

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に関する研究

加藤 誠巳

上智大学 理工学部

電気・電子工学科 教授

1. まえがき

階層型ニューラルネットをパターン認識問題へ適用した例は多数見受けられ、そのうち特に文字認識、音声認識への期待は益々高まりつつある。しかし、これまでの報告の多くは、比較的単純な特徴ベクトルを用い、単一のニューラルネットで学習及び認識を行い、また認識に際しては単純に最大値を出力した出力ニューロンに相当するカテゴリを認識結果とするような構成であった⁽¹⁾⁽²⁾。本稿では、階層型ニューラルネットとその学習則であるBP法⁽³⁾を用い、パターン認識問題の一例として手書き数字認識を取り挙げ、複数の特徴ベクトルを個別に学習した複数のニューラルネットの出力値を統合的に判定する複合ニューラルネット回路と名付けた方法により認識率の改善を図ることが出来ることを示す⁽⁴⁾。

2. 手書き数字の特徴ベクトルの抽出

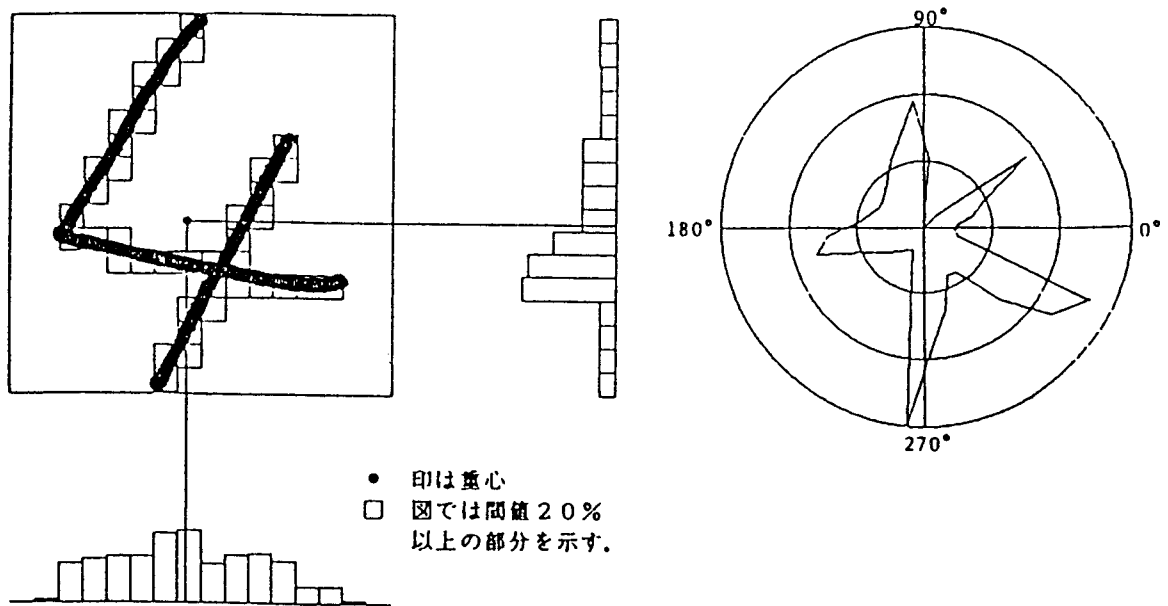
ここで用いた手書き数字データはタブレットのペンより入力された点列情報及びペンアップ/ダウン情報を基に点列間を均一の太さを有する線分で結んだものを用いた。更に、大きさ及び位置に依存しないようにするため枠取り・正規化処理を行う。すなわち外接長方形を用いた枠取り処理を行った後、長辺の長さを一定の値(256ドット)に正規化し、さらにこの正規化された長方形が外接正方形の中央になるように配置する。

手書き数字データは、12人により作成された計960個を用いた。各作成者は0~9までの数字をそれぞれ8個ずつ計80個の数字を入力し、ファイルに保存した。認識に利用する特徴ベクトルとして種々のものが考えられるがここでは次の4種類の情報を採用した。

- (1) 直交軸方向面積投影情報
- (2) 重心の周りの極方向面積投影情報
- (3) 原イメージ情報
- (4) 原イメージの2次元FFTのパワー情報

図1に、手書き数字‘4’に対するこれら投影情報及び原イメージ抽出の様子を示す。すなわち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2に、この数字‘4’の16×16の原イメージに対する2次元FFTのパワーの例を示す。(但し直流分は0としている)。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た129次元ベクトルを2次元FFT情報として使用している。



(a) 直交軸方向面積投影情報

(b) 重心の周りの極方向面積投影情報

図1 面積投影

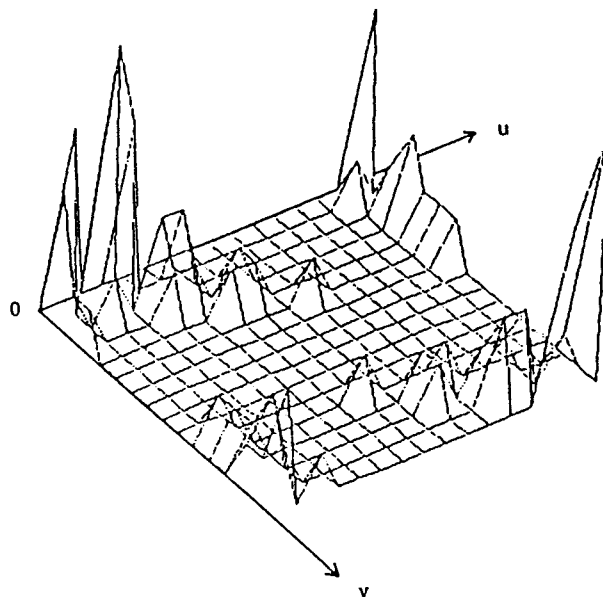


図2 2次元FFTのパワー

3. 個別ニューラルネットの学習とその結果

以下に示す5種類の特徴ベクトルを入力として、通常のBP法を用いて個別に3層のニューラルネットの学習を行い、オープン認識率を求めた。5種類の特徴ベクトルは次の通りである。

- (1) 直交軸方向面積投影情報
- (2) 重心の周りの極方向面積投影情報
- (3) 直交軸方向面積投影情報+重心の周りの極方向面積投影情報
- (4) 原イメージ情報
- (5) 原イメージの2次元FFTのパワー情報

この場合、各ニューラルネットの学習に用いた手書き数字データベースは条件を等しくするため同一のものを用いており、前述の12人の作成した960個のデータの内の240個（各数字一人当たり2個ずつ計24個）の手書き数字を用いた。また、各ニューラルネットの未知データに対する認識率（オープン認識率）算出用の手書き数字データベースに対しても条件を等しくするため同一のものを用いており、同じ12人の作成した残りの720個（各数字一人当たり6個ずつ計72個）の手書き数字を用いた。

表1に、個別ニューラルネットの構造及びオープン認識率を示す。これから直交軸方向面積投影+重心周りの極方向面積投影の如く、異種の特徴ベクトルを同時に入力して学習させると、それぞれ単独で学習させた場合よりも大幅に認識率が向上することが明らかになった。

4. 複合ニューラルネット回路による総合判定

前節で述べた5種類の特徴ベクトルを用いて個別に学習させたニューラルネットのそれぞれのオープン認識率は満足すべき値ではない。認識率改善の一方法として個別ニューラルネットの判定結果の多数決により最終結果を決定することも考えられるが必ずしも認識率の向上には結び付かない。これは個別ニューラルネットにおける認識判定には最大値を出力したニューロンのみ注目し、その他の出力ニューロンの値を有効に活用していないためと考えられる。本稿では、複数の個別ニューラルネットの出力ニューロンの各出力値を有効に活用するため、未知(unknown)の確率を考慮にいれた Dempster および Shafer の確率理論^{(5) (6)}を用いて最終判定を行う、図3に示すような総合判定回路を設けた複合ニューラルネット回路と名付けた方法を提案する。

ここでは、独立に学習を行った複数個の個別ニューラルネットの各出力ニューロンの出力値を統合する方法として次に述べる2種の手法について検討を行った。

(1) 方法1

本方法では個別ニューラルネットの出力ニューロンの出力値（これはシグモイド関数の出力であるので0と1の間の値をとる）はそれぞれ独立である（厳密には独立ではない）と仮定し、それぞれ別個の基本確率と考え、Dempster の結合則を適用する。

(2) 方法2

方法1では個別ニューラルネットの各出力ニューロンの出力値は独立であるとしたのに対し、ここでは独立と考えないで1つの個別ニューラルネットの出力ニューロンの出力値の組を予め定めた一定の未知量 μ を加味して正規化したものを基本確率と考える。

表1 個別ニューラルネット回路の構造及び認識率

入力に加えた特徴ベクトル	入力層ニューロン数	中間層ニューロン数	出力層ニューロン数	オープン認識率
直交軸方向面積投影	32	10	10	89.7%
重心周りの極方向面積投影	36	10	10	88.6%
直交軸方向+重心周りの極方向面積投影	68	15	10	95.0%
原イメージ	256	10	10	89.9%
原イメージの2次元FFTのパワー	129	10	10	93.9%

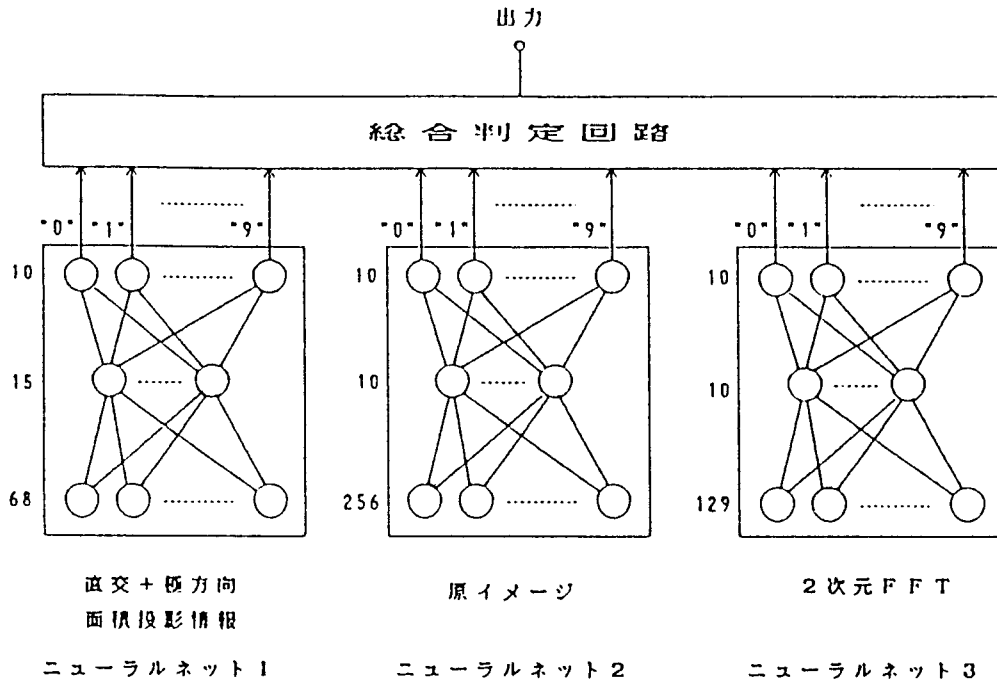


図3 総合判定回路を有する複合ニューラルネット回路

以上の2つの方法に対し、図3に示すように、
 ①直交軸方向面積投影情報+重心周りの極方向面積投影情報、②原イメージ情報、③原イメージの2次元FFTのパワー情報の3つの特徴ベクトルを用いて、それぞれ独立に学習させた3種の個別ニューラルネットを総合判定回路に加えて認識実験を行った。その結果最終認識率は、方法1では97.9%、方法2では図4に示すように $\mu \geq 0.5$ で認識率は98.1%となり、本方法の有効性が確認された。

認識率 (%)

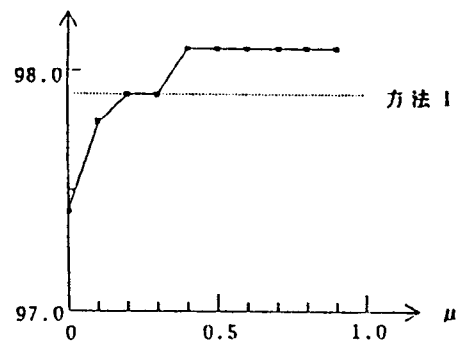


図4 方法2による未知量 μ に対するオープン認識率の変化

5. むすび

相異なる特徴ベクトルを用いて独立に学習させた複数個の個別ニューラルネットの出力を Dempster & Shafer の確率理論を用いて統合的に判定を行う複合ニューラルネット回路を提案し、手書き数字認識に適用してその有効性を確認した。複数個の相異なる特徴ベクトルを同時に入力して学習させる場合に比べ、ここで提案した手法は、ハードウェアの複雑さ、学習時間の短縮の面で一般に有利であると考えられる。

今後、ここで採用した以外の特徴ベクトル、例えば x 軸、y 軸周りのモーメント、重心の座標、Hadamard 変換係数、Hough 変換信号等を入力として用いることにより認識率の更なる改善を意図している。

最後に、本研究の遂行に当り多大な御支援をいただいた(財)高柳記念電子科学技術振興財団に厚くお礼申し上げますとともに、故高柳先生のご冥福を心よりお祈り致します。

参考文献

- [1] 香田、他：“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，pp.71-77 (昭63).
- [2] 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，pp.79-86(昭63).
- [3] D.E. Rumelhart et al.：“Parallel Distributed Processing”，Vol.1 & Vol.2 MIT Press (1986).
- [4] 加藤、高木：“複数の特徴ベクトルを利用したニューラルネットによる手書き数字認識”，信学技報，PRU88-151(平01).
- [5] 石塚：“Dempster & Shafer の確率理論”，電子通信学会誌，第66巻，9号，pp.900-903 (昭58).
- [6] 石塚：“曖昧な知識の表現と利用”，情報処理学会誌，Vol.26，No.12，pp.1481-1486 (昭60).