

ニューラルネットを利用した文字認識実験の一検討

4N-1

二宮孝之、米田政明、長谷博行、酒井充
(富山大学工学部電子情報工学科)

1. はじめに

文字認識においてニューラルネットワークを利用した研究が盛んに行われている。筆者等は、文字の局所的な方向成分の特徴抽出を階層型のニューラルネットで行い(特徴抽出部)、それを Kohonen の量子化法で分類し(量子化部)、学習データから求めた確率辞書をもとに認識を行う(認識部)という、3つの部分からなる文字認識方式を考案した⁽¹⁾。この方式に対して、手書き仮名文字を用いて詳細な認識実験を行った。

2. 文字認識方式

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の各画素に対し、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示す。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の小領域から線方向を抽出するように図2に示すような8個の特徴パターンを学習させる。学習方法には誤差逆伝搬法を用いている。したが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。

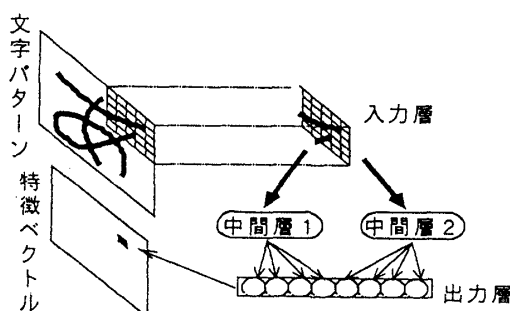


図1: 特徴抽出部

2.2 量子化部

量子化部ではすべての画素について、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

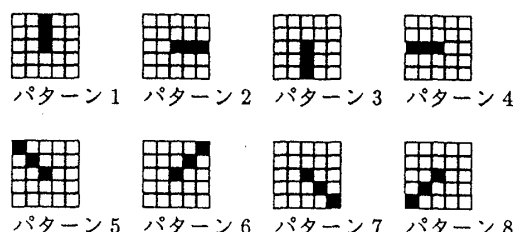


図2: 学習に使用した特徴パターン

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 はすべての画素で同じである。学習が終了した時点で、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が、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合わせることで各文字カテゴリの確率が求まる。この確率の最大となる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 x_k を Kohonen のネットワークに入力する。そのとき参照ベクトル i が活性化したとき、カテゴリ j の下での x_k の条件付き確率は、

$$P(x_k | c_j) = \frac{a_{ij}}{\sum_i a_{ij}} \quad (1)$$

となる。ここで、 c_j はカテゴリ j である。

Bayes の法則より、

$$P(c_j | x_k) = \frac{P(x_k | c_j) \cdot P(c_j)}{\sum_n P(x_k | c_n) \cdot P(c_n)} \quad (2)$$

ただし、 m はカテゴリ数である。ここで、各カテゴリは等しく $\frac{1}{m}$ の確率で生起し、すべての画素について x_k が独立であると仮定すると、

$$P(c_j | X) = P(c_j | x_1) \cdot P(c_j | x_2) \cdot \dots \cdot P(c_j | x_l) \quad (3)$$

A Consideration on Character Recognition by Layered Neural Network and Vector Quantization Method
Takayuki Ninomiya, Masaaki Yoneda, Hiroyuki Hase, Mitsuru Sakai
Faculty of Engineering, Toyama University

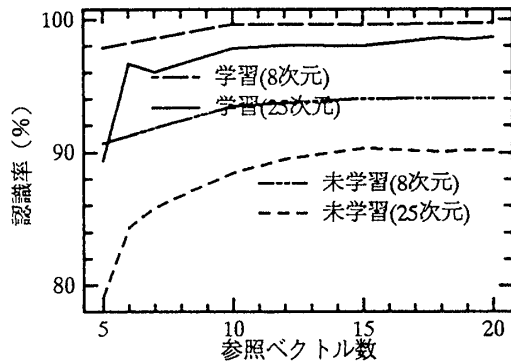


図 3: 特徴抽出の有効性

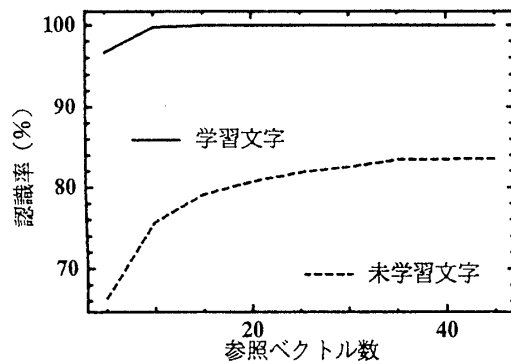


図 4: 参照ベクトル数による認識率

となる。ただし、 l は画素数である。

3. 実験および考察

文字データは 180dpi のイメージスキャナで入力した小学生 135 人の手書き平仮名 71 文字種を使用する。その内、一部を学習データ、残りをテストデータとした。

3.1 特徴抽出の有効性

本文字認識方式において、階層型ニューラルネットワークを用いた特徴抽出の有効性を調べるために、特徴抽出を施さない文字データによる認識実験を行う。すなわち、特徴抽出部に入力として利用した文字の 5×5 の小領域をそのまま 25 次元の特徴ベクトルとして量子化部で分類し、認識実験を行う (図 3)。なお、この実験については都合により手書き片仮名文字を使用した (92 サンプル/文字種、46 文字種)。

3.2 認識実験

本文字認識方式で、量子化部の参照ベクトル数を 5 ~ 45 の間で変化させ認識率を比較した。この結果を図 4 (50 人を学習データとした) に示す。

3.3 累積認識率

本文字認識方式の認識部では (3) 式が最大となるカテゴリを認識結果としているが、ここでは真のカテゴリが何位までに属しているかと言う累積認識率を示す。

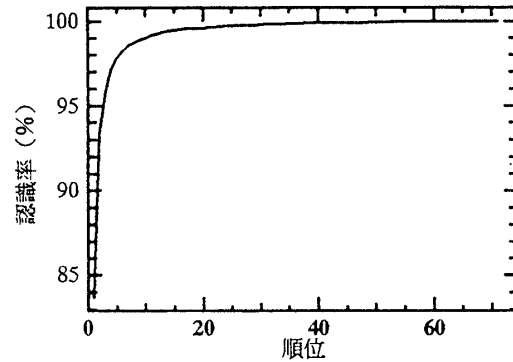


図 5: 累積認識率

これは上の実験で未学習データにおける、参照ベクトル数 40 のときの累積認識率を示す (図 5)。

特徴抽出の有効性を調べる実験では、特徴抽出を施した場合と比較して、未学習文字に対して約 3% 程度認識率が低下している。また学習文字に対しても若干認識率が低下している。したがって、本文字認識において、特徴抽出は有効であり、25 次元のベクトルを 8 次元に圧縮する効果もあると言える。

参照ベクトル数を変えた時の実験では、参照ベクトル数を 35 以上にしてもあまり認識率は変らなかった。このことから参照ベクトル数を増やし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あまり貢献していないと考えられる。

累積認識率について、第 1 位では約 83% の認識率であるが、第 5 位までについて考えると約 97% の認識率が得られている。なお、誤認識は主に濁点や半濁点のみ異なる類似文字への誤認識が多かった。

4 まとめ

今回手書き平仮名文字に対して、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 Kohonen の量子化法によって分類し、学習データをもとにした確率辞書で認識する手法によって、学習文字に対して 100 %、未学習文字に対して 83.6 % の認識率を得た。誤認識は主に類似文字への誤認識であるので、今後文字の局所的な変動についての検討をし、類似文字の認識率の向上を計りたいと思う。

参考文献

- (1) 鈴木保春：ベクトル量子化法を用いた手書き文字認識, 平成 4 年度電気関係学会北陸支部連合大会
- (2) T.Kohonen : Self-Organization and Associative Memory, Springer-Verlag(1984)