

SNS 上の画像群からのユーザー嗜好の抽出と観光広告への応用

石野 淳一[†] 中田 洋平[‡] 日吉 久礎[‡]

青山学院大学大学院理工学研究科[†] 青山学院大学理工学部[‡]

1. 序論

広告は、消費者の購買意欲に影響を与えるという意味で企業にとって重要である。近年、広告を出すターゲットを絞り、ターゲットの嗜好にあった商品を広告する手法が注目を集めている。例えば、インターネットの検索連動型広告などの広告手法は効果的である。このような広告手法は、行動ターゲティング広告と呼ばれ、現状ではテキストマイニング技術により、検索ワードなどを解析することで、消費者の嗜好を抽出している。

また、近年ではスマートフォンやタブレット端末の普及により、SNS の利用人口は増加し続けている[3]。SNS は新しい形のコミュニケーションの手段であり、趣味などを通じてつながったユーザー同士がコミュニケーションをとる場である。また、SNS の特性上 SNS ユーザーは、コメントや画像といった形で、他のユーザーとコミュニケーションをとる。これらのコミュニケーションの記録であるコメントや画像には各ユーザーの感情、趣味、嗜好が多分に反映されていると考えられる。つまり、これらのコメントや画像を俯瞰的にとらえ、分析することで、各ユーザーの嗜好や趣味を抽出することが可能であると考えられる。例えば、どの様な食べ物が好きか、どの様な服が好きか、どの様な場所が好きか、などといったユーザーの嗜好が抽出できるのではないだろうか。

そこで、本研究では SNS 上の各個人の画像群を解析することにより、各個人の嗜好を抽出することを目的とする。特に本研究では、風景画像群から所有者がどのような観光地を好むか、という嗜好抽出方法を提案する。今回提案する方法は、旅行代理店や、総合旅行サイトなどでの観光広告への応用を念頭においている。

2. 画像群からの観光地に対する嗜好抽出方法

本研究では、画像を用いてユーザー同士が観光情報の共有を行える SNS があること想定する。そして、この仮想 SNS ユーザーの観光地に対する嗜好を各ユーザーの画像群から抽出する。そのためにまず、観光地を古風、ビーチ、都会、大地の 4 個のカテゴリーに分類する。抽出するユーザーの嗜好は、各カテゴリーをどれだけ好むかという、カテゴリーポイントによってあらわせるものとする。カテゴリーポイントは、古風、ビーチ、都会、大地ポイントである。最終的に、得られたカテゴリーポイントをレーダーグラフの形で表示する。カテゴリーポイントの算出手順は、以下の 5 つの手順のとおりである。

1. 次の 2 種類の画像群をそれぞれカテゴリーフォルダ、ユーザーフォルダと呼ぶ。1 つ目の画像群は、各カテゴリーに属する画像の集合である。特に本研究では、古風フォルダ（京都の画像）、ビーチフォルダ（沖縄の画像）、都会フォルダ（東京の画像）、大地フォルダ（北海道の画像）の 4 種類を用意する。2 つ目の画像群は、ユーザーが SNS 上で所有している画像の集合である。
2. 次に、ユーザーフォルダ内の各画像と、各カテゴリーフォルダ内の画像ごとの HSV 色モデルの色合いの差を算出する。各画像間の HSV の色合いの差は、画像ごとに HSV 空間におけるヒストグラムを作成し、ヒストグラム間のユークリッド距離を計算し、これを差とする（OpenCV を使用[1][2]）。
3. ユーザーフォルダ内の各画像と各カテゴリーフォルダの距離を計算する。ユーザーフォルダ内の全画像の、2 で求めた各カテゴリーフォルダの全画像に対する距離の中で下位 k 個の距離の平均値をとる。そして、この平均値をカテゴリーフォルダとユーザーフォルダ内の 1 枚の画像との距離とする。

From pictures on SNS, sampling user's taste and apply sightseeing advertisements.

[†]Junichi Ishino Aoyama Gakuin University

[‡]Youhei Nakata Aoyama Gakuin University

[‡]Hisamoto Hiroyoshi Aoyama Gakuin University

4. ユーザーフォルダ内の各画像に対し、カテゴリポイントをつける。ポイントの付け方は、画像ごとに、各カテゴリに属するかの判定を行い、属するものにカテゴリポイントを与え、それ以外にはポイントを与えない。カテゴリに属するかは、3 で求めた距離に対して、閾値処理を行うことによって判定する。画像判別を行うにあたり、判別しやすいカテゴリと、しにくいものが存在したため、判別しやすいカテゴリと、しづらいカテゴリの間で、ポイントの重みを変えることにより差を持たせた。つまり、判別しやすいカテゴリのポイントは小さく、判別しづらいカテゴリのポイントを大きくした。
5. ユーザーフォルダに対して、カテゴリポイントをつける。4 でユーザーフォルダ内の各画像につけたカテゴリポイントをユーザーフォルダ内の全画像で、カテゴリごとに平均し、ユーザーフォルダのカテゴリポイントとする。

以上の手順で算出した、ユーザーフォルダに対する各カテゴリのカテゴリポイントをレーダーグラフで表示することによって、ユーザーがどのようなカテゴリの観光地を好むか、という嗜好を可視化できる。

3. 検証

提案手法の有効性を示すための実験を行った。実験方法は、ユーザーフォルダとして、各カテゴリのイメージに合う画像群から成る 4 個のテストフォルダを用意した。テストフォルダとして用意したのは、奈良、グアム、横浜、ニュージーランドの画像群である。それぞれ順に、古風、ビーチ、都会、大地ポイントが高くなることを期待した。結果として、4 個のテストフォルダは、すべてイメージしたカテゴリのポイントが最も高くなり、本研究で提案した手法の有効性を示す結果が得られた（図 1、2）。しかし一方で、テストフォルダの中には高くなってほしくないポイントが高い値を示したフォルダも存在した。つまり、色情報だけを特徴量とするポイント算出では精度が低く、限界があるということも分かった。

4. 結論

本研究では、風景画像群から観光地に対する嗜好抽出方法を提案した。そして、風景画像群

からどのような観光地を好むのか、というユーザーの嗜好を抽出することが可能であることを示した。しかし一方で、単純な色情報のみを用いたポイント化では、精度に限界があることも分かった。将来研究として、精度の高いユーザー嗜好の抽出法を開発することが挙げられる。具体的には、被写体の位置情報を考慮することや、被写体の形状などの特徴量を考慮する必要があると考えられる。また、多様なニーズへの対応を目指し、カテゴリ数を現状より増やすことも必要である。

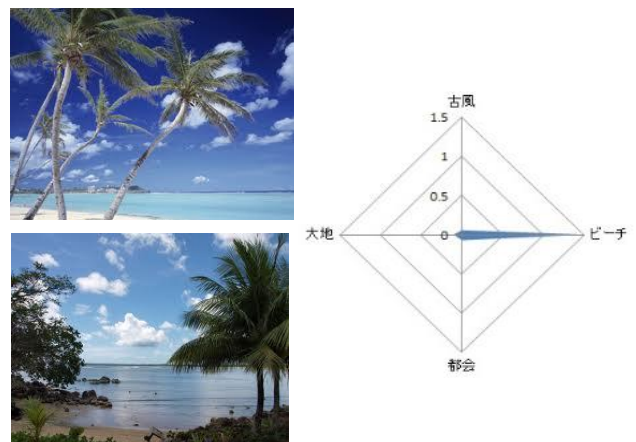


図 1：グアムフォルダの画像の例と得られたレーダーグラフ

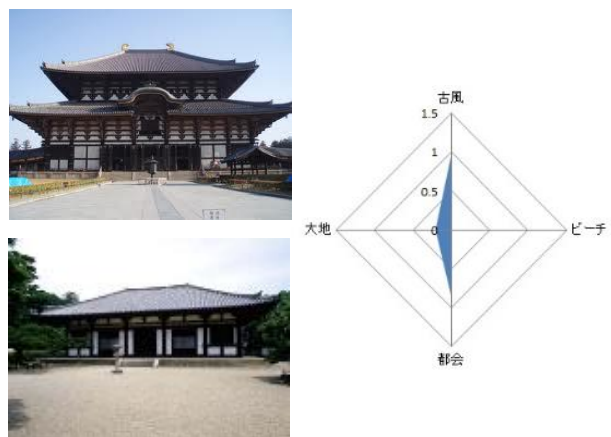


図 2：奈良フォルダの画像の例と得られたレーダーグラフ

5. 参考文献・資料

- [1] 奥富正敏：「デジタル画像処理」，財団法人画像情報振興協会，(2004)。
- [2] 永田雅人：「実践 OpenCV 映像処理&解析」，株式会社カットシステム，(2009)。
- [3] ICT 総研，SNS 利用動向・広告活用状況に関する調査，
http://www.ictr.co.jp/topics_20111227.html。