

# 色情報を用いた SIFT による類似画像検索精度の向上

## Improving Precision of Similar Image Retrieval Using SIFT with Color Features

平山 力地<sup>\*1</sup>  
Hirayama Rikichi

新美礼彦<sup>\*2</sup>  
Niimi Ayahiko

<sup>\*1</sup> 公立はこだて未来大学  
Future University Hakodate #1

<sup>\*2</sup> 公立はこだて未来大学  
Future University Hakodate #2

SIFT is a typical technology for extracting features; it can search image data even if the image data has changed in terms of size and angle. SIFT only uses the luminance of an image to extract its features, which suggests that SIFT eliminates almost all color features from consideration. While gray-scale image data is depicted only in terms of luminance, there are cases where we can discriminate color image data, but not gray-scale image data. If we use SIFT for Similar-Image Retrieval on a computer, the same effect is observed. Therefore, this research aims to improve Similar-Image Retrieval by using SIFT with color features.

### 1. はじめに

近年、画像データの扱いが一般的に浸透していくのに伴い、大量の画像データを効率的に扱う方法が求められている。類似画像検索はその方法の一つとして注目されており、インターネット上のサービスも現れ始めている。類似画像検索ではコンピュータに画像を分析させることにより画像を分類したり、自分の好きな画像と似たような画像を探することができる[大坂 2005]。

また、類似画像検索は画像認識にも利用されている。画像認識とは、例えば写真画像が与えられた時、そこに写っている物を答えるような技術のことである。画像認識には大きく分けて二種類、一般画像認識と特定画像認識がある。例を挙げると、車の写真画像が与えられた時、一般画像認識は「車が写っている」と答え、特定画像認識は「A 社の B という車が写っている」と答える。技術的な難易度は一般画像認識の方が高いとされ、特定画像認識を利用して一般画像認識に近づけようというのが最近のアプローチである[黄瀬 2010]。

類似画像検索を行うためには、画像から特徴を抽出するための方法が必要であり、近年盛んに研究されている[D. G. Lowe 1999] [阿倍 2008]。SIFT は類似画像検索に用いる代表的な特徴量の抽出方法であり、画像の局所的な特徴に着目することにより拡大縮小や回転変化に頑強という性質を持っている[藤吉 2007]。SIFT は特徴量を抽出する際、グレースケール画像を使用する。グレースケール画像を使用するという事は色情報がほとんど切り捨てられているということである。そこで本研究では、SIFT に色情報を組み込み検索精度、特に特定画像認識の精度の向上を図った。

### 2. SIFT の問題点

SIFT は特徴量を抽出する際に、画像の輝度情報のみを利用している。これは、元の画像がカラー画像であってもグレースケール画像であっても、得られる特徴量は変わらないということである(図1参照)。図1では、入力にカラー画像を使用した場合と、グレースケール画像を使用した場合の SIFT 特徴量の位置を示している。グレースケール画像の方は、あらかじめ外部でグレースケール画像へ変換処理を行っているため、若干結果が異なっ

ているが、ほぼ同一の特徴量が得られていることがわかる。

カラー画像の特徴量

グレースケール画像の特徴量

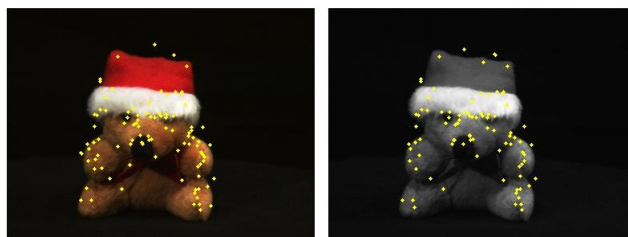


図1 特徴点の記述例

だが、グレースケール画像を利用しているということは、カラー画像に比べると色情報が少なくなるということであり、そのため SIFT 特徴量は単純な色の変化に弱いという欠点がある。色が全く違っていても特徴点周辺が似ていれば、類似していると判断する事があるということである。具体的には、人の目でも似ていると判断してしまう二つのグレースケール画像は、それぞれの画像の持つ SIFT 特徴量も似通うことが多くなるということである(図2参照)。

グレースケール化

グレースケール化

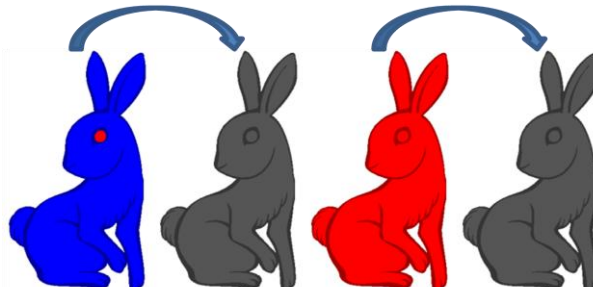


図2 グレースケール画像

類似画像検索において検索に使用する画像に写っている物と、色が全く違う物が検索結果として出力されることを期待するケースは少ないと考えられるので、このようなケースを防止するために色情報は必要であると思われる。

### 3. 関連研究

色情報を SIFT に組み込む手法の報告は過去にもなされている。Alaa E らによる研究では色情報を利用することにより既存の SIFT と比べ、同一の画像から多くの特徴量ベクトルを得ることに成功している[Alaa E 2006]。また 1 枚の画像から得られる特徴量ベクトルが増えたことにより、同じオブジェクトが写っている画像間で対応する特徴量も増え、検索精度が上がったとの報告がある。しかし特徴量ベクトルの数が増えるということは、検索時の計算量も増えるということである。そこで本研究では特徴量ベクトルの入力データ数は増やさず、特徴量ベクトルの次元数を増やし、計算量の増加を抑えつつ検索精度を向上させたい。また先行研究では実際の類似画像検索においてどのような影響が出たか言及していなかったため、実験を行い SIFT に色情報を組み込むことの有用性について示す。

### 4. 提案手法

本研究では、形は似ているが色の違う物、を区別するために、SIFT の特徴量ベクトルに色情報を新たに追加する。具体的には、色情報無しでは似ていると判断してしまう画像は、特徴点の記述される位置も非常に似通うことを利用する。このように、似たような位置に記述された特徴点を見分けるために、特徴点の記述された位置の周辺の色情報を SIFT の特徴量ベクトルに追加する(図3参照)。

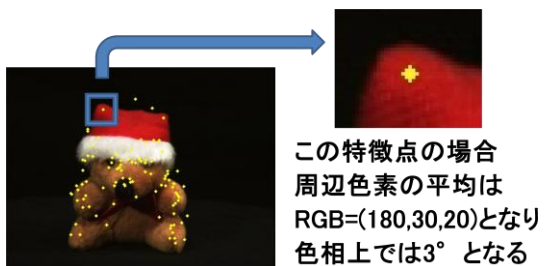


図3 色情報

ただし、RGB の値をそのまま利用すると直感的に似た色がわかりにくい上、色以外に明暗によって大きく数値が変わってしまうので、ここでは色相を利用する(図4参照)。



図4 色相

色相は  $0^\circ \sim 360^\circ$  の角度で定義され、似た色は近い角度になる性質がある。色相は RGB の値を利用して、次の式で求めることができる。なお RGB とは、それぞれ赤・緑・青の色を表しており、各 8bit で表されているとする。

$$\tan \text{Hue} = \frac{\sqrt{3}(G - B)}{2R - G - B}$$

この式を変形して

$$\text{Hue} = \tan^{-1} \frac{\sqrt{3}(G - B)}{2R - G - B}$$

また、以上により求められた色相の値は角度であるため、そのままでは特徴量ベクトルとして利用することが出来ない。2 つの  $n$  次元ベクトル、 $x(x_1, x_2, \dots, x_n), y(y_1, y_2, \dots, y_n)$  の距離  $d(x, y)$  は

$$d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2}$$

で表されるため、仮に角度をそのままの数値で 3 つの 1 次元ベクトル、 $x(0), y(359), z(180)$  として利用した場合、角度ではベクトル  $x$  とベクトル  $y$  が近くなるはずだが、ユークリッド距離では

$$d(x, y) = \sqrt{(0 - 359)^2} = 359$$

$$d(x, z) = \sqrt{(0 - 180)^2} = 180$$

となり、実際とは異なった結果が出てしまう。そこで角度を  $\sin$  と  $\cos$  で表す。先ほどの 1 次元ベクトルを  $\sin$  と  $\cos$  で表した場合、 $x(0), y(-0.01, 0.99), z(0, -1)$  となり

$$d(x, y) = \sqrt{(0 - (-0.01))^2 + (1 - 0.99)^2} = 0.01$$

$$d(x, z) = \sqrt{(0 - 0)^2 + (1 - (-1))^2} = 2$$

となり、正しい結果が得られる。よって本研究では既存の SIFT 特徴量である 128 次元ベクトルに、 $\sin$  と  $\cos$  の値を加えた 130 次元ベクトルを提案する。また、 $\sin$  と  $\cos$  の値をそのまま利用すると、既存の SIFT 特徴量ベクトルの値に比べ非常に小さくなり検索結果に影響が出ないので、この値を定数倍することにより重み付けを行っている。

### 5. 実験

提案手法の有効性を評価するため、画像データを利用した類似画像検索の実験を行う。実験は、実験用の画像データセットを対象に一回、実データを対象に一回、計二回行った。前者はインターネット上で公開されているデータセットなので、類似研究と性能比較をする際の指標となる。また後者は実データなので、実際に提案手法を用いてシステムを構築し運用した際の性能指標になるとと思われる。

#### 5.1 実験 1

実験には isis が公開している画像データセットである、Amsterdam Library of Object Images を用いる[isis 2010]。このデータセットは、照明の当たり具合などが違うオブジェクト各 24 枚・1000 種類、合計 24000 枚の画像からなるデータセットである。全て手作業による撮影なので若干の誤差はあるが、比較的似たような条件で撮られている画像データセットとなる。本研究の提案手法は類似画像検索、特に特定画像認識の精度を上げることを目的としているので、同一のオブジェクトが写っている画像が多数含まれているデータセットである Amsterdam Library of Object Images を用いた(図5参照)。また背景が全て黒で統一されているので、背景による検索結果への影響は少なく抑えられていると考えられる[柄原 2010]。

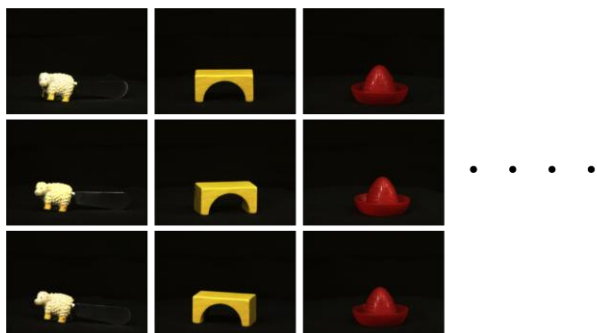


図5 データ例

今回はこの画像データセットから、同条件で撮影された 1000 種類のオブジェクトの画像各 1 枚、合計 1000 枚をデータベースに登録した。また、違う条件で撮影された 1000 枚の画像を検索に使用する画像として用いた。これらのデータセットを利用して、既存の SIFT と、色情報を組み込んだ SIFT で、どのような結果の違いが出るか実験を行う。検索画像 1 枚につき、1 枚の検索結果が得られるとし、検索に使用した画像と同じオブジェクトが写っている画像を出力できた場合、正解とする。この条件では、検索結果が正解となるのは同一オブジェクトを検出した時だけとなるので、正解率が上がれば特定画像認識の精度が向上したと考えられる。

## 5.2 実験 1 結果

5 章で述べた条件で実験を行ったところ、次の結果が出た。表の正解・不正解の項目は、横軸が既存の SIFT での結果、縦軸が色情報を加えた SIFT での結果となっている。この表では同一の実験データの条件を変えることによって、結果がどのように変化したか見ることが出来る。今回の実験では両条件で正解だった実験データ、不正解だった実験データではなく、条件を変えることにより結果が変わった実験データに着目する。よって、既存の SIFT では正解だったが色情報を加えたことにより不正解となった 5 件、また既存の SIFT では不正解だったが色情報を加えたことにより正解となった 44 件が重要となる。

表1 実験結果

		SIFT+色情報		合計
		正解	不正解	
SIFT	正解	759	5	764
	不正解	44	192	236
合計		803	197	1000

既存の SIFT では正解が 764 件だったのに対し、色情報を組み込んだ SIFT では 803 件となっている。また SIFT では正解だったケースが、色情報を追加したことにより不正解となったケースは 5 件だけとなっている。この結果をマクネマー検定にかけると有意であると出たので、類似画像検索結果に色情報の影響があったと言える。

表と検定結果により、色情報の追加により検索精度が向上したと言える。色情報を追加したことにより不正解になった 5 件については、いずれも元の画像から得られる特徴量が極端に少ないという共通点が見受けられた。また、似た色のオブジェクトが写っている画像を類似画像として出力しているケースが多いことから、元の特徴量が少ないために追加した色情報の影響を大き

く受け、構造ではなく色が似ている画像を類似していると判断し出力してしまったのだと考えられる(図6参照)。

このように今回の実験結果では、得られる特徴量が少ない画像ほど、検索結果に影響が出る傾向が見られた。このことから、既存の 128 次元ベクトルに 2 次元の色情報ベクトルを単純に追加するだけでは、検索結果全体、特に得られる特徴量の多い画像の検索結果に与える影響は少ないのだと思われる。だが、実験結果が良くなっていることから、今回の提案手法は既存の SIFT に比べ性能が良くなったと言える。

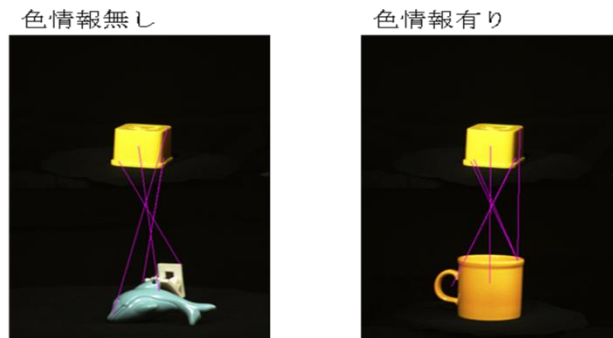


図6 検索結果例

## 5.3 実験 2

次に実データでの実験を行った。データセットと異なり撮影条件が統一されていないので、実際にシステムを構築して運用した場合でも提案手法が有効に働か実験することができる。実験には厚生労働省の公開している食中毒事件が多いとされるキノコの画像を利用した(図7参照)。毒キノコ 3 種類と、それと間違えやすい食用キノコ 10 種類、計 13 種類のキノコの画像を 2 枚ずつ、合計 26 枚を Google イメージから探し、画像の真偽を確かめた上で、用意しデータベースとして利用した。検索画像にはデータベースに登録している画像とは違う、毒キノコ 3 種類の画像を 1 枚ずつ用意した。この条件で類似画像検索を行い、検索画像と同じ種類の画像を出力した場合は正解、似ているとされる種類の画像を出力した場合は類似種が見つかったとし、それ以外の種類の画像を出力した場合は不正解とする。実データでの実験では実験用データセットを利用した時ほど正確に正解を検出することが出来ず、またデータ数が少なくなることで検索回数も少なくなるので、結果に差を出すために正解だけをカウントせず類似種もカウントした。なお検索結果は、検索画像 1 枚につき 3 枚出力する。



図7 食中毒事件の多いキノコの例

## 5.4 実験 2 結果

7章で述べた条件で実験を行ったところ、次の結果が出た。表中の数字は各検索で検出することの出来た件数である。

表2 実データでの実験結果

	SIFT			SIFT+色情報		
	正解	類似種	不正解	正解	類似種	不正解
クサウラベニタケ	0	1	2	1	1	1
ツキヨタケ	0	1	2	1	1	1
ニガクリタケ	0	1	2	0	0	3
合計	0	3	6	2	2	5

正解数が 0 件から 2 件に増えたことから、既存の SIFT に比べると、色情報を組み込むことで検索精度が上がっていることがわかる。しかし、ニガクリタケに関しては類似種を検出できず、検索精度が下がっている。この類似種は誤食事件が多いことから分かるように、非常に形が似ている(図8参照)。そのため既存の SIFT では検出できたが、色情報を加えた SIFT では色が違うこの類似種を検出できなくなったのだと考えられる。このように色情報を加えたことにより検索精度が下がってしまったと思われるケースがあったが、どのような条件だと色情報が悪い方向に働くのか、詳しい分析は引き続き行う予定である。結果として実データでの実験では、全体としては既存の SIFT より性能は上がったが、ニガクリタケのケースでは性能が下がってしまった。

ニガクリタケ

ナメコ



図8 既存の SIFT で検出できた類似種

## 6. まとめ

実験結果により、色情報を SIFT に付加することは有効だとわかった。しかし、SIFT のアルゴリズムからわかるように、SIFT というのは特徴の少ない部分・画像からは特徴量を抽出しないという性質がある。これは特徴点のローカライズ、特に低コントラストの特徴点を消すことにより起こる。それが影響したのか、一般的な SIFT での実験結果では簡単な形の物体ほど検索精度が悪くなっていた。このような、得られる特徴量が少ない画像ほど色情報を使うことの利点があると思われる。しかし、色情報を追加したことにより不正解となったケースもあるように、色の重みを増しすぎると検索結果に悪影響を及ぼすと考えられる。

今回は実験用のデータセットの性質上、データベース内には検索画像と同じオブジェクトが写っている画像はあっても、似ている違うオブジェクトが写っている画像はほぼ無いので、色によって見分ける必要性のある検索ケースが少なかった。しかし、そのようなケースにおいても検索精度の向上が見られたので、色情報を付加することは有効であったと言える。

## 6.1 今後の展望

今回の実験では、実験に使用したデータ量が少なく、実システム上でも有効に働くか未知の部分がある。よってまずは、さらなる大規模データセットにて実験を行い、同様の結果が得られるか実験する必要があるだろう。

また先に述べたように、既存の SIFT 特徴量のベクトルに追加するという形を取るならば、色情報ベクトルの重みを決定する必要がある。ある程度の重みを付けなければ色情報を付加する効果が出ないのだが、重みを付けすぎること悪影響の原因となる。この重みの適正な数値を、実験により決める必要がある。

## 参考文献

- [藤吉 2007] 藤吉弘亘: Gradient ベースの特徴抽出 -SIFT と HOG-, 情報処理学会研究報告, CVIM.160, pp.211-224(2007).
- [D. G. Lowe 1999] D. G. Lowe: Object recognition from local scale-invariant features, IEEE International Conference on Computer Vision, pp.1150-1157(1999).
- [Alaa E 2006] Alaa E: CSIFT: A SIFT descriptor with color invariant characteristics, IEEE Conference on Computer Vision and Pattern Recognition, pp.1978-1983(2006).
- [阿倍 2008] 阿倍拓也, 原克幸: 局所特徴量の位置関係に基づいた類似画像検索, 社団法人映像情報メディア学会技術報, Vol.32, NO.34, pp.37-40(2008).
- [大坂 2005] 大坂充: 画像コンテンツのクラスタリングによる画像類似検索システム, 公立ほこだて未来大学卒業論文(2005).
- [栢原 2010] 栢原康介, 中川祐治: 背景領域削除による類似画像検索の検索制度向上, 第 3 回楽天研究開発シンポジウム(2010).
- [黄瀬 2010] 黄瀬浩一: 局所特徴量を用いた画像照合による特定物体認識, 人工知能学会誌, Vol.34, No.1(2010).
- [isis 2010] <http://staff.science.uva.nl/~aloi/>