

手書き筆記体からの文字切り出しおよび認識

小泉 和也¹・西村 広光²・富川 武彦²

1 大学院工学研究科電気電子専攻(s052806@cce.kanagawa-it.ac.jp)

2 情報メディア学科({nisimura, tomikawa}@ic.kanagawa-it.ac.jp)

Hand-written Cursive Character Segmentation and Recognition

Kazuya KOIZUMI¹, Hiromitsu NISHIMURA², Takehiko TOMIKAWA²

Abstract

The purpose of this paper is an improvement of a handwritten cursive character segmentation method for its recognition system. The paper was evaluated segmentation from human decision in the past. Now, this paper is evaluated segmentation from character recognition without the help of human decision. Our segmentation method is an alphabet cursive character recognition system in cooperated with segmentation and evaluated segmentation result. Constructed character recognition system is able to use word slant correction. The same way, it was compared character slant correction. Both recognition performances are compared in order to evaluate from recognition experiments. As a result of the evaluation, segmentation precision has been improved by using character slant correction.

Key Words: hand-written character, segmentation, cursive character, character recognition

1. はじめに

近年、名刺の読み取りソフトや携帯電話における写真からの文字読み取りなど、文字認識の実用化・普及が進んできている。しかし、実用化されているものの殆どが印刷文字を対象としている。これは、手書き文字、特に筆記体や行書のように続き書きされた文字は認識性能が十分でないことが原因であり、現在も認識性能向上に向けた技術開発が期待されている。現在、手書きオフライン文字認識での単文字の認識率は約 97%であるが、筆記体文字の認識は約 80%程度である。不十分な認識性能の原因には続き書きされたパターンでは、どこまでが 1 文字かを認識することが難しいことが挙げられる。そこで本研究では手書き文字認識の精度向上に不可欠と言える、続き書きされた文字、特に英字筆記体文字の切り出しについて検討を行った。

手書き筆記体文字を認識する際には、続き書きされた文字を一文字ずつ切り出して認識する方法が一般的である。現在最も広く利用されている手法は、多重切断仮説を利用して認識する手法である¹⁾。これは複数の切断候補を探索し、複数の切断領域を組み合わせながら認識を行う認識手法である。この手法を

利用すれば、理論的には過剰切断があっても認識には影響を与えないが、認識速度の面から考えると、過剰な切断仮説は認識時の組み合わせパターンの指数関数的増加を生み、認識速度低下に繋がるため、可能な限り過剰切断を行わないようにする必要がある。

筆者らは先に過剰切断を抑制した続け書きパターンからの文字切り出し手法を報告し、人間の目による切り出しの正当性評価を行った。本論文では、多重切断仮説を利用する文字認識システムに提案切り出し法を統合し、提案切り出し手法が実際の単語から切り出した文字の認識においてどれだけ有効であるか実験を通して検討した。

また、切り出し・認識時に利用する傾き補正においては、過剰に補正してしまうことが認識性能低下に繋がると考えられる。そこで本研究では、認識の際に利用する傾き補正の処理に関して、単純に単語全体で一律の補正する場合と、過度の傾き補正をしないような制限下において、単語全体での一段階の補正を行い、切り出したパターンを統合して文字認識する際に再度の補正を行う方法について、どちらが文字認識システムで良好に

働くか比較検討を行った。

2. 文字切り出しシステムの概要

筆者らは文字の切り出しを行い、人間の目による切り出し正当性の評価を過去に報告している²⁾。以降、本論文ではこの切り出しシステムを前切り出しシステムと記載する。

前切り出しシステムでは図1の流れで行った。始めに与えられた濃淡画像を最大階調値の半分を閾値として単純に2値化処理を行った。次に、ノイズの除去を行った後、単語全体での傾き補正によるパターンへの補正を行った。最後に、補正した画像の射影を利用して文字分割を行い、文字の切り出しを実現し、人間の目によって正しく文字ごとに切断できているかを評価した。

文字切り出しシステムの新たな改良部分について、2.1節から解説を進める。

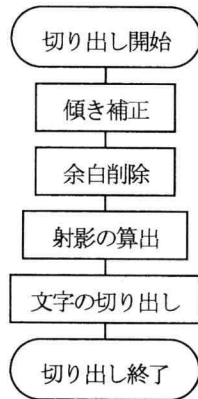


図1 従来のシステム

2.1. 前システムの改良

前切り出しシステムでは文字の切り出し位置を探し、その部分で切り出した目印として白い線を引いていただけだった。そのため、切り出しの評価実験においても人間の目による切断方法の正当性評価にとどまっていた。

文字パターンを一文字ごとに切り分けた処理結果が、人間の判断では正しく見えても、実際に文字認識システムと協調させて利用した場合に切り出しが認識性能向上に有効であるかは保障されない。そこで認識と協調したシステムを構築し、切り出しが文字認識性能向上に有効であるかの検討を進めた。

具体的には図2に示すように分割されたパターンがある。これを複数組み合わせることで文字として認識できるかを試行しながら単語認識を行うシステムを構築し、文字切り出し方法と

認識性能に関する評価検討を行うこととした。

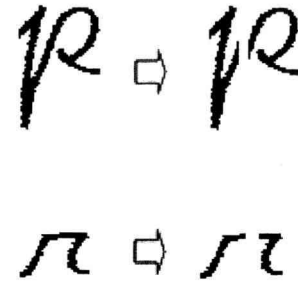


図2 「p」「r」の分割例

さらに、前システムではパターンへの傾き補正を単語単位で行っている。しかし、認識部と協調させる場合には傾き補正を複数の分割パターンを組み合わせた後に行うことも考えられる。しかし、これはパターン全体を考慮した補正にならないために、過度の補正になる恐れもある。そこで本研究では、切り出し段階での傾き補正のみを利用した場合、切り出したパターンを組み合わせた後にも傾き補正を行った場合について認識性能の比較検討を行うこととした。

2段階の傾き補正は、図3に示すように、単語全体での字形の傾きがあり、いくつかの文字では、さらに全体とは異なる傾きをしている場合に対して有効であると考えられる。しかし、2段階の傾き補正を行うことで過剰に傾きを補正し、認識に悪影響を与えることも考えられる。そこで本研究では、補正する角度を15度にし、2段階目に関しては横幅12画素以下の微小な領域に対しては傾き補正を行わないこととした。微小なパターンの傾き補正では画像を文字として分割したときに図4のような極端に縦長の画像ができてしまうためである。しかしこの多くの補正は、傾き補正を行ってもほとんど補正されず、認識に悪影響を与えることは無いと考えられるためである。

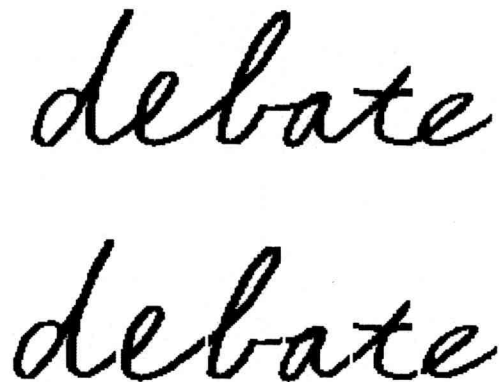


図3 実際の筆記体文字の切り出し

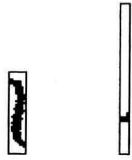


図4 分割文字の微小領域

2.2. 切断されたパターンの組み合わせ

切断したパターンを組み合わせる場合、文字の字形を構成する文字ストローク部分と、隣接文字間を結ぶ筆の運びのために描かれたリガチャ部分を切り分けて考える必要がある。図5において、灰色に示した部分をリガチャとよぶ。



図5 ストローク部分とリガチャ部分

複数に切り分けられたパターンを組み合わせる場合、リガチャではない部分が一文字にどれだけの数が含まれる可能性があるかを考える必要がある。本研究では、英文字小文字を対象とした効果的な切り出し方法として、縦方向ストロークの周辺箇所でのパターンの切断を行っている。そのため、1つの文字に縦方向のストロークが複数あるパターンの場合、理論的には最も過剰にパターンを分割する。すなわち、図6に示すように「m」や「w」といった縦方向のストロークにより、3つ以上に切り分けられる文字パターンである。



図6 「m」の分割例

図6において切断されたパターンを結合して「m」とするには6つの分割文字を組み合わせなければならない。ここでリガチャ部は文字と文字をつなげる部分のことを指すので、それ以外の文字本体であるリガチャでない部分を着目すると3つである。このリガチャでない部分を3つまで組み合わせることができるとした条件で分割文字の組み合わせをおこなっていくこととした。連結アルゴリズムの詳細は3章で後述する。

2.3. パターンマッチング

2.2で作成した文字領域の文字を組み合わせることで1つの単語を作成するが、このままだと何を基準で単語として組み合わせるといかわからない。そこで単体の文字領域での文字認識を行い、その類似度を利用している。パターンマッチングの流れを図7に示す。

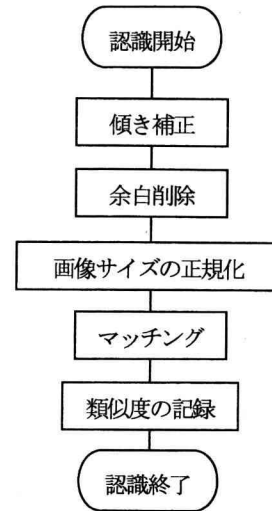


図7 パターンマッチングの流れ

この文字認識は文字切り出し部分で利用している。マッチングを除くすべての項目は以前の論文²⁾を参照されたい。本研究でのマッチングは重ね合わせ類似度算出を利用した単純なパターンマッチング法³⁾を用いた。

重ね合わせ類似度の算出は、2つのパターンの類似性を数値化したものである。文字認識においては、(1)式に示す類似度算出式を用いてパターンの類似度を求めることが多い。本研究でも、この類似度算出式を利用することとした。

$$\text{文字類似度} = \frac{\sum |x_{di} - x_i|}{\text{画像の画素数}} \quad (1)$$

x_{di} :辞書画像の画素の濃度値

x_i :入力画像の画素の濃度値

パターンマッチングを行う際には、対象パターンの大きさを揃え、傾きを補正し、位置ずれを考慮しながら最適な類似度を求める。

大きさを揃える際には、画像を拡大、縮小を行い、重ね合わせて類似度を向上させる。本研究ではあらかじめ大きさを指定し、そのサイズに合わせるようにした。

本論文で傾き補正とは、対象画像の傾きを変化させることを

指し、これを行うことで類似度を向上させることができる。本来は画像を回転させることで行うが、本研究では単語の傾きを平行四辺形のブロックとして捉え、長方形に直した場合にその画素に対応する場所を平行四辺形のブロックから求める処理を行い、単語が斜めに描かれないようにしている。

摂動法とは対象画像を縦、横に微小な平行移動を行いながら類似度を向上させる。摂動の変化量は対象画素の8近傍とし、8近傍内を移動して更に高い類似度が現れなくなるまで繰り返す処理を行う。類似度の求め方は(1)式と同様のものを用いている。

3. 切り出しパターンの統合

単語から文字ごとに切り出す際に多重切断仮説を利用する。多重仮説を利用して高い認識性能を実現するためには、その単語の文字数以上の分割数で切り出されることが多い。しかし、過剰な切り出しは文字認識するパターンが増え、類似度が低下する問題がある。そこで切断されたパターン同士で組み合わせる⁴⁾。

具体的に説明すると図8のようなものがある。図にある灰色に示している部分はリガチャを表している。この中の「e」と「b」に注目する。このようにリガチャを「e」の一部として見ることもでき、「b」の一部として見ることもできる。この場合は両方の切り出しパターンに適用することによって複数候補の高い類似度を出した文字を求めることができ、認識結果として適切な候補をだすことができると考える。

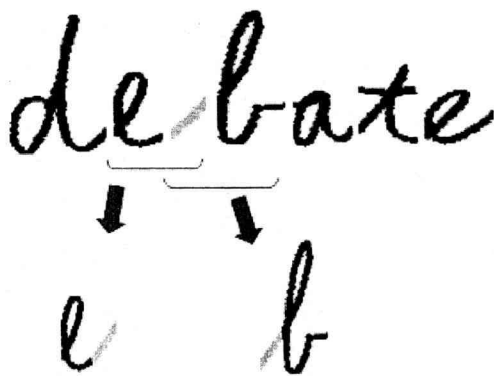


図8 リガチャを共有する場合

また図9のような組み合わせ方もある。この単語の中の「e」と「c」に注目する。切り出しパターンを組み合わせる場合にリガチャを含めて認識するよりも含めないで認識するほうが類似度が良いことがある。このように、リガチャ部に対して、文字の認識に利用しない箇所を許容することで複数候補の高い類似

度を出した文字を求めることができ、認識結果として適切な候補をだすことができると考えた。

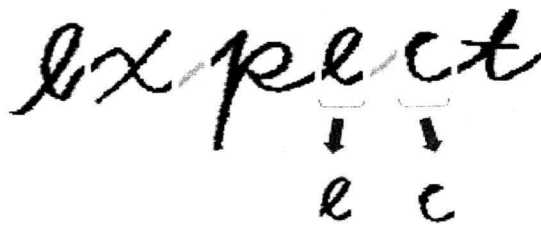


図9 リガチャを利用しない場合

4. 切り出し結果の認識システムによる評価実験

本章では2章で説明したシステムを用いて、切り出し結果からパターンを統合し、文字認識を行い、単語全体と文字ごとのパターン類似度がどのように変化するかを調べた。

さらに、傾き補正に関しては、2.1節に示した1段階に傾きを補正した方法、2段階に傾きを補正した方法についてその効果の差異を検討した。

4.1. 実験内容

この実験での文字認識は a~z ままでを別々に書いてもらった標準パターンを4組、あらかじめ文字の切り出したものから a~z ままで選んだものを1組の計5組を使って行った。これらは本学の5人の日本人学生に筆記させたものである。

実験で使用した画像はサイズが縦 152pixel×横 520pixel、保存方法を 256 階調の濃淡画像のものを用いた。このサンプルは前システムで利用したもので、A4のコピー用紙にPILOT SUPER プチ<中>耐水性というペンで書いてもらったものである。これも本学の1人の日本人学生に筆記させたものである。

以上のサンプルデータを30単語用意し、各単語を1~2単語の計55単語を用意した。このデータを用いて文字の傾き補正の有効性の検証と前システムと本システムの比較の2つの実験を行った。



図10 筆記単語例

4.2. 切り出し評価実験

切り出し評価は2つの方法で行った。

1つ目の評価方法として2.3節にある認識文字の傾き補正を利用した場合と利用しなかった場合を比較する。比較対象は単語全体の類似度と文字ごとの類似度の2つを比較し、この結果で評価を行う。単語類似度はその単語のつづり通りの文字認識が行われているかを表す度合いとする。また切り出し文字数が本来の単語の文字数と比べ正しいかの評価も行うこととする。単語類似度では類似度1位の結果を用いることとした。文字ごとの類似度では1位類似度のみを正解した場合と5位類似度までを正解した場合の結果から評価を行う。



図 11 文字領域の文字判定

表 1 「ask」の認識例

		a	s	k	
文字 似 の 度 順 位	1	o 90%	s 91%	l 85%	
	2	a 88%	c 85%	d 82%	
	3	c 85%	x 82%	h 79%	
	4	e 80%	v 77%	j 72%	
	5	u 70%	r 75%	z 68%	

単語類似度を求める例として、図 11 に示した「ask」を文字領域として切り出されたものと、この文字領域ごとの文字認識結果例を表 1 としたものをあげる。文字の1位類似度のみを正解とする場合は表 1 の1位類似度だけに注目する。本来の単語の文字と比べて正解のものは「s」の1つだけである。この正解数から本来の単語の文字数である3で割ると、単語類似度は33%となる。文字の5位類似度までを正解とする場合は表 1 の1位から5位までに注目する。本来の単語の文字と比べて正解のものは「a」と「s」である。この正解数から本来の単語の文字数である3で割ると、単語類似度66%となる。このように、切り出し結果を結合した状態のパターンの類似度が類似順位の条件を満たした数から本来の単語の文字数で割ることで単語の正しい文字数の割合を求める。単語類似度を求めた式は(2)式のとおりである。

$$\text{単語類似度} = \frac{\text{正解文字数}}{\text{本来の単語の文字数}} \times 100 \quad (2)$$

2つ目の評価方法として先に報告した人間の判断による切り出し評価と同じパターンに対して、パターンの結合・認識部を

利用した今回の評価結果とを比較する。今回の評価結果として利用するのは1つ目の評価実験の結果で高い結合文字認識性能を示した方式を用いる。

前システムでは文字の切り出し結果を「良い」、「判断困難」、「悪い」の3段階でおこなっていた。本システムでは切り出したパターンを結合して類似度を求めたとき、85%以上の類似度を出すパターンを切り出し・結合できた箇所を「良い」とし、80%~85%未満の類似度を出すパターンを切り出し・結合できた箇所を「判断困難」、それ以外の80%未満の類似度を切り出し・結合を「悪い」として判断を行うこととした。

4.3. 実験結果

4.2の評価方法に従って1つ目の評価を行った結果を図 12、図 13 に記す。これは各類似度範囲の該当数を求め、それに対する全サンプル数の割合を示している。

類似度の求め方は(2)式を用いた。この際、過剰分割または不足分割した数だけのペナルティは与えないで評価を行った。1過不足分割数ごとに1%、5%を類似度から引くことも試したが、ペナルティを与えないほうがはっきりと比較ができたためである。

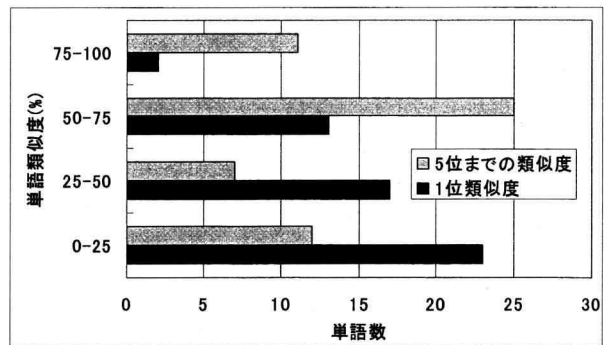


図 12 文字傾き補正なしの類似度分布

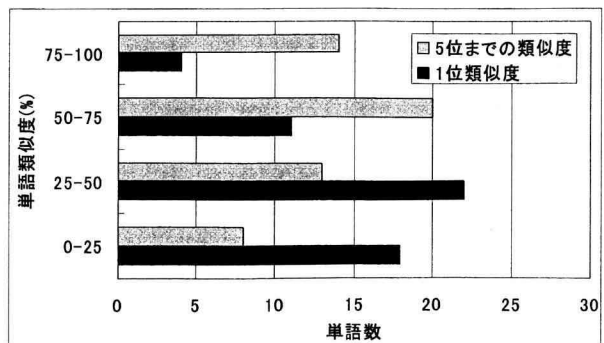


図 13 文字傾き補正ありの類似度分布

また単語の文字数に対する切り出し文字数の割合を表 2 に記す。「+」は過剰分割、「0」は過不足なし、「-」は不足分割を表す。

表 2 単語の過不足分割の割合

	+	0	-
1段階の傾き補正	50.91	27.27	21.82
2段階の傾き補正	50.91	32.73	16.36

この結果より 2 段階の傾き補正については 2 段階の傾き補正ありのほうが高い類似度を示した。これより 5 位類似度までが 50%を超えているものが半数以上あることから全体のおよそ 3、4 割程度は単語認識ができるものと考えられる。しかし、一般的な文字認識では類似度 90%以上で正解判定していることを考えると、今回の実験における類似度はそれほど高くはない。これは表 2 にある不足分割の割合が 16%あることが原因と考えられる。

また、単語の文字切り出し数については、2 段階の補正を行った方式のほうが補正なしの方式と比較して不足分割数が減少し、判断困難なところまで向上している。切り出し数についても 2 段階の補正があるほうが優れていることがわかる。

単語全体ではなく、文字ごとの類似度の結果を表 3 に記す。表 2 からわかるように、2 段階目の傾き補正を行うことで 1 位類似度が向上している。この結果から 2 段階の傾き補正が文字認識部の高精度化に有効であるといえる。

表 3 文字の類似度

	1位類似度	5位類似度まで
1段階の傾き補正	43.02	75.42
2段階の傾き補正	47.24	70.35

次に 2 つ目の評価である前システムと本システムの評価を行った結果を表 4 に記す。表の数値はその評価に対する全文字数の割合を示している。

表 4 前切り出し評価と本システムでの切り出し評価の比較

	良い	判断困難	悪い
前システム	74.17	7.99	17.84
本システム	81.22	17.26	1.52

この結果より 3 つの評価それぞれにおいて、本システムが優れていることから、目視での判断よりも良い結果と判断されたことになる。

また「悪い」の評価が格段に減ったことにより文字認識を行う上で問題が少なくなったと言える。

5. まとめ

今回の実験結果では 2 段階の傾き補正を利用した単語の類似度に関しては 2 段階の傾き補正を行ったほうが優れていることがわかった。しかし一般的なシステムと比べ、本システムでの単語類似度が低くなっているが、単語辞書を用意し照らし合わせていくことで文字の組み合わせを絞ることができるので類似度を向上することができる。また前システムの判断が適切であり、現システムによって精度が向上したことがわかった。

今回、文字ごとの類似度は 2 段階の傾き補正を使用したほうが優れていると結論を出したが、一般的に言われている文字認識の精度は 90%以上である。類似度が低くなった原因として文字認識を行う際、図 14 のような誤認識の問題点が挙げられる。

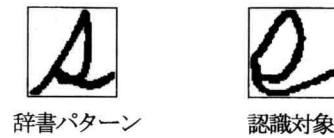


図 14 文字認識の問題点

図 14 は認識の辞書パターンにある「s」と認識対象の「s」である。認識対象の文字の左のリガチャが切れてしまっている。これは前システムで行った過剰分割の抑制方法として文字の左側で切り出すようにしていたためと考えられる。そのため、類似度の低下につながっている。この改善方法として文字の左側のみの切り出しに加え、右側でも切り出すことにし、リガチャとストローク部分をはっきり分けるような現在の認識システムに合うようにする必要がある。

今後は上記を含めたさらなる改良を行い、さらにサンプル数を増やし実験を行き、結果を出していきたい。

参考文献

- 1) H. Fujisawa, Y. Nakano and K. Kurino: Segmentation Methods for Character Recognition: Proc. IEEE, Vol.80 [7], pp.1079-1092(1992)
- 2) 小泉和也,西村広光,富川武彦: 手書き筆記体からの文字切り出し,神奈川県工科大学研究報告 B 理工学編 第 29 号(2004)
- 3) 舟久保登: 資格パターンの処理と認識,啓学出版(1990)
- 4) 丸山健一: 複数手法の統合による文字認識の高精度化に関する研究,信州大学博士(工学)学位論文(2003)