

ランレングス補正を用いた画質劣化にロバストな特徴抽出法

森 稔[†] 澤木美奈子[†] 萩田 紀博^{†*} 村瀬 洋[†]
武川 直樹[†]

Robust Feature Extraction Method Based on Run-Length Compensation for Degraded Character Recognition

Minoru MORI[†], Minako SAWAKI[†], Norihiro HAGITA^{†*}, Hiroshi MURASE[†],
and Naoki MUKAWA[†]

あらまし 文字認識では、幾何学的特徴抽出を基本とする変形にロバストな手法、及びテンプレートマッチングを基本とする画質劣化にロバストな手法が提案されている。しかし、両手法は個別に検討されており、画質劣化と変形の両方にロバストな手法は実現されていない。本論文では、変形にロバストな文字線のランレングスに基づく幾何学的特徴抽出において、画質劣化に対するロバスト性を向上させる手法を提案する。提案手法は、ランレングスを計測する際、画像の劣化状態の推定結果に基づいてランレングスを補正し、劣化のない状態で得られるランレングスの近似値を求めることを特長とする。人工的に劣化させた手書き文字を用いた認識実験により、提案手法が変形に対するロバスト性を保持しつつ、従来手法より画質劣化に対するロバスト性を向上可能であることを示す。

キーワード OCR, 特徴抽出, ランレングス補正, 劣化推定, 方向寄与度

1. ま え が き

文字認識において、認識性能を低下させる主な原因には、画質劣化と変形がある。しかし、従来から提案されている認識手法の多くは、どちらか片方の問題に絞って検討されている。例えば、画質劣化に対処した手法として、対象文字画像の劣化状態やフォント形状を反映させたテンプレートの獲得法が提案されている [1]~[4]。また、ノイズやテクスチャにより劣化した文字を認識可能とする識別尺度も提案されている [5]~[7]。しかし、上記手法は画像間のテンプレートマッチングを基本としているため、変形が少ない印刷文字等に適用範囲が限られる。

一方、マルチフォントや手書き文字における変形に対しては、文字線の幾何学的情報、特に文字線の方向情報を抽出した特徴（方向特徴）が有効であることが

知られている [8]。例えば、方向特徴の一つに、文字線のランレングスに基づいた方向寄与度特徴 [9] 及び LSD 特徴 [10], [11] がある。しかし、幾何学的特徴を用いる認識手法では、文字画像に画質劣化が生じた場合、幾何学的情報を正しく抽出することが困難になるため、認識率が急激に低下するという問題がある。画質劣化の影響を低減するため、背景ノイズ等の除去 [12]~[14] やかすれた部分の穴埋め [15] を行う前処理が提案されている。しかし、これらの手法では本来残すべき文字線をノイズと誤判定して除去したり、逆にノイズを強調してしまうなど、必ずしも有効に機能するわけではない。

本論文では、変形にロバストな方向特徴を基本とし、画質劣化に対するロバスト性を向上することにより、画質劣化と変形の商品質低下にロバストな特徴抽出法を提案する。ここで、画質劣化の種類は汚れ及びかすれを対象とする^(注1)。画質劣化へのロバスト性を高めるため、入力文字画像の劣化状態を推定し、推定結果に応じて文字線のランレングスを補正することにより、

[†] 日本電信電話株式会社 NTT コミュニケーション科学基礎研究所, 厚木市

NTT Communication Science Laboratories, NTT Corporation, 3-1 Morinosato-Wakamiya, Atsugi-shi, 243-0198 Japan

* 現在, ATR メディア情報科学研究所

(注1): ただし、汚れとかすれが同時に生じたものは考慮しない。

劣化のない場合に得られるランレングスの近似値を求める“ランレングス補正法”を提案する．以下，ランレングス補正法のアルゴリズムを，2. で説明する．3. で，ノイズ種類の判定法を述べる．上記技術を用いた画質劣化にロバストな特徴抽出処理を，4. で説明する．5. で，手書き漢字データ ETL-9 [16] を人工的に劣化させた文字画像を用いた認識実験において，提案手法により変形へのロバスト性を保持しつつ，従来手法より画質劣化に対するロバスト性が向上されていることを示す．まとめと今後の課題を，6. で述べる．

2. ランレングス補正法

2.1 基本的な考え方

一般に，幾何学的情報は汚れまたはかすれの程度が大きくなるにつれ，正しく抽出することが困難になる．幾何学的情報の一つである文字線のランレングスについても，汚れやかすれにより本来計測されるべき値が大きく変化する．そこで，画質が劣化した文字から，文字線のランレングスを安定して抽出可能とする“ランレングス補正法”を提案する．提案手法の基本的な考え方は，観測領域の劣化程度を推定し，推定結果に応じて劣化によりひずんだランレングスを補正することにより，劣化のない状態で計測されるランレングスの近似値を求めることにある．

観測領域として，文字画像上に方形窓を設定する．劣化状態を推定するための情報として，方形窓内の黒画素及び白画素の連続性に着目する．認識対象となる文字カテゴリーが限定されると，文字の複雑さはある範囲に限定される．したがって，方形窓内で本来観測されるべき範囲以上に黒画素及び白画素が複雑に分布していれば，方形窓内の領域は劣化している（汚れている若しくはかすれている）ととらえることができる．

以下，横方向のランレングスを計測及び補正する場合について説明する．2 値の入力文字画像 G (横 W 画素 \times 縦 W 画素) 上に，観測用の方形窓として，横幅が I 画素，高さが 1 画素の領域を設定する．着目している座標 (x, y) の画素の値を， $g_{x,y}$ とする． x は画素の横方向の座標である． y は縦方向の座標である．画素値 $g_{x,y}$ は，黒画素なら 1，白画素なら 0 の値をとる．ここで，方形窓内の画素を左端から右端まで走査することにより得られる 4 種のパラメータ a, b, c, e を，以下の式で定義する．

$$a = \sum_{k=0}^{I-2} g_{x+k,y} \cdot g_{x+k+1,y} \quad (1)$$

$$b = \sum_{k=0}^{I-2} (1 - g_{x+k,y}) \cdot g_{x+k+1,y} \quad (2)$$

$$c = \sum_{k=0}^{I-2} g_{x+k,y} \cdot (1 - g_{x+k+1,y}) \quad (3)$$

$$e = \sum_{k=0}^{I-2} (1 - g_{x+k,y}) \cdot (1 - g_{x+k+1,y}) \quad (4)$$

a は黒画素同士の連結数， b は白画素から黒画素への変化数， c は黒画素から白画素への変化数， e は白画素同士の連結数，を各々表す．図 1 に，文字画像上に観測用の方形窓を設定し，パラメータを計測する様子を示す．

以下，上記 4 種のパラメータを用いた具体的なランレングスの補正アルゴリズムを，汚れ及びかすれのノイズ種類別に分けて説明する（ノイズ種類の判定処理については 3. で述べる）．

2.2 汚れに対する補正アルゴリズム

方形窓内で観測される黒画素数を r_o ，その中で本来文字線を構成するランレングスを r_t ，ノイズ成分を r_n とすると，

$$r_o = r_t + r_n \quad (5)$$

となる．したがって，求めるべきランレングス r_t は，

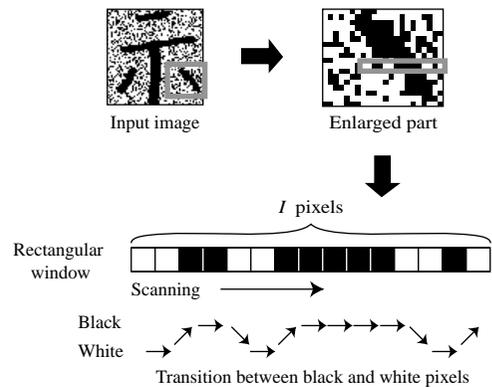


図 1 入力画像からのパラメータ計測例
Fig.1 Extraction of parameters from the input image.

$$\begin{aligned} r_t &= r_o - r_n \\ &= (1 - r_n/r_o) \cdot r_o \end{aligned} \quad (6)$$

である。 r_n/r_o は、観測された黒画素数中にノイズが含まれている割合（劣化程度）を表しており、汚れが生じていない場合は 0、汚れているほど 1 に近づく。

$$0 \leq r_n/r_o \leq 1 \quad (7)$$

この r_n/r_o を、2.1 で述べたパラメータ a, b, c, e を用いて表すことを考える。ここで、 r_n/r_o は、黒画素数における平均値からの増加率と、複雑度における平均値からの増加率との比で近似できると仮定する。ただし、平均値とは認識対象文字集合から得られる値であり、具体的には学習データでの平均値である。複雑度は、黒画素から白画素への変化及び白画素から黒画素への変化の頻度数を表す。増加率とは、平均値から増加した画素数の比率、及び増加した頻度数の比率である。つまり、 r_n/r_o は汚れたことによって平均的な値からどの程度黒画素数が増え、どの程度画素分布が複雑になったかの割合である、と仮定する。

ノイズが増加したことによって生じる、黒画素数における平均値からの増加率を ΔK_{add} 、複雑度における平均値からの増加率を ΔV_{add} とすると、黒画素数の増加率と複雑度の増加率の比 β_{add} は、

$$\beta_{\text{add}} = \Delta K_{\text{add}}/\Delta V_{\text{add}} \quad (8)$$

となる。 β_{add} は、汚れのない状態で 1 になる。また、汚れがあまりひどくない段階では、ノイズ量が増大するにつれ、 ΔK_{add} の増加率以上に ΔV_{add} の増加率が大きいため、 β_{add} は 0 に近づく。したがって、上記仮定のもとで、

$$r_n/r_o \simeq 1 - \beta_{\text{add}} \quad (9)$$

とおくことができ、式 (9) を式 (6) に代入することにより、

$$\begin{aligned} r_t &= (1 - (1 - \beta_{\text{add}})) \cdot r_o \\ &= \beta_{\text{add}} \cdot r_o \end{aligned} \quad (10)$$

が得られる。ここで、 β_{add} は方形窓内の黒画素数を補正し、ランレングスの近似値を算出する補正係数となっている。

具体的な ΔK_{add} 及び ΔV_{add} の定義はいくつか考えられる。本論文では、画質劣化のない全学習データ

から計測される a, b, c, e の平均値を、各々 $\bar{a}, \bar{b}, \bar{c}, \bar{e}$ としたとき、

$$\Delta K_{\text{add}} = (a + b)/(\bar{a} + \bar{b}) \quad (11)$$

$$\Delta V_{\text{add}} = (b + c)/(\bar{b} + \bar{c}) \quad (12)$$

と定義する。黒画素数 r_o は、

$$r_o = a + b \quad (13)$$

であるので、最終的なランレングスの近似値 r'_t は、式 (8)、(10)~(13) より、

$$r'_t = \frac{(a + b)/(\bar{a} + \bar{b})}{(b + c)/(\bar{b} + \bar{c})} \cdot (a + b) \quad (14)$$

により算出される。

2.3 かすれに対する補正アルゴリズム

汚れの場合と同様に、方形窓内で観測される黒画素数、本来の文字線のランレングス、ノイズ成分を、各々 r_o, r_t, r_n とすると、

$$r_o = r_t - r_n \quad (15)$$

となる。求めるべきランレングス r_t は、

$$\begin{aligned} r_t &= r_o + r_n \\ &= (1 + r_n/r_o) \cdot r_o. \end{aligned} \quad (16)$$

である。汚れにおける黒画素を基準とした仮定を、白画素を基準とした仮定に置き換えて検討する。白画素の平均値からの増加率を ΔK_{sub} 、複雑度の平均値からの増加率を ΔV_{sub} 、白画素量の増加率と複雑度の増加率の比を β_{sub} とすると、各々

$$\beta_{\text{sub}} = \Delta K_{\text{sub}}/\Delta V_{\text{sub}} \quad (17)$$

$$\Delta K_{\text{sub}} = (e + c)/(\bar{e} + \bar{c}) \quad (18)$$

$$\Delta V_{\text{sub}} = (b + c)/(\bar{b} + \bar{c}) \quad (19)$$

と定義できる。

$$r_n/r_o \simeq 1 - \beta_{\text{sub}} \quad (20)$$

とおくことにより、最終的なランレングスの近似値 r'_t は、

$$r'_t = \left(2 - \frac{(e + c)/(\bar{e} + \bar{c})}{(b + c)/(\bar{b} + \bar{c})} \right) \cdot (a + b) \quad (21)$$

により算出される。

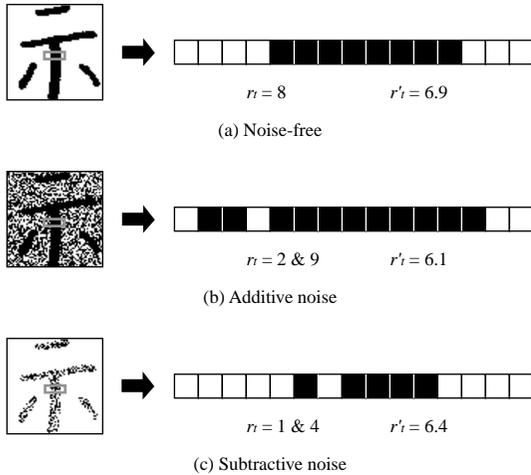


図2 観測ランレングス及び補正ランレングスの抽出例． $\bar{a} = 7.4$, $\bar{b} = 0.9$, $\bar{c} = 0.9$, $\bar{e} = 4.8$ は実験値から算出．

Fig. 2 Extraction of observed run-length and compensated run-length. $\bar{a} = 7.4$, $\bar{b} = 0.9$, $\bar{c} = 0.9$, and $\bar{e} = 4.8$ are calculated from training data.

2.4 補正ランレングスの算出例

従来の黒画素連結数を計測して得られるランレングス（以後、観測ランレングスと称す）及び補正ランレングスについて、劣化のない文字・汚れた文字・かすれた文字の各々から算出した例を、図2に示す．図2より、補正ランレングス r'_t は、本来得られるべき値である観測ランレングス $r_t = 8$ より、やや小さい値が算出されている．しかし、観測ランレングスが汚れに対し $r_t = 2 \& 9$ 、かすれに対し $r_t = 1 \& 4$ と変動が激しいことに比べ、補正ランレングスは、汚れに対し $r'_t = 6.1$ 、かすれに対し $r'_t = 6.4$ と、より安定した算出が可能であることがわかる．

また、2.2 及び 2.3 と同様の処理を、横方向のみならず縦方向・右斜め方向・左斜め方向にも適用することにより、各方向での補正ランレングスを算出することができる．

3. ノイズ種類の判定

本章では、入力文字画像におけるノイズ種類（汚れまたはかすれ）の判定処理について説明する．2. で述べたランレングス補正法は、汚れまたはかすれに応じて異なる処理をする．したがって、事前に入力文字画像が汚れているか、若しくはかすれているかを判定す

る必要がある．ここでは、2.1 で用いた画素の連続性に関する性質を利用する．具体的には、ランレングス観測用の方形窓を文字画像全体の走査 ($I = W$) に拡張する．文字画像を横方向に走査することにより得られる a, b, c, e を用い、以下の式で定義される射影値 $p(y)$ ($y = 1, \dots, W$) [6] を、各行ごとに算出する．

$$p(y) = \frac{a \cdot e - b \cdot c}{\sqrt{(a+b) \cdot (c+e) \cdot (a+c) \cdot (b+e)}} \quad (-1 \leq p(y) \leq 1) \quad (22)$$

上記定義式において、汚れている場合には、 $a \cdot e$ が大幅に減少し、かつ $b \cdot c$ が大幅に増加する．かすれている場合には、わずかの $a \cdot e$ の減少と $b \cdot c$ の増加が起こる．したがって、射影値の増減を検出することにより、入力文字画像が汚れているか、またはかすれているかの判定ができる．上記処理を横方向に加え縦方向にも適用することにより、最終的にノイズ種類判定用の $2W$ 次元からなる特徴ベクトル T を算出する．

事前学習では、ノイズの種類及び劣化程度別に、全カテゴリーの学習用劣化文字画像から $2W$ 次元の特徴ベクトル T を抽出し、平均ベクトル M を作成する．ノイズ種類判定時には、入力文字画像から同様に抽出した特徴ベクトル T と各劣化程度の特徴ベクトル M 間の距離値を計算し、距離値が最小となる M の属するノイズ種類を、入力文字画像のノイズ種類と判定する．

4. 画質劣化にロバストな特徴抽出処理

方向特徴の一つとして、各方向の分布程度を表す方向寄与度特徴 [9] 及び LSD 特徴 [10], [11] がある．本論文では、方向寄与度特徴を用いる．方向寄与度は、文字線のランレングスから計算することができ、以下の d_i ($i = 1, \dots, 4$) により求められる．

$$d_i = \frac{l_i}{\sqrt{\sum_{j=1}^4 l_j^2}} \quad (23)$$

ここで、 l_1, l_2, l_3, l_4 は、各々横、右斜め、縦、及び左斜め方向の文字線のランレングスである．方向寄与度の抽出例を、図3に示す．

今回、画質劣化文字の認識において、ランレングス補正法及びノイズ種類判定処理を用いた特徴抽出処理を、以下のように行う．

Step 1: 入力文字画像のノイズ種類を判定する．

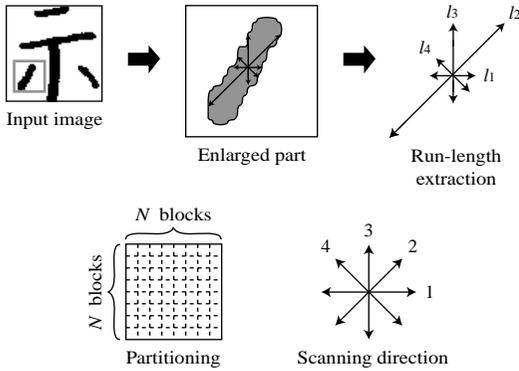


図 3 方向寄与度の抽出例

Fig. 3 Feature extraction of direction contributivity.

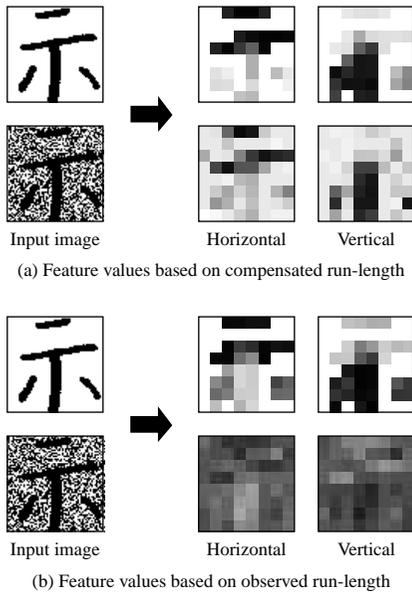


図 4 補正ランレングス及び観測ランレングスにより得られた特徴値の抽出例．横方向及び縦方向の特徴値を濃淡値で可視化．

Fig. 4 Examples of the feature values based on compensated run-length and observed run-length. Horizontal and vertical feature values are visualized.

Step 2: 入力文字画像を、 $N \times N$ 個の粗い局所領域に分割する．

Step 3: 各黒画素において、Step 1 の結果に基づき、補正ランレングス $r'_{t,i}$ ($i = 1, \dots, 4$) を求める．

Step 4: 局所領域ごとに、補正ランレングス $r'_{t,i}$ を平均化する．

Step 5: 式 (23) において、 l_i の代わりに平均化された補正ランレングス $r'_{t,i}$ を使い、局所領域ごとに方向寄与度 d_i を計算する．

ここで、方向寄与度をランレングスの平均化後に算出するのは、ランレングス補正によっても補正しきれないランレングスの変動を抑制し、より安定した方向寄与度の特徴値を得るためである．また、本論文では 1 辺当りの局所領域数を $N = 8$ とする．つまり、文字画像は $8 \times 8 = 64$ 個の局所領域に分割される．

図 4 (a) に補正ランレングスに基づく方向寄与度の特徴値を、図 4 (b) に観測ランレングスに基づく方向寄与度の特徴値を、各々劣化のない文字画像及びノイズを重ねた文字画像から抽出し、濃淡画像として図示した例を示す．図 4 より、補正ランレングスを用いた特徴は、背景ノイズの影響を抑制し文字線の方向情報をよく表現していることがわかる．

5. 認識実験

5.1 データ及び実験条件

提案手法を評価するため、ノイズ種類の判定性能評価、及び補正ランレングスに基づいた特徴と観測ランレングスに基づいた特徴との比較を行った．

実験データには、変形に対するロバスト性を評価するため、手書き漢字データ ETL-9 (3,036 カテゴリー、200 サンプル/カテゴリー)[16] を用いた．学習データには奇数番 (1, ..., 199) の 100 サンプル/カテゴリーを、評価データには偶数番 (2, ..., 200) の 100 サンプル/カテゴリーを用いた．更に、画質劣化に対するロバスト性を評価するため、上記学習用及び評価用データから、汚れ及びかすれノイズ文字画像を以下の手順で生成した [6]．かすれノイズ文字を作成するためのノイズモデル画像として、すべて黒画素からなる画像において、 $|\alpha|\%$ ($\alpha < 0$) の黒画素を白画素に反転させた画像を作成する．また、汚れノイズ文字を作成するために、すべて白画素からなる画像において、 $\alpha\%$ ($\alpha \geq 0$) の白画素を黒画素に反転させたノイズモデル画像を作成する．ここで、 $|\alpha|$ は画素の白黒を反転させる割合であり、ノイズモデル画像における劣化程度を表す．本論文では、 $-70 \leq \alpha \leq 70[\%]$ の範囲を 10 刻みで用いた．かすれノイズ文字画像 G^α を、ノイズのない文字画像 G とノイズモデル画像 Z^α 間の AND 演算により生成する．汚れノイズ文字画像 G^α を、 G と Z^α 間の OR 演算により生成する．すなわち、

$$g_{x,y}^\alpha = \begin{cases} g_{x,y} \wedge z_{x,y}^\alpha & \text{if } \alpha < 0, \\ g_{x,y} \vee z_{x,y}^\alpha & \text{otherwise.} \end{cases} \quad (24)$$

ここで、 $g_{x,y}^\alpha$, $g_{x,y}$, $z_{x,y}^\alpha$ は、各々 G^α , G , Z^α における座標 (x, y) の画素値 (0 または 1) を表す。汚れノイズモデル画像、かすれノイズモデル画像、及び各劣化文字画像の生成例を、図 5 に示す。学習用データから生成した劣化文字画像はノイズ種類判定用の学習用データに、評価用データから生成した劣化文字画像は各実験の評価用データとした。この劣化文字画像を用いて得られた評価結果は、筆者の一部により実劣化データ適用時でも有効であることを確認しており [6]、本論文でも提案手法の評価に用いる。認識辞書の学習データには、劣化のない学習データのみを用いた。

また、実験条件としての各パラメータは以下のとおりとした。正規化後のパターンサイズ W は、 64×64 画素とした。したがって、ノイズ種類判定用特徴ベクトルの次元数は、 $128 (= 64 \text{ 画素} \times 2 \text{ 方向})$ 次元となる。認識用特徴ベクトルの次元数は、 $256 (= 8 \times 8 \text{ 局所領域} \times 4 \text{ 方向})$ 次元である。また、ランレングス観測用の方形窓の大きさ I は、方形窓内に平均して約 1

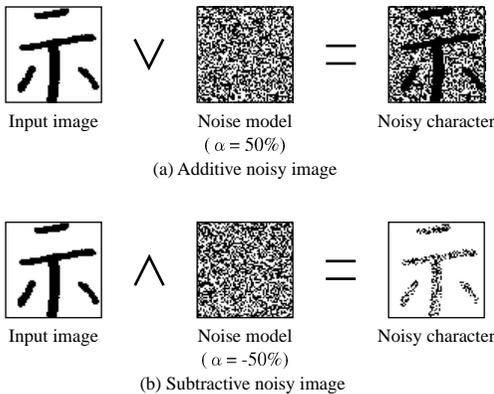


図 5 汚れノイズモデル画像、かすれノイズモデル画像、及び各劣化文字画像の生成例
Fig. 5 Additive noise model, subtractive noise model, and each noisy character image.

本の文字線が含まれる (b, c が約 1 となる) 大きさとした。その理由は、着目した文字線のランレングスを記述するには、その文字線全体 (1 本) が含まれる大きさが必要であること、及び複数の文字線が含まれると、どの文字線のランレングスを記述しているかあいまいになるからである。学習データにおいて横方向の方形窓を用い調べた結果、 $I = 15$ 画素 (着目画素より ± 7 画素) となった。したがって、横方向及び縦方向に $I = 15$ 、各斜め方向に $I = 11$ (着目画素より ± 5 画素) と設定した。そのときの各パラメータの平均値 $\bar{a}, \bar{b}, \bar{c}, \bar{e}$ は、各々 7.4, 0.9, 0.9, 4.8 である。識別関数には、ノイズ種類判定処理及び認識処理ともに、ユークリッド距離を用いた。

5.2 ノイズ種類判定の実験結果

まず、ノイズ種類の判定性能を評価した。劣化程度には、上述のように $-70 \leq \alpha \leq 70$ の 15 段階を用いた。各 α は、 $-70 \sim -10$ がかすれ、0 が劣化のない状態、 $10 \sim 70$ が汚れを表す。学習用データから、劣化程度別に 15 パターンずつ、計 3,036 カテゴリー $\times 100$ サンプル $\times 15$ 劣化程度の劣化文字画像を生成した。また、評価用劣化文字画像も同様に評価用データから作成した。ただし、ノイズモデル画像を生成するための乱数の seed は、学習用データ及び評価用データでは別の値を用いた。

表 1 に、劣化程度別の入力文字画像に対するノイズ種類判定率を示す。 $-10 \leq \alpha \leq 0$ を除いた全劣化程度において、99.8% 以上という高い判定ができています。劣化程度が $\alpha = 0$ の場合に 4.1% がかすれに、 $\alpha = -10$ では 5.9% が汚れに誤判定された。誤判定の理由には、単純な構造をもつ文字がかすれによりやや複雑さを増したものと、元来複雑な構造をもつ文字との間で複雑さの差が小さくなったことが考えられる。

5.3 認識実験結果

次に、提案した補正ランレングスに基づく方向寄与度特徴と、従来の観測ランレングスに基づく方向寄与度特徴 (ノイズ除去処理無し、及び 3×3 画素のメジアンフィルタを使用) を比較した。各劣化程度 α の評

表 1 ノイズ種類の判定率 [%]
Table 1 Noise type detection accuracy [%].

Output	Input														
	Subtractive							Additive							
	-70	-60	-50	-40	-30	-20	-10	0	10	20	30	40	50	60	70
Subtractive	99.9	99.9	99.9	99.9	99.9	99.8	94.1	4.1	0.01	0	0	0	0	0	0
Additive	0.05	0.07	0.05	0.03	0.06	0.2	5.9	95.9	99.9	100	100	100	100	100	100

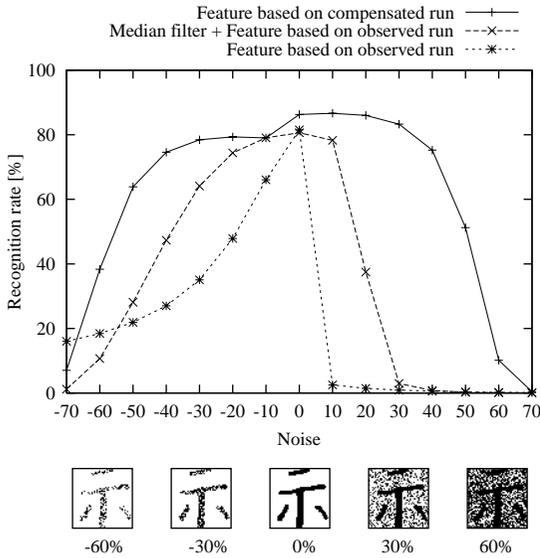


図 6 各特徴の劣化文字画像に対する認識率

Fig. 6 Recognition rates for patterns with subtractive or additive noise.

評価データに対する両特徴の認識率を、図 6 に示す。補正ランレングスに基づく特徴は、 $-60 \leq \alpha \leq 70$ という劣化のほぼ全範囲において、観測ランレングスに基づく特徴より優れた結果を得ている。汚れノイズが生じた場合、従来の観測ランレングスに基づく特徴は、本来特徴値が 0 であるべき背景部より、0 以上の特徴値が抽出される。したがって、特徴ベクトルが平均ベクトルより大きくずれ、誤認識となりやすい。しかし、提案手法ではランレングス補正により複数の背景部でランレングス及び特徴値が 0 となり、ノイズの影響が抑制され正しく認識できている。したがって、 $\alpha \geq 10$ の汚れノイズに対し、補正ランレングスに基づく特徴と観測ランレングスに基づく特徴では、認識結果に大きな差が生じている。一方、かすれノイズが生じた場合、背景部の特徴値は 0 のままで、文字部の特徴値が変動するだけである。したがって、観測ランレングスにおいても、かすれにおける認識率の低下は汚れの場合より穏やかとなり、補正ランレングスを用いた特徴との認識結果の差も小さい。

また、補正ランレングスに基づく特徴において、背景ノイズの大きさが認識性能に及ぼす影響について調べた。評価データとして、3. での劣化文字画像の作成において、反転させる画素の大きさを $1 \times 1 \sim 5 \times 5$ に変化させたノイズモデル画像を用い、劣化文字画像を

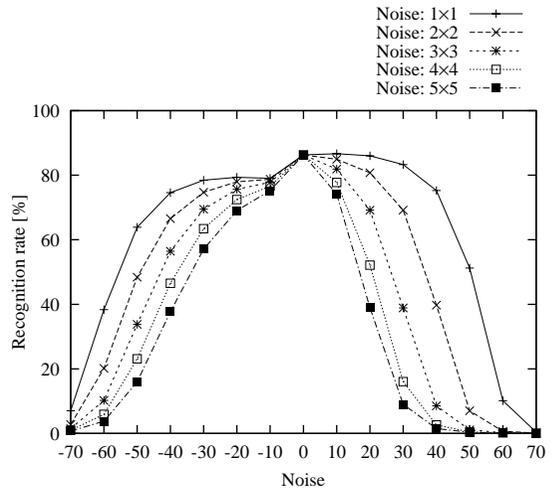


図 7 各大きさの背景ノイズをもつ劣化文字画像に対する認識率

Fig. 7 Recognition rates for patterns with each size of subtractive/additive noise.

作成した。図 7 に、各大きさのノイズに対する認識率の変化を示す。背景ノイズが大きくなるにつれ、徐々に補正が困難になり、認識率が低下している。この結果は、画素の分布状態を観測することのみでは、本質的にノイズの影響を抑制することが困難であることを示唆している。例えば、 5×5 画素の大きさは、ほぼ文字線の太さと同じであり、ランレングスの長さのみでは、文字線かノイズかを判別できない。この問題に対しては、他の有効な情報の導入など、より柔軟に対処する必要がある。

5.4 考察

図 6 より、補正ランレングスを用いた特徴では、劣化程度が大きい ($\alpha \leq -60$ 及び $\alpha \geq 60$) 場合、急激に認識率が低下する。原因は、汚れまたはかすれの程度が大きい場合、補正処理での仮定とは逆に複雑さが減少するため、ランレングスを正しく補正できないからである。特にかすれの $\alpha \leq -60$ では、黒画素数を 2 倍以上に増加する必要がある。しかし、式 (21) では最大でも観測黒画素数を 2 倍にしかできず、補正に限界がある。ただし、上記の劣化状態では人でも認識は容易でなく、実用上問題となることは少ないと思われる。

表 1 において、主に $-10 \leq \alpha \leq 0$ でノイズ種類の誤判定が生じているが、この誤判定に伴う認識率低下の影響を調べるため、ノイズ種類の判定が 100% 正し

いと仮定した場合の認識率を調べた。 $\alpha = -10$ において、0.1%の差が認められたが、 $\alpha = 0$ を含むその他の劣化程度での認識率の差はほぼ0%に近く、 $\alpha = -10$ 及び0における誤判定の影響はそれほど大きくないことがわかった。その理由には、 α が小さい(劣化が少ない)ため、誤った増減方向に補正が行われても絶対的な補正量が少なく、劣化がない場合に得られるランレングス値との差が小さいこと、及び5.3における図6の結果に対する考察で述べたように、かすれの場合は汚れの場合に比べて悪影響が小さいことが挙げられる。

提案手法により、従来では誤読であったものが正読になった例を、図8に示す。複雑な文字においても、提案手法が適切にノイズを抑制したことにより、正しく認識できたことがわかる。また、提案手法を用いても誤読であった例を、図9に示す。図9(a)ではノイズを抑制しきれず、類似文字に誤読している。図9(b)では、逆にノイズと同時に文字線のランレングスも、過剰に抑制している。

ランレングス補正法は、文字線のランレングスに基づく他の特徴にも、そのまま適用可能である。また、提案手法は識別関数の補正による画質劣化への対処[17]~[19]とは異なり、画質劣化の補正処理が特徴抽出処理内に閉じている。したがって、後段の識別関数の選択に制約がなく、より高性能な識別関数を用いることにより、更に高い認識率を得ることができる。

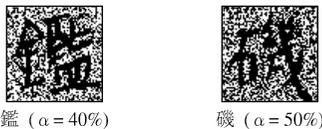


図8 提案手法により正しく認識された文字画像例
Fig. 8 Examples of images recognized correctly by the proposed method.

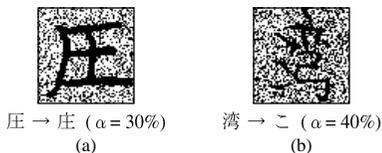


図9 提案手法でも誤認識された文字画像例(正解カテゴリ → 誤認識結果)
Fig. 9 Examples of images recognized erroneously by the proposed method (Correct category → Erroneous category as the recognition result).

以上の結果から、提案したランレングス補正法を用いた特徴は、従来では認識困難であった手書き等の変形が生じた文字において、汚れやかすれが生じた場合の認識にも有効であるといえる。しかし、5.3の図7の結果に対して述べたように、本論文の提案手法で対処可能な劣化は、ガウスノイズ及びそれに類似したノイズであり、提案手法だけで現実には生じる多くの画質劣化に対処できるわけではない。今後、他の様々な画質劣化にも対処できるよう拡張する必要がある。

6. む す び

本論文では、画質劣化と変形の両品質低下にロバストな特徴抽出法を検討した。画質劣化が生じた文字画像から文字線の方向情報を抽出するため、文字線のランレングスを劣化状態の推定結果に応じて補正する“ランレングス補正法”を提案し、ランレングスに基づく特徴である方向寄与度特徴に適用した。ETL-9を人工的に劣化させた文字画像を用いた認識実験において、提案手法を用いた特徴は、汚れ及びかすれにおける幅広い劣化程度において、従来の観測ランレングスに基づいた特徴より高い認識率を得た。この結果から、提案手法により、特徴抽出において変形に対するロバスト性を保持しつつ、画質劣化に対してもロバスト性が向上可能であることが明らかになった。

今後の課題には、潰れなど他種の画質劣化への対処、及び実環境の画質劣化文字画像に対する評価がある。

謝辞 ETL-9を提供して頂いた産総研(旧電総研)の皆様、日ごろ御世話頂く石井健一郎所長、議論頂いた研究グループのメンバー、及び有益な御助言をして頂いた査読者に感謝致します。

文 献

- [1] G.E. Kopec, “Supervised template estimation for document image decoding,” IEEE Trans. Pattern Anal. Mach. Intell., vol.19, no.12, pp.1313-1324, Dec. 1997.
- [2] Y. Xu and G. Nagy, “Prototype extraction and adaptive OCR,” IEEE Trans. Pattern Anal. Mach. Intell., vol.21, no.12, pp.1280-1296, Dec. 1999.
- [3] T.K. Ho, “Bootstrapping text recognition from stop words,” Proc. 14th ICPR, vol.1, pp.605-609, Brisbane, Australia, Aug. 1998.
- [4] M. Sawaki, H. Murase, and N. Hagita, “Automatic acquisition of context-based images templates for degraded character recognition in scene images,” Proc. 15th ICPR, pp.15-18, Barcelona, Spain, Sept. 2000.
- [5] 飯島泰蔵, パターン認識理論, 森北出版, 東京, 1989.
- [6] M. Sawaki and N. Hagita, “Text-line extraction and

- character recognition of document headlines with graphical designs using complementary similarity measure,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.20, no.10, pp.1103–1109, Oct. 1998.
- [7] A. Sato, “A learning method for definite canonicalization based on minimum classification error,” *Proc. 15th ICPR*, vol.2, pp.199–202, Barcelona, Spain, Sept. 2000.
- [8] M. Umeda, “Advances in recognition methods for handwritten Kanji characters,” *IEICE Trans. Inf. & Syst.*, vol.E79-D, no.5, pp.401–410, May 1996.
- [9] 萩田紀博, 内藤誠一郎, 増田 功, “外郭方向寄与度特徴による手書き漢字の識別,” *信学論 (D)*, vol.J66-D, no.10, pp.1185–1192, Oct. 1983.
- [10] J. Zhu, T. Hong, and J.J. Hull, “Image-based keyword recognition in oriental language document images,” *Pattern Recognit.*, vol.30, no.8, pp.1293–1300, Aug. 1997.
- [11] S.N. Srihari, T. Hong, and G. Srikantan, “Machine-printed Japanese document recognition,” *Pattern Recognit.*, vol.30, no.8, pp.1301–1313, Aug. 1997.
- [12] H. Ozawa and T. Nakagawa, “A character image enhancement method from characters with various background image,” *Proc. 2nd ICDAR*, pp.58–61, Oct. 1993.
- [13] S. Liang, M. Ahmadi, and M. Shridhar, “A morphological approach to text string extraction from regular periodic overlapping text/background images,” *CVGIP*, vol.56, no.5, pp.402–413, Sept. 1994.
- [14] M.Y. Yoon, S.W. Lee, and J.S. Kim, “Faxed image restoration using Kalman filtering,” *Proc. 3rd ICDAR*, vol.2, pp.677–680, Montreal, Canada, Aug. 1995.
- [15] Z. Shi and V. Govindaraju, “Character image enhancement by selective region-growing,” *Pattern Recognit. Lett.*, vol.17, no.5, pp.523–527, May 1996.
- [16] 斎藤泰一, 山田博三, 山本和彦, “JIS 第 1 水準手書漢字データベース ETL9 とその解析,” *信学論 (D)*, vol.J68-D, no.4, pp.757–764, April 1985.
- [17] 大町真一郎, 阿曾弘具, “つぶれを考慮した低品質印刷文字の高精度認識,” *信学論 (D-II)*, vol.J79-D-II, no.9, pp.1534–1542, Sept. 1996.
- [18] S. Omachi, M. Inoue, and H. Aso, “A noise-adaptive discriminant function and its application to blurred machine-printed Kanji recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.22, no.3, pp.314–319, March 2000.
- [19] 森 稔, 倉掛正治, 杉村利明, 塩 昭夫, 鈴木 章, “背景・文字の形状特徴と動的修正識別関数を用いた映像中テロップ文字認識,” *信学論 (D-II)*, vol.J83-D-II, no.7, pp.1658–1666, July 2000.
- (平成 14 年 10 月 4 日受付, 15 年 3 月 4 日再受付)



森 稔 (正員)

平 5 東工大・工・電気電子卒。同年日本電信電話(株)入社。以来,文字認識・文書処理の研究・開発に従事。現在,同社 NTT コミュニケーション科学基礎研究所メディア情報研究部研究主任。平 11 本会学術奨励賞受賞。



澤木美奈子 (正員)

平元慶応大・工・電気卒。同年日本電信電話(株)入社。以来,文字認識の研究に従事。現在,同社 NTT 先端技術総合研究所企画部担当課長。平 9 本会学術奨励賞受賞。平 13 より本会編集委員。映像情報メディア学会会員。



萩田 紀博 (正員)

昭 51 慶応大・工・電気卒。昭 53 同大大学院修士課程了。同年日本電信電話公社(現 NTT)入社。以来,主に文字認識・画像認識の研究に従事。平 13 年 10 月より,ATR メディア情報科学研究所長。昭 56 本会論文賞受賞。工博。情報処理学会,人工知能学会,IEEE 各会員。



村瀬 洋 (正員)

昭 53 名大・工・電子卒。昭 55 同大大学院修士課程了。同年日本電信電話公社(現 NTT)入社。以来,文字・図形認識,コンピュータビジョン,マルチメディア認識の研究に従事。平 4 から 1 年間米国コロンビア大客員研究員。現在,日本電信電話(株)NTT コミュニケーション科学基礎研究所メディア情報研究部長。昭 60 本会学術奨励賞,平 4 電気通信普及財団テレコムシステム技術賞,平 6 CVPR 最優秀論文賞,平 7 情報処理学会山下記念研究賞,平 8 ICRA 最優秀ビデオ賞,平 13 高柳記念電子科学技術振興財団高柳記念奨励賞,本会情報システムソサイエティ論文賞,平 14 本会業績賞受賞。工博。情報処理学会,IEEE 各会員。



武川 直樹 (正員)

昭 49 早大・理工・電子通信卒。昭 51 同大大学院修士課程了。同年日本電信電話公社(現 NTT)入社。画像認識,画像符号化,及びコンピュータビジョンの研究・開発に従事。平 6~12(株)NTT データ。平 12 より日本電信電話(株)NTT コミュニケーション科学基礎研究所環境理解研究グループリーダー。工博。情報処理学会,IEEE 各会員。