

5V-06 オンライン自由筆記文字列認識方式の検討

岡野 祐一 川又 武典 依田 文夫
三菱電機(株) 情報技術総合研究所

1. はじめに

従来のオンライン文字認識は文字枠内に筆記するという条件のもとに、携帯端末等における入力手段として用いられてきた。しかし最近では、ユーザへの筆記条件の緩和、また適用分野の拡大のために文字枠なしで筆記した文字列を読み取る方式が検討されている^{[1],[2]}。これらは、文字切出し、文字認識の結果に対して文字単位の接続情報 (n-gram 確率) を用いて最適な文字列を求めるものである。しかし、日本語文章は、形態素単位で文章が構成されており、文字単位の接続確率だけでは、形態素間の接続確率が反映できないため、十分な性能が得られないという課題があった。そこで、今回は形態素単位での接続情報を用いた文字列認識実験を行った。

2. 文字列認識処理の概要

図1に文字列認識処理の流れを示す。

2.1 文字切出し矩形候補作成

正しく行切り出しされた文字列データに対して、各ストロークのx軸方向への投影の重なり具合から基本切出し矩形を求め、次にこの基本切出し矩形を複数結合した矩形を文字切出し矩形候補に追加する。今回の実験では、x軸方向の重なりのない位置をもとに基本切出し矩形を抽出し、さらに結合後の矩形の横幅が、高さの2倍以内に収まる範囲内で文字切出し矩形候補を追加した。また、求めた各文字切出し矩形に対して、矩形の縦横比、前後矩形とのスペース幅等から、矩形評価値を算出している。

2.2 文字認識処理

文字切出し矩形候補に対して文字認識処理を行う。文字認識では、文字筆記過程の時系列情報及び線分間の相対的な関係を反映させたオンライン特徴を抽出し、パターンマッチング手法により文字を認識する^[3]。

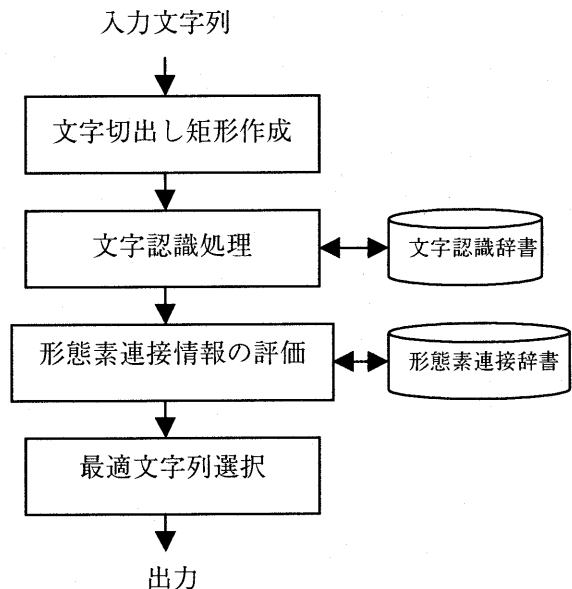


図1. 文字列認識処理のフロー

2.3 形態素接続情報の評価

2.3.1 形態素接続辞書について

1年分の新聞記事データベースに対して、形態素解析処理を行い、出現頻度の高い上位約4万5千語の形態素を抽出する。形態素解析の結果、この上位の形態素に属さないものに対しては、全て1文字毎に分離し1文字形態素としている。次に、新聞記事データベースから求めた各形態素間の出現頻度から形態素単位のbigram 確率を求める。なお、出現頻度の低い形態素に対してはバックオフ・スムージング処理によりbigram 確率を推定している。

2.3.2 形態素接続情報評価

入力文字列の先頭から、各文字切出し矩形に対する文字認識結果をもとに形態素辞書との照合を行い、形態素と一致する文字認識結果の候補文字の組み合わせを求める。形態素が照合された文字切出し位置において、前接する形態素の候補とのbigram 確率を求め、これを形態素の接続評価値とする。

2.3.3 最適文字列選択

形態素が照合された文字切り出し位置 n での文字

列評価値 $C(n)$ を以下のように定義した。

$$C(n) = C(m) + \sum W_i \times P_i$$

ここで、 $C(m)$ は前接する形態素が照合された文字切出し位置 m での文字列評価値、 $P_i (i=1 \sim 3)$ はそれぞれ現在評価している形態素を構成する切出し矩形の評価値の合計、文字認識結果の評価値の合計、及び前接する形態素接続情報の評価値である。また、 W_i はそれぞれの評価値に対する重み係数である。

各切出し位置での文字列評価値をもとに動的計画法により最適な文字列を探索する。

なお最適文字列を求める際に、文字列候補の組み合わせの増大を防ぐために、文字列評価を行った切出し位置において、それまでの文字列評価値が大きい上位候補 10 個の文字列候補のみを残すようにしている。最終的に最後の切出し位置において文字列評価値が最も高い文字列候補を最適文字列として選択する。

3. 認識実験

3.1 実験データ

今回文字列認識実験に用いたデータは以下のものである。文例：東京農工大で収集されたオンライン手書きデータベース^[注1]に使用された文章の一部。筆者：4名。筆記条件：“普段どおりに筆記。文字の大きさ、改行位置等は自由”。収集したデータは 968 行、約 17000 文字パターンである。

図2に収集した自由筆記手書きデータの一例を示す。

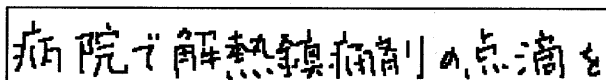


図2 自由筆記手書きデータの例

3.2 実験方法

収集した手書きデータを正しく 1 行単位に切出し、各 1 行の文字列パターンに対して文字列認識実験を行った。また、このとき文字列評価の際に用いる文字認識結果の候補文字数をいくつか変えて最終的な文字認識率を評価した。

4 実験結果

表1に実験結果を示す。表中の文字切出し率、文字認識率、枠無し認識率は以下の式で定義した。なお、ここでの文字認識率は、形態素接続情報による評価を含めた最適文字列選択後の文字認識率である。

$$\text{文字切出し率} = 100 \times L / N (\%)$$

$$\text{文字認識率} = 100 \times M / L (\%)$$

$$\text{枠無し認識率} = 100 \times M / N (\%)$$

N : 全入力文字数、 L : 正しく切出された文字数、 M : 正しく切出された文字のうち正しく認識できた文字数。

表1 文字列認識結果 (%)

候補文字数	1	3	5	7	10
文字切出し率	93.7	93.6	92.9	92.0	90.2
文字認識率	91.1	94.1	94.6	94.6	94.0
枠無し認識率	85.4	88.1	88.0	87.0	84.8

また、表2, 3に文字単位の bigram 及び trigram 確率を用いた場合の文字列認識結果を示す。

表2 文字単位の接続情報 (bigram) の結果 (%)

候補文字数	1	3	5	7	10
文字切出し率	93.7	93.4	93.4	93.2	92.9
文字認識率	91.0	93.8	94.5	94.7	94.9
枠無し認識率	85.2	87.6	88.2	88.3	88.1

表3 文字単位の接続情報 (trigram) の結果 (%)

候補文字数	1	3	5	7	10
文字切出し率	93.9	95.0	95.0	95.0	94.8
文字認識率	91.1	94.6	95.6	95.8	96.1
枠無し認識率	85.6	89.9	90.9	91.0	91.2

個々の認識結果からは、文字単位の接続情報だけでは正しく認識できない文字列に対する改善効果がみられた。しかし、表1の平均の枠無し認識率を見ると、候補文字数3の時に88.1%となった後、候補文字数が大きくなるに従って認識率が低下する結果となった。また、文字単位の接続情報を用いた結果に比べて候補文字数が大きい場合に、認識率が低くなった。これは、文字認識結果の候補文字に正解が含まれない場合や、他の候補文字の組み合わせにより誤った形態素が上位に選択された場合の誤認識の影響が大きいものと考えられる。

5 今後の課題

今後は、より多くの人数による手書き文字列データでの評価を行うとともに、文字認識結果に正解文字が含まれない場合に対する対処、及び誤った形態素が上位候補になる誤認識に対する対処を行う予定である。

[参考文献]

- [1] 仙田他 “切り出し・認識・言語の確信度…” PRMU98-138
- [2] 福島他 “確率モデルに基づく…” PRMU98-139
- [3] 川又他 “スローク間…” 第61回情処学全大, 2000, 5V-05
- [注1] TUAT Nakagawa Lab. HANDS-kuchibue_d-97-6