

形状抽出とパターン照合の組合せによる顔特徴点抽出

福井 和広[†] 山口 修[†]

Facial Feature Point Extraction Method Based on Combination of Shape Extraction and Pattern Matching

Kazuhiro FUKUI[†] and Osamu YAMAGUCHI[†]

あらまし 本論文では、顔認識に向けて動画像中から瞳、鼻孔、口端などの顔特徴点を高速かつ高精度に抽出する方法を提案する。これらの特徴点を基準とした顔の切出し精度は、テンプレートマッチングに基づく顔認識法の性能に大きく影響する。また処理速度に関しては、動画像に対して識別の試行回数を増やす観点あるいは顔辞書の実時間生成という実用的な観点から高速性が要求される。これまでにさまざまな抽出法が提案されているが、個人差、表情変化、顔向き、照明変化などの影響のために安定な抽出が難しく、抽出精度と処理速度の面で十分には満足できるとは言えなかった。提案する方法は、形状抽出とパターン照合の組合せにより、少ない計算コストで高い位置精度を実現している。具体的には、分離度フィルタにより瞳、鼻孔、口端などの特徴点の候補を抽出する。次に部分空間法を使ったパターン照合により候補から正しい特徴点を選択する。評価システムを構築してさまざまな条件の顔画像に対して実験を行った結果、合計 1700 枚の静止画像に対して 99% の特徴点の抽出率、また動画像に対してハードウェアを用いずに 10 回/秒の抽出速度、9880 フレームに対して 97% の抽出率を達成できた。

キーワード 顔特徴点抽出, 顔認識, 部分空間法, 分離度

1. ま え が き

本論文では、顔認識に向けて動画像から瞳、鼻孔、口端など顔特徴点を、顔向き、照明変化の影響に対してロバストかつ高速・高精度で抽出する方法を提案する。特徴点の安定な抽出は、その形状および輝度が、個人差、表情変化、頭部の動き、照明変化などの影響によりさまざまに変化するため難しい課題である。

顔特徴点の抽出精度は、テンプレートマッチングに基づく顔識別法 [1], [2], [11]~[13] の性能に大きく影響する。これらの方法では、瞳、鼻孔、口端などの特徴点を基準とした 2 次元 affine 変換により位置、サイズを正規化した顔画像を切り出し、あらかじめ登録されている辞書画像とのパターン類似度を計る。従って、正規化画像の切出し精度つまり正規化の基準となる特徴点の抽出精度が強く識別性能に影響することになる。

また特徴点の抽出速度に関しては、動画像に対する顔の切出し・識別の試行回数を増やす観点あるいは顔

辞書の実時間生成のための顔サンプル画像の大量切出しという実用的な観点から高速性が不可欠である。例えば監視などで被識別者に意識させないで顔を識別する場合には、動画像列から識別可能なフレームを選択し利用することになるため、顔切出し回数が多いほど有利となる。辞書生成に関しては、従来の手書き漢字認識では、数百枚の学習サンプル画像から辞書を生成していたが、本論文で対象としている顔のような 3 次元物体は、文字に比べて形状や輝度の変動が大きく、更に大量の学習サンプル画像が必要となってくる [3]。従って従来の人手による顔切出し [1], [2] や正面向きなどの限られた条件下での自動切出しでは実用化への対応が難しかった。

これまでに提案されている顔特徴点抽出法は、大きく分けると“形状情報に基づく方法”と“パターン情報に基づく方法”に分類できる。

形状情報に基づく方法の代表的な例は、エッジ情報に基づく方法 [9], [16] である。これらの方法では、顔領域内のエッジに対して、エネルギー最小化を適用して構造モデルを当てはめ特徴点を抽出する。しかし初期パラメータの適切な設定が容易でない点や解像度、

[†] (株) 東芝関西研究所, 神戸市
TOSHIBA Kansai Research Laboratories, 8-6-26 Motoyama-minami-cho, Higashinada-ku, Kobe-shi, 658 Japan

照明条件により明瞭なエッジが得られない場合には安定な抽出が難しい点などの課題があった。

これに対してパターン情報に基づく方法は、エッジ情報を使わずに輝度やカラー情報を用いる方法である。例えば固有空間を適用した Eigen feature 法 [12] はエッジに基づく方法に比べてノイズや照明変化、形状変化にロバストで実画像に対する有効性も示されている。しかし画像間の内積計算を複数回行う必要があり処理量が膨大になる点やエッジに基づく方法に比べて位置検出の精度が低い点などの問題があった。カラー情報 [4] に基づく方法は濃淡情報に比べて照明条件の変化に対して比較的ロバストであるが、照明の位置や顔の向きによってはカラー情報が失われる場合も多く常に有効であるとは限らない。

処理速度に関しては、単純な特徴点抽出処理による特徴点のリアルタイム追跡法 [14] が提案されているが、追跡している点がどの顔特徴点に対応しているかわからないために正規化基準点としては使えない。

本論文では、従来のパターン情報に基づく方法の上記課題を解決するために形状情報とパターン情報を組み合わせた方法を提案する。基本的な考えは、位置精度の高い形状情報により特徴点の候補を求め、それをパターン照合で検証するというものである。具体的には領域ベースの分離度フィルタにより円形状あるいは端点に対応する特徴点の部分形状を求め、それを部分空間法 [10] で検証する。本方法は形状情報により位置決めを行うので高い位置精度が期待できる。更に候補群からの正しい特徴点の選択に、部分空間法によるパターンマッチングを用いるために、特徴点の形状、輝度の変動に対して比較的ロバストである。

処理速度に関しては、計算コストの少ない分離度フィルタで絞り込んだ候補に対してのみパターン照合を行うので全体に対してパターン照合を行う方法に比べ計算量の大幅な削減が実現できる。

更に、本論文では抽出された特徴点の幾何学的関係が正しいかを検証する方法について述べる。通常、短時間で大量に収集された学習パターンには、まばたきをした瞬間などに発生する抽出誤りが含まれてくる。そこで複数フレームにおける抽出済みの正しい顔特徴点の 2 次元座標から新しく抽出された特徴点と正しく抽出されているかを検証する。

以下、本論文では、動画像中から瞳、鼻穴、口端などの顔特徴点を照明変化や顔向きの変化に対して安定に抽出する方法、更にさまざまな条件下で撮影した静

止画像および動画像を用いた評価実験の結果について述べる。

2. 提案する特徴点抽出法

2.1 形状情報とパターン情報の組合せ

個人差、表情、顔向きなどに影響を受けにくい特徴点の部分抽出するという考えは、文献 [8] で有効性が述べられている。通常これらの抽出された部分特徴は誤った部分特徴も多数含んでいるために、実験的に得られた幾何拘束により正しい候補の選択を行う [8], [17]。しかし幾何拘束だけでは、顔向きや照明状態の加減により目尻を瞳と間違えるような微妙な位置誤りを検出することは難しかった。

これに対して、本論文では幾何拘束に加えてパターン照合を適用する。図 1 に示すように、特徴点の部分形状の情報（例えば、目における瞳の位置、直径）を基準に特徴点近傍の部分画像を位置とサイズに関して正規化して切り出す。ここで正しく部分形状が抽出されていれば、これを基準に正規化・切り取られた特徴点近傍の部分画像は、あらかじめ正しい形状情報に基づいて切り取られた辞書画像（以下、 $d1$ と表記）と類似度が高くなるはずである（図中 $f3$ の場合）。逆に形状情報が誤って抽出されている場合は、正しく正規化が行われず、辞書 $d1$ との類似度が低くなる（図中 $f1$, $f2$, $f4$ の場合）。なお本論文では瞳を円形で近似した場合について述べるが、更に正確にだ円近似まで行えば正規化の精度が向上してパターン照合の有効性が増すと思われる。

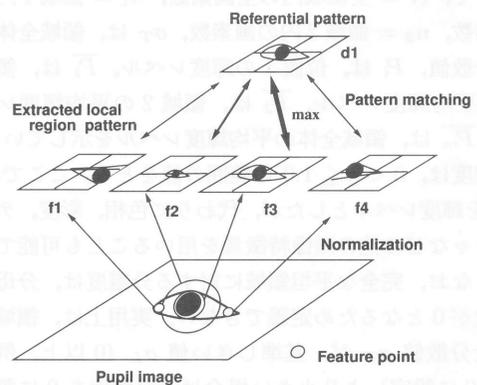


図 1 形状抽出とパターン照合の組合せ
Fig. 1 Combination of shape extraction and pattern matching.

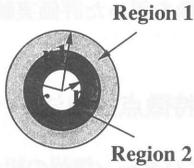


図2 分離度フィルタ
Fig.2 Separability filter.

2.2 分離度フィルタによる形状情報の抽出

従来のエッジベースの形状抽出法はノイズに弱いという欠点があった。この問題に対して、筆者らは領域ベースのエッジ[15]および輪郭抽出法[5],[6]を提案している。本論文では、瞳、鼻孔、口端などの特徴点を抽出するために“分離度フィルタ”を提案する。図2に示すような二つの同心円領域1 (r1-r) と領域2 (r-r2) からなるテンプレートマスク (rをマスク半径と定義) を考える。ここで線形判別法より領域1と2の領域情報(例えば、輝度)の分離程度を表す量“分離度”を導入する。この分離度は正規化された値で領域1と2の情報が完全に分離されている場合は最大値1.0となる。分離度は、ノイズの影響を受けにくく特徴点の明暗差(領域1と2の輝度差)に依存しないで求まる[15]。分離度フィルタは、以下の式から求まる分離度 η を出力する。

$$\eta = \frac{\sigma_b^2}{\sigma_T^2} \tag{1}$$

$$\sigma_b^2 = n_1(\overline{P_1} - \overline{P_m})^2 + n_2(\overline{P_2} - \overline{P_m})^2 \tag{2}$$

$$\sigma_T^2 = \sum_{i=1}^N (P_i - \overline{P_m})^2 \tag{3}$$

ここで、 N = 全領域内の全画素数、 n_1 = 領域1内の画素数、 n_2 = 領域2内の画素数、 σ_T は、領域全体の全分散値、 P_i は、位置 i の輝度レベル、 $\overline{P_1}$ は、領域1の平均輝度レベル、 $\overline{P_2}$ は、領域2の平均輝度レベル、 $\overline{P_m}$ は、領域全体の平均輝度レベルを示している。分離度は、 $0 < \eta \leq 1.0$ の範囲の値をとる。ここでは、 P_i を輝度レベルとしたが、代わりに色相、彩度、テクスチャなどの他の画像特徴量を用いることも可能である。なお、完全な平坦領域に対する分離度は、分母の分散が0となるため定義できない。実用上は、領域内の全分散値 σ_T が、基準しきい値 σ_L (0 以上、例えば10に設定) より小さい場合は、分離度を0に設定する。

この分離度フィルタの基本的な考えは、文献[16]の

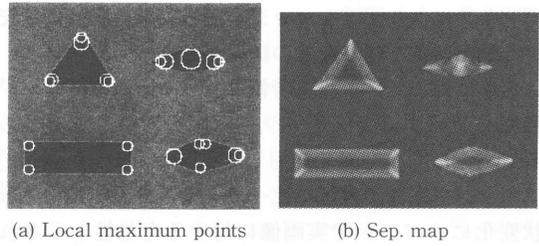


図3 出力例 (a)分離度の局所最大点、
(b)分離度分布(出力を255倍)
Fig.3 An example of output with sep. filter.

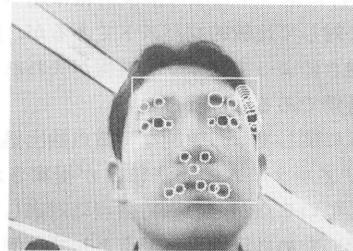


図4 分離度の局所最大点
Fig.4 Local maximum points of sep. for a real face image.

Blob detectorと同じであるが、Blob detectorと比較して計算式が非常にシンプルであり、経験により決定すべき余分なパラメータが不要である。また式(1)の分母 σ_T^2 は、 $\sigma_T^2 = \sum_{i=1}^N P_i^2 - (n_1 + n_2)\overline{P_m}^2$ と単純化できる。ここで P_i^2 の計算をあらかじめ用意した2乗テーブル参照に置き換えれば計算の大部分を和演算のみで構成でき高速化やハードウェア化に都合が良い。

更に文献[16]では円形の瞳を抽出するという単純な適用についてしか言及がないが、図3に示すように分離度フィルタは円形特徴点に加えて端点においても出力が局所最大(図中で円の中心は局所最大点、半径はマスク半径)となる。図4は、顔画像に対して分離度フィルタを適用した結果である。瞳、鼻孔、口端だけではなく、眉端、唇端、目頭、目尻などでも局所最大となっていることがわかる。本論文ではこの特性を利用して口端も抽出する。

2.3 パターン照合

特徴点候補から正しい特徴点を選択する。文献[16]ではBlob detectorで検出された瞳候補に対して、目全体のフレームモデルをエネルギー最小化を適用して当てはめて正しい瞳か否かを検証している。しかしエッジの存在が前提になっており、照明変化や顔向き



図5 瞳領域画像の例
Fig.5 Examples of left pupil image.

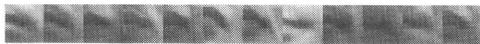


図6 口端画像の例
Fig.6 Examples of left mouth edge image.

にロバストであるとは言えなかった。また収束計算コストが大きい点や検出が初期パラメータに依存する点などの欠点もあった。これに対して本論文では、部分空間法 [10] を用いたパターン照合により正しい特徴点か否かを検証する。パターン照合ではエッジを抽出する必要がなく、パターン全体の情報を用いるのでノイズにもロバストである。また部分空間法を適用しているので、パターンの変形に対しても許容度が高い。ここで部分空間法について簡単に説明する。画像データを $n \times n$ 次元のベクトルとみなす。この $n \times n$ 次元の特徴空間中で特徴点が表す部分空間を複数の固有ベクトルが張る線形部分空間で表す。ここで入力ベクトル (パターン) を部分空間に投影したときの射影角度 θ は入力ベクトルが特徴点ベクトルにどれだけ類似しているか表している。識別では入力ベクトルと各特徴点の部分空間との θ を求め、最も高い θ が得られた部分空間に該当する特徴点として入力ベクトルを分類する。具体的には、射影成分 $\cos \theta$ は以下の式で求まる。

$$\cos^2 \theta = \frac{\sum_{i=1}^L (\text{Vect}, \text{Eigenvect}_i)^2}{|\text{Vect}|^2} \quad (4)$$

ここで L は部分空間を形成する固有ベクトルの数、 $(\text{Vect}, \text{Eigenvect}_i)$ は入力ベクトルと第 i 番目の固有ベクトルとの内積である。

各特徴点の部分空間を張る固有ベクトル (辞書画像 $d1$) を求めるために、さまざまな人物、照明条件、顔向きにおける左右の瞳、鼻孔、左右の口端の学習パターンを収集する。学習パターンは、分離度の局所最大点を中心に分離度フィルタの半径を基準に一定サイズの方角画像として切り出す。図 5 と図 6 に切り出された瞳と口端近傍の学習パターンの一部を示す。これらの学習パターンに対して主成分分析を適用して固有ベクトル (辞書画像 $d1$) を求める。しかし $d1$ のみでは、高い信頼度のパターン照合を期待できない。例えば、瞳の場合には、図 7 に示すような目尻、目頭、眉端などの特徴点に対しても類似度が高くなってしま



図7 非瞳領域画像の例：分離度フィルタが検出した特徴点候補の中で瞳領域に類似したパターン
Fig.7 Examples of image similar to pupil image.

う。これは部分空間法が、どの程度似ているかに基づいており、他のカテゴリとの関係を考慮していないことによる。そこで特徴点候補の中で辞書 $d1$ との類似度がしきい値より高い目尻、目頭、眉端などの誤りパターンを選別収集して“負の辞書 (以下、 $nd1$ と表記)”を生成した。最終的な類似度は、先に求めた類似度から負の類似度を差し引いた値となる。これにより約 3% の抽出率の向上が見られた。

2.4 検出位置の検証

正射影モデルを仮定すると、3次元モデルをもたずに、任意の方向から見た顔特徴点の2次元座標を、式 (5)、(6) に示すように異なる方向から見た4枚の画像上で対応づけられた特徴点の2次元座標の線形結合により表現できる [18]。この特性を利用して動画列列において既に抽出されている4枚のフレーム ($f = 1 \sim 4$) における正しい特徴点 i の2次元座標値 (x_{if}, y_{if}) から新しく検出された特徴点の2次元座標値 (\bar{x}_i, \bar{y}_i) を検証する。この関係は瞳、鼻孔などの特徴点 i について成り立つ。4フレームの画像に対する最低四つの特徴点 i の2次元座標と新しい座標 (\bar{x}_i, \bar{y}_i) から線形結合係数 a_i, b_i ($i = 1 \sim 4$) を最小2乗法により求める。次に得られた結合係数から近似座標を (x'_i, y'_i) を計算して (\bar{x}_i, \bar{y}_i) との誤差 Res を計算する。

$$\bar{x}_i = a_1 x_{i1} + a_2 x_{i2} + a_3 x_{i3} + a_4 x_{i4} \quad (5)$$

$$\bar{y}_i = b_1 y_{i1} + b_2 y_{i2} + b_3 y_{i3} + b_4 y_{i4} \quad (6)$$

$$Res = \sum_{i=1}^4 \sqrt{(\bar{x}_i - x'_i)^2 + (\bar{y}_i - y'_i)^2} \quad (7)$$

ここで、誤差 Res がしきい値より大きい場合は抽出位置に誤りが生じていることを示している。

3. 処理の流れ

本方法では、瞳、鼻孔、口端はそれぞれ単独で抽出可能であるが、ここでは特徴点がすべて見えている前提で処理の流れを図 8 に沿って説明する。

(1) 部分空間法による顔領域抽出

顔領域抽出には部分空間法 [2] を適用する。瞳と鼻

孔が含まれている 30×30 ピクセルの顔辞書パターン (以下、d2 と表記) を画像全体にわたって移動させながら類似度を逐次求めていき、類似度の局所最大点を顔領域として抽出する。これを入力画像を3段階にスケール変化させて行う。

(2) 特徴点候補の抽出

口端も抽出するために、抽出された顔領域の下部を30%拡大した領域に対して分離度フィルタを適用して顔特徴点候補を抽出する。分離度フィルタの適用範囲は、顔領域全体にp-タイル法 (比率を25%に設定) を適用して暗い領域に限定する。限定された領域の各画素に対して分離度フィルタの半径 r を3~6と変化させて出力の最大値をその画素に対する出力値、マスク半径 r を特徴点の半径とする。この後ガウシアン ($\sigma = 1.0$) を掛けて平滑化した後に局所最大点を求め特徴点の候補とする。

(3) 特徴点の空間配置による絞り込み

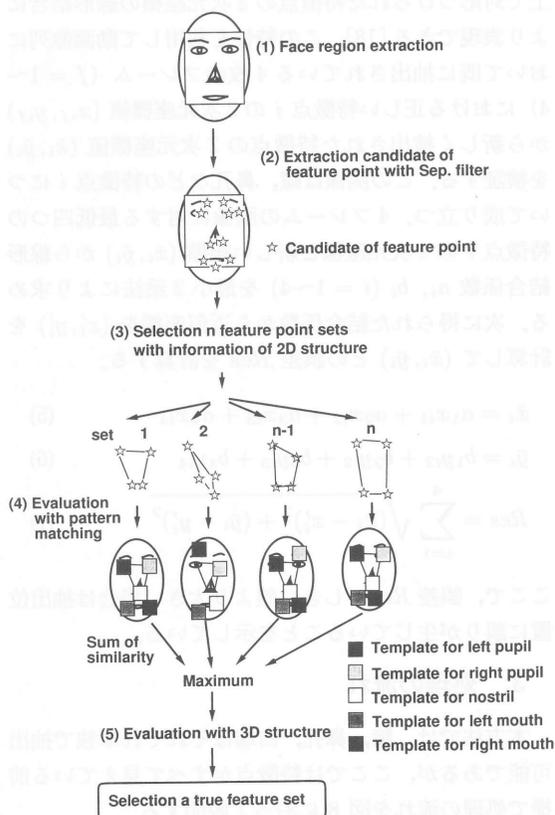


図8 特徴点抽出処理の流れ
Fig.8 Extraction process flow.

パターン照合の計算コストを下げるために、顔領域全体の大きかな制約により両瞳, 両鼻孔, 両口端からなる特徴点の組合せ候補を絞り込む。顔領域の幅 $L0$ から, 両瞳間, 両鼻孔間, 両口端間の長さ $c1L0, c2L0, c3L0$ を制限する。ここで $c1, c2, c3$ は実験的に決まる変数である。

(4) パターン情報による検証

特徴点の組合せに対してパターン照合を行って特徴点に対する類似度の合計を求める。最も類似度合計が高い特徴点の組合せを正しい組合せとする。

(5) 学習画像の選別

式 (7) から求まる誤差 Res が, しいき値より大きい場合は, サンプル画像から除外する。

4. 評価実験および考察

4.1 評価システム

評価実験システムは, 図9に示すようにSGI製ワークステーションIndy (CPU R5000 180 MHz), カラーCCDカメラ (東芝製IK-C40) からなる。CCDカメラヘッドは焦点距離6.5mmのレンズが装着されてモニタの下部に設置されている。カメラ入力された画像は, Indy内蔵の画像取込みボードにより360ピクセル \times 240ピクセルのサイズにデジタル化され, その後の顔領域抽出, 顔特徴点抽出, 顔切出しの一連の処理は, IndyのCPUによりソフトウェアで処理される。

評価画像として, 1人の人物のみが含まれている4種類の画像を準備した。データ1は, 計算機の指示に従って顔をディスプレイ上の左右, 上下の合計17方向に向ける様子を撮影した静止画像である。50人の対象者に対して照明条件の異なる居室と実験室でそれぞれ日時を変えて2回ずつ撮影を行い, 合計 $17 \times 50 \times 4 = 3400$ 枚を収集した。データ2は, データ1とはカメラ配置と照明条件を変えて収集した124枚の静止画像である。データ3はデータ1を収集する際に顔向きを変えていく様子を撮影した動画画像である。データ4は自然な状

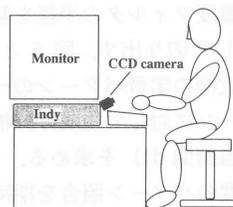
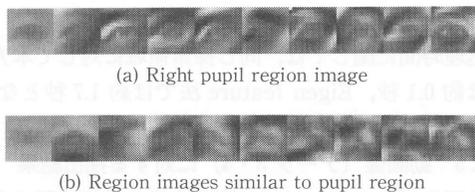


図9 実験システム (data1, 3, 4を収集)
Fig.9 Experimental system for data1, 3, 4.



(a) Right pupil region image

(b) Region images similar to pupil region

図 10 瞳抽出のための辞書画像

Fig.10 Reference images for pupil extraction.

態で計算機を操作中に撮影した顔の動画像である。

4.2 辞書生成

顔領域抽出の辞書 d_2 は、データ 1 (1 回目) の 1700 枚の画像から切り出した顔領域画像を用いて生成した。瞳、鼻孔、口端抽出の各辞書 d_1 , nd_1 も、同様にデータ 1 の 1700 枚から分離度フィルタにより切り出された特徴点近傍画像 (15 ピクセル \times 15 ピクセル) を用いて生成した。例として図 10 (a) に右瞳, (b) に非瞳領域に対する固有画像を固有値が高い方から順に 10 枚示す。なお各辞書で使用する固有ベクトル数は実験的に 10 に設定した。

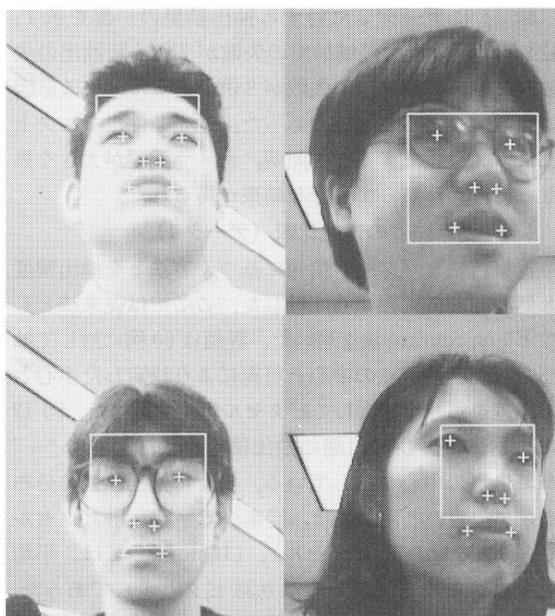
4.3 静止画像 (データ 1・2) に対する抽出結果

データ 1 の非学習データ 1700 枚に対して評価実験を行った。顔領域抽出は抽出された領域が瞳と鼻穴をすべて含む場合に成功と判断した。1700 例中、横顔、笑顔、首を傾けた顔などの 8 例に対して抽出に失敗した。

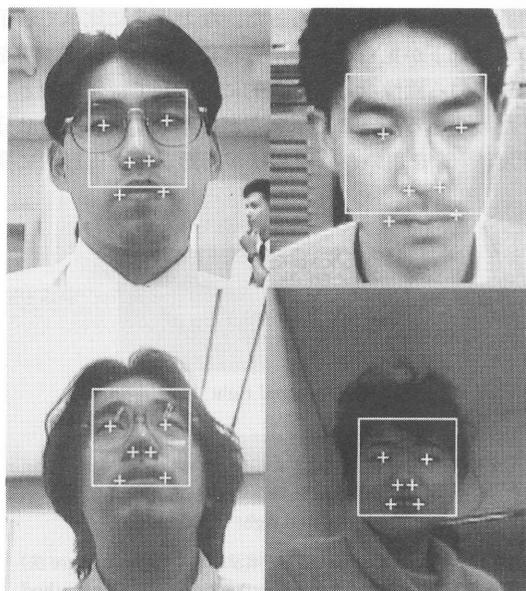
顔領域抽出が成功した 1692 例に対して瞳、鼻孔、口端の抽出を行った。図 11 (a) に抽出結果の一例を示す。図中の“+”が両瞳、両鼻孔、口端を示している。目視により抽出された特徴点が正しい特徴点の 1 ピクセル以内にある場合に抽出成功と判断した。瞳抽出の失敗は 3 例、抽出率は 99.8%、鼻孔抽出率は 100.0%、口端抽出は 14 例失敗で抽出率は 99.1%であった。

瞳抽出の失敗の原因は、メガネのレンズ反射により瞳領域がハイライト領域になったことが挙げられる。この問題は部分画像だけでは解決が難しく更に顔全体モデルを使う必要があると思われる。口端抽出の失敗の原因は、唇表面の反射により輝度値が高くなりモード法で決まる分離度フィルタの適用範囲に口端が入らなかったことや、笑ったために口端近傍のパターンが大きく変形し辞書との類似度が低下したことが挙げられる。口端に関しては、 p -タイル法の比率変更や、学習パターンを増やすことでかなり改善できると思われる。

しかし瞳、鼻孔、口端の 6 点のうちいずれかの 2 組 4 点が抽出されていれば顔切出しは可能である。こ



(a) Extracted feature points (data1)



(b) Extracted feature points (data2)

図 11 特徴点抽出結果

Fig.11 Extracted feature points.

の観点に立てば顔切出しの失敗は 1 例のみで、ディスプレイ下部にカメラを設置するという限定された状況であるが、本方法の有効性を十分確認することができた。

図 11 にデータ 2 に対する抽出結果の一部を示す。124 例に対して顔領域抽出の失敗が 7 例、瞳抽出の失敗が 2 例、口端抽出の失敗が 3 例であった。失敗原因はほぼデータ 1 と同様であった。逆光条件、モニタ上部にカメラ設置するなど照明、カメラ配置が大きく異なるにもかかわらず良好な結果が得られた。

4.4 従来法との性能比較 (データ 1)

提案方法とパターン情報に基づく方法とを抽出精度と処理速度の両面から比較した。ここでは従来法として Eigen feature 法を用いた。評価は 60 例に対して抽出された各特徴点の中心と目視により決定した中心との誤差を画像を拡大してピクセル単位で調べた。図 12 に、右瞳および右口端の抽出例を示す。図中で“+”が特徴点の中心を示している。表 1 に抽出誤差の分布を示す。瞳の抽出誤差は、提案法で平均 0.15 ピクセル (従来法では 0.77 ピクセル)、口端に関しては、平均 0.16 ピクセル (従来法では 0.33 ピクセル) であった。この抽出精度の差は、顔を横に向けた場合や視線を横にそらした場合に、特徴点近傍パターンが大きく変化するために特に顕著であった。なお 4.3 に比べて若干の精度低下が見られるが、これは比較のために、横を向いている画像などを意識的に選択して評価対象としたためである。ここで筆者らが提案している組合せ分離度法 [6] を用いれば瞳形状が円形からだ円形状に変



(a) Extracted right pupil



(b) Extracted right mouth edge

図 12 抽出結果の例 (左: 提案法, 右: Eigen feature 法)

Fig. 12 Results of feature extraction (left: Our method, right: Eigen feature method).

表 1 抽出誤差の比較結果 (a) 右瞳 (b) 右口端 (60 例)
Table 1 Distribution of extraction error (a) Right pupil.
(b) Right mouth edge (60 samples).

Error (pixel)	0	1	2	3	4	5	6
(a) Our method	76%	21%	-	3%	-	-	-
(a) Eigen feature	31%	26%	24%	12%	-	3%	2%
(b) Our method	88%	6%	6%	-	-	-	-
(b) Eigen feature	64%	23%	11%	-	-	-	-

形しても高い位置精度で抽出可能である。

処理時間に関しては、同じ探索領域に対して本方法では約 0.1 秒、Eigen feature 法では約 1.7 秒となり、大幅な計算量削減を実現できた。

4.5 動画像 (データ 3・4) に対する抽出結果

動画像データ 3 に対して瞳と鼻孔の抽出実験を行った。動画像に対する評価では、対象フレーム枚数が膨大になるために、目視による評価は困難である。そこで式 (7) より求める誤差 Res の変化を調べることにより正しく特徴点が抽出されているかを判断した。図 13 は Res の推移の一部を示している。図中で正常に抽出されている場合には誤差が小さくなっているが、まばたきなどにより瞳が抽出できずに誤って眉を抽出した場合には A のように誤差が大きくなる。これを利用して誤差 Res がしきい値 (実験的に 80 に設定) より大きいフレームは処理結果を蓄えておいて後で目視による評価を行った。約 10 フレーム/秒の速度で処理された 9880 フレームに対して顔領域抽出に成功したフレームは 9786 枚 (抽出率は 99.0%)、この 9786 枚のフレームに対して 9675 枚で 4 点すべての特徴点抽出の成功し抽出率は 98.8% であった。

更にデータ 4 を用いて顔切出しから辞書生成完了までの時間を調べた。225 枚の学習パターン (15 ピクセル \times 15 ピクセル) の切出しから辞書生成完了までの 5 回の平均処理時間は 45 秒であった。図 14 は 0.25 秒ごとの抽出結果を示す。この処理時間は実時間での辞書生成や顔認識の実現可能性を示している。しかし、無意識の動作を対象としているため収集した学習パターンに偏りが見られると考えられる。より少ない学習パターンからいかに高精度の辞書を生成するかが今後の課題である。これに対しては、いくつかの視点か

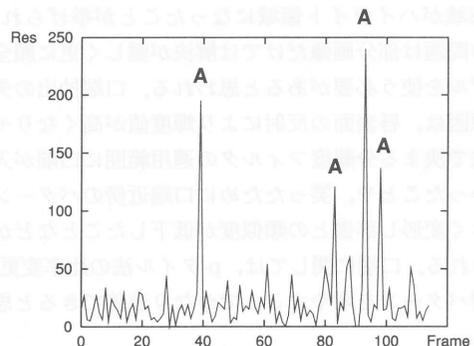


図 13 誤差 Res の変化
Fig. 13 Change of Res .

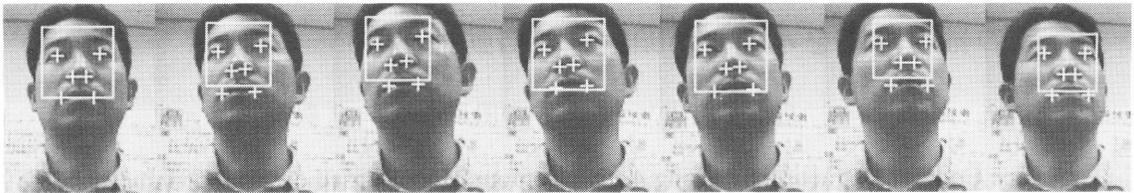


図14 動画画像からの抽出結果 (frame1~7)
Fig.14 Results of feature extraction from sequence images (frame1~7).

らの学習顔画像から任意の視点の顔画像を生成する方法 [7],[18] が適用可能であると思われる。

5. むすび

画像から顔領域を抽出し瞳、鼻孔、口端などの顔特徴点を高速かつ高精度に抽出する方法を提案した。提案した方法では分離度フィルタによる形状情報と部分空間法によるパターン照合の組合せにより、少ない計算量で高い位置精度を実現している。計算機ディスプレイ下に設置されたカメラから得られた静止画像に対して正規化顔画像を抽出し可能な少なくとも4点の特徴点が得られる率は99.5%であった。更に動画像に対しても、98.8%という高い抽出率となった。今後は動画像を用いた人物認証に適用を進めていく。

文 献

- [1] 赤松 茂, 佐々木努, 深町映夫, 末永康仁, “濃淡画像マッピングによるロバストな正面顔の識別法,” 信学論 (D-II), vol.J76-D-II, no.7, pp.1363-1373, July 1993.
- [2] 小松良江, 有木康雄, “部分空間法を用いた向きによらない顔の切り出しと認識,” 信学技報, PRU95-191, 1996.
- [3] 村瀬 洋, シェリー ナイヤー, “2次元照合による3次元物体認識,” 信学論 (D-II), vol.J77-D-II, no.11, pp.2179-2187, Nov. 1994.
- [4] 佐々木努, 赤松 茂, 末永康仁, “顔画像認識のための色情報を用いた顔の位置合わせ法,” 信学技報, IE91-2, 1991.
- [5] 福井和広, “領域間の分離度を用いた輪郭抽出,” 情処研報, CV88-2, pp.9-16, 1994.
- [6] 山口 修, 福井和広, “分離度特徴を用いた顔画像解析一目瞳の検出,” 情処学会, 第52回全国大会 (2), pp.187-188, 1996.
- [7] 向川康博, 中村裕一, 大田友一, “2枚の顔写真を用いた任意方向の顔画像の生成,” 情処論, vol.37, no.4, pp.635-644, April 1996.
- [8] 宋 欣光, 李 七雨, 徐 剛, 辻 三郎, “部分特徴テンプレートとグローバル制約による顔器官特徴の抽出,” 信学論 (D-II), vol.J77-D-II, no.8, pp.1601-1609, Aug. 1994.
- [9] 坂本静生, 宮尾陽子, 田島譲二, “顔画像からの目の特徴点抽出,” 信学論 (D-II), vol.J76-D-II, no.8, pp.1796-1804, Aug. 1993.
- [10] エルッキ オヤ著, 小川英光, 佐藤 誠訳, “パターン認識と部分空間法,” 産業図書, 1986.
- [11] M. Turk and A.P. Pentland, “Face recognition using eigenfaces,” Proc. CVPR 11, pp.453-458, 1993.
- [12] A. Pentland, B. Moghaddam, and T. Starner, “View-based and modular eigenspaces for face recognition,” CVPR '94, pp.84-91, 1994.
- [13] R. Brunelli and T. Poggio, “Face recognition: Features versus templates,” IEEE PAMI, vol.15, no.10, pp.1042-1052, 1993.
- [14] A. Gee and R. Cipolla, “Adaptive visual tracking by temporal consensus,” ACCV '95, I, pp.44-48, Singapore, Dec. 1995.
- [15] K. Fukui, “Edge extraction method based on separability of image features,” IEICE Trans. on Information and Systems, vol.E78-D, no.12, 1995.
- [16] A.L. Yuille, P.W. Hallinan, and D.S. Cohen, “Feature extraction from faces using deformable templates,” IJCV, vol.8, no.2, pp.99-111, 1992.
- [17] H.P. Graf, T. Chen, E. Petajan, and E. Cosatto, “Locating faces and facial parts,” International Workshop an Automatic Face- and Gesture- Recognition, pp.41-46, 1995.
- [18] S. Ullman and R. Basri, “Recognition by linear combinations of models,” IEEE Trans. PAMI, vol.13, no.10, pp.992-1006, 1991.

(平成8年12月25日受付, 9年3月14日再受付)



福井 和広 (正員)

昭61九大・工・機械卒。昭63同大学院工学研究科了。同年(株)東芝入社。現在、関西研究所勤務。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。



山口 修 (正員)

平4岡山大・工・情報卒。平6同大学院工学研究科了。同年(株)東芝入社。現在、同社関西研究所勤務。画像認識、ヒューマンインタフェースの研究開発に従事。