

オンライン文字認識による日本語入力

吉田 和永, 亘理 誠夫, 迫江 博昭, 清水 洋, 上原 和彦

(日本電気㈱)

1. まえがき

計算機による日本語情報処理が一般化されてきており、日本語入力方式に関してさまざまな方式が提案されているが¹⁾、いまだに決定的なものがないのが現状である。日本語ワードプロセッサの入力方式では、かな(又はローマ字)キーボードとかな漢字変換による方式が主流となりつつあるようである²⁾。しかし現実にはキーボードに不慣れな人やキーボードを好まない人も多く存在するので、日本語入力方式としてすべてこれで十分であるとはいえない。また、窓口等における客によるデータ入力のように、全くの初心者がデータを入力する場合にはキーボードによる入力は無理があると思われる。

オンライン文字認識は、文字を書くときのペン先の動きをもとに文字を認識する技術である³⁾。文字を書くことはだれでも可能であるので、オンライン文字認識はだれにでも扱える日本語入力方式とすることができる。また、考えながら文章を作る場合は文字を書きながら作るのが普通である。これは文字を書くことが日本人にとって極めて自然な行為であるので、文字を書くことにより思考がさまたげられないためと思われる。

また、オンライン文字認識はタブレット上に文字を書くため、文字情報と同時に、文字が書かれた位置の情報も入力できる。さらに、このタブレットから図形等を入力することも可能である。これらの特徴により、オンライン文字認識を用いれば、使いやすい校正システムや図形入力システムを構成できる可能性がある。

オンライン文字認識の入力速度は文字を書く速度によって制限されるため、熟練者のタッチ・タイプによるキーボードほどには高速に入力することは難しい。しかしキーボード入力の場合でも原稿は手書きにより作成されることを考慮すると、原稿作成の段階で直接オンライン

文字認識を用いて入力するならば全体としての入力時間を短縮させることも可能である。

このように、オンライン文字認識は情報の発生時点における情報の入力に適しており、日本語入力手段として極めて高い可能性を持っていると言える。本論文では、日本語入力手段としてみた場合のオンライン文字認識の特徴に関して考察したのち、我々が開発したオンライン文字認識装置⁴⁾について述べる。この認識装置はスタックDPマッチング法⁵⁾と呼ばれるアルゴリズムを基本的な認識処理に用いており、漢字・ひらがな・カタカナ・英数字・記号を認識することができる。最後に本装置の応用例について述べる。

2. オンライン文字認識の特徴

2.1 オンライン文字認識の位置づけ

ここでは、オンライン文字認識を、他の入力方式との比較により位置づける。現在使用されている日本語入力方式には次のようなものがある。¹⁾

- a) 漢テレ方式
- b) 2ストローク方式
- c) ペン・タッチ方式
- d) かな(ローマ字)漢字変換方式
- e) 光学的文字認識方式(OCR)
- f) オンライン文字認識方式
- g) 音声入力方式

これらの入力方式のうち、漢テレ方式、及び、2ストローク方式は高速入力が可能である。しかしこれらを使いこなすためには、熟練を要するので専任オペレータ向きの入力方式であるといえる。そこで、初心者でも扱えると考えられる(C)のペン・タッチ方式以下の方式に関してオンライン文字認識方式と比較する。

(1) ペン・タッチ方式、かな漢字変換方式との比較

ペン・タッチ方式やかな漢字変換などのキーボードを用いた入力方式と比較すると、オンライン文字認識の特徴は、人間にとって自然な入力方式であるという点にある。

ペン・タッチ方式では文字を入力する場合、多数の文字の中から目で見ることにより必要な文字を探さなければならない。この「探す」という作業は、使用者にとって心理的に大きな負担となる。

かな漢字変換方式では使用するキーボードはカナ・キーボードでよいので、タッチ・タイプ⁶⁾することも可能である。しかしタッチ・タイプができるようになるためには、ある程度の練習が必要であるので、初心者や、まれにしか使用しないという人はカナ・キーの中から「探す」作業が必要になる。この「探す」作業を少しでも軽減するために五十音配列のキーボードやローマ字入力がある。五十音配列は、はじめからタッチ・タイプは諦めている方式といえる。ローマ字入力は入力にカナ・キーボードより普及している英文キーボードが使えるが、ストローク数はおよそ2倍必要となる。また、かな漢字変換方式ではキーボードによる入力のほかに、変換結果の中から同音異字を選択するという特別な作業が必要である。

この点、オンライン文字認識では「探す」という作業は必要ない。複雑な漢字などを書く場合は手元を見るために目を使うこともあるが、注視する必要はないので心理的にそれほど負担とはならない。ただ、認識率が低い場合は認識結果の確認という作業が必要であり、使用者の負担となる可能性もある。

(2) 光学的文字認識装置（OCR）との比較

オンライン文字認識も（手書き）OCRも共に、文字を書くという人間にとって自然な動作により文字を入力するという点においては似ているが、その性格は異なる。OCRは、人間が文字を書く場所とは異なる場所でも時間的にも遅延をもって動作するという点においてオ

プライン的な入力装置である。これに対しオンライン文字認識は、キーボードなどと同様に入力作業と同期して動作する入力方式である。言い換えれば、OCRがセンター・バッチ型の入力方式であるのに対して、オンライン文字認識は分散処理型の入力方式であるといえる。このような性格の違いから、具体的には次のような特徴の差が生じる。

OCRでは大量のデータを処理することが多いので高速な動作が要求される。これは文字を書くための時間のほかに、認識処理時間が別に必要であるためである。一方、オンライン文字認識では文字を書く動作に追隨して認識処理ができればよいので動作は比較的低速でよい。既に清書された原稿がある場合、入力時間はOCRのほうが早い可能性があるが、原稿作成のための時間もふくめた全体ではオンライン文字認識のほうが効率がよいといえる。

センター・バッチ処理という性格からOCRは比較的大型の装置であってよい場合も多いが、分散処理型の使用形態が多いと考えられるオンライン文字認識は、キーボードなみの小型化が要求される。また、オンライン文字認識は入力のために小型のタブレットがあればよいが、OCRでは光学系や紙送り機構などのメカが必要であるため小型化は難しい。オンライン文字認識は電気回路の小型化が実現すれば超小型化も可能である。

OCRが文字データを機械に入力するという一方通行的なマンマシン・インタフェースの性格が強いのに対して、オンライン文字認識は質問応答システムや学習システムなどの会話型マンマシン・インタフェース用としても高い適性を持っている。また、オンライン文字認識では入力にタブレットを用いているので、図形入力との複合処理が可能である。

認識技術的な立場からみると、オンライン文字認識は、細線化処理が不用なこと、筆順情報がある程度は使えることなどOCRと比べて有利な点も多い。このため実現性はオンライン文字認識のほうが、高いといえる。

(3) 音声入力方式との比較

音声は人間にとって最も自然な情報伝達手段であるが、日本語文の入力手段として考えた場合、オンライン文字認識と比較して次のような特徴を持つ。

音声入力方式の本質的な欠点は、入力のために声を発しなければならないという点にある。この声は周囲に対しては騒音となり、かつ機密の内容の入力に際しては盗聴の問題も出てくる。逆に音声入力は、現在の技術では周囲の騒音により誤認識が生じやすくなるという欠点をもつ。オンライン文字認識では、このような問題は全く生じない。

音声入力の入力速度は、オンライン文字認識と比較して高いと考えられている。しかし現在日本語文の入力手段として使用できる音声入力方式は単音節認識の技術を用いたものであるため⁷⁾、入力する場合は単音節単位に区切って発声しなければならないので、入力速度はオンライン文字認識と比較して大差ないとも言える。連続発声した日本語文が認識できるようになれば入力速度の点で極めて優位になるが、連続音声の認識は現在まだ研究段階にすぎない。⁸⁾

また、音声により入力できるのは「読み」の情報のみであるので、日本語文を入力するためには、かな漢字変換が必要である。オンライン文字認識では漢字を直接入力できるので、かな漢字変換の必要はない。

このように、オンライン文字認識は他の入力方式にない数々の利点をもつ入力方式といえる。

2. 2 比較評価実験

入力方式の使いやすさは、実際に使用してみなければわからない面が多い。我々は、初心者を対象とした各入力方式の比較評価実験を行ったので⁹⁾、以下に結果を述べる。

(1) 実験対象

被験者は日常データ入力を行っている女性5名で、一部に英文タイプの素養のある人をふくむが、カナ・タイプの経験者はいない。評価用文章は100文字からなる

新聞の社説文である。

(2) 入力速度および入力誤り数の比較実験

実験結果を図1に示す。①～⑤及び⑩、⑫は前記文章をかな文字入力した場合（入力誤りを許す）であり、⑥～⑨及び⑪は前記文章を漢字かなまじり文で入力した場合（誤り訂正を行う）である。⑤～⑨の入力方式では、かな漢字変換を併用した。⑩～⑫は模擬的な実験であり、完全に認識できることを想定している。

②、⑦はオンラインひらがな認識であり、筆跡中の認識対象文字の分割は1文字入力することに終了キーを押す方式を用いている。この方式では、次の文字を入力するためには認識結果がでるまで（約500msec）待たなければならない。⑩、⑪もオンライン文字認識を想定したものであり、認識結果を待たずに文字を書いた場合を示す。たとえば、実装置に1個の文字枠中に書かれた筆跡を1つの認識対象文字として分割する方式を用いれば、その入力速度は⑩、⑪の場合と等しくなる。

この実験は、全くの初心者を対象としたものであるが、音声入りに次いでオンライン文字認識が高速であるという結果が得られている。

(3) 主観評価実験

前記実験において、かな文字入力の場合（①～⑤）の実験の終了時に、被験者に原稿の見易さ、使い易さの2項目に関して順位（1～5）をつけてもらった。この結果によると（表1）、オンライン文字認識は音声入りに次いでローマ字入力と同順位となっている。

ローマ字入力が高い順位を得ているのは、被験者が英文タイプに慣れているためと思われる。漢字かなまじり文の入力の場合は、かな漢字変換がローマ字入力では必要なので、オンライン文字認識のほうが優位になる可能性がある。

また音声入力との比較では、この実験では音声入力のほうが好まれているという結果がえられたが、音声入力が周囲環境の影響を受けやすいことを考慮すると、実際のオフィス環境での評価は異なる可能性がある。

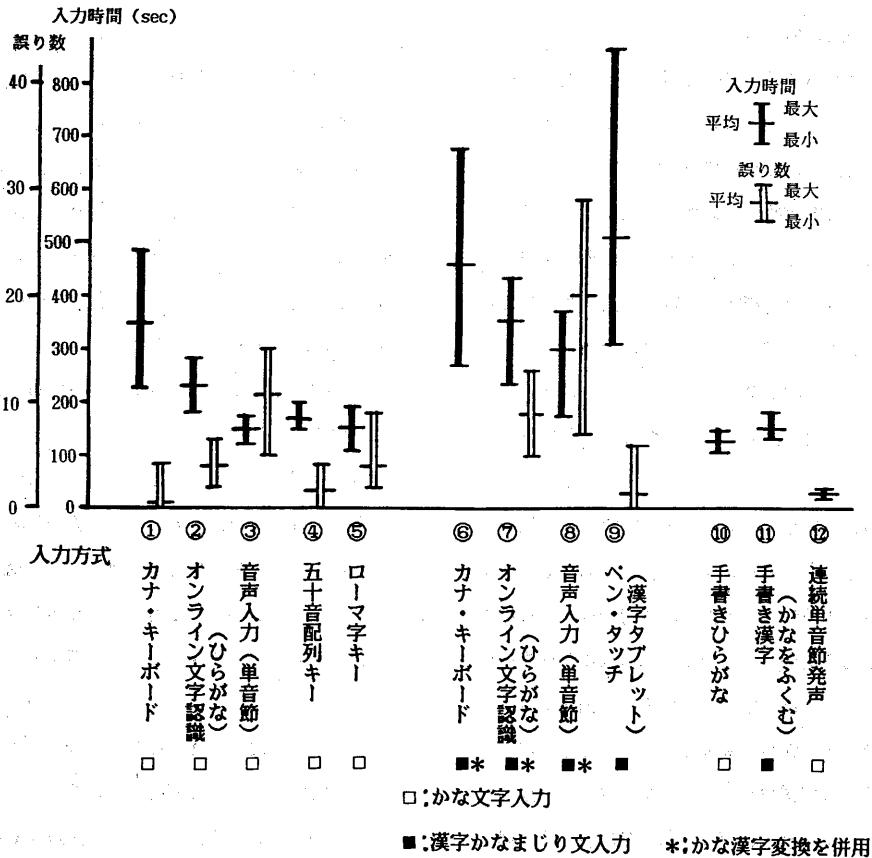


図1 文書入力時間と誤り数

表1 主観評価結果

入力方式	原稿の見易さ	使い易さ
① カナ・キーボード	5.0	4.8
② オンライン文字認識	2.6	2.8
③ 音声入力	1.2	1.6
④ 五十音配列キー	3.0	3.0
⑤ ローマ字キー	2.6	2.8

(5人の平均)

3. オンライン手書き文字入力ターミナルの開発思想

これまでに検討してきたオンライン文字認識の特徴をふまえて、日本語入力用にオンライン手書き文字入力ターミナル⁴⁾を開発した。この装置は次に述べる開発思想に基づいたものである。

(1) 認識対象字種

認識可能な字種は多ければ多いほどよい。しかし一般に、字種を増やすと認識率の悪化や認識処理量の増大をまねくので、ただ増やせばよいというものではない。そこで、本装置で認識可能な漢字は、新聞で使用される漢字の99%以上を含む常用漢字(1945字)を基本とし、外字登録によりJIS第1水準の文字(2965字)まで拡張できるようにした。この他に、ひらがな・カタカナ・英数字・記号も認識可能とした。

(2) 筆記者制限

オンライン文字認識においても、特定話者の音声認識で行われているように、あらかじめ使用者個人の筆跡を標準パターンとして登録しておくという方式もあり得る。しかし、本装置では認識対象字種が多いため、あらかじめ全ての文字を登録する特定筆記者方式は実用的でない。そこで、本装置は誰でもすぐに使用できる不特定筆記者方式とした。

(3) 文字品質に対する許容度

文字の品質に関しては主観がはいることが多く、一概に論ずることは難しい。まず、字体が異なると文字の形状も大きく異なる。本装置では一般的な楷書体を対象とした。しかし、このような制限を加えても実際は、画数や筆順が正しく守られるとは限らない。

楷書は1画（ストローク）ごとに区切って書くのが正しい書き方であるが、特に文字を早く書いた場合、正しくは別れるべきストロークが連続してしまうことが多い。入力方式としては、早く入力できることは重要であるので、本装置ではストロークが連続した文字も認識できるように考慮した。

また、筆順に関しても、正しい筆順の基準は存在するが、漢字などでは実際はさまざまな筆順で書かれる。しかし、利用者が協力的であるなら極端な筆順変動は除外できると考えられるので、本装置では常識的な筆順変動に対処できれば一応十分であるとした。

(4) 認識時間

認識時間は、筆記速度に十分追従できる程度に短くなければならない。少なくとも2秒/文字以下を目標とした。

本装置では、以上述べた開発思想を満足するような認識方式及び装置構成を用いた。その内容については、次章以下で述べる。

4. 認識方式

本装置の認識方式に関して述べる⁵⁾。本装置では、字

形の歪みの他に筆順の変動やストローク間の連続が生じた文字をも効率よく認識するために開発した、スタックDPマッチング法とよばれるアルゴリズムを使用している。図2に認識処理手順を示す。以下に個々の処理に關し順に説明する。

4.1 前処理

タブレット上に書かれた文字の筆跡は、ペン先の位置座標およびペン先のON/OFF状態のデータとして等時間間隔でサンプリングされる。このサンプリングは、早く書かれた文字の細部の形状を保存するために比較的短い間隔で行われるが、直線部分やゆっくり書かれた部分においては冗長な点を多く含むことになる。認識処理の効率を上げるため、これらの冗長なサンプル点を除いて文字の形状を保存する最少のサンプル点を選択する。このようなサンプル点の選択は、ダイナミック・プログラミング（DP）法を用いて効率よく行うことができる。

選択されたサンプル点列は、隣接するサンプル点間を結ぶベクトルの時系列に変換される。この時、ペン先がOFF状態の部分も1つのベクトルとして変換することにより、文字全体を一筆書きで書かれたようなパターンと

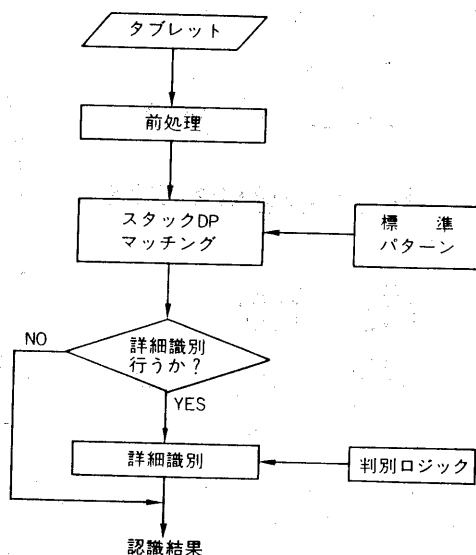


図2 認識処理手順

して表現する。これにより、ストロークの連続により変形した文字に対処することができる。以下では、このベクトルの時系列を入力パターンと呼ぶ。

4.2 スタックDPマッチング

前処理の結果得られた入力パターンは、認識装置内にあらかじめ保持されている標準パターンとの間でパターン・マッチング処理が行われることにより認識される。筆順変動などの文字の大きな変形は、この標準パターンの中にパターンの枝分かれとして表現されている。

図3に、「希」という文字の標準パターンの表現の一例を示す。通常、「希」という文字を書く場合「メ」、「ナ」、「巾」という順序で書かれ、この順序が異なる筆順はまず起こり得ない。しかし、各部分の筆順は第3図に示すように、おのおの2通りずつある。これらの筆順変動をおのおの別の標準パターンとして用意すれば、8通りの標準パターンが必要となるが、このように枝分かれパターンとして表現することにより、全体としてコンパクトに表現することができる。

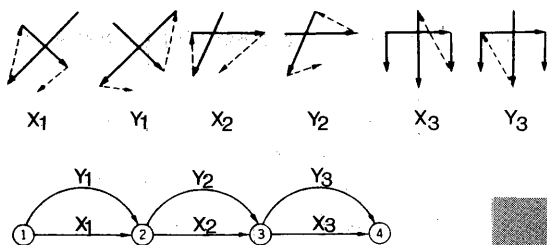


図3 「希」の標準パターン表現

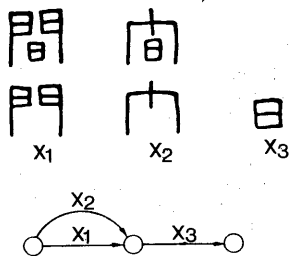


図4 略字の標準パターン例

また、図4に示すように、筆順変動のみでなく略字による変形も枝分かれパターンで表現することにより効率よく表現することができる。

枝分かれパターンとして表現された標準パターンと入力パターンとの間のパターン・マッチングは、スタックDPマッチング法を用いることにより効率よく実行できる。この方法はDPマッチング法を発展させたものであり、DPマッチング法自身のもつ時間軸正規化能力により文字の微細な歪みなども吸収できる。

4.3 詳細識別

前記のスタックDPマッチング法においては、筆点の移動方向ベクトルのみ注目して認識を行っている。このため「人」と「入」や「土」と「土」などの文字は、この方法だけでは識別が困難である。そこで、このような文字については、文字中の特徴部分に注目した詳細識別をトップ・ダウン的に適用することにより認識する。

5. 装置概要

ここでは本装置のハードウェアについて述べる。本装置のハードウェアは、タブレット、マイクロ・プロセッサとその周辺回路、外部インターフェース回路、表示器、電源から構成されている。図5に本装置の外観を示す。この装置の中には、認識処理部も含まれている。

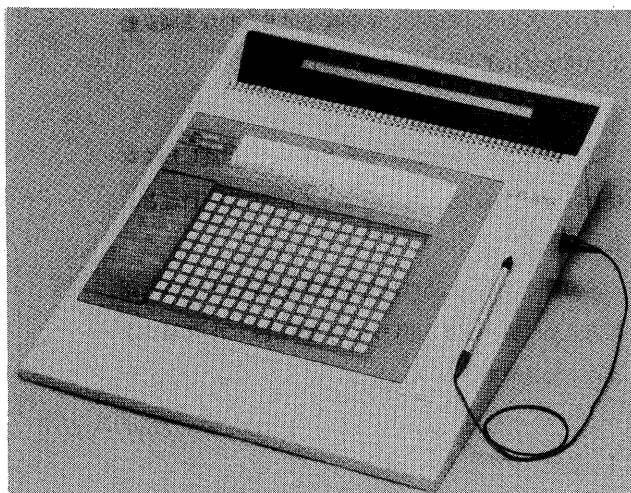


図5 オンライン手書き文字入力ターミナル

5.1 装置構成

本装置の構成を図6に示す。各部分について以下に説明する。

(1) タブレット

このタブレットから、文字、図形および各種ファンクション情報が入力される。座標検出方式は電磁結合位相検出方式であり、文字等は専用ボールペンを用いて入力する。記入面は、およそA4版の大きさである。

記入面上は、文字（または図形）を記入するための文字記入エリアと、各種ファンクション情報を入力するためのファンクションキーエリアとに分割されている。文字の入力は、文字記入エリア内の記入枠中に一文字ずつ記入することにより行う。記入枠の大きさは約10mm×10mmであり、この記入枠は文字記入エリア内に16×11個ならんでいる。ファンクションキーエリアには、最小5mm×5mmのキーが32×8個ある。このキーからは、それぞれ異なるコードが出力される。

(2) マイクロプロセッサとその周辺回路

マイクロプロセッサには、16ビット系のものを1個使用している。認識処理をはじめ、タブレットの制御や表示器の制御も、この一つのマイクロプロセッサで行っている。認識処理は1文字あたり、最大2秒以下で実行できる。

(3) メモリ

使用目的により、プログラム用メモリ（認識処理プログラム、通信制御プログラム等のアプリケーション・プログラムを含む）、標準パターン用メモリ、ワーク用メモリの3種のメモリに分かれている。プログラム用メモリおよび標準パターン用メモリはROMを、ワーク用メモリはRAMを用いている。

(4) 表示器

表示器は液晶ディスプレイであり、16×16ドットの文字を1行40文字表示できる。表示器は必要に応じて、本体からの着脱が可能である。

(5) インタフェース

インタフェースは、EIA RS232C（200～19200bps）およびセントロニクス・インタフェースをおのおの1ポートずつ持っている。RS232Cインタフェースは、半二重調歩同期式で、ホストとの通信ができる。セントロニクス・インタフェースはプリンタ用の出力ポートで、最大40行の編集された文書を出力できる。

5.2 ユーザ・インタフェース

ここでは、認識の本質的な部分以外で、実際に使用する上で装置の操作性などを左右するいくつかの点について述べる。

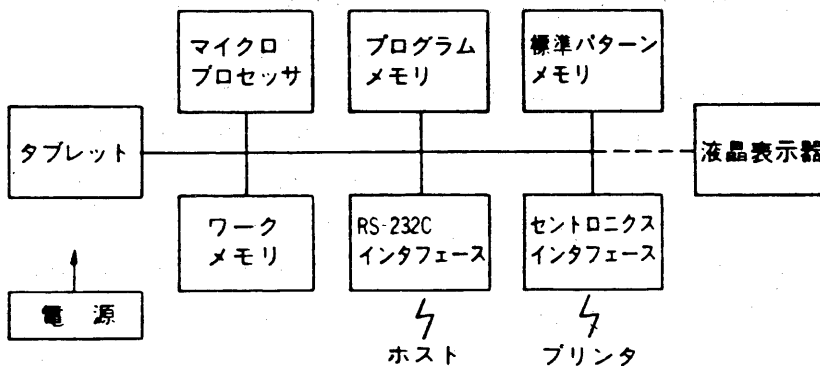


図6 装置構成

(1) 入力モード

本装置で同時に認識できる字種は、以下の4種に分かれている。

- ① 漢字・ひらがな（「、」、「。」を含む）
- ② カタカナ（「ー」長音記号を含む）
- ③ 英数字
- ④ 記号（31種類）

これらの字種は、モード切り替えキーにより選択する。このように字種を分けたのは、漢字の「力」とカタカナの「カ」などのように文字の形だけでは区別が困難な文字があるためである。入力したあとで字種を指定する方法もあるが、操作を単純にするためにモード切り替え方式とした。字種は、モード切り替えキー以外にも、ホストからのコマンドにより選択することもできる。

ここに示した4種の字種を認識する文字入力モード以外に、本装置ではイメージ・モードなるモードがある。これは文字を認識するかわりに、ペン先の位置のデータを出力するモードである。これにより、図形や位置の情報をタブレットから入力することが可能となる。

(2) 文字セグメンテーション方式

筆跡を1文字毎に分割することを文字のセグメンテーションと呼ぶ。文字の認識は、分割された筆跡毎に行われる。一般には次のような方式がある。

- ① 文字枠方式： タブレット上の文字枠内に文字を書き、ペンが次の文字枠に移動した場合、前の文字を切りだす方式。
- ② 終了キー方式： 1文字が書き終わった場合、終了キー（タブレットのファンクションキーエリアにある）を押すことにより文字を切りだす方式。
- ③ タイム・アウト方式： ペンが、ある一定時間以上タブレットから離れた場合、文字を切りだす方式。

本装置では以上の3種の方式を併用している。

(3) 出力データ

本装置から出力されるデータには、次に示すものがある。

- ① タブレット上に筆記された文字の認識結果。JIS 2 バイトコードで出力される。
- ② 文字の書かれた位置。文字記入エリア内の記入枠のアドレスとして出力される。
- ③ タブレット上のペンの位置。このデータは、イメージ・モードの状態の時、出力される。
- ④ ファンクションキーエリア内の、ペンで示されたファンクションキーのコード。

5. 3 使用形態

本装置は、日本語入力の様々な分野で使用できるよう、次に示すような3種類の使用形態がとれるようになっている。

- ① ディスプレイを有するインテリジェント端末の入力部として使用する。
- ② ホスト・コンピュータに接続して、端末として使用する。この場合、表示部をディスプレイとして用いる。
- ③ 単体で、簡易ワードプロセッサとして使用する。本装置は、簡単な編集機能を有している。作成された文書は、セントロニクス・インタフェースに接続されたプリンタに出力することができる。

6. 応用

本装置の持つ機能により、本装置はたとえば次のような分野への応用が可能である。

(1) 日本語ワードプロセッサ

日本語ワードプロセッサにおいては、入力方式の使い勝手以外にも、編集のしやすさが全体の使いやすさをきめる。最近ではカーソルの移動やファンクション情報の入力に、マウスなどの座標入力手段を使用したものがある。このようなシステムを実際に使用してみると、テキスト入力時と座標入力時で、キーボードとマウスを使い分けなければならない、煩わしく感じることもある。

この点、入力方式にオンライン文字認識を用いればテキスト入力と座標入力が同一のタブレットからできるの

で、このような煩わしさはない。

(2) データ・エントリ業務

文書入力とは異なるが、人名や地名などのデータ入力にもオンライン文字認識は有効である。

たとえば、電話による顧客からの注文情報を計算機に入力する場合、キーボードからの入力はキーボードに対して神経を集中させる必要があるため、顧客と会話しながらの入力は難しい。自然な入力方式と言われている音声入力も、入力するためには声を出さなければならないので電話を使用しながらの入力は不可能である。このような場合は、オンライン文字認識が有効である。

窓口などにおいて、顧客より直接データを入力してもらう場合も、熟練を必要としない自然な入力方式であるオンライン文字認識が有効と思われる。

7. あとがき

情報の発生時点における入力手段に適しており、日本語入力手段として極めて高い可能性を持っていると思われるオンライン文字認識について、その特徴を論じ、我々が開発したオンライン手書き文字入力ターミナルについて述べた。本装置のように様々な分野に応用可能な装置を示すことにより、新しい応用分野が開ける可能性がある。

計算機による日本語情報処理の進展に伴い、入力手段としてオンライン文字認識が広く使われることが期待される。今後は、認識率の向上など装置性能のより一層の向上を進めていきたい。

謝辞 日頃御指導いただく当研究所メディアテクノロジー研究部の千葉部長、ならびに装置の開発で御協力いただいた当社伝送通信事業部第一開発部の皆様に感謝致します。

参考文献

- (1) 高橋：「日本語入出力装置」，情報処理，Vol.20，No.6，p.933-940（昭54-10）。
- (2) 佐藤，小橋：「日本語ワードプロセッサに関する技術動向」，情処学会日本文入力方式研究会資料，7-1，（昭57-10）。
- (3) 迫江，吉田：「日本語文書入力に対するオンライン文字認識技術の応用」，電気学会情報処理研究会資料，IP-83-1，（昭58-02）。
- (4) 吉田，伊東，小林，佐藤：「オンライン手書漢字入力ターミナルCR-100」，NEC技報，Vol.37，No.5，p.14-17（昭59-05）。
- (5) 吉田，迫江：「スタックDPマッチングによるオンライン手書き文字認識」，電子通信学会技術研究報告，PRL83-29（昭58-09）。
- (6) 山田：「日本語テキスト入力法の人間工学的比較」情処学会日本語情報処理シンポジウム1（昭53-07）。
- (7) 井上：「音声日本語ワードプロセッサVWP-103Nモデル2の音声入力方式」，情処学会日本文入力方式研究会資料，14-3，（昭59-03）。
- (8) 渡辺：「連続音節認識の一方式」，日本音響学会講演論文集，2-7-3，（昭57-10）。
- (9) 吉田：「日本語ワードプロセッサに対する各種入力手段の比較評価」，情処学会第24回全国大会，2G-3，（昭57-03）。