

外郭局所的輪郭線特徴を用いた 統合型ニューラルネットワークによる類似文字認識

堀 桂太郎[†] 菅原 広喜[†] 伊藤 彰義^{††}

外郭局所的輪郭線 (P-LOVE) 特徴は、文字を周囲 8 方向から走査して得られる外郭点を起点に、文字の輪郭線の連結情報を抽出するものである。本論文では、P-LOVE 特徴を外郭深度別に 3 分割し、各特徴ベクトルごとに個別のニューラルネットワーク (NN) により識別を行い、さらにそれらの結果を統合処理する NN システムを構成した。このシステムで ETL9B を用いて、詳細分類で対象となる類似文字集合の識別実験を行った結果、外郭深度を分割せずに一括して 1 つの NN で識別するよりも高い識別率を得た。また、統合処理する NN システムにおいて、1 段目の NN の出力結果を多数決判定してリジェクト処理を行うことについて検討した。

An Unified Neural Network for Handwritten Similar Characters Recognition by Using Peripheral Local Outline Vector

KEITARO HORI,[†] HIROKI SUGAWARA[†] and AKIYOSHI ITOH^{††}

We have proposed the feature extraction method: peripheral local outline vector (P-LOVE) already. P-LOVE extracts a feature at a peripheral point on a character which is scanning from 8 directions around the character. In this paper, we propose the unified neural network (NN) system in which each NN classifies patterns depending on the P-LOVE features for the different peripheral depth. Experimental results show that the system achieves higher recognition rates than a single NN for handwritten similar characters of ETL9B. In addition, we consider a rejecting method by using a majority decision of NNs in the 1st stage.

1. はじめに

コンピュータの普及にともなって、手書き漢字認識システムの高精度化がますます大きな課題となってきた。認識システムでは、特徴抽出部において対象文字の特徴ベクトルを抽出した後、識別部において候補字種を特定するのが一般的である。したがって、認識システムを高精度化するためには、有効な特徴ベクトルを抽出することが重要な一要因となる。

手書き漢字認識に有効な特徴抽出法として、先に我々が提案した外郭局所的輪郭線特徴¹⁾は、文字を周囲 8 方向から走査して得られる外郭点²⁾を起点にして、文字の輪郭線を局所的に観測し、その連結情報を特徴ベクトルとするものである。

本論文では、外郭局所的輪郭線特徴を外郭深度別に 3 分割し、それぞれの特徴ベクトルごとに個別のニューラルネットワーク (以下、NN と呼ぶ) で識別を行い、さらにそれらの結果を統合処理する NN を用いた認識システム³⁾を提案する。

これまで、複数の特徴量を合成して文字認識を行う手法がいくつか報告^{4),5)}されている。文献 4) は手書き類似漢字、文献 5) は手書き数字を認識対象としたものである。いずれの報告でも既存の複数の特徴抽出法を採用して実験を行っている。各報告の効果をさらに向上させるためには、将来、異なった構造情報を有効に抽出する複数の高性能な特徴抽出法の開発が期待される。

統合処理する NN を文字認識に応用する試みも、これまでに報告^{6)~8)}されている。文献 6) はマルチフォントにおける追加フォントの学習方法、文献 7) は大規模な NN の構成方法であり、文献 8) は類似文字集合を対象にして複数の識別手法によって得られた結果を統合処理することで識別の信頼性向上をねらったものである。

[†] 日本大学大学院理工学研究科
Graduate School of Science and Technology, Nihon University

^{††} 日本大学理工学部電子工学科
Department of Electronics Engineering, College of Science and Technology, Nihon University

一方、本論文は、単独の特徴抽出法で得られた特徴ベクトルを効果的に利用することを課題とし、そのために統合型 NN を使用して識別率の向上を目的とするものであり、上記の報告とは着目点および課題を異にする。

また、文献 9) では、改良型方向線素特徴¹⁰⁾を用いた場合に、特徴量の要素の相関を考慮してマハラノビス距離の近似値をより正確に求める手法が報告されている。本論文は、単独の特徴抽出法で得られた特徴量の有効な利用法という観点においては文献 9) と同じ立場をとるが、用いた特徴抽出法や特徴ベクトルの分割の仕方、また使用する識別関数が異なる。

今回提案する認識システムで類似文字集合の識別実験を行った結果、特徴ベクトルを外郭深度別に分割して統合処理を行う認識システムは、特徴ベクトルを分割せず一括して 1 つの NN で識別するよりも高い識別率が得られることを確認した。

また、統合処理を行う NN システムにおいて、1 段目および 2 段目の NN の識別結果を検討し、最終的な識別率が向上した理由について考察する。

認識システムで誤認識の数を減らすための有効な一手法にリジェクト処理がある。誤認識する可能性の高い文字が入力された際には、その文字をリジェクト処理して、他の識別方法などを適用して正認識へ導くことで、システム全体の識別率を向上させることが期待できる。

そこで本論文では、1 段目の NN の認識結果を多数決判定してリジェクト処理を行った場合の、リジェクト率や識別率についても考察する。

2. 外郭局所的輪郭線特徴

従来から手書き漢字認識に有効とされている特徴抽出法に、外郭方向寄与度 (peripheral direction contributivity) 特徴²⁾ (以下、P-DC 特徴と呼ぶ) がある。P-DC 特徴は、文字線の複雑さ、文字線の方向、文字線の接続関係、文字線の相対位置関係の構造特徴を反映する特徴抽出法と考えられ広く使われている¹¹⁾。しかし、P-DC 特徴は文字線の内部において 8 方向に触手を伸ばすことで特徴ベクトルを得る手法であるため、文字の変形の仕方によっては、特徴ベクトルが大きく変動してしまう場合もある。我々は、文字の潰れ、太さ変動、位置変動などの変形に対しては、文字線の内部情報よりも文字の輪郭線情報に着目した方がより有効な特徴ベクトルの抽出が行えるとして、外郭局所的輪郭線 (peripheral local outline vector) 特徴¹⁾ (以下、P-LOVE 特徴と呼ぶ) を考案した。

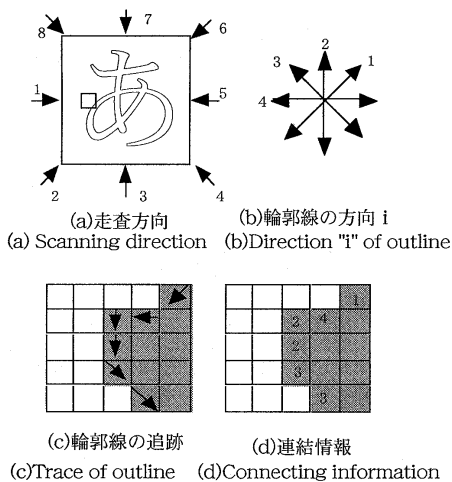


図 1 P-LOVE 特徴

Fig. 1 The feature extraction of P-LOVE.

次にその原理を示す。

2.1 P-LOVE 特徴の原理

- (1) 文字のすべての輪郭線を追跡し、輪郭線どうしがつながる連結方向 i ($i = 1, 2, 3, 4$) を保持しておく (図 1 (b), (c), (d))。
- (2) 文字を縦、横、斜めの 8 方向から走査して文字の外郭点を得る (図 1 (a))。
- (3) 得られた外郭点を中心にある大きさのウィンドウを開き、(1) で得ている連結情報を参照する (図 1 (d))。
- (4) ウィンドウ内に含まれる輪郭点の総和に対する、各連結方向成分の和を特徴とする。

つまり、 i 方向の P-LOVE 特徴 f_i は次式のようになる。

$$f_i = d_i / (d_1 + d_2 + d_3 + d_4) \quad (1)$$

d_i : i 方向の連結情報を持つ輪郭点の和

漢字では各走査方向で外郭深度を 3 までとれば文字線の 96% 以上を含むことが報告²⁾されているため、外郭深度は 3 までとし、得られた特徴量を各走査方向ごとに 8 区間に分割して平均して次元圧縮を行う。これにより、P-LOVE 特徴の次元数は、 $8 \times 3 \times 4 \times 8$ (走査方向 \times 外郭深度 \times 連結方向 \times 平均区間) = 768 次元となる。

2.2 P-LOVE 特徴の分割

単独の特徴抽出法で得られた特徴ベクトルの有効利用という観点で P-LOVE 特徴をとらえた場合、外郭深度に着目して特徴ベクトルを分割して使用することが考えられる。たとえば、全 768 次元の特徴ベクトル

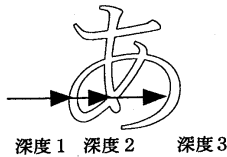


図2 外郭深度
Fig. 2 Peripheral depth.

を3つの外郭深度ごとに分割した場合、異なる外郭深度ごとで特徴を抽出したという意味を持つ3個の特徴ベクトルブロックができ、この場合の各外郭深度の特徴ベクトルは、 $768 \div 3 = 256$ 次元となる。

3. 類似文字の識別

対象カテゴリーの多い漢字認識では、識別に要する処理量を低減するために、大分類、中分類、詳細識別と、候補字種を徐々に絞り込んでいく階層的な認識システム¹²⁾を構成することが考えられる。この場合、詳細識別では類似文字集合が対象となる。したがって、詳細識別までの候補字種選択が正確に行われれば、最終的な識別率の向上のためには類似文字の識別が大きな鍵となる。

今回は、特に類似文字集合では深度ごとの外郭点が字種の差異を表すのに有効ではないかと考え、P-LOVE 特徴を外郭深度ごとに分割して使用することにした。

外郭深度は、文字を周囲8方向から走査して得られる最初の外郭点を外郭深度1、次に得られる外郭点を外郭深度2の画素というように定義している(図2)。図3に手書き漢字データベース ETL9B¹³⁾の見本文字を用いて、深度ごとに抽出した外郭点の例を示す。外郭深度が増すにつれて、文字の内部にある外郭点が選択されている様子が見てとれる。

4. 認識システム

NNの文字認識への応用は近年さかんに研究されている^{6)~8)}。NNを用いて文字の識別を行う場合、抽出した文字の特徴ベクトルを用いて学習を行い、識別では出力層で最大の反応値を示したカテゴリーを候補字種とする認識システム(図4)が考えられる。

一方、外郭深度に着目して、それぞれの深度ごとにNNを用いて識別し、さらにそれらの識別結果を統合処理するNNによる認識システムを図5に示す。図5の認識システムでは、たとえ1段目で3つの識別結果が一致しなかったとしても、それらの出力パターン規則を2段目のNNが学習することで、最終的に正認識

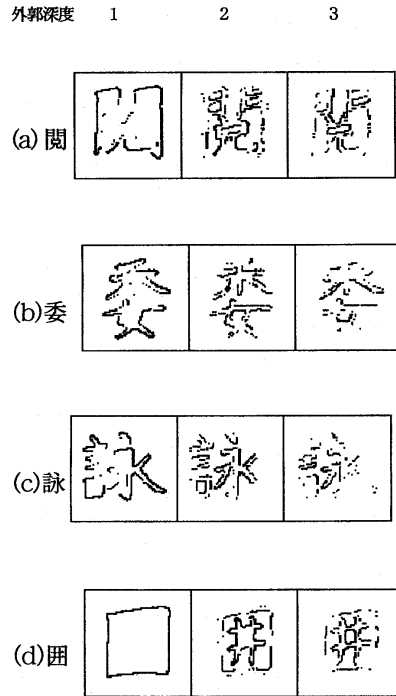


図3 深度ごとの外郭点
Fig. 3 The outline of each peripheral depth.

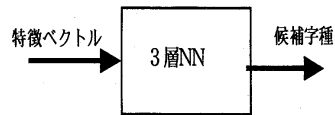


図4 NNによる認識システム
Fig. 4 The recognition system by using a neural network.

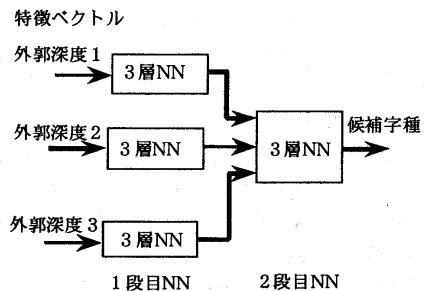


図5 外郭深度ごとの識別結果を統合処理する認識システム
Fig. 5 The unified neural network which utilized the features in peripheral depths.

表 1 類似文字集合
Table 1 The sets of similar characters.

A	闊園園開開開開開開開開開開開開開開開開
B	委萎嘉姦喜享孝妻姿妾姿姿姿姿姿姿姿姿姿姿
C	詠諫該諫諫諫諫諫諫諫諫諫諫諫諫諫諫諫諫
D	開因園回園固園困因凶因團團四田日目

が行えることが期待できる。このような統合処理する NN 認識システムでは、1 段目 NN の出力結果をどのような形で 2 段目 NN に入力するかについていくつかの方法が考えられるが、今回用いる方法については後述する。

5. 実 験

電子技術総合研究所提供の、手書き漢字データベース ETL9B (全 3036 字種) を外接枠が 64×64 画素になるように大きさの正規化を行って使用した。1 字種 200 サンプルのうち奇数番号を学習文字、偶数番号を未知文字とした。また、P-LOVE 特徴においては、文字の輪郭線を局所的に観測するのが有効であることを筆者らの一部¹⁾はすでに報告しており、これに基づき外郭点で開くウインドウの大きさは 3×3 画素とした。

5.1 類似文字集合の作成

字種ごとの学習文字の平均特徴を用いて、K-means 法によるクラスタリングを行った。恣意的にクラスタ数は 300 とし、その中からメンバ数が 16 の 14 クラスタについて内分散の小さい順に 3 クラスタを選択し類似文字集合とした。また、周囲が文字線で囲まれた「くがまえ」を持つ字種は、外郭深度 1 では識別が特に困難だと思われる。そこで、メンバ数が 12 であった「くがまえ」のクラスタに、それらの字種と同様に周囲が文字線で囲まれた 4 字種 (四田日目) を追加した類似文字集合を用意した。これらの類似文字集合を表 1 に示す。

5.2 1 段の NN による識別

図 4 に示した、3 層の NN 1 段により識別を行う。NN は類似文字集合ごとに用意する。NN の入力層の素子数を低減するために、P-LOVE 特徴 768 次元を主成分分析により次元圧縮³⁾した。圧縮後の次元数と累積寄与率との関係を図 6 に類似文字集合 A を例として示す。主成分数 200 で累積寄与率は 94% を超え、400 ではほとんど飽和している。この傾向は他の類似文字集合でも同様であった。これらの結果から、圧縮

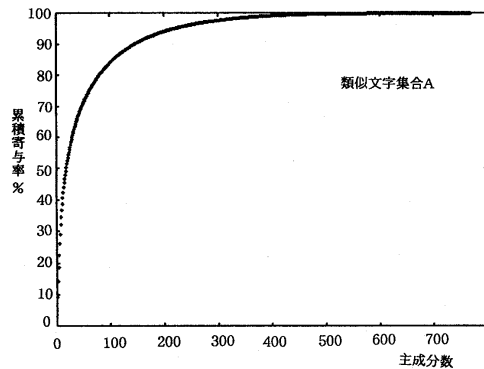


図 6 類似文字集合 A に対する圧縮後の主成分数とその累積寄与率
Fig. 6 The number of principal component and the accumulated contribution ratio after condensed for similar character set A.

後の次元数は各類似文字集合とも 400 次元とした。このときの累積寄与率はすべての類似文字集合とも約 99% である。出力層の素子は各字種と 1 対 1 に対応させ、識別では、最も大きな値を出力した素子で候補字種を決めた。NN の各層の素子数は、入力層 400、中間層 32、出力層 16 とした。中間層の素子数を決めるために類似文字集合 A を用いて識別実験を行った結果、3 万回の学習では未知文字に対する識別率が 16 素子で 84.94%、32 素子で 85.75% となり、さらに 48 素子に増すと識別率は 84.56% に低下した。また、中間層の素子数の増加による NN の学習時間の増加 (64 素子のとき 32 素子のときの 2 倍となる) を考慮し、ここでは 32 素子を採用した。各素子の荷重およびしきい値は乱数で初期化し、慣性モーメント項を考慮したバックプロパゲーションによる一括学習法を用いた。類似文字集合 A を用いた予備実験の結果、学習回数を 2 万回から 3 万回に増すと未知文字に対する識別率が 0.94% 向上したが、4 万回まで学習を続けても、識別率は 0.06% しか向上しなかった。したがって、ここでは各類似文字集合とも学習回数を 3 万回として実験を行った。

次に、P-LOVE 特徴 (全 768 次元) を 3 種の外郭深度ごとの 256 次元に分割後、それぞれを主成分分析によって次元圧縮した。特徴分割を行わない場合と比較するため、使用する特徴次元数の合計が 400 次元以下となるようにした。すなわち圧縮後の次元数は、各外郭深度ごとに 130 次元とし (130 次元 \times 3 外郭深度 = 390 次元)、図 4 に示した 3 層の NN 1 段により識別を行った。このときの累積寄与率は、いずれの類似文字集合でもすべて 97% を超えた。3 万回の学習を

表2 1段のNNでの識別率(%)

Table 2 Recognition rate of the single neural network.

類似文字集合	文字	分割なし	深度別分割
		400次元	390次元
A	学習	99.69	99.81
	未知	85.75	85.06
B	学習	99.94	99.94
	未知	95.38	94.81
C	学習	99.88	99.69
	未知	91.94	91.38
D	学習	100.00	99.94
	未知	94.13	93.44

行った後の識別結果を表2に示す。

5.3 統合処理するNNによる識別

図5に示した、外郭深度ごとの識別結果を統合処理するNNによって識別を行う。P-LOVE特徴全768次元を3種の外郭深度ごとの256次元に分割したものを主成分分析によって130次元に圧縮した。

1段目の各NNの素子数は、入力層130、中間層32、出力層16とした。2段目NNの入力層には1段目NNの出力層からのシグモイド関数の0~1のアナログ出力値をそのまま入力し、素子数は入力層48(16×3)、中間層32、出力層16とした。識別や学習の条件は前の実験と同様である。類似文字集合Aを用いた予備実験の結果、1段目のNNでは、各外郭深度とも学習回数が8000回を超えると未知文字の識別率が飽和あるいは低下してくる傾向があったので、1段目各NNの学習回数は8000回とした。また、2段目のNNでは、学習回数を1000回から5000回まで増すと未知文字に対する識別率が類似文字集合Aでは0.13%低下したが、その他の類似文字集合では識別率が最高で0.38%向上した。しかし、8000回まで学習を続けても、識別率は飽和あるいは低下する傾向があったので、2段目NNの学習回数は5000回とした。これらの識別結果を表3に示す。

また、外郭深度別に特徴ベクトルを分割したことの有効性を確かめるために、特徴ベクトルをランダムに3分割して、前の実験と同様の条件で統合処理するNNによって識別を行った結果を表4に示す。

6. 考察

図4に示したNN1段を用いた場合と、図5の統合NNを用いた場合を比較し、識別率の向上した理由やリジェクト処理などについて考察する。また、類似文字集合によっては、深度情報を選択することで統合NNによる認識システムをより効果的に構成できるこ

表3 外郭深度別NN, 統合NNでの識別率(%)

Table 3 Recognition rate of the unified neural network which utilized the features in peripheral depths (%).

類似文字集合	文字	深度1	深度2	深度3	統合
A	学習	97.00	99.44	98.81	99.94
	未知	60.94	83.38	73.31	88.06
B	学習	99.56	99.56	99.56	99.94
	未知	92.19	91.94	87.44	96.81
C	学習	99.56	99.63	99.13	100.00
	未知	88.19	86.63	82.13	94.63
D	学習	79.50	99.81	99.81	99.44
	未知	20.13	92.19	90.94	94.50

表4 特徴をランダムに3分割した、統合NNでの識別率(%)

Table 4 Recognition rate of the unified neural network which utilized the features in random (%).

類似文字集合	文字	識別率(%)
A	学習	100.00
	未知	69.19
B	学習	99.94
	未知	87.00
C	学習	99.94
	未知	82.75
D	学習	99.94
	未知	84.81

とを確認する。

6.1 統合処理するNNの識別率

1段のNNでは、特徴を外郭深度ごとに分割して使用すると、分割しない場合に比べて、未知文字の識別率(表2)が、0.56~0.69%低下した。しかし、外郭深度別の特徴ベクトルによる統合NNを用いた場合(表3)は、表2と比較すると、すべての類似文字集合で識別率が向上し、未知文字では、0.37~2.69%の向上が確認できた。また、特徴ベクトルをランダムに分割した統合NN(表4)では、外郭深度別の統合NNを用いた場合(表3)に見られたような識別率の向上は認められなかった。これらより、特徴ベクトルを外郭深度に着目して分割利用した統合NNが識別に有効だったと考えられる。

次に、統合型NNを用いた場合の、各段のNNの識別結果を調べたものを表5に示す。統合処理システムでは、たとえ1段目の識別で誤認識したNNの数が多いとき(表5の正認識:誤認識が、1:2および0:3)でも、2段目のNNによって最終的に正認識される場

表5 各段のNNの識別結果

Table 5 Recognition rate of the neural network in each stage.

文字	類似文字集合	A			B			C			D			
		1段目NN (深度1,2,3)	1段目	2段目 正認識数	識別率 (%)	1段目	2段目 正認識数	識別率 (%)	1段目	2段目 正認識数	識別率 (%)	1段目	2段目 正認識数	識別率 (%)
学習	3:0	1539	1539	100.00	1587	1587	100.00	1580	1580	100.00	1271	1271	100.00	
	2:1	15	15	100.00	6	6	100.00	12	12	100.00	43	43	100.00	
	1:2	4	4	100.00	0	0	—	1	1	100.00	37	37	100.00	
	0:3	42	41	97.62	7	6	85.71	7	7	100.00	249	248	99.60	
未知	3:0	858	850	99.07	1400	1396	99.71	1293	1287	99.54	297	294	98.99	
	2:1	270	245	90.74	111	98	88.29	157	138	87.90	381	378	99.21	
	1:2	195	145	74.36	53	39	73.59	86	60	69.77	240	229	95.42	
	0:3	277	169	61.01	36	16	44.44	64	29	45.31	682	611	89.59	

合が確認できる。たとえば、未知文字の識別において、類似文字集合Aでは、1段目NN3個のすべてが誤認識した場合でも、最終的には61.01%が正認識として救済されており、また類似文字集合Bでは、同様の場合44.44%、Cでは45.31%、Dでは89.59%が救済されている。

また、外郭深度1の特徴ベクトルだけでは、識別が特に困難だと考えられた周囲が文字線で囲まれている類似文字集合Dは、外郭深度1の情報を入力した1段目NNにおいて、学習文字で79.50%、未知文字で20.13%の識別率が得られている(表3)。

これは手書きの場合、筆者が「くがまえ」の内部の文字を考えながら、「くがまえ」を筆記するため、同一筆者によって書かれた文字でも、字種によって「くがまえ」の形状にある種の規則性が生じることと、今回類似文字集合として追加した4字種(四田日目)が、「くがまえ」の字種とは異なり、縦長あるいは横長の文字であったことが原因だと推測される。

6.2 リジェクト処理

表6に、統合型NNを用いた場合の、1段目3個の深度別NNの判定結果と、2段目1個の統合NNの識別結果の関係を示す。これより、1段目NNの判定がすべて一致したときには、いずれの類似文字集合においても最終的な識別率が高いことが分かる。そこで、1段目NNの判定結果、つまり第1候補として出力された字種がすべて一致したとき以外をリジェクトすることにすれば効果的なリジェクト処理が行えるのではないかと考えた。リジェクト率 R_e とその際の識別率 R_c を次式のように定義する。

$$R_e = (n - S1_S) / n \times 100 \quad (\%) \quad (2)$$

$$R_c = S2_C / S1_S \times 100 \quad (\%) \quad (3)$$

n : サンプル数

$S1_S$: 1段目NNの判定がすべて一致した個数
(リジェクトされずに残った個数)

$S2_C$: 2段目NNの正認識した個数

たとえば、表6(a)類似文字集合A・未知文字では、 $(1600 - 927) \div 1600 \times 100 = 42.06\%$ のリジェクト率で、識別率は、 $882 \div 927 = 95.15\%$ となる。同様にして計算したリジェクト率と識別率を表7に示す。

これらの結果より、類似文字集合B、Cの未知文字識別では、10%台のリジェクト率で、識別率をそれぞれ約2.6%および3.8%向上させられることが確認できた。リジェクトした文字は、他の識別方法によって正認識できれば、システム全体の識別率向上が期待できるが、今回は、リジェクト後の救済処理については言及しない。

また、類似文字集合A、Dの未知文字識別については、リジェクト処理を行うことで識別率をそれぞれ、約7.1%および0.7%向上することができるものの、リジェクト率はそれぞれ約42%および76%にも達している。これは、類似文字集合A、Dでは、統合型NNシステムの外郭深度別に識別を行う1段目NNにおいて、外郭深度1でのNNの識別率が他の深度を使用した場合に比べて大きく低下している(表3)ことが原因だと考えられる。この対策については、次に考察する。

6.3 外郭深度特徴の選択

表3で、類似文字集合Dにおいては、深度1での識

表 6 各段の NN の判定結果

Table 6 The decision of the neural network in each stage.

(a) 類似文字集合 A

(a) Set A

文字	1段目NN (深度1,2,3) a		2段目NNの 正認識個数b	b/a(%)
	判定	個数		
学習	全て一致	1543	1543	100.00
	2つ一致	39	38	97.44
	全て不一致	18	18	100.00
	計	1600	1599	99.94
未知	全て一致	927	882	95.15
	2つ一致	515	410	79.61
	全て不一致	158	117	74.05
	計	1600	1409	88.06

(c) 類似文字集合 C

(c) Set C

文字	1段目NN (深度1,2,3) a		2段目NNの 正認識個数b	b/a(%)
	判定	個数		
学習	全て一致	1580	1580	100.00
	2つ一致	17	17	100.00
	全て不一致	3	3	100.00
	計	1600	1600	100.00
未知	全て一致	1315	1294	98.40
	2つ一致	234	184	78.63
	全て不一致	51	36	70.59
	計	1600	1514	94.63

(b) 類似文字集合 B

(b) Set B

文字	1段目NN (深度1,2,3) a		2段目NNの 正認識個数b	b/a(%)
	判定	個数		
学習	全て一致	1587	1587	100.00
	2つ一致	10	10	100.00
	全て不一致	3	2	66.67
	計	1600	1599	99.94
未知	全て一致	1410	1402	99.43
	2つ一致	150	119	79.33
	全て不一致	40	28	70.00
	計	1600	1549	96.81

(d) 類似文字集合 D

(d) Set D

文字	1段目NN (深度1,2,3) a		2段目NNの 正認識個数b	b/a(%)
	判定	個数		
学習	全て一致	1301	1301	100.00
	2つ一致	178	177	99.44
	全て不一致	121	121	100.00
	計	1600	1599	99.94
未知	全て一致	377	359	95.23
	2つ一致	773	730	94.44
	全て不一致	450	423	94.00
	計	1600	1512	94.50

表 7 リジェクト率と識別率 (%) : 未知文字

Table 7 Rejection and recognition rates (%): unknown set.

類似文字集合	リジェクト率	識別率
A	42.06	95.15
B	11.88	99.43
C	17.81	98.40
D	76.44	95.23

表 8 外郭深度 2, 3 を用いた NN の識別結果 (%)

Table 8 Recognition rate of the neural network by using the 2nd and 3rd peripheral depths.

類似文字集合	文字	深度 2	深度 3	統合
D	学習	99.81	99.81	99.94
	未知	92.19	90.94	95.00

別率が他の深度の場合と比べて、学習文字で約 20%も低い値となっている。これは、類似文字集合 D は、周囲を文字線で囲まれている文字が集まっているため、外郭深度 1 では類似した特徴ベクトルが抽出されてしまうことが原因だと考えられる。そこで、類似文字集

合 D において、深度 1 を除外して、つまり 1 段目の NN は、深度 2 と深度 3 の 2 個の構成として統合処理する NN で識別を行った実験結果を表 8 に示す。この場合、未知文字で約 0.5% 識別率が向上した。また、同様に深度 1 の NN を除外した場合の各段の NN の

表9 類似文字集合Dで深度1を除外した場合の各段NNの判定結果

Table 9 Recognition rate of the set D in each stage excepting the 1st peripheral depth.

文字	1段目NN(深度2.3) a		2段目NNの 正認識個数b	b/a(%)
	出力	個数		
学習	一致	1595	1595	100.00
	不一致	5	4	80.00
	計	1600	1595	99.67
未知	一致	1480	1446	97.70
	不一致	120	74	61.67
	計	1600	1520	95.00

判定結果を表9に示す。表9において、1段目NNの判定が一致しなかったとき、つまり第1候補として出力された字種が一致しなかった場合をリジェクトすることにすれば、リジェクト率は7.5%、その場合の識別率は97.70%となる。これは、表7に示した3つの深度すべてを使用した場合と比べると、リジェクト率は約70%、その際の識別率は約2.5%改善されている。

このように、類似文字集合によっては識別に不利益な深度情報を除外することで、より効果的なリジェクト処理を行えると同時に、高い識別率を得られることが確認できた。今回は、類似文字集合1つだけを対象に外郭深度特徴の選択を行ったが、他の類似文字集合にも同様の効果が期待できる。

7. む す び

先に提案したP-LOVE特徴を外郭深度別に3分割し、特徴ベクトルごとに個別のニューラルネットワーク(NN)により識別を行い、さらにそれらの結果を統合処理するNNシステムを構成した。この認識システムで類似文字集合の識別実験を行った結果、外郭深度を分割せずに一括して1つのNNで識別するよりも高い識別率を得られることを確認した。統合処理するNNシステムは、図4に示した1段のNNによるシステムと比べ、個々のNNの規模が小さくて済む。また、複数のCPUが使用できる環境があれば、複数のNNを並列して学習させることが可能となるため、短い時間でシステムを構築できる利点がある。

また、今回提案した外郭深度に着目した統合処理システムでは、たとえ1段目の識別で誤認識したNNの数が多い場合でも、2段目のNNによって最終的に正認識される場合があることを確認した。これは、たとえ1段目のNNが誤認識をしている場合であっても、1段目のNNの各出力値パターンを2段目のNNが与えられた教師信号を用いて学習したことが有効だった

と考えられる。

次に、1段目のNNの認識結果を多数決判定してリジェクト処理を行えば、識別率を向上させることができることを確認した。

さらに、類似文字集合によっては識別に不利益な深度情報を除外すれば、認識システムを簡素化できると同時に、より効果的なリジェクト処理が行え、かつ高い識別率を得られることを確認した。

次元圧縮時に用いる主成分数、NNの中間層の素子数、学習回数などの最適数については、一般には議論の対象である。本論文では、外郭深度に着目した統合型NNの効果を確認することが主な目的であるため、すべての類似文字集合において同一の条件で実験を行った。しかし、各類似文字集合ごとにより適切な条件を見つけて実験を行えば、さらに高い識別率が得られる可能性が高いと考えられる。

異なった構造情報を有効に抽出する複数の高性能な特徴抽出法の開発が容易でない現状では、今回提案したように単独の特徴抽出法で得られた特徴ベクトルを効果的に利用し、識別率の向上を実現する方法も興味深い。

今回は、4つの類似文字集合を作成し、統合型NNシステムの効果を確認したが、今後はさらに類似文字集合の数を増やして実験を行う予定である。また、特徴ベクトルの分割の仕方¹⁴⁾についても検討を行いたい。

さらに、リジェクトした文字の具体的な救済方法について検討し、システム全体の識別率の向上を図りたい。

参 考 文 献

- 堀桂太郎, 根本孝一, 伊藤彰義: 文字の輪郭線に着目した手書き漢字の特徴抽出法, 信学論(D-II), Vol.J82-D-II, No.2, pp.188-195 (1999).
- 萩田紀博, 内藤誠一郎, 増田 功: 外郭方向寄与度特徴による手書き漢字の識別, 信学論(D), Vol.J66-D, No.10, pp.1185-1192 (1983).
- 堀桂太郎, 菅原広喜, 伊藤彰義: 外郭深度に着目した統合型ニューラルネットワークによる類似文字認識, 信学技報, PRMU98-145 (Dec. 1998).
- 島田恭宏, 大倉 充, 塩野 充, 橋本禮治: 複数特徴部分空間法による手書き類似漢字識別, 信学論(D-II), Vol.J78-D-II, No.10, pp.1460-1468 (1995).
- 西川修史, 若林哲史, 木村文隆, 三宅康二, 堤田俊夫: 手書き数字認識における特徴量の合成, 信学技報, PRU95-114 (Sep. 1995).
- 山田敬嗣, 上 博行, 溝口正典, 天満 満: ニューラルネットを用いた文字認識, 信学技報, PRU88-58 (1988).

- 7) 森 吉弘：大規模ニューラルネットの構築に向けて，信学技報，PRU88-59 (1988).
- 8) 木村義政，近藤利夫，曾根原登：複数手法における候補カテゴリーの識別値を統合処理する神経回路による手書き漢字認識，信学論 (D-II)，Vol.J77-D-II，No.4，pp.791-800 (1994).
- 9) 孫 方，大町真一郎，加藤 寧，阿曾弘具：特徴量の要素の相関を考慮した高速・高精度な識別閾数と文字認識への応用，信学論 (D-II)，Vol.J81-D-II，No.9，pp.2027-2034 (1998).
- 10) 孫 寧，安部正人，根元義章：改良型方向線素特徴および部分空間法を用いた高精度な手書き漢字認識システム，信学論 (D-II)，Vol.J78-D-II，No.6，pp.922-930 (1995).
- 11) 荒井正之，奥田健三，渡辺博芳，宮道壽一：手書き文字認識における個人辞書の作成法，信学論 (D-II)，Vol.J80-D-II，No.7，pp.2023-2026 (1997).
- 12) 伊藤彰義，遠藤 武，堀桂太郎，島村 徹：階層的印刷漢字認識システムにおける字種を複数クラスタに登録する辞書構成法，信学論 (D-II)，Vol.J78-D-II，No.6，pp.896-905 (1995).
- 13) 斉藤泰一，山田博三，山本和彦：JIS 第 1 水準手書き漢字データベース ETL9 とその解析，信学論 (D)，Vol.J68-D，No.4，pp.757-764 (1985).
- 14) 堀桂太郎，菅原広喜，伊藤彰義：走査方向別の特徴を用いた統合型ニューラルネットワークによる類似文字認識，信学総合大会，D-12-30 (Mar. 1999).

(平成 11 年 4 月 30 日受付)

(平成 11 年 9 月 2 日採録)



堀 桂太郎 (学生会員)

昭和 59 年千葉工業大学工学部電子工学科卒業。平成 3 年日本大学大学院博士前期課程修了。現在，同大学院博士後期課程在学中 (社会人学生)。文字認識の研究に従事。東京都立蔵前工業高等学校電気科教諭。



菅原 広喜

平成 10 年日本大学理工学部電子工学科卒業。現在，同大学院博士前期課程在学中。文字認識の研究に従事。



伊藤 彰義

昭和 41 年日本大学理工学部電気工学科卒業。昭和 43 年同大学院修士課程修了。昭和 46 年同大学院博士課程修了。同年同大学理工学部助手。昭和 62~63 年 CMU 客員助教授。平成元年日本大学電子工学科教授。平成元年~5 年日本応用磁気学会理事。平成 7~9 年応用物理学会常務理事。日本応用磁気学会論文賞受賞。矢崎学術賞受賞。文字認識・3D レーダ画像処理・光磁気記録・薄膜の研究に従事。工学博士。