

VisualRank における位置情報活用の検討

川久保秀敏[†] 柳井 啓司[†]

[†] 電気通信大学大学院 情報工学専攻
東京都調布市調布ヶ丘 1-5-1

E-mail: [†]kawaku-h@mm.cs.uec.ac.jp, ^{††}yanai@cs.uec.ac.jp

あらまし 本論文では VisualRank を用いて位置情報付き画像をランク付けする手法を提案する。提案手法では、カラーヒストグラムと SIFT 記述子 [1] による Bag-of-Features 表現 [2] を用いて画像の類似度を計算する。ランキングを求めるときに、地理上の一点を指定し、位置情報を用いて VisualRank のバイアスペクトルを作成する。代表的な視覚特徴を持ち、指定した地点の近くで撮影された画像ほど上位にランキングされる。本論文では提案手法を用いて 350 語の単語について実験を行った。名詞 250 語と形容詞 100 語について Web 上から位置情報付き画像を収集し、各単語について代表的画像のランキングを求めた。

キーワード VisualRank 位置情報付き画像 Web 画像 画像ランキング

Inquest of VisualRank for Geotagging images

Hidetoshi KAWAKUBO[†] and Keiji YANAI[†]

[†] Department of Computer Science, The University of Electro-Communications

E-mail: [†]kawaku-h@mm.cs.uec.ac.jp, ^{††}yanai@cs.uec.ac.jp

Abstract In this paper, we propose a image ranking system for geotagging images. In proposal system, we use VisualRank using color histograms and Bag-of-Features representations of SIFT descriptors to calc similarities of images. Proposal system need a geographical coordinate as a parameter. The parameter will be used in making geotaged-base bias vectors. We tested the system using 250 noun concepts and 100 adjective concepts. We collected images from web about each word concept. Then we calced image ranking about each concept.

Key words VisualRank, geotagging image, web images, image ranking

1. はじめに

本論文では、VisualRank [3] を用いた位置情報付き画像のランキング手法を提案する。提案手法では、VisualRank の中で位置情報にもとづくバイアスペクトルを用いることによって画像の視覚特徴と位置情報の両方に基づいたランキングを生成する。提案手法ではランキングを求めるときに、地理上の一点を指定する。代表的な視覚特徴を持ち、指定した地点の近くで撮影された画像ほど上位にランキングされる。

1.1 研究の背景

現在、Web 上の画像検索エンジンではテキストベースによる手法が主に使われており、画像特徴に基づく検索手法の研究が行われている。テキストベースの手法による検索結果に、画像特徴を用いた修正を加えることで検索精度の向上を目指す研究があり、そうした研究の 1 つに VisualRank がある。VisualRank を用いると、画像集合に対してランキングを生成し、類似画像の多い代表画像を示すことができる。

本研究では、VisualRank を用いて位置情報付き画像に対するランク付けを行う。位置情報付き画像とは、撮影された場所の緯度・経度（ジオタグ）が付与された写真画像である。現在、Web 上には大量の位置情報付き画像が存在し、画像の位置情報を一般物体認識 [4] に生かすことを目指す研究がある [5], [6]。本研究は、そうした研究で用いる位置情報付き画像データセットの自動作成に役立てることを目的の 1 つとしている。

2. 関連研究

He らは画像データベースの分析手法として、ImageRank [7] を提案した。ImageRank では、まず画像の近傍関係をグラフ化し、画像から画像への遷移確率を求める。そして遷移確率行列の固有ベクトルを用いて、代表画像を決定している。

Jing らによって提案された VisualRank [3] は PageRank アルゴリズム [8] を画像に適用したものであり、画像間の類似度をもとに画像のランキングを得る手法である。Jing らは、VisualRank をテキストベースの画像検索結果を修正するため

に用いている [3] . 画像の類似度には SIFT 特徴 [1] のマッチ数を使い, 商品画像やランドマーク画像を対象にした実験が行われている .

VisualRank では, しばしば上位画像の見た目が似かよ, 検索結果の多様性が失われる . そこで安倍らは, 検索結果の多様性を高めた Multiclass VisualRank を提案している [9], [10] . Multiclass VisualRank では, ランク付けの前に画像をクラスタリングしておき, 個々のクラスタに対して別々に VisualRank を実行する . そして各クラスタ内でのランキング上位画像を並列に提示する . 事前にクラスタリングを行うことで, 検索結果の多様性を確保している .

3. VisualRank

各画像の VisualRank 値からなる列ベクトル R は式 (1) を反復し, 収束させることで得られる .

$$R = S \times R \quad (1)$$

式 (1) 中の S は, 類似度行列の各列を正規化したものである . 各列の合計が 1 になるように正規化することで, VisualRank 値の合計が変化しないようにする .

式 (1) に示した反復ステップに, 補正ベクトル P を加えたものが式 (2) である .

$$R = \alpha(S \times R) + (1 - \alpha)P, \quad (0 \leq \alpha \leq 1) \quad (2)$$

補正ベクトル P として一様なベクトルを与えると, 各画像の VisualRank 値を均等化させる方向へ補正がかかる . 一方, 不均一なベクトルを与えると, 一部の画像を強調するバイアスがかかる . 補正の強度はパラメータ α によって調整する . 多くの場合, $\alpha \geq 0.8$ を満たす値が設定される .

4. 提案手法

本研究では位置情報付き画像を対して VisualRank を使用し, 画像特徴と位置情報の両方に基づくランキングを生成する手法を提案する . 提案手法ではランキングを生成する際に, 地理上の一点を注目点として指定する . 代表的な視覚特徴を持ち, 撮影された地点が注目点に近い画像ほど上位にランキングされる .

画像の類似度行列はカラーヒストグラムと SIFT 記述子 [1] による Bag-of-Features 表現ベクトル [2] を用いて作成し, 補正ベクトルは位置情報を用いて作成する . それぞれについての説明を以下に述べる .

4.1 画像の類似度

[3] では, 画像の類似度に SIFT 特徴のマッチ数を用いている . これは商品やランドマークといった特定の見た目を持つ物体の画像について有効である . 本研究ではより多様な見た目を持つ単語に対応するため, カラーヒストグラムと SIFT 記述子による Bag-of-Features 表現を用いた . これらの特徴表現手法は一般物体認識の分野で良く使われているものである .

類似度にはヒストグラムインタセクションを用いた . カラーヒストグラムと Bag-of-Features 表現それぞれで類似度行列を

作っておき, それらを合成することで, 特徴量を混合した類似度行列を得ることができる (式 (3)) .

$$S_{combine} = \beta S_{color} + (1 - \beta) S_{BoF}, \quad (0 \leq \beta \leq 1) \quad (3)$$

S_{color}, S_{BoF} は各特徴を用いて作成した類似度行列を示し, それらを合成して得られる類似度行列が $S_{combine}$ である . β を変化させることで, 特徴量の重みを設定することが可能である .

4.2 補正ベクトル

提案手法では地理上の一点の座標をパラメータとして与え, 注目点とする . 注目点に近い地点で撮影された画像の VisualRank 値をより大きくするために, 不均一な補正ベクトル (バイアスベクトル) を用いた . バイアスベクトルによる補正は各反復ステップでかかるため, 『注目点付近の画像によく類似した画像』の VisualRank 値も間接的に大きくなる .

位置情報によるバイアスベクトル P_i^{geo} は, まず式 (4) のようにして作成し, VisualRank ベクトル R と同じ 1-norm になるよう正規化して使用した (式 (5)) . D_i は画像 i の位置情報と注目点の距離である . 地球を半径 1 の完全な球体と近似して, 球面三角法で計算した (式 (6)) .

$$p_i^{geo} = \frac{D_i}{\pi} \quad (4)$$

$$P_i^{geo} = \frac{p_i^{geo}}{\|p_i^{geo}\|_1} \|R\|_1 \quad (5)$$

$$D_i = \cos^{-1}(\sin(lat_i) \sin(lat_A) + \cos(lat_i) \cos(lat_A) \cos(long_i - long_A)) \quad (6)$$

$lat_i, long_i$ は画像 i の緯度・経度であり, $lat_A, long_A$ は注目点の緯度・経度である .

5. 実験

5.1 位置情報付き画像の収集

オンラインアルバムサービス Flickr [11] では, 位置情報付き画像のみを対象にした画像検索を行うことが可能である . 本研究では位置情報画像の収集に, Flickr の検索機能を利用した .

名詞 250 語と形容詞 100 語の合計 350 語について, Flickr 上で検索を行い, 位置情報付き画像を 2000 枚ずつ収集した . 実験に用いた単語を表 1, 表 2 に示す . 画像を収集するために Flickr が提供している WebAPI [12] を利用した .

Flickr では, 一部のユーザーが非常に類似した画像を大量に投稿していることがある . そのため画像を収集するに当たって, 同一ユーザーが投稿した画像の枚数を制限した .

5.2 VisualRank の算出

画像を収集したのち, 単語ごとに VisualRank を算出した . 提案手法で VisualRank を求める際に与える注目点パラメータには, 世界の 10 都市の座標を用いた . 画像特徴量のカラーヒストグラムは RGB 空間での 64 次元とし, SIFT 記述子による BoK 表現は 500 次元とした .

画像特徴量の混合パラメータ β や, 位置情報によるバイア

表 1 250 語の名詞

africa, airplane, alexander, alligator, america, ant, arc de triomphe, arm, asia, bach, backpack, banana, barbeque, battle, beach, bear, beauty, beaver, bee, beer, beetle, big ben, board, boat, bob, book, box, bread, brother, buddha, bug, building, burger, bus, butterfly, cactus, cake, california, canada, candy, canoe, car, castle, cat, cedar, chair, chalk, chicken, china, circle, city, coffee, coke, color, computer, cookie, coral, crow, dandelion, daughter, desert, desk, dessert, deutschland, dice, dish, doctor, dog, dolphin, dragon, dragonfly, dream, duck, eagle, edison, eel, egg, egypt, eiffel tower, election, elephant, elevator, erica, europe, face, father, fern, field, fireworks, fish, flea, flower, fly, fork, france, frog, fruit, game, gates, giraffe, goat, goose, gorilla, grape, grass, grasshopper, gun, half, ham, hawk, head, height, helicopter, hibiscus, hornet, horse, hospital, house, ice cream, india, insect, italia, ivy, japan, jellyfish, jump, kangaroo, kayak, lamp, lavender, lawn, leaf, leg, lemon, level, library, light, lincoln, line, lion, lizard, love, machu picchu, mangrove, manta, mantis, marriage, mars, milk, mint, monkey, moon, mosquito, moss, moth, mother, mountain, mouse, mozart, museum, mushroom, napoleon, new york, niagara, octopus, olive, owl, oyster, palm, paris, park, parrot, party, pasta, pen, penguin, people, phone, pine, pizza, plant, playstation, pool, pope, potato, president, pride, pyramid, rabbit, rainbow, rice, rome, rose, salad, salmon, salt, santa claus, school, sea, shakespeare, shark, ship, shrimp, sister, sky, skyscraper, snail, snake, socks, son, sound, spider, sport, square, starfish, statue of liberty, steak, sugar, sun, sushi, swan, sword, tea, teacher, temple, test, thomas, tiger, toad, tokyo, tool, town, train, tripod, tuna, turtle, uluru, usa, valley, village, watch, waterfall, wave, whale, wii, wine, worm, xbox, zoo

表 2 100 語の形容詞

aerial, ancient, antique, bad, beautiful, best, better, big, black, blue, botanical, bottom, bright, brown, cherry, classic, clean, clear, cold, colourful, concrete, cool, crazy, cute, dark, digital, dry, electric, empty, famous, female, first, general, good, grand, gray, great, green, happy, hard, heavy, high, historic, holy, hot, human, iced, interior, international, large, latest, long, male, medieval, military, mobile, modern, more, most, national, natural, nautical, new, nice, old, open, orange, outdoor, pink, present, public, purple, rainy, red, rich, rural, rusted, scenic, second, sexy, short, small, special, strong, sunny, sweet, top, traditional, tropical, twin, underwater, urban, vintage, warm, welcome, white, wide, wild, wooden, yellow

表 3 実験で用いた注目点の都市名と座標（北緯・東経が正值）

都市名	緯度	経度
tokyo	35.689506	139.691701
beijing	39.904667	116.408198
sydney	-33.867139	151.207114
delhi	28.635308	77.22496
cairo	30.064742	31.249509
paris	48.8566667	2.3509871
cape town	-33.9237762	18.4233455
new york	40.714269	-74.005973
san francisco	37.7749295	-122.4194155
rio de janeiro	-22.9035393	-43.2095869

表 4 実験で用いたパラメータ値

	0.80	0.00
	0.85	0.25
α	0.90	β 0.50
	0.95	0.75
	1.00	1.00

スの強さを決定するパラメータ α はそれぞれ 5 種類の値を試した。

実験で使用した注目点とパラメータの値を、表 3、表 4 に示す。

6. 結果

本研究では、350 語の単語について提案手法を用いて実験を行った。それらの結果は、<http://mm.cs.uec.ac.jp/kawaku/geovisualrank/> にて閲覧できるようにする予定である。

ここでは結果の一部を示す。

図 2、図 3 に 'house' をクエリとした場合の実験結果を示す。図 2 は位置情報によるバイアスを使わずに、類似度のみでランク付けを行った結果である。ヨーロッパやアメリカで撮影され

た西洋の家の画像が上位に来ている。

一方、図 3、図 3 は位置情報によるバイアスを用いて注目点に sydney を設定した場合の結果である。シドニー付近で撮影された家の画像のとシドニー・オペラハウスの画像が上位 10 枚に入っており、注目点付近で撮影された画像の代表画像が得られている。パラメータ α によってバイアスの強さを設定することで、地域性と視覚の世界的代表さのどちらを重視するかを調整することができる。図 4 は図 3 に比べて視覚の代表さをより優先した場合の結果であり、上位 10 枚の中に日本で撮影された画像が 1 枚入っている。

図 5、図 6、図 7 は 'pyramid' をクエリとした結果である。図 5 はカイロを注目点とした結果であり、エジプトのピラミッドの画像が上位画像になっている。図 6 は注目点がパリであり、ルーブル美術館の前にあるピラミッド型の建築物が上位に表示される。図 7 はリオデジャネイロを注目点としているもののメキシコのピラミッドが主な上位画像となっている。これは、リオデジャネイロ付近には 'pyramid' に対応する画像が無く、エジプトよりもメキシコの方がリオデジャネイロに近いためであると考えられる。図 8、図 9、図 10、図 11 の図は、'traditonal' に関する結果であり、各地の民族的な人物画像が上位画像となっている。

図 12 から図 14 は 'napoleon' の結果である。パリでは歴史上のナポレオンに関する画像が得られシドニーではナポレオンフィッシュの画像が得られている。シドニーの結果においては BoF のみを使った方が上位にナポレオンフィッシュの画像が多い。注目点によって、代表画像に含まれる物体の種類が大きくなることのある都分かる。そのため今後、最適な特徴量やパラメータを自動的に選択する手法を提案手法へ取り入れる場合には、単語と注目点の組に関して求めるようにすべきと考えられる。

6.1 テキスタグによる類似度を用いた場合

Flickr 上の画像にはユーザによってテキスタグが付与されている。タグに基づく類似度を用いた実験を追加して行い、結果を求めた。

タグベースの類似度は以下のようにして算出した。

step.1 クエリに関する画像に付与されることが多かった上位 500 種類のタグをコードブックとする。ただし、クエリ自体はコードブックに入れない。

step.2 各画像はコードブックに関する 500 次元のベクトルで表現する。各次元に対応するタグが付与されていれば、その次元の値を 1 とし、付与されていなければ 0 とする。

step.3 step.2 で求めたベクトルのコサイン類似度を画像間の類似度とする。

図 1 はタグによる類似度を使った場合の結果である。類似度以外のパラメータは、図 13 や図 14 と同じである。タグによる類似度を使うと上位画像ほとんどが魚の画像となった。上位画像に付与されていたタグを確かめたところ、"fish", "sea", "underwater" などのタグが共通していた。タグは人手でつけられたものであり、内容に直接関係がある場合が多く、検索結果を向上させるために有用であると考えられる。そのため今後

は、視覚特徴だけでなくタグも活用して類似度を求めることを検討している。



図 1 タグによる類似度を使用した場合．'napoleon' での上位画像と分布（注目点：sydney . $\alpha = 0.85$ ）

7. おわりに

本論文では，VisualRank [3] を用いた位置情報付き画像のランキング手法を提案した．提案手法では，VisualRank での類似度の計算にカラーヒストグラムと，SIFT 記述子による Bag-of-Features 表現を用いた．また，パラメータとして与えた地理的座標の近くで撮影された画像を強調するバイアスペクトルを用いた．

提案手法を用いて 350 語の単語について実験を行った．バイアスペクトルによって注目点付近で取られた画像のランクが上がることや，注目点を変化させることで結果が変わることを確認した．

今後の課題として，提案手法によるランキングの定量的評価，他手法との比較，適切なパラメータの自動選択を可能にすることが挙げられる．

文 献

- [1] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol.60, no.2, pp.91–110, 2004.
- [2] G. Csurka, C. Bray, C. Dance, and L. Fan, "Visual categorization with bags of keypoints," Proc. of ECCV Workshop on Statistical Learning in Computer Vision, pp.59–74, 2004.
- [3] Y. Jing and S. Baluja, "Visualrank: Applying pagerank to large-scale image search," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.30, no.11, pp.1870–1890, 2008,.
- [4] 柳井啓司, "一般物体認識の現状と今後," 情報処理学会論文誌: コンピュータビジョン・イメージメディア, vol.48, no.SIG16 (CVIM19), pp.1–24, 2007.
- [5] 川久保秀敏, 柳井啓司, "Bag-of-features 表現を用いたエントロピーによる単語の視覚性の分析," 情報処理学会研究会報告: コンピュータビジョン・イメージメディア研究会, vol.2009, no.29, pp.267–274, 2009.
- [6] H. Kawakubo and K. Yanai, "An analysis of the relation between visual concepts and geo-locations using geotagged images on the web," IEEE ICME WS on Internet Multimedia Search and Mining, pp.1644–1647, 2009.
- [7] X. He, W.Y. Ma, and H. Zhang, "ImageRank: spectral techniques for structural analysis of image database," IEEE International Conference on Multimedia and Expo, pp.25–28, 2003.
- [8] S. Brin and L. Page, "The anatomy of a large-scale hyper-textual Web search engine," Computer networks and ISDN

systems, vol.30, no.1-7, pp.107–117, 1998.

- [9] 安倍 満, 吉田悠一, "Visualrank の多クラスへの拡張: 画像特徴量を用いた類似画像の自動分類とランキング付け手法," 電子情報通信学会技術研究報告, PRMU2008-178, pp.183–188, Dec. 2008.
- [10] M. Ambai and Y. Yoshida, "Multiclass VisualRank: image ranking method in clustered subsets based on visual features," 2009.
- [11] "Flickr," <http://www.flickr.com/>.
- [12] "Flickr API," <http://www.flickr.com/services/api/>.



図 2 'house' での上位画像 10 枚と、上位 100 枚の分布 (位置情報によるバイアス無し, color50%+BoF50%)



図 3 'house' での上位画像 10 枚と、上位 100 枚の分布 (注目点:sydney . $\alpha = 0.85$, color50%+BoF50%)



図 4 'house' での上位画像 10 枚と、上位 100 枚の分布 (注目点:sydney . $\alpha = 0.95$, color50%+BoF50%)



図 5 'pyramid' での上位画像 10 枚 (注目点: cairo . $\alpha = 0.85$, color50%+BoF50%)



図 6 'pyramid' での上位画像 10 枚 (注目点: paris . $\alpha = 0.85$, color50%+BoF50%)



図 7 'pyramid' での上位画像 10 枚 (注目点: rio de janeiro . $\alpha = 0.85$, color50%+BoF50%)



図 8 'traditional' での上位画像 10 枚 (注目点: tokyo . $\alpha = 0.85$, color50%+BoF50%)



図 9 'traditional' での上位画像 10 枚 (注目点: sydney . $\alpha = 0.85$, color50%+BoF50%)



図 10 'traditional' での上位画像 10 枚 (注目点: rio de janeiro . $\alpha = 0.85$, color50%+BoF50%)



図 11 'traditional' での上位画像 10 枚 (注目点: delhi . $\alpha = 0.85$, color50%+BoF50%)



図 12 'napoleon' での上位画像 10 枚 (注目点: paris , $\alpha = 0.85$, color50%+BoF50%)



図 13 'napoleon' での上位画像 10 枚 (注目点: sydney . $\alpha = 0.85$, color50%+BoF50%)



図 14 'napoleon' での上位画像 10 枚 (注目点: sydney . $\alpha = 0.85$, BoF のみ)