

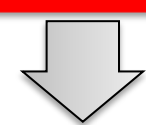
IS1-54 位置情報を考慮したVisualRankによる地域別代表画像の選出

電気通信大学大学院 情報工学専攻 川久保秀敏, 柳井啓司

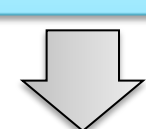
概要

- 画像の撮影位置情報の活用

単語に関する, 地域ごとの代表画像を得る



位置情報付き画像のデータセットの自動作成



認識分野や自動アノテーションへの活用

手法

VisualRank

- PageRankアルゴリズムを画像に適用したもの
- Jingらによって提案された
 - Y.Jing and S.Baluja, "VisualRank: Applying pagerank to large-scale image search", IEEE Trans PAMI, 2008

- リンク構造の代わりに, 画像の類似関係を利用



VisualRankの更新式

- 更新式を反复し, Visualrank値を収束させる

$$R \leftarrow \alpha(S \times R) + (1-\alpha)P \quad (0 \leq \alpha \leq 1)$$

$$\alpha \begin{bmatrix} \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ \\ \circ & \circ & \circ & \circ \end{bmatrix} \times \begin{bmatrix} \circ \\ \circ \\ \circ \\ \circ \end{bmatrix} + (1-\alpha) \begin{bmatrix} \circ \\ \circ \\ \circ \\ \circ \end{bmatrix}$$

S R P

R : 各画像のVisualRank値が並んだベクトル

S : 画像の類似度行列を, 列ごとに正規化したもの

P : 補正ベクトル. 各画像への補正値が並ぶ.

α : 補正の強さを調整するパラメータ

- VisualRankを用いた位置情報付き画像のランキング手法を提案

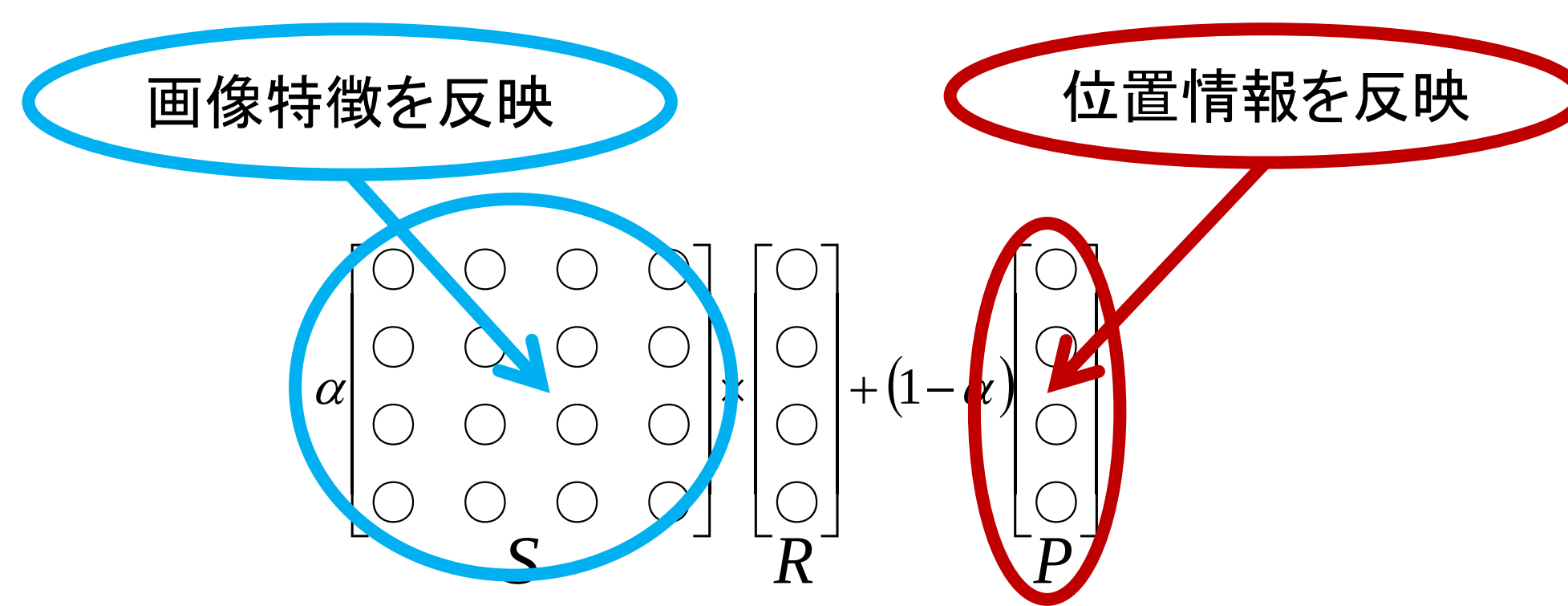
- 画像特徴と位置情報の両方を考慮してランク付けを行なう

- 上位にランキングされる画像は,
 - 代表的な視覚特徴を持つ
 - 撮影地点による評価点が高い

- 画像特徴と位置の重みをパラメータで調整可能

提案手法

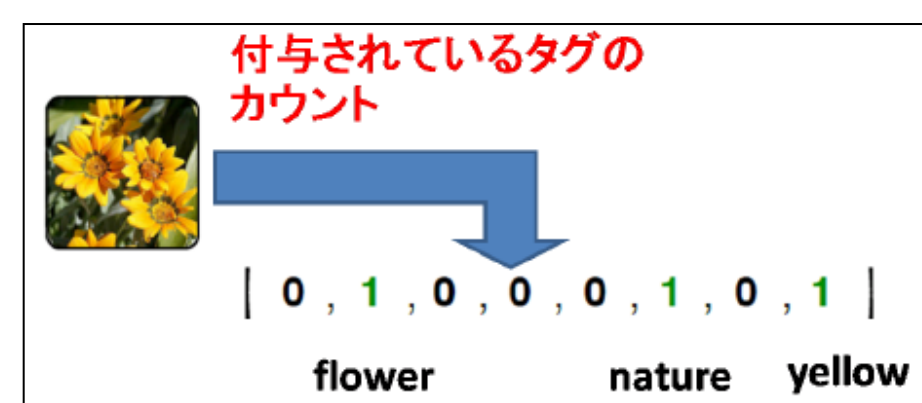
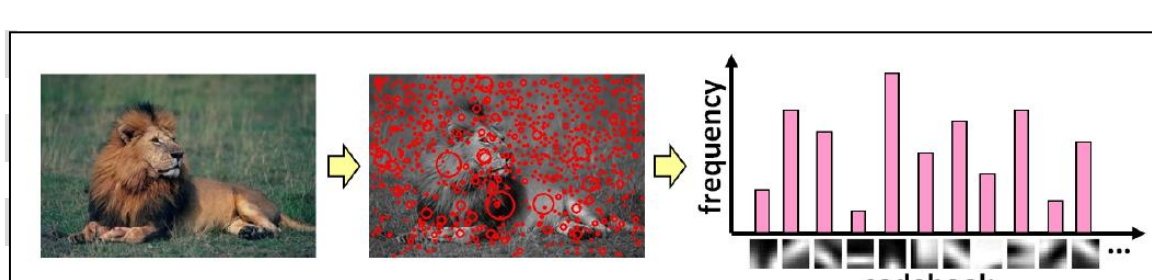
- 類似度行列 S を画像特徴から生成する
- 補正ベクトル P を位置情報から生成し, 不均一なバイアスペクトルとして利用する



画像の特徴量

- JingらはSIFT特徴点のマッチ数を類似度を使用
- 同一物体がよく撮影される, 商品画像やランドマーク画像に適しているものの, 一般的な画像セットには不向き

- 64次元のカラーヒストグラム (RGB色空間)
- 500次元のBag-of-Features表現 (SIFT記述子)
 - 類似度はヒストグラムインタセクションで計算
- 500次元のBag-of-Tags表現
 - 類似度はコサイン類似度で計算



世界各地域の house 画像

アジア



アメリカ

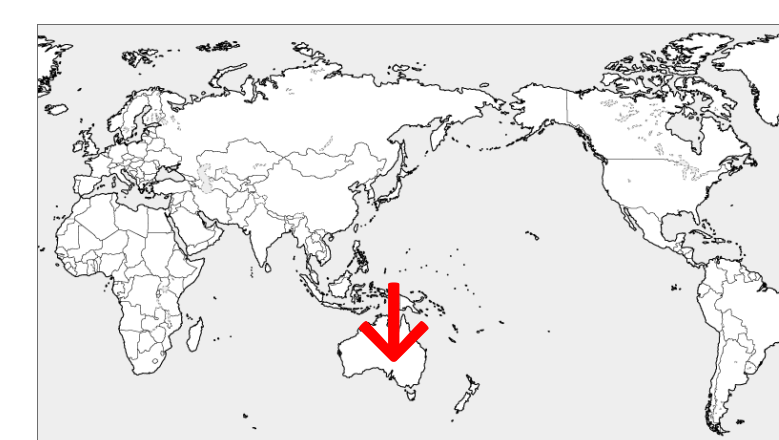


オーストラリア



位置情報によるバイアスペクトルの生成

- ランキング生成時にユーザが注目する地点の座標を注目点パラメータとして与える。



- 単純なバイアスペクトルとして, 注目点と撮影地点の距離を用いるものを試した
 - 球面三角法によって距離を求めた
 - 地球を半径1の球とする

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

D_i : 求める距離

$lat_i, long_i$: 画像 i の撮影位置情報
 $lat_A, long_A$: 注目点の座標

注目点に近い画像を優先 (ポジティブなバイアス)

$$p_i^{geo} = \frac{\pi - D_i}{\pi}$$

注目点から遠い画像を優先 (ネガティブなバイアス)

$$p_i^{geo} = \frac{D_i}{\pi}$$

バイアスペクトルは正規化して使用する

$$P_i^{geo} = \frac{p_i^{geo}}{\|P_i^{geo}\|_1} \|R\|_1$$

実験・結果

実験

- クエリ単語: 名詞250語と形容詞100語
- 各クエリについてFlickr画像を2000枚収集
 - 同一ユーザの投稿画像は10枚まで
- 注目点パラメータには, 以下の10都市の座標を使用

東京	リオデジャネイロ
北京	パリ
シドニー	カイロ
サンフランシスコ	ケープタウン
ニューヨーク	デリー

結果例: 各地のpyramid画像

- $\alpha=0.85$, 特徴量: color, BoF

- 注目点: カイロ エジプトのピラミッド



- 注目点: パリ ルーブル美術館のピラミッド型建物



- 注目点: ニューヨーク



- 注目点: 東京



traditional

- シドニー



- 東京



- リオデジャネイロ



game

- 東京



- ケープタウン



パラメータ α による変化

- house, 注目点: シドニー



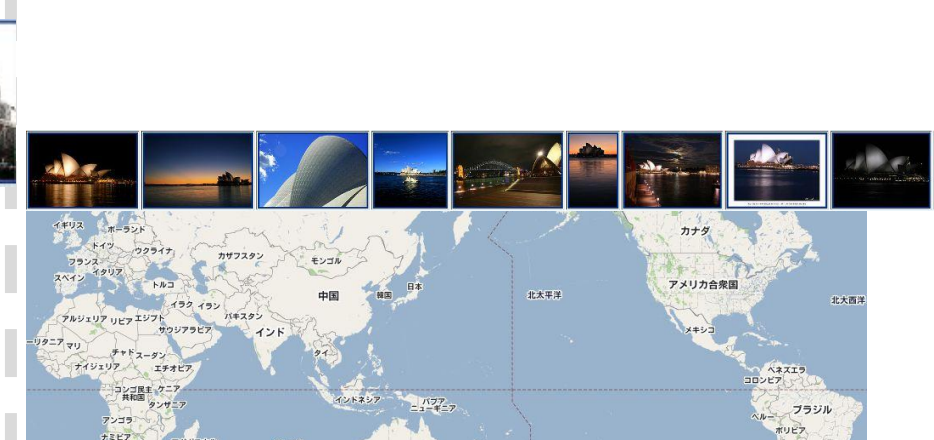
- $\alpha=1.0$
- 通常のVisualrank



- $\alpha=0.90$



- $\alpha=0.80$



- $\alpha=0.0$
- 注目点に近い順



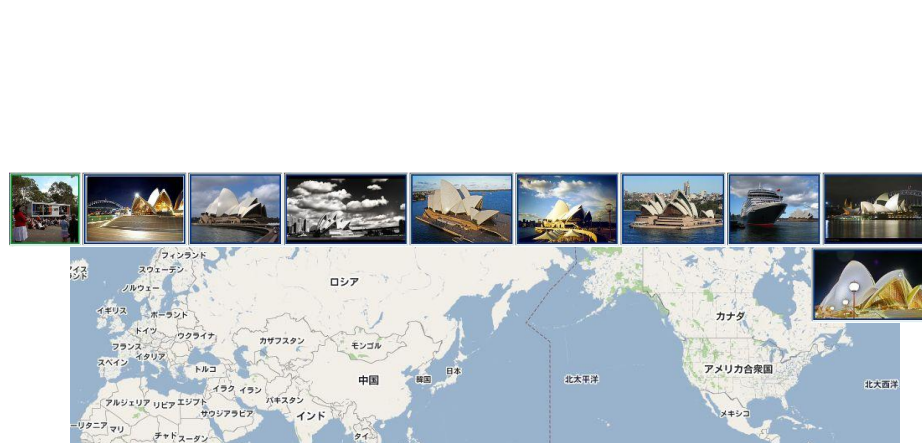
- 1~10位
- 11~25位
- 26~50位
- 51~100位



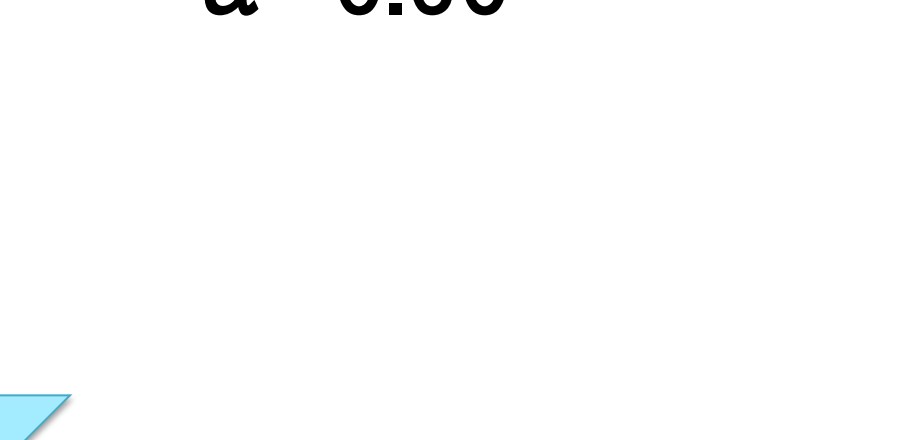
- $\alpha=0.95$



- $\alpha=0.85$



- $\alpha=0.50$



視覚類似度とタグ類似度

- cat 注目点: 東京

タグでは "pets" "cute" などが共通したペット画像が主になり室内画像が主



視覚では 背景部分の類似性から, 室外画像が主

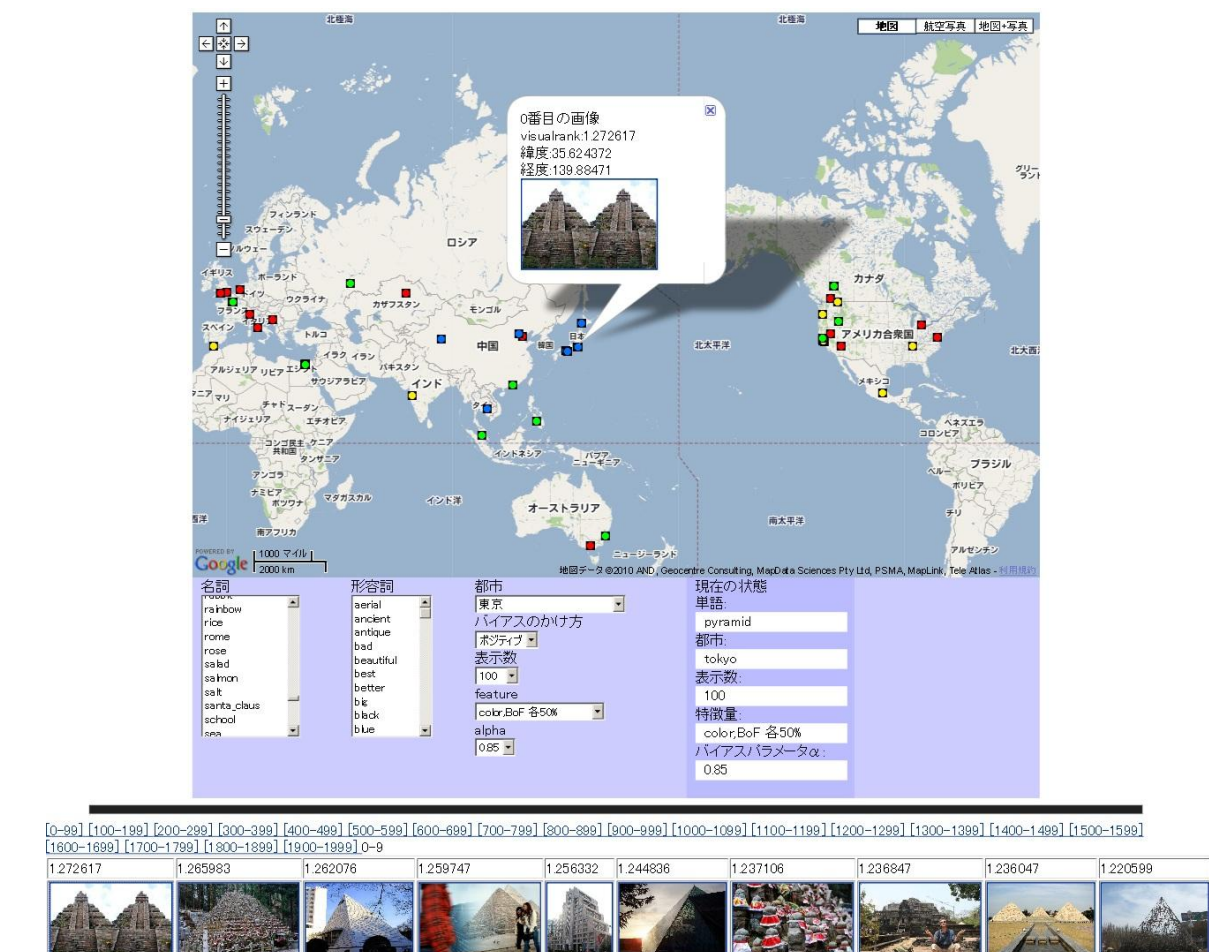


タグ+視覚では, 屋内でのペット画像と, 屋外での撮影画像の両方が上位画像に入り, 上位画像の撮影シーンの多様性が向上



実験結果の公開

<http://mm.cs.uec.ac.jp/kawaku/geovisualrank/>



今後

- 提案手法の評価
 - 被験者に評価させる (Amazon Mechanical Turk ?)
- 他手法との比較
 - 位置情報での絞り込んだ画像でのVisualRank
 - テキスト検索時のクエリに地域名を足した場合
 - 位置情報によるバイアスの作成方法を変えたもの
- 評価・比較をもとに, 提案手法の改良を行ない, 位置情報付き画像データセットの自動作成へ活用