ファジィ推論とニューラルネットワークを用いた 人物画像の認識に関する研究*

福田 敏男^{*1}, 伊藤 茂則^{*2}, 新井 史人^{*1}

Recognition of Human Face Using Fuzzy Inference and Neural Network

Toshio FUKUDA, Shigenori ITO and Fumihito ARAI

This paper proposes an image extraction technique of the human face by color image processing, fuzzy inference, and neural network. The image processing technique provides a method for extraction of skin color and measurement of detailed information of the human face. The fuzzy inference estimates the face direction based on the configuration of the extracted face candidates, while the neural network distinguishes faces and other objects by detailed information and face direction. Compared with the previous works which extract data on the human full face, this method makes it possible to recognize not only a full face but also profiles. Further, we show how this method works for the illusion problem. The extraction accuracy of this system is 80 percent at this moment. According to the experimental results, this system will be useful for the vision systems of autonomous robots.

Key Words: Digital Image Processing, Measurement, Fuzzy Inference, Neural Network, Human Face, Recognition, Illusion

1. はじめに

近年, 人手不足問題の解消と危険な仕事の自動化を目的に, 生産業界をはじめとするさまざまな分野でロボットの導入および開発が進められている⁽¹⁾. ロボットが作業対象について, より詳細な情報を得る手段として CCD カメラなどの視覚装置を用いることは有用であり, シーン認識の研究は現在さかんに行われている⁽²⁾.

これまで, 人物の顔の認識に関する研究は多数報告されている. 顔の認識の研究の中心分野は, 計算機に顔だけ写っている画像を入力して人物の喜怒哀楽を認識する表情認識⁽³⁾⁽⁴⁾と, 入力された顔の特徴を判断して誰であるかを認識する個人識別⁽⁵⁾⁽⁶⁾の二つが代表的にあげられるが, その前処理として必要不可欠な, シーンからの人物の抽出の研究はあまり行われていない.

本研究が対象としている人物の自動抽出が可能となれば, カメラから得られた画像より, 人物の位置が認識できるので, 例えば看護用ロボット⁽¹⁾などの視覚部

に応用できる. 現在, 日本では高齢化社会の到来に伴い, その高齢者を介護する人材の不足は重大な問題であり, 看護用ロボットは人材不足を解消するための手段として期待されている.

自然画像からの人物の自動抽出に対する研究は, 従来, 写真の中から正面を向いた人物に対して髪の毛と目に着目して抽出する方法⁽⁷⁾, 濃淡画像をニューラルネットワーク(以下 NN と記す.)に入力し人物を抽出する方法⁽⁸⁾が報告されている. しかし, これらの方法では正面を向いた人物しか抽出できない. さらに前者においては濃淡画像を用いて髪と目の黒い部分を探しているため, 実際の自然画像には髪の毛の形の下に二つの孤立点というパターンの画像は多数存在し, この方法では背景が複雑な画像では抽出能力に限度がある. また, 後者においては入力画像の中から一定の大きさのウィンドウを切出し, それを NN に入力することで人物である可能性を評価しているため人物の顔の大きさが最初からわかっていないと抽出できないという欠点がある.

そこで本研究では色情報を用いた自然画像からの人物の自動抽出を目的とする. さらに, コンピュータなどでシーン認識を行う際, 画像処理によって情報劣化が発生し, 実際の認識対象が持つ意味と異なって認識

* 原稿受付 平成4年4月30日.

*¹ 正員, 名古屋大学工学部 (〒464-01 名古屋市中千種区不老町1).

*² 学生員, 名古屋大学工学部.

してしまう錯覚問題が発生することがあるが、このような錯覚問題に対して有効なシステムの開発を目指す。自動抽出の手法としては、色情報を利用して、まず肌色、黒色抽出などの画像処理を行い人物の候補を抽出し、ファジィ推論で「顔の向き」の推定を行い、さらに「人物判定」のため NN を用いる。

従来の研究と比較し本研究の特徴は、色情報を用いたこと、ファジィ推論を用いて「顔の向き」を推定したこと、人物の正面の顔だけでなく横顔の認識も可能としたこと、シーン認識で問題となる錯覚問題を取り上げ、錯覚問題に対して検討したことがあげられる。これらの結果、良好な結果を得られたので報告する。

2. 人物自動抽出システム構成

図 1 にファジィ推論、NN を用いた人物自動抽出システムの構成を示す。このシステムは以下に示す三つの部分から構成されている。

- (1) 画像撮影部
- (2) 画像処理部
- (3) 認識部

画像撮影部では、原画像を CCD カメラによってカラー画像を取り込む。本研究では画像の入力条件を以下のように定めている。

- (1) 人物は服を着ていて服は肌色でないもの
- (2) 髪の毛は存在し黒であるもの
- (3) 人物の周囲に肌色に近い色がないもの
- (4) 適当な露出条件であるもの

画像処理部では肌色、黒色抽出のカラー画像処理や、肌色領域縦横比、両目間の距離、目と口の高さの差、肌色領域に隣接する髪の毛の部分の幅の画像計測を行っている。

認識部では画像処理部から得られる情報で、まず肌色領域と黒色領域の相関関係や肌色領域の縦横比につ

いて評価し、次にファジィ推論によって「顔の向き」を推定し、あらかじめ学習した NN を用いて人物かどうかの最終評価を行っている。

3. 人物候補の抽出

3・1 抽出過程 本システムでは、画像入力後、まずある画素における濃淡値を 4 近傍画素の平均濃淡値に置き換える平滑化処理を行い、続いて肌色抽出を行う。この肌色情報をもとに、肌色領域とその近傍画素をとってウィンドウを作成し、さらにウィンドウ内で黒色の抽出を行う。

3・2 肌色抽出 本研究は、人物特有の色である肌色を人物の候補の重要なパラメータとして最初に肌色抽出を行い、その肌色情報に基づき肌色抽出以後の処理を行っている。

コンピュータによる、色の表現方法としては CIE の XYZ 表色系やマンセルの色空間がありそれらを用いて色を取扱った研究⁽⁹⁾⁽¹⁰⁾があるが、これらの方法では、視感による色の判断と差があると言われており⁽¹¹⁾、また計算機への入力で R, G, B 値に微妙な誤差が生じてくるので実際にこれらを用いて抽出するのは困難である。そのため、本研究では人物の肌色の R, G, B 値を室外で得られた画像より 1000 個採集し

図 1 システム構成図

(a) R-G 直線

(b) R-B 直線

図 2 肌色直線

ラフにプロットした結果、図2のように直線の周辺に点が集まっていることがわかる。ただし、図2(a)の横軸は画素の R 成分濃度値、縦軸は G 成分濃度値、図2(b)の横軸は画素の R 成分濃度値、縦軸は B 成分濃度値を示し、それぞれ濃度値は0~255の値をとる。そこで、最小2乗法により1000個のデータを近似した結果、 R, G, B 値の関係は実験的に以下の式(1)、(2)のように求めた。ただし、肌色と黒色の区別のため肌色は $R > R_0$ と定義し、本研究では経験的に $R_0 = 70$ と決定した。

$$G = 0.803R + 13.9 \dots\dots\dots (1)$$

$$B = 0.715R + 12.1 \dots\dots\dots (2)$$

次に、肌色領域の決定方法は、図3に示すように肌色抽出を2段階に分割し、1回目は G, B がそれぞれ式(1)、(2)の $\pm m_1$ の間の値を肌色として採用し、2回目は $\pm m_2$ の間の値を肌色として採用した。さらに二つの画像をラベリングし、両ラベルを比較し2回目のラベルのついた部分で、どこかに1回目のラベルがついていれば、2回目のラベル全体を肌色領域とした。ただし、本研究では経験的に $m_1 = 10, m_2 = 20$ と決定した。

3.3 処理ウィンドウの作成 肌色抽出で抽出された画素のまわりを以後の処理で扱いやすくするため、原画像にウィンドウを設定した。ウィンドウの設定方法としては、まずある二つの肌色領域間の距離が P 画素以内の時、両肌色領域は同一の肌色領域であると定義する。次に、各領域に外接する長方形より縦横上下、4方向に T 画素分拡大して一つのウィンドウを設定している。本研究では経験的に $P = 5, T = 20$ と

決定した。

3.4 黒色抽出 肌色領域周辺の髪の毛の抽出を行うためにウィンドウ内で黒色抽出を行った。黒は実験的に R, G, B 値がすべて B_0 未満の値でかつ R, G, B 値の平均値の $\pm m_0$ の間の値をとると定義し、本研究では経験的に $B_0 = 70, m_0 = 20$ と決定した。

4. 人物候補の評価

4.1 評価方法 3章で作成した一つのウィンドウを一つの候補と考え、この章ではそれぞれの候補について人物かどうかの評価を行う。評価の方法はまず三つの「人物である条件」について評価することにより候補を限定し、残った候補についてファジィ推論により「顔の向き」の推定を行い、NNで人物かどうかの最終評価を行う。ただし、「顔の向き」は正面から真横までを対象としている。

本研究では人物候補の最終評価をファジィ推論とNNとに分けて行った。「顔の向き」を含めた認識をNNのみで行うと、かなりの学習回数を必要とする。そこで、人間の知識を導入できるファジィ推論によって人物の「顔の向き」という認識しやすいパラメータに変換し、NNの入力とする。

4.2 人物の認識条件 ウィンドウ内で得られた肌色領域と黒色領域が人物のものであるかどうかを評価するため、以下の三つの条件について評価し人物候補の数を絞る。

- 1 肌色領域と黒色領域が接しているか
- 2 黒色領域が肌色領域よりも上にあるか
- 3 肌色領域の縦横比が妥当であるか

上の条件1, 2は黒色領域と肌色領域の相関関係が人物の髪の毛と肌の関係と合っているか評価しているもので、条件3で肌色部分の縦横比が大きすぎたり、小

(a) 肌色の2段階抽出結果

(b) 肌色領域

図3 肌色領域決定方法

図4 ファジィ, NNで用いるパラメータ

さすぎたりしてないかそれぞれ評価し、具体的に「横の長さ/縦の長さ」が0.3~1.2のものを人物の候補として妥当と判断した。

4.3 ファジィ推論を用いた人物の顔の向きを検出

本処理では、簡略化ファジィ推論⁽¹²⁾を用い「顔の向き」の推定を行っている。入力図4に示す四つのパラメータ ($r1, r2, b1, b2$) を用いて以下のように表し、それぞれ A, B, C と定義する。

$$1. r2/r1 \equiv A$$

$$2. b1/r2 \equiv B$$

$$3. b2/r2 \equiv C$$

$r1$: 肌色領域縦の長さ

$r2$: 肌色領域横の長さ

$b1$: 肌色領域の左側に隣接する黒色部分の長さ ($r1$ の上から 1/5 の位置)

$b2$: 肌色領域の右側に隣接する黒色部分の長さ ($r1$ の上から 1/5 の位置)

図5にファジィメンバーシップ関数、表1にルールマップを示す。ただし、メンバーシップ関数、ルールマップは、多数の顔データから経験的に決定した。ファジィ推論、各ルールの適合度 u_{ijk} は、式(3)のように表せる。

$$u_{ijk} = \mu_{Ai} \mu_{Bj} \mu_{Ck} \dots \dots \dots (3)$$

μ_{Ai} : パラメータ ($r2/r1$) に対しての i 番目のメンバーシップに対する適合度

μ_{Bj} : パラメータ ($b1/r2$) に対しての j 番目のメンバーシップに対する適合度

μ_{Ck} : パラメータ ($b2/r2$) に対しての k 番目のメ

図5 メンバーシップ関数

ンバーシップに対する適合度

u_{ijk} : A, B, C それぞれの i, j, k 番目のメンバーシップに対する各ルールの適合度

この式から式(4)によって顔の向き y が推定できる⁽¹²⁾。

$$y = \frac{\sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 u_{ijk} y_{ijk}}{\sum_{i=1}^3 \sum_{j=1}^3 \sum_{k=1}^3 u_{ijk}} \dots \dots \dots (4)$$

y_{ijk} : A, B, C それぞれの i, j, k 番目の山に対する後件部の実数値

ここに y の値は $-90 \sim 90$ の値をとり、正面のとき0、向かって右真横の時 -90 、向かって左真横の時 90 となる。

4.4 ニューラルネットワークを用いた最終評価

本処理ではバックプロパゲーションアルゴリズムで学習したフィードフォワード形NN⁽¹³⁾を用い、人物であるかどうかの最終の評価を行う。NNの入力層は、以下の4ユニットで構成される。

I_1 : 顔の向き (ファジィの出力値)/90

I_2 : 肌色部分の面積/ $r2^2$

I_3 : $d1$ (両目間の距離)/ $r2$

I_4 : $d2$ (目と口の高さの差)/ $r2$

また出力層は以下の4ユニットで構成される。

O_1 : 右横顔

O_2 : 正面の顔

O_3 : 左横顔

O_4 : 人物でない

また中間層は2層各16ユニットで行った。両目間の距離 $d1$ 、目と口の高さの差 $d2$ の求め方はまず肌色領域で横に2個並んだ黒の孤立点を目とみなし $d1$ を求め、それより下にあつてかつ両目間の中点から水平方向のほぼ真下にある孤立点を口と考えその点と両目間の中点との距離を $d2$ とした。ただし、片目しか見えない横顔では両目間の距離 $d1$ は0となる。

NNの入出力関係を数式で表すと式(5)、(6)のよ

表1 ルールマップ

		CS	CM	CB
AS	BS	0	-60	-90
	BM	60	0	-60
	BB	90	60	0
AM	BS	0	-45	-60
	BM	45	0	-45
	BB	60	45	0
AB	BS	0	-30	-45
	BM	30	0	-30
	BB	45	30	0

うになる。

$$I_i = \sum_{j=1}^J W_{ij} O_j \quad (j=1, \dots, J) \quad \dots\dots\dots (5)$$

$$O_i = f(I_i + \alpha_i) \quad (j=1, \dots, J) \quad \dots\dots\dots (6)$$

I_i : ユニット i の入力値

O_i, O_j : ユニット i , 前層ユニット j の出力値

$f(x)$: シグモイド関数

W_{ij} : ユニット i とユニット j の結合係数

α_i : ユニット i のオフセット値

誤差関数 E を式(7)とおき、誤差 E が減少するように学習が進められる。

$$E = \sum_{i=1}^I (T_{pi} - O_{pi})^2 / 2 \quad (i=1, \dots, j) \quad \dots\dots\dots (7)$$

O_{pi} : ユニット i の出力値

T_{pi} : ユニット i の目標値

本論文では室内、室外の画像を画像処理し得られたデータから50個の教師信号をNNに与え、学習を10000回行った。図6に本研究で用いたNNの構造を示す。

NNの出力層のうち、最大となるユニットをもって入力画像の認識結果とする。

5. 実験方法および効果

図7を代表とする入力条件に合った画像を室内10枚、室外10枚、合計20枚撮影し、本システムに入力し人物の抽出をさせた。ただし、女性の画像は室外の物に3枚含まれていた。

表2に示すように、認識率は人物かどうかは室内で90%、室外で70%、顔の向きまで認識できたのは室内で70%、室外で60%であり、室内の画像は室外のものに比べ肌色と区別しにくい色が少ないことからかなり良好な成果が得られている。また、女性の例については3枚とも正解が得られ、本システムが男女を問わず抽出できることを示している。ただし、顔の向きの正答率については、向かって右真横顔を-90度、左真横顔を90度とし、-90~-30度の向きの顔が入力され

図6 本研究で用いたNN

たときに右横顔、-30~+30度の向きの時正面顔、+30~+90度の向きの時左横顔と認識されればそれぞれ正解とした。

例として、図7~図10に原画像（図7）を本システムに入力したときの一連の処理結果を示す。まず肌色抽出によって図8が得られ、その肌色情報に従ってウィンドウを設定し、ウィンドウ内で黒色を抽出したものが図9である。さらに各ウィンドウについてファジィ推論、NNで評価した結果、図10が得られる。本例では図8、図9のように肌色、黒色抽出がかなりうまく行われたため、図10に示すように良好な結果が得られている。

また、今回の実験で誤答であった原因は、色抽出での失敗、詳細情報計測での失敗の二つに分けられる。色抽出の失敗例としては、背景が全面に土であったため、その土の部分が肌色抽出で抽出されてしまったものや、髪の毛が光っていて黒色として抽出されなかったものである。今後、さらに微妙な色の差を識別できるようにシステムへと、本手法の改善が要求される。また、詳細情報計測で失敗した例は髪の毛が目にかかっていた場合に、両目間の距離が計測できなかったものである。本手法では前処理において画像処理を行っているが、この際、ある種の画像劣化が生じると、認識結果に大きく影響する。このような画像劣化に対し、いかに対応するかが今後の課題である。ここでは、画像劣化が発生した場合を想定し、以下に錯覚問題を例にとって本システムの判断能力について調査する。

6. 錯覚問題に対する検討

6.1 錯覚とは 人間が視覚によって対象物を認識する場合、目に入力されるさまざまな視覚情報を利用し対象物を認識する⁽¹⁴⁾。しかし、目に入力される画像自体の持っている情報量が対象物本来の意味を認識させるに不十分である場合、人間は錯覚をおこすことが知られている⁽¹⁵⁾。

このような錯覚問題は計算機で認識問題を扱う場合についても同様に情報量不足が原因で起こりうる。計

表2 認識結果

	Whether Person or Not	Facial Direction
Indoor	90 %	70 %
Outdoor	70 %	60 %

算機で対象物を認識する際、まず画像が計算機に入力される時画像劣化は避けられないし、また認識に必要なパラメータの計測のために画像を画像処理などにより変換すればそのたびに情報劣化が伴う。ここでは、ある種の画像劣化を想定し、そのような状況での本システムの認識能力について検討する。

6・2 錯覚問題への対応 図 11 は心理学でよく錯覚の例として取り上げられるフィッシャーの「男と少女」⁽¹⁵⁾である。この絵は②から⑦の絵が、①の男性の顔と⑧の女性の裸体のどちらともとれるという問題である。ここでは、②～⑦のような、男性の顔か女性の裸体か判断がむずかしいような抽象的な画像が画像処理

図 7 原画像

図 8 肌色抽出結果

(a) 候補 1

(b) 候補 2

図 9 ウィンドウ作成結果

(a) 左横顔

(b) 正面顔

図 10 最終結果

図 11 フィッシャー「男と少女」

によって得られたとき、本システムが②～⑦のどこまでを男性の顔と認識するかを調べる。ただし、本研究では、裸の人物を対象としていないので女性の裸体かどうかの評価はできない。

まず、図 11 では顔の輪郭線の内部は肌色領域であると考えれば、①の男性の顔は画像処理が成功した例と考えられ、また②～⑦は画像処理の結果①が情報劣化したものと考えられる。そこで②～⑦を本システムに入力し、両目間の距離などの詳細情報を用い NN で判断した。ただし、ここでは錯覚の画像から得られるデータを NN の教師信号に用いていない。この結果、③までが人物の顔であると判断された。この理由としては④～⑦は目と鼻がつながっていることにより、目が認識されなかったことが原因である。また、実際に我々人間がこの図をみて明らかに男性の顔であるとわかるのは②、③であろう。以上のことより、本システムでは錯覚問題を考えた場合、人間の視覚判断とほぼ近い結果が得られていることがわかる。

7. ま と め

本研究では自然画像の中から、色情報を基に、ファジィ推論、ニューラルネットワークを用いて人物を自動的に抽出する手法を提案し、従来の研究では困難であった真横を向いた人物の抽出も人物の「顔の向き」

を考慮することで可能とした。

また本研究では錯覚問題について取り上げ、本システムに抽象的な絵を入力することによって錯覚問題に関する検討を行った。

文 献

- (1) 中野, ロボット工学の歴史と展望, 日本ロボット学会誌, 8-58 (1990), 58.
- (2) 黒木・長田, 視覚ロボットによる位置決め, 日本ロボット学会誌, 9-3 (1991), 109.
- (3) 小林・原, ニューラルネットワークによる人物の表情認識—心情カテゴリーによる基本表情認識—, 日本ロボット学会講演会予稿集, (1991), 751.
- (4) Mase, K., Recognition of facial expression from optical flow, *JEICE Trans.* E-74, 10, (1991), 3474.
- (5) 南, 顔画像を用いた個人識別, 日本機械学会オーディオ・ビジュアルシンポジウム講演論文集, (1992), 15.
- (6) Kerin, M. A. and Stonham, T. J., Face Recognition Using a Digital Neural Network with Self-organising Capabilities, *Proc. of IEEE 10th International Conference on Pattern Recognition*, 1 (1990), 739.
- (7) 安居院・ほか 2 名, 人物の写っている画像からの顔の抽出, 電子情報通信学会春期全国大会講演論文集, (1991), 7.
- (8) 安居院・ほか 2 名, 神経回路網を用いた白黒集合写真からの顔領域の抽出, 電子情報通信学会春期全国大会講演論文集, (1992), 7.
- (9) 堀田・ほか 2 名, 均等色空間に基づくカラー画像の領域分割, 電子情報通信学会論文誌 D-II, J74-D-II, 10 (1991), 1370.
- (10) 石橋・岸野, 領域分割を用いた人物像の色彩テクスチャ分析・合成, 電子通信学会論文誌 D-II, J74-D-II, 11 (1991), 1546.
- (11) 尾崎・谷口, 画像処理 (第 2 版)—その基礎から応用まで—, (1988), 15, 共立出版.
- (12) 市橋・田中, PID と FUZZY のハイブリッド型コントローラー, 第 4 回ファジィシステムシンポジウム講演論文集, (1988), 97.
- (13) Rumelhart, D. E., McClelland, J. L. and The PDP Research Group, *Parallel Distributed Processing*, Vol. 1 (1986), 547; Vol. 2 (1986), 611. The MIT Press.
- (14) Lam, D.M.-K Gilbert, C.D., *Neural Mechanisms of Visual Perception*, (1989), 157, Portfolio Publishing Company.
- (15) アトニー・F., 物の見え方と知覚システム, 別冊サイエンス, 特集視覚の心理学イメージの世界, (1975), 61, 経サイエンス社.