

H-018

稜線特徴量の二値手書き文字認識への適用と考察

Examination of Handwritten Characters Recognition for Binary Image by Ridge Features

岩永 晋輔† 林 千里† 鈴木 道孝‡ 伊藤 彰義‡
Shinsuke Iwanaga Chisato Hayashi Michitaka Suzuki Akiyoshi Itoh

1. まえがき

我々は多値画像の稜線に着目した特徴量—稜線特徴量—を提案した[1]. 稜線特徴量はストロークの中心線が最も濃いと考えられることから, ストロークの濃淡値を二次曲面で近似し, その曲面の極大点を結んだ稜線の方向を特徴としている. そのため入力画像は多値画像でないといけない. しかし二値画像に対して平滑化を行い, 多値画像を得ることにより稜線特徴量を適用できる. これを実際にデータベースを使って応用した.

本報告では平滑化を行った ETL9B に稜線特徴量[1]を適用し, 稜線特徴量が二値・多値の双方に適用できることを示す.

2. 稜線特徴量

2.1 稜線特徴量

画素点の付近で稜線が直線とみなせる領域でその直線の方程式が,

$$\mathbf{r} \cdot \mathbf{n}_{ij} + c_{ij} = 0 \quad (1)$$

で与えられているとする. ここで, $\mathbf{r} = (x, y)$ は稜線上の点であり, \mathbf{n}_{ij} は稜線の法線ベクトル, c_{ij} は定数である. 画素点 \mathbf{r}_{ij} は, 離散的に分布する点であるので, この方程式を正確に満たすわけではない. そこで, 1画素以内の距離に収まれば稜線上の点とし, 距離に応じた画素点の稜線強度を定義する. すなわち,

$$|\mathbf{r}_{ij} \cdot \mathbf{n}_{ij} + c_{ij}| \leq 1 \quad (2)$$

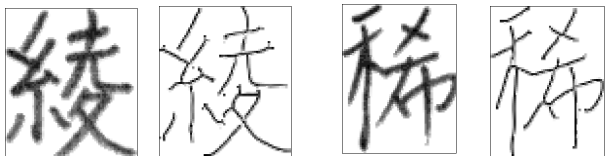
のとき, \mathbf{r}_{ij} を稜線上の点とし, その稜線強度を次のように定義する.

$$f_{ij} = 1 - |\mathbf{r}_{ij} \cdot \mathbf{n}_{ij} + c_{ij}| \quad (3)$$

この稜線の方程式を得るために, 近傍画素の画像濃淡値 g_{ij} を x, y の二次曲面で近似する.

$$g(x, y) = a_{xx}x^2 + a_{xy}xy + a_{yy}y^2 + a_x x + a_y y + a_0 \quad (4)$$

ただし近似精度を良くする為に背景領域に対して距離変換を行う. 実際に多値画像から得た稜線画像を図1に示す.



(a) 原画像 (b) 稜線画像 (c) 原画像 (d) 稜線画像

図1 抽出した稜線特徴量

識別のためには次元の小さな特徴ベクトルを作る必要がある. 特徴ベクトルを作るために, 小領域に分割をし, 小領域ごとに重みをつけて平均を取る. 小領域は縦・横それぞれストロークの数になるように, かつ領域が半分ずつ重なるように決定する. 各領域の重みはすべての画素が均等になるように重みを付ける.

このようにして, 稜線の方向4方向と, 小領域を縦・横7個ずつに分割し, 合計196次元($4 \times 7 \times 7$)の特徴を作る.

2.3 輪郭特徴量との比較

輪郭特徴は距離変換していない画像から Sobel フィルタで抽出した勾配ベクトルの強度を式(3)の代わりに使って求める. 図2に稜線特徴と輪郭特徴の比較を示す.

図2の輪郭特徴では, 「鳥」の第4画の水平ストロークには十分な情報が抽出されず, 「鳥」に誤読される可能性があるが, 稜線特徴では明瞭に抽出されていることがわかる.



(a) 原画像 (b) 輪郭強度 (c) 稜線強度

図2 稜線特徴と輪郭特徴の比較

3. 二値手書き文字への応用

稜線特徴を二値手書き文字画像に応用する. 図3に示すように, ETL9B データベースをそのまま稜線を抽出した場合, 文字の中心部をうまく抽出することができない(図3(a), (b)). これは二値画像の場合, 文字領域が平坦であり, 二次曲面で近似して, その極大を求めることができない為である. 2×2 の平均フィルタを2回かけて平滑化を行うことでそれが改善される(図3(c), (d)).



(a) (b) (c) (d)

図3 二値画像における稜線特徴量 (a)原画像 (b)二値画像から抽出した稜線画像 (c)平滑化後 (d)平滑化した稜線画像から抽出した稜線画像

† 日本大学理工学研究科

‡ 日本大学理工学部

4. 認識実験

稜線特徴量を文字画像データベース ETL9B, および ETL9G のデータを用いた認識実験を行う. 比較する為に輪郭特徴量の実験も行う.

データベースは各字種 200 個のサンプルを 20 個ずつ, 10 個のデータグループに分け, 9 グループを学習サンプルとして用い, 残りの 1 グループを未知サンプルとして用いる. これをデータグループの役割を代えて 10 通りの認識実験を行い, その平均の認識率を, 全体の認識率とする.

識別には階層ベイズ判別分析(HBDA)を用いる[2]. 事後確率を正規分布で近似することにより(5)式の識別関数を得る. この式が最小となる字種と判別する.

$$g(\mathbf{x}, c) = (\mathbf{x} - \mathbf{m}_c)^T \Sigma_c^{-1} (\mathbf{x} - \mathbf{m}_c) + \ln |\Sigma_c| \quad (5)$$

$$\tilde{c}(\mathbf{x}) = \underset{c}{\operatorname{argmin}} g(\mathbf{x}, c) \quad (6)$$

特徴ベクトル \mathbf{x} , 字種 c とし, 各字種の分布のパラメータの期待値を平均ベクトル \mathbf{m}_c , 共分散行列を Σ_c とする. ここで, 各字種の事前確率は等しいものとする.

HBDA では共分散行列を, 次のような形で推定する.

$$\Sigma_c = \frac{n_c S_c + \nu_1 \operatorname{diag} S_D + \nu_2 S_D}{n_c + \nu_1 + \nu_2} \quad (7)$$

n_c は字種 c のサンプル数である. S_D は各字種の級内共分散をすべての字種の集合 D の要素にわたり平均したものである. ν_1, ν_2 は正のハイパーパラメータであり, 認識率が最も良い値を用いる.

5. 結果, 考察

認識結果を表 1 に示す. 結果より, ETL9B で稜線特徴を用いた認識率は 99.11% となり, ETL9G と比較して 0.25% 減少するが, 99% 以上の認識率を挙げている. なお [2] では, さらに二段階識別を適用することによって, ETL9G に対して現在最高の 99.47% の認識率を達成している.

表 1 認識率

	稜線特徴量	輪郭特徴量
ETL-9G	99.36±0.07	99.09±0.14
ETL-9B	二値	98.90±0.18
	平滑化後	99.11±0.14

※±**は 10 回の実験における認識率の標準偏差を示す

5.1 二値画像と多値画像における稜線の違い

ETL9B での稜線特徴量の認識率を見てみると二値のままではなく 2×2 の平均化フィルタを 2 回をかけることで精度が上がっていることがわかる.

図 4 より平滑化された ETL9B から抽出された稜線画像は ETL9G と比較して同程度の稜線が抽出されているのが確認できる.

ETL9B と ETL9G の違いを比較する. 図 5 において平滑化した ETL9B から抽出した稜線画像では孤立点が多く, 稜線強度が小さいことがわかる. このため ETL9B では ETL9G よりも認識率が低くなったと考えられる. 孤立点や稜線強度が小さくなる原因としては, 原画像が二値化されている為, ストロークの中心部分が必ずしも最も濃い濃淡値であるとは限らないためと考えられる. このため, 二値化の方法や平滑化の方法を改良すれば, 更に認識率が上がると考えられる.

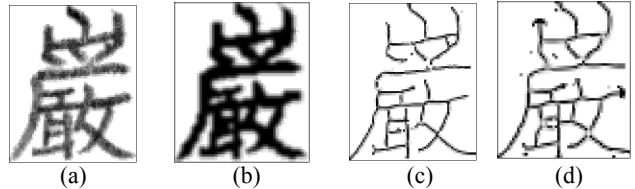


図 4 各データベースでの稜線画像 (a)ETL9G 原画像 (b)ETL9B 平滑化後 (c)ETL9G 稜線画像 (d)ETL9B 平滑化後稜線画像

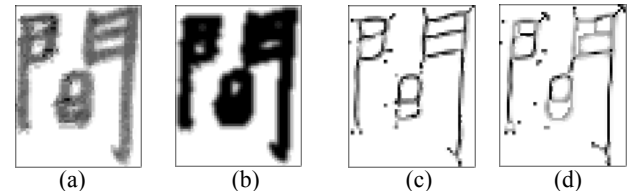


図 5 認識率低下原因サンプル「問」 (a)ETL9G 原画像 (b)ETL9B 平滑化後 (c)ETL9G 稜線画像 (d)ETL9B 平滑化後稜線画像

5.2 二値画像における稜線特徴と輪郭特徴の違い

ETL9B を用いた場合, 稜線特徴量は輪郭特徴量に比べ 0.09% 認識率が高い. 図 6 に示すような類似文字において輪郭特徴を用いた場合だと潰れが起きてしまうが, 稜線特徴では中心部分まで抽出されている.

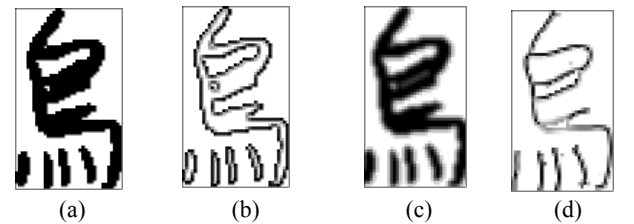


図 6 ETL9B における輪郭特徴と稜線特徴「鳥」 (a)ETL9B 原画像 (b)ETL9B 輪郭特徴 (c)ETL9B 平滑化後 (d)ETL9B 平滑化後稜線画像

6. むすび

稜線特徴量を二値手書き文字に適用した. 二値画像を平滑化することにより多値手書き文字と同様に稜線を抽出できた. また, これにより二値画像に対しても 99% 以上の精度で識別でき, 二値画像に対しても稜線特徴量が応用できることがわかった.

文 献

[1] 鈴木道孝, 林千里, 伊藤彰義, "稜線特徴量による多値画像手書き文字の認識," 信学技法, PRMU2006-269, pp.85-90, March 2007.
 [2] 鈴木道孝, 伊藤彰義, "階層ベイズモデルによる二段階識別とその手書き文字認識への応用," 情処学論に投稿中