領域を限定したSIFT特徴の抽出

古畑俊一郎† 亀田能成†† 大田友一††

† †† 筑波大学 大学院 システム情報工学研究科 〒 305–8573 茨城県つくば市 天王台 1–1–1
E-mail: †furuhata@image.esys.tsukuba.ac.jp, ††{kameda,ohta}@iit.tsukuba.ac.jp

あらまし SIFT(Scale-invariant feature transform) は画像中の特徴的な小領域を抽出する有用な手法であり、画像処理分野において幅広く用いられるようになりつつある。しかしながら、Lowe らの提案する SIFT では、その特徴量算出において画像全体を処理する必要があり、前景や一部の不定形領域内に限定した SIFT 特徴量を得ることはできない。それに対する解決法の一つとして、我々は不定形のマスク領域内で SIFT 特徴量を算出する方法を提案する。これにより、SIFT の演算を前景抽出処理等の画像処理と組み合わせやすくなる。我々の手法は B-SIFT と異なり、演算対象のマスク領域を確定的に限定するため、背景領域の画像情報には影響を受けないという特徴を持つ。

キーワード SIFT、対応点探索、特徴点追跡、特徴検出

SIFT Features Extraction in Selected Regions

Shunichiro FURUHATA†, Yoshinari KAMEDA††, and Yuichi OHTA††

† ††Graduate School of Systems and Information Engineering, University of Tsukuba
Tenoudai 1–1–1, Tsukuba, Ibaraki, 305–8573 Japan.
E-mail: †furuhata@image.esys.tsukuba.ac.jp, ††{kameda,ohta}@iit.tsukuba.ac.jp

Abstract SIFT(Scale-invariant feature transform) is a region-of-interest detection operator which is widely used in computer vision. However, the original SIFT approach needs a whole image as input and it cannot produce SIFT keys which are grown inside arbitrary shape of foreground regions. We propose a new SIFT calculation method in which SIFT keys are defined by the textures inside mask regions with arbitrary shapes. Our method can be easily coupled with other region extraction methods. Different from B-SIFT, we just use the texture information inside the mask regions so that the produced SIFT features are not affected by any textures outside the mask regions.

Key words SIFT, corresponding point search, feature point tracking, detecting features

1. はじめに

Lowe らにより提案された SIFT(Scale-invariant feature transform) [1] は、画像特徴の取得に際して良く用いられる手法の 1 つである。原論文の手法では、画像全体の特徴を使用して SIFT key を算出する (図 1)。そのため、対象とする物体中の背景に近い領域に画像特徴が存在する場合、背景のテクスチャに影響された key が算出されてしまう。このような key は、対象の背景が変わった時、異なる key になることが考えられ、その後に行う key のマッチングの際に、各々の画像特徴をそれらの key からマッチングさせることができない。

本稿では、画像中の特定の領域のみから SIFT key を求める手法を提案する。前提として、取得される画像は固定カメラで撮影されたものとし、背景差分法等の手法により対象領域を確定できるものとする。この領域を以後、マスク領域と呼ぶ。提案する手法では、SIFT key の算出アルゴリズムに改良を加えることにより、マスク領域内の特徴のみから SIFT key を算出する。これにより、背景の影響を受けない SIFT key 算出を実現できる。

2. 関連研究

画像中の特徴的な小領域に注目して画像特徴量を求める研究は、カメラ幾何を求めるための対応点探索や、物体追跡、または Bag of words に基づく画像認識 [2] 等に広く利用され、コンピュータビジョンの基礎となる分野である。

小領域の輝度分布に注目して対応関係を求める手法としてはテンプレートマッチングが古典的であるが、照明変動や対象物体の見え方の変動に弱く、そのままでは 3 次元空間中の物体を扱うのは難しい。また、最初に必要なテンプレートをどのように与えるかという問題が残されているため、様々な状況で容易に利用可能であるとは言い難い。

一方で、Harris detector [3] や SUSAN [4]、KLT tracker [5]

図 1　SIFT 適用の例

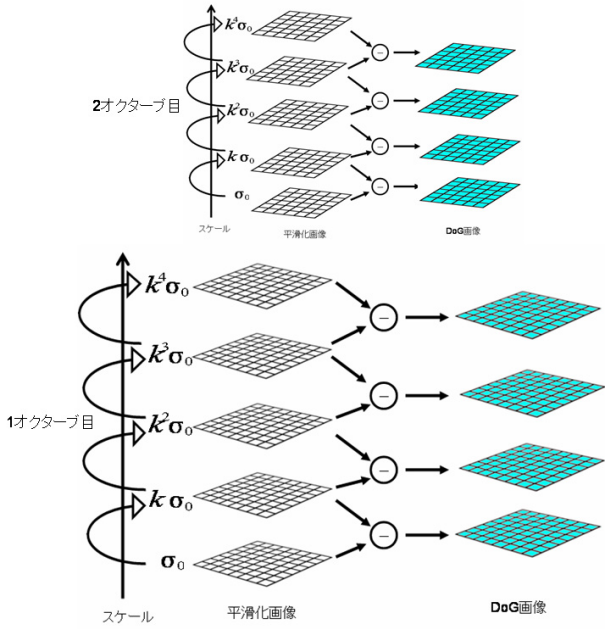


図 2　DoG ピラミッドの概念図

のように、極小領域の特異な輝度分布に着目して特徴点を得る手法が提案され、これらは対応点探索等によく用いられる。しかし、これらの手法では、各抽出点における周辺小領域の記述能力が高くないため、複数の画像中で多くの点が求められる場合、その対応付けには何らかの拘束が必要である。

これらの問題を同時に解決する手法として、近傍領域の輝度分布の特異性を記述子としてもつ、SIFT [1] や SURF [6] のような領域型の特徴量算出方法が注目されている。特に、SIFT は近傍領域の輝度分布を濃度勾配として捉え、かつ大きさ・回転に不変な記述を得ることができるため、注目を集めている。

Lowe らの提案する SIFT は、画像内の輝度の勾配を利用することにより、画像特徴を取得する手法であり、画像からその特徴を表現した SIFT key を生成する。SIFT key は、その座

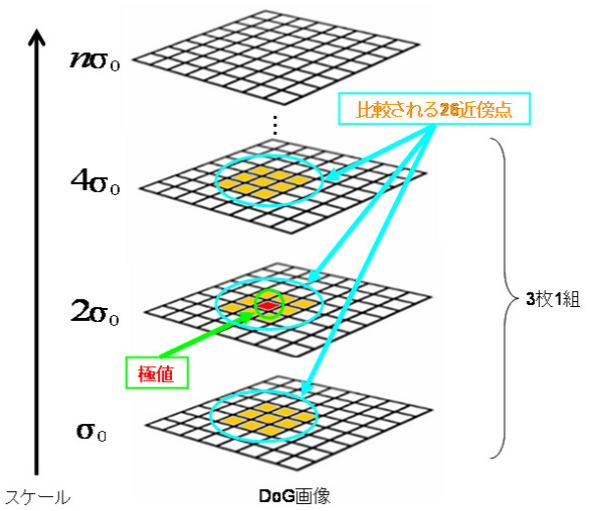


図 3　特徴発生位置の決定方法

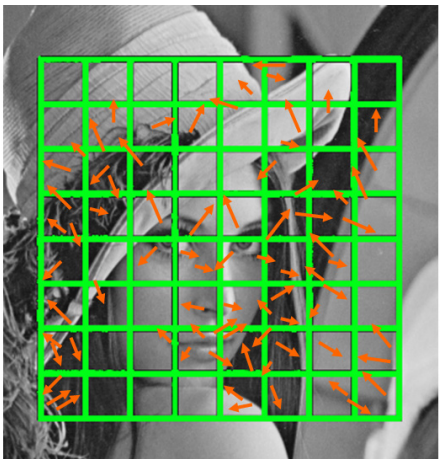


図 4　特徴ベクトルの生成

標、方向、大きさで局所的画像の特徴を表すベクトルであり、以下の方法により生成される [7]。

まず、入力画像をグレースケール化し、何段階かに縮小する。その次に、それらの縮小した画像に対して、ガウシアンフィルタにより平滑化を行う。この時点で縮小された画像とガウスフィルタにより図 2 のような、大きさの違う画像によるガウシアンピラミッドが構成できる。ガウシアンピラミッドの画像の大きさ毎の階層、オクターブは、事前に階層を規定する。つぎに、各オクターブの中で、平滑化された画像間で差を取る。この処理を DoG（Difference of Gaussian）と呼ぶ。DoG により構成されたピラミッドより、各オクターブにおいて特徴点になりそうな画素を探す。具体的には、ある画素の 8 近傍と、同じオクターブのピラミッドの前後の画像における 9 近傍について DoG 値を参照し、合計で 26 近傍に対して、注目点が極値と見なしうる時、同じピクセルが極値と比較し、条件を満たすものを特徴ベクトルとして採用し、その位置を key point として取る (図 3)。その後、それら特徴ベクトルの勾配方向のヒストグラムから、最大値の 80 パーセントを超える特徴ベクトルを選別し（図 4)、それらを統合することで、SIFT key を算出する。

しかしながら、SIFT の手法は画像全体から特徴をとるため、

特定の対象からのみ特徴を取得するという事はできない。そこで前景となる領域を中心として SIFT key を取得するべく提案されたのが B-SIFT [8] である。B-SIFT では、背景の影響を低減することができる。まず、何らかの手法で画像内のエッジを算出し、背景と前景の境界を規定する。その後、前景と判断された領域内で発生した key に対しては、重心に近いものには大きな重みを、また重心から離れたものには小さな重みをつける。また、背景と判断された領域内のものに対しては、限りなく 0 に近い重みをつける。以上の処理により、背景の影響を受けると考えられる対象と背景の境界や背景領域中の SIFT key の発生を抑制することができる。その後のマッチング等の処理において、背景の影響を低減した結果を求める事ができる。

この手法に対し本論文では、マスク画像により対象領域を規定し、マスク領域外の特徴ベクトルは使用せず、マスク範囲内の特徴ベクトルでのみ key を生成する。これにより、B-SIFT とは違い、マスク領域外の影響を完全に排除した SIFT key の算出を実現する。

3. マスク領域内での SIFT の算出

本手法は、与えられたマスク領域に対し、マスク領域内の DoG による極値の算出を一旦行い、そのあと、マスク領域内のもののみで、特徴ベクトルを構成する。以後の処理も Lowe らの SIFT と同様に行うが、各オクターブでの特徴ベクトルを求める際にマスク領域内かどうかを参照するところが、Lowe らとは異なる点である（図 5）。

なお本稿では、固定カメラにより取得された画像を使用するものとする。そのため、マスク領域の確定においては、背景差分法を用いて切り出す（図 6）。この時、一定の大きさ以下の領域は背景差分時のノイズと考えられるので、ラベリングを行い、マスク領域を絞り込み、マスク領域とする。

3.1 DoG での極値の算出

Lowe らの SIFT アルゴリズムと同様に DoG を実行し、極値を算出する。その後、平滑化の度合いが異なる画像間で、特徴ベクトルを算出し、key point を得る。Lowe の SIFT アルゴリズムでは、続いて $2\sigma_0$、$4\sigma_0$、$8\sigma_0$ と、σ_0 の値の大きな画像に同様の処理を行い、最終的な特徴ベクトルを得るが、本手法では、各オクターブの key point 算出後に、マスク領域を用いて、判定処理を行う。

3.2 オクターブ毎の極値判定

算出された key point に対し、事前に与えられたマスク領域内にあるか、判定を行う。領域内と判定された key point は、その後の処理に使用し、領域外と判定された key point は、破棄する (図 7)。

3.3 最終的な特徴ベクトルの構成

この後、以降の σ の違う DoG 画像に対しても同様に処理を行う。これにより、マスク領域内の特徴のみを使用し、特徴ベクトルを構成する（図 8）。こうして求められた特徴を使用し、SIFT アルゴリズムと同様に SIFT key を求める事で、マスク領域内のみからなる SIFT key の取得を実現する。

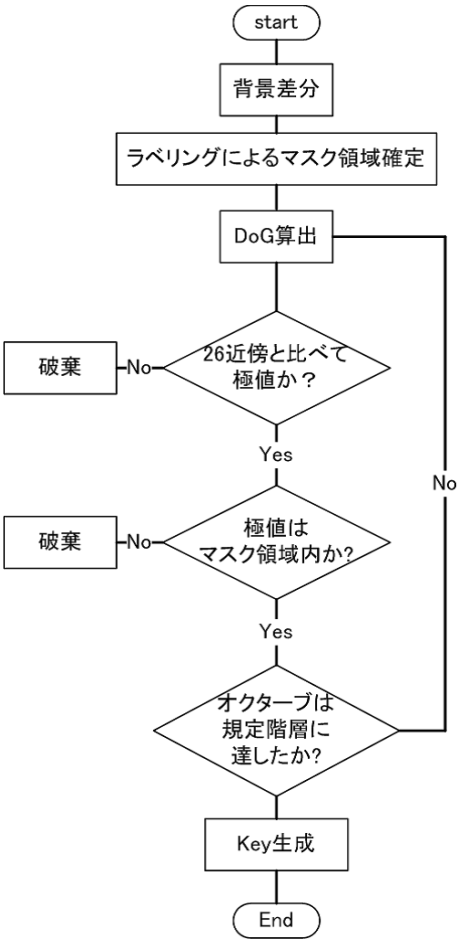


図 5　提案手法の処理の流れ

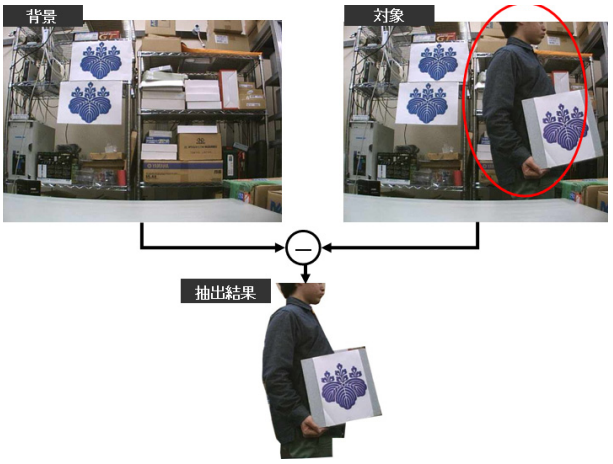


図 6　マスク領域

4. 実　　験

提案手法による SIFT オペレータにより作成された key 同士のマッチングを行い、画像間で同一物体の判定を行う。マッチングのアルゴリズムは、文献 [1] の第 6 章に準じ、best-bin-first アルゴリズムを利用してふるい分けを行い、アフィン変換を行うことで、key 同士のマッチングを確認した。

実験では、様々な画像に対して、Lowe らの SIFT と本手法を適用した。それらによる SIFT key の算出状況と、SIFT key 利

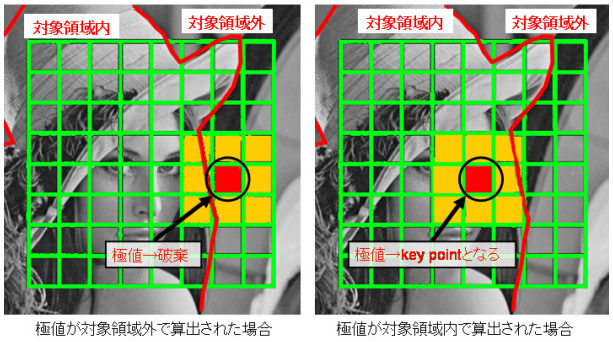


図 7　極値の判定

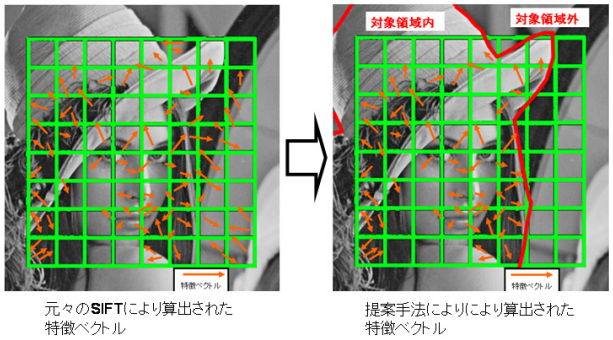


図 8　特徴ベクトルの削除

用したマッチング結果の比較を行った。なお、Lowe らの SIFT の実装に当たっては、Hess らにより作成されたプログラム [9] を用いた。

4.1　人工的に加工した画像に対する実験

まず、実験 1 として、図 9 の入力画像のような、ある程度変化がある、640 × 480 サイズの背景画像に、人工的に小領域画像を貼り付け、SIFT key を算出し、小領域同士のマッチングを行った。まず、SIFT key の算出結果は、Lowe らの手法では 2919 点取得されたのに対し、提案手法では、198 点となった。また、取得された SIFT key も対象とする小領域内からのみ発生している事が確認できた。

次に、上記で求められた SIFT key に対して、白色背景の小領域画像で取得された SIFT key とマッチングを実行した (図 10)。元の手法で算出された SIFT key 同士では、93 点のマッチングが見られたのに対し、提案手法を適用した key 同士では、92 点のマッチングが確認出来た。本手法適用により、求められる SIFT key の数が大幅に減少している。それに対し、マッチング数はほぼ変化がない。これは、本手法により、マッチングとは関係のない対象以外の SIFT key が算出されず、対象のみから SIFT key を求める事ができたためであると考えられる。

4.2　自然画像に対する実験

次に、実験 2 として、図 11 の入力画像ような、机上を撮影した画像に対して SIFT key を算出し、小領域同士のマッチングを行った。まず、SIFT key の算出結果は、Lowe らの手法では、図 11 左画像で 1031 点、図 11 右画像で 338 点取得されたのに対し、提案手法を適用した結果では、図 11 左画像で 361 点、図 11 右画像で 333 点に減少した。両画像とも key が減少

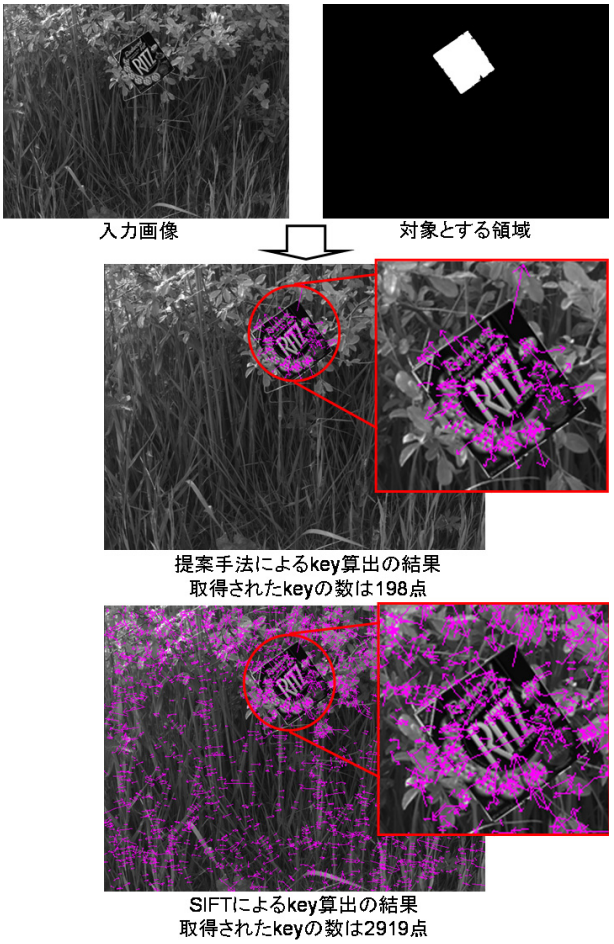


図 9　実験 1:貼り付けた小領域に対する実験結果

しており、特に図 11 左画像では key の減少が顕著であるが、これは、実験 1 と同様に、背景からの SIFT key の算出が行われていないためである。

次に、上記で求められた SIFT key に対して、マッチングを実行した (図 12)。Lowe らの手法で算出された SIFT key では 135 点、提案手法を適用した SIFT key では 94 点のマッチングが確認できた。Lowe らの SIFT を適用した結果では、マスク領域外からも key が算出されているため、背景領域中に同じような物体（本実験における左側に置かれている桐模様の印刷されたノート）が存在するとその影響を受け、マスク領域以外ともマッチングが起こっている。それに対し、提案手法では、key の取得する領域を限定しているため、背景にある対象に似た特徴は対応は取らず、マスク領域内同士でマッチングが行われている。そのため、マッチング数が提案手法では減少しているが、対象同士のマッチングが実現できている事が確認できる。

5.　お わ り に

本研究では、取得した画像中の対象領域のみから SIFT 特徴を取得する方法を提案した。これは、マスク領域を設定することにより、対象領域を規定し、その範囲内で SIFT key を算出することで実現した。実験の結果、マスク領域内でのみ SIFT key を取得することを実現でき、マッチング時には、背景に同

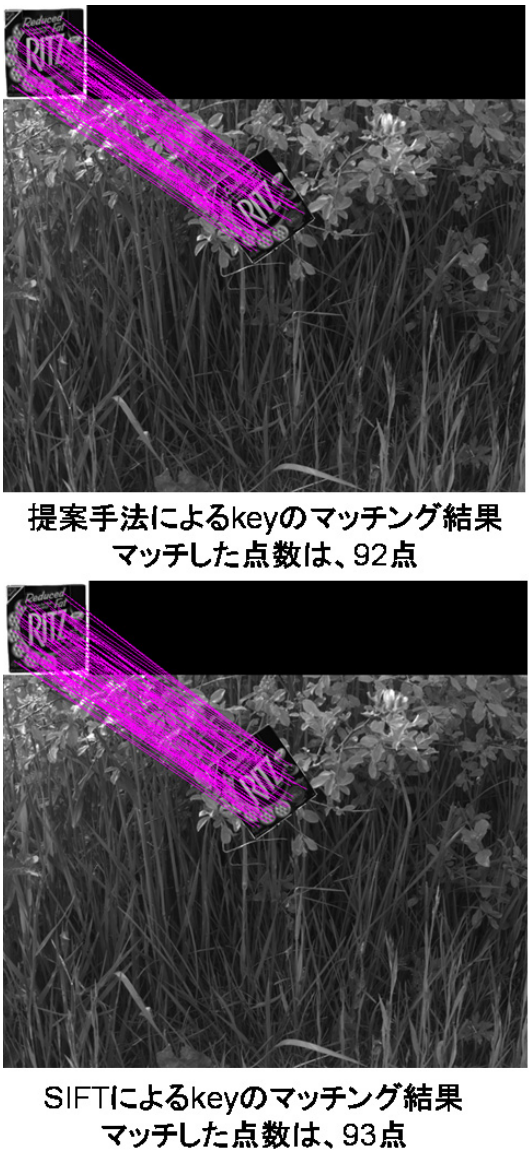


図 10　実験 1:貼り付けた小領域に関するマッチング結果

じような模様があっても、それに影響されず、対象領域内のみでのマッチングが実現できた。

文　　献

[1] David G. Lowe,“Object Recognition from Local Scale-invariant Features,” International Conference on Computer Vision (ICCV’99), Volume:2, Corfu, Greece, pp.1150-1157 Volume: 2, September.1999

[2] Fei-Fei Li, and Pietro Perona “A Bayesian Hierarchical Model for Learning Natural Scene Categories,” Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05), Volume 2, pp.524 - 531, San Diego,USA, June.2005

[3] Chris Harris, and Mike Stephens, “A Combined Corner and Edge Detector”, Proceedings of The Fourth Alvey Vision Conference, Manchester, UK, pp 147-151. 1988

[4] Stephen.M.Smith and J.Michael.Brady, “SUSAN A New Approach to Low Level Image Processing,” International Journal of Computer, Volume 23, Number 1, pp 45-78, May 1997.

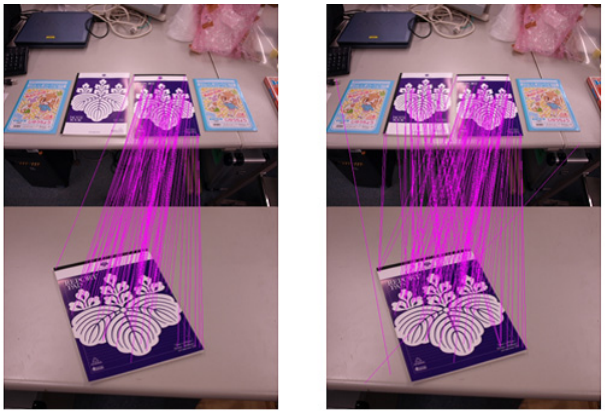
[5] Jianbo Shi and Carlo Tomasi,,“ Good Features to Track,” Computer Vision and Pattern Recognition, 1994. Proceedings (CVPR ’94), pp 593-600,June.1994

図 11　実験 2:机上スペースに対する実験結果

[6] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool, “SURF: Speeded Up Robust Features,” Proceedings of the ninth European Conference on Computer Vision, Vol. 1, pp. 404-417., May 2006

[7] 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

[8] Andrew Stein and Martial Hebert, “Incorporating Background Invariance into Feature-based Object Recognition, Workshop on Applications of Computer Vision (WACV), Volume:1, pp. 37-44, Jane. 2005.



左:提案手法によるマッチングの結果、マッチした点数は**94**点
右:**SIFT**によるマッチングの結果、マッチした点数は**135**点

図 12　実験 2:机上スペースに対するマッチング結果

[9]　Rob Hess, “SIFT Feture Detector,” http://web.engr.oregonstate.edu/ hess/