

文脈適応型手書き文字認識システム

岩山尚美

富士通研究所

naomi@flab.fujitsu.co.jp

一人のユーザが継続して使い続けることを仮定したオンライン手書き文字認識システムは、そのユーザにとっての認識性能が高くなければならない。本論文では、ユーザの特別な操作を必要とせずに個人への適応を行う手書き文字認識システムを提案する。このシステムでは、ユーザが使用した文字列を自動的に登録し、それをそのユーザの文脈として字形認識後の文脈処理を行う。本システムの有効性を実験により検証した。繰り返しが多い入力対象に対して、個々のユーザにとっての認識性能が向上した。また、同じ誤りを繰り返さないことにより、ユーザのストレスを解消し、満足度を向上させた。

A Handwriting Character Recognition System with Adaptive Context Processing

Naomi Iwayama

Fujitsu Laboratories Ltd.

naomi@flab.fujitsu.co.jp

A user who regularly uses an on-line handwriting character recognition system will expect that the system comes to recognize his own handwriting with a high degree of accuracy. In this paper, we present a handwriting character recognition system with adaptive context processing. This system automatically collects the strings inputted by a user and regards the collected strings as the context of the user. Based on the context, it chooses the string among the candidate strings returned by the pattern recognition module. We conclude that this method is effective for the following two reasons: it improves recognition accuracy for each user; and it reduces the stresses from the same misrecognitions being repeated and enhances user's satisfaction.

1. はじめに

オンライン手書き文字認識は、文字枠内に書かれた入力パターンに対して候補文字を出力する字形認識部と、字形認識部から出力された候補文字の中から文脈を考慮して最適な文字列を出力する文脈処理部から成る。字形認識では、オンライン認識とオフライン認識の統合手法などにより高精度な認識性能が実現されるようになった¹⁾。文脈処理では、数文字単位での処理が可能でコンパクト

な文脈処理が報告されている²⁾。字形認識も文脈処理も平均的な性能を表す指標としての評価用データの認識率が上がったことを根拠として、有効性が論じられてきた。

オンライン手書き文字認識は、PDAなどの携帯情報端末への文字入力手段として利用されることが多い。携帯情報端末は、一人のユーザが継続して使い続けるものであり、一人のユーザを仮定した手書き文字認識の観点からみると、今まで取

り扱われていなかった問題が存在する。例えば、あるユーザの「岩山」という入力に対して、手書き文字認識システムの結果が「若山」となったとする。これに対して、ユーザは、候補文字の入れ替えや書き直しなどの修正操作を行うことによって「岩山」という正しい文字列を入力する。次回、同じユーザが再度、「岩山」を入力した場合、前回の入力時とほとんど同じ筆跡が入力されることになり、それに対して、再び、「若山」という誤った結果が返ることになる。ユーザがよく使用する文字列に対してこのような現象が起きることは、そのユーザにとっての認識性能が低くなるというだけでなく、ユーザがシステムに対する不信感を抱く原因になると考えられる⁹⁾。一人のユーザを仮定すると、そのユーザにとっての認識性能は、そのユーザ自身が使用した文字列によって評価されるものであり、そのユーザが使用する文字列の認識性能を向上させることを目的としなければならない。個々のユーザにとっての文字認識システムの使い易さとは、そのユーザの書いた文字列が正しく認識できることであり、評価用データの認識率が高いことではない。

個々のユーザの認識性能を向上させるために、ユーザの字形を辞書に登録して、個人に適応した字形認識を行う方式が有効であることが知られている¹⁰⁾。しかしながら、先の例において、ユーザの字形登録を行った結果、字形認識部において「岩」が正しく認識されたとしても、その後の文脈処理において、「岩山」よりも「若山」の方が確からしいという判断がなされれば、やはり、「若山」という誤った結果を返してしまうことになる。したがって、個々のユーザの認識性能を向上させるためには、字形認識における適応だけでは不十分である。

このような観点から、本論文では、字形認識における適応ではなく、文脈処理において個人への適応を行う文脈適応型手書き文字認識システム(以下、文脈適応システムと略す)を提案し、本システムが個々のユーザの認識性能を向上させることを実験的に示す。以下の節では、まず、文脈

処理における適応について説明する。次に、実験結果について述べ、最後に本研究の結論を述べる。

2. 文脈処理における適応

2.1. 文脈処理の定式化

本節では、手書き文字認識システムにおける文脈処理について説明する。

図1に手書き文字認識システムの概要を示す。

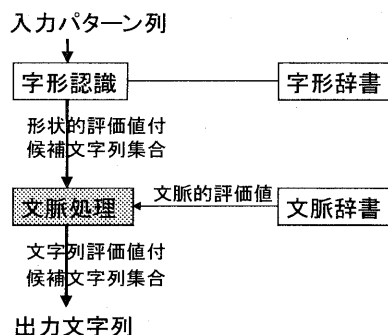


図1：手書き文字認識システム

手書き文字認識システムにおいて、字形認識部は、字形辞書との照合により、入力パターン列 X の各パターンに対して候補文字とその形状による評価値を出力するものであり、 X を候補文字列 S_x に形状的评价値 $L_{\text{shape}}(S_x)$ のついた候補文字列集合 $\{(S_x, L_{\text{shape}}(S_x))\}$ に変換するものであるとみなせる。文脈処理部は、文脈辞書から算出される S_x の文脈的评价値 $L_{\text{context}}(S_x)$ と $L_{\text{shape}}(S_x)$ をもとに、文字列评价値 $L_{\text{string}}(S_x)$ のついた候補文字列集合 $\{(S_x, L_{\text{string}}(S_x))\}$ を出力する。文脈処理結果 $\{(S_x, L_{\text{string}}(S_x))\}$ の $L_{\text{string}}(S_x)$ 順による1位候補文字列 S_x をシステムの出力文字列とする。

$L_{\text{shape}}(S_x)$ は S_x を個々の文字の形状から評価したものであり、 $L_{\text{context}}(S_x)$ は S_x を文脈辞書で定義される文脈から評価したものである。では、文字入力における文脈とは何であろうか。それは、入力される文字列の特性である。したがって、文脈辞書は、その特性を持った文字列の集合として表現することが可能である。ユーザが不特定多数であり、入力対象が限定されていないならば、日本語であるということが文脈である。入力対象が住所

であると限定されていれば、住所であるということが文脈である。一人のユーザを仮定した場合の文脈とは、そのユーザが、いつ、どんな文字列を使用したかということに他ならない。

2.2. 文脈適応システム

本節では、文脈適応システムについて述べる。

従来の文脈処理は、文脈辞書として2文字間の遷移確率辞書などを用いて、言語的な確からしさとしての言語的評価値($L_{\text{language}}(S_x)$)を $L_{\text{context}}(S_x)$ としていた。このような文脈処理の結果は、遷移確率を得るための学習データに依存し、入力された文字列が未学習であった場合には、第1節で述べたように何度も誤った結果を返すことになる。

この問題を解決する手段として、文脈処理においてユーザへの適応を行う文脈適応システムを提案する。文脈処理において適応を行うということは、 $L_{\text{context}}(S_x)$ をユーザに適応させることであり、 $L_{\text{context}}(S_x)$ は文脈辞書から算出されることから、文脈辞書をユーザに適応させることに帰着される。

第2.1節で述べたように、個々のユーザにとっての文脈を定義する文脈辞書は、そのユーザが使用した文字列集合として表現できる。したがって、文脈辞書をユーザに適応させる実現方法として、ユーザが確定した文字列をユーザ用語辞書として自動的に登録する。ユーザの確定とは、手書き文字入力インタフェースにおいて、他のアプリケーションに認識結果を転送することなどにより、認識結果を確定したとみなせる状態を意味する。

ここで重要なことは、ユーザ用語辞書に登録するために、ユーザは何ら特別な操作を必要としないことである。オンライン手書き文字認識の特長はインタラクティブ性にあり、ユーザが確定した文字列が何であるかは、ユーザの明示的な指示なしにシステムは知ることができるのである。

システムの入力は文脈処理結果の1位候補文字列である。システムは2位以下の候補も表示可能であるが、ユーザは出力される1位候補文字列のみからシステムの認識性能を評価する。したがって、ユーザにとっての認識性能を向上させる観点

からは、文脈処理は1位候補文字列を正解とすることにのみ注力すればよい。

図2に文脈適応システムの概要を示す。

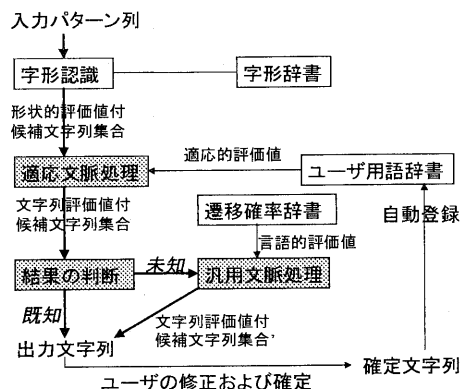


図2： 文脈適応型手書き文字認識システム

個々のユーザにとっての認識性能を向上させることを目的として、以下の3つの方針に基づき文脈処理を設計した。

方針 1: 過去に入力されたことのある文字列に対しては、適応の効果がある結果を出力する。

方針 2: 適応を行うことにより、文脈処理の効果を低下させない。

方針 3: 過去に入力されたことのない文字列に対しても、文脈処理効果のある結果を出力する。

これらの方針に基づき、文脈処理を以下のように実装した。

方針 1 により、ユーザの文脈をユーザ自身が使用した文字列であるとして、ユーザ用語辞書から算出される適応的評価値($L_{\text{adaptation}}(S_x)$)を $L_{\text{context}}(S_x)$ とする適応文脈処理を行う。

3つの方針を満たすための $L_{\text{adaptation}}(S_x)$ の要件は以下の4つである。

要件 1: $L_{\text{adaptation}}(S_x)$ は確率値の log をとったものとみなせる。

要件 2: ユーザ用語辞書に存在することによる適応の効果が表れる。

要件 3: 候補文字列の中でユーザ用語辞書に存在するものが複数ある場合は、その競合を解消する。

方針 4: ユーザ用語辞書に存在することによる適応の効果が過度でない。

要件 1 の必要性は次の理由による。我々の手書き文字認識システムでは、字形認識部が出力する $L_{\text{shape}}(S_x)$ は確率値の \log をとったもので近似できる。したがって、 $L_{\text{adaptation}}(S_x)$ を確率値の \log をとったものとみなせれば、

$$L_{\text{string}}(S_x) = L_{\text{shape}}(S_x) + L_{\text{adaptation}}(S_x)$$

と定義できることになる。

上記の 4 つの要件を満たすものとして $L_{\text{adaptation}}(S_x)$ を以下のように実装した。

要件 1 と要件 3 により、ユーザ用語辞書で定義される文脈において、確からしいものほど、 $L_{\text{adaptation}}(S_x)$ の値が大きくなる必要がある。これを満たすものとして、 $L_{\text{adaptation}}(S_x)$ を用語の長さ、頻度に応じて値が大きくなり、最近使用したものが優先されるように与えた。また、 $L_{\text{adaptation}}(S_x)$ の値のオーダーを $L_{\text{shape}}(S_x)$ に合わせ、
if $S_x \in$ ユーザ用語辞書 then $L_{\text{adaptation}}(S_x) > 0$
else $L_{\text{adaptation}}(S_x) = 0$ とした。

要件 2 により、ユーザ用語辞書に存在する文字列 S の入力パターン X に対して、 $S \in \{S_j\}$ であるとき、 S が適応文脈処理による 1 位候補文字列となる必要がある、つまり、 $L_{\text{string}}(S)$ のうちで、 $L_{\text{string}}(S) = L_{\text{shape}}(S) + L_{\text{adaptation}}(S)$ が最大となることが必要である。

要件 4 により、ユーザ用語辞書に存在しない文字列 S の入力パターン X に対して、 $S \in \{S_j\}$ であるとき、ユーザ用語辞書に存在する文字列 S' ($S' \in \{S_j\}$) があっても適応文脈処理による 1 位候補文字列が S' とならないことが必要、つまり、 $L_{\text{string}}(S)$ のうちで、 $L_{\text{string}}(S) = L_{\text{shape}}(S) + L_{\text{adaptation}}(S)$ が最大とならないことが必要である。

要件 2 と要件 4 を満たすように $L_{\text{adaptation}}(S_x)$ の最小値と最大値を字形認識部の出力結果の $L_{\text{shape}}(S_x)$ から統計的に算出して決定した。

$L_{\text{adaptation}}(S_x)$ を上記のように実装することにより、方針 1 と方針 2 を満たすことができる。さらに、方針 3 を満たすために、適応文脈処理による 1 位候補文字列がユーザ用語辞書に存在しない場合は、適応文脈処理の結果を破棄し、 $L_{\text{language}}(S_x)$ を $L_{\text{context}}(S_x)$ とする汎用文脈処理を行い、その結

果を文脈処理結果とする。適応文脈処理では、ユーザ用語辞書に存在する候補文字列 S_x があっても、1 位候補文字列がユーザ用語辞書に存在しないものとなる場合がある。これは、 $L_{\text{shape}}(S_x)$ と $L_{\text{adaptation}}(S_x)$ の両方から評価すると、入力された文字列が、適応文脈処理にとっては未知の文字列、つまり、過去に入力されたことのない文字列であると判断されたことを意味する。未知の文字列 ($L_{\text{adaptation}}(S_x) = 0$) である 1 位候補文字列は、ユーザ用語辞書に存在しない他の候補文字列と文脈的評価値が同じであり、これを文脈処理結果とすると、文脈処理効果のない結果を出力することになる。これを回避するために、適応文脈処理にとって未知の文字列であると判断された場合は、汎用文脈処理 (ユーザを特定しない文脈における文脈処理) の結果を文脈処理結果とすることとした。

一方、適応文脈処理による 1 位候補文字列がユーザ辞書に存在するものである場合は、入力された文字列が、適応文脈処理にとっては既知の文字列であると判断されたことを意味する。この場合は、方針 1 により、適応文脈処理の結果を文脈処理結果とする。

3. 実験結果

本節では、上述した文脈適応システムの有効性を評価するための実験とその結果について述べる。

3.1. 実験計画

次の 3 項目の検証を目的として実験を計画した。

- 1: システムが出力する 1 位正解率が向上する。
- 2: 1 位正解率の向上はユーザ自身が適応したことによるものではない。
- 3: 1 位正解率の向上をユーザが体感できる。

ユーザはシステムの性能を出力文字列により評価することから、システムの認識性能を 1 位正解率により評価することとした。文脈適応システムの 1 位正解率は、入力対象に依存するものである。なぜならば、文脈適応システムは、同じ文字列が繰り返して入力される場合に効果が発揮され、同じ文字列が繰り返されることのない入力対象に対しては効果がないからである。入力したすべての

文字のうちで、繰り返しの割合を示す指標として、

反復文字列出現率=

過去に入力済み文字列の文字数 / 入力文字数

を定義する。反復文字列出現率は、適応の効果による文脈適応システムの1位正解率の期待値を表すものである。

上記3項目を検証するために、次のような実験を行った。入力データが書かれた用紙の内容を手書き文字入力インタフェース部より入力を行う。本実験では、ある個人の3ヶ月分のスケジュールデータと個人データの2種類のデータを入力データとした。スケジュールデータの文字数は400文字であり、反復文字列出現率は32.5%である。個人データは、ユーザが頻繁に使用する入力対象という位置づけとし、1回の実験につき同じ内容を3セット入力することとしたので、反復文字列出現率は66.7%である。

各被験者は、1回の実験につきスケジュールデータと個人データを入力し、それを3回繰り返す。1回めの入力時は、適応のないシステムによる認識を行う。したがって、反復文字列出現率が0%の入力対象を文脈適応システムで認識したものとみなせる。2回め、3回めの入力時は、文脈適応システムによる認識を行う。3回めの入力時は、2回めの入力により、入力すべき文字列すべてをユーザ用語辞書に登録済みの状態で認識を行う。したがって、スケジュールデータ、個人データともに反復文字列出現率が100%の入力対象が仮想的に作り出されたことになる。

3回の実験後、アンケートにより認識性能の違いをユーザが体感できたかどうかを調査する。

14人の被験者の性別、年齢、職種、ペン入力の経験による人数構成は以下である。

性別	男性 9人	女性 5人	
年齢	20代 3人	30代 10人	40代 1人
職種	研究職 10人	事務職 4人	
経験	あり 9人	なし 5人	

2人の被験者は、1回の実験につき個人データを1セットしか入力していなかったため、個人データに関しては12人のデータを評価対象とした。

3.2. 文字入力におけるユーザ自身の適応

本節では、文字入力においてユーザ自身の適応が起こらなかったことを検証する。

表1に3回の実験における文脈処理前1位正解率(%)の平均値を示す。

	1回め	2回め	3回め
スケジュール	81.72	81.84	80.51
個人	80.78	79.23	78.63

表1：文脈処理前1位正解率の平均値

対応がある場合の分散分析の結果、5%の危険率で3回の実験における文脈処理前1位正解率の平均値に、スケジュールデータ(F(2,13)=0.78, p<0.05)、個人データ(F(2,11)=0.42, P<0.05)ともに、有意差は確認できなかった。つまり、文字入力におけるユーザ自身の適応は起こらなかったと言える。なぜならば、ペン入力をを用いて文字入力を行うことに慣れて、ユーザ自身の適応が起こるならば、回を追うごとに文脈処理前1位正解率が向上するはずだからである。

3.3. 文脈適応システムの性能面での効果

第3.2節で、3回の実験において文脈処理前1位正解率に差がないことが結論づけられた。したがって、適応のないシステムと文脈適応システムの文脈処理後1位正解率に有意差があるならば、文脈適応システムが個々のユーザの認識性能向上に寄与したことが検証できる。

図3にスケジュールデータの文脈処理後1位正解率の平均値(%)を示す。

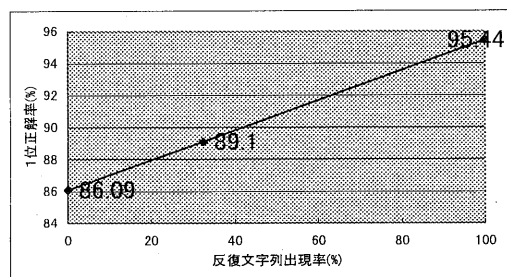


図3：文脈処理後1位正解率の平均値

対応がある場合の分散分析の結果では、5%の危険率で、反復文字列出現率 32.5%の入力対象における文脈適応システムの効果($F(1,13)=2.49$, $p<0.05$)は確認できなかったが、反復文字列出現率 100%の入力対象における文脈適応システムの効果 ($F(1,13)=24.8$, $p<0.05$)が確認できた。

次に、図4に個人データの文脈処理後1位正解率の平均値(%)を示す。

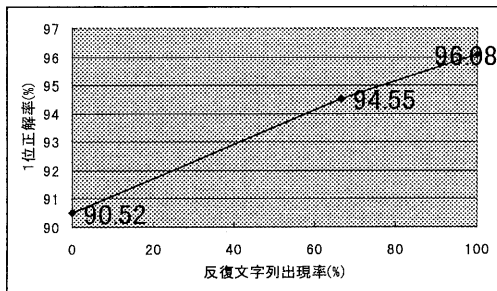


図4：文脈処理後1位正解率の平均値

対応がある場合の分散分析の結果により、5%の危険率で、反復文字列出現率 66.7%の入力対象($F(1,11)=15.46$, $p<0.05$)、反復文字列出現率 100%の入力対象 ($F(1,11)=12.70$, $p<0.05$)ともに、文脈適応システムの効果が確認できた。

つまり、文脈適応システムにより、反復文字列出現率が 32.5%の入力対象では認識性能の向上は確認できないが、66.7%、100%の入力対象では認識性能が向上したと言える。

3.4. 文脈適応システムの心理面での効果

本節では、入力実験後のアンケートにより、文脈適応システムのユーザの心理面での効果について分析した結果について述べる。

文脈適応システムを使用した場合に、適応を行わないシステムと比べて認識性能の向上をユーザが体感として感じる場合が多いことがわかった。

適応を行わないシステムに対する不満点として、「同じ文字に対して同じ誤り結果を何度も繰り返す」や「特定文字が認識できない」が挙げられた。この回答から、誤った結果を返した文字が何であり、誤った結果が何であったかがユーザの意識に

残るストレスになっていることが推察できる。

文脈適応システムの効果を感じた点として、「認識性能が高くなったこと」や、「前回誤ったものについて誤らなくなったこと」が挙げられた。この回答から、意識に残っている誤った文字が正しく認識されるようになることで、ユーザの満足度が向上していることが推察できる。

4. まとめ

本論文では、一人のユーザが使い続けることを仮定して、文脈適応型手書き文字認識システムを提案した。このシステムでは、字形認識後の文脈処理においてユーザへの適応を行う。ユーザは適応を行うために、何ら特別な操作を行う必要がない。普通にシステムを使い続けるだけで、繰り返しが多い入力対象に対して、個々のユーザにとっての認識性能を向上させることが実験により検証された。また、同じ誤りを繰り返さないことにより、ユーザのストレスを解消し、満足度を向上させる効果があるという結果が得られた。

謝辞

実験用の手書き文字入力インタフェース部を実装していただいた芝雅則氏、熱心に議論していただいた岩山登氏に感謝します。また、実験に協力していただいたすべての方に心から感謝します。

参考文献

- [1] 田中宏, 中島健次, 他. オンライン認識とオフライン認識の候補統合によるハイブリッド型ペン入力文字認識エンジン. 信学技法, PRMU, 98-140, 1998
- [2] 伊東伸泰. Bigram によるオンライン漢字認識の文脈後処理手法. 情報処理学会研究報告, Vol. 93-NL-97, 1993
- [3] D.A. ノーマン. (野島久雄訳). 誰のためのデザイン?. 新曜社, pp.67, 1990
- [4] 木村義政, 小高和己. ハンディ型パーソナルペン入力における個人字形登録の効果. 信学技法, PRMU, 97-5, 1997