

背景領域細線化を用いた手書き文字切出しの改良手法

Improvement of Handwritten Character Segmentation Method with Background Thinning

中山 英久[†]

Hidehisa NAKAYAMA

藤原 勇太[‡]

Yuta FUJIWARA

加藤 寧[‡]

Nei KATO

1. まえがき

日本語手書き文字認識に関する研究は、主に個別文字を対象として、数多く行われてきた。その結果、手書き文字データベース ETL9B[1] のひらがな・漢字 3036 字種に対しては、99% 以上の認識率が報告されている [2]。しかしながら、手書き文字列に対する認識率は、個別文字に比べ、まだ十分な精度であるとはいえない。この要因には、手書き文字列からの個別文字への切出しが困難であることにある。手書き文字列は文字パターンが一定ではなく、文字幅、高さ、文字のピッチ (文字と文字の間隔) が不定であることに加えて、分離・接触・入込み文字の存在があるため、正確に文字を切出すことは非常に困難である。

上記の理由から、手書き文字列切出しでは、文字切出し位置に冗長性を持たせ、過剰な切断が生じることを前提として候補ラティス法 [3] を用いて最終的な切出し位置を決定する方法を取ることが多い [4],[5]。すなわち各候補ラティスの文字識別結果から言語情報を用いた知識処理によって最も妥当な文字列を選択することで切出し位置を補正し、最終的な文字の切出し位置を決定する。しかし、知識処理に頼る余り、過剰な切断が必要以上に多くなった場合、候補ラティス上での探索パスの組合せが増加し、知識処理における負荷の増大を招くことになる [6]。このことから、前処理としての文字切出しにおいては、接触・入込み文字を確実に切断しつつ、過剰な切断を抑制することが求められる。この両方を達成できれば、切出し後の個別文字認識の精度が向上し、それに伴って知識処理への負荷が軽減されることとなり、結果として認識システム全体の性能向上が期待できる。

従来、手書き文字切出しにおいては、黒画素領域に着目した手法として、ラベリング処理による手法 [7],[8]、射影ヒストグラムを用いた手法 [9],[10]、またそれら両方を用いた手法等が報告されている [11]。また近年では、接触文字や入込み文字をより正確に切出すため、文字領域ではなく、その背景領域に着目し、その細線化パターンを利用した手法が提案されている [12],[13],[14]。これらの手法では背景細線化パターンの利用によって非線形な線分が得られるために、自然なパスで文字を切出すことが可能である。そのため、手書き漢字等の複雑な接触・入込み文字が発生する文字切出しに対して非常に有効な方法であると考えられる。

いずれの従来手法でも接触文字の切出しにおいては、文字の黒画素分布やサイズ等の形状的特徴に着目し、接触文字の切出し位置であるかどうか判定を行っているが、手書き文字の形状は書き手に大きく依存すると

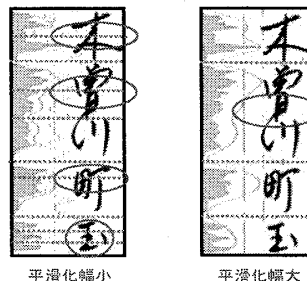


図 1: 射影ヒストグラム手法における切出し失敗例

いう問題がある。つまり、その特徴に条件を設けて切断位置を決定する方法では、接触文字の切断位置であるのか、もしくは個別文字を過剰に切出してしまおう位置であるのかどうかの判定についての信頼性が低く、先に述べた接触・入込み文字の確実な切断及び個別文字の過剰な切断の抑制は困難である。

そこで、本論文では、文献 [13] の背景領域細線化による文字切出し手法をベースとし、切出し判定を高精度化した改良手法を提案する。すなわち、切出し判定に認識処理を導入し、書き手に大きく依存する文字の形状的特徴にあまり重きを置かず、文字情報である認識処理を利用することで、対象の文字が接触文字か個別文字の判定を正確に行い、その後、接触文字に対して背景領域の細線化パターンを利用した詳細な切出しを行う。これにより、接触文字の確実な切断と個別文字の過剰な切出しの抑制の両方の性能向上を図る。実際に手書き郵便宛名を用いた認識実験の結果、提案手法が従来手法よりも認識率が向上し、かつ候補ラティス上の探索パス数が削減されていることから、その有効性を確認した。

2. 従来の文字切出し手法

本論文では、接触文字の切断位置であるのかどうかの判定を切出し判定と定義し、切出し判定の観点から従来手法について述べる。

2.1 黒画素射影ヒストグラムによる文字切出し手法

文献 [11] の手法について説明する。この手法では、ラベリング処理後、黒画素の射影ヒストグラムの谷部を切断位置に設定することで、接触・続け字の切出しを行う。この手法の問題点は、すべての文字画像に対して適用可能な最適なパラメータの決定が困難なことである。例えば図 1 では、射影ヒストグラムの平滑化幅が小さい場合、ヒストグラムの変化に対して過敏に処理することになるため、切出しミスは減少するものの、個別文字の過剰な切断が発生する。一方、平滑化幅が大きい場合、射影ヒストグラムの変化を大域的に捉え

[†]東北工業大学 工学部 知能エレクトロニクス学科, Department of Electronics and Intelligent Systems, Faculty of Engineering, Tohoku Institute of Technology

[‡]東北大学 大学院 情報科学研究科, Graduate School of Information Sciences, Tohoku University

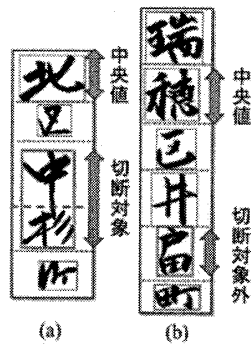


図2: 推定文字幅による切出し判定

るため、個別文字の過剰な切断を抑制できる反面、接触・入込み文字の切出しミスが生じることになる。つまり、黒画素の分布が書き手に大きく依存するため、一定の制約条件によって接触・入込み文字の正確な切出し位置を決定することは困難である。

2.2 背景領域細線化による文字切出し手法

ここでは文献 [13] の手法を取り上げる。この手法は2段階の切出しによって構成される。まず第1段階においては、黒画素射影ヒストグラムによって、画素の無い位置を切断位置として文字を分離する。この時、分離された領域を1ブロックとし、そのブロックの幅の中央値を推定文字幅として算出する。続いて、各ブロックにおいて、推定文字幅以上の幅を持つブロックに対して、背景領域細線化を用いた非線形パスの探索を行い、入込み文字を切り出す。そして、再度ブロックの幅の中央値を算出し、推定文字幅を更新する。このとき、各ブロックの幅が推定文字幅の定数倍以上ならばそのブロックを接触・続け字を含むブロックとして判定する。第2段階では、接触・続け字を含むブロックとして判定されたブロックについて、背景領域細線化及び文字画素の細線化を用いて、より詳細に切出しを行う。

以上のように従来手法では、背景領域細線化による接触文字の詳細切出しを行う前段階において推定文字幅を越えるブロックを接触・続け字を持つブロックとして判定しており、書き手が規則正しい大きさとで文字を書いた場合には正しく動作する(図2(a))が、書き手に制約条件を課さない自由書記においては正常な動作を保証することはできない(図2(b))。つまり、推定文字幅は書き手に大きく依存するため、接触文字の判定に用いるには、信頼性が不十分であると言える。実際に文献 [13] では、全接触文字の約17%が推定文字幅による切出し判定で誤判定しており、この誤判定は直接文字列の誤認識に繋がるため致命的である。

3. 認識処理導入による改良

3.1 提案手法の概要

切出し判定精度向上のためには、書き手に極力依存しない特徴を利用することが有効であると考えられる。文字画像においては、その画像が「文字である」という情報すなわち「文字情報」は不変であり、書き手に依存しないと考えられる。よって、提案手法では文字情

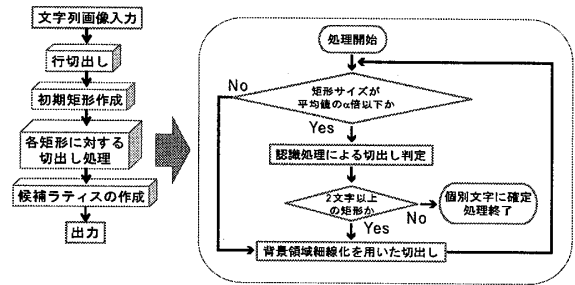


図3: 提案手法全体の構成

報を切出し判定の特徴として用いること、すなわち認識処理の導入により切出し判定精度の向上を図る。

3.2 一括認識に採用する文字列クラス

提案手法では接触・続け字・入込み文字を n 文字の組み合わせと考え、 n -gram と呼ぶことにする。郵便宛名データベース [15] における接触・続け字の接触文字数を目視により予備調査した結果、約90%が2文字での接触であった。本論文では、手書き文字列学習データとして、 $n=2$ すなわち bi-gram のみを考慮し、その学習データは、個別文字学習データから擬似的に作成することにする。また、すべての bi-gram クラスの学習データを用意した場合、処理量が膨大になるため、地名に出現する頻度によって絞り込みを行い、クラス数の削減を行った。具体的には、カウント数600をしきい値とし、カウント数の制限の無い73,549クラスから8,609クラスへ削減を行った。

3.3 一括認識による切出し判定

提案手法における切出し判定の詳細を述べる。以下の手順で行う。

1. 初期矩形の作成
2. 各矩形に対し、一括認識処理を行い、全学習パターンによる類似度を算出
3. K-NN(K-Nearest Neighbor) 法の考え方をを用いて切出し判定

まず、予め行切出しされた文字列に対し、ラベリングによって連結成分を囲う矩形を作成する。その後、黒画素の射影ヒストグラムを作成し、画素が存在しない区間の中点を切出し位置に決定する。その後、各区間内にあるラベル(矩形)を統合し初期矩形を作成する。

続いて、各矩形に対し、一括認識処理を行い、個別文字及び bi-gram の全学習パターンとの類似度を算出する。提案手法においては、判定の確度を高めるため、類似度の高い上位第 K 位までの結果を用いて、K-NN法の考え方による判定を行う。また、結果を参照する際、認識結果の字種を用いるのではなく、その字種の文字数を用いる。これは、個別文字から擬似的に作成した場合の文字列一括認識の精度が低いため、その確度の向上を図るためである。

例として $K=3$ の場合を示す(図4)。まず、初期矩形を作成し、各矩形の類似度を算出する。ここで、「名古」

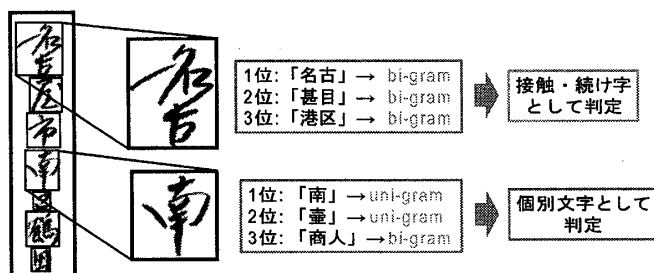


図 4: K-NN 法を用いた一括認識による判定

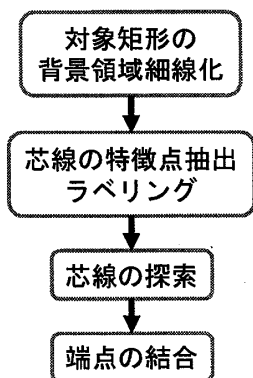


図 5: 背景領域細線化による切出し処理の流れ

の矩形に対しての上位 3 位の類似度は、すべて bi-gram クラスが占めているため、この矩形は接触・続け字として判定し、切出しを行う。また、「南」の矩形の上位 3 位の類似度は、uni-gram が多数を占めているため、この場合は個別文字に確定し、切出しを行わない。

提案手法の切出し処理の一連の流れを図 3 のフローチャートにまとめる。まず、各矩形に対し、矩形サイズが全矩形サイズの平均値の α 倍よりも大きい場合は、認識処理による切出し判定処理を行うことなく、切出し処理へ移行する。これは、3 文字以上の矩形のほとんどは矩形サイズによる文字数の判定が容易であることによる。また、切出し処理後、分離された各矩形に対しては、再び一連の処理を行い、認識処理による切出し判定においてすべての矩形が個別文字に確定するまで、繰り返すことで文字切出しを行う。

3.4 背景領域細線化を用いた文字切出し

文字切出しの流れを図 5 に示す。以下、各処理について図を用いて説明する。

3.4.1 対象矩形の背景領域細線化

高速性を重視し、鈴木らによる方法 [16] を背景領域細線化処理に採用し、対象矩形の芯線を得る。

3.4.2 芯線の特徴点抽出・ラベリング

本論文では、芯線が 3 本以上交差する交差点と芯線の端点を、文献 [17] で用いられている $3 \times 3 \text{dots}$ のマス

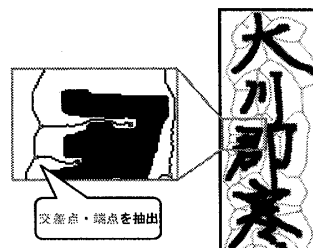


図 6: 芯線の端点・交差点の抽出

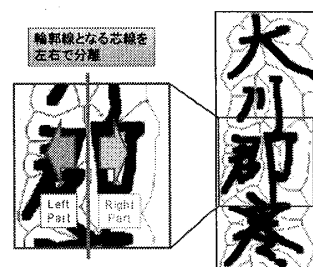


図 7: 芯線のラベリング

クを用いて抽出する (図 6)。

次に、画像の中央を境に、左部を Left Part、右部を Right Part とし、各 Part の輪郭線となる芯線をそれぞれ Left Part Segment, Right Part Segment と定義し、ラベリングによって分離する。図 7 のようにラベリングを行う。

3.4.3 芯線の探索

芯線の探索は、Left から Right へ探索後、Right から Left へ探索を行う。本論文で探索対象となる芯線は、各 Part の Segment 上の交差点を始点とし、端点または逆 Part の交差点まで伸びる芯線とする。

この段階において、Left Part Segment または Right Part Segment 上の交差点から逆の Part の Segment 上の交差点まで探索が成功した場合、そのパスを切出しパスとして確定する。これにより入込み文字を切断することができる。

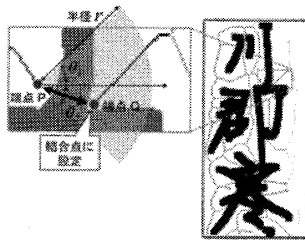


図 8: 端点同士の結合

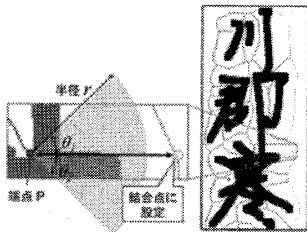


図 9: 端点と Right Part Segment との結合

3.4.4 端点の結合

芯線の探索時、端点まで探索が成功した場合、端点の結合を行う。まず、探索が成功した端点から次に述べる探索範囲内に逆 Part の端点があるかどうか調べる。探索範囲は、注目する端点を中心として、探索開始交差点の Part と逆 Part 方向へ水平線を引いた時、その水平線から仰角 $\theta_1 = 45^\circ$ 、俯角 $\theta_2 = 45^\circ$ となるように中心角 $\theta_1 + \theta_2 = 90^\circ$ をとった時の、半径 r の扇形とする (図 8,9)。以下、探索範囲に逆 Part の端点が存在する場合と存在しない場合の結合方法について説明する。

- 逆 Part の端点が探索範囲内に存在する場合
探索範囲内に逆 Part の端点が存在した場合、そのうち最近傍の端点を結合点とし、注目する端点と結合する (図 8)。
- 逆 Part の端点が探索範囲内に存在しない場合
まず、注目する端点から逆 Part 方向へ逆 Part の Segment と交差するまで水平線を引く。この時、水平線が文字ストロークと 1 回のみ交差する場合、その水平線を切出しパスとして採用し、水平線と逆 Part の Segment との交差点を結合点とする (図 9)。

4. 認識実験

4.1 手書き郵便宛名文字列を対象とした認識実験

手書き文字列認識における提案手法の有効性を評価するため、郵便宛名文字列認識実験により、従来手法 (射影ヒストグラムによる切出し判定手法 [11]、文字幅による切出し判定手法 [13]) との比較を行った。ただし、文字幅による切出し判定手法は、提案手法における切出し判定部のみを文字幅による判定処理に切り替えたものであり、切出しパスの生成は提案手法と同一の処理で行う。

4.2 実験環境

■ 認識対象データ

郵政研究所主催、第 3 回文字認識技術コンテストにおける宛名データ 1,840 件を用いる。画像は幅 $720[\text{pixel}] \times$ 高さ $1680[\text{pixel}]$ の 256 階調の濃淡画像であり、記載するのに用いられた筆記具は、ボールペン・万年筆・サインペン・毛筆の 4 種類である。

■ 認識システム

本論文における認識システムは、前処理、文字切出し、個別文字認識、知識処理から構成される。
[前処理] 前処理として、各文字列画像に対しコントラスト変換、エッジ保存平滑化を適用した後、2 値画像を生成する。

[文字切出し] 文字切出しでは、提案手法または従来の切出し判定手法を用いて文字列画像における切出し位置を決定し、候補ラティスを生成する。なお、提案手法における一括認識処理では、特徴量として 420 次元改良型方向線素特徴量 [2] を抽出し、識別関数は部分空間法 (CLAFIC) を用いる。

[個別文字認識] 個別文字認識システムには、196 次元改良型方向線素特徴量 [2] を抽出し、大分類では、ユークリッド距離、細分類では、改良型マハラノビス距離 [2] を用いて識別する。

[知識処理] 候補ラティス法を用いた段階的探索法 [18] による知識処理によって、最も評価値の高いものを認識結果とする。知識処理に用いる地名辞書は、全国規模の住所 169,439 件から 958,946 地名のリストを生成した。

4.3 評価項目及び評価指標

評価項目は文字切出し性能、知識処理への負荷、認識率である。評価指標を以下に示す。なお、文字切出し性能評価指標は、文献 [10] のものを用いる。

- 候補矩形数: 切出し後の矩形数
- 真の矩形数: 理想的に切出された正解の矩形数
- 正解矩形数: 統合後に得られる真の矩形数
- 候補ラティス数: 候補ラティスにおける全矩形数
- 切出し成功率: 式 (1) で算出される、どれだけ正しい矩形を得られたかの精度を表す指標
- 切出し効率: 式 (2) で算出される、どれだけ過剰な切出しを抑制しているかの効率を表す指標

$$\text{切出し成功率} = \frac{\text{正解矩形数}}{\text{真の矩形数}} \times 100 \quad (1)$$

$$\text{切出し効率} = \frac{\text{正解矩形数}}{\text{候補矩形数}} \times 100 \quad (2)$$

例として図 10 では、候補矩形数は a,b,c,d の 4 つ、正解矩形数は e,c,d の 3 つ、候補ラティス数は a~h の 8 つとなる。

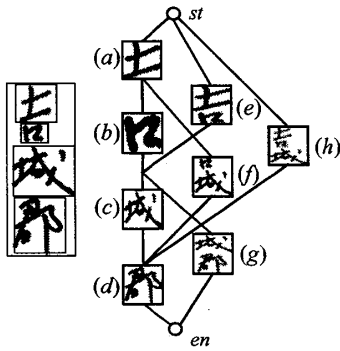


図 10: 候補ラティスの例

表 1: 文字切出し性能

文字切出し手法	切出し成功率	切出し効率
射影ヒストグラム手法	94.96	71.15
推定文字幅手法	91.69	75.92
提案手法	95.69	84.40

表 2: 知識処理への負荷

文字切出し手法	候補矩形数	候補ラティス数
射影ヒストグラム手法	22106	53230
推定文字幅手法	20914	39402
提案手法	19963	37688

表 3: 認識率

文字切出し手法	正解データ数	認識率
射影ヒストグラム手法	1561	84.84
推定文字幅手法	1515	82.34
提案手法	1602	87.07

右側が提案手法の結果である。これらにより、提案手法による切出し精度の改善が確認できる。

4.4 実験結果

■ 文字切出し精度に関する評価

実験結果を表 1 に示す。この結果より、切出し成功率と切出し効率はほぼトレードオフの関係にあることが分かる。この理由は、候補矩形数の増加を考慮せずに、とにかく切出せる可能性のある位置で切出しを行った場合、当然切出し効率は下がるが、候補ラティスの生成段階において多数の統合矩形が生成されるために、正解矩形数が作成される可能性が増えるためである。また、実験結果からいずれの評価値においても提案手法は高い値を達成しており、高精度な文字切出しの実現が確認できる。

■ 知識処理負荷に関する評価

実験結果を表 2 に示す。提案手法が最も候補矩形数が少なく、それから生成される候補ラティス数も従来手法に比べ、大幅に削減されていることが分かる。最大では、15442 個の候補ラティスの削減を達成した。以上から、提案手法では知識処理への負荷が達成されていると言える。

■ 認識システム性能に関する評価

表 3 より、提案手法が従来手法に比べ、高い認識率を達成したことが分かる。よって、文字切出しの高精度化及び知識処理負荷の軽減による認識システム全体の性能向上が確認できる。

4.4.1 切出し成功サンプル

切出し成功画像例を図 11,12 に示す。図 11 の左側が射影ヒストグラムによる従来手法、右側が提案手法であり、図 12 の左側が推定文字幅による切出し判定手法、

4.4.2 誤認識サンプル

提案手法における誤認識の原因について考察する。まず、提案手法で誤認識したデータを図 13(a) に示す。このデータは、切出し判定の失敗により誤読した例である。「ツ木」の矩形は、一括認識において個別文字として判定されてしまったために誤読したと考えられる。これは、提案手法で採用した bi-gram クラスとして「ツ木」の組み合わせクラスが存在しないことが原因であり、bi-gram クラスの採用方法の再検討、及び一括認識処理の高精度化によって改善可能と考えられる。

また図 13(b) のデータでは、提案手法において正しく切出し判定されたが、初期の切出しパスが不適切な位置に生成されてしまったために、切出し後の矩形の切出し判定が失敗し、結果として誤読してしまっている。解決のためには、切出し位置の補正などの処理が必要であり、今後の課題となる。

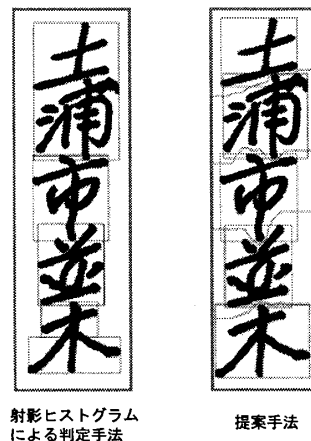


図 11: 切出し成功例 1

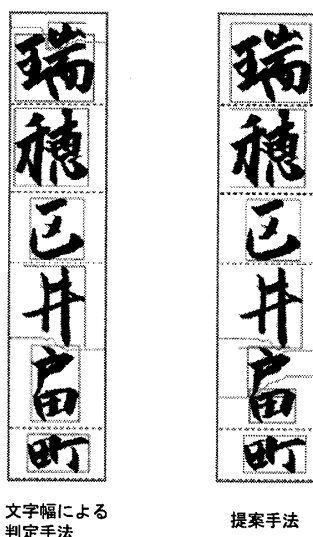


図 12: 切出し成功例 2

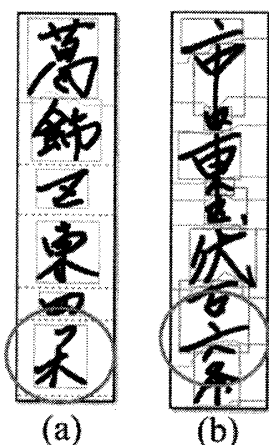


図 13: 誤認識データ例

5. むすび

文字切出しにおいては、接触文字を確実に切断しつつ、過剰な切断を抑制することが求められるが、それには接触文字かどうかの切出し判定の精度が重要である。しかし従来手法では、書き手に大きく依存する特徴を用いていたため、切出し判定精度が不十分であり、誤判定によってシステム全体の性能低下を招いていた。そこで、書き手に極力依存しない認識処理を導入し、切出し判定の高精度化を図った。手書き郵便宛名を用いた評価実験の結果、提案手法が従来手法よりも高い文字切出し性能を達成し、かつ候補ラティス数削減による知識処理負荷の軽減によって認識率の向上を実現することができた。

謝辞

本研究を行うにあたり、手書き文字データベースを提供して頂いた産業総合研究所と、手書き郵便宛名サンプルを提供して頂いた郵政研究所に感謝致します。

参考文献

- [1] 斉藤泰一, 山田博三, 山本和彦, “JIS 第 1 水準手書き漢字データベース ETL9B とその解析,” 信学論 (D), vol.J68-D, no.4, pp.757-764, Apr. 1985.
- [2] N. Kato, M. Suzuki, S. Omachi, H. Aso, and Y. Nemoto, “A handwritten character recognition system using directional element feature and asymmetric mahalnobis distance,” IEEE Trans. on Pattern Analysis & Machine Intelligence, vol.21, no.3, pp.258-262, March 1999.
- [3] 村瀬洋, 新谷幹夫, 若原徹, 小高和己, “言語情報を利用した手書き文字列からの文字切出しと認識,” 信学論 (D), vol.J69-D, no.9, pp.1292-1301, Sept. 1986.
- [4] 鈴木雅人, 安部裕之, 加藤寧, 根元義章, “候補字種の動的抽出を用いた手書きあて名の 2 段階認識手法,” 信学論 (D-II), vol.82, no.11, pp.1895-1902, Nov. 1999.
- [5] 梅田三千雄, 本庄大介, “完全一致法を用いた手書き住所文字列の認識,” 情処学論, vol.44, no.1, pp.114-123, Jan. 2003.
- [6] 山口輝幸, 吉川大弘, 篠木剛, 鶴岡信治, “接触した手書き漢字に対する個別文字の分離システム,” 信学技報, PRMU, pp.83-90, Jan. 2002.
- [7] 馬場口登, 塚本正敏, 相原恒博, “手書き日本語文字列からの文字切出しの基礎的考察,” 信学論 (D), vol.J68-D, no.12, pp.2123-2131, Dec. 1985.
- [8] 鈴木雅人, 加藤寧, 阿曾弘具, 根元義章, 市村洋, “キー文字駆動型地名推論に基づく手書きあて名認識,” 信学論 (D-II), vol.J80-D-II, no.5, pp.1077-1085, May 1997.
- [9] 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出しの一改良,” 信学論 (D), vol.J69-D, no.11, pp.1774-1782, Nov. 1986.
- [10] 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出しに関する一手法,” 情処学論, vol.38, no.2, pp.280-289, Feb. 1997.
- [11] 中山英久, 鈴木雅人, 加藤寧, 根元義章, “手書き郵便宛名に対して効果的な文字切り出しに関する一考察,” 電気関係学会東支部連合大会, p.224, 2001.
- [12] Z. Liang, and P. Shi, “A metasynthetic approach for segmenting handwritten Chinese character strings,” Pattern Recognition, vol.26, no.26, pp.1498-1511, Jan. 2005.
- [13] S. Zhao, Z. Chi, P. Shi, and H. Yan, “Two-stage segmentation of unconstrained handwritten chinese characters,” Pattern Recognition, vol.36, no.36, pp.145-156, Nov. 2003.
- [14] 梅田三千雄, 橋本智広, “背景領域の細線化に基づく古文書の文字切り出しと認識,” 情処学論, vol.45, no.4, pp.1188-1197, Apr. 2004.
- [15] 川又文男, 若原徹, 松井敏弘, 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,” 信学'94 秋大, Sep. 1994.
- [16] 鈴木智, 阿部圭一, “距離変換の結果を利用した二値画像の逐次型細線化,” 信学論 (D), vol.J68-D, no.4, pp.473-480, Apr. 1985.
- [17] A. Elnagar, and R. Alhaji, “Segmentation of connected handwritten numeral strings,” Pattern Recognition, vol.36, no.36, pp.625-634, March 2003.
- [18] 中山英久, 鈴木雅人, 加藤寧, 根元義章, “候補ラティスの段階的探索法を用いた高精度な手書きあて名認識,” 信学論 (D-II), vol.J87-D-II, no.2, pp.544-553, Feb. 2004.