

個人の旅行写真の一般物体認識に基づく 観光地推薦のためのユーザインタフェース

理学専攻 情報科学コース 1740653 北村 理紗 (指導教員：伊藤 貴之)

1 概要

近年の急速なデジタル端末の普及に伴い、誰もが手軽に写真を撮影することが可能となった。今や写真が生活や行動履歴を記録するものとなりつつあり、ライフログの一種といっても過言ではない。そこで、このような写真を分析することで個人の嗜好を推測できるのではないかと考えた。本論文では、過去の旅行写真の被写体情報に着目して撮影者の旅行の嗜好を推測し、その嗜好に合った観光地を推薦する手法を提案する。

また、過去に撮影した大量の写真についてその全貌を必ずしもユーザが詳細に把握しているとは限らない。そこで、過去の旅行写真を被写体ごとに分類提示することで、過去の旅行履歴を振り返りながら旅行プランを考え、次の旅行で訪問する観光地の決定を支援するユーザインタフェースを提案する。

なお、本論文では、旅行の目的地となる都市や地域は決まっているが、その都市や地域の中で具体的にどのような観光名所を訪れるか決まっていないという状況を想定して、その都市や地域における観光名所を推薦するシステムを提案する。

2 関連研究

写真を利用した旅行推薦システムに関する研究は数多く発表されている [1, 2]。Crandall ら [1] は、写真に付与されたタグとジオタグ情報からランドマークや観光スポットの抽出手法を提案している。また、Gao ら [2] は写真に付与されたタグとジオタグ情報、旅行サイトの口コミ情報からランドマークをランキングし、ユーザの興味を考慮した旅行ガイドシステムを提案している。これらの手法はいずれも写真の位置情報を前提とした推薦手法であり、位置情報が付加されていない写真を利用できないという問題がある。

3 提案手法

本手法では、写真に付与されたジオタグ情報を使わずに、旅行写真に一般物体認識を適用することで、よく撮影する被写体のキーワードを集める。このキーワードを検索ワードとし、特定の都市や地域における観光情報を抽出する。また、物体認識結果より得られたキーワード間の共起関係を、ノードをクラスタリングして配置するグラフ可視化手法、Koala[3]を用いてグラフ構造化する。このグラフに写真を追加した状態で表示することで、被写体ごとに旅行写真を整理し、ユーザ操作によって訪問地の決定を支援するユーザインタフェースを提供する。観光地の推薦では、ユーザインタフェース上で関心のあるノードをクリック操作で選択してもらい、そのキーワードに関連する観光地を地図上に表示する。

3.1 被写体情報の取得

本手法では Microsoft の画像認識 API である Computer Vision API を用いることで一般物体認識の情報

を取得する。この API では一般物体認識結果を物体名のリストで返し、その各々について範囲 $[0,1]$ の実数で確信度を返す。

3.2 物体認識結果のグラフ構造化

本手法のユーザインタフェースでは、キーワードをノードとしたグラフを構築し、Koala に搭載されたクラスタリング処理を適用して構造化する。 i 番目の写真の j 番目のキーワードの確信度を c_{ij} 、キーワードの総数を m 、写真の総数を n とし、以下の処理によってノードのクラスタリングを実行する。

- j 番目のキーワードに対応するノードを n 次元ベクトル (c_{1j}, \dots, c_{nj}) で表現する。
- 任意の 2 ノード間についてベクトルの内積を算出し、内積が閾値以上であれば 2 ノードをエッジで接続する。この処理は同一写真での共起度の高いキーワードをエッジで接続することに相当する。
- 同一キーワードと共起する傾向にある 2 つのキーワードをできるだけ同一クラスタに所属させる。

Computer Vision API では、上位概念と考えられる一般性の高いキーワードと、より具体的な、下位概念と考えられるキーワードが同時に付与される。そのため、上位概念となる同一のキーワードに共起する下位概念のキーワード群が同一クラスタに所属されるようなクラスタリング結果を得る。

3.3 ユーザインタフェース

本手法では、Koala により構造化されたグラフを JSON 形式で出力し、D3.js の Force Layout を利用してユーザインタフェースを提供する。D3.js はグラフ描画できる JavaScript のライブラリで、Force Layout では、ノード間の距離や画面の中心に働く力などの制約を加えたグラフを描画することができる。本手法では、Koala で得られた 1 つのクラスタを 1 つのノードとし、以下のようにノードを定義する。また各ノードを表示する際には、ノードに属するキーワードをあわせて表示する。

- エッジ数の最も多いクラスタを「ルートノード」とする。
- 閾値以上のエッジ数でエッジ数の最も多いクラスタとエッジが張られているクラスタを「ハブノード」とする。

エッジ数が多いノードであればあるほど、そのノードに対応するキーワードは多様なキーワードと共起して写真に付与される確率の高いキーワードであり、上位概念に属するキーワードであると考えられる。そこで、クラスタのエッジ数で表示するノードを制限し、木構造のように表示することで、ユーザの旅行目的や

訪問地を絞り込むのに適したユーザインタフェースを提供する。これにより、まず旅行目的に近い意味を有する上位概念のノードをユーザに選択してもらい、それに接続されている、下位概念のキーワードに対応するノードを同時に表示する。このような操作により、ユーザの旅行目的に沿った形でキーワードを絞り込む。

3.4 観光地の検索

本手法では、観光情報の抽出に Google Maps Platform の Places API を使用する。この API では、様々なパラメータを指定してプレイス（場所）を検索することができ、本手法では場所名や住所、ユーザのクチコミなどとキーワードを照合して検索する Text Search requests を使用する。本手法ではユーザインタフェース上で関心のあるノードもしくは写真ノードを選択させ、選択されたノードに含まれるキーワード（もしくは選択された写真に付与されたキーワード）を観光地検索に使用するキーワードとする。また、本手法では、Maps JavaScript API を利用して取得した観光地を Google マップ上に表示する。

4 実行結果

本章では、旅行写真 2,581 枚に対して本手法を適用した例を紹介する。物体認識により得られたキーワードの総数は延べ 145,000 個である。

まず、Koala でグラフ構造化し、ハブノードにあたるキーワードを抽出した例を図 1 に示す。ハブノードはエッジ数が多いクラスタであるため、一般性の高い単語が抽出されていることがわかる。図 1 の (d) は最もエッジ数の多いクラスタでルートノードにあたるものであり、outdoor や sky など、極めて一般性の高い単語が抽出されている。

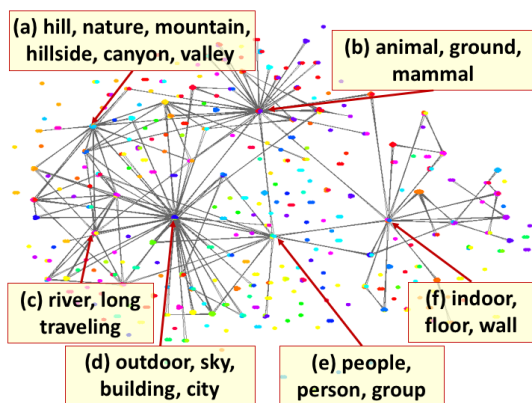


図 1: ハブノードのキーワード

次に、ユーザインタフェースの例（初期画面）を図 2 に示す。この結果では、最もエッジ数の多いルートノードを中央に配置する同時に、ルートノードからエッジが張られているハブノードに対して、各ハブノードのキーワードが付与された写真を 4 枚選んでノードとして配置している。ノードの横に表示されたキーワードは、各ハブノードを代表すると考えられるキーワードを手動で選んでいる。

図 2 の初期画面から nature が属するハブノードを選択し、その子ノードを表示した例を図 3 に示す。nature の子ノードとして、nature の下位概念のキーワードと

考えられる forest や lake などのキーワードが表示されていることがわかる。



図 2: ユーザインタフェースの例

5 まとめと今後の課題

本論文では、旅行写真の被写体情報を利用した観光地推薦手法と被写体ごとに写真を一覧表示するユーザインタフェースを提案した。今後の課題として、観光地の検索処理にて、キーワードによっては推薦情報の満足度が低い情報もみられるため、有用なキーワードに改善する手法を検討したい。

参考文献

- [1] D. J. Crandall, L. Backstrom, D. Huttenlocher, J. Kleinberg, Mapping the world's photos, In Proceedings of the 18th International Conference on World Wide Web, pp. 761-770, 2009.
- [2] Y. Gao, J. Tang, R. Hong, Q. Dai, T. S. Chua and R. Jain, W2Go: a travel guidance system by automatic landmark ranking, In Proceedings of the 18th ACM international conference on Multimedia, pp. 123-132, 2010.
- [3] T. Itoh, K. Klein, Key-node-Separated Graph Clustering and Layout for Human Relationship Graph Visualization, IEEE Computer Graphics and Applications, Vol. 35, No. 6, pp. 30-40, 2015.

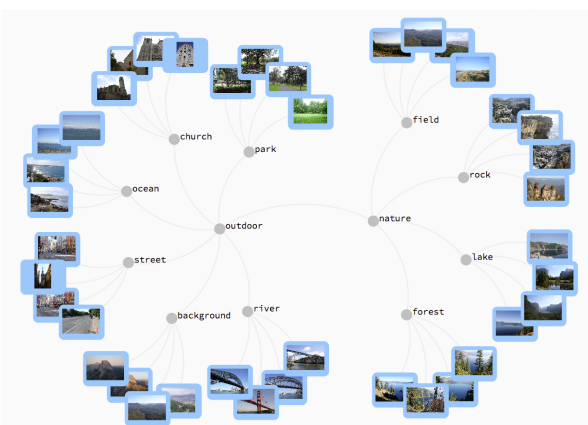


図 3: キーワードノードをクリックした例