

OCR手法を用いたオンライン手書き文字認識の改良

2L-5

川又 武典 依田 文夫
三菱電機(株) 情報技術総合研究所

1. はじめに

PDA, パソコンコンピュータ等の携帯情報端末における入力手段の1つとして、オンライン文字認識方式の研究が進められている。従来のオンライン文字認識方式は、構造解析的な手法によるものが主流であったが、近年OCR（オフライン）手法を融合した認識方式の研究が行われており^{[1][2]}、我々も時系列情報を利用したオンライン特徴を抽出し、OCR手法を用いて認識する方式^[3]の検討を行った。今回は、前記方式におけるオンライン特徴の評価を行い、その結果を基に特徴の改良を行った。

2. オンライン特徴の評価

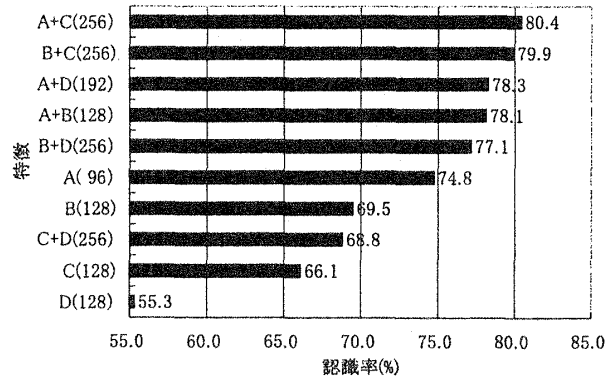
表1に示す学習・評価データを用いて、線分内の特徴である、8方向コード特徴、累積方向コード特徴、線分間の特徴である、相対位置特徴、相対線密度特徴の相補性評価及び分布特徴抽出時の領域分割に関する評価を行った。

表1 学習・評価データ

	データベース名称	データ数
学習	当社オンライン設計データ	250パターン/文字
評価	HANDS_kuchibue_d-97-06[4]	120人分

2.1 特徴間の相補性評価

単独の特徴における認識率と2つの特徴を組み合わせた場合の認識率から、特徴間の相補性評価を行った。なお、分布特徴を求める際の領域分割は、4×4（16領域）とし、正準判別分析法により抽出した圧縮特徴のうち、単独特徴の場合は、64, 96, 128次元、2つの特徴を組み合わせた場合は、128, 196, 256次元の圧縮特徴を用いて識別を行い、最も高い認識率となる次元数の結果（図1中、括弧内が次元数）により比較した。また、識別は入力パターンの圧縮特徴と標準パターンの圧縮特徴とのユークリッド距離を求めることにより行った。



A: 8方向コード特徴, B: 累積方向コード特徴
C: 相対位置特徴, D: 相対線密度特徴

図1. 認識率（相補性評価）

図1に示すように、線分内の特徴と線分間の特徴を組み合わせた場合に認識率の向上効果が高く、線分内の特徴と線分間の特徴には比較的大きい相補性があることがわかる。また、線分内の特徴は単独でも比較的高い性能を示しているが、線分間の特徴、特に相対線密度特徴は単独では性能が十分でないことが判る。

2.2 領域分割数の評価

分布特徴抽出時の領域分割を、4×4, 6×6, 8×8にした場合の認識率の評価を行った。なお、識別に用いた特徴次元数は、それぞれ、128, 256, 256次元とした。

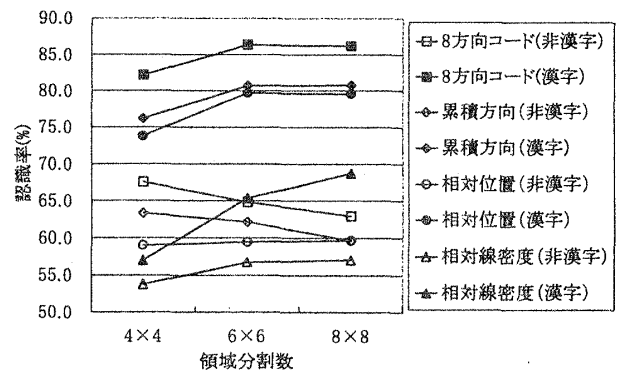


図2. 認識率（領域分割数評価）

図2に示すように、非漢字においては、8方向コード特徴及び累積方向特徴は、領域分割を細かくすると認識率が低下し、相対位置特徴及び相対線密度特徴は、逆に向上することが判る。また、漢字の場合は、相対線密度特徴以外は6×6領域分割で最高の性能になることが判った。

3. オンライン特徴の改良

3.1 領域分割数の変更

相対位置特徴及び相対線密度特徴に関して、分布特徴抽出時の領域分割数を4×4から6×6に変更した。

3.2 新規特徴の追加

相対位置特徴及び相対線密度特徴に対して、以下の改良を行い、線分間の新たな特徴として追加した。

(1) 抽出点における運筆方向の反映

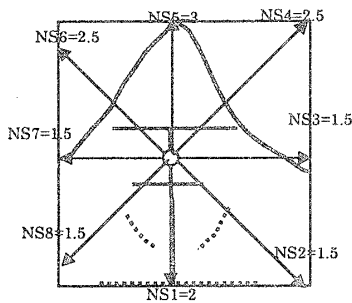
抽出点における運筆方向を基準とし(図3では抽出点の運筆方向が下向きなので、その方向をNS1とする)、探索方向を決定することにより、運筆方向を反映すると共に、文字の傾きに強い特徴を抽出する。

(2) 筆順(線分間の時系列情報)の反映

相対特徴抽出時の探索対象の線分を、抽出点における線分より時間的に前に筆記された線分のみ(図3の実線線分)に限定することにより、線分間の筆順を反映させた。

(3) 線分間の角度の反映

相対線密度抽出時に探索方向と交差線分との角度を考慮した線分数(交差角度90度の場合は線分数1、45度の場合は線分数0.5、0度の場合は線分数0)を抽出することにより、線分間の角度関係を反映させた。



$$FSD_j = \frac{NS_j}{\sum_{n=1}^8 NS_n}$$

$j = 1, 2, \dots, 8$

図3. 改良相対線密度特徴の例(文字「金」)

4. 改良方式の評価結果

改良方式について、表1に示す学習・評価データを用いて評価した結果を示す。

(1) 次元数を变化させた場合の認識率、分類率

表2に次元数別の認識率、分類率を示す。

表2. 次元数別の認識率、分類率(%)

次元数	漢字		非漢字		全体	
	第1位	第10位	第1位	第10位	第1位	第10位
256	94.04	99.19	76.26	95.44	84.75	97.23
320	94.18	99.22	76.36	95.49	84.87	97.27
384	94.25	99.23	76.40	95.53	84.93	97.30
448	94.32	99.25	76.46	95.56	84.99	97.32

表2に示すように、次元数を増加させるほど、認識率、分類率共に向上するが、320次元を超えると向上の割合は小さくなる。

(2) 改良前との比較

表3に次元数320における、改良後と改良前の認識率、第10位分類率の比較結果を示す。

表3. 改良前との比較(%)

方式	漢字		非漢字		全体	
	第1位	第10位	第1位	第10位	第1位	第10位
改良前	93.15	99.17	75.25	95.39	83.80	97.20
改良後	94.18	99.22	76.36	95.49	84.87	97.27

上記結果より、改良前に比べて、認識率、第10位分類率ともに向上することが判った。

(3) オンライン方式との比較

改良方式(320次元)と当社オンライン方式^[5]との認識率及び分類率の比較結果を表4に示す。比較は、評価データの内、比較的品質の良いmdb0055、mdb0066データと、筆順違い・続け書きの多いmdb0051データについて行った。

表4. オンライン方式との比較

データ	方式	漢字		非漢字	
		第1位	第10位	第1位	第10位
mdb0051	本方式	91.8	99.3	67.8	90.7
	オンライン	88.1	96.6	69.8	87.7
mdb0055	本方式	96.9	99.9	86.9	98.3
	オンライン	96.7	99.0	87.9	96.9
mdb0066	本方式	96.8	99.7	84.2	98.7
	オンライン	96.3	98.5	88.2	96.5

表4に示すように、漢字においてはすべてのデータでオンライン方式を越える性能を示しているが、非漢字の認識率は、オンライン方式を下回っている。

5. まとめと今後の予定

オンライン特徴の評価・改良を行い、改良特徴が認識率向上に有効なことを確認した。但し、非漢字における認識率は、従来のオンライン認識方式に比べ低い。したがって、今後は、改良方式におけるエラー解析を行うと共に、非漢字における認識率向上のために、変形、位置ずれに強い特徴の検討を行う予定である。

参考文献

[1] 岡本他: "Direction...", ICPR'98, vol. 2, pp. 1747-1751
 [2] 田中他: "オンライン認識とオフライン認識...", PRMU98-140, pp. 31-38, (1998. 12)
 [3] 川又他: "OCR手法を用いた...", 第58回情報学全大 vol. 4D-2, pp. 21-22, 1999
 [4] 中川他: "文章形式...", PRU95-110, pp. 43-48 (1995. 9)
 [5] 亀代他: "方向コード特徴と...", 信学総大D-12-90 (1997)