

## 学習者認識のための顔画像検出と顔の向き認識の基礎的実験

武岡 さおり・尾崎 正弘・川田 博美・足達 義則\*

### Basic Experiments in Computer Recognition of Direction of the Human Face

Saori TAKEOKA, Masahiro OZAKI, Hiromi KAWADA and Yoshinori ADACHI

#### はじめに

コンピュータを用いて顔画像の自動認識を行う技術は、セキュリティやマンマシン・インタフェースなどの観点から、個人識別や人物の行動管理など広い範囲での応用が期待され、重要な技術として注目されている。

我々はこれまで、インターネットを利用して、学習者の理解度を考慮した CAI システムのプロトタイプモデルを試作して学習実験<sup>1)～5)</sup>を行ってきた。このシステムでは、学習者個別の学習履歴データ、現在の学習状況、学習課題への理解度などをデータベースとして保持しており、それを利用して次の学習を進めていく。このため、学習を始める際に、学習者の個人認証が必要となる。

本研究では、学習者の認証方法として、顔画像を用いた個人認証を採用し、その基礎的な実験を行った。

顔画像を用いて個人認識を行うためには、まず、カメラに映し出された画像から対象とする人物画像を検出する処理が必要となる。つまり、顔画像認識を全自動化するためには、顔認識処理の前に、原画像から顔領域を切り出す処理が必要となる<sup>6)～12)</sup>。我々は、皮膚の肌色情報を用いて、原画像中から顔領域の自動検出を行った。

また、より正確に人物の顔を把握するためには、人物の顔の向きを認識する処理が必要となる。そこで、検出された顔領域画像を用いて、顔の方向認識実験を行った。これらの結果について報告する。

#### 肌色領域検出処理

顔画像認識を行う場合、対象とする画像には、動画画像と静止画像が考えられる。動画画像から顔領域の検出を行う場合、前フレーム画像と現フレーム画像との差分処理により、リアルタイムで比較的容易に顔の位置を検出することができる。しかし、静止画像の場合は、動画画像と異なり、使用できる情報量が少ないため、顔の位置をすばやく正確に検出することは容易ではな

---

\* 中部大学

い。また、動画像の場合でも対象とする人物の動きが極端に小さい場合には、前後画像フレームの差分を取ることで場所が特定できず、静止画像と同様の処理が必要になる。したがって、本研究では、静止画像を対象とし、顔画像を抽出する手法について検討した。

顔領域を検出するために、本研究では、YCC 表色系を用い肌色領域の検出を行った。

YCC 表色系はカラーテレビの信号として利用されており、白黒テレビとの互換性を保つために、輝度信号と色信号を分離して別々に取り扱っている。

輝度YはRGBの値から、人間の比視感度を用いて、次のように定義される。

$$Y = 0.299R + 0.587G + 0.114B$$

YにおいてR、G、Bに対する係数は、比視感度曲線を考慮して作られており、 $R = G = B$ のとき、すなわち無彩色のときは $Y = R = G = B$ となる。

この時、色情報のCr、Cbは、それぞれR、Bから輝度信号Yを差し引いた形になっており、色差信号と呼ばれる。また、デジタルテレビやJPEG圧縮では色差のダイナミックレンジをせまくした次の変数が用いられている。

$$Cr = 0.713(R - Y) = 0.500R - 0.419G - 0.081B$$

$$Cb = 0.564(B - Y) = -0.169R - 0.331G - 0.886B$$

さらに、人間の目が輝度Yに比べて色の解像度に鈍感であることにもとづいて、Cr、Cbに関するサンプリングはYの半分の空間密度で行い、それによりデータ量を小さくしている。

YCC表色系における利点は、一見すると肌の色が違う人物でも、その個人差によらず、抽出が可能である点が挙げられる。また、Cr、Cb値の範囲による抽出法のため、画像の大きさにか



原画像



肌色領域抽出結果



原画像



肌色領域抽出結果

図1 YCC表色系を用いた肌色領域抽出結果の例

かわらず、一定の抽出結果が得られる。本研究では、

$$133 \leq Cr \leq 173 \quad \text{かつ} \quad 77 \leq Cb \leq 127$$

を満たす Cr、Cb 値を持つピクセルを顔の肌色領域とした<sup>9)</sup>。

YCC 表色系を用いて、肌色領域の抽出を行った結果を図 1 に示す。

### 顔領域検出実験

顔領域を検出するためには、まず原画像中の顔の位置を知る必要がある。そこで、肌色領域画像に対して、x 軸方向と y 軸方向の肌色画素のヒストグラムを作成し、顔が存在する位置を決定した。

まず、肌色領域画像から得られたヒストグラムを用いて、顔の左右両端の探索を行った。図 2 のように、肌色領域画像の各列ごとに肌色ピクセル数をカウントし、ヒストグラムを描いた。そして最大ピクセル数を元に閾値を決定し、閾値よりも大きな値を持つ領域において、最も大きな幅を持つ領域の左右両端を顔の左右両端として決定した。

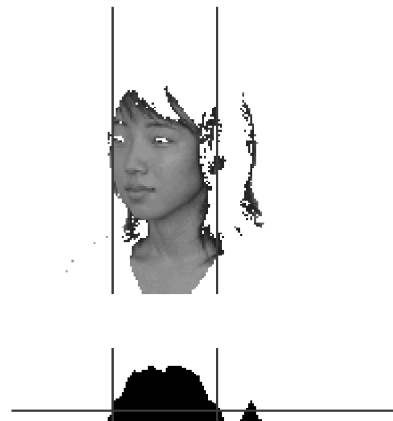


図 2 顔の左右両端の決定

次に、頭頂部の探索を行った。頭頂部探索には、肌色領域画像と Sobel フィルタによって検出したエッジ画像を用いた。図 3 のように、肌色領域画像の上端より上部に向かってエッジを探索し、頭頂部とした。

さらに、目・顎の探索を行った。これには、肌色領域から得られたヒストグラムと、エッジ画像から得られたヒストグラムとを用いた。図 4 に示すように、肌色ピクセルのヒストグラムは、目の位置ではくぼみ、頬に向かって同程度以上に盛り上がる。またエッジ画像から得られたヒストグラムでは、目の付近を超えると大きくくぼんでいる。この位置を目の位置とした。

顎の位置は、肌色ピクセルのヒストグラム上で、目の位置から下方向に大きくくぼんだ位置を探索し、そこを顎とした。



図 3 エッジ画像を用いた頭頂部探索

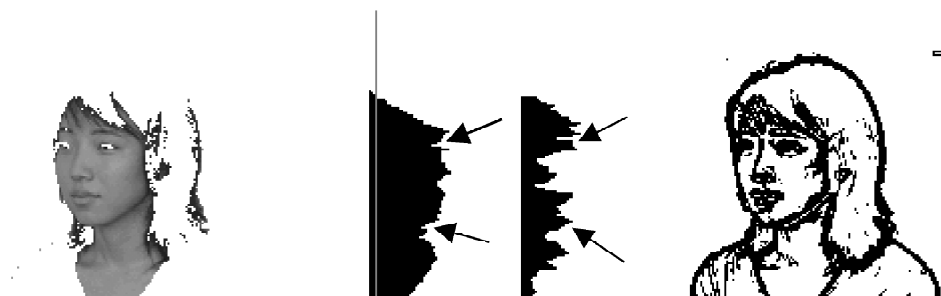


図4 目・顎の探索



図5 顔領域検出結果

以上のような手法で得られた頭頂部と顎の位置を顔の上下両端とし、顔領域を検出した。図5に顔領域の検出結果を示す。

### 顔の向き認識実験

検出した顔画像を用いて、顔の向きの認識実験を行った。

まず、辞書画像作成のために、人物の顔の右側面を $0^\circ$ 、正面を $90^\circ$ 、左側面を $180^\circ$ となるように設定し、 $15^\circ$ 間隔で合計13パターンの画像を、13名分用意した。そして、顔領域検出と同様の処理を行い、顔領域画像を切り出し、全ての画像のサイズを $77 \times 60$ ピクセル（縦横比1.3:1）となるように正規化した。図6にその例を示す。

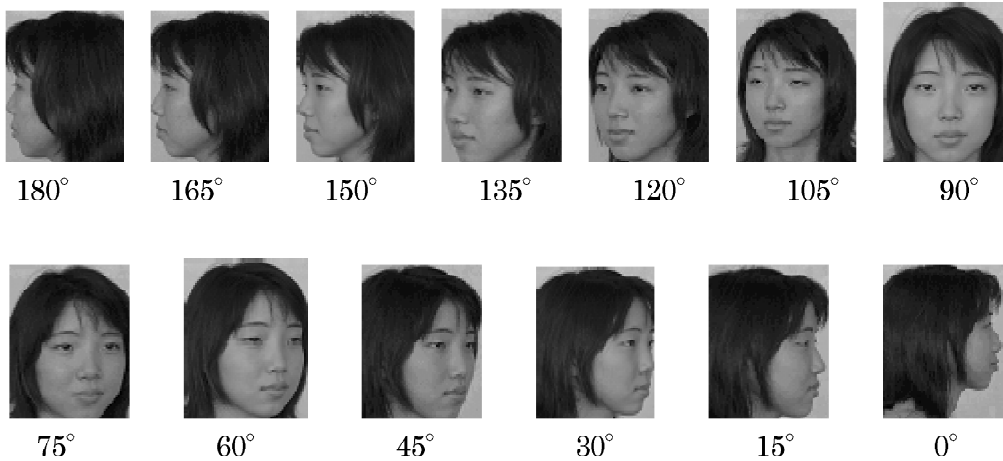


図6 0°から180°の画像（例）

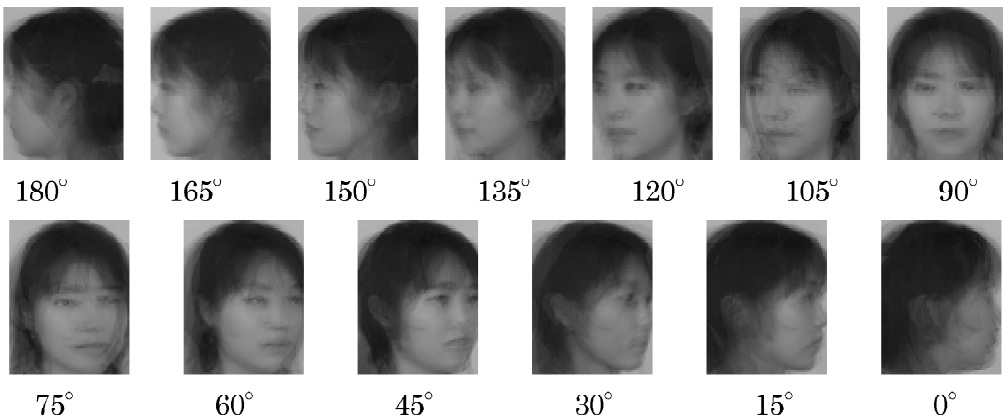


図7 辞書画像として用いた平均顔画像

13パターンの各角度について、12名分の画像をもとに平均顔画像を1枚ずつ作成し、辞書画像とした。このとき、1名分の画像は認識の対象画像として省いた。このようにして作成した13パターンの辞書画像を主成分分析で得られる固有空間に投影した。図7に作成された辞書画像を示す。固有空間の次元は、寄与率90%として3次元空間となった。

次に、辞書画像を作成する際に用いなかった1名分の顔画像13パターンそれぞれについて顔の向きの認識実験を行った。固有空間へ投影した認識対象画像とそれぞれ13枚の辞書画像とのユークリッド距離を求めた。そして距離の近い2枚の画像の角度と距離を用いて角度補間を行い、顔の向きを決定した。

図8は、顔の向きの認識対象画像として用いた画像である。

## 実験結果

表1に、顔の向きの認識実験を行った結果、予測値と実測値の誤差を示す。

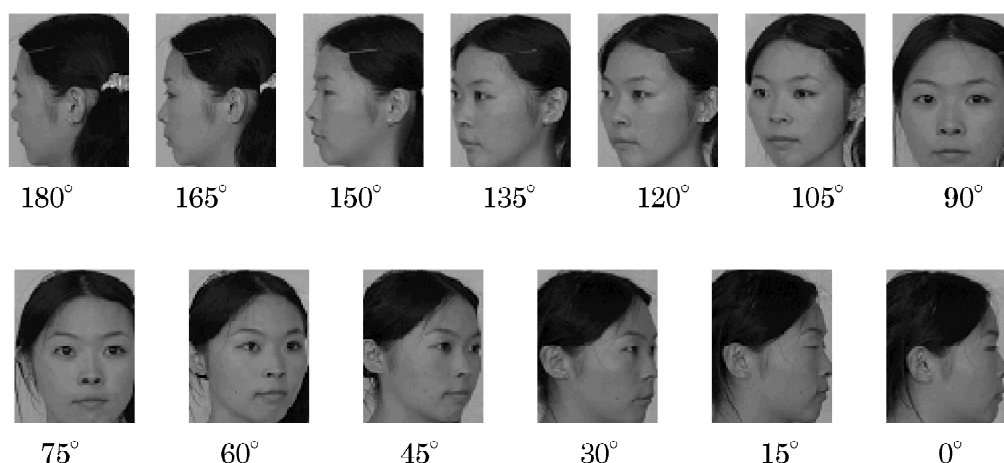


図8 顔の向きの認識に用いた顔画像

表1 顔の向きの認識実験結果

| 実際の角度 (°) | 得られた角度 (°) | 誤差 (°) |
|-----------|------------|--------|
| 0         | 6          | 6      |
| 15        | 10         | -5     |
| 30        | 13         | 17     |
| 45        | 37         | -8     |
| 60        | 100        | 40     |
| 75        | 79         | 4      |
| 90        | 88         | -2     |
| 105       | 108        | 3      |
| 120       | 114        | -6     |
| 135       | 133        | -2     |
| 150       | 150        | 0      |
| 165       | 175        | 10     |
| 180       | 177        | -3     |

13パターンの画像を認識させた結果、誤差の絶対値の平均は8.3°であった。60°のパターンの画像を認識させた際に誤差が40°と最大で、逆に誤差が最小となったのは150°のパターンであった。

## 考察

### (1) 顔領域の検出について

今回、YCC表色系を用いて肌色領域を抽出することで画像中の顔領域の検出を行った。YCC表色系は、肌色の個人差にあまり影響を受けることなく抽出でき、画像の大きさに関わらず一



図9 頭髪部を肌色領域として検出した例

定の肌色領域が抽出できるという利点がある。しかし一方で、肌色に似た色は全て抽出してしまうため、髪の色を染めたり、光線の反射で茶色くなったりすると、図9のように、頭髪部までも肌色として抽出してしまうという欠点もある。

また、髪型の影響によって顔領域が狭くなると、顔の領域の上下の範囲を決定する際に使用している目や口、顎などの顔部品を捉えることができなくなる場合がある。また、顔を上または下に向けるなど、顔の向きによっても、肌色領域の範囲や目、口などの部品の位置が変化し、現状では正しく顔領域を検出することができない。

今後、このような場合の顔領域検出手法についても検討していく必要がある。

## (2) 顔の向きの認識について

認識実験において、角度 $60^\circ$ での認識誤差が $40^\circ$ と大きかった理由は不明であるが、髪型の個性などにより、辞書画像に用いた顔画像に偏りが生じたことも原因のひとつとして考えられる。今回は、実験に用いた画像の人数が少ないため、いくつかの問題点が残されたが、今後実験を重ね、これらの問題を改善していく必要があると考える。

また、今回は $15^\circ$ 間隔で13パターンの辞書画像を作成した。しかし、角度が $0^\circ$ から $30^\circ$ 、 $75^\circ$ から $105^\circ$ 、 $150^\circ$ から $180^\circ$ の間では、顔領域画像にあまり変化がみられない。逆に $30^\circ$ から $60^\circ$ 、 $120^\circ$ から $150^\circ$ の間では、目や口、鼻などの部品の位置や顎のラインなどが著しく変化している。そこで、顔の変化が少ない角度付近については辞書画像を少なくし、変化の大きい角度付近は、細かな角度で辞書画像を作成して、認識の精度が上がるか検討していきたい。

## おわりに

顔画像を用いて個人認識を行うためには、撮影された原画像から顔領域を検出する処理が必要である。今回は、静止画像を対象とし、皮膚の色情報を用いて、画像中から顔領域の自動検出を行った。また、検出された顔画像を用いて顔の向きの認識実験を行った。

本研究では、顔画像領域を抽出するために、顔の肌色領域のヒストグラムを用いた。また、顔の向きの認識では、顔の向きを表す固有空間を生成し、抽出した顔領域画像を空間内に射影することによって、辞書画像との距離を求め、顔の角度の推定をした。その結果、平均認識誤差 $8.3^\circ$ で顔の向きを認識できた。

今後は、顔の向きや髪型などに左右されることなく、より正確に顔領域の検出ができるよう、

検討を重ねていきたい。

#### 参考文献

- 1) 武岡さおり、尾崎正弘、岩下紀久雄、江島徹郎、足達義則 “学習者の理解度を考慮したハイパーテキスト型 CAI 教材の開発に向けて” 日本教育情報学会年会論文集、16、pp. 178-181、2000
- 2) 江島徹郎、尾崎正弘、足達義則、武岡さおり “インターネットブラウザの機能を利用・拡張したオーサリングシステムの試作” 日本教育情報学会年会論文集、16、pp. 168-171、2000
- 3) 武岡さおり、尾崎正弘、川田博美、岩下紀久雄、江島徹郎、足達義則 “学習者の理解度を考慮したハイパーテキスト型 CAI 教材の試作” 名古屋女子大学紀要 (人文・社会編)、48、pp. 177-186、2002
- 4) 尾崎正弘、武岡さおり、川田博美、小山幸治、足達義則 “個別学習によるハイパーテキスト「シスアドブック」の開発” 教育システム情報学会第27回全国大会講演論文集、pp. 305-306、2002
- 5) 小山幸治、武岡さおり、川田博美、尾崎正弘、足達義則、 “理解度向上支援総合ネットワーク型教育システムの構築——データ構造に着目した DB の構築——” 日本教育情報学会第18回年会論文集、pp. 254-257、2002
- 6) 本郷仁志、羽飼直記、山本和彦 “4 方向面特徴と色面特徴の統合による顔および顔部品の抽出” 画像の認識・理解シンポジウム (MIRU'98)、II、pp. 307-312、1998
- 7) 羽飼直記、本郷仁志、加藤邦人、山本和彦 “UV 平面の 2 次元色ヒストグラムを用いた顔部品抽出” 画像の認識・理解シンポジウム (MIRU'98)、II、pp. 391-396、1998
- 8) 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抽出” 電気学会論文誌、Vol. 119-C、No. 11、pp. 1351-1358、1999
- 9) 堀田一弘、栗田多喜夫、三島健稔 “Log-Polar 画像から抽出したスペクトル特徴を用いた対象の大きさの変化に影響を受けにくい認識法——顔検出と顔識別への応用” 信学技報 PRMU99-110、pp. 69-76、1999
- 10) 足達義則、今井昭宏、尾崎正弘、石井直弘 “肌色領域抽出手法の検討” 電気学会論文誌 C、Vol. 120-C、No. 12、pp. 1927-1932、2000
- 11) 武岡さおり、小山幸治、川田博美、尾崎正弘、足達義則 “顔による行動追跡を付加した教育システムの検討” 日本教育情報学会第18回年会論文集、pp. 236-239、2002
- 12) 武岡さおり、尾崎正弘、足達義則 “肌色領域を用いた顔部品抽出による顔の方向認識の検討” 電気関係学会東海支部連合大会講演論文集、p. 320、2002