

# 位置情報を考慮した VisualRank による地域別代表画像の選出

川久保秀敏<sup>†</sup> 柳井 啓司<sup>†</sup>

<sup>†</sup> 電気通信大学 電気通信学研究科 情報工学専攻

〒 182-8585 東京都調布市調布ヶ丘 1-5-1

E-mail: <sup>†</sup>kawaku-h@mm.cs.uec.ac.jp, <sup>††</sup>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、代表的な視覚特徴を持ち、ランク付け時に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。提案手法での VisualRank では、画像の視覚特徴を用いて類似度行列を作成し、位置情報からバイアスペクトルを作成する。多様なクエリ単語に対応するため、画像の類似度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] から求める。また、バイアスペクトルを作成するために、ランキングを求める際に地理上の一点を注目点として指定する。バイアスペクトルでは、指定した注目点の近くで撮影された画像ほど大きい値が与えられる。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 上からそれぞれ 2000 枚ずつの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。ランク付け時に指定する注目点には世界の 10 の都市の座標を用いた。その結果、世界の各地域に関する代表画像ランキングが得られた。キーワード VisualRank, 位置情報付き画像, web 画像, 画像ランキング

## Selecting representative photographs using VisualRank for Geotagging images

Hidetoshi KAWAKUBO<sup>†</sup> and Keiji YANAI<sup>†</sup>

<sup>†</sup> Department of Computer Science, The University of Electro-Communications

1-5-1 Chofugaoka, Chofu-shi, Tokyo, 182-8585, Japan

E-mail: <sup>†</sup>kawaku-h@mm.cs.uec.ac.jp, <sup>††</sup>yanai@cs.uec.ac.jp

**Abstract** In this paper, we propose a new method to rank geotagged images which is an extension of the VisualRank method. The proposed method generates ranking of geotagged images considering both visual similarity and proximity of locations where photos are taken. In the proposed method, a similarity matrix is constructed based on visual features, and a bias vector is generated based on the degree of the proximity of images to the given reference points. The difference to the original VisualRank is that our method takes into account the proximity to the given reference locations by giving larger/smaller scores to the images taken at the places which are close to the reference locations. In the experiments, we prepared seven hundred thousand images gathered from Flickr regarding 350 words including 250 nouns and 100 adjectives. As features to calculate the similarity matrix, we used the bag-of-features representation, a color histogram and tag textual features. In the experiments, we obtained representative images regarding various reference cities, which cannot be obtained by VisualRank and other existing methods. From the experimental results, we confirmed the effectiveness of the proposed methods.

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

### 1. はじめに

近年、計算機の性能が向上し、これまでテキストベースの手法が主であった画像検索の分野で、画像の視覚特徴を用いる手法が研究されている。Jing らが提案した VisualRank [3] は、画像の類似度行列に PageRank アル

ゴリズムを適用した手法である。Jing らは VisualRank を、商品画像やランドマーク画像に対して用いる実験を行っている。

一方、web 上のアルバムサービスや GPS 機能付きカメラの普及に伴い、撮影地点の緯度経度情報がメタデータとして付与された画像が web 上に増加した。オンライ

ンアルバムサービスの Flickr [4] 上では、大量の位置情報付き画像が公開されており、それらの画像に対して検索を行うことが可能である。

本論文では、VisualRank を位置情報付き画像向けに用いて、視覚特徴と位置情報の両方を考慮した画像ランキング手法を提案する。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。提案手法によるランキングでは、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なる。

提案手法は商品やランドマーク以外の画像に対しても適用することを目指している。そのため画像の類似度を求めるための特徴量として、カラーヒストグラムと SIFT 記述子 [1] による Bag-of-Features 表現 [2] を用いた。また、ランキングを行う際に指定した地点との距離を考慮するために、位置情報にもとづくバイアスペクトルを用いる。

本論文では、名詞 250 語と形容詞 100 語について Flickr [4] 上から位置情報付き画像を収集し、提案手法によって代表的画像のランキングを求めた。

## 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」[5] に生かすことを目指す研究がある [6], [7]。

「一般物体認識」とは、制限の無い実世界画像に対して、画像中の物体を認識することである。一般物体認識の目的の一つとして、画像に対して内容を意味する単語を付与することがある。計算機による一般物体認識では、画像の視覚特徴量やメタデータといった情報をもとに付与する単語を決定する。精度の高い一般物体認識を実現するためには、学習データセットの質が重要である。多様な単語概念について、「代表的な視覚特徴」を示す画像データセットが望ましい。しかし単語概念に対応する

「代表的な視覚特徴」は地域や文化によって、異なる場合がある。例えば、日本の家とヨーロッパの家、アフリカの家、エスキモーの家は、それぞれ見た目が大きく異なり、「家」に対応する画像特徴量は地域によって異ってくる。視覚に地域差のある「家」などの単語概念では、地域別のデータセットを作成することで、認識精度の向上を図れる。

本研究は、一般物体認識のための地域別データセットの自動作成に役立てることを目的としている。地域別代表画像の選出を自動で行えば、人手で行うよりも、多くの概念についてデータセットを作成することが可能になる。

## 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 [8] 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 [9] 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り、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 [10], [11]。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

VisualRank は PageRank アルゴリズム [9] を画像に適用したものである。PageRank は web ページのランク付けを行うためのアルゴリズムであり、まず web ページ間のリンク構造を行列で表現する。そして反復計算によって各ページのランク値を求める。より多くのより重要なページからリンクされているページが上位にランク付けされる。

VisualRank では、リンク構造を表現する行列の代わ

りに、画像の類似度行列を用いる。VisualRank の最も基本的な求め方は、式 (1) に示す更新式を反復することである。式 (1) では、各画像の VisualRank 値を並べた列ベクトル  $R$  を、類似度行列の各列を正規化した行列  $S$  にかけている。VisualRank の前に重要な画像が予想可能な場合は、重要と予想される画像の VisualRank 値の初期値を大きくすることが可能である。

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

更新時に VisualRank 値の合計が変化しないようにするために、 $S$  の各列は値の合計が 1 になるように正規化する。式 (1) は、べき乗法の更新式と見ることができ、反復計算によって  $R$  は  $S$  の絶対値最大固有値に対応する固有ベクトルへと収束する。

類似度行列の大きさは画像数の 2 乗となるため、計算量を抑えるためにはランク付けの対象となる画像の数を減らす必要がある。対象画像を絞り込むためには、コンテンツベースな手法よりも高速な、テキストベースな手法が用いられる。テキストベースな画像検索での上位画像を対象画像とし、それらの再ランク付けに VisualRank を用いるようにすることで、計算量を抑えることができる。また、ランク値の初期値を画像によって変え、Jing ら [3] は、テキストベースな画像検索での上位 1000 枚を、VisualRank を用いて再ランク付けしている。

式 (1) に示した更新式に、補正ベクトル  $P$  を加えたものが式 (2) である。反復ステップごとに、式 (1) での結果と  $P$  が合成される。

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

PageRank での補正ベクトル  $P$  は、リンクを使わずに web ページにアクセスすることをモデル化するためのものである。VisualRank でも更新ステップごとに補正をかけるために、 $P$  が用いられる。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。更新時に VisualRank 値の合計が変化しないようにするために、 $P$  の値の合計は  $R$  のものと等しくする。補正の強度はパラメータ  $\alpha$  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に対して 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。

提案手法での VisualRank では、反復計算に式 (2) の更新式を用いる。画像の視覚特徴を用いて類似度行列を作成し、位置情報からバイアスペクトルを作成する。VisualRank 値の初期値は全ての画像について等しく 1 とした。バイアスペクトルを用いることで、結果に対する視覚特徴と位置情報の影響の強さを 1 つのパラメータで調整できる。

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注目点に近い画像に大きな値が与えられるように作成する。類似度計算とバイアスペクトルの作成についてのそれぞれの説明を以下に述べる。

##### 4.1 画像の類似度

[3] では、画像の類似度を求めるために、局所特徴量である 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も対応するため、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、幅広いカテゴリの画像を対象とした場合の表現性の高さから、一般物体認識の分野で良く使われているものである。

類似度計算には計算量の少ない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。 $\beta$  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 (バイアスペクトル) を用いた。バイアス

表 1 250 語の名詞

africa, airplane, alexander, alligator, america, ant, arc de triomphe, arm, asia, bach, backpack, banana, barbecue, 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, tulip, tuna, turtle, uluru, usa, valley, village, watch, waterfall, wave, whale, wii, wine, worm, xbox, zoo
--

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))。要素の合計値を  $R$  と同じにすることで、反復計算時に VisualRank 値の合計が変化しないようにしている。 $D_i$  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))。注目点に近い地点で撮影された画像ほど、 $D_i$  が小さくなり、 $P_i^{geo}$  が大きくなる。

$$p_i^{geo} = 1 - \frac{D_i}{\pi} \quad (4)$$

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

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

$lat_i, long_i$  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。

複数のバイアスベクトルの平均を求めることで、複数の地域に関するバイアスベクトルを作成することが出来る。例えば、アメリカに関するバイアスとヨーロッパに関するバイアスを平均することで、欧米で撮影された画像を優先するようなバイアスベクトルを作成することが可能である。

また、式 (4) の代わりに式 (7) を用いることで、注目点から遠い地点で撮影された画像を優先するネガティブなバイアスベクトルを作成することが可能である。

$$p_i^{geo} = \frac{D_i}{\pi} \quad (7)$$

## 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 [4] 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。名詞は、身の回りの物体の名称、生物名、固有名詞などを

人手で設定した。形容詞は、単語概念と視覚特徴の関係性を調査した、柳井らの研究 [12], [13] を参考にして設定した。画像を収集するために Flickr が提供している webAPI [14] 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  $\beta$  や、位置情報によるバイアスの強さを決定するパラメータ  $\alpha$  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。

ここでは結果の一部を示す。

### 6.1 'house' についての結果

図 1、図 2、図 3 に 'house' をクエリとした場合の実験

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
$\alpha$	0.90	$\beta$ 0.50
	0.95	0.75
	1.00	1.00

結果を示す。図 1 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 2、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  $\alpha$  によってバイアスの強さを設定することで、地域性と視覚的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 3 は図 2 に比べて視覚的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。

## 6.2 注目点によって上位画像に変化が見られた例

図 4、図 5、図 6 は 'pyramid' をクエリとした結果である。図 4 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 5 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 6 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' に対応する代表的画像が無く、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。今後、上位画像と注目点の距離を評価することで結果の妥当性を評価し、パラメータの自動調整に活用することが考えられる。

図 7、図 8、図 9、図 10 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 11、図 12、図 13 は 'rainbow' をクエリとした結果である。図 11、図 12 では、それぞれの地域で撮影された虹の写真が上位画像となっている。図 13 は注目点を東京とした場合の結果であり、虹の画像の他にレインボーブリッジの画像が上位画像に入っている。クエリ単語自体が示す物体に地域性が乏しかったとしても、特定地域で固有名詞と結びつき地域性を持つ場合があると分かる。

図 14、図 15、図 16、図 17 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ことなることがあると分かる。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。

## 6.3 テキスト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ト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の多かった上位 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



- 1~10位
- 11~25位
- 26~50位
- 51~100位

図 1 '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%）



- 1~10位
- 11~25位
- 26~50位
- 51~100位

図 2 '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%）



- 1~10位
- 11~25位
- 26~50位
- 51~100位

図 3 'house' での上位画像 10 枚と，上位 100 枚の分布（注目点：sydney,  $\alpha = 0.95$ ，color50%+BoF50%）

像間の類似度とする。

図 18, 図 19 は 'house' をクエリとしてタグベースの類似度を用いた結果である。タグのみによる類似度であるため、視覚特徴を用いた場合よりも多様な見た目の画像が上位画像となっている。

図 18 はバイアスを用いなかった場合の結果である。上位画像に付与されていたタグには 'architecture', 'sky', 'clouds' などといった画像内容を示すものがあつた一方で、撮影に使われたカメラのメーカー名といった画像内容に関係のないタグもあつた。今回は、クエリの関する画像に多く付与された上位 500 種類のタグを使用した。今後、タグと画像内容の関係性を考慮した使用タグの選別を行うようにできれば、結果の向上が図れる物と考える。図 19 はシドニーを注目点とした場合の結果である。視覚特徴量による類似度を用いた場合と同様に、シドニー・オペラハウスの画像が上位に入っている。

タグは人手によって付与された物であり、画像内容を示すタグは結果を向上させるために有用であると考えられる。Flickr では各ユーザがタグを付与するため、ユーザごとにタグの付与の仕方が異なっている。画像内容を表すタグを少数付与するユーザもいれば、撮影位置情報やカメラ情報などのメタデータを含んだ多数のタグを付与するユーザも存在する。こうした多様な基準で付与されたタグ情報を、類似度計算により良く活用するのが今後の課題である。

#### 6.4 バイアスのバリエーション

図 20 は 'insect' をクエリとし、シドニー・デリー・ケープタウンの 3 都市についてのバイアスペクトルの平均ベクトルを用いた場合の結果である。インド洋を囲む地域で撮影された画像が上位画像となっている。

図 21 は 'arc de triomphe' をクエリとしパリについてネガティブなバイアスペクトルを用いた場合の結果である。パリから遠い地点で撮影された凱旋門の画像として、ラオスのパトゥーサイの画像や、中国やラスベガスにある凱旋門を模した建築物の画像が選ばれ上位画像となっている。

#### 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. Csuska, 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] "Flickr," <http://www.flickr.com/>.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] S. Brin and L. Page, "The anatomy of a large-scale hypertextual Web search engine," Computer networks and ISDN systems, vol.30, no.1-7, pp.107-117, 1998.
- [10] 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.
- [11] M. Ambai and Y. Yoshida, "Multiclass VisualRank: image ranking method in clustered subsets based on visual features," Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval ACM, pp.732-733 2009.
- [12] 柳井啓司, K. Barnard, "一般物体認識のための単語概念の視覚性の分析," 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.SIG10 (CVIM17), pp.88-97, 2007.
- [13] K. Yanai and K. Barnard, "Image region entropy: A measure of "visualness" of web images associated with one concept," Proc. of ACM International Conference Multimedia, pp.420-423, 2005.
- [14] "Flickr API," <http://www.flickr.com/services/api/>.



図 4 'pyramid' での上位画像 10 枚  
(注目点: cairo,  $\alpha = 0.85$ , color50%+BoF50%)



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



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



図 7 'traditional' での上位画像 10 枚  
(注目点: tokyo,  $\alpha = 0.85$ , color50%+BoF50%)



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



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



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



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



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



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



図 14 'napoleon' での上位画像 10 枚  
(注目点: paris,  $\alpha = 0.85$ , color50%+BoF50%)



図 15 'napoleon' での上位画像 10 枚  
(注目点: sydney,  $\alpha = 0.85$ , color50%+BoF50%)



図 16 'napoleon' での上位画像 10 枚  
(注目点: sydney,  $\alpha = 0.85$ , color のみ)



図 17 'napoleon' での上位画像 10 枚  
(注目点: sydney,  $\alpha = 0.85$ , BoF のみ)



図 18 'house' での上位画像 10 枚  
(位置情報によるバイアス無し, tag のみ)



図 19 'house' での上位画像 10 枚  
(注目点: sydney,  $\alpha = 0.85$ , tag のみ)



図 20 3つの都市についてのバイアスを混合した場合  
(クエリ: insect, 注目点: sydney, delhi, cape town,  
 $\alpha = 0.85$ , color50%+BoF50%)



図 21 注目点に関してネガティブなバイアスを用いた場合  
(クエリ: arc de triomphe, 注目点: paris,  
 $\alpha = 0.85$ , color50%+BoF50%)

