階層型ニューラルネットとベクトル量子化法を用いた文字認識実験

2G−6

鈴木保春、米田政明、長谷博行、酒井充、二宮孝之
(富山大学工学部電子情報工学科)

1 はじめに

近年ニューラルネットワークを利用した文字認識が盛んに研究されている。学習法としては誤差逆伝搬法とベクトル量子化法が主流となっているが、今までのベクトル量子化法を利用した文字認識は、文字パターンそのものをベクトル量子化法により大分類し、その結果を直接認識に利用するか、またはさらに後段で階層型ニューラルネットを利用して認識する方法が多かった[1]。しかし手書き文字を対象とした場合は変動が大きいためベクトル量子化法でカテゴリ分類した場合、1カテゴリのサンプル群が複数の参照ベクトルに適合することが大いに考えられる。そこで筆者等は、特徴抽出を階層型のニューラルネットで行い、それをKohonenの量子化法[2]で分類し、学習データから求めた確率辞書をもとに認識を行うという、3つの部分からなる文字認識方式を考案した。この方式に対して、手書き片仮名文字を用いて認識実験を行った。この結果について報告するとともに、本文字認識方式を考察する。

2 文字認識方式

本文字認識方式は特徴抽出部、量子化部、認識部の3つの部分からなっている。

2.1 特徴抽出部

特徴抽出部では図1のように16×16画素からなる文字データに対し、各画素においてその画素を中心とする5×5の領域を3層の階層型ニューラルネットの入力層に呈示する。このニューラルネットの入力層には25個のユニットが出力層には8個のユニットがあり、5×5の小領域から線の方向を抽出するように図1に示すような8個の特徴パターンを学習させる。またその時の学習方法は誤差逆伝搬法を用いている。従ってこの特徴抽出部では、各画素において8次元の特徴ベクトルを出力することになる。

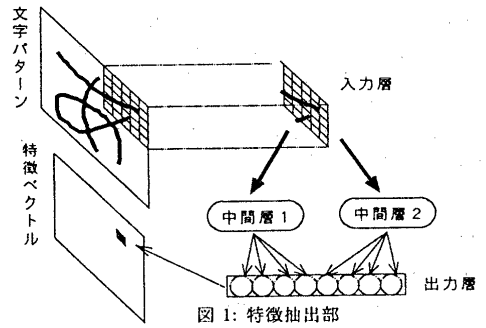

図1: 特徴抽出部

Character Recognition by Layered Neural Network and Vector Quantization Method

Yasuharu Suzuki, Masāki Yoneda, Hiroyuki Hase, Mitsuru Sakai, Takayuki Ninomiya

Toyama Univ.

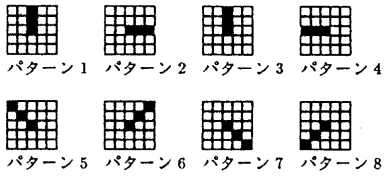

図2: 学習に使用した特徴パターン

2.2 量子化部

次段の量子化部では各画素において、前段で得られた8次元特徴ベクトルをKohonenの自己組織化アルゴリズムによって分類を行う。このとき参照ベクトルの数nは全画素について同じである。学習が終了した時点では、第k画素の参照ベクトルiには第jカテゴリのサンプルがa_{ij}個属したとすると、これが即ち生起確率$a_{ij}/\sum_i a_{ij}$となるわけである。これらの生起確率を認識に利用するため、1カテゴリが1つの参照ベクトルに対応しなくてもよい。

2.3 認識部

最後の認識部では、量子化部の参照ベクトル毎に各文字カテゴリが生起する確率辞書を持っているので、後述のように、すべての画素についてこれらをかけあわせることによって各文字カテゴリの確率が求まる。この確率が最大のカテゴリを認識結果とする。

今、任意のパターンXのk番目の画素の特徴ベクトル$\mathbf{x}_k$をKohonenのネットワークに入力する。$\mathbf{x}_k$が参照ベクトルiに属しているとしたとき、カテゴリjの下での$\mathbf{x}_k$の条件付き確率は、

$$P(\mathbf{x}_k \mid c_j) = \frac{a_{ij}}{\sum_i a_{ij}} \tag{1}$$

となる。ここで、c_jはカテゴリjである。

Bayesの法則より、

$$P(c_j \mid \mathbf{x}_k) = \frac{P(\mathbf{x}_k \mid c_j) \cdot P(c_j)}{\sum_m P(\mathbf{x}_k \mid c_m) \cdot P(c_m)} \tag{2}$$

ただし、mはカテゴリ数である。各カテゴリは等しく$\frac{1}{m}$の確率で生起すると仮定すると、

$$P(c_j \mid \mathbf{x}_k) = \frac{P(\mathbf{x}_k \mid c_j)}{\sum_l P(\mathbf{x}_k \mid c_l)} \tag{3}$$

となる。

すべての画素について$\mathbf{x}_k$が独立であると仮定すると、

$$\begin{aligned} P(c_j \mid \mathbf{X}) &= \frac{P(c_j \cap \mathbf{X})}{P(\mathbf{X})} \\ &= \frac{P(c_j \cap (\mathbf{x}_1 \cap \mathbf{x}_2 \cap \ldots \cap \mathbf{x}_l))}{P(\mathbf{x}_1 \cap \mathbf{x}_2 \cap \ldots \cap \mathbf{x}_l)} \\ &= \frac{P((c_j \cap \mathbf{x}_1) \cap (c_j \cap \mathbf{x}_2) \cap \ldots \cap (c_j \cap \mathbf{x}_l))}{P(\mathbf{x}_1 \cap \mathbf{x}_2 \cap \ldots \cap \mathbf{x}_l)} \end{aligned}$$

$$= \frac{P(c_j \cap \mathbf{x}_1) \cdot P(c_j \cap \mathbf{x}_2) \cdot \ldots \cdot P(c_j \cap \mathbf{x}_l)}{P(\mathbf{x}_1) \cdot P(\mathbf{x}_2) \cdot \ldots \cdot P(\mathbf{x}_l)}$$
$$= P(c_j \mid \mathbf{x}_1) \cdot P(c_j \mid \mathbf{x}_2) \cdot \ldots \cdot P(c_j \mid \mathbf{x}_l) \quad (4)$$

となる。ただし、lは画素数である。

3 実験

入力データは180dpiのイメージスキャナで入力した92人分の手書き片仮名46文字種を使用する。その内、一部を学習データ残りを未学習データとした。実験では、まず量子化部の参照ベクトル数を5〜20の間で変化させ認識率を比較した。この結果を図3(50人分を学習データとした)に示す。

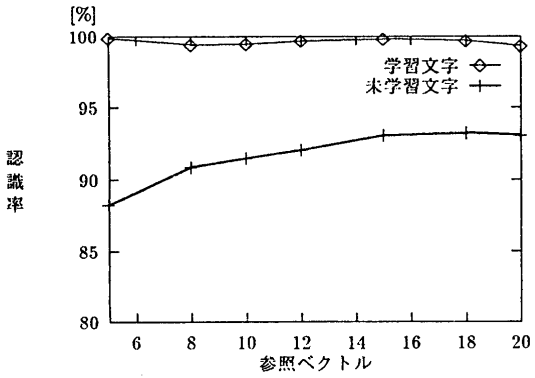

図 3: 参照ベクトル数による認識率

この実験では未学習文字について参照ベクトル数が18のとき認識率がもっともよかったので、参照ベクトル数を18に固定して、学習人数を変えた結果を図4のグラフに示す。

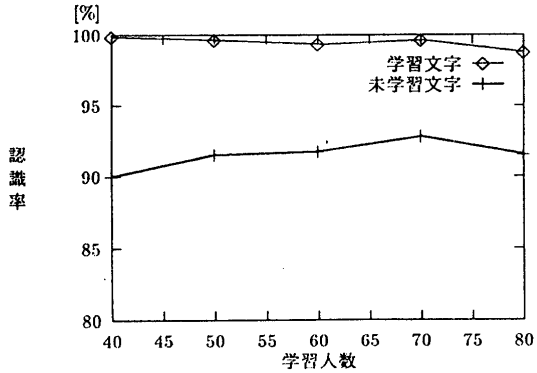

図 4: 学習人数による認識率

4 考察

最初の実験では、参照ベクトル数を増やすことによりある程度まで認識率を向上させることができた。しかし、参照ベクトル数を15以上にしても認識率はあまり変化しなかった。これは、参照ベクトル数が増えても似たようなベクトルが増えるだけで、特徴ベクトルの分類に貢献していないと考えられる。特に、周辺部では類似した特徴ベクトルが多く、設定しただけの参照ベクトルを作ることができない画素もあった。これらのことから参照ベクトル数はある程度以上増やしても無駄であると考えられる。

次の学習人数を変えた実験では、人数を増やしてもあまり認識率には影響しなかった。学習文字と未学習文字の認識率の間に7%以上の差があることから学習に使用する人数が不足しているのではないかと考えられる。

また、カテゴリ別の認識率(図5)をみると未学習データにおいて“ク”と“フ”や、“ノ”と“ン”のような類似文字で認識率が悪くなっていることがわかる。

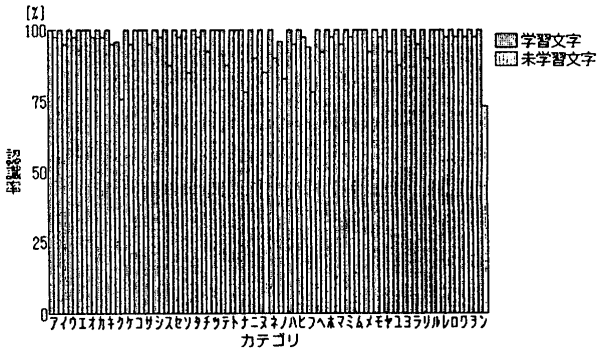

図 5: カテゴリ別認識率

5 まとめ

今回手書き文字に対して、3層の階層型ニューラルネットによって抽出した特徴ベクトルをKohoneの量子化法によって分類し、学習パターンをもとにした確率辞書で認識する手法を提案した。そして、学習文字に対して99.61%、未学習文字に対して92.82%の認識率を得た。今後はさらに学習人数を増やすことにより、未知文字の認識率の向上を計りたいと思う。また、認識部にニューラルネットを用いる方式も考えてみたい。

参考文献

[1] 岩田他：大規模4層ニューラルネット“Comb NET”, 信学論 D-II,J73-D-II,8,pp1261-1267(1990)

[2] T.Kohonen : Self-Organization and Associative Memory,Springer-Verlag(1984)