

視覚障害者用文字認識システム

4T-7

菅原一秀

日本アイ・ビー・エム株式会社 東京基礎研究所

1 はじめに

視覚障害者が日常的な印刷情報に簡単にアクセスできるためのシステムの研究開発を行なっている。出力は音声又は点字によるので不要な情報は飛ばして、必要な情報に効率的にアクセスできることが特に重要である。そのためにスキャナから入力された文書画像のレイアウト解析を行ない、論理構造を抽出して利用する。

本稿では雑誌を読む場合及び請求書を読む場合の手順を想定し、効率的なアクセスのために必要な技術を検討する。また、レイアウト解析から得られる情報と、文字認識の内容から得られる情報を組み合わせて文字列をグループ化し、文書の論理構造を抽出する手法を提案する。

文書画像から論理構造を抽出する研究は文書処理の効率化を目指して行なわれてきた[1][2][3]。これらの論文ではいずれも対象を技術論文や、特許公報などの専門的文書に限っており日常的な、雑誌、請求書などの文書についてはこれまで対象外であった。

2 文書へのアクセス

以下に本システムを使用して雑誌を読む場合及び請求書を読む場合の具体的な手順を示す。

2.1 雑誌の場合

1. 視覚障害者は、雑誌の目次と思われるページをスキャナから入力する。
2. システムが目次のページと判断できない場合は、視覚障害者にその旨を伝え、別のページを入力するよう音声で指示する。また、本の上下が間違っていればその旨も伝える。
3. 目次であれば、システムは書かれている内容（タイトル、ページ番号）を視覚障害者に提示する。システムは目次の内容を保存し、視覚障害者の呼び出しに常に応じる。

4. 視覚障害者は、自分の読みたいページが決まれば、該当ページと思われる部分を開きスキャナで入力する。システムは、入力されたページにつき、ページ探索機能によりページ番号を認識し、それを視覚障害者に音声で提示する。
5. 目的のページと視覚障害者が確認した場合には、見出しや本文の内容を提示するようシステムに指示し、音声あるいは点字により読み上げさせる。

2.2 請求書の場合

1. 視覚障害者は、自分宛にきた請求書をスキャナで入力する。
2. システムはあらかじめ準備された標準的なレイアウトモデルから、入力されたものに適合するレイアウトモデルを選択し、請求書であれば設定された順位どおりに、例えば、発行元の名前、日付、請求金額を読み上げる。（注：適合しない場合は暗黙者のボランティアが、その文書のレイアウトを設定したレイアウトモデルを作成する。その際、視覚障害者の希望に応じ、提示する内容の優先順位を付けておく。）
3. 視覚障害者が、他の細かい部分について知りたい場合には、あらかじめコマンド（指示ボタン）を決めておくことにより、その指示に従って読み上げさせる。

これらを実現するために以下のことができればいけない。

目次であることの判定：目次らしさの判定を行なう。目次の論理構造の抽出が成功したか否かによって判定する。

目次の論理構造の抽出：記事のグループ化、記事のタイトル、著者、ページ番号の対応づけを行なう。

ページ番号の認識：ページ番号を見つけだし、数値を得る。

記事内容の論理構造の抽出：見出し、本文の抽出とそれらの関連づけを行なう。

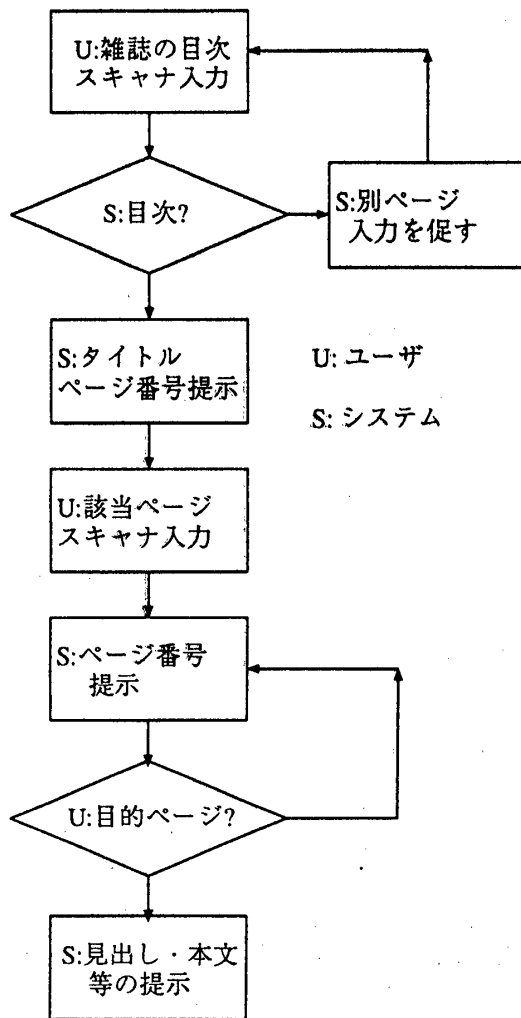


図 1: 雑誌へのアクセス

レイアウトモデルの選出: いくつかの標準的なレイアウトモデルから入力画像に適合するものを選び出す。

ユーザ定義レイアウトモデルの作成支援: エンドユーザが請求書などのレイアウトモデルを容易に作成できるようにする。

次節では上記の目次、記事などの論理構造の抽出に利用できる文字列のグループ化について述べる。

3 グループ化のための手法

目次はタイトルとページ番号がならんでいるだけではなく、大抵はあるまとまりを持った項目ごとにグループ化されている。それらは上下(左右)の空白の幅の違い、インデント、字体の違い、あるいは枠で囲まれることなどによって表されている。記事の場合は本文はパラグラフごとにまとめられ、いくつかのパラグラフをまとめて見出しが振られる。請求書の場合は請求項目ごとに

日付、内容、請求金額などが並べられていて、個別の請求内容を表すグループをなし、最終行に合計金額が記載されている。論理構造を抽出するにはこれらのグループをなすものをまとめて扱うことが必要である。

簡単のため以下では文書は横書きであり段組は一段になっているとする。一行ごとに文字列ブロックの位置と文字認識の結果が与えられているとする。ただし目次の場合には直線や直線状の点線も扱う。文字をローマ数字+位取り記号(ページ番号や金額用)とその他の文字に分類しておく。ただし左端の文字列ブロックではすべての文字を区別しないで扱う。これは目次のタイトル—ページ番号の対応と、記事からパラグラフをまとめる場合を同様に扱うためである。隣合う行の類似度を左右端の位置の一致度、分類された文字及び直線又は直線状の点線の対応の具合により定める。いき値を定めて類似度の高いものを同じグループに属するとする。このいき値や要素ごとの重みづけは実験的に求める。この手法ではパラグラフの先頭のインデントのある行やパラグラフの最後の短い行末からなる行は独立のグループとされる可能性が高いが、それぞれグループの先頭又は最後尾につくことができるとすることで解決する。

この様にして作られたグループを使うことにより、不要な箇所をまとめてスキップする、見出しを先に読み、興味のある所に対応するパラグラフの内容を読むなどの効率的なアクセスが可能となる。

4 まとめ

視覚障害者用文字認識システムにおいて、ユーザが読みたい個所に速やかにアクセスすることを可能にする仕組みを検討してきた。今後の課題は各種の文書を対象にこの手法の実験を行ない、本アプローチの有効性を実証し、更には視覚障害者が実際に操作できるプロトタイプを作成することである。

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。

参考文献

- [1] 小山乃り子、黄瀬浩一、高松忍: “適用性を高めた印刷文書の論理構造化の一手法”, 電子情報通信学会総合大会講演論文集, D-581, p. 307, March, 1995
- [2] 山下品夫、天野富夫: “モデルに基づいた文書画像のレイアウト理解”, 電子情報通信学会論文誌, D-II Vol.J75-D-II No.10 pp.1673-1681, Dec., 1992
- [3] 山田満: “文書画像の ODA 論理構造化文書への変換方式”, 電子情報通信学会論文誌, D-II Vol.J76-D-II No.11 pp.2274-2284, Nov., 1993