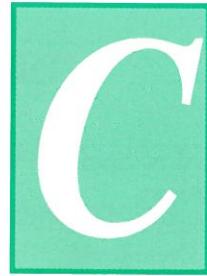


電気学会論文誌

電子・情報・システム部門誌



Vol.129 No.5 2009

特集：ビジョン技術の新たな潮流

IEEJ Transactions on Electronics, Information and Systems



社団法人 電気学会

The Institute of Electrical Engineers of Japan

方向選択性局所特徴を用いた統計学習による 多視点物体認識システム

非会員 横野 順* 非会員 Tomaso Poggio**

Object Recognition Using Boosted Oriented Filter Based Local Descriptors

Jerry Jun Yokono*, Non-member, Tomaso Poggio**, Non-member

Object recognition system based on local descriptors is increasingly used recently because of their perceived robustness with respect to occlusions and to global geometrical deformations. Such a descriptor—based on a set of oriented Gaussian derivative filters—is used in our recognition system. In this paper, we explore the multiview 3D object recognition and multiview face identification. Basic idea is to find discriminant features to describe an object across different views. Boosting framework is used to select features out of huge feature pool created by collecting the local features from the positive training examples. We conduct experiments on 3D objects and face images and get excellent recognition rate. Comparison to SVM is also noted in the paper.

キーワード：物体認識, 任意視点顔識別, 局所特徴量, 特徴量選択, ガウシアン局所特徴

Keywords : object recognition, face identification, local features, feature selection, AdaBoost, Gaussian descriptors

1. はじめに

実環境下での 3 次元物体の認識は画像認識において困難な課題の一つである。近年の報告では、照明変化や物体の回転、大きさ、ゆがみなどに対して頑健な局所特徴量が提案され、実環境で動作するシステムが増えてきている。

Lowe⁽¹⁴⁾は、SIFT 特徴を用いて局所的なパターンの変化に強い局所特徴量を提案し、特定物体の認識を一枚の画像から実現して見せた。Viola-Jones⁽³⁾は、Poggio ら⁽²⁴⁾の提案した Haar 特徴を高速に計算し、更に AdaBoost を用いた統計学習器で高性能かつ高速なシステムを実現した。以降、多数のシステムが局所特徴量の統計学習による認識手法を報告してきた^{(1)(5)(11)(15)~(17)(20)(21)}。例えば、Csurka⁽¹⁾らの Bag-of-Features 法では、局所特徴量のヒストグラムを SVM で識別することで、幾何学的拘束なしに高性能な認識が出来ることを報告した。Leibe ら⁽¹⁶⁾の Implicit Shape Model では、局所特徴量のマッチングを最初に行い、幾何学的な関係をアフィンパラメータ空間で投票することで “implicit” にモデル化し、物体の検出とセグメンテーションを同時にを行う方法を提案した。Dalal ら⁽¹¹⁾は、HOG と呼ばれるヒストグラムベースの局所特徴を使い SVM で識別器を構成することで高精度の人検

出システムが実現できることを示した。

だが、これらの提案システムでは、学習に用いたある特定の視点における物体モデルを学習しており、任意の視点からの認識にフォーカスをおいているわけではない。視点の変化による物体の見えの変化に対してはまだ課題が多い。本論文では、様々な視点からの物体画像を集めて、3 次元物体を記述するための特徴量を自動的に抽出し、物体モデルを学習するためのフレームワークを提案する。幾何学的な拘束条件は、局所特徴量のマッチング閾値に含ませることで位置情報を有効に用いたシステムを構築した。本論文では、任意視点三次元物体、視点・位置がある程度正規化された物体、任意視点顔識別の実験を行い、提案システムの有効性を示した。

2 章で本論文で用いる局所特徴量であるガウシアン微分方向フィルタの紹介、3 章では幾何学的拘束によるマッチングを行って識別器を学習するフレームワークの紹介、4、5 章では、我々が行った認識実験について報告する。

2. ガウシアン微分方向フィルタ

<2.1> Steerable Filters ガウシアン微分方向フィルタ (Oriented Gaussian derivatives) は、方向と周波数に選択性を持つフィルタで、照明変化や視点変化に対しても不变性が強い。1 次微分はエッジ特徴、2 次微分はバー特徴を抽出しており、より高次になるとより複雑な特徴を抽出できる。ガウシアン微分方向フィルタは、Freeman ら⁽⁴⁾により提案された SteerableFilters を用いて効率的に計算することが出来た。

* ソニー（株）システム技術研究所
〒141-0001 東京都品川区北品川 5-1-12

Sony Corporation System Technologies Lab,
5-1-12 kitashinagawa, shinagawa-ku, tokyo 141-0001

** Dept. of Brain and Cognitive Sciences, Massachusetts Institute of Technology
43 Vassar Street, Cambridge, MA 02124

て、いくつかの基底フィルタで求めた値を係数付き線型和をとることで、任意の角度のフィルタ出力が得られる。

ある方向 θ の n 次ガウシアン微分は、幅 σ のガウシアン関数 G を用いて、以下のように表現できる。

$$G = \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right)$$

$$G_1(\theta) = \cos(\theta)G_1(0^\circ) + \sin(\theta)G_1(90^\circ)$$

$$G_2(\theta) = k_{21}(\theta)G_2(0^\circ) + k_{22}(\theta)G_2(60^\circ) + k_{23}(\theta)G_2(120^\circ)$$

$$k_{2i}(\theta) = \frac{1}{3}\{1 + 2\cos(2(\theta - \theta_i))\}$$

$$G_3(\theta) = k_{31}(\theta)G_3(0^\circ) + k_{32}(\theta)G_3(45^\circ) + k_{33}(\theta)G_3(90^\circ) + k_{34}(\theta)G_3(135^\circ)$$

.....(1)

ここで、 $k_m(\theta)$ は基底係数である。局所特徴量は、特定の局所パターンに選択性が求められる一方で、回転や位置ずれ、スケールの変化などのアフィン変化に対して不变性があることが望ましい。一般には、選択性が強ければ不变性が弱くなり、不变性が強くなれば選択性が弱くなるという、トレードオフの関係にある (selectivity-invariance tradeoff)。ガウシアン微分方向フィルタがこれらのトレードオフをうまく表現していることは、例えば文献⁽⁶⁾⁽⁹⁾で実験的に示されている。本論文の実験では、微分次数は3次まで、4方向、3スケール（ガウシアンの幅 $\sigma=1,2,4$ ）を用いている。これにより、ある画素点において36次元の特徴量が計算されることになる。各フィルタが畳み込み演算を行う領域は小さいため、図1上段に示すように5画素離れた点におけるジエットを4つ結合して、より複雑なパターンを抽出出来るように180次元の特徴量ベクトルとした。回転に対する不变性は、Steerable Filtersの回転性により、任意の角度にフィルタ出力を回転できる（図1下段参照）。まず、ある画素点でのガウシアンの幅が最も大きいフィルタの一次微分を使って勾配方向を計算し、その方向を主方向とする。

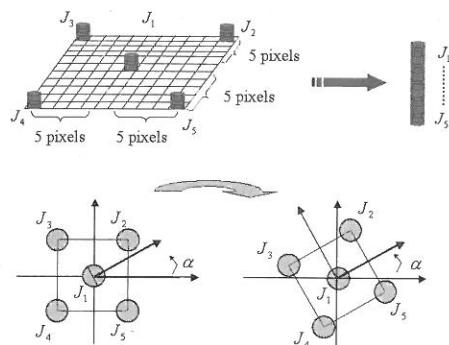


Fig. 1. Rotation invariant Gaussian derivative descriptor is realized by steering filter responses and neighboring locations.

ここで、 Gx, Gy は、一次微分ガウシアンフィルタで、 σ_{\max} はガウシアンの幅の最大値 ($\sigma_{\max} = 4$) である。この方法の利点は、フィルタの出力から直接、主方向を計算できることである。次に、この方向を用いて、以下の式にしたがって、n 次微分のフィルタ出力を回転し、同時に近傍の 4 画素点の位置とフィルタ出力を回転する。

$$G_n = G_n(\theta_i + \alpha), \theta_1 = 0, \theta_2 = \frac{pi}{4}, \theta_3 = \frac{pi}{2}, \theta_4 = \frac{3pi}{4} \dots \dots \dots \quad (3)$$

今後、この記述子を RIGD(Rotation Invariant Gaussian Descriptor) と呼ぶ。

3. 弱い幾何拘束による局所特徴量の Boosting

〈3・1〉 システムの概要

〈3・1〉 システムの概要 提案システムの概要を図 2 に示す。

処理の流れは、(1)認識対象物体から局所特微量を大量に集めて特微量プールを構築する、(2)プール特微量を用いて学習サンプルに対する識別用特微量を計算する(識別用特微量にはプール特徴との近さをはかる“相関特微量”を用いており、詳細は後で述べる。)(3)相関特微量を用いて教師あり統計学習を行い、認識に有効な特微量だけを選択しながら識別器も学習する、というように大きく3つに分かれている。

〈3・2〉 特徴量プールの構築

〈3・2〉 特微量プールの構築　特微量プールの作り方を説明する。このプールは物体を記述するための“辞書”である。

対象物体画像(positive samples)から、局所特微量 p_1, p_2, \dots, p_N を抽出し、特微量プールを構築する $P = \{p_1, p_2, \dots, p_N\}$ 。ただし、 $p_i(\vec{v}_i; x, y, s)$ は、局所特微量ベクトル \vec{v}_i を、画素点 (x, y) においてスケール s で求めたものである。通常、対象物体画像から抽出される局所特徴の数 N はとても多い。

“bag-of-features”⁽¹⁾ のアプローチでは、位置の情報を使わないが、我々の方法では、局所特徴が抽出された画素点を保持し認識に用いる。この時、特徴量は、オーバーラップする領域から抽出し、冗長な表現になっている。特徴量を

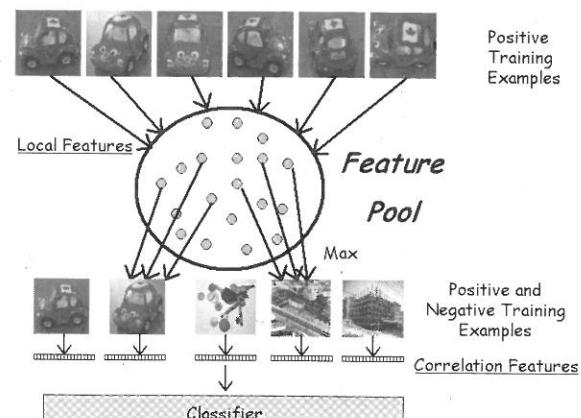


Fig. 2. System overview

抽出する特徴点としては、例えば、Harris コーナー⁽¹⁸⁾などが使えて、その点での局所特徴量を計算する。コーナー一点は、形に関して高い情報を持つと一般的に言われている⁽¹⁹⁾。Leibe ら⁽¹⁶⁾は、特徴プールを更にクラスタリング手法を用いて codebook と呼ばれる小数の特徴量で代表させているが、我々のシステムでは、特に特定物体の表現には抽出された特徴量をそのまま用いて、次節で示すように最終的に少数の特徴量を選択する。

〈3・3〉 相関特徴と幾何学的拘束 特徴量プールを構築後、全ての positive と negative サンプルを用いて、幾何学的な拘束をかけながら相関特徴を計算し、識別器が用いる特徴量とする。その手順を以下に示す。

(i) あるサンプル画像から特徴量 $S = \{s_1, s_2, \dots, s_L\}$ を計算する。

(ii) 特徴量プール P 内の各特徴量 $p_i(x_i, y_i) i \in [1, 2, \dots, N]$ について、以下の条件に合うようなサブセット特徴 $S_{neighbor} = \{s_j(x_j, y_j)\} \subset S$ を求める。

$$|x_i - x_j| < r_x \text{ and } |y_i - y_j| < r_y \quad \dots \quad (4)$$

ここで、 x と y は特徴量が抽出された画素点である。 r_x, r_y は、認識タスクにより設定するパラメータで、幾何学的拘束の条件となり、大きいほどよりデフォーマブルな物体を表現でき、小さいほど厳しい条件となる。

(iii) $S_{neighbor}$ の中で、 $p_i(x_i, y_i)$ と相関が最大になるものを探し、相関特徴 C の要素 $c(i)$ とする。 C の次元数は N である（プール P 内の特徴量の数）。

$$\begin{aligned} C(i) &= \max_{s_j \in S_{neighbor}} (\text{similarity}(p_i, s_j)) \\ &= \max_{s_j \in S_{neighbor}} \frac{(p_i - \bar{p}_i)(s_j - \bar{s}_j)}{\sqrt{(p_i - \bar{p}_i)^2(s_j - \bar{s}_j)^2}} \quad (\text{norm corr}) \\ &\dots \quad (5) \end{aligned}$$

この相関特徴が識別器への入力である。

繰り返しになるが、 $S_{neighbor}$ での最大値を計算することは、弱い幾何学的拘束をかけることに相当する。

〈3・4〉 特徴量選択 我々は、認識するための少数の特徴量 $Q = \{q_1, q_2, \dots, q_M\}$ を特徴量プール P の中から見つけたい。この少数特徴量セット Q を用いて、認識システムの最終的な出力スコア F は、additive models の形で、以下のように書ける。

$$F = \sum_{j=1}^M f(q_j) \quad (M \ll N) \quad \dots \quad (6)$$

f は、特徴量 q_j が物体を認識するための弱い仮説を立てる。 M は、 Q の中の特徴量の数を表し、 P 内の特徴量の数に比べて少ない。我々は、この問題を解く識別器として、GentleAdaBoost⁽²⁾⁽⁵⁾（アルゴリズム 1 参照）を用いている。GentleAdaBoost は、(1)特徴量選択が可能、(2)DiscreteAdaBoost より少数サンプルからの学習で高性能、(3)ランタイムでは選択された特徴量だけを計算すればよいので SVM など他手法に比べて高速、という利点がある。

Boosting の学習では、毎ラウンド、弱識別器である

1. Initialize weights $w_i = 1/N, i = 1, 2, \dots, N$

where N is the number of samples.

2. Repeat for $m = 1, 2, \dots, M$

(a) Fit the regression stump $f_m(x) = a(x_j > th) + b$

(b) Update the function by $F(x) \leftarrow F(x) + f_m(x)$

(c) Update weights by $w_i \leftarrow w_i e^{-y_i f_m(x)}$

3. Final strong classifier output is

$$\text{sign}[F(x)] = \text{sign}\left[\sum_{m=1}^M f_m(x)\right]$$

Algorithm 1 Gentle AdaBoost (Friedman, et al. [2])

regression stump ($f(x) = a(x > th) + b$) の学習をしながら、データ重みつき誤差が最小になる相関特徴の特徴量次元を一つ選択する。各相関特徴量次元が一つの局所特徴量を一次元に縮退したものであることを考えると、Boosting は特徴量プールの中から識別に有効な特徴量を弱識別器として選択していることに相当する。また、それと同時に、その特徴量の閾値のチューニングとどれだけ確からしいかという信頼度も学習していることになる。認識時には、学習で選択された特徴量のみを使って相関特徴を計算し、スコアを計算するため、高速な計算が可能である。幾何拘束の条件を厳しくしていくと、マッチングする局所特徴量の数が減少することで、更に高速な処理が実現できる。

Bag-of-features⁽¹⁾では、まずローカル特徴量をクラスタリングすることでいくつかのプロトタイプに分類し、それらプロトタイプが画像中にどのくらい存在するかをヒストグラムという形で表現している。そのため、各特徴量がどの程度ディスクリミネイティブかは気にしないような表現になっている。我々のシステムでは、ある物体を記述する局所特徴量は、認識対象物体とその他を分けるためのディスクリミネイティブな特徴量であるべきだとの考え方から、局所特徴量の存在の確からしさを学習することができる。また、bag-of-features では局所特徴の位置情報は持たないで“何個存在するか”という表現であるのに対して、我々のシステムでは、局所特徴量の位置も認識に使う。

4. 物体認識

〈4・1〉 単一識別モデルによる多視点物体認識 3 次元物体を任意視点から認識するシステムを学習する場合、2 次元上に投影された様々な画像の中から基準画像を決めて、ある決められた角度毎にモデルを学習して認識器を構成するというアプローチがよく取られる⁽²⁵⁾。我々のシステムでは、視点変化に強い特徴や、ある視点固有の特徴量などを統計的に学習する枠組みになっている。ここでは、Bag-of-features 法との比較も行い、任意視点物体の認識に対する相関特徴の有効性と、CalTech 物体認識データベースにおける幾何拘束の有効性を示す。

〈4・2〉 ミニカーの認識 図 3、図 6 のようにおもちゃのミニカーを用いて、様々な角度から撮影を行い、100×100 ピクセルに正規化された合計 74 枚の画像から Harris コーナー



Fig. 3. Some examples from the 74 training images of toy car object. Images are taken from all the viewpoints and used for training the multiview recognition system.

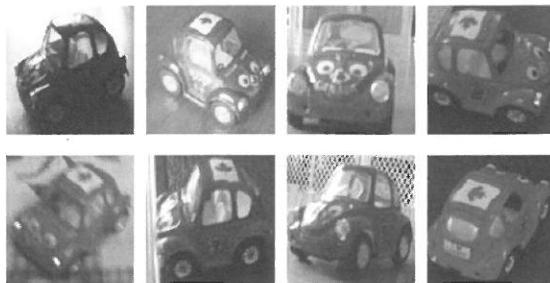


Fig. 4. Some examples from the 100 test images used in the experiment. Note that images are taken under different viewpoints, different light conditions, different backgrounds.

点を抽出し、その点における RIGD (Rotation Invariant Gaussian Descriptor)を計算した。合計 7089 個の特徴量がプールに抽出された。positive, negative 両方のサンプルに対して 7089 次元の相関特徴量を計算し、GentleAdaBoost で学習した。Negative の学習サンプルは、自然画像 (CalTech101⁽²⁸⁾ “BACKGROUND_Google”) からランダムに切り抜いた 4000 枚を用いた。テストは、全く異なる照明環境で様々な角度から撮影された画像を 100 枚(図 4 に画像の例を示す), negative サンプルは、学習時とは異なる自然画像 (CalTech256⁽²⁹⁾ “257.clutter”) から切り抜かれた 9000 枚の画像を用いた。図を見ると分かるように、物体の大きさや回転が変化したり、背景が複雑になっているため、難しい認識タスクになっている。比較のために、SIFT 特徴⁽¹⁴⁾+Boosting, 5 周波数 8 方向 Gabor 特徴⁽²⁸⁾+ Boosting, RIGD + SVM の組み合わせでも同じ実験を行った。結果の ROC 性能曲線を図 5 に示す。図中，“no geometry” とあるのは、 $r_x = r_y = \infty$ (ピクセル) とした場合で、幾何学的拘束なし，“weak geometry” とあるのは、 $r_x = r_y = 40$ (ピクセル) とした場合である。“AUC” は Area Under the Curve (ROC 曲線の面積) で、大きいほど性能が良いことを示している。また、“kmeans” とあるのは、初期特徴量プールを Bag-of-features などでやられるように kmeans クラスタリングにより 200 プロトタイプにし、それを新たな特徴量プールとして使ったことを示す。図 5 から以下のことがわかる。

(1) RIGD 特徴は従来提案されている SIFT や Gabor などの局所特徴量に比べて高い性能を示している。

(2) GentleBoost と SVM は同等の性能を示している。Boosting では、もともとの次元 7089 次元のうち、選ばれた 512 特徴だけを認識時に用いている。SVM の場合は、全ての次元 (7089 次元) を使うため、Boosting がいかに高速に

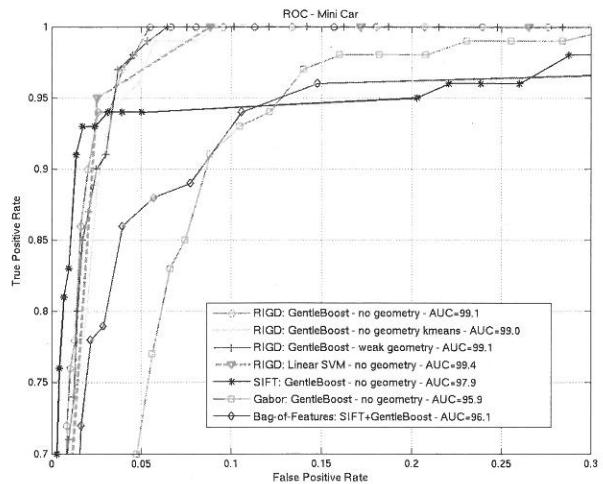


Fig. 5. ROC of the toy car object. RIGD descriptor outperforms other existing descriptors.

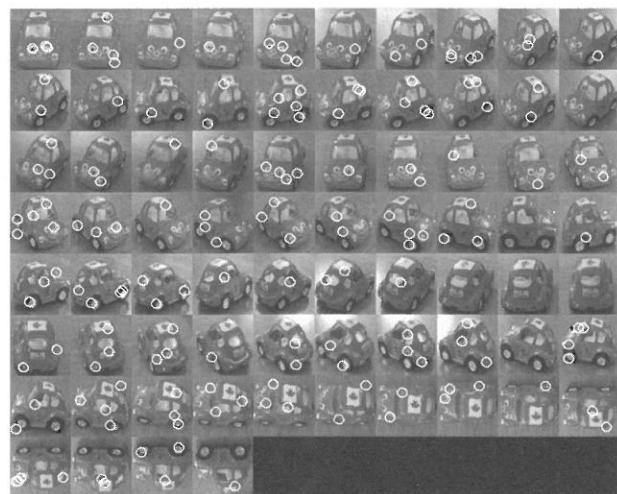


Fig. 6. First 200 selected features overlaid on all the 74 training images. Boosting procedure selects local features from initial 7089 features in the pool. Radius of the white circles indicates support region of the local descriptor.

計算できるかがわかる。

(3) この認識タスクでは3次元の回転により位置の一貫性が保てないため、幾何学的拘束はかけなくとも性能は変わらない。

(4) Bag-of-features⁽¹⁾法に比べて大幅に性能が高い。同じ統計学習器 (GentleBoost) を用いているので、ヒストグラム特徴と相関特徴の差異が出ていると考えられる。

(5) 特徴量プールをクラスタリングすることで性能は大きく変化しないが、若干下がる。

図 6 に、Boosting により選ばれた特徴量の例を示した。

4.3 CalTech101 カテゴリ認識 前節では任意視点物体認識の例をあげたが、ここでは単一視点物体の認識実験を報告する。

実験は、一般に公開されている CalTech101 物体カテゴリデータベース⁽²⁸⁾から airplanes カテゴリを用いて行った。

データベースの最初の 40 枚を学習に使い (図 7 参照), 残り 760 枚でテストを行った。学習とテストに使う negative サンプルは, それぞれ, CalTech101 の “BACKGROUND_Google”, CalTech256 の “257.clutter” からランダムに切り抜いた画像 1000 枚ずつを用いた。実験では, 幾何学的拘束をかけない設定 (グラフ中 “no geometry”), 幾何学的拘束をかけた設定 (グラフ中 “weak geometry” $rx=ry=40, 20, 10$) で行った。また, 比較のために, SIFT+GentleBoost の Bag-of-features 法でも全く同じ実験設定を行った。結果の ROC 曲線を図 8 に示す。グラフから, 幾何学的拘束をかけない方法ですでに Bag-of-features を上回り, 幾何学的拘束を強くしていくと性能が上がっていくのがわかる。これは, この認識タスクでは物体の位置が画像中である程度正規化されているためであると考えられる。Boosting によって選択された特徴量を使って作成したコレージュ画像を図 9 に示す。コレージュは, 選択特徴量が抽出された画像名から, 特徴量の座標周辺の輝度画像を切り抜き配置することで作成した。輝度画像の大きさは, RIGID



Fig. 7. Sample images of “airplanes” categories from CalTech 101 Object Categories database. We use the first 40 images for training and the remaining images for testing.

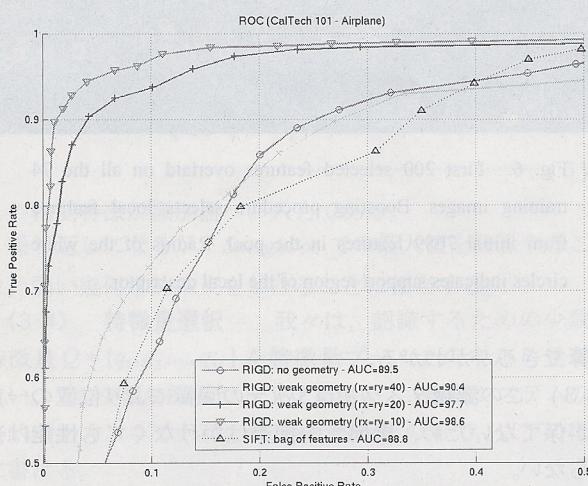


Fig. 8. ROC curve of “airplanes” category. Performance significantly improves with weak geometric constraints.

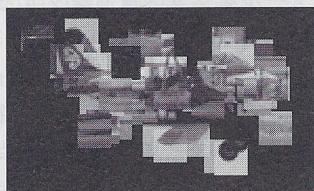


Fig. 9. Collage image of “airplanes” category. Collage is created by overlaying selected local features.

特徴量の支持領域と一致するようにした。ウイングや車輪などの特徴的なパーツが様々な飛行機の一部から選択され, 背景の局所領域が選択されていないのがわかる。

5. 任意姿勢顔識別

〈5・1〉 マルチスケール局所特徴による任意視点顔識別
我々はすでに⁽¹⁰⁾において, 幾何学的な拘束なしで報告を行ったが, 本実験では拘束をかけることでより高い性能を確認した。局所特徴量もスケールスペースの中で計算を行い, 画像もフィルタもマルチスケールの表現になっている。顔画像は 80×80 , 56×56 , 40×40 ピクセルの 3 スケールが表現されている。図 10 に(4),(5)式で記述された “max” 演算を示している。実験では $r_x = 40, r_y = 20$ とした。データベースは 9 人の被験者の各約 250 枚で構成され, 様々な角度から, 2 日かけて収集を行った。オンライン顔検出器⁽²⁶⁾が切り出した輝度画像を 80×80 ピクセルにリサイズしている。図 11 に示すように, 顔画像のアラインは行わず, 顔検出の枠をそのまま保存する。我々は, 1 日目に撮影された顔画像 50 枚をランダムに用いて特徴量プールを 9 つ構築し, 9 つの one-vs-all Gentle AdaBoost 識別器 (200 regression stump 弱識別器) を学習した。2 日目の撮影では, これらの識別器の結果を画面中に表示したこと, 被験者は出来るだけ認

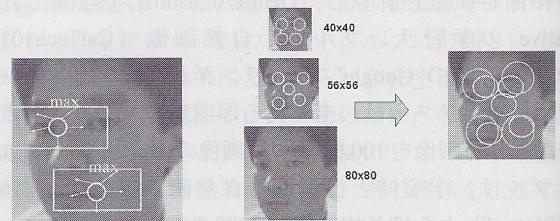


Fig. 10. Multiscale local feature representation on the face captures the various frequencies of the facial parts.



Fig. 11. Multiview face DB contains 9 subjects with approximately 250 images for each.



Fig. 12. Colored circles indicate the selected facial features during the boosting procedure. Radius of the circles represents the support regions of the local descriptor.

Table 1. Identification performance on our database.

Methods	#correct / #total	Accuracy
No Geometric Constraints [10]	1196 / 2269	52.71 %
Weak Geometry (single scale)	1457 / 2269	64.21 %
Weak Geometry (multi scales)	1543 / 2269	68.00 %
Laplacian patches with geometric constraints (multi scales)	1345 / 2269	59.28 %

識されないように表情を変化させたり、顔の姿勢を変化させたりしたため、認識困難なデータセットになっている。識別は、9つの識別器の最大スコアを出すものを採用した。表1に結果を示す。最も良い結果は、弱い幾何学的拘束を用いたマルチスケール局所特徴量を用いた時である。図12にいくつかの選択された特徴量の場所を示す。

6 まとめ

本論文では、ガウシアン微分方向フィルタを Steerable Filters を用いて効率的に計算することで局所特微量を算出し、それらを統計学習することで、リアルワールドで動作する多視点の物体認識システムを提案した。それぞれの局所特微量の存在の確からしさを学習するために、相関特徴の regression stump による Boosting 学習を行った。これにより、ディスクリミナントな特微量の選択と学習を同時に行えるようになった。システムの有効性を確かめるために、3 次元物体、そして任意視点顔識別の実験を行い、良好な結果を得た。特に Bag-of-features 法との比較では、任意視点認識の場合の相関特徴の有効性と、CalTech カテゴリ認識データベースの場合の幾何拘束の有効性を示した。また、RIGD 特微量の有効性も、SIFT や Gabor などの従来局所特微量と比較して示された。選択された特微量だけを用いて認識をすればよいので、ランタイムでは高速に処理することが可能である。少数サンプル (CalTech の実験では 40 枚) から学習できることも特色である。

(平成 20 年 9 月 12 日受付、平成 20 年 11 月 25 日再受付)

文 献

- (1) G. Csurka, C. Bray, C. Dance, and L. Fan : "Visual categorization with bags of keypoints", ECCV (2004)
 - (2) J. Friedman, T. Hastie, and R. Tibshirani : "Additive Logistic Regression: a Statistical View of Boosting", Tech Report, Dept. of Statistics, Stanford Univ. (1998)
 - (3) P. Viola and M. Jones : "Robust Real-time Object Detection", International Journal of Computer Vision (2001)
 - (4) W. Freeman and E. Adelson : "The design and use of steerable filters", PAMI, Vol.13, No.9, pp.891-906 (1991)
 - (5) A. Torralba, K. Murphy, and W. Freeman : "Sharing features: efficient boosting procedures for multiclass object detection", CVPR (2004)
 - (6) J. J. Yokono and T. Poggio : "Oriented filters for Object Recognition: an empirical study", FG (2004)
 - (7) J. J. Yokono and T. Poggio : "Boosting a Biologically Inspired Local Descriptor for Geometry-free Face and Full Multi-view 3D Object Recognition", *AI Memo*, MIT (2005)
 - (8) J. J. Yokono and T. Poggio : "Rotation Invariant Object Recognition from One Training Example", *AI Memo*, MIT (2004)
 - (9) J. J. Yokono and T. Poggio : "Evaluation of Sets of Oriented and Non-oriented Receptive Fields as Local Descriptors", *AI Memo*, MIT (2004)
 - (10) J. J. Yokono and T. Poggio : "A Multiview Face Identification Model With No Geometric Constraints", FG (2006)
 - (11) N. Dalal and B. Triggs : "Histograms of oriented gradients for human detection", CVPR (2005)
 - (12) K. Grauman and T. Darrell : "The pyramid match kernel: Discriminative classification with sets of image features", ICCV (2005)
 - (13) R. Fergus, P. Perona, and A. Zisserman : "Object class recognition by unsupervised scale invariant learning", CVPR (2003)
 - (14) D. Lowe : Distinctive image features from scale-invariant keypoints.

- International Journal of Computer Vision (2004)

 - (15) J. Shotton, A. Blake, and R. Cipolla : "Multi-Scale Categorical Object Recognition Using Contour Fragments", PAMI, 1270-1281 (2008)
 - (16) B. Leibe and B. Schiele : "Interleaved object categorization and segmentation", BMVC (2003)
 - (17) T. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio : "Object Recognition with Cortex-like Mechanisms", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.29, No.3, pp.411-426 (2007)
 - (18) C. Harris and M. Stephens : "A combined corner and edge detector", Alvey Vision Conference, pp.147-151 (1988)
 - (19) C. Schmid, R. Mohr, and C. Bauckhage : "Evaluation of interest point detectors", IJCV, Vol.37, No.2, pp.151-172 (2000)
 - (20) C. Wallraven, B. Caputo, and A. Graf : "Recognition with Local features: the kernel recipe", ICCV (2003)
 - (21) S. Ullman, M. Vidal-Naquet, and SE. ali : "Visual features of intermediate complexity and their use in classification", Nature Neuroscience, Vol.5, No.7, pp.1-6 (2002)
 - (22) D. Crandall and D. Huttenlocher : "Weakly Supervised Learning of Part-Based Spatial Models for Visual Object Recognition", ECCV (2006)
 - (23) P. Yang, S. Shan, W. Gao, S. Li, and D. Zhang : "Face Recognition Using Ada-Boosted Gabor Features", IEEE FG2004, pp.356-361
 - (24) C. Papageorgiou and T. Poggio : "A Trainable System for Object Detection", International Journal of Computer Vision, Vol.38, No.1, pp.15-33 (2000)
 - (25) A. Pentland, B. Moghaddam, and T. Starner : "View-based and modular eigenspaces for face recognition", CVPR (1994)
 - (26) 佐部浩太郎・日台健一：「ピクセル差分特徴を用いた実時間任意姿勢顔検出の学習」，大10回画像センシングシンポジウム (SSII 2004)
 - (27) M. Lades, J. Vorbruggen, J. Buhmann, J. Lange, and C. Malsburg : "Distortion Invariant Object Recognition in the Dynamic Link Architecture", *IEEE Transactions on Computers*, Vol.42, pp.300-311 (1993)
 - (28) http://www.vision.caltech.edu/Image_Datasets/Caltech101/
 - (29) http://www.vision.caltech.edu/Image_Datasets/Caltech256/

横野 順 (非会員) 1974年6月20日生。1999年3月早稲田大学応用物理学専攻卒業。同年ソニー(株)入社。2003-2004 マサチューセッツ工科大学CBCL(Center for Biological and Computational Learning)客員研究員。現在、ソニー(株)システム技術研究所勤務。物体認識や顔認識などの画像認識の研究に従事。



Tomaso Poggio



Tomaso Poggio (非会員) Tomaso A. Poggio, is the Eugene McDermott Professor at the Department of Brain and Cognitive Sciences; Co-Director, Center for Biological and Computational Learning; Member for the last 27 years of the Computer Science and Artificial Intelligence Laboratory at MIT; since 2000, member of the faculty of the McGovern Institute for Brain Research. Born in Genoa, Italy in 1947 (and naturalized in 1994), he received his Doctor in Theoretical Physics from the University of Genoa in 1971 and was a Wissenschaftlicher Assistant, Max Planck Institut für Biologische Kybernetik, Tübingen, Germany from 1972 until 1981 when he became Associate Professor at MIT. He is an honorary member of the Neuroscience Research Program, a member of the American Academy of Arts and Sciences and a Founding Fellow of AAAI. He received several awards such as the Otto-Hahn-Medaille Award of the Max-Planck-Society, the Max Planck Research Award (with M. Fahle), from the Alexander von Humboldt Foundation, the MIT 50K Entrepreneurship Competition Award, the Laurea Honoris Causa from the University of Pavia in 2000 (Volta Bicentennial) and the 2003 Gabor Award.