

ペン入力手書き文字認識Active-X部品

Pen-Input Handwriting Recognition Modules

あらまし

近年のインターネットのブロードバンド化, 携帯電話・PDAなどの普及によりオフィス外で情報機器を操作する業務が増え, キーボードに代わる情報入力手段としてペン入力が目立っている。著者らは高精度なオンライン文字認識機能を容易にアプリケーションに組み込めるよう, 筆跡の表示や認識候補の修正などの対話的な操作を組み込んだソフトウェア部品をWindowsのActive-Xコントロールとして実装した。また単語検索や住所検索といった様々な機能を文字認識機能とともに利用できる部品も作成した。本稿では, まずはじめに著者らが開発した手書き文字認識技術と文脈処理技術の説明をする。続いて今回実装した4種類の手書き認識部品「汎用認識部品」「単語検索部品」「住所検索部品」「連続数字認識部品」の紹介をする。最後に, ペン入力インタフェースの長所と短所についてまとめた上で, これらの認識部品がどのような分野で有効に活用できるかについて考察する。

Abstract

The recent popularization of the broadband Internet, cellular phones, PDAs, and e-Japan project has created the need for a new, easy-to-use, data-input technology for conveniently entering information in mobile situations. The pen-input interface, which can easily be used by beginners, is one such interface that is likely to be widely used instead of the conventional keyboard. We have developed handwriting recognition software modules as Windows Active-X controls, with which programmers can easily build high-precision handwriting online recognition into their products. These modules enable interactive handwriting recognition operations such as pen stroke display and recognition candidate selection, as well as useful operations such as word and address searches. This paper explains our new handwriting recognition technology and context processing technique. This paper also describes four software modules: general-purpose recognition module, handwritten word search module, handwritten address search module, and handwritten number (continuous digits) recognition module. We summarize the advantages and disadvantages of the pen-input interface, and discuss effective pen-input applications using these modules.



田中 宏 (たなか ひろし)
パーソナル&サービス研究所メディア
ソリューション研究部 所属
現在, オンライン手書き文字認識シ
ステムの研究開発に従事。



中島健次 (なかじま けんじ)
パーソナル&サービス研究所メディア
ソリューション研究部 所属
現在, ペン入力サイン認証システム
およびオンライン手書き文字認識シ
ステムの研究開発に従事。



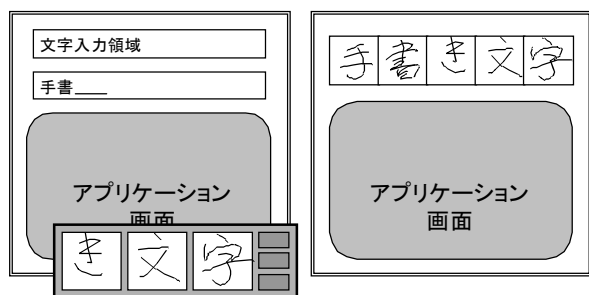
石垣一司 (いしがき かずし)
パーソナル&サービス研究所 所属
現在, オンライン手書き文字認識シ
ステム, ユーザビリティ評価技術,
バイオメトリクス認証システムの研
究開発に従事。

まえがき

近年のインターネットの普及や機器の小型化により、オフィス外で情報機器を操作する機会が増えている。外出時に情報を入力する機会が多いPDAや、小売店内や倉庫、顧客先などで立ったまま利用される業務用端末などではキーボードが使えないため、これらの多くはペン入力を採用している。従来は単に画面上のボタンを選ぶだけのマウスの代わりとしてペンを用いることが多かったが、最近では顧客の住所氏名を入力したり、在庫の商品名を検索したりと、ペンを使ってテキストを入力したい場面も増えてきた。

ペンを用いたテキスト入力にはソフトキーボードを用いるのが簡単だが、両手が使え通常のキーボードに比べてペン1本で操作するソフトキーボードは使い勝手が悪い。そこで、自然な操作でテキストが入力できる手書き文字認識技術が必要とされている。手書き文字認識によってテキストを入力する手書きソフトキーボードの画面を図-1(a)に示す。ソフトキーボードは通常のキーボードの代わりであり、どんなアプリケーションに対しても通常のキーボードと同様に入力できるが、ソフトキーボードの入力領域がアプリケーション画面を隠してしまうためアプリケーションそのものの操作を邪魔してしまうという問題がある。一方で、アプリケーション画面に埋め込んで使われるソフトウェア部品は画面を隠すことはないが{図-1(b)}、アプリケーションをペン入力のために設計して作り直さなければならない。

業務用機器では専用アプリケーションしか使わないことが多いため、様々なアプリケーションに対応可能な汎用のソフトキーボードよりも専用アプリケーションに合わせて設計できるソフトウェア部品を利用の方が望ましい。著者らは、特定業務におけるペン入力の使い勝手



(a) 手書きソフトキーボード (b) 手書き文字認識部品

図-1 手書きソフトキーボードと認識部品
Fig.1-Software keyboard and software module.

の向上と業務効率の改善を目指し、専用アプリケーションへの組み込みを前提としたペン入力手書き文字認識ソフトウェア部品を開発した。

著者らが開発した認識部品はMicrosoft Windows上で利用できるActive-Xにより実装し、プロパティ設定によって容易に動作がカスタマイズできる。また富士通研究所が独自に開発した高度な手書き文字認識技術とユーザインタフェースを実装し、文字認識と一体化した単語検索や住所検索などの便利な機能を持つ4種類の認識部品を用意することで業務に合わせた有効活用が可能となっている。

本稿では、まずペン入力手書き文字認識技術について紹介する。続いてこの技術を基本として開発した4種類の手書き文字認識部品について説明し、最後に認識部品を用いた応用について述べる。

手書き文字認識技術

手書き文字認識技術は、ユーザが筆記した文字列を認識して文字コード列に変換する技術である。入力された筆跡はまず1文字ずつ字形認識が行われ、続いて文脈処理によって最適な認識結果文字列が得られる(図-2)。

字形認識は、著者らが開発したハイブリッド型文字認識方式を用いている^{(1),(2)}。ハイブリッド型文字認識では内部に性質が異なる2種類の文字認識方式を用いており、それぞれの認識方式によって得られた文字認識結果の候補を統合することによって単独の文字認識では実現が困難な高い認識性能を実現する(図-3)。

文脈処理は、文章中に文字や二つの文字の並びが現れる頻度をあらかじめ大量の文章(新聞記事DBなど)で学習しておき、日本語としてもっともらしい文字の並びが優先的に選択されるように文字認識結果を修正する。

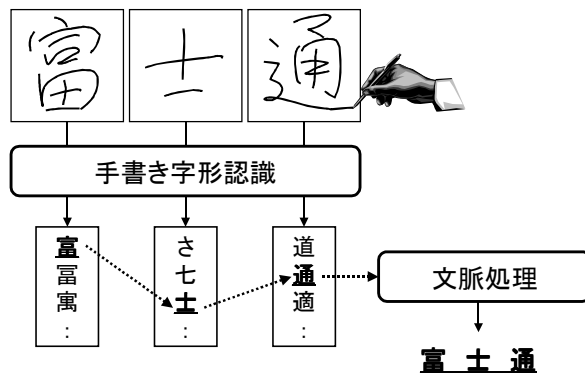


図-2 手書き文字認識技術
Fig.2-Handwriting recognition.

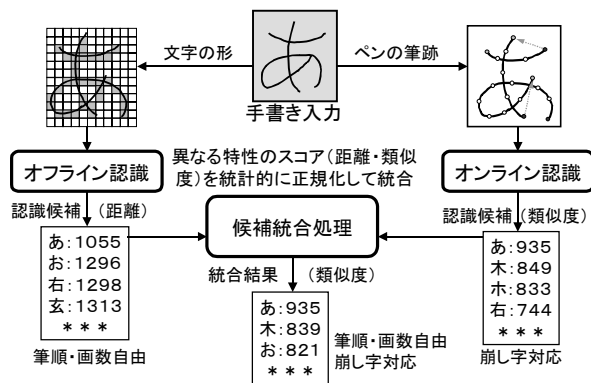


図-3 ハイブリッド型文字認識
Fig.3-Hybrid character recognition.

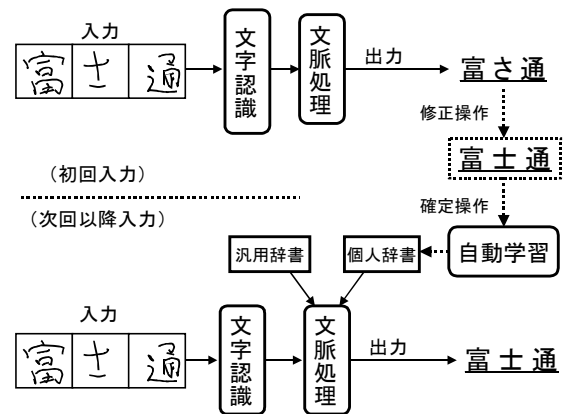


図-5 ユーザ適応技術(文脈適応)
Fig.5-User adaptation (context base adaptation).

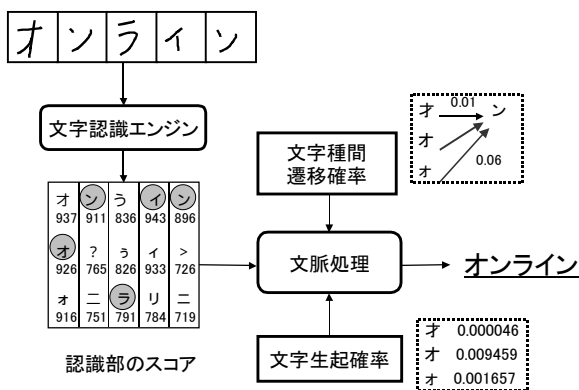


図-4 文脈処理
Fig.4-Context processing.

例えば図-4の例では、同じような形の文字(「オ」「オ」や「う」「ラ」)に誤って認識された部分が、前後の文字の並びを参照することにより正しい結果(「オンライン」)に修正されている。

以上の文字認識技術により、著者らの手書き文字認識システムは業界最高水準の高い認識性能を実現している(公開手書き文字データベースHANDS_kuchibue_d⁽³⁾で95.1%という高い認識率を達成)。しかし、ユーザが書く文字の形は同じ文字でも個人ごとに大幅に異なることがあるので、すべてのユーザに共通の認識方法では認識精度に限界がある。そこで著者らは、ユーザごとの書き癖を学習して更に高い認識精度を実現するため、ユーザ適応技術を開発した⁽⁴⁾⁻⁽⁶⁾。

ユーザ適応技術には、個人の文字の形状を学習する字形適応と個人がよく使う文字列を学習する文脈適応とがあり、著者らの文字認識技術では両者とも実装されている。その一例として図-5に文脈適応の例を示す。この例では、ユーザが書いた「富士通」という文字列が初回入

力では「富さ通」に誤っているが、それをユーザによる誤り訂正操作で訂正すると、このユーザが「富士通」という文字列を使ったという情報が記憶される。これにより次回からは「富士通」という文字の並びが出力されやすくなるため、同じ誤りを繰り返さない。

ユーザ適応技術には単に認識誤りを減らすだけでなく、同一の誤りを繰り返さないという効果がある。これは、特定のユーザがよく使う文字や文字列の認識精度を集中的に高め、ユーザの満足度を向上させる効果がある。

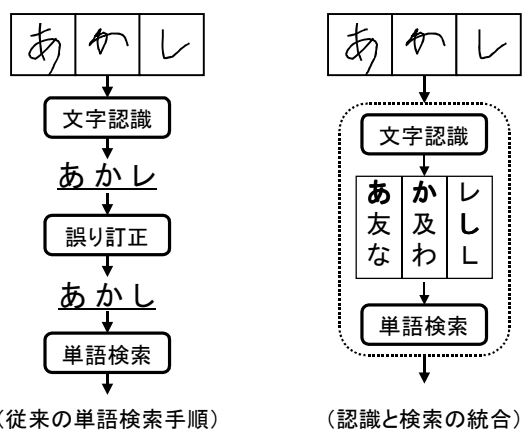
手書き単語検索・住所検索技術

業務分野でのテキスト入力では、駅名や商品名のように入力文字列が単語リストの形であらかじめ与えられていることが多い。そのような場合、ユーザが入力した文字列を検索キーとして単語リストの中から目的とする単語を選ぶ単語検索技術が有用である。

通常の単語検索では入力したい単語に含まれる文字列を指定して検索するため、手書き入力した検索文字列に認識誤りがあると検索の効率が落ちる。それを避けるために検索前に誤り訂正を行うと、認識・訂正・検索という三つのステップが必要となり、操作が煩雑になる。

著者らのシステムでは文字認識結果の候補も含めて単語検索処理を行うため、文字認識の結果が少々誤っていても精度の良い単語検索が可能である。また文字認識と単語検索を一体化することにより認識対象文字を検索単語の中に含まれる文字に限定することができ、文字認識の精度も向上する(図-6)。

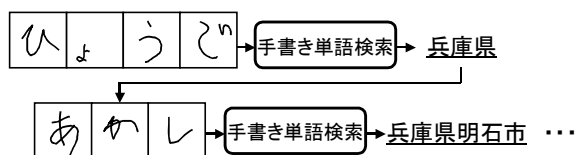
以上は単一の単語リストから手書き入力で目的単語を選ぶ場合の説明だが、住所や組織名のように検索対象データが階層構造になっている場合も手書き単語検索の



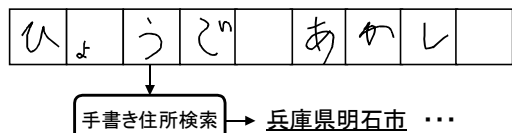
(従来の単語検索手順) (認識と検索の統合)

図-6 手書き単語検索

Fig.6-Word search using handwriting recognition.



(a) 単語検索を用いた住所検索 -住所階層ごとに単語検索-



(b) 手書き住所検索 -住所階層の区切りを空白で指示-

図-7 手書き住所検索

Fig.7-Address search using handwriting recognition.

技術を応用することができる。例えば住所入力の場合、都道府県名や市町村名を階層ごとに入力すれば複数回の単語検索で目的とする住所を入力することができる。しかし、最上位の都道府県名から順番に入力しなければならない点、途中の階層を省略できない点、各階層ごとに検索結果を確認しなければならない点など、必ずしも便利な入力方法とは言えない。

著者らは、住所のように階層構造を持ったデータから効率良く目的とする文字列を入力する技術を開発した。図-7の例にあるように、従来の手書き単語検索を用いると各階層ごとに別々に単語検索を呼ばなければならなかったが、著者らのシステムでは複数の検索キー（市町村名など）を空白で区切って指定することにより目的とする住所を一度で検索できる。途中の住所階層を省略して検索することも可能であり、検索キーはひらがなでも漢字でも入力できる。本方式は、各階層の地名検索では前方一致による文字列検索を用いているため、地名をこ

く一部だけ入力しただけでも検索が可能となっている。例えば「兵庫県明石市大久保町（ひょうごけん・あかし・おおくぼちょう）」という地名を入力する際、「ひょう・明・おお」などのように先頭の短い文字列を書くだけで、その組み合わせに一致する住所が検索できる。

以上2種類の検索技術により筆記量の削減が可能となり、業務効率の向上へ寄与できると著者らは考えている。

文字認識のActive-X部品化

著者らは文字認識技術が業務アプリケーション開発者にとっても手軽に使えるものになるよう、これまで述べてきた認識技術、検索技術に加えて誤り訂正などのユーザインタフェースもカプセル化したソフトウェア部品をMicrosoft Windows上で動作するActive-X部品の形で実装した。これにより文字認識技術のアプリケーションプログラム上での利用が容易になるとともに、我々ペン入力の開発者がこれまでに研究を積み重ねて得た、ペン入力に適した使いやすいインタフェースのノウハウをも容易に利用することが可能となった。

著者らが実装した文字認識Active-X部品は現時点で4種類存在する。すべての認識部品は既に述べたハイブリッド型文字認識技術を基本としており、任意の文字列の手書き入力を行うことができる認識部品として、文脈後処理技術とユーザ適応技術、および誤り訂正インタフェースを実装した汎用認識部品“HandWrite”がある。さらに、手書き単語検索技術を実装した単語検索部品“HandSearch”，手書き住所検索技術を提供する住所検索部品“HandAddress”，汎用文字認識技術を数字認識に特化し、認識対象文字は数字だけであるが、文字間を続けて書かれた場合でも認識できる連続数字認識部品“HandNumber”が存在する。

それぞれWindows上（9x/NT/2000など）で動作する。プログラマはVisual BasicやVisual C++といった開発環境を用い、画面上に認識部品を貼り付けるだけでこれらの認識機能やインタフェースが利用できる。またInternet Explorer上のJavaScriptあるいはVBScriptから呼び出すことにより、Webページ上での利用も可能である（現在はHandWriteとHandNumberのみだが、ほかの二つも対応可能）。

汎用文字認識部品“HandWrite”

汎用認識部品は任意の文字列を手書き入力できる認識部品である。図-8の画面例にあるように、手書き入力領域に文字を筆記し、その認識結果が同じ領域に表示され

る。もし認識結果が誤った場合にはその文字をペンで選択すると、認識候補のリストがポップアップで表示される。ユーザはその中から正解文字を選んで修正すればよい(図-9)。

図中に三つの領域があるが、それぞれが別個の汎用認識部品である。個々の領域ごとに認識対象文字種(英字, 数字, 平仮名, 漢字など)や動作モード(認識候補選択/上書き, 編集ベルトの有無, タイムアウト設定など), 色(筆跡色, 文字色, 背景色など)など様々な属性の設定が可能である。

単語検索部品 “HandSearch”

単語検索部品はあらかじめ設定した単語リストの中から単語を検索する部品である。図-10の例のように、鉄道の駅名を入力する場合や大量の商品名の中から目的とする商品を探し出す場合など、入力したい単語の集合が

あらかじめ分かっている場合に利用できる。文字認識の段階から検索機能を利用しているため、入力文字列の一部が認識できない場合でも目的単語を見つけることができる。

住所検索部品 “HandAddress”

住所検索部品は内部に全国の住所辞書を持っており、住所の階層構造を利用して効率良く住所を入力する部品である。市町村などの各階層の地名を一部入力するだけで目的とする地名がリストから選択できる様子を図-11に示す。これにより、長い地名をすべて筆記しなくても住所が入力できる。

手書き入力は初心者に分かりやすい反面、長い文字列を書くとき手が疲れるという問題がある。またキーボードが使えない場合によく利用されるため立った状態など不安定な姿勢での利用も多い。そのため、短い筆記で迅速



図-8 汎用文字認識部品

Fig.8-General purpose handwriting recognition module.



図-10 単語検索部品

Fig.10-Handwriting word search module.



図-9 認識候補選択

Fig.9-Recognition candidate selection.

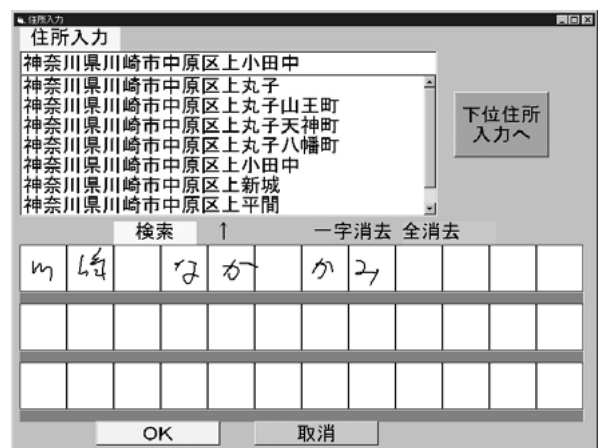


図-11 住所検索部品

Fig.11-Handwriting address search module.

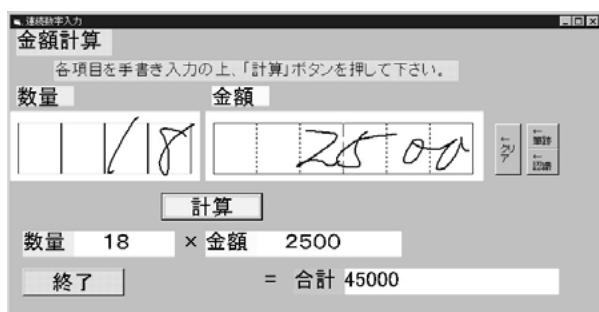


図-12 連続数字認識部品

Fig.12-Handwriting number recognition module.

に入力できるインターフェースが求められる。以上のような意味で、住所検索部品はペン入力にとって重要な特徴を有していると言える。

連続数字認識部品 “HandNumber”

連続数字認識部品は、連続数字を含む数字文字列が高い精度で認識可能な認識部品である。これは図-12にあるように複数の数字が連続して書かれたり、文字枠をはみ出して書かれた場合でも高精度で認識する。

現在提供している4種類の認識部品はいずれも文字枠内に一文字ずつ筆記するタイプの認識部品だが、このような筆記制限（枠をはみ出して書いてはいけない）は文字認識の精度を高めるために有効である反面、使い勝手の低下にもつながる。とくに領収書に金額を書く場合に見られるように、数字は続けて書かれたり記入枠をはみ出して書かれることが多い。そこで、認識文字種を数字に限定することで高い精度を保ちながら、筆記制限を大幅に緩和したのがこの認識部品である。

文字認識部品を用いたペン入力の応用

ペン入力インターフェースにはもちろん長所も短所もある。文字認識部品がどのように業務に活用できるかを考えるためには、この長所と短所を踏まえた上で適切な応用を考えるのが望ましい。本章では、まずペン入力の長所・短所について述べ、さらに具体的な適用分野について考察する。

すでに述べたようにペン入力は片手で操作できるため、機器を固定できないような場面、例えば立ったまま使わなければならないような場合でも使うことができる。またマウスのようなポインティング操作とテキストを入力する操作が同じ1本のペンで行われるため、キーボードとマウスのように持ち替える手間が省け、スムーズな操作が可能である。このようなポインティングとテキスト入力のシームレスな結合は複数の入力領域にランダムに

情報を入力するような場合（領域選択と入力が頻繁に行われる）に有用である。さらに、ペン入力はキーボードに比べて操作の習得が簡単であり、初めて使う操作者でもわずかな時間の練習で利用できるようになる。

以上の長所に対し、手書き筆記には手が疲れるため長いテキストの入力には向かない、文字認識の誤り訂正が必要である、文字を筆記するためある程度の広さの領域が必要である、などの欠点が存在する。また、ペンでの操作はキーボードに比べて習得が容易だとはいうものの、初めて使う人はディスプレイ上にペンで書くという経験がないため、力の入れ具合などの感覚がつかめず若干戸惑うこともある。さらに、世の中にペン入力が十分に普及しておらず、標準的なインターフェースが存在しないことも初心者が戸惑う理由の一つとして挙げられる。以上をまとめると次のようになる。

- 長所：
 - ・ 機器を手にとって片手で使える
 - ・ 持ち替え不要（シームレスな操作性）
 - ・ 初心者が習得しやすい
- 短所：
 - ・ 手が疲れる。長いテキストには不向き
 - ・ 誤り訂正操作が必要
 - ・ 画面上に筆記領域が必要
 - ・ 初めて使う人には少し違和感がある

片手で操作できるという特徴は、オフィス外での作業中にその場で情報を入力するような業務に活用できる。例えば倉庫で商品を見ながら在庫確認をする場合や、顧客先で見積り伝票の入力・修正を行う場合などが考えられる。

シームレスな操作性は複数の入力項目を持つ伝票入力に有用である。

初心者にも習得しやすいという特徴は、例えばコンビニエンスストアやスーパーのように、アルバイトなどの非熟練者が頻繁に入れ替わる業務では非常に重要である。

将来的には、一般利用者が自分で情報を入力するような場合にも便利に使うことができるのではと期待しているが、上記で述べたように初心者には戸惑いもあり、現状では少し難しい。さらに使い勝手の改善を進める必要がある。

む す び

本稿では、キーボード・マウスに代わる入力手段として新たな可能性を持つペン入力手書き文字認識技術について紹介した。

ペン入力はモバイル環境での利用に適しており、また

初心者にも分かりやすいと言われている。しかしペン入力をキーボードに代わる入力手段として利用するためには高い精度の手書き文字認識技術が必要である。また手書き文字認識技術を有効に活用するためには、分かりやすいインタフェースと開発環境の整備が必要となる。著者らはこのような視点からペン入力手書き文字認識部品の開発を行っている。

本稿で述べたように、著者らは4種類の文字認識部品をWindows上のActive-X部品として実装した。これにより、アプリケーション開発者は高精度な手書き認識機能を容易に活用することができる。手書き認識機能は、とくにオフィス外で情報を入力する業務やアルバイトなどの非熟練者が多い業務に向けた業務アプリケーションにおいて有用であるが、一般ユーザが直接扱うようなKIOSK端末やATMなどで活用できるまでには至っていない。

今後はこれら四つの文字認識部品の改良を行うとともに、様々な機能を持つ認識部品の開発も必要だと思われる。一例を挙げれば、最近では文字枠を使わないで文字を手書き入力する枠なし認識技術が注目されており、枠なし認識機能を持った認識部品の開発も必要であろう。それに加え、本稿で述べたような業務目的だけでなく一般ユーザにも便利なペン入力技術の開発も進めていきた

いと考えている。

参考文献

- (1) 秋山勝彦ほか：オンライン手書き文字日本語認識のための線形処理時間伸縮マッチングアルゴリズム. 信学論, Vol.J81.D, No.2, 電子情報通信学会, p.651-659 (1998.4).
- (2) Hiroshi Tanaka et al.: Hybrid Pen-Input Character Recognition System Based on Integration of Online-Offline Recognition. Proc. ICDAR'99, p.209-212 (1999.8).
- (3) 中川正樹ほか：文章形式字体制限なしオンライン手書き文字パターンの収集と利用. 信学技報, PRU95-110, 電子情報通信学会, 1995.9.
- (4) 秋山勝彦ほか：オンライン手書き文字認識のためのテンプレートキャッシングによる筆者適応手法. 信学技報, PRMU 2000-210, 電子情報通信学会, p.69-76 (2001.3).
- (5) Naomi Iwayama et al.: Adaptive Context Processing in on-line Handwriting Character Recognition. Proc. 7th IWFHR, p.469-474 (2000.9).
- (6) 岩山尚美ほか：誤認識した字形と使用した文字列を学習する個人適応型オンライン手書き文字認識. 情処技報, 2001-HI-96-8, 情報処理学会, p.55-62 (2001.11).