色情報および動き情報に基づく人物顔領域と顔部品の抽出 およびその動きの検出

非会員 志 賀 泰 博 (工学院大学)

非会員 海老根 秀 之 (工学院大学)

非会員 池 田 美 鈴 (工学院大学)

非会員 中 村 納 (工学院大学)

Extraction of facial area and facial parts based on color and motion information and detection of its movements

Yasuhiro Shiga, Non-member, Hideyuki Ebine, Non-member, Misuzu Ikeda, Non-member, Osamu Nakamura, Non-member (Kogakuin University)

In this paper, a new algorithm for extracting a human face and facial parts from images based on skin color and motion information, and the extracting method of motion information of the eye and mouth regions are presented. For the efficient extraction of the face area, two cases are assumed. If the moving area is large, the facial area is extracted based on moving area and color information. On the contrary, it is judged that no movement exists when the moving area is small, the facial area is extracted based on the facial area which has already been extracted in the previous frame. For precise recognition of facial expressions, we propose a new scheme for extracting an opening and closing ratio of the eyes and mouth, and the efficacy of the proposed method is examined experimentally using several images.

キーワード：顔領域抽出，顔部品抽出，特徴抽出，表情認識，修正 HSV 表色系

1. はじめに

本論文では、動画画像からの顔領域および顔部品の抽出法および表情認識のための有効な特徴量の抽出法について述べる。画像中からの顔領域の抽出は、個人認証および表情認識などの分野で非常に重要であり、盛んに研究が行なわれている⁽¹⁾。従来手法では色情報を基にして顔領域の抽出を行っていたが、背景に肌色と同等な色の物体が存在する場合に正確な抽出ができない場合や処理時間に問題があった⁽¹⁰⁾。顔領域の抽出に関して、肌色情報の他に動き情報を利用する手法が提案されており⁽⁴⁾⁽⁵⁾ 安定に顔領域を抽出できるとの報告がされている。提案手法では色情報の他に画像中の人物の動きも考慮することとした。

表情を推定するための特徴抽出に関する報告としては、周波数変換やオプティカルフローを用いたもの⁽⁸⁾⁽⁹⁾などが報告されている。これらの報告では、全ての顔部位から同一の手法で特徴を抽出しているが、表情表出に伴う顔表面の二次元的な色分布の変化や三次元的な形状変化を捉える上では必ずしも全ての顔部位に適している手法とは言えない。提案手法では、それぞれの部位に適したアルゴリズム

を用いて特徴量を抽出することとしている。目領域では、2つの色情報からその動きを検出する。口領域の開口度については、上下左右4つの特徴点の周辺分布をフレーム毎に追跡することで捉える。

2. 色 情 報

本手法では、色情報に基づいて顔画像からの顔領域と顔部品の抽出、および顔部品の動きを検出する。

〈2・1〉 修正 HSV 表色系 顔領域の抽出には、修正 HSV 表色系を用いる。RGB 表色系の入力画像から修正 HSV 表色系への変換式を式(1)に示す。ここで、 H は色の属性を表す色相(図1(b))、 Q_c は色の鮮やかさを表す修正彩度(図1(c))、 I は色の明るさを表す強度(図1(d))である。ただし、これらの画像では色相値、彩度値、強度値はそれぞれ256階調の濃淡値で表示されている。

一般に人物肌色は、人種および民族の多様性に関わらず色相にはあまり大きな違いは見られないという報告がされている⁽²⁾。また、人物肌色の彩度は背景に比べ比較的高いという特徴がある。そこで、本手法では修正 HSV 表色系の修正彩度成分 Q_c および色相成分 H の2つの色情報を適

用して肌色領域を抽出する。

修正彩度成分は人物肌色を強調できるだけでなく、黒に近い色を低彩度として表すことが可能なため目領域の抽出処理にも利用することとした。

$$H = \cos^{-1} \left\{ \frac{(R - G) + (R - B)}{2\sqrt{(R - G)^2 + (R - B)(G - B)}} \right\}$$

$$Qc = \sqrt{\left(\frac{2R - G - B}{2} \right)^2 + \left(\frac{\sqrt{3}(G - B)}{2} \right)^2} \quad (1)$$

$$I = \frac{R + G + B}{3}$$

図1 修正 HSV 表色系
Fig. 1. Modified HSV color system.

図2 YIQ 基底
Fig. 2. YIQ basis.

〈2・2〉YIQ 基底 唇領域の抽出には、YIQ 基底の Q および Y 成分を用いる。肌色領域内の Q および Y 成分の分布を図2に示す。ただし、これらの画像では Q および Y 成分はそれぞれ 256 階調の濃淡値で表示されている。ま

た、 Q 成分はそのダイナミックレンジが狭いため、図2(a)ではその最大が 255、最小が 0 になるよう線形変換したものである。 Q 成分は顔領域内においては唇領域でピーク値を示すことが知られている⁽⁶⁾。一方、唇は肌に比べ暗い色をもつことから、図2(b)の Y 成分で見た場合その値が低くなる。

$$Q = 0.21R - 0.52G + 0.31B \quad (2)$$

$$Y = 0.30R + 0.59G + 0.11B \quad (3)$$

本論文ではこれらの特徴的な分布特性の違いに基づいて唇領域を抽出する。

3. 顔領域と顔部品の抽出

〈3・1〉画像入力 入力画像(図1(a))は一般的な屋内環境から横 256 × 縦 240 画素の RGB 各色 256 階調のカラー画像で面光源と CCD カメラを用いて動画像として入力される。

〈3・2〉顔領域抽出の概要 顔領域の抽出は次の3つの段から成る。

- (1) 頭部の動きの判定
- (2) 動きがある場合の抽出
- (3) 動きが無い場合の抽出

〈3・2・1〉頭部の動きの有無の判定 3枚の連続する入力画像 (f_{t-1} , f_t , f_{t+1}) のフレーム間差分を用いて変化領域を検出し、その変化領域の面積に基づいて動きの有無を判定する。面積が大きければ動きはあり、また小さければ動きは無しと判定する。しかしながら、その面積は画面に含まれる人物領域の大きさによって大きく異なる。従って、この差分量をそのまま判定の基準に用いた場合、画面に占める人物領域の割合に従った閾値を全て予め設定しておく必要があり、実際には極めて困難である。そこで本手法では画像中の人物領域の大きさが閾値に影響を与えないように、前フレーム、即ち f_{t-1} で抽出された顔領域の面積(画素数)を基準として用い、動きのある領域(画素数)をこの顔領域の面積に対する相対的な比として表すこととした。

具体的手順は次の通りである。

- (1) 3枚の入力画像 (f_{t-1} , f_t , f_{t+1}) から2枚の差分画像 ($S_1(f_t - f_{t-1})$, $S_2(f_{t+1} - f_t)$) を作成する。
- (2) 差分画像に含まれる光源のちらつき等に起因する雑音を除去するために、閾値 T_z として2値化処理を施す(2値化後のそれぞれの画像を S_{b1} , S_{b2} とする)。本手法では予備実験に基づき $T_z = 5$ とした。
- (3) S_{b1} , S_{b2} の面積 R_{s1} および R_{s2} を求める。
- (4) 既に抽出されている前フレーム f_{t-1} の顔領域 K_{t-1} の面積 R_{t-1} を求める。
- (5) 次の不等式が満たされるとき、動きありと判定

する。

$$\left(\frac{R_{s\lambda}}{R_{t-1}} \times 100 \right) \geq T_s \quad (\lambda = 1, 2) \quad (4)$$

本手法では閾値 T_s を 40 と設定した。 T_s の値は意図的に動くように指示した被験者 2 名の動画像を用いた予備実験から設定した。図 4 の横軸はフレーム番号、縦軸は前フレームで切り出した顔領域の大きさに対する現フレームでの動きのある領域の割合 (式 (4) の左辺) である。“▲” の付いたフレームでは大きく顔領域に欠落が確認された。2 名の被験者共に、この割合が 40% 程度の閾値で顔領域に欠落なく抽出が行われたため、この値を閾値として設定した。

図 3 人物候補領域

Fig. 3. Candidate areas of a person.

図 4 変化領域の割合と抽出結果

Fig. 4. Ratio of differential areas and extraction results.

〈3・2・2〉動きがある場合 S_{b1} と S_{b2} の共通領域、すなわち人物候補領域 (図 3) を抽出した後、2 つの色情報を基に肌色領域を抽出する。人物肌色の色相値はある範囲内に分布するという報告⁽²⁾⁽³⁾に基づき、一般的な屋内環境下で予備実験として JIS 標準色票の 10R~7.5YR の範囲の色相値を測定したところ、人物肌色での色相値の範囲はほぼ 6~31 間と確認された。また、顔領域は屋内環境の場合、背景に比べ比較的彩度が高いという特徴がある。そこで、本手法では検出した人物候補領域に対して修正彩度成分 Q_c および色相成分 H の 2 つの色情報を適用して肌色領域を抽出する。

顔領域抽出の処理手順は次の通りである。

- (1) 2 枚の差分画像 S_{b1}, S_{b2} の論理積

図 5 肌色領域の抽出

Fig. 5. Extraction of skin area.

図 6 修正彩度成分のヒストグラム

Fig. 6. The histogram of Quasi-chroma.

$$F_L = S_{b1} \cap S_{b2} \quad (5)$$

を人物候補領域 (図 3) として検出する。

- (2) 画像 f_t における色相画像 (図 5(a)) の各画素のうち、色相値が 6~31 の範囲の画素を抽出する (図 5(b))。
- (3) 上記処理 (2) で検出された画素から修正彩度成分のヒストグラムを作成し、肌色領域抽出のための閾値 T_c を探索する (図 6)。

色相成分で限定された範囲の修正彩度成分のヒストグラムの形状は、背景に含まれる彩度値の低い

肌色に近い色と高い肌色領域(斜線部)の双峰性を示すと考えられるため、 T_c を2つのピーク値の間の谷に設定した。

- (4) 人物候補領域内の画素と閾値 T_c 以上の修正彩度値を持つ画素を統合して肌色候補領域を求める(図 5(e))。
- (5) 肌色候補領域の欠落した領域を黒画素で埋め、 5×5 画素のマスクを用いた拡大・縮退処理を施すことにより輪郭付近を平滑化する。
- (6) 最大領域を肌色領域として抽出する(図 7(a))。
- (7) 唇の中心を基準として楕円を用いて首領域の除去を行なう。

唇領域の抽出法は後の〈3・3〉で詳しく説明する。

- (8) 最後に首領域を除去した顔領域に対して、楕円を用いて頭部の補正を行なう(図 7(b))。

この処理により、髪型による眉や目の欠落を補正することができる。ただし、楕円の中心点は首領域を除去した顔領域の中心とした。

図 7 顔領域の抽出

Fig. 7. Extraction of facial area.

〈3・2・3〉動きが無い場合 頭部の動きが少ない場合、現フレームの顔領域の位置は、前フレームで切り出された顔領域とほぼ同じ位置であると判断できる。また、動きの少ない場合は切り出される領域が若干ずれていても正確に顔部位が抽出されていれば、表情の認識には影響ない。従って、本手法ではあらかじめ切り出されている前フレームの顔領域と現フレームの原画像を対応させ、2枚の画像の論理積から顔領域を抽出する。

〈3・3〉唇領域の抽出 顔領域の正規化および表情の認識には特に必要のない首領域の除去を行なう際に基準となる唇領域を抽出する。

唇領域の抽出手順は次の通りである。

- (1) 首を含む顔領域内(図 7(a))の Q 成分と Y 成分においてその最大が 255、最小が 0 になるように線形変換し、 $F_Q(x, y)$ (図 2(a)) および $F_Y(x, y)$ (図 2(b)) を作成する。
- (2) $F_Q(x, y)$ から $F_Y(x, y)$ を引き、さらに次のような条件を満足する差分画像 $F_{Q-Y}(x, y)$ (図 8(a)) を作成する。

$$F_{Q-Y}(x, y) = \begin{cases} F_d(x, y) : F_d(x, y) > 0 \\ 0 : F_d(x, y) \leq 0 \end{cases} \quad (6)$$

$$F_d(x, y) = F'_Q(x, y) - F_Y(x, y) \quad \dots\dots\dots (7)$$

$$F'_Q(x, y) = \begin{cases} F_Q(x, y) : F_Q(x, y) \geq T_q \\ 0 : F_Q(x, y) < T_q \end{cases} \quad (8)$$

顔領域内で唇領域の Q 成分の値はピーク値をもつこと⁽⁶⁾ および顔領域内におけるほとんどの領域である肌色部は唇領域に比べ低い値を示すことから、閾値 T_q は Q 成分のヒストグラムのピーク値と最大値 255 の間の平均値として設定した。

- (3) 最後に唇領域の大きさを考慮して、 $F_{Q-Y}(x, y)$ を閾値 1 で 2 値化した後、面積が最大の領域から 3 つの領域を唇候補領域として抽出し、その中で平均の Q 値が最大の領域を唇領域として抽出する(図 8(b))。

図 8 唇領域の抽出

Fig. 8. Extraction of the lip area.

〈3・4〉目領域の抽出 唇領域と同じように、肌色領域に比べ目領域の Y 成分は低い。この処理では、唇領域の抽出の際に使用した差分画像 $F_{Q-Y}(x, y)$ を再度用いる。さらに、図 9(a) に示すように修正彩度成分も同じように目領域で低い値を示すため、この 2 つの画像の特徴を用いて目領域を抽出する。

目領域の抽出手順は次の通りである。

- (1) 顔領域内(図 7(b))の修正彩度成分の最大が 255、最小が 0 になるように線形変換し、 $F_{Q_C}(x, y)$ (図 9(a)) を作成する。
- (2) $F_{Q_C}(x, y)$ に対して、閾値 T_{qc} として 2 値化処理を行ない、この画像に対して 5×5 画素の拡大・縮退処理を施し欠落部を穴埋めする(図 9(b))。
- (3) 閾値 1 で差分画像 $F_{Q-Y}(x, y)$ を 2 値化する(図 9(c))。
- (4) 図 9(d) に示すように、 $F'_{Q_C}(x, y)$ (図 9(b)) と

$F'_{Q-I}(x, y)$ (図 9(c)) の共通領域を検出することにより、目の候補領域が得られる。

- (5) 唇領域の中心を原点とし、式 (9) を満たす領域を目領域探索範囲として設定し、下から探索して最初に見つかった目領域候補を目領域として検出する (図 9(e))。

$$y = ax \begin{cases} 3/2 \leq a \leq 3 & (x \leq p_r) \\ -3 \leq a \leq -3/2 & (x \geq p_l) \end{cases} \dots (9)$$

ここで、 p_r と p_l はそれぞれ唇領域の左右の端点の x 座標である。

図 9 目領域の抽出

Fig. 9. Extraction of eyes area.

〈3・5〉 顔領域の正規化 顔部位から正確に表情変化を捉えるためには、個人間による顔の大きさや顔部品の相対的な位置関係の違いを整える必要がある。そこで、本手法では両目の中心、鼻の頂点および口の両端を参照点として、入力画像を平均的な顔画像、すなわち基準顔へアフィン変換することにより、これらの違いを除去することとした⁽¹¹⁾。基準となる座標 (図 10) は、100 人の平常時の顔画像の平均を取った画像の両目の中心、鼻の頂点および口の両端の座標値から求めたものである。アフィン変換を実施する際に、表情が表出している顔画像では、表情の表出の

影響で大きく口が動く場合があるため、口からの抽出する参照点は初期フレーム (平常時の顔画像) で抽出した点を使用する。

手順は次の通りである。

- (1) 初期フレームと対象フレームの顔画像から両目の中心と鼻の頂点の 3 点の座標を特徴点として抽出する。
- (2) これらの座標が基準座標へ移動するように初期フレームおよび対象フレームの顔画像をそれぞれアフィン変換する。
- (3) 上記 (2) で変換された顔画像内の、初期フレームの顔画像から新たに両目の中心、鼻の頂点、口両端の 5 点の座標を特徴点として抽出する。
- (4) これら 5 点の座標が基準座標へ移動するように上記 (2) で変換された初期フレームと対象フレームの顔画像をさらにアフィン変換する。

目と口の特徴点の抽出には分離度フィルタ⁽⁷⁾を用いている。分離度フィルタは判別分析法に基づき 2 つの領域の濃度値の分散比を求めるもので、この分散比が最大になった点の特徴点として抽出するものである。ただし、目の中心は目尻と目頭の座標の平均の値としている。鼻の頂点に関しては他の点に比べ明るい点であり、鼻の下は暗いと考えられる。そこで、唇領域の重心と両目の目頭の中心を結ぶ線分上の濃度分布から判別分析法に基づいて鼻の頂点を抽出する。

図 10 基準座標

Fig. 10. Standard face.

4. 顔部品からの動きの検出

表情を認識するために、目および口領域から特徴量を抽出する。一般に、単一の手法で全ての顔部位から認識に必要な特徴を統一的に抽出することは極めて困難である。本論文では、それぞれの顔部位に適した手法を用いて特徴量を抽出することとした。ただし、特徴量を抽出する際に基準となる初期フレームは平常時の顔画像である。

〈4・1〉 目領域の特徴量 提案手法では顔領域を抽出する際に使用した修正彩度成分と強度成分を利用して目領域の動きを検出することとした。修正彩度画像は HSV 表色系の彩度画像と比べ肌部を強調し、黒目やまつ毛等の黒

に近い色の領域を低彩度で表すことができる(図 11(b))。一方、強度画像は輝度画像に似た分布をもつ。ただし、図 11(a) に示すように、目領域の特徴抽出部では輝度値で表示している。従って、修正彩度画像から強度画像を差し引くことによって得られる負の値をもつ領域は、黒に近い色を有する領域と考えられる。この差分量の増減に着目して目領域の開口度を検出する。

図 11 目領域

Fig. 11. Eye region.

具体的手順は次の通りである。

- (1) 図 11(a),(b) に示す修正彩度成分および強度成分の最大が 255, 最小が 0 になるよう線形変換する。
- (2) 注目フレーム t での修正彩度画像と強度画像の差分量 M_t を

$$M_t = \frac{1}{S} \sum_{y=Y_1}^{Y_2} \sum_{x=X_1}^{X_2} \begin{cases} |F(x,y)| & (F(x,y) < 0) \\ 0 & (F(x,y) \geq 0) \end{cases} \quad (10)$$

として求める。ここで、

$Q_c(x,y)$: (x,y) における修正彩度値
 $I(x,y)$: (x,y) における強度値
 $F(x,y)$: $Q_c(x,y) - I(x,y)$
 $X_1, X_2; Y_1, Y_2$: 目領域の固定ウインドウ内の座標
 S : ウインドウ内の総画素数
 である。

- (3) 対象フレームでの開口度 M_e は次式で算出する。

$$M_e = M_t - M_i \quad (11)$$

ただし、 M_i は初期フレーム(平常時の開眼状態の目領域)での修正彩度画像と強度画像の差分量である。 M_e は負の値を示せば目が閉じられたことを表し、逆に正の値を示せば驚きの表情のように目が大きく開かれていることを表す。

〈4・2〉 口領域の特徴量 口領域の開口度は、唇領域から求められた 4 つの特徴点の周辺分布を追跡することによって求められる。

具体的手順は次の通りである。

- (1) 前記 3・3) で抽出された初期フレーム i の唇領域(図 8(b)) から 4 つの特徴点 $h_1(i), h_2(i), v_1(i), v_2(i)$ を抽出する(図 12(a))。

ここで、初期フレームは平常時の口が閉じられた画像である。

- (2) 得られた各特徴点に対し、図 12(b), (c) に示すようなウインドウを設定する。

表情表出により口が大きく変化した場合、図 12(a) では上唇が欠落する場合があるため、このことを考慮してウインドウは差分画像 $F_{Q-I}(x,y)$ に対して設定することとした。

- (3) 初期フレーム i から特徴点周辺部のウインドウ $S_W(i)$ 内の面積を求める(フレーム t でのウインドウを図 13 に示す)。
- (4) 次フレーム $i+1$ において、4 つのウインドウを左右、上下にそれぞれ移動し、ウインドウ $S_W(i+1)$ に含まれる特徴点周辺部の分布を同様に検出する。
- (5) 新しい特徴点 $h_1(i+1), h_2(i+1), v_1(i+1), v_2(i+1)$ を探索する。

ただし、フレーム $i+1$ での特徴点の位置 $p^*(x)$ は

$$p^*(x) = \{p(x) | S_W(i+1) = S_W(i)\} \quad (12)$$

を満足する点とする。

- (6) 上記 (2)~(5) までの処理を繰り返し、対象フレームまでの動きを検出する。
- (7) 最後に、口領域の水平および垂直方向への初期フレームからの動き量 V_q, H_q を次式で算出する。

$$V_q = (v_{t,2} - v_{t,1}) - (v_{i,2} - v_{i,1}) \quad (13)$$

$$H_q = (h_{t,2} - h_{t,1}) - (h_{i,2} - h_{i,1}) \quad (14)$$

ここで、

$v_{t,2} - v_{t,1}$: 対象フレームでの $v_1 \sim v_2$ の距離

$v_{i,2} - v_{i,1}$: 初期フレームでの $v_1 \sim v_2$ の距離

$h_{t,2} - h_{t,1}$: 対象フレームでの $h_1 \sim h_2$ の距離

$h_{i,2} - h_{i,1}$: 初期フレームでの $h_1 \sim h_2$ の距離

とする。

5. 実験結果

〈5・1〉 顔領域抽出実験 提案手法の有効性を評価するため、被験者 22 名(男性 18 名 女性 4 名、眼鏡無し)の画像(総フレーム数 4912 枚)に対して顔領域の抽出実験を行った。実験に用いた画像は、図 1 のような人物が 1 人映っており、背景には動物体がないものを使用した。抽出実験の評価は、顔領域を抽出できているもの(図 14(a))、顔領域に欠落や他の領域が含まれているもの(図 14(b))、顔領域以外の領域が顔領域として抽出されたもの(失敗)の 3 つに分けて行なった。結果を表 1 に示す。実験結果から、顔以外の領域が抽出された画像はなく、良好に抽出されることがわかる。正確に抽出できていないものが 4.3% あったが、その主な原因は顔領域とともに髪の毛の領域(茶色系の

図 12 口領域

Fig. 12. Mouth region.

図 13 フレーム間での特徴点の追跡

Fig. 13. Feature point tracking between successive frames.

髪の部分)も抽出されてしまい、頬ではなく髪が首領域を除去する際の楕円の端点と判断されたためと考えられる。また、髪などの領域を含んで顔領域が抽出され、かつ動き無しと判定された場合、前フレームを基に顔領域の抽出を行なうため以後のフレームでも髪などの領域が依然として残ってしまうことがあった。この点に関しては、何フレームかおきにリフレッシュ機能をつけることで改善できると考えられる。

表 1 顔領域抽出結果

Table 1. Results of extraction of facial area.

Success	Success (including hair)	Failure
4702/4912 (95.7%)	210/4912 (4.3%)	0/4912 (0.0%)

処理時間に関しては、Pentium II 400MHz 搭載の PC で色情報のみで抽出していた従来手法では 2.33sec/frame であったのに対して、提案手法では動き有りと判定されたフレームで平均 0.53sec/frame、動き無しと判定された場

図 14 顔領域の抽出例

Fig. 14. Examples of extraction of facial area.

合は 0.14sec/frame と高速に顔領域の抽出を行うことができた。従来手法では人物の動きを考慮していなかったことや等輝度領域から得られた等濃線を用いて首領域を削除していたことにより、計算コストの増大を招いていた。これらを考慮した結果、本手法では大幅に処理時間を短縮することができたと考えられる。顔領域の抽出精度については従来手法でも同様の動画像に対して実験を行なったところ、顔領域の正抽出率は 92.7%(4553/4912)、顔領域に大きく欠落があったものが 7.3%(359/4912)、顔領域以外の領域が顔領域として抽出されたものが 0.0%(0/4912)であった。顔領域の欠落の主な原因は、図 15(a)のように前髪が眉にかかることによるものであった。眉の欠落は表情を認識するための重要な手掛かりを減らす他、顔領域の欠落は顔認識を行なう際に悪影響をおよぼすと考えられる。提案手法では顔の形状を考慮することにより、顔領域の欠落を補正することが可能となった(図 15(b))。

図 15 従来手法との比較

Fig. 15. Comparison with previous method.

〈5・2〉顔部品抽出実験 顔領域の抽出を行なった後、唇および目領域の検出実験を行なった。この際、上記の実験で髪や首を含む図 14(b)のような画像に対しても顔部品は対象の画像内に含まれているため実験に用いた。目領域に関しては、両目が抽出できたものを成功とした。結果を表 2 に示す。

提案手法では唇領域の抽出の後に目領域を抽出しているため、唇領域を誤抽出したほとんどの画像に対して目領域を誤抽出してしまった。その結果、唇領域の抽出率に比べ

表 2 顔部品抽出結果
Table 2. Results of extraction of facial parts.

eyes	lip
4768/4912 (97.1%)	4904/4912 (99.8%)

て目領域の抽出率が若干低くなっている。
〈5・3〉特徴抽出実験 提案した手法を用いて、被験者 10 名に対し、瞬きの検出実験を行なった。例として図 16 に瞬き時の目領域の変化を示す。この被験者は 3 回の瞬きを行なったが、その動きを良好に捉えている。口領域では表情表出に伴って大きく動くと考えられる驚きおよび幸福の表情に関して、被験者 10 名に対して実験を行なった。例として図 17 に驚きの表情時の垂直方向の口領域の変化を示す。驚きの表情では口が縦方向に開かれることによる垂直方向への開口度が増加する様子が捉えられている。また、喜びの表情 (図 18) では水平方向への開口度が増加する特徴が捉えられた。

実験に用いた画像の初期フレームでは確認されなかったものの、顔全体が赤みを帯びている被験者は図 19 に示すように唇以外の領域も抽出されてしまう画像も確認された。特に、赤みを帯びた肌部が唇の上部に付着しているものが確認された。このような画像において口領域の位置を決定するには問題はないが、初期フレームでの上唇における特徴点が大きくずれてしまう可能性が考えられる。その結果、後のフレームにおける特徴量の抽出精度にも影響を与える可能性があるため、特徴点の抽出精度の向上を図ることが今後必要であると考えられる。

図 16 目領域の時系列変化 (瞬き)
Fig. 16. Movements of eye region in time series (blink).

6. ま と め

本論文では、顔領域および顔部品の抽出法と目と口領域からの特徴抽出手法について述べた。実験では、提案手法を用いて良好な結果を得ることができた。今後は頬領域での特徴量等⁽¹²⁾を考慮した表情認識システムの構築と提案手法を更に改良し、眼鏡、人種、髪などに対する有効性を確認する予定である。

(平成 12 年 7 月 4 日受付, 同 13 年 1 月 4 日再受付)

図 17 口領域の時系列変化 (驚きの表情)
Fig. 17. Movements of mouth region in time series (surprise).

図 18 口領域の時系列変化 (幸福の表情)
Fig. 18. Movements of mouth region in time series (happiness).

図 19 初期フレームでの唇領域
Fig. 19. Lip area in first frame.

文 献

(1) 金子正秀, 長谷川修, “顔画像処理技術とその応用に関する最近の研究動向”, 電学大, MC4-1, pp.491-494, 1999.
(2) 池田光男, “色相工学の基礎”, p.14, 朝倉書店.
(3) 田口卯三郎, “色彩学”, p.128-129, 共立出版.
(4) 浅沼克紀, 大西正輝, 小島篤博, 福永邦雄: “色情報と領域追跡情報を用いた人物の顔と手の領域の抽出”, 電学論 C, Vol.119-C, No.11, pp.1351-1358, 1999.
(5) 本郷仁志, 山本和彦, “動領域内の肌色推定による顔領域および顔部品抽出”, 映像学誌, Vol.52, No.12, pp.1840-1847, 1998.
(6) 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, pp.9-15, 1991.
(7) 福井和広, 山口修, “形状抽出とパターン照合の組み合わせによる顔特徴点抽出”, 信学論, D-II, Vol.80-D-II, No.8, pp.2170-2177, 1997.
(8) 坂口竜己, 森島繁生, “画像の 2 次元離散コサイン変換を利用した実時間顔表情認識”, 信学論, D-II, Vol.80-D-II, No.6, pp.1547-1554, 1997.

-
- (9) ズデネク ブロハースカ, 伊藤崇之, 岡本敏雄, “動き情報を用いた表情特徴の自動抽出手法”, 信学論, D-II, Vol.J81-D-II, No.6, pp.1150-1159, 1998.
  - (10) 松橋聡, 藤本研司, 中村納, 南敏, “顔領域抽出に有効な修正HSV表色系の提案”, テレビ学誌, Vol.49, No.6, pp.787-797, 1995.
  - (11) 海老根秀之, 中村納, “顔部位の個人的差異を排除した顔表情の認識”, 電学論 C, Vol.119-C, No.4, pp.474-481, 1999.
  - (12) 志賀泰博, 海老根秀之, 中村納, “表情認識のための顔器官および表面からの特徴量の抽出”, 情処研, HI84-17, pp.97-104, 1999.

志賀 泰博 (非会員) 1999年3月 工学院大学工学部電子工学科卒業。2001年3月 同大学院情報工学専攻博士前期課程修了。現在, 凸版印刷(株)勤務。画像処理, パターン認識の研究に従事。

海老根 秀之 (非会員) 1997年3月 工学院大学工学部電子工学科卒業。1999年3月 同大学院情報工学専攻博士前期課程修了。現在, 同大学院情報工学専攻博士後期課程在学中。画像処理, パターン認識の研究に従事。電子情報通信学会学生会員。

池田 美鈴 (非会員) 2000年3月 工学院大学工学部電子工学科卒業。現在, 同大学院情報工学専攻修士課程在学中。画像処理, パターン認識の研究に従事。電子情報通信学会学生会員。

中村 納 (非会員) 1969年3月 工学院大学工学部電子工学科卒業。1971年3月 同大学院修士課程修了。同年, 同大学助手となる。以来, 学習理論, 画像処理, 画像符号化等の研究に従事。1988年4月~1989年3月, ケベック大招聘教授。1992年 工学院大学教授。工学博士。画像電子学会, 電子情報通信学会, 情報処理学会, 人工知能学会, 映像情報メディア学会, IEEE, Pattern Recognition Society, 日本顔学会各会員。

Society, 日本顔学会各会員。