

一般物体認識における色情報の活用手法の検討

瓶子 義晴[†] 山中 高夫[‡]

[†] 上智大学大学院理工学専攻科 〒102-8554 東京都千代田区紀尾井町7番1号

E-mail: [†] y.heishi.88@gmail.com, [‡] takao-y@sophia.ac.jp

あらまし 効率的な画像検索を実現するために、画像中に映っている物体やシーン内容を自動的に認識する一般物体認識として、これまでに SIFT 特徴量と Bag of Features を利用した手法が提案されている。SIFT 特徴量は画像の輝度から計算される局所特徴量であり、物体の色情報は全く含まれていない。今までに、色情報を活用する手法として、Color SIFT と呼ばれる色情報を含んだ SIFT 特徴量の利用が提案されているが、他にも様々な形態が考えられる。本研究では、このような一般物体認識における色情報の活用手法について検討し、Color SIFT 特徴量を利用するよりも、カラーチャンネル毎に SIFT 特徴量から Bag of Features のベクトル量子化ヒストグラムを求め、その後に情報を統合する方が高い認識精度を得られることが分かった。

キーワード 一般物体認識、SIFT、Bag of Features、色情報、Color SIFT

1. はじめに

コンピュータやインターネットの普及により、画像のデジタル化が進み、大量の画像がデジタルデータとして保存されるようになった。このような大量の画像データを効率的に取り扱うためには、画像の内容を人手に頼ってラベル付けするだけでなく、自動的にその内容を把握する手法が求められる。このような背景から、画像に映っている物体の自動認識手法の研究が盛んに行われており、「一般物体認識」と呼ばれている[1]。

一般物体認識の研究は古くから行われているが、近年のコンピュータの発展と新しい手法の提案により、その認識精度は急速に向上した。物体を表現する特徴量として、part-based と呼ばれる表現方法が有効であることが分かり、特に Scale-Invariant Feature Transform (SIFT)[2] という物体の大きさや回転に影響されない特徴量がよく利用されている。さらに、SIFT 特徴量を計算した位置の情報は利用せず、画像中に似たような SIFT 特徴量がどのくらいあるかを表すベクトル量子化ヒストグラムによって画像を特徴づける Bag of Features (BoF) という手法が提案され[3]、それにより物体認識精度が大きく向上することが分かった[3-5]。

しかし、上記のような BoF による手法は画像の色情報を全く用いていない。画像データは、通常、人間の色覚特性に合わせた 3 チャンネルのカラーデータであり、色は物体を認識するために有益な情報を含んでいることが多い。そのため、一般物体認識の手法において色情報を活用する手段を導入することで識別率を向上することが期待される。これまでも様々な色情報を組み込んだ SIFT 特徴量が提案された (Color SIFT)。例えば、HSV-SIFT[6]、W-SIFT[7]、rgSIFT[8]、Transformed Color SIFT[8]などが例として挙げられる。これらは、画像中の

ある一点において、各カラーチャンネルで SIFT 特徴量を計算し、それを一つに連結した特徴量である。これらの特徴量を用いた一般物体認識は、輝度から計算される SIFT 特徴量を用いた一般物体認識よりも高い識別性能を示すことが報告されている[8]。

一般物体認識に色情報を活用する方法は、上記のように画像中の一点で色情報を組み込んだ SIFT 特徴量 (Color SIFT) を計算する手法だけでなく、他にも様々な方法が考えられる。本研究では、カラーチャンネル毎に SIFT 特徴量のベクトル量子化ヒストグラムを計算し、その後各チャンネルのヒストグラムを連結して活用する手法を提案し、その識別性能を評価した。一般物体認識のアルゴリズムにおいて、色情報の活用方法によって識別性能が変化するため、色情報の有効な活用方法を検討することが本研究の目的である。

2. 一般物体認識の従来手法

初めに、一般物体認識のアルゴリズムとして現在標準的な手法として用いられている Scale-Invariant Feature Transform (SIFT)特徴量と Bag of Features (BoF)について説明する。さらに、色情報を利用するために提案された Color SIFT 特徴量についても説明する。

2.1 Scale-Invariant Feature Transform (SIFT)

Scale-Invariant Feature Transform (SIFT)は 1999 年に Lowe が提案した局所特徴量[2]であり、物体のスケールや回転に対して不変の性質を持つ。これは、当初、特定の物体が画像中のどこに映っているかを認識する「特定物体認識」の手法として提案されたが、後述する BoF の手法と組み合わせることで一般物体認識に対して非常に有効な特徴量表現であることが分かっている[3-5]。

SIFT の計算プロセスは、「特徴点検出」と「特徴量記

述」の2段階に分けられる。「特徴点検出」は画像中の物体を記述するために最も良い特徴点の場所を複数検出し、さらにその特徴点を含む物体の表現にふさわしいスケール(物体の大きさに対応)を求める。「特徴量記述」では、検出された全ての特徴点に対して、主に勾配情報を利用した128次元のベクトルとして特徴量を表現する。このとき、各特徴点において主となる方向成分を検出してその軸方向を基準とするため、物体の回転に対して不変な特徴量を求めることができる。

近年の一般物体認識では、SIFT特徴量計算の2段階プロセスのうち、後者の「特徴量記述」部分だけを利用して「特徴点検出」はSIFTで提案されたものではなく独自の手法を利用することも多い[1]。本研究では、特徴点検出手法の優劣ではなく、色情報の活用手法が着眼点であるため、特徴点検出もLoweが提案した手法をそのまま利用した。

2.2 Bag of Features (BoF)

Bag of Features (BoF)は、統計的言語処理において利用されていたBag-of-Words Model [9]を画像に適用した手法であり、2004年の発表[3]以降、一般物体認識に対して有効であることが示された[3-5]。Bag of Wordsでは、語順を無視して文章を単語の集合と考え、単語の出現頻度で文章の特徴量を記述する手法である。これと同じように、BoFは、画像の特徴点の場所は無視して局所特徴量の集合と考え、その局所特徴量の出現頻度で画像の特徴量を記述する手法である。

BoFによる一般物体認識の流れを図1に示す。はじめに、多くの学習画像に対して、多くの特徴点に対する局所特徴量を計算する(局所特徴量検出)。局所特徴量の出現頻度を効率的に表現するためには、多くの特徴量を代表する特徴量ベクトルを用意して、その特徴量ベクトルに離散化する必要がある(ベクトル量子化)。そのため、局所特徴量検出で得られた大量の局所特徴量に対してクラスタリングを行い、各クラスを中心ベクトルを代表的な特徴量ベクトルとする(コードブック)。このコードブックに基づいて、各画像中に含まれる全ての特徴点の局所特徴量をベクトル量子化し、そのヒストグラムを構成する(ベクトル量子化ヒストグラム)。このベクトル量子化ヒストグラムが画像を特徴付ける特徴ベクトルとなる。実際に画像の識別を行うときには、その対象画像に対して局所特徴量を計算し、コードブックを利用してベクトル量子化ヒストグラムを求める。このベクトル量子化ヒストグラムに基づいて識別器が対象画像を物体カテゴリに識別する。

本研究では、局所特徴量としてSIFT特徴量を用いたBoFに対して、色情報の活用手法を検討した。

2.3 Color SIFT

Color SIFTとは、画像中の一点に対して、各カラーチ

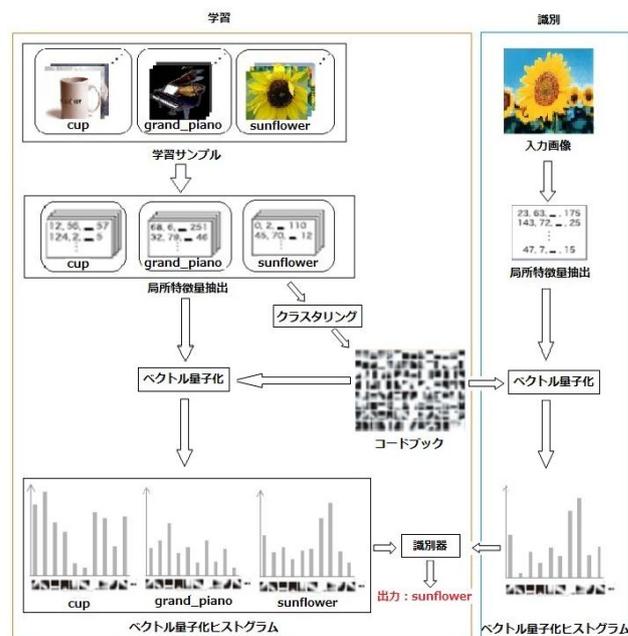


図1 Bag of Features による一般物体認識

ャネルでSIFT特徴量を計算し、それらを連結して一つの特徴量として表現するものである。通常の画像はRGBの3チャンネルでデータが保存されているが、Color SIFTでは、そのRGBデータに対して適切な色変換を行ってからSIFT特徴量を計算する。その色変換として、様々な手法を用いることが提案されている。前述で例に挙げたHSV-SIFT[6]、W-SIFT[7]、rgSIFT[8]、Transformed Color SIFT[8]は、それぞれ、HSV色空間、補色の色空間において輝度を正規化した空間、RGB色空間において輝度を正規化した空間、RGB色空間において各カラーチャンネルの分布を正規化した空間においてSIFT特徴量を計算したものである。これらのColor SIFTの特徴とそれらを利用した一般物体認識の識別性能が文献[8]にまとめられており、その文献で調査したデータセットに対してはColor SIFTが通常のSIFTより高い識別性能を示すことが報告されている。

本研究では実装を容易にするため、色変換を行わずRGB色空間の各カラーチャンネル(R, G, B)においてSIFT特徴量を計算し、それを連結して一つのColor SIFT特徴量ベクトルとして表すRGB-SIFTを用いた。

3. 一般物体認識に対する色情報の活用方法

本研究では、Color SIFTのうち、最も単純なRGB-SIFTによるBoFを利用した手法(RGB-SIFT-BoF)と、カラーチャンネル毎にSIFT特徴量からBoFのベクトル量子化ヒストグラムを計算した後に情報を統合する手法(RGB-Channel-BoF)の比較を行った。ここでは、それらの手法について具体的に説明する。

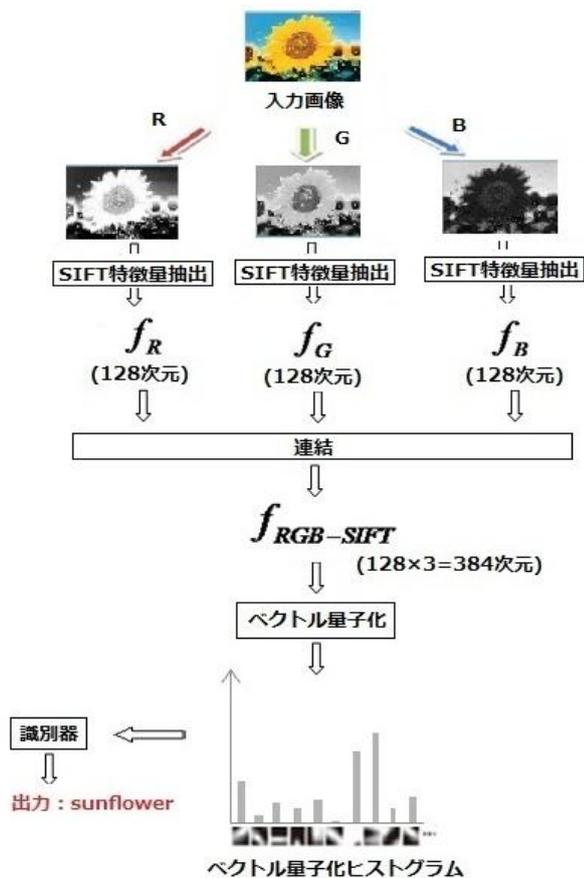


図2 Color SIFT に基づいた Bag of Features を利用する手法 (RGB-SIFT-BoF)

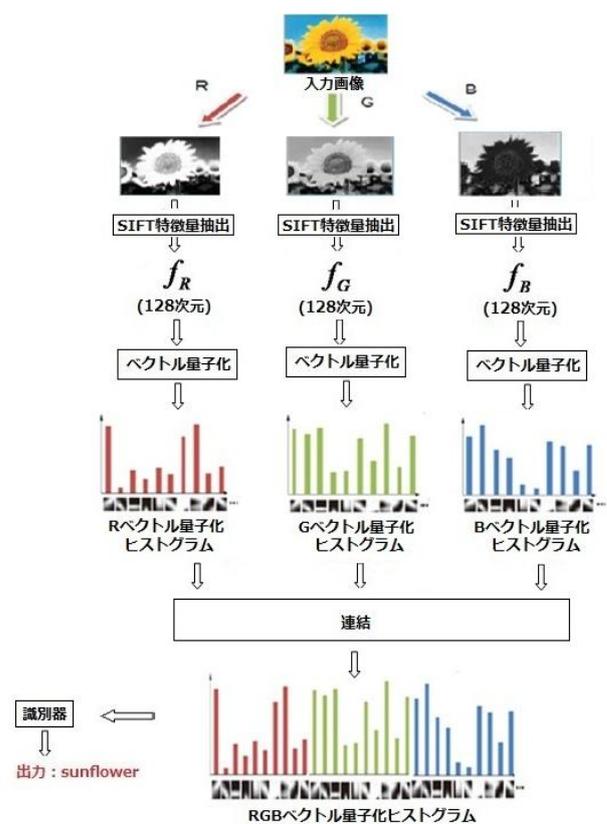


図3 カラーチャンネル毎の Bag of Features を利用する手法 (RGB-Channel-BoF)

3.1 Color SIFT による BoF を利用する方法 (RGB-SIFT-BoF)

RGB-SIFT による BoF を利用する手法の処理の流れを図2に示す。本論文では、この手法を RGB-SIFT-BoF と呼ぶ。与えられる画像データは RGB の3チャンネルの値を持ったカラー画像とする。

初めに、RGB-SIFT 特徴量を計算するための特徴点を求める必要がある。与えられたカラー画像から輝度値を計算し、その輝度画像に通常の SIFT 特徴量の特徴点検出手法を適用して特徴点を求める。

次に、求めた各特徴点に対して、RGB-SIFT の特徴量記述を行う。画像は、図2に示すように、3つのカラーチャンネルに分けられ、それぞれのカラーチャンネルで SIFT 特徴量を計算する。各カラーチャンネルから計算される SIFT 特徴量を次式のベクトルで表す。

$$f_R = [f_{R1}, \dots, f_{R128}]^T$$

$$f_G = [f_{G1}, \dots, f_{G128}]^T$$

$$f_B = [f_{B1}, \dots, f_{B128}]^T$$

ただし、 f_R 、 f_G 、 f_B はそれぞれ R、G、B チャンネルの SIFT 特

徴量ベクトルとする。RGB-SIFT の特徴量ベクトルを構成するため、次式のようにこれらの3つの特徴量ベクトルを連結する。

$$f_{RGB-SIFT} = [f_{R1}, \dots, f_{R128}, f_{G1}, \dots, f_{G128}, f_{B1}, \dots, f_{B128}]^T$$

したがって、SIFT 特徴量は通常 128 次元のベクトルとして求められるが、RGB-SIFT ではその3倍の384次元のベクトルとなる。

与えられた画像の全特徴点に対して RGB-SIFT 特徴量を求め、ベクトル量子化によりベクトル量子化ヒストグラムを計算する。これが画像中の物体を表す特徴量ベクトルとなる。最終的な特徴量ベクトルの次元は、ベクトル量子化ヒストグラムの次元となるので、コードブックのサイズと等しくなる。

3.2 カラーチャンネル毎の BoF を利用する方法 (RGB-Channel-BoF)

色情報を一般物体認識に活用する方法は、Color SIFT を利用する手法だけではなく、他にも様々な手法が考えられる。本研究では、この節で説明するように、カラーチャンネル毎に BoF のベクトル量子化ヒストグラムを計算し、その後で情報を統合する手法を提案する。

表1 データセットにおける画像カテゴリと画像枚数

データセットA ランダムカテゴリ		データセットB カラーカテゴリ		データセットC 白黒カテゴリ	
カテゴリ	枚数	カテゴリ	枚数	カテゴリ	枚数
Accordion	55	Sunflower	85	Accordion	55
Anchor	42	Strawberry	35	Car_side	123
Ant	42	Stop_sign	64	Dalmatian	67
Barrel	47	Lobster	41	Grand_piano	99
Bass	54	Rooster	49	Headphone	42
Beaver	46	Garfield	34	Panda	38
Bonsai	128	Dolphin	65	Schooner	63
Brain	98	Dollar_bill	52	Snoopy	35
Brontosaurus	43	Flamingo	67	Soccer_ball	64
Buddha	85	Pizza	53	Yin_yang	60

本手法の処理の流れを図3に示す。本論文では、この手法を RGB-Channel-BoF と呼ぶ。この手法では、特徴点を検出する前に、与えられたカラー画像を3つのカラーチャンネルに分ける。その各カラーチャンネルの画像に対して、SIFTの特徴点検出及び特徴量記述を行う。各カラーチャンネルで特徴点検出を行うので、3つのカラーチャンネルで異なる特徴点を利用される。各カラーチャンネルで求めた SIFT 特徴量に対して、ベクトル量子化ヒストグラムを求める。最後に、各カラーチャンネルで求めたベクトル量子化ヒストグラム連結して、画像中の物体を表す特徴量ベクトルとする。最終的な特徴量ベクトルの次元は、各カラーチャンネルのベクトル量子化ヒストグラムの次元の3倍になる。つまり、コードブックサイズの3倍に等しくなる。

4. 評価実験

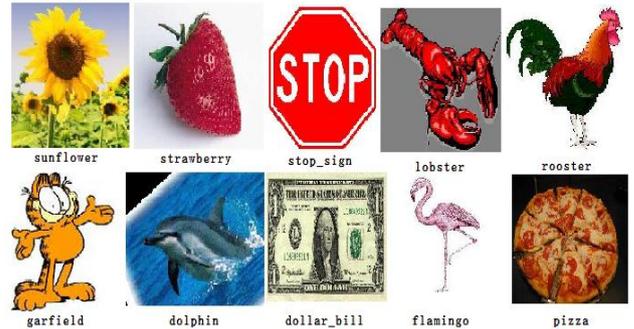
4.1 実装

本研究では、標準的な SIFT 特徴量と BoF を利用した一般物体認識アルゴリズム [4] と、3.1 節で説明した RGB-SIFT-BoF、3.2 節で説明した RGB-Channel-BoF を実装し、それらの性能を評価した。開発環境として Microsoft Visual Studio 2008 を利用し、C++ で実装した。画像処理ライブラリとして、OpenCV2.1 を利用した。また、SIFT 特徴量の計算には、Robin Hess のコードを利用した [10]。識別器として、誤り訂正出力符号方式の Multiclass Support Vector Machines (SVM) [11] を利用した。

BoF の計算において、コードブック構築のためのクラスタリング手法には k-Means Clustering を利用し、各データセットの全画像から求めた。後述の実験1では、コードブックのサイズを、標準的な SIFT + BoF と RGB-SIFT-BoF で 300 に設定し、RGB-Channel-BoF では各カラーチャンネルで 100 ずつとし、連結後のベクトル量子化ヒストグラムの次元が 300 となるように設定した。実験2では、コードブックのサイズを変化させてその影響を検討した。



(a) データセット A (ランダムカテゴリ) の画像例



(b) データセット B (カラーカテゴリ) の画像例



(c) データセット C (白黒カテゴリ) の画像例

図4 各データセットにおける画像例

4.2 データセット

実験に用いる画像データとして、一般物体認識のアルゴリズム評価によく用いられる Caltech101 [12] を利用した。これはカリフォルニア工科大学が 2005 年に発表した 101 種類 (9144 枚) の画像からなるデータセットである。そのまま利用すると 101 カテゴリという非常に困難な問題設定となるので、本研究ではこの 101 カテゴリから 10 カテゴリを選択して実験用のデータセットを作成し、10 カテゴリの識別問題として利用した。10 カテゴリの選択方法として、(1) ランダム (データセット A: ランダムカテゴリ)、(2) 色情報が物体認識に対して重要な要素となっていないようなカテゴリ (データセット B: カラーカテゴリ)、(3) 色情報があまり含まれていないカテゴリ (データセット C: 白黒カテゴリ) の 3 種類を用いた。各デー

タセットのカテゴリ名とそのカテゴリに含まれている画像の枚数を表 1 に示す。また、各データセットにおける各カテゴリの代表画像を図 4(a)~(c)に示す。この 3 種類のデータセットに対して、それぞれ評価実験を行った。

データは学習データとテストデータに分け、学習データで識別器を構成し、テストデータ中の各画像に対してカテゴリを推定した。その推定結果が実際のカテゴリと一致する場合を正解とし、その正解率を識別率として計算した。学習データとして、各カテゴリからランダムに 30 画像を選択し、残りの画像をテストデータとした。この識別を 5 回繰り返す、その識別率の平均値を最終的な識別率とした。

4.3 実験 1：各手法に対する識別性能の比較

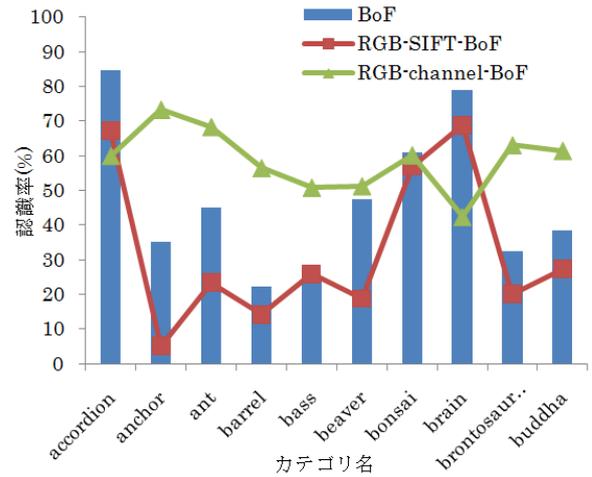
標準的な SIFT+BoF 手法と RGB-SIFT-BoF、RGB-Channel-BoF の 3 種類の手法に対する一般物体認識の認識精度を比較するため、前述のデータセット A (ランダムカテゴリ)、データセット B (カラーカテゴリ)、データセット C (白黒カテゴリ) に対して識別率を調べた。各手法における最終的な特徴ベクトルの次元は 300 となるように設定した。

識別率をまとめた結果を表 2 に示し、それぞれのデータセットに対する各カテゴリの識別率を図 5 (a)~(c)に示す。表 2 から分かるように、データセット A (ランダムカテゴリ) とデータセット C (白黒カテゴリ) に対しては、RGB-Channel-BoF が最も識別率が高くなり、データセット B (カラーカテゴリ) に対しては標準的な BoF の識別率が最も高くなった。また、Color SIFT を利用した手法である RGB-SIFT-BoF の識別率は全般的に低くなることが分かった。

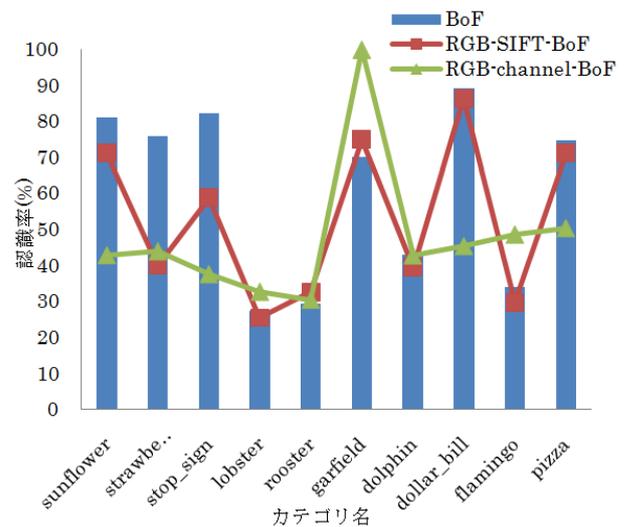
図 5 におけるカテゴリ毎の識別率を見ると、データセット A (ランダムカテゴリ) やデータセット C (白黒カテゴリ) において、BoF や RGB-SIFT-BoF ではカテゴリによって識別率の高いものと低いものがあるが、RGB-Channel-BoF では比較的どのカテゴリでも高い識別率となることが分かった。また、データセット B (カラーカテゴリ) において、BoF で識別率の高いカテゴリに

表 2 各データセットに対する平均識別率

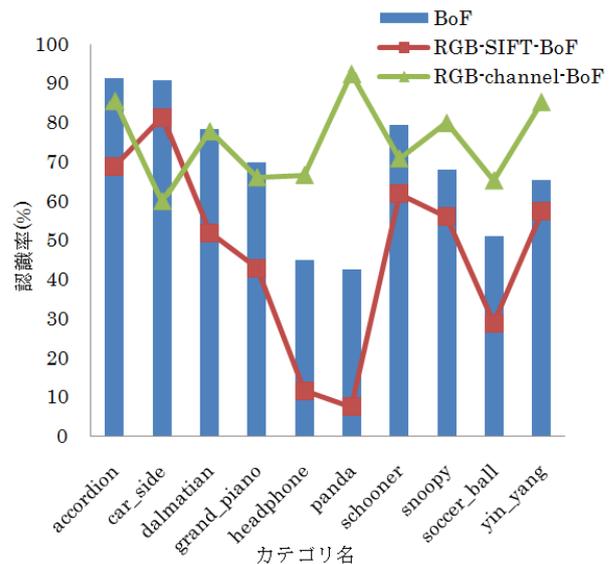
	データセット A ランダムカテゴリ	データセット B カラーカテゴリ	データセット C 白黒カテゴリ
BoF (従来手法)	46.95%	60.70%	68.16%
RGB-SIFT-BoF (Color SIFT による BoF)	32.73%	53.00%	46.80%
RGB-Channel-BoF (カラーチャンネル毎の BoF を連結)	55.83%	47.52%	75.02%



(a) データセット A (ランダムカテゴリ) に対する識別率



(b) データセット B (カラーカテゴリ) に対する識別率



(c) データセット C (白黒カテゴリ) に対する識別率

図 5 各データセットに対する識別率

対して、RGB-SIFT-BoF や RGB-Channel-BoF の識別率が低くなる場合のあることが分かった。

RGB-SIFT-BoF と RGB-Channel-BoF は色情報を活用した認識手法であるため、当初データセット B (カラーカテゴリ) においてその効果を見ることができないのではないかと予想していたが、それとは反対の結果が得られた。詳細な考察は後述するが、特徴的な色を持つ画像は撮影条件や撮影機材の変化により同じ物体の色でも異なる色として画像に表れることも多く、そのカラーバランスの変化が識別に大きく影響を与えていると思われる。

4.4 実験 2 : コードブックサイズの影響

次に BoF のコードブックサイズが色情報を用いた一般物体認識に与える影響を検討するため、コードブックサイズを変化させて識別率を調べた。標準的な SIFT+BoF と RGB-SIFT-BoF では、コードブックサイズがベクトル量子化ヒストグラムの次元に相当するので、最終的な画像中の物体を表す特徴量ベクトル (識別器の入力ベクトル) の次元と一致する。それに対して、RGB-Channel-BoF では、コードブックサイズの 3 倍が最終的な特徴量ベクトルの次元となる。

最終的な特徴量ベクトルの次元を、100 次元 (RGB-Channel-BoF は 99 次元)、300 次元、900 次元に設定した時のデータセット A (ランダムカテゴリ) に対する全カテゴリ平均識別率を表 3 に示す。特徴量ベクトルの次元が変化しても全般的な傾向は変わらず、このデータセットに対しては RGB-Channel-BoF が最も識別率が高く、Color SIFT を用いた RGB-SIFT-BoF が最も低くなった。この結果から、調べた範囲では、特徴量ベクトルの次元は識別率に大きな影響を与えないことが分かった。

5. 考察

実験結果から、データセット A (ランダムカテゴリ) やデータセット C (白黒カテゴリ) に対しては、RGB-Channel-BoF の識別率が、標準の SIFT+BoF や RGB-SIFT-BoF の識別率と比べて高い傾向にあった。この理由は以下のように推測される。RGB-SIFT-BoF は画像中の一点における RGB 値から計算した SIFT 特徴量を連結しているので、カラーバランスが少しでもずれると大きく影響を受ける。それに対して、RGB-Channel-BoF はカラーチャンネル毎にベクトル量子化ヒストグラムを構成するので、ある一点におけるカラーバランスの変化はそれほど影響しない。RGB-Channel-BoF は、RGB-SIFT に比べて比較的ゆるやかに色情報を活用していると考えられる。そのため、ある程度のカラーバランスの変化に対してロバスト性があり、識別率が高くなったと考えられる。標準的な SIFT+BoF に対しては、色情報を活用することで情報量が増加し、識別率が増加したと推測される。

表 3 コードブックサイズに対する平均識別率
(データセット A : ランダムカテゴリ)

最終的な画像中の物体を表す特徴量ベクトルの次元	100 次元 ^(注1)	300 次元	900 次元
BoF (従来手法)	39.96%	46.95%	41.64%
RGB-SIFT-BoF (Color SIFT による BoF)	32.04%	32.73%	37.17%
RGB-Channel-BoF (カラーチャンネル毎の BoF を連結)	52.20%	55.83%	58.73%

(注 1) RGB-Channel-BoF に対しては 99 次元 (3×33 次元)

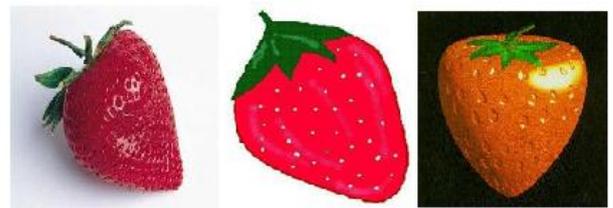


図 6 Strawberry カテゴリの画像例

一方、データセット B (カラーカテゴリ) では、RGB-SIFT-BoF と RGB-Channel-BoF は、標準の SIFT+BoF に比べて識別率が低い結果となった。このデータセットの画像を詳しく見てみると、同じカテゴリでもカラーバランスが異なっている例が多く見られた。例えば、Strawberry のカテゴリでは、図 6 に示すように、実際のイチゴやイラストのイチゴなどが含まれており、3 つの画像で異なるカラーバランスになっていることが分かる。このようなカテゴリに対して色情報を用いると、異なる色と判断してしまうので、識別率が低下したと考えられる。対策としては、人の色彩知覚を考慮した $L^*a^*b^*$ 表色系などを利用して近い色と遠い色を判断し、その情報を利用することが考えられるので、今後検討したい。

データセット C (白黒カテゴリ) に対して、RGB-Channel-BoF が標準の SIFT+BoF や RGB-SIFT-BoF に比べて識別率が高くなった。図 4 (c) から分かるように、図中に映っている物体自体は白黒の物体が多いので、色情報はそれほど重要ではないように感じる。しかし、図 5 (c) のカテゴリ毎の識別率を見ると、Headphone や Panda のカテゴリで RGB-Channel-BoF の識別率が高くなっているので、背景の色情報を活用して識別している可能性も考えられる。今後、さらに様々な白黒物体のカテゴリに対する識別性能を調べることで原因を明確にしたい。

rgSIFT や W-SIFT、Transformed Color SIFT など、本研究で用いた RGB-SIFT 以外にも様々な Color SIFT が提案されており、文献[8]ではこれらの Color SIFT は標準の SIFT よりも識別性能が高くなることが報告されている。

本研究では、RGB-SIFT は標準の SIFT よりも低い識別性能を示したが、この原因が Color SIFT の種類によるものなのか、それともデータセットに依存するものなのかはまだ不明であり、今後検討が必要である。rgSIFT や W-SIFT は輝度変化に対して不変である特性を持ち、Transformed Color SIFT は照明変化に対して不変である特性を持つので、輝度変化にも照明変化にも不変性を持たない RGB-SIFT よりも高い識別性能を示す可能性もある。一方、輝度変化に対して不変性を持つ Color SIFT は、識別に輝度情報を用いることができないので、識別性能が低くなる可能性もある。このような点について、さらに様々なデータセットに対して様々な Color SIFT の識別性能を調べ、本論文で提案する RGB-Channel-BoF と比較する必要があるだろう。また、RGB-Channel-BoF も RGB のカラーチャネルを用いるのではなく、Color SIFT と同じように色変換をした後のカラーチャネルを利用する方法も考えられる。この点についても今後検討したい。

RGB-Channel-BoF が RGB-SIFT-BoF と比べて識別率が高い傾向はデータセットに依存する可能性もあるので、さらに様々なデータセットに対して検討する必要がある。ただし、図5を見ると、少なくともほとんどのカテゴリで RGB-Channel-BoF は RGB-SIFT-BoF と同等以上の識別率を示しているため、他のデータセットでも RGB-SIFT-BoF を用いるよりも RGB-Channel-BoF を用いた方が、識別率が上がる傾向になると考えている。

最後に、実験2で調べた特徴量ベクトルの次元数の影響について考察する。BoF のベクトル量子化において、コードブックサイズが小さいと、SIFT 特徴量ベクトルの多少の変化に対するロバスト性は高くなるが、量子化誤差が大きくなる。逆に、コードブックサイズが大きいと量子化誤差が小さくなるが、変化に対するロバスト性は低くなる。また、識別器にとっても入力の特徴量ベクトルの次元数が大きくなると識別が難しくなるという一般的な問題 (Curse of Dimensionality) がある。これらの理由のため、特徴量ベクトルの次元数が識別に影響する可能性があるが、調べた範囲の次元ではそれほど影響のないことが分かった。この傾向は通常の SIFT+BoF に対して次元数の影響を調べた論文の結果とも一致している [4]。

6. まとめ

本研究では、一般物体認識において、色情報を活用する手法の検討を行い、Color SIFT を用いた認識手法よりも、カラーチャネル毎に Bag of Features のベクトル量子化ヒストグラムを作成し、その後に情報を統合する認識手法の方が識別性能が高くなる可能性を示した。今後さらに様々なデータセットに対する調査が必要であるが、調べた範囲のデータセットに対しては上記のことを確認

することができた。また、Bag of Features のコードブックサイズのパラメータは認識結果の傾向に大きな影響を与えないことも確認できた。

本論文では、実装を簡単にするため、比較対象として RGB-SIFT を利用したが、今後、輝度に対して不変である rgSIFT や W-SIFT、照明変化に対して不変である Transformed Color SIFT などと比較する必要がある。また、L*a*b*表色系など人の色彩認知をよく表現している色空間を利用して色情報を活用する手法が有望であると考えている。

本論文の実験において、色情報を SIFT にどのように組み込むかという取り組みだけではなく、Bag of Features の枠組みを含めて色情報をどのように活用していくかということが重要な検討課題になることが分かった。今後さらに研究を進展させて、一般物体認識に色情報を活用する枠組みを確立したい。

文 献

- [1] 柳井啓司, “一般物体認識の現状と今後,” 情報処理学会論文誌 コンピュータビジョンとイメージメディア, vol.48, no.16, pp. 1-24, 2007.
- [2] D. G. Lowe, “Object recognition from local scale-invariant features,” Proc. IEEE International Conference on Computer Vision, pp.1150-1157, 1999.
- [3] G. Csurka, C. R. Dance, L. Fan, and C. Bray, “Visual categorization with bags of keypoints,” ECCV International Workshop on Statistical Learning in Computer Vision, pp. 1-22, 2004.
- [4] E. Nowak, F. Jurie, and B. Triggs, “Sampling strategies for bag-of-features image classification,” European Conference on Computer Vision, pp. 490-503, 2006.
- [5] H. Zhang, A. C. Berg, M. Maire, and J. Malik, “SVM-KNN: Discriminative nearest neighbor classification for visual category recognition,” IEEE Conference on Computer Vision and Pattern Recognition, pp. 2126-2136, 2006.
- [6] A. Bosch, A. Zisserman, and X. Munoz, “Scene classification using hybrid generative/discriminative approach,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 30, no. 4, pp. 712-727, 2008.
- [7] J. M. Geusebroek, R. van den Boomgaard, A. W. M. Smeulders, and H. Geerts, “Color invariance,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 12, pp. 1338-1350, 2001.
- [8] E. A. Koen, van de Sande, T. Gevers, and C. G. M. Snoek, “Evaluating color descriptors for object and scene recognition,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 9, pp. 1582-1596, 2010.
- [9] C. D. Manning and H. Schtze, Foundations of Statistical Natural

Language Processing, MIT Press, 1999.

- [10] R. Hess, "An Open Source SIFT Library," Proc. ACM Multimedia, 2010.
- [11] C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
- [12] L. Fei-Fei, R. Fergus, and P. Perona. "Learning generative visual models from few training examples: an incremental Bayesian approach tested on 101 object categories," IEEE CVPR Workshop on Generative-Model Based Vision, 2004.