

旅行者の感情に基づいた観光スポット推薦 Tourist spot recommendation based on travelers' emotions

篠田 広人
Hiroto Shinoda

広島市立大学大学院 情報科学研究科
Graduate School of Information Sciences, Hiroshima City University

概要

本研究では、旅行者の感情という観点から観光スポットを推薦するシステムの構築を目指している。感情に基づく観光スポット推薦とは、例えば「驚き」の観光名所やレストランなど、感情を軸として旅行者に観光スポットを提示するものである。本研究では旅行ブログに着目し、旅行ブログ中のテキストと画像を対応付けた後、感情推定することで、感情に関する観光スポットを自動的に収集する。

1 はじめに

近年、訪日外国人旅行者は増加の一途をたどっており、来たる 2020 年の東京オリンピックに向けて、少なくとも今後数年はこの傾向が続くものと考えられる。訪日外国人旅行者は、旅先の経験を、ツイッターやブログなどのソーシャルメディア上で写真と共に紹介することも少なくない。こうした情報はリアルタイム性が高く、他の旅行者が旅行計画を立てる上で有用なことが多い。そこで本研究では、ソーシャルメディア、特に旅行ブログを情報源とした観光推薦システムの構築を目指す。

従来では、例えば「ショッピング」、「宿泊」といった旅行者の目的別に観光スポットの推薦を行う手法があった。そうした手法に対して、本研究では、感情に基づく観光スポット推薦を行うことを目指している。感情に基づく観光スポット推薦とは、例えば「驚き」の観光名所やレストランなど、感情をひとつの軸として旅行者に観光スポットを提示するものである。

感情という観点から実際の旅行ブログを眺めると、旅行ブログ中の画像から、旅行者本人の驚きや喜びの表情を読み取れることがある一方で、ブログ本文中に感情に関する言語表現が使われることもあることが分かる。そこで、まず旅行ブログ中の画像とブログ本文を対応付け、次に画像またはテキスト中の感情を推定することで、感情に関する観光スポットを自動的に収集する。

2 画像とブログ本文の対応付け

まずは、システム概要について説明する。図 1 にシステム概要を示す。図 1 のシステムにおいては①の「画像解析」と②の「キャプション抽出」、そしてそれらと本文のマッチングを行うことで③の「画像と対応文の抽出」ができる。そしてその後、④の「画像または、テキストの感情推定」を行うことで感情に関する観光スポットを自動的に収集できると考えられる。

次に、画像に写っている観光スポットの検出について述べる。TravelBlog 中から、観光地などを写した画像に着目し、Google Cloud Vision API を使って解析すると図 2 のよ

うになる。図 2 の解析結果から、「amusement park」、「resort」が図 2 の画像と一致していると考えられる一方で、「amusement ride」や「games」といったように図 2 の画像にそぐわないものもあった。

さらに、画像に写った人物の表情認識について述べる。実際に、Google Cloud Vision API で解析した結果を図 3 に示す。図 3 にあるように、喜びの度合いを表す「joyLikelihood」が「VERY_LIKELY」、つまり喜んでいる可能性が非常に高いことが分かる一方、悲しみ、怒り、驚きを表す「sorrowLikelihood」、「angerLikelihood」、「surpriseLikelihood」が「VERY_UNLIKELY」であることから悲しんでいる、怒っている、驚いている可能性は低いことが分かる。

ここで、Google Cloud Vision API の表情認識とカテゴリ検出の解析結果を用いて提案システムの動作例を説明する。動作例の一つとして、例えば、画像のカテゴリ検出結果が本文中の単語と一致した場合について説明する。図 4 に解析結果を示す。図 4 から右に画像があり、左にカテゴリ検出の解析結果を示している。図 4 から下の文が解析結果に対応する本文中の文で、解析結果の Moraine lake と本文中の Moraine lake が一致していることが分かる。

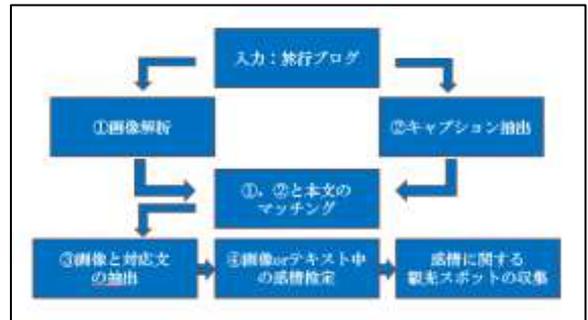


図 1: システム概要



図 2: カテゴリ検出の解析結果

また、本研究では、ブログ中の一連の文がある画像に関するものであるかどうかの判定を行った。画像に対応する文には「○」ラベル、それ以外の文は「×」ラベルを付与した。表 1 にラベルの付与例を示す。また図 5 に表 1 に関

連する画像を示す。図5にはキャプションがあり、その中に Dinner という単語がある。表1の一番目の文、二番目の文にも dinner という単語がある。よって画像と対応している文と考えられるので、ラベルを「○」としている。また、表1の三番目の文は、図5で表情認識をした結果、喜びの度合いが非常に高かった。よって、表1の三番目の文で驚きや喜びを表す amazing や dinner に関連があると思われる単語である food があるので、ラベルを「○」としている。ブログ中の一連の文がある画像に関するものであるかどうかを判断するタスクを、本研究では系列ラベリング問題として解く。石野[1]らは、ある Twitter ユーザが投稿した一連のツイートから、旅行計画中および旅行中のものを、畳み込みニューラルネットおよび Long short-term memory (LSTM)を用いて抽出する手法を提案している。本研究では、石野らの手法に基づき、画像に対する対応文であるかどうかを LSTM で抽出する。



図5：表1に関連する画像

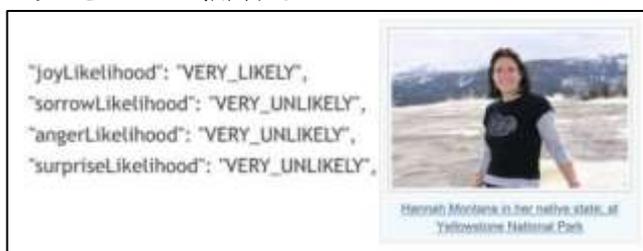


図3：感情推定の解析結果

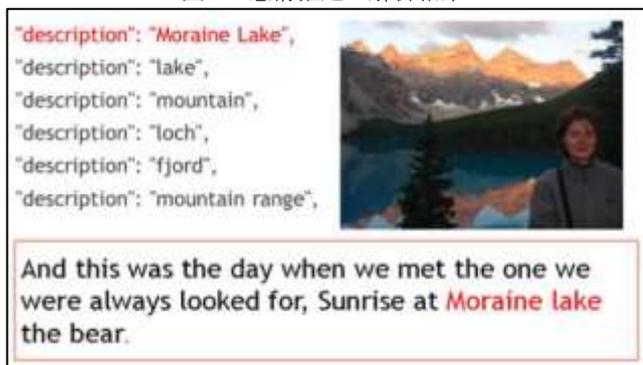


図4：提案システム動作例

表1：ラベルの付与例

ブログ本文	ラベル
Today we were real tourists and with video camera in to went up and had dinner on the Canadian National tower.	○
We got given a voucher for Christmas to have a swish dinner at the 360 degree revolving restaurant.	○
It was amazing- both the view and the food!	○
Sunday we will head down with the relies to Niagara Falls.	×

3 実験

提案手法の有効性を調べるために実験を行った。

データ

最大の旅行ブログサイトのひとつである TravelBlog(travelblog.org)のデータを実験に用いる。478,034件のブログエントリーおよびこれらのブログに含まれる933,997件の画像を収集した。これらの画像データに対し、Google Cloud Vision APIを用いて物体検出、ランドマーク検出、顔検出を行った。このデータの中で任意に選択した256画像に対し、画像と関連する文を人手で対応付けたものを訓練および評価に用いる。

実験条件

ブログ中の各文を、まず100次元の実数ベクトルに変換し、次に畳み込み関数を適用して300次元のベクトルにする。このベクトルに対し、各次元の最大値を計算した結果として得られる300次元のベクトルをLSTM関数に入力する。LSTM関数の出力に線形関数を適用し、最も値の大きい次元に対応するラベルを予測ラベルとする[1]。評価尺度には精度、再現率、F値を用いる。

実験結果

実験の結果、精度0.452、再現率0.229、F値0.304が得られた。再現率が低い原因のひとつに、物体検出の結果得られた言語表現(ラベル)とブログ中の表現が一致しない点が挙げられる。

4 おわりに

本研究では、旅行ブログ中の画像と本文を対応付ける手法を提案した。今後は、画像と本文を対応づけた後、感情推定をすることで、感情に関する観光スポットを自動的に収集する手法を考えていく必要がある。

参考文献

[1] 石野亜耶, 難波英嗣, 竹澤寿幸(2017). Twitter を利用した旅行計画者の行動分析. 観光情報学会第16回研究発表会.