

An aerial photograph of a city, likely Tokyo, taken during sunset or sunrise. The sky is filled with soft, golden light and scattered clouds. The city below is densely packed with buildings of various heights and colors, with a prominent skyscraper visible in the distance. The overall atmosphere is calm and scenic.

マルチカーネル学習を用いた 画像特徴と航空写真特徴の 重要度の推定

電子情報通信学会

パターン認識・メディア理解研究会2009

電気通信大学大学院 情報工学専攻

八重樫恵太 柳井啓司

2009年11月27日

背景

大量の位置情報付き画像

様々な画像認識手法



撮影位置のついた写真：
世界で4000万枚以上
1ヶ月で10万枚増加



写真を閉鎖する
効率的にカ

文を画像認識し
する方法はな



航空写真の利用

- 認識対象 ⇔ 位置情報：密接な関係
– Ex. 海岸

~~非効率~~

~~世界中の海岸の
位置を学習?~~

撮影位置

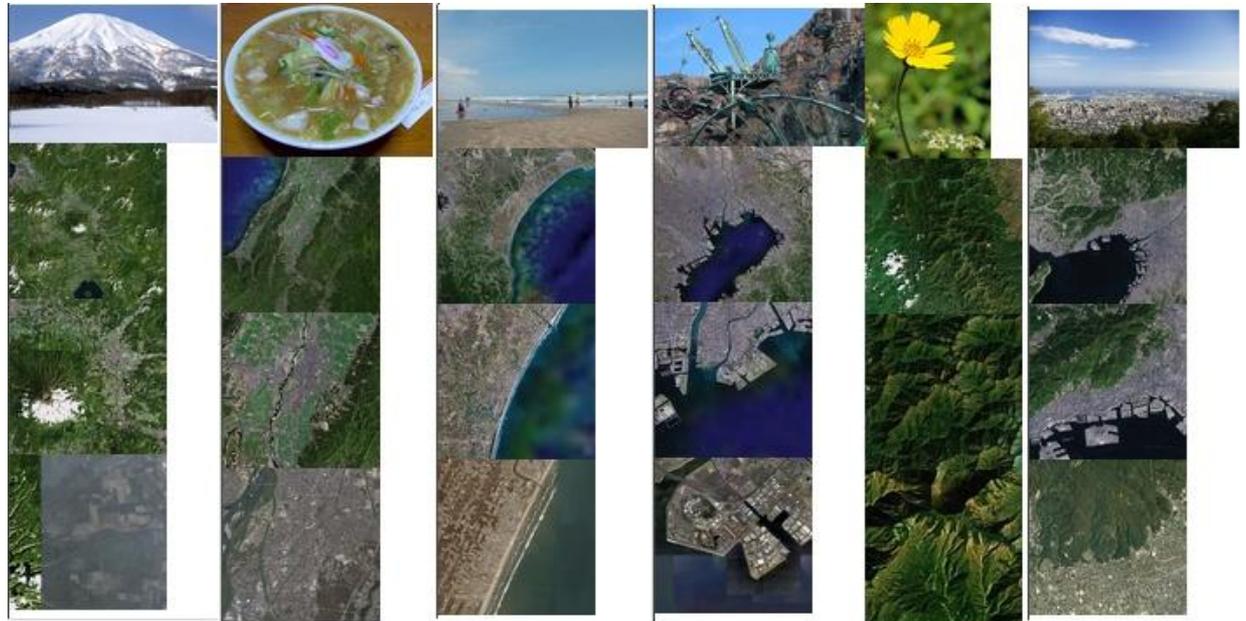
地理的コンテキスト



本研究の意義

- 位置情報付き写真の認識に、対応する位置の航空写真を利用する
- 画像 \leftrightarrow 位置情報の特徴をマルチカーネル学習 (MKL)で推定

どの
特徴が
どれだけ
有用?
↓
MKLで
推定



実験の手順

データ収集



&



特徴抽出

画像特徴

画像特徴

MKL=
マルチカーネル
学習

機械学習

MKL-SVM

実験結果

データ収集

データ収集



特徴抽出



機械学習



&



画像特徴



画像特徴



実験結果

データ収集と航空写真の利用

カテゴリで検索

“Mountain”

API Search Request

日本国内の画像

flickr®

Photos with a geotag

latitude=42.844202
longitude=140.852899



Level 4

497m x 497m
256x256

Level 3

1.91km x 1.91km
256x256

Level 2

7.64km x 7.64km
256x256

Level 1

30.8km x 30.8km
256x256

Web地図サービス

Google™
Maps

Google™
Earth

特徴抽出

データ収集



特徴抽出



機械学習



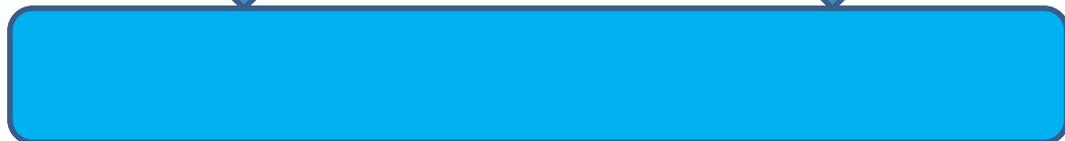
&



画像特徴



画像特徴



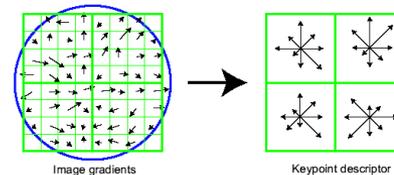
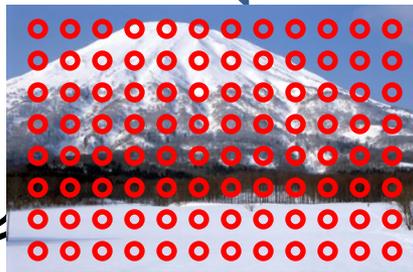
実験結果

特徴抽出全体の流れ

元画像

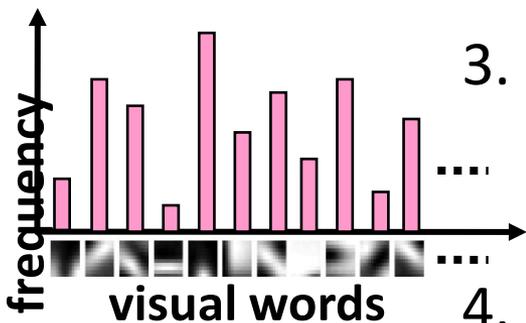


10x10px
格子点抽出



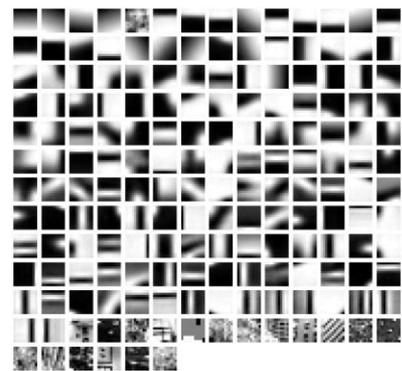
SIFT特徴

1. 特徴点を抽出
2. SIFT記述子により特徴点周辺パターンを抽出
SIFT特徴ベクトルとして抽出
3. 特徴ベクトルをクラスタリングして求めた
コードブックに基づいてSIFT特徴ベクトル
をベクトル量子化
4. 画像毎にヒストグラムを作成



これを
Bag of Keypoints
表現に変換
するための

Bag of Keypoints 表現



画像・航空写真からの特徴抽出

画像から

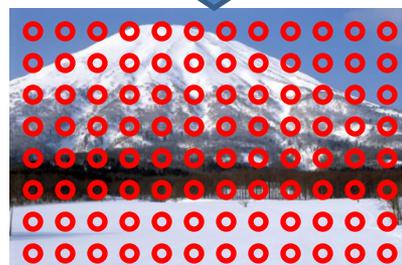
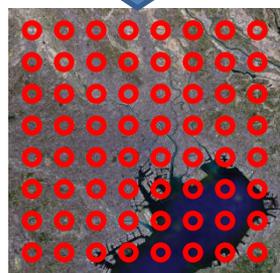
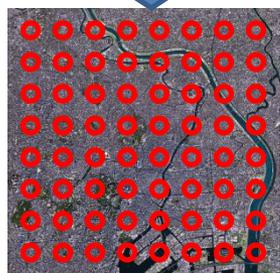


Image
Codebook

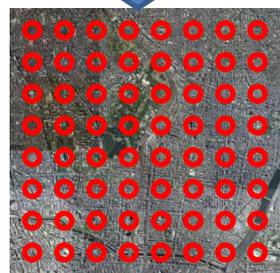
各レベルの航空写真から



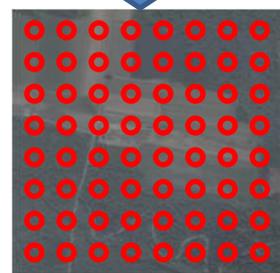
Level 1
Codebook



Level 2
Codebook

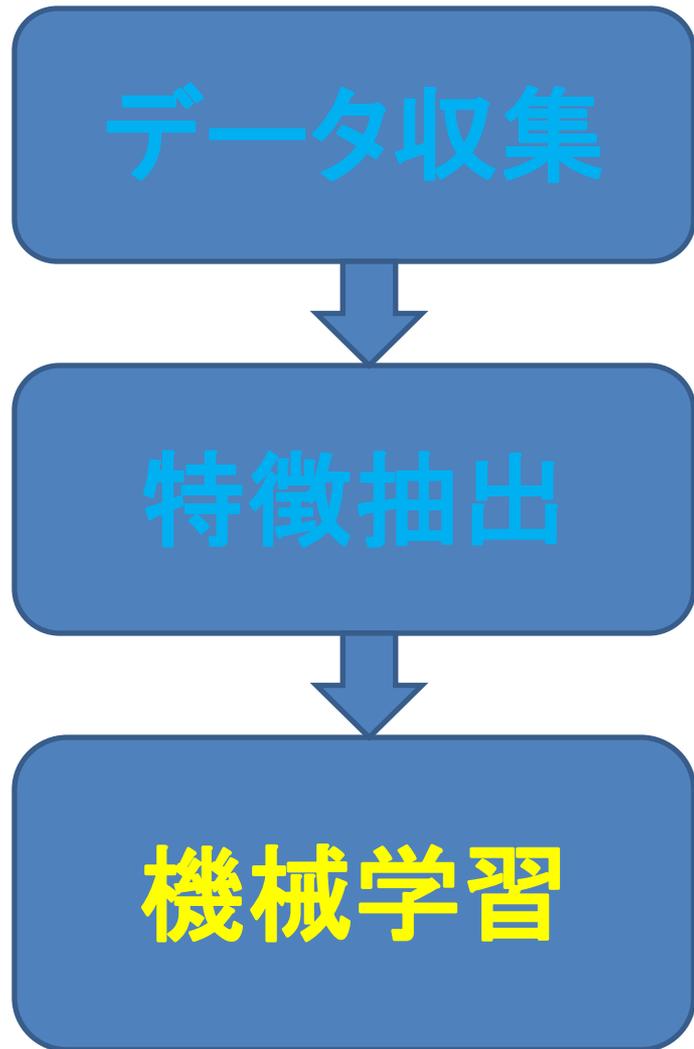


Level 3
Codebook

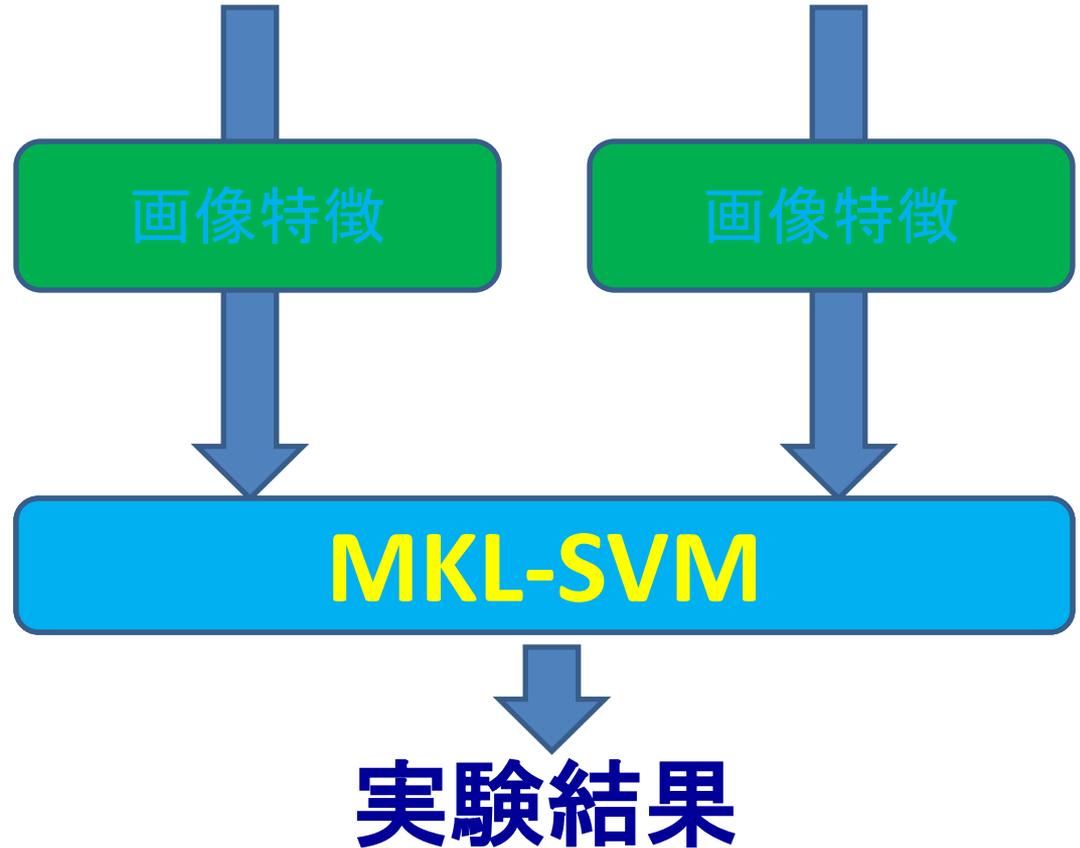


Level 4
Codebook

機械学習

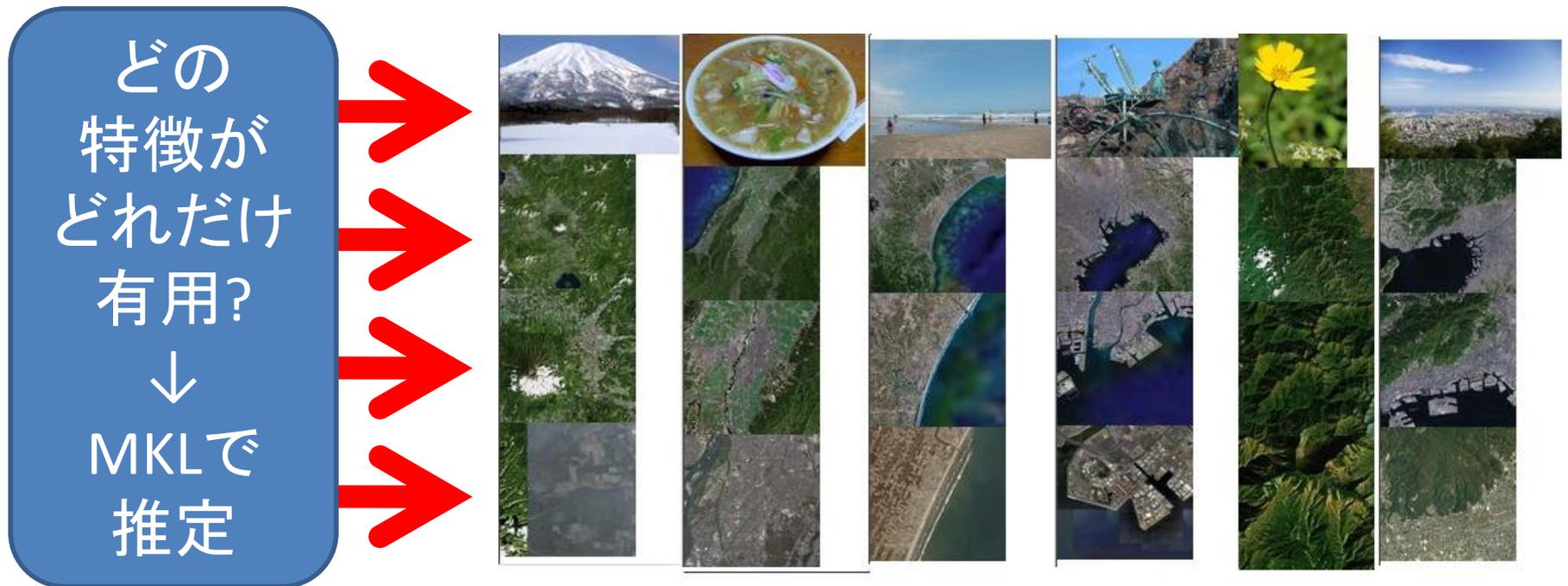


&



特徴の統合と画像分類方法

- 特徴の融合 : Multiple Kernel Learningで重み推定



- 分類手法 : MKL-SVM (2クラス分類)

Multiple Kernel Learningの利用

- 複数のSVMのカーネルを線形結合することにより特徴を統合する

$$K_{combined}(\mathbf{x}, \mathbf{x}') = \sum_{j=1}^K \beta_j k_j(\mathbf{x}, \mathbf{x}') \quad \text{with } \beta_j \geq 0, \quad \sum_{j=1}^K \beta_j = 1.$$

MKLを使用する利点

様々な特徴の
最適な重み



種類ごとの
重要な重み

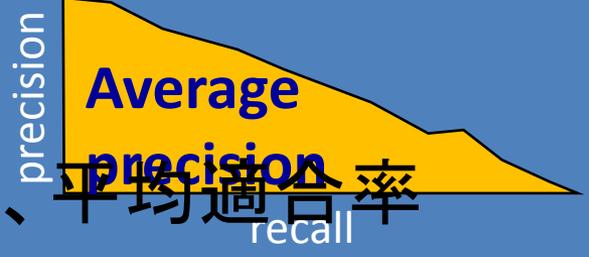
特徴の組み合わせと評価

baseline



2. 評価方法:

5-fold クロスバリデーション
 = クロスバリデーション、平均適合率で評価 (AP).



• 画像の組み合わせ

MKL

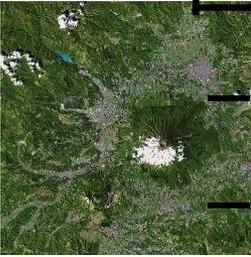


Image + Level 1

Image + Level 2

Image + Level 3

- 画像のみ
- 画像 + 航空写真レベル1 (MKL)

- 画像 + 航空写真レベル2 (MKL)

- 画像 + 航空写真レベル3 (MKL)

MKL



MKL



Image + Level 4

Image + Multi levels

位置情報付き画像の実験カテゴリ

東京タワー

ディズニーリゾート

城

典型的なランドマークなど



典型的なランドマークなど

橋

鉄道

海岸

公園

航空写真から見える地理的物体



航空写真から見える地理的物体

景色

夕日

桜・花見

紅葉

祭

航空写真で見えないが風景



航空写真で見えないが風景

位置情報付き画像の実験カテゴリ

花

神社

猫

屋外の物体

屋外の物体

ラーメン

寿司

自動販売機

概ね屋内の物体

概ね屋内の物体

ランダムな画像

200枚のpositive画像と200枚のnegative画像を手動で選択



ラーメン



ラーメンではない

報告書p.163表2：お詫びと訂正

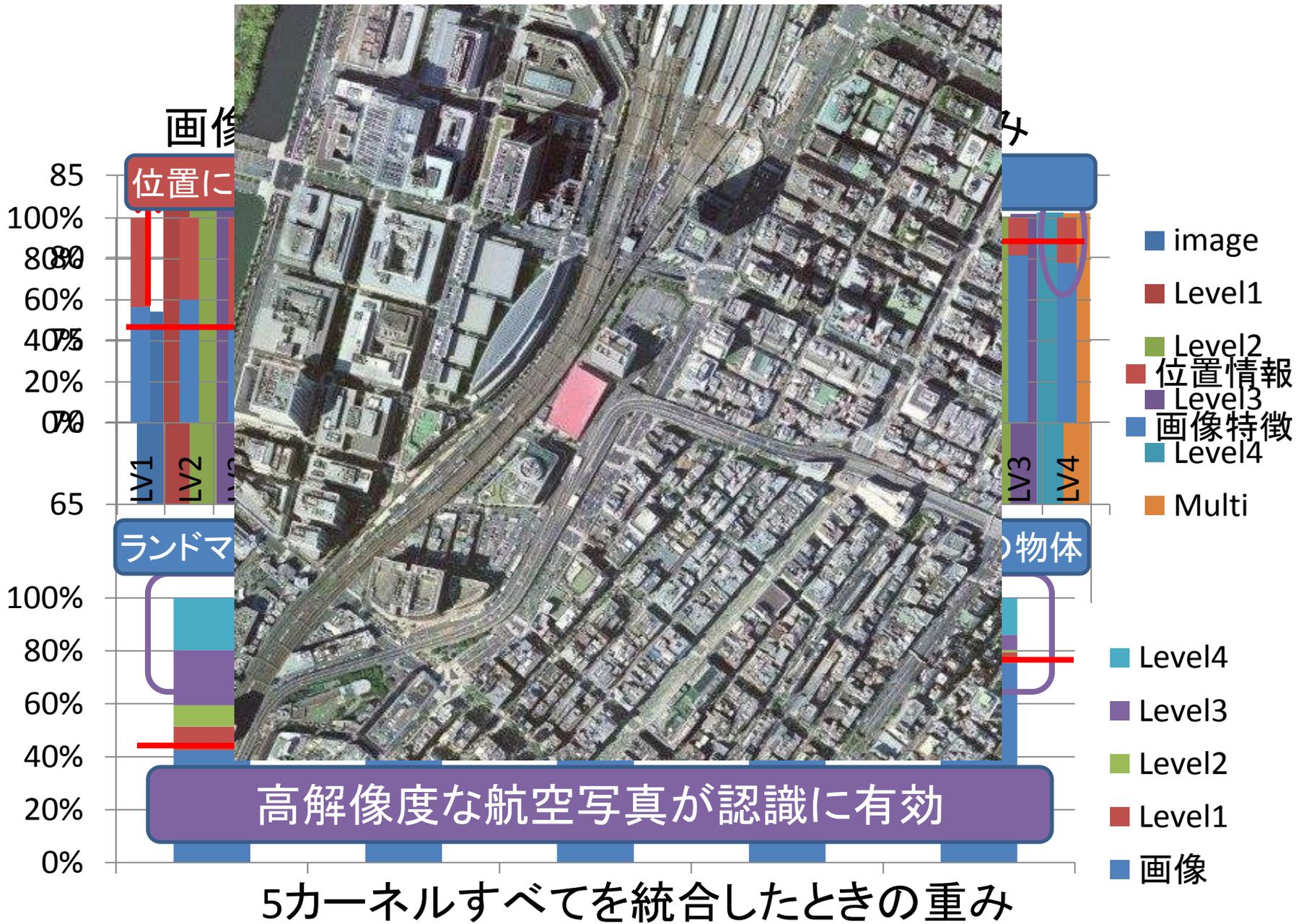
- Imageカラムの数値は、
画像+5カーネルを統合したときの重みの数値
- ベースラインの比較がありません。

カテゴリ	Level1	Level2	Level3	Level4	Image	
カテゴリ	Image	Level1	Level2	Level3	Level4	Multi

Diagram illustrating the table structure correction. The top row shows the original table structure with columns: カテゴリ, Level1, Level2, Level3, Level4, Image. A blue arrow points from the 'Image' column to the 'Multi' column in the bottom row, labeled '追加' (Addition). The bottom row shows the corrected table structure with columns: カテゴリ, Image, Level1, Level2, Level3, Level4, Multi. A blue arrow points from the 'Image' column in the bottom row to the 'Image' column in the top row, labeled '同じ数値' (Same numerical value).

詳細な数値はWebでご覧下さい。御手数をお掛け致します

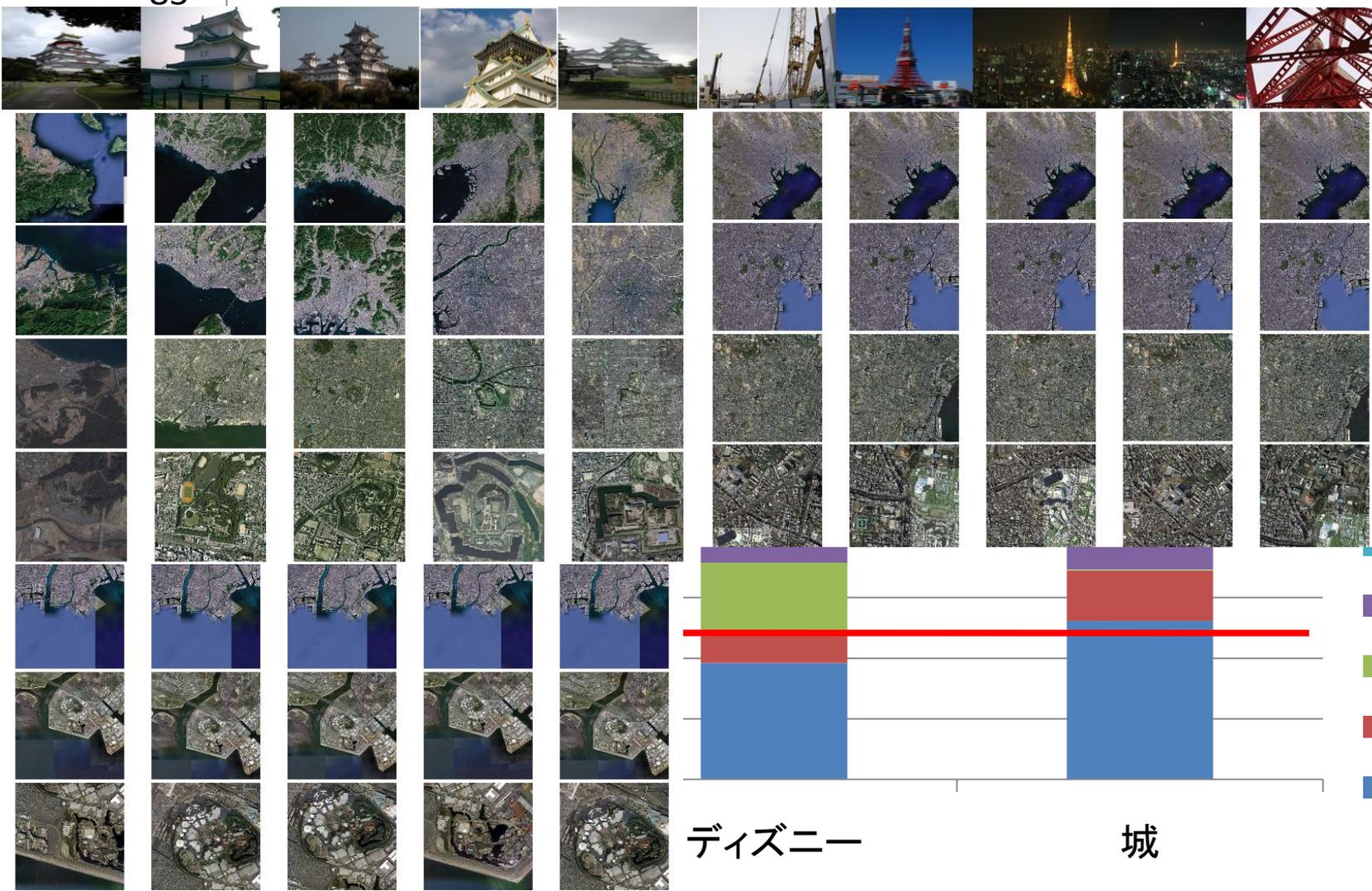
[http://mm.cs.uec.ac.jp/yaegashi/
prmu2009eratta.html](http://mm.cs.uec.ac.jp/yaegashi/prmu2009eratta.html)



実験結果(2) - ランドマーク

位置に特有

85

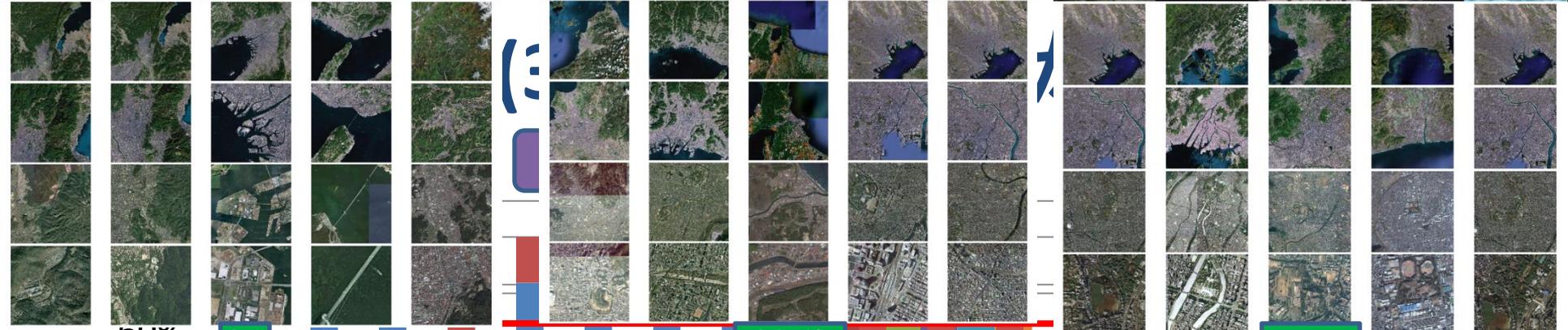


- image
- Level1
- Level2
- 位置情報
- Level3
- 画像特徴
- Level4
- Multi
- Level4
- Level3
- Level2
- Level1
- 画像



ディズニー

城



橋

水

公園

60%
40%
20%
0%

橋

鉄道

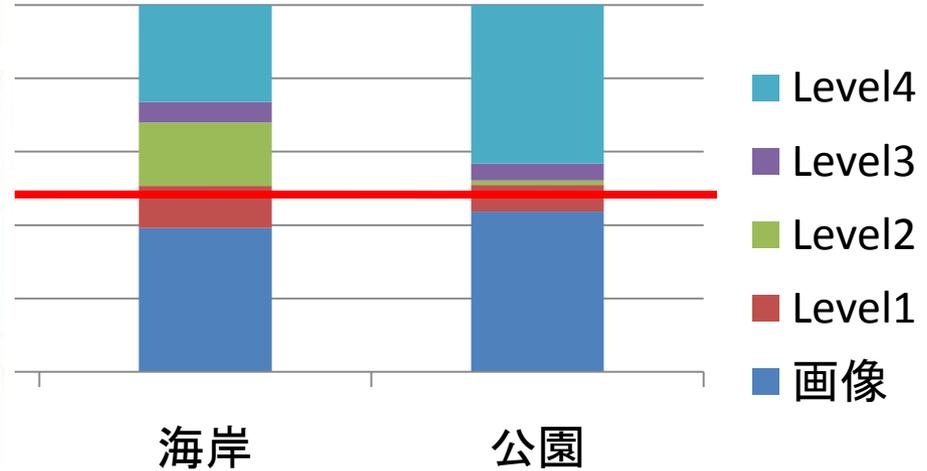
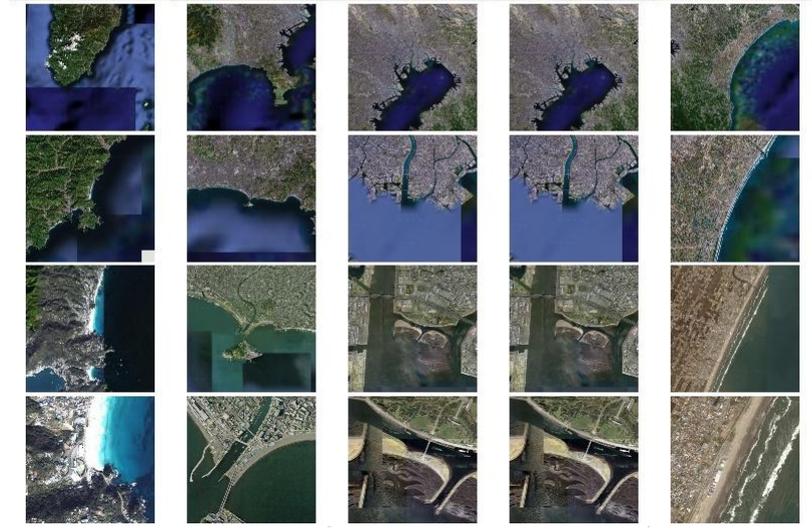
公園

- Level2
- 位置情報
- Level3
- 画像特徴
- Level4
- Multi



海岸

公園



海岸

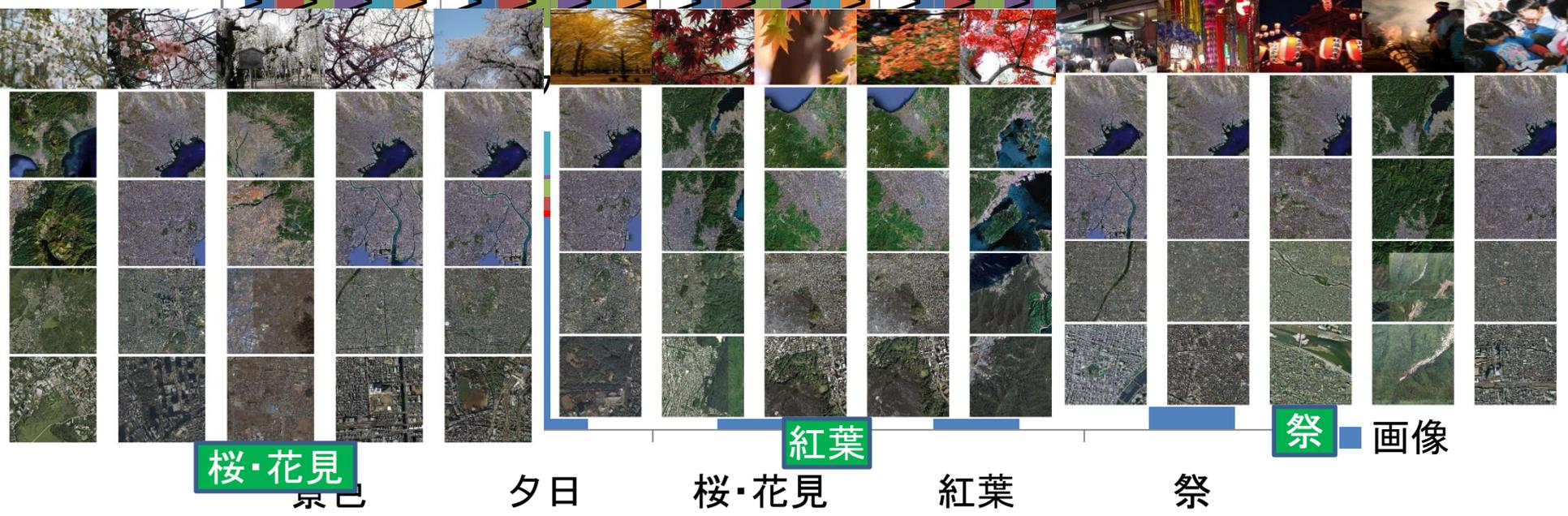
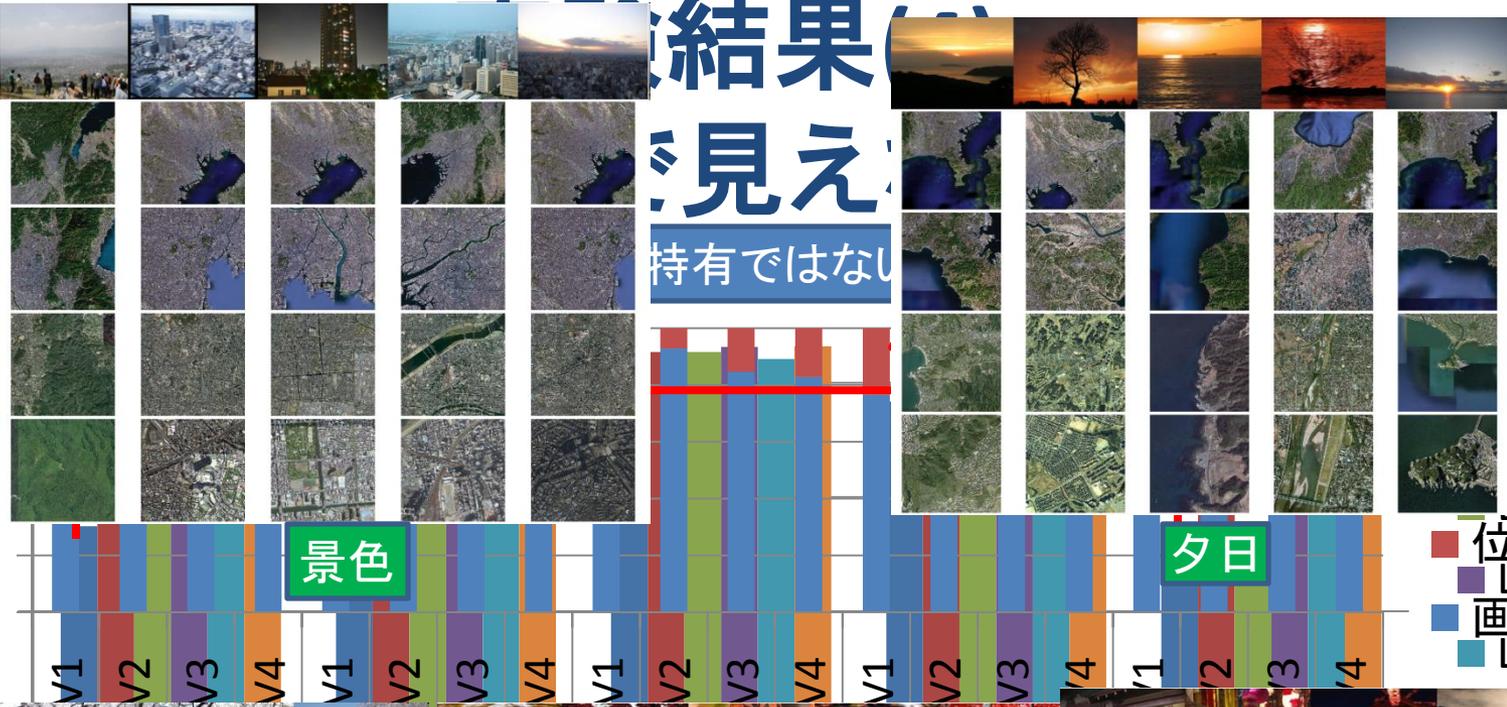
公園

- Level4
- Level3
- Level2
- Level1
- 画像

結果 見え

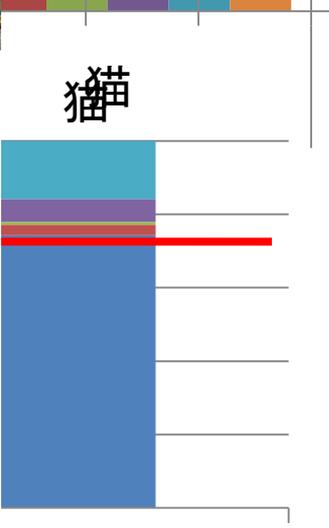
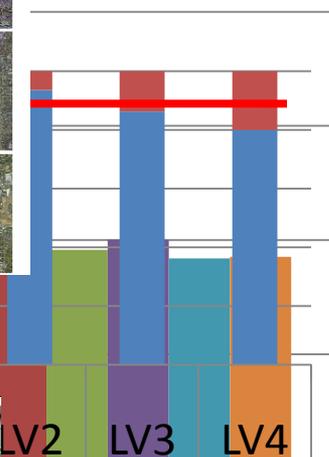
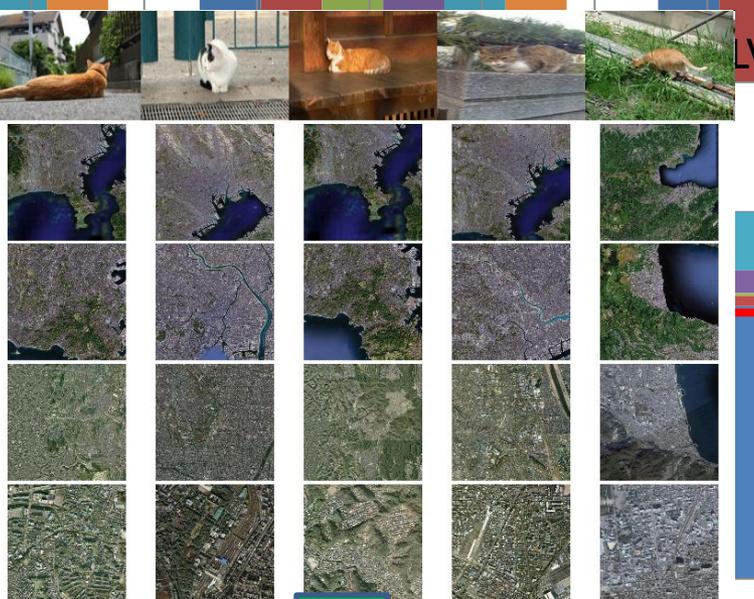
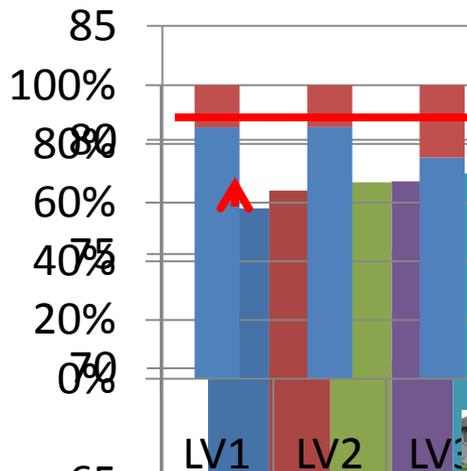
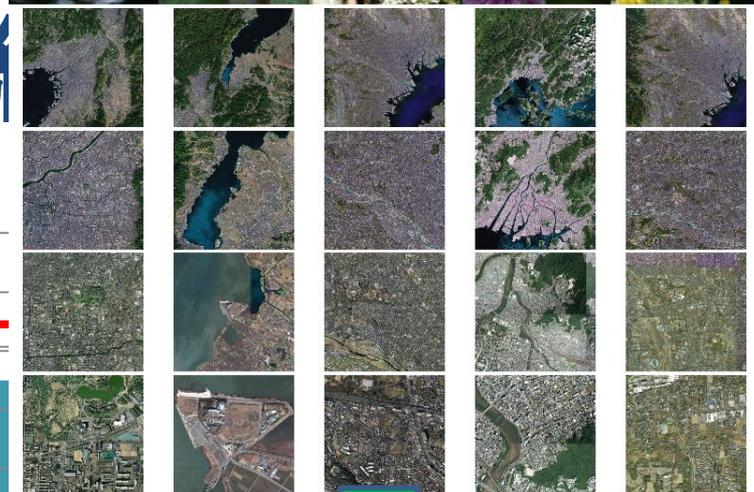
特有ではない

85
100%
80%
60%
40%
20%
0%



実験結果

物体



- image
- Level1
- Level2
- Level3
- Level4
- Multi

- Level4
- Level3
- Level2
- Level1
- 画像



猫

猫

猫

花



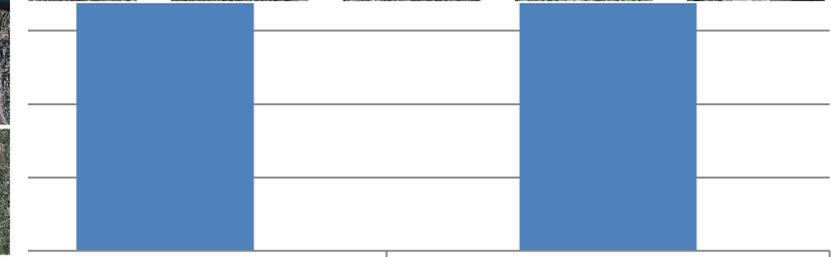
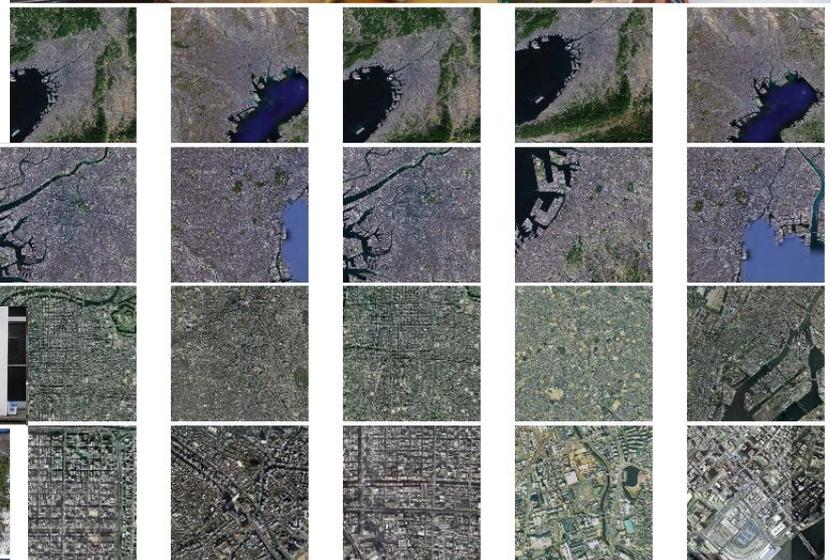
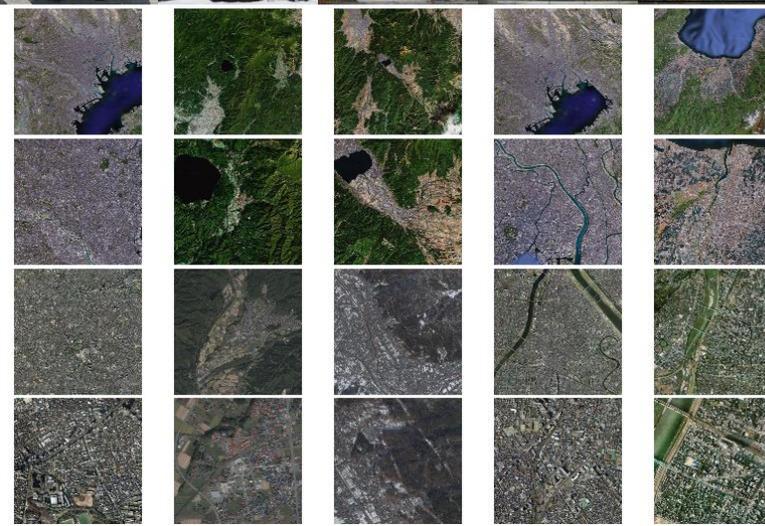
概ね屋内の物体



に位置に特有ではない



- image
- Level1
- Level2
- 位置情報
- Level3
- 画像特徴
- Level4
- Multi



ラーメン

自販機

- Level4
- Level3
- Level2
- Level1
- 画像

結論

- 航空写真の利用の可能性を確認

– 重みとベースラインとの比較を以て、航空写真が有効なカテゴリを定量的に確認

位置情報が有効

ディズニーリゾート



東京タワー



城



位置情報が若干有効

橋



鉄道



海岸



公園



位置情報があまり有効でない

景色



夕日



桜・花見



紅葉



祭



花



神社



猫



位置情報が明らかに有効でない

ラーメン



寿司



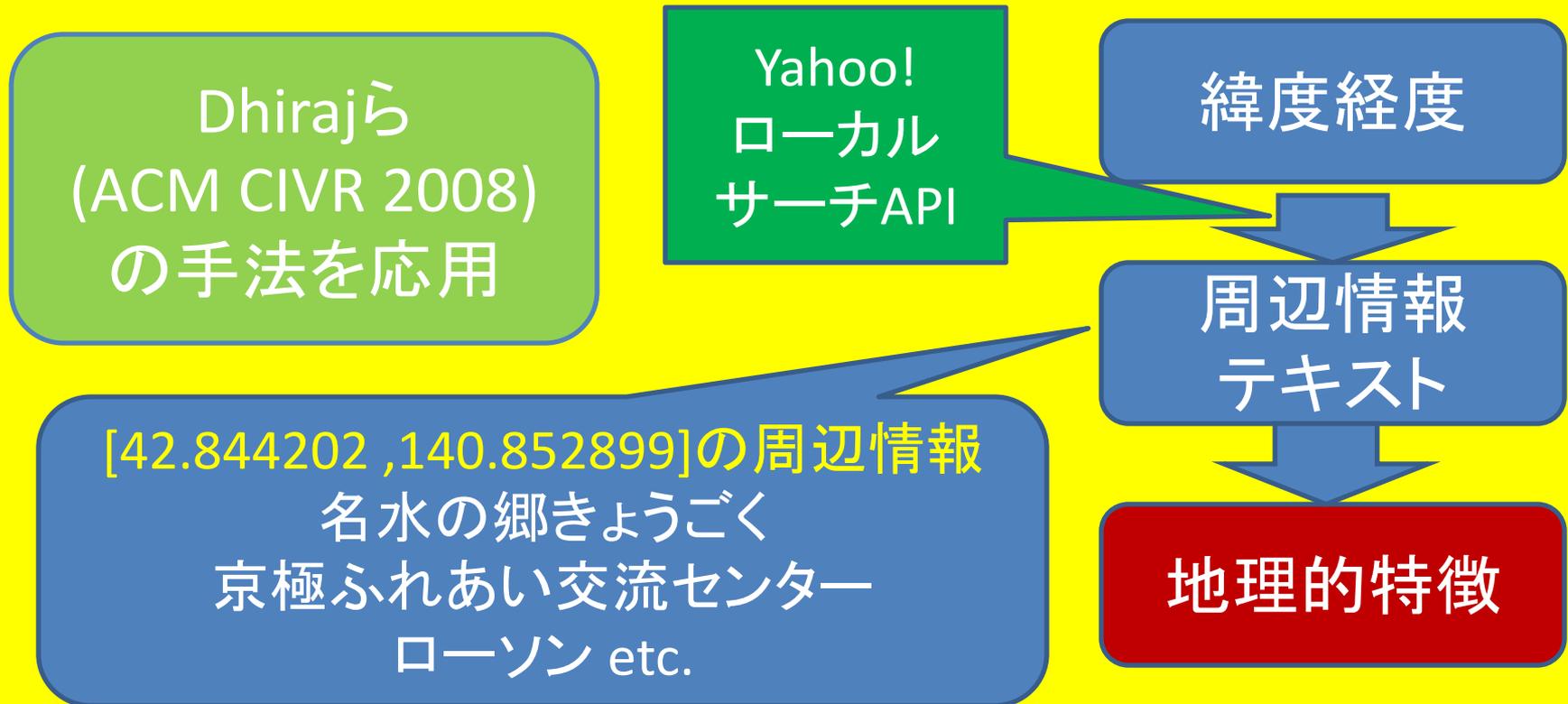
自動販売機



- 解像度の高い航空写真が認識に有効である可能性を示唆

今後の課題

～ 周辺情報の利用 ～



実験結果：表2の訂正後の数値

[http://mm.cs.uec.ac.jp/yaegashi/
prmu2009eratta.html](http://mm.cs.uec.ac.jp/yaegashi/prmu2009eratta.html)

実験データセット

[http://mm.cs.uec.ac.jp/yaegashi/
flicker2/annotate.cgi?mode=viewset](http://mm.cs.uec.ac.jp/yaegashi/flicker2/annotate.cgi?mode=viewset)



ご静聴ありがとうございました