

パターンハッシング：部分画像と不変量索引を用いた分散アピアランスモデル

山口 修[†] 福井 和 広[†]

本論文では、幾何学的特徴ベースの手法にアピアランス空間の考え方を導入した物体認識の枠組みを提案する。モデル画像を特徴点を基準にして、複数の部分画像に分割する。次に大きさと向きを正規化した各部分画像を、基準とした特徴点に付随する幾何学的不変量をインデックスとしてハッシュテーブルに登録する。認識時には、テーブルに登録された膨大な部分画像モデルをハッシングにより高速に選択し、抽出された部分画像とのマッチングを効率的に行う。そして、この部分的なマッチング結果を投票手続きにより集積し、入力画像から対象の位置を検出する。検証実験により、隠れのある複数物体を安定に検出できることを確認した。

Pattern Hashing: Distributed Appearance Model Using Local Images and Invariant Indexing

OSAMU YAMAGUCHI[†] and KAZUHIRO FUKUI[†]

In this paper, we propose “Pattern Hashing” as a new scheme for object recognition by effectively introducing appearance-based approach in the framework of geometric feature-based approach. We divide the model object image into multiple local images based on the position of feature points. Those local images are normalized in terms of their size and orientation, and then registered to a hash table by which we define a hash function using geometric invariants. In recognition stage, fast model selection is efficiently executed by hashing technique. The result of local appearance pattern matching is accumulated through voting procedure by which we extract the target object in the input image. We demonstrate that partly occluded object region or multiple object positions can indeed be detected by the proposed algorithm.

1. はじめに

物体認識は、コンピュータビジョンの中心的な問題の一つである。多くの物体認識方法は、幾何学的特徴点ベースの研究¹⁾とアピアランス空間の研究^{11),12)}の2つのタイプに分類できる。本研究の目的は、幾何学的特徴点ベースの手法に、アピアランス空間のアプローチを効果的に統合した新しいモデル表現法を構築し、そのモデルを用いて部分隠れやノイズに対してロバストな認識方法を実現することである。

2つのアプローチを結合した方法として、Schmidら²⁾は、Harrisらのコーナー検出法⁴⁾によって求めた着目点 (interest points) において、微分ガウシアンフィルタに基づいた濃淡値に不変な局所的な記述を作成し、マッチングに利用した。Matasら³⁾は、不変点集合記述法 (Invariant Point-Set Signature) を提

案し、複数個の着目点から構成される凸包領域内で、着目点の間の線分上の濃淡情報を用いてモデルを記述した。Lowe⁵⁾は、着目点におけるスケールと位置に不変な特性を持つ SIFT (Scale Invariant Feature Transform) 特徴を提案し、部分的な隠れを有する画像の認識を行った。大場ら⁶⁾は、着目点の周辺画像 (ウインドウ) を局所固有空間として表現し、1つの物体を部分的な要素の集合として投票手続きを用いて認識を行った。着目点に対し各種の信頼性評価を行い、計算量の削減と誤認識の軽減を目的とし、有効なウインドウを選択する手法を示した。いずれの方法も、局所的な情報の記述法が中心的な課題であった。

図1は提案手法の枠組みの概念を示している。人間は複数の断片的な画像だけから顔の位置を推測でき、たとえ1つの部分パターンが欠けたとしても、複数の見え方の可能性によって正しい顔位置を求めることができる。図1といくつかの従来例は、モデル表現のための以下のヒントを与える。

[†] 株式会社東芝研究開発センター
TOSHIBA Corporate Research and Development Center

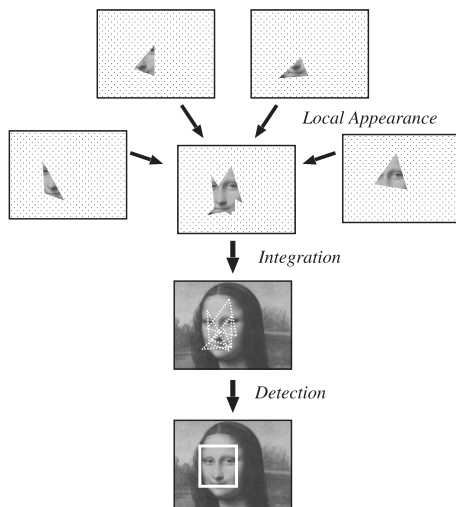


図 1 提案法のモチーフ：人間は部分画像を複数用いることで、正しく顔の位置を求められる

Fig. 1 Motif of proposed method: Human can detect the correct face position from some couple of occluded images.

- 局所的な見え方の表現

ノイズや隠れを持つ画像から目的の物体の正しい位置を見つけるためには、物体に対する複数の部分画像によるモデルの表現法が必要となる。

- 部分情報と全体情報の関連づけ

各部分画像パターンは、物体全体領域のどの部分であるかという関係を、効率良く関連づけて記憶する必要がある。

本論文では、複数の部分画像マッチングを用いた物体認識の新しい方法“パターンハッシング”を提案する。物体のモデル画像は、複数の部分画像パターンに分割してハッシュテーブルに登録する。それぞれの部分画像パターンは、特徴点の組合せによって構成される基底ベクトルに基づいて画像から切り出される。そして、基底ベクトルの成す幾何学的関係から不変量を計算し、その不変量によって定められた位置に部分画像モデルを格納する。この不変量は、類似した部分画像モデルをハッシュテーブルから検索するための索引として効果的に使われる。認識時には、テーブルに登録された膨大な部分画像モデルをハッシングにより高速に選択し、入力画像から抽出された部分画像とのマッチングを行う。部分画像どうしが類似している場合、対象とする物体が入力画像中に存在すると仮定することができる。モデルの登録時には、部分画像と物体の存在領域との相対的な位置関係を記述しておき、部分画像のマッチングの結果から対象物の位置に関する存在領域が求まる。これらのパターンマッチン

グを繰り返し、複数の局所的な見え方を集積する。物体の存在領域の仮説を統合するためには投票手続きを用い、その投票結果から大きな投票値を選択することで、物体の位置、方向を検出する。部分画像のパターンマッチングを行うためにモデルパターンを“切り刻み (hash)”，モデル検索のためにハッシングテクニックを利用することから、本手法をパターンハッシング (Pattern Hashing) と呼ぶ。

2. パターンハッシング (Pattern Hashing)

2.1 分散アピアランスモデル

提案法は、Geometric Hashing¹⁾ と関連する。Geometric Hashing では、モデル記述の基本要素は主として点であった。提案法では、基本要素として線や点といった幾何学的特徴の代わりに、新たに部分画像パターンを導入する。モデル画像を部分画像パターンに分割するには、(1) 分割に関する基準、(2) マッチングの際の画像パターンの正規化、を考慮する必要がある。どのように部分画像パターンを作成するかという分割に関する基準は、対象固有の知識に依存しないことが望ましい。よって、モデル画像に対して特徴点を求め、それらの組合せによって各部分領域を決定する。一般に画像マッチングでは、対象領域のサイズと方向の正規化が必要となる。そこで、正規化のための基底の選択方法も、対象モデル上の特徴点の組合せに基づいて決定すれば、上述の (1)、(2) を共通化して考えられる。

しかしながら、特徴点の組合せを用いた部分パターンへの分割は、特徴点の数に応じて大量のパターンが生成されることが明らかである。これはモデルの選択も複雑になり、さらなるマッチングのコストの増大を招く。たとえば、文献(6)では、すべての部分モデルと照合を行う必要があるため、モデルの数が増加して登録ウィンドウの数が増加すると大きな計算コストが必要となる。よって、マッチングを行うモデルの検索のために、効率の良い索引づけが必要となる。Geometric Hashing では、2 つの特徴点から基底を構成し、ハッシュテーブルを用いたモデルの検索、マッチングを行った。そこで提案法においても、ハッシュテーブルを用いることとし、モデルの部分画像パターンを効率良く選択する方法を与える。具体的には 3 つの特徴点から定義される 2 つの基底ベクトルから、相似変換に対する幾何学的不変量を計算し、これをハッシュ関数に適用する。これらの不変量は、部分画像パターンが入力画像の中でさまざまな位置、サイズ、方向で存在しても、不変な性質を持つため、類似したパターンを検索

するキーとして都合がよい。

以上の考察から、大量の複数の部分画像パターンをハッシュテーブル上に分散して配置し、不変量を用いることで高速な検索を行うことが可能なモデル表現である、分散アピアランスモデルが構成できる。

一方、このような認識を実現するためには、特徴点抽出法についても、上述した変換に対する不変性が必要となる。物体認識では、Harrisらの特徴点検出⁴⁾のような角点を抽出する方法が利用される。しかしながら、Harrisらの角点検出は、変換に対する不変性は十分ではなく、これを改良するための報告として文献⁷⁾がある。文献⁷⁾では、特徴点の“再現性”(repeatability)が議論されており、再現性の高い特徴点抽出を導入することの重要性が再認識されている。また、変換に不変な特徴としては、円形領域を検出し利用する方法が考えられ、顔認識を中心によく用いられてきた^{8),10)}。大きさの変化に対応したblob抽出法⁹⁾なども存在し、これらの特徴点抽出法を採用、またはさらに改良していくことで、変換に不変な認識が実現できる。

2.2 モデルの登録

図2は分散アピアランスモデルの登録方法を示している。説明のために顔画像(図2(a))を登録する場合を例にする。まず、モデル画像に対して特徴点検出を行う。対象物の全体領域を表す矩形 G_p (図2(b))と、その矩形中の特徴点が着目点として選択される。次に、着目点集合の中から任意の3点を選び、ある1点を原点として、2つのベクトルを設定し、これを基底ベクトルとする(図2(c))。この基底ベクトルは、部分画像パターンの設定と、索引づけのための幾何学的不変量を計算するために使われる。部分画像パターンは図2(d)のように、基底ベクトルに平行な四辺形の領域を設定し、基底ベクトルを正規直交ベクトルに変換することで正方形パターンに正規化する。

これらの正規化部分画像パターンは図2(c)に示すように、3点の組合せに応じた数だけ対象モデルから切り出す。このとき、モデルを登録する際にはパターンの冗長性を考え、2つの基底ベクトルに順序をつけ、第1ベクトルと第2ベクトルの長さの比が、1.0以上となる組合せのみを登録する。3点 p_1, p_2, p_3 の組合せを、{原点, 第1ベクトルの端点, 第2ベクトルの端点}と書くと、原点を同じとする基底ベクトルの組 $\{p_1, p_2, p_3\}, \{p_1, p_3, p_2\}$ については、図2(d)に示すように基底ベクトル周辺の画像パターンが同一となるために、ベクトルの順序を規定することで、一方のみを選択すれば十分となる。ただし3点が二等辺三角形を構成する場合には、同じ長さのベクトルが2つ

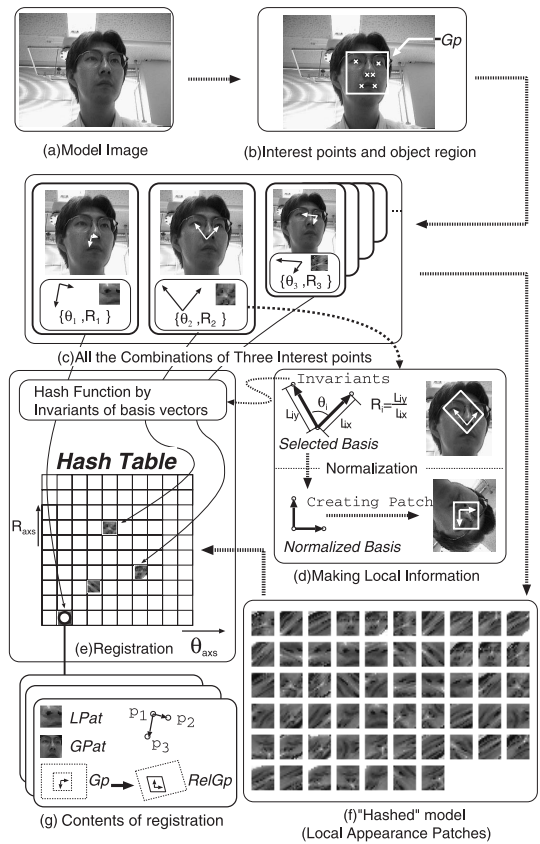


図2 分散アピアランスモデルの生成

Fig. 2 Creation of a Distributed Local Appearance Model.

現れるため、それぞれの順序を区別して登録する。また、2つの基底ベクトルのなす角度が0度、180度に近い角度の場合は、部分画像がテクスチャとしての意味を持たないため、パターンの登録を行わない。よって、通常3点の組合せでは、原点と他の2点の組合せによる6つのパターンが考えられるが、基底ベクトルの長さの組合せによって、実際には4つ以下のパターンが登録されることになる。

次に、正規化部分画像パターンは、基底ベクトルから計算される幾何学的不変量に基づいたハッシュ関数により、テーブルの所定の位置に格納される。図2(g)に示すように、テーブル上に登録されるデータは、正規化部分画像パターン $LPat$ に加え、部分画像が検出対象領域 G_p のどの部分に位置するかという相対的な位置関係 $RelGp$ 、また、対象物全体の正規化画像パターン $GPat$ がともに記憶される。これらの詳細については、後述する。図2(f)は、モデル画像に対して切り出された顔画像の正規化部分画像パターン集合を示す。これらがハッシュテーブル(図2(e))上に自動的に分散配置される。

2.3 ハッシュ関数とハッシングにおける衝突処理

ハッシュテーブルにおける格納位置を決定するためのハッシュ関数のパラメータには、基底ベクトルを構成するために選択された3点 p_1, p_2, p_3 を用いる。基底ベクトルは、 p_1 を原点、 $\vec{p}_2 - \vec{p}_1$ 、 $\vec{p}_3 - \vec{p}_1$ を2つのベクトルとする。このときハッシュ関数 HF は以下のように定義する。

$$HF(p_1, p_2, p_3) = \{\theta_{axs}, R_{axs}\} \quad (1)$$

$$\theta_{axs} = \cos^{-1} \frac{(\vec{p}_2 - \vec{p}_1, \vec{p}_3 - \vec{p}_1)}{\|\vec{p}_2 - \vec{p}_1\| \|\vec{p}_3 - \vec{p}_1\|}$$

$$R_{axs} = \frac{\|\vec{p}_3 - \vec{p}_1\|}{\|\vec{p}_2 - \vec{p}_1\|}$$

ここで、2つのベクトルがなす角度 θ_{axs} と2つのベクトルの長さの比 R_{axs} は、平行移動、回転、拡大縮小といった相似変換に対しての不変量となる。図2(e)に示したハッシュテーブルは2次元であり、一方が θ_{axs} 、他方が R_{axs} の軸を表している。 θ_{axs} と R_{axs} はともに量子化され、2つの不変量によって、索引づけが行われる。

複数の部分モデルパターンの登録を逐次行くと、3点の組合せによって同じハッシュ関数の値を持つことがあり、ハッシングにおける衝突 (collision) が発生する。ハッシュテーブルの各位置では、チェーン法により、複数のパターンをリスト構造として持つようにする。すなわち、基底ベクトルの成す幾何学的関係が類似した複数の部分モデルは同じ位置に記憶される。

2.4 部分情報と全体情報の関係記述

ハッシュテーブルに登録される正規化部分画像 $LPat$ は、対象物の局所的な部分情報であり、部分画像パターンのマッチングにおいて得られる情報は、対象物の部分的な位置情報でしかない。その部分的な位置情報から、対象物の全体の存在領域を計算するためには、部分パターンと検出対象領域との相対的な位置関係の記述が必要となる。

図3の例で部分と全体の関係づけを説明する。図3(b)は図2(g)と同じ登録情報を表し、モデル画像(図3(a))から得られた図3(b)の登録情報 M_j は次のように表す。

$$M_j = \{p_1, p_2, p_3, LPat, GPat, RelGp, Label\} \quad (2)$$

p_1, p_2, p_3 は基底ベクトルを構成する3点の座標位置、 $LPat$ は $M \times M$ pixelの正規化部分画像、 $GPat$ は検出対象領域 Gp の部分パターンを表す $M' \times M'$ pixelの正規化画像である。 $Label$ は物体の種類、カテゴリなどを表す付帯情報で、複数種類の物体を検出を行う際に、どの物体であるかというラベルを与える。 $RelGp$

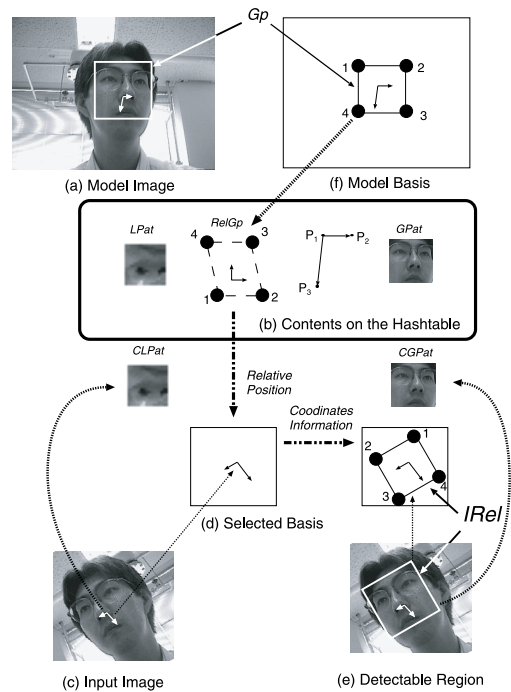


図3 部分-全体関係の記述

Fig. 3 Description of Part-Whole relation.

は、検出対象領域 Gp を表す四辺形の4点の座標について、図3(f)の座標系を正規直交系にした場合の相対座標で記述したものであり、対象物の全体の位置を求める際に次のように用いる。図3(c)の入力画像の3点の特徴点を選択し、基底ベクトルを構成したとする。その部分画像パターン $CLPat$ がモデル部分画像パターン $LPat$ とマッチした場合、入力画像上での部分パターンの位置が求まる。そこで、図3(d)の選択された基底ベクトルで張られる座標系を用いて $RelGp$ の表す4点の座標位置を変換する。これにより入力画像(図3(c))における各4点の座標が決定し、図3(e)に示すように、対象物の領域 Gp に対応する物体全体の領域位置 $IRel$ が求まる。この領域 $IRel$ を用いて、入力画像中の検出対象領域のパターン $CGPat$ を抽出でき、 $GPat$ と比較することで、全体的な領域の整合性をパターンマッチングによりチェックすることが可能となる。

3. パターンハッシングによる認識

3.1 特徴点検出とモデル選択

図4は、パターンハッシングによる認識の流れを示している。この図では、部分的に隠れている物体を画像から検出する例を示している。

認識の方法は、最初に入力画像(図4(1))に対し

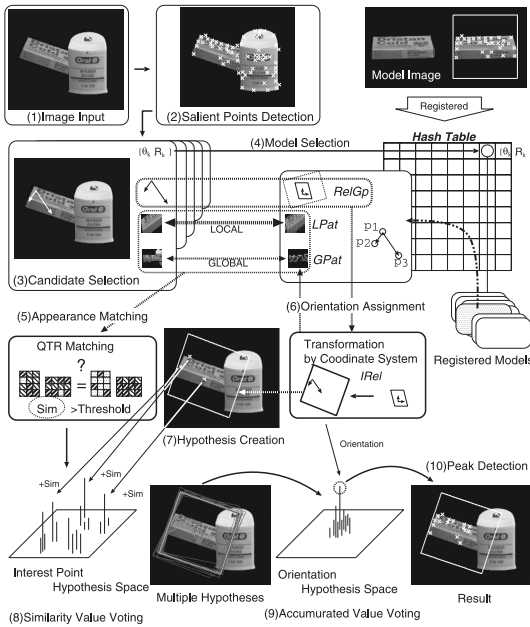


図 4 パターンハッシングによる認識処理の流れ
Fig. 4 The flow of recognition stage.

て特徴点検出を行う(図 4(2)). これらの特徴点は基底ベクトルを構成するために使われる. 一般に隠れやノイズが存在するので, これらの点は, 必ずしもモデル登録と与えられた着目点をすべて含んでいる必要はない.

モデル登録と同様に, 3 つの特徴点 p_{1k}, p_{2k}, p_{3k} の組合せが選択され(図 4(3)), 基底ベクトルのなす不変量が計算される. 不変量 θ_k, R_k により, マッチングの対象となるモデルをハッシュテーブルから選択し(図 4(4)), それぞれのモデルパターンとマッチングを行う.

3.2 マッチングのための類似性測度

候補として抽出した部分画像パターンは登録されたモデルの部分画像パターンと比較する. マッチングのための類似性測度には, 正規化相関や残差の 2 乗和(SSD)がよく使われるが, これらの測度は, 線形濃度変換, ヒストグラム平坦化などの濃度変換処理により, 類似度の値が影響を受けてしまう. 本研究では, マッチング測度として, 定性的 3 値表現(Qualitative Trinary Representation: 以下 QTR)を導入する¹³⁾. この表現では, 隣接画素値の大小関係, 同値関係を利用しており, 一般の類似性測度とは異なる性質の類似度を定義する.

QTR は, 画像 I_s に対して, 2 つの 3 値画像 Q_h, Q_v で定義する.

$$QTR(I_s) = \{Q_h, Q_v\} \quad (3)$$

なお, Q_h, Q_v は以下の式で定義される.

$$Q_h(x, y) = \text{sgn}(I_s(x, y) - I_s(x + 1, y)) \quad (4)$$

$$(0 \leq x \leq M - 2, 0 \leq y \leq M - 1)$$

$$Q_v(x, y) = \text{sgn}(I_s(x, y) - I_s(x, y + 1)) \quad (5)$$

$$(0 \leq x \leq M - 1, 0 \leq y \leq M - 2)$$

$I_s(x, y)$ は画像の (x, y) での輝度値, $M \times M$ pixel の画像が与えられた場合, 2 つの 3 値画像 Q_h, Q_v は, $M - 1 \times M$ pixel 画像と $M \times M - 1$ pixel 画像となる. なお, sgn 関数は,

$$\text{sgn}(x) = \begin{cases} 1 : x > 0 \\ 0 : x = 0 \\ -1 : x < 0 \end{cases} \quad (6)$$

である.

図 4(5) に 3 値画像の例を示す. 2 つの画像 G_1, G_2 の類似度 $QSim(G_1, G_2)$ は, 3 値画像のそれぞれの画素の一致する回数 D_h, D_v によって定義される.

$$QSim(G_1, G_2) = D_h + D_v \quad (7)$$

$$QTR(G_1) = \{Q_{h1}, Q_{v1}\} \quad (8)$$

$$QTR(G_2) = \{Q_{h2}, Q_{v2}\} \quad (9)$$

$$D_h = \text{Count}(Q_{h1}, Q_{h2}) \quad (10)$$

$$D_v = \text{Count}(Q_{v1}, Q_{v2}) \quad (11)$$

$$\text{Count}(Q_1, Q_2) =$$

$$\sum_k \sum_l \begin{cases} 1 : Q_1(k, l) = Q_2(k, l) \\ 0 : Q_1(k, l) \neq Q_2(k, l) \end{cases} \quad (12)$$

この類似度計算には, 乗算が必要ではなく, 加減算と比較演算のみで構成され, 高速な計算が可能である. また, 部分的な輝度の変化, 隠れに対してもロバストであり, 輝度の単調増加変換に対しては, 類似度は不変的な性質を持つ.

3.3 仮説の生成と投票による統合

登録モデル M_j が選択されると, 対象となる 3 点 p_{1k}, p_{2k}, p_{3k} と $RelGp$ から, 相対的な検出対象領域 $IRel$ が計算され(図 4(6)), 部分画像パターン, $CLPat_k, CGPat_k$ が求まる. この検索対象候補 C_k

$$C_k = \{p_{1k}, p_{2k}, p_{3k}, CLPat_k, CGPat_k, IRel\} \quad (13)$$

がマッチングの候補となる.

まず, 検索対象候補 C_k と登録モデル M_j の部分画像パターンどうしの類似度 L_{Sim}, G_{Sim} を求める(図 4(5)).

$$L_{Sim} = QSim(CLPat_k, LPat_j) \quad (14)$$

$$G_{Sim} = QSim(CGPat_k, GPat_j) \quad (15)$$

それぞれの類似度が設定した値を超えた場合に, その場所の部分的な見え方を有する物体が存在すること

を仮定でき、これを仮説情報として生成する．仮説情報 H_i は、以下のように記述する．

$$H_i = \{p_{1k}, p_{2k}, p_{3k}, IRel, L_{Sim}, G_{Sim}, Label\} \quad (16)$$

ここで、 $IRel$ は物体の存在仮説領域として利用され、入力画像上での物体の位置、サイズ、方向を表す．複数の仮説が画像マッチングにより生成され、部分的な隠れがある場合でも、さまざまな見え方の可能性が仮説情報として生成され、投票による仮説の集積によって対象の位置を検出できる．また、 $Label$ は、物体の種類、カテゴリを表し、式 (2) の M_j から引用する．

図 4 下は、仮説の投票と統合の様子を模式的に表している．図 4(8) に示す着目点仮説空間は、各特徴点が着目点としてどの程度支持されているかを求めるために用意される．この空間の各セルは物体上の着目点候補と対応し、各仮説情報の 3 点の位置に対応するセルに対して、類似度値 L_{Sim} を投票する (図 4(8))．ここで、ある特徴点が複数の領域の着目点としての可能性がある場合が存在する．たとえば、図 5 の例では、 \times で示した特徴点を実線、破線、双方の領域の特徴点となっている．このような場合を考慮して、同じ位置の特徴点であっても、異なる存在領域の場合には、区別して投票を集積しておく．

すべての仮説について投票を行った後、次に、着目点仮説空間において蓄積された類似度値を用いて、姿勢仮説空間への投票を行う．図 4(9) に示す姿勢仮説空間は、画像中のどの位置に物体が存在するかを表現する．仮説情報の存在仮説領域 $IRel$ と対応する姿勢仮説空間のセルに対し、先の着目点仮説空間に蓄積された値を用いて投票を行う．これは、ある領域仮説に関連するすべての特徴点の投票値が集積されることになる．

姿勢仮説空間に集積された投票結果を評価し、物体の存在位置を求める．1 つの対象領域の位置のみが必要な場合は、各セルの投票を数え、最も大きな投票を持つ位置を結果とする (図 4(10))．複数の候補を認識する場合は、姿勢仮説空間に投票されたピーク値の大きな順に評価を行い、部分的に重なっている領域の排他処理などの後処理を行う．たとえば、図 5 では、複数の実線で示された領域仮説が、同じセルに投票され、大きな投票値を持つこととなる．この場合、共有されていた点の投票値を、破線部分が示す領域仮説のセルから減じたのち、再評価を繰り返すこととなる．

3.4 計算量とメモリ量の考察

3 点の組合せによって生成される部分モデルのマッチングに関する計算量についてまとめる．登録されて

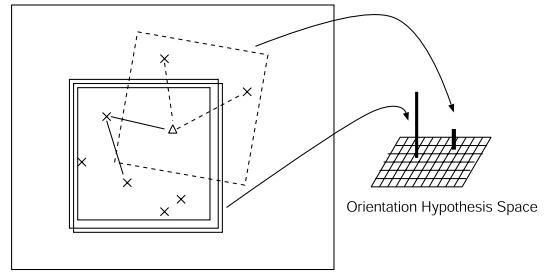


図 5 複数の領域仮説の例
Fig. 5 Example of multiple hypotheses.

いるモデルの数を λ 、各モデルについての着目点の数を m 、画像中の特徴点の数を n とした場合、すべての部分モデルどうしのマッチングを行う場合の回数は、 $O(\lambda m^3 n^3)$ となる．一方、本手法ではあらかじめモデルの登録がなされているため、認識時のマッチング回数は、ハッシュテーブルの 1 つの格納場所に登録されている部分モデルの平均数を p とすると、 $O(pn^3)$ となる．

次に、ハッシュテーブルのメモリ量について考察する．ハッシュテーブル内の格納場所の数は、各不変量の量子化サイズによって決定される．しかしながら、メモリ量は部分モデルの登録数によって大きく変化し、ハッシュテーブルにおける各部分モデルの数のオーダは、 $O(\lambda m^3)$ となる．これに登録情報 M_j で示した構造体を格納するだけのメモリ量を乗じたメモリが必要となる．

4. 実験

本章では、提案法を用いた以下の実験を行う．まず、単一種類の物体認識の具体例として、画像中から複数の顔領域を検出する動作検証実験を行う．次に、複数種類の物体を登録し、同時に認識可能であることの動作検証を行う．最後に、ロバスト性を評価するために顔画像データを用いた評価実験を行う．

4.1 顔領域検出

4.1.1 顔画像モデルの登録

アルゴリズムの検証のために、顔画像を利用した実験を行う．収集した顔データベースには 25 名分、950 枚の画像 (320×240 pixel) で、顔の向きをわずかに変えさせながら画像を収集した．この画像を用いてモデルの登録を行った．各画像に対し特徴点検出を行い、瞳、鼻孔、口角の 6 点を、着目点として手動で選択して与えた．同時に、顔領域を囲む対象領域 G_p を設定した．特徴点の検出法には、分離度フィルタによる円形領域検出¹⁰⁾ を利用した．円形領域は平行移動、回転変換に対しては不変であるが拡大縮小変換を考慮す

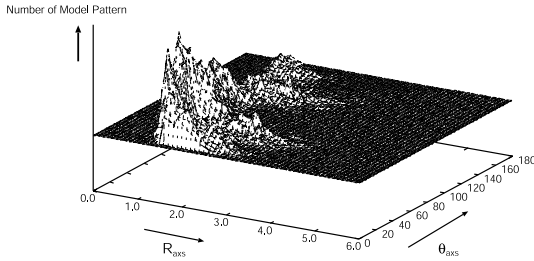


図 6 ハッシュテーブル上の部分モデルの登録数

Fig. 6 The number of registered partial model on the hash table.

る必要がある．そこで，分離度フィルタのマスク形状を数段階の半径にあわせて用意し，複数の半径の円形領域を抽出することで，拡大，縮小に対応できる．また，部分画像マッチングのためのテンプレートの大きさは， 16×16 pixel とした．

950 枚の顔画像データを登録した場合のハッシュテーブルについて，部分モデルの登録数をグラフにした．図 6 が登録後の状態を示し，各軸は，不変量 $\theta_{\alpha x s}$ と $R_{\alpha x s}$ を表している．この量子化の刻み幅は，角度 $\theta_{\alpha x s}$ 方向は 1 度刻み，長さ比 $R_{\alpha x s}$ 方向は 0.1 刻みとした．これらの刻み幅は検出精度にも関連し，今後検討が必要である．モデルの冗長性を考慮し，長さ比の関係や角度制限によって選択されたもののみが登録され，部分モデル M_j の数は，100,195 個となった．ここで，1 つ以上の登録がある格納場所に関する平均登録数は約 32 個であった．各部分モデルが不変量による索引によって，テーブル上に分散されている様子が分かる．

4.1.2 実験結果

図 7 は，複数人の正面顔を同時に検出している様子を示したものである．画像平面中での回転したいくつかの顔の領域と，それらの顔のいくつかの着目点を検出している．図 7 左上は，部分的な隠れとなる眼帯をしている人物を含む画像例を示す．図 7 左中の特徴点（円形領域）の候補数は 74 個であり，3 点の組合せの総数は 388,944 である．しかしながら，認識時のマッチングのコストを抑えるために，先のモデル登録の条件（長さ比，角度条件）による制限に加え，さらに別の条件を用いて削減すると効率が良い．ここで，図 6 から分かるようにモデルには登録されていないベクトルの長さ比の上限値（5.0 以上は登録されていない）を設けることや，ベクトルの長さに対する条件として画像を横断するような長さのベクトルを候補に含まないようにすることで候補数を削減できる．本実験では，パターンの切り出しが行われた基底ベクトル候補は，69,792 個であった．



図 7 顔認識の例．左上：入力画像，右上：計算された複数の仮説，左中：円領域検出の結果，右中：検出された顔領域と顔特徴点，下段：その他の例

Fig. 7 Examples of face recognition. Top left: Input image, Top right: Generated multiple hypothesis, Middle left: Result of circular region detection, Middle right: Detected facial parts points, Bottom left: For faces with different scales, Bottom right: For the upside-down image.

図 7 右上の四辺形は，それぞれのマッチングにおいて，類似性が高いと判断された位置候補の I_{Rel} を示し，線分はそれらの基底ベクトルを表している．複数の見え方を統合した結果，3 つの顔領域を正しく検出している．また，左下のように大きさの異なる対象の場合や，右下のように部分的な隠れを持つ逆さ画像の場合でも検出できている．

試作したプログラムは，Windows XP 上で，Visual C++6.0 で実装され，Pentium4 2.0 GHz を CPU とする PC で動作させた．図 7 左上の画像を処理するために，約 4 秒の時間を要した．

4.2 複数種類の物体検出

次に，複数種類のモデルを同時に検出する実験を行った．実験には，Columbia 大学物体画像データベース (COIL)⁴⁾ の画像を利用し，あらかじめ 10 個のオブジェクトを選び，ある一方向の見え方の画像のみを使用し，ハッシュテーブルに登録した．特徴点検出には Harris らの角点検出を用いている．図 8 は，部分的な隠れや回転した 5 つの物体の合成画像から，登録されている 3 つの物体を正しく検出することを示している．このとき，式 (16) に示した， $Label$ を用いて，物

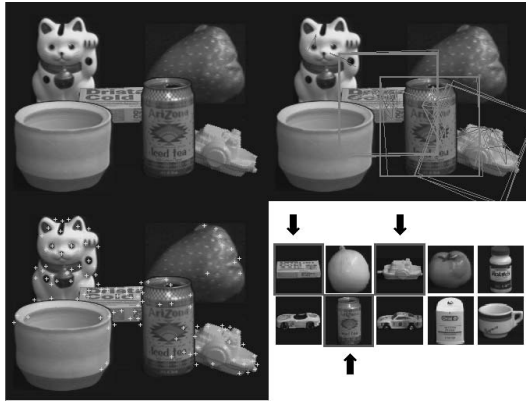


図 8 複数種類のモデルの検出．左上：入力画像，右上：計算された複数の仮説，左下：特徴点検出の結果，右下：登録されている 10 個の物体と見つかった物体（矢印）

Fig.8 Detection of multiple model. Top Left: Input image, Top Right: multiple hypotheses, Down Left: Result of feature points detection, Down Right: Detected objects and ten registered objects.

体の種類を特定することが可能となる．

4.3 ロバスト性の評価実験

4.3.1 実験データ

次に提案法の基本的な頑健性について調べるために，テスト画像に回転，拡大縮小，一部隠蔽を施し，検出評価実験を行う．実験のテストデータには，財団法人ソフトピアジャパンで開発された顔画像データベース を使用し，さまざまな年代，性別の人物データでほぼ同じ角度で撮影された画像セット（640×480 pixel）300 枚を用いた．画像データは 320×240 pixel に縮小し，カラー情報は用いないため，白黒濃淡画像に変換して利用した．各画像には目鼻 4 点の正解点位置を入力し，検出領域がその 4 点を含む場合に正解と判定する．ハッシュテーブルへの登録にはテストデータとは異なるデータで，4.1.1 項で用いた顔画像データをそのまま用いる．

4.3.2 部分的な隠蔽に対する評価

部分的な隠蔽に対する耐性を調べる方法は，さまざまなアプローチが考えられ，一般的な定量評価は難しい．本論文では隠蔽条件として，図 9 のように画像の左から 40%，45%，50%の位置までのノイズ画像を作成し，各テスト画像に掛け合わせた画像を作成し，評価を行った．撮影されている顔位置は各個人で異なるため，各画像ごとの顔領域の隠蔽状態がそれぞ

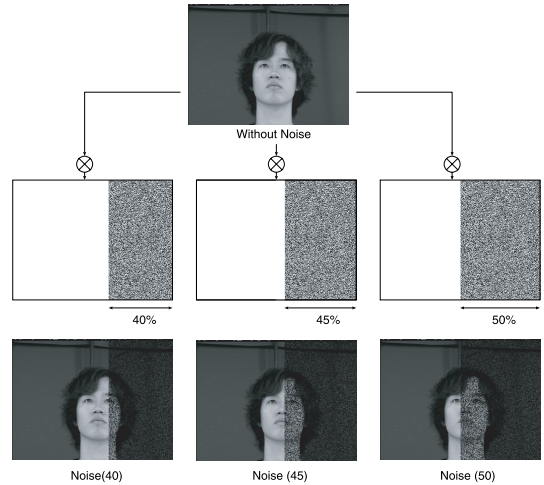


図 9 隠蔽ノイズの混入

Fig.9 Occluding noise.

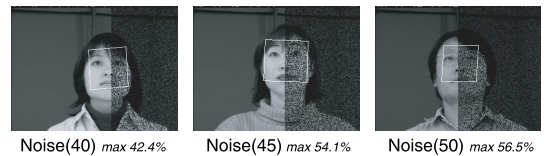


図 10 隠蔽ノイズ画像に対する検出結果：検出結果に対する最大隠蔽面積率を持つ画像を示す

Fig. 10 Detection result of occluding noise image.

表 1 部分隠蔽ノイズ画像に対する検出結果

Table 1 Results of detection for partial occlusion.

	隠蔽なし	Noise(40)	Noise(45)	Noise(50)
正解率	292/300 (97.3%)	285/300 (95.0%)	230/300 (76.6%)	77/300 (25.6%)
隠蔽面積				
平均	0.0%	11.2%	29.7%	42.5%
(最小)	—	(0.0%)	(0.0%)	(12.8%)
(最大)	—	(42.4%)	(52.7%)	(56.1%)
着目点				
検出点数 (平均)	6.00	5.69	4.38	3.36
着目点が 3 点未満の画像	0/300 (0.0%)	0/300 (0.0%)	11/300 (3.6%)	142/300 (47.3%)

れ異なる．

隠蔽ノイズ画像を用いた検出例は図 10 のようになり，その結果を表 1 に示す．表 1 には，それぞれの検出正解率と検出された領域に対する隠蔽ノイズの面積率，また，6 点の着目点のうち検出される特徴点の平均検出点数を示し，最後に，着目点に対応した特徴点が隠蔽され，3 点未満になる場合の画像データ数をまとめた．

Noise(40) の場合を隠蔽ノイズがない場合と比較す

本論文に使用した顔画像データは，財団法人ソフトピアジャパン 研究開発部地域結集型共同研究推進室から使用許諾を受けたものです．権利者に無断で複写，利用，配布などを行うことは禁じられています．

ると、検出正解率は大きく影響を受けていない。しかし、Noise(50)では、25%程度まで低下がみられる。遮蔽による影響として、(1) 遮蔽ノイズの部分がパターンマッチングに影響を及ぼし、低いマッチング類似度となる、(2) 着目点に対応する特徴点が検出できない、という2つの原因があげられる。検出された領域に対する遮蔽ノイズの面積を調べると、それぞれ平均11.2%、29.7%、42.5%となった。このとき、図10右のように、最大で56.1%の面積を占めたものでも検出に成功している例も存在する。

次に、特徴点の有無により、どの程度影響を受けるかを調べる。図10に示した顔をほぼ半分を覆う例では、着目点に対応する特徴点は3点だけが検出されている。着目点に対応する特徴点が3点未満の場合は認識が困難であり、Noise(50)の場合には、平均検出点数は3.36と3に近く、また47.3%の画像で3点未満となっていた。Noise(45)の場合においては、特徴点が3点未満となっている例は少ないものの、一部の特徴点が隠れた画像の約1/4が検出失敗となり、検出率の低下を招いている。この実験から隠蔽面積率が30%を超えると急激に認識率に影響がでることが推測される。

この結果は、モデルの着目点の配置、数、さらに隠蔽の仕方により、性能が変化する。従来の特徴点ベースの方法と比較すると、特徴点検出の性能に加えて、パターンマッチングに関連した評価も行う必要がある。具体的には、パターンの類似性と隠蔽率との関係などがあげられ、今後も評価方法について考察しなければならない。

4.3.3 回転、拡大縮小変換に対する評価

次に、回転、拡大縮小変換に対しての性能を調べるために、Noise(40)の画像に対して、 -90 度から $+90$ 度の回転変換、ならびに80%から120%の拡大縮小変換をランダムに施し、検出を行った結果の一部を図11に示す。図中の数値は元画像に対する(拡大率、回転角度)を表す。特徴点検出は、回転、拡大縮小変換に影響を受けておらず、安定した検出が行われている。また、表2に示した検出正解率も、ほぼ同程度であり、回転、拡大変換に対する本手法の有効性を示している。

しかしながら、今回の実験ではテスト画像を画像変換させたにすぎない。実際に3次元物体(顔)が回転した場合は、陰影の状態などが変化するため、特徴点抽出やパターンの類似性に影響がでてくる。今後、実環境での大量データ評価を行う必要がある。

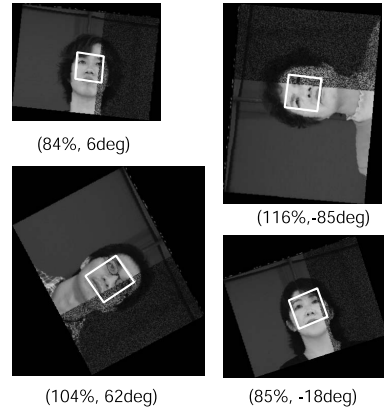


図11 回転、拡大縮小画像に対する検出結果画像

Fig. 11 Results of detection for rotation and scaling.

表2 回転、拡大縮小画像に対する検出結果

Table 2 Results of detection for rotation and scaling.

Noise(40)	変換なし	変換あり
正解率	285/300	283/300
	(95.0%)	(94.3%)

5. まとめ

本論文では、新しい物体認識のモデルの表現法と検出方法“パターンハッシング”を提案した。不変量インデクシングを導入したハッシュテーブルは分散アピアランスモデルの表現に適している。パターンハッシングによる認識の基礎実験を行い、任意の位置、サイズ、方向の複数の物体を安定に認識できることを示した。

今後は、3次元物体の複数の見え方の登録に関する定量的な精度評価と誤差評価、効果的なモデル探索と検証の戦略、テーブル上でのモデルの簡約化などの考察を扱っていく。

参考文献

- 1) Lamdan, Y. and Wolfson, H.J.: Geometric Hashing: A general and efficient model-based recognition scheme, *Proc. International Conference Computer Vision*, pp. 238–249 (1988).
- 2) Schmid, C. and Mohr, R.: Local Grayvalue Invariants for Image Retrieval, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.19, No.5, pp.530–535 (1997).
- 3) Matas, J., Buriánek, J. and Kittler, J.: Object Recognition using the Invariant Pixel-Set Signature, *Proc. British Machine Vision Conference*, Vol.2, pp.606–615 (2000).
- 4) Harris, C.J. and Stephens, M.: A combined corner and edge detector, *Proc. 4th Alvey Vision Conference*, Manchester, pp.147–151

- (1988).
- 5) Lowe, D.G.: Object recognition from local scale invariant features, *Proc. IEEE International Conference on Computer Vision*, pp.1150–1157 (1999).
 - 6) 大場光太郎, 池内克史: 局所固有空間手法による金属物体の安定認識, 電子情報通信学会論文誌 D-II, Vol.J80-D-II, No.12, pp.3147–3154 (1997).
 - 7) Mikolajczyk, K. and Schmid, C.: Indexing based on scale invariant interest points, *Proc. International Conference on Computer Vision*, pp.525–531 (2001).
 - 8) Yuille, A.L., Hallinan, P.W. and Cohen, D.S.: Feature extraction from faces using deformable templates, *International Journal of Computer Vision*, Vol.8, No.2, pp.99–111 (1992).
 - 9) Lindeberg, T.: Feature Detection with Automatic Scale Selection, *International Journal of Computer Vision*, Vol.30, No.2, pp.77–116 (1998).
 - 10) Fukui, K. and Yamaguchi, O.: Facial feature point extraction method based on combination of shape extraction and pattern matching, *Systems and Computers in Japan*, Vol.29, No.6, pp.49–58 (1998).
 - 11) Murase, H. and Nayar, S.K.: Visual Learning and Recognition of 3D Objects from Appearance, *International Journal of Computer Vision*, Vol.14, pp.5–24 (1995).
 - 12) Rowley, H.A., Baluja, S. and Kanade, T.: Rotational Invariant Neural Network-Based Face Detection, *Proc. Computer Vision and Pattern Recognition*, pp.38–44 (1998).
 - 13) 山口 修, 福井和広: 定性的 3 値表現による画像マッチング, 電子情報通信学会技術研究報

告, PRMU2002-34, Vol.102, No.156, pp.23–30 (2002).

- 14) Nene, S.A., Nayar, S.K. and Murase, H.: Columbia Object Image Library (COIL-100), *Columbia University Technical Report CUCS-006-96* (1996).

(平成 14 年 7 月 11 日受付)

(平成 15 年 1 月 16 日採録)

(担当編集委員 尺長 健)



山口 修 (正会員)

1969 年生 . 1994 年岡山大学大学院工学研究科情報工学専攻修士課程修了 . 同年 (株) 東芝入社 . 現在 (株) 東芝研究開発センターマルチメディアラボラトリー研究主務 . 顔画像認識, ヒューマンインタフェースの研究開発に従事 . 1996 年本会全国大会優秀賞受賞, 2002 年電子情報通信学会論文賞受賞, 電子情報通信学会会員 .



福井 和広

1963 生 . 1988 年九州大学大学院工学研究科機械工学専攻修士課程修了 . 同年 (株) 東芝入社 . 現在 (株) 東芝研究開発センターマルチメディアラボラトリー主任研究員 . コンピュータビジョン, ヒューマンインタフェースの研究開発に従事 . 2002 年電子情報通信学会論文賞受賞, 電子情報通信学会会員 .