

7F-4

複数の特徴ベクトルを用いたニューラル ネットによる手書き数字認識

加藤誠巳 高木啓三郎
(上智大学理工学部)

1. まえがき

近年、階層型神経回路網の学習則にBP法^[1](Back Propagation Method)が導入されて以来、BP型ネットワークによるパターン認識に関する数多くの研究がなされてきた。特に、画像認識、音声認識等においては従来手法とほとんど大差ない程度の認識率が容易に得られ、今後さらに改善されて行くものと思われる。

BP型ネットワークの手書き数字認識への適用に関しては、既にいくつか報告されているが^{[2][3][4]}、それらはいずれも単独のネットワークで構成されており、しかもネットワークの出力値の判定、解釈も比較的単純なものを用いていた。本稿では、手書き数字のイメージデータから複数の認識用入力層特徴ベクトルデータを作成し、それらを用いてそれぞれ別個に学習させたBP型ネットワークにより独立・並行的に認識を行った後、各ネットワークの出力値パラメータを統合的に用いて最終認識結果を判定する手法について検討を行った結果について述べる。

2. 認識に用いた入力層データ

ここでは、手書き数字データはすべてペン入力方式のタブレットより入力された細線を均一に肉付けしたイメージデータを用いている。認識に用いる入力層データとして、直交座標軸および極座標方向への面積投影、2次元FFT信号、原イメージの3種を用いた。図1に、手書き数字'4'に対するこれら投影情報および原イメージ抽出の様子を示す。肉付けされたイメージデータは、x軸、y軸方向へそれぞれ面積投影され、各々16次元ベクトルデータで表現される。この処理と同時に求められた重心を中心とする極座標方向への面積投影データは10°ずつ、計36次元ベクトルデータとして表現される。さらに、原イメージは16×16のメッシュに相当する256次元ベクトルデータで表現される。

図2aに16×16の原イメージに対する2次元FFTの例を示す(ただし直流分は0としている)。図2bはこの256個のパワーベクトルから直流成分並びに対称成分を取り除いて最大値で正規化し、一次元化した129次元ベクトルであり、これを入力層データの1つとして利用する。

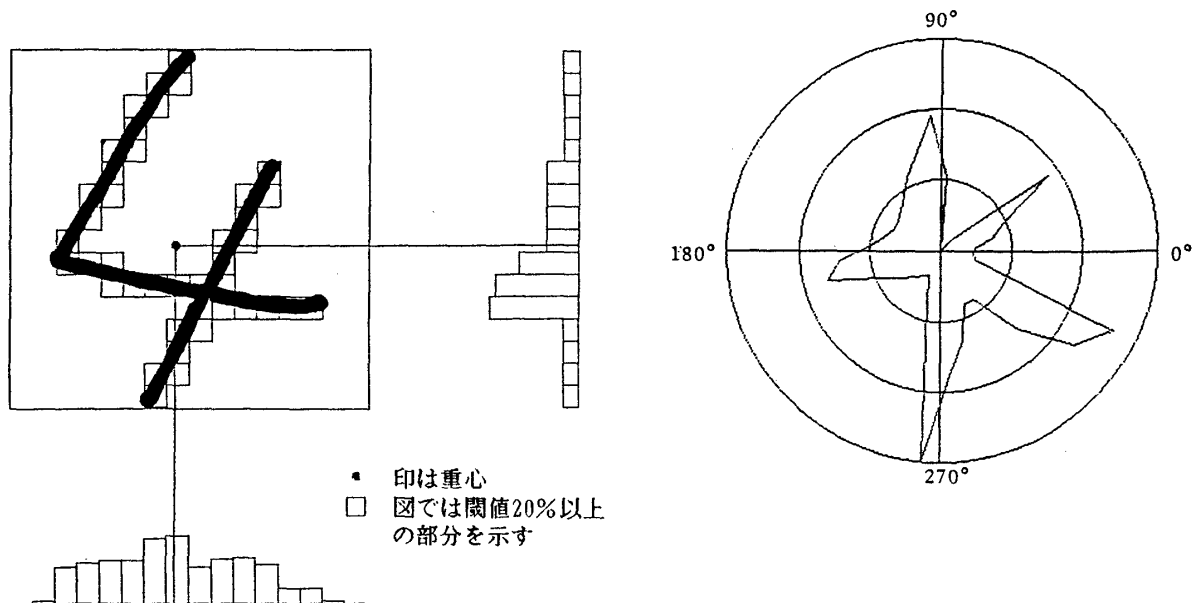


図1 面積投影

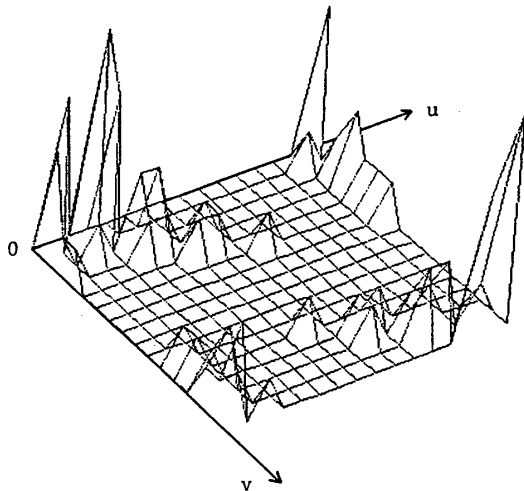


図2 a 2次元FFTのパワー

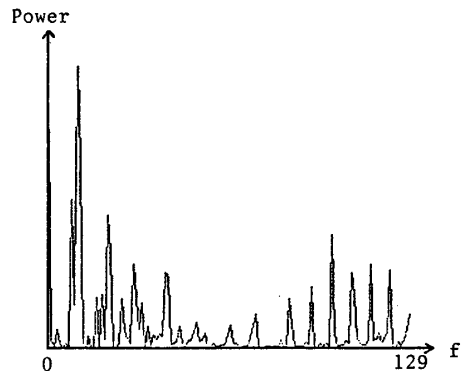


図2 b 1次元化した2次元FFT信号

3. ネットワークの構造

図3に、ここで採用した3種の入力層データを用いてそれぞれ独立に認識を行うネットワークの構造を示す。

4. 最終結果の統合的判定手法

認識すべきテスト入力パターンが既に学習済みのパターンと良く類似している場合、ネットワークは正解に相当する出力ノードのみが大きな出力を出す、類似度が低くなるにつれてネットワークの複数の出力ニューロンが活性化したり、あるいは出力の最大値が低下する場合が多い。このような状況にあつては3つのネットワークの判定結果の多数決で最終判定をすることは得策ではない。ここでは、それらの曖昧な出力状態に

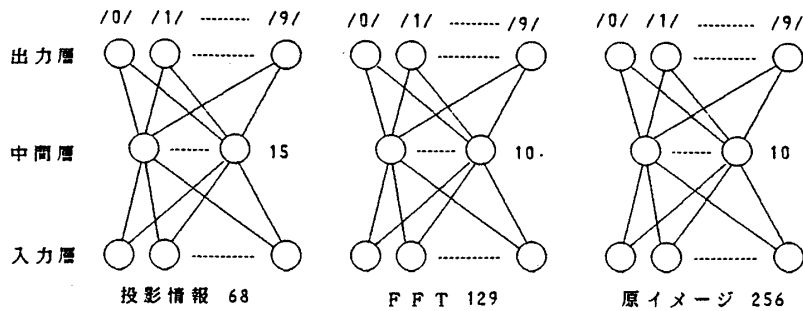


図3 ニューラルネットの構造

対してDempster & Shaferの確率理論^{[5][6]}を適用することにより3つのネットワークの出力を総合的に考慮して各事象の基本確率を算出し、その最大値を与えるものを最終結果として判定するようにした。

5. 認識結果

手書き文字データとして、12人が作成した0~9までの数字960サンプルを用意し、そのうち240サンプルを各ニューラルネットの学習に用い、残り720サンプルを認識実験に用いた。まず、各ネットワークをそれぞれ独立に単純な最大値を用いて認識させた場合の正答率は面積投影、原イメージ、FFTに対しそれぞれ95.0%、89.9%、93.9%であった。一方、本手法による統合的判定手法を用いて総合判定した場合の認識率は97.9%になった。

6. 結び

複数の特徴ベクトルを用いてニューラルネットの学習を個別に行い、認識に際してはそれぞれ別個に認識したネットワークの出力パラメータを統合的に判定することにより認識率を向上させる手法について述べた。実験の結果、最終認識率97.9%が達成されたが、これは個別ネットワークの認識率を大きく上回ることが明らかとなった。

ここでは、複数の特徴ベクトルとして面積投影情報、原イメージ、FFTの3種を用いたが他のさらに有効な特徴ベクトルに置き換えたり、または追加することでさらに認識率が向上することが期待される。終わりに、有益な御助言を戴いたNTT基礎研寛一彦、小高和己、萩田紀博の各氏並びに東大生研石塚満氏に謝意を表す。

参考文献

[1] D. E. Rumelhart, J. L. McClelland: "PARALLEL DISTRIBUTED PROCESSING", M. I. T. press (1986).
 [2] 津久井, 平井: "ニューロ・パターンマッチングによる文字認識", 信学技報, PRU88-56 (1988-9).
 [3] 香田, 高木, 鈴木: "ニューラルネットによる手書き英数字認識", 信学技報, PRU88-57 (1988-9).
 [4] 山田他: "ニューラルネットを用いた文字認識", 信学技報, PRU88-58 (1988-9).
 [5] 石塚: "Dempster & Shaferの確率理論", 信学誌, Vol. 66, No. 9, pp. 900-903 (1983).
 [6] 石塚: "不確かな知識の取り扱い", 計測と制御, Vol. 22, No. 9, pp. 774-779 (1983).