築
2. ユーザからのテキストまたは画像検索
3. システムから観光地画像提
4. ユーザの嗜好分析による新たな観光地推薦

3.1 データベースの構築

観光地に関する情報をインターネットから収集し、抽出した特徴を格納するデータベースを構築する。2.3節を参考にし、感情の観点からも類似度を評価することを目指す。そのため収集するデータは、観光地に関するGoogleMapの画像や口コミ、メタデータや観光サイトで記載されているテキストを想定している。画像は図2のような画像である。また、口コミは“きれい”や“楽しかった”などの感情ベースのテキストを多く含んでいるため収集する。メタデータは、場所やカテゴリなどの属性情報を指す。函館山であれば、函館、自然などのデータを収集する。観光サイトとして、例えば函館市の観光サイトである“はこぶら[7]”というサイトがある。函館山であれば、「函館市と近郊を一望できる展望台。両側に海に囲まれた独特の地形が演出する眺望、特に凛とした空気にきらめく夜景が満喫できる。」など記載されているテキストを収集する。収集したテキストに対してBERT^{†1}を活用し、テキストの意味や文脈情報を数値化した高次元ベクトル、感情的なトーンや感情の種類を数値化した感情ベクトルを抽出する。高次元ベクトルとは、テキストをトークンに分解し、各トークンに対してBERTであらかじめ決められた768次元のベクトルの平均ベクトルである。例えば、“函館山の夜景はきれいです”というテキストからは、[-4.72808421e-01, -2.84938753e-01,, 6.40003234e-02]という768次元のベクトルを得ることができる。感情ベクトルとは、BERTに登録されている感情辞書を基にテキストに含まれる怒り・

嫌悪・恐れ・喜び・中立・悲しみ・驚きの七種類の感情割合算出し、それらを一つのベクトルにまとめたものである。例えば、“函館山の夜景はきれいです”というテキストから、[0.00969428, 0.03076578, 0.12267866, 0.66109705, 0.14224996, 0.02053896, 0.01297538]というベクトルを得ることができ、これは喜びの割合が大きいことを示している。画像情報からはCNN^{†2}を用いることで画像のピクセル情報を数値化したベクトルを抽出する。このベクトルを用いると、画像の視覚的観点から類似度を求めることが可能になる。具体的には、二つのベクトルの内積を測ることで、類似度を評価している。内積の値が大きいということは、二つのベクトルが同じ方向を向いていることを示しており、その二つのベクトルは類似しているといえる。例えば、函館付近の観光スポットである、函館山の夜景(図2)と五稜郭の景色(図3)に対し、城岱牧場展望台(図4)がどちらに類似しているかCNNを用いて求めたところ、城岱牧場展望台と函館山の夜景の内積は19683.505859375、五稜郭との内積は9690.763671875であり、函館山の夜景の方が類似しているという結果を得ることができる。また、画像から生成したキャプションテキストをBERTを用いることで高次元ベクトルに変換する。例えば、城岱牧場展望台(図4)に対してキャプション生成を行うと、“a city at night with a sky background”というテキストを得ることができる。このテキストに対して高次元ベクトルを算出し、データベースに格納する。まとめると、データベースへは、テキストから抽出したメタデータ、高次元ベクトル、感情ベクトル、画像、画像から抽出したCNNベクトル、キャプション、キャプションから抽出する高次元ベクトルを格納する。

3.2 ユーザからのテキストまたは画像検索

ユーザからはテキスト入力または画像アップロードによる検索を行う。入力された情報から、3.1節で述べた手法を活用し、ベクトルやキャプションテキストなどの情報を抽出する。ユーザの入力情報をデータ

†1 「自然言語処理モデルBERT(Bidirectional Encoder Representations from Transformers)」, <https://arxiv.org/abs/1810.04805>

†2 「畳み込みニューラルネットワークCNN(Convolutional Neural Networks)」, <https://www.tensorflow.org/tutorials/images/cnn?hl=ja>

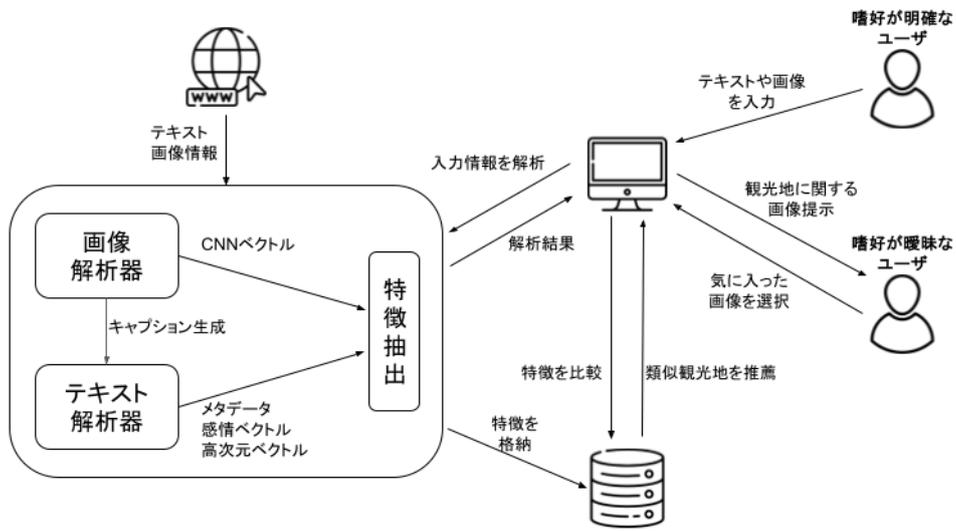


図1 システムの概要図



図2 函館山



図4 城岱牧場展望台



図3 五稜郭

ベースの情報と照らし合わせ、場所などのメタデータの一致度やベクトル同士の内積を求めることで類似度を評価し、ユーザーの嗜好に合わせた観光地を推薦する。

3.3 システムから観光地画像提案

ユーザーの嗜好が曖昧な場合、その要求を表現することはユーザーにとって負担になるだけでなく、システムがユーザーの嗜好を汲み取ることが難しくなる。そのため、システム側から観光地に関する画像の選択肢を提示し、気に入った画像をユーザーに選択させ、

ユーザの情報要求を汲み取る。選択した画像をユーザからの入力とみなし、3.2節で述べた手順で検索する。画像の選択肢を提示し、ユーザに直観的に選択してもらうことでユーザの検索時の負担の軽減と、システム側がユーザの嗜好を汲み取ることを容易にすることが可能になると考える。

3.4 ユーザの嗜好分析による新たな観光地推薦

ユーザが過去に行った観光地に対してお気に入り度を評価してもらい、その内容を分析することによってユーザが興味を持ちそうな観光地を推薦する機能を検討する。

4 評価実験

システムの有用性を評価する方法に関しては、提案システムを用いて比較実験を行い、検索結果に対してのユーザの満足度をアンケートにより測ることで、どの程度ユーザの嗜好を反映できたのか検証する。データの収集や保持のコストの観点から最初は対象地域を道南に限定し、著者らの所属大学の学生に対し実験を行う予定である。具体的には、まず、本システムを使った場合とそうでない場合でそれぞれ観光地を検索する。そうでない場合というのは、既存の地図アプリやブラウザ等、普段ユーザが使用するサービスを想定している。また、システムから観光地に関する画像を提案する機能についても実際に被験者に使ってもらう、アンケートによってユーザの満足度を測り、有効性を検証する。

5 まとめ

本研究では、テキストと画像情報を解析し、その特徴を抽出することで、観光地検索を支援するシステムを提案した。今後は、提案するシステムの開発に着手し、プロトタイプの開発を予定している。プロトタイプ完成後には、4節で述べた比較実験を実施し、システムの有効性を検証する。

参考文献

- [1] Guo, W., Wang, J., and Wang, S.: Deep multi-modal representation learning: A survey, *Ieee Access*, Vol. 7(2019), pp. 63373–63394.
- [2] 観光庁: 「コロナ禍の生活におけるインターネットやSNSからの”情報”に対する意識と旅行」に関する意識調査, <https://www.jtbcorp.jp/jp/newsroom/2021/04/sns.html>, 2021. アクセス 2024年7月18日.
- [3] 経済産業省: アフターコロナの中で、どこまで回復したか - 旅行・観光 -, https://www.meti.go.jp/statistics/toppage/report/minikaisetsu/hitokoto_kako/20230512hitokoto.html, 2023. アクセス 2024年7月18日.
- [4] 住友千将, 岳五一: 観光者の嗜好性を考慮した観光経路構成アルゴリズムの実装と実証実験, *パーソナルコンピュータ利用技術学会論文誌*, Vol. 15, No. 1(2021), pp. 13–20.
- [5] 中野広貴, 荒澤孔明, 渡邊稜平, 服部峻: 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, *電子情報通信学会技術研究報告*, Vol. 118, No. 408(2019), pp. 45–50.
- [6] 長尾将宏, 草野大智, 園田泰子, 本橋洋介, ほか: 自由記述文にもとづく観光スポットレコメンドシステムの提案, *研究報告コンシューマ・デバイス & システム (CDS)*, Vol. 2020, No. 14(2020), pp. 1–5.
- [7] 函館市観光部観光誘致課: はこぶら, <https://www.hakobura.jp/>. アクセス 2024年8月18日.