
画像認識性能を改善する高精度な特徴量抽出手法の検討

A Study on Feature-Extraction Methods for Improvement of Image-Recognition Performance

井上 俊明

Toshiaki Inoue

要旨 各種のカメラ搭載機器の急速な普及に伴い、撮影・蓄積された画像を有効に活用する画像認識技術への期待が高まっている。特に近年、画像中のさまざまな物体を認識する、一般物体認識技術の研究が盛んに行われるようになった。一般物体認識では、画像の見えの変化に比較的頑強なSIFTやSURFなどの特徴量抽出手法が広く用いられているが、高い認識性能を得るためには、これらのさらなる高精度化が効果的である。そこで本報告では、SIFTに注目した高精度化手法を提案する。従来のSIFTでは、画像から特徴点を検出するためにDoGフィルタを用いていたのに対し、本提案では、DoGフィルタの代わりにウインドウサイズを固定した単純なラプラシアンフィルタを用いる。これにより検出精度を改善でき、かつ再現性のあるより多くの特徴点を検出できることがわかった。公開画像を用いて本提案の有効性を評価した結果、従来のSIFTやSURFと比較して再現性が改善され、また画像照合に応用した場合にも照合精度が改善されることを確認した。

Summary As rapid spread of various devices with built-in cameras, it is strongly expected to develop advanced image recognition technologies. Robust feature extraction methods such as SIFT and SURF are widely used in many kinds of generic object recognition tasks, and further improvements on these methods are essential to boost recognition performance up effectively. In this paper, we describe an improvement of SIFT on keypoint detection. In original SIFT, DoG filter is used to detect keypoints from input images, whereas in our simple method, fixed window laplacian filter is adopted to improve both detection error and repeatability. Repeatability test results using benchmark datasets show that our proposed method outperforms original SIFT and SURF. Image-verification test results also show that the method is effective to improve discriminative performance.

キーワード：画像認識, 特徴量, SIFT, ラプラシアンフィルタ, 再現性

1. はじめに

近年、デジタルカメラをはじめ携帯情報端末、車載機器、ネットワークカメラなど、撮像機能を搭載した機器の急速な普及に伴い、これらの機器で撮影・蓄積された画像を、日常生活の「便利・快適」や「安全・安心」のために有効に活用する技術への期待が高まっている。画像認識技術はその主要候補であり、対象を人物、顔、自動車、道路などに限定した物体検出については既に実用化の段階にある。一方、対象を限定しない一般物体やシーンの認識など、より困難な問題に関する研究も盛んに行われている⁽¹⁾。

画像認識では、画像から抽出したさまざまな特徴量を利用する手法が一般的であるが、認識性能を改善するためには、撮影条件による物体の見え（方向、大きさ、明るさ、色調など）の変化に対して、高精度で頑

強な特徴量抽出手法を用いることが効果的である。たとえばSIFT⁽²⁾やSURF⁽³⁾は、画像の拡大縮小や回転に対して不変な局所特徴量を抽出する手法として、画像認識の研究でよく用いられている。近年提案された一般物体の高精度な認識手法であるBoF (Bag-of-features) でも、SIFTがよく用いられている⁽⁴⁾。

そこで本報告では、画像認識の性能を改善できるSIFTの高精度化手法を提案する⁽⁵⁾。以降、2章でSIFTの概要を説明し、3章でSIFTの精度を改善する手法を提案する。4章で実験結果を踏まえて提案手法の有効性を考察し、5章でまとめと今後の課題を述べる。

2. SIFTによる特徴量抽出手法の概要

SIFTは、D. Loweによって提案された局所特徴量（以後、特徴量と記す）の抽出手法であり、一般物体

認識に関する研究でよく用いられている。実験用のソフトウェアも多く公開されており、さまざまな精度改善手法が提案されている^(6,7)。

SIFTのアルゴリズムは、図1に示すように特徴点検出（ディテクタ）と特徴量記述（ディスクリプタ）で構成される⁽⁸⁾。ディテクタは、画像から特徴点の候補（輝度変化が最大となる画像上の位置）とスケール（特徴点における画像の拡大率に対応する値）を検出し、ローカライズ（有効な特徴点の候補を絞り込むこと）によって、画像の相似変化に不変な特徴点を検出する。一方、ディスクリプタは、ディテクタによって検出された特徴点のそれぞれについて、オリエンテーション算出（特徴点に「向き」を定義）および特徴量記述（特徴点周辺の4×4の矩形領域が持つ輝度勾配の情報から128次元のヒストグラムを生成）によって、画像の回転変化に不変な特徴量（特徴ベクトル）を生成する。

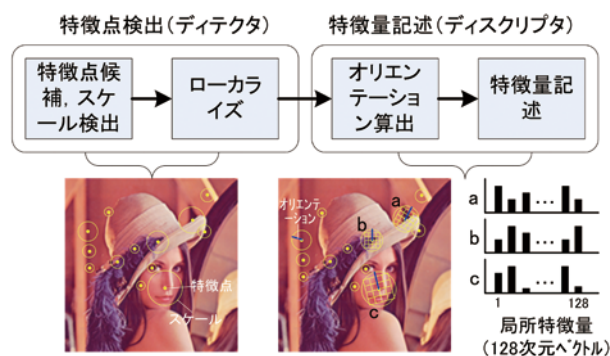


図1 SIFTのアルゴリズム

3. SIFTの高精度化の原理

ディテクタの処理で特徴点の候補を検出する部分（極値検出）は、SIFTのアルゴリズムの初段に位置しており、その検出精度はディスクリプタに至る後の処理に大きく影響すると考えられる。そこで極値検出に注目したSIFTの高精度化を提案する。

従来の極値検出

図2に、DoG(Difference of Gaussian)フィルタによる従来の極値検出方法を示す⁽⁸⁾。増加率 k で平滑化度合いを変えた複数のガウシアンフィルタ（ σ , $k\sigma$, $k^2\sigma$, ...）で平滑化された画像群（スケールスペース画像）に対して、隣接する平滑化画像どうして差分画像を求め。次に差分画像の注目画素と隣接する26画素とで画素値の比較を行い、最大値または最小値をとる場合に、その画素値を極値として検出する。以上を全ての

差分画像に対して行う。

ここで k を1に近づければ、DoGフィルタで検出された極値は理論値に近く（誤差が小さく）なるが⁽²⁾、SIFTでは、極値検出の計算コストが大きなウェイトを占めているため⁽⁹⁾、 k を1に近い値に設定すると、スケールスペース画像の増加により不利になることが懸念される。

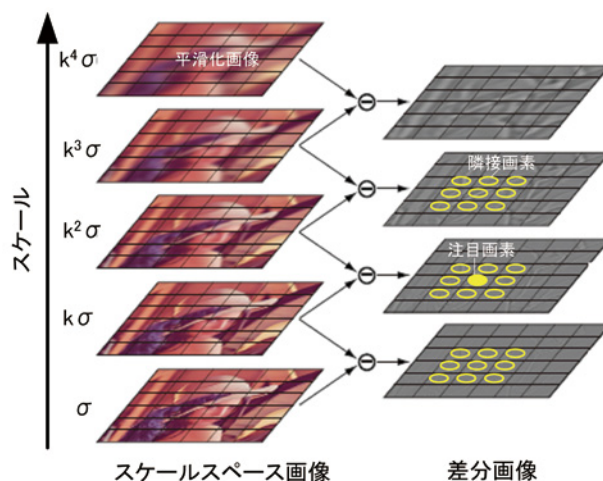


図2 DoGフィルタによる極値検出方法

極値検出の高精度化

そこでスケールスペース画像を増加させずに極値検出を高精度化するため、我々はDoGフィルタの代わりに、ウィンドウサイズを固定したラプラシアンフィルタに係数を乗じたものを用いることを提案する⁽⁵⁾。図3に、ラプラシアンフィルタによる極値検出方法を示す。差分画像は、各々の平滑化画像にラプラシアンフィルタを施して生成される。たとえば、ウィンドウサイズを、注目画素の周囲の隣接画素を含む3×3画素の領域に設定すれば、計算コストの少ない最小サイズのラプラシアンフィルタで実現できる。差分画像からの極値の検出手順は、従来のDoGフィルタによる場合と同様である。

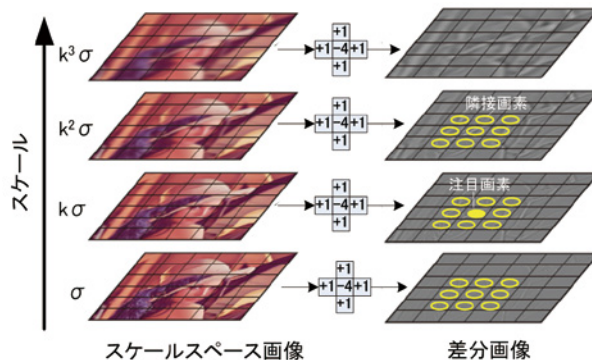


図3 ラプラシアンフィルタによる極値検出方法

精度の比較

DoGフィルタとラプラシアンフィルタは、ともに理論値に対する近似であるため、近似誤差が極値検出の精度に影響することが懸念される。そこで両者の精度を、

$$\text{相対誤差} = \frac{\text{極値の理論値に対するラプラシアンフィルタの誤差}}{\text{極値の理論値に対するDoGフィルタの誤差}}$$

で計算される相対誤差で比較する。

図4に、 σ と k をパラメータとした相対誤差の計算結果を示す。たとえば従来のSIFTで最適な特徴点を得ることができることとされる⁽⁸⁾、 $\sigma=1.6$ 、および $k=2^{1/3}$ の条件で相対誤差は0.318となり、 σ や k の増加に伴い急速に小さくなる。すなわちDoGフィルタをラプラシアンフィルタに置き換えることで極値検出の精度が改善されることがわかる。

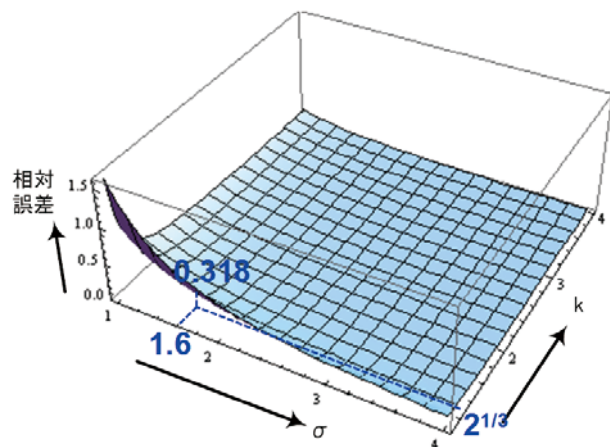


図4 極値検出の相対誤差の計算結果

4. 実験結果と考察

ここでは、実験により特徴点の検出数と再現性、画像照合精度の3種を比較して、本提案の有効性を確認した結果を述べる。

特徴点検出数の比較

図5に、本提案と従来のSIFTによる特徴点の検出例を示す(図中の点が特徴点)。本提案では、輝度変化部分により多くの特徴点が発見されている。また図6に、特徴点の検出数の比較結果を示す。横軸(増幅率)は、DoGフィルタやラプラシアンフィルタによって得られた差分画像の画素値に意図的に乗じた定数で、大きくするほど検出数を増加させる効果がある。2章で述べたローカライズ処理によって、特徴点の候

補を絞り込まれた後の結果であるが、増幅率によらず、本提案ではより多くの特徴点が発見されている。



(a) 従来SIFT (b) 提案SIFT

図5 特徴点の検出例

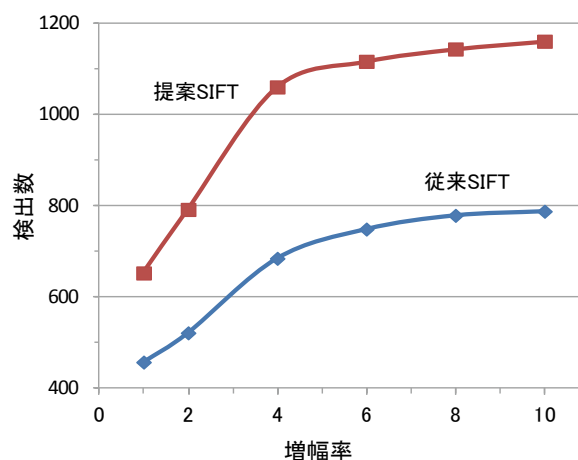


図6 特徴点検出数の比較

再現性の比較

特徴量抽出手法の精度を評価する指標として、再現性(repeatability)がよく用いられている⁽¹⁰⁾。再現性は、幾何学的変換が既知の関係にある2つの画像について、特徴点の検出数(2つの画像のうち検出数が少ない方の特徴点の総数)に対する、対応点の検出数(一方の画像で検出された特徴点の位置を幾何学的変換によって他方の画像上に写像した位置付近で検出することができた特徴点の総数)の比率で定義される。すなわち再現性が大きいほど、その幾何学的変換に関して不変性に優れた特徴量抽出手法である。また評価画像として、Leuven大学の公開画像⁽¹¹⁾がよく用いられている。これは8種類の元画像のそれぞれに対して、拡大縮小、回転、並行移動など5種類の幾何学的変換を施した変換画像、および変換行列で構成されるデータセットである。

図7に、再現性の比較結果を示す。ひとつの元画像に対して各幾何学変換に対応した5通りの再現性が得

られが、ここでは特に本提案の再現性が最も小さい幾何学変換で比較した。参考のため、もう一つの代表的な特徴量抽出手法であるSURF (Speeded Up Robust Features)の再現性も併せて示す。SURFはオープンソースの画像処理ライブラリ(OpenCV⁽¹²⁾)で提供されているものを用いたが、本提案では、いずれに対しても再現性が改善されている。また図8に、本提案で増幅率を1および10とした場合の比較結果を示す。増幅率によって特徴点の検出数だけでなく、再現性も改善されている。

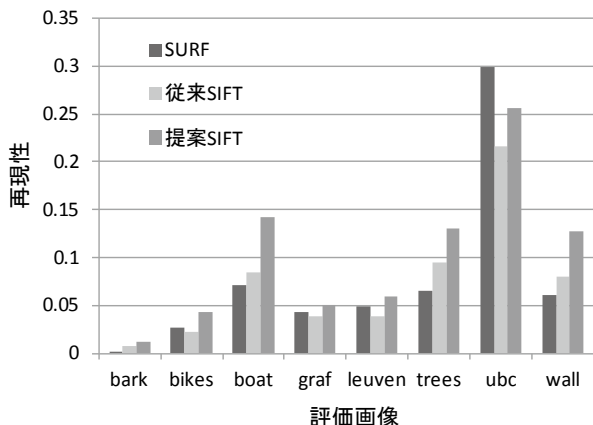


図7 再現性の比較

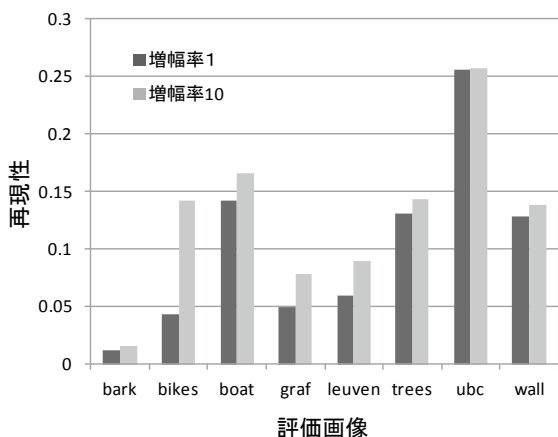


図8 増幅率による再現性の改善効果

画像照合性能の比較

本提案を画像照合に応用した場合の有効性を確認する実験を行った。図9に実験の構成を示す。Amsterdam大学の公開画像(ALOIデータセット)⁽¹³⁾を用いた。登録画像として、同一の照明条件で撮影された500種類の物体の画像をランダムに選択し、テスト画像として、各登録画像と同種類の物体だが異なる照明条件で撮影された画像を選択し、テスト画像と登録画像の照合を行って正解率を評価した。具体的には、

SIFTで抽出された、テスト画像と各登録画像の全ての特徴量(1特徴量あたり128次元ベクトル)を比較して対応点(値の最も近い特徴量のペア)を求め、登録画像の中で最も対応点数の多いものを照合結果とし、照合結果がテスト画像と同種の物体である場合を正解とした。

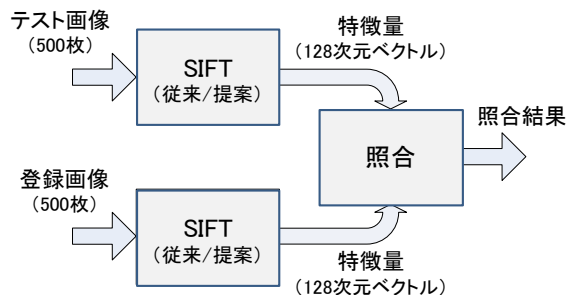


図9 画像照合実験の構成

図10に、正解率の比較結果を示す。本提案では正解率が改善され、さらに増幅率も改善に寄与している。また表1に、本提案(増幅率10)と従来手法の、正解数と不正解数の関係を示す。本提案で不正解となり、従来手法で正解となる画像数が極めて少ないことから、本提案が効果的に正解率に寄与していると考えられる。表1の22画像については分析の余地はあるが、ALOIデータセットは、小さな物体で背景の無い画像が集められているため、対応点数が比較的少なく、かつ誤対応が多く含まれた評価画像による結果であると考えられる。

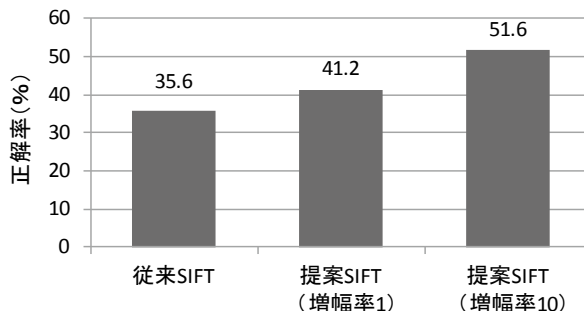


図10 正解率の比較

表1 正解数と不正解数の関係

		提案SIFT (増幅率10)	
		正解	不正解
従来 SIFT	正解	156	22
	不正解	102	220

5. まとめ

本報告では、画像認識の分野で広く用いられている特徴量抽出手法であるSIFTに注目し、その高精度化手法を提案した。SIFTの極値検出の部分、DoGフィルタからウインドウサイズを固定した単純なラプラシアンフィルタに置き換えることで、極値検出の精度を改善するとともに、再現性のある特徴点をより多く検出できることを確認した。また本提案を画像照合に用いることで、照合性能を効果的に改善できることを確認した。

ラプラシアンフィルタを用いることで計算コストの増加が懸念されるが、これは局所的な演算の繰り返しであり、かつ単純な加算とシフト演算で実現できるため、GPUなどを用いて並列処理を行うことにより、容易に高速化が可能である。とはいえSIFT全体の演算量は非常に多く、特にリアルタイム処理が必要な組み込みシステムなどへの応用では、精度を維持しつつ演算量を削減することは今後も課題であり続ける。また実験用の公開画像を用いた評価だけでなく、目的に合わせてさまざまな条件で撮影した自然画像を用いての評価も必要である。さらに実用的な画像照合への応用では、より大規模なデータセットに対する識別精度の評価も必要である。

今回は、SIFTのディテクタに限定した改善であったが、その他の処理、例えばディスクリプタの工夫などを組み合わせることで、さらなる高精度化が期待できる。

参 考 文 献

- (1) 川他, 「パターン認識・メディア理解のグランドチャレンジ」, 電子情報通信学会誌, Vol. 92, No. 8, pp. 639-675, 2009年.
- (2) D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints", International Journal of Computer Vision, Vol. 60, Issue 2, pp. 91-110, Nov. 2004.
- (3) H. Bay, A. Ess, T. Tuytelaars, and L. V. Gool, "SURF: Speeded Up Robust Features", Computer Vision and Image Understanding, Vol. 110, No. 3, pp. 346-359, 2008.
- (4) G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray, "Visual categorization with bags of keypoints", Proc. of ECCV Workshop on Statistical Learning in Computer Vision, pp. 59-74, 2004.
- (5) 井上, 太田, 「ラプラシアンフィルタを用いたSIFTの高精度化に関する一検討」, 電子情報通信学会技術研究報告, Vol. 111, No. 499, pp.7-12, 2011年3月29日.
- (6) Y. Ke and R. Sukthankar, "PCA-SIFT: A More Distinctive Representation for Local Image Descriptors", Conference on Computer Vision and Pattern Recognition Vol. 27, No. 10, pp. 511-517, July, 2004.
- (7) 平山, 新美, 「色情報を用いたSIFTによる類似画像検索精度の向上」, 第25回人工知能学会全国大会, 111-2, 2011年6月.
- (8) 藤吉, 「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.
- (9) 梅本, 黄瀬, 「SIFTにおける特徴点候補検出の複素一次系による高速化」, 画像の認識・理解シンポジウム(MIRU2008), IS2-15, 2008年7月.
- (10) C. Schmid, R. Mohr, and C. Bauckhage, "Evaluation of Interest point Detectors", International Journal of Computer Vision, Vol. 37, No. 2, pp.151-172, June 2000.
- (11) <http://www.robots.ox.ac.uk/~vgg/research/affine/>
- (12) http://opencv.jp/opencv2-x-samples/surf_extraction
- (13) <http://staff.science.uva.nl/~aloi/>

著 者 紹 介

井上 俊明 (いのうえ としあき)

研究開発部に 所属。

LSI回路・アーキテクチャ技術の研究, プラズマディスプレイ製品向け信号処理LSIの開発を経て, 現在, 映像処理技術の研究に従事。