

ニューラルネットによる顔の部位認識アルゴリズム

4 P-6

吉田 博明* 蚊野 浩* 西條 淳夫* 人見 正明* 梅崎 太造** 大隅 正人*

*三洋電機メカトロニクス研究所 **中部大学工学部

1. はじめに

我々は、赤外画像と可視画像のセンサフュージョンによる画像処理技術[1]に着目し、顔領域の抽出、部位認識を行い、さらに特定部位の皮膚温を計測するシステムの開発を行っている。

今回、フィルタネットというニューラルネットワークの一種を用いて、シーンから抽出した顔領域に対して、両眉・唇の認識を行うアルゴリズムを開発したので報告する。

2. フィルタネットによる部位認識

フィルタネットは3層構造のフィード・フォワード型ニューラルネットであり、認識結果を領域として出力するタイプのものである。ネットワークの学習には、通常のバックプロパゲーション法を用いる。フィルタネットは文字認識における文字領域の切り出しなどに有効であると報告されている[2]。本報告で用いるフィルタネットは、顔画像を入力し、両眉と唇に相当する領域を出力するものである。ネットワークモデルを図1に示す。各層の構成ユニットは図1の通りである。

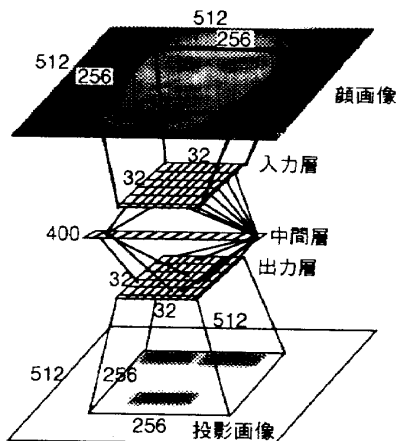


図1 フィルタネット

3. フィルタネットによる両眉・唇領域の認識

今回我々は、顔画像からの両眉及び唇領域の同時認識を試みた。そこで、認識に用いるフィルタネットは、両眉と唇の3領域を含むような一つの大きな

フィルタネットを用いた。このような大きな領域を1つのネットワークとして学習させれば、各部位の形状、位置の情報だけでなく、3領域の位置関係も一度に学習することができるので、左右の眉の誤認識などが防止できると考えられるからである。

3. 1 フィルタネットの学習

学習画像の作成は以下の通りに行う。

- 512×512の顔画像を入力する。
- 256×256の領域を乱数を用いて切り出す。
- 切り出した領域を平滑化、間引きにより32×32の画像に圧縮する。
- ヒストグラム平坦化により濃淡レベルを正規化する。(図2(b)(1),(c)(1))

以上の処理により32×32の学習画像を作成する。c.において切り出した画像を256×256から32×32へと縦横各1/8に圧縮を行っている。顔画像の圧縮は、圧縮率を大きくしていくに従い、眉や唇といった顔面内部位の特徴が消失していく。そこで、実験により顔面内部位の特徴が残り、個人の特徴がなくなるような圧縮率を導き出した結果縦横1/8の圧縮率となった。

次に、上記処理で作成した学習画像に対する教示画像は、各部位の外接四角形に対応するユニットを1、その他のユニットを0とする。また、両眉・唇領域が完全には含まれない学習画像の学習を抑制学習と呼んでいるが、この抑制学習時の教示画像は全ユニットを0として学習を行う。(図2(b)(2)(c)(2)参照)このような教示画像を作成するために、512×512の顔画像に対して、あらかじめ目視で、両眉・唇領域を指定する。

フィルタネットの学習は以下の手順で行う。

- 顔画像から両眉、唇の3領域が含まれるように、学習画像及び教示画像を作成する。

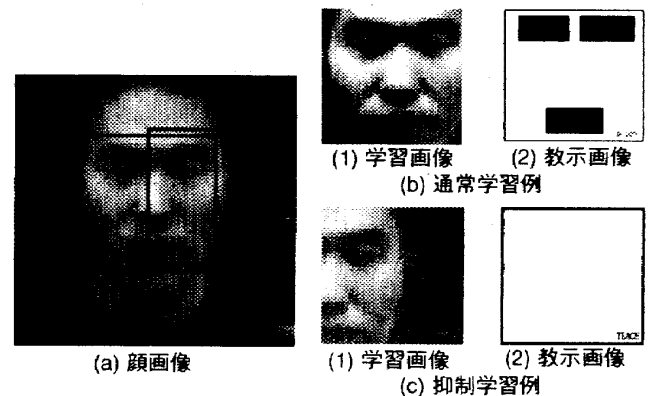


図2 フィルタネットの学習

Evaluation of facial parts using neural networks
Hiroaki Yoshida*, Hiroshi Kano*, Atsuo Saijo*, Masaaki Hitomi*, Taizo Umezaki**, Masato Osumi*
*SANYO Electric Co., Ltd. 1-18-13, Hashiridani, Hirakata-shi, Osaka 573 Japan
**Faculty of Engineering, Chubu University 1200, Matsumoto-cho, Kasugai-shi, Aichi 487 Japan

- b. 入力(32×32)・中間(400)・出力(32×32)の3層のネットワークを用い、バックプロパゲーション法により学習を行う。
- c. 顔画像の任意領域から、抑制学習画像および抑制教示画像を作成する。
- d. 上記bと同じネットワークで学習を行う。
- e. a～dをニューラルネットの出力値の二乗誤差がしきい値以下になるか、学習回数がある一定回数に到達するまで学習を行う。

3. 2 部位認識方法

領域の認識は以下の手順で行う。

- a. 顔画像から学習画像作成と同様な前処理を施した画像を切り出し、フィルタネットに入力して出力値を計算する。(図3(b)(c))
- b. 出力値を顔画像の大きさに拡大投影し、ガウス窓をかけて加算する。
- c. a, bの処理を16dotのシフト幅でx, y方向に全画面走査する。
- c. 投影加算した中で出力値の最大値を求め、出力値が最大値の50%以上の領域を各特徴部位とする。(図3(d)(e))

4 認識アルゴリズムの評価

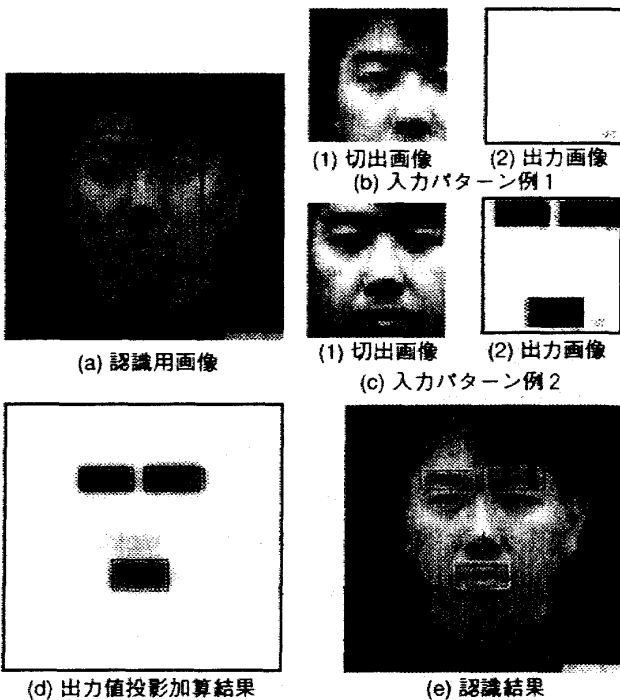


図3 両眉・唇領域認識例

4. 1 評価実験

本アルゴリズムの検証を行うために以下の評価実験を行った。フィルタネットは、16枚(男女各8枚)の画像を用いて学習し、学習に用いていない242枚(男性24名女性12名各7枚ないし6枚)の画像を用いて認識結果の評価を行った。

認識の正否判定は、以下の2条件が共に満たされたときに正しく認識されたとする。この場合、抽出された領域は両眉・唇にほぼ正しく対応していた。

1) 人間の教示した領域と、フィルタネットにより切り出した領域の重なり合う領域の面積が、教示領域の75%以上である。

2) フィルタネットによる切り出し領域の面積が、教示領域の150%未満である。

4. 2 実験結果

評価実験結果を表1に示す。表1より98.3%の成功率で3領域の認識に成功した。認識成功例を図4に示す。また、本アルゴリズムとの比較のために、3領域それぞれを個別に認識する3つのフィルタネットを作成し、同様の認識実験を行った。その認識結果を表2に示す。表2より3領域を個別に認識するフィルタネットでも95.0%の認識率で両眉・唇の3領域が認識されているが、今回提案している手法ではさらに高い認識率が得られている。これは、本手法が両眉・唇の3領域全て含んだ画像を用いて学習しているため、3領域の位置関係も学習しているためである。

表1 評価実験結果

	認識成功率(%)
3領域認識	98.3(238/242)
唇領域認識	100(242/242)
左眉領域認識	98.3(238/242)
右眉領域認識	100(242/242)

表2 比較実験結果

	認識成功率(%)
3領域認識	95.0(230/242)
唇領域認識	100(242/242)
左眉領域認識	97.5(236/242)
右眉領域認識	96.6(234/242)



図4 両眉・唇領域認識成功例

5. まとめ

今回、フィルタネットによる顔面内特徴部位認識アルゴリズムを提案した。また、本手法を検証するための評価実験を行い、未学習画像に対して98.3%の認識率が得られた。本手法で、両眉及び唇領域の3領域の形状や、位置関係の情報が、ネットワークに学習させることができた。

謝辞

なお、この研究は通商産業省・工業技術院の産業科学技術研究開発プロジェクト「人間感覚計測応用技術」の一環として実施したものである。

参考文献

- [1] 広野豊, 河田宏: "赤外・カラー画像による人物画像処理"第131回画像電子学会研究会, pp.13-18(1992)
- [2] 藤吉弘巨, 梅崎太造, 金出武雄: "ナンバープレート内の一連番号の切り出しと認識について"信学技報 PRU95-67, pp.103-108(1995)