2 段階文字認識と後処理を融合した 手書き住所文字列の認識

正員 梅田 三千雄 (大阪電通大)
非会員 濱 裕治郎 (大阪電通大)

Recognition of Hand-Written Character Strings of Mail Addresses by Fusing Two-Stage Character Recognition and Post-Processing

Michio Umeda, Member (Osaka Electoro-Communication University)

Yuujiro Hama, Non-member (Osaka Electoro-Communication University)

In this paper, we first recognize individual character in the prefecture/city part of mail addresses, followed by inferring the entire prefecture/city parts. Then we extend the character category set and recognize entire mail addresses. As pre-processing steps, correction of the characters' stroke widths and orientations is also carried out. This method attempts to improve the recognition rate for individual character as well as the recognition rate for mail address character strings. By using the proposed method, a rate of 85.4% of correctly recognized addresses was achieved in experiments utilizing data distributed by the Research Institute of the Ministry of Postal Affairs.

キーワード：手書き住所文字列認識，アフィン変換，2 段階文字認識，後処理，地名辞書

1. はじめに

近年，文字認識の研究分野で文字列や文章を対象とするものが盛んになりつつある。手書きの住所文字列を認識する研究は，その代表的な研究の 1 つである。しかし，手書きの文字は，活字文字に比べて，文字の大きさが均一でなく，個々の文字が文字間で接触していたり，続け文字が存在したりする可能性があるため，それぞれの文字を一意に切り出すことが難しい。また，葉書に書かれている住所文字列には，ペン，毛筆など様々な筆記具が用いられており，筆記時の制約もない。このような条件下では，文字の変形が大きくなる。その上，略字，崩し字，つぶれ文字，かすれ文字が含まれているため，個別文字認識を高精度に実現できないなどの問題がある。

個別文字認識率の向上をめざして，住所文字列の画像データに含まれる文字をそのまま文字認識の標準パターン作成に利用する方法もある⁽¹⁾⁽²⁾⁽³⁾。しかし，手書き住所文字列のような文字では，変形が大きく，高精度に認識できない。さらに，文字列認識の後処理法として主流となっているのは，文字位置に複数通りの可能性を持たせて文字認識し，後処理の段階で文字位置を確定する候補文字ラティス法⁽⁴⁾である。この手法では，個別文字認識の信頼度が低いとき，単語検索であらゆる可能性を求めていこうとすると，単語辞

書の全検索や候補文字の組合せの大幅な増加がおこる。この問題を解決するため，文字単位で後処理を行う文字タグ法⁽⁵⁾が提案されている。また，後処理に使用する地名辞書の住所件数は，全国規模の辞書であっても文献 (1) では約 9000 件，文献 (2)(3) では約 2000 件と全般的に小規模なものが用いられており，実用レベルにはほど遠いのが現状である。

本論文では，手書き住所文字列を高精度に認識するために，まず，個々の変形の大きい文字の認識を試みる。提案されている文字認識手法に加えて，文字列中に存在する変形の大きい文字に対応するため，文字認識の前処理の段階で文字線の太さと文字の傾きを推定し，推定値に応じて文字パターンに変形を加え画一化を施す。これにより，個別文字認識率の向上を図る。次に，住所文字列認識における後処理について検討する。手書きの漢字認識では，認識対象とする文字種が多いことも問題となる。そこで，文字認識と後処理を融合した住所文字列認識を提案する。

住所に使われている文字の種類を調べると，行政区域名部分，すなわち，都道府県名，市名，郡名，町村名，区名の住所に使われている文字種は，地域名を含めた住所全体に使用される文字種に比べて半数以下である。このことに着目し，言語情報を用いて個別文字認識の認識対象文字種を削減する。初めに，行政区域名部分の推定を目的として，

行政区域名部分の住所の認識に必要な文字種のみを認識対象として個別文字認識する。その後、行政区域名部分の推定結果より、認識対象とする字種を増加させて再び個別文字認識し、地域名部分を認識する。これにより、個別文字認識率の向上をめざし、住所文字列の認識精度の向上を試みる。後処理法としては、候補文字ラティス法を用いる。まず、行政区域名部分を検索し、検索の終了した時点で、認識結果としての住所文字列の候補を仮に評価し、住所候補文字列数の削減を行う。その後、残った候補の下に存在する地域名のみを対象にして検索する。単語検索の途中で住所候補の削減を行い、検索対象を削減することで、単語検索の効率化を図る。住所の表記には、途中の地名を省略したものや、市内、都内などのあいまいな表記もある。本手法は、このような表記にも対応可能な認識法となっている。

2. 手書き住所文字列認識システム

手書き住所文字列認識の処理の流れを図1に示す。まず、この認識システムの処理の概要について述べる。

前処理部では、住所文字列の画像データの2値化と雑音除去を施す。次に、文字切り出しを行う。文字切り出しには、行切り出しと、1つ1つの文字を取り出す文字切り出しが含まれる。この後、初期の個別文字認識を行う。その文字認識結果を用いて、都名、道名、府名、県名、市名、郡名、区名、町名、村名の行政区域名部分の検索をする。ここで、行政区域名部分までの住所の候補文字列が100件以上得られたとき、候補以下に続く地域部分の検索量を削減するために、これらの候補を仮に評価し、住所候補数を約100個に絞る。仮評価の終わった段階で、2度目の個別文字認識を行う。この個別文字認識は、仮評価によって残った住所候補の下部に続く地域名部分を推定するために行う。次に、2度目の文字認識結果を用いて、地域名部分の検索を行う。地域名部分の検索によって複数の住所候補を得たとき、

図 1. 処理の流れ
Fig.1. Stream of processing.

これらの住所候補を評価し、最も確からしい住所候補1つに絞り込む。これを住所文字列の認識結果とする。

本研究では、個別文字認識の標準パターンの作成には、電子技術総合研究所の作成した手書き漢字データベース ETL9B を、文字認識後処理の地名辞書には、全国規模の約 17 万件の住所件数を有するものを使用した。

また今回、後処理の段階で使用した地名辞書において、JIS 第 1 水準内の漢字および平仮名の出現する文字種数を調べた。その結果、行政区域名部分までで 936 字種、住所全体で 2143 字種出現することがわかった。この結果より、初期の個別文字認識では、行政区域名部分を認識するのに必要な 936 字種をカテゴリとして認識し、2 度目の個別文字認識は、行政区域名部分の検索結果に応じて、候補以下の地域部分を認識するのに必要な文字種を追加して行う。以後、この 2 度に分けて実行する個別文字認識を 2 段階文字認識と呼ぶことにする。それぞれの個別文字認識では、文字認識の特徴として方向線素特徴⁽⁶⁾を用いる。初めに、100 次元の方向線素特徴を用いて大分類し、次いで、分類結果上位 30 位を 452 次元の特徴を用いて細分類する。個々の処理については、これ以降で説明をする。

3. 文字切り出し

文字切り出しの例を図2に示す。文字切り出しでは、まず、行の検出を行う。行の検出は、住所文字列データを構成する黒画素連結成分を縦方向に見たときの黒画素連結成分の分離状態により、行数と各部がどの行に含まれるかを検出する。次に、個々の文字の切り出しを行う。これは、検出された行ごとに処理する。黒画素の連結成分を囲む外接矩形の形状や、外接矩形の位置関係を用い、これらを統合または分離することで文字を切り出す。統合、分離は、基本的に日本語の文字が正方形に近いと考えて、繰り返し実行する。住所に含まれる文字は、大きさが均一でなく、文字が分離、接触している場合もある。

図 2. 文字切り出し例
Fig.2. Example of character segmentation.

従って、それぞれの文字に対する切り出しを1つに定めるのは困難である。そこで、外接矩形を統合、分離して求めた基本矩形を1つ、または幾つか組み合わせた複数の切り出し候補矩形を求め⁽⁴⁾、切り出し正解率の向上を図る。縦横比が正方形に近くない基本矩形は、その前後の矩形と合わせて1字をなす可能性があり、基本矩形どうしが接触しているときは、その矩形全部で1字をなすことが多い。予備実験により、基本矩形を組み合わせたときの縦横比が以下の条件を満たすとき、基本矩形を統合して切り出し候補矩形とした。

- 統合した後の矩形の縦横比が1.7:1.0未満
- 統合した後の矩形の縦横比が1.7:1.0以上、2.0:1.0以下であり、次のどれかを満たすとき
 - － 統合する矩形の中に縦横比が1.2:1.0以上、または0.65:1.0以下のものが含まれる。
 - － 統合する矩形の平均幅の半分以下の幅の矩形が存在する。
 - － 統合する矩形がすべて重なっている(接触している)。
- 上記以外で、統合した後の矩形の縦横比が3.0:1.0以下で、かつすべての矩形が重なっている(接触している)。

4. 個別文字認識

文字認識処理は、前処理、特徴抽出、大分類、細分類によって構成される。まず、前処理として文字パターンの大きさの正規化、文字の太さと傾きの補正⁽⁷⁾、文字の形状を整える非線形正規化⁽⁸⁾、スムージングをする。次に、この処理により得られた文字パターンから文字の特徴を抽出する。まず、低次元特徴を用いた大分類により、高速に候補となるカテゴリをある程度の数に絞り込む。その後、この絞り込まれたカテゴリのみを対象にして、高次元特徴を用いた細分類をする。文字認識の特徴には、方向線素特徴を用いた。方向線素特徴とは、文字線の輪郭部の方向性に着目した特徴である。抽出方法は、図3に示すように、文字パターンの輪郭部の任意の点における形状が図に示すマスクパターンに一致するときのみ線素化を行う。次に、それぞれの方向別に画像領域を分割し、その領域内での線素数を数えて、これを特徴とする。領域の分割の仕方により次元数を変更する。識別関数には、大分類、細分類それぞれにシティブロック距離、重み付きシティブロック距離を用い、大分類した結果上位30位の文字候補を細分類する。重み付きシティブロック距離を、

$$D_i = \sum_{j=1}^n | \max(0, |q_j - p_{ij}| - a_{ij}) | \quad (1)$$

$$\max(x, y) = \begin{cases} x & (x > y) \\ y & (x \leq y) \end{cases}$$

q_j : 入力パターン n : 特徴の次元数
 p_{ij} : 標準パターン i : 文字種
 a_{ij} : 標準パターンの標準偏差

と定義する。

図 3. 方向線素特徴

Fig.3. Directional element feature.

< 4・1 > 文字の太さと傾きの推定 手書き住所文字列は種々の筆記具で記述されているので、文字線の太さは様々で、ETL9 Bのような標準的な文字に比べて、右上がり文字も多く存在する。各々の文字を画一化するため、非線形正規化を施す前に、入力パターンに対して、文字の太さと傾きを推定し、文字の太さと傾きの補正という文字の形状変換を施す。文字の太さの推定は、縦横96画素に大きさを正規化したパターンに対して行う。推定方法は、次に示すような手法を用いた。

1. 2値画像の距離変換⁽⁹⁾により、文字内部のすべての点において白画素からの距離値を求める。
2. 距離変換した入力画像のすべての距離値の和を黒画素数で割った値を文字の太さとする。

次に、以下に示す処理により文字の傾きを推定し、右上がり文字を見つけ出す。傾きの推定も同様に、大きさを正規化したパターンに対して行う。

1. 入力パターンの重心より左半分、右半分それぞれで重心を求める。
2. 左半分と右半分の重心の高さの差を文字の傾きとする。

< 4・2 > 文字の太さと傾きの補正 この後に、文字の太さと傾きの推定値により、文字の太さと傾きを補正する。文字の太さ C_1 未満の文字画像に対して1回膨張処理、文字の太さ C_2 以上の文字に1回収縮処理、太さ C_3 以上の文字には、もう一度収縮処理をする。予備実験により、 C_1 、 C_2 、 C_3 は各々1.5, 3.0, 4.0と定めた。傾きの補正では、文字

の傾きが3以上のときに、右上がり文字として、傾き補正を行った。右上がり文字と推定された入力パターンに

$$\begin{aligned} X &= x \\ Y &= 0.02 \times (g_l - g_r) \times (x - 48) + y \end{aligned} \tag{2}$$

g_l ：左半分の重心 y 座標
 g_r ：右半分の重心 y 座標
(X, Y)：補正後の黒画素の座標
(x, y)：補正前の黒画素の座標

のアフィン変換をして傾き補正する。文字の太さと傾きの補正例を図4に示す。

図 4. 文字の太さと傾きの補正例
Fig.4. Example of correction of characters' stroke widths and slants.

5. 後処理

後処理の流れを図5に示す。地名辞書を用いて地名推定を行う場合、個別文字認識の精度が低いと、すべての文字が誤っている可能性があるものとして、言語処理を認識結果全体に適用する必要がある。処理を大きくせざるを得ない。また、処理を大きくし候補地名が多くなると、誤った結果を導いてしまう可能性が高くなる。そこで、比較的単語数の少ない行政区域名部分をまず推定し、その結果に応じて推定された住所の下につながる地域名単語のみを検索する。これにより、処理量を削減でき、かつ認識率の向上が期待できる。さらに、行政区域名部分の住所を推定するためと地域名部分を推定するための個別文字認識で、それぞれ対象文字種を限定することで、認識率の向上を図る。

<5.1> 地名辞書構成 住所は、階層的な構造をしている。階層ごとに区切られた住所の要素すべてを総称して単語と呼ぶ。後処理で利用する辞書は47都道府県を含み、登録されている単語件数は、168550件である。行政区域名の単語数は、4017件であり、地域名の単語数は、164533件である。市内、都内などのあいまいな住所表記に対応するため、地名辞書には、市内、都内などの単語も含む。これらの単語は、以後、行政区域名単語として扱う。辞書は、図6

図 5. 後処理の流れ
Fig.5. Stream of post-processing.

単語番号	単語の綴り	階層	上位階層の単語番号
⋮	⋮	⋮	⋮
98423	大阪府	府	
98424	大阪市	市	98423
98425	都島区	区	98423, 98424
98426	網島町	地域	98423, 98424, 98425
⋮	⋮	⋮	⋮
100795	寝屋川市	市	98423
⋮	⋮	⋮	⋮

図 6. 地名辞書
Fig.6. Address dictionary.

に示すように、単語が1つの単位として構成される。これらは、単語の綴りとともに、固有の単語番号と上位階層の単語番号を持つことにより、辞書中での階層構造が実現されている。

<5.2> 行政区域名部分の単語検索 住所は地域名部分を除くと、多くとも3つの階層以内で構成されている。後処理では、この知識を利用する。住所認識で重要な文字である都、道、府、県、市、郡、区、町、村、内の10種類の文字をキー文字と呼ぶ。図7に示すように、キー文字を含む単語で、単語からキー文字を除いた部分を地名要素、キー文字を含まない地域名部分の単語を地域要素と呼び、住所文字列画像の先頭、基本矩形と基本矩形の間、画像の下部をそれぞれノードと呼ぶことにする。検索は、ノード間ごとに行う。あいまいな住所記述に対応するために、住所の階層を途中省略している住所文字列も認める。以下に後処理の手順を示す。

1. まず、住所画像の先頭部分から、辞書中のすべての行政区域名単語と認識結果を比較し、条件に合えば住所候補として保存する。
2. 次に、その候補の下部に存在する行政区域名の単語を地名辞書から探し出す。探し出した単語と住所候補の下に続くところの認識結果を比較し、条件に合えば単語候補とする。
3. 住所候補と単語候補を組合わせて住所候補とし、もとの住所候補とは別に保存する。
4. 住所文字列の地域名部分を除くと、日本の住所は3階層以内で構成されているので、この下部についても画像の最下部まで検索していなければ、階層が3階層になっていないかぎり2, 3を繰り返す。同じノード間で、同じ単語が検索された場合、最も平均個別認識順位のよいものを残し、その他は削除する。住所候補以下に単語候補が検索されなかった場合は、最下部まで検索済みと、その住所候補に記録する。

このようにして、行政区域名部分の単語検索を行い、住所候補を作成していく。単語検索で住所候補、単語候補にする条件を以下に述べる。

図 7. 住所の階層

Fig.7. Hierarchy in addresses.

< 5.3 > 単語検索条件 文字種を 936 字種に限定して個別に文字認識した初期の認識結果に基づいて単語検索を行う。認識候補の第 300 位までに含まれない文字候補を非一致文字とする。比較している単語の文字長と認識結果の比較対象部分の文字長が一致し、以下の条件のどちらかを満たす場合、単語候補として保存する。

- 単語の地名要素の文字長が 2 文字以上のとき、一致文字中にその地名要素を構成する文字が対応する位置で 1 文字でも存在し、かつ単語のキー文字の認識順位が 200 位以内の場合
- 単語の地名要素の文字長が 1 文字のとき、その地名要素を構成する文字が非一致文字でなく、かつ単語

のキー文字の認識順位が 200 位以内の場合

この条件で、キー文字の認識順位の制限のみが条件を満たさなかった場合、単語の地名要素の平均認識順位が 30 位以内であれば、候補として保存する。地名要素が 1 文字のときは、単語検索の条件を緩和している。また、住所候補として、都内、市内などが得られたときも他の住所候補が得られたときと同様に、その候補の下に存在する可能性のある行政区域名の単語すべてを辞書から探し出して認識結果と比較する。

< 5.4 > 住所候補の仮評価 行政区域名部分の地名の検索を終えた時点で、住所候補の仮評価を行う。住所候補が 100 個以上存在する場合、仮評価により評価の高い住所候補上位 100 個に限定し、その住所候補を含む住所候補に絞り込む。評価としては、評価式

$$R = \frac{n1 + n2 \times 100 + n3 \times 200 + n4 \times 100}{n5} \quad (3)$$

$n1$: 一致文字数

$n2$: カテゴリ内で一致文字、非一致文字でない文字数

$n3$: 非一致文字数

$n4$: カテゴリ外の文字数

$n5$: 住所候補を構成する文字数

の値が小さい住所候補を優先する。R の値は、住所候補の文字中に一致文字の占める割合が多ければ、多いほど小さくなる。また、非一致文字に対しては、他の文字に比べて大きな重みを付けた。これにより、住所候補中に非一致文字の占める割合が多いほど R の値は大きくなる。この R の値が同一の住所候補が出現し、候補が 1 つに定まらない場合は、各文字を認識したときの順位の平均値が小さいものの順に高い評価を与える。

< 5.5 > 2 度目の個別文字認識 住所候補の仮評価により残った約 100 個の住所候補の下部の地域名を推定するため、2 度目の個別文字認識をする。初期の個別文字認識のカテゴリである 936 字種に地域名部分を認識する上で足りない文字種をカテゴリに追加し、すべての切り出し候補矩形を再び認識する。以下にその手順を説明する。

1. 残った住所候補の下部に存在する地域名を地名辞書より調べる。
2. その地域名を構成する文字種で初期の文字種である 936 字種以外の文字を調べる。
3. 行政区域名部分を認識するのに使用した 936 字種に 2 で調べた字種を加えた文字種をカテゴリとし、すべての切り出し候補矩形の認識を行う。

< 5.6 > 地域名部分の単語検索 住所候補の仮評価を終えた後、残った住所候補の地域名部分の単語検索を行う。仮評価により残った候補の下部に存在する単語を地名辞書から探し出す。探し出した単語と住所候補の下に続くところの認識結果を比較し、条件に合えば単語候補とする。

住所候補と単語候補を組合わせて住所候補とする。地名の文字長と認識結果の文字長が一致し、以下の条件を満たす場合、住所候補として保存する。

- 地域要素の文字長が2文字以上のとき、一致文字中にその地域名を構成する文字が対応する位置で1文字でも存在する場合
- 地域要素の文字長が1文字のとき、その地域名を構成する文字が非一致文字でない場合

画像の最下部まで検索できたものを住所候補として保存する。最下部まで到達していない住所候補は棄却する。地域名部分の検索時間の短縮のため、都道府県の階層からの地域名部分に直接つながる住所文字列は認めない。また、キーワード内に続く地域名部分の検索は、対象とする住所候補の行政区域名部分の文字認識順位が、すべて第10位候補以内に含まれるものについてのみ検索した。

<5・7> 住所候補の絞り込み 求められた複数の住所候補から1つの住所候補に絞り込む。複数の住所候補の個々を評価し、最もよい評価を持つ住所候補を結果とする。住所候補の評価としては、住所候補の仮評価のときと同様に、式(3)の値が最も小さいものを優先し、1つに定まらない場合は、各文字を認識したときの順位の平均値により1つに絞り込む。これにより、認識結果としての住所文字列が得られる。なお、住所候補が1つも検索されなかった場合、このときの住所の推論は不可能であると判定して棄却する。

6. 後処理例

図8の例題をもとに、このアルゴリズムを用いた後処理例を示す。図には、文字認識結果として、一部の切り出し候補矩形に対し第4位候補までを記す。また四角で囲まれた文字は、正解文字中の一致文字である。以下で正解に至るまでの過程を述べる。まず、切り出し候補矩形すべてをカテゴリ936字種で認識し、画像の先頭より、この認識結果と辞書中の行政区域名部分のすべての地名と比較する。その結果、住所候補として摂津市、安芸市、久喜市、大分県などを得る。次に、それぞれの住所候補に対して、その下に存在する行政区域名部分の地名と比較する。大分県以外の住所候補の下部には、行政区域名部分の地名が存在しなかったため、単語候補は抽出されない。大分県の場合、中津江村、上津江村が単語候補として抽出された。それまでの住所候補と抽出した単語を組み合わせて新たに住所候補とする。この場合、大分県上津江村、大分県中津江村が新たに住所候補として加わった。その後、住所候補の下部について検索を行ない、検索すべき住所候補がなくなった時点で行政区域名部分の単語検索を終える。

行政区域名部分の単語検索を終えた時点で、住所候補を一度仮に評価し、住所候補を約100個にする。これにより、

521個あった住所候補が100個になった。ここで残った住所候補の下部に存在する地域名単語を調べる。そして、その地域名を構成する文字で初めの936字種以外の字種を調べ、これを追加して、再び切り出し候補矩形のすべてを認識しなおす。認識対象とする字種数は、1227字種となった。この後、地域名部分の推定を行う。その結果、安芸市大井乙、久喜市太田袋、摂津市新在家、大分県上津江村川原などの住所候補が出力された。評価した結果、住所文字列として安芸市大井乙を得る。2段階文字認識をせずにカテゴリを固定して認識した場合、安芸市は住所候補として検索されず正解を得られなかった。

図 8. 後処理例
Fig.8. Example of post-processing.

7. 結果と考察

実験には、郵政省の第3回文字認識技術コンテストで配

布されたデータ 1471 件を使用した。これらのデータには、手書き年賀状の縦書き宛名住所のパターンが収められており、筆記具は、サインペン、ボールペン、毛筆などが使用されている。これらは、宛名部分の横 45mm × 縦 105mm の一定領域を 400dpi、濃淡 256 階調で量子化してある。評価実験として、以下の実験を行なった。ここでは、JIS 第 1 水準漢字と平仮名の文字種の中で住所を構成する文字である 2143 字種を文字認識のカテゴリとする。

実験 1：正しく文字切り出しできた割合を求める。文字切り出し時に文字位置を確定していないので、目視により、正しい文字位置に切り出し候補矩形が存在すればよいことにする。

実験に用いた住所文字列には 9721 個の文字が存在し、そのうち 9337 個の文字切り出しに成功したので、切り出し率は 96.0 % となった。また、正しく切り出せた文字の中で、JIS 第 1 水準の漢字と平仮名の文字種である文字数を調べると 9270 文字であった。個別文字認識実験では、この文字を実験の対象とする。なお、標準パターンの作成には、ETL9B を用いた。

実験 2：正しく文字切り出しできた文字について個別文字認識率を求める。認識対象文字種を限定し、2 度個別文字認識したときの認識実験結果を表 1 に、比較対象実験として、認識対象文字種を固定して 1 度だけ個別文字認識したときの結果を表 2 に示す。また、文字の太さと傾きの補正が、どの程度個別文字認識において効果があるか調べるために、認識対象文字種を固定し、文字の太さと傾きの補正を施さなかったときの認識実験結果を表 3 に示す。なお、認識対象文字種を限定したとき、行政区域名部分では 1 度目の文字認識結果を、地域名部分では 2 度目の文字認識結果を用いる。認識対象文字種を固定したときは、地名辞書中の単語を構成する第 1 水準漢字および平仮名の 2143 字種を文字認識の対象とする。

実験 3：2 段階文字認識を用いたときと、対象文字種を限定せず一定の文字種で後処理前に 1 度だけ個別文字認識したときの後処理結果を表 4 に記す。認識できた住所文字列の平均個別文字認識順位、仮評価前後の平均住所候補数、2 度目の個別文字認識時の平均カテゴリ数も併せて記す。

表 1. 2 段階文字認識の累積個別文字認識率
Table 1. Cumulative recognition rate of two-stage character recognition.

候補順位	1	2	3	4	5	6
認識率	70.3	79.5	83.3	86.1	87.6	88.7
候補順位	7	8	9	10	30	300
認識率	89.6	90.3	91.0	91.5	94.5	99.3

表 2. 認識対象文字種固定の累積個別文字認識率
Table 2. Cumulative recognition rate when character category set is fixed.

候補順位	1	2	3	3	5	6
認識率	66.3	75.6	79.7	82.7	84.4	85.7
候補順位	7	8	9	10	30	300
認識率	86.7	87.6	88.3	89.0	92.2	98.6

表 3. 文字種固定、補正なしの累積個別文字認識率
Table 3. Recognition rate when shape re-formation is not performed.

候補順位	1	2	3	3	5	6
認識率	63.0	72.8	77.4	80.1	82.1	83.4
候補順位	7	8	9	10	30	300
認識率	84.6	85.4	86.1	86.9	90.4	98.0

表 4. 後処理結果
Table 4. Result of post-processing.

	2 段階文字認識	認識対象文字種固定
件数正解率	85.4 %	84.0 %
棄却率	0.1 %	0.2 %
誤答率	14.5 %	15.8 %
正解パターンの平均文字認識順位	23.0 位	28.6 位
仮評価前の平均住所候補数	430.6 個	150.5 個
仮評価後の平均住所候補数	98.9 個	77.1 個
2 度目の個別文字認識時の平均カテゴリ数	1243.2	—

評価実験の結果、個別文字認識実験の比較対象として行った実験より、文字の太さと傾きの補正は、個別文字認識率を向上させるのに有効であるとわかった。第 1 位分類率では、認識対象文字種を固定して文字補正を行う認識実験結果と比較して 3.3 % 向上している。後処理結果としては、2 段階文字認識を用いることにより、85.4 % の件数正解率を得た。本手法を用いることで、高い件数正解率が得られた。また、仮評価を行うことで、2 段階文字認識を用いたとき、住所候補数は平均で 430.6 個から 98.9 個に減少した。これにより、地域名部分の検索が大幅に減少していることがわかる。2 段階文字認識を用いた方が、認識対象文字種を固定したときよりも仮評価前の住所候補数が多い。これは行政区域名部分の検索で、単語検索の条件が認識対象文字種を固定するときより緩和されるからである。仮評価を行うことで、住所候補数の差は減少し、処理時間への影響は少なくなっている。

正解が得られなかったサンプルについて、その原因を調べると、誤って記述された住所や、切り出しの段階で失敗し

ているもの、個別文字認識誤りによるものが多かった。正解を得られなかったパターンを例を図9に示しておく。失敗例1では、出力結果、正解の文字列とも評価式のRの値は等しい。しかし、平均文字認識順位は出力結果の方が良く、誤認識となっている。切り出し候補矩形[2-3]の“島”と[6-6]の“島”の認識順位が誤認識原因である。失敗例2では、地域名部分の文字認識結果中に、正解の単語候補の一致文字が存在しない。これにより、正解の住所候補は得ていない。基本矩形8と9を統合して1つの文字候補とし、非一致文字でない“栄”が単語候補となり出力結果となった、地域名部分の文字認識順位が誤認識の原因である。

図9. 失敗例
Fig. 9. Example of failure samples.

8. おわりに

本論文では、手書き住所文字列の認識において、行政区域名部分と地域名部分を認識する際にそれぞれ認識対象とする文字種を限定し、個別文字認識することで、個別文字認識率の向上を試みた。これにより、個別文字認識率では、1位で4.0%向上し、しかも累積個別文字認識率は30位で94.5%となった。この結果、本後処理を加えることで、住所文字列の件数正解率でも1.4%向上し、85.4%の正解を得た。行政区域名部分の検索を終了した時点で、一度住所候補を絞り込むので、単語候補にする条件を緩和しても地域名部分の検索の高速化が図れた。単語辞書に市内、都内などの単語を加えることで、これらを含む住所表記にも対応できた。しかし、2段階文字認識を用いて個別文字認識の対象文字種を限定することで、正解でない単語が検索される可能性が高くなり、結果として誤ってしまう可能性がある。今後の課題は、このことについて、詳細に誤認識の

原因を調べることである。また、効率よく単語検索を行うため、手書き住所文字列のような大きな変形を含んだ文字の認識精度をさらに向上させる必要がある。

謝辞 本研究を行うにあたって、郵便葉書の手書きの住所文字列データを提供して下さった郵政研究所に感謝します。また、ETL9Bの手書き漢字データベースを提供して下さった電子技術総合研究所に感謝します。

(平成9年2月26日受付)

文 献

- (1) 鈴木 雅人, 孫 寧, 阿曾 弘具: “キー文字駆動型地名推論による手書き宛名認識アルゴリズム”, 信学技報, PRU95-5(1995).
- (2) 井野 英文, 孫 寧, 根元 義章: “ストローク幅に着目した閾値可変型手書き宛名の切り出し・認識アルゴリズム”, 信学技報, PRU95-109(1995).
- (3) 井野 英文, 猿田 和樹, 加藤 寧, 根元 義章: “ストローク情報に基づく手書き郵便宛名の切り出しに関する一手法”, 情処論, Vol.38, No.2, pp.280-289(1997).
- (4) 村瀬 洋, 新谷 幹夫, 若原 徹, 小高 和巳: “言語情報を利用した手書き文字列からの文字切り出しと認識”, 信学論, Vol.J69-D, No.9, pp.1292-1301(1986).
- (5) 福島 俊一, 下村 秀樹, 森 義和: “手書き文字列読み取りのための単語列探索アルゴリズム—文字タグ法—”, 情処論, Vol.37, No.4, pp.500-509(1996).
- (6) 郭 軍, 孫 寧, 根元 義章, 佐藤 利三郎: “整形変換を用いた手書き文字データベース ETL9B の高精度認識”, 信学論, Vol.J76-DII, No.5, pp.1015-1022(1993).
- (7) 濱 裕治郎, 梅田 三千雄: “2段階個別文字認識を用いた手書き住所文字列の認識”, 信学技報, PRMU96-150(1997).
- (8) 吉田 収志, 本郷 保夫: “非線形正規化を用いた背景特徴パターンマッチング法による手書き漢字認識”, 信学技報, PRU92-34(1992).
- (9) 高木 幹夫, 下田 陽久: “画像解析ハンドブック”, 東京大学出版会, pp.576-577(1991).

梅田 三千雄 (正員) 1945年5月13日生まれ。昭和43年大阪大学卒業。同年日本電信電話公社(現NTT)入社。平成元年大阪電気通信大学工学部教授。現在、同情報工学部教授。工学博士。文字認識、画像処理、認知科学等の研究に従事。電子情報通信学会、情報処理学会、テレビジョン学会、画像電子学会会員。

濱 裕治郎 (非会員) 1972年3月31日生まれ。平成7年3月大阪電気通信大学工学部経営工学科卒業。同年4月同大学大学院工学研究科博士課程前期情報工学専攻入学。現在在学中。文字認識、特に手書き住所文字列の認識に関する研究に従事。情報処理学会会員。