

An aerial photograph of a city at sunset, with buildings and streets visible under a hazy, golden sky. The text is overlaid on the center of the image.

# ジオタグ画像認識における 周辺テキスト情報の有効性の検証

情報処理学会  
第171回コンピュータビジョン  
・イメージメディア研究会  
電気通信大学大学院 情報工学専攻  
八重樫恵太 柳井啓司  
2010年3月19日

# 背景

大量の位置情報付き画像



撮影時に自動的に記録：  
緯度・経度、方角、日付

位置情報と時間特徴は  
自動的に取得可能

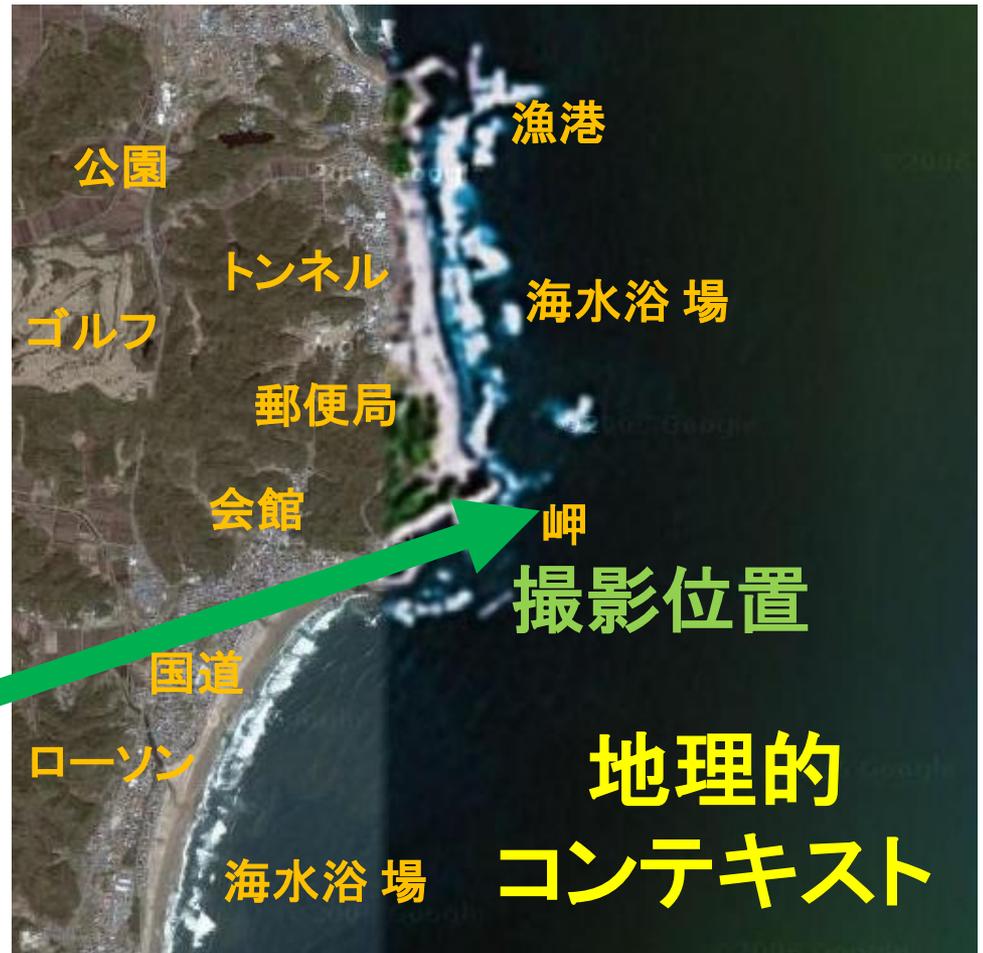
画像認識精度向上のために  
メタデータ、位置情報特徴を利用



# 位置テキスト特徴の利用

認識対象  
⇔ 位置情報  
密接な関係

(36.995768,140.98246)



# 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

# 画像収集と航空写真の利用

カテゴリで検索

“Mountain”

API Search Request

flickr®

日本国内の画像

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

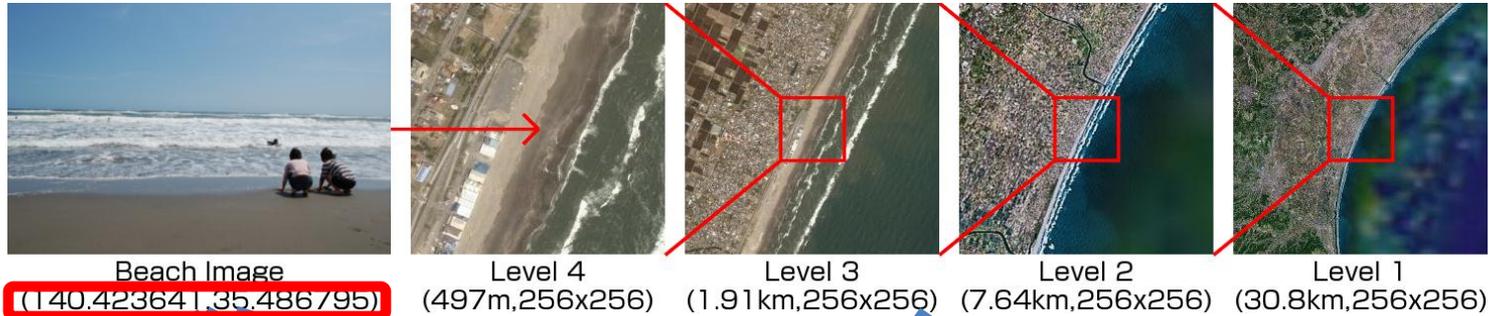
30.8km x 30.8km  
256x256

Web地図サービス

Google™  
Maps

Google™  
Earth

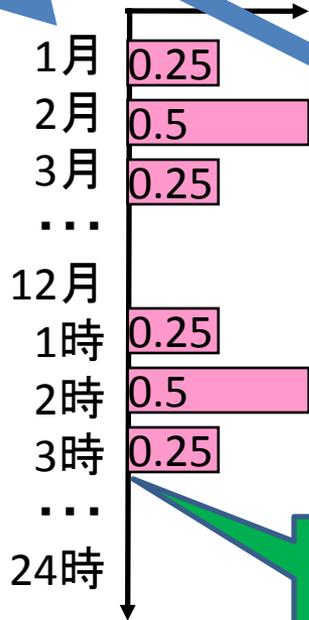
# 特徴量の種類



Location

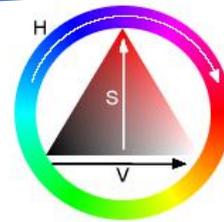
Geo-text

北今泉海水浴場  
 九十九里IC  
 波乗り道路etc.



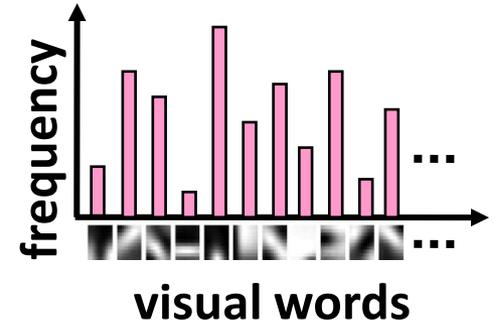
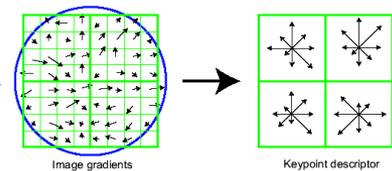
Date

月・時の該当箇所に投票



HSV

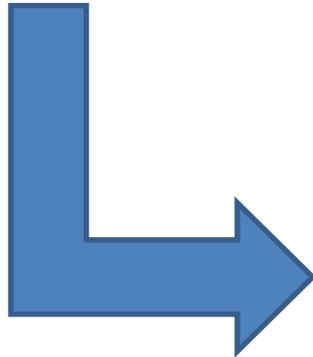
BoK



# Yahoo! ローカルサーチAPI



(34.687752,  
138.972456)



白浜荘  
白浜神社  
ホテル伊豆急  
白浜大浜海水浴場  
アオギリ自生地  
下田プリンスホテル  
下田白浜郵便局  
白浜保育園  
市立白浜小学校  
白浜中央海水浴場  
下田ビューホテル  
国道135号線  
ジヨモ  
市立下田東中学校  
柿崎保育園  
臨海荘  
玉泉寺  
下田聚楽ホテル  
下田柿崎郵便局  
外浦海水浴場



(31.591148,  
130.558997)



「のぼる屋の  
ラーメン」

南日本銀行本店  
ローソン  
リッチモンドホテル  
鹿児島交通  
鹿児島東郵便局  
山形屋  
鹿児島信用金庫本店  
市役所別館  
鹿児島銀行本店  
宝山ホール  
鹿児島市教育総合センター  
さつまや  
西本願寺  
鹿児島県産業会館  
鹿児島市役所  
市立名山小学校  
中央公民館  
ホテル&レジデンス南洲館  
照国神社前通り郵便局  
鹿児島相互信用金庫

# 位置テキスト特徴



Yahoo!  
ローカル  
サーチAPI

緯度経度

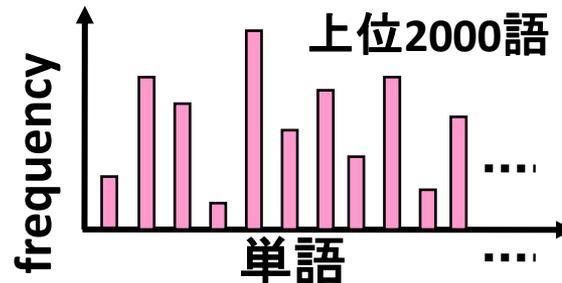
(35.486795,140.423641)

[42.844202,140.852899]の周辺情報

白里中央海水浴場  
白里郵便局  
北今泉海水浴場  
グラシラスケア九十九里etc.

ビル 局  
郵便 寺  
小学校  
...

メタデータには  
含まれない情報  
(Flickrのテキスト  
タグではない)



周辺情報  
テキスト

単語分解  
Chasenを使用

ヒストグラム

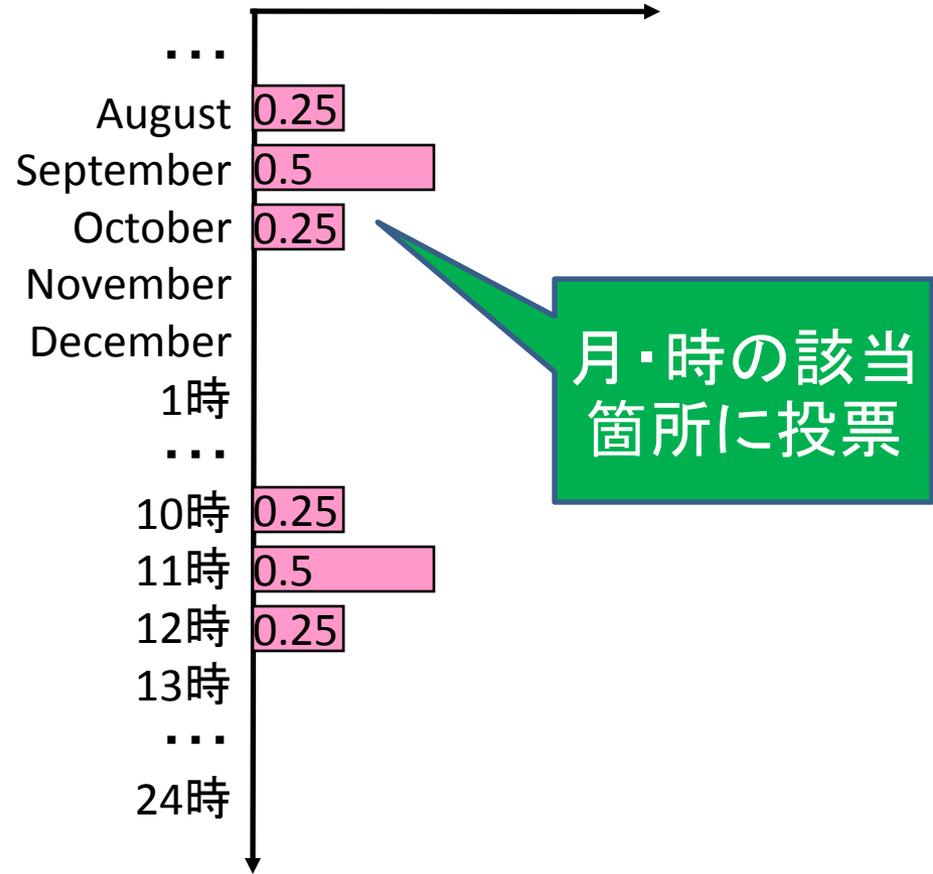
# 時間(時間帯・季節)特徴



2007-09-02 11:41:03



撮影日時

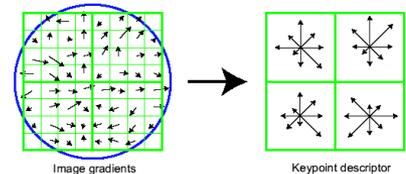
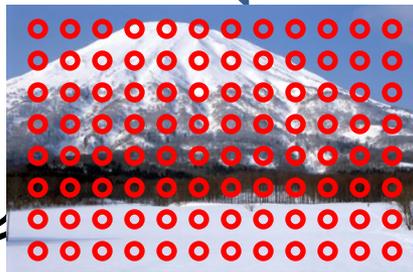


# Bag of keypoints (1)

元画像

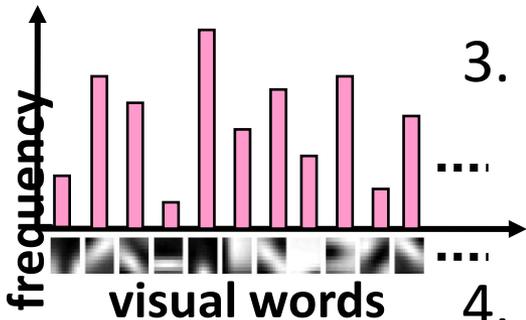


10x10px  
格子点抽出

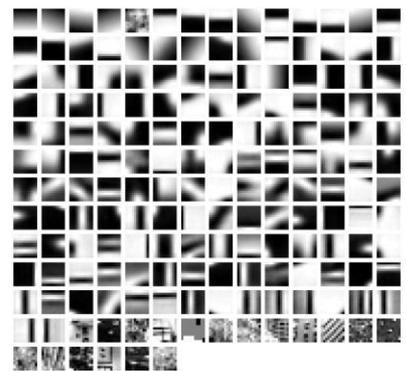


SIFT特徴

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. これを Bag of Keypoints 表現に変換
3. SIFT特徴ベクトルをクラスタリングして求めた visual words をベクトル量子化.
4. 画像毎にヒストグラムを作成.



Bag of keypoints 表現



# Bag of keypoints (2)

画像から

各レベルの航空写真から

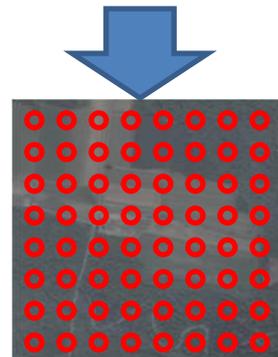
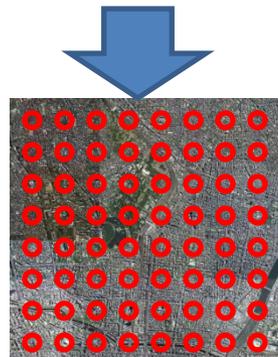
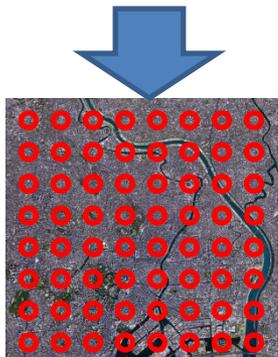
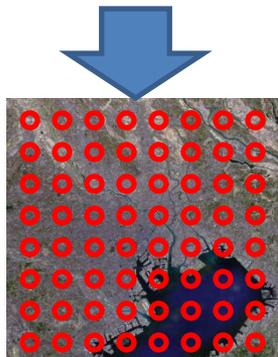
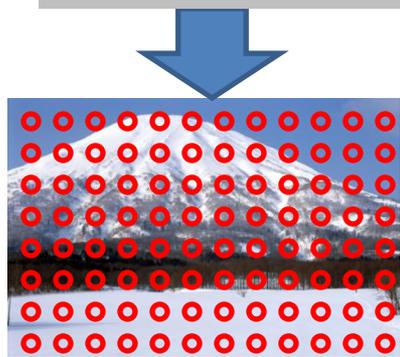


Image  
Codebook



Level 1  
Codebook



Level 2  
Codebook



Level 3  
Codebook



Level 4  
Codebook

# 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

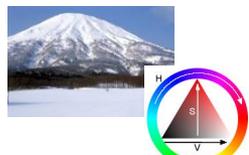
様々な特徴の  
最適な重み



種類ごとの  
重要な重み

MKL-SVM : 緯度経度 はRBF, それ以外は $\chi^2$

## 特徴の組合せと評価



緯度  
経度

周辺情報  
テキスト



	BoK	HSV	Aerial	Geo	Text	Time
BoK	●					
HSV		●				
視覚	●	●				
視覚+航写	●	●	●			
視覚+位置	●	●		●		
視覚+テキ	●	●			●	
視覚+航写+テキ	●	●	●		●	
視覚+航写+テキ+時間	●	●	●		●	●

### 2クラス分類:

5-fold クロスバリデーション  
平均適合率で評価(AP):

precision



recall

# 位置情報付き画像の実験カテゴリ



位置に特有なランドマーク  
東京タワー，ディズニーリゾート



狭い範囲の  
地理構成物  
橋，神社，建物，  
城，鉄道



広い範囲の地理構成物  
公園，庭園，風景



地形  
湖畔，川，海岸

# 位置情報付き画像の実験カテゴリ



屋外の人工物  
像・モニュメント, 自動車,  
自転車, 落書き,  
自動販売機



時期依存的要素  
紅葉, 桜・花見,  
夕日, コスプレ, 祭



天然の物体

200枚のpositive画像と200枚のnegative画像を手動で選択

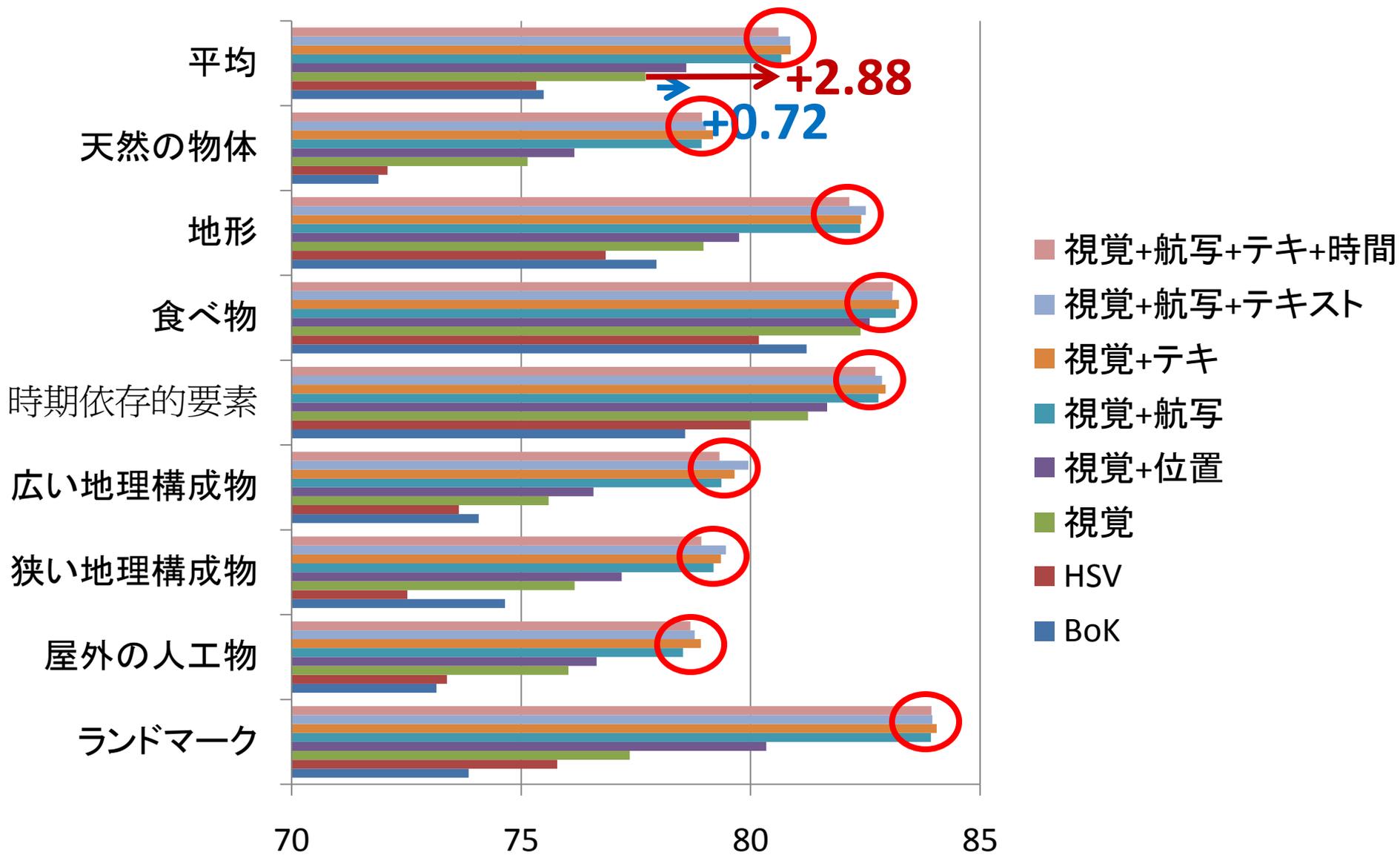


ラーメン

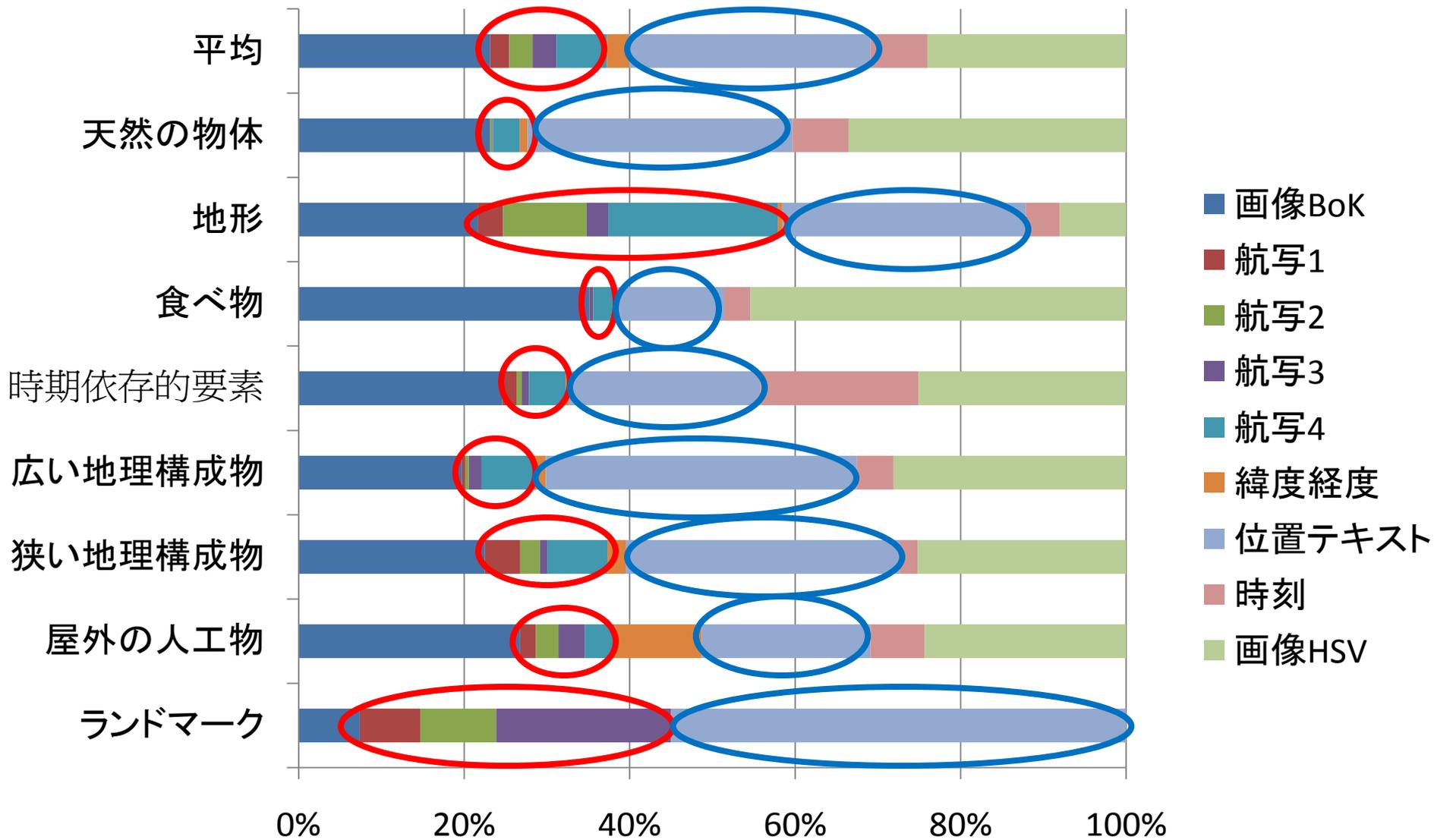


ラーメンではない

# 実験結果(1)



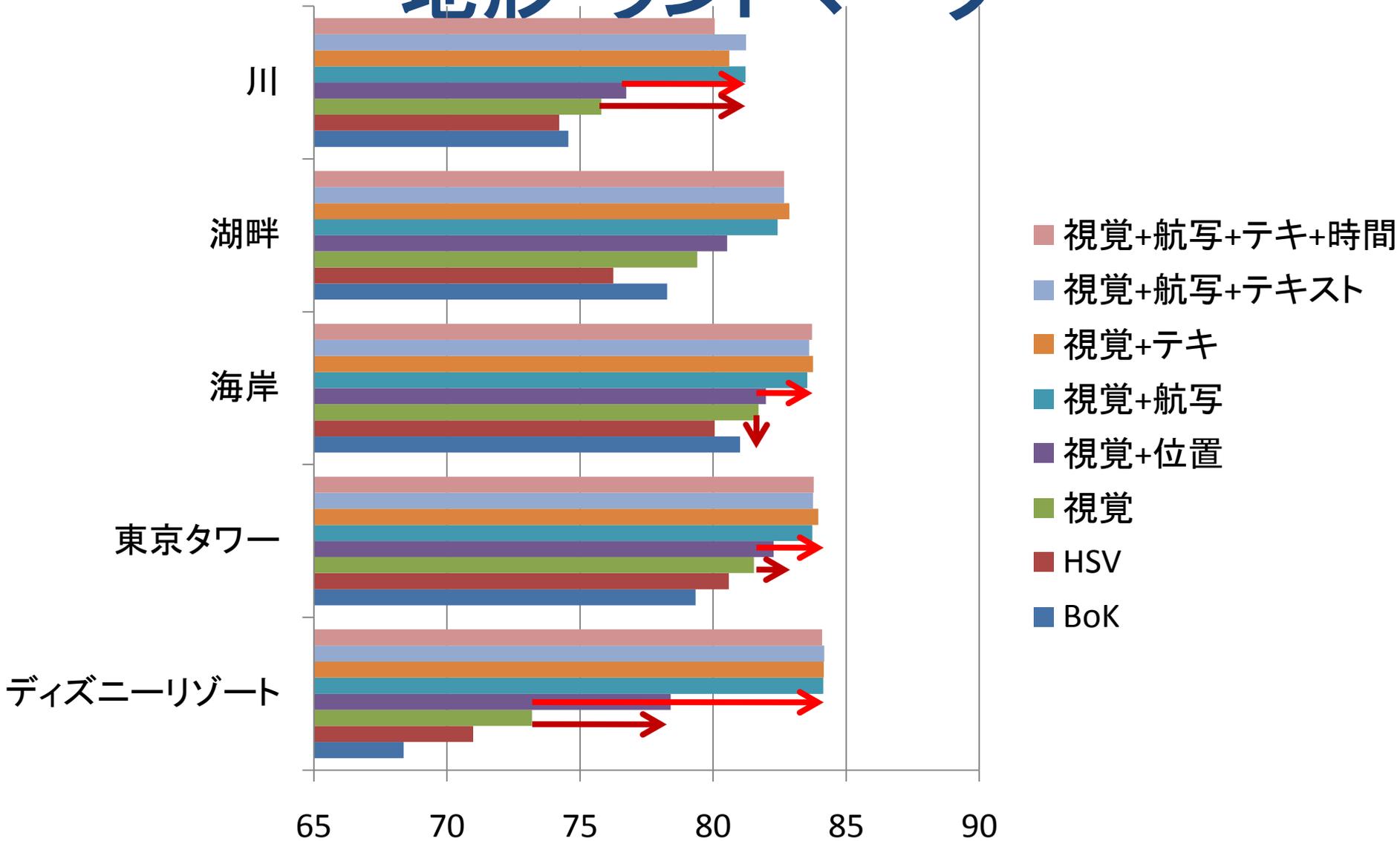
# 実験結果(2)



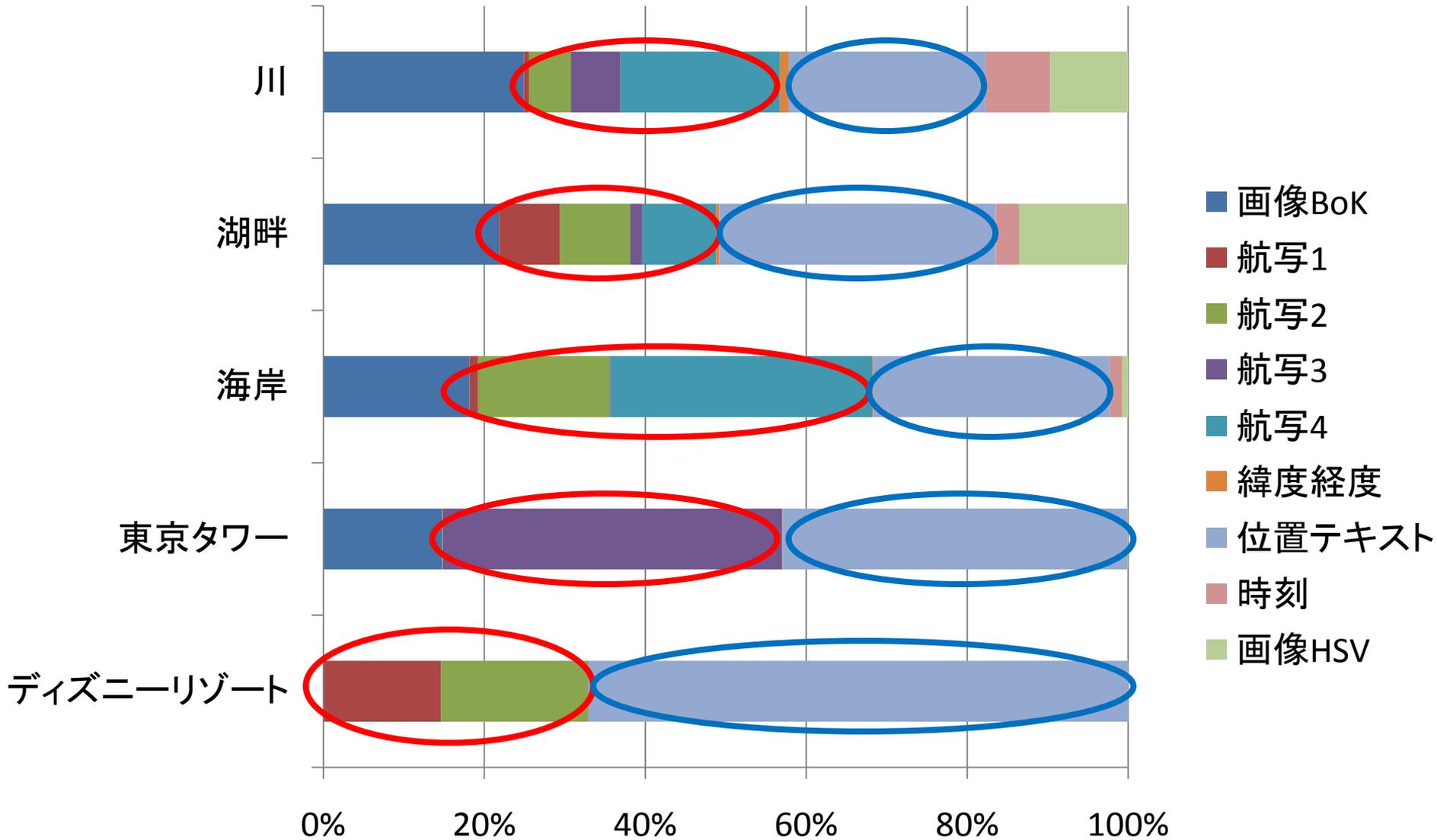
# 位置テキストの分布(上位20語)

東京タワー		ディズニーリゾート		花		夕日	
ビル	0.1089	場	0.117	ビル	0.0326	線	0.027
大使館	0.0395	駐車	0.1074	局	0.0186	公園	0.0187
会館	0.0255	東京	0.0649	郵便	0.0178	市立	0.0175
六本木	0.0211	舞浜	0.0424	寺	0.0163	小学校	0.0166
神谷町	0.0199	リゾート	0.0416	小学校	0.0161	局	0.0166
日本	0.0186	パーキング	0.0347	市立	0.0158	ビル	0.0165
麻布	0.0183	ホテル	0.0321	新宿	0.0136	郵便	0.0158
聖	0.0183	南葛西	0.024	幼稚園	0.0121	センター	0.0141
虎ノ門	0.0181	ベイ	0.0197	東京	0.0117	橋	0.0139
飯倉	0.018	入口	0.0139	センター	0.0116	寺	0.0127
寺	0.017	東急	0.012	線	0.0112	保育園	0.0124
タワー	0.0164	区立	0.0115	保育園	0.0111	横浜	0.0107
館	0.014	シアター	0.0108	学校	0.0103	東京	0.0101
泉	0.0111	公園	0.0105	公園	0.0098	場	0.0094
局	0.0109	ランド	0.01	院	0.0094	幼稚園	0.0092
郵便	0.0109	野毛	0.01	セブン	0.0092	セブン	0.0087
芝	0.0101	立体	0.0092	イレブン	0.0091	イレブン	0.0087
殿	0.0097	クリスタル	0.0089	中学校	0.0088	ホテル	0.0084
教会	0.0093	保育園	0.0089	橋	0.0085	中学校	0.0077
振興	0.0092	小学校	0.0087	会館	0.0082	国道	0.0074

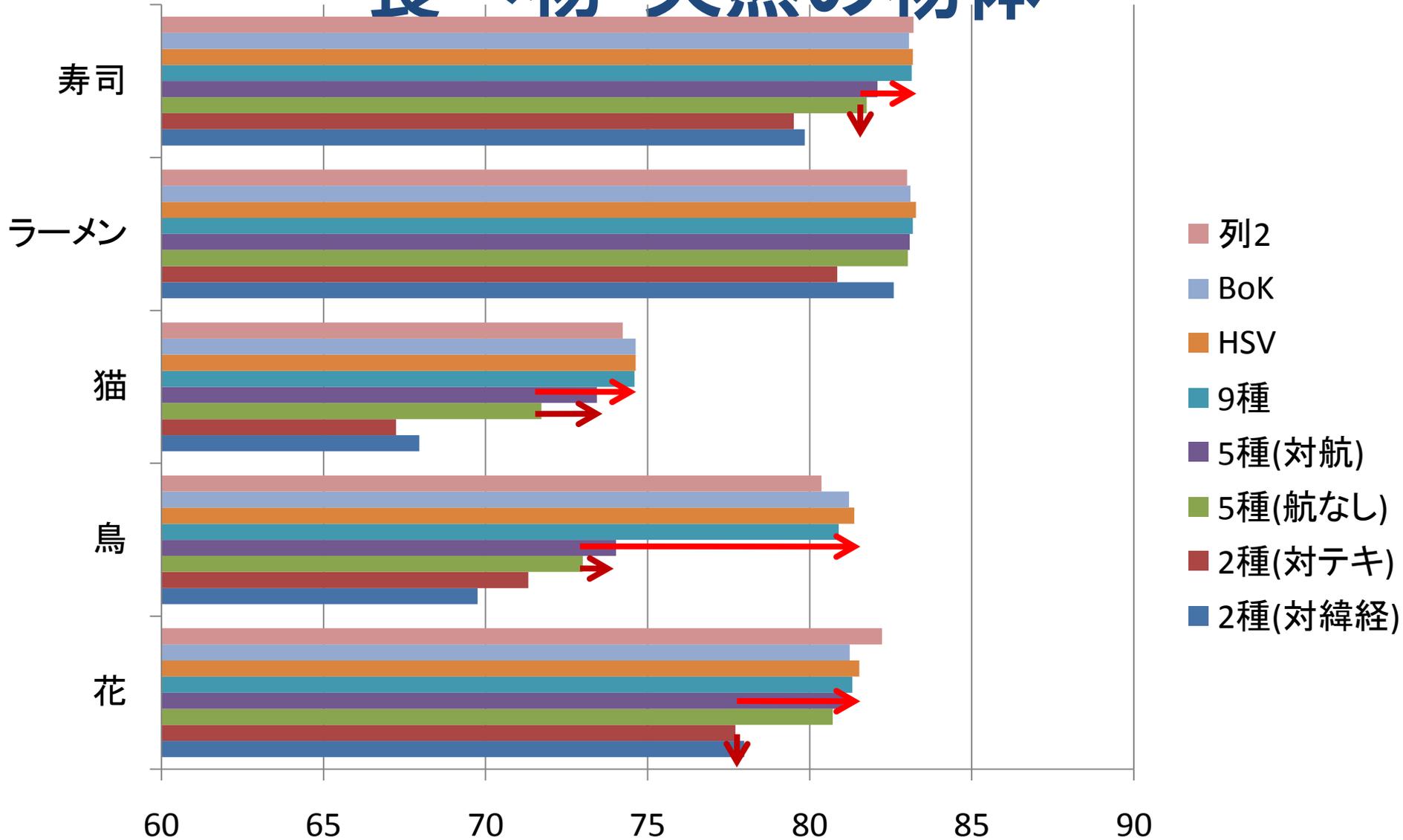
# カテゴリごとの結果 地形・ランドマーク



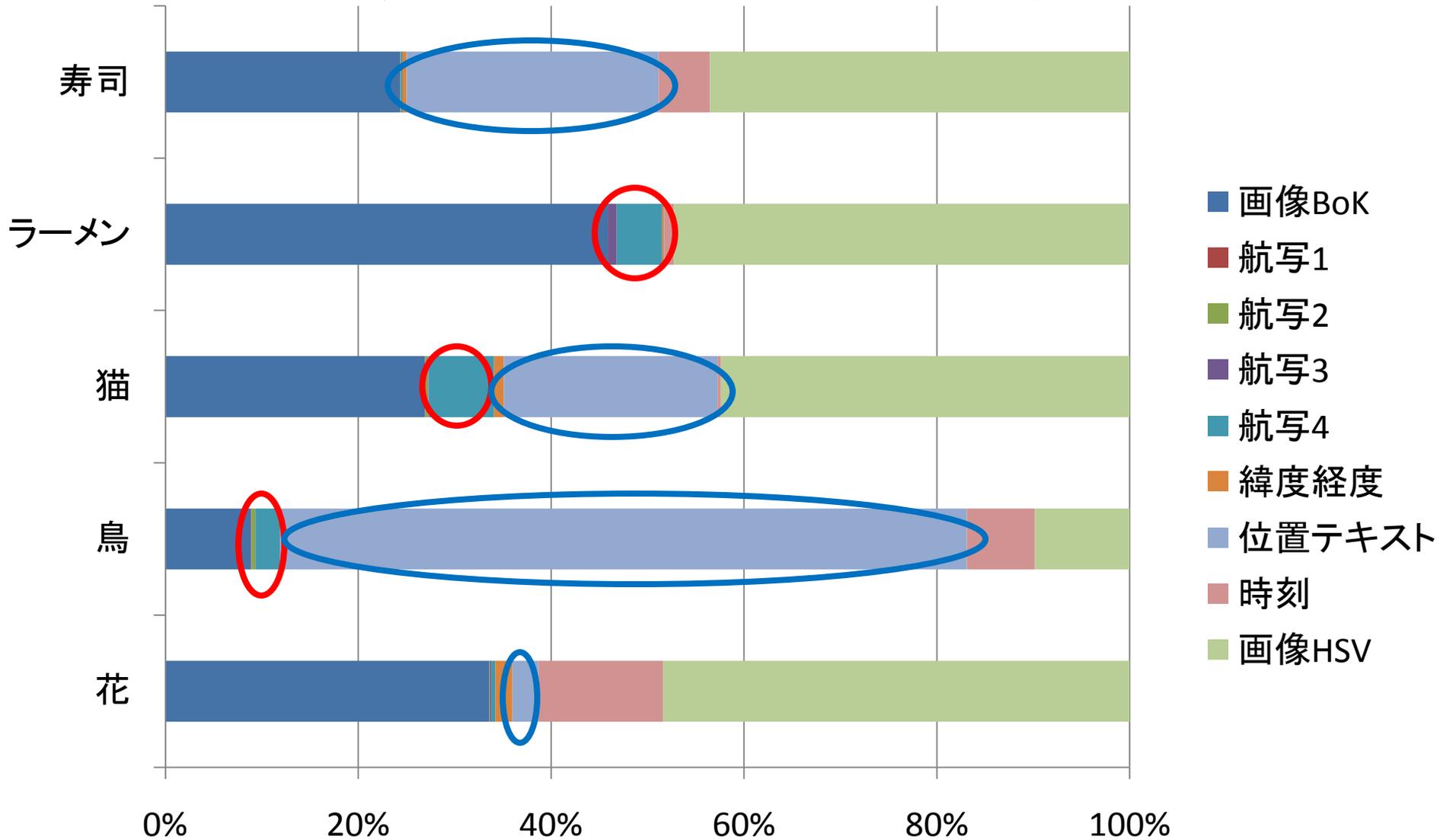
# カテゴリごとの結果 地形・ランドマーク



# カテゴリごとの結果 食べ物・天然の物体



# カテゴリごとの結果 食べ物・天然の物体



# 結論と今後の課題

- 結論

- 航空写真→位置テキストの利用
- いずれの組み合わせもBoK+HSVより約3%のgain
- 重みからテキストの有用性は認められるが、航空写真を超える有意な精度向上は認められない

- 今後の課題

- 単語抽出方法の再検討
  - 固有名詞(地名など)の辞書強化、過剰分割の抑制
- 国内画像28カテゴリ→世界規模でより多くのカテゴリ

An aerial photograph of a densely populated city, likely Tokyo, taken during the golden hour of sunset. The sky is filled with soft, warm light and scattered clouds. The city below is a vast expanse of buildings, with a mix of low-rise residential structures and taller commercial buildings. The Japanese text 'ご静聴ありがとうございました' is overlaid in the lower center of the image.

ご静聴ありがとうございました