

視覚障害者用パーソナル OCR システムの設計

小場 久雄, 下平 博, 木村 正行
北陸先端科学技術大学院大学 情報科学研究科

我々が過去に設計した知的自動点訳システムの使い勝手に関する成果と課題を分析し、その結果を踏まえた新しい個人用小規模 OCR システムの設計方針について検討する。設計当時の技術の制約からくる問題点を踏まえつつ、視覚障害者に家庭で気軽に使用してもらえるようなシステムの条件について検討する。また既に市販 PC の環境に慣れ親しんだユーザが導入し易いように、独自インタフェースだけではなく、我々の OCR エンジンに HI として既存の HTML ブラウザと音声合成ソフトを組み合わせる利用方法についても検討する。

Design of Personal OCR System for the Visually Disabled

Hisao Koba, Hiroshi Shimodaira, Masayuki Kimura
School of Information Science
Japan Advanced Institute of Science and Technology
E-mail: koba@jaist.ac.jp

The purpose of this study is to examine the design of a small OCR system for visual disability, which is intended for using without any help of sighted people. The priorities of the design factors of the system, which should be easy to use, are determined by analyzing the problems of some technical restrictions of our former OCR system. For the users who are familiar with today's PC system, we also examine the possibility of the use of the existing html browser and the voice synthesys software as an alternative interface to the original one.

1 はじめに

本研究では、視覚障害者が晴眼者の助けを借りずに自分自身で気軽に利用できるようなパーソナルOCR(Optical Character Recognition)システムの設計について検討する。

従来我々が開発してきた知的自動点訳システム[1]は視覚障害者の利用を考慮したインタフェースが実装されていた。しかしながら設計時のコンピュータ能力の問題から、

- システムが巨大である(図1)
文字認識専用装置とその他の処理用 workstation、システム操作用 PC(Personal Computer) など複数装置を組み合わせた大規模システムなので、価格と維持管理の両面でパーソナルな使用は事実上不可能だった。
- 処理速度が遅い
設計当時の最速コンピュータと専用に設計された装置で分散処理しているにもかかわらず、1ページ当たり1分程度の処理時間がかかり、複数文書の interactive 処理の使い勝手に問題があった。

等の課題の解消が困難であった。

一方、今日のPCの急激な進歩により、従来専用装置で行っていた処理を全てソフトウェアで代替する事が可能になった。また大量生産の恩恵を受けたPCは価格面でもパーソナルなシステムにふさわしいものになっている。その結果として、視

覚障害者が家庭や出先で手軽に墨字情報を処理できるようなOCRシステムの実現も技術的に十分可能になっている。

その上、今日普及しているPCの大多数はOS(Operating System)としてMicrosoftのWindows95があらかじめ導入されている事が期待できる。よって開発面でも普及の面でもこのOSをターゲットにすることによって、特にコンピュータシステムに精通している訳では無い一般ユーザのPCへのシステム導入が抵抗感無く行われる事も期待できる状況になっている。

実際にPC上で動くOCRシステムを利用した視覚障害者向け文書処理システムは既に幾つか商品として売り出されている。例えばリコーの「点図君」と「読み取り物語EX」とのコンビネーションシステムや、アメディアの「ヨメール」などがある。前者は点図・点訳システムに特に視覚障害者を対象にしている訳ではないOCRソフトを組み合わせたものでしかなく、後者は点字出力が無い。

そこで、以下では我々の知的自動点訳システムの経験に基づく、視覚障害者用パーソナルOCRシステムに必要な条件について検討する。本システムは将来的には実際の視覚障害者に広く利用してもらう事も視野に入れている。また、視覚障害者の利用を設計時から考慮した独自システムだけではなく、HTMLブラウザと音声合成ソフトを組み合わせたシステムについても検討を行なう。これは既存PCのユーザが操作に慣れ親しんでいると思われる高普及度のソフトウェアと我々のOCRエンジンなどの機能を利用したものである。

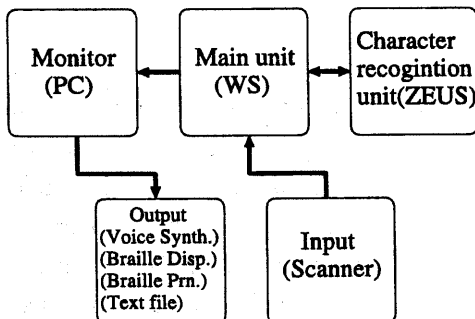


図 1: 知的自動点訳システム

2 OCR システムに必要な機能

本節では知的自動点訳システムにおける問題点の分析に基づいて、パーソナルなOCRシステムに求められる機能を検討する。

パーソナルなソフトウェアの普及を考える上で考慮しなければならない点として、ソフトウェアそのものの機能と共に、そのソフトウェアのPCへのインストールがある。しかしこの点に関してはあらかじめPCにインストールしておくなどの方法も考えられるので、以下ではシステムが使える状態になった後に必要な機能についてのみ検討する。

OCR システムは文字通り印刷された活字を文字認識して plain text にするものなので、第一に認識の速度と精度が問題となる。

まず認識速度については、当然ながら速ければ速い程良いと言える。認識速度の速い遅いについて特に明確な基準は存在しないが、比較的普及しているソフトウェア（いわゆるワープロや表計算）の操作を考えると、やはり数秒が待つ限度ではないかと思われる。視覚障害者の利用も考慮した場合、認識速度が遅い場合の問題点としては、

1. 待ち時間が長いと操作に対するユーザの不安感が増長される
2. interactive 処理に支障が生じる

の 2 点が考えられる。

1 は、ユーザの操作に対する適切なフィードバックが返って来ないために生じる問題である。このことは視覚障害者にとっては特に問題となる。処理に時間がかかる場合はその間何らかのフィードバックをシステムから返し続ける事で、ユーザにシステムが正しく動作している事を知らせる事になる。一般的には画面に処理状況を表示する方法が主流であるが、視覚障害者にとってはこの手法は無意味なので、音声あるいは点字で出力する事になる。しかし、

- 常に音声を出力し続けるとうるさい
システムの動作状況を常に出力し続けると、場合によっては騒音となり兼ねない。また、システムからのフィードバック情報に対するユーザの注意を妨げるので、重要な情報まで無視され兼ねない事になる。
- 邪魔にならない音声情報は情報量が少ない
「ピープ音」のような単純音をシステム動作中に出し続ける場合、それほど邪魔にならないと思われる。しかしそれだけでは情報量が少ないので、システムが動作している以上の事はユーザには分からない。

等の問題がある。知的自動点訳システムにおいては、ユーザからの要求に応じて随時システム情報を音声で提供する手法を採用した。必要な情報量

はユーザによって異なる事も考えられるので、カスタマイズ機能をつけるのが有効だと考えられる。

2 については、複数枚数の文書を続けて処理する場合、処理の待ち時間が長いと次の認識結果が出て来るまでユーザは待ち続けるしか無い。この結果、システムの使用感が大きく損なわれる事となる。知的自動点訳システムでは、処理を担当する workstation と操作を担当する PC を並行動作させ、次の文書を workstation で処理している間 PC で認識済の文書をユーザが処理できるようにしてこの問題の解消を計った。しかしこの場合ユーザが認識済の文書を操作している間に次の文書の処理が終了する事になる。そのことをどのようにしてユーザに通知するかが難しく、結局、

- 次の文書の処理状況をユーザが自発的に確認する
- 特定の操作時に PC が次の文書の処理状況を自動的に確認し、処理済なら自動的に PC 側に転送する

方法を採用したが、操作体系の一貫性が失われ、ユーザの操作に混乱も見られた。この問題は、OCR エンジンが十分な認識速度を持っていれば解決する事である。

次に問題となるのが文字認識の精度である。晴眼者の場合は認識誤りがあるかどうか原稿と照合して調べる事ができ、訂正も容易に行う事ができる。しかし視覚障害者の場合、原稿との照合は晴眼者の助け無しに行う事は不可能であり、訂正も容易ではない。よって、認識精度の向上は晴眼者用以上に必要である。

認識した文字情報の主な出力としては、音声、点字、ファイルの 3 種類がある。

- 音声は聴覚が正常であれば特別な訓練無しで利用できる。その上、後述するように触覚より情報量が豊富なので、視覚障害者用システムでは最も有用だと考えられる。しかし、現在の PC 用日本語音声合成システムの音声は肉声以上に癖があり、話す速さや音色をユーザ毎に調整しないと聞き取りにくい場合がある。したがって、リアルタイムに調整できる機能が必要である。

- 点字を利用できるようになるためにはある程度の期間の訓練が必要である。このために、点字利用が可能な視覚障害者は少数派にとどまっている。その上、現状では点字出力装置（点字ディスプレイやプリンタ）は大きく高価なものが大半であり、残念ながらパーソナルなシステムには不向きである。しかし、視覚だけでなく聴覚にも障害がある場合は音声は無意味であるし、認識内容の詳細を検討する場合などには有効である。したがって、視覚障害者用システムには点字出力を備えておくべきであると考ええる。

- 文書が必要な度に一々認識しなおすのはあまりに非効率なのは明らかである。よって、認識結果をファイルに保存したり既存文書ファイルを読み込んで音声や点字で出力する機能は必須である。また、システムに既存の文書ファイルを読み込むことで、文書リーダーとしても利用出来る。

知的自動点訳システムのように複数のハードウェアや OS を必要とする複雑なシステムを一般ユーザの家庭に導入することは、物理的にも技術的にも心理的にも事実上不可能である。しかし今日の PC の発達は、これまで複数装置で行っていた処理を 1 台の PC 上で全てソフトウェアで処理することが既に可能になっている。したがって、知的自動点訳システムと同等のシステムを一般家庭の PC 上に実現することに技術的障害は無いだろう。

3 システム概要

本研究で設計する OCR システムは図 2 のような構成を取る。まずハードウェアとしては、

- PC
- フラットベッドスキャナ
- 点字ディスプレイ
- 点字プリンタ

がある。本システムでは処理の大部分をソフトウェアで行うので、PC は十分な処理能力を持つものが

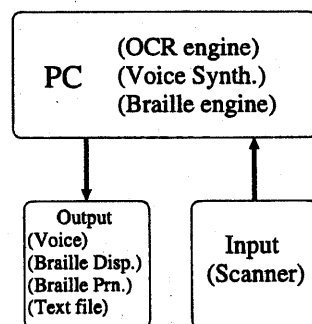


図 2: パーソナル OCR システム

必要である。ソフトウェアの構成は次のようになっている。

- OCR エンジン
- 点訳エンジン
- 音声合成ソフト
- インタフェース部

OCR エンジンとは現時点で毎分数十文字の認識速度があるので、複数原稿の interactive な読み取りにも対応できる性能を持つ。

4 HI 設計

視覚障害者を対象としたシステムの HI (Human Interface) を設計する場合に考慮しなければならない項目としては、

- 操作の手間の削減
- 音声または触覚によるシステムからの適切なフィードバックの提供

の 2 点がある [2]。前者が重要な理由は、システムの操作において、視覚障害者は晴眼者に比べて外部記憶の利用が困難なため、記憶負荷がより大きいと考えられる点にある。また操作を誤った場合の訂正がしにくい（誤ったかどうか判断しにくい）事もあるので、操作体系はできるだけ誤りを生じにくいシンプルなものにする必要がある。

後者はユーザの心理的不安の解消と外部記憶としてのシステムの活用のために必要である。

実際のシステムの具体的なインタフェースとしては以下のようなものを採用する。

- 1 キー操作

原稿を読み取ってから出力するまでや、読み取り結果の保存、終了など基本的な操作はメニューから選ばなくても1つのキーを押すだけでできるようにする。

- 10 キー操作

操作は基本的に10キーで行えるようにする。10キーだけで操作できるようにした場合の主な利点は、

- － キー配列の認識が容易
- － 扱うキーの数が少ない

であり、操作に必要なキーを探す負担を大きく削減する効果が期待できる。

- On demand help

ユーザの外部記憶であるシステム情報の活用を補強するために、システム操作中の任意の時点でユーザがヘルプキーを押すことで、その操作時点で有効なキー操作についての情報を出力する。

またもう一つの考え方として、独自インタフェースだけでなく、現在かなり広く普及していると思われるHTMLブラウザと市販の音声読み上げソフトを利用したHIを使用できるようにする。この方法の利点としては、

- 昨今のインターネットブームにより、多くのPCユーザが既に持っている
- ユーザが好きなブラウザや読み上げソフトウェアを選択できる
- ユーザが使い慣れたインタフェースをそのまま活用できる
- Web browsing と OCR をシームレスに活用できる可能性がある

- 文書の構造(章・節など)やその他付加情報(強調や注など)をOCRエンジンが認識し、それをHTML化することで原稿の情報の多くをそのまま認識文書に残すことが可能になった場合、ブラウザを使うことによってplain textより豊富な情報を提供出来る可能性がある

などがあり、システムを開発する側にとっても新たなHIを開発しなくても良いという利点がある。しかしながら、

- ブラウザはOCRシステムのために開発されたものではないので、専用HIより使い勝手が劣る可能性がある
- 視覚障害者にとっての操作のし易さ(カスタマイズなど)は使用するソフトに依存する

などの問題点も考えられる。

逆にOCR部分とHI部分を切り離すことで、本システムのHI部分をPCの汎用文書音声リーダーとして使うことも可能となる。

5 まとめ

視覚障害というある種特殊な条件を持ったユーザを対象に以前設計した大規模システムの知見を元に、個人使用のための小規模システムの設計を行った。

現時点では視覚障害者の実際の使用における本設計の有効性がまだ検証されていないので、実際に試作したシステムの実験によって今後その有効性を確認する予定である。

6 謝辞

本研究を進めるにあたり、プログラムの提供や数々の有益な助言を頂いた静岡県立大学の石川准先生と(株)PFUの角谷浩様に感謝いたします。なお本研究の一部は、平成9年度科学技術振興事業団「地域研究開発促進拠点支援事業」に関する再委託研究事業の補助を受けている。

参考文献

- [1] 小場 久雄, 下平 博, 木村 正行 (1996), 視覚障害者向け OCR システムの HI 設計, 第 12 回 ヒューマン・インタフェース・シンポジウム
- [2] 小場 久雄, 視覚障害者向けシステム設計方法の検討, 情報処理学会研究報告, HI66-3(1996)
- [3] 浅川 智恵子, 菅原 一秀, 加藤 和彦, 小出 昭夫, 視覚障害者用文書処理システムにおける情報伝達法, 情報処理学会研究報告, HI62-5(1995)
- [4] 斎藤 正夫, 視覚障害者支援ソフトウェアの製作, 情報処理, Vol.36, No.12, pp1116-1121(1995)
- [5] 石川 准, GUI 用スクリーンリーダの現状と課題-北米と欧州の取り組みを中心に-, 情報処理, Vol.36, No.12, pp1133-1139(1995)