

# 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

八重樫恵太<sup>†</sup> 柳井 啓司<sup>††</sup>

**あらまし** 本稿では、我々が提案した、位置情報付き画像の認識において、写真の撮影位置に対応する航空写真を付加的な画像特徴量として利用する際に、分類カテゴリによって、どの程度、航空写真の特徴量の有効性に違いがあるか分析を行う。具体的には、Multiple Kernel Learning(MKL)を用いて、写真と航空写真の重みを推定することにより、分類における写真と位置情報の有効性の度合いを評価する。実験では、様々なカテゴリについて重みの推定を行い、その結果、航空写真が有効であるカテゴリと、そうでないカテゴリを分類することが可能となった。

**キーワード** 一般画像認識, 位置情報付き画像, 航空写真, Bag of Features, Multiple Kernel Learning

## Estimation of the Significance of Geotagged Photos and Corresponding Aerial Photos Using Multiple Kernel Learning

Keita YAEGASHI<sup>†</sup> and Keiji YANAI<sup>††</sup>

**Abstract** We have been conducting study of exploiting aerial photos as additional image features for visual recognition of geotagged photos. In this paper, we introduce multiple kernel learning (MKL) to evaluate contribution of both features for recognition by estimating the weights of geotagged photo and aerial photo. The experimental results demonstrate effectiveness of usage of aerial photos for recognition of consumer photos.

**Key words** General Object Recognition, Geospatial Information, Aerial Image, Bag-of-Features, Multiple Kernel Learning

### 1. はじめに

今日では、デジタルカメラの普及により、撮影位置の情報を持つ位置情報付き写真はWWW(World Wide Web)上に大量に存在している。また、Flickrなどの写真共有を行うソーシャルサイトの普及により、個人が多くの位置情報付き写真を収集、整理し、共有することが容易になった。一方で、Google Mapsなどのオンラインマッピングサービスも普及し、地図情報の検索機能と共に日々高度化している。地図情報から、デジタル写真や航空写真、Google Street Viewに代表されるような周辺写真情報へアクセスすることはネットユーザーにとってもはや当然のこととなった。

利用者は自宅から地図を利用した位置の検索だけでなく、航空写真の閲覧など高度なサービスを利用できるようになった。また、写真の収集に関しても、位置情報を積極的に活用しようとする試みが行われてきた。その例としてFlickrでは、写真に位置情報を付加して投稿し、地図(Flickr Map)を用いて写真を検索、整理できるようになっている。

大量の写真情報の普及にもかかわらず、それらを自動で整理・分類し、ユーザーの手間を省くことは未だ困難な課題であり続けている。単純な画像分類へのソリューションとして、現状ではタグ(内容を表現する複数の単語の集合)やタイトルなど、テ

キストベースのメタデータが主流になっている。写真を分類するための一般画像認識の基礎技術が高度化する中で、認識精度の向上を図るにあたり、写真と関連する多様な情報をいかに効率的に組み合わせるかが求められる。

我々は、位置情報付き写真の一般画像認識において写真の撮影位置に対応する航空写真を付加的な画像特徴量として利用する研究を行ってきた[6],[7]。画像認識における位置情報の有効性が高い認識カテゴリと、そうでないカテゴリを明確に区別することを目標として位置付けている。

一般に、認識対象と位置は密接な関係があり、例えば、海岸は海と陸の境目にしかなく、海や陸の真中には存在し得ない。しかしながら、海岸の写真の認識において位置情報を役立てるには、世界中の海岸の位置を学習データに持つておかないといけない。これには、膨大なデータを用意する必要がある。そこで、我々は、写真の撮影位置の地理的な状況を表す情報源として、航空写真に注目している。写真の画像特徴量に合わせて、航空写真から抽出した画像特徴量をを認識に用いることで、写真の撮影場所の地理的なコンテキストを反映した認識が可能となると考えている。

本研究では、分類カテゴリによって、どの程度航空写真の特徴量の有効性に違いがあるか分析を行う。従来の我々の研究と同様に、Flickrから収集した位置情報付き写真と、それぞれの



図 1 研究の全体的手順

位置情報に対応する複数の縮尺 (レベル) の航空写真を用いる. 航空写真の特徴量の有効性を明確に判定するにあたり, 機械学習の段階においてマルチカーネル学習 (MKL) を用いて, 写真と航空写真の重みを推定し, 写真と位置情報の有効性を評価する.

本稿は以下のように構成される. まず実験の全体的手順と本研究の方針について第 2 節で触れる. 後述するように, 実験には画像収集, 特徴抽出, 機械学習の手順を踏むが, 特徴抽出で適用する手法の詳細は, 第 3 節で説明する. また, 機械学習に用いる手法については第 4 節で述べる. 実験方法と評価方法, 結果を第 5 節で考察し, 第 6 節で結論付ける.

## 2. 手順と方針

我々の認識実験は全体的に, 図 1 に示す要領で行われる. 全体の流れとしては, Flickr より収集した位置情報付き画像から特徴抽出したものを, 機械学習することで認識精度を検証するものである.

機械学習の段階において, 画像と位置情報がどのように有効に利用されているかどうかを考察する必要がある. これにあたり従来の我々の手法 [6], [7] では, 位置情報と航空写真 (ないしは位置情報) を単純に結合した特徴量を学習させ, 結合する特徴の種類組み合わせによりこれを検証していた. ただし組み合わせごとの精度を比較するのみでは, 画像と位置情報との有効性を柔軟に判断する上では不十分であった.

本研究では, 第 4.2 節で説明する MKL-SVM を用いて, 認識精度のほかに画像と航空写真の特徴の有効性を重みの推定により検証する. 各手順において用いる手法の詳細については後述する.

## 3. 画像からの特徴抽出

本節では, 画像から認識に必要な特徴を抽出する方法について述べる. 画像の特徴を記述する手法としては, 本実験では, 特徴抽出のために局所特徴の一種である SIFT 特徴を用いる. また, この局所特徴を簡潔に記述するために Bag of Keypoints 手法を用いてデータをベクトル量子化する. 画像と航空写真のいずれについても同じ特徴を抽出する.

### 3.1 SIFT 特徴

SIFT(Scale Invariant Feature Transform) [2] とは, David Lowe によって提案された特徴点とそれに付随する特徴ベクトルの抽出法である. 画像からいくつかの特徴点を抽出し, 特徴

点周りの局所画像パターンを勾配方向ヒストグラムから成る 128 次元特徴ベクトルで表現する. SIFT 特徴は, 回転や拡大縮小に対してロバストであるので, 類似画像に対しては類似する特徴点を多く保持することができる.

特徴点の抽出は, 自動で行う方法 (ガウシアン差分による選定) と, 手動で指定する方法 (主に格子状あるいはランダム点指定) がある. 自動で行う手法は, 主に特定の物体同士の対応点を検出するのに有用であるが, 本実験では, 特定の物体を認識することに固執しないので, 10 ピクセルの格子状に点を抽出する方法を採用する.

### 3.2 Bag of visual words 表現

Bag of Keypoints モデル [1] とは, 画像を局所特徴の集合と捉えた手法である. 局所特徴の特徴ベクトルをベクトル量子化し, Visual Words と呼ばれる特徴ベクトルを生成する. それらの集合をコードブックと呼び, それを記述子として画像の特徴ベクトルを生成する. これにより画像を Visual Words の集合 (bag) として表現する. ベクトル量子化は, ユークリッド距離を尺度とし, k-Means 法を用いて行う. 本研究では, 処理速度や結果における精度の差異を考慮した上で, クラスタ数を 1000 に固定した.

## 4. 学習と分類

本節では, 機械学習で用いる手法の詳細を説明する. 認識精度の検証に当たっては, 我々の従来の手法と同様にサポートベクタマシン (SVM) を用いる. 画像と航空写真の有効性について判定するために, マルチカーネル学習を導入する.

### 4.1 サポートベクタマシン

サポートベクタマシン (SVM) はニューロンのモデルとして最も単純な線形しきい素子を用いて, 2 クラスのパターン識別器を構成する手法である. カーネル学習法と組み合わせると非線形の識別器になる. 本実験では, カーネル関数として非線形の  $\chi^2$  カーネルを用いる.

### 4.2 マルチカーネル学習

本研究では, 特徴を統合して画像を認識するために, 複数の特徴量のカーネルを線形結合することにより統合カーネルを作成し, それをサポートベクタマシン (SVM) に適用して特徴統合による画像認識を実現する. 最適なカーネル (カーネルを重みつきで線形結合したカーネル) のサブカーネルに対する重み  $j$  を学習する. これはマルチカーネル学習 (Multiple Kernel Learning, MKL) [3] 問題と呼ばれ, 統合カーネルは以下で定式化される.

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

$$\beta_j \geq 0, \sum_{j=1}^K \beta_j = 1$$

最近の研究では, この MKL 問題を凸面最適化問題として効果的に解く方法が提案されている [4]. マルチカーネル学習は SVM のみを前提としたものではないが, SVM のフレームワークで解く方法が一般的で, MKL-SVM と呼ばれることもある. 本研究

表 1 認識カテゴリの定義

カテゴリ	定義の詳細
ラーメン	食べられる状態にあるラーメン
東京タワー	東京タワーが目立つように映っている写真
景色	遠くが見える景色で、近くに物体がないもの
花	花がアップで撮影されている写真か、写真の殆どが花で占められているような写真
寿司	食べられる状態にある寿司が目立つように映っている写真
自動販売機	自動販売機が目立つように映っている写真
夕日	太陽が映っていて、空が夕日の色に染まっている情景
神社	実際の神社の社殿や、鳥居などの建築と思われるもの
城	日本の伝統的な城の建築物が目立つように映っている写真
祭	日本の伝統的な祭りの情景や、縁日の風景
橋	橋を横、前、下から明らかに確認でき、かつ目立つように映っている写真
桜・花見	桜の風景や花見の風景を映したもの
紅葉	紅葉の情景、または写真の殆どが紅葉で占められているような写真
鉄道	電車又は気動車が目立つように写っている写真
ディズニーリゾート	位置情報などから、ディズニーリゾートで撮影されたと推測できる写真
海岸	実際の海岸が映っている写真
猫	動物としての猫が映っている写真
公園	遊具やベンチ、噴水など公園としての設備が映っている写真

では、SHOGUN [5] ツールキットを用いて実装した MKL-SVM を使用して実験を行う。

## 5. 実験方法

実験は、各画像から抽出した特徴量をサポートベクタマシン (SVM) で学習させ、分類結果により精度を判定することにより行う。学習と分類は正解画像と不正解画像の 2 クラスで行う。正解画像については、画像の内容に応じていくつかのカテゴリのデータセット用意し、不正解画像は正解画像に用いないデータをランダムに選択することで構成される。また、SVM のフレームワークの下で、マルチカーネル学習で種類ごとの重みを推定する (MKL-SVM) ことにより、画像と航空写真のどちらがどれだけ有用であるかを判断する。

### 5.1 データセット

後述する複数のカテゴリを準備し、カテゴリ毎の精度検証を行う。Flickr では大量の画像を収集可能であり、位置情報付き写真は少なくとも 40 万枚は存在する。ただし、それらの写真は Flickr ユーザーの主観的な観点から撮影されたものであり、あるカテゴリにおいてそのカテゴリの写真そのものを描写しているとは限らない。例えばタワーに関する写真はタワーそのもののほかにタワーからの眺め、電車に関する写真は、電車自体のほかに車窓や駅舎などが含まれることがある。

データソースから得られる実験データセットを、より客観的なものに近付けるべく、我々は Flickr 画像を用いて認識の精度検証を行うにあたり、適切な画像を手動で選定することによりデータセットを作成する。

正解データセットについては、表 1 に示すような 18 種類のカテゴリを準備した。これらのデータは、タグなどのメタデータである程度分類したものの中から、カテゴリごとに写真の外観、位置情報ともに適切なものを手動で日本国内で撮影された 200 枚を選定した。正解画像の一例として、図 3 を挙げる事ができる。

カテゴリの選定に当たっては、主に次のような観点から選択した。また、それによる認識精度に関する仮説を示す。

- 東京タワー、ディズニーリゾートは、特定の位置に関するカテゴリである。これらの画像は必然的に位置情報が偏るので、航空写真の特徴量に類似の外観ものが集中するので、航空写真の特徴量に依存する形で認識精度に改善がみられると考えられる。

- 航空写真からその形状を確認できる可能性があるものとして、城、橋、鉄道、海岸、公園が挙げられる。航空写真がある程度詳細なものであれば、これらは把握可能であり、それらの特徴量によって認識精度に若干の改善がみられると考えられる。

- 景色、夕日、桜・花見、紅葉は、航空写真で見えないが風景として捉えられるカテゴリである。景色については、航空写真と関連することも考えられるが、撮影位置と視界の位置が異なるので、航空写真による精度はさほど変化しないと考えられる。夕日、桜・花見、紅葉については時系列的な相関があるが、特定の情報に偏在しないと考えられるので、同様に位置情報による変化は考えにくいであろう。

- 花、神社、猫は、主に屋外で見られるが、航空写真では確認しにくい可能性が高い。これらも同様に位置情報による変化は考えにくいと思われる。

- 寿司、自動販売機、ラーメンは、屋内の物体として挙げられる。これらの物体は、撮影位置が極端に偏向している場合を除いて、撮影位置の航空写真が認識精度の改善に貢献するとは考えにくいと思われる。

不正解データセットは、Flickr から収集した画像データの中から、正解データに用いられていないものをランダムに選定することで 200 枚準備した。

いずれのデータセットも、一般性を保証するため写真の撮影者が偏向しないように選定した。画像は縦と横のうち長い方の画素が 500 ピクセルになるようにリサイズしてある。

カテゴリごとのデータセットの枚数を決定するにあたり、各カテゴリにおいて Flickr から得られる適切な画像の枚数の平均を考慮した。正確な実験結果を保証するためには、一般に大量のデータセットを構築する必要があるが、本研究においては身近なデータからデータセットを構築することを重視した上で、今後万データソースを見直す場合は認識精度の比較検討と共に考慮することになる。

### 5.2 航空写真

航空写真を利用するにあたり、それぞれの画像に対応する位



図 2 位置情報付き写真と航空写真の対応付け

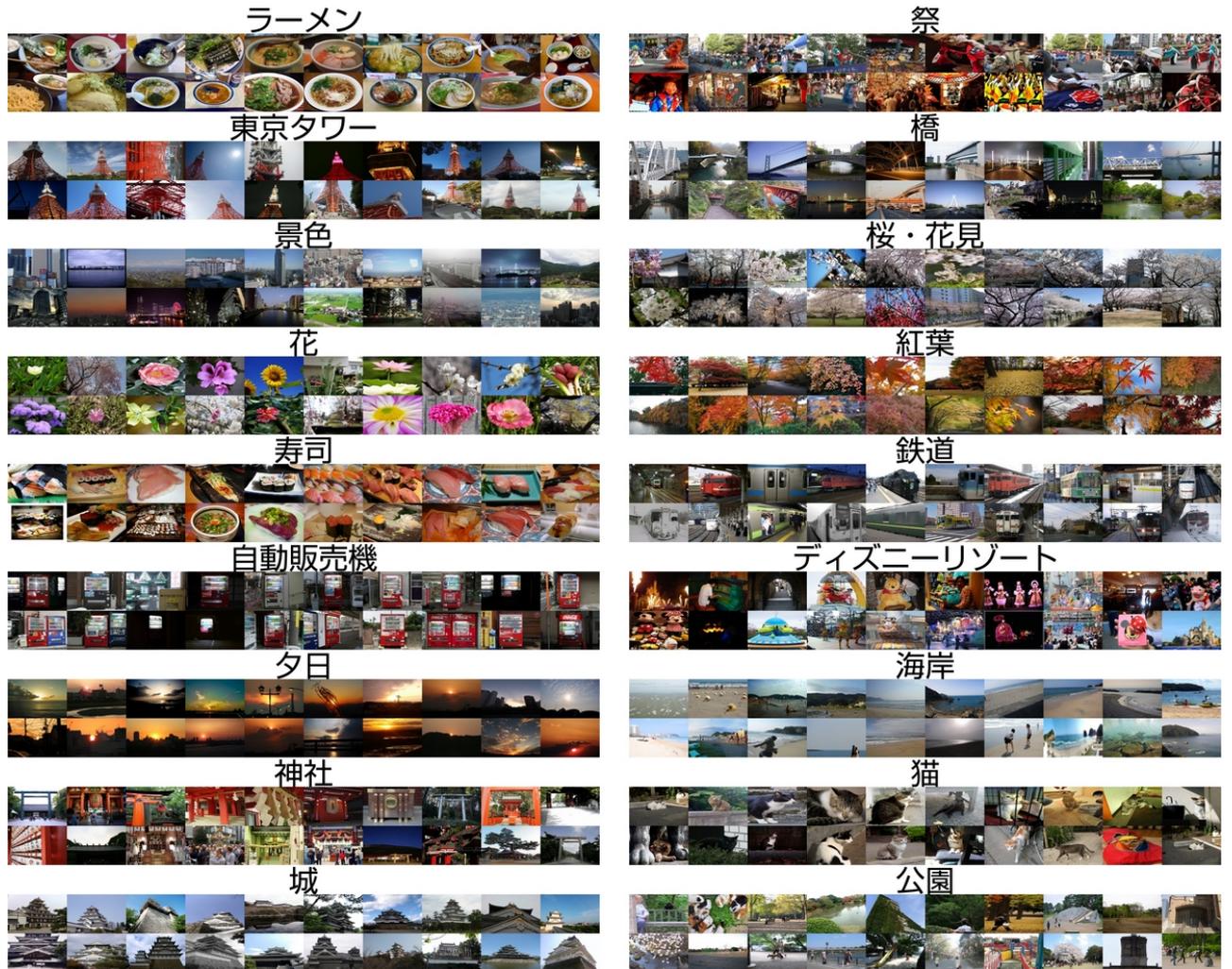


図 3 実験に用いる正解データセットの例。上からディズニーリゾート、花、景色、山、ラーメンを、右列はそれぞれ上から道路、海岸、神社、東京タワー、鉄道を示す。

置情報を地図サイトで表示し、スクリーンキャプチャしたものを特徴抽出に利用する。特徴抽出の方法はデータセットの画像と同様である。写真の内容によっては、より詳細な縮尺を持つ航空写真の方が検証に際し有効であると考えられるが、撮影位置によっては、詳細な航空写真を入手できないこともある。広範なデータセットに対応するため、図 2 に示すような 4 種類のスケールを採用する。図 2 に示すように、航空写真は撮影位置が中央にくるような 1 辺 256 ピクセルの正方形に加工したものを利用する。

### 5.3 評価

認識実験に際しては、後述のように写真の内容を予め各カテゴリに分類し、カテゴリごとに、画像と航空写真をペアにしたものを MKL-SVM を用いて学習・分類・重みの推定を行う。

学習と分類はクロスバリデーションで行う。すなわち、データセットを 5 分割し、1 つを分類用に、残り 4 つを学習用に充てることを、分類用に相当する各 fold について 5 通り繰り返す。

認識精度は各分類結果における SVM の出力値に対して、以下のように平均適合率を計算することで評価する。SVM による出力値に基づいて、テストデータをソートする。ソートした

表 2 MKL-SVM における平均適合率の計算結果. Level1~4 は、それぞれ画像とそのレベルの航空写真を MKL カーネルとしたときの結果である. Image は画像のみの分類結果である.

カテゴリ	Image	Level1	Level2	Level3	Level4	Multi
城	81.28	83.28	83.48	83.46	83.42	<b>83.53</b>
猫	73.98	74.52	74.58	<b>75.01</b>	74.20	74.26
ディズニーリゾート	68.00	83.97	83.97	<b>84.12</b>	83.88	84.06
橋	69.51	71.77	73.56	73.47	74.23	<b>74.89</b>
花	77.00	77.78	78.15	78.18	<b>78.53</b>	78.48
紅葉	79.18	81.64	81.93	81.92	82.12	<b>82.45</b>
自動販売機	83.17	83.44	83.26	83.41	83.40	<b>83.43</b>
祭	72.32	74.34	73.94	76.90	<b>77.84</b>	77.75
桜・花見	80.94	81.36	81.35	81.57	81.05	<b>81.61</b>
公園	67.43	71.99	70.13	72.97	75.69	<b>76.04</b>
景色	73.71	77.88	77.49	<b>78.45</b>	77.29	78.37
ラーメン	82.34	82.54	82.59	82.63	82.67	<b>82.70</b>
神社	72.79	75.37	74.71	76.86	<b>78.96</b>	78.53
寿司	80.11	81.84	81.63	81.91	<b>82.11</b>	81.93
夕日	83.01	82.94	83.13	83.09	83.07	<b>83.11</b>
海岸	80.10	83.73	<b>84.01</b>	83.79	83.36	83.85
東京タワー	80.84	83.67	83.77	<b>83.89</b>	83.85	83.84
鉄道	74.70	75.76	74.96	75.17	<b>76.24</b>	76.20

データを最初から順番に読み込み、Positive のデータが出現した時点で、それまでの読み込んだデータの数  $m_i$  と、Positive データの出現頻度  $i$  を記録する. ここで  $p_i = \frac{i}{m_i}$  とおく. 最後までテストデータを読み込んだときのすべての Positive データの出現数を  $n$  とすると、平均適合率 (Average Precision) は、

$$P = \frac{1}{n} \sum_{i=1}^n p_i \quad (2)$$

で計算される.

#### 5.4 実験結果

画像と航空写真を MKL-SVM で分類した結果について示す. ここでは、画像と航空写真の 2 組について MKL-SVM で分類、重み推定を行った. 結果については、5 回のクロスバリデーションの平均値を示す.

平均適合率の結果を表 2 に示す. 表 2 における Level1~4 は、画像と、航空写真のいずれかの縮尺のものを MKL カーネルとして入力した結果で、比較のため、画像のみを分類させたものをベースラインとして右端の Image に示す. この結果から、概ね殆どのカテゴリにおいて、MKL の利用によって、航空写真が認識精度に改善に影響を及ぼしたことがわかる. 位置情報が単なる付加的情報ではなく、適切な重みにより分類されていることを確認するために、表 4 に画像と位置情報の 2 カーネルで MKL を行った時の重みの比較を示す.

表 4 における LV1~4 は、画像と、航空写真のいずれかの縮尺のものを MKL カーネルとして入力した場合のものを示す. また、位置情報として 4 種類の航空写真を画像と併せて 5 カーネルで行った時の重みの比較を表 3 に示す.

いずれも全体として、画像特徴としての位置情報を適度に利用した上で認識精度を上げたものであると考えてよい.

以下、各カテゴリについて結果を考察する.

- 東京タワー、ディズニーリゾートは、特定の位置に関するカテゴリである. 航空写真で比較的レベルの低いものが効いているのは、レベルの低い航空写真の方がランドマークの外観をよくあらわしているからであると考えられる. すなわち、詳細な航空写真では撮影位置で共通の航空写真の特徴量が得にくいことが言える.

- 航空写真からその形状を確認できる可能性があるものとして、海岸の特徴量は、比較的レベルの低い航空写真で特徴が現れやすい. これはレベルの低い航空写真で、海面の画像特徴を多く学習しやすいためであると考えられる. 一方、橋、鉄道、海岸、公園については、線路やグラウンド、道路の特徴といった、レベルの高い航空写真で確認できる特徴量が効いていると推測される.

- 景色、夕日、桜・花見、紅葉は、航空写真で見えないが風景として捉えられるカテゴリである. これらは位置情報による変化は考えにくいものと仮定していたが、実際には若干航空写真の重みが付いており、2 つのカーネルで比較した場合、レベルの高い航空写真ほど重みが高くなっている. これについては、例えば高層ビル街で壮大な夕日や紅葉が見られないように、撮影位置に関する相関があるものと推測される.

- 花、神社、猫のような屋外で見られるものについても前項と同様に、レベルの高い航空写真ほど重みが高くなっているため、撮影位置と相関があると考えられる.

- 寿司、自動販売機、ラーメンは、屋内の物体として挙げられる. これらの重みは他の位置情報とは関係の薄いカテゴリと比べて若干高い. これらは位置情報による改善というよりは、むしろ位置情報がある程度偏向していることを示している. この種のカテゴリは撮影者の趣向に依存する傾向があるため、よりの確な精度を得るには、収集の段階からデータセットの作成を見直す必要があると考える.

Flickr から得られる写真は個人により撮影されているものである以上、位置情報の偏向については常に考慮しなければならない問題であるが、これについてはデータの大量収集と効率的なデータ・カテゴリの選定に対して包括的に配慮することで、引き続き対処していきたいと考える.

## 6. おわりに

本研究では、位置情報付き写真の一般画像認識において写真の撮影位置に対応する航空写真を付加的な画像特徴量として利用する研究手法を改良した. 撮影位置の情報は画像認識の精度向上に対していかに貢献するかを検討するにあたり、SVM による機械学習において得た認識精度と、マルチカーネル学習 (MKL) での重み推定による写真と位置情報の有効性の度合いを評価した. 本実験により、航空写真を利用した分類が有効であるカテゴリが存在することが示された.

今後予定する取り組みとしては、現状より更に多様なカテゴリにおいて分類を行い、画像認識における位置情報の有効性についてより明確にする必要があることである. これを実現するにあたり、豊富でかつ明確に区別されていない画像ソースの中

表 3 MKL による重み推定結果. 位置情報付き写真と位置情報 (航空写真 (Level1~4)) のすべ  
てのカーネルで重みを推定した結果を示す.

カテゴリ	画像	Level1	Level2	Level3	Level4	カテゴリ	画像	Level1	Level2	Level3	Level4
猫	<b>0.742998</b>	0.028288	0.008282	0.062544	0.157888	公園	<b>0.436722</b>	0.072962	0.01194	0.045496	0.43288
ディズニー	<b>0.384238</b>	0.095286	0.235888	0.131498	0.153086	ラーメン	<b>0.87288</b>	0.002266	0	0.036764	0.088094
橋	<b>0.582488</b>	0.077	0.044412	0.07038	0.225718	景色	<b>0.604498</b>	0.078468	0.024184	0.093446	0.199406
城	<b>0.523238</b>	0.166148	0.00366	0.098968	0.207982	夕日	<b>0.728526</b>	0.055446	0.05768	0.016434	0.141914
花	<b>0.658232</b>	0	0.042462	0.050614	0.248698	海岸	<b>0.392038</b>	0.115284	0.172952	0.055222	0.264498
祭	<b>0.518292</b>	0.05837	0.000702	0.185068	0.237564	神社	<b>0.507376</b>	0.032762	0.008748	0.060598	0.390518
紅葉	<b>0.522908</b>	0.140746	0.005912	0.061672	0.268764	寿司	<b>0.59562</b>	0.062254	0.014692	0.061694	0.265736
自販機	<b>0.793696</b>	0.058212	0.008928	0.074274	0.064892	東京タワー	0.363714	0.00788	0.002154	<b>0.395538</b>	0.230714
桜・花見	<b>0.773604</b>	0.037922	0.005624	0.092982	0.089868	鉄道	<b>0.604114</b>	0.105688	0.014196	0.052092	0.223914

表 4 MKL による重み推定結果. 位置情報付き写真と位置情報 (航空写真 (LV1~4)) の 2 つの  
カーネルで重みを推定した結果を示す.

カテゴリ	レベル	画像特徴	位置情報	カテゴリ	レベル	画像特徴	位置情報	カテゴリ	レベル	画像特徴	位置情報
城	LV1	<b>0.6744</b>	0.3256	自販機	LV1	<b>0.882315</b>	0.117685	神社	LV1	<b>0.790665</b>	0.209335
	LV2	<b>0.679645</b>	0.320355		LV2	<b>0.911495</b>	0.088505		LV2	<b>0.828445</b>	0.171555
	LV3	<b>0.547845</b>	0.452155		LV3	<b>0.85224</b>	0.14776		LV3	<b>0.6403</b>	0.3597
	LV4	<b>0.5388</b>	0.4612		LV4	<b>0.8602</b>	0.1398		LV4	<b>0.524385</b>	0.475615
猫	LV1	<b>0.917985</b>	0.082015	祭	LV1	<b>0.807035</b>	0.192965	寿司	LV1	<b>0.763045</b>	0.236955
	LV2	<b>0.936535</b>	0.063465		LV2	<b>0.852605</b>	0.147395		LV2	<b>0.814605</b>	0.185395
	LV3	<b>0.8626</b>	0.1374		LV3	<b>0.5972</b>	0.4028		LV3	<b>0.690875</b>	0.309125
	LV4	<b>0.80013</b>	0.19987		LV4	<b>0.54077</b>	0.45923		LV4	<b>0.60764</b>	0.39236
ディズニー	LV1	0.486945	<b>0.513055</b>	桜・花見	LV1	<b>0.913235</b>	0.086765	夕日	LV1	<b>0.84347</b>	0.15653
	LV2	<b>0.53557</b>	0.46443		LV2	<b>0.92839</b>	0.07161		LV2	<b>0.85067</b>	0.14933
	LV3	0.386825	<b>0.613175</b>		LV3	<b>0.846985</b>	0.153015		LV3	<b>0.876925</b>	0.123075
	LV4	0.174835	<b>0.825165</b>		LV4	<b>0.829465</b>	0.170535		LV4	<b>0.786185</b>	0.213815
橋	LV1	<b>0.772925</b>	0.227075	公園	LV1	<b>0.78473</b>	0.21527	海岸	LV1	0.44391	<b>0.55609</b>
	LV2	<b>0.783125</b>	0.216875		LV2	<b>0.86537</b>	0.13463		LV2	<b>0.516065</b>	0.483935
	LV3	<b>0.694935</b>	0.305065		LV3	<b>0.649275</b>	0.350725		LV3	<b>0.5154</b>	0.4846
	LV4	<b>0.60599</b>	0.39401		LV4	0.437185	<b>0.562815</b>		LV4	0.461555	<b>0.538445</b>
花	LV1	<b>0.85663</b>	0.14337	景色	LV1	<b>0.80023</b>	0.19977	東京タワー	LV1	<b>0.534385</b>	0.465615
	LV2	<b>0.857575</b>	0.142425		LV2	<b>0.8355</b>	0.1645		LV2	<b>0.58798</b>	0.41202
	LV3	<b>0.75346</b>	0.24654		LV3	<b>0.66089</b>	0.33911		LV3	0.48116	<b>0.51884</b>
	LV4	<b>0.66366</b>	0.33634		LV4	<b>0.601115</b>	0.398885		LV4	0.281265	<b>0.718735</b>
紅葉	LV1	<b>0.765895</b>	0.234105	ラーメン	LV1	<b>0.96695</b>	0.03305	鉄道	LV1	<b>0.80733</b>	0.19267
	LV2	<b>0.744415</b>	0.255585		LV2	<b>0.968095</b>	0.031905		LV2	<b>0.874365</b>	0.125635
	LV3	<b>0.68315</b>	0.31685		LV3	<b>0.90939</b>	0.09061		LV3	<b>0.754885</b>	0.245115
	LV4	<b>0.571985</b>	0.428015		LV4	<b>0.87075</b>	0.12925		LV4	<b>0.677985</b>	0.322015

から必要ないしは有効なものを効率的に収集・選定し, データセット作成を強化することを考慮する. また, 航空写真においても, より明確な特徴を持つ画像を利用する必要があるために, より詳細な航空写真を用いて実験することを視野に入れる.

## 文 献

- [1] 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.
- [2] D. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91–110, 2004.
- [3] G. R. G. Lanckriet, N. Cristianini, P. Bartlett, L. E. Ghaoui, and M. I. Jordan. Learning the kernel matrix with semidefinite programming. *Journal of Machine Learning Research*, Vol. 5, pp. 27–72, 2004.
- [4] S. Sonnenburg, G. Rätsch, C. Schäfer, and B. Schölkopf. Large scale multiple kernel learning. *Journal of Machine Learning Research*, Vol. 7, pp. 1531–1565, 2006.
- [5] Shogun - A Large Scale Machine Learning Toolbox. <http://www.shogun-toolbox.org/>
- [6] K. Yaegashi and K. Yanai. Can Geotags Help Image Recognition?. *Proc. of the Pacific-Rim Symposium on Image and Video Technology*, pp. 361–373, 2009.
- [7] 八重樫 恵太, 柳井 啓司: 撮影位置の情報を用いた一般画像認識の可能性の検討, 情報処理学会 CVIM 研究会, pp.15–22, CVIM163-3, (2008)