異なる特徴抽出法を併用した手書き類似文字の識別

堀 桂太郎*

Techniques of Similar Handwritten Characters Recognition by Using Different Feature Vectors

Keitaro HORI

ABSTRACT

To extract unique feature vectors of handwritten characters and thereby identify them efficiently, we now propose and make available a means to use feature vectors known as P-DC and P-LOVE. These feature vectors tend to be used separately in general systems. However, we sought to achieve a more effective way of recognition by simultaneously using both sets of unique feature vectors. After some experiments using an integral type of neural network (NN), we found that we were able to truly accomplish a better way of recognition if we used both feature vectors together rather than using each of them separately. Based on the judgment obtained from such a recognition system using an integral NN, we also found that it became possible to process rejections in an effective manner.

KEW WORDS: handwritten characters, feature vector, P-DC, P-LOVE, neural network

1.はじめに

アルファベットと比べて対象字種数が非常に多 い漢字の自動認識では , 徐々に候補字種を絞り込ん でいく階層的な認識システム1)を使用するのが一般 的である.この場合,最終段階においては,類似し た文字集合内での識別が対象となる. つまり, 類似 文字の認識精度を向上させることは,システム全体 の性能向上には欠かせない要素となる.一方,手書 き文字の認識に有効な特徴抽出法として,外郭方向 寄与度特徴²⁾や外郭局所的輪郭線特徴³⁾が提案され ている.外郭方向寄与度特徴は文字線内部の情報を 抽出するのに対して,外郭局所的輪郭線特徴は文字 の輪郭線から情報を抽出する.これらの特徴は,単 独で使用されるのが一般的である.しかし,着目点 の異なる両特徴を併用することで,より効果的な識 別が行えるのではないかと考えた、そこで、識別関 数にニューラルネットワーク(以下, NN と呼ぶ) を用いた場合の両特徴の併用方法を検討した.実験 の結果,統合型のNN用いて併用した場合に,最も

高い識別率を得られることが確認できた.本論では, 異なる2種の特徴抽出法を効果的に併用することで, 各特徴を単独で使用した場合を上回る識別率が得 られることを報告する.また,統合型 NN 認識シス テムの判定結果に着目することで,有効なリジェク ト処理が行えることについても言及する.

2.特徵抽出法

2 · 1 P-DC 特徵

外郭方向寄与度(peripheral direction contributivity)特徴(以下,P-DC 特徴と呼ぶ)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であり広く使われている.P-DC 特徴は,文字の各黒画素に与えられる方向寄与度を抽出して,それを特徴ベクトルとするものである.例えば,図 1(a)に示す文字線内の黒画素点 P_1 の方向寄与度 d_{P_1} を次のように表す.

 $d_{P1} = (d_{1P1}, d_{2P1}, d_{3P1}, d_{4P1})$ 各要素 $d_{mp1} (m = 1,2,3,4)$ は点 P_1 から図 1(b) に示す8方向に触手を伸ばして求まる黒点連結長 $l_i(i)$

^{*} 電気情報工学科

= 1,2,3,・・・,8) を用いて,式(1)のように定義さ れる.

$$d_{mp1} = \frac{l_{m} + l_{m+4}}{\sqrt{\int_{j=1}^{4} (l_{j} + l_{j+4})^{2}}}$$
(1)

図1(a)の黒点 P₁における方向寄与度 d₂1の実測 値は, d P 1 = (0.24, 0.30, 0.87, 0.30)となり, 点P₁は垂直方向の度合いの強い文字線上にあるこ とがわかる、このように方向寄与度は文字線の方向 と接続関係をベクトルの形で表すことができる。方 向寄与度を抽出する点は,文字を外部から 45 度お きに8方向から走査し,横切る各文字線の外郭点と する.また,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 94%以上を含むことを実験で確 認しているため,外郭深度は3までとし,得られた 特徴量を走査方向毎で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-DC特徴の次元数は, 8×3×4×8(走查方向×外郭深度×方向寄与度×平 均区間)=768次元となる.

2·2 P-LOVE 特徵

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.このため,筆者 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,外郭局所的輪郭線(peripheral local outline vector)特徴(以下, P-LOVE 特徴と呼ぶ)を考案し た.次にその原理を示す.

- (1) 文字の全ての輪郭線を追跡し,輪郭線どうし が繋がる連結方向i(i=1,2,3,4)を保持して おく(図2(a),(b),(c)).
- (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2(d)).

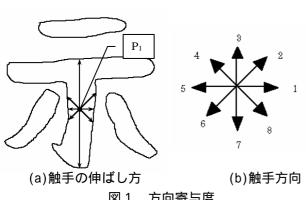


図 1 方向寄与度

- (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2(c),(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 は式(2)のようにな

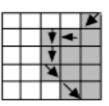
$$f_{i} = d_{i} / (d_{1} + d_{2} + d_{3} + d_{4})$$
 (2)
 $d_{i} : i 方向の連結情報を持つ輪郭点の和$

得られた特徴量を走査方向毎で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,8×3×4×8(走査方向×外郭深度× 連結方向×平均区間)=768次元となる.

3.実験

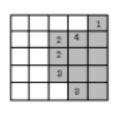
産業技術総合研究所提供の手書き漢字データベ ース ETL-9B⁴ 全 3036 字種の各字種 200 サンプルの うち奇数番目を学習文字,偶数番目を未知文字とし て使用する.そして, k-means 法によりクラスタリ ングを行い表1に示す類似文字集合A,B,Cを作成し た. 各類似文字集合のメンバ数は,16である.

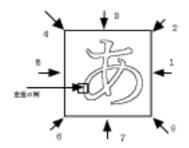
両特徴とも,特徴次元数を768とし,さらに主成 分分析によって次元圧縮した.いずれの類似文字集 合においても,両特徴とも主成分数を400まで用い ると累積寄与率が 98%を超えたため,圧縮後の次元 数は 400 を基準とした. 識別には 3 層から成る NN を使用し、図3に示す4種類の実験を行った.表2 に,類似文字集合 A の場合の累積寄与率を示す.



(a)輪郭線の追跡

(b)輪郭線の方向 i





(c)連結方向

(d)走查方向

図 2 P-LOVE 特徴の原理

表 1 類似文字集合

			_	•					•				
集合	メンバ												
Α	闇	閏	閲	開	閣	間	閑	関	閤	閃	闘	閥	聞
	閉	問	悶										
В	委	萎	嘉	姦	喜	享	孝	妻	姿	妾	凄	妥	毒
	婆	妄	婁										
С	詠	課	該	諌	議	訣	謙	誤	護	譲	設	談	諜
	謬	謀	訳										

表 2 類似文字集合 A の累積寄与率(%)

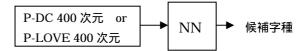
実験					
	P-DC	P-LOVE	P-DC	P-LOVE	
次元数	400	400	200	200	400
累積寄与率	99.30	99.27	94.09	94.16	97.07

表3 未知文字の識別率(%)

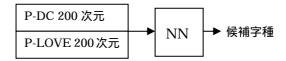
実験			_	_	
	P-DC	P-LOVE			
A	84.88	85.75	73.56	86.31	90.94
В	94.62	95.38	93.25	94.88	97.12
C	91.38	91.94	90.25	92.62	95.12

- ・実験 : 各特徴の 400 次元を使用して ,個別の NN で識別する .
- ・実験 : 各特徴を個別に 200 次元に圧縮した後, 合わせて 400 次元を使用し NN で識別する.
- ・実験 : 両特徴を合わせて(768+768=1536 次元) 圧縮した 400 次元を使用し NN で識別する.
- ・実験 : 各特徴の 200 次元, 計 400 次元を使用して統合型 NN で識別する.

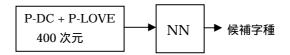
実験 ~ に使用した NN の各層の素子数は,入 力層 400, 中間層 32, 出力層 16 とした. 中間層の 素子数を決めるために類似文字集合 A に P-LOVE 特 徴を用いた予備実験を行った結果,3万回の学習で は未知文字に対する識別率が 16 素子で 84.94%とな リ,32 素子で85.57%となり,さらに48 素子に増や すと84.56%に低下した.また,中間層の素子数の増 加による NN の学習時間(64 素子のとき,32 素子の 2 倍となる)の増加を考慮して,ここでは中間層を 32 素子とした.また,各素子の加重及び,しきい値 は乱数で初期化し,慣性モーメント項を考慮したバ ックプロパゲーションによる一括学習法を用いた. 学習回数については、2万回から3万回に増すと識 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 とした.



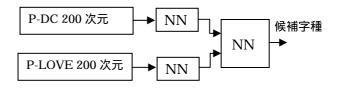
(a)実験 :特徴を単独で使用(各々400次元)



(b)実験 : 個別に圧縮して計 400 次元使用



(C)実験 : 合わせて圧縮して 400 次元使用



(d)実験 : 各々200 次元を統合型 NN で識別 図 3 NN による識別実験

実験 に使用した 1 段目 NN の各層の素子数は, 入力層 200,中間層 32,出力層 16 とした.2 段目 NN は,入力層 32,中間層 32,出力層 16 とした. また,学習回数は,1 段目 NN が 6000 回,2 段目 NN が 2000 回とした.

これらの学習回数は,学習回数と識別率を測定する予備実験を行った結果から定めたものである.表3に,識別実験を行った結果を示す.

4. 結果と考察

各特徴を単独で用いた実験 では、全ての類似文字集合において P-LOVE 特徴の方が高い識別率を得ている。両特徴を主成分分析によって個別に圧縮した実験 では、いずれの類似文字集合とも識別率が低下している。つまり、実験 の特徴使用法では、各々の特徴の異なる利点を活かすことができていないと考えられる。一方、両特徴を合わせてから圧縮した実験 では、類似文字集合 A,C においてP-LOVE 特徴単独使用時よりも高い識別率が得られている。さらに、統合型 NN による識別を行う実験では、全ての類似文字集合において、実験 では、全ての類似文字集合において、実験 の最大識別率よりも、さらに 1.74~4.63%の向上が確認できた。

類似文字集合	A			В			С		
1 段目 NN	1 段目	2 段目の	識別率	1 段目	2 段目の	識別率	1 段目	2 段目の	識別率
正認識:誤認識	の数	正解数	(%)	の数	正解数	(%)	の数	正解数	(%)
2:0	1339	1330	99.33	1494	1492	99.90	1436	1431	99.65
1:1	127	84	66.14	69	51	73.91	98	72	73.47
0:2	134	41	30.60	37	11	29.73	66	19	28.79

表 4 1 段目 NN の判定結果と識別率の関係

表 5 1 段目 NN の認識結果が一致した場合

類似文字集合	A	В	C	
1段目の数	1473	1531	1502	
2段目の正解数	1371	1503	1450	
識別率(%)	93.01	98.17	96.54	
リジェクト率(%)	7.94	4.31	6.13	

表4に、実験 における1段目 NNの判定結果と2段目 NNの識別率の関係を示す.この表からは、例えば、類似文字集合Aの場合、1段目 NNにおいて2つの NNともが正認識をした未知文字は全1600個中の1339個であり、2段目 NNでは1339個中の1330個が最終的に正認識されたことを示している.また、この際の識別率は、1330÷1339=0.9933=99.33%である.この実験結果から、いずれの類似文字集合においても、1段目 NNの両方が正認識をした場合に、最終的な識別率は99%以上になっていることがわかる.一方で、1段目 NNの両方が誤認識をした場合であっても、最終的な識別率が30%程度得られている.これは、誤認識をした1段目 NNの各出力値パターンを2段目 NNが学習したことが有効であったと考えられる.

正認識であれ誤認識であれ、1 段目 NN の認識結果が一致した場合の識別率などを計算すると表 5 のようになり、いずれも表 3 (実験)の識別率を上回る結果が得られることがわかる. 但し、リジェクト率 R_j は、未知文字のサンプル数(1600個)を samp、1 段目 NN の認識結果が一致した場合の数を agr として、式(3)のように定義した.

$$R_{j} = \frac{\text{samp-agr}}{\text{samp}} \times 100 \text{ (\%)}$$

表 5 より, 1 段目 NN の認識結果が一致しなかった場合を, リジェクトと判定することにすれば,

6.13~7.94%のリジェクト率で,高い識別率を得ることができる.

5.まとめ

識別が困難な手書き類似文字集合を対象として, 異なる特徴を併用して高い識別率を得る方法を検 討した.各種実験の結果、統合型の NN を用いれば, 特徴を単独で使用する場合よりも高い識別率が得 られることを確認した.この統合型 NN 認識システムでは,たとえ1段目 NN が誤認識をした場合であっても,その出力パターンを2段目 NN が学習することで,最終的に正認識が行われる場合があり,これが識別率の向上に寄与していると考えられる.また,1段目の判定結果を利用することで,リジェクトを行い,低いリジェクト率で,より高い識別率を得られることを確認した.今後は,リジェクトされた文字の救済処理を検討し,最終的な識別率を向上させる手法を考案する予定である.

参考文献

- 1) 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).
- 2) 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).
- 3) 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).
- 4) 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).