

エッジ画像と色情報を用いた SURF アルゴリズムの検討

A Study of SURF Algorithm using Edge Image and Color Information

○佐々木栄裕*, 今野峻一*, 恒川佳隆*

○Yoshihiro Sasaki*, Syunichi Konno*, Yoshitaka Tsunekawa*

*岩手大学

*Iwate University

キーワード: SURF (Speeded Up Robust Features), 画像認識 (image recognition), エッジ画像 (edge image), 色情報 (color information), 特徴点検出 (interest point detection)

連絡先: 〒 020-8551 盛岡市上田 4-3-5 岩手大学 工学部 恒川佳隆,

Tel.&Fax.: (019)621-6468, E-mail: tsune@iwate-u.ac.jp

1. はじめに

近年, 画像認識や特徴点追跡が様々な分野で利用されている. 例としては, デジタルカメラの顔認識, 防犯カメラの人認識・追跡, 指紋認証などが挙げられる. これらの画像認識や特徴点追跡では局所特徴量抽出処理が行われている.

局所特徴量とは, 特徴点検出法により画像中の濃淡の変化が大きい特徴点を検出し, その特徴点周りの領域を画素値や微分値により特徴ベクトルにしたものである.

画像認識や特徴点追跡に局所特徴量を利用する手法として Lowe の Scale-Invariant Feature Transform (SIFT)¹⁾ がよく用いられている. SIFT は画像認識に必要な特徴点の検出と特徴量の記述を行う手法であり, その特徴量は画像の回転, スケールの変化, 照明変化に対して不変性を持っている. また, 処理には画素の輝度を用いている. しかし, SIFT は計算コストが高く, SIFT を用いたアプリケーションではリアルタイム処理

が必要なものが多いため, 処理の高速化が重要なテーマとなっている. そこで, SIFT を拡張した手法として, 必要な処理を近似処理に置き換え高速化した Speeded Up Robust Features(SURF) が Herbert Bay らによって提案されている²⁾. SURF は SIFT を高速化したものであるが, その処理の精度は SIFT に対してあまり向上していない. 原因として, 認識したい物体上に検出される特徴点が少ないことや, 物体の背景の変化により, その特徴が変化してしまうからだと考えられる.

本報告では, 高速な処理が可能な SURF に注目し, 入力画像のエッジ画像と色情報を用いることで高性能化した SURF を提案し, 道路標識認識を例に, エッジ画像と色情報を用いることの有用性を検討する. 高精度化の手法としては, エッジ画像を用いて標識上に検出される特徴点を増やし, 色情報により標識外の部分に現れた特徴点を削除する. この手法により標識認識に使

用される点数を増やし、認識精度の向上を図る。

2. SURF

SURF は照明や回転、スケールの変化に不変な特徴量を記述するため、画像を用いたトラッキングや一般物体認識、画像のより合わせなどに用いられる。SURF の処理の概要は以下の通りである。

- 1) 特徴点の検出
- 2) 特徴点の絞り込み
- 3) オリエンテーションの決定
- 4) 特徴量記述

2.1 特徴点の検出

特徴点とは特徴抽出に適した点であり、その点の周りの輝度の勾配情報から特徴量と呼ばれる情報を記述する。ヘッセ行列によって DoG 画像を作成し、DoG 画像を用いて極大値を検出することで特徴点を検出する。

2.1.1 DoG 画像の作成

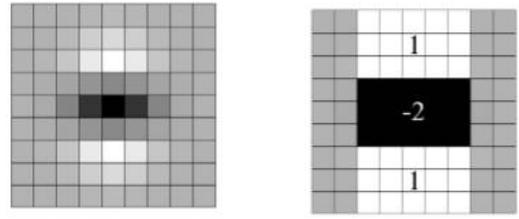
DoG 画像の作成はヘッセ行列を用いる。ヘッセ行列は式(1)となり、 $G(x, y, \sigma)$ はガウス関数、 $L(x, y, \sigma)$ はガウス関数の二階微分、 $I(x, y)$ は入力画像である。

$$H(x, y, \sigma) = \begin{bmatrix} L_{xx}(x, y, \sigma) * I(x, y) & L_{xy}(x, y, \sigma) * I(x, y) \\ L_{xy}(x, y, \sigma) * I(x, y) & L_{yy}(x, y, \sigma) * I(x, y) \end{bmatrix} \quad (1)$$

$$L(x, y, \sigma) = G''(x, y, \sigma) \quad (2)$$

$$G(x, y, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (3)$$

ここで、式(2)を Fig.1(a) のように 2次元のフィルタ形式にしたものを LoG(Laplacian Of



(a)LoG filter(L_{yy}) (b)D filter(D_{yy})

Fig. 1 LoG filter と D filter (y 方向)

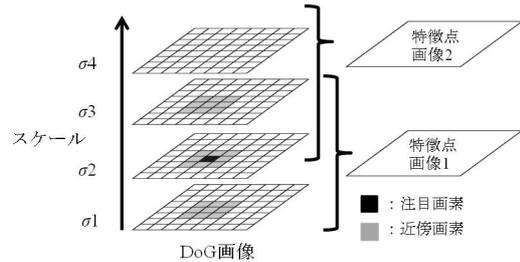


Fig. 2 極値検出

Gaussian) フィルタと呼ぶ。Fig.1は y 方向のフィルタの例である。LoG フィルタと入力画像の畳みこみでは処理時間がかかってしまうので、Fig.1(b) のように近似したフィルタを用いて高速処理する。近似したフィルタを D フィルタと呼ぶ。その後、式(4)によってヘッセ行列を計算する。ここで、式(4)の 0.9 は、LoG フィルタを D フィルタに近似した際の誤差を補う係数である。

$$\det(H_{approx}) = D_{xx} * I(x, y) \times D_{yy} * I(x, y) - (0.9 \times D_{xy} * I(x, y))^2 \quad (4)$$

式(4)を計算すると Fig.2 のような DoG 画像が一つ作成される。次にガウス関数の σ を大きくしていき、すなわち D フィルタのサイズを大きくしていき、式(4)を再び計算し DoG 画像をまた一つ作成する。この処理を繰り返し行い、DoG 画像を最低 3 枚作成する。その後、作成した DoG 画像を用いて極値検出を行い、特徴点を検出する。

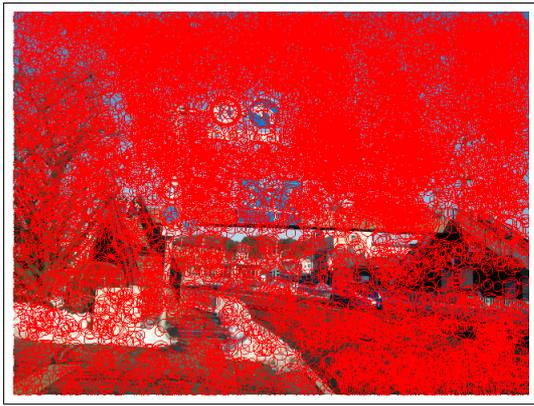


Fig. 3 検出された特徴点



Fig. 4 絞り込み後の特徴点

2.1.2 極値検出

特徴点の検出は Fig.2 のように DoG 画像 3 枚一組で行う。DoG 画像の注目画素と、その周りの近傍 26 点を比較し、極値であった場合、その画素を特徴点として検出する。この極値検出は、 σ の値の小さい DoG 画像から行う。一度極値が検出された画素は、より大きなスケールで極値が検出されても特徴点としない。この処理をスケールの異なる DoG 画像の全画素に対して行う。

この処理を行うと、DoG 画像 n 枚に対して特徴点画像が $n-2$ 枚作成される。作成された特徴点画像をすべて重ね合わせると、Fig.3 のようにすべての特徴点が検出された画像が作成される。Fig.3 の円の中心にあたる部分が特徴点である。円の大きさは、特徴点が検出されたスケールの大きさを表している。

2.2 特徴点の絞り込み

Fig.3 のままでは特徴点が多く、DoG 値が小さい点やエッジ上の点が含まれている。これらの点はノイズや背景の影響を受けやすく特徴抽出に不向きな特徴点である。そこで、コントラストによる絞り込みと主曲率による絞り込みで不要な特徴点を削除する。

まずコントラストによる絞り込みでは DoG 値によって特徴点の削除を行う。DoG 値とは式

(4) のヘッセ行列を計算した後の値であり、Fig.2 の画素値の事である。コントラストが低いとノイズによる影響を大きく受けてしまう。DoG 値はコントラストが低い領域では小さくなるため、DoG 値が閾値以下の特徴点を削除する。微分は値の変化量を見ているので、輝度変化が大きい部分で値が大きくなる。すなわち DoG 値が小さい特徴点は空や道路といった局所的に輝度変化が少ない部分に現れやすく、コントラストによる絞り込みでは主にその部分が削除される。次に主曲率による絞り込みを行う。主曲率による絞り込みはヘッセ行列の固有値から、エッジ上にある特徴点を判断し削除する。ヘッセ行列から求められる第 1 固有値を λ_1 、第 2 固有値を λ_2 ($\lambda_1 > \lambda_2$) とすると、 $\lambda_1 \gg \lambda_2$ または $\lambda_1 \ll \lambda_2$ の場合、その点はエッジ上にあると判断できる。これらの絞り込みを行った結果を Fig.4 に示す。

2.3 オリエンテーションの決定

検出された特徴点に対して特徴点を比較する要素である特徴量の記述を行う。まず、検出された特徴点のオリエンテーションを求める。オリエンテーションとは特徴点における方向を表し、特徴量記述の際にオリエンテーションにより向き正規化を行なうことで、回転不変な特徴量となる。オリエンテーションの決定は、検

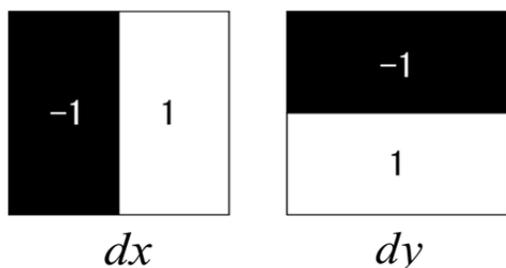


Fig. 5 Haar 型フィルタ

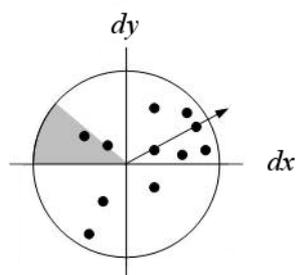


Fig. 6 オリエンテーションの決定

出された特徴点の位置から、直径が検出された特徴点のスケールサイズ $\sigma \times 6$ の円領域内で勾配方向 (dx, dy) を Haar 型フィルタにより求め、それらを用いて決定する。Haar 型フィルタとは Fig.5 に示すようなフィルタである。求めた勾配 (dx, dy) を Fig.6 のようにプロットし、灰色の部分に 15 度ずつ回転させながら、灰色の中の点を式 (5) で計算する。そして M_n の値が一番大きい時の角度をオリエンテーションと決める。

$$M_n = \sum dx^2 + \sum dy^2 \quad (5)$$

2.4 特徴量記述

特徴点のスケールとオリエンテーションを元に、64 次元の特徴量の記述を行う。特徴量の記述は検出された特徴点の周辺領域の勾配で計算する。周辺領域の範囲は、一辺が検出された特徴点のスケールサイズ $\sigma \times 20$ の正方領域である。そしてこの領域をオリエンテーション方向に回転させ、この向きにあわせた勾配を計算す



Fig. 7 特徴点マッチング結果

る。これにより回転に不変な特徴を得る。次に、定めた周辺領域を 4×4 のブロックに分割し、各ブロック内で勾配 (dx, dy) を計算する。そして各ブロック内で $\sum dx$, $\sum |dx|$, $\sum dy$, $\sum |dy|$ を求めることにより、 $4 \times 4 \times 4 = 64$ 次元の特徴量が得られる。

3. 特徴点のマッチングと道路標識認識

道路標識を認識する場合、カメラから得られる入力画像から特徴点を検出し、特徴量の記述を終えたら、認識したい画像 (テンプレート画像) にも同様の処理を行い、それぞれの特徴点の比較を行う。他の特徴点より際立って同じような特徴量を持った点であれば、Fig.7 のように一致させる (特徴点のマッチング)。そして、一致した特徴点が認識したい標識に多く集まっていれば、テンプレート画像が入力画像中に存在すると判断する。

4. エッジ画像と色情報を用いた SURF

SURF による道路標識認識では、標識とのマッチング点数が多いほど判断材料が増え、標識認識・識別が行いやすくなるため、マッチング点数を増加させることが課題となる。そこで、特徴点検出にエッジ画像と色情報を用いた SURF

を提案し、性能の向上を図る。また、SURFを用いて道路標識認識を行う場合、道路標識が回転していることは希なため、オリエンテーションの決定の処理は省略する。

4.1 エッジ画像を用いることによる特徴点増加

画像認識は検出したい特徴点が多いほど、テンプレート画像と一致しているかの判断要素が増えるので精度が向上する。そこでそのままの入力画像から特徴点検出するのではなく、Fig.8(b)のような入力画像からエッジを検出した画像に対して特徴点検出を行うことで、特徴点を多く検出させる。

特徴点検出処理ではヘッセ行列が使用され、DoG画像の作成が行われる。ここで、近似したヘッセ行列の要素 D_{xx} の働きを Fig.9 を基に考える。 D_{xx} フィルタは Fig.1(b) の D_{yy} フィルタを90度回転したものである。Fig.9の灰色部は画素値"0"であり、白い部分はエッジを表している。 D_{xx} フィルタのサイズを変更しエッジにフィルタリングすると、結果は Fig.10 のようになる。フィルタサイズが大きいほど、早い段階でフィルタがエッジ上を通るので、より左側から微分の値が出始める。またフィルタサイズが大きいほど、多くのエッジの値を取り込むのでより大きな値となる。Fig.2のように特徴点はDoG画像3枚一組で、2枚目の注目画素が極値である場合に検出される。Fig.10においても四角い影の部分のように、前後のフィルタ出力結果より大きい値をとる場所(極値)が存在する。式(4)よりこの D_{xx} の値がDoG値に影響するので、四角い影の部分の位置が極値となりやすい。すなわち、エッジとなる局所領域で特徴点検出を行えば、エッジの両側に特徴点検出される。

道路標識は視認性を高めるように目立つ色で着色しており、背景との境界がはっきりと現れる。つまり背景と標識の間にはエッジが存在す

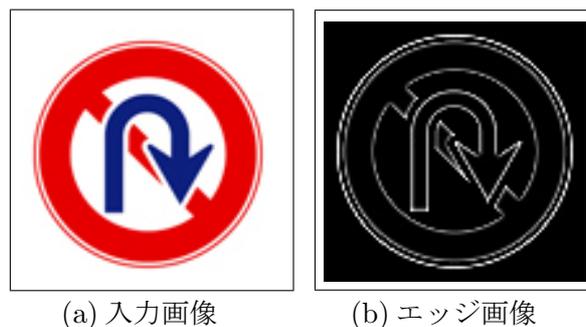


Fig. 8 入力画像とそのエッジ画像

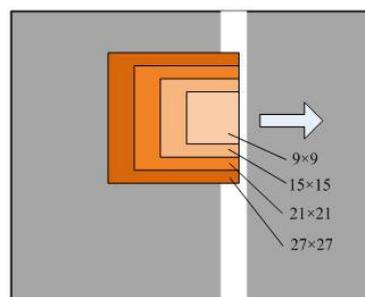


Fig. 9 D_{xx} フィルタリング処理

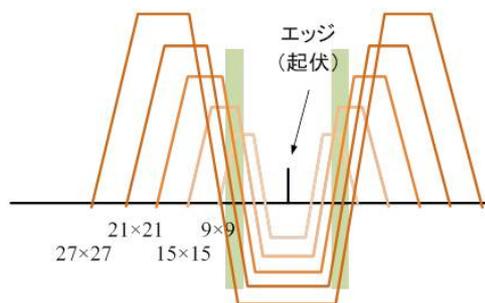


Fig. 10 フィルタリング結果

る。また、標識に描かれるシンボルもはっきりと色分けされているためエッジが検出しやすい。ゆえに、標識上に多数の特徴点を検出できるのである。例として、Fig.8(a)からは39点の特徴点が検出されるが、Fig.8(b)からは472点もの特徴点が検出される。

4.2 色情報による特徴点の絞り込み

エッジ画像を用いることにより特徴点数が増え、比較対象は増加したが、認識したい物体以外の部分、例えば道路標識以外の部分に現れる特



Fig. 11 入力された風景画像

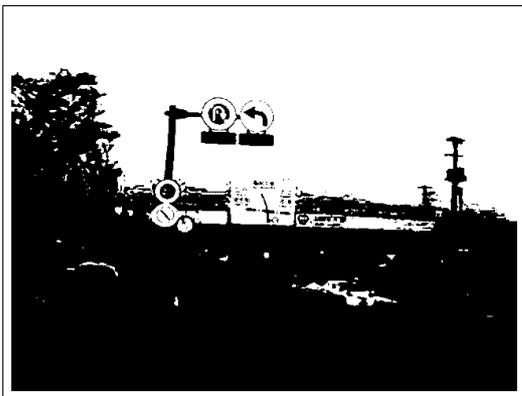


Fig. 12 赤青抽出画像

徴点の数も増加している。この余分な特徴点が多いと特徴点の比較に多くの時間がかかってしまう。そこで色情報を用いた特徴点の絞り込みを新たに行う。

道路標識は主に、赤、青、白から構成されている。したがって画像から赤と青を検出し、その部分の特徴点以外を削除する。白を検出しないのは、風景画像には雲などの白いオブジェクトが多く存在するからである。

画素を一つ一つ注目し、RGB成分のうち、RがGとBそれぞれよりも画素値が30離れている場合は赤、BがRとGそれぞれよりも画素値が30離れている場合は青と判断する。画像Fig.11から赤と青のみを抽出した画像をFig.12に示す。Fig.12の白色以外に現れた特徴点を削除すれば、標識だけに特徴点が見れる。

しかし画像から赤と青のみを抽出したため“転

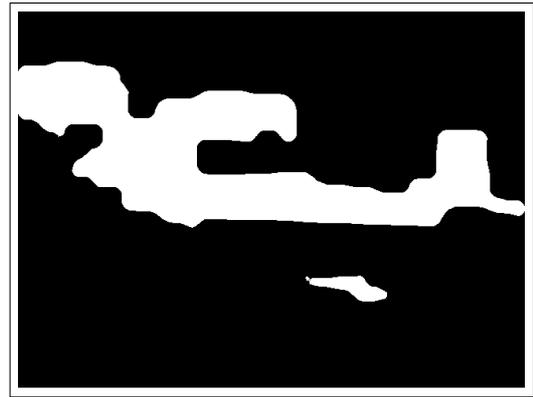


Fig. 13 青空の削除・クロージング処理・メディアンフィルタ適用後

回禁止 (Fig.8(a))”のような白い部分が含まれる標識では隙間ができてしまい、このままでは標識の中に現れた特徴点まで削除してしまう。そこでこの隙間を埋めるため、クロージング処理を行う。そして、ごま塩ノイズを除去するメディアンフィルタを適用した結果がFig.13である。

また、画像から青を抽出するため青空まで検出してしまふ。抽出された青空部分に現れた特徴点を削除するために、赤と青を抽出したFig.12の画像の処理段階で、標識の大きさよりも大きい局所的範囲で、すべて白であったら空と判断しその点の特徴点は削除する、という処理を追加する。

このように色情報を用いた処理の追加により、標識外の部分に現れた特徴点を削除し、標識上に多くの特徴点を残すことが可能となるのである。

5. 提案型SURFでの処理結果

エッジ画像と色情報を用いた提案型SURFでの処理結果を示す。Fig.14は入力画像のエッジ画像から特徴点を検出した後に、色情報によって絞り込まれた特徴点である。各円の中心に特徴点が存在する。Fig.15は提案型SURFによるマッチング結果である。Fig.7と比較すると、標識とマッチングがとれている数が増加していることがわかる。



Fig. 14 色による特徴点絞り込み後の特徴点

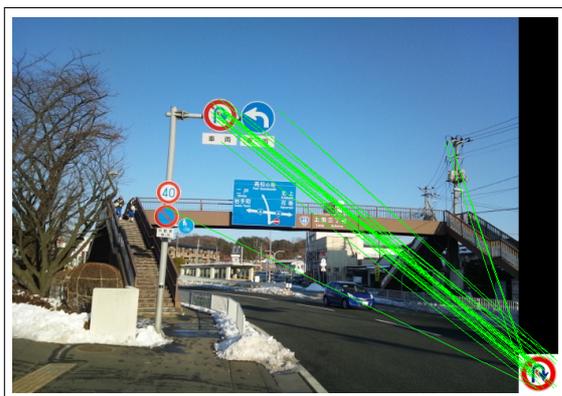


Fig. 15 提案型 SURF によるマッチング結果

従来の SURF と提案型 SURF の性能の比較を示す。評価項目は画像の認識率と処理時間である。標識の認識基準は、認識したい標識とのマッチングが 3 点以上とれている場合である。評価は Intel Core i7-3770 3.4GHz CPU, Math Works 社 Matlab を用いて行った。用意した入力画像は 26 枚であり、画像サイズは 480×640 である。これらの画像と標識のテンプレートとの間でマッチングを行った。処理時間は平均値である。評価結果を Table1 に示す。

提案型 SURF は従来の SURF に比べ、認識率を大きく向上することができた。さらに、エッジ画像に加え、色情報も用いることにより、処理時間の増加が抑えられた。

Table 1 性能比較

	SURF	SURF +エッジ画像	提案型 SURF
認識率 [%]	25.0	58.3	58.3
処理時間 [s]	14.86	47.05	29.93

6. まとめ

本報告では SURF の高性能化を図るために、エッジ画像と色情報を用いた SURF を提案した。そして、道路標識認識を例に性能評価を行った。エッジ画像を使用することで、検出される特徴点の数を増加させ、色情報を用いることで不要な特徴点を削除する。これにより処理の増加、特徴点数の増加に伴う処理時間の増加を抑えつつ、認識率を高めることができた。

今回は道路標識認識を例に処理を行ったが、道路標識以外を認識する場合、色による判断が有効であるとは限らない。しかし、検出したい物体以外の部分に現れた特徴点を削除する手法は有効であり、特徴点を物体上に多く残すことで、認識率の向上が可能であると考えられる。今後の課題としては、物体上に多くの特徴点を残す手法の検討やハードウェア化による処理の高速化が挙げられる。

参考文献

- 1) David G.Lowe "Object Recognition from Local Scale-Invariant Features", Proc. of the International Conference on Computer Vision, Corfu Sept. 1999.
- 2) Herbert Bay, Tinne Tuytelaars, Luc Van Gool "SURF: Speeded Up Robust Features", computer vision-ECCV Lecture Notes in Computer Science, 2006
- 3) 今野峻一, 恒川佳隆, "SURF 特徴点検出を用いた道路標識検出アルゴリズムの検討", 平成 24 年度第 3 回情報処理学会東北支部研究会