

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.

KEY 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、識別回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、有効なりジェクト処理が行えることについても言及する。

2. 特徴抽出法

2.1 P-DC 特徴

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。

$$d_{P_1} = (d_{1P_1}, d_{2P_1}, d_{3P_1}, d_{4P_1})$$
各要素 d_{mP_1} ($m = 1, 2, 3, 4$) 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.

$$d_{m+1} = \frac{l_m + l_{m+4}}{\sqrt{\sum_{j=1}^4 (l_j + l_{j+4})^2}} \quad (1)$$

図1(a)の黒点 P_1 における方向寄与度 d_{P_1} の実測値は, $d_{P_1} = (0.24, 0.30, 0.87, 0.30)$ となり, 点 P_1 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角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, $8 \times 3 \times 4 \times 8$ (走査方向 \times 外郭深度 \times 方向寄与度 \times 平均区間) = 768 次元となる.

2.2 P-LOVE 特徴

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)).

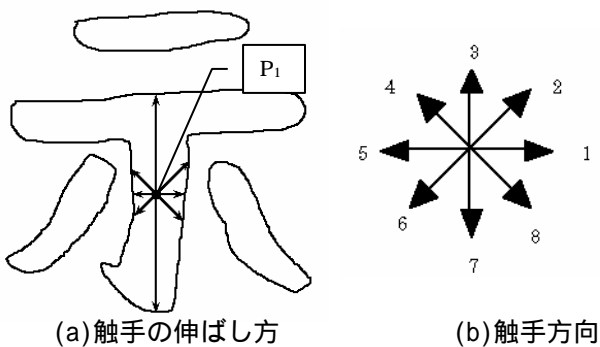


図1 方向寄与度

- (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線の総和に対する, 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)のようになる.

$$f_i = d_i / (d_1 + d_2 + d_3 + d_4) \quad (2)$$

d_i : i 方向の連結情報を持つ輪郭線の和

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, $8 \times 3 \times 4 \times 8$ (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.

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.

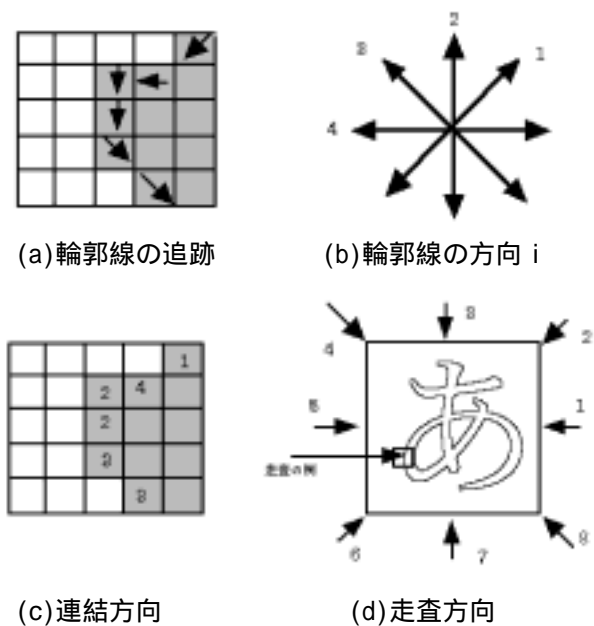


図2 P-LOVE 特徴の原理

表4 1 段目 NN の判定結果と識別率の関係

類似文字集合	A			B			C		
	1 段目の数	2 段目の正解数	識別率 (%)	1 段目の数	2 段目の正解数	識別率 (%)	1 段目の数	2 段目の正解数	識別率 (%)
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5 1 段目 NN の認識結果が一致した場合

類似文字集合	A	B	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、 $1330 \div 1339 = 0.9933 = 99.33\%$ 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。

$$R_j = \frac{\text{samp} - \text{agr}}{\text{samp}} \times 100 (\%) \quad (3)$$

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ー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).