

## 品質を考慮した印刷文字認識手法

大町 真一郎<sup>†</sup> 阿 曾 弘 具<sup>†</sup>

印刷文字認識において、認識率を下げる原因は読み取られた文字の品質によって異なると考えられる。具体的には、つぶれのある低品質の文字では文字パターンの形状が変わってしまい、通常の文字パターンから特徴量を抽出して評価値を求める認識手法では誤認識を起こしやすくなる。一方、高品質文字の認識の場合の最大の難点としては、類似文字の識別が考えられる。品質の悪い文字の場合は文字画像そのものを用い、つぶれの程度によって決まる重みを用いた手法によって認識率が改善されるが、つぶれの少ない高品質の文字に適用した場合は認識率改善の効果は薄く、時間の浪費につながる。また、高品質文字では部分パターンを用いた手法で類似文字識別が可能になるが、つぶれの多い文字ではその文字本来の特徴が得られず、効果的に働かない。このように異なる品質の文字を同様の手法で扱うことは非常に困難であるが、現実には1つの実用的な認識システムに対して、異なる品質の文字でも効率的かつ高精度に認識を行うことが要求されている。本論文では以上の問題を解決するため、読み取られた文字の品質を判断して品質に応じた手法を適用することで活字文字を効率的かつ高精度に認識する新しいシステムを提案し、実験によりその有効性を示す。

## Recognition of Printed Characters by Considering Quality

SHIN'ICHIRO OMACHI<sup>†</sup> and HIROTOMO ASO<sup>†</sup>

In printed character recognition, the factors of the decline in recognition rate are different due to the quality of images. In this paper, an efficient algorithm to recognize printed characters by selecting the optimum method according to the quality of each character is proposed. In character recognition, usually feature vector is extracted from an input pattern and evaluation value is calculated. But, for poor quality characters, since the reliable feature vector can not be extracted from those blurred patterns, traditional methods are not efficient. For characters of good quality, discrimination of similar characters is the most difficult problem. In this paper, we propose a new method for poor quality character recognition by using pattern itself with weighting factors determined by degree of blur. Also, to discriminate similar characters, a method using partial image is proposed. The proposed new recognition system first distinguishes the quality of an unknown pattern, and applies the optimum method according to the quality of the pattern. The experimental results show the proposed system is practical and useful with high recognition rate.

### 1. はじめに

最近、様々な文書作成ツールが開発され、多様な印刷文書がより簡単に作成できるようになった。パソコンでもマルチフォント・マルチサイズの印刷文書が比較的容易に作成でき、1ページの中に大きなタイトル文字から小さなルビや脚注などの文字まで、多種の文字が混在する文書が数多く出回るようになった。本文に使われる文字と比べて、ルビや脚注などの文字は極端に小さく低品質になりやすい。また、複写機により何度も複写した文書やファクシミリによる転送(量子化)で得られる文書はもとの文書に比べて大変劣化し

ているが、これらの文書もできる限り正しく読みたいという要求がある。したがって、明瞭に印刷された高品質の文字からつぶれの多い低品質の文字まであらゆる文字に対して有効な認識技術が実用的な文字認識装置には不可欠である。

しかし、印刷文字認識においては読み取られた文字の品質によって認識率を下げる要因が異なるため、どのような品質の文字にも効率的かつ高精度な認識システムの開発は困難である。低品質の文字、特につぶれのある文字画像では文字のパターンが本来のものと大きく変わってしまい、通常の文字パターンから特徴量を抽出して評価値を求める認識手法では誤認識を起こしやすくなる。高品質の文字は低品質の文字と比べて認識率は高いが、類似文字において誤認識することが多く、さらに認識率を上げるためには類似文字間の識

<sup>†</sup> 東北大学大学院工学研究科  
Graduate School of Engineering, Tohoku University

別が最も重要な課題になる。

つぶれのある画像の場合にも、1文字のイメージの全領域がつぶれていることは少なく、つぶれた部分のみが全体に悪影響を及ぼすことが多い。著者らは、文字の各部分ごとにつぶれの程度（つぶれ度）を定義し、それによって信頼度を求め、整合性の評価において信頼度に応じた重み付けを行うことで低品質文字を高精度に認識する手法を提案した<sup>1)</sup>。本論文の低品質文字の認識部では、さらにこの重みを単純類似度法に適用し、重み付きの単純類似度とすることで低品質の文字の認識率が向上することを確かめる。一方、高品質の文字の認識における一番の難点は類似文字の高精度識別である。類似文字として、日本語の場合平仮名と片仮名の清音・濁音・半濁音の文字が典型的である。本論文の高品質文字の認識部では、高品質文字を対象とし、部分パターンを用いた類似文字識別法および類似文字か否かの判定、類似文字識別のための領域を選出する方法を提案する。

このように高品質文字と低品質文字ではその特徴に違いがあり、認識時に品質に対する考慮が不可欠となる。どちらか一方に有効な手法であっても、他のもう一方の品質の文字が認識対象となった場合に無駄な計算時間がかかるうえに、認識率の低下も避けられないと思われる。本論文では、読み取られた文字の品質を判断して品質に応じた手法を適用することで活字文字を効率的かつ高精度に認識する新しいシステムを提案し、実験によりその有効性を示す。

なお、本論文の実験で用いたデータはすべてノイズ除去・平滑化の機能が組み込まれている光学式イメージスキャナで読み込んで二値化した。読み取り精度は400 dpiとした。

## 2. 文字認識アルゴリズム

### 2.1 特徴量とアルゴリズムの流れ

本手法においては、文字認識のための特徴量として方向線素特徴量<sup>2)</sup>を用いる。方向線素特徴量は以下のように求める。

まず、入力された文字画像に対して大きさの正規化・細線化・線素化処理を施す。正規化の大きさは64ドット×64ドットとした。また、線素化とは細線化した後のそれぞれの黒画素に縦（|）、横（—）、右上斜め（/）、左上斜め（\）の4種類の線素を割り当てる処理である。細線化は細め回数を制限して行い、線幅が1にならない部分は輪郭線を線素化する。入力画像と線素化画像の例を図1に示す。そして、線素化後の画像を8×8の小領域に分割し、図2の太線で示す重

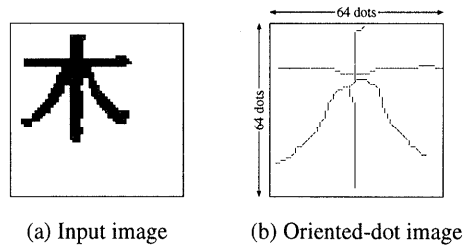


図1 文字画像の例  
Fig. 1 Character image.

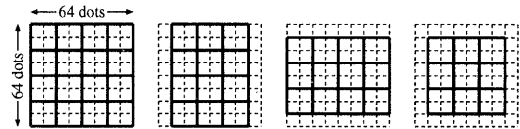


図2 49の領域  
Fig. 2 49 regions.

なりのある16ドット×16ドットの領域ごと（計49領域）に、その領域に含まれる各線素を中心部に大きい重みをおいてガウスフィルタ的な重み付きで計数する。これにより得られる4×49=196次元のベクトルを特徴量とする。候補選出時には、あらかじめ学習パターンから作成しておいた認識対象となる各々の文字の平均ベクトル（辞書）とこのベクトルとのユークリッド距離を評価値とし、その値の小さい順に候補文字とする。この辞書のことを本論文では便宜上「ベクトル辞書」と呼ぶ。

方向線素特徴量は活字認識に有効な特徴量として提案され、分割統治的な手法（連想整合法）を用いることで高速処理が可能である<sup>3)</sup>。ただしつぶれのある低品質の文字画像では、細線化することによりパターンが大きく変わり、本来の特徴が失われてしまうため誤認識が生じやすい特徴量でもある。この性質は、現在文字認識において良い成果が得られている、細線化や輪郭線抽出を行った後方向線分を抽出することにより求める特徴量<sup>4)~6)</sup>について一般にいえることである。この点は低品質文字認識用の特徴量としては大変不利である。しかし、細線化することは線の太さによらない特徴量であることを意味し、また、文字の正規化で小さな文字の線幅が太くなくても細線化で細くなり文字の大きさによらないという特徴を持つ。さらに、細線化後に線素化処理を行うことで、線分の途切れ・ボイドの発生などパターンの変形をともなうかすれに対する影響が少ない。すなわち、高品質の文字に対しては、細線化を行うことは文字の変形を吸収するための有効な手段であるといえる。提案する手法は以上のこ

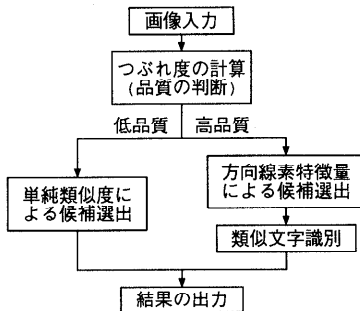


図3 文字認識アルゴリズム

Fig. 3 Algorithm for character recognition.

とに着目し、低品質文字であれば方向線素特徴量を用いる代わりに文字画像そのものを用いて重み付きの単純類似度により候補選出を行い、高品質文字であれば方向線素特徴量を用いた識別を大分類としてさらに詳細識別（類似文字識別）を行うものである。提案する文字認識のアルゴリズムを図3に示す。

## 2.2 品質の判断

印刷文字認識において、認識率を下げる原因は読み取られた文字の品質によって異なっている。したがって、どのような品質の文字にも有効な認識システムを開発するためには、まず文字の品質を判断し、高品質の文字と低品質の文字のそれぞれに最適な認識手法を適用する必要がある。

本論文では、品質を判断するために文字画像の品質を表す尺度であるつぶれ度<sup>1)</sup>を用いる。文字パターンの小領域ごとのつぶれを表す尺度であるつぶれ度は、入力パターンを正規化・細線化した画像から求める。細線化の方法としては、Hilditchの方法<sup>7)</sup>とほぼ同様の結果が得られる10近傍法<sup>8)</sup>を用いた。これは、線幅が2ドット以上のときそれを輪郭から1ドットずつ削っていくものであり、通常の処理では最終的に線幅が1のパターンを得る。ここでは、印刷文字が本来一定の幅の線で構成されていることに注目し、削る回数に上限（以下「細線化回数」と呼ぶ）を設け、線幅が1にならずに残った部分をつぶれていると見なす。適当な細線化回数 $\star$ を設定しておき、49の各小領域ごとに残った黒画素のうち輪郭線を除く黒画素数を求め、 $\lfloor (\text{黒画素数})/32 \rfloor$ の値をつぶれ度とする。各小領域は $16 \times 16 = 256$ ドットの大きさであるのでつぶれ度は0から8までの値をとり、0のときはほとんどつぶれがなく、値が大きくなるに従ってつぶれが大きくなる。

さらに、49領域のつぶれ度の平均を平均つぶれ度と

定義し、文字画像全体の品質を表す尺度とする。平均つぶれ度は実際の文字認識実験に適用することで文字の品質の判断基準として有効であることが確認されている<sup>1)</sup>。未知入力パターンの平均つぶれ度を求め、その値がある閾値 $\theta_1$ 以上であれば低品質の文字と見なし、 $\theta_1$ より小さければ高品質の文字と見なす。

## 3. 低品質文字の高精度認識法

つぶれのある低品質文字の場合、文字画像から特徴量を求めるのではなく文字画像そのものを用い、単純類似度法<sup>9)</sup>により認識する手法がパターンの変形に強い手法として知られている。しかし、複写機による複写やファクシミリで転送することにより生じるつぶれは偶発的であり、その偶発性ゆえにあらかじめ十分な学習パターンを準備して辞書を作成することが難しい。本論文では、文字パターンの各部分領域ごとのつぶれ度を利用し、つぶれの影響を軽減するための動的重み付けを導入する。

低品質文字の認識には、つぶれ度による重みを用い、正規化画像を用いた単純類似度を重み付きで計算するように拡張し（以下「重み付き単純類似度」と呼ぶ）、これを用いる。辞書としては大きさの正規化をほどこした学習パターンの同じ字種をすべて重ね合わせ、黒画素数の平均をとったもの（以下「イメージ辞書」と呼ぶ）を用いた。字種 $k$ の辞書パターンを $g_k$ とし、未知入力パターンを正規化した二値画像パターンを $f$ とすると、字種 $k$ の辞書パターンとの間の評価値 $S_k$ を、

$$S_k = \frac{\left( \sum_{x,y} w(x,y) f(x,y) g_k(x,y) \right)^2}{\sum_{x,y} w(x,y) f(x,y)^2 \sum_{x,y} w(x,y) g_k(x,y)^2} \quad (1)$$

と定義し、 $S_k$ の値が大きい順に候補文字とする。ここで、 $w(x,y)$ がつぶれ度による重みである。図2から分かるように49領域は重なりがあり、画像上のある1点 $(x,y)$ に対応する重みは最大4領域分存在するので、これらの平均値を $w(x,y)$ とする。

## 4. 高品質文字の詳細識別法

### 4.1 認識の流れ

入力画像が高品質であると判断された場合、まず、方向線素特徴量を用い、ユークリッド距離を評価値として認識を行う。ここではこの処理を大分類として位置付け、大分類の結果得られた1位候補と2位候補が

$\star$  認識実験では細線化回数は6回とした（予備実験により定めた）。

類似文字かどうかを調べる。これらが類似文字であれば再認識処理（詳細識別）を行い、類似文字でなければ大分類の1位候補をそのまま最終候補として出力する。3位以下の候補についても考慮する場合は、1位候補と2位候補に対してこれらの処理を行った後、改めて新しい1位候補と3位候補、4位候補…、に対して同様の処理を行う。

詳細識別は、類似文字を識別するための領域（以下、「識別領域」と呼ぶ）のみを用いてイメージ自体の整合性を調べる。以下、4.2節では類似文字であるか否かの判定と識別領域の決定法について、4.3節では識別領域における整合性の評価法について述べる。

#### 4.2 類似文字の判定と識別領域の自動設定

2つの類似文字は局所的にのみ異なっており、その異なっている部分に識別のための情報があり、識別するのに適した領域は文字によって異なる。本節では、適切な領域を設定する手法について検討する。

まず、ここでいう「類似文字」の定義を明確にする。類似文字とは、

- 誤認識の可能性のある文字の組、
- 全体の形が似ているが一部が異なる。つまり、

部分的な比較により識別が可能、

であるものとする。したがって、「ば」と「𠂔」、 「王」と「玉」などは類似文字であるが、「や」と「𠂔」、 「へ」（平仮名）と「𠂔」（片仮名）などは同形の文字であって類似文字ではない（正規化を行うので大きさの違いによる識別はできない）。

適切な識別領域の設定法として、2つの文字の異なる部分のみを抽出し、同形の部分は抽出しないというのは1つの選択である。この考えに基づき、文字*i*と文字*j*が類似文字であるか否かを判定し、類似文字である場合にそれらの識別領域を同時に求める方法を考案した。まず、ベクトル辞書から*i*、*j*の辞書パターンを取り出し、49の各領域ごとに*i*と*j*の間の距離を計算する ( $d_1(i, j) \sim d_{49}(i, j)$  とする)。 *i*と*j*が類似文字で一部のみ異なる場合、 $d_1(i, j) \sim d_{49}(i, j)$ のうち異なる部分のみ値が大きく、それ以外は値が小さくなる。 *i*と*j*が同形の文字の場合は全領域で値が小さくなる。それ以外の場合は、 $d_1(i, j) \sim d_{49}(i, j)$ は様々な値をとる。実際にベクトル辞書を用いて、「か」と「が」（類似文字）、「あ」と「い」（異なる文字）、「や」と「𠂔」（同形の文字）のそれぞれについて、49の各領域ごとに距離を求め、大きい順にソートした結果を図4に示す。上で述べた特徴が現れていることが分かる。

そこで、 $d_1(i, j) \sim d_{49}(i, j)$ の平均を  $m(i, j)$  とし、

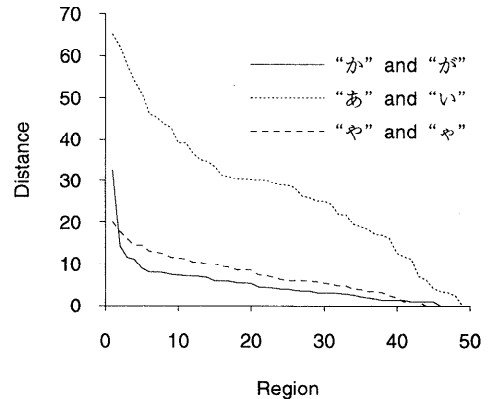


図4 領域ごとの距離の分布

Fig. 4 Distribution of distance for each region.

閾値  $\theta_2$  を用い、次の集合  $E(i, j)$  を考える。

$$E(i, j) = \{k | d_k(i, j) > \theta_2 \cdot m(i, j)\} \quad (2)$$

$E(i, j) = \phi$  のとき *i* と *j* は類似文字ではないものとする。  $E(i, j) \neq \phi$  のとき *i* と *j* が類似文字であると判断し、 $E(i, j)$  を識別領域の集合として詳細識別を行う。なお、予備実験より  $\theta_2 = 1.9$  が適切であることを確認した。

#### 4.3 整合性の評価

類似文字識別を行うため、大分類に用いる辞書のほかに詳細識別用の辞書を用意しておく。詳細識別用の辞書としては、3章で用いた「イメージ辞書」を利用する。整合性の評価を行うときはイメージ辞書を用い、識別領域のみを用いてイメージ自体の整合性を調べる。

まず、各識別領域について未知入力文字と辞書パターンとの整合性を評価値（以下「相違量」と呼ぶ）として求める。そして、識別領域として用いたすべての領域について相違量の和をとったものを最終的な評価値とする。

各識別領域の相違量は次のようにして求める。未知入力文字の識別領域は方向線素特徴量を求める際に分割された16ドット×16ドットの領域のうちの数領域である。辞書パターンの方は、各識別領域の場所を中心とする32ドット×32ドットの領域を対象とし、その中で最もよく未知パターンと整合する16ドット×16ドットの部分パターンを選び、そのパターンとの評価値を相違量とする。例を図5に示す。図で、未知パターン(a)に対する1位候補として「ぶ」が、2位候補として「ふ」が得られたとする。これらの識別領域は(a)の網かけをした領域である。(b)と(c)が1位候補と2位候補の辞書パターン（イメージ辞書）であり、各位置の点の濃度は値の大きさに比例している（最大値は1）。(b)と(c)のそれぞれについて、網か

けをした領域の中で未知パターンの識別領域と最もよく整合する部分パターンを求めて相違量を計算する。相違量のより具体的な求め方を次に示す。16ドット×16ドット未知パターンを

$$f(x, y) \quad (0 \leq x \leq 15, 0 \leq y \leq 15) \quad (3)$$

とし、32ドット×32ドット辞書パターンを

$$g(x, y) \quad (0 \leq x \leq 31, 0 \leq y \leq 31) \quad (4)$$

とする。ここで、 $f(x, y)$  の値は0 (白画素) または1 (黒画素) であり、 $g(x, y)$  は0から1までの任意の値をとりうる。辞書パターンの

$$\alpha \leq x \leq \alpha + 15, \quad \beta \leq y \leq \beta + 15 \quad (5)$$

の領域に対して、辞書パターンを  $k$  倍したイメージと未知パターンとの二乗誤差が最小となるように  $k$  を定めたときの二乗誤差の値を  $e(\alpha, \beta)$  とする。つまり、

$$e(\alpha, \beta) = \min_k \sum_{\substack{0 \leq x \leq 15 \\ 0 \leq y \leq 15}} \{f(x, y) - k \cdot g(x + \alpha, y + \beta)\}^2 \quad (6)$$

とする。式(6)の右辺を展開して二乗誤差が最小となる  $k$  を解析的に求め、これを式(6)に代入すれば次の式が得られる。

$$e(\alpha, \beta) = \{1 - S(\alpha, \beta)\} \times \sum_{x, y} f(x, y)^2 \quad (7)$$

$S(\alpha, \beta)$  は単純類似度<sup>9)</sup>で、この場合は

$$S(\alpha, \beta) = \frac{\left( \sum_{x, y} \{f(x, y)g(x + \alpha, y + \beta)\} \right)^2}{\sum_{x, y} f(x, y)^2 \sum_{x, y} g(x + \alpha, y + \beta)^2} \quad (8)$$

である。ただし、式(8)の分母が0でないときに式(7)が成立する。式(8)の分母のうち、 $\sum_{x, y} g(x + \alpha, y + \beta)^2 = 0$  のときは、 $S(\alpha, \beta) = 0$  として式(7)が成立し、 $\sum_{x, y} f(x, y)^2 = 0$  のときは式(6)で  $k = 1$  とおき、

$$e(\alpha, \beta) = \sum_{x, y} \{g(x + \alpha, y + \beta)\}^2 \quad (9)$$

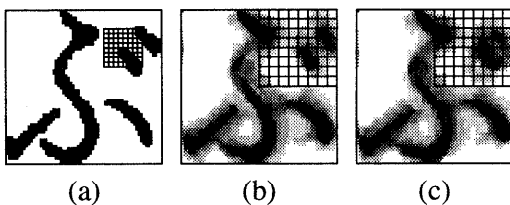


図5 評価値を求める領域

Fig. 5 Region for calculating evaluation value.

とする。

この識別領域の相違量  $e$  を、

$$e = \min_{\substack{0 \leq \alpha \leq 16 \\ 0 \leq \beta \leq 16}} e(\alpha, \beta) \quad (10)$$

と定める。大分類における1位候補と2位候補の評価値をそれぞれ  $e_1, e_2$  とし、識別領域のみに限定して求めた相違量を  $e'_1, e'_2$  とすると、

$$\frac{e'_1}{e'_2} > \frac{e_2}{e_1} \quad (11)$$

のとき1位候補と2位候補を交換する。

## 5. 認識実験

提案手法の有効性を確かめるために認識実験を行った。認識対象字種として、JIS第1水準の漢字2,965字種と、平仮名と片仮名144字種(人間でも区別できない大文字の「あ」と小文字の「あ」などは一方を除いた)、計3,109字種を用いた。学習サンプル(イメージ辞書・ベクトル辞書とも)を、レーザプリンタで出力した表1に示す5ポイントから25ポイントまでの11種類の明朝体とゴシック体とした。1種類あたり1文字を用いたので、学習文字数は1字種あたり11文字である。

学習サンプルとして複写機による複写やファクシミリによる転送で得られるつぶれた文字を用いることにより、ある程度つぶれに強い辞書を作成することも可能であると考えられる。しかし、つぶれは偶発的に発生することが多く、可能なすべてのつぶれに対する学習サンプルを集めることは不可能であり、学習サンプルに依存した辞書となって他のつぶれに対処しきれないと思われる。さらに、このようなつぶれた画像を学習サンプルとして用いることで、逆に明瞭な文字画像の認識に悪影響を与えることが考えられる。本論文は1つの辞書であらゆる品質の文字画像を高精度に認識する手法を提案しているため、プリンタから出力したそのままのきれいなデータを学習サンプルとし、つぶれた文字の認識にはつぶれ度を用いた補正を行うという立場をとる。人間が文字を覚えるときにもきれいな正確なパターンを最初に学習し、つぶれはつぶれとして認識していると考えられる。本提案手法は、このことに着目した手法ともいえる。

表1 学習サンプルとして用いたデータ

Table 1 Data for learning.

フォント	ポイント
明朝体	5, 6, 8, 10, 12, 14, 20, 25
ゴシック体	6, 12, 25

表2 各データの平均つぶれ度

Table 2 Average degree of blur for each data.

フォント	平均つぶれ度
10ポイント明朝体	0.005
6ポイント明朝体	0.514
6ポイントゴシック体	1.061

認識対象の入力データとしては、5.1, 5.2, 5.3 各節のすべての実験において、品質の比較的良好な10ポイント明朝体、中程度の6ポイント明朝体、品質の悪い6ポイントゴシック体の3種類のフォントをレーザープリンタで学習サンプルとは別に出力したものを利用した。

これらのサンプルの品質を評価した。各データのすべての文字画像について平均つぶれ度を求め、その平均値を計算したものを表2に示す。6ポイントの文字は比較的品质が悪く、特にゴシック体の方は文字を構成する線が太いこともあってつぶれの多い低品質のパターンとなっている。

なお、本章では、方向線素特徴量を用いてユークリッド距離を評価値として候補を選出する手法のことを従来法と呼ぶことにする。

### 5.1 重み付き単純類似度法の有効性

重み付き単純類似度法の有効性を確認するため、重み付き単純類似度法、重みなしの単純類似度法、従来法の3種類について評価実験を行った。

結果を図6に示す。図より、6ポイントのデータについては従来法よりも重み付き単純類似度法の方がはるかに良い結果が得られており、10ポイントのデータについては同じ認識率が得られた。すなわち、品質の悪い小さな文字については提案手法を用いた方が良いが、比較的大きな品質の良い文字については提案手法よりもはるかに高速な従来法で十分な認識率が得られることが分かる。また、単純類似度法との比較でも、より品質の悪い6ポイントゴシック体では重み付きの方が良い結果が得られ、品質が中程度の6ポイント明朝体では若干重みなしの方が良い結果が得られた。これらのことから、提案した重み付き単純類似度法は特に品質の悪い文字に対して有効に働くことが確認された。

### 5.2 相違量の有効性

相違量を用いた詳細識別の有効性を確認するための評価実験を行った。ただし、上位3位までの候補を対象とした。実験は、提案する手法と従来法のほか、相違量そのものの有効性を確認するために相違量の代わりに単純類似度  $S(\alpha, \beta)$  を用いて詳細識別をする実験も行った。単純類似度はよく整合しているほど値が大

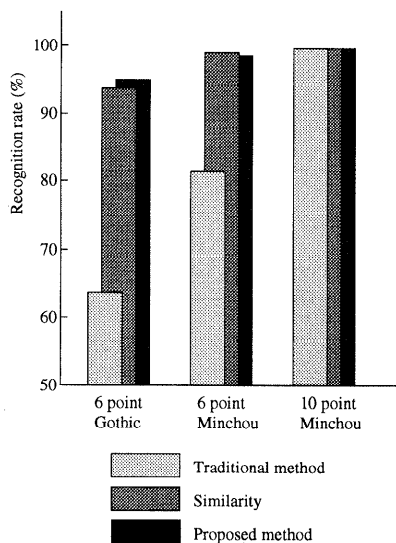


図6 重み付き単純類似度法の性能評価実験結果

Fig. 6 Experimental results for evaluating Similarity with weighting factor.

きくなるので、式(10)の代わりに

$$e = \max_{\alpha, \beta} S(\alpha, \beta) \quad (12)$$

と定義して領域ごとの評価値とし、式(11)の代わりに

$$\frac{e'_2}{e'_1} > \frac{e_2}{e_1} \quad (13)$$

が満たされたときに1位候補と2位候補を交換する。これは、ずらし類似度<sup>10)</sup>の考え方を局所的なパターンに適用し、その値を評価値として類似文字識別を行うものである。

結果を図7に示す。図より、いずれのデータでも従来法と比較して提案した手法を用いた方が良い結果が得られている。また、単純類似度を用いた詳細識別では、詳細識別を行わない場合よりもむしろ認識率が低下していることが確認された。

また、最も認識率の高い10ポイント明朝体について、従来法では誤認識されていたが相違量による詳細識別を行うことにより正しく認識されるようになった字種の従来法における1位候補から3位候補までを表3に示す。逆に詳細識別を行うことにより誤認識に転じる字種は存在しなかった。また、この場合の認識対象となった文字画像を図8に示す。図中で、網かけをした領域が1位候補と2位候補の詳細識別時に識別領域として用いられた。表3と図8から、1位候補と2位候補の文字を識別するための適当な領域が4.2節で述べた手法により抽出されていることが分かる。

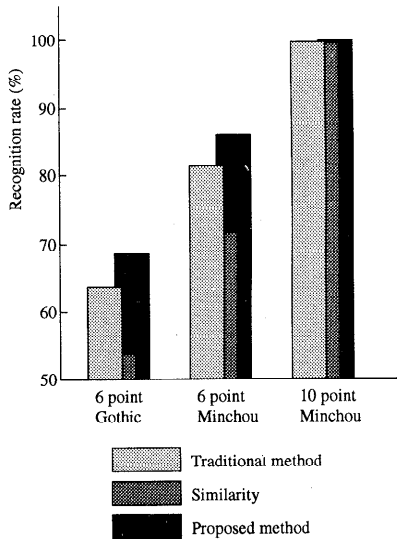


図7 相違量の性能評価実験結果

Fig. 7 Experimental results for evaluating amount of difference.

表3 詳細識別により正しく認識された字種

Table 3 Correctly recognized characters by fine classification.

正解	従来法による候補		
	1位	2位	3位
窟	筍	窟	窟
卵	卵	卵	叩
鱗	鱗	鱗	鱗

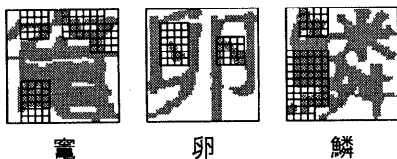


図8 文字画像と識別領域

Fig. 8 Character images and discriminating regions.

### 5.3 総合評価

以上の結果をふまえ、提案手法の有効性を確認する実験を行った。本手法で重要な役割を持つ品質の判断については、予備実験より平均つぶれ度の値が0.2以上であるか否かを基準とした。すなわち、平均つぶれ度が0.2より小さければ方向線素特徴量を用いてユークリッド距離を評価値として候補を選出し、その後詳細識別を行う。0.2以上であれば低品質文字であるとして重み付き単純類似度で識別を行う。

また、比較のため、従来法のほか、つねに重み付き単純類似度を用いる手法、重みなしの単純類似度を用

表4 認識実験の結果

Table 4 Experimental results.

手法	認識率	平均処理時間
本手法	97.9%	5.8 秒/文字
重み付き単純類似度法	97.7%	7.2 秒/文字
単純類似度法	97.5%	6.3 秒/文字
従来法 + 詳細識別	84.9%	4.9 秒/文字
従来法	81.7%	1.9 秒/文字

表5 本手法により正しく認識された字種

Table 5 Correctly recognized characters by proposed method.

正解	重み付き 単純類似度法	従来法 + 詳細識別	本手法	平均つ ぶれ度
ば	(ば)	ば	ば	0.00
バ	(バ)	バ	バ	0.00
ブ	(ブ)	ブ	ブ	0.00
竿	(芋)	竿	竿	0.00
土	(上)	土	土	0.00
侮	(海)	侮	侮	0.00
鶯	鶯	(鶯)	鶯	0.45
鷹	鷹	(鷹)	鷹	0.39
膚	膚	(膚)	膚	0.71

( ) は誤認識であることを表す。

いる手法、従来法で大分類を行った後つねに相違量を用いた詳細識別を行う手法(「従来法 + 詳細識別」と表記)についても実験を行った。平均認識率と SPARC Station 20 を用いた場合の平均処理時間を表4に示す。なお、実験にはC言語を使用した。表より、本手法は他のどの手法よりも高い認識率が得られていることが分かる。また、処理時間についても、認識率が2番目に高い重み付き単純類似度法の約8割程度で済むことが確認された。

また、10ポイント明朝体について、重み付き単純類似度法や従来法 + 詳細識別では誤認識されていたが本手法を用いることにより正しく認識されるようになった字種、およびその入力画像の平均つぶれ度を表5に示す。なお、本手法による10ポイント明朝体の認識率は100%であった。また、表5において、従来法 + 詳細識別では誤認識されていたが本手法を用いることにより正解となった字種「鶯」「鷹」「膚」の入力画像を正規化したものを図9に示す。図中、網かけをした領域はつぶれ度が1以上の領域を表す。10ポイントの文字の画像でもこのようにつぶれがひどい領域が存在するが、つぶれていない部分の情報と全体の概形が保たれていたために本手法により正しく認識された。

表より、重み付き単純類似度法では線の密度が比較的低いためつぶれにくい類似文字が存在する字種が誤認識されていることが分かる。本手法を用いること

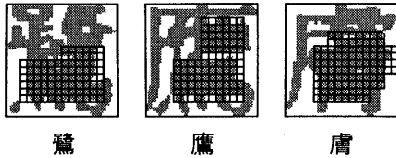


図9 文字画像とつぶれ度が1以上の領域

Fig. 9 Character images and regions whose degree of blur is greater than 0.

表6 誤認識された文字の例

Table 6 Examples of missed characters.

		(a)	(b)	(c)	(d)	(e)
入力画像		ば	ぶ	駅	堰	襖
平均つぶれ度		0.02	0.00	0.86	1.80	1.37
正解		ば	ぶ	駅	堰	襖
候補	1位	ば	ぶ	駅	堰	換
	2位	ば	ぶ	駅	堰	襖
	3位	ぼ	ぶ	銀	塩	携

により、これらは高品質の文字と判断され、方向線素特徴量を用いて正しく認識されるようになった。一方、従来法 + 詳細識別で誤認識されたのは、線の密度が比較的高く、つぶれやすい字種であることが分かる。これらの字種はつぶれにより本来の方向線素特徴量が抽出されず、文字パターンそのものを用いる重み付き単純類似度法が効果的に働いたものと考えられる。以上の結果から、文字認識において品質を判断することの重要性が確認された。

表6には、6ポイント明朝体を対象とした場合に、本手法を用いても誤認識となったJISコード順で最初の5文字の文字画像の例を示す。平均つぶれ度、正解の字種、得られた候補文字(1位から3位まで)を示してある。表において、(a)と(b)は平均つぶれ度が0.2未満であるため、方向線素特徴量を用いてユークリッド距離を評価値として候補を選出し、その後詳細識別を行った。これらは正解文字と1位候補文字が類似文字であり、相違量を求めるための識別領域がつぶれてしまい、相違量による詳細識別が有効に働かなかったために誤ってしまったものと考えられる。これらはいずれも2位候補として正解の文字が得られている。また、(c)~(e)は平均つぶれ度が0.2以上であり、重み付き単純類似度法で識別を行ったものである。表中の画像から分かるように、これらの文字の概形はつぶれにより1位候補文字の概形とほぼ同じになっている。重み付き単純類似度法では概形が重要な判断基準となるため、正しく認識できなかったものと考えられる。なお、これらも2位候補としては正解の文字が得られている。

このように、識別に有効な構造情報がつぶれにより失われてしまった場合は、人間でもその認識は難しい。しかし、1位候補が正解でなかった場合にも、上位の候補文字として正解の文字が得られれば、言語情報を用いた後処理と組み合わせることで識別が可能になると期待される。そのような後処理と組み合わせ、実際の認識システムを構築することは今後の課題である。

## 6. おわりに

本論文では、様々な品質の文字が混在する文書を高精度に認識することを目的とし、品質を判断して品質に応じて手法を変える新しいアルゴリズムを提案した。まず、低品質文字認識のために、文字画像そのものを用い、つぶれ度による動的重み付けを行った単純類似度法を用いる手法を提案した。そして、これが特につぶれのある低品質文字の認識に有効であることを確認した。また、高品質文字の認識率を上げるためには類似文字の詳細識別が不可欠であると考え、類似文字間の異なる領域のみを参照して詳細識別を行う手法を示した。これは、類似文字を高精度に識別するための新たな評価値(相違量)を用い、類似文字を識別するための領域(識別領域)を自動的に定め、従来法による大分類と組み合わせることで類似文字識別を行うものである。そして、これが詳細識別に有効であることを確認した。

さらに、品質を判断して様々な品質の文字を高精度に認識する手法を提案した。品質の判断には、繰返し回数を制限した細線化を行うことにより入力画像から文字のつぶれを表す尺度であるつぶれ度を求め、これを用いた。そして、品質の異なる3種類の文字パターンを用いた認識実験により提案手法の有効性を確認した。

謝辞 日頃からご指導いただく北陸先端科学技術大学院大学木村正行教授、東北大学大学院情報科学研究科丸岡章教授に感謝します。

## 参考文献

- 1) 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, 電子情報通信学会論文誌(D-II), Vol.J79-D-II, No.9, pp.1534-1542 (1996).
- 2) 孫寧, 田原透, 阿曾弘具, 木村正行: 方向線素特徴量を用いた高精度文字認識, 電子情報通信学会論文誌(D-II), Vol.J74-D-II, No.3, pp.330-339 (1991).
- 3) 孫寧, 阿曾弘具, 木村正行: 連想整合法に基づく高速文字認識アルゴリズム, 情報処理学会論



文誌, Vol.32, No.3, pp.404-413 (1991).

- 4) 江島俊朗, 中村洋介, 木村正行: 構造情報を含む手書き漢字認識のための特徴量, 電子情報通信学会論文誌 (D), Vol.J68-D, No.4, pp.789-796 (1985).
- 5) 若林哲史, 鶴岡信治, 木村文隆, 三宅康二: 特徴量の次元数増加による手書き数字認識の高精度化, 電子情報通信学会論文誌 (D-II), Vol.J77-D-II, No.10, pp.2046-2053 (1994).
- 6) 孫 寧, 安倍正人, 根元義章: 改良型方向線素特徴量および部分空間法を用いた高精度な手書き文字認識システム, 電子情報通信学会論文誌 (D-II), Vol.J78-D-II, No.6, pp.922-930 (1995).
- 7) Hilditch, C.J.: Linear Skeleton from Square Cupboards, *Machine Intelligence*, Vol.6, pp.403-420, Edinburgh University Press, Edinburgh (1969).
- 8) 阿曾弘具: 並列処理向き細線化アルゴリズム, 電子情報通信学会論文誌 (D-II), Vol.J76-D-II, No.9, pp.2148-2150 (1993).
- 9) 飯島泰蔵: パターン認識理論, 森北出版 (1989).
- 10) 山田博三, 齊藤泰一, 森 俊二: 類似度法の一改良—ずらし類似度—, 電子情報通信学会論文誌 (D), Vol.J64-D, No.10, pp.970-975 (1981).

(平成9年3月21日受付)

(平成9年9月10日採録)



大町真一郎 (正会員)

1988年東北大学工学部情報工学科卒業. 1993年同大学院博士課程修了. 同年同大情報処理教育センター助手. 1996年同大工学部助手. 1997年同大学院工学研究科助手, 現在に至る. 博士(工学). パターン認識, 並列処理, 高速高精度文字認識システムの開発等の研究に従事. IEEE, 電子情報通信学会, 人工知能学会各会員.



阿曾 弘具 (正会員)

1968年東北大学工学部電気工学科卒業. 1974年同大学院博士課程修了. 1973年同大工学部助手. 1979年名古屋大学工学部講師. 1982年同大助教授, 1986年東北大学工学部助教授, 1991年同大教授を経て, 現在, 同大学院工学研究科電気・通信工学専攻教授. 工学博士. その間, 学習オートマトン, セル構造オートマトン, シストリックアルゴリズム設計理論, 文字認識, 文書認識システムの開発, 音声認識などの研究に従事. 平成3年度電子情報通信学会業績賞受賞. IEEE, ACM, EATCS, 電子情報通信学会, 人工知能学会, 日本認知科学会, LA 各会員.