

視覚障害者用文書処理システムにおける情報伝達法

浅川 智恵子、菅原 一秀、加藤 和彦、小出 昭夫

日本アイ・ビー・エム株式会社 東京基礎研究所

視覚障害者の情報源を拡大するため、印刷文書へのアクセスを可能にすることは重要な課題である。しかし、最近の印刷文書には飾り文字・図表・絵や写真・カラーページなどの視覚情報が多く含まれている。このようなタイプの文書に非視覚的にアクセスするためには、従来のOCRとは異なる視覚障害者にとってきめの細かいユーザーインターフェースが必要となる。本稿では視覚障害者用文書処理システムにおけるユーザーインターフェースについて、雑誌を読むという点を中心に考察を行ったのでその結果を報告する。

Document Processing System for the Visually Disabled

Chieko ASAKAWA, Kazuhide SUGAWARA, Kazuhiko KATOU, Akio KOIDE

IBM Research, Tokyo Research Laboratory
1623-14, Shimotsuruma, Yamato-shi,
Kanagawa-ken, 242, Japan
chie@trl.ibm.co.jp

It is very important for the visually disabled to be able to directly access printed documents in order to increase their information resources. But nowadays these documents contain a lot of visual information such as decorative lettering, diagrams, illustrations, photographs, colored pages, and so on. Since it is impossible to access all the information contained in such documents through the current user interfaces of OCRs, a more friendly user interface should be specially designed for the visually disabled. In this paper, we will report the results of our investigation into the ideal user interface for a visually disabled person reading a magazine.

1 はじめに

視覚障害者の情報源は、従来、点字本やカセットテープに録音された録音図書のみであった。これらの図書は、点字図書館等から貸し出されるがその数は決して充分とはいえず、すべての視覚障害者のニーズに応えることはできないのが現状である。視覚障害者が社会参加していく上で、このような情報源の不足は大きな障害の1つである。

近年のパーソナルコンピュータの普及と視覚障害者用周辺機器の開発により、視覚障害者もコンピュータを利用すれば情報源が拡大されるのではないかと期待されるようになった。実際、ある程度の制限はあるものの、画面情報を点字や音声で確認できるソフトウェアを使えば、独力で電子ブックや電子辞書の検索が可能である。また、ネットワークを介して必要な情報を入手することも可能になってきた。

しかし、我々の身辺には印刷文書の形で提供される情報が多く存在する。広告/雑誌/機器の取扱い説明書/パンフレット/カタログ/各種請求書など、今後電子ファイル化される見込みの少ない印刷文書へのアクセスを非視覚的に可能にすることは、大変重要な課題である。

スキャナーやビデオカメラ等の入力装置が急速に普及している現在、OCR技術も向上してきた。これらの機器を使えば、視覚障害者も印刷文書を読めるのではないかと期待される。しかし、これらのほとんどの機器はGUI環境化で稼動するため視覚障害者には利用できない。また、上記で述べた印刷文書には図や絵などOCRが不得意とする画像が多く含まれているため、現状の仕様ではその期待に応えることはできない。

筆者らは、現在このような問題を解決するため、光学的メディアリーダ(画像情報抽出伝達システム)の研究を進めている。これは、文字情報だけでなく膨大な画像情報の中から有効と思われる情報を抽出し、非視覚的メディアに変換し、その情報を効率良く視覚障害者に伝達するためのシステムである。

本稿では、視覚障害者が雑誌を読むという点に注目し、読みたい記事に素早く到達し、内容を理解するためにはどのような機能やユーザーインターフェースが提供されるべきか考察したのでその結

果を報告する。

2 雑誌を読む

文字情報中心の単行本や論文誌等と異なり、OCRで絵や写真、カラーページ等が多く含まれる雑誌を読み取ることは大変難しい。文字認識に失敗した場合、現状のOCRシステムでは領域分割のやり直し等の機能があるが、原本照合のできない視覚障害者には使用困難である。そこで、より正確なレイアウト解析、画像情報を聴覚情報に変換、文字情報を構造化ファイルに変換するなどの機能が付加できれば、以下のようなユーザーインターフェースが提供でき、視覚障害者も独力で印刷文書へアクセスできると考える。

ここで雑誌とは、マンスリーやウィークリーに発行される一般的な情報誌、あるいは日常生活の身近な話題を取り上げた生活情報誌などを指す。これらの大部分は、晴眼者が一目で見てその雑誌の特集や記事の注目点などが理解できるよう視覚的にさまざまな趣向が凝らされている。このような視覚的な情報を非視覚的に伝達することは大変難しい問題である。

筆者は、自らが視覚障害者であることから、まず何冊かの雑誌を実際に読み、読みたい記事を探し出しそのページを素早く開くためにはどうすればよいか、また、その記事の内容を素早く理解するためにはどのような情報が伝達されるべきかについて考察した。

以下では、この考察結果に基づき、雑誌名の特定から始まり、読みたい記事の内容を効率良く理解するまでの過程に起きる問題点、及びそれらを解決するために必要な機能やユーザーインターフェースの提案を行う。

尚、以下で「メッセージを出す、読み上げる、伝達する」といった場合は、点字/音声両方の出力を意味する。

2.1 表紙

表紙から得られる重要な情報は、表題と発行日である。雑誌名が表紙から確認できなければ内容の確認が必要になるので、その時間を短縮するためにも、雑誌名の確認は表紙から行うことが望ま

しい。この他、表紙にはその号の特集記事の見出し等が書かれていることも多く、表紙の文字情報が確認できればスムーズに雑誌の選択ができる。

2.1.1 表題

雑誌の表題は、大量にある他の雑誌と容易に区別できるよう、視覚的には文字と判断できるが、コンピュータで機械的に文字認識できない飾り文字で書かれていることが多い。表紙の絵や写真は毎回変わるが表題の字体はほとんど変わることがない。そこで、文字認識の結果に頼らず、表題を1つのパターン化されたイメージとして処理すればこの問題は解決できるのではないかと考えた。まず、雑誌の表題をその背景と分離し1つのテンプレートにする。このテンプレートに対応する表紙が入力された時、これとのパターンマッチングを行うことで雑誌名の特定が可能になる。

実際のユーザーインターフェースは以下の通りである。

1. 自動探索モード

システム側で作成したテンプレートに対応する雑誌名を表示する。

例

1. クロワッサン
2. ミセス
3. Hanako
4. 暮らしの手帖
5. DOS/Vマガジン

手元にある雑誌の中から「Hanako」を探した場合、表紙をスキャナーに乗せ、「3」の位置にカーソルを置き、表題探索ボタンを押す。ここでスキャンが開始される。スキャンが終了し、パターンマッチングの結果、以下のようなメッセージを出す。

Hanakoである場合 「これははなこです」
Hanakoでない場合 「次のいずれか」

- テンプレートに登録済みの雑誌の場合
— 「これは…です」
- テンプレートに登録されていない場合
— 「未登録です」

2. テンプレート作成/登録機能

表題がテンプレートに登録されていない場合や、すでに登録済みの表題に字体の変更があった時など、柔軟に対応するための機能である。ここでの操作は晴眼者のボランティアの助けが必要である。

1. スキャナーに表紙を乗せ、テンプレート作成ボタンを押す。
2. 入力された表紙の画像が画面に表示される。
3. マウスを使って晴眼者のボランティアが表題の書かれている領域を指定し、表題抽出ボタンを押す。
4. 抽出された表題が画面に表示される。
5. 背景が表題の中に一部残されている場合には、晴眼者のボランティアに背景の余分な情報を消してもらおう。この時はグラフィックスエディターを使用することになる。
6. 背景と表題がきれいに分離できれば、次にテンプレート登録ボタンを押す。その表題に対応する雑誌名も登録する。

上記の処理を一度行うと、以後は自動探索モードで処理できる。

このテンプレート作成/登録のインターフェースは、コンピュータの知識がない晴眼者のボランティアが容易に操作できるものになることが望ましい。これにより、多くのボランティアのサポートが期待できるので、表題だけではなく特定の雑誌で使われている絵文字などの処理にも利用でき有効性が増す。

2.1.2 発行日

雑誌の発行日は、表紙のある一定の位置に毎回決まって印刷されることが多い。しかし雑誌ごとに書かれている位置は異なる。そこで、個々の雑誌別にそのレイアウト情報を記述しシステムに知らせることができれば発行日の伝達が容易になると考えた。ここではこの機能をレイアウトモデルと呼ぶ。

実際のユーザーインターフェースは以下の通りである。

1. 自動探索モード

表題確認後、発行日探索ボタンを押す。システムはすでにこの雑誌名を認識しているので、これ

に対応するレイアウトモデルを使って発行日を探
索する。以下のいずれかのメッセージで結果を伝
達する。

- 発行日は nn 月 nn 日 nn 号です
- 発行日は見つかりませんでした

2. 検索モード

読みたい雑誌の発行日がわかっている場合、発
行日検索ボタンを押す。ここで調べたい発行日
を入力し、再度このボタンを押す。指定した発行
日が検索された場合、「成功 nn 月 nn 日 nn 号」とメ
ッセージを出す。そうでない場合は「失敗 nn 月
nn 日 nn 号」とメッセージを出し、別の雑誌の表紙を
再度入力するよう要求する。

2.1.3 その他の文字情報

雑誌の表紙には、表題や発行日の他に、その号
の特集記事が一目で解るよう、目次の一部が抜粋
して書かれていることが多い。表紙をスキャンし
た後、表紙の文字情報読み上げボタンを押すと、そ
のページ内の文字情報をすべて読み上げる。これ
により雑誌の特集記事が始めにわかるので雑誌の
選択が効率良く行える。

2.2 目次

視覚障害者が雑誌を読む場合、目次は重要な役
割を果たす。雑誌の場合、単行本のようにその第
1 ページから順に読むのではなく、読みたい記事
が書かれているページを直接開くことが多いから
である。

ここでの課題は、まずスキャンされたページが
目次であることをどう認識するかである。幸い目
次にはその書き方に一定のルールがある。

例

1 項目 タイトル、著者名、区切り記号、ページ番
号の順に記述されている。
グループ 個々の項目がその内容別に分類されて
いる。

このようなスタイルをさらに分析すれば目次の認
識は可能である。

目次ページの抽出は以下のようなユーザーイン
ターフェースで行われる。

1. 目次であると思われるページをスキャナーに
乗せ、目次探索ボタンを押す。
2. 目次でない場合「目次ではありません。別の
ページを入力してください。」とのメッセ
ージが出される。
3. 目次である場合「目次です。次のページを入
力してください。」とのメッセージが出され
る。これは、目次が複数ページに渡っている
かどうか確認するためである。次のページ
が目次でなければ、ここで「目次のスキャン
が終了しました。登録ボタンを押してくださ
い」とメッセージが出されるので、目次登録
ボタンを押す。これでスキャンした目次ペ
ージがすべて目次として登録される。

登録された目次は、後で効率良く目次情報の参照
ができるよう構造化ファイルに変換される。

この結果、以下のようなユーザーインターフェ
ースで目次の参照ができる。

全項目の読み上げ 全項目読み上げボタンが押され
たら目次のレイアウトや構造を考えず第 1 行
目から順に読み上げる。読みたい記事のタイ
トルが読まれた時、ページ確定ボタンを押す。
グループ名の読み上げ グループ名の読み上げボタ
ンが押されたら、目次内のグループの見出し
を読み上げる。この時、見出しの先頭に「グ
ループ 1」「グループ 2」と名前を付けて読
み上げる。

項目の読み上げ あるグループの見出しが読み上げ
られている時、項目読み上げボタンを押すと
そのグループ内の各項目を読み上げる。読み
上げが停止している時、このボタンを押すと
読みたいグループ名を指定してその中の各
項目を読むことができる。

文字列検索 検索したい文字列を入力し、再度この
ボタンを押すと目次の検索を行う。該当の
文字列が存在した場合はその項目を読み上
げる。

読みたい記事が決まったらページ確定ボタンを
押す。ページ探索モードに入るには、読みたい記
事の項目が読み上げられている時にページ確定ボ
タンを押すか、このボタンを押した後、直接その
ページ番号を入力するかのどちらかである。

2.3 ページ番号

ページ探索モードではページ番号のみをスキャンする。これによりスキャンに要する時間を節約できる。

ここでの問題点は、雑誌のページの振り方が曖昧なため、せっかくそのページをスキャンさせてもページ番号が伝達されないことがあるという点である。広告ページや写真のみのページにはページ番号が振られていないことが多い。このようなページが何ページも続くと大変厄介である。また、背景に絵や写真があるとページ番号の抽出や認識に失敗することも多いと予想される。

実際のユーザーインターフェースは以下のようになる。

1. 雑誌の1ページをスキャナーに乗せ、ページ探索ボタンを押す。
2. そのページが指定ページではない場合「失敗 nn ページ」とメッセージを出す。ページ番号が認識できなかった場合は「失敗」とメッセージを出す。
3. 別のページをスキャナーに乗せ、ページ探索ボタンを押す。
4. 該当ページの場合「成功 nn ページ」とメッセージを出す。

2.4 本文

本文のレイアウトには上記で述べたような一定のルールは全く見られない。そこで本章では、図1から4までの場合を取り上げこの記事を視覚障害者が理解するためにどのようなユーザーインターフェースが提供されるべきか検討する。

図1の記事に対応する目次の見出しは「料理・食べもの 夏の日のランチ 青い器に白い料理… 18」である。ページ探索機能を使って18ページを開くが実際にはこのページ番号を認識することは困難であるため、その前後のページをキーにして18ページを開く。このページは写真のみである。この場合、システムは「文字情報がありません」とのメッセージを出す。次のページをスキャナーに乗せ、スキャン開始ボタンを押す。スキャン終了後直ちに文字情報の読み上げが開始される。この時システムは文書の論理構造に基づき構造化

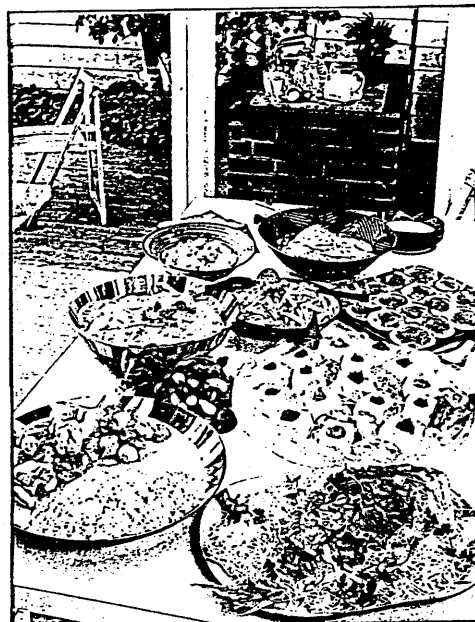


図1: 目次で指示されたページ、写真のみ (暮しの手帖社刊「暮しの手帖」1994 8.9月号より)

ファイルを作成する。ページの先頭から順に読むのではなく、構造化された情報をキーに文書を読む場合には以下のようなユーザーインターフェースの提供が望ましい。(図2参照)

1. タイトル読み上げボタンを押す。この場合横書きで書かれた「夏の日のランチ… 白い料理」と読む。

しかし現実には飾り文字で書かれたこのタイトルを読むことはできない。システムは「タイトルはありません」とのメッセージを出すことになる。

2. 見出しの読み上げボタンを押す。ここには見出しは特にないので「見出しはありません」とのメッセージを出す。
3. 段落確認ボタンを押す。ここには5つの段落があるので「5つの段落があります」とのメッセージを出す。
4. 段落別内容読み上げボタンを押す。「段落1」という段落番号に続いて第1番目の段落を読む。読み飛ばして次の段落を読む時は再度このボタンを押す。「段落2」と読み上げた後、2つ目の段落の読み上げが開始される。本文の内容からユーザーはこのページが目次に対応していることを確認できる。



図2: 文字情報のあるページ (暮しの手帖社刊「暮しの手帖」1994 8.9月号より)



図3: 他のページへの参照がある場合 (暮しの手帖社刊「暮しの手帖」1994 8.9月号より)

5. レイアウト読み上げボタンを押す。ページのレイアウトが伝達されれば内容の理解や疑問点の解決に役立つと思われる。この場合の疑問とは、文字情報が予想以上に少ないという点かもしれない。ここでいうレイアウト情報とは例えば以下のような内容を意味する。

- (a) ページの上1/3に横長の模様がある。下2/3は左半分に写真、右半分に縦書き2段で文字情報がある。ここには5つの段落がある。
- (b) ページ全体の1/3が文字情報である。

次に図3を効率良く読むためのユーザーインターフェースを考える。

1. 見出し読み上げボタンを押す。「見出し1 かわり冷奴」「見出し2 水ぎょうざ」と読む。
2. 見出し別内容読み上げボタンを押す。「見出し1 かわり冷奴」と読み上げた後、その説明文を読む。段落の場合と同様、次の見出しの説明文に飛びたい時は、もう1度このボタンを押す。「見出し2 水ぎょうざ」と読み上げた後、その説明文を読む。ここでは「作り方は139ページ」とあるので次に139ページを開く。

(注意) ここでは水ぎょうざの写真の上に「つるとした口当たりで栄養満点の水ぎょうざ…」という写真の説明がある。見出し別の内容読み上げだけを行うと、この情報は誤って「かわり冷奴」の説明の一部となる可能性がある。このような場合を考慮し、読み上げにおける次候補伝達機能を準備する必要があると考える。この機能により、水ぎょうざの写真の説明を、次に正しく「水ぎょうざ」の説明として読むことができる。

ここで139ページを開くため、ページ探索ボタンを押す。139ページが開いたらスキャン開始ボタンを押す(図4参照)。このページは文字情報だけである。以下のようなユーザーインターフェースが提供されれば効果的である。

1. タイトル読み上げボタンを押す。「夏の日の…」とタイトルを読み上げる。ただしここでは認識できないので「タイトルはありません」と読む。
2. 見出し読み上げボタンを押す。「見出し1 水ぎょうざ、見出し2 鯛の中国ふう…」と読む。
3. 見出し別内容読み上げボタンを押す。「見出し1 水ぎょうざ」に続いてその説明文を読み始める。再度、このボタンを押すと次の見



図4：料理の作り方のページ（暮しの手帖社刊「暮しの手帖」1994.8.9月号より）

出しに飛ぶことができる。

4. 小見出し読み上げボタンを押す。見出しの内容が読まれている時このボタンを押すと「小見出し1 材料, 小見出し2 作り方」と読む。
5. 小見出し別内容読み上げボタンを押す。「小見出し1 材料」の後にその内容を読む。次に飛びたい時は再度このボタンを押す。

2.5 ナビゲーション機能

表紙の確認から始まり本文の内容を理解するまで、視覚障害者が自力で雑誌を読むためには、晴眼者とは比較できないほどの時間と根気が必要である。そこでシステム側ではできる限りの情報をユーザーに伝達する必要がある。ナビゲーション機能はこの目的達成のために有効に利用できる機能である。ナビゲーション機能とは、システムが現在どのような状況下でありユーザー側からどのようなリクエストが出せるかをユーザーが指定すればいつでも伝達してくれる機能である。

例1

現在表紙が入力されています。ここでは表題の探索、テンプレートの作成/登録、発行日の探索または検索、その他の文字情報の確認ができます。それぞれの機能を選択するため以下のボタンを押し

てください。

メニュー表示及びテンキーのキーアサインを表示する。

例2

現在「Hanako」を読んでいます。あなたが探索しているのはnnページのxxという記事です。

この他にもさまざまなナビゲーションが考えられるが、このようなフィードバックを行うことでユーザーが道に迷った時、あるいはなんらかの理由で処理を中断し再開する時など大変有効に利用できる機能である。

3 ボタン操作

雑誌を読むための操作はすべて機能ボタンで行うことにする。本稿で使用したボタンを機能別に分けると表1のようになる。これらのボタンは基本的にテンキーに割り付けられる。これにより、キーボードに不慣れな視覚障害者でもテンキーだけでシステムを利用することができる。しかし、すべての機能ボタンの配置をユーザーが容易に記憶できるように割り付けることは大変難しい。そこで、すべてのボタンを一旦テンキーに割り付けるがユーザーの指定により自由に変更できるようにする。また、初心者の場合を考慮しボタンの説明を行うヘルプ機能なども準備する。この他、機能を階層化し、それぞれの階層別にメニューにして表示すればカーソル移動により音声/点字で現在選択できる機能の確認ができるので有効であると考ええる。

また本稿では、個々の機能別にボタンを設定したが、全体の機能を階層化し機能別ではなく階層別にボタンを設定すれば、少ないボタンで容易に操作できる可能性もあると考える。いずれにしても、1つの方法を提供するのではなく、ユーザーのレベルや好みに合わせて選択できるように、きめの細かいユーザーインターフェースの提供が必要である。

• 共通

- スキャン開始ボタン
- 全ページ読み上げボタン
- 読み上げ停止ボタン

- 表紙
 - 表題探索ボタン
 - テンプレート作成ボタン（表題抽出ボタン、テンプレート登録ボタン）
 - 発行日探索ボタン（発行日検索ボタン）
 - 表紙内文字情報読み上げボタン
- 目次
 - 目次探索ボタン（目次登録ボタン）
- 目次参照
 - グループ名読み上げボタン
 - 項目読み上げボタン
 - 文字列検索ボタン
- ページ番号
 - ページ確定ボタン
 - ページ探索ボタン
- 本文
 - タイトル読み上げボタン
 - 見出し読み上げボタン
 - * 見出し別内容読み上げボタン
 - * 小見出し読み上げボタン
 - ・ 小見出し別内容読み上げボタン
 - 段落確認ボタン
 - * 段落別内容読み上げボタン
 - レイアウト読み上げボタン

新の認識エンジンが本システムで利用できるよう考慮しなければならない。

次にパフォーマンスも重要である。1枚の印刷文書を読むために待てる時間には限界がある。本システムでは、OCRシステムが従来行ってきた処理に、さらにいくつかの付加がかかるためパフォーマンスは落ちると予想される。システム開発上、この点についても考慮し、より早くアウトプットが提供できるようにしたいと考えている。

今回提案したユーザーインターフェースは視覚障害者が独力でスキャナーを操作し雑誌を読むためのものである。そのため、独力で操作できる反面、読みたい記事の書かれているページを開くまでに多少時間がかかる可能性がある。今後、スキャナー入力の処理をボランティアなどに任せ、ネットワークを介してその情報を入手し、検索処理だけを独力で行うようにすれば雑誌を読むという点については効率化が計れるのではないかと考える。今後は上記で述べた問題を解決していくと共に、雑誌だけではなくその他の印刷文書にアクセスするためにはどのような機能やユーザーインターフェースが提供されるべきか、さらに研究し、視覚障害者にとって使いやすい文書処理システムを実現したいと考えている。

なお、本研究は医療福祉機器技術研究開発制度の一環として、新エネルギー・産業技術総合開発機構（NEDO）からの委託により実施したものである。

4 おわりに

今回の考察の結果、視覚障害者が直接雑誌を読むためには、文字に変換された情報をそのまま伝達するだけでなく、文書を構造化ファイルに変換しその構造情報に基づき内容の読み上げができるようなユーザーインターフェースの提供が必要であることが明らかになった。

しかし、このようなユーザーインターフェースを実現するためにはいくつかの問題がある。第1にOCRの精度である。現在のOCRは、定形パターン処理という範囲内でその技術が実用化されている。しかし、視覚障害者はあらゆる文書をOCRで読みたいと考えているため、今後常に最