

7F-2

複数カメラによる顔の向きへの識別*

三谷純司† 外川文雄†

RWCP 新機能シャープ研究室‡
シャープ(株) 応用システム研究所内§

1 はじめに

我々は、画像や音声などを用いた人にやさしいマルチモーダルヒューマンインタフェースの実現を目指し、人と人との対話の解析を行っている [1]。この解析の結果、ユーザの注意の対象を知ることが重要であることを確認した。このユーザの注意の対象を理解する方法の一つとして顔の向きへの識別が有効であると考え、顔の向きへの識別に関する実験を行っている。本稿では、エッジの分布パターンを特徴として高速処理が可能な識別手法と、識別精度の向上のためにカメラを複数用いる方法について述べる。

2 顔の向きへの識別手法

2.1 手法の概略

顔の三次元モデルとのマッチングに基づいて画像からの顔の向きを識別する方法は実時間での処理が困難である。そこで、本手法では、顔の各向きについて、それぞれ二次元モデルを用意し、この各モデルとのマッチングに基づく方法を用いている。マッチングの特徴については、目や口などの顔の特徴となる部品を正確に画像から抽出することは困難であり、顔にマークを付ける方法 [2] などが考えられている。しかし、自然なヒューマンインタフェースとしては、あまり望ましくない。ここでは、目や口などは画像上ではエッジとして現れることに着目し、顔画像におけるエッジの分布パターンを特徴として用いた。さらに、識別精度を向上させるために、カメラを複数用いている。

2.2 特徴抽出

本手法では、顔領域を囲む矩形内 (図 1) におけるエッジの分布パターンを特徴として用いる。この矩形内においてエッジの強さを求め、エッジの強さが閾値以上なら 1 に、エッジの強さが閾値未満なら 0 とするパターンを作成する (図 2)。マッチングのために、モザイク処理により、矩形を一定の大きさになるように正規化し、さらにノルムが 1 になるように正規化して、特徴パターンとする (図 3)。



図 1: 顔領域画像

図 2: エッジの分布パターン

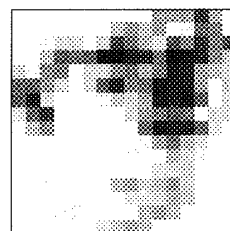


図 3: 特徴パターン

2.3 マッチング

顔の各向きに対して、あらかじめ複数の顔画像データから特徴パターンを求めて、顔の向き d_i ごとに平均して標準パターン R_{d_i} を作成する。マッチングは、パターンをベクトルとみなして、特徴パターン F と標準パターン R_{d_i} との内積 $F \cdot R_{d_i}$ により求められる類似度 S_{d_i} によって評価する。

カメラが 1 台の場合、顔の向きごとに標準パターンとの類似度 S_{d_i} を計算し、最大類似度 $S_{d_m} = \max_{d_i} S_{d_i}$ となる方向 d_m を顔の向きとする。

*Discrimination of Facial Direction Using Multiple Cameras

†Junji MITANI, Fumio TOGAWA

‡Real World Computing Partnership Novel Functions Sharp Laboratory

§in Integrated Media Laboratories, SHARP Corp.

2.4 複数カメラを用いた識別法

予備実験を行い、カメラ1台の場合の識別結果について調べたところ、正しく識別した場合、正しい向きの類似度は他の方向の類似度より比較的高い値であった。一方、誤った識別の場合、誤った向きの類似度は正しい向きの類似度に比較してかなり近いことがわかった。

そこで、カメラを複数用いた場合、各カメラ c_j の顔の向き d_i ごとに類似度 S_{c_j, d_i} を計算し、各カメラ c_j の1位の類似度と2位の類似度の差 ΔS_{c_j} を求める。 $\Delta S_{c_n} = \max_{c_j} \Delta S_{c_j}$ となるカメラ c_n において最大類似度 $S_{c_n, d_m} = \max_{d_i} S_{c_n, d_i}$ となる向き d_m を顔の向きとする。

3 実験

カメラ2台を用いて、被験者の前方より左右 30° の角度から、室内の通常の照明で、撮影した。入力画像は 640×468 画素、RGB 各8ビットのカラー画像から顔領域を切り出し、 150×150 程度の大きさとなっている。被験者には、正面、および正面に対して左右 30° の3方向を向いてもらい、顔画像の撮影を行った(図4)。被験者は21人で、このうち3人のデータ(内眼鏡使用者2人)を用いて標準パターンを作成し、テストデータには残りの18人分のデータ(内眼鏡使用者8人)を使用した。特徴パターンの大きさは 16×16 とした。

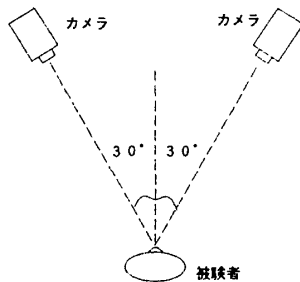


図4: 撮影条件

4 結果

左右それぞれのカメラによる顔の向きの識別の結果を表1に示す。1台のカメラで 30° 差の3方向についてどちらを向いているかを約94%の精度で識別でき、本手法の有効性を示している。

また、上記の識別結果を被験者の眼鏡の有無別に示すと表2のようになる。眼鏡の有無による結果の差がほとんどないことから、本手法は眼鏡による影響をほとんど受けないことがわかる。

一方、2台のカメラを用いた識別実験の結果は表3のようになった。カメラを複数用いることにより識別

の精度が向上することが確かめられた。

表1: 1台のカメラによる顔の向きの識別
(正解数/全体の数)

	左カメラ	右カメラ
左向き	18/18	16/18
正面	15/18	18/18
右向き	18/18	16/18
合計	51/54	50/54

表2: 眼鏡の有無別の識別結果
(正解数/全体の数)

眼鏡あり	眼鏡なし
45/48	56/60

表3: 2台のカメラによる顔の向きの識別

(正解数/全体の数)

左向き	18/18
正面	18/18
右向き	18/18
合計	54/54

5 おわりに

高速処理が可能な手法として、エッジの分布パターンを特徴として用い、複数の標準パターンとのマッチングにより顔の向きを識別する手法が有効であることを確認した。さらに、複数カメラを用いることにより、精度が向上することを確認できた。今後は、より多方向の識別を行い、実際のヒューマンインタフェースへの応用へ進めていく予定である。

謝辞

本研究を行うに際して、日頃ご指導いただく応用システム研究所中島隆之所長に感謝いたします。また、ご討議いただいた新機能シャープ研究室の皆様、および実験にご協力いただいた皆様に感謝いたします。

参考文献

- [1] Watanuki, Sakamoto, Togawa: Analysis of Multimodal Interaction Data in Human Communication, ICSLP94, 17.8, 1994
- [2] 大村, 伴野, 小林: 単眼視画像による顔の向き検出法の指示入力への応用, 信学論 D-II, J72-D-II, No.9, pp1441-1447, 1989-9