

G-9

手書き文字認識の稜線特徴量における稜線幅と稜線強度形状の評価と検討

Investigation of Ridge Width and Ridge Intensity Form in Ridge Features for Handwritten Character Recognition

○長谷川周<sup>1</sup>, 鈴木道孝<sup>2</sup>, 塚本新<sup>3</sup>, 伊藤彰義<sup>2,4</sup>

\*Shu Hasegawa<sup>1</sup>, Michitaka Suzuki<sup>2</sup>, Arata Tsukamoto<sup>3</sup>, Akiyoshi Itoh<sup>2,4</sup>

The ridge feature is effective for handwritten character recognition. The ridge feature intensity is defined as non-zero within the vicinity of the ridge lines in multivariate character images, and its width was reported to be optimized as 1.5 pixel. In this report, the variance of the feature due to a small displacement of a character stroke is shown to be non-zero in the conventional scheme, and therefore a new improved scheme with zero variance is proposed. Experiments on the ETL9G database show that the proposed scheme yields higher recognition rates than the conventional one.

1. はじめに

我々は多値手書き文字の認識における特徴量として、稜線特徴量<sup>[1]</sup>を用いている。稜線特徴量では稜線半幅をパラメータとして用いているが<sup>[2]</sup>、稜線半幅を変化させた際に、誤認識の原因となり得る特徴量の変動が起こる。本稿では特徴量の変動を評価し、それを改善する稜線強度の決定方法を提案する。

2. 稜線特徴量

稜線特徴は文字ストロークの濃淡値が高い部分の連続した方向を特徴としている。稜線の方程式を得るために、注目画素を中心に5×5画素の濃淡値をみる。注目画素と、注目画素からx方向, y方向それぞれn画素離れた点での画像濃淡値 $g_{ij}$ をx, yの二次曲面

$$g(x, y) = a_{xx}x^2 + a_{xy}xy + a_{yy}y^2 + a_xy + a_0 \quad (1)$$

で近似する。これにより、文字画像の濃淡値を局所的に二次曲面で近似することができる。

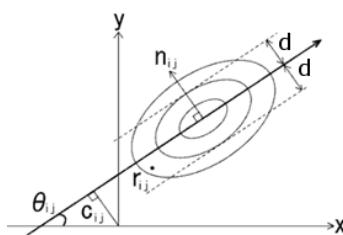


Fig. 1 Extracting method of ridge line.

図1のように、画素点 $r_{ij}$ の付近において稜線が直線とみなせる領域でその直線の方程式が、

$$\mathbf{r} \cdot \mathbf{n}_{ij} + c_{ij} = 0 \quad (2)$$

で与えられるとする。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 $r_{ij}$ は、離散的に分布する点のため、式(2)を正確に満たすわけではない。[2]では、稜線半幅を $d$ として、 $d$ 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が次のように定義される。

$$f_{ij} = \begin{cases} d - |\mathbf{r}_{ij} \cdot \mathbf{n}_{ij} + c_{ij}| & (0 \leq |\mathbf{r}_{ij} \cdot \mathbf{n}_{ij} + c_{ij}| < d) \\ 0 & (d \leq |\mathbf{r}_{ij} \cdot \mathbf{n}_{ij} + c_{ij}|) \end{cases} \quad (3)$$

稜線が画素点 $r_{ij}$ を通るとき、その稜線強度は $d$ となり、

稜線が二つの画素点の間を通るとき、二つの画素点は稜線との距離に応じた $d$ より小さい稜線強度を持つ。距離が $d$ 以上の場合、稜線強度は0となる。

推定された稜線を画素ごとに45°おきに4方向に量子化し、文字画像を縦・横7個ずつの小領域で分割し稜線強度の重み付け平均をとる。4方向×縦7×横7領域の合計196次元のベクトルを特徴量として用いる。

3. 従来の稜線強度

従来の稜線強度は、式(3)で定義されている。これは、稜線として推定された直線が、注目画素点と完全に重なった場合に稜線強度を最大とし、直線と画素点との距離が大きくなるにつれ、正比例的に稜線強度を減少させるものである。この稜線強度の決定方法を図2に示す。このように、各画素点における稜線強度を図示したものを稜線強度形状と呼ぶ。

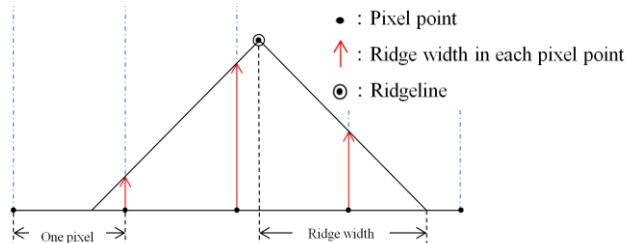


Fig. 2 Pyramid ridge intensity form.

図2のような従来の稜線強度形状をピラミッド形稜線強度と呼ぶことにする。ピラミッド形稜線強度における、稜線半幅 $d$ を変化させた時の誤読文字数のグラフを図3の赤の棒グラフに示す。

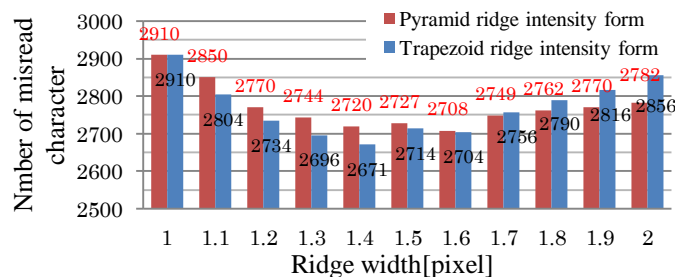


Fig. 3 Number of misread character in two ridge intensity.

図 3 を見ると、稜線半幅 1.0~1.4 の範囲では誤読文字数が減少しており、1.6~2.0 の範囲では増加している。稜線半幅 1.4 以下では、元画像における濃淡値を持つ画素数に比べ、稜線強度が 0 となり特徴量を持たない画素数が多くなるため誤読文字数が増加していると考えられる。稜線半幅 1.7~2.0 の範囲では、二次曲面近似の際に、5×5 画素の窓内において実際の濃淡値分布をよく近似できていない曲面の稜線まで特徴量に含めるため誤読文字数が増加すると考えられる。

従来の稜線強度を評価するために、文字ストロークの垂直（または平行）方向成分の、書き手による“ずれ”を考える。ストロークにずれがある場合、ピラミッド形稜線強度では稜線幅内の画素における稜線強度の合計に差がでる。この差は稜線半幅とずれの大きさに依存しており、特徴量にばらつきが生じる原因となるため、誤読文字数が増加すると考えられる。

稜線強度のばらつきを評価するために変動係数  $V$  を計算する。変動係数を式(4)に示す。

$$V = \sigma/m \quad (4)$$

ストロークのずれに対する稜線強度（稜線方向における単位長さ当たりの積分値）の標準偏差  $\sigma$  を、平均  $m$  で割った値が変動係数となる。これは、ばらつきを表す標準偏差を平均で除して無単位としたもので、稜線強度のばらつきを表す。図 4 に稜線半幅を変えた時の変動係数のグラフを示す。稜線半幅 1.4 付近で変動係数が大きくなっていることが分かる。したがって稜線半幅 1.4 付近では、稜線強度のばらつきが大きく、認識精度が悪くなると考えられる。

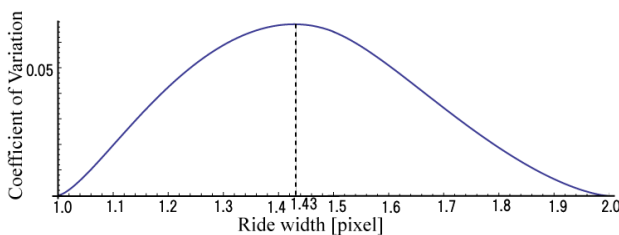


Fig. 4 Coefficient of Variation of pyramid ridge intensity.

#### 4. 台形稜線強度

稜線強度のばらつきにより誤読文字数が増加すると考えられるため、稜線強度のばらつきを表す変動係数を 0 にすることで、誤読文字数が減少すると考えられる。変動係数を 0 にするための新たな稜線強度形状として、台形稜線強度を提案する。台形稜線強度を次のように定義する。

$$f_{ij} = \begin{cases} 1 & (0 \leq |r_{ij} \cdot n_{ij} + c_{ij}| < d - 1) \\ d - |r_{ij} \cdot n_{ij} + c_{ij}| & (d - 1 \leq |r_{ij} \cdot n_{ij} + c_{ij}| < d) \\ 0 & (d \leq |r_{ij} \cdot n_{ij} + c_{ij}|) \end{cases} \quad (5)$$

台形稜線強度ではストロークのずれによる、稜線幅内の稜線強度の差が出ないため、(4) 式における標準偏差  $\sigma$  が 0 となり、変動係数も 0 となる。変動係数がストロークのずれ、稜線半幅に依存せずに常に 0 となるため、稜線半幅 1.4 付近において誤読が軽減されると考えられる。ストロークのずれに対する稜線強度のばらつきの例を図 5 に示す。

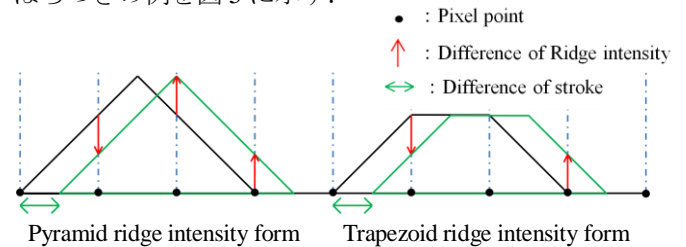


Fig. 5 Variation of ridge intensity if ridge width is 1.5.

#### 5. 実験条件

文字データベース ETL-9G(3036 字種、各 200 文字)を対象に、提案方法を用いた認識実験を行う。文字の特徴量には稜線特徴量、識別には階層ベイズ判別分析を用いる。各字種 180 文字を学習文字、残り 20 文字を未知文字とし、それを 1 セットとして、10 回の交差検証により認識実験を行う。

#### 6. 実験結果

図 3 の青の棒グラフに、台形稜線強度における稜線半幅を変化させた時の誤読文字数のグラフを示す。稜線半幅 1.6 以下の場合に誤読文字数が減少していることが分かる。またピラミッド形稜線強度における最少誤読文字数は、稜線半幅 1.6 の時の 2708 文字であったが、提案手法では稜線半幅 1.4 における誤読文字数が最も少なく、2671 文字まで減少した。稜線半幅 1.43 で最大値をとる変動係数を、0 とする稜線強度形状を適用した結果、稜線半幅 1.4 での誤読文字数が最も減少したことが分かる。

#### 7. まとめ

稜線強度のばらつきを評価するために、ピラミッド形稜線強度の変動係数を算出し、変動係数を 0 にするような台形稜線強度を提案した。結果、変動係数が最大となる稜線半幅 1.4 における誤読文字数が、最も減少し、従来の稜線強度では 2910 文字だった誤読文字数を 2671 文字まで減少させることができた。

#### 参考文献

- [1] 鈴木道孝, 林千里, 伊藤彰義: “稜線特徴量による多値手書き文字認識”, 信学技報, PRMU, vol.106(606), pp.85-90, March 2007.
- [2] 王偉: “稜線特徴量における稜線強度の検討” 卒業論文, 2011