

スケールに依存しない局所特徴量の誤差評価と画像検索への適用

寺沢 憲吾^{†a)} 長崎 健[†] 川嶋 稔夫[†]

Error Evaluation of Scale Invariant Local Descriptor and Its Application to Image Indexing

Kengo TERASAWA^{†a)}, Takeshi NAGASAKI[†], and Toshio KAWASHIMA[†]

あらまし 画像のインデクシングを行う手法の一つに、画像から特徴点を抽出し、その特徴点の近傍を記述した特徴量ベクトルにより点対点对応を求めようとする方法がある。本論文では、この点対点对応の精度を向上させるため、特徴量ベクトル間の距離尺度として従来より用いられているマハラノビス距離に代わる新しい距離尺度を導入する。人工的に作成した誤差を含む画像を用いて特徴量の観測誤差の従う分布を求め、これに基づいて距離尺度を修正することで、マハラノビス距離による対応付けにおいて生じやすい誤対応を削減することができる。また特徴点の属性のうち固有スケール (characteristic scale) の再現性が比較的高いことに着目し、従来から用いられている Local Jet に加えてこれを特徴量として積極的に活用する手法を提案し、インデクシングに有効であることを示す。

キーワード インデクシング, 特徴点, 局所特徴量, 固有スケール

1. ま え が き

画像データベースから、あるクエリー画像と類似度の高い画像を検索する手法には、主に大局的特徴量に着目する方法と、主に局所的特徴量に着目する方法とがある。局所的特徴量に着目する方法の特長は、ある特定の点の近傍の狭い領域について、特徴的な量を記述することである。この方法は部分的オクルージョンや背景の変化に対してロバストであり、また、点対点对応の構造を求めることが可能であるため、位置決めや、広域画像の中から複数の目的物を検索するというタスクにも適している。

局所的特徴量を用いる画像検索はおおむね、まず画像から情報が集約された点 (特徴点) を抽出し、次にその特徴点の近傍を低次元の特徴量で記述し、その特徴量をもとに点同士の対応を求める、という段階を踏むが、これらの各ステップにおいて様々な手法が提案され、評価・検証されている。

特徴点の抽出法としては Harris function [1] 等を用いてコーナを検出するもの、ラプラシアン等を用いて斑点を検出するもの、ウェーブレットによるもの [2] などがある。これらはいずれもガウス導関数やウェーブレット関数を用いて画像中のある点の近傍を展開することにより実装されるが、こうした展開はガウス関数やマザーウェーブレットのスケールパラメータに依存するため、同一画像であっても解像度が異なれば出力結果が異なるという問題を抱えている。この問題に対し Lindeberg [3] は、問題空間を scale-space に拡張することによって、画像のスケールに依存しない近傍展開を可能とした。また、これに基づき、スケールに依存しない特徴点抽出のための様々な手法が提案されている [4] ~ [6]。

抽出された特徴点の近傍は低次元の特徴量ベクトルによって記述される。この特徴量ベクトルとしては、Local Jet [7] が用いられることが多い。Local Jet はある点の近傍をガウス導関数で展開することにより得られる表現であるが、ここでもまたガウス導関数がスケールの変化や画像の回転に対して不変でないことが問題となる。回転に対する不変性をもたせるための方法としては、回転不変量を構成するもの [2], [8] ~ [10], Steerable Filter [11] を用いて主方向に正規化するも

[†] 公立はこだて未来大学大学院システム情報科学研究科, 函館市 School of Systems Information Science, Future University-Hakodate, 116-2 Kamedanakano-cho, Hakodate-shi, 041-8655 Japan

a) E-mail: g3103004@fun.ac.jp

の [5], [12] などがある．また，スケールの変化に対する不変性をもたせるための方法としては，複数のスケールで特徴量を記述する方法 [9], [13] や固有スケールで近傍半径を定める方法 [5], [12] などが用いられている．Local Jet 以外の特徴量としては，最も単純な image vector を直接用いる方法や，Lowe [6] による SIFT などが代表的である．

特徴量をもとに点同士の対応を求める段階においては多くの手法でマハラノビス距離を用いることとしている．この方法は計算が簡略であるという長所がある一方で，本来各特徴点の種別ごとに別個の共分散行列が必要になるところを全特徴点に対して同一の共分散行列を適用することで代用していたり，特徴量のもつ誤差に正規性を仮定していたりするなどの不完全な点がある．

また，これらの手法はいずれも 1 枚の平面画像のみをもとに画像検索を行うものであるから，三次元物体の認識に用いる場合には自ずと限界が生じる．画像の拡大縮小や回転に対しては前述の手法で対応可能であるが，物体の回転や視点の移動等により物体の見かけが大きく変わってしまう場合はほとんど対応不可能である．その中で，本の表紙やポスターのように二次元構造をもつ物体が三次元空間内で回転する場合のみについては，このときの見かけ上の変形がアフィン変換で近似できることを利用して，アフィン不変な特徴量を構成して対応付けを可能とする研究がある [14]．ただしこのように対応可能な変形を増加させることは，一方で検索精度の低下も同時にもたらす．

本論文で提案する手法は，拡大縮小及び平面内の回転までの変形に対応可能な検索手法である．本論文ではまず，微小な誤差を含む画像を人工的に大量に生成することにより，特徴量の誤差のもつ性質を明らかにし，その性質を踏まえた新しい距離尺度を提案するとともに，それが点同士の対応付けに有効であることを示す．次いで，従来特徴量を記述するための近傍半径を定めるのにとどまっていた固有スケールを画像単位でスケール比を推定するのに活用する方法を提案し，それが画像同士の対応付けに有効であることを示す．

2. Local Jet とその scale invariant 化

本論文で用いる特徴点抽出手法及び特徴量記述手法はいずれも Local Jet を用いる．この章では Local Jet に関する簡単な説明と，そのスケールに対する正規化法について述べる．

Local Jet はある点の近傍をガウス導関数により展開したものであり，具体的には式

$$L_{i_1 \dots i_n}(\mathbf{x}, \sigma) = G_{i_1 \dots i_n}(\mathbf{x}, \sigma) * I(\mathbf{x}) \quad (1)$$

で定義される．ここで $I(\mathbf{x})$ は画像の濃度値， $G(\mathbf{x}, \sigma)$ はガウス関数，添字はその方向の微分を表している．Local Jet による近傍表現は比較的 low order の部分のみで概形を表現することができ，必要に応じて high order の部分を用いることで詳細を記述することができるという便利な性質がある一方で，このままでは画像の解像度とガウス関数のパラメータ σ に依存しているため，異なるスケールの画像検索に用いるには不都合であるという側面ももつ．そこでまずこの依存性を排除するため，Lindeberg [3] や Mikolajczyk and Schmid [5] に倣い固有スケールを導入する．

まず，スケールに対して正規化された m 次微分 D を以下のように定義する．

$$D_{i_1 \dots i_m}(\mathbf{x}, \sigma) = \sigma^m L_{i_1 \dots i_m}(\mathbf{x}, \sigma) \quad (2)$$

この D がスケールに対して正規化されていることは以下のように分かる．スケールの異なる 2 枚の画像 I, I' を考え，これらは $I(\mathbf{x}) = I'(\mathbf{x}')$ ，ただし $\mathbf{x}' = t\mathbf{x}$ ，で関連づけられているものとする．ここでガウス微分を考えると

$$\begin{aligned} \sigma^m G_{i_1 \dots i_m}(\mathbf{x}, \sigma) * I(\mathbf{x}) \\ = t^m \sigma^m G_{i_1 \dots i_m}(\mathbf{x}', t\sigma) * I'(\mathbf{x}') \end{aligned} \quad (3)$$

となり，

$$D_{i_1 \dots i_m}(\mathbf{x}, \sigma) = D'_{i_1 \dots i_m}(\mathbf{x}', t\sigma) \quad (4)$$

が得られ，適切に σ を選べば D の値は画像のスケールに依存しないことが示された．

次に適切に σ を選ぶ方法であるが，画像 I' に対しては $\sigma' = t\sigma$ としなければならないが，この t は未知量であり，事前に知ることはできない．そこで，各画像の各点に対して固有スケールを求めることを考える．固有スケールは， D を用いて適当に定義された特徴量（例えば square gradient, laplacian, Harris function など）に対し， σ 方向に極大を与えるスケールとして定義される．この固有スケールは，画像のスケールの変化に対応して変化するという特徴をもつため，式 (2) の σ の値にこれを用いることにより Local Jet の値を画像のスケールに対して正規化することが

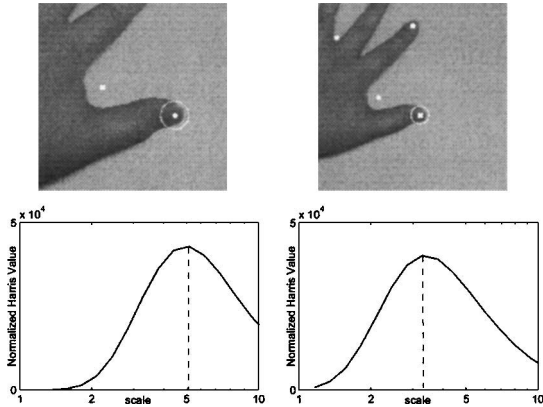


図 1 解像度の異なる 2 枚の画像と、それぞれの対応する点における固有スケール．上段の円の半径と下段のグラフで極大を与えるスケール（破線で示されている）が対応している．

Fig.1 Two images of different resolution and the characteristic scale of corresponding point. Characteristic scale is represented as dashed line in the bottom row, and also represented as the radius of the circle in the top of row.

できる．図 1 は固有スケールの例を示したもので，解像度の異なる 2 枚の画像の対応する点（親指の先）に対し，それぞれについて次章で述べる Harris function の値を求め，その極大を与えるスケールを固有スケールとしたものである．Harris function はコーナ検出器の機能をもつものであるから，ここで求めた固有スケールとはつまりどの近傍半径をとったときに最も強いコーナパターンが観測されるかを表している．上段の画像においては半径 2σ の円を表示しており，この近傍領域内で強いコーナパターンが観測されていることが見てとれる．また，2 枚の画像の固有スケールを比較することにより，確かに画像のスケール比と固有スケールの比が対応していることが確認できる．

3. 特徴点抽出法

前章で定めた固有スケールを用いてスケール不変な特徴点を抽出する方法について述べる．

特徴点の抽出法としては様々な手法が提案されているが，中でも Harris and Stephens [1] によるものが再現性に優れているとされている [15]．ここではこの手法を scale-space に拡張したものについて述べる．

σ を適当な段階で離散化し， x, y, σ の三次元格子空間内の各格子について次式の 2×2 行列 M を求める．

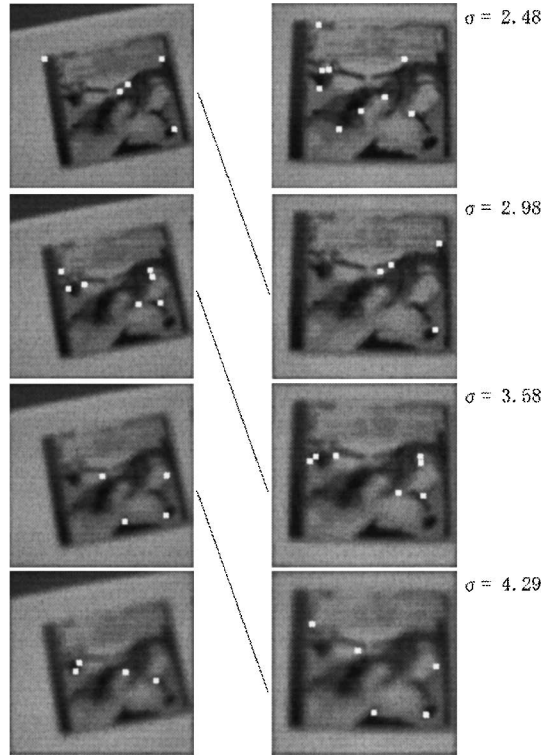


図 2 スケールの違う 2 枚の画像の特徴点

Fig.2 Interest points of two images with different scale.

$$M = \exp\left(-\frac{x^2 + y^2}{2\tilde{\sigma}^2}\right) \otimes \begin{bmatrix} D_x^2 & D_x D_y \\ D_x D_y & D_y^2 \end{bmatrix} \quad (5)$$

ここで $\tilde{\sigma}$ は σ に比例するように定める．本研究の実験においては $\tilde{\sigma} = \sigma$ とした．

次に，求めた行列 M について以下のように定義される R (Harris function と呼ぶ) を求める．

$$R = \det(M) - k \text{trace}(M)^2 \quad (6)$$

k は定数であり，今回の実験では一般に広く使われている $k = 0.06$ とした．この R が x, y, σ の三次元空間内で極大値をとる点を特徴点とする．具体的には，Harris function の値が 26 近傍のいずれよりも大きく，かつ適当なしきい値 t 以上の点，すなわち

$$\forall i, j, k \in \{-1, 0, 1\}$$

$$R(x_n, y_m, \sigma_l) \geq R(x_{n+i}, y_{m+j}, \sigma_{l+k})$$

かつ

$$R(x_n, y_m, \sigma_l) \geq t \quad (7)$$

を満たす点 (x_n, y_m, σ_l) を特徴点として抽出する．今回の実験では 1 画像当たり 50 ~ 100 点程度が抽出されるよう、 $t = 4000$ をしきい値として採用とした．

このようにして特徴点を求めることで、原画像のスケールに依存しない特徴点群を得ることが可能となる．図 2 はその例を示したもので、スケールの違う 2 枚の画像の特徴点を (x, y, σ) の三次元空間内の点として図示したものである． σ の各切断面における特徴点のみを比較すると特徴点是对応しないが、これを三次元空間内の点群としてとらえれば対応するスケール比で対応する特徴点が出現している様子を見てとることができる．

4. 特徴量記述

次に、この特徴点について低次元の特徴量を記述することを考える．特徴量は、光源やカメラ位置、撮影条件等のノイズに対してロバストであることが望ましい．

式 (2) の D において、 σ として固有スケールを採用すれば、これはスケールの変化に対して不変となる．固有スケールを定める際の特徴量の候補は前述のようにいくつか考えられるが、ここでは特徴点算出の際に既に Harris function を求めているため、これをそのまま用いる．したがって固有スケールは特徴点の σ 座標自身である．こうして得られたスケール不変な Local Jet から更に回転に対する不変量を構成するために、

$$\nu = \left[\nu[0] \quad \nu[1] \quad \nu[2] \quad \nu[3] \quad \nu[4] \right]^T$$

$$= \begin{bmatrix} D \\ D_x D_x + D_y D_y \\ D_x D_{xx} D_x + 2D_x D_{xy} D_y + D_y D_{yy} D_y \\ D_{xx} + D_{yy} \\ D_{xx} D_{xx} + 2D_{xy} D_{yx} + D_{yy} D_{yy} \end{bmatrix} \quad (8)$$

を計算する．この ν は、画像平面内の回転に対して不変である [8], [9] ．

更に、照明の変化に対してロバストにするために、 D_x, D_{xx}, \dots の代わりに $D_x/D, D_{xx}/D, \dots$ のように濃度値 D で割った値を用いることにすれば、これらは濃度値の線形変換に対して不変になる．以下ではこの正規化を用いて議論を進める（したがって $\nu[0]$

は用いない）．

なお、ここでは二次の項までを示したが、より高次の項を用いることにより、更にベクトルの次元を増やすことも可能である．しかし、高次の項を用いることは点对点对応の精度を向上させる可能性がある一方で、ノイズに対して敏感になる上、計算量が増大する等のデメリットもある．したがって本研究では二次までの項のみを用いて実験を行うこととした．

5. 特徴量ベクトル同士の対応付け

5.1 距離尺度の導入

このようにして各特徴点について周辺情報を低次元の特徴量で記述できたので、どの特徴点が対応しているかを調べるために、特徴量ベクトル間に距離尺度を導入する．既往の方法においては二つの特徴量ベクトル ν_i, ν_j 間の距離をマハラノビス距離を用いて

$$d_M^2(\nu_i, \nu_j) = (\nu_i - \nu_j)^T \Lambda^{-1} (\nu_i - \nu_j) \quad (9)$$

(Λ^{-1} は共分散行列の逆行列) として定義されることが多い [5], [9] が、この方法は計算が簡略であるという長所がある一方で、本来各特徴点の種別ごとに別個の共分散行列が必要になるところを全特徴点に対して同一の共分散行列を適用することで代用していたり、特徴量のもつ誤差に正規性を仮定していたりするなどの不完全な点がある．そのため、特徴量 ν のある成分の絶対値が大きいような特徴点において、本来ならば対応する点同士であるにもかかわらず大きな距離が算出され、結果的に誤った点对点对応が与えられるケースがしばしば見受けられる．

この問題を解決するため、特徴量ベクトルの各成分のもち得る誤差を評価することを考える．

ある特徴量 ν をもつ特徴点について、誤差を含む特徴量の観測値 $\tilde{\nu}$ を

$$\tilde{\nu}[i] = \nu[i] + e[i] \quad (10)$$

と書くことにする．ここで e が誤差を表している．また、 $\nu[i]$ はベクトル ν の第 i 成分である．マハラノビス距離を用いる方法においては、ベクトル e が多次元正規分布に従うということを仮定している．しかし、 ν が D_x, D_y, D_{xx}, \dots から計算により求められることを考えればこの仮定は不自然であり、むしろ

$$\tilde{D}_* = D_* + \varepsilon_* \quad (11)$$

の ε_* が誤差分布に従うと考える方が自然である． \tilde{D}_*

がこのような誤差をもつと考えると、例えば $\widetilde{D}_x \widetilde{D}_x$ は

$$\begin{aligned}\widetilde{D}_x \widetilde{D}_x &= (D_x + \varepsilon_x)(D_x + \varepsilon_x) \\ &= D_x D_x + 2\varepsilon_x D_x + \varepsilon_x^2\end{aligned}\quad (12)$$

となり、 D_x の値に応じて誤差が拡大することとなり、前述の現象が説明できる。

以上のことから、 $\nu[i]$ の観測誤差 $e[i]$ の分散は D_x 、 D_y 、 D_{xx} 、 \dots 等の値に依存しており、多次元正規分布とはみなせないことが分かる。この値に対する誤差の依存性を正規化するため、本研究では以下のように距離尺度の修正を行う。

D_* の観測誤差 ε_* が正規分布 $N(0, \sigma)$ に従う場合、例えば $\nu[1]$ の誤差は以下のように評価される。

$$\begin{aligned}e[1] &= \widetilde{\nu}[1] - \nu[1] \\ &= (\widetilde{D}_x \widetilde{D}_x + \widetilde{D}_y \widetilde{D}_y) - (D_x D_x + D_y D_y) \\ &\equiv 2\varepsilon_x D_x + 2\varepsilon_y D_y \\ &= 2D_x N(0, \sigma) + 2D_y N(0, \sigma)\end{aligned}\quad (13)$$

独立な確率変数の分散の和の加法性により、右辺を確率変数と見たときの分散は $4\sigma^2(D_x^2 + D_y^2)$ で評価される。よって左辺を $\sqrt{4(D_x^2 + D_y^2)}$ で割れば、誤差の分散は一定の値 σ^2 で評価されることとなる（実際には D_x 、 D_y の値は不明であるので、その最ゆう推定量である \widetilde{D}_x 、 \widetilde{D}_y の値を用いる）。 $\nu[2]$ 、 \dots 、 $\nu[4]$ についても同様の計算を行い、

$$\begin{aligned}e[2] &\equiv 2\varepsilon_x(D_{xx}D_x + D_{xy}D_y) \\ &\quad + 2\varepsilon_y(D_{yy}D_y + D_xD_{xy}) \\ &\quad + \varepsilon_{xx}D_x^2 + \varepsilon_{yy}D_y^2 + 2\varepsilon_{xy}D_xD_y\end{aligned}\quad (14)$$

$$e[3] \equiv \varepsilon_{xx} + \varepsilon_{yy}\quad (15)$$

$$e[4] \equiv 2\varepsilon_{xx}D_{xx} + 4\varepsilon_{xy}D_{xy} + 2\varepsilon_{yy}D_{yy}\quad (16)$$

などを用いて同様の操作を行えば、 ν の各成分について、誤差の分散が σ^2 で正規化されたことになる。

以上をまとめると、特徴量ベクトル ν_i 、 ν_j 間の距離 $d(\nu_i, \nu_j)$ を

$$d^2(\nu_i, \nu_j) = \sum_k (\nu_i[k] - \nu_j[k])^2 / \alpha[k]\quad (17)$$

ただし

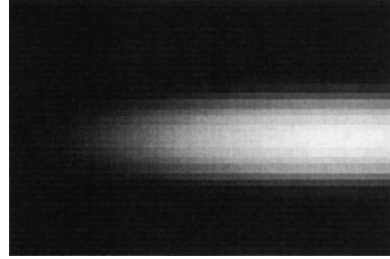


図3 実験に用いた人工画像（拡大図）
Fig. 3 Synthetic image used in our experiment.

$$\begin{aligned}\alpha[1] &= 4(D_x^2 + D_y^2) \\ \alpha[2] &= 4(D_{xx}D_x + D_{xy}D_y)^2 + 4(D_{yy}D_y + D_xD_{xy})^2 \\ &\quad + D_x^4 + D_y^4 + 4D_x^2D_y^2 \\ \alpha[3] &= 2 \\ \alpha[4] &= 4(D_{xx}^2 + 4D_{xy}^2 + D_{yy}^2)\end{aligned}$$

で定義すればよいことが分かる。本研究ではこの $d(\nu_i, \nu_j)$ を、二つの特徴量ベクトル間の距離として定義することとした。

5.2 誤差評価に関する実験

前節で述べた内容を確認するために、人工的な画像に対し様々なノイズを与えて、 D_x 、 D_y 、 D_{xx} 、 D_{xy} 、 D_{yy} の分布を調べた。ここではノイズの発生要因として、 x 、 y 及び σ 座標が離散化されたことに伴うもの、並びにカメラのノイズを考え、それらの人工的再現を行った。

まず

$$I(x) = \frac{1}{1 + e^{-(x^T A x - r^2)}}\quad (18)$$

で楕円画像を作成し（図3）、これを x 方向、 y 方向に微小の平行移動を加えながら、スケールも微小に変化させ、更に10%の白色ノイズを加え、その各々について特徴量を計算し、分布度合を調べた。その結果のうち、 D_x についてのものを図4に示す。左右ほぼ均等に裾の広い分布となっており、おおよそ正規分布であることが確認できる。 D_y 、 D_{xx} 等についてもほぼ同様の分布であり、このことから、 \widetilde{D}_* の誤差が正規分布に従うとの仮定は無理のないものと考えられる。

誤差の発生要因を詳細に探るために x 、 y 、 σ 、10%白色ノイズの各誤差要因に対しての個別の照査も行った。 x 、 y 、 σ の各変位に関する影響は主に特徴点の検出位置が格子点に制限されることにより生じており、変位が1格子間隔に達すると特徴量はもとの値に戻ると

いう周期性をもっている．この周期の範囲内で特徴量はおおむね一様に分布している．各成分について一様分布であるが，それらが三つ重ね合わされることにより，分布に正規性が出てくる．また，白色ノイズの影響はもとより正規的である．これらを複合した結果，すべてを総合した分布は正規分布状になるものと考えられる．

次に，前節で述べた距離尺度の修正による効果を検証する．検証のため，実画像 22 枚から 1153 点の特徴点を抽出し，それらの間の距離 ($1153 \times 1152 / 2 = 664128$ 通り) を計算する．これと，上記で作成したノイズを含む人工画像の対応する特徴点同士の距離を計算したものを比較し，その距離が全組合せの中で何位に相当するかを見た．人工画像の特徴点同士の距離としては，最大距離 (worst) と，平均距離 (mean) の 2 種類を用いた．表 1 にその結果を示す．平均的誤差の場合，従来法においては対応する点は 664128 通り中 13544 位で，これは対応点候補のうちおおむね上位 2.0% 以内に正しい対応点が含まれることを示している．一方改良法においては順位は 2424 位となり，正しい対応点は上位 0.36% 以内に含まれることになり，検出力は約 5 倍となっている．また，最悪の場合においても約 2 倍超の検出力を示しており，本手法の有効性が示されている．

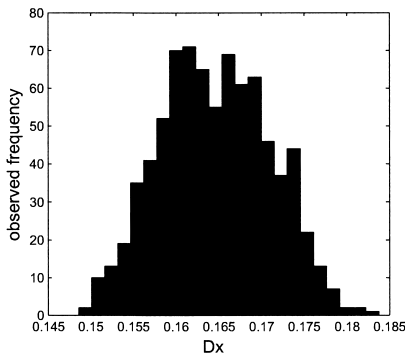


図 4 D_x の分布
Fig. 4 Distribution of D_x .

表 1 ノイズを含む対応点の検出順位
Table 1 Rank of corresponding point with noise.

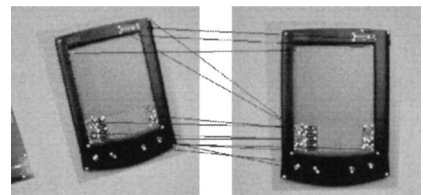
	normal	arranged
worst	127199	54788
mean	13544	2424

6. 自動スケール選択による絞込み

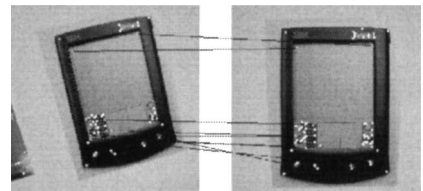
各特徴点における固有スケールは画像の解像度に比例して変化するため，二つの画像間における対応する点同士の固有スケールの比は，常に一定の値をとるはずである．これを用いれば 2 枚の画像間のスケール比をある程度推定することができ，極端に固有スケール比が異なる点同士は対応点候補から除外できることになる．ここでは，この考えに立脚して，点对点対応の候補点を大幅に絞り込む手法を提案する．

第一段階としてまず，クエリー画像の全特徴点からデータ画像の全特徴点への距離を求め，各点に対して最短距離を与える点を対応点候補として仮決定する (図 5(a))．

次に，2 枚の画像間で対応すると仮決定された点の固有スケール比をもとに，投票法により画像間のスケール比を推定する．表 2 にその例を示す．2 枚の画像に対し，仮対応する点の組合せを表上にプロットしていくと，2 枚の画像が対応していれば，上段の表のように対応スケールテーブルの中に，対角線方向に大きい値が集中する部分が形成される．この例の場合，対角線より 1 ブロック右上の対角線方向に大きい値が集中していることが分かる．これを画像間のスケールの比の推定値として採択する．より正確には，下段の表のようにスケール比ごとに組合せ数をまとめて集計し，最大の得票を獲得しているスケール比を調べる．この例では，スケール比 1 : 1.2 が最大の投票を獲得しており，2 画像間のスケール比をおおよそ 1 : 1.2 と推



(a) Result of simple nearest neighbor method



(b) Result after elimination of error matching

図 5 点对点対応の改良

Fig. 5 Improvement of point to point matching.

表 2 固有スケールの対応表
Table 2 Match table of characteristic scale.

Query Image	Database Image							
	1.44	1.72	2.07	2.48	2.98	3.58	4.29	5.15
1.44	3	2	1	2				1
1.72			3	1			3	
2.07			1	4				
2.48					4			
2.98				1		4	1	
3.58							4	
4.29								2
5.15	1					1		

対角線方向に集積した値
Diagonally accumulated count

...	1.44:1	1.2:1	1:1	1:1.2	1:1.44	1:1.72	...
	1	1	4	23	3	2	

定することができる。

これにより 2 枚の画像間のスケール比が推定できたので、対応点候補が大きく絞り込める。すなわち、表 2 のうち、推定された画像スケール比から大きく離れたスケール比をもつ対応点は誤対応であると考えられるため、推定された画像スケール比に対応する点の中から改めて候補点を探しなおせばよい。つまり、第一段階では画像内の全特徴点を対象にして最小距離の点を仮決定したが、ここで改めて対応する固有スケールをもつ点のみを対象にして最小距離の点を候補点として決定し直す。この操作により、第一段階の点对点対応から誤対応を切り落とし、対応の精度を向上させることができる(図 5(b))。

7. 実験結果

本手法がインデクシングに有効であることを示すため、7 種類の物体を撮影した画像データベース (A1, B1, ..., G1) に対し、それと角度やスケールを変えて撮影した画像 (A2, B2, ..., G2) をクエリー画像として、画像検索を行った。その一部を図 6 に示す。計算過程は下記のとおりである。

0. データベース内の画像に対し、あらかじめ特徴点及び局所特徴量(特徴量ベクトル)を算出しておく。画像は 256 階調グレースケールで画素数は 256×256 とし、スケール方向のステップは 10 段階 ($\sigma = 1.2^n$, $n = 1, 2, \dots, 10$) とする。また、特徴点抽出の際に用いる Harris function のしきい値は $t = 4000$ とした。

1. クエリー画像に対し、同様に特徴点及び局所特徴量を算出する。

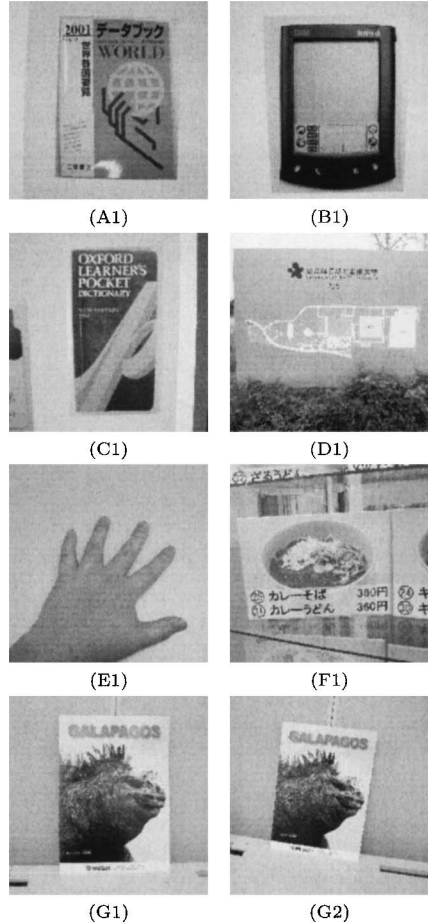


図 6 実験に用いた画像の一部
Fig. 6 Images used in our experiment.

2. データベース内の各画像とクエリー画像で、すべての特徴点間の距離を計算し、最小距離の点を当てはめる。

3. 当てはめた点をもとに、スケール比を推定する。推定は投票法で行い、投票結果で最大を与えるスケール比を推定スケール比として採用する。

4. 推定されたスケール比に ± 1 段階の誤差のみ許容して、それ以外のスケール比の対応点候補は切り捨てる。

5. 最終的に残った対応点候補とクエリーの点との最短距離が一定値(ここでは 0.04 とした)以下であれば、クエリーの点との対応点として確定する。

6. 確定した対応点の個数が最も多い画像を、検索すべき画像として抽出する。ただし、最小スケールと最大スケールで検出された特徴点は、真の意味の固有

表 3 画像検索の実験結果 .
Table 3 Results of image indexing.

(a) Our method.

Query Image	Database Image							推定結果	正誤
	A1	B1	C1	D1	E1	F1	G1		
A2	26	1	2	4	0	7	3	A	
B2	1	9	3	5	0	2	1	B	
C2	3	4	12	2	0	5	4	C	
D2	2	2	0	8	1	7	6	D	
E2	1	0	1	2	6	4	5	E	
F2	5	2	2	2	3	23	14	F	
G2	11	0	0	13	4	9	14	G	

(b) Without elimination with scale.

Query Image	Database Image							推定結果	正誤
	A1	B1	C1	D1	E1	F1	G1		
A2	27	4	15	8	3	19	11	A	
B2	4	13	9	8	1	14	8	F	×
C2	9	5	15	4	0	12	6	C	
D2	3	2	2	14	2	11	13	D	
E2	7	0	1	6	7	6	6	A or E	
F2	16	3	7	24	10	32	27	F	
G2	24	6	12	26	14	36	32	F	×

スケールではないと考えられるので、個数算出から除く。

実験結果を表 3 に示す。比較対照のため、3.、4. のステップを除いて行った実験結果も併せて示してある。表示されている数字は確定した対応点の個数である。

これらの結果を比べると、3.、4. のステップを除いて行った場合はいくつかのケースにおいて誤った画像を検出しているのに対し、提案手法においてはすべてのケースにおいて正しい画像を検出していることが分かる。また、誤った画像に対する対応点の個数も従来法に対して大幅に削減されており、安定性が向上している。

もちろん、3.、4. のステップを除いて行った場合においても、これに更に幾何情報を用いた検定を加える等により、安定した結果を得ることは可能である。しかし、この種の検定は特徴点の数に対して組合せ論的に計算量が増大するという欠点を有している。これに対し、本手法においては計算量のオーダを増大させることなく、結果の安定性を向上させることができる。また、必要に応じて、これに更に幾何情報を用いた検定を追加することにより更に結果を安定させることも可能であるが、この場合においても候補点が従来法より絞り込まれている分だけ、計算量が大きく軽減されたことになる。

8. むすび

本論文では局所特徴量を用いた画像検索における点対点の対応付けの高精度化の手法として、特徴量ベクトル間の距離尺度の改善及び自動スケール選択による対応点の絞り込み手法を提案した。また、これらを実画像の検索に対して適用することにより、この手法の有効性を確認した。

この手法を用いた画像検索を更に高精度化するには、従来手法と同じように特徴量ベクトルを高次元化するが幾何情報による検定を加える等すればよい。しかしそのような従来手法を用いる際にあっても、今回提案した手法により候補点がかかり絞り込まれたため、合計の計算負荷も大幅に縮小されるものと見込まれる。

文 献

- [1] C. Harris and M. Stephens, "A combined corner and edge detector," Proc. Alvey Vision Conf., pp.147-151, 1988.
- [2] N. Sebe and M.S. Lew, "Comparing salient point detectors," Proc. IEEE Int. Conf. on Multimedia and Expo, pp.65-68, 2001.
- [3] T. Lindeberg, "Feature detection with automatic scale selection," Int. J. Comput. Vis., vol.30, no.2, pp.77-116, 1998.
- [4] D.G. Lowe, "Object recognition from local scale-invariant features," Proc. IEEE Int. Conf. on Computer Vision, vol.2, pp.1150-1157, 1999.
- [5] K. Mikolajczyk and C. Schmid, "Indexing based on scale invariant interest points," Proc. IEEE Int. Conf. on Computer Vision, vol.1, pp.525-531, 2001.
- [6] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," Int. J. Comput. Vis., vol.60, no.2, pp.91-110, 2004.
- [7] J.J. Koenderink and A.J. van Doorn, "Representation of local geometry in the visual system," Biol. Cybern., vol.55, pp.367-375, 1987.
- [8] B.M. ter Haar Romeny, L.M.J. Florack, A.H. Salden, and M.A. Viergever, "Higher order differential structure of images," Image Vis. Comput., vol.12, pp.317-325, 1994.
- [9] C. Schmid and R. Mohr, "Local grayvalue invariants for image retrieval," IEEE Trans. Pattern Anal. Mach. Intell., vol.19, no.5, pp.530-534, 1997.
- [10] 長崎 健, 戸田真志, 川嶋稔夫, "日常生活における行動記録映像の構造化," 信学技報, PRMU2002-164, 2002.
- [11] W.T. Freeman and E.H. Adelson, "The design and use of steerable filters," IEEE Trans. Pattern Anal. Mach. Intell., vol.13, no.9, pp.891-906, 1991.
- [12] O. Chomat, V.C. de Verdière, D. Hall, and J.L. Crowley, "Local scale selection for Gaussian based description techniques," Proc. European Conf. on

Computer Vision, LNCS 1842, vol.1, pp.117-133, 2000.

- [13] Y. Dufournaud, C. Schmid, and R. Horaud, "Matching images with different resolutions," Proc. IEEE Conf. on Computer Vision and Pattern Recognition, vol.1, pp.612-618, 2000.
- [14] K. Mikolajczyk and C. Schmid, "An affine invariant interest point detector," Proc. European Conf. on Computer Vision, LNCS 2350, vol.1, pp.128-142, 2002.
- [15] C. Schmid, R. Mohr, and C. Bauckhage, "Comparing and evaluating interest points," Proc. Int. Conf. on Computer Vision, pp.230-235, 1998.

(平成 16 年 10 月 8 日受付, 17 年 2 月 1 日再受付)



寺沢 憲吾 (学生員)

平 10 東大・工・土木卒。平 12 同大学院修士課程了。現在, 公立はこだて未来大学大学院博士課程在学中。物体抽出, 画像認識に関する研究に従事。



長崎 健 (正員)

平 4 北大・工・情報工卒。平 6 同大学院修士課程了。平 9 同大学院博士課程単位取得退学。平 9 (株) ビーユージー入社。平 6~9 まで日本学術振興会特別研究員。平 12 よりはこだて未来大講師。コンピュータビジョンの研究に従事。日本ロボット学会, 情報処理学会各会員。



川嶋 稔夫 (正員)

昭 55 北大・工・電子卒。昭 57 同大学院修士課程了。同年苫小牧高専助手。昭 59 北大工学部助手。平 2 同講師。平 4 同助教授。平 12 はこだて未来大教授となり現在に至る。コンピュータビジョン, センサフュージョンに関する研究に従事。工博。情報処理学会, 日本ロボット学会各会員。

情報処理学会, 日本ロボット学会各会員。