

## グラフ構造表現による一般物体認識

堀 貴博<sup>†</sup> 滝口 哲也<sup>††</sup> 有木 康雄<sup>††</sup>

<sup>†</sup> 神戸大学大学院システム情報学研究科 〒 657-8501 兵庫県神戸市灘区六甲台町 1-1

<sup>††</sup> 神戸大学自然科学系先端融合研究環 〒 657-8501 兵庫県神戸市灘区六甲台町 1-1

E-mail: †horitaka@me.cs.scitec.kobe-u.ac.jp, ††{takigu,ariki}@kobe-u.ac.jp

あらまし コンピュータによる一般物体認識は、近年、ロボットビジョンや画像検索といった分野で、実現が強く求められている。従来研究には、画像から局所的特徴である SIFT(Scale-Invariant Feature Transform) を抽出し、クラスタリングにより出現頻度ヒストグラム表現に変換する Bag of Features があるが、画像全体を対象とするために位置情報や特徴点間の関係性が失われるという問題がある。この問題に対応するため、提案手法では、SIFT 特徴点を線でつなぎ、グラフ化することで特徴点間に関係性を持たせ、位置情報を伴う構造表現を実現する。しかし、グラフ表現は統計処理に適していない。そこで、グラフをグラフ編集距離によりベクトル表現に変換してクラス分類を行う。複数の画像データセットを用いた評価実験により、提案手法は従来手法と比較して認識精度が向上することを示す。

キーワード 一般物体認識, グラフ, SIFT, グラフ編集距離, グラフ-ベクトル変換

## Generic Object Recognition by Graph Structural Expression

Takahiro HORI<sup>†</sup>, Tetsuya TAKIGUCHI<sup>††</sup>, and Yasuo ARIKI<sup>††</sup>

<sup>†</sup> Graduate School of System Informatics, Kobe University 1-1 Rokkodai, Nada-ku, Kobe, 657-8501 Japan

<sup>††</sup> Organization of Advanced Science and Technology, Kobe University 1-1 Rokkodai, nada, Kobe  
657-8501, Japan

E-mail: †horitaka@me.cs.scitec.kobe-u.ac.jp, ††{takigu,ariki}@kobe-u.ac.jp

**Abstract** This paper describes a method for generic object recognition using graph structural expression. In recent years, generic object recognition by a computer is strongly requested in the field like the robot vision and the image retrieval. Conventional methods use bag-of-features (BoF) that express the image as an appearance frequency histogram of visual words by quantizing SIFT (Scale-Invariant Feature Transform) features. However, there is a problem that the location information and the relations between keypoints are lost that are important as structural information. To deal with this problem, in the proposed method, the graph is constructed by connecting SIFT keypoints by the line. As a result, the keypoints have the relation, and then structural representation with location information is achieved. Since the graph representation is not suitable for the statistical work, the graph is embedded into a vector space according to the graph edit distance. As a result of two image datasets, the proposed method has improved the recognition rate compared with conventional methods.

**Key words** generic object recognition, graph, SIFT, graph edit distance, vector-space embedding

### 1. ま え が き

一般物体認識とは、制約のない図 1 のような実世界シーンを撮影した画像に含まれる物体を、計算機が一般的名称で認識することを指し、コンピュータビジョンの分野において困難な研究課題の一つである。機械による人間の高次視覚機能の実現という観点から、ロボットビジョンへの応用などが期待されている。また近年、デジタルカメラの普及、ハードディスクの大容量

化、Web の発達に伴いデジタル画像が爆発的に増大し、人手による管理が困難になっている。そこで計算機による画像の検索や分類が求められており、一般物体認識の重要性がますます高まってきている。

一般物体認識に関する従来研究のアプローチには、2 つの代表的な方法がある。ひとつは、領域に基づく方法である。これは、領域分割された画像領域に対して自動アノテーションを行う手法であり、Barnard らによる word-image-translation

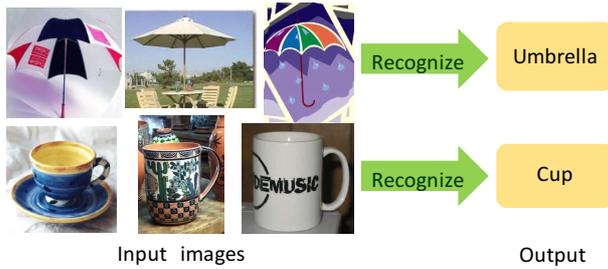


図 1 一般物体認識

Fig.1 Generic Object Recognition

model [1] [2] [3] などが有名である。しかしこの方法は、オクルージョンがある場合や、画像が複雑で領域分割が上手くいかない場合には、対処することが難しい。

そこで、この問題を解決するために、局所パターンに基づく方法が提案されている。これは、画像の局所的な特徴の組み合わせによって、画像の照合を行う手法である。これには、図 2 のような Bag of Features という局所の特徴の出現頻度ヒストグラムを求め、画像全体を特徴付ける手法がよく用いられる [4]。しかし、画像全体を対象とするために、位置情報や特徴間の関係性が失われるという問題がある。

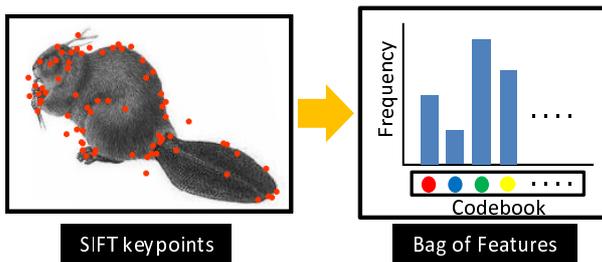


図 2 Bag of Features

Fig.2 Bag of Features

本研究ではこの問題に対処するために、図 3 に示すように特徴点間を線で結び、局所の特徴の集合をグラフとして表現する。また、そのグラフをグラフ編集距離を用いて統計的学習が容易なベクトル表現に変換することで、物体構造と局所の特徴を統合した認識性能の高い手法を提案する。

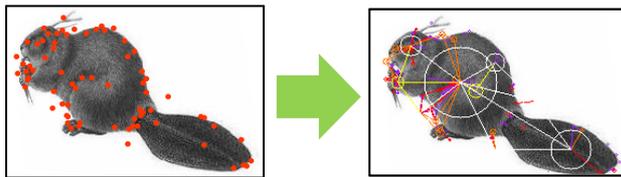


図 3 グラフの作成

Fig.3 Graph construction

本論文の構成は以下の通りである。2 章で、提案手法の流れについて述べ、3 章で画像をグラフとして表現する方法について述べる。4 章ではグラフ編集距離について述べ、5 章でグラフをベクトル表現に変換する方法について述べる。6 章では、提案手法の有効性を複数の画像データセットで評価した結果に

ついて述べ、7 章で、本論文のまとめと今後の方針について述べる。

## 2. 提案手法の流れ

提案手法の流れを図 4 に示す。まず、全ての画像について SIFT 特徴 [5] を抽出する。抽出された特徴点をつないで各画像のグラフを作成する。学習画像から作成したグラフをトレーニンググラフ、テスト画像から作成したグラフをテストグラフとする。次に、トレーニンググラフの中から  $n$  個のプロトタイプグラフを選び出し、各グラフ (トレーニンググラフ及びテストグラフ) と  $n$  個のプロトタイプグラフとの間で、グラフ編集距離 GED (Graph Edit Distance) を計算する。こうすることで、グラフを  $n$  次元ベクトル空間に埋め込む。得られた学習データの  $n$  次元ベクトルを用いてクラス分類器を学習する。学習したクラス分類器により、テストデータを分類して認識結果を出力する。

次に、提案手法の各処理について、それぞれ詳細に述べる。

## 3. 画像のグラフ構造表現

本研究では [6] で提案したグラフの記法、構造表現を採用する。

### 3.1 グラフの定義

グラフ  $G$  を  $G = (V, E, X)$  と定義する。 $V$  はノードの集合、 $E$  はエッジの集合、 $X$  はノードの関連単項測定値の集合を表す。ノードは SIFT によって検出された特徴点であり、その関連単項測定値は、対応する特徴点に関する 128 次元の特徴量である。また、2 つのノード  $u_\alpha$  と  $u_\beta$  間のエッジは  $e_{\alpha\beta}$  と表す。以降では、プロトタイプグラフを  $G^p$ 、それ以外のグラフをシーングラフと呼び  $G^s$  として区別する。

### 3.2 近接グラフ

画像から抽出した特徴点をエッジで接続してグラフを作成するが、全ての点を相互に接続すると完全グラフとなり、通常それは計算上適していない。また、遠方の特徴点との関係性は薄いので、近傍の特徴点のみと接続するのが望ましい。そこで、遠方の特徴点に関連していないグラフを近接グラフと定義し、以下の式により近接グラフを構成するエッジを作成する。

$$E = \left\{ e_{ij} \mid \forall i, j, \frac{\| \mathbf{p}_i - \mathbf{p}_j \|}{\sqrt{\sigma_i \sigma_j}} < \chi \right\} \quad (1)$$

ここで、 $\mathbf{p} = (p_x, p_y)$  は特徴点の座標、 $\sigma$  はスケール、 $\chi$  は定数を表す。この式により、スケールの大きい特徴点ほどより遠方の特徴点と接続し、定数  $\chi$  以上の場合、エッジは作成されない。余分なエッジが作成されないので、作成された近接グラフはコンピュータの負担をかなり軽減し、同時に検出性能を改良する。プロトタイプグラフ、シーングラフともに近接グラフとして作成する。

### 3.3 擬似階層グラフ

一般的に、SIFT 特徴点は、スケールが大きいと高い信頼度を示す。そこで、特徴点のスケールの大きさをグラフを階層に分割し、信頼度の高い大きいスケールの階層からグラフマッチ

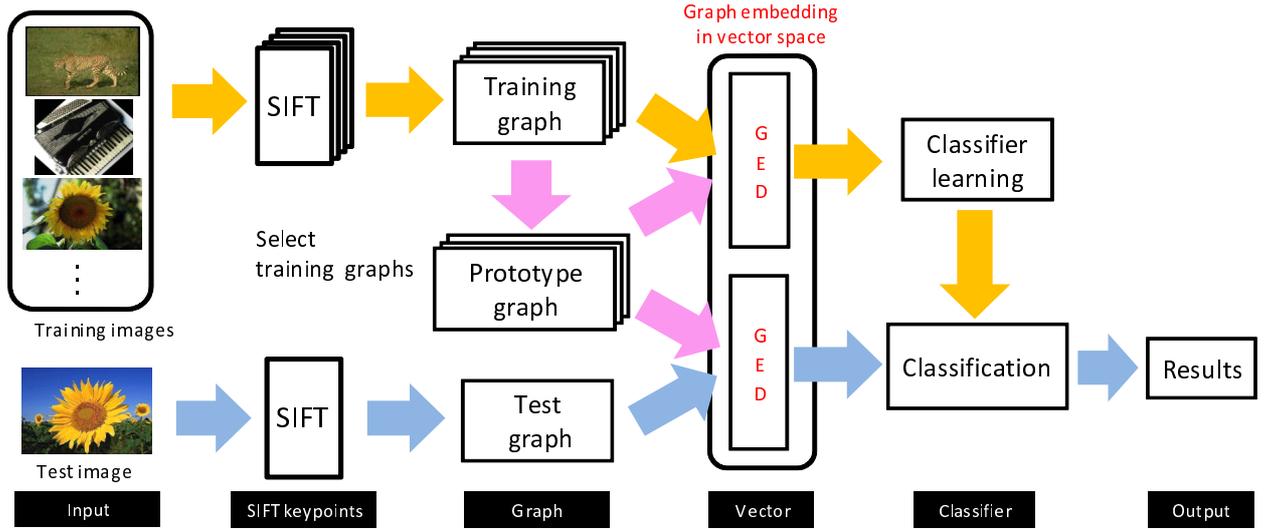


図 4 提案手法の流れ

Fig. 4 System overview

ングを開始して、徐々に階層を下げていくことにより認識性能の向上を図る。これを擬似階層グラフと定義する。また、上位の階層でマッチングしない場合には、早期に処理を終了することで、全てのノードを比較する必要がなくなり、計算時間が短縮できる。グラフの階層数を  $L$ 、階層レベルを  $l$  とする。特徴点のスケールに基づいてグラフを部分グラフ  $\{G_l\}_{l=1}^L$  の集合に分解するには、次式に示す閾値  $s_l$  を用いる。

$$s_l = \sigma_{min} \left( \frac{\sigma_{max} - \sigma_{min}}{\sigma_{min}} \right)^{\frac{L-l}{L-1}} \quad (2)$$

ここで、 $\sigma_{max}$ 、 $\sigma_{min}$  は各グラフにおける特徴点のスケールの最大値と最小値である。各階層レベル  $l$  において、スケール  $\sigma$  が閾値  $s_l$  より大きい特徴点のみを保有する。2つのグラフをグラフマッチングする際、両方ではなく片方のグラフのみを擬似階層グラフとする。これは、各グラフのマッチングすべきノードが、同じスケール階層にあるとは限らないからである。本研究では、プロトタイプグラフを近接グラフとして作成し、擬似階層グラフに変換する。階層数 3 に分割したプロトタイプグラフ  $\{G_l^p\}_{l=1}^3$  の例を図 5 に示す。

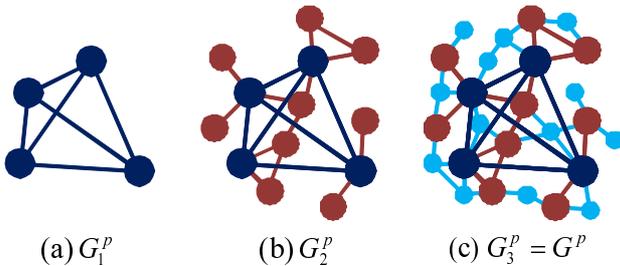


図 5 プロトタイプグラフ

Fig. 5 Prototype Graph

#### 4. グラフ編集距離

2つのグラフの構造的類似性を評価する過程は、一般的にグ

ラフマッチングと呼ばれる。この問題に対して多くの研究が行われている [7]。本論文ではその中でグラフ編集距離 [8] [9] を使用する。これは、2つのグラフの相違を計算するためによく用いられる方法であり、多くの最適近似アルゴリズムが提案されている [10]。それらの論文で主に問題とされているのはグラフ編集距離の計算量であり、それを改善するために本論文では擬似階層グラフを用いる。

グラフ編集距離の基本的な考え方は、1つのグラフを他のグラフに変換するのに必要である最小の編集数として、2つのグラフの相違を定義することである。つまり、ノードとエッジの挿入、削除、置換で構成される編集操作  $e_d$  の数で定義される。全てのグラフの組  $G_1$  と  $G_2$  において、特定の編集操作で  $G_1$  を  $G_2$  に変換する編集経路の集合  $h(G_1, G_2) = (e_{d1}, \dots, e_{dk})$  (ここで各  $e_{di}$  が編集操作を表す) が存在する。図 6 で 2つのグラフ  $G_1$  と  $G_2$  間の編集経路の例を与える。この編集経路は、(a) 1つのエッジの削除、(b) 1つのノードの置換、(c) 1つのノードの挿入、および (d) 2つのエッジの挿入から成る。一般的に、2つのグラフ間の編集経路は複数存在する。どの編集経路が最良かを評価するために、編集コストを導入する。基本的な考え方は、それぞれの編集操作によって起こるひずみを、編集コスト  $c$  で表すことである。2つのグラフ  $G_1$  と  $G_2$  間の編集距離  $d$  は、 $d(G_1, G_2)$  で表され、1つのグラフを他のグラフに変換する最小コストである編集経路のコストとなる。それは以下の式によって表される。

$$d(G_1, G_2) = \min_{(e_{d1}, \dots, e_{dk}) \in h(G_1, G_2)} \sum_{i=1}^k c(e_{di}) \quad (3)$$

ここで、 $h(G_1, G_2)$  は  $G_1$  から  $G_2$  へ変換する編集経路のセット、 $c(e_d)$  は編集操作  $e_d$  の編集コストを示す。編集コストは全て定数として定義する。

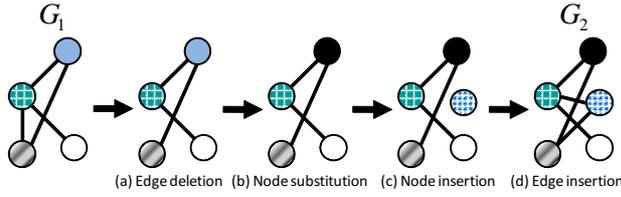


図 6 グラフ間の編集経路

Fig. 6 An edit path between two graphs  $G_1$  and  $G_2$

## 5. グラフ-ベクトル変換

本研究で用いる変換手法は、[11] で提案された手順に従う。この手法は、メディアングラフの導出などに使用され、その有効性が示されている [12] [13]。以下でそのアプローチについて述べ、その概略を図 7 に示す。

まず、トレーニンググラフのセット  $T = \{G_1, G_2, \dots, G_n\}$  を用意する。次に、 $m$  個のプロトタイプグラフのセット  $P = \{G_1^p, \dots, G_m^p\} \in T$  を、 $T (m \leq n)$  から選ぶ。その後、シーングラフ  $G^s$  と全てのプロトタイプグラフ  $G_k^p \in P$  の相違を計算する。従って、各シーングラフについて、 $m$  個の相違評価値  $d_1, \dots, d_m$  (ここで  $d_k = d(G^s, G_k^p)$ ) が求められる。この評価値を  $m$  次元ベクトルとする。このように、プロトタイプグラフセット  $P$  を用いて、全てのシーングラフ  $G^s$  を  $m$  次元ベクトルに変換できる。次式のように埋め込み手順を定義する。

$$\psi : G^s \rightarrow \mathbf{R}^m \quad (4)$$

$$\psi \rightarrow (d(G^s, G_1^p), d(G^s, G_2^p), \dots, d(G^s, G_m^p)) \quad (5)$$

ここで、 $d(G^s, G_i^p)$  はグラフ相違評価値、つまりグラフ編集距離である。

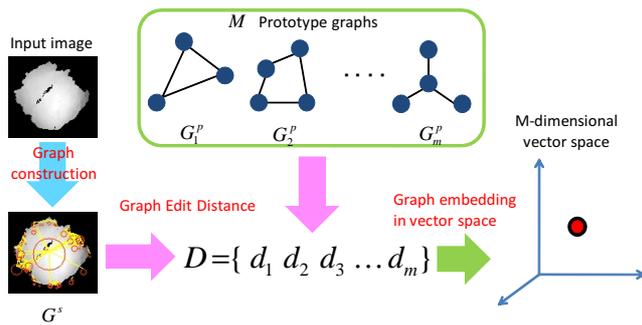


図 7 グラフ-ベクトル変換

Fig. 7 Graph embedding in vector spaces

本研究では、この定義に従って、グラフの埋め込みを実行する際に、プロトタイプグラフ  $P$  のセットを、トレーニンググラフ  $T$  のセットから選択する。その選択方法には、グラフ編集距離の和による画像の順位付けを用いる。ある物体クラス  $c$  に属するトレーニンググラフのセットを  $T_c$ 、 $\exists G_i \in T_c$ 、 $\tilde{T}_c = T_c \setminus G_i$  として、次式により、 $G_i$  と同一のクラスに属するトレーニンググラフ間のグラフ編集距離の和  $D(G_i)$  を求める。

$$D(G_i) = \sum_{G_j \in \tilde{T}_c} d(G_i, G_j) \quad (6)$$

$T_c$  において、この  $D(G_i)$  の小さいほうが上位になるように順位付けを行う。これにより、クラス内のグラフ間の類似度がより高いものが上位に順位付けされる。よって、この順位の上位からグラフを選ぶことで、そのクラスで代表的な形状を持つグラフをプロトタイプグラフとして選択することができる。

## 6. 評価実験

本章では、10 クラスの物体の分類実験を行い、提案手法であるグラフ構造表現を用いた一般物体認識と、従来手法である BoF を比較することで、その有効性を評価する。

### 6.1 実験条件

実験には、Caltech-101 データベース [14] と PASCAL Visual Object Challenge (VOC) 2007 データセット [15] を用いる。

Caltech-101 データベースは、一般物体認識を対象とするデータベースで、101 クラスで構成され、同一クラスの物体でも色や形が大きく異なる。本実験では、クラス内の画像枚数が 50 枚以上かつ 200 枚以下のクラスからランダムに 10 クラスを選択して用いる。今回実験で用いたクラスの画像の例を、図 8 に示す。



図 8 Caltech-101 データベース

Fig. 8 Caltech-101 dataset

各クラスの画像から、それぞれランダムに 30 枚選んで学習画像とし、残りをテスト画像とする (Caltech-101 はクラスごとに画像数が異なる)。それを 1 セットとして、学習画像の異なる 10 セットのデータセットを作成し、それぞれの分類実験による認識率の平均を、実験結果として評価する。10 クラスの画像 841 枚から選んだ学習画像 300 枚、テスト画像 541 枚のデータセット 10 セットを用いて実験を行う。

PASCAL VOC データセットは、20 個の物体クラスのいずれかに属する物体が含まれる画像で構成される、一般物体認識を目的としたベンチマークである。1 枚の画像に複数の物体が存在し、画像が複数のクラスに属している場合もある。さらに、物体の向き、見た目、姿勢が大きく異なっているので、PASCAL は Caltech よりも難易度が高い。本実験では、画像内の物体の位置を示す方形領域の情報を用いて、画像中の物体

領域を切り出し，対象画像として用いる．ランダムに 10 クラスを選択し，学習画像 4986 枚，テスト画像 4863 枚を用いて実験を行う．画像の分類実験による認識率を，実験結果として評価する．PASCAL の対象画像の例を，図 9 に示す．



図 9 PASCAL Visual Object Challenge (VOC) 2007 データセット  
Fig. 9 PASCAL Visual Object Challenge (VOC) 2007 dataset

また，提案手法で用いる閾値の定数  $\chi, \varepsilon_1, \varepsilon_2$  や擬似階層グラフの階層数  $L$ ，グラフ編集距離の編集コスト  $c$  は実験的に最適な値を用いた．

### 6.2 実験概要

実験は，まず 3 章で述べたように，全ての画像から SIFT 特徴を抽出し，グラフ構造で表現する．次に，トレーニンググラフの中からプロトタイプグラフを選び擬似階層グラフへと変換する．この際，5 章で述べたように，プロトタイプグラフは，トレーニンググラフのセットからグラフ編集距離の和の順位付けにより選択する．次に，4 章，5 章で示したように，シーングラフとプロトタイプグラフの間のグラフ編集距離を求めて，グラフをベクトル空間に埋め込む．Caltech-101 の実験では，学習画像が 300 枚と少ないので，プロトタイプグラフの選択を行わず，プロトタイプグラフはトレーニンググラフと同一のものを用いる．PASCAL を用いた実験では，各クラス画像の順位の上位から 50 枚の画像を選択し，プロトタイプグラフとする．また，従来手法では，BoF の Codebook サイズを 1000 とし，1000 次元でベクトル量子化を行う．本実験では，比較のため，それぞれの手法で作成されたベクトルをクラス分類するための識別器として，k-NN 法 (k-Nearest Neighbor algorithm) と SVM (Support Vector Machine) を用いる．SVM は，基本的には 2 クラスの識別にしか対応していないが，複数の SVM を組み合わせることでマルチクラスに対応できる．

従って，k-NN 法 (Caltech-101 ( $k = 10$ ), PASCAL ( $k = 100$ )), マルチクラス SVM (線形カーネル linear), マルチクラス SVM (非線形カーネル radial basis function (RBF)) の 3 つの識別器によって，画像のクラス分類を行い精度を比較する．

### 6.3 実験結果と考察

Caltech-101 の実験に関する認識率を図 10 に，PASCAL の実験に関する認識率を図 ?? に示す．また，Caltech-101 の実験において，マルチクラス SVM(RBF) を識別器に用いた際の，

各クラスごとの認識率を図 11 に示す．

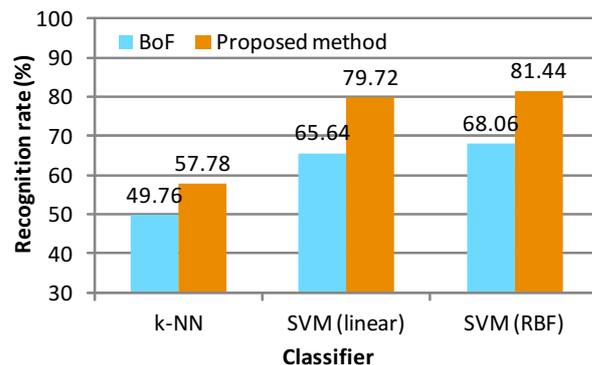


図 10 認識結果 (Caltech-101)  
Fig. 10 Recognition result (Caltech-101)

図 10 と図 ?? より，グラフ構造表現を用いた提案手法を導入することで，二つのデータセットにおける全ての識別器において，従来手法 (BoF) より認識率が向上した．これより，グラフ構造表現をベクトル化した提案手法の有効性を示すことができた．認識率が最も高かったのは，Caltech-101 の実験では，SVM (RBF) を用いた提案手法で，81.44%，PASCAL の実験では，SVM (RBF) を用いた提案手法で，38.25%であった．また，全ての実験の中で，従来手法 (BoF) と比較して最も認識率が向上したのは Caltech-101 の実験での SVM (linear) であり，14.08 ポイントの向上を示した．図 11 において，Caltech-101 の実験での認識率をクラスごとに見ると，leopards, soccer\_ball, stop\_sign, umbrella の認識率が，BoF に比べて大幅に向上している．

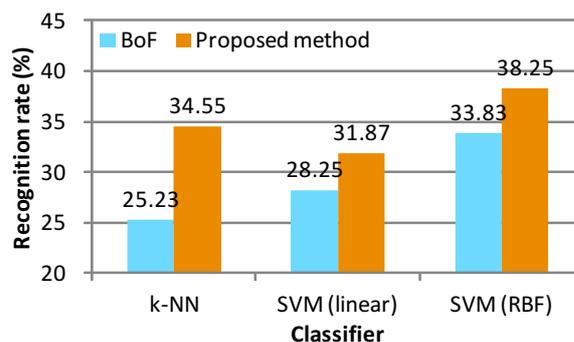


図 11 認識結果 (PASCAL)  
Fig. 11 Recognition result (PASCAL)

このように，従来手法 BoF よりも認識率が向上した理由について，以下に考察する．本研究で用いた SIFT 特徴量は，画像の輝度勾配情報から求められる．従って，輝度勾配が少ない，あるいは同一クラス他画像と著しく輝度が異なると抽出した特徴量の値が適切ではないので，従来手法の BoF では認識率が低下する．これに対して形状によって物体を比較する提案手法では，特徴量の値が適切でなくても，特徴点さえ抽出できていればグラフを作成できるので，認識率の低下は起こらない．図 12 は，提案手法で認識が成功し，従来手法で認識が失敗し

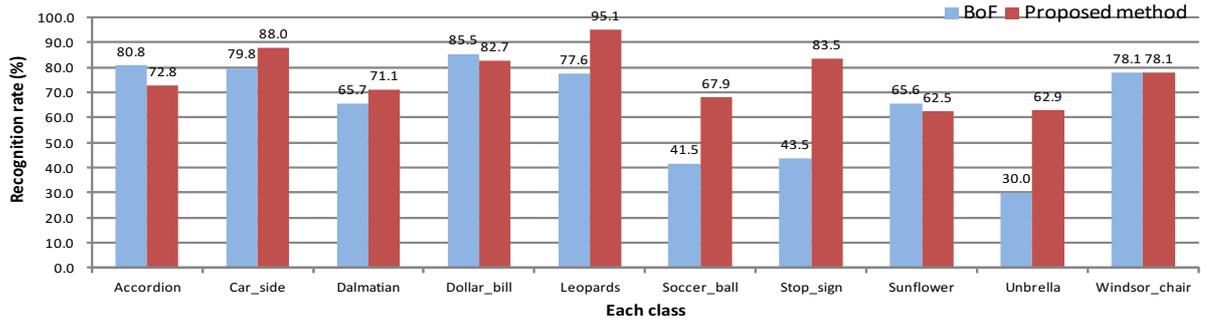


図 12 各クラスの認識結果 (SVM(RBF))(Caltech-101)

Fig. 12 Recognition result of each class (SVM(RBF))(Caltech-101)

た画像である。下段右から 2 番目の傘画像には、ほとんど輝度情報が含まれていないが、提案手法では正しく認識できていることが分かる。

以上より、グラフ構造表現を一般物体認識に導入することは、有効であることが分かった。しかし、導入方法には改善の余地があると思われる。特にグラフ編集距離の計算は、擬似階層グラフや近接グラフを用いることで軽減したとはいえ、まだかなりの時間がかかる。また、今回はプロトタイプグラフの選択に単純な順位付けを用いて実験を行ったが、トレーニンググラフからプロトタイプグラフを選択する手法を工夫することで、より画像の特徴を明確に表現できるようになり、認識率が向上すると考えられる。



図 13 認識が提案手法で成功し、従来手法で失敗した画像

Fig. 13 Images that was recognized by proposed method, and was not recognized by conventional method

## 7. ま と め

本論文では、ベクトル空間へのグラフの埋め込みを用いたグラフ構造表現による、新しい一般物体認識手法を提案した。識別器としてマルチクラス SVM (RBF) を用いた場合、提案手法は従来手法の BoF に比べて、複数のデータセットで認識率が向上した。今後は、プロトタイプグラフの選択手法の改良や、グラフ編集距離の計算時間を更に短縮する予定である。また、この提案手法を、3 次元画像から作成された 3 次元グラフに拡張し、一般物体認識に応用することも検討していく予定である。

### 文 献

[1] K. Barnard and D. A. Forsyth, “Learning the Semantics of Words and Pictures,” IEEE International Conference on Computer Vision, Vol. 2, pp. 408–415, 2001.  
 [2] K. Barnard, P. Duygulu, N.d. Freitas, D. Forsyth, D. Blei

and M. Jordan, “Matching Words and Pictures,” Journal of Machine Learning Reserch, Vol. 3, pp. 1107–1135, 2003.  
 [3] P. Duygulu, K. Barnard, N.d. Freitas and D. Forsyth, “Visual categorization with bags of keypoints,” European Conference on Computer Vision, pp. IV:97–112, 2002.  
 [4] G. Csurka, C.R. Dance, L. Fan, J. Willamowski, C. Bray, “Object Recognition as Machine Translation: Learning aLexicons for a Fixed Image Vocabulary,” ECCV Workshop on Statistical Learning in Computer Vision, pp. 1–22, 2004.  
 [5] D. G. Low, “Distinctive image features from scale-invariant keypoints,” Journal of Computer Vision, Vol. 60, No. 2, pp. 91–110, 2004.  
 [6] J. Revaud, Y. Ariki and A. Baskurt, “Scale-Invariant Proximity Graph for Fast Probabilistic Object Recognition,” Conference on Image and Video Retrieval (CIVR), pp. 414–421, 2010.  
 [7] D. Conte, P. Foggia, C. Sansone and M. Vento, “Thirty years of graph matching in pattern recognition,” International Journal of Pattern Recognition and Artificial Intelligence, Vol. 8, No. 3, pp. 265–298, 2004.  
 [8] H. Bunke and G. Allerman, “Inexact graph matching for structural pattern recognition,” Pattern Recognition Letters, Vol. 1, pp. 245–253, 1983.  
 [9] A. Sanfeliu, and K. Fu, “A distance measure between attributed relational graphs for pattern recognition,” IEEE Transactions on Systems, Man, and Cybernetics, Vol. 13, No. 3, pp. 353–362, 1983.  
 [10] K. Riesen and H. Bunke, “Approximate graph edit distance computation by means of bipartite graph matching,” Image and Vision Computing, Vol. 27, Bo. 7, pp. 950–959, 2009.  
 [11] K. Riesen, M. Neuhaus and H. Bunke, “Graph Embedding in Vector Spaces by Means of Prototype Selection,” Graph-based representations in pattern recognition (GbrRPR), ed. F. Escolano et al, pp. 383–393, Springer-Verlag Berlin, Heidelberg, 2007.  
 [12] E. Valveny and M. Ferrer, “Application of Graph Embedding to solve Graph Matching Problems,” CIFED, pp. 13–18, 2008.  
 [13] M. Ferrer, E. Valveny, F. Serratos, K. Riesen and H. Bunke, “Generalized median graph computation by means of graph embedding in vector spaces,” Pattern Recognition, Vol. 43, No. 4, pp. 1642–1655, 2010.  
 [14] Caltech-101 Database, [http://www.vision.caltech.edu/Image\\_Datasets/Caltech101/](http://www.vision.caltech.edu/Image_Datasets/Caltech101/)  
 [15] M. Everingham, L. Gool, C. Williams, J. Winn, and A. Zisserman, “The PASCAL Visual Object Classes Challenge 2007 (VOC2007) Results,” <http://pascal.in.ecs.soton.ac.uk/challenges/VOC/voc2007/index.html>.