

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2005-316888

(P2005-316888A)

(43) 公開日 平成17年11月10日(2005.11.10)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
GO6T 1/00	GO6T 1/00 340A	5B050
GO6T 7/00	GO6T 1/00 200A	5B057
	GO6T 1/00 315	5L096
	GO6T 7/00 350C	

審査請求 未請求 請求項の数 4 O L (全 25 頁)

(21) 出願番号	特願2004-136394 (P2004-136394)	(71) 出願人	503360115 独立行政法人科学技術振興機構 埼玉県川口市本町4丁目1番8号
(22) 出願日	平成16年4月30日 (2004.4.30)	(74) 代理人	100105371 弁理士 加古 進
		(72) 発明者	呂 建明 千葉県千葉市稲毛区轟町4-6-23-301
		(72) 発明者	谷萩 隆嗣 東京都文京区千駄木5-19-10
		(72) 発明者	関屋 大雄 東京都墨田区緑2-13-7 アーバンハイツ両国1010
		Fターム(参考)	5B050 BA12 CA02 EA05 EA07 EA12 EA13 EA18 EA27 GA08

最終頁に続く

(54) 【発明の名称】 顔認識システム

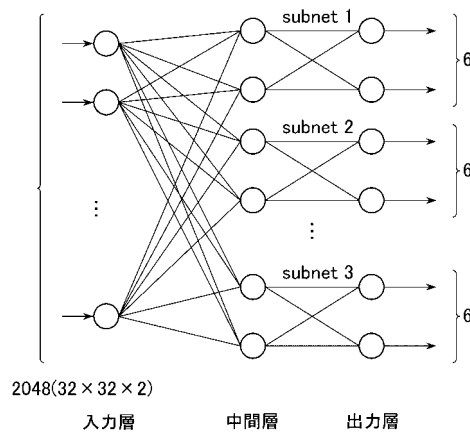
(57) 【要約】

【課題】 顔の位置・角度の変動や照明の変動を吸収でき、識別対象となる人数が増加しても効率のよい学習と認識が可能な顔認識システムの提供。

【解決手段】 本発明は、図1に示すようにニューラルネットワーク(NN)を複数のサブネットに分割し、サブネット単位で並列に学習と認識を行なう。また並列NNの構成にあたって、パターンのおいまいさを考慮してファジィ的に複数のサブネットに同一顔パターンが帰属することを許している。これにより、比較的近いパターン同士は必ず同じサブネットで効率よく認識させることができる。

また、本発明では識別対象人物の顔パターンを、3次元形状計測装置により3次元映像として取得している。このため、人物の位置、姿勢の変化や照明の変動などの影響を極力排除・補正した入力顔パターンを得ることが可能となり、顔の回転にロバストな顔認識システムを実現している。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

ニューラルネットワークを利用した顔認識システムであって、

3次元形状計測装置から取得した顔画像データ及び3次元形状データから鼻の頂点を含む特徴点を抽出する特徴点抽出手段と、

前記抽出した特徴点の回転に対して顔画像データの補正を行い正面を向いた画像とする回転補正手段と、

前記鼻の頂点を用いて、一定の領域を切り出す顔領域切り出し手段と

前記一定の領域として切り出された顔画像データの平均値を一定に補正する照明補正手段と、

前記照明補正された顔画像データおよび対応する3次元形状データを、ファジィクラスタリングにより、1人のデータが複数のクラスタに帰属されるように、各々のクラスタに学習した並列ニューラルネットワークと

を備え、前記照明補正された顔画像データおよび対応する3次元形状データを前記学習した並列ニューラルネットワークに対して与えて認識することを特徴とするニューラルネットワークを利用した顔認識システム。

10

【請求項 2】

請求項 1 に記載したニューラルネットワークを利用した顔認識システムにおいて、

特徴点抽出手段は、特徴点として、前記鼻の頂点以外に、鼻の両側の点、目の位置の点を抽出することを特徴とするニューラルネットワークを利用した顔認識システム。

20

【請求項 3】

請求項 1 又は 2 に記載したニューラルネットワークを利用した顔認識システムにおいて、

前記並列ニューラルネットワークの学習過程で、顔画像をクラスタに分けて学習させるときに、振り分けを最適化するために、1人のデータをいくつのクラスタに帰属するか、1つのクラスタに最大何人のデータを帰属するかを調整することを特徴とするニューラルネットワークを利用した顔認識システム。

【請求項 4】

請求項 1 ~ 3 のいずれかに記載したニューラルネットワークを利用した顔認識システムの機能を、コンピュータ・システムに実現させるプログラム。

30

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ニューラルネットワークを利用した顔認識システムに関するものである。

【背景技術】

【0002】

情報化社会の急激な変化につれて、セキュリティに関する要求が高まっている。その中でも、人間は顔を見て自然に相手を判断できることから、機械においても顔認識を利用したセキュリティシステムを構築しようという研究が盛んに行なわれている（非特許文献 1）。

40

また、顔認識を利用したセキュリティシステムは、指紋パターンや眼底パターンを用いた個人識別に比べ、認識される側の心理的抵抗感が少ないというメリットもある。撮影した顔画像から正面顔パターンを切り出して個人識別を行なう方法はこれまでに多数提案されている。これらはいずれも、入力パターンを特徴ベクトルで表現し、あらかじめ登録したそれぞれの人物を代表する標準ベクトルとの距離を求め、もっとも距離が小さい人物を同一人物と判定する方法を用いている。これらを大きく分類すると、顔パターンの幾何学的特徴を用いる方法と大局的な濃淡パターンを用いる方法の 2 つに分けられる。

顔パターンの幾何学的特徴を用いる方法としては、顔の幅、両目の中心間の距離、頭頂からの目の高さなどの顔の特徴点をパラメータ化して、その特徴量を用いて認識を行なう方法や、横顔の輪郭線などの形状を用いて認識を行なう方法（非特許文献 2）などがある

50

。しかし、顔の回転などによって変動した顔画像から各特徴点や形状を正確に抽出することは困難であるという問題がある。

一方、大局的な濃淡パターンを用いる方法は、顔画像を各画素の濃淡値の特徴ベクトルとして考え、マッチングによって識別を行なう方法である。しかし画像内の全画素をマッチングに用いてしまうと、次元数が膨大となり、計算量が大きくなるという問題がある。

【0003】

そこで、入力ベクトルを低次元に圧縮する方法が検討されてきた（非特許文献3）。この方法の代表として、Turkらによる固有顔法（非特許文献4）や、顔画像をモザイク化してニューラルネットワーク（以降「NN」という）に学習させる方法などがある（非特許文献5）。

しかしながら、これらの顔認識方法では、顔の位置・角度の変動や照明の変動によって認識率が低下するという問題がある。これに対応する方法としては、あらかじめ複数方向から撮影した顔画像を用意しておき、入力パターンと、その角度に近い登録パターンとをマッチングする方法（非特許文献6）などが提案されている。しかし、この方法によっても、2次元入力画像の特徴点などから顔画像の角度を求めるのが困難であるという問題があり、回転・照明変動を吸収できる顔認識方法への要望は高い。

また、非特許文献3, 5では、個人認識にニューラルネットワーク（NN）を適用した結果が示され、NNが顔画像の認識にも有用であることが述べられているが、識別の対象となる人数の増加と共にNNが膨大なものになってしまうという問題が指摘されている。すなわち、従来のバックプロパゲーションを利用したNNでは、人数が増えると、NNの規模が大きくなる。NNの規模が大きくなるにつれて、ローカルミニマムなど種々の原因によって、学習の効率が急に下がり、認識率も急速に落ちてしまう。

このような欠点を克服するための方法として、複数の小規模なNNを作り、それらを統合して認識を行なうことのできるCombNET-IIが提案されている（非特許文献7）。CombNET-IIは大規模なシステムにも適用可能な効率の良いアルゴリズムである。

しかしながら、CombNET-IIは各クラスすなわち各小規模なNNに属するパターンが重複しない手法（以降「ハードクラスタリング」とよぶ）を用いているため、近いパターンでもNNで学習されないものが存在する。このため、これを顔画像の認識に応用する際に、パターンの変動が大きいときには、認識率が大幅に低下すると考えられる。

【0004】

【非特許文献1】赤松茂 “コンピュータによる顔の認識 - サーベイ - , ” 電子情報通信学会論文誌 A vol.J80-A, no.8, pp.1215-1230, Aug.1997

【非特許文献2】阿倍亨 阿曾弘具 木村正行 “3次元データより得られる横顔輪郭線を用いた人間の顔の自動識別, ” 電子情報通信学会論文誌 D-II vol.J73-D-II, no.9, pp.1468-1476, Sep.1990

【非特許文献3】小杉信 “モザイクとニューラルネットを用いた顔画像の認識, ” 電子情報通信学会論文誌 D-II vol. J76-D-II, no.6, pp.1132-1139, Jun.1993

【非特許文献4】M. Turk and A. Pentland "Eigenfaces for Recognition," Journal of Cognitive Neuroscience vol. 3, no.1, pp.71-86, Mar.1991

【非特許文献5】谷萩隆嗣 高野裕昭 “カテゴリーを組み合わせたNNによる顔画像認識, ” 電子情報通信学会論文誌 D-II vol.J77-D-II, no.11, pp.2151-2159, Nov.1994

【非特許文献6】平野晃昭 中村納 “顔の回転にロバストな個人識別方式, ” 電気学会論文誌 C vol.120-C, no.6, pp.800-808, Sep.2000

【非特許文献7】堀田健一, 岩田彰, 松尾啓志, 鈴村宣夫 “大規模ニューラルネット “CombNET-II”, ” 電子情報通信学会論文誌 D, Vol.J-75-D, No.3, pp.545-553, 1992

【発明の開示】

【発明が解決しようとする課題】

【0005】

本発明の課題は、上記の問題を解決して、顔の位置・角度の変動や照明の変動を吸収で

10

20

30

40

50

き、識別対象となる人数が増加しても効率のよい学習および認識が可能な顔認識システムを提供することである。

【課題を解決するための手段】

【0006】

上記の課題を解決するために、本発明は、ニューラルネットワークを利用した顔認識システムであって、3次元形状計測装置から取得した顔画像データ及び3次元形状データから鼻の頂点を含む特徴点を抽出する特徴点抽出手段と、前記抽出した特徴点の回転に対して顔画像データの補正を行い正面を向いた画像とする回転補正手段と、前記鼻の頂点を用いて、一定の領域を切り出す顔領域切り出し手段と前記一定の領域として切り出された顔画像データの平均値を一定に補正する照明補正手段と、前記照明補正された顔画像データおよび対応する3次元形状データを、ファジィクラスタリングにより、1人のデータが複数のクラスタに帰属されるように、各々のクラスタに学習した並列ニューラルネットワークとを備え、前記照明補正された顔画像データおよび対応する3次元形状データを前記学習した並列ニューラルネットワークに対して与えて認識することを特徴とするニューラルネットワークを利用した顔認識システムである。

10

ここで、特徴点抽出手段は、特徴点として、前記鼻の頂点以外に、鼻の両側の点、目の位置の点を抽出することを特徴とすることもできる。

さらに、前記並列ニューラルネットワークの学習過程で、顔画像をクラスタに分けて学習させるときに、振り分けを最適化するために、1人のデータをいくつかのクラスタに帰属するか、1つのクラスタに最大何人のデータを帰属するかを調整することを特徴とすることもできる。

20

上記のいずれかに記載したニューラルネットワークを利用した顔認識システムの機能を、コンピュータ・システムに実現させるプログラムも、本発明である。

【発明の効果】

【0007】

本発明の顔認識システムを用いることによる効果は、大きく下記の2つである。

第一に、本発明では大規模なNNを複数の小規模NN(サブネットワーク)に分割し、サブネットワーク単位で並列に学習と認識を行なうことで、ローカルミニマムなどに起因する大規模NNでの認識率低下を回避することができる。並列NNの構成にあたっては、従来、登録者の顔パターンで相互に類似性の高いものを集めてひとつのサブネットワークとする方法が採用されており、異なるサブネットワーク間では同じパターンが存在することはなかったが、本発明ではパターンのあいまいさを考慮してファジィ的に複数のサブネットワークに同一顔パターンが帰属することを許しているのが大きな特徴である。これにより、比較的近いパターン同士は必ず同じサブネットワークで効率よく認識させることができる。

30

第二に、本発明では識別対象人物の顔パターンを、3次元形状計測装置により3次元映像として取得した。すなわち、従来の2次元映像にはない顔面の高さ情報等が利用できる。このため、人物の位置、姿勢の変化や照明の変動などの影響を極力排除・補正した入力顔パターンを得ることが可能となり、顔の回転にロバストな顔認識システムを実現することができた。

40

【発明を実施するための最良の形態】

【0008】

本発明では、上述の非特許文献7のCombNET-IIと同様に、複数の小規模なNN(以降「サブネットワーク」と呼ぶ)を作って図1に示すような並列NNを構成し、サブネットワーク単位で並列に学習と認識を行なうことで、大規模な認識システムにおいても効率良く学習や認識ができるようにする。なお、図1に示す並列NNは、入力層に2048ノード、中間層に32ノード、出力層に6ノードを持った3階層NNの例であるが、これらの値は、後のシミュレーションで例として用いる値である。

しかしながら、従来のCombNET-IIでは、同じパターンが1つのサブネットワークにしか属することができないハードクラスタリングの手法を用いているため、近いパタ

50

ーンでもNNで学習されないものが存在する。このため、これを顔画像の認識に応用する際に、パターンの変動が大きいときには、認識率が大幅に低下すると考えられる。

この問題を解決するため、本発明では、ハードクラスタリングではなく、ファジィクラスタリングを利用して顔パターンを分類する方法を提案する。ファジィクラスタリングは帰属度にあいまいさを許して、1つのパターンをいくつかのサブネットワークに帰属させることを認めている。したがって、各クラスすなわち各サブネットワークに属するパターンが重複しているので、比較的近いカテゴリ同士は必ず同じNNで効率良く認識させることができる。

また、従来の顔認識システムにおいては、人物の位置及び姿勢の変化によって回転した顔画像に対しては、識別精度が若干低下するということが問題となっている。この人物の位置および姿勢の変化による画像変動は、本質的に顔の3次元形状によるものであり、これを解決するためには顔の3次元形状を利用することが有効であると考えられる。そこで本発明では、3次元形状計測装置を用い、3次元形状を利用して顔画像認識を行なう。

10

【0009】

本発明の顔認識システムのシステム構成の例を図23に示す。なお、図23には顔の学習と認識の両方を行なう場合のシステム構成を示しているが、実際にセキュリティシステムとして利用する際には、学習と認識を別々のコンピュータで行なう場合も考えられる。

図23において、コンピュータ120は、本発明の顔認識システムで顔の学習および認識の処理を行なうコンピュータである。コンピュータ120は、例えば通常のパソコンなどを用いることができるが、他のコンピュータであってもよい。記憶領域122には、顔の学習および認識の処理を行なうプログラムが格納されており、CPU123が処理を行なう。

20

3次元形状計測装置110は、顔認識システムに登録される者(登録者)の顔を撮影し、その3次元情報を得る装置である。また、入力装置130は、キーボードやマウスなどの入力装置であり、顔認識システムに顔を登録する(学習させる)際に、登録者の氏名などの情報を入力したり、顔認識システムの操作を行なったりするための入力装置である。3次元形状計測装置110および入力装置130からの入力は、入力I/F121を介してコンピュータ120に入力され、処理される。

コンピュータ120で処理された結果は、出力I/F124を介して出力される。学習過程においては、例えば処理の途中で生成されたデータや学習が終わった旨の信号などであり、必要に応じてディスプレイなどの出力装置140に出力するようにしてもよい。また、認識過程においては、認識または拒否の信号をセキュリティ機器などに出力する。

30

【0010】

以降、本発明の顔認識システムの実施形態を詳細に説明する。

まず、1.で本実施形態の顔認識システムで用いるファジィクラスタリングの手法を説明し、2.でファジィクラスタリングを用いて顔画像を振り分ける手順を説明する。

次に、3.でNNによる顔の学習過程を、4.でNNによる顔の認識過程を説明する。

次に、5.で顔画像の3次元情報からNNに入力するデータを作成する過程を説明する。

最後に、6.で学習と認識のシミュレーションを行ない、本実施形態の手法の有効性を示す。

40

【0011】

<1.ファジィクラスタリング>

パターン認識では認識対象に対して、人間の主観や個性など複雑な要素が関係してくる。すなわち、扱う対象の性質に、真偽の2値のみでは説明できないあいまいさが存在する。また、認識側にも主観や個性などに関係するあいまいさが存在する。

パターン認識の分野においては、扱うべきデータ集合のそれぞれのデータ・ベクトルをあらかじめ決められたどのクラスタに属しているかを決定するファジィクラスタリング法が多用されている。そこでFCM(fuzzy c-means)法を利用して顔画像の認識を行ない、パターンの振り分けを最適化することで、認識率を向上させることができると考えられ

50

る。

ここではまず、従来のハードクラスタリングの手法について簡単に説明し、本実施形態で用いるファジィクラスタリングの手法および、FCM (fuzzy c-means) 法を用いて分割行列を決定する方法を説明する。

【0012】

(1) ハードクラスタリング

各データを分割する際、ただ1個のクラスタにのみ帰属させる方法はハードクラスタリングと呼ばれる。

データ集合を $X = \{x_1, x_2, \dots, x_n\}$ ($x_j \in R_s, j = 1, 2, 3, \dots, n$) とした場合、データ x_j がクラスタ C_j に属するかどうかを行列の (i, j) 要素で表現する。 n 個のデータを c 種類のクラスタに分割する場合、行列 U は $c \times n$ の行列となる。そして要素 u_{ij} は0または1の値を取り、 $u_{ij} = 1$ の場合データ x_j がクラスタ C_j に属さないことを示す。各クラスタには少なくとも1個以上のデータが必ず属し、各データはどれか1個のクラスタに属している。

10

【0013】

(2) ファジィクラスタリング

しかし、クラスタは本来明確に定義できるものではなく、パターン集合をカテゴリ分解する際に、境界部分のパターンはいくつかのカテゴリに関連を持つ。ハードクラスタリングでは、いずれか1つのカテゴリに帰属させることになるが、ファジィクラスタリングでは帰属度にあいまいさを許して、複数のカテゴリに帰属することを認める。

20

ファジィクラスタリングの手法を定式化すれば、以下のようになる。

取り扱うデータ集合 n 個の s 次元の分類対象を

【数1】

$$X = \{x_1, x_2, \dots, x_n\} (x_j \in R_s)$$

とする。ファジィクラスタリングによる分割行列を

【数2】

$$U = \{u_{ij} \mid (i \in [1, c], j \in [1, n], u_{ij} \in [0, 1])\} \quad (1)$$

30

とすると、 u_{ij} は以下の2つの条件を満足する。

【数3】

$$\sum_{i=1}^c u_{ij} = 1 \quad (2)$$

【数4】

$$0 < \sum_{j=1}^n u_{ij} < n \quad (3)$$

40

ただし、 c は $2 \leq c \leq n$ で、クラスタの数である。 u_{ij} は閉区間 $[0, 1]$ の値を取り、 x_j のクラスタへの帰属度を表す。

【0014】

(3) FCM 反復法

次に、FCM (fuzzy c-means) 法を用いて分割行列 U を決定する方法を述べる。

FCM 法は、

【数 5】

$$J_m(U, V) = \sum_{i=1}^c \sum_{j=1}^n (u_{ij})^m \|x_j - v_i\|^2 \quad (4)$$

として与えられる目的関数の最小化により、分割行列 U を決定する反復方法である。ここでは $m \geq 1$ である。 $V = \{v_1, v_2, \dots, v_c\}$ ($v_i \in R_s, i \in [1, c]$) はクラスタ中心ベクトルの集合である。 v_i は x_j のメンバシップ度の m 次重み付き平均である。このことは、より小さいメンバシップ度の x_j より、より大きいメンバシップ度の x_j が v_i により大きい影響を与えることを示している。この傾向は、 m を大きくすればするほど、より大きくなる。

10

さらに、 G を正定対称行列として、一般化ノルム

【数 6】

$$\|x_j - v_i\|_G^2 = (x_j - v_i)^T G (x_j - v_i) \quad (5)$$

を導入すれば、ファジィクラスタリング問題は次のように一般化される。

【数 7】

$$J_m(U, V) = \sum_{i=1}^c \sum_{j=1}^n (u_{ij})^m \|x_j - v_i\|_G^2 \quad (6)$$

20

【0015】

< 2 . 顔画像の振り分け >

上述の 1 . で述べたように、ファジィクラスタリングを利用してパターンの振り分けを最適化することで、認識率を向上させることができる。本実施形態においても、この方法で顔画像の振り分けを行なう。

パターンの振り分けは、まず (1) 各パラメータを初期化して、(2) 反復収束法により分割行列を得て、(3) 分割行列の各要素を非ファジィ化することにより行なう。以降、これらを順に説明する。

30

【0016】

(1) パラメータの初期化

まず、クラスタ数、クラスタ合併の調節係数、重み、収束判定値、計算回数などの各パラメータの初期値を設定する。

【0017】

(2) 反復収束法の計算

(a) まず、クラスタの平均ベクトル $v_i^{(1)}$ を

【数 8】

$$v_i^{(l)} = \frac{\sum_{j=1}^n (u_{ij})^m x_j}{\sum_{j=1}^n (u_{ij})^m} \quad (7)$$

40

から求める。ただし、1 回目の反復における U を U^1 としたとき、式 (7) の u_{ij} は U^1 の要素を表す。

【0018】

(b) クラスタの合併

50

【数 9】

$$D_{ik} = \left\| v_i^{(l)} - v_k^{(l)} \right\|, (i \in [1, c], k \in [1, c])$$

において、もし $D_{ik} < \min_d$ (クラスタ間の最小距離) しかも $D_{ik} = 0$ であれば、上述の式 (2), (3) の条件のもとで、以下の合併作業を行なう。

【数 10】

$$u_{ij} = u_{ij} + u_{kj} \quad (8)$$

10

【数 11】

$$u_{ij} = 0 \quad (9)$$

ただし $j \in [1, n]$ である。

【0019】

(c) $x_j \in v_i^{(l)}$ の時は、以下の式

【数 12】

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|x_j - v_i^{(l)}\|_G^2}{\|x_j - v_k^{(l)}\|_G^2} \right)^{1/(m-1)}} \quad (10)$$

20

により U^l を U^{l+1} に更新する。

$x_j \in v_i^{(l)}$ の時は、

30

【数 13】

$$u_{ij} = \begin{cases} 1, j = i \\ 0, j \neq i \end{cases}$$

とおく。

【0020】

(d) 与えられた収束判定値 ϵ に対して、

40

【数 14】

$$\|U^{(l+1)} - U^{(l)}\|_G \leq \epsilon$$

となれば終了。そうでなければ、 $l = l + 1$ として上記 (a) に戻る。

【0021】

(3) メンバシップ関数の非ファジィ化

以上の計算で得た分割行列 U の各要素を非ファジィ化する。ここで、閾値の調節によって、カテゴリの重複度を調節し、顔画像の振り分けを最適化する。

(a) 行の非ファジィ化

50

横方向は1つの要素がいくつのカテゴリに属するかを決める。1つの要素は最大n個のカテゴリに属することができる。これ以上になると、冗長度が大きくなり、ネットワークの学習に大変時間がかかってしまう。

(b) 列の非ファジィ化

縦方向は1つのカテゴリにいくつの要素を含むかを決定する。1つのカテゴリに最大c個の要素を包括できる。これ以上となると、ネットワークの認識率が著しく下がってしまう。しかも、学習時間も急に長くなる。重複の行を消して、1つのカテゴリにする。

本実施形態では、例として、1人は同時に5つのサブネットワークに帰属することができるように、また、1つのサブネットワーク中に最大6人が帰属することができるように、メンバシップ関数を非ファジィ化するとき、1つのパターンが帰属できるクラスタ数を決めるクラスタへの最大帰属数を5に設定する。また、クラスタに含まれる要素数(パターンの数)を決めるクラスタの最大冗長度(クラスタ内の最大要素数)を6に設定する。

10

【0022】

< 3 . NNの学習過程 >

統合型のネットワークでは、個々のネットワーク(サブネットワーク)での学習の負荷を低減するために、1つのネットワークが扱うカテゴリ数を減らしている。その際、学習用の各パターンを各サブネットワーク(すなわち各クラスタ)にどのように割り振るかが重要なポイントとなる。

例えば、上述の非特許文献5の方法では、少なくともN番目までの近いカテゴリの対がどこかのサブネットワークに属していることを保証するクラスタリングを用いている。

20

本実施形態では、次の手順で学習を行なう。

最初に、各カテゴリの標準パターンを求める。次に、各標準パターン間のユークリッド距離を求める。例えば、いま、R種類のパターンの認識を考えているとしよう。カテゴリ数はRなので、最初にR個のサブネットワークを用意する。サブネットワークiでは、カテゴリiから近い順にN番目までのカテゴリが認識対象になっている。ここで、サブネットワークiのカテゴリiをメインカテゴリ、それに近いN番目までのカテゴリをサブカテゴリと呼ぶことにする。

あるサブネットワークを作っていく過程で、そのメインカテゴリとサブカテゴリがすでに前のサブネットワークで一緒になっていたら、そのサブカテゴリを削除する。サブカテゴリが全て削除されてしまったらそのサブネットワークも削除する。

30

このような方法を用いると、カテゴリの数以下のサブネットワークで少なくとも、N番目までの近いカテゴリの対がどこかのサブネットワークに属していることが保証される。

また、各サブネットワークの規模はそろっている方が望ましい場合が多い。そのため、扱うカテゴリ数が少ないサブネットワークの結合を行なう。

【0023】

上述の学習過程をまとめると、図21に示すフローチャートのようなになる。なお、各サブネットワークは、バックプロパゲーションアルゴリズムで学習を行なう。

(1) 各カテゴリの標準パターンを作る(S310)。

(2) $i = 1$ と置く(S320)。

(3) 標準パターンP(i)からN番目までに近い標準パターンP(j) ($j > i$)を探す(S330)。すなわち、P(j)はN個探索される。

40

(4) すでにP(i)とP(j)が同じサブネットワークに属しているかを調べ、属していたら、そのP(j)を削除する(S340)。

(5) 上記(4)の結果、P(j)が0個になったら(S350でYesの場合)、サブネットワークは作らない。一方、P(j)が0個でなければ、標準パターンP(i)とP(j)の合計(N+1)個のカテゴリが属するサブネットワークを作成する(S360)。

(6) $i < R$ (Rはカテゴリ数)ならば、 $i = i + 1$ として上記(3)のステップ(すなわち、図21のS330)に戻る。

同様に上記(3)~(5)のステップを繰り返して、iがRに達したら(S350でN

50

o の場合)、次に、扱うカテゴリ数が少ないサブネットワークの結合を行なう (S 3 9 0)。

【0024】

サブネットワークの結合処理 (図 2 1 の S 3 9 0) のフローチャートを、図 2 2 に示す。

(7) $k = 1$ と置く (S 4 1 0)。

(8) $n = k + 1$ と置く (S 4 2 0)。

(9) 2 つのサブネットワーク k, n を考え、それらが扱うカテゴリ数の合計が $(N + 1)$ 個以下であれば (S 4 3 0 で Yes の場合)、ネットワーク k と n を統合する (S 4 4 0)。

10

(10) $n < M$ (M は上述の (6) のステップが終了したときのサブネットワーク数) ならば (S 4 5 0 で Yes の場合)、 $n = n + 1$ として上述の (9) のステップに戻る。

(11) $n \geq M$ の場合 (S 4 5 0 で No の場合)、 $k < M$ ならば (S 4 7 0 で Yes の場合)、 $k = k + 1$ として上述の (8) のステップに戻る。

これを、 $k \geq M$ になるまで (すなわち、S 4 7 0 で No になるまで) 繰り返す。

【0025】

< 4 . NN の認識過程 >

学習した後の並列 NN の重みを利用して、認識の対象となる顔画像データを NN の構造に従って計算する。次に、FCM 法で得たサブネットワークの帰属情報によって、認識の判断を行なう。基本的には、近いもの同士はサブネットワークによって行なわれ、遠いもの同士は類似度法により判別されることになる。また、1 つのサブネットワークに否定されたパターンは、他のサブネットワークにある場合にも否定される。

20

本実施形態では次の (1) ~ (2) の手順で認識を行なう。

前提として、上記の < 3 . NN の学習過程 > で述べた方法で、登録者の学習用顔パターンを並列 NN に学習させ、登録する。そして、登録者の認識用顔パターンと未登録者の顔パターンを、学習を行なった並列 NN に認識させる。

【0026】

(1) 並列 NN の否定能力による排除

まず、ある NN (あるサブネットワーク) で学習しているにも関わらず出力が最大とならなかったパターンを排除する。また、あるサブネットワークでは、すべての要素の値が閾値より低い場合、このサブネットワークのすべてのパターンを排除する。本実施形態ではこの閾値を 0 . 5 にした。さらにあるサブネットワークで一度否定された顔パターンを他のサブネットワークの答えから排除する。このとき、すべての顔パターンが排除されたら、この入力顔パターンを未登録者と判断する。

30

(2) 類似度法による判断

(1) が行なわれた後、答えが残った場合、類似度法による判断を行なう。答えとなる要素 (登録者) の標準顔パターン (学習用顔パターン) と入力顔パターンの類似度

【数 1 5】

$$\text{similarity}_c = \frac{1}{K^{(l)}} \sum_{k=1}^{K^{(l)}} \frac{(x, x_k^{(l)})}{\|x\| \cdot \|x_k^{(l)}\|} \quad (11)$$

$$(l = 1, 2, \dots, L)$$

40

を計算し、類似度が一番大きい要素ユニットを最後の答えとする。ここで、 K は標準パターンの数で、 L は人数である。

ただし、類似度がある閾値より小さいとき、この入力顔パターンを未登録者と判断する。

50

【 0 0 2 7 】

< 5 . 3 次元情報の利用 >

上述では、ファジィクラスタリングと NN を用いて、あいまいな顔画像を効率よく処理する方法を示した。

しかし、人物の位置及び姿勢の変化によって回転した顔画像に対しては、識別精度が若干低下するということが問題となっている。この人物の位置および姿勢の変化による画像変動は、本質的に顔の 3 次元形状によるものであり、これを解決するためには顔の 3 次元形状を利用することが有効であると考えられる。そこで本実施形態では、3 次元形状計測装置を用い、3 次元形状を利用して顔画像認識を行なう。

以降、本実施形態の顔認識システムにおける、顔画像の 3 次元情報から NN に入力するデータを作成する過程を説明する。なお、本実施形態のシステム構成例は上述の図 2 3 に示したとおりである。 10

【 0 0 2 8 】

(1 . 画像の撮影)

撮影装置としては、図 2 に示すような非接触型 3 次元形状計測用レンジファインダを用いる。後述する本実施形態のシミュレーションでは、3 次元形状計測用レンジファインダとして Danae - R (http://www.necsan-ei.co.jp/general/spec/danae_r/danae_rs.htm) を用いた。この装置は、左右のフラッシュ 2 2 0 からスリット光を投影して左右の CCD カメラ 2 1 0 で被写体 2 3 0 を撮影する。その際、視差情報から物体の奥行きを計算し、被写体 2 3 0 の 3 次元形状を得るものである。 20

3 次元形状計測装置で撮影した画像は、図 3 に示すように、2 次元濃淡データと奥行情報からなる。

【 0 0 2 9 】

(2 . 入力データの生成)

撮影した画像には、撮影時の姿勢変化による顔の回転変動や、照明の明るさによる変動が加わっている。また、画像データをそのまま NN の入力データとして用いるとデータ量が膨大となり、計算時間が大きくなってしまふ。そこで、撮影した画像に対して下記の (1) ~ (6) の手順で前処理を行なって回転と照明の補正、データ量の削減を行なう。

【 0 0 3 0 】

(1) グレースケールに変換 30

撮影した 2 次元濃淡データはカラー画像として得られるが、本研究ではグレースケール画像のみを用いる。これは、カラー画像を扱うことによってデータ量が多くなるためと、照明の色や皮膚の色など、変動要因が増えるためである。

【 0 0 3 1 】

(2) 特徴点の抽出

取り込んだ顔画像から、正面を向いた顔領域を切り出すために、3 次元形状とテクスチャ情報から、下記の (a) ~ (e) の特徴点を抽出する。なお、抽出した特徴点を図 4 に示す。

(a) 高さが最小となる点から鼻頂点座標を取得

最初に、顔の中心となる鼻頂点座標を取得する。鼻頂点は顔の z 座標が最小となる点であることを利用し、座標を決定する。(図 4 点 1) 40

(b) 鼻頂点から上の極点を取得

鼻頂点から y 軸方向に走査し、傾きが正から負、負から正に変化した部分の y 座標を取得し、左右の目の中間点、左右の眉の中間点の座標とする。(図 4 点 2)

(c) 鼻の傾きを取得

鼻の頂点と目の間の点を結んだ直線上の点を取り、最小二乗法により近似直線を計算する。この直線の傾きを x 軸の回転補正用に用いる。(図 4 直線 3)

(d) 鼻の両側の点を取得

鼻頂点から左右一定距離にある点を取得する。この周囲の領域の z 座標を y 軸の回転補正用に用いる。(図 4 点 4) 50

(e) 濃度から目の位置を取得

目については、高さ情報では正確な位置を取得するのが難しいため、2次元濃淡データを用いて目の位置を決定する。まず目の間の点から左右に探索し、黒い部分を検出する。次に、目の濃淡データは左右の目で対称となっていることを利用し、先ほど取得した両目の位置の周囲で左右それぞれの画素値を比較し、対称になっている部分を探索する。得られた画像を最終的な目の位置とする。(図4点5)

【0032】

(3) 回転補正

次に、上述の(2)で取得した特徴点から顔画像の回転に対しての補正を行なって、正面を向いた画像を生成する。

【0033】

(a) x 軸方向の回転補正

上述の(2)で取得した鼻の傾きがすべての画像で等しくなるようにx軸方向に回転補正を行なう。図5において、回転前の鼻の直線の式を $z = a_1 y + b_1$, 回転後の鼻の直線の式を $z' = a_2 y' + b_2$, 補正する回転角度を θ_x とすると、 $\tan \theta_x$ は、

【数16】

$$z' = y \sin \theta_x + z \cos \theta_x \quad (12)$$

【数17】

$$y' = y \cos \theta_x - z \sin \theta_x \quad (13)$$

より、

【数18】

$$y \sin \theta_x + z \cos \theta_x = a_2 (y \cos \theta_x - z \sin \theta_x) + b_2$$

$$z = \frac{a_2 \cos \theta_x - \sin \theta_x}{\cos \theta_x + a_2 \sin \theta_x} y + \frac{b_2}{\cos \theta_x + a_2 \sin \theta_x}$$

$$a_1 = \frac{a_2 \cos \theta_x - \sin \theta_x}{\cos \theta_x + a_2 \sin \theta_x}$$

$$\frac{\sin \theta_x}{\cos \theta_x} = \tan \theta_x = \frac{a_2 - a_1}{1 + a_1 a_2} [\text{rad}] \quad (14)$$

となる。

【0034】

(b) y 軸方向の回転補正

上述の(2)で取得した鼻の左右の点のz座標が左右で等しくなるように、y軸方向に回転補正を行なう。図6の鼻の左右の点のx, z座標をそれぞれ (x_l, z_l) , (x_r, z_r) , 回転後の座標をそれぞれ (x'_l, z'_l) , (x'_r, z'_r) , 補正する回転角度を θ_y とすると、 $\tan \theta_y$ は、

【数19】

$$z'_l = x_l \sin(-\theta_y) + z_l \cos(-\theta_y) \quad (15)$$

10

20

30

40

50

【数 2 0】

$$z'_r = x_r \sin(-\theta_y) + z_r \cos(-\theta_y) \quad (16)$$

$z'_l = z'_r$ より、

【数 2 1】

$$\frac{\sin \theta_y}{\cos \theta_y} = \tan \theta_y = \frac{z_l - z_r}{x_l - x_r} [\text{rad}] \quad (17)$$

10

となる。

【0 0 3 5】

(c) z 軸方向の回転補正

上述の(2)で取得した左右の目の位置の y 座標が左右で等しくなるように z 軸方向に回転補正を行なう。図 7 の左右の目の点の x, y 座標をそれぞれ (x_l, y_l) , (x_r, y_r) , 回転後の座標をそれぞれ (x'_l, y'_l) , (x'_r, y'_r) , 補正する回転角度を θ_z とすると、 $\tan \theta_z$ は、

【数 2 2】

$$y'_l = x_l \sin \theta_z + y_l \cos \theta_z \quad (18)$$

20

【数 2 3】

$$y'_r = x_r \sin \theta_z + y_r \cos \theta_z \quad (19)$$

$y'_l = y'_r$ より、

【数 2 4】

$$\frac{\sin \theta_z}{\cos \theta_z} = \tan \theta_z = \frac{y_r - y_l}{x_l - x_r} [\text{rad}] \quad (20)$$

30

となる。

【0 0 3 6】

ここで、特徴点の抽出と回転補正については、1度だけで正確な補正を行なうのは困難であるので、上述(2)の特徴点の抽出、および、(3)の回転補正の処理を、補正した角度が閾値以下になるまで、繰り返し行なうとよい。

【0 0 3 7】

(4) 顔領域の切り出し

上記(3)で回転補正を行ない正面を向いた画像に対して、鼻頂点座標を中心として、一定の距離を切り出し、図 8 に示すような顔領域とする。

40

【0 0 3 8】

(5) 簡単な照明の補正

画像撮影中の照明の明るさに多少変動があるため、各画像の輝度の平均値を一定にする処理を行なう。

【0 0 3 9】

(6) データ量の削減

上記(4)で切り出しを行なった画像をそのまま使用すると、NNによる学習、認識時間が膨大になってしまうので、画像を縮小して使用する。本実施形態では、例えば 32×32 画素に縮小して使用する。

NNに inputs データとしては、 32×32 画素の 2 次元濃淡データ(図 9)と対応す

50

る 32×32 画素の 3 次元形状データ (図 10) とを合わせて、 $32 \times 32 \times 2$ 個のデータを用いる。ここで、3 次元形状データとしては、2 次元濃淡データの各画素に対応する位置の z 軸上の値を用い、高さの基準として鼻の頂点の高さを用いる。また、2 次元濃淡データに合わせるため、画素値を 8 bit 量子化し、鼻頂点側を 255、後頭部側を 0 として表現する。

【 0 0 4 0 】

< 6 . シミュレーション >

最後に、本実施形態の顔認識システムを用いて学習・認識のシミュレーションを行ない、その結果を示す。

【 0 0 4 1 】

(1 . 画像の撮影)

シミュレーションに用いる画像として、上述の 3 次元形状計測装置 (D a n a e - R) によって撮影した画像を用いる。顔画像は正面を向いた画像を取り込んだが、1 回 1 回の撮影の間に被写体が顔を動かしているため、実用上考えられる程度に斜めを向いた画像が含まれている。被写体からカメラまでは 60 cm、室内の照明は撮影装置のストロボのみとした。また、撮影装置の制限上、今回の実験では眼鏡は外して撮影を行なった。

今回の実験では、1 人につき 5 枚の 2 次元濃淡データ・3 次元形状データを撮影し、登録者用として 30 人、未登録者用として 5 人分の画像を用意した。このうち登録者の画像は 1 人につき 4 枚を学習用に、1 枚を認識用に用いる。すなわち、登録する画像としては $30 \times 4 = 120$ 枚、認識する画像としては登録者 $30 \times 1 = 30$ 枚、未登録者 $5 \times 5 = 25$ 枚を用いた。

ここで、今回実験に用いた画像の枚数は全部で $35 \times 5 = 175$ 枚と数が少ない。そこでシミュレーションの結果の精度を高めるため、leave-one-out法を用いる。すなわち 1 回のシミュレーションでは 1 人分の画像 5 枚のうち 4 枚を選び学習を行ない、残った 1 枚を用いて認識を行なう。この操作を、選ぶ画像を 5 通り変化させることで 5 回のシミュレーションを行ない、平均をとることによってシミュレーションの信頼性を高める。

【 0 0 4 2 】

(2 . 入力データの生成)

顔画像の学習・認識処理を行なう前処理として、NN への入力データの生成を行なう。入力データの生成方法は、上述の < 5 . 3 次元情報の利用 > で説明したとおりである。ここで、特徴点の抽出、顔領域の切り出しに用いるパラメータについてはいくつかの値を試し、実験的に決定した。

(1) グレースケールに変換

取り込んだ画像は RGB 各 8 bit の画像なので、グレースケール 8 bit の画像に変換する。

(2) 特徴点を抽出

2 次元濃淡データと 3 次元形状データから特徴点を抽出する。

(3) 回転補正

抽出した特徴点から x y z 軸それぞれについて回転補正を行なう。ここで、x 軸については、すべての画像で鼻の傾きが y 軸から $0.6 [rad]$ になるように補正を行なった。

(4) 顔画像の切り出し

顔領域については鼻頂点を中心に、顔の大きさが一定になるように切り出し処理を行なった。ここで今回の実験では、すべての画像において眉の上部からあごの部分が入る大きさとして 100 画素を顔の大きさとした。

(5) 簡単な照明の補正

画像撮影時の照明の明るさに多少の変動があるため、各画像の輝度の平均値を一定にする処理を行なった。

(6) データ量の削減

学習、認識処理の計算量を抑えるため、2 次元濃淡画像、3 次元形状画像をそれぞれ 3

10

20

30

40

50

2 × 3 2 画素に縮小処理を行なった。

ここまでの処理を行なった同一人物の画像 5 枚を例として図 1 1 に示す。この後の学習・認識処理では、2 次元濃淡画像、3 次元形状画像それぞれ 3 2 × 3 2 画素を同時に用いる。また、すべての画像について実際に回転補正を行ない、回転補正を行なった量の平均値と最大値を図 1 2 に示す。なお、x 軸の回転補正については、シミュレーション時に角度を指定しているため、平均を 0 ° として計算した。

【 0 0 4 3 】

(3 . ファジィクラスタリング)

(1) まず、最初のクラスタ (サブネットワーク) 数を、登録者の人数と等しい 3 0 にする。一人の登録者につき学習用パターン 4 枚を用いて、個人別の平均ベクトルを求める。

クラスタ分割に用いるパラメータは、以下のように設定した。

- ・クラスタへの最大要素数 : 6
- ・要素の最大帰属数 : 5
- ・収束判断値 : 0 . 0 1
- ・データ集合の個数 : 3 0 人
- ・データの次元 : 2 0 4 8 (3 2 × 3 2 × 2)
- ・一人あたりの顔画像総枚数 : 4
- ・行の非ファジィ化の調節係数 : 1 . 1 5

【 0 0 4 4 】

(2) F C M 法による計算

上述の < 2 . 顔画像の振り分け > で定めたメンバシップ関数及び式 (7) と式 (1 0) により繰り返し計算し、分割行列 U を更新する。U の更新量が収束判断値以下になったら計算を終了する。

ここで、メンバシップ関数の分割行列の例を図 1 3 に示す。最上行の数字は小規模 N N (サブネットワーク) の番号であり、各行の和はおよそ 1 となる。また、最左列の数字は登録者の番号である。最終的なクラスタ (サブネットワーク) 数は 2 3 となった。すなわち、パターン 1 , 2 , 4 , 6 , 7 , 8 , 9 , 1 0 , 1 1 , 1 2 , 1 3 , 1 5 , 1 6 , 1 7 , 1 8 , 1 9 , 2 0 , 2 2 , 2 3 , 2 5 , 2 6 , 2 7 , 3 0 を中心に 2 3 個のクラスタ (サブネットワーク) が形成された。

【 0 0 4 5 】

(3) 非ファジィ化

(a) 行の非ファジィ化

要素の最大帰属数を 5、要素数を減らすための非ファジィ化の閾値を (1 / クラスタ数) × 行の非ファジィ化の調節係数 = 0 . 0 5 に設定する。各要素がこの閾値より小さければ、帰属度を 0 とする。メンバシップ値が閾値より大きくても、要素の最大帰属数 5 より多いため、メンバシップ値の大きさを順に上位 5 個のクラスタ (サブネットワーク) だけに入れる。行の非ファジィ化を行った後の分割行列は図 1 4 のようになった。

(b) 列の非ファジィ化

クラスタ (サブネットワーク) の最大要素数を 6 に設定し、0 ではない要素を大きさの順に上位 6 個を保存し、他の要素を除外する。ここで除外作業を行うとき、除外対象となる要素は少なくとも 1 つのクラスタに帰属させることを保証するため、たとえ同要素が上位 6 個の中に入らなくても、同じ行で他のクラスタへの帰属が全部 0 であれば、この要素は保留しなければならない。列の非ファジィ化を行った結果は図 1 5 のようになった。また、非ファジィ化を行った結果を図 1 6 に示す。

(c) 非ファジィ化による合併

サブネットワーク数を減らすため、これを更に合併すると、最終的に図 1 7 のような分割行列が得られる。最終的なサブネットワーク数は 9 個となった。

ここで注意すべきことは、サブネットワークを合併しなくても、本方法を適用することができるが、図 1 7 に示すようにサブネットワークを統合すれば、計算量をさらに減らすことができ、効率が良くなることである。

10

20

30

40

50

【0046】

(4) 並列NNによる学習と認識)

上記で作成したサブネットワーク(図17)を用いて、小規模な並列NNによって学習と認識を行う。

【0047】

(1) 学習

まず、30名の登録者の120枚学習用顔パターンを、並列NNに学習させる。

(a) NNの構造

今回用いたNNの構造は、図1に示すように入力層に2048ノード、中間層に32ノード、出力層に6ノードを持った、3階層NNとする。

(b) パラメータの設定

NNの各パラメータを以下のように設定し、学習を行った。学習に要した学習回数を図18に示す。

- ・データ集合の個数：30
- ・入力層データの次元： $32 \times 32 \times 2 = 2048$
- ・一人あたりの学習顔画像の総枚数：4
- ・入力層のノード数：入力層データの次元 = 2048
- ・中間層のノード数：32
- ・出力層のノード数：クラスタへの最大要素数 = 6
- ・モーメント：0.01
- ・モーメント更新量：0.01
- ・学習係数(結合係数) = 0.13
- ・学習係数(オフセット値) = 0.1
- ・誤差収束判定値：0.01
- ・最大計算回数：10000

【0048】

(2) 認識

学習の終わったNNに登録者30名の未学習顔パターン30枚と、未登録者5名の顔パターン25枚を以下のステップに従って並列NNに認識させた。

また、認識率、拒否率についてはシミュレーション結果の信頼性を高めるため、leave-one-out法によって5通りのシミュレーションを行い、その平均を計算した。

(a) NNの出力

NNの出力のうち、最も大きい値を出力したニューロンに対応するパターンを、そのサブネットワークが出した答えとする。ただし、学習が行われなかったユニットに顔パターンを認識させるとその出力は0.5になったので、本実験では0.5以下の値を出力したニューロンは、最も大きい値を出力していても無視することにする。

(b) 並列NNの否定能力による排除

あるサブネットワークで一度拒否された要素ユニットを他のサブネットワークの答えから排除する。この時、全ての要素ユニットが排除されたら、この入力顔パターンを未登録と判断する。

(c) 類似度法による判断

答えが1つ以上残った場合、類似度法による判断を行う。答えとなる要素(登録者)の標準顔パターン(学習顔パターン)と入力顔パターンの類似度を計算する。その差が一番大きい要素ユニットを最後の答えとする。ただし、この差がある閾値より小さいと、この入力顔パターンを未登録者と判断する。

類似度法の閾値を0.001ずつ変えたときの認識率と拒否率の変化を図19のグラフおよび図20の表に示す。なお、図19の横軸は類似度法に用いる閾値を、縦軸は認識率と拒否率を表している。

【0049】

(3) シミュレーション結果の考察

10

20

30

40

50

(a) 入力データの生成結果

本実施形態では、学習・認識処理を行う前に、前処理として回転変動が加わった画像の補正を行ない、補正後の画像を NN の入力データとした。図 12 に示した表より、x 軸方向に対して最大 17.4° 、y 軸方向に対し最大 10.4° 、z 軸に対し最大 7.4° の回転が加わった画像に対しても、正面を向いた画像を作成することができたことがわかる。これは、今回の実験中、複数の顔画像を撮影する間に位置・回転変動が起こっていることを考えると、正面顔を撮影した場合に限定される回転変動としては十分といえる。

(b) 認識結果

図 19 のグラフを見ると、横軸の、類似度法で用いる閾値を変動させると認識率と拒否率が大きく変わることがわかる。認識率と拒否率のどちらかを優先するかは応用場面によって変わってくるため、応用場面によって閾値を選ぶ必要がある。

今回の実験では、認識率と拒否率が同じくらいになるように閾値を選んだ場合（閾値 = 0.992 ）、認識率 97.3% 、拒否率 97.6% となった。但し、ここで認識率、拒否率それぞれの意味の定義は次のようである。

認識率は、登録者の未学習画像を入力したときに、正しい登録者と認識できた確率である。拒否率は、未学習画像を入力したときに、未登録者と認識した確率である。この場合、入力が登録者の場合は不正解、未登録者の場合は正解となる。

また、シミュレーションで使用した顔画像（斜め顔画像を含む場合）に対して、上述の非特許文献 5 による従来のクラスタリング手法を利用した方法の認識率は 76.2% であり、通常の 3 層ニューラルネットワークにより学習して、認識を行ったときの認識率は 71.8% であることを確認した。本実施形態のように、ファジィクラスタリング法と顔の 3 次元形状を利用すれば、認識率は 97.3% であるので、本実施形態の手法は非特許文献 5 の考えを大幅に発展させたものであるといえる。さらに、本実施形態の顔認識システムにおいて、放射基底関数（RBF: Radial Basis Function）を中間層の入出力関数に用いた RBF ネットワークまたはウェーブレット関数を中間層の入出力関数に用いたウェーブレットネットワークを利用すれば、認識率および拒否率は 99% 以上を達成できることを確認した。

【図面の簡単な説明】

【0050】

【図 1】本実施形態の並列ニューラルネットワークの構成例を示した図である。

【図 2】本実施形態で用いる 3 次元形状計測装置を示した図である。

【図 3】3 次元形状計測装置で撮影した顔画像の例である。

【図 4】図 3 の画像の 3 次元形状とテクスチャ情報から抽出した特徴点を示した図である。

【図 5】図 3 の画像に x 軸方向の回転補正を行なう様子を示した図である。

【図 6】図 3 の画像に y 軸方向の回転補正を行なう様子を示した図である。

【図 7】図 3 の画像に z 軸方向の回転補正を行なう様子を示した図である。

【図 8】回転補正を行なった画像から切り出した顔領域を示した図である。

【図 9】NN に入力する 32×32 画素の 2 次元濃淡データである。【図 10】NN に入力する 32×32 画素の 3 次元形状データである。

【図 11】本実施形態を用いたシミュレーションにおいて、同一人物を撮影した 5 枚の顔画像から生成した NN への入力データの例を示した図である。

【図 12】上記シミュレーションにおいて、回転補正を行なった量の平均値と最大値を示した表である。

【図 13】上記シミュレーションにおいて、メンバシップ関数の分割行列を示した表である。

【図 14】上記シミュレーションにおいて、行の非ファジィ化を行った後の分割行列を示した表である。

【図 15】上記シミュレーションにおいて、列の非ファジィ化を行った後の分割行列を示した表である。

10

20

30

40

50

【図16】上記シミュレーションにおいて、非ファジィ化を行った結果を示した表である。

【図17】上記シミュレーションにおいて、サブネットワークの合併を行った結果を示した表である。

【図18】上記シミュレーションにおいて、学習に要した学習回数を示した表である。

【図19】上記シミュレーションにおいて、類似度法の閾値を変化させたときの認識率と拒否率の変化を示したグラフである。

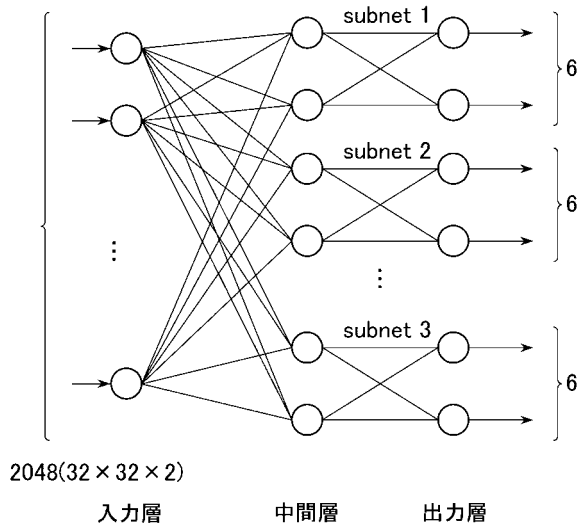
【図20】上記シミュレーションにおいて、類似度法の閾値を変化させたときの認識率と拒否率の変化を示した表である。

【図21】本実施形態の学習過程のフローチャートである。

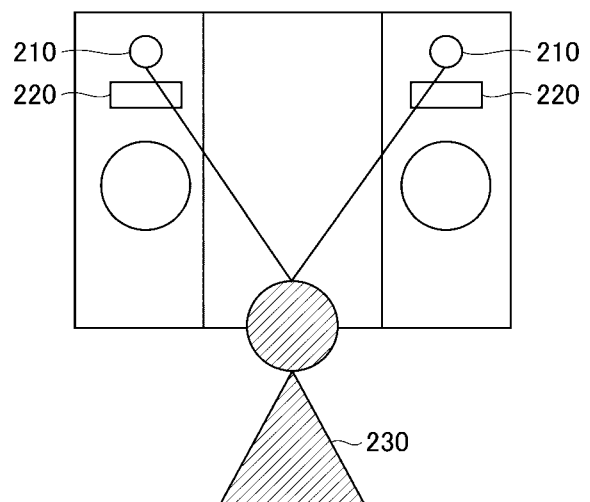
【図22】サブネットワークの結合処理のフローチャートである。

【図23】本実施形態のシステム構成例である。

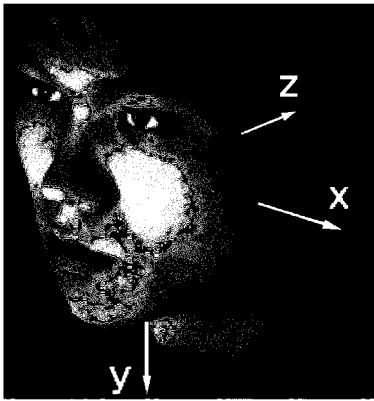
【図1】



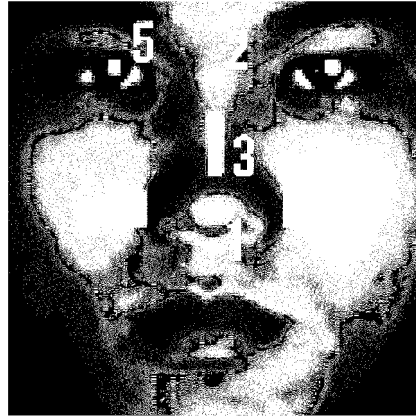
【図2】



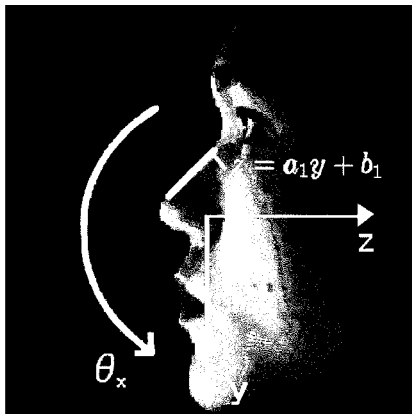
【 図 3 】



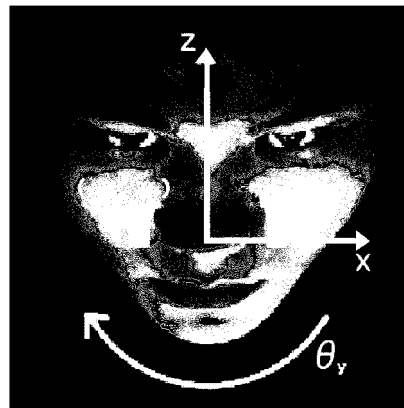
【 図 4 】



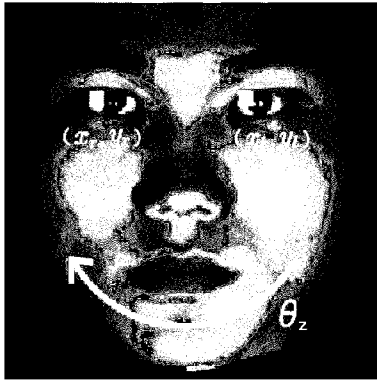
【 図 5 】



【 図 6 】



【 図 7 】



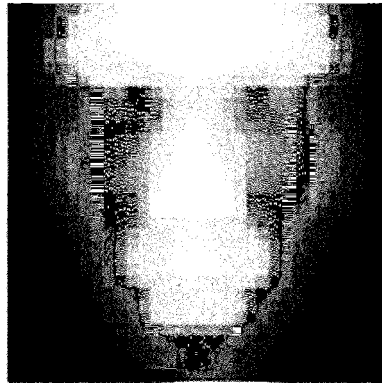
【 図 8 】



【 図 9 】



【 図 10 】



【 図 1 5 】

登録番号	1	2	4	6	7	8	9	10	11	12	13	15	16	17	18	19	20	22	23	25	26	27	30	
1	0.996																							
2		0.987																						
3		0.168				0.085											0.073							
4			0.999																					
5		0.088	0.088			0.071												0.090						
6				0.998																				
7					0.999																			
8						0.993																		
9							1.000																	
10								0.999																
11									0.999															
12										0.992														
13											0.987													
14												0.078												
15													1.000											
16														1.000										
17															1.000									
18																1.000								
19																	0.996							
20																								
21																								0.072
22																								
23																								
24																								
25																								
26																								
27																								
28																								
29																								
30																								0.987

【 図 1 6 】

サブネット	顔パターン	サブネット	顔パターン
1	1,5,14,21,24,29	13	16
2	2,3,5,14,21,24	14	17
3	4,28	15	18
4	6,28	16	19
5	7	17	3,5,14,20,28,29
6	3,5,8,14,21,29	18	22,28
7	9	19	5,23
8	10	20	25
9	11	21	26
10	3,12,24	22	27,29
11	3,13,14,21,24	23	21,30
12	15		

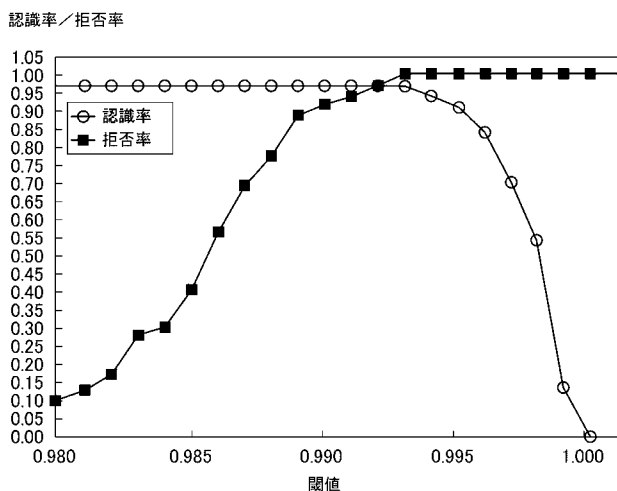
【 図 1 7 】

サブネット	顔パターン
Subnet 1	1,5,14,21,24,29
Subnet 2	2,3,5,14,21,24
Subnet 3	4,28,6,7,9,10
Subnet 4	3,5,8,14,21,29
Subnet 5	11,3,12,24,15,16
Subnet 6	3,13,14,21,24,17
Subnet 7	18,19,22,28,5,23
Subnet 8	3,5,14,20,28,29
Subnet 9	25,26,27,29,21,30

【 図 1 8 】

サブネット	学習回数
Subnet 1	3217
Subnet 2	2202
Subnet 3	4476
Subnet 4	1757
Subnet 5	3077
Subnet 6	1549
Subnet 7	2825
Subnet 8	1371
Subnet 9	1625

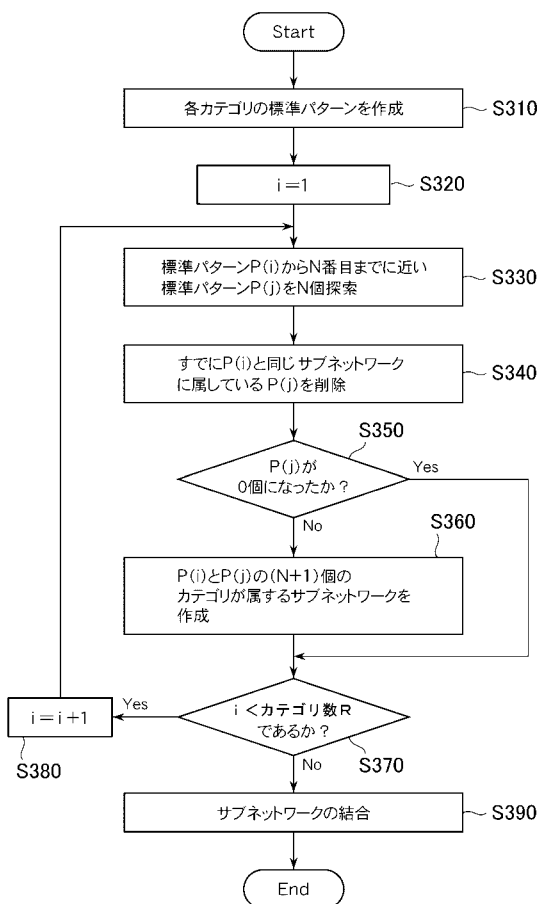
【 図 1 9 】



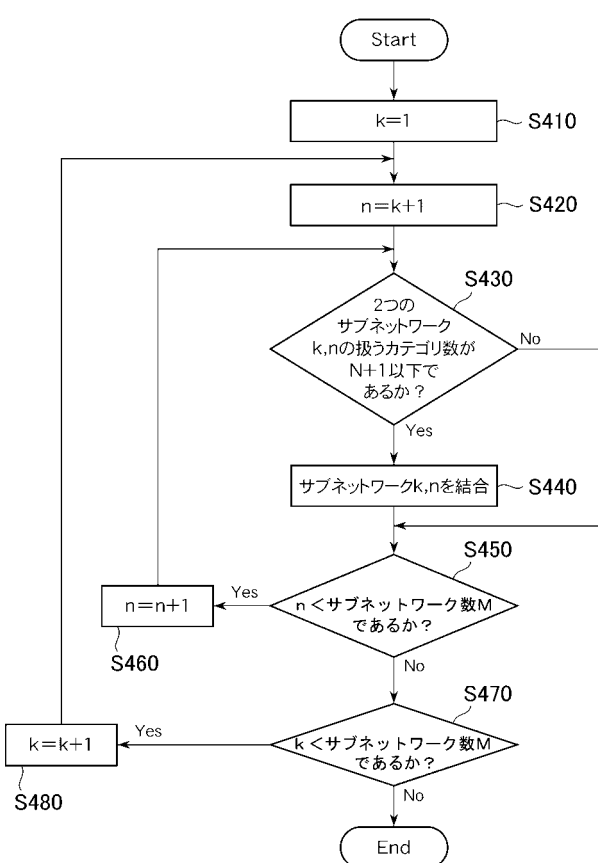
【 図 2 0 】

閾値	認識率 (平均認識枚数/全枚数)	拒否率 (平均拒否枚数/全枚数)
1.000	0.000(0.0/30)	1.000(25.0/25)
0.999	0.133(4.0/30)	1.000(25.0/25)
0.998	0.540(16.2/30)	1.000(25.0/25)
0.997	0.700(21.0/30)	1.000(25.0/25)
0.996	0.840(25.2/30)	1.000(25.0/25)
0.995	0.907(27.2/30)	1.000(25.0/25)
0.994	0.940(28.2/30)	1.000(25.0/25)
0.993	0.967(29.0/30)	1.000(25.0/25)
0.992	0.973(29.2/30)	0.976(24.4/25)
0.991	0.973(29.2/30)	0.944(23.6/25)
0.990	0.973(29.2/30)	0.920(23.0/25)
0.989	0.973(29.2/30)	0.888(22.2/25)
0.988	0.973(29.2/30)	0.776(19.4/25)
0.987	0.973(29.2/30)	0.696(17.4/25)
0.986	0.973(29.2/30)	0.568(14.2/25)
0.985	0.973(29.2/30)	0.408(10.2/25)
0.984	0.973(29.2/30)	0.304(7.6/25)
0.983	0.973(29.2/30)	0.280(7.0/25)
0.982	0.973(29.2/30)	0.168(4.2/25)
0.981	0.973(29.2/30)	0.128(3.2/25)
0.980	0.973(29.2/30)	0.096(2.4/25)

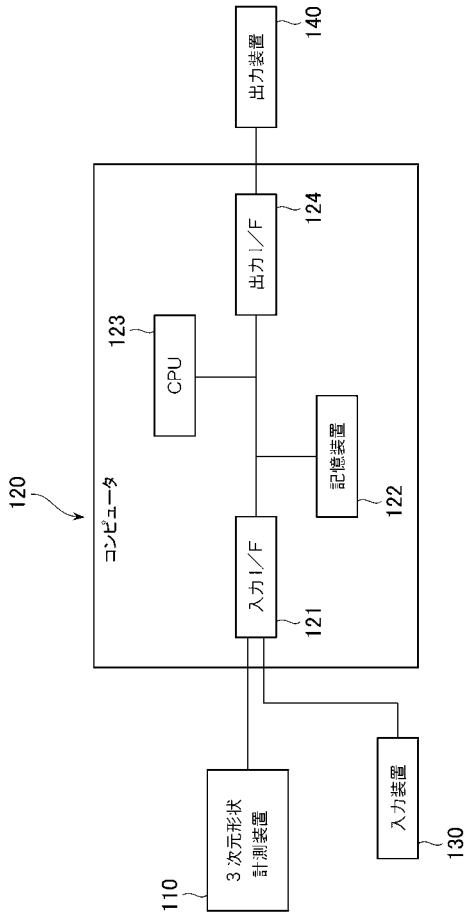
【 図 2 1 】



【 図 2 2 】



【 図 2 3 】



フロントページの続き

Fターム(参考) 5B057 AA19 BA02 CA08 CA13 CA16 DA12 DB03 DB09 DC07 DC08
DC36
5L096 AA06 AA09 BA02 CA02 EA16 FA12 HA11 JA04 LA13