

5C-1

手書きユーザインタフェースにおける  
遅延認識方式の提案曾谷俊男, 福島英洋, 中川正樹, 高橋延匡  
東京農工大学

## 1. はじめに

本稿では、文書の手書き入力を実現する上で重要な手書きユーザインタフェースの構築手法について述べる。手書きユーザインタフェースの実現のためには、オンライン手書き文字認識系が高い認識率を達成していることが重要ではある。しかし、これと同等に、ユーザに無理のない入力環境を実現するためには、手書きを用いたユーザインタフェースをいかに構築するかが重要な課題である。

日本語文は英文などに比べ、非常に多くの文字を用いて表現される。このため、その入力がかさばり問題となる。現在、仮名漢字変換による方法が日本語文入力手法の主流を占め、ほとんどのワードプロセッサなどの文書入力システムとして用いられている。しかし、より高い変換率を得るためには、文節の切れ目などで文を区切って入力する必要がある。それでも同音異義語によって正しい変換結果を得られるとは限らない。また、この入力手法がユーザに与えるストレスも社会的に認知され、研究されつつある。

このような背景から、筆者らは手書きユーザインタフェースの研究を行っている。手書きユーザインタフェース実現のためには、次に示す2つの主要な問題を解決する必要がある。

- (1) いかにして手書きパターン(文字、図形など)を認識(または処理)するか
- (2) いかにしてマンマシンインタフェースの全体構造を構成するか

本稿の主眼は、後者に置かれている。しかし、筆者らは前者についても研究を行っている。これらは互いに影響するものであるからである。そこで、筆者らが研究・開発を行ったJOLIS-2Eの特長について述べる。

## 2. JOLIS-2Eの特長

筆者らの属する研究室では、JOLIS(Japanese On-Line Input System)計画と名付けられたオンライン手書き文字認識に関する研究を行っている。その認識手法は、構造解析的手法と統計的手法という代表的な2つの認識手法を併用したものである。これによりJOLISの認識手法は、それぞれの手法の特長を合わせ持つものとなっている。

文字認識系を実際に用いることを考えれば、すべてのユーザのあらゆる変形を含む手書き文字を正認識することは不可能と言える。現実的に考えれば、特定ユーザに認識系が適応する能力は重要である。さらに、認識処理、学習による作用はユーザに理解しやすいものでなくてはならない(認識系の透明性)。透明性が高ければ、たとえ使い始めて性能(認識率)が思わしくなくても認識系の学習能力を信頼して使い続け、最終的に優れた性能を得ることができる。これが筆者らの基本思想のひとつである「パーソナル指向」の根拠である。

多くの漢字パターンでは共通の部分パターン(サブパターン)をもつ。したがって、これを用いて漢字パターンを表現できる。JOLIS-2Eの字体表現辞書では、標準パターンをサブパターンから

構成される2進木として表現される。サブパターンは他のサブパターンを用いて表現することもできる。

手書きユーザインタフェースから見たJOLIS-2Eの認識手法の特長を次に示す。

- (1) 構造解析的手法はユーザに理解されやすい認識手法であること。ユーザから信頼感を得るために誤認識やリジェクトの原因はわかりやすくなければならない
- (2) 確率的相違度を用いているため、JOLIS-2Eはパラメトリックな学習により、自動的にユーザに適応が可能なこと
- (3) 相違度を確率という無次元量で表現しているため、その他の情報(例えば、文脈的情報)を認識モデルに組み込むことができること
- (4) 部分パターンを複数の文字カテゴリで共用するため、パターン追加による学習の効果が複数カテゴリに波及すること

## 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を実現するために手書き文字認識技術は必須である。しかし、それだけで実際の文書入力について役立つわけではない。手書きに適したユーザインタフェースの設計が必須なのである。

まず、実際の文書作成方法について調査を行った。その結果、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ことが判明した。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。紙を用いる理由のひとつとして、仮名漢字変換などによる思考の中断がないことが考えられる。

手書き入力は、次のような目的に優れた特長を持つ。

- (1) 初期の思考段階から文書作成を支援すること
- (2) 最終出力にいたる前の中間原稿を操作すること
- (3) 数式や線図形を入力すること

一方で、手書き入力は、原稿が用意された状態で入力(いわゆるコピータイプ)には向かないと考えられる。しかし、最終目標は文書作成時の様々な段階、様々な入力対象に適した複数の入力手段を与えることにある。それぞれの目的、入力するデータに合わせて最も使いよい手段を用いればよいのである。

文書作成システムの研究においては単に効率を重視するのではなく、文書の質を求めることも意義深い。そこで、筆者らの手書きユーザインタフェースの基本思想を「思考の中断のない自然な思考環境」をユーザに提供することとした。これは思考の中断を避けられない仮名漢字変換では望めないことである。文書作成に今でも紙とペンを用いる理由はここにある。しかし、単に実物の紙とペンでできることだけを計算機上で実現しても、あまり意味はない。当然そこには $\alpha$ が求められる。そこで設計目標は、文書作成時の「手書きの自由さと計算機の便利さ」を統合することとした。手書きの自由さ、表現力を生かすため、基本方針は「手書きのパターンを大事にする」とした。

The Lazy Recognition Scheme for Handwriting Interfaces.

Toshio SOUYA, Hidehiro FUKUSHIMA, Masaki NAKAGAWA, Nobumasa TAKAHASHI

Tokyo University of Agriculture And Technology.

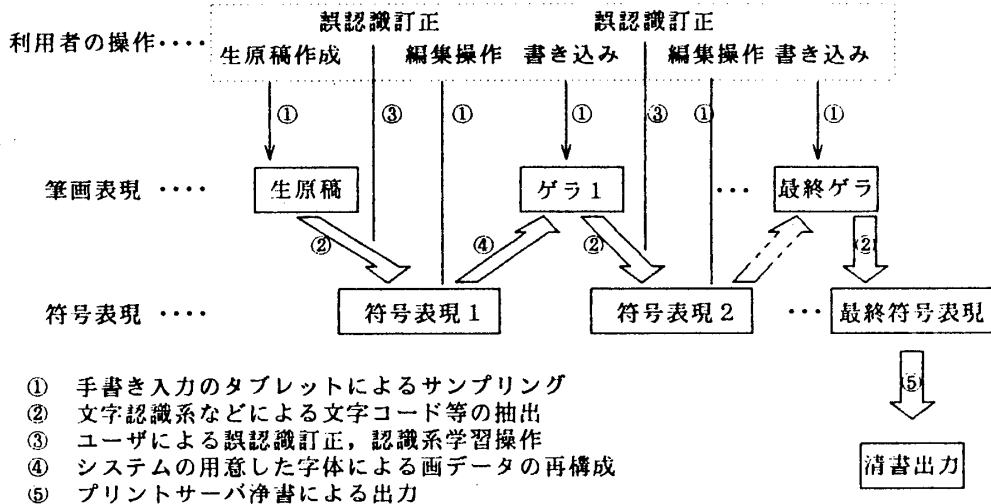


図2 文書作成作業のながれ

その基本思想から、文字認識処理は遅延認識方式(Lazy Recognition)で行われる。現在まで報告されてきた、オンライン手書き文字認識系を用いたすべての手書き入力システムでは、文字が入力されると即座に認識され、結果が示された(以下、リアルタイム認識方式と呼ぶ)。ここでは、ユーザは認識結果の表示やその確認、誤認識の訂正などで思考の中断を強いられる。思考環境として考えれば、仮名漢字変換方式と変わりはない。遅延認識方式では、ユーザが認識結果を求めるまで認識結果を示されない。ユーザの思考が自発的に途切れるのを待つわけである。認識処理はリアルタイムでもよいが、結果は表示しない。手書きパターンから認識結果へ表示が切り替わることも、思考を乱すからである。手書きパターンはそのまま人間には容易に理解できるため、表示を変える必要はない。

文書は2つ文書表現：筆画表現と符号表現が相互に変換可能な状態で管理される。遅延認識は筆画表現から符号表現への変換として行われる。

基本的に筆画表現はタブレットからの生筆点データである。符号表現は現在の計算機上の文書表現と同様の表現で、パターン識別結果などのコード列によって表現される。ここで重要なのは、符号表現が次に示す理由により、認識処理終了後も保管される点である。

- (1) その配置、大きさ、筆速速度など、文字コードより多くの情報を手書きパターンは持っていること
- (2) 誤認識時に、ユーザが元の手書きパターンを参照できるようになること
- (3) 符号表現を残すことにより、入力に制限を課す必要がなくなること

ユーザはまず、筆画表現を用いた文書作成を行う。このため、正認識、誤認識はユーザの思考を乱すことはない。符号表現を操作することも可能である。

事実上、完璧な文字認識は不可能なため、筆画表現から符号表現への変換はユーザからの支援を必要とする。いつ、どのようにユーザからの支援を得るかは、ユーザインタフェースの設計上、重要である。

ユーザによる原稿入力(筆画表現による文書作成)が終了すると、認識結果が示される。この時、ユーザは認識結果の訂正に集中することができる。ユーザによる訂正は認識系に学習用データを提供する。

筆者らの手書きユーザインタフェースを用いた文書作成作

業は、現在、実際に行われている文書作成作業、つまり、初期原稿を書き、計算機に入力し、仮出力を用いて推敲するなどに対応づけることができる(図1)。ここで、遅延認識による筆画表現から符号表現への変換は、計算機への入力(タイピング)に対応する。タイピングの代替と考えれば、誤認識の訂正作業は(認識系の性能にもよるが)不当な手間を要求しているとは言えない。

一方、リアルタイム認識方式では、後から訂正する手間は無い。しかし、手書き入力時に誤認識やリジェクトにより、ユーザの思考は中断させられる。後から訂正する手間は無いとはいえ、訂正の手間自体は減っているわけではない。遅延認識方式では、完成した文章のパターンが認識系に渡されるため、前後関係を参照する文脈処理による大分類、候補選択、結果確認などの処理を取り入れる自由度が高い。

入力と訂正の分離はユーザに創造的環境を提供するだけでなく、システム作成者にも利益をもたらす。訂正が分離されているので、文字認識系本体とは分離して学習機構を構築できる。それらの精度にも寄与する。また、筆画表現作成系、筆画表現から符号表現への変換系など、それぞれの作業段階用ソフトウェアを別々に作成できる。したがって、不必要に単一プログラムが大きくなることもない。ある部分(例えば文字認識)で技術が進歩した場合も容易に組み込みが可能である。

参考文献

- [1] 曾谷, 他, “手書きユーザインタフェース”, 第30回プログラミングシンポジウム報告集, pp.1-15(1990)
- [2] 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回プログラミングシンポジウム報告集, pp.69-76(1991)
- [3] Nakagawa, M., et al., “Non-keyboard Input of Japanese Text On-line Recognition of Handwritten Characters as the Most Hopeful Approach”, Journal of Information Processing, Vol.13, No.1, 15-34, IPS Japan(1990)
- [4] Souya, T., et al., “Handwriting Interface for a Large Character Set”, Proc. of 5th Handwriting Conf. of IGS, pp.166-168(1991)