

ストローク情報に基づく手書き郵便宛名の 切出しに関する一手法

井野 英文[†] 猿田 和樹^{††}
加藤 寧^{†††} 根元 義章^{†††}

手書き郵便宛名の高精度な認識には、文字切出し、個別文字認識、郵便宛名辞書を用いた知識処理が必要になる。従来、これらの処理を組み合わせた手法が主に報告されているが、そのほとんどの場合、単一の筆記具によって書かれた文字を対象としている。文字の平均的な線幅（以下、ストローク幅）がほぼ一定の場合、接触・入込み文字に対処する手法として、射影ヒストグラムに対し、実験的に求めた閾値により切出し位置を決定する手法が有効であった。しかし、実際の郵便宛名には、筆・サインペン・ボールペン等、様々なストローク幅の文字が存在し、これらの手法をそのまま適用するには限界があると考えられる。本論文では、筆記具に依存しない切出し手法として、ストローク情報に基づく切出しアルゴリズムを提案する。提案法は、ストローク幅、水平方向走査時の文字ストローク交差回数および射影ヒストグラムの変化量という文字ストロークに関する3つの情報を効果的に組み合わせることにより、切出し候補位置を適切に決定する手法である。認識実験では、郵政研究所が配布した手書き宛名サンプル画像を用いて提案法の有効性を示す。また、切出し部に関する評価実験では、提案法が従来法より優れていることを示す。

Handwritten Address Segmentation Algorithm Based on Stroke Information

HIDEFUMI INO,[†] KAZUKI SARUTA,^{††} NEI KATO^{†††}
and YOSHIAKI NEMOTO^{†††}

In handwritten address recognition, it's important to segment characters correctly. Conventional methods mainly used fixed threshold to segment characters from handwritten address image. The problem of those methods is that they can not cope with variable stroke width of handwritten characters. In this paper, we propose a new algorithm in which adaptive threshold is used depending on the stroke information of character image. As a result, this algorithm obtained a high recognition rate for the test samples, distributed by Institute for Posts and Telecommunication Policy, MPT.

1. はじめに

近年、文字認識の新たな研究分野として自由書式の手書き文書認識に関する研究がさかに行われており¹⁾、特に手書き郵便宛名の高精度な認識への期待が高まっている^{2)~4)}。しかし、手書き郵便宛名から文字を切り出す場合、(1)活字に比べ文字の大きさや太さが不定、(2)文字のピッチ（文字と文字の間隔）が不

定、(3)接触・入込み文字の存在、等の問題点があれば、切出し候補位置を推定することは必ずしも容易ではない。

従来、手書き郵便宛名の切出し・認識アルゴリズムとして、ラベリングによる手法^{5),6)}と射影ヒストグラムに着目した手法^{7)~9)}等が報告されている。文献5),6)では、宛名のサンプル画像に対してラベリングを行うことによって切出し位置を決定し、それに従って認識および郵便宛名辞書を用いた知識処理を行っている。この手法は、分離文字の場合は有効であるが、文字どうしが接触している場合、誤りが生じる可能性が高い。特に分離文字の一部が隣り合う文字と接触している場合、ラベリング法は連結成分を追跡するがゆえに、つながっている部分を1つの候補矩形として見なす問題

[†] 松下電器産業株式会社
Matsushita Electric Industrial Co., Ltd

^{††} 山形大学
Yamagata University

^{†††} 東北大学大学院情報科学研究科
Graduate School of Information Sciences, Tohoku University

点があり、初期の候補図形からいかにして切出し位置を決定するかが課題といえる。

一方、文献7), 8)では、切出し候補位置として、射影ヒストグラム of 切れ目あるいは谷の部分に着目し、実験的あるいは経験的に固定閾値を設定して切出しを行っている。この手法は、対象とする宛名サンプル画像のストローク幅がほぼ一定である場合に有効性が確認されている。しかし、実際の郵便宛名は、筆・サインペン・ボールペン等、様々な筆記具で書かれており、そのため文字のストローク幅は異なり、射影ヒストグラムにも影響を及ぼす。このような場合、基準となる閾値を決定するのは困難であるため、十分な効果が得られない場合がある。

固定閾値の改良法として、文献9)が報告されている。この手法は、宛名サンプル画像から生成した射影ヒストグラムから度数分布を算出する。次に、その階級値の小さい方から見て、最初にピークが存在する位置をもとに閾値を決定する。この手法の問題点として、縦線1本分に相当する階級の度数が最初のピークとならない場合、ストローク幅を正しく抽出するのは困難となり、より正確なストローク幅の抽出手法および適切な閾値の設定法が望まれている。また、ストローク幅の推定法として、文献10)では線分上の任意の点から4方向(45度間隔)を眺望し、各点における最も短い幅の平均を文字のストローク幅として定義している。この方法の場合、計算量がやや大きい、いくつかの測定点が適切に選択できれば少ない計算量で正確にストローク幅を推定できる。

本論文では従来の問題点を考慮し、射影ヒストグラム法をベースに文字のストローク情報に基づく切出しアルゴリズムを提案する。提案法は、郵便宛名から抽出できるストローク幅、水平方向走査時に文字ストロークを横切る回数、射影ヒストグラムの変化量の3つの情報を効果的に組み合わせて利用する手法である。これにより、筆記具の異なる郵便宛名にも対応することが可能となる。また、接触文字や入込み文字が多く含まれる郵便宛名に対しても正確に切出し候補位置を推定できるようになり、宛名認識システムの高精度化を図ることが可能となる。

認識実験では、郵政研究所が配布した第3回文字認識コンテストの手書き宛名サンプル画像を対象に、セット1からセット9までのうち算用数字を含まない321サンプル画像に対して提案法の評価を行った(321サンプル画像のうち、筆、サインペン、ボールペンによって書かれたサンプル画像の数はそれぞれ114, 102, 105であり、平均的なストローク幅は40ドット

から4ドット程度までの範囲にある)。その結果、システム全体としては、94.4%の高い正解率が得られた。また、比較実験においては、従来の固定閾値を用いた手法での最も良い場合に比べ、切出し成功率では8.8ポイントの改善、正解率では6.2%(20サンプル)の改善となり、提案法の有効性が明らかになった。

本論文の構成は以下のとおりである。2章では提案する切出しアルゴリズムについて述べる。3章で提案法を用いた手書き宛名認識システムの概略について述べる。4章で認識実験を行い、その結果を示す。5章で提案法の有効性を確認し結論を述べる。

2. ストローク情報に基づく切出しアルゴリズム

本章では、提案するストローク情報に基づく切出しアルゴリズムについて述べる。なお、本アルゴリズムは縦書きを想定しており、横書きの宛名については別途検討する必要がある。

2.1 閾値可変の必要性

射影ヒストグラムを用いた文字の切出しを考えると、接触・入込み文字の発生頻度が低い場合、切出し候補位置として考えられるのは射影ヒストグラムの切れ目の部分である。しかし、文字どうしが接触あるいは入り込んでいる場合、射影ヒストグラムには切れ目がなく、候補文字図形を切り出すには適切な閾値を設定する必要があると考えられる。

閾値可変の必要性について、以下の具体例を用いて説明する。

図1はそれぞれサインペンと筆で書かれたサンプル画像であり、文字のストローク幅が異なる。宛名サンプル画像の右側には水平方向に射影して得られる射影ヒストグラムを示す。図1に示した閾値はそれぞれの宛名サンプル画像の接触・入込み部を切り出すために最低必要な閾値を示してあり、それぞれ異なる値となっている。図1から、文字のストローク幅に応じて閾値を設定する必要があるといえる。

次に、実験に用いない任意の50サンプル画像を用い、接触・入込み文字の状態について調査した。調査結果を表1に示す。表1から接触・入込み文字の9割弱が縦線1本分の接触となっていることが分かる。図2(a)に縦線1本による接触文字の例、(b)に縦線1本による入込み文字の例を示す。

以上のことにより、宛名サンプル画像ごとにストローク幅を推定し、適切な閾値を設定することができれば、縦線1本分で接触・入り込んでいる文字を効果的に切り出すことが可能であると考えられる。

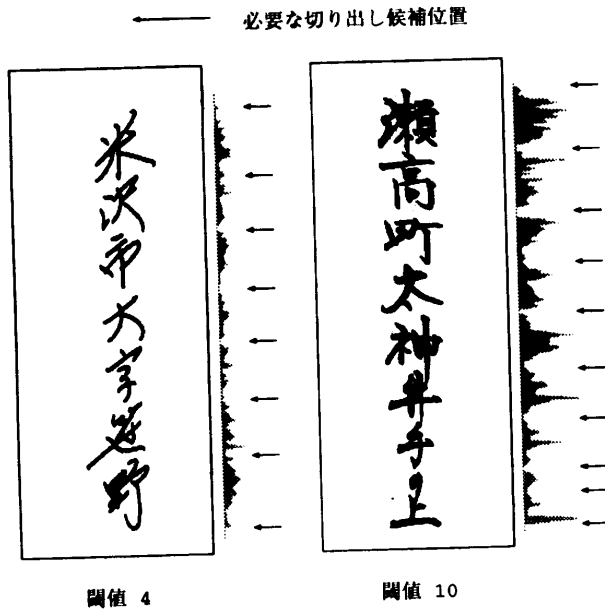


図1 接触・入込み文字を切り出すための閾値について
Fig. 1 Segmentation threshold for touching and overlapping characters.

表1 接触・入込み文字の分類
Table 1 Categorize for touching and overlapping characters.

縦線1本分での接触・入込み	87.1 %
その他のケース (1本以外)	12.9 %

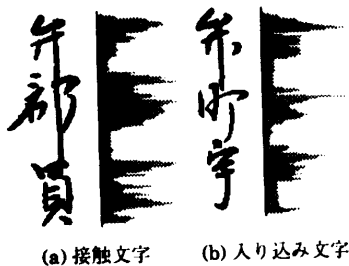


図2 接触・入込み文字の例

Fig. 2 Examples of touching and overlapping character.

本論文で提案するストローク情報に基づく切り出しアルゴリズムは、2.2節で求められるストローク幅を射影ヒストグラムに対する閾値とし、2.4節に示す条件を満たすものを切り出し候補位置として決定する。以下、その詳細について述べる。

2.2 ストローク幅の推定

ストローク幅の推定は、画像処理で使用される収縮法¹¹⁾を用いて行う。入力2値画像を f_{ij} 、出力2値画像を g_{ij} とすると、収縮法は次式で表される。

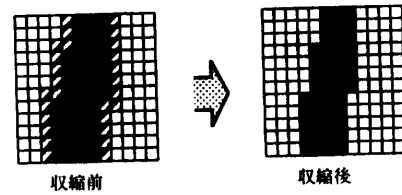


図3 収縮法の適用
Fig. 3 Reduction of pixels.

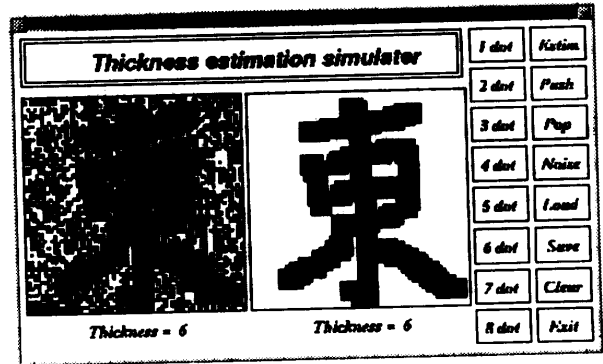


図4 ストローク幅推定シミュレータ
Fig. 4 Thickness estimation simulator.

$$g_{ij} = \begin{cases} 0 & f_{ij} \text{ あるいはその8近傍の} \\ & \text{いずれかが0の場合} \\ 1 & \text{その他の場合} \end{cases} \quad (1)$$

ただし、 i, j はそれぞれ水平方向、垂直方向の座標を表す。収縮の様子を図3に示す。網かけの部分は近傍に白い画素が存在することを示しており、この部分が収縮の対象となる画素である。文字のイメージは、1回の収縮で線幅が2ドット小さくなる。この処理を、原画像の黒画素数と収縮後の黒画素数の比が以下の条件を満たすまで繰り返し、収縮回数 n を決定する。そして、この収縮回数を2倍した値をストローク幅とする。

$$\frac{n \text{ 回収縮させた後の黒画素数}}{\text{原画像の黒画素数}} \leq 0.05 \quad (2)$$

ここで、定数0.05は、収縮回数が増えノイズや潰れによって影響されるのを防ぐためのものである。これによって、たとえば図4の左右どちらの画像に対しても、同じ収縮回数を得ることができる。なお、図4は筆者らが作成したストローク幅推定シミュレータである。マウスにより任意のストローク幅で文字を作成し、さらにランダムノイズも加えられるようになっている。画面上で真のストローク幅、ノイズを加えたイメージ、提案法を用いた推定値等を確認できる仕組みになっている。ランダムノイズをかける際に、2値画像内の背景部分の各画素に対し、一様乱数により白黒を決定し

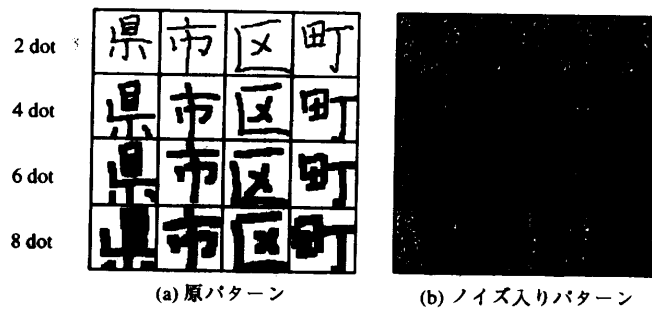


図5 評価使用データ
Fig. 5 Simulation data.

ている。この例では原データの黒画素数の約96%のランダムノイズを加えており、潰れの割合はこの範囲以内なら正しい収縮回数を得ることができる。ストローク幅が奇数の場合、誤差は+1dot発生するが、ストローク幅よりやや大きくなっているため、後述する式(3)による切出し候補位置の決定にはほとんど影響しない。

2.3 収縮法によるストローク幅推定法の評価

前節で述べたストローク幅推定法の評価を行う。宛名住所への使用頻度が高い、県・市・町・区の4文字に対し、ストローク幅が2・4・6・8の4通りのデータを作成し、ノイズを加えて推定を行った。評価に利用したデータを図5に示す。図5(a)が原データ、図5(b)がそれぞれにランダムノイズを加えたデータである。シミュレーションの結果、いずれのデータに対しても正確なストローク幅の推定が可能であることが確認できた。なお、図5(b)は推定可能範囲内で最大量のノイズを加えたデータであり、その範囲は原データの黒画素数の50% (8dotの県) から590% (2dotの市) までである。傾向として、ストローク幅が小さく、構造が単純な文字ほどより多くのノイズが加えられた場合でも正確な推定が可能である。

続いて文献9)で提案されているストローク幅推定法との比較を行う。文献9)の手法では、まず宛名サンプル画像から生成した射影ヒストグラムから度数分布を算出する。次に、その階級値の小さい方から見て、最初にピークが存在する位置をもとに閾値を決定する。この手法では、たとえば図6に示す例のように、筆で書かれた文字の場合、平均的なストローク幅より小さい部分に度数分布の最初のピークが生じてしまい、正解が得られないことがあると考えられる。図6のグラフは、射影ヒストグラムをもとに算出された度数分布をグラフにしたものである。また、文献9)の手法は、原データの黒画素数の1%程度のノイズにも影響されるという問題点も考えられる。これに対し、縮退法は

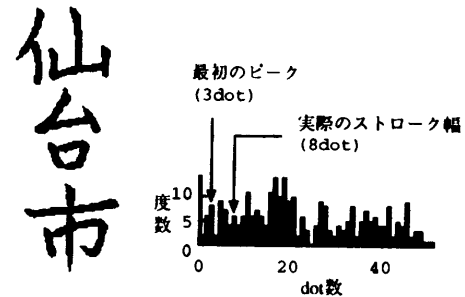


図6 筆によって書かれた文字列の例
Fig. 6 Example of writing brush of address data.

図6のデータに対して正しい推定が可能であり、またランダムノイズに耐えられる上限は321%である。

2.4 切出し候補位置の決定法

切出し候補位置は、以下の3つの条件(ストローク幅条件、交差条件、ヒストグラム変化量条件)を順に適用し決定する。

条件1 ストローク幅条件:

水平方向に射影して得られる射影ヒストグラムの値 $h(j)$ に対し、次式が成立する場合、

$$h(j) \leq 2 \cdot n \quad (3)$$

(j は宛名サンプル画像における垂直方向の座標、 n は収縮回数)

条件2 交差条件:

条件1が成立した部分において、水平方向走査時に文字ストロークと交差する回数が1である場合、

条件3 ヒストグラム変化量条件:

文字の境目は、射影ヒストグラムの変化量が大きいという性質を利用し、射影ヒストグラムの値 $h(j')$ が次式を満たす場合、 j' を切出し候補位置として決定する。

$$\exists \alpha (0 < \alpha \leq n) \quad 2 \cdot h(j') \leq h(j' + \alpha) \quad (4)$$

(j' は条件1, 2を満たす部分、 α は整数)

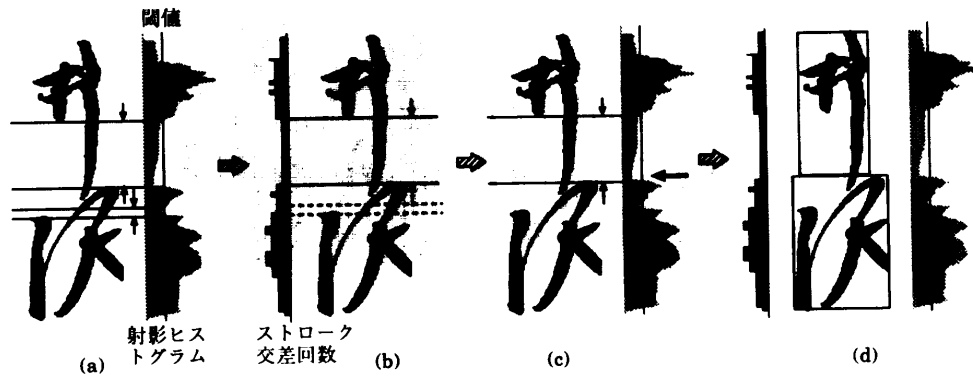


図7 切出し候補位置を決定する流れ
Fig. 7 Procedure of determining segmentation point.

係数の2であるが、これは、注目している文字が次の文字と接触・入り込んでいる場合、次の文字のストロークを検出するためのものである。また、 α の値を収縮回数で制限した理由は、

- (1) 注目点のすぐ下の位置に水平方向のストロークが存在する場合、射影ヒストグラム上、そのピークがストローク幅のほぼ半分の位置(α 付近)に存在するのは一般的であり、その範囲内で走査すれば接触部分の検出ができる、
 - (2) 収縮回数を用いることによってストローク幅が異なるサンプル画像にも対応できる、
- という理由による。

提案法の具体的な処理の流れを図7を用いて説明する。図7において、宛名サンプル画像の左側に水平走査時の文字ストロークとの交差回数、右側に射影ヒストグラムを示す。図7(a)は条件1により閾値を設定した様子を表している。射影ヒストグラムの値が $2n$ 以下の部分として、図7(a)の \uparrow で示した範囲に切出し候補位置を限定する。図7(b)では水平走査時の宛名サンプル画像のストロークとの交差回数により切出し候補位置をさらに限定する。この場合、破線で示した部分は文字ストロークとの交差回数が2であるため、切出し候補位置から除外する。図7(c)では射影ヒストグラムの変化量に着目し、切出し候補位置を決定する。条件3を満たすのは \leftarrow が指している点であり、これが最終的な切出し候補位置となる。図7(d)は、提案法により切出し位置が決定された宛名イメージを表している。なお、図7(d)から分かるように入込み文字、あるいは接触文字を切断した場合、隣の文字の一部が混入する場合があるが、ここでは、無理な分離をせずに各枠内に入ったイメージをそのまま認識系に送ることとする。その理由として、図7(d)のような単純な入込み文字の場合、切断位置の近辺をスキヤ

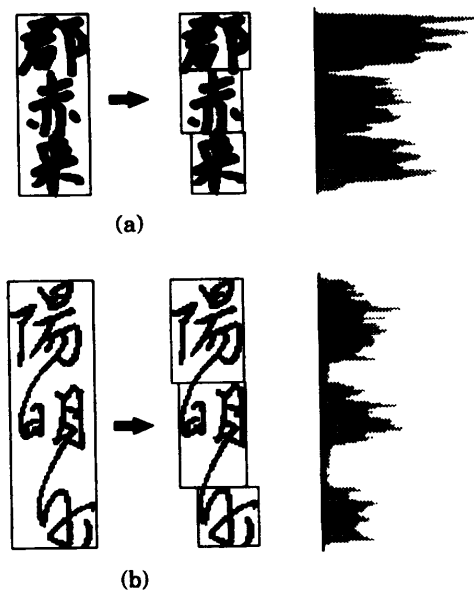


図8 提案法を用いて切出し候補位置を決定した例
Fig. 8 Segmentation examples using proposal method.

ンし、白画素と黒画素の境目を見つけることにより分離することは可能であるが、文字どうしが接触している場合、分離する位置を見つけるのが困難であり、また、混入部分が少ない場合、認識系に与える影響が小さいことがあげられる。

次に、図8に提案法を用いて接触・入込み文字を切り出した例を示す。図8(a), (b)の左側の切出し結果は、射影ヒストグラムをもとに切り出した例である。その右隣は提案法を用いて切り出した結果である。このように、ストローク幅、ストロークとの交差回数、ヒストグラムの変化量に着目する提案法は、ストローク幅の異なる宛名サンプル画像にも対応できることが分かる。

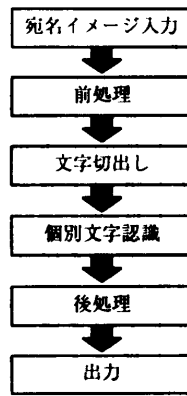


図9 宛名認識システム

Fig.9 Recognition system of handwritten address.

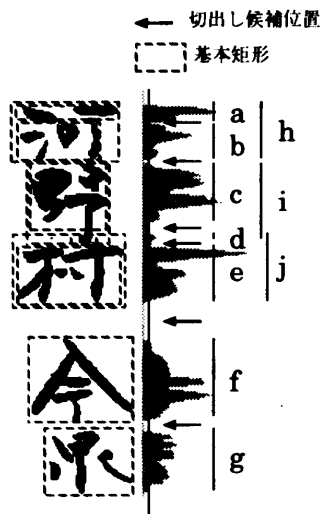


図10 切り出し候補位置および基本矩形

Fig.10 Segmentation position for characters.

3. 手書き宛名認識システムの概略

2章で述べた切り出しアルゴリズムを用いた宛名認識システムの概略を図9に示す。システムは、前処理部、切り出し部、個別文字認識部、後処理部の4部で構成されている。

3.1 前処理部

前処理部では、宛名サンプル画像の2値化とノイズ除去処理を行う。

宛名サンプル画像の解像度は400DPI、階調は256であり、まず輝度値50を閾値として2値化を行う。

次に、2値化した宛名サンプル画像に対して3×3の論理マスクによりノイズ除去処理を行う。これによって孤立画素と、幅が2ドット以下の線ノイズが除去される。

3.2 切り出し部

切り出し部では、提案法を用いて宛名サンプル画像に

表2 生成する候補文字ラティス
Table 2 Generate character lattice.

番号	候補文字ラティス
1	a → b → c → d → e → f → g
2	h → c → d → e → f → g
3	h → i → e → f → g
4	h → c → j → f → g
5	a → b → i → e → f → g
6	a → b → c → j → f → g

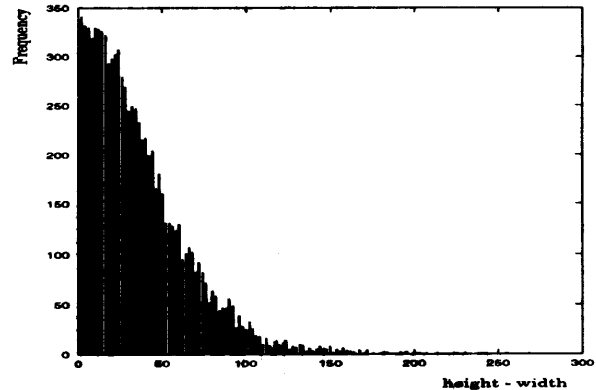


図11 文字図形の縦と横の長さの差の度数分布

Fig.11 Frequency distribution of difference between length and width of characters.

対して切り出し候補位置を決定する。次に、切り出し候補位置から候補文字ラティスを生成する。

候補文字ラティスの生成について図10を用いて説明する。図10では、切り出し候補位置をもとに統合処理を行い、文字となりうる基本矩形を決定する様子を表している。図10におけるa~jは基本矩形の縦の長さを表している。また、表2は、a~jから生成される文字ラティスを表している。候補文字ラティスの生成条件として、統合処理によって生成される基本矩形の縦と横の差を200ドット以下とする。200ドット以下に制限した理由は、実験に使用しない1,508宛名サンプル画像を用い、各文字の縦と横の長さの差を調べた結果、その差が200ドット以下の文字矩形が99.5%存在することによる(図11)。

3.3 認識部

認識部では、候補文字ラティスの中の個々の基本矩形に対して認識を行う。距離尺度として、文献12)で提案されている改良型マハラノビス距離を用いる。なお、標準パターンは、ETL9B手書き文字データベース全サンプルと第3回文字認識コンテストで配布されたサンプルのうち、実験に使用しない1,508画像サンプルから抽出した出現頻度の高い上位100字種を用いて作成する。これはETL9Bが高品質な手書き文字データベースであるうえ、複数の筆記具で書かれた

表3 100字種の内訳

Table 3 Number of 100 categories.

市 838	西 121	高 87	下 65	沢 51	久 41	県 36	口 33	間 28	美 24
町 709	中 117	宮 86	三 63	城 51	平 39	岩 36	富 32	池 27	塚 24
区 306	南 106	谷 85	八 62	浜 50	福 39	安 35	尾 32	代 27	手 24
田 235	字 106	上 85	井 61	江 49	台 39	多 34	石 32	根 26	古 24
郡 201	島 103	小 84	日 60	京 46	子 39	横 34	賀 32	海 26	金 24
大 192	原 102	津 77	戸 56	水 43	豊 38	都 33	河 32	一 26	名 23
川 188	北 100	新 73	千 55	神 43	寺 38	倉 33	浦 32	白 25	馬 23
東 159	本 99	木 68	長 54	見 43	橋 38	泉 33	前 31	港 25	国 23
山 144	内 97	松 66	岡 53	吉 43	佐 37	生 33	保 29	玉 25	丘 23
野 140	村 96	和 65	崎 52	桑 42	屋 37	阪 33	黒 28	郷 25	部 22

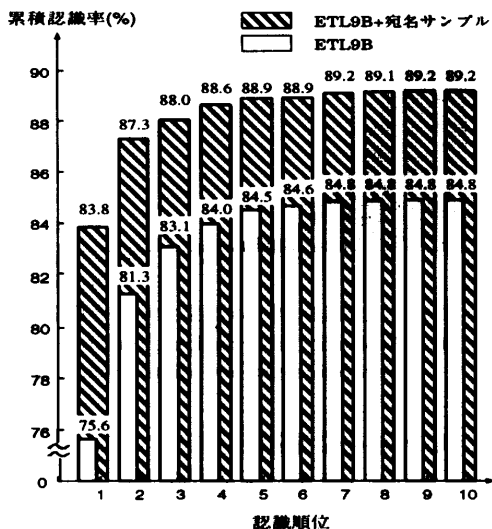


図12 異なる標準パターンによる個別文字認識結果

Fig. 12 Result of character recognition with different standard patterns.

サンプルが含まれていないという理由による。なお、100字種についての内訳を表3に示す。

図12に、ETL9Bのみから作成した標準パターンによる認識結果と、今回作成した標準パターンによる認識結果を示す。認識実験は、実験対象となる宛名サンプル画像に含まれる文字を手作業により正確に切り出したイメージに対し個別文字認識を行った。図12から、実サンプルを混ぜた標準パターンの方が、手書き宛名認識に有効であることが分かる。

3.4 後処理部

後処理部では、3.3節の認識結果を用い、郵便宛名辞書を利用した知識処理を行い、得点の高い宛名候補を結果として出力する。なお、後処理部で使用する郵便宛名辞書は、郵政研究所が配布した町域までの1,840件の宛名リストである。

宛名候補の絞り込みにおいて以下の得点評価を行う。この場合、各候補文字の認識結果は第3位まで用いる。

生成された k 番目の候補文字ラティスを C^k 、 C^k

の i 番目の文字に対する n 位候補字種 ($1 \leq n \leq 3$) を $C^k(i, n)$ とする。郵便宛名辞書における k' 番目の宛名を $M^{k'}$ 、 $M^{k'}$ の j 番目の文字を $M^{k'}(j)$ とする。得点 $f(i, j, k, k')$ を $C^k(i, n)$ と $M^{k'}(j)$ の整合により以下のように決定する。

$$f(i, j, k, k') = \begin{cases} 4 - n & \text{if } C^k(i, n) = M^{k'}(j) \\ 0 & \text{if } C^k(i, n) \neq M^{k'}(j) \end{cases} \quad (5)$$

ただし、 $1 \leq k \leq P$ 、 $1 \leq k' \leq Q$ とし、 P 、 Q はそれぞれ生成した候補文字ラティスの数および郵便宛名辞書に含まれている宛名を表す。式(5)において、文字列の順序関係を考慮し $i \leq j$ とし、一度対応関係がとれた i 、 j については再び評価の対象としない。

次に、 $M^{k'}$ に対する C^k の得点 $p_k^{k'}$ を次式により決定する。

$$p_k^{k'} = \sum_{i,j} f(i, j, k, k') \quad (6)$$

$p_k^{k'}$ が最大となる宛名候補を選択し、結果として出力する。

$$k^* = \arg \max_{k'} p_k^{k'} \quad (7)$$

4. 実験および考察

提案法の有効性を確認するため、宛名認識システム全体および切出し部に関する評価実験を行う。比較のため、従来法として文字切出し部に射影ヒストグラムに対して固定閾値を用いた手法についても実験を行う。

4.1 実験データ

認識実験に用いた宛名サンプル画像は、郵政研究所が配布した第3回文字認識コンテストのデータの一部である。サンプル数は、セット1からセット9までの332個のうち算用数字を含まない321である。数字を含むサンプル画像を外した理由として、ETL9Bには数字サンプルが含まれていないことと、これらのサン

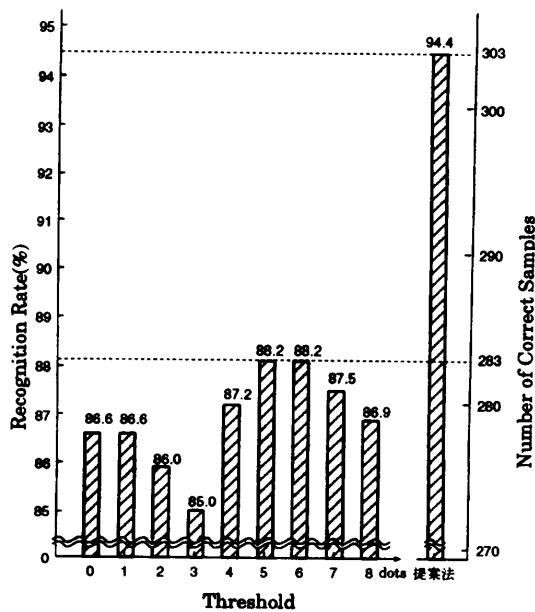


図 13 宛名の認識結果

Fig. 13 Recognition result of handwritten address.

ブル画像において文字どうしの接触、あるいは入込みがほとんど発生しておらず、提案法を検証するうえで、外しても支障がないと判断したことによる。宛名サンプル画像は、横 720 ドット縦 1680 ドットの大きさである。

4.2 宛名認識システムに関する評価

図 13 に提案システムによる結果と、固定閾値による切出しを行った結果を示す。図から分かるように、従来法では、閾値が 5 か 6 のとき、最も良い結果が得られている。その理由として、閾値を低く設定した場合（閾値 4 以下）、接触・入込み文字を十分に切り出すことができず、高く設定した場合（閾値 7 以上）、過剰な候補位置の生成とイメージの切落しが発生したことがあげられる。以上のことから、固定閾値を用いた手法では、すべてのサンプル画像に対して適切な閾値を与えることができず、高精度な切出しが困難であることが分かる。これに対し、提案システムでは、従来法の最も正解率が高い閾値 5 の場合に比べ、20 サンプルの改善となり、94.4 % の正解率が得られている。つまり、提案システムでは宛名サンプル画像ごとに適切な閾値の設定が行えるため、筆記具が異なる場合にも対応でき、正解切出し位置をより高精度に獲得できたといえる。

4.3 切出し部に関する評価

ここでは、切出し処理能力の評価を行う。まず、真の矩形、候補矩形、正解矩形を以下のように定義する。

表 4 切出し能力の評価

Table 4 Performance of segmentation.

	固定閾値による切出し*	提案法による切出し
候補矩形数	2,324	2,525
正解矩形数	1,841	2,027
真の矩形数	2,108	2,108
切出し効率	79.21 %	80.27 %
切出し成功率	87.33 %	96.15 %

* 閾値が 5 の場合

真の矩形*とは、正しい切出し位置から得られる矩形である。候補矩形**とは、切出し候補位置によって生成される矩形である。正解矩形***とは、統合処理を行った後に得られる真の矩形である。

また、評価式としては、切出し成功率と切出し効率を以下のように定義する。

$$\text{切出し成功率} = \frac{\text{正解矩形数}}{\text{真の矩形数}} \quad (8)$$

$$\text{切出し効率} = \frac{\text{正解矩形数}}{\text{候補矩形数}} \quad (9)$$

表 4 に評価結果を示す。固定閾値を用いた場合の結果は、図 13 で最も正解率の高い閾値 5 の結果を用いる。固定閾値による切出し結果と比較した場合、提案法では候補矩形と正解矩形はそれぞれ 201, 186 の増加となっている。提案法では、従来法より候補矩形を多く生成するものの、増加分の 92.5 % が正しい切出し矩形となっている。そのため、切出し効率では 1.1 % の向上、切出し成功率としては、8.8 ポイントの向上という結果が得られている。

4.4 誤認識の原因

宛名認識システムとして誤認識したサンプル画像は 18 であった。誤認識の原因を分析すると、正しい切出し候補位置を割り出せたものの、認識部の誤りが原因であるサンプル数は 12 であり、切出し部の誤りが原因であるサンプル数はわずか 6 であることが分かった。認識部の誤りにより不正解となったサンプル画像については、今後、認識部の性能向上により、救済される可能性が残されている。図 14 に切り出すことができなかったサンプル画像を示す。原因としては、文字どうしの接触・入込みが同時に発生していることがあげられる。

* 図 10 において、h, i, e, f, g により生成される 5 つの矩形。

** 図 10 において、a, b, c, d, e, f, g により生成される 7 つの矩形。

*** 図 10 において、h, i, e, f, g により生成される 5 つの矩形。



(a) 正解は「岐阜」 (b) 正解は「国町」

図 14 切出しができなかったサンプル画像の例
Fig. 14 Examples for error segmentation.

5. 結 論

郵便宛名は筆・サインペン・ボールペン等によって書かれており、筆記具によって記述される文字のストローク幅は異なり、射影ヒストグラムにも影響を与える。本論文では、まず郵政研究所が配布した宛名サンプル画像を用いて調査を行った。その結果、接触あるいは入込み文字の約 9 割が縦線 1 本分で発生していることが判明した。この調査結果をふまえ、本論文では宛名サンプル画像から文字を切り出す際に、文字のストローク幅、水平方向走査時の文字ストロークとの交差回数、射影ヒストグラムの変化量を利用したストローク情報に基づく切出しアルゴリズムを提案した。射影ヒストグラムに対する従来の固定閾値を用いた手法は、宛名サンプル画像のストローク幅にばらつきがある場合、適切な閾値を見つけるのが困難であったり、ストローク幅の細い文字に対するイメージの欠落やストローク幅の太い文字に対する接触・入込み文字の誤切出しといった問題点が存在しているが、提案法は、宛名サンプル画像ごとに適切な閾値の設定が可能であり、筆記具が異なる場合にも対応できる。郵政研究所が配布した 321 の手書き宛名サンプル画像を用いて認識実験を行った結果、94.4% の高い正解率が得られた。この結果は、従来の固定閾値を用いた最も良い場合に比べ 20 サンプルの改善となった。

今後の課題として、低品質文字に対する認識精度を向上させる必要があると考えられる。また、郵便宛名辞書をより大規模に構築し、汎化能力を高める一方、より多くのサンプル画像について実験を行う必要もあると考えられる。

謝辞 本研究を行うにあたって、手書き宛名のサンプル画像を提供して下さった郵政研究所に感謝します。また ETL9B 手書き文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。

参 考 文 献

- 1) 西野文人：文字認識における自然言語処理，情報処理，Vol.34, No.10, pp.1274-1280 (1993).
- 2) 松井俊弘，山下郁生，若原 徹，吉室 誠：文字認識アルゴリズムの複合化手法の検討・第 1 回文字認識コンテストの結果より，信学技報，Vol.PRU92-33, pp.65-72 (1992).
- 3) 能見 正，松井俊弘，山下郁生，若原 徹，吉室 誠：手書き数字認識における誤読・リジェクトパターン分析・第 2 回文字認識コンテストの実施結果より，信学技報，Vol.PRU93-46, pp.25-32 (1993).
- 4) 川又文男，若原 徹，松井俊弘，能見 正，山下郁生，堤田敏夫：手書き漢字宛名認識に関する第 3 回文字認識技術コンテストの実施，1994 年度，電子情報通信学会秋期大会，D-321, p.329 (1994).
- 5) 馬場口登，塚本正敏，相原恒博：認識処理の導入による手書き文字切りだしの一改良，電子情報通信学会論文誌，Vol.J69-D, No.11, pp.1774-1782 (1986).
- 6) 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，Vol.PRU95-5, pp.33-40 (1995).
- 7) 小林弥生，津雲 淳：文字接続情報を利用した手書き文字列認識，信学技報，Vol.PRU91-67, pp.39-46 (1991).
- 8) 郵便自動化 (Postal Mechanization)，東芝レビュー，Vol.48, No.7, pp.527-550 (1993).
- 9) Ohta, K., Kaneko, I., Itamoto, Y. and Nishijima, Y.: Character Segmentation of Address Reading/Letter Sorting Machine for the Ministry of Posts and Telecommunications of Japan, *NEC Research & Development*, Vol.34, No.2, pp.248-256 (1993).
- 10) 山田博三，森 俊二：手書き文字データベースの解析 (II)，電子技術総合研究所技報，Vol.40, No.6, pp.513-529 (1976).
- 11) 高木幹雄，下田陽久：画像処理ハンドブック，東京大学出版会 (1991).
- 12) 加藤 寧，安倍正人，根元義章：改良型マハラノビス距離を用いた高精度な文字認識，電子情報通信学会論文誌，Vol.J79-D-II, No.1, pp.45-52 (1996).

(平成 8 年 3 月 25 日受付)

(平成 8 年 11 月 7 日採録)



井野 英文

昭和 45 年生。平成 6 年東北大学工学部情報工学科卒業。平成 8 年同大大学院情報科学研究科博士前期課程修了。同年松下電器(株)に入社。現在同社マルチメディア開発センターに勤務。マルチメディアオーサリングに関する研究開発に従事。電子情報通信学会会員。



猿田 和樹

昭和 43 年生。平成 3 年秋田大学鉱山学部電子工学科卒業。平成 5 年同大大学院鉱山学研究科修士課程修了。平成 8 年東北大学大学院情報科学研究科博士後期課程修了。同年山形大学人文学部講師。現在に至る。博士(情報科学)。ニューラルネットワーク、文字認識に関する研究に従事。電子情報通信学会会員。



加藤 寧(正会員)

昭和 37 年生。昭和 63 年東北大学大学院工学研究科情報工学専攻修士課程修了。平成 3 年同後期課程修了。工学博士。同年同大大型計算機センター助手。平成 7 年同大大学院情報科学研究科助手。平成 8 年同助教授。現在に至る。文字認識、ニューラルネットおよびコンピュータネットワークの構築、管理に関する研究に従事。電子情報通信学会、IEEE 各会員。



横元 義章(正会員)

昭和 20 年生。昭和 43 年東北大学工学部通信工学科卒業。昭和 48 年同大大学院博士課程修了。工学博士。同年同大助手。昭和 59 年同大電気通信研究所助教授。平成 3 年同大大型計算機センター教授。平成 7 年同大大学院情報科学研究科教授。現在に至る。マイクロ波伝送路回路、衛星利用ネットワーク、情報伝達システム、手書き文字認識、画像処理等の研究に従事。昭和 56 年 IEEE・MTT・Micro Wave Prize 受賞。IEEE、電子情報通信学会各会員。