

形状歪み補正を利用した表及びグラフを含んだ書籍の音読

鈴木 良典[†] 山下 淳[‡] 金子 透[†] 浅間 一[‡]

[†]静岡大学工学部 〒432-8561 静岡県浜松市中区城北 3-5-1

[‡]東京大学大学院工学系研究科 〒113-8656 東京都文京区本郷 7-3-1

E-mail: [†]{f0810089, tmtkane}@ipc.shizuoka.ac.jp [‡]{yamashita, asama}@robot.t.u-tokyo.ac.jp

あらまし 書籍などの紙媒体の印刷物は我々の生活において重要な情報源の1つであるが、視覚障害者にとって紙媒体の印刷物から情報を得ることは容易なことではない。そこで本研究では、既存の製品で読み上げの対象とされていない表、グラフを対象とした音読システムの構築を目的とする。また、スキャナ装置には印刷物において綴じ目付近に形状歪みが生じるという問題点がある。この問題を解決するためにステレオカメラによる形状歪み補正を画像に行う。実験により、本手法が書籍の表、グラフの読み上げに有効であることを示した。

キーワード ステレオカメラ、音読、形状歪み補正

Distortion Correction of Documents and Reading Out of Their Tables and Graphs

Ryosuke SUZUKI[†] Atsushi YAMASHITA[‡] Toru KANEKO[†] and Hajime Asama[‡]

[†]Faculty of Engineering, Shizuoka University 3-5-1 Johoku, Naka-ku, Hamamatsu-shi, Shizuoka, 432-8561, Japan

[‡]School of Engineering, The University of Tokyo 7-3-1 Hongou, Bunkyo-ku, Tokyo, 113-8656, Japan

E-mail: [†]{f0810089, tmtkane}@ipc.shizuoka.ac.jp [‡]{yamashita, asama}@robot.t.u-tokyo.ac.jp

Abstract Print media like books is an important source of information. However, it is very difficult for blind people to obtain the information from print media. This research aims at making a system to read out tables and graphs which cannot be read out by existing products. In addition, there is a problem that scanners yield distortion at the seam of books. In order to solve this problem, we correct the image distortion by using stereo camera. Experimental results showed the effectiveness of the proposed method.

Keyword Stereo Camera, Reading Out, Correction of Shape Distortion

1. 序論

書籍などの紙媒体の印刷物は我々の生活において重要な情報源のひとつであるが、視覚障害者にとって紙媒体の印刷物から情報を得ることは容易なことではない。

以前より、活字を音読することにより視覚障害の手助けをする試みが行われている。

視覚障害者のための既存製品としては、イメージスキャナとパソコンを組み合わせた音声読み上げ読書器などが製品化されている。これらの製品は文書画像をスキャナ装置によって読み取るため、書籍のような印刷物において綴じ目付近に形状歪みが生じ、文字の誤認識につながるという問題点がある。

文章の自動読み上げの研究として、文書画像取得の手段にカメラを用いたもの[1]もあるが、取得される画像に形状歪みが見られ、文字認識において誤認識が起

こるという問題がある。

上記の文書画像の歪みの問題に対して、我々はステレオカメラを用いて3次元計測した紙面を平面に展開して形状歪み補正し、これを文字認識する方式を従来から提案している[2][3]。そしてこの方式を用いて文字認識した文書の音読システムの構築を行っている[4]。

また、上記の音読システムでは文章のみを音読の対象としており、表またはグラフ等には対応していない。これでは視覚障害者が書籍等の紙媒体から情報を得る際の機能として不十分である。

そこで本研究では、これまでのシステムが対象としていなかった表、グラフを含んだ書籍を音読するシステムの検討を行った。

2. 処理手順

本研究では書籍の文章、表、棒グラフ、折れ線グラフを音読の対象とする。

以下に本研究の処理手順を示す(図1)。システムの構築に必要な処理は5段階に分けられる。

1. 2台のカメラで書籍画像の取得を行う。
2. 書籍画像の形状歪み補正を行う。
3. 補正された書籍画像に文字認識と表、棒グラフ、折れ線グラフの認識を行う。
4. 表、棒グラフ、折れ線グラフと判別された領域に2.4節~2.6節で説明する処理を行い、表またはグラフの情報を音読できる形に整理する。
5. 文章、表、グラフの情報を音読して使用者に情報を伝える。

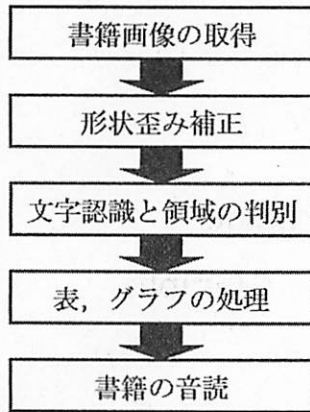


図1 処理手順

2.1. 書籍画像の取得

書面が歪曲している書籍からでも画像が取得できるように、本研究では2台のカメラにより画像を取得する。

実際の撮影の際には見開きの書面全体が写る様に書籍を設置し、書面に影ができないように照明を当て、固定された2台のカメラによってステレオ撮影を行う。

図2に実際に書籍画像を取得する際に使用する装置を示す。

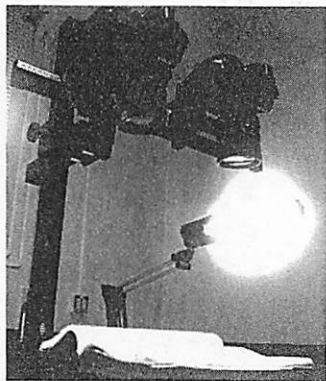


図2 実験環境

2.2. 形状歪み補正

カメラによる取得画像は書面形状の歪みにより、そのままでは良い文字認識結果が得られない。そこで本研究では、ステレオ画像を用いて書籍の形状歪み補正を行う。

形状歪み補正の手法としては、秋本らの手法[3]を用いる。[3]では、まず取得したステレオ画像から書面の形状を推定し、得られた形状を元に推定形状を引き伸ばすことにより形状歪みを補正し、最終的な補正画像を生成する(図3)。

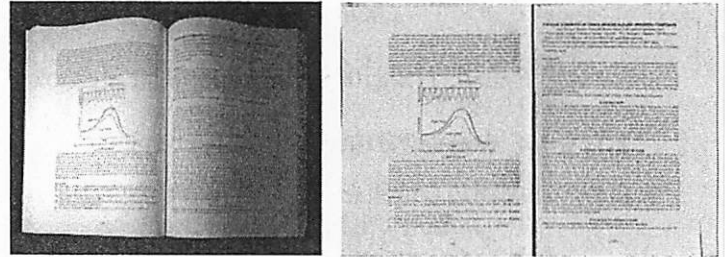


図3 秋本らの手法[3]による補正結果

2.3. 文字認識と領域の判別

生成された補正画像に対して文字認識ライブラリによる処理を適用して文字認識を行う。

また、図4のように文字、表、グラフの領域の判別も行い、文字認識ライブラリの文字、表、グラフの識別はそれぞれを矩形領域に自動分別する。

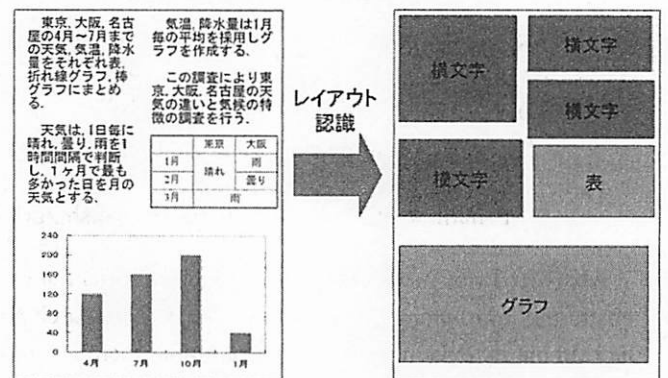


図4 レイアウト認識結果例

さらに、画像中における文字、表、グラフそれぞれの矩形左頂点上の座標と幅、高さのデータも取得する。

2.4. 表の処理

表の場合、最も左側の列に存在する文字を項目1、最も上段の行にある文字を項目2とする(図5)。ここで、表の前提条件として表の項目1, 2がそれぞれ存在するものとする。

まず、最も左の列のセルの数 m と最も上の行のセルの数 n とし、表を $m \times n$ に分割する(図6)。

	項目2	東京	大阪
項目1	1月	晴れ	雨
	2月	晴れ	曇り
	3月	雨	

図5 分割前の表

	項目2	
項目1	東京	大阪
	1月	晴れ
	2月	晴れ
	3月	雨

図6 分割後の表

分割の際、図5で色が付いた領域に存在するセルで

は図6で示すように同じ文字情報が格納される。

表の文字を音読する際には上の行から図7の矢印で示した順番で読み上げる。

	項目2	東京	大阪
項目1	1月	→	→
	2月	→	→
	3月	→	→

図7 分割前の表と読み上げ順序

読み上げる際には、セル毎に横の項目1、上の項目2を探索し、セル内の文字を音読する前に読み上げる。

2.5. 棒グラフの処理

棒グラフの横軸、縦軸の目盛りに記された内容をそれぞれ項目と定義する。ここではグラフの各棒が示す値に項目が必ず存在するグラフを対象とする。

まず、グラフとして判断された領域にハフ変換を行い、棒の輪郭の左右の直線を検出する。この検出した2本の直線から端点のx、y座標を取得する。

次に図8で示すように取得したx、y座標を基に各棒の項目を探索し、特定する。項目の探索は、取得した座標が項目の矩形領域の範囲に存在するかどうかで決定する。

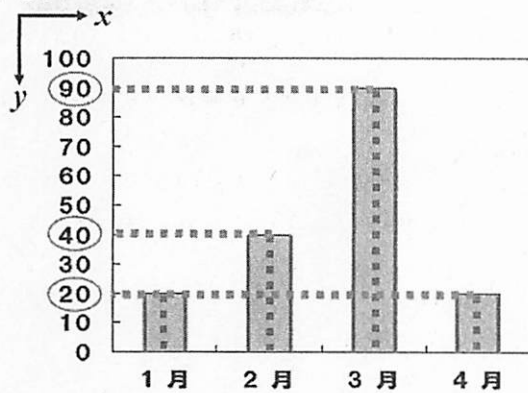


図8 棒グラフのデータ判別例

ここでの矩形領域の範囲とは、図9で示すような項目の座標と幅、高さから得られる範囲である。横軸の項目の場合は(a)で示すようにx座標と幅、縦軸の項目の場合は(b)で示すようにy座標と高さから項目の探索範囲を取得する。

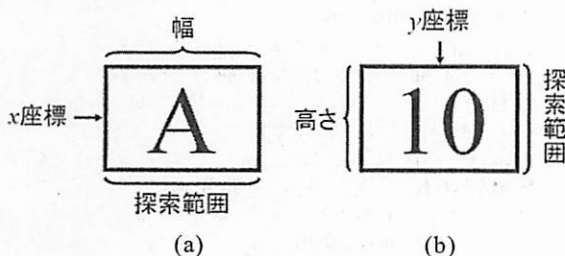


図9 項目の探索範囲

横軸の場合は、項目の探索範囲に取得した各棒のx座標の中心が存在する場合に棒の項目として特定する。縦軸の場合も、取得した各棒の探索範囲にy座標がある場合に棒の項目として特定する。

各棒グラフを読み上げる際には、横軸の項目を読み上げた後に縦軸の項目を読み上げることで音読を行う。

2.6. 折れ線グラフの処理

折れ線グラフでもグラフの横軸、縦軸の目盛りに記された内容をそれぞれ項目と定義する。ここでもグラフの各棒が示す値に項目が必ず存在するグラフを対象とする。

まず、ハフ変換を用いて折れ線グラフの直線を検出し、グラフの矩形領域の座標と高さ、幅から横軸の項目を特定する。

その後、折れ線グラフの横軸のデータのx座標と検出された直線の交点からy座標を導出する。取得したy座標を基に図10で示すように棒グラフの項目の特定方法と同じ手法で項目を特定する。

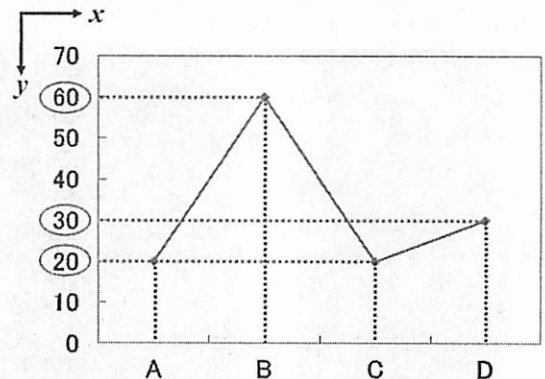


図10 折れ線グラフのデータ判別例

データの読み上げの際には棒グラフと同じように横軸のデータを音読後に縦軸のデータを音読する。

2.7. 書籍情報の音読

横書き文章の場合には、左上から音読を行い、縦書き文章の場合には右上から音読を行う。

音読する順番としては、文字情報を音読した後に表、グラフの情報を音読する。

3. 実験

文章部分とそれ以外が同じレイアウトの書面を使用して、表、グラフの認識および音読を行った。

3.1. 実験環境

実験は図2で示した環境で行った。テーブルに開いた書籍を置き、上方から2台のデジタル一眼レフカメラによって書面を撮影した。カメラはNikon社製のD700を2台使用し、画像サイズは左右それぞれ4256×2832ピクセルである。また文字認識ライブラリはPanasonic社製の活字認識ライブラリVer13を使用した。

今回の実験で使用する表は、表1に示したものを使用した。

表 1 実験で使した表

	東京	大阪	名古屋
4日	晴れ	曇り	晴れ
5日	雨		曇り
6日	晴れ	雨	
7日	雨	晴れ	

また棒グラフ,折れ線グラフは図 11 で示したものを
使用した。

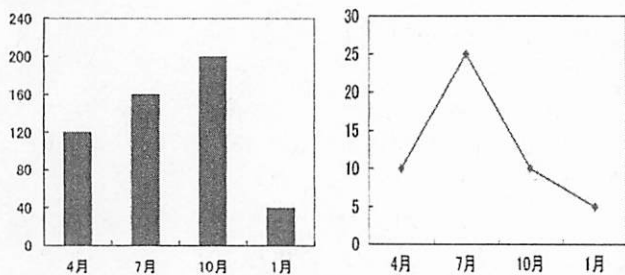


図 11 実験で使したグラフ

3.2. 結果

図 12 は書籍の形状歪み補正の結果である。

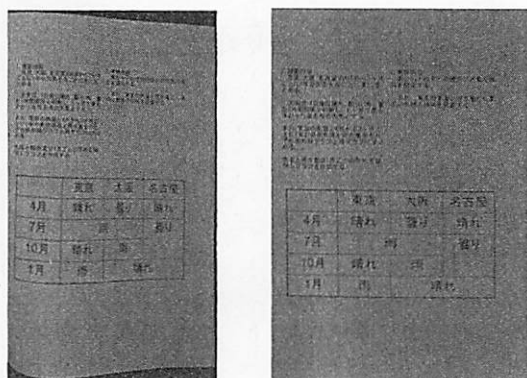


図 12 形状歪み補正結果

図 13 は実験で使した表またはグラフを含んだ書
面全体での音読順序を示した結果であり, 数字順に文
字群を音読した。

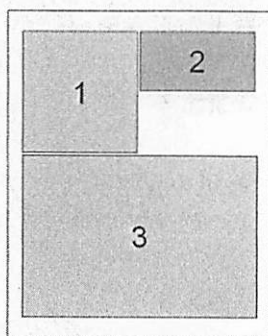


図 13 書籍全体の音読順序

表の音読結果を下記の括弧内に示す。

- 1 行目:「4 日, 東京, 晴れ, 4 日, 大阪, 曇り, 4
日, 名古屋, 晴れ」
2 行目:「5 日, 東京, 雨, 5 日, 大阪, 雨, 5 日,
名古屋, 曇り」

3 行目:「6 日, 東京, 晴れ, 6 日, 大阪, 雨, 6 日,
名古屋, 曇り」

4 行目:「7 日, 東京, 雨, 7 日, 大阪, 晴れ, 7 日,
名古屋, 晴れ」

図 14, 図 15 はそれぞれ補正した画像の棒グラフと
折れ線グラフにハフ変換と音読を実行した結果である。
図の白色の直線が直線検出結果であり, それぞれの項
目の音読順序を数字で示した。

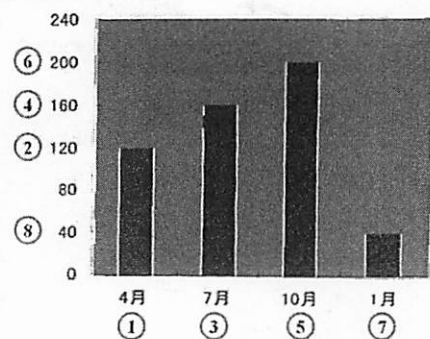


図 14 棒グラフの直線検出結果と音読順序

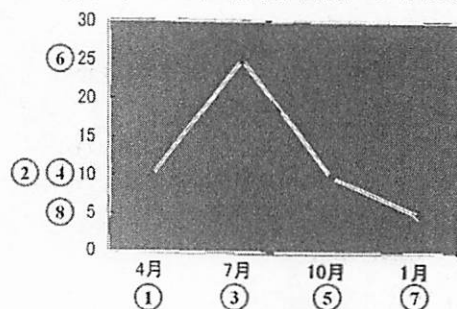


図 15 折れ線グラフの直線検出結果と音読順序

4. 結論

文章, 表, グラフが混在する印刷物の内容を音読す
るシステムの構築を行った。今後の課題としては, 現
在は手動で行っている表またはグラフの判別を自動化
する必要がある。また, グラフの示す値が項目の目盛
りに一致しない場合でも音読を可能にすることでシス
テムの利便性を向上させる必要もある。

文 献

- [1] 大西淳児, 小野東:“汎用ウェブカメラを用いた
墨字文書自動読み取りソフトウェアの試作”, 筑
波技術大学テクノレポート, Vol.16, pp.18-21,
2009.
- [2] 川原子淳, 山下淳, 金子透, 三浦憲二郎:“ステ
レオビジョンシステムによる表面湾曲書籍の文
字認識”, 映像情報メディア学会誌, Vol.56, No.5,
pp.858-862, May 2002.
- [3] 秋本鍊太郎, 山下淳, 金子透:“ステレオカメラ
を用いた歪曲形状書籍画像補正の高速化”, 第 15
回画像センシングシンポジウム講演論文集,
pp.IS1-14-1-IS1-14-8, 2009.
- [4] 大石将太, 山下淳, 金子透:“ステレオ画像を用
いた書籍画像の形状歪み・陰影補正による文字認
識および読み上げ順序を考慮した音読システム”,
精密工学会画像応用技術専門委員会サマーセ
ミナー2010 テキスト, Vol.19, pp.13-16, 2010.