

顔画像認識に基づくセキュリティシステムの研究

谷 萩 隆 嗣 千葉大学大学院教授

1. まえがき

情報化社会の急激な発展につれて、セキュリティに関する要求が高まり、セキュリティシステムでも身体属性を利用した個人認識の研究が盛んに行われている。指紋パターンや眼底パターンを用いた個人識別に比べて、顔画像認識は、認識される側の心理的抵抗感が少ないというメリットがある。

1990年代に、米国で FARET (Face Recognition Technology) 開発プログラムが開始されてから、顔画像認識の研究が再び注目され、コンピュータによる顔画像処理や顔画像識別の研究を行っている研究者が増えている。FARET は犯罪捜査などを目的とする顔画像データベース検索、セキュリティのためのアクセスコントロールなどに、顔画像認識技術を実用化することを目指している。

従来の顔画像認識の代表的な方法として、顔の幅、両眼の中心間の距離、頭頂からの目の高さ、目と鼻の距離などの顔の特徴点をパラメータ化して、その特徴量を使って認識を行う方法や、横顔の輪郭線を波形解析して認識を行う方法、3次元表面形状による方法などがある。このような特徴抽出をベースにしたパターンマッチング (MP: Matching of Pattern) による認識は、パターンの変動に弱いので、認識率を上げることは難しい。これに対して、ニューラルネットワークのような非線形性の特徴を持つシステムによるパターン認識を構築する方法が提案された [1]。

パターン認識問題では、いろいろなニューラルネットワークモデルが使われている。代表的なものは、LVQ (Learning Vector Quantization) 法と MLP (Multilayer Perceptron) 法である。本研究では3層パーセプトロン型ニューラルネットワークで、バックプロパゲーション学習法を使用する。

一般に、ニューラルネットワークの規模が大きくなるにつれて、ローカルミニマムなどの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いく。そこで、大規模なニューラルネットワークをいくつかの小規模なニューラルネットワーク (サブネット) に分割するという並列ニューラルネットワークが提案された [1]。

従来のハードクラスタリング法では、同じパターンは1つのサブネットにしか属することができないため、顔画像認識に应用する際に、顔画像の各パターン間の境界面の非線形性の度合いが大きいために、誤認識率が低くなる。

本研究では、ファジィクラスタリングに基づいて顔パターンを分割し、幾つかの小規模なネットワークに属させる方法を考えた。さらに認識率を向上させるために、顔全体の認識結果と目の部分の結果を併用し、2次認識を行った。また、未登録者の排除と新しいメンバーの追加方法について検討した。

2. 顔画像の前処理

2.1 顔の特徴と視覚情報

顔画像の認識については顔の全体的特性と顔に含まれている個々の特徴についての情報の両方が重要である。顔の知覚や認知において髪型や顔の輪郭、目の相対的重要性と、下半分の内部特徴、特に鼻の相対的な重要性の低さが指摘されている。

Fraser と Parker の特徴の目立ちやすさを調べる実験では、被験者は線描の顔の各部分を素速く続けて示される [2]。Fraser と Parker は各要素欠如を検出する相対的な困難度を比較したが、被験者は顔の輪郭の欠如にもっとも気づきやすく、続いて目、口の順に難しくなって、鼻はもっとも気づきにくかった。また、鼻が相対的に重要度が低いのは、正面向きの顔写真では鼻の知覚が困難であることによるのかもしれない。正面向きの顔は、目や口といった特徴次元を特定するにはおそらく最適の角度であるが、鼻の特定にとっては最悪の角度である。

Haig は、特徴をあらかじめ決めて操作するのではなく、ランダムな位置にある複数の小さな正方形の窓を通して顔の一部分を見えるようにし、その窓の位置と数を操作して顔を認識させるという手続きを用いている [3] [4]。こうした顔の認識課題でも、髪と目の部分の重要性が確認された。本研究では顔の外部特徴 (顔の全体的特徴、すなわち顔の輪郭) と内部特徴 (目の部分) を組み合わせて、認識を行うこととする。

2.2 画像データの取り込み

CCD カメラで取り込んだモノクロ濃淡画像を、実際に画像データとしてニューラルネットワークで学習、認識させるまで、種々の前処理を行った。本研究は真正面から見た顔画像を認識対象としている。顔の切り出しを容易にするため、CCD カメラを用いて同一条件のもとで撮影を行った。本研究では、表情が変わっても個人識別ができるようにするために、1人に付き、“幸福(happiness)”、“怒り(angry)”、“驚き(surprise)”、“悲しみ(sadness)”の4つの表情でそれぞれを6枚ずつを撮影した。

調査により、人間は日常で無表情と微笑のほうが多く、特にセキュリティシステムにおいては常に無表情の顔をしているため、セキュリティを対象とした研究では、無表情と微笑の2種類の画像を使うこととする。

(1) 画像データの2値化

取り込んだ画像データは、そのままでは背景も画像に含まれているため顔の位置が不明である。そこでまず、この画像の中から背景の部分を取り除く。背景領域は濃度が低いことを利用し、画像データを2値化することにより顔部分を切り出す。ここでは画像処理のヒストグラムを利用した。

2値化の処理を行った画像で、まず大まかな位置として、顔の上端、左右の端点を求めておく。そして、左右の端点によって顔幅を求め、その1.6倍を顔の高さとすることによって下端を探す。それから、鼻の直下のy座標、耳を除去した顔の左右の端点のx座標を求める。前処理で得た顔のパラメータを用いて、もとの画像から顔を切り出す。本研究では、目の部分のパターンも利用するため、目の状態を調べる必要がある。目が開いているか、閉じているかを確認することも必要となる。目の位置を探すとき、2値化した画像を利用する。瞳の濃度が顔全体の中で一番低いことを利用し、顔の原画像に対し、低いしきい値を設けて、2値化を行う。人間の顔のもつ大局的な特徴である左右対称性を利用し、顔の中心軸に対称性がある2つの点は瞳である。その付近の濃度が高い部分の幅によって、目の状態を確認し、目を閉じている顔画像を廃棄する。

(2) 切り出された画像の正規化

個人によって、切り出された顔画像の大きさが違うため、画素数を正規化する。まず、変換係数を算出し、幾何学的に変換する。その次に濃度補間を行う。認識率を向上させるために、この濃度補間には、精度の高い3次の補間法を用いることとする。

(3) 画像のモザイクと目の部分の切り出し

正規化後でもデータ数が多いため、画素にモザイク化を施す。この画像で、目の位置を探索し、目の部分を切り出す。これらのデータは顔画像認識の原データとなる。

3. ファジィクラスタリング法 (FCM 法)

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、認識対象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・ベクトルが、あらかじめ決められたどのクラスに属しているかを決定する方法として、ファジィクラスタリング法が知られている。FCM (fuzzy c - means) 法による顔画像の認識において、パターンの振り分けを最適化することで、認識率の向上に繋がると考えられる。

ニューラルネットワークの学習則およびその高速化の研究は盛んに行われているが、学習そのものへの負荷を減らすためには、1つのニューラルネットワークで学習するカテゴリーの数を減らし、複数の小規模なサブネットから並列ネットワークを作る。サブネットの要素とネット数を決めるとき、本研究ではファジィ理論を利用して顔パターンを分割する方法を用いる。本研究で最適化した振り分けアルゴリズムは、大規模なニューラルネットワークを幾つかの並列小規模ニューラルネットワークに分割する際に重要な役割を担っており、認識率の向上に大いに関係がある。

以下、ファジィクラスタリング手法を導入して、振り分けアルゴリズムを導く。

3.1 ファジィ理論

ファジィ理論は1965年にザデー (L. A. Zadeh) が発表した論文 “Fuzzy Sets” で提唱された。これは、あいまいさを伴う情報処理という、ヒューマンファクタなどを考慮した複雑なシステム制御からの発想とも言える。一般には、従来の反復分割最適型手法とファジィ理論の考えを結びつけた、FCM (fuzzy c - means) 法が多く使われている。本研究でも FCM 法を利用する。

3.2 ファジィクラスタリング

画像処理やパターン認識へのファジィ理論の適用については、ファジィクラスタリングが重要である。与えられた多次元データ集合を、そのデータ集合の構造のみから“似ている”データ同士を同一のクラスターにまとめ、指定された任意の数のクラスターに分割することがクラスタリングである。一つのクラスターは、与えられたデータ集合内の一つの部分集合であり、分割とはクラスターの族を形成することである。

(1) ハードクラスタリング

各データを、その分割のうちただ1個のクラスターにのみ帰属させることがハードクラスタリングである。データ集合を X とした場合、データがクラスターに属するかどうかを行列 U の (i, j) 要素で表現する。 n 個のデータを c 種類のクラスターに分割する場合、行列 U は $c \times n$ の行列となる。そして要素は、0 または 1 の値を取り、1 の場合データはクラスターに属し、0 の場合はクラスターに属さないことになる。各クラスターには少なくとも1個以上のデータが必ず属し、各データはどれか1個のクラスターに属している。ハードクラスタリングの分割手法として、最も代表的なのが HCM であり、これはクラスター内の2乗を最小化する方法である。

(2) ファジィクラスタリング

クラスターは、本来明確に定義できるものではなく、あいまいなものである。そのあいまいな情報を生かしたまま扱うのがファジィクラスタリングである。クラスタリングにより、パターン集合をカテゴリー分割する際に、境界部分のパターンはいくつかのカテゴリーに関連を持つ。従来のクラスタリング（ハードクラスタリング）では、いずれか1つのカテゴリー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、いくつかのカテゴリーに少しずつ帰属することを認めている。境界部の属性があいまいな事柄を表現しようとしたファジィ集合の最も自然な応用の1つと言えよう。ハードクラスタリングにおける HCM をファジィに拡張したクラスタリング手法が FCM である。

3.3 顔画像の FCM 法による振り分けアルゴリズム

(1) パラメータの初期化

主にクラスターの合併に使われるクラスター間の最小距離は経験によってパターンの平均距離とする。非ファジィ化に使われるクラスター内の最大要素数と要素の最大冗長度は経験によってそれぞれ6と5に設定する。また、分割行列、クラスター数、クラスターの中心ベクトル、標識ベクトル、重み、収束判定値、計算回数などの初期値も設定する。

(2) 反復収束法の計算

ステップ1：クラスターの平均ベクトルを求める。

ステップ2：クラスターの合併を判断する。

ステップ3：クラスターの更新を判断する。

ステップ4：与えられた収束判定値以下となれば終了。そうでなければ、ステップ1に戻る。

(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の各要素を非ファジィ化する。実際には、しきい値の調節によってカテゴリーの重複度を調節し、顔画像の振り分けを最適化する。

a) 行の非ファジィ化 横方向は1つの要素がいくつかのカテゴリーに属するかを決める。1つの要素は最大5つのカテゴリー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つのカテゴリーにいくつかの要素を含むかを決定する。1つのカテゴリーに最大6つの要素を包含できることとする。これ以上にすると、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。また、重複する行があれば、それを消して1つのカテゴリーにする。

4. ニューラルネットワーク

パターンマッチング法（MP 法）は N 次元特徴区間の距離に基づいた最も単純なパターン認識の方法であり、いろいろな分野で幅広く用いられている。しかし、この方法は画像の変動に弱く、顔画像の認識に用いた場合には、認識率がなかなか上がらない。変動の大きい顔画像の認識に対応できるのは非線形分離能力を持つニューラルネットワークである。

4.1 ネットワークモデル

ニューロンモデルを単体で用い、教師信号による学習を行うものは、基本パーセプトロンと呼ばれている。この基本パーセプトロンは入力されたパターンを線形分離する能力しか持たない。現実には線形分離可能な問題は非常に少ないため、基本パーセプトロンを層状に組み合わせて、非線形分離能力を持たせた3層パーセプトロンが実際に用いられる。ニューロンを層状に並べ、第1層より、外部からの入力をそのまま第2層に伝える。第2層より、入力信号を変換し、中間信号として第3層に伝える。また、第1層と第2層の結合を無作為に行ったネットワークを単純パーセプトロンと呼び、素子数が十分に多ければどのような入力パターンも分類できる能力を有している。

基本パーセプトロン（入力、結合係数、教師信号、出力、関数は2値）の一般的な学習則は、次式のように結合係数を変化させていく。これをデルタルールと呼ぶ。さらに、パーセプトロンの学習則を求めるには、素子の働きを一般化した学習則を求める必要がある。あるパターンがネットワークに入力されたときの出力層のある番目の素子への入力は、その前の層の素子の出力と素子の結合係数との積の総和である。その出力は関数で変換された値である。

このとき、誤差を表す評価関数 E を最小にする最小2乗平均誤差法を用いる。評価関数の極小値を求めるには、この評価関数が結合係数の関数となっていることから、それぞれの結合係数を少しずつ変化させる最急降下法を用いる。

すべての学習パターンがネットワークに入力されてから結合係数が更新される方法は一括修正法と呼ばれている。その他、入力パターンが入力される度に、結合係数を更新させる方法も用いられる。この方法は、逐次修正法と呼ばれる。逐次修正法は収束速度が遅いので、本研究では一括修正法を用いた。

一方、単純パーセプトロンは、複雑な問題を解くために、天文学的な数の素子数を必要とする。そこで、現実的な数の素子数の中間層を持ったネットワークに複雑な問題を学習させるバックプロパゲーションアルゴリズムが1986年に提案された。本研究では、パターン分離問題に有効な3層パーセプトロン型ネットワークでバックプロパゲーションアルゴリズムを用いた。

4.2 アルゴリズム

ステップ1：ネットワークの状態を決める結合係数を乱数値で初期化する。目標とする総合誤差値や学習回数などのパラメータを決める。

ステップ2：登録者の学習用パターン集合から学習するパターンを選択し、ネットワークの入力とする。

ステップ3：（順伝播計算）学習パターンの値を入力層に入れ、入力層から中間層への結合荷重係数をかけて、シグモイド関数により中間層の出力を求める。中間層の出力を出力層の入力として、同じように出力層の出力を計算する。

ステップ4：（逆伝播計算）出力層への重みの更新および中間層への重みの更新を行う。

ステップ5：学習の打ち切り条件が満たされるまで、ステップ3からステップ4を繰り返す。

4.3 並列ニューラルネットワークによる顔画像認識

並列ニューラルネットワークによる顔画像認識は、訓練した後で並列ニューラルネットワークの重みなどを利用して、認識の対象となる顔画像のデータをニューラルネットワークの構造に従って計算することである。それから、FCM法で得たサブネットの帰属情報によって、認識の判断を行う。基本的には、近いもの同士の判別はサブネットによって行われ、遠いもの同士は類似度（MP）法によって判別され、目の部分の認識と組み合わせて判断する。また、1つのサブネットでは否定されたパターンでは、ほかのサブネットにある同じパターンも否定されるというサブネットの否定能力は、認識率向上の要とも言える。

次に、認識アルゴリズムを示す。

ステップ1：登録者の学習用顔パターンをこの並列ニューラルネットワークに学習させ、登録する。

ステップ2：登録者の認識用顔パターンと未登録者の顔パターンを、学習を行った並列ニューラルネットワークに認識させる。

ステップ3：入力顔パターンのサブネットの要素ユニットへの類似度を計算する入力顔パターンの平均距離を計算し、サブネットの要素ユニットへの類似度を計算する。類似度が一番大きい要素ユニットはそのサブネットが出した答えとする。ただし、その類似度はあるしきい値を超えないと無視される。

ステップ4：並列ニューラルネットワークの否定能力による排除能力を適用して、あるサブネットで一度否定された要素ユニットを他のサブネットの答えから排除する。このとき、すべ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者と判断する。

ステップ5：ステップ4が行われた後、答えが残った場合、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（学習用顔パターン）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、この差がある平均類似度しきい値より小さいと、この入力顔パター

ンを未登録者と判断する。

ステップ6：顔全体の認識を行った上、目の部分の認識を行う。2つの答えが一致しない場合には、それぞれ出した答えと入力パターンと類似度で再認識させ、認識の結果を最後の答えとする。

5．シミュレーション結果

5.1 画像データベースの構築

顔画像認識では顔の全体的特性、顔に含まれている個々の特徴についての情報の両方が重要である。顔画像認識のセキュリティへの応用の場合には、撮影条件の変動に対する認識のロバスト性が必要である。しかし、コンピュータによる顔画像認識技術に対するもう1つの重要な評価の視点として、人間による顔認識特性との整合性が挙げられる。与えられた顔画像と人間の目にもよく似て見える顔パターンがコンピュータによる認識結果として選ばれることが望ましい。人間による顔の認識過程のモデルには、視覚モデルと認知モデルの2つがある。視覚モデルとは、顔の類似性に代表されるように人が顔から読み取っている様々な情報と顔パターンの物理的特性との関係を明らかにするものである。

これに対して認知モデルとは、入力された顔パターンはすでに特定の顔を表すラベルに対応付けられているものとし、その顔に対応する人物の情報がどのような過程を経て検索され、人物の同定が行われるのかを明らかにしようとするものである。本研究では、視覚モデルを検討の中心とする。

5.1.1 撮 影

(1) 登録者の場合

撮影の対象は合計25人で、1人に付き、“幸福”、“怒り”、“驚き”、“悲しみ”の4つの表情でそれぞれを6枚ずつ撮影した。1人あたりについて24枚の画像データを取り込んだ。

本研究では、無表情と微笑の2種類の顔画像を認識に使うこととした。当初の登録者20人、追加登録者5人、合計25人で、1人に付き24枚の顔画像を撮影した。

(2) 未登録者の場合

未登録者20人をランダムに撮影し、合計300枚の顔画像を認識に使う。

5.1.2 前処理

- (1) 取り込んだ画像データ(ドット)は、縦、横方向ともに0.75倍に(ドット数)の圧縮を行う。
- (2) 画像データを2値化する。
- (3) 顔の大まかな位置を探索し、さらに各部分を探索する。
- (4) 切り出された画像を正規化する。
- (5) 画像のモザイク化と目の部分の切り出しを行う。
- (6) 撮影画像で、目の位置を探索し、目の部分を切り出す。

5.2 ファジィクラスタリング法

まず、集合数を20にする。1人の登録者に付き学習用パターン12枚により、個人別の平均カテゴリーを求める。

5.2.1 初期化とFCM法の計算

クラスター数、クラスター中心ベクトル、クラスターへの最大帰属度、クラスターへの最小帰属度、クラスターの最大冗長度、合併用の距離ベクトル、重み、収束判断値、繰り返し計算回数、データ集合の個数、データの次元、1人当たりの顔画像の総枚数、非ファジィ化のしきい値、クラスター合併の調節係数、列の非ファジィ化の調節係数、クラスター間の最小距離、分割行列などを設定する。さらに、メンバシップ関数の初期化分割行列を設定する。

5.2.2 非ファジィ化

(1) 行の非ファジィ化

要素の最大帰属度を5に設定する。要素数を減らすため、非ファジィ化のしきい値を0.055とする。各要素の値がこのしきい値より小さければ、帰属度0とする。メンバシップ値がしきい値より大きくても、要素の最大帰属度5より多いとき、メンバシップ値の大きさ順に、上位5個のクラスターだけを入れる。

(2) 列の非ファジィ化

縦方向は1つのクラスターにいくつの要素を含むことができるかを定める。ここでは、クラスター内の最大要素数を6に設定する。0ではない要素を、大きさ順に上位から6個保存し、他の要素を除外してしまう。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つのクラスター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の中に入っていないくても、もし同じ行で他のクラスターへの帰属度が全部0であれば、この要素は保存しなければならない。

(3) 非ファジィ化による合併

サブネット数を減らすため、これを更に合併する。これに基づいて、大規模なニューラルネットワークを小規模なサブネットに分割する並列ニューラルネットワークを利用して、学習と認識を行うことができる。

5.3 パターン照合

5.3.1 最短距離法

パターン識別において、入力パターンと標準パターンを比較する場合、パターン照合法は、両パターンをそのままの形で比較するもっとも基本的な方法である。これには、パターンをベクトルによって表現した場合、2つのベクトル間の距離を利用する最短距離法と、2つのベクトルがなす角度に注目した類似度法の2種類がある。

最短距離法は N 次元特徴区間の距離に基づいた最も単純なパターン認識の方法であり、計算には時間もかからないので、いろいろな分野で幅広く使われている。また、最短距離法は、複雑な識別法の要素手段にもよく使用される。ただし、この方法はパターンの変形には弱く、特に、認識のカテゴリー数が多くなると、認識率は著しく低下する。ここでは、学習用として25人（初期登録者20人、追加登録者5人の合計25人）の300枚、認識用として20人の300枚の顔画像データを用いて、最短距離法とニューラルネットワーク法（NN法）で認識を行った。それぞれの認識率を求めた結果、最短距離法では70.67%であった。最短距離法の認識率の低さは、最短距離法が非線形分離の能力を持っていないためである。したがって、最短距離法は非線形性の度合いが強い顔画像の認識に向いていないことがわかる。

5.3.2 ニューラルネットワーク

3階層ニューラルネットワークは、非線形分離の能力を持ち、誤差逆伝播法を武器に、学習後の汎化能力は変形の大きい顔画像に対応できる。

まず、従来のニューラルネットワークで認識実験を行う。最短距離法と同じ顔画像データを利用した25人（初期登録者20人と追加登録者5人）の場合は、認識率は90.3%であった。結果から見れば、認識率は確かに最短距離法より大幅に向上した。また、認識人数を減らした6人の場合は、98.6%の高い認識率を得た。

この結果によって、多くの人数を持つ大規模なニューラルネットワークをこのようないくつかの小規模なニューラルネットワークに分割すれば、高い認識率の得られることが示された。

5.3.3 並列ニューラルネットワーク

以上の結果から、1つのニューラルネットワークにおいては、学習と認識には小人数のほうが有利であることが分かる。そこで、多くの人数を持つ大規模なニューラルネットワークを、いくつかの小規模なニューラルネットワーク（サブネット）に分割する方法が、すでに筆者によって提案されている[1]。その方法は、1つの顔パターンをいくつかのサブネットに属させる考えに基づいている。ここで、ファジィ理論を用いて、サブネットを分割する。

メンバシップ関数より、クラスタリングを繰り返し計算する。さらに、非ファジィ化を行った後、サブネットを合併する。その結果、登録者に対して90.11%の認識率が得られ、未登録者に対して、99.01%の高い拒否率が得られた。誤認識は顔画像を撮影するときの条件と大いに関係があるので、撮影するときの影響をいかに少なくするかということがこれからの課題となる。また、以上の実験では4種類の表情を持った顔画像を使用しているため、認識率を下げていると思われる。

5.4 認識率を向上させるための提案

撮影するときの光の強さが n 倍に変化するとともに、パターンの濃度は変わる。この濃度変化はパターンを構成する画素の値をおのおの n 倍するものである。それより作られるすべてのベクトル要素を n 倍、そして最終的に元のパターンベクトルの大きさも n 倍になる。元のパターンベクトルを x とすると、最終的なパターンベクトルは nx となる。また、カメラとの距離変動から生じる顔の大きさの変化は、もとのパターンベクトルに同じような影響を与える。

顔画像認識のセキュリティへの応用の場合には、撮影条件の変動に対する認識のロバスト性を評価する必要がある。しかし、コンピュータによる顔画像認識技術に対するもう1つの重要な評価の視点として、人間による顔の認識特性と

の整合性が挙げられる。すなわち、与えられた顔画像と人間の目にもよく似て見える顔パターンがコンピュータによる認識結果として選ばれることが望ましい。

5.4.1 従来の並列ニューラルネットワークの欠点

n 倍されたパターンと複数の標準パターン間の距離を再び計算した際に、距離の長さの順序が変わってしまうことがある。この場合には、入力パターンが誤認識されてしまう。そこで、最短距離法の代わりに、複合類似法を用いることとする。

従来の認識法 [1] では、誤認識された結果がこの方法でかなり改善された。認識率は以前の93.3%から95%に上昇した。

5.4.2 学 習

この実験では、まずモザイク化しない画像データ（顔全体のパターン）を使用した。20名の登録者の240枚の学習用顔パターンを、改善した並列ニューラルネットワークに学習させた。

ニューラルネットワークの構造としては、入力層に1024 ノード、中間層に128 ノード、出力層に 6 ノード持つ 3 階層ニューラルネットワークを用いた。

5.4.3 認 識

20名登録者の240枚認識用顔パターンと、20名の未登録者の300枚の顔パターンを並列ニューラルネットワークに認識させた。登録者の場合には、認識率が91.7%であった。未登録者の場合には、96.3%の拒否率を得た。

（認識過程）

ステップ1（入力顔パターンのサブネットの要素ユニットへの類似度の計算）：入力顔パターンの平均距離を計算し、サブネットの要素ユニットへの類似度を計算する。類似度が一番大きい要素ユニットはそのサブネットが出した答えとする。ただし、その類似度があるしきい値を超えないと無視される。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、類似度が0.5になったので、本実験ではこのしきい値を0.5にした。

ステップ2（並列ニューラルネットワークの否定能力による排除）：あるサブネットで一度否定された要素ユニットを他のサブネットの答えから排除する。このとき、すべ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者と判断する。

ステップ3（類似度法による判断）：ステップ2が行われた後、答えが残った場合、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（学習用顔パターン）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、この差がある平均類似度しきい値より小さいと、この入力顔パターンを未登録者と判断する。本実験では、このしきい値を0.97にした。

5.4.4 新しいメンバーの追加

従来法 [1]：従来の方法で、新しいメンバーを追加する場合には、FCM 法を用いて再振り分けをしなければならない。そして、すべてのサブネットを学習させる。サブネットが多くなると共に、学習の時間は長くなる。本実験の場合には、5人の新しいメンバーを追加した。従来の方法では、25人を9個のサブネットに分割した。9個のサブネットを学習させ、認識に用いた。

提案法：従来法の結果から、登録者の人数が多いとき、新しいメンバーを追加するには、時間が掛かることがわかった。そこで、すべてのサブネットの再学習の代わりに、新たなサブネットだけを学習すれば、新しくメンバーを追加する場合に有利であると考えられる。まず、5人の追加登録者をひとつの新しいサブネットに入れて学習する。

学習を終わったら、認識を行う。そして、誤認識を調べ、サブネットを調節する。すなわち、誤認識されたパターンを再分割し、学習および認識を再び行う。これによって認識率は91.7%から95%に上昇した。

5.4.5 目の部分との併用

顔全体のパターンを使ったとき、登録者については95%の認識率を得た。また、未登録者については96.33%の拒否率を得た。セキュリティシステムの場合には、未登録者が入ってくるのは非常に危険なことであるため、未登録者については、さらに高い拒否率が得られることが望ましい。

Haig, Fraser と Parker の実験により、輪郭と目の部分の重要性が確認された。そこで、顔全体の認識を行った上、目の部分の認識を行う。2つの答えが一致しない場合には、それぞれ出した答えと入力パターンと類似度で再認識させ、認識の結果を最後の答えとする。

このようにすることによって、登録者に対する認識率は97%（誤認識：9枚）に上がった。未登録者の場合には、99.33%（誤認識：2枚）の拒否率を得た。

5.4.6 考 察

（１）シグモイド関数の傾き

この傾きが大きいと、ステップ関数と同様な働きとなり学習が進みにくだけでなく、最終的な認識率も低くなってくる。逆に、傾きが小さいと、各ユニットの出力は0や1の状態になり難しく、常にニューラルネットワークが修正されるので、学習回数を増やすと過剰学習により認識率が低下すると考えられる。本研究では、傾きを2.0にして実験を行った。

（２）中間層のユニット数

中間層のユニット数を変えると、認識率も変わる。中間層のユニット数を増やすと、認識率は向上する。しかし、ユニット数が128を超すと、認識率はあまり上がらない。それは、ユニット数がある値を超すと、認識率の向上が飽和してくるのが理由と思われる。本研究では、中間層ユニットは128ノードに設定した。

（３）認識率の向上

顔画像認識をセキュリティシステムに適用する場合には、撮影条件の変動に対する認識のロバスト性を重視して評価している。しかし、コンピュータによる顔認識技術に対するもう1つの重要な評価の視点として、人間による顔の認識特性との整合性があげられる。与えられた顔画像と人間の目にもよく似て見える顔パターンがコンピュータによる認識結果として選ばれることが望ましい。

本研究では、2種類の表情の顔パターンを利用して多数のシミュレーションを行った。高い認識率が得られたが、3回だけやや低い認識率の結果もあった。それは、顔の変形しやすい特徴に対して、情報が少ないためであると考えられるので、さまざまな表情で訓練して、また、カラー情報を含めることで認識のロバスト性を高めることが必要であると考えられる。本研究では、サブネットの振り分け方法としてファジィクラスタリングを使ったが、学習と認識のネットワークモデルとして、ニューロ・ファジィモデルを導入することが考えられる。

6．むすび

本研究では、顔画像認識をする際に、登録者についての認識率を向上させると共に、未登録者の排除も行った。まず、認識の基礎となる顔画像データの前処理の段階で、精度の高い3次補間法を用いて、顔と目の部分画像をそれぞれ32×32画素と64×64画素でモザイク化してから、切り出した。また、最適なクラスタリングを行うため、あいまいさを許すアルゴリズム（FCM）法を用い、n次元空間における似通った顔パターン（ただの空間距離法で判断すると、間違いやすい近似パターン）を同じサブネットに帰属させ、ニューラルネットワークで認識させた。

顔の視覚と認識には顔の全体的特性、顔に含まれている個々の特徴についての情報の両方が重要である。顔の視覚や記憶において、髪型や顔の輪郭、目の相対的重要性と、下半分の内部特徴、特に鼻の相対的重要性の低さが指摘されている。

本研究では顔の外部特徴（顔の全体的特徴、すなわち顔の輪郭）と内部特徴（目の部分）を組み合わせ、認識を行う方法を提案した。従来の方法[1]では、未登録者が入っている場合には、全体の認識率が下がった。本研究では、この問題を大幅に改善することができた。また、新しいメンバーを追加する場合には、従来の方法では、登録者全員をもう一度ファジィクラスタリングで振り分け、新たにサブネットを作成する必要があるが、本研究では、自律調節法を利用し、新しいメンバーを含むサブネットだけを学習させて済ませることができた。

参考文献

- [1] 谷萩隆嗣、高野裕昭：カテゴリーを組み合わせたニューラルネットワークによる顔画像認識、電子情報通信学会論文誌 D- , Vol. J77-D- , No.11, pp. 2151-2159, 1994.
- [2] I. Fraser and D. Parker : Reaction time measures of feature saliency in a perceptual integration task, Aspects of Face Processing (H. D. Ellis, M. A. Jeeves, F. Newcomb and A. Young, Eds.), Dordrecht, Martinus Nijhoff, 1986.
- [3] N. D. Haig : The effect of feature displacement on face recognition, Vol. 13, pp. 505-512, 1984.
- [4] N. D. Haig : Exploring recognition with interchanged facial features, Perception, Vol. 15, pp. 235-245, 1986.