

# ニューラルネットワークの2段階構成による手書き漢字認識

梅田 三千雄<sup>†</sup> 横田 宏<sup>†</sup>

本論文では、ニューラルネットワーク(以下 NN と略す)による文字認識において、自己想起型 NN と対判定型 NN の2段階構成による手書き漢字認識システムを提案する。大分類処理には、自己想起型 NN を使用し、まず多数のカテゴリから少量の認識候補に絞る。次に、自己想起型 NN では識別しきれない認識候補を対象として、対判定型 NN により認識候補を1つに絞ることで認識結果を得る。自己想起型 NN は、各カテゴリに対して独立したネットワークが構成され、学習時に他のカテゴリの影響を受けない。したがって、対象カテゴリが変化しても、容易に対応することができる。一方、細分類処理に使用する対判定型 NN は、認識対象を2種類に限定することにより、形状の類似したカテゴリに対して高精度な認識が期待できる。形状の類似した文字に限定して対判定型 NN を作成したが、必ずしもすべての類似文字には対応していない。そこで対判定型 NN が存在しない場合の処理方法についても検討した。本システムにより、教育漢字 80 字種、240 字種、1,006 字種を対象とした認識実験において、98.71%、98.31%、96.47%という高い認識率が得られた。

## Handwritten Chinese Character Recognition by Two-stage Composition of Neural Networks

MICHIO UMEDA<sup>†</sup> and HIROSHI YOKOTA<sup>†</sup>

This paper proposes a handwritten Chinese character recognition method by two-stage composition of autoassociative neural networks and alternative neural networks. Autoassociative neural networks are used as a rough classification processing to obtain a small number of recognition candidates from all categories. Alternative neural networks are used as a discrimination processing to obtain an only recognition result from several candidates selected by the prior processing. Since autoassociative neural networks are the network which is made up of each category, the learning of networks can be executed without any influence from other categories. Therefore, even if categories to be recognized change, networks can correspond to them easily. On the other hand, alternative neural networks used in a discrimination processing can expect highly precise recognition by limiting a recognition object to two kinds of categories. The recognition rates of 98.71%, 98.31%, and 96.47% were obtained from the experiment on educational Chinese characters of 80, 240, and 1,006 categories.

### 1. はじめに

ニューラルネットワーク(以下 NN と略す)による文字認識において、多数の研究結果が報告されている。その多くは対象カテゴリ数の少ない数字など<sup>1)</sup>の認識であり、高精度な認識結果が得られている。しかし、多くの文書では漢字が使われているため、漢字を対象とした文字認識システムが不可欠となる。対象となるカテゴリ数が増加すると NN の規模が大きくなり、計算量が増大するなどの問題点がある。その回避方法として、学習時には複数の小規模なネットワークに分割しておき、認識時にそれぞれのネットワークを統合す

る認識システムが報告されている<sup>2)~6)</sup>。

NN を用いた漢字認識の例として、HoneycombNET<sup>3)</sup>や CombNET-II<sup>4)</sup>、ELNET<sup>5)</sup>などがあげられる。HoneycombNET、CombNET-II は大分類処理にベクトル量子化を用い、細分類処理では対象を少数に限定した小規模なネットワークを複数用いて、一方 ELNET は細分類処理にカテゴリごとに1つの排他的ネットワークを用い、そのカテゴリに対して発火、それ以外に対しては抑制を行うことで出力層を1つに限定することにより、ともに漢字の持つカテゴリの多さに対処している。

しかし、HoneycombNET、CombNET-II では細分類処理に使用する NN の対象をいかに設定するか、ELNET では1つのネットワークを学習するための抑制用データに複数の文字種を用いているが、その文字種

<sup>†</sup> 大阪電気通信大学大学院工学研究科  
Graduate School of Engineering, Osaka Electro-  
Communication University

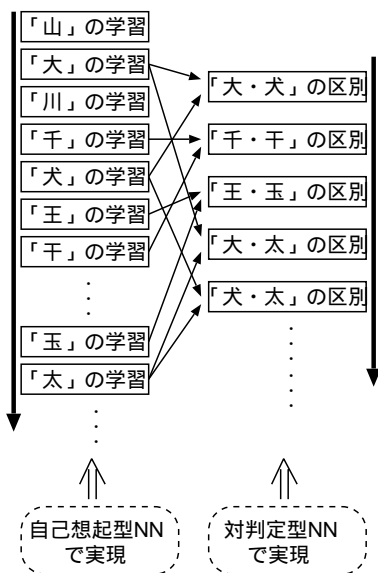


図1 学習モデル  
Fig. 1 A learning model.

をどのように選択するかで、システム全体の認識精度が大きく左右される。また、これらのネットワークでは、対象となるカテゴリが変化すると、ネットワークを一から学習して再構成しなければならないなどの問題点もある。

本論文では、この問題点を解決する方法として、大分類処理にはカテゴリごとに独立したネットワークを形成する自己想起型 NN<sup>7)</sup>、細分類処理には対象を2つのカテゴリに限定した対判定型 NN を用いる2段階構成の手書き漢字認識システムを提案する。

この認識システムは、人間の文字学習・認識過程を参考にしたものである。人間は図1に示すように1つずつ新しい漢字を繰り返し学習しながら覚えていく。そして、形状の類似した文字に出会うと、それぞれの違いを見つけ出して区別できるように覚えていく。前者を自己想起型 NN で、後者を対判定型 NN で実現しようとしたものがこの認識システムである。

自己想起型 NN とは、1つのネットワークが1つのカテゴリの文字パターンのみを学習するものであり、認識対象カテゴリが変化しても、変化したカテゴリのネットワークを新たに作成するだけで容易に対応することができる。また、これは砂時計型 NN とも呼ばれ、入力層と出力層のユニット数が等しく、中間層のユニット数はそれより少ない個数で構成し、入力値がそのまま理想出力値となるように学習するものである。したがって、入力パターンに対する誤差の大小によって文字認識を行うことができる。

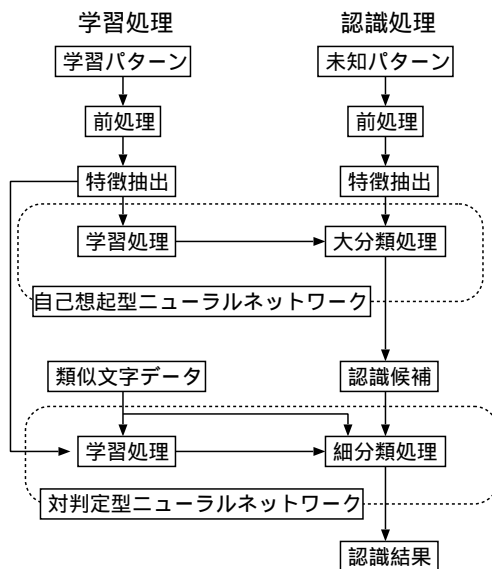


図2 認識システムの流れ  
Fig. 2 Flow of recognition system.

対判定型 NN とは、出力層のユニット数が2個で、それぞれにカテゴリを対応させたものである。したがって、各出力ユニットの出力値によって、2種類のカテゴリを分類することができる。

以下本論文では、2章で提案するシステム構成を示し、3章で前処理と特徴抽出、4章で大分類処理である自己想起型 NN、5章で細分類処理である対判定型 NN、6章で ETL9B に対する認識実験の結果と考察、7章で結論を述べる。

## 2. システム概要

文字認識システムの処理の流れを図2に示す。学習処理では、まず前処理として大きさの正規化、太さ変換、非線形正規化、スムージングの各処理を施す。次に、正規化後のパターンから特徴抽出を行う。特徴抽出には、文字認識においてすでに有効性の確認されている加重方向指数ヒストグラム特徴<sup>9)</sup>を用いた。得られた特徴量を自己想起型 NN および対判定型 NN に入力して、誤差逆伝播法(バックプロパゲーション法)により出力値と理想出力値との誤差が小さくなるように学習し、NN を作成する。

自己想起型 NN は対象カテゴリの数だけ作成すればよいが、そこから2つのカテゴリを選択して区別する組合せ総数は膨大となる。そこで、漢字の形状が類似している組合せのみに限定した類似文字データを作成し、その類似文字データを基に対判定型 NN を作成することにした。

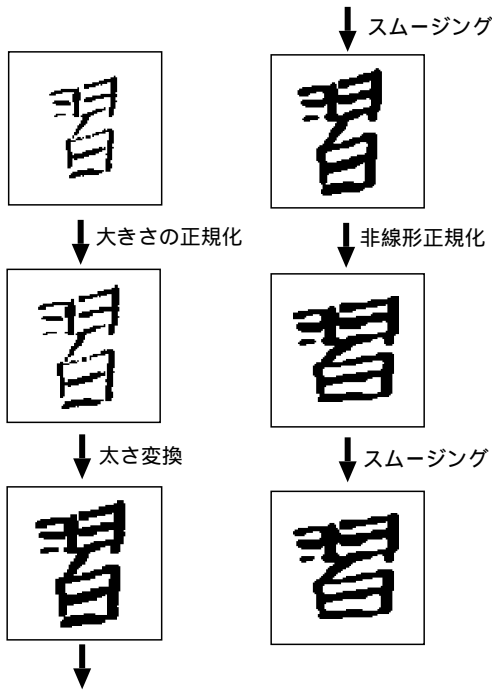


図3 前処理による文字パターンの正規化

Fig. 3 Normalization of character pattern by pre-processing.

認識処理では、入力パターンに対して学習時と同様の前処理を施し、特徴抽出して、自己想起型 NN により少量の認識候補に絞る大分類処理を行い、さらに対判定型 NN により少量の認識候補を対象とした細分類処理を行って、認識結果を得る。なお認識候補が 3 つ以上の場合には、対判定型 NN を複数回使用し、認識候補を 1 つずつ削除していく消去法的な手法を用いた。

### 3. 文字パターンの前処理と特徴抽出

前処理では、文字パターン  $f(x, y)$  の重心回りの 2 次モーメント

$$r_m = \frac{\sum f(x, y) \cdot \sqrt{(x - X_m)^2 + (y - Y_m)^2}}{\sum f(x, y)} \quad (1)$$

$r_m$  : 2 次モーメント  $X_m, Y_m$  : 重心

を一定にする大きさの正規化、かすれ字やつぶれ字を補正するための太さ変換、文字線の配置を均一化する非線形正規化<sup>8)</sup>、文字線の輪郭部分を平滑化するスムージングの各処理を施す。なお、非線形正規化により文字の凹凸部分が強調されてしまうため、スムージングは非線形正規化の前後で行うことにした。前処理による文字パターンの正規化の例を図 3 に示す。

特徴抽出法には加重方向指数ヒストグラム特徴を用

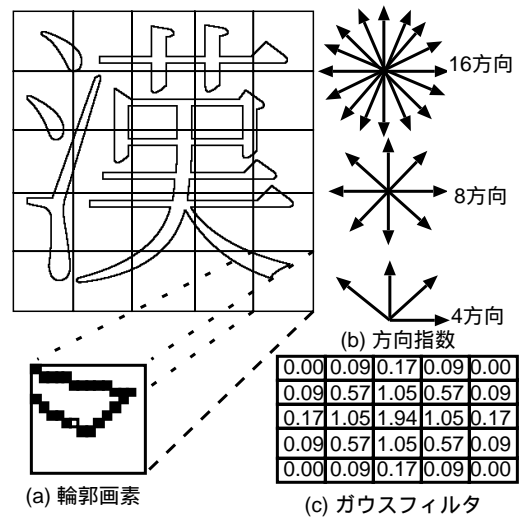


図4 加重方向指数ヒストグラム特徴

Fig. 4 Weighted direction index histogram feature.

いた。文字パターンに対して、図 4 (a) のように輪郭を追跡し、輪郭画素それぞれにおいて図 4 (b) の方向指数を得る。方向指数とは、対象画素とそれととなりあう輪郭画素の位置関係により算出される。これにより得られた 16 方向の方向指数を主要となる偶数番目の方向を中心に 1:2:1 の重みを付け 8 方向に圧縮する。さらに反対方向を同一視することにより 4 方向まで圧縮する。64 × 64 画素の文字パターン領域を 16 × 16 個の小領域に分割し、それぞれの領域ごとに方向指数のヒストグラムを算出し、さらに図 4 (c) に示すガウスフィルタを 1 つおきにかけることにより、8 × 8 個の領域に圧縮する。以上の手順により 4 (方向) × 64 (領域) = 256 次元の特徴量が抽出される。なお、本実験における特徴次元数は、256 次元のものと小領域への分割を 4 × 4 個に変更することで得られる 64 次元の 2 種類の特徴量を使用し、256 次元を高次元、64 次元を低次元と呼ぶことにする。

### 4. 自己想起型ニューラルネットワーク

自己想起型 NN とは、入力層と出力層のユニット数が等しく、中間層のユニット数がそれより少ない砂時計型 NN である。また、出力層の理想出力には入力値そのものを与えて学習させることにより、入力層と中間層の間には圧縮・符号化処理、中間層と出力層の間には復号化処理のネットワークが形成される。したがって、この NN をパターン認識に応用する場合には、1 つの NN に 1 つのカテゴリのパターンを学習させて、対象カテゴリの分だけ構成し、入力値と各 NN の出力値との誤差の大きさを比較することにより、入力

パターンがどのカテゴリに属するかを判断する。つまり、この NN の学習には他のカテゴリは関係なく、カテゴリごとに独立したネットワークを形成し、学習を容易に実現することができる。

#### 4.1 構成と学習

自己想起型 NN は、3 層階層構成とした。ユニット数は、特徴次元数に合わせ、入力層および出力層では 256 個、中間層では 128 個のものと、入力層および出力層では 64 個、中間層では 32 個からなる 2 種類を用い、学習回数は 1,000 回とした。対象カテゴリの文字パターンの組を入力して、誤差逆伝播法により、理想出力値（教師信号）と出力値の誤差

$$D = \frac{1}{N} \sum_{i=1}^N (T_i - O_i)^2 \quad (2)$$

$D$ : 誤差  $N$ : 特徴次元数

$T_i$ : 教師信号  $O_i$ : 出力値

が小さくなるように学習する。

自己想起型 NN は、認識対象とするカテゴリが増加しても、その増加したカテゴリの NN だけを新たに学習して追加するだけでよく、それまでのネットワークをそのまま利用できるのが特長である。

#### 4.2 大分類処理

自己想起型 NN は、カテゴリごとに独立したネットワークを形成し、各ネットワークはそのカテゴリのみの特徴を学習する。その結果、ある文字パターンを入力すると、そのパターンの属するカテゴリを学習したネットワークにおける入力値と出力値の誤差は、他のカテゴリを学習したネットワークに比べて一般に小さくなる。しかし、異なるカテゴリであっても形状の類似した文字に対して、誤差が小さくなることもあり、誤認識の可能性がある。そこで、自己想起型 NN における認識処理では、結果を 1 つに絞るのではなく、複数の候補を抽出する大分類処理と位置付けて使用する。

大分類処理では、まず認識しようとする文字パターンに対して、前処理を施して特徴抽出し、認識対象となるカテゴリのネットワークそれぞれに入力し誤差を求める。得られた誤差の最小のカテゴリを第 1 認識候補とし、最小の誤差からの相対距離がある一定範囲内にカテゴリが存在すれば、そのカテゴリを認識候補とする。

たとえば、ある文字パターンの入力に対して、対象カテゴリのすべてのネットワークで生じる誤差を算出し、それを距離値としてソートした結果が図 5 のようになったとして、仮に相対距離 30 をしきい値として大分類すれば、「官」から「言」までの 5 カテゴリが認

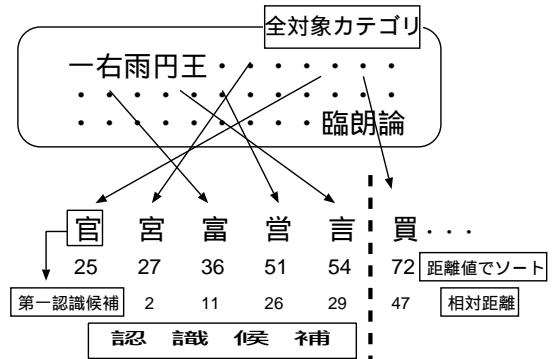


図 5 大分類処理の例

Fig. 5 Example of rough classification processing.

識候補となる。

### 5. 対判定型ニューラルネットワーク

自己想起型 NN だけでは認識の困難な 2 種類の文字を区別するために、出力層のユニットに各カテゴリを対応させた対判定型 NN を用いる。つまり、対判定型 NN の出力層のユニット数は 2 個であり、それぞれの出力値の大小により入力文字がどちらのカテゴリに属するかを判定する。また、認識対象を 2 種類に限定することにより、高速な学習ができ、認識精度は高くすることが期待できる。

#### 5.1 類似文字

対象カテゴリすべてに対してネットワークを作成するのが理想であるが、対象カテゴリ数が増えるとその組合せ総数は爆発的に増大するため、組合せすべてに対して対判定型 NN を作成することが困難となる。そこで、形状の似ている文字に限定して類似文字を選出し、その類似文字に対してのみ対判定型 NN を作成することにする。

類似文字の設定には自己想起型 NN を用いた。その選出には、各カテゴリの平均パターンあるいは平均としての類似性に着目する方法や、個々の文字パターンのうちの最も近いものに着目する方法などが考えられる。ここでは、自己想起型 NN で生じる誤差を尺度として、後者の方法により設定した。具体的には、まずすべての学習パターン（1,006 字種、各 100 パターン）を 1 つずつすべての自己想起型 NN に入力して、それぞれのネットワークでの誤差を算出し、入力パターンのカテゴリに対応したネットワークでの誤差との差を相対距離として求める。次に、この相対距離でカテゴリ対をソートし、その小さいものから順に重複を削除しながら 5,000 種類に達するまで選出して類似文字データとした。したがって、1 つの学習パターンから

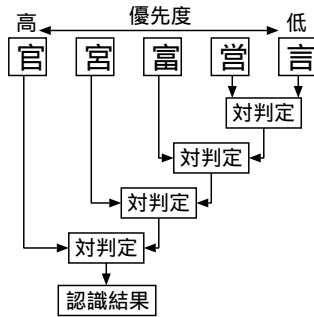


図 6 細分類処理の例

Fig. 6 Example of discrimination processing.

複数の類似文字の組合せが設定されることもある。

### 5.2 構成と学習

対判定型 NN も 3 層階層構成とした。ユニット数は入力層では 256 個，中間層では 128 個，出力層では 2 個とし，学習回数は 1,000 回とした。この NN では，類似文字データをもとにして，2 種類のカテゴリを区別するように学習する。つまり，それぞれのカテゴリを出力層に対応付けて，入力パターンのカテゴリに対応する方のユニットに 1，他方に 0 を教師信号とし，その教師信号との誤差が小さくなるように誤差逆伝播法を用いて学習する。

対判定型 NN でも類似文字ごとに独立したネットワークが形成される。

### 5.3 細分類処理

大分類処理によって得られた複数のカテゴリから 1 つのカテゴリに絞るのが細分類処理である。そのため，細分類処理では大分類処理に比べて高精度の処理能力が必要となる。ここでは，認識対象を 2 つのカテゴリに限定することにより，高精度な認識が可能になると考えられる対判定型 NN を用いる。自己想起型 NN による大分類処理において，認識候補が複数個ある場合の処理例を図 6 に示す。大分類処理により得られたカテゴリに，誤差の小さい順に優先度を与える。次に，優先度の低い 2 つのカテゴリに対応する対判定型 NN に入力して，出力値の大きいカテゴリを候補として残り，小さいカテゴリを削除する。これを候補が 1 つになるまで繰り返し，最終的に残ったカテゴリを認識結果とする。

しかし，対判定型 NN は類似文字データにより少量しか作成されていない。そのため，細分類処理時に対象となる対判定型 NN が存在しないことがある。そこで，この問題に対処するために，次の 3 種類の方法について検討する。

手法 1：優先度の高いカテゴリを選択し，低い方のカ

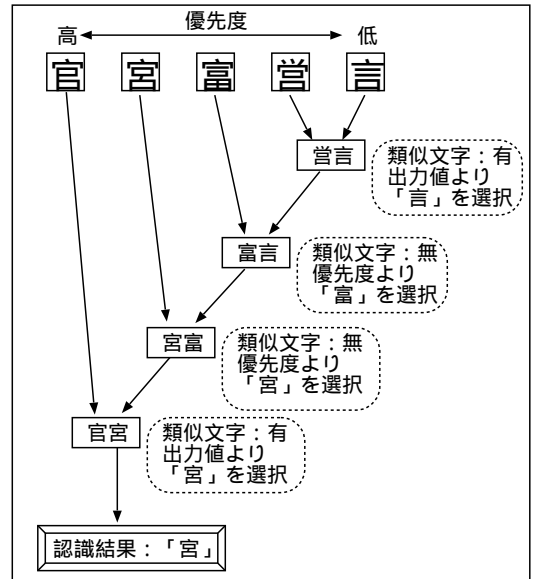


図 7 手法 1 による動作例

Fig. 7 Example of operation by technique 1.

テゴリを消去する。この動作例を図 7 に示す。1 回目の対判定では，対判定型 NN が作成されているので，出力値により「言」を残し「宮」を削除する。2 回目の対判定では，対判定型 NN が作成されていないので，優先度の高い「富」を残し「言」を削除する。以下同様の手順により「宮」が認識結果となる。

手法 2：大分類処理により得られたカテゴリは形状が類似しているものを多く含んでおり，かつその中に正解のカテゴリも含まれている可能性が高い。そこで対判定の対象となる 2 種類のカテゴリそれぞれについて，認識候補内での類似文字の数により判定する。その類似文字の多い方のカテゴリを選択し，少ない方のカテゴリを消去する。なお，類似文字が同数の場合は手法 1 を用いる。この動作例を図 8 に示す。1 回目の対判定は，手法 1 と同じである。2 回目の対判定では，対判定型 NN が作成されていないので，認識候補内での類似文字をそれぞれ選出し，類似文字の多い「言」を残し「富」を削除する。

手法 3：手法 2 に加え，それまでの対判定型 NN により選択された回数を考慮する。つまり，手法 2 における類似文字の個数に対判定型 NN により選択された回数を加算して判定値とし，その大きい方のカテゴリを選択し，小さい方のカテゴリを削除する。なお，判定値が同数の場合は手法 1 を用いる。この動作例を図 9 に示す。この

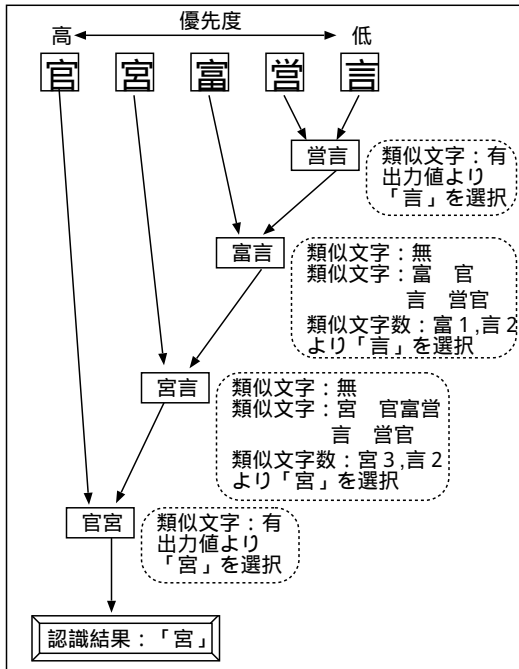


図8 手法2による動作例  
Fig. 8 Example of operation by technique 2.

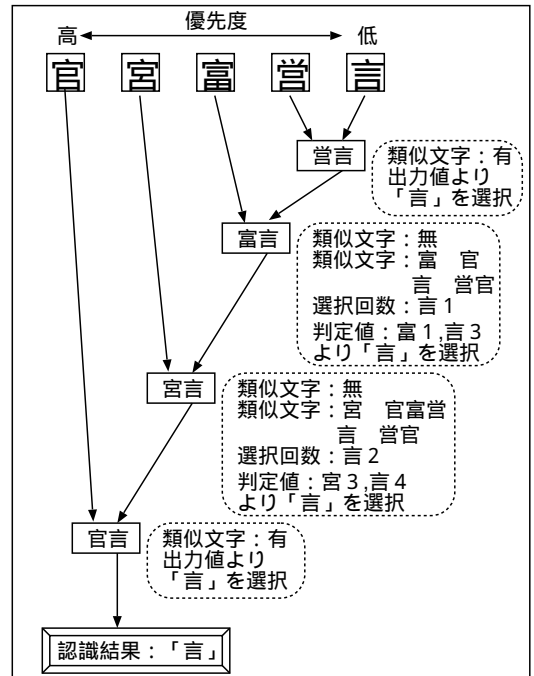


図9 手法3による動作例  
Fig. 9 Example of operation by technique 3.

方法では、たとえば2回目の対判定において、対判定型 NN が作成されていないので、認識候補内での類似文字をそれぞれ選出し、さらに「言」が選択された回数を加算したものを判定値とし、判定値の大きい「言」を残し「富」を削除する。

### 6. 結果と考察

認識実験には、電子技術総合研究所において作成された JIS 第 1 水準手書き文字データベース ETL9B を用いた。ETL9B には各文字種について 200 個の文字パターンがあり、その奇数番目を NN の学習に用いる学習パターン、偶数番目を認識システムの性能評価のための未知パターンとして使用した。認識対象は小学校 1 年生で習う漢字 80 字種、さらに 2 年生で習う漢字 160 字種を合わせた 240 字種、そして小学校の 6 年間で習うすべての漢字 1,006 字種と変化させて認識実験を行った。対判定型 NN は 1,006 字種を認識対象として 5,000 種類用意した。この 5,000 種類とは、1,006 字種から 2 種類の文字を選択するすべての組合せが約 50 万通りであるため、全体の約 1.0% に相当する。なお、この 5,000 種類のうち 1 年生で習う 80 字種により構成されているものが 122 種類、1 年生と 2 年生で習う 240 字種により構成されているものが 437 種類存

在し、それぞれ組合せ総数の 7.8%、1.5% に相当する。

自己想起型 NN を単独で用いたときの認識結果を表 1 に示す。累積認識率とは、最小誤差から一定の相対距離の範囲に含まれるカテゴリ群を認識候補とし、その中に正解のカテゴリが含まれる割合である。また、そのときの認識候補数の平均値を平均候補数として示した。したがって、相対距離 0 のときが通常の認識率に相当する。この結果より、認識候補を第 1 候補からの相対距離によって設定することにより、少量の候補数で高い累積認識率の得られることが確認できた。しかし、低次元の場合には、同じ相対距離であっても候補数は多くなることも明らかになった。これは、高次元に比べて特徴量が減少したことにより、それぞれのカテゴリの特徴が類似したためではないかと考えられる。

次に、対判定型 NN のみによる認識結果を表 2 に示す。この結果は、3 種類の認識対象それぞれに存在する類似文字に対する認識率の平均値と、その最低値、最高値である。最低値だけを見るとあまり高い認識率ではないが、特定の組合せ数種類を除くと、ほぼすべての類似文字において 100% 近い値であった。

なお、最低値となる類似文字は、対象カテゴリ数 80 字種では「石・右」、240 字種では「間・聞」、1,006 字種では「千・干」であった。これらの文字は、小領域

表 1 自己想起型 NN による結果

Table 1 Recognition result by autoassociative neural networks.

累積認識率				
特徴次元数	相対距離	80 字種	240 字種	1,006 字種
256	0.000	98.03%	97.43%	94.61%
	0.001	99.20%	98.85%	97.67%
	0.002	99.56%	99.45%	98.91%
	0.003	99.75%	99.70%	99.46%
	0.004	99.86%	99.84%	99.70%
64	0.000	97.04%	95.08%	87.94%
	0.001	98.96%	98.18%	95.77%
	0.002	99.44%	99.10%	98.06%
	0.003	99.70%	99.56%	99.02%
	0.004	99.78%	99.71%	99.46%
平均候補数				
特徴次元数	相対距離	80 字種	240 字種	1,006 字種
256	0.000	1.00	1.00	1.00
	0.001	1.05	1.06	1.18
	0.002	1.15	1.19	1.61
	0.003	1.32	1.44	2.63
	0.004	1.59	1.89	4.71
64	0.000	1.00	1.00	1.00
	0.001	1.12	1.24	1.99
	0.002	1.33	1.89	5.73
	0.003	1.72	3.33	14.39
	0.004	2.31	5.76	28.89

表 2 対判定型 NN による認識率

Table 2 Recognition rate by alternative neural networks.

対象数	類似文字数	平均値	最低値	最高値
80	122	99.26%	95.00%	100.00%
240	437	99.34%	93.50%	100.00%
1,006	5,000	99.39%	82.00%	100.00%

での文字線の長さや角度が少し異なるだけである。そのため粗領域に分割し、方向圧縮する加重方向指数ヒストグラム特徴では、顕著な違いが表現されないのが原因であろう。

自己想起型 NN による大分類処理において、相対距離を 0.002, 0.003, 0.004 に設定し、対判定型 NN との 2 段階構成による認識実験を行った。なお、対判定型 NN が存在しない場合には、5 章で提案した 3 種類の手法を用いた。それぞれの手法による認識結果を表 3, 表 4, 表 5 に示す。

この結果より、手法 3 において最も高い認識率が得られた。その認識率は、高次元では、それぞれ 98.71%, 98.31%, 96.47%, 低次元では、98.39%, 96.95%, 92.28% となる。これは、自己想起型 NN を単独で用いた場合に比べて、高次元で、0.63%, 0.88%, 1.86%, 低次元で、1.35%, 1.87%, 4.34% の認識率の向上になっている。手法 1 では、正解カテゴリの優先

表 3 手法 1 による認識率

Table 3 Recognition rate by technique 1.

特徴次元数	相対距離	80 字種	240 字種	1,006 字種
256	0.002	98.58%	98.14%	96.01%
	0.003	98.59%	98.11%	95.98%
	0.004	98.58%	98.10%	95.96%
64	0.002	98.04%	96.33%	90.68%
	0.003	98.03%	96.38%	90.71%
	0.004	97.96%	96.36%	90.70%

表 4 手法 2 による認識率

Table 4 Recognition rate by technique 2.

特徴次元数	相対距離	80 字種	240 字種	1,006 字種
256	0.002	98.56%	98.17%	96.12%
	0.003	98.59%	98.15%	96.10%
	0.004	98.56%	98.15%	96.16%
64	0.002	98.09%	96.44%	91.18%
	0.003	98.05%	96.47%	91.24%
	0.004	97.99%	96.46%	91.23%

表 5 手法 3 による認識率

Table 5 Recognition rate by technique 3.

特徴次元数	相対距離	80 字種	240 字種	1,006 字種
256	0.002	98.66%	98.31%	96.47%
	0.003	98.71%	98.28%	96.37%
	0.004	98.65%	98.25%	96.18%
64	0.002	98.39%	96.95%	92.28%
	0.003	98.33%	96.80%	91.81%
	0.004	98.25%	96.53%	91.43%

度が低い場合に類似文字が存在しないと、簡単に削除されてしまうことが認識率の向上をさまたげたのであろう。また、手法 3 では、類似文字だけで選択する手法 2 に比べて、選択回数を考慮したことにより、正解カテゴリに類似文字が少量である場合でも、簡単に削除されないことが、認識率の向上に結び付いたと考えられる。しかし、相対距離の違いによる認識率の変化はあまり見られない。これは認識候補を多く設定しても十分対応できる類似文字を用意していないためであろう。

さらに、自己想起型 NN 単独での認識と本システムとを比較するために、各手法における誤認識訂正率を求めたものを表 6, 表 7, 表 8 に示す。誤認識訂正率とは、本システムにおける誤認識数が自己想起型 NN 単独での誤認識数からどれだけ減少したかの割合である。

この結果より、3 種類の手法による違いはほとんど見られないが、対象カテゴリによる違いは大きく、80 字種では、240 字種や 1,006 字種に比べて、かなり高い値となっている。これは、作成した対判定型 NN の数が、それぞれの組合せ総数の 7.8%, 1.5%, 1.0% であり、80 字種は 240 字種や 1,006 字種に比べて、多

表6 手法1における誤認識訂正率

Table 6 Correction rate of mis-recognized characters by technique 1.

特徴次元数	相対距離	80 字種	240 字種	1,006 字種
256	0.002	27.22%	27.66%	26.02%
	0.003	28.48%	26.30%	25.30%
	0.004	27.85%	25.97%	25.08%
64	0.002	33.76%	25.42%	22.72%
	0.003	33.33%	26.36%	22.94%
	0.004	31.22%	25.93%	22.91%

表7 手法2における誤認識訂正率

Table 7 Correction rate of mis-recognized characters by technique 2.

特徴次元数	相対距離	80 字種	240 字種	1,006 字種
256	0.002	27.22%	28.90%	27.99%
	0.003	28.48%	28.08%	27.59%
	0.004	27.22%	27.92%	26.92%
64	0.002	35.44%	27.54%	26.85%
	0.003	34.18%	28.22%	27.40%
	0.004	32.07%	27.97%	27.30%

表8 手法3における誤認識訂正率

Table 8 Correction rate of mis-recognized characters by technique 3.

特徴次元数	相対距離	80 字種	240 字種	1,006 字種
256	0.002	32.28%	34.28%	34.42%
	0.003	34.81%	33.12%	32.61%
	0.004	31.65%	31.66%	29.14%
64	0.002	45.57%	38.05%	36.01%
	0.003	43.46%	34.83%	32.05%
	0.004	40.93%	29.32%	28.91%

くの対判定型 NN が作成されていたためであると考えられる。

また、提案システムの性能を比較するために、出力層の各ユニットに対象カテゴリを対応付けた3層の階層型 NN による認識実験を行った。入力層のユニット数は256個とし、高次元の加重方向指数ヒストグラム特徴を入力とした。中間層は一律に128個とし、出力層は認識対象カテゴリに対応させてユニットを配置し、カテゴリに対応したユニットに1、他には0を教師信号として学習させた。学習回数は、前述の実験と同じく1,000回である。

実験結果は、80、240、1,006字種の認識対象それぞれについて、98.30%、96.56%、89.61%の認識率であった。いずれの認識対象の場合でも、本システムの認識率の方が高く、提案システムの有効性が示唆される。特に、認識対象を拡大するにつれて、その差が顕著になる。逆に、カテゴリ対応型 NN において、1,006字種を対象としたときの認識率が大きく低下しているのは、認識対象の多少にかかわらず、中間層のユニ

表9 本システムのネットワーク規模

Table 9 Network size of this recognition system.

	自己想起型 NN		対判定型 NN
	低次元	高次元	
ユニット数	入力層	64	256
	中間層	32	128
	出力層	64	256
NNの個数	1,006		5,000

ト数を一律に128個としたことによるのではないかと考えられる。

1,006字種を認識対象としたときの本システムのネットワーク規模を表9に示しておく。また、1つのネットワークの学習に要する時間は、ワークステーション(Origin 2000)の1プロセッサを使用して、高次元の自己想起型 NN で約40分、対判定型 NN は約10分であった。学習は、ネットワークごとに独立して実行できるため、複数のプロセッサを使用することにより、全体の学習時間は大幅に短縮することができる。1文字の認識時間は約0.8秒であった。これも、システムとして最適化したり、並列処理を導入することによって、高速化することが可能である。

## 7. おわりに

本論文では、自己想起型 NN による手書き文字認識に着目し、自己想起型 NN と対判定型 NN の2段階構成により、人間の学習・認識過程をモデル化した認識システムを提案した。大分類処理では、自己想起型 NN を使用して少量の認識候補を選出し、細分類処理では、自己想起型 NN で認識しきれない候補カテゴリを対象として、対判定型 NN によって1つずつ認識候補を削除していき、最終的に1つの認識結果を得る。本システムは、認識精度の低下につながる形状の類似したカテゴリに対して、対判定型 NN を用いることにより、認識精度の低下を抑えた。しかし、組合せすべてに対して対判定型 NN を作成することは困難であり、ここでは類似文字データを作成し、その類似文字に対して対判定型 NN を作成することで対処した。

その結果、認識対象が80字種では、98.71%、240字種では98.31%、1,006字種では96.47%と高い認識率を得ることができた。しかし、それぞれの相対距離における累積認識率には到底及んでいない。その原因として、対判定を行うときに類似文字として存在しないため、本来正解として出力されるべきカテゴリが削除されてしまうこと、また2種類の NN に同一の特徴を使用したことにより、同じような文字パターンで誤認識するところにある。



今後は、対象字種の拡大にともなう対判定型 NN の爆発的增加への効果的な対処方法、類似文字が存在しない場合の回避方法や類似文字の選出方法などについて検討する必要がある。さらに、それぞれの NN に異なる特徴抽出法を用いる認識の高精度化についても検討する必要がある。

謝辞 手書き文字データベース ETL9B を提供して下さった電子技術総合研究所の皆様に深く感謝いたします。

### 参 考 文 献

- 1) 岩田 彰, 高橋秀和, 堤田敏夫: ニューラルネットワークによる郵便番号手書き数字認識, 信学技報, PRU95-2 (1995).
- 2) 石井 力, 和泉勇治, 加藤 寧, 根元義章: ニューラルネットワークを用いた二者択一式の手書き文字認識システム, 信学論, Vol.J83-D-II, No.3, pp.988-995 (2000).
- 3) 荒井正之, 奥田健三, 宮道壽一: Honeycomb ネット II による多字種の手書き漢字認識, 信学論, Vol.J77-D-II, No.9, pp.1708-1715 (1994).
- 4) 猪野 浩, 岩田 彰: CombNET-II を用いた JIS 第 1 水準手書き漢字の認識, 信学技報, PRU93-151 (1994).
- 5) 猿田和樹, 孫 寧, 安部正人, 根元義章: 排他的学習ネット (ELNET) を用いた ETL9B の認識, 信学技報, PRU94-93 (1994).
- 6) 萩原義裕, 小畑秀文: 大規模神経回路網の統合と手書き文字認識への応用, 信学論, Vol.J82-D-II, No.11, pp.1940-1948 (1999).
- 7) 井上 聡, 若林哲史, 鶴岡信治, 木村文隆, 三宅康二: 自己想起回路による手書き数字認識, 情

報処理学会論文誌, Vol.39, No.8, pp.2476-2484 (1998).

- 8) 若林哲史, 鄧 嚴, 鶴岡信治, 木村文隆, 三宅康二: 非線形正規化と特徴量の圧縮による手書き漢字認識の高精度化, 信学論, Vol.J79-D-II, No.5, pp.765-774 (1996).
- 9) 鶴岡信治, 栗田昌徳, 原田智夫, 木村文隆, 三宅康二: 加重方向指数ヒストグラム法による手書き漢字・ひらがな認識, 信学論, Vol.J70-D, No.7, pp.1390-1397 (1987).

(平成 13 年 3 月 26 日受付)

(平成 13 年 9 月 12 日採録)



梅田三千雄 (正会員)

昭和 20 年生。昭和 43 年大阪大学卒業。同年日本電信電話公社 (現 NTT) 入社。平成元年大阪電気通信大学工学部教授。現在、同総合情報学部教授。工学博士。文字認識, 画像処理, 認知科学等の研究に従事。電子情報通信学会, 映像情報メディア学会, 画像電子学会会員。



横田 宏 (学生会員)

昭和 51 年生。平成 11 年大阪電気通信大学情報工学部情報工学科卒業。同年同大学大学院工学研究科博士課程前期情報工学専攻入学, 現在在学中。文字認識, 特にニューラルネットワークによる手書き漢字認識に関する研究に従事。