

遅延認識方式による 手書き原稿作成プロトタイプシステム

曾谷俊男[†], 福島英洋^{**}, 高橋延匡, 中川正樹
東京農工大学 工学部
東京都小金井市中町2-14-16

考えながらの入力(創造入力)を念頭においたオンライン手書き入力による文書作成システムの基本設計とそのプロトタイプについて述べる。従来見られるほとんどの報告では入力を即座に認識している。しかし、これでは認識結果の検討, 誤認識の訂正などにより思考の中断が生じる。筆者らはこれを避けるため, 入力パターンを即座には認識しない(あるいは表示を遅らせる)遅延認識方式を考案した。また, 手書きパターンを一時的な表現とは考えず, 手書きパターンをも文書表現とするため, パターンとコードの2つの文書表現形式を採用した。本基本設計により, 思考の中断のない入力環境を実現できる。

A prototype of a document preparation system through handwriting employing lazy recognition scheme.

Toshio Souya[†], Hidehiro Fukushima^{**}, Nobumasa Takahashi, Masaki Nakagawa
Faculty of Technology Tokyo University of Agriculture and Technology.
2-14-16 Naka-cho, Koganei, Tokyo, 184, Japan.

A design of a document preparation system through handwriting with consideration of creative writing and its prototype are described. All the conventional handwriting recognition interfaces employ the real-time recognition scheme which displays the recognition result immediately after a pattern is written. This scheme, however, imposes interruptions of user's thinking caused by confirmation and correction. To avoid these interruptions, we propose the lazy recognition scheme which delays (display of) recognition until required. Also, to fully utilize handwritten patterns, pen-trace representation and code representation of a document are maintained in mutually convertible forms.

[†] 現在 日本アイ・ビー・エム 東京基礎研究所(IBM Japan Co.Ltd. Tokyo Research Lab.)

^{**} 現在 日立製作所 システム開発研究所(Hitachi Co.Ltd. System Development Lab.)

1. はじめに

オンライン手書き入力によるユーザインタフェースをどのように構築すればよいかを追究することは、オンライン手書き文字認識システムの研究と同様に重要である。我々はオンライン手書き文字認識系をJOLISと名付け研究している[1]。しかし、いかに手書き文字認識系が認識率を向上させようとも、認識系単体ではユーザインタフェースの問題は解決されない。

過去に報告された多くのオンライン手書きによる文書入力システムでは、オンライン手書き入力のリアルタイム性に注目し、入力されたパタン（文字・校正記号）を即座に認識している（以下、リアルタイム認識方式）。我々はリアルタイム認識方式がユーザに思考の中断を強いると考え、遅延認識方式を考案した。これにより、ユーザは文書入力中は文書内容に完全に集中できる。又、我々の「手書きユーザインタフェース」では、パタン認識を行うことにより、計算機による文書処理の恩恵（清書出力、自然言語処理による誤り訂正など）を受けることを意図している。この点でストロークエディタ [2] とは一線を画する。

本稿では、手書きによる文書作成システムの基本アーキテクチャである「手書きユーザインタフェース」の設計とプロトタイプについて述べる。なお、「手書きユーザインタフェース」は我々の設計した手書きによるユーザインタフェースシステムの名称として本稿では用いる。

2. 遅延認識方式

ここではまず、創造入力に対する支援の必要性、その支援対象である文書作成作業の現状から多くの人が紙を用いる理由を考察する。そこから創造入力支援に必要な要件を導き、解決策である遅延認識方式について説明する。

2.1 創造入力支援の必要性

従来のユーザインタフェース手法では、主に入力効率や速度、すなわち打鍵回数、マウスや目線の移動量などに注意が払われてきた。これらの評価尺度は入力作業、特に入力すべき原稿がすでに出来上がった状態からの入力（以下、コピー入力）に伴う肉体的疲労を

問題としていると言える。しかし、文書作成という高度な知的作業においては肉体的疲労だけが唯一重要な評価尺度であるのだろうか。特に創造的な作業における文書作成では、発案者とタイピストといった分類は成り立たない。文書にすることにより初めて考えたことが明瞭になり文書作成中に発想を刺激されることがままあるからである。つまり、コピー入力だけを問題にするのではなく、内容を考えながらの入力（以下、創造入力）におけるユーザインタフェースの研究が必要であるといえる。そこで、我々の設計する手書きユーザインタフェースでは、高度な知的作業である創造入力の支援を中心に考えた。

2.2 文書作成作業形態の現状に関する考察

調査によれば、文書作成初期の発案から最終出力にいたるまですべてを計算機上で行う人は少数派である [3]。多くの人々は、文書内容や構造を考えたり、推敲や再構成などの段階で紙（原稿用紙や仮出力など）とペンを用いて作業をしている。紙とペンを用いる理由を次に示す。

- (1) どこでも作業できること
- (2) 紙を置く場所さえあれば広い範囲を一度に見渡せること
- (3) CRT画面や液晶表示面より紙面出力の方が見やすいこと
- (4) 書き込みが自由、手軽であること
- (5) 落ち着いて作業できること

このうち(1)～(3)は紙の物理的特性によるものである。計算機の表示デバイスがこれらの点で紙のレベルに達するのは容易ではない。しかし、大型化、高分解能化、無視差化、軽量化の方向で確実に進歩しているので、紙との格差は小さくなりつつある。

残る(4)、(5)に関しては、創造入力用手書き入力に取り入れるべきであり、しかもそれは、手書き入力ならではの特徴である。

2.3 なぜ紙を用いると落ち着くか

さて、なぜ紙を用いると落ち着いて作業ができるのかについて考えてみる。認知心理学の教えるところによれば、単に人間が外界からの刺激（音声、文字など）

を（視覚、聴覚などにかかわらず）「知覚する」だけでも「注意」が必要である[4]。「注意」していない事柄に関しては、ほとんど無視され、急速に短期記憶から消失する。創造的作業である文書作成（創造入力にこれにあたる）について、心理学的にその詳細が明らかになっているわけではない。しかし、知覚よりさらに高度な精神作業であることから、なおさら「注意」をその内容に向ける必要があると推測される。

一般に「注意」は、エネルギーのようなもので、その「量」は限られていると考えられている。したがって、2つの作業に「注意」を分散すれば、作業の質、効率は低下するだけでなく心理的負担も増大する。作業対象を我々の問題に近いキーボードによる日本語入力に限定し、同時に2つの作業を行う場合の干渉について文献[5]に示されている。そこで論じられている作業効率の低下も2次作業中の「判断と出力」が「注意」を必要とする作業であるためと考えられる。

一方、まったく「注意」を必要としない作業も存在する。高度に熟練した作業がそれであり、「自動化」された作業と呼ばれる。一部の英文タイピストにそのいい例が見られる。原稿を見て、タイプライタを操作する作業に高度に熟練し、「自動化」されている。そのためタイピング作業にはまったく「注意」を必要とせず、タイピングと同時にまったく関係のない世間話が可能である。つまり、創造入力において文書内容に「注意」を集中するためには、計算機への入力作業が「自動化」されていけばよいのである。しかし、作業が「自動化」されるためには、大量の時間をかけた過度の練習が必要である。

ここで手書き作業について考える。少なくとも日本においては、ほとんどすべての人が文字、式、図の手書きを初等教育で学び、日常生活で用いている。滅多に使用しない文字などを除き、手書きするという作業は日常生活において特に意識されることなく過度に練習され、「自動化」されている。したがって、紙とペンをを用いた作業では、常にその内容に「注意」を向けられる。つまり、思考の中断がない。これが「紙とペンをを用いると落ち着いて考えられる」理由と考えられる。

2.4 創造入力用手書き入力の要件

紙とペンをを用いる理由に関する考察から我々は手書きの計算機入力環境は次の特徴を備えるべきであると考える。

- ・思考の中断のない思考環境を構築する
- ・入力の自由さを確保する

特に重要な特徴は思考の中断がないこと、つまり「人の思考を妨げない」ことである。入力の自由さもユーザの思考内容を容易に表現させることで思考を妨げる要因を取り除く。

そこで「人の思考を妨げない」を我々の手書きユーザインタフェースの基本思想とする。

2.5 遅延認識方式

オンライン手書き文字入力は、オフライン手書き文字入力に比べて次の点が優れているといわれる。

- (1) ストロークの分離が容易であること
- (2) 時系列のデータが得られること
- (3) 誤認識をすぐその場で訂正できること

このうち、我々は(3)に疑問を感じた。認識系にもよるが、オフライン入力で大量の文書を入力した場合、その誤認識の訂正は大変な手間となることがある。

「オンライン入力では、書いた文字を即座に認識できる。認識結果を即座にユーザに示し、誤認識であれば訂正してもらえばよい。」というのが(3)の根拠である。事実、現在発表されているオンライン手書き文字認識を用いた入力系ではすべて、個々の文字ボタンを入力した直後に認識を行い、認識結果を表示する方式（リアルタイム認識方式）となっている。

確かにオフライン手書き文字入力(OCR)を比較対象とすれば、特徴と考えられなくはない。しかし、入力方式全般を比較対象とすれば、仮名漢字変換方式であってもこれと同様の特徴はある。つまり、これはオンライン手書き文字入力固有の特徴ではなく、「リアルタイムで入力の解釈を行う入力方式」の特徴なのである。

さて、リアルタイムで入力の解釈を行う入力方式、例えば仮名漢字変換による入力方式では、その入力操作に高度に熟練していないかぎり、変換結果の確認作業や誤変換の訂正作業が、ユーザの思考の中断を招く。

これは、文書内容とともに確認・訂正作業にも「注意」(2.3節参照)が必要なためである。充分な「注意」が必要な2つの作業を交互に行わなくてはならない。したがって、手書き入力であってもリアルタイム認識方式では、仮名漢字変換とまったく同様に思考の中断が起こってしまう。単に、誤変換の確認作業が誤認識の確認作業に置き代わるだけである。

そこで、上の(3)をオンライン手書き文字入力の特徴と考えることはやめる。ここから遅延認識方式の概念が生まれた。

遅延認識方式では、個々の文字が書かれた直後には認識処理を行わない。1つの文書の入力がすべて終わった時点や、ユーザが「これでひと区切り」と考えた時点で文字認識処理を行う。あるいは、バックグラウンドで認識処理を行っても、その結果表示は必要になるまで行わない。つまり、人間の思考が自発的に途切れた時、はじめて認識(あるいは表示)を行うのである。これにより、計算機システム側の都合で人間の思考をさえぎることがなくなる。手書き入力では、認識処理を施す以前のパターンでも(字の汚さは別として)容易に読める。むしろ、表示の切り替えにより思考を乱すことを考えれば、手書きパターンの方が汚くても思考の継続性はよい。入力直後に認識を行う必要はないのである。

遅延認識方式では、文書の内容を入力する時は、文書内容だけを考えていけばよい。その後、認識処理結果が示され、確認作業となる。認識結果の確認、誤認識の訂正を行っている時には、その作業だけに集中すればよい。

文字認識系の認識率があまり高くない場合、文書入力後の訂正作業が膨大な量になる恐れがある。しかし、同じ認識系を使用している場合、リアルタイム認識方式であっても、訂正作業が膨大になることに変わりはない。単にそれが細分化されているだけである。つまり、訂正の作業量は、認識系の問題である。むしろ、遅延認識方式を用いることにより、入力時に訂正作業が混入しないことで、入力時の思考環境を守ることができる。

3. 手書きユーザインタフェースの基本設計

手書きユーザインタフェースは次の二つの基本概念から構成される。

- (1) 遅延認識方式
- (2) 二つの文書表現形態

遅延認識方式についてはすでに述べた。これは、概念上、1回の入力セッション(文書入力作業の全部、または一部)が終了するまで認識処理を行わない方式である。二つの文書形態は、手書き入力パターンそのままの表現である「筆画表現」と、通常の計算機上での文書表現と同等の「符号表現」が用意される。遅延認識は筆画表現から符号表現への変換として実現される。

3.1 筆画表現と符号表現

手書き入力された文字等は、認識結果が出てしまえばもはや不必要なものと我々は考えていない。手書きの原稿には記入者の思考の過程が挿入や削除の記号の形で残っている。考えをまとめるため、余白にラフな図を書くこともよくある。これらは重要な情報であると我々は考えている。

そこで、手書きデータと符号化されたデータは分離する。分離することにより、手書きデータの自由さを保証するのが次に示す2つの表現による取扱い方法である。

(1) 筆画表現

筆画表現とは基本的にタブレットに書かれた筆点データとそのタイムスタンプで構成される。筆点データとは、タブレットからサンプルリングされた時系列の座標値の集合である。したがって書き込むデータにはなんら制限がなく、あらゆる記入が許される。

(2) 符号表現

符号表現とは文書本文の部分で考えると、文字列である。つまり、筆点入力されたデータを符号化し、これにレイアウト情報を加えれば清書出力が可能な段階に達したものである。文章なら文字列になるが、図形であれば図形記述言語、数式であればその記述言語という内部表現になる。これには符号化が不可能なストローク情報、ビットイメージ情報も含まれる。これらの集合で一つの文書の符号表現が構成される。

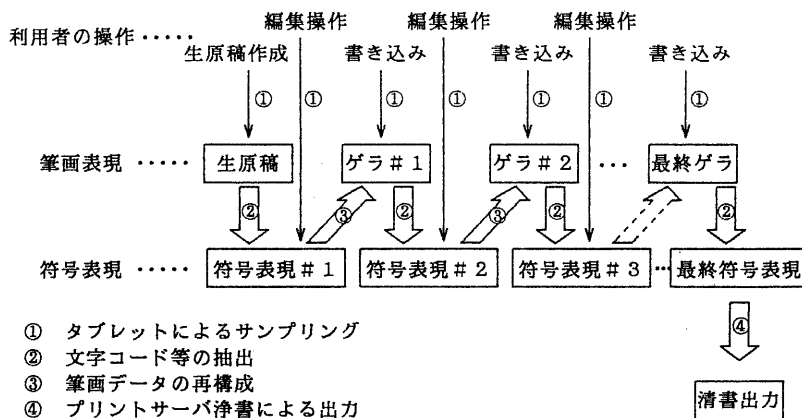


図1 文書作成作業の流れ

3.2 筆画表現と具体表現の相互変換

筆画表現と符号表現はそれぞれ独立した文書である。そこでは文面が同一であっても、手書きと計算機コードという表現形式の違いから内容には差がある。人間から見れば、筆画表現の自由さが重要であるが、計算機の便利さを享受するためには、筆画表現と符号表現の間で変換が必要である。次にその相互の変換について述べる。

(1) 筆画表現から符号表現

筆画表現から符号表現をつくる時、次に示す処理が必要になる。

- (a) データの切出し
- (b) 文字、校正記号などのシンボルの認識
- (c) レイアウト情報の抽出

まず、図形領域、生データのままで出力系に渡すデータなどの切出しを行わなくてはならない。その後、文字であれば1文字単位の切出し、図形であればプリミティブなどの処理単位による切出しが必要である。文字認識、図形認識はこの切出し結果に対して行われている。認識処理はこの時点まで遅延されることになる。

レイアウト情報の抽出は領域の切出しで得られた分布や、文章内の文字の位置関係などを解釈する処理である。

(2) 符号表現から筆画表現

符号表現から生成した筆画表現（再生筆画表現）は、その源となった筆画表現を忠実に再現する必要はない。再生筆画表現はさらに訂正等の書き込みを加えること、その書き込みを思考過程として残すことがその目的である。

3.3 手書きユーザインタフェースを用いた文書作成作業形態

通常、ワードプロセッサ(以下ワープロ)で文書を作る場合、次のステップによる場合が多い。

- (1)手書きの生原稿を考えながら作る
 - (2)生原稿をワープロに入力する
 - (3)仮出力する
 - (4)仮出力にミスの訂正、推敲の書き込みをする
 - (5)書き込みを参照しながらワープロ上の文書をエディットする
- 以降、(3)から(5)を任意回数繰り返し、最終稿に至る。

これを本方式の作業の流れと対応づけると次のようになる。

- (1')手書きの筆画表現(生原稿)を考えながら作る
- (2')筆画表現から符号表現に変換する
- (3')符号表現から筆画表現に変換する
- (4')再生された筆画表現にミスの訂正、推敲の書き込み

をする

(5') 書込み済みの筆画表現を符号表現に変換する以降、(3') から(5') を任意回数繰り返し、最終稿に至る。

このように現在、実際に行われている文書作成ステップとまったく同様に文書作成ができ、手間の軽減を図れる。

3.4 手書きユーザインタフェースの特徴

これまで述べたように手書きユーザインタフェースには次のような特徴がある。

(1) ユーザが文書入力時には文書内容に集中できること(自然な思考環境の実現)

これは、本システムの最大の特徴である。認識結果はここでは表示されないため、誤認識の訂正はもちろん、手書きパターンから認識結果へ表示が切り替わることから、ユーザは邪魔されることはない。

(2) 認識結果の確認・訂正時には、その作業にだけ集中していればよいこと

これらの作業は文書内容の作成に区切りがつき、自然に文書内容に対する集中が途切れた後で行われる。この時には、文書内容にわずらわされることなく、認識系の手助けをすることだけを考えればよい。入力文字パターンから筆画表現として残されるため、誤認識時などいつでも手書きパターンを参照できる。

(3) 教示データとなる、手書き文字認識系への操作を違和感なく組み込む

手書き文字認識では誤認識は本質的に避けられないので、それにできるだけ対処し2度3度と同じ間違いを繰り返さないようにすることが必要である。このため、文字認識系の学習機能は重要である。文字認識系に効率のよい学習をさせるには、正確な教師データ(学習動作の元となるデータ)が必要である。教師データは、誤認識の訂正操作や字体表現辞書への明示的パターン追加などから得られる。本システムでは誤認識の訂正時にユーザを訂正作業に集中させられるため、より正確な教師データが期待できる。また、字体表現辞書へのパターン追加など明示的辞書操作も訂正時に自然に組み込むことができる。

4. 手書きユーザインタフェースのプロトタイプ

手書きユーザインタフェースの基本設計は、手書きで文書を扱うための枠組みである。

そこで、現状の技術水準を確かめ、基本設計の有効性を確認する目的から筆画表現エディタのサブセットである筆画表現作成ツールと筆画表現から符号表現への変換系について実現可能な仕様を考察し、筆画表現作成ツールの実現を行った。

4.1 プロトタイプ的设计

筆画表現の作成用にエディタを用意する。しかし、筆画表現のデータ単位が筆画であり、文字の区切り等は考慮されていない。したがって現在の計算機エディタのように文字列の操作をした後の自動的な再配置などは基本的に行わない。現在の紙とペンに対応する。

筆画表現を処理し符号表現にした後再び筆画表現とすると、符号表現のデータ(文字の切りだし結果、文字認識結果)を捨ててしまうことは得策ではない。再び符号表現に変換する時に無用の手間が生ずることになる。これは、手書きパターンの処理系に学習操作を加えた場合、処理の再現性が保証されなくなることによる。そこで、符号データを(持てる場合は)持てるようグループデータを設ける。

グループは、ストロークまたはグループの集合として定義される。このグループに属性を持たせる。文字であれば、文字を構成するストロークでグループを構成し、その属性として文字コードを持てばよいことになる。

筆画表現ではなにを書くことも許されるが、あらかじめ記入用紙を設定し、それに書き込む手段は必要と思われる。これは、用紙を設定することにより、計算機による記入データの解釈を助けることになる。記入者にとってもガイドラインを示すことになる。

このように、3種類のデータ(筆画、グループ、用紙)の集合として筆画表現を表現する。このとき、これらをまとめて管理するデータが必要である。そこで管理用データを加え、四つのデータから筆画表現を構成する。その内容を表1に示す。

さて、今回のプロトタイプでは、入力対象として文字を中心に考えることとする。あらゆる文書作成にお

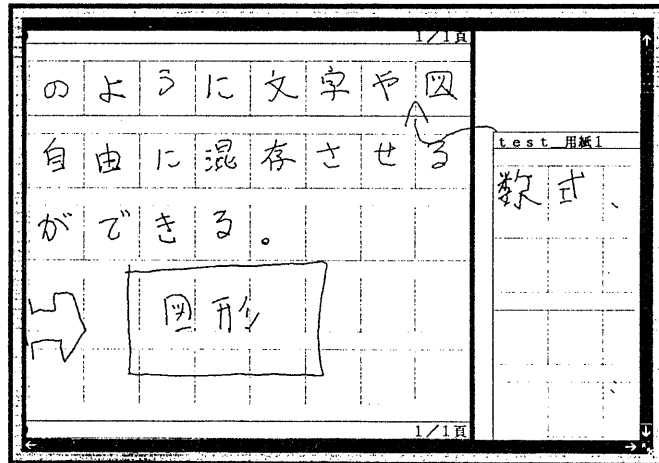


図2 筆画表現作成ツール画面例

いて文字の持つ重要性が、その理由である。通常でも紙とペンを用いて原稿を作成する場合、原稿用紙がしばしば使われる。原稿用紙の使い方は初等教育に含まれており、古くから存在するためその利用者は今でも多い。そこで原稿用紙の形式を記入用紙として採用することとする（原稿用紙メタファ）。

原稿用紙には1マスに1文字書くことが基本である。しかし、実際に原稿用紙を用いて原稿を書く場合は、必ずしもそうではない。行間や欄外の余白にも書き込みがなされることが多い。これを入力として認めると結果的にはすべて枠なしで文字切出し処理をすることと変わりが無い。そこで、余白への書き込みは今回対処しないこととする。その代用として、新たな原稿用紙（子原稿用紙）を指示された場所に関係付け、追加する機能を設ける。

4.2 筆画表現作成ツール

筆画表現作成ツールでは、前節に述べた原稿用紙メタファに基づく筆画表現を作成する。基本的にはジェスチャの筆跡もすべて筆画表現に残す。しかし、原稿用紙のページめくりなど、原稿用紙そのものに関するジェスチャは、即座に解析、実行されるので筆画表現には含まれない。

その他は実物の原稿用紙と（鉛筆ではないので筆跡の消去はできない）ペンを用いて原稿を作成する環境をそのまま再現するよう設計を行った。筆画表現作成ツールの作業時表示例を図2に示す。

筆画表現作成ツールでは、思考の中断のない入力環境の魅力をかいまみることができた。しかし、それよりも感じるのは、入力筆跡表示におけるリアルタイム性の重要性である。通常はまったく表示の遅れなどない現実の紙とペンを用いる操作を再現している。これが当然となっているので、少しの表示の遅れでも如実に感じられる。マウスとマウスカーソルの位置関係と異なり、入力ペンとエコーバックは物理的にほとんど一致しているためと考えられる。速度についてはアプリケーションプログラムで対処することには限界がある。

5. おわりに

手書きならではの特徴を生かした手書きユーザインタフェースの基本設計について述べた。その基本思想は「人の思考を妨げない」ことにある。遅延認識方式と二つの表現形式による文書管理について述べた。遅延認識方式の趣旨は入力された手書きパターンを即座には認識「しない」ことにある。また、二つの文書表現

形式を設け、手書きパタンをそのまま残す。これは手書きパタンを大切にするためである。

さらに、筆画表現から符号表現への変換について我々の現在の技術水準にそった考察を加え、より現実的な設計を行った。また、原稿用紙メタファによる筆画表現作成ツールの実現を行った。

参考文献

- [1] 本間, 他: オンライン手書き文字認識システムJOLIS-2Eの開発, 情報処理学会第40回全国大会講演論文集, 6E-3, pp.358~359(1990)
- [2] 守屋, 森田, 稲井, 清水: ストロークエディタと直接指示・操作方式, 情報処理学会論文誌, Vol.32, No.8, pp.1022-1029(1991)
- [3] 下村, 他: 形態素解析を利用した日本語スペルチェッカ, 情報処理学会第31回プログラミング・シンポジウム報告集, pp.117~125(1990)
- [4] J.R.アンダーソン(富田他訳): 認知心理学概論, p.552, 誠信出版, 東京(1982)
- [5] 岡留, 小野, 山田: タイプ入力作業の構成要素間に起こる干渉, 情報処理学会論文誌, Vol.27, No.3, pp.304-311(1986)
- [6] 早川, 他: 表示一体型タブレットを用いた計算機システムの構想, 情報処理学会第33回プログラミングシンポジウム報告集, pp.125~134(1992)
- [7] 曾谷, 他: 手書きユーザインタフェース, 情報処理学会第31回プログラミング・シンポジウム報告集, pp.1~10(1990)
- [8] 曾谷, 他: オンライン手書き文字認識系の使い勝手に関する考察, 第32回情報処理学会プログラミング・シンポジウム報告集, pp.69~76(1991)
- [9] Souya, Fukushima, Takahashi, Nakagawa: handwriting Interface for a Large Character Set, Proc. of 5th handwriting Conf. of The Int'l Graphonomics Society, pp.166~168(1991)
- [10] 花木, 天満, 吉田, 鈴木, 石, 菊池: オンライン実時間文字認識システムの開発, 電子通信学会論文誌, Vol.J64-D, No.5, pp.411-418(1981)
- [11] 小高, 若原, 橋本: オンライン手書き文字認識装置, 電子通信学会論文誌, J65-D, No.8, pp.951-958

(1982)

[12] 福永, 葛貫, 三瓶, 横井: 高度マンマシンインタフェース付きワークステーションの開発, 日立評論, Vol.28, No.2(1986)

[13] 正嶋, 福永, 平沢, 篠崎, 松田: 手書きマンマシンインタフェースの開発, 電子通信学会技術研究報告, OS88-48(1988)

[14] 寺井, 中田: オンライン手書き文字認識を用いた原稿校正システム, 情報処理, Vol.15, No.6, pp.16~19(1974)

[15] 葛貫, 横山, 正嶋, 福永: JIS校正記号準拠のオンライン手書き編集方式, 情報処理学会論文誌, Vol.27, No.10, pp.1027-1035(1986)

[16] 児島, 戸井田: 手書き編集記号を用いたオンライン文字図形編集法, 情報処理学会論文誌, Vol.29, No.3, pp.242-248(1988)