

画像の“説得度”を考慮した地域情報検索

岩木 祐輔† アダム ヤトフト† 田中 克己†

† 京都大学大学院情報学研究科 社会情報学専攻

1. はじめに

現在, Web 上には多くの地域情報が存在している。しかし, 旅行者が調べ物をする際などに, リンク構造 (PageRank や HITS など) に基づく従来の検索では, 必ずしも「見て行きたくなる」コンテンツを発見できるとは限らない。

そこで, 本研究では, Web アルバムの実世界情報に関する集合知をもとに, 写真の撮影対象や場所・撮影者の属性から, 画像を見たときに行きたくなる度合い (= 説得度) を数値化し, それをもとに検索を行う方法を提案する。

2. 提案手法

「見たら行きたくなる」度合いを定式化するため, 説得工学[1]の手法を参考にした。説得工学とは, 人を(強制ではなく)説得させて動かす要素を, 分析的に追究している分野である。

説得の要素には大きく分けて「信用性」と「専門性」の 2 つがあり, これを画像検索に照らし合わせると, 図 2 のようになる。従来の画像検索においては, 「画像は周辺テキストとの関連が強い」という仮定が成り立つもとでは偽りのなさや検証容易性を保証しているが, 客観性や専門性については一切考慮されていない。



図 1: 「屋久島 観光」の画像検索の例
屋久島で人気の海の写真などは見られない

Regional Information Retrieval Considering the degree of Persuasiveness of Images

†Yusuke Iwaki, Adam Jatowt, and Katsumi Tanaka.

Department of Social Informatics, Graduate School of Informatics, Kyoto University

そこで, 本研究では Web アルバムなどの体験に基づくコンテンツをもとにして, 客観性を保証するための指標 Position Rank と Support Rank を提案する。これによって従来検索では保証されなかった信頼性を補完することができる。なお, 専門性に関する指標は Photo Quality[2] や Attractiveness[3]などの従来研究にて言及されており, 本研究では詳細に触れない。

2.1. Position Rank : 撮影場所の Popularity

ある地域において典型的な被写体や撮影場所は, 客観性が高いと考えられる。本節では, 対象の写真が Web アルバムでどのくらい典型的に撮影されているかを数値化する指標 Position Rank を提案する。

Web アルバムにおける写真列はすべて位置情報を持っていると仮定すれば, Web グラフと同様, グラフ構造として表すことができる。そこで HITS アルゴリズムを適用し, その Hub スコアを Position Rank とする。これは多くの人が旅行において高い確率で訪れて写真撮影をしている場所を抽出していることになる。

なお, Authority スコアや PageRank ではなく Hub スコアを採用したのは, 予備実験での評価に基づく。

2.2. Support Rank : 撮影対象の Sociality

前節で提案した Position Rank は, 「電車」や「駅」などのように一部のマニアばかり撮影しているものであってもハイスコアとなりうる。一般的には, さまざまなジャンルを万遍なく撮影している「多様な撮影者」がより多く撮影している対象物のほうが, 客観性が高いと考えられるため, ここでは地物の撮影者層の偏りのなさを数値化する Support Rank を提案する。

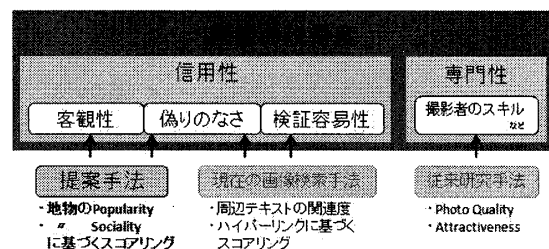


図 2: 画像検索における説得の要素

Step 1. 撮影者層の一様性を評価

それぞれの写真の支持が一様分布に基づくとして、その有意性を検定する、有意に一様といえるものほどよい被写体とする。

Step 2. 撮影者の一般性を評価

それぞれの撮影者が、よい被写体ほど多く撮影すると仮定して、有意性を検定する。よくない被写体ほど多く撮影する撮影者はおそらくマニアであり、このステップでスコアが低くなる。

Step 3. Support Rank の算出

それぞれの被写体が、よい撮影者ほど多く撮影していると仮定して、有意性を検定する。これを正規化して Support Rank とする。

表 1 に図 3 の例におけるそれぞれのステップでの正規化スコアを順に示した。

3. 検索精度の評価

前章で提案した 2 つの指標の導入によって、「見て行きたくなる」画像を取得できるかの検証を行った。

実験に用いた写真集合は、「京都 紅葉」などのクエリで Web アルバムから収集した。これは、現在の Web 上の画像の多くには位置情報が付加されておらず、確実に位置情報が付加されている写真を利用するためである。

ベースライン手法として、テキストベースの従来の検索のランキングを用いて、本研究の手法との適合率/再現率の比較を行った。

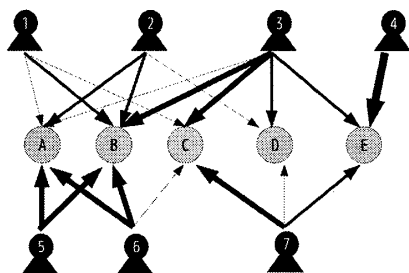


図 3: 支持グラフの例

表 1: 図 3 のグラフにおけるスコア

Step 1		Step 2		Step 3	
id	score	id	score	id	score
A	0.1352	1	0.1915	A	0.3420
B	0.2985	2	0.1890	B	0.3422
C	0.1007	3	0.1193	C	0.1524
D	0.4201	4	0.0259	D	0.1627
E	0.0455	5	0.1884	E	0.0006
		6	0.1914		
		7	0.0944		

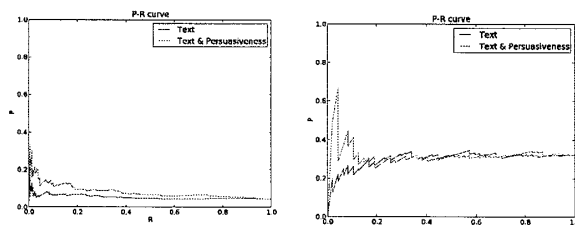


図 4: 適合率/再現率の評価
(左:「京都 紅葉」 右:「潮岬 観光」)

「京都 紅葉」「潮岬 観光」それぞれの例での結果を図 4 に示した。いずれの例においても従来のテキストベースの手法のみの場合と比較して精度の向上が見られた。

この評価では「専門性」の要素を組み入れておらず、単純にきれいな写真で人が正解とした写真でもシステムは不正解と判定した写真が、数多く存在すると思われる。そのため、今後は専門性も要素として組み入れて評価をする必要がある。

4. まとめと今後の課題

本研究では、地域画像検索において「見て行きたくなる」写真を効率よく発見できるように、画像の説得度を提案し、その分析と評価を行った。主に写真の対象物の客観性を考慮することで、従来のキーワードベースの手法と比較して精度向上が見られた。

今後の課題としては、画像の専門性を尺度として追加した際の説得度の評価が挙げられる。

謝辞

本研究の一部は、京都大学 GCOE プログラム「知識循環社会のための情報学教育研究拠点」、および、文部科学省科学研究費補助金特定領域研究「情報爆発時代に向けた新しい IT 基盤技術の研究」、計画研究「情報爆発時代に対応するコンテンツ融合と操作環境融合に関する研究」（研究代表者：田中克己、A01-00-02、課題番号 18049041）によるものです。ここに記して謝意を表します。

参考文献

- [1] Fogg B. J.: *Persuasive Technology: Using Computers to Change What We Think and Do*, Morgan Kaufmann Pub(2002). ISBN: 9781558606432.
- [2] Ke Y., Tang X. and Jing F.: *The Design of High-Level Features for Photo Quality Assessment*, CVPR' 06: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington DC, USA, IEEE Computer Society, pp.419-426 (2006).
- [3] Pedro J.S. and Siersdorfer S.: *Ranking and Classifying Attractiveness of Photos in Folksonomies*, 18th International World Wide Web Conference (WWW2009) (2009).