

# 照明環境に適応的な色情報を用いた 正規化顔画像の抽出法の一検討

近畿職業能力開発大学校

附属京都職業能力開発短期大学校 殿村正延

A study on the extract method of a normalized face image using the color information for adaptation to lighting environment

Masanobu TONOMURA

**要約** 顔認証とは、生体認証(バイオメトリクス)の一つであり、コンピュータに取り付けられているカメラで撮影した顔画像とあらかじめ登録してある顔画像の特徴を照合して個人を特定する技術のことである。この認証技術はまだ新しく、照明や背景など周りの環境に影響を受けやすいなどの問題が指摘されている。照明は、認証処理の前処理として行う正規化画像の抽出過程とその後の認証処理過程それぞれに影響を与える。本論文では、特に認証率を向上させる上で重要な前処理過程に限定し、以下を議論する。

同一照明下において、被写体のカメラからの遠近で顔への照明の照射角度は大きく変わる。これにより輪郭線による特徴量は大きく劣化するが、色による特徴量はそれほどの影響は受けない。しかし、照明の角度変化に不変というわけではなく、色分布は変化する。さらに人により顔、瞳、唇の色も異なる。この変化を多層パーセプトロンの学習・適応能力で吸収し正規化画像を自動抽出する方法を検討し、その効果を評価する。この手法により、認証システムが設置場所の照明環境に適応的になる。

## I はじめに

近年、バイオメトリクス技術が注目されている。バイオメトリクスとは、「指紋」「虹彩」「声紋」「顔」「静脈」など「その人にしかない固有の特徴」を使用し、人物を認識する技術のことであり、生体認証とも呼ばれる。顔認証とは、この生体認証の一つであり、コンピュータに取り付けられているカメラで撮影した顔画像と予め登録してある顔画像の特徴を照合して個人を特定する技術のことである。顔認証のメリットは<長所>

① パスワード認証による忘却や漏洩、カード認証による盗難、防失、複製などの心配がないこと

② 非接触認証で生理的嫌悪感が少ないこと

などがある。①はバイオメトリクス全般にいえるが、

②は指紋認証、虹彩認証にはない特徴である。

顔認証技術は1973年に京都大学が行った研究が始まりである。デジタル画像処理技術の進歩とともに発展し、1993年米国の顔認証アルゴリズム・コンテストによって共通の評価データベース基盤を持つに至り、急速にアルゴリズムが発展した。1997年頃からこの顔認証技術を用いた商品が市場に出回るようになった。

日本では、グローリー工業株式会社の「多重変動分析法による局所特徴比較方式<sup>(1)(2)</sup>(2003.10)」を用いた顔照合技術や東芝情報機器株式会社の「顔deろぐおん<sup>(3)</sup>(2003.11)」やオムロン株式会社の顔認識入退室管理システム「Face Key<sup>(4)</sup>」などが商品化されている。

この認証技術はまだ新しく、問題点としては、

<短所>

① 一卵性双子などの厳密な識別

- ② 照明変化
- ③ 顔の向きの変化
- ④ 表情変化
- ⑤ サングラス、マスク、経年変化

に弱いなどが指摘されており、今日も研究・開発が活発に行われている。

本論文では、上記の「短所②」の問題に焦点をあて、照明環境に適応的な正規化画像の抽出法について検討する。正規化画像の抽出は認証処理の前処理過程で行われる。正規化画像の抽出が正しく行われなければ、その後の認証処理は失敗する。つまり、この処理はシステム全体の認識率を向上させるための重要な過程である。著者らは、2003年に正面顔認証システムを試作し認識率の評価を行っている<sup>(5)</sup>が、このシステムの中に組み込んだ前処理過程に限定し詳細に記述する。

## II. 顔認証システムの概要

### 1 パターン認識システム

パターン認識とは、観測されたパターンを予め定められた複数のクラスのうちの一つに対応させる処理である<sup>(6)</sup>。パターン認識で扱う対象はさまざまなセンサを用いて状況を判断する視覚、聴覚、臭覚、触覚などがある。機械でパターン認識系を構成する場合、どのような対象であろうと一般に図1のような形をとる。パターンが入力されると、まず前処理部でノイズ除去、正規化などの処理を行う。続いて特徴抽出部では、膨大な情報を持つ原パターンから識別に必要な本質的な特徴のみを抽出する。この特徴をもとに識別部では識別処理を行う。識別処理は入力パターンに対して複数のクラスのうちのひとつに対応させることによって行われる。そのため、予め識別辞書を用意し、抽出された特徴をこの辞書と照合することにより入力パターンの所属するクラスを出力する。図1は手書き文字認識の例であるが、顔認証の場合は、原パターンが人の顔を含むカラー画像データとなり、出力が本人の名前となる。

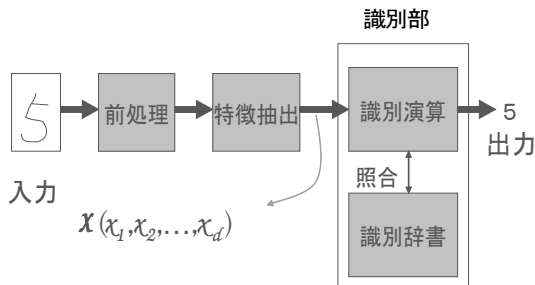


図1 認識系の構成

### 2 顔認証システム

著者らが構築した顔認証システムのアプリケーションウィンドウを図2に示す。以下にこのシステムの処理フローを簡単に示す。

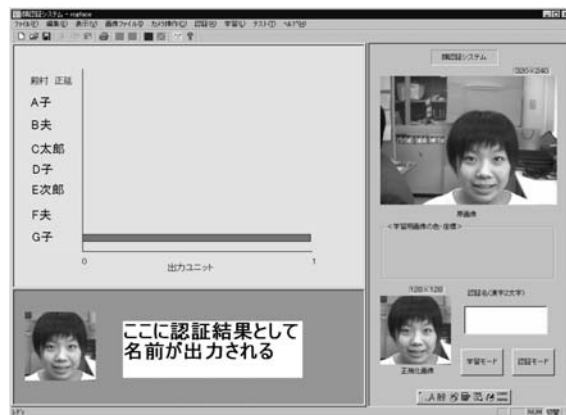


図2 顔認証システムの実行画面

#### <顔認証の処理フロー>

##### (1) 320×240カラーの顔画像の取込み

Webカメラのキャプチャ動画(207レム/sec)から適当なタイミングで静止画を切り出す(図3)。



図3 320×240画素カラー画像

##### (2) 前処理

被写体のカメラからの遠近、顔の多少の回転にかかわらず常にまっすぐな同じ大きさの顔を得るために、移動、回転、拡張などの前処理を行い、128×128のモノクロ正規化画像を得る(図4)。



図4 128×128画素モノクロ画像

##### (3) 特徴抽出

前処理で得た正規化画像から特徴を抽出する。特徴量

としては、一般的にM特徴（モザイク化）、KI特徴、KF特徴、平均顔の主成分分析、判別分析<sup>(7)</sup>を使用するケースが多いが、著者らが試行した認証処理は表情変化に影響を受けにくい場所（今回は両目の中央、おでこ）の局所的な濃度勾配ヒストグラムを特徴量として使用する。その一例を図5に表示する。いずれの手法を使用するにせよ、特徴量が一人一人違うため顔を識別することができる。



図5 特徴抽出

#### (4) 識別部

パターン識別法としては、①パターンの生起を統計的・確率的モデルで捉え、入力パターンがそのモデルに合致するかで識別する統計的・確率的パターン識別法と、②パターンの特徴の相互関連をルールで記述し、標準的なテンプレートと合致するかどうかで識別する構造的・構文的パターン識別法の2通りに大別されるが、本システムでは①の手法を使用している。認証結果は[0,1]の実数で出力し、0.95より大きい値を返した人物を本人として認証する。図6に認証結果の一例を表示する。

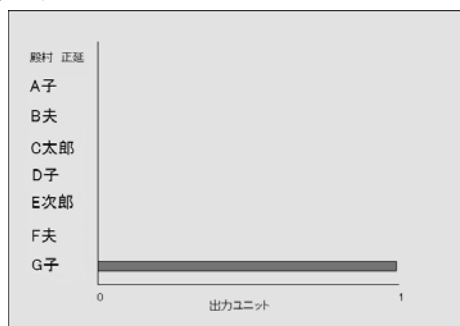


図6 認証結果例

### III. 照明環境に適応的な正規化画像の抽出法

#### 1 処理の着想

図4のように原画像の中から顔を抜き出すためには、

①顔、瞳、口、鼻等の輪郭線に着目する方法<sup>(7)(8)</sup>

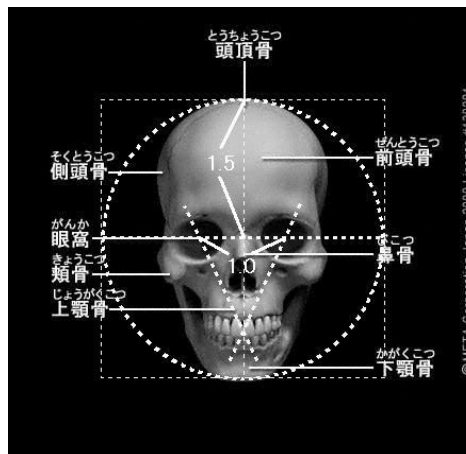
②顔、瞳、唇等の色に着目する方法

がある。

①は照明の照度、角度が常に一定になるよう固定できれば顔の各部の抽出が可能であるが、これは被写体がカメラに向かうときに常に同じ位置に顔がくること

を強いることとなり、精神的なストレスを与えることになる。また、この手法は照明のあたり方(角度)によっては位置が少しずれただけで輪郭が大きく変化するため急激に特徴量が劣化する。さらに、輪郭線は個人差が大きく各部の共通特長を取り出すことが一般的に難しいといわれている。

②は人が一般的に持っている常識、つまり、顔は肌色、瞳は黒、唇は赤と言った認識と一致する。これは被写体が多少カメラに近づきすぎたり、離れすぎたりしても、また顔を多少回転させても①の方法ほど照明角度の変化により大きく変化しない。このことは被写体に撮影時の位置について自由度を増やし、ストレスを与えないことにつながる。図7に標準的な人間の頭部の骨格を示す。図から、もし両瞳の位置と唇の位置が特定できれば、両瞳を結んだ線分の中心からその線分の1.5倍の半径を持つ円はほぼ顔の外接円となる。さらに、その円の中心と唇を結ぶ線分を画像の垂直方向と一致するように補正すれば、多少の拡張・回転に対応できる正規化画像の抽出が行えることがわかるであろう。これは著者らの独自のアイデアである。実際、図4、図13の結果からうまく抽出できることが確認できる。



情報処理推進機構(IPA)「教育用画像素材集」より

図7 両瞳と唇の位置に着目した正規化画像の抽出

以上のことから、本論文では②の手法を採用する。しかし、顔、瞳、唇の色は人によって違い、さらに照明のあたり方(角度)の変化によっても微妙に変化する。つまり、RGBの3次元色空間における顔、瞳、唇それぞれの色特徴ベクトルはある形状を持った分布であり、これは照明の角度変化に伴い形、位置を変える可変的な分布となる。図8に実際に観測した各部の分布を示す。

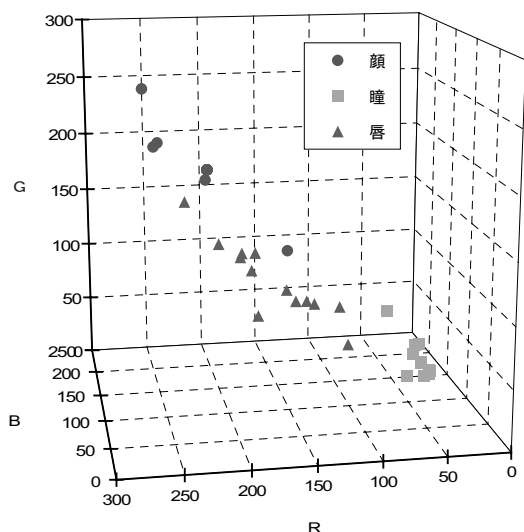


図8 顔・瞳・唇の色空間分布

この問題に対処するため、それぞれの部分の色情報を学習対象となる複数人の画像データから取り出し、色識別器としてニューラルネットワークを使用することで、被写体の照明角度の変化に対して顔、瞳、唇の部位の色の変動を吸収し、正規化画像の抽出を行うことができる。

さらに、顔認証システムを設置する場所を変えた場合、照明の色、照度、角度も大きく変化するが、その場所を取り込んだ画像を使って色の学習を行うことにより、照明環境に適応的となる。これは、音声認識システムにおいて、認識率を向上させるために個人の音声の特徴とマイクの設置された部屋の音響環境の特徴をエンロールにより学習させるのと類似しており<sup>(9)(10)</sup>、パターン認識システムとしては、自然な対応であるといえる。

## 2 色識別器

色識別器として、統計的識別器であるニューラルネットワークを使用する。ニューラルネットワークには、多くのモデルが提案されているが、本論文では多くの応用に適用されている多層パーセプトロン(階層型ニューラルネットワーク)を使用する。このモデルは、閾値関数の非線形処理により理論的には任意の入出力関係を学習により獲得できることが示されている<sup>(11)</sup>。学習法としては一般的に誤差逆伝搬法を使用する<sup>(12)</sup>。

図9に示すような、入力ユニット3個、第1中間層ユニット19個、第2中間層ユニット5個、出力ユニット1個の4層パーセプトロンを使用する。ただし、ユニット数にオフセットは含めていない。このモデルを色識別器として顔、瞳、唇のそれぞれに対して準備する。

入力ユニットには[0,1]に正規化したRGB値を使用する。学習時には出力にそれぞれの部位の色かどうかを{0,1}で割り当て、想起時には0.5を閾値として判定する。

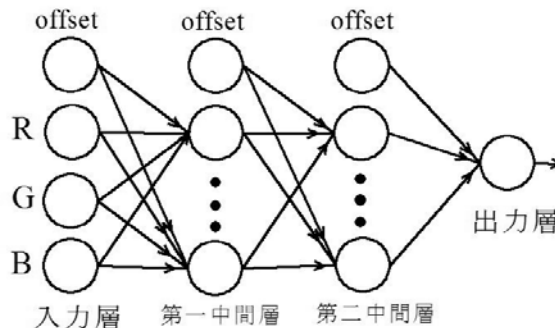


図9 多層パーセプトロン

## 3 正規化画像抽出法

本手法は、次の3つを前提条件とする。

- ① 顔が完全に原画像内にあること
- ② 顔に表情を入れないこと
- ③ 顔の傾きは25度以内であること

③の条件は、後述する(3)の処理において左右瞳を仮の顔の中心を境として画像の左右どちら側にあるかで判別するための制約である。

この処理は、図10のように4つの過程からなる。それぞれの過程の詳細を以下に示す。

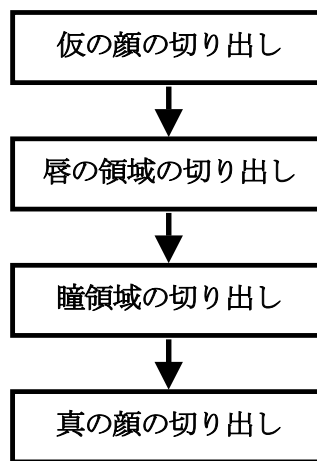


図10 正規化画像抽出過程

### (1) 仮の顔切り出し処理

・想起した顔部分を取り出す(図11の2値化された顔領域を参照)

\* 出力ユニットの出力値が0.5以上なら1、0.5未満なら0に2値化

- ・メディアンフィルタによる雑音除去
- ・ラベリングと各領域の基本特徴量計算
- ・顔領域の候補確定

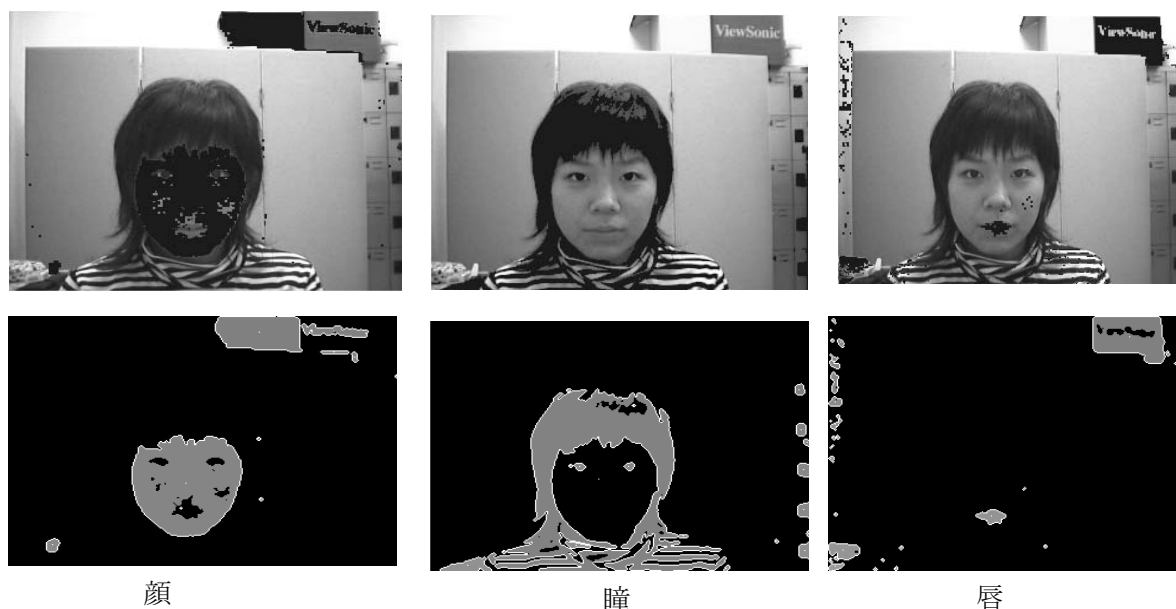


図 1 1 顔・瞳・唇の想起画面

- \* 面積(以降、 $S_f$  とする)×円形度が最大のもの。その領域の中心を顔の仮中心とする (図 1 2 の白い点)
- \* 顔領域の面積から半径を計算し、円を表示 (図 1 2 の白い内側の円)
- \* 上記半径の1.2倍の円を表示 (図 1 1 の白い外側の円)

(2) 唇の領域の切り出し処理

- ・想起した唇部分を取り出す (図 1 1 の 2 値化された唇領域を参照)
- ・メディアンフィルタによる雑音除去
- ・ラベリングと各領域の基本特徴量計算
- ・唇領域の確定
  - \* 顔領域の1.2倍の半径内で、かつ、顔の仮中心座標より下にあり、かつ、面積×円形度が最大のもの (図 1 2 の唇の灰色の点)

(3) 瞳領域の切り出し処理

- ・想起した瞳領域を取り出す (図 1 1 の 2 値化された瞳領域を参照)
- ・メディアンフィルタによる雑音除去
- ・ラベリングと各領域の基本特徴量計算
- ・左瞳の複数候補算出
  - \* 顔半径内  $\wedge$  円形度  $> 0.1$   $\wedge$  面積  $< 0.003 \times S_f$   $\wedge$  左にある
- ・右瞳の複数候補算出
  - \* 顔半径内  $\wedge$  円形度  $> 0.1$   $\wedge$  面積  $<$

$0.003 \times S_f$   $\wedge$  右にある

- ・両瞳の確定
  - \* 左右瞳候補の全ての組み合わせの中から、唇と両瞳の中点を結んだ線分から唇を中心とする各瞳の角度を求める。この値を図 7 で示した骨格構造と誤差許容分を考慮して、18度～26度の範囲にある組み合わせに限定する。
  - \* 唇を基準とした両瞳の左右対称性を判定するために、①唇から各瞳の距離の差の絶対値と②唇を通る両瞳の中線から各瞳への角度の差の絶対値を求める。①、②が最小となる組み合わせを瞳と確定する (図 1 2 の両瞳の黒点)

(4) 真の顔の切り出し処理

- ・確定した両瞳を結んだ線分の中点が真の顔の中心となる(図 1 2 の両瞳の中央にある黒い点)。両瞳の距離の1.5倍を半径とする円(図 1 2 の黒円)で、唇が真下にくるように回転、拡縮、モノクロ化を行うことで図 4 のような正規化画像を得る。



図 1 2 正規化画像抽出過程の各領域

## IV. 正解抽出率の測定

### 1 学習・テスト用データと学習方法

情報技術科の学生、男女10人ずつ合計20人の顔を1人につき10枚のカラー画像として撮影する。10枚のうち5枚を正規化画像抽出のための色学習用に使用し、残り5枚を正解抽出率測定のためのテスト用とする。撮影方法は背景、照明などの撮影条件を統一し、カメラから離れたり近づいたり、左右に首を傾げる等パターンを変える。

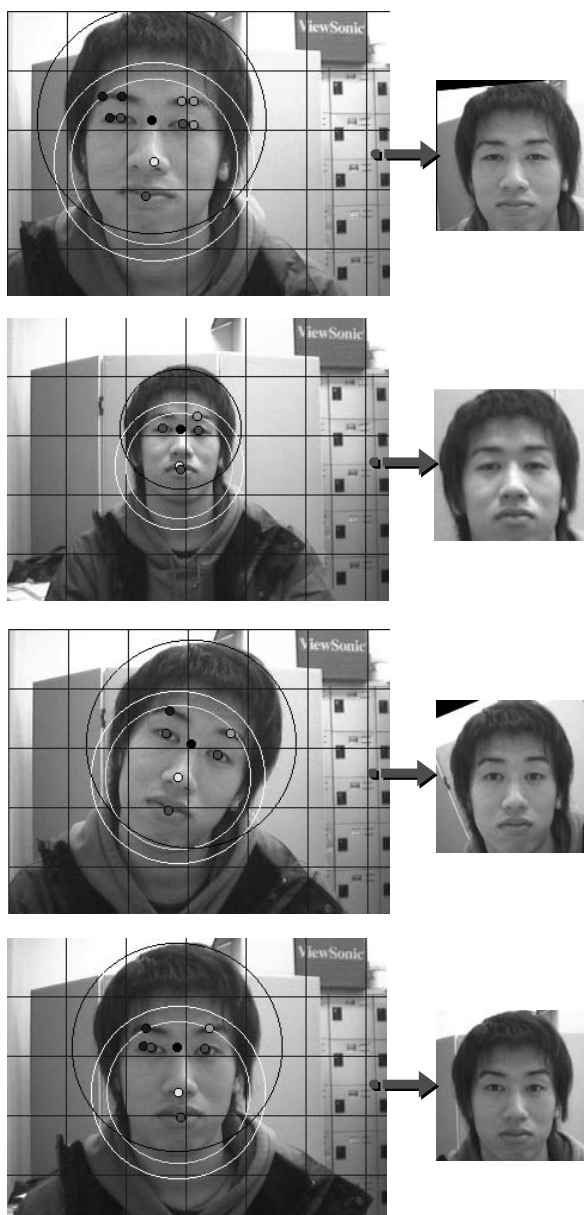


図13 正規化画像抽出の成功例

正規化画像を得るための学習データセットは、顔、瞳、唇のそれぞれの領域についてRGB値を教師データ（入力ユニット値：RGB値、出力ユニット値：1）、そ

れ以外の領域のRGB値を反教師データ（入力ユニット値：RGB値、出力ユニット値：0）として1人につき5枚、1枚につき平均15点取る。つまり、各部位につき1500セット取る。教師データと反教師データを交互にサンプルし、データの偏りを少なくする。学習は一括修正法で5万回打ち切りとする。

### 2 測定結果・考察

正解抽出率は、99%（99枚/100枚）となり、よい結果を得た。図13に正規化画像を正しく抽出できた例を4つ示す。被写体のカメラからの距離や顔の傾きによって変化する照度、角度にかかわらず抽出できているのがわかる。

抽出に失敗した1枚を図14に示す。仮顔の切り出し、唇の切り出しに成功しているが、瞳の切り出しに失敗している。これは眉毛を瞳として確定したためであり、Ⅲ. 3 (3) の「顔半径内、かつ、円形度 > 0.1、かつ面積 < 0.003 × Sf」の条件を満たし、かつ、「唇を基準とした両瞳の左右対称性」の条件を正しい瞳のペアより満たしたためである。これを解決するためには、例えば「円形度 > 0.4」の条件に変更してやる

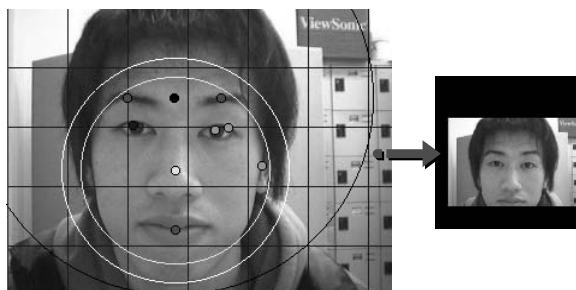


図14 正規化画像抽出の失敗例

ことで正しく抽出されることを確認したが、このような単純な操作では細目になっている他の画像の正規化がうまくいなくなる可能性がある。

なお、今回表示していない他の19名の正規化画像の抽出は全て正しく行われた。

### 3 今後の課題

測定結果より、本手法が有効であることが確認できたが、問題点も明らかになった。

#### (1) パラメータチューニングの労力

Ⅲ. 3で述べた、顔、瞳、唇の切り出し条件のパラメータが経験則で決定される。さらに多くの画像を抽

出する場合、新たなパラメータチューニングを必要とする可能性がある。

この問題を改善するためには、①切り出す部位の設定条件をより詳細に分析することや、②他の画像処理手法と組み合わせることで経験則に基づくパラメータの数を減らすことが考えられる。さらに、③パラメータチューニングの自動化のために、ニューラルネットワーク、遺伝的アルゴリズム、ファジー推論等のソフトコンピューティングの各種手法を組み合わせることで学習能力を持たせることも考えられる。現在、①および③を組み合わせた各手法を比較検証中である。

## (2) 学習データ収集の労力

顔、瞳、唇の領域を正しく捉えるために、新たなデータの追加や、各領域の色が競合するようなデータを誤って登録した場合に削除する必要があるなど、各領域の抽出状況をみながら試行錯誤する必要がある。さらに、この作業はある程度熟練を要する。

この問題を改善するためには、データ登録者がストレスを感じないように登録・削除・データの検証が簡単に行えるアプリケーションインターフェースを導入することが考えられる。

さらに、他の正規化画像抽出法との比較を行い、本手法の有効性を検証する必要がある。

## V. おわりに

照明環境に適応的な正規化画像抽出法について検討した。本手法は、顔、瞳、唇の色情報に着目し、可変的な分布を持つ各部位の色を識別するために多層パーセプトロンを使用する。これにより、照明の角度変化に強く、認証対象者のカメラ位置に対する自由度を増やすことができる。また、顔認証システムの設置場所の照明環境にも適応的となる。20人の被験者で測定したところ、有効な結果が得られた。さらに、本手法の問題点を明らかにし、改善の方向性も示した。

なお、本内容はセミナー「ニューラルネットワークの利用技術」の中で、ニューラルネットワークの適用事例の一つとして紹介している。

### [参考文献]

- (1) <http://www.glory.co.jp/press/img-press/j-031021.pdf>
- (2) <http://www.hitachi.co.jp/New/cnews/>

031106a.html

- (3) <http://www.toshiba-tie.co.jp/press/031006.html>
- (4) <http://www.njp.co.jp/koukoku/koukoku3.htm>
- (5) 山崎 可林、塩川 友美、殿村 正延、顔認証システムの構築、京都職業能力開発短期大学校総合制作実習論文、Mar 2003,P1-15.
- (6) 石井健一郎他、わかりやすいパターン認識、オーム社、1998、P1-11
- (7) 中川聖一、パターン情報処理、丸善、1999、P251-255
- (8) RBrunelli and T.Poggio, Face recognition featured versus templates, IEEE Trans. PAMI,vol.15,No.10,Oct 1993,P1042-1052.
- (9) 殿村 正延、音声認識システムの実用化可能性の検証—IBM ViaVoiceの不特定話者に対する認識率・問題点の評価—、京都職業能力開発短期大学校紀要17号、Dec 2003、P21-26.
- (10) 殿村 正延、不特定話者に対する音声認識システムの認識率評価、職業能力開発報文誌、第17巻第1号、2005、P45-49.
- (11) 坂和 正敏・田中 雅博、ニューロコンピューティング入門、森北出版、1997、P39-42
- (12) D.E. Rumelhart, G.E. Hinton and R.J. Williams, Learning representation by back-propagation errors, Nature, Vol.323, 1986, P533-536.

