

## OCR 手法を用いたオンライン手書き文字認識

4D-2

川又 武典      岡野 祐一      平野 敬      依田 文夫  
 三菱電機(株) 情報技術総合研究所

### 1. はじめに

PDA, ペ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、我々も大分類に OCR 手法を取り入れた認識方式を日本語及び中国語向けに開発した [1][2]。しかし、単純に時系列情報を 2次元のパターン情報に変換し、OCR 手法を適用する方式では、識別に十分な特徴が得られないため、オンライン方式に比べ認識精度が劣り、両者を融合することによる効果が十分でないという課題があった。そこで、特徴空間レベルでオフライン特徴とオンライン特徴の融合を図る方式 [3] の検討が行われており、今回我々も時系列情報及び線分情報を利用したオンライン特徴を抽出し、OCR 手法を適用する認識方式の検討を行った。

### 2. 特徴抽出及び識別

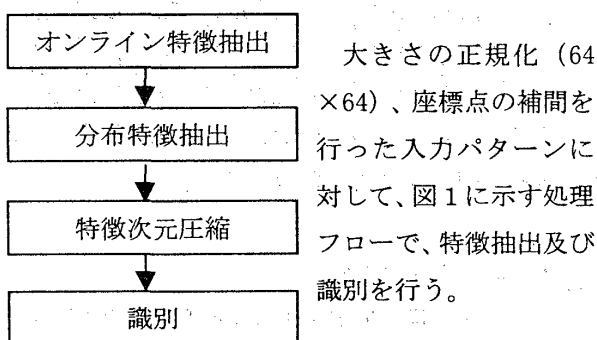


図 1. 処理フロー

#### 2. 1 オンライン特徴抽出

時系列情報及び線分情報を用いて以下のオンライン特徴を抽出する。

#### (1) 方向コード特徴

ペンの運筆方向を反映させるため、サンプル点における運筆方向を 8 方向に量子化した特徴を抽出する。「ソ」と「ン」のように形が似ているが、線分の運筆方向が異なるパターンの識別が可能になる。

#### (2) 始点、終点、交点特徴

線分における特徴点情報を反映させるため、始点、終点及び交点を特徴点として、位置及び点数を各特徴点毎に抽出する。

#### (3) 累積方向コード特徴

線分内の接続関係を反映させるため、抽出対象のサンプル点が属する線分の始点から、サンプル点までの方向コードの累積数を、線分全体における累積方向コードで正規化した特徴 (FC) を抽出する。

$$FC_{jm} = \sum_{k=1}^m D_{jk} / \sqrt{\sum_{n=1}^M \left( \sum_{k=1}^M D_{nk} \right)^2}$$

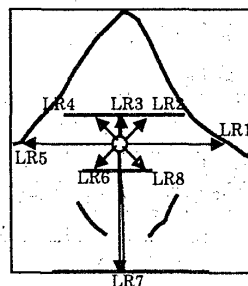
j: 方向 (1, 2, ..., 8), m: 線分上の抽出点 (1, 2, ..., M)

M: 抽出点が属する線分のサンプル点数

D<sub>jk</sub>: 抽出点 k における方向 j の方向コード

#### (4) 線分間の相対位置特徴

線分間の相対的な位置関係を反映させるため、各サンプル点から、8 方向に探索し、最初に他の線分あるいは外接矩形と交差するまでの長さ (LR<sub>j</sub>) を、8 方向のすべての長さで正規化した特徴 (FR) を抽出する。



$$FR_j = \frac{LR_j}{\sqrt{\sum_{n=1}^8 LR_n^2}}$$

j = 1, 2, ..., 8

図 2. 線分間の相対位置特徴

#### (5) 線分間の相対線密度特徴

複数の線分との関係を反映させるため、各サンプル点から、8 方向に探索し、外接矩形に到達するまでに横切る線分数 + 1 (NS<sub>j</sub>) を 8 方向の総線分数で正規

化した特徴(FSD)を抽出する。

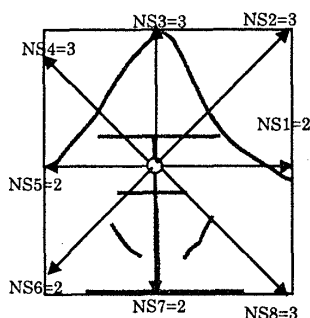


図3. 線分間の相対線密度特徴

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

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

### 2. 2 分布特徴抽出

文字の外接矩形を4×4の小領域に分割し、小領域内の各特徴値を統合することにより、35×4×4=560次元の分布特徴を得る。

### 2. 3 特徴次元圧縮

続け字、崩し字等による特徴変動の影響を除去するため、識別に有効な特徴(圧縮特徴)を抽出する。圧縮特徴の抽出には、正準判別分析法を用いた。

### 2. 4 識別

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、識別を行う。

## 3. 認識実験

表1に示す、学習・評価データを用いて辞書設計、認識実験を行った。なお、辞書設計に用いた認識対象文字は、非漢字348文字、漢字3693文字(第1水準漢字2965文字+第2水準漢字728文字)の合計4041文字、1テンプレート/文字である。但し、評価3のデータに含まれる認識対象外の非漢字32文字は除いた。

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1	当社オンライン評価データ(楷書体)	100パターン/文字
評価2	当社オンライン評価データ(自由筆記)	150パターン/文字
評価3	HANDS_kuchibue_d-97-06[4]	120人分

#### (1) 次元数を変化させた場合の分類率

表2に圧縮特徴の次元数を変化させた場合の評価3のデータにおける分類率を示す。

表2. 次元数別の分類率の変化

	128次元	192次元	256次元	320次元
第1位	82.9	83.4	83.7	83.8
第10位	96.9	97.1	97.2	97.2

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、256次元を超えると向上の割合は小さくなる。

#### (2) 各データにおける分類率

表3~4に評価データの分類率を示す(次元数256)。

表3. 分類率(漢字)

	第1位	第10位	第20位	第30位	第40位	第50位
評価1	97.9	99.8	99.9	99.9	100.0	100.0
評価2	97.7	99.9	99.9	100.0	100.0	100.0
評価3	93.0	99.1	99.5	99.6	99.7	99.7

表4. 分類率(非漢字)

	第1位	第10位	第20位	第30位	第40位	第50位
評価1	86.9	98.1	98.9	99.2	99.4	99.5
評価2	84.6	97.9	98.7	99.0	99.2	99.3
評価3	74.4	94.9	96.7	97.5	98.0	98.3

上記結果より、漢字では、すべてのデータにおいて、認識率、分類率共に高い性能を示しており、本方式が有効であることが判った。しかし、非漢字における認識率は十分でない。次に、文字の品質による性能を明らかにするために、評価3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データにおける本方式とオンライン方式<sup>[5]</sup>との認識率、分類率の比較結果を表5に示す。

表5. 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	90.4	99.1	67.0	92.1
	オンライン	88.1	96.6	69.8	87.7
mdb0055	本方式	96.0	99.9	85.8	98.4
	オンライン	96.7	99.0	87.9	96.9
mdb0066	本方式	96.1	99.6	83.8	98.8
	オンライン	96.3	98.5	88.2	96.5

表5に示すように、漢字では、品質に依らずオンライン方式と同程度の性能を示している。

## 5. まとめと今後の予定

オンライン方式で抽出可能な特徴を用いたOCR手法による認識方式の検討を行い、性能評価を行った。今後は、領域分割数を変化させた場合の性能評価、非漢字に有効な特徴の検討及びオンライン方式との併用による文字認識精度の向上を行う予定である。

### 参考文献

- [1]川又他:"大局的整合法と...", 信学総大D-482(1996)
- [2]南部他:"On-line...", ICP'98, vol. 2, pp. 1145-1149
- [3]岡本他:"Direction...", ICP'98, vol. 2, pp. 1747-1751
- [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)