

顔画像の大きさと個人性を考慮した顔位置推定

平山 高嗣[†] 岩井 儀雄[†] 谷内田 正彦[†]

顔画像の自動認識において、顔位置の推定は最も重要な過程である。顔の形状モデルを用いて顔位置を推定する場合、最適なモデルが用いられなければ、その精度は低下する。この問題を解決するにあたり、推定精度と処理時間のトレードオフが生じる。本研究では、顔画像の大きさと個人性を考慮した顔位置推定手法を提案する。この手法は顔の位置と大きさと個人モデルを並列的に効率良く推定する。顔の位置と大きさは、ダイナミックリンクアーキテクチャとスケール変換に基づく繰り返し処理によって推定される。個人モデルの推定は、クラスタリングによって階層化されたモデルデータベースを利用する。この推定は、顔の位置と大きさの推定が進むにつれ、その階層構造に従って最適なモデルを含むクラスタを選択していく。最終的に選択された個人モデルは個人識別結果とみなすことができる。

Facial Position Estimation adapted to variations of facial scale and individuals

TAKATSUGU HIRAYAMA,[†] YOSHIO IWAI[†] and MASAHIKO YACHIDA[†]

The facial position estimation is the most important in processes of automatic face recognition. If facial position is estimated by using a geometrical model, the estimation accuracy deteriorates unless the model is optimal for facial appearance. In this work, we deal with variations of facial scale and individuals, and we propose an efficient method that estimates facial position in parallel with facial scale and a personal model. Facial position and scale are iteratively estimated by the process based on the Dynamic Link Architecture and the Scale Conversion. The proposed method estimates a personal model by using a hierarchical face model database. As the facial position estimation proceeds, the method selects the optimal model according to the hierarchical structure of the database. We can regard the selected personal model as a person identification result.

1. はじめに

顔画像認識を応用したヒューマンインターフェイスやセキュリティシステムが有効視され、近年盛んに研究されている。その応用システムは主に、顔検出、顔器官(特徴)検出、そして個人識別、あるいは表情認識から構成される。この構成の中で、最も重要とされるのが顔検出である。顔検出に関する研究は、顔領域検出と顔位置推定に大別され、後者が特に重要とされる。なぜなら、その推定精度がその他の認識精度に影響を及ぼすからである。また、顔認識システムのユーザがカメラに向けて顔位置を合わせるという作業は望ましくなく、システムが顔位置を追跡する必要があることから、顔位置推定は重要な役割を持っている。

従来の顔検出手法としては、肌色や顔輪郭のエッジを利用した特徴ベースの手法¹⁾²⁾、PCA やニューラ

ルネットワークを応用したアピアランススペースの手法³⁾⁴⁾⁵⁾⁶⁾⁷⁾、顔器官の配置を考慮した幾何学的特徴ベースの手法⁸⁾⁹⁾¹⁰⁾ などがある。特徴ベースの手法は、顔領域検出を目的に提案されることが多い。計算コストの小さな処理で大まかな顔領域を検出できるため、実時間での顔発見や顔位置推定の前処理に適している。アピアランススペースや幾何学的特徴ベースの手法は、顔に関する知識を複合的に利用するため、詳細に顔位置を推定できる。

近年の顔認識研究の議論の中心は、顔の見え方の多様性に対するロバストな認識である。顔はその姿勢や照明に代表される撮影条件の変化、表情や時間経過に伴う容貌の変化といったそれ自体の変形などによって、その見え方が様々に変化する。顔位置を正確に推定するためには、顔の見え方の変化を推定し、最適なモデルを用いなければならない。一般的には、この問題は様々な見え方のモデルを用意し、その全てを顔画像との照合に用いることで解決できる。しかし、多数のモデルを照合に用いれば、処理時間が増大する。顔認識の応用システムはリアルタイム性が要求されるため、

[†] 大阪大学大学院基礎工学研究科
Graduate School of Engineering Science, Osaka University

この推定精度と処理時間のトレードオフを解決しなければならない。

我々はこれまでに、顔グラフ照合に基づく顔認識手法に関する研究を行ってきた¹¹⁾¹²⁾。顔グラフのノードは、顔の特徴点に配置され、その位置情報と特徴点の周辺から抽出されるガボールウェーブレット特徴量¹³⁾を保持する。この顔グラフを利用した顔認識手法は、現在最も高い認識性能を示すことができるとされている¹⁴⁾。顔グラフ照合の代表としては、Elastic Bunch Graph Matching¹⁵⁾と我々が提案したフレキシブル特徴照合法¹¹⁾がある。これらの照合手法の問題点は、平均モデルグラフを利用したラスタスキャンによって顔位置推定を行うことにある。つまり、平均モデルと顔画像の大きさに差があるほど顔位置推定の精度が低下する。一般的に、アピアランススペースの手法や幾何学的特徴ベースの手法で顔位置を推定する場合、顔の大きさが正確に推定されなければ、顔位置も正確に推定されない。また、この逆も成り立つ。この依存性の解決は、計算コストの制約がある場合、つまり、あらゆる大きさのモデルを用いてラスタスキャンできない場合に難題となる。本研究では、これを解決することを目指す。さらに、平均モデルとの個人性に関する類似度が低い顔画像に対しても、顔位置の推定精度が低下すると考えられるため、顔の個人性の変化に対してロバストな手法を開発することも目的とする。顔の大きさと同様に、最適な個人モデルが推定されなければならない。

そこで本研究では、顔の位置と大きさと個人モデルの推定を並列的に効率良く行う手法を提案する。顔位置の推定には、ダイナミックリンクアーキテクチャ(DLA)¹⁶⁾の基本概念に基づく計算効率の良い手法を、顔の大きさの推定には、我々が提案したスケール変換¹²⁾という高効率な統計的手法を適用する。この2つの手法を交互に繰り返すことで、並列的に双方の推定精度を高めていく。さらに、この並列的な処理に個人モデルの推定が加えられる。この推定が常に全ての個人モデルから行われることは非効率であるため、クラスタリングによって階層化されたモデルデータベースを用いる。つまり、顔の位置と大きさの推定が進むにつれ、その階層構造に従って最適なモデルを含むクラスタが選択されていく。最終的に選択された個人モデルは個人識別結果とみなすことができる。本稿では、まず、個人モデルデータベースの階層的クラスタリングについて述べる。次に、顔の位置と大きさと個人モデルの推定を並列的に行う手法について述べる。そして、この手法に基づき、顔画像の大きさと個人性の変化にロバストな顔位置推定システムを構築し、評価実験によってその有効性を検証する。

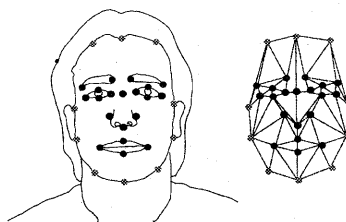


図1 顔の表現

2. 顔の表現

本研究では、目、唇、鼻といった顔の30個の特徴点上に配置されたノードを結ぶグラフによって顔を表現する。図1に顔グラフの例を示す。各ノードは対応する特徴点の位置におけるガボール特徴量を保持する。この特徴量は照明変動にロバストな特徴量で、ガボールウェーブレット変換によって求められる¹³⁾¹⁵⁾。ガボールウェーブレット変換は、画像とガボールカーネルの畳み込み演算によって行われる。本研究では、我々が開発したガボールカーネル¹²⁾を使用する。このカーネルは、顔の大きさに対してガボール特徴量を正規化することに対して優れた性能を持つ。また、この正規化により認識精度が向上し、計算コストが削減されることを我々は確認している。

3. 個人モデルの階層的クラスタリング

本手法では、個人モデル推定の効率化を図るために、個人モデルデータベースの構造を階層的クラスタリング¹⁷⁾によって階層化する。階層化されたモデルデータベースは図2のような構造となる。階層的クラスタリングは、ボトムアップ手法とトップダウン手法がある。ボトムアップ手法は、 j 個の1要素クラスタから連続的にクラスタを融合し、各階層でのクラスタを形成していく。一方、トップダウン手法は、 j 個全ての要素を含む1個のクラスタから連続的にクラスタを分割していく。ボトムアップ手法で全ての要素を完全に融合するまでクラスタリングを続けた場合と、トップダウン手法で全ての要素を完全に分割するまでクラスタリングを続けた場合に構成される階層的クラスタは、融合と分割の評価が同じであれば同等となる。 j 個のクラスタを k ($k \neq 1$)個のクラスタへと段階的に融合する場合にはボトムアップ手法が、1個のクラスタを k ($k \neq j$)個のクラスタに段階的に分割する場合にはトップダウン手法が適用される。つまり、クラスタリングを途中で打ち切る場合には、適用する手法が異なる。本研究では、クラスタリングを途中で打ち切らないので、どちらの手法も適用することができるが、アルゴリズムがより簡単なボトムアップ手法を適用することにする。

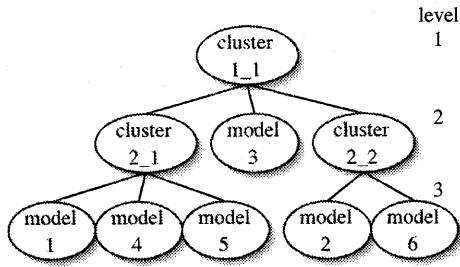


図2 階層的モデルデータベースの例

次に、クラスタの融合の評価方法について述べる。まず、評価に用いる尺度についてである。一般的には、クラスタリングは特徴空間のユークリッド距離を用いるが、本手法では顔グラフの類似度の計算にガボール特徴量の相関値を用いているので、それを評価の尺度とする。式(1)に顔グラフ X と Y の相関関数を示す。式(1)の m は顔グラフのノードを表す。評価のアルゴリズムは、最近傍アルゴリズムと最遠近傍アルゴリズムが代表的である。最近傍アルゴリズムは、クラスタ間の最大相関値がある閾値以上であればクラスタを融合する。これは単連結アルゴリズムとも呼ばれ、クラスタ内の要素が互いに必ずしも類似しない細長いクラスタを形成する可能性がある。一方、最遠近傍アルゴリズムは、クラスタ間の最小相関値がある閾値以上であればクラスタを融合する。これは完全結合アルゴリズムとも呼ばれ、形成されたクラスタは、クラスタ内の要素が互いに閾値以上の類似度を持つことになる。本手法では、クラスタ内の要素が互いに類似していることが望ましいので、最遠近傍アルゴリズムを利用してクラスタを形成する。階層構造は、本研究ではボトムアップ手法を利用するので、閾値を大きな値から徐々に小さくすることで構築できる。

$$R = \frac{1}{30} \cdot \frac{\sum_{m=1}^{30} X_m Y_m}{\sqrt{\sum_{m=1}^{30} X_m^2 \sum_{m=1}^{30} Y_m^2}} \quad (1)$$

個人モデルの推定は、このように構築された階層構造を上位階層からたどることで行う。そのためには、各クラスタの代表データを定義しなければならない。本研究では、クラスタ内の要素が互いに類似していることから、代表データをクラスタ中心とする。これは、個人モデルグラフが持つガボール特徴量の平均から求められ、我々はこれをクラスタ平均モデルグラフと呼ぶ。このグラフが顔の位置と大きさを推定するための参照モデルグラフとなる。

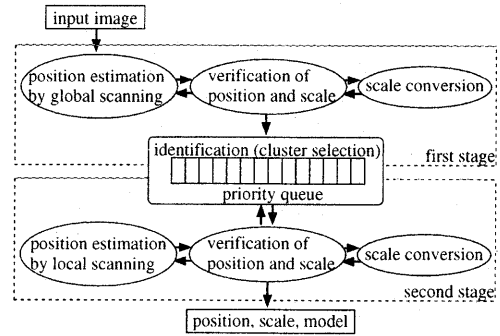


図3 提案手法の状態遷移

4. 顔画像の大きさと個人性を考慮した顔位置推定

4.1 提案手法の概要

提案手法の概略を図3に示す。本手法は、大域探索による位置推定、局所探索による位置推定、推定位置と大きさの検証、スケール変換、個人モデル推定の5つの状態から構成され、これらの遷移の繰り返しによって顔の位置と大きさ個人モデルの推定を行う。大域探索による位置推定は顔の位置を小さな計算コストで推定することができる。しかし、その反面、安定した推定を行えない。そこで、次にこの推定位置を中心に局所的に詳細な探索を行う手法を適用する。その手法が局所探索による位置推定である。この推定法は、大域探索によるものより大きな計算コストを要するが、より正確な推定を行える。つまり、顔位置の推定は粗密法に基づいて行われる。検証状態は推定した位置と大きさから生成された顔グラフの妥当性を検証する。スケール変換は我々が開発した手法で、顔の大きさを推定する。この手法は、他の顔の大きさ推定の手法より計算効率が良い。そして、個人モデル推定は階層化されたモデルデータベースと優先度付き待ち行列を用いて行う。本手法は、この待ち行列の上位に存在する顔グラフの位置と大きさを更新し、その顔グラフが参照するクラスタ平均モデルグラフから階層構造をたどって最適なモデルを選択していくという処理を繰り返す。

5つの状態のそれぞれについてと、各状態間の遷移について次節より詳しく順に述べる。

4.2 大域探索による位置推定

大域探索による位置推定は、ダイナミックリンククアーキテクチャ(DLA)¹⁶⁾の基本概念に基づく。まず、画像全体に d_G 画素間隔で矩形格子状に $n_G \times n_G$ 個のサンプリング点を配置し、それらの点上からガボール特徴量を抽出する。そして、参照モデルグラフの各ノードに対して次の特徴量類似度 ϕ_{ml} を計算する。

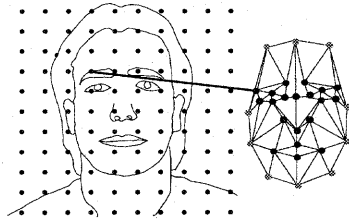


図4 大域探索による位置推定

$$\begin{aligned} \phi_{ml} &= 1 - \|V_m - D_l\|, & (2) \\ m &= 1, \dots, 20, \\ l &= 1, \dots, n_G^2. \end{aligned}$$

この類似度を局所特徴量類似度と呼ぶことにする。ここで、 D_l はサンプリング点 l のガボール特徴量、 V_m は参照モデルグラフのノード m のそれである。また、この参照モデルグラフは、最上位階層のクラスタ平均モデルグラフ、つまり、全ての個人モデルグラフの平均グラフである。位置の推定は、最も局所特徴量類似度が高いサンプリング点 l を探索し (図4)、ノード m がその点 l と一致するように参照モデルグラフを配置することによって行われる。この一連の処理が、参照モデルグラフの顔輪郭部以外の20個のノードに対して行われる。ゆえに、大域探索による位置推定は20個の推定を行う。

4.3 推定した位置と大きさの検証

推定した顔の位置と大きさの検証は、大域特徴量類似度 π を用いて行う。顔の大きさの推定を行うスケール変換については4.4節で示す。

$$\begin{aligned} \pi(G(x, y, s_{opt})) &= \frac{1}{30} \cdot \\ & \frac{\sum_{m=1}^{30} V_m W_m(G(x, y, s_{opt}))}{\sqrt{\sum_{m=1}^{30} V_m^2 \sum_{m=1}^{30} W_m(G(x, y, s_{opt}))^2}}. & (3) \end{aligned}$$

ここで、 $G(x, y, s_{opt})$ は参照モデルグラフから生成されたシャドウグラフ G である。シャドウグラフは推定された位置 (x, y) と大きさ (s_{opt}) を持つ。また、 W_m はシャドウグラフのガボール特徴量である。このシャドウグラフが個人モデル推定の状態へ入力される。

4.4 スケール変換

スケール変換は顔グラフを入力画像に対して適切な大きさに効率的に伸縮させる役割を果たす。これには、スケール辞書というデータベースを利用する。スケール辞書はあらかじめシステム構築時に作成される。

まず、スケール辞書の生成について説明する。顔の大きさが平均グラフとほぼ等しい N_t 枚の教師顔画像を用意する。この平均グラフは全ての個人モデルグラフから生成されたものである。次に、平均グラフを伸縮率 s_T で T 段階に伸縮させたグラフを、各教師顔画像の顔位置 (x_{true}, y_{true}) に配置する。このグラフの妥当性は大域特徴量類似度 π によって求められる。

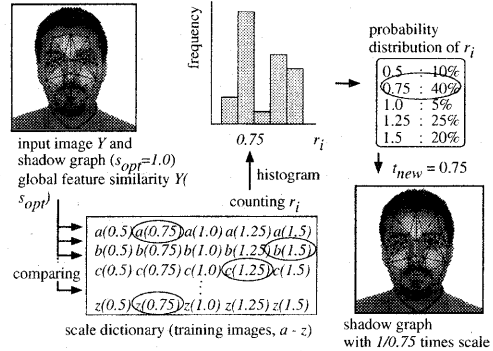


図5 スケール変換

つまり、式(3)の V_m を平均グラフのガボール特徴量、 W_m を伸縮グラフのそれとして求められる。次に、この大域特徴量類似度がスケール辞書と呼ばれるデータベースに学習データとして登録される。

入力画像の顔の大きさの推定は、この辞書を用いて行われる。まず、検証状態から遷移してきたシャドウグラフ G の大域特徴量類似度をスケール辞書の学習データと比較する。比較は教師顔画像ごとに行い、最も誤差の小さい学習データをそれぞれ検出し、そのデータに対応した平均グラフの伸縮率 r_i を導き出す。

$$\begin{aligned} r_i &= \arg \min_{s_T} [\pi(G(x, y, s_{opt})) \\ & - \pi(G_{ave}(x_{true}, y_{true}, s_T))] & (4) \\ & \text{for each training image,} \\ & i = 1, \dots, N_t, \quad T = 1, \dots, T. \end{aligned}$$

ここで、式(4)の $\pi(G_{ave}(x_{true}, y_{true}, s_T))$ はスケール辞書の学習データである。次に、全ての教師顔画像から求めた r_i のヒストグラムを生成する。このヒストグラムから求まる r_i の確率分布に従い、入力画像に対するグラフの大きさの比率 t_{new} を推定する。シャドウグラフの適切な大きさへの伸縮は、 t_{new} によって更新された伸縮率 s_{opt} を求めて行われる ($s_{opt} \leftarrow s_{opt} \times s_{new}$ ($s_{new} = 1/t_{new}$)). また、伸縮率 s_{opt} は、画像から抽出されるガボール特徴量を顔の大きさに対して正規化するために用いられる。図5にスケール変換の処理の流れを示す。

4.5 個人モデル推定

個人モデルの推定は、検証状態から遷移してきたシャドウグラフ G がどのクラスタに属するかを判定する。シャドウグラフが参照するクラスタ平均モデルグラフが、レベル L の階層のあるクラスタの代表とする。まず、このクラスタに属する下位階層 ($L+1$) の全てのクラスタ平均モデルグラフとシャドウグラフの照合を行う。照合には大域特徴量類似度 π を用いる。つまり、式(3)の V_m をクラスタ平均モデルグラフのガボール特徴量とする。次に、照合後の全てのシャドウグラフをその大域特徴量類似度を優先度とし

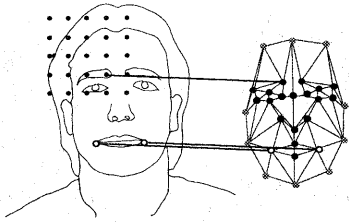


図6 局所探索による位置推定

て優先度付き待ち行列に格納する。この際、シャドウグラフの参照モデルグラフを、照合に用いたそれぞれのクラスタ平均モデルグラフに変更する。そして、優先度付き待ち行列の上位に存在するシャドウグラフの参照モデルグラフがモデル推定の結果となる。

4.6 局所探索による位置推定

局所探索による位置推定も、DLAの基本概念に基づく。この推定法は、大域探索による位置推定と類似するが、探索領域の大きさと局所特徴量類似度の計算に用いられるガボール特徴量の数に関してそれと異なる。

まず、検証状態から遷移してきたシャドウグラフが持つ推定位置のまわりに、 $d_L \times s_{opt}$ 画素間隔で矩形格子状に $n_L \times n_L (< n_G \times n_G)$ 個のサンプリング点を配置する。状態遷移の初期には、 $d_L \times s_{opt} \times n_L$ は顔の大きさとほぼ等しくなる。なお、状態遷移の過程でサンプリング点と同じ位置に配置されることを防ぐために、 d_L を単調減少させる。次に、シャドウグラフが参照するクラスタ平均モデルグラフのノード m に対して式(5)の局所特徴量類似度が計算される。

$$\phi_{ml} = 1 - \|V_m - D_l(s_{opt})\|, \quad (5)$$

$$m = 1, \dots, 20,$$

$$l = 1, \dots, n_L^2.$$

この式はガボール特徴量が顔の大きさと正規化される点で式(2)と異なる。そして、最も高い局所特徴量類似度を持つ点 l を探索し、参照モデルグラフをそのノード m が点 l と一致するように移動させる。それから、その移動されたグラフ上で n_D 個のガボール特徴量(R_n)を抽出する(図6)。図6中の白色ノードは n_D 個のガボール特徴量を抽出するためのサンプリング点の例である。そして、次の式で表される局所特徴量類似度の合計 ψ を計算する。

$$\psi = \phi_{ml} + \sum_{n=1}^{n_D} [1 - \|V_n - R_n(s_{opt})\|]. \quad (6)$$

この一連の処理が、参照モデルグラフの顔輪郭部以外の20個のノードに適用され、最も高い ψ を持つグラフの位置を局所探索による位置推定の結果とする。

4.7 状態遷移

前節までに示したそれぞれの処理の遷移について述べる。図3に示すように、本手法は2つのステージ

で構成される。第1ステージでは、大域探索による位置推定とスケール変換が、検証状態を介して交互に行われる。第2ステージは第1ステージに類似するが、局所探索による位置推定が第1ステージの大域探索による位置推定の代わりに行われる。それぞれのステージにおいて、スケール変換への遷移は、検証時に求まる大域特徴量類似度がスケール辞書の学習データの範囲内となった時に行われる。また、終了状態への遷移は、モデルの推定が階層的クラスタの最下位層まで進み、かつ大域特徴量類似度がある閾値を越えた場合に起こる。その時のシャドウグラフの参照モデルグラフが個人モデルグラフとなっているので、それを個人識別の結果とみなすことができる。なお、未登録人物の顔画像を入力した時などは、終了状態への遷移の条件を満たすことができない可能性があるため、状態遷移の回数がある閾値を越えた場合にも終了状態への遷移が起こるようにする。

本手法では、顔の位置と大きさと個人モデルの推定精度を向上させるために、ビーム探索¹⁸⁾によってこれらの更新を行う。我々は、優先度付き待ち行列を利用してこれを実現する。つまり、この待ち行列から優先度の高いシャドウグラフを c 個選択し、それぞれから並列的にさらなる推定を行う。そして、これを繰り返す。ビーム探索は計算効率が c の数に依存するが、 c 個からの更新はそれぞれ独立に行うことができるため、 c 個のプロセスに負分散させるという並列処理に展開することが可能である。すなわち、この並列処理により、顔認識の精度と処理時間のトレードオフの解決が期待できる。

5. 評価実験

5.1 実験環境

提案手法の有効性を検証するために実験システムを構築し、評価実験を行った。比較に用いるシステムは、我々が提案した従来システム¹²⁾である。このシステムは、5画素間隔のラスタスキャンによって顔位置を検出し、その後にスケール変換によって顔の大きさを推定する。そして我々が開発したフレキシブル特徴照合法を用いて顔の特徴点を検出し、その特徴量とモデルデータベースとの比較により個人識別を行うというシステムである。それぞれのシステムの実装はSGI製ワークステーション Origin300 (SCPU R14000, 500 MHz, memory 4 GB, OS IRIX6.5.16m) 上で行った。画像はPurdue大学のAR Face データベース¹⁹⁾のもの(256 × 256画素, 8bit濃淡)を利用した。個人モデルデータベースには、ほぼ等しい大きさの50人の顔画像とそれぞれのモデルグラフが登録されている。画像は1人につき1枚で、全てが無表情の正面顔である。ただし、眼鏡をかけた画像や照明変動のある画像も含まれる。モデルグラフの特徴点は我々が目視に

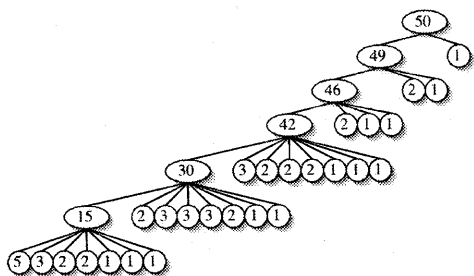


図7 構成された階層的クラスタ

よって画像中で定義した. ガボールウェーブレット変換は2つの大きさと8方向のカーネルを用いて行われた. 提案システムで用いられるパラメータは以下のように設定された: $n_G = 12, d_G = 20, n_L = 5, d_L = 30, n_D = 2, N_t = 20, T = 11, s_T = 0.5 \sim 1.5$ (0.1刻み), $c = 5$. また, n_D 個のガボール特徴量は目の位置から抽出された. テスト画像は, モデルデータベースに登録された人物のモデル画像, 登録されていない人物の画像がそれぞれ50枚(テストセット100%)と, それらを80%と120%に伸縮した画像(テストセット80%, 120%)200枚と, テストセット100%の画像から顔の大きさは変わらずに表情がわずかに変化した画像100枚(テストセット expression)で, 計400枚とした. なお, 画像の伸縮はMacOS X上でAdobe Photoshop7.0を用いてbi-cubic内挿法で行われた.

5.2 構成された階層的クラスタ

本研究では, 階層的クラスタリングを行うために設定する相関値の閾値を0.9から0.02刻みで減少させた. 個人モデルグラフ間の相関値の最大は0.92, 最小は0.78であった. 構成された階層的クラスタを図7に示す. 図中の楕円と円はクラスタを表す. 数字はクラスタ内の要素(個人モデルグラフ)の数である. この階層的クラスタはレベル数が8, 最大分岐数が8である. 我々がこの階層的クラスタから得る見解は, 本研究で用いた個人モデルデータベースの特徴量空間は, いくつかの大きなクラスタに分割できるような分布ではなく, 正規分布のような分布をしているということである.

5.3 位置と大きさの推定精度

図8に提案システムを用いて顔位置と大きさを推定した結果を示す. この結果から, 顔位置と大きさが変化した画像だけでなく, 照明が眼鏡に反射し, 目が隠れている顔画像や照明条件の異なる顔画像, そして表情がわずかに変化した顔画像に対しても顔位置と大きさの推定が正しく行われていることがわかる. また, 個人モデルの推定によって最適なモデルグラフを選択しているので, 表情が変化した顔画像以外で正確に特徴点を検出できていることもわかる. 表情が変化した顔画像の特徴点を正確に検出するには, フレキシブル

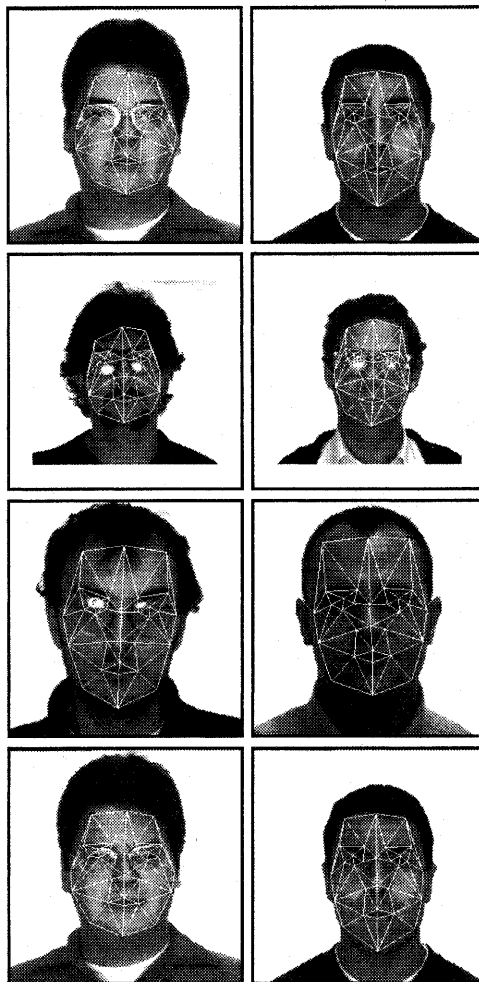


図8 実験結果

特徴照合法を適用する必要がある. この場合, 個人モデルの推定を行っている提案システムは, 従来システムより小さな計算コストで精度良く特徴点を検出できる可能性がある. 従来システムには平均モデルグラフを利用したフレキシブル特徴照合法が導入されている.

表1にテストセット100%, 80%, 120%, expression内の登録人物の顔画像に関して推定位置と真値の誤差, 推定した大きさ, 計算コストを求めた結果を示す. 計算コストはガボールウェーブレット変換を行ったノードの数である. なお, 従来システムの計算コストにフレキシブル特徴照合法で費された計算コストは含まれない*. この表から, 提案システムは, 顔位置

* フレキシブル特徴照合法で費された計算コストは, およそ2000であった.

表 1 顔の位置と大きさの推定の精度と計算コスト

system	test set	position error (pixel)	estimated scale (%)	cost
proposed	100%	1.61	100+1.0	6986
	80%	1.71	80+3.0	6774
	120%	2.13	120+2.2	7132
	expression	3.91	100+2.4	6918
previous	100%	3.14	100+1.6	30944
	80%	7.52	80+4.2	30944
	120%	19.5	120+10	30944
	expression	4.46	100+3.8	30944

と大きさの推定精度に関して、従来システムより向上し、真値に近い値を得たことがわかる。また、計算コストに関しては、従来システムの 1/5 程度に軽減された。提案システムは、 c 個のプロセッサを用いて並列処理を行うことができる。本実験では $c = 5$ としているので、提案システムの計算コストは、表中の値よりさらに 1/5 程度の軽減が可能である。ゆえに、提案システムは推定精度と計算効率のトレードオフを解決できるといえる。なお、本実験では、1 個のノードに対してガボールウェーブレット変換を施すのにかかる処理時間は約 6 ミリ秒であった。

5.4 個人識別の精度

提案システムが最終的に推定した個人モデル、つまり終了状態に遷移した個人モデルは、個人識別の結果とみなされる。我々は、個人識別の精度を本人受入率と他人拒否率から求める。本人受入率は、モデルデータベースに登録されている人物の顔画像が入力された場合に、その人物の個人モデルが正しく推定された割合である。ただし、推定結果の終了状態への遷移が、設定された閾値を越えずに行われた場合は、システムはその結果を受理しないものとする。この受理しないケースが未登録の人物の顔画像が入力された場合に起きる割合を他人拒否率とする。図 9 にそれぞれの実験システムの認識率を示す。この図は、提案システムでは、終了状態への遷移の判定基準の閾値を変化させ、従来システムでは、個人識別処理で設定する必要がある閾値を変化させ、それに対する本人受入率と他人拒否率を ROC 曲線の形で表したものである。この曲線はグラフの右上に位置するほど、その精度が良いといえる。従って、テストセット 80%、100%、120%、expression の全てに対して、提案システムの精度向上が確認できる。提案システムは、登録人物の顔画像に対しては個人モデルを推定できるため、顔位置の推定精度が高くなり、本人受入率も高めることができる。一方、未登録人物に対しては個人モデルを推定できないため、顔位置の推定精度は幾分低くなり、高い他人拒否率が得られると考えられる。また、テストセット expression の精度が他のものより低くなっていることも確認できる。この精度を向上させるためには、提案システムにフレキシブル特徴照合法を追加導入して特徴点を正確に検出するなどの改良が必要である。

提案システムの問題点としては、データベースに登

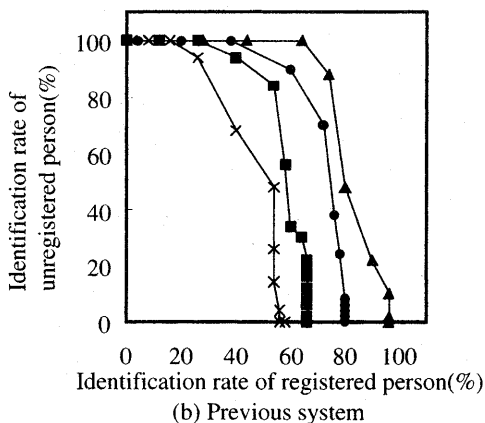
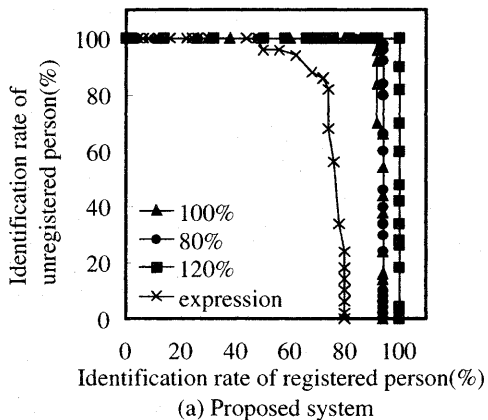


図 9 個人識別率

録される個人モデルが増加した場合に、1つのクラスターから分岐するクラスター数も増加するということがある。この場合、本実験結果の精度を維持するためには、ビーム探索の c を大きくしなければならない可能性がある。分岐数は、階層的クラスタをトップダウン手法と K 平均アルゴリズムで構築し、 K 分木とすることで抑えることができるが、逆にレベル数が増加し、計算コストも増加してしまう。この問題の解決は今後の課題である。我々は、まず登録人物を増やして提案システムの性能を検証しなければならない。

6. おわりに

本稿では、顔の位置と大きさと個人モデルを並列的に効率良く推定する手法を提案した。提案手法は、大域探索による位置推定、局所探索による位置推定、顔の大きさを推定するスケール変換、推定された位置と大きさの検証、個人モデル推定の5状態の遷移をビーム探索によって繰り返す。効率的な個人モデル推定を行うために、階層的クラスタリングによって個人モデルデータベースの構造を階層化した。我々は顔認識実験によって提案システムの性能を評価し、顔の位置と大きさと個人モデルの並列的な推定が顔認識の精度と計算効率のトレードオフを解決できるという着目点の正しさも実証した。

今後の課題としては、個人モデルデータベースの大きさによって提案システムの性能がどう変化するかを検証することや、我々が提案している従来システム以外との比較を行うこと、そして表情が変化した画像に対する顔位置と大きさと個人モデルの推定精度を向上させることなどが挙げられる。

謝辞 本論文で述べた研究の一部は日本学術振興会未来開拓研究推進事業「感性的ヒューマンインタフェース」(プロジェクト番号 JSPS-RFTF99P01404) の援助を受けた。

参考文献

- 1) 呉海元, 陳謙, 谷内田正彦: ファジィパターン照合を用いた色彩画像からの顔検出システム, 電子情報通信学会論文誌 D-II, Vol. 80, No. 7, pp. 1774-1785 (1998).
- 2) C. Kotropoulos, A. Tefas and I. Pitas: Morphological elastic graph matching applied to frontal face authentication under well-controlled and real conditions, *Pattern Recognition*, Vol. 33, pp. 1935-1947 (2000).
- 3) M. Turk and A. Pentland: Eigenface for recognition, *Journal of Cognitive Neuroscience*, Vol. 3, No. 1, pp. 71-86 (1991).
- 4) H. Rowley, S. Baluja and T. Kanade: Neural network-based face detection, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 203-208 (1996).
- 5) P. Penev and J. Atick: Local feature analysis: A general statistical theory for object representation, *Neural Systems*, Vol. 7, No. 3, pp. 477-500 (1996).
- 6) P. Belhumeur, J. Hespanha and D. Kriegman: Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 711-720 (1997).
- 7) P. Viola and M. Jones: Rapid object detec-

- tion using a boosted cascade of simple features, *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 511-518 (2001).
- 8) R. Brunelli and T. Poggio: Face recognition through geometrical features, *Proceedings of the 2nd European Conference on Computer Vision*, pp. 792-800 (1992).
- 9) A. Lanitis, C. J. Taylor and T. F. Cootes: Automatic interpretation and coding of face images using flexible models, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 743-756 (1997).
- 10) K. C. Yow and R. Cipolla: Enhancing human face detection using motion and active contours, *Proceedings of the 3rd Asian Conference on Computer Vision*, pp. 515-522 (1998).
- 11) D. Pramadihanto, Y. Iwai and M. Yachida: Integrated Person Identification and Expression Recognition from Facial Images, *IEICE Transactions on Information and System*, Vol. E84-D, No. 7, pp. 856-866 (2001).
- 12) 平山高嗣, 岩井儀雄, 谷内田正彦: 顔の大きさの変化にロバストな個人識別システム, 情報処理学会論文誌, Vol. 44, No. 6, pp. 1625-1634 (2003).
- 13) J. G. Daugman: Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters, *Journal of the Optical Society of America A*, Vol. 2, pp. 1160-1169 (1985).
- 14) O. Ayinde and Y.-H. Yang: Face recognition approach based on rank correlation of Gabor-filtered images, *Pattern Recognition*, Vol. 35, pp. 1275-1289 (2002).
- 15) L. Wiskott, J. M. Fellous, N. Krüger and C. von der Malsburg: Face recognition by Elastic Bunch Graph Matching, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 775-779 (1997).
- 16) M. Lades, J. C. Vorbrüggen, J. Buhmann, J. Lange, C. von der Malsburg, R. P. Würtz and W. Konen: Distortion invariant object recognition in the Dynamic Link Architecture, *IEEE Trans. on Computers*, Vol. 42, pp. 300-311 (1993).
- 17) R. O. Duda, P. E. Hart and D. G. Stork: Pattern classification, Wiley-Interscience (2001).
- 18) A. Barr and E. A. Feigenbaum: The Handbook of Artificial Intelligence, Addison-Wesley, Menlo Park, CA (1981).
- 19) A. M. Martinez and R. Benavente: The AR face database, CVC Technical Report24 (1998).