

複数画像変換による認識

1D-4

— 形態情報・ストローク情報による —

外園 哲也 岩崎 智 平出 卓也 宮崎 孝之 天沼 博 高橋 邦夫

神奈川大学 工学部 電気工学科

1. はじめに

複数の変換画像を用いる事によって、認識率は向上する。さらに、この中の1種類の変換画像の類似度より選択された小数の辞書のみを利用し複数の変換画像の類似度計算を行う事によって、処理の簡略化が可能である。

2. 予備的事項

漢字データとして「手書漢字データベース ETL9B」(JIS 第一水準漢字 2965 字種・200 サンプル、計 593000 データ収録)を用い、「方向別ストローク類似度による分類」^[1]を一次分類として採用した。また、各字種のサンプルの内、奇数番目を辞書漢字、偶数番目を未知漢字としている。なお、入力画像は 64×63 ドットの 2 値画像である。

3. 画像変換

3.1 画像変換 1 - 正規化^[1]

入力画像の文字の大きさや癖、文字の太さによる誤認識を軽減するために、入力画像を 64×64 ドットの領域いっぱいに拡張する。この拡張度合は正規化パラメータ n で調整できる。本研究では、正規化パラメータ $n=0.8, 1.0$ を用いた。

3.2 画像変換 2 - 距離変換

3.2.1 領域拡大

距離変換処理において画像のはみ出しを防ぐため、64×64 ドットの正規化画像を、72×72 ドットに拡大する。

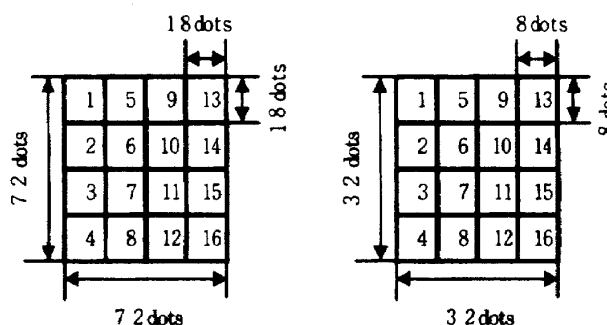
3.2.2 距離変換

形態情報抽出の前処理として、3.2.1 で得られた画像に対し 8 近傍距離変換を行う。原画像をレベル 1 と定義し、レベル 2 およびレベル 3 を形態情報抽出画像(等高線画像)として採用する。

4. 類似度

4.1 領域分割

距離変換画像を 18×18 ドットを 1 領域とする 16 領域に分割する。以後、各領域を「領域(b)」(b=1~16)と呼ぶことにする。



形態情報抽出用 特徴点情報抽出用
図1 領域分割

4.2 形態情報の抽出

3.2.2 で得られた等高線画像に対し、3×3 ドットのマスクパターン (256 種類) をかけて縦、横、右斜め、左斜め、不定の 5 パターンを抽出し、領域 (b) ごとに分類する。

4.3 形態情報による類似度

各領域 (b) での 5 種類のパターンの個数により、80 次元の特徴ベクトルができる。領域 (b) における候補漢字 k の辞書図形の特徴ベクトル $NP_i(b)$ と、入力図形の特徴ベクトル $NP_i(b)'$ とを次式で比較し、形態情報類似度 $DMI(b)^{(kd)}$ を算出する。(i: パターン数)

$$DMI(b)^{(kd)} = \sum_{i=1}^5 \left| NP_i(b)^{(kd)} - NP_i(b)' \right|$$

Recognition of Handwritten Chinese Characters by A Number of Image Transform - By Morphological Information and Stroke Information

Tetsuya Hokazono, Satoshi Iwasaki, Takuya Hiraide, Takayuki Miyazaki, Hiroshi Amanuma and Kunio Takahashi

Department of Electrical Engineering
Kanagawa University

4.4 圧縮細線化

特徴点を抽出するために、64×64ドットの正規化画像を32×32ドットに圧縮し、さらに線幅を1ドットに縮める細線化を行う。この圧縮細線化画像を8×8ドットを1領域とする16領域に分割する。(図1右図参照)

4.5 特徴点の抽出

4.4で得られた画像に対し、29種類の特徴点マスクパターン(3×3ドット)をかけ、端点・コーナー点・T字点・十字点という4種類の特徴点を抽出し、領域(b)ごとに分類する。

4.6 特徴点情報による類似度

各領域(b)での4種類の特徴点パターンの個数により、4×16=64次元の特徴ベクトルができる。領域(b)における候補漢字kの辞書図形の特徴ベクトル $TP_i(b)$ と入力図形の特徴ベクトル $TP_i(b)'$ とを次式で比較し、特徴点類似度 $DM2(b)^{(kd)}$ を算出する。

$$DM2(b)^{(kd)} = \sum_{i=1}^4 |TP_i(b)^{(kd)} - TP_i(b)'|$$

(k:候補漢字 d:学習サンプル i:パターン数)

4.7 類似度算出

次式より、候補漢字kの類似度 DM^k を算出する。

$$DM^k = \min_d \left[\sum_{b=1}^{16} \{ DM1(b)^{(kd)} + DM2(b)^{(kd)} \} \right]$$

5. 辞書選択

多数の変換画像による認識では処理速度の低下を招く原因となるために、1つの変換画像により辞書の選択を行う。1種類の変換画像の類似度をソートし、類似度の低い方から10個間隔で5個の辞書を取り出す。本研究では、正規化パラメータ1.0の距離変換レベル2の画像を使用して辞書の選択を行った。残りの画像は、この辞書に対してのみ処理を行い、類似度を算出する。

6. 総合化

算出された類似度の総合化を行う。正規化画像(n=0.8, 1.0)に対し、距離変換のレベルごとに算出した類似度 $(DM(0.8)_2^k, DM(0.8)_3^k)$ 、 $(DM(1)_2^k, DM(1)_3^k)$

を偏差値化した値をそれぞれ $(S(0.8)_2^k, S(0.8)_3^k)$ 、 $(S(1)_2^k, S(1)_3^k)$ とする。これらの値と一次分類で得られた類似度 AI^k とを総合化して得られた値 F^k を候補漢字に対する評価値とする。評価値 F^k は次式で与えられる。

$$F^k = \frac{\alpha AI^k + \beta(S(0.8)_2^k + S(0.8)_3^k) + \gamma(S(1)_2^k + S(1)_3^k)}{\alpha + \beta + \gamma}$$

$$(\alpha = 1, \beta = 1, \gamma = 1)$$

$$S^k = \frac{(\overline{DM} - DM^k)}{\sqrt{\sum_i (DM^k - DM^i)^2 / i}} \times 10 + 50$$

$$\overline{DM} = \sum_i \frac{DM^i}{i} \quad (i: \text{候補数})$$

7. 結果

1次候補列として利用した「方向別ストローク類似度による分類」は、正解率99.77%、平均候補数1.55個、認識率98.16%である。この候補列を使用して、辞書選択を行わずに処理を行った結果、正規化パラメータ0.8、1.0の距離変換レベル2、3の複数変換画像を用いて、認識率98.57%を得ている。ところが、辞書選択を行った結果、同様の複数変換画像を用いて、認識率99.11%と認識率の向上が認められた。また、辞書選択を行うと、行わない時と比較して処理時間を約3/5に抑える事ができ、処理速度の向上がみられた。

8. あとがき

辞書選択を行う事により、認識率および処理速度の向上がみられた。

[参考文献]

- [1] 熊谷憲二、金子真輝、新井武、栗原雅明、天沼博、高橋邦夫：「交叉情報・ストローク類似度による分類—正規化画像利用」電子情報通信学会パターン認識・理解研究会、PRU95-184、1995-12
- [2] 若林 哲史、Y. DENG、鶴岡 信治、木村 文隆、三宅 康二：「非線形正規化と特徴量の圧縮による手書き漢字認識の高精度化」、電気情報通信学会論文誌(D-2) Vol. J79-D-2 No. 5 pp. 765-774 1996-5