

# 書棚画像からの書籍の背文字領域の抽出について

石川 幸博・大倉 充\*・塩野 充\*・橋本 禮治\*\*

岡山理科大学大学院工学研究科修士課程電子工学専攻

\*岡山理科大学工学部情報工学科

\*\*岡山理科大学工学部電子工学科

(1996年10月7日 受理)

## 1. ま え が き

情報化社会の現代において、出版される書籍の数は膨大なものである。その書籍を個人が有効に活用するために図書館などの公共施設においてはしばしば書籍のデータベースが用いられており、今後、このデータベースの需要はますます増えると考えられる。しかしながら、現在、書籍のデータベースを構築するには膨大な手作業が必要であり、書籍のデータベースの自動作成が有用であると考えられている。そのためには、カメラから入力された画像情報からの各書籍のタイトル、著者名、出版社名などの抽出及び認識を行う必要があると考えられる。本研究では、その第一段階として撮影した書棚画像から各書籍領域を決定し、背文字領域を抽出する手法について検討する。

## 2. 使用データ

本研究では、カメラを用いて書棚を真正面から撮影し、得られた写真をイメージスキャナによって、解像度180dpi, RGB 各256階調でコンピュータに入力したデータを使用した。画像サイズは550×700画素である。図1に入力画像の一例を示す(濃淡画像表示)。なお、カメラにより撮影したデータを画像処理する際には、照明の有無、照明法及びその配置が問題となる場合があるが、本研究では特殊な照明器具を用いず、部屋の蛍光灯の照明のみを用いて撮影した。

## 3. 書籍の領域抽出

各書籍を構成するエッジを求めるためには、画像の濃度勾配により分割された(セクタ分割と呼ぶ)画素にラベリングを行い、その領域を最小2乗法によって書籍のエッジを直線近似する方法が提案されている<sup>1)</sup>。しかし、この方法ではエッジ候補領域が途切れている場合に抽出が困難となる。本研究では、セクタ分割した画素に対して Hough 変換<sup>3)</sup>を用いた直線検出によって書籍のエッジを求め、書棚画像からの書籍の1冊ごとの領域抽出を行う。



図1 原画像

### 3.1 白黒濃淡画像の作成

各書籍の領域抽出では、白黒濃淡画像を用いたため。入力画像に HSI 変換<sup>3)</sup>を施し、その明度値を256階調に正規化して白黒濃淡画像の作成を行った。

### 3.2 濃度勾配による画素のセクタ分割

書棚画像における書籍の画像的特徴として、書籍の背の丈方向や書籍間に直線的なエッジ成分が検出できると考えられる。画像からエッジ成分を検出するために、図2に示すマスクと式(1)~(3)を用いて各画素の濃度勾配ベクトルの方向 $\theta$ を求めた後、その勾配方向により各画素を3つのセクタに分割する<sup>1)</sup>(図3参照)。本研究では、処理を簡素化するために対象書籍の傾きを $-47^\circ < \theta < 47^\circ$ に制限する。また、セクタの角度に重複があるのは、セクタ間にまたがったエッジ領域を抽出するためである。

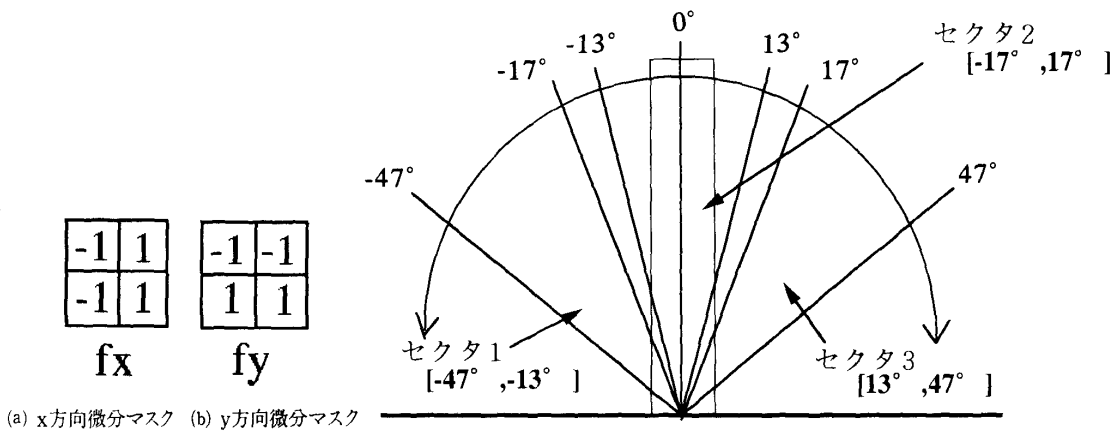


図2 使用マスク

図3 セクタ分割

$$f_x(x, y) = f(x+1, y) + f(x+1, y+1) - f(x, y) - f(x, y+1) \quad (1)$$

$$f_y(x, y) = f(x, y+1) + f(x+1, y+1) - f(x, y) - f(x+1, y) \quad (2)$$

$$\theta(x, y) = \tan^{-1}(f_y(x, y) / f_x(x, y)) \quad (3)$$

### 3.3 Hough 変換で用いる画像データの作成

濃度勾配によって分割された各セクタの画像 (図4 参照) に対して, Hough 変換により直線, すなわち書籍のエッジ候補検出を行うのであるが, 画像中には書籍のエッジ成分以外の画素も入っている。これらの画素数を減少させるために, 書籍のエッジは比較的 低い濃度値を持っていることを利用して, この画像からあるしきい値以上の濃度値を持つ画素を削除する。また, マスク処理により雑音成分と考えられる孤立画素を削除する。図5に Hough 変換用データの一例を示す。

### 3.4 書籍のエッジ検出

上述した処理によって求められた画像から, Hough 変換<sup>3)</sup>によって書籍を構成するエッジを検出する。この処理に用いる  $\rho$ - $\theta$  パラメータ空間における  $\theta$  の間隔は  $1^\circ$  とした。あまり  $\theta$  を細かくとると雑音成分の影響を受けやすいと考えられ, 予備実験においても  $\theta$  の間隔  $1^\circ$  の際に最も良好な結果が得られた。また,  $\rho$ - $\theta$  平面から  $x$ - $y$  平面に直線を投影する際のしきい値は最高カウント数の43(%) とした。この値も予備実験により求められた値である。しかし, 書棚画像には多数の書籍があると考えられ, 書籍によりそれぞれ Hough 変換用画像中のエッジ成分の画素の抽出率も一様ではない。つまりシャープなエッジは, 濃度勾配方向が一定値に集中するのでより密に抽出されており, そうでないものは粗く抽出されている。このためしきい値までの全ての値を投影してしまうと, 同一エッジ付近に何本もの直線が現れる場合がある。これを防ぐために以下の処理によって直線検出を行った。最高カウント数から順に直線検出を行い, 検出された直線の  $x$ - $y$  平面上での傾きと切片を

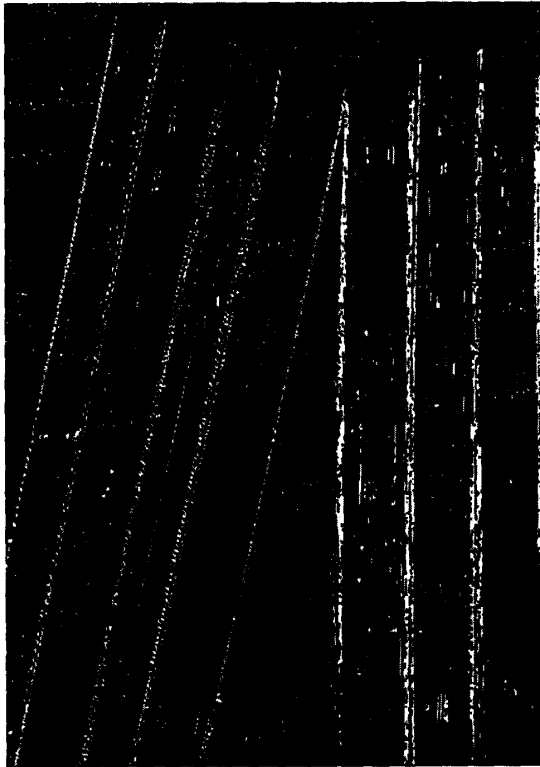


図4 セクタ分割画像の一例（セクタ2）

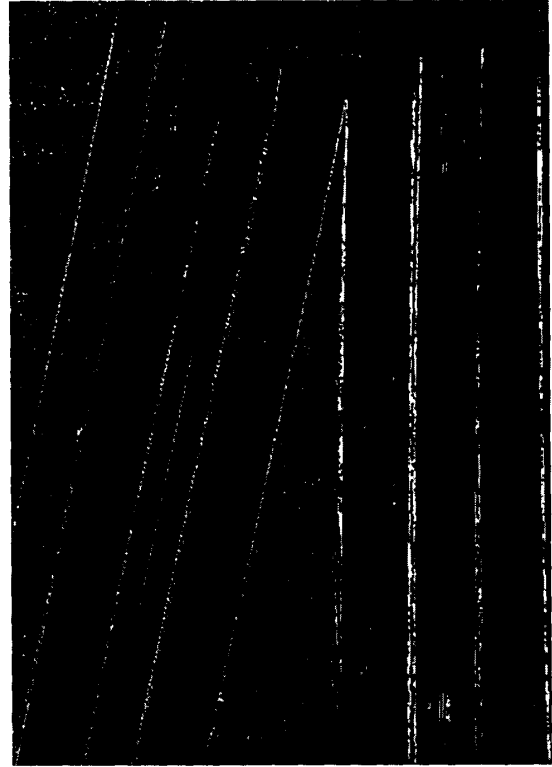


図5 Hough 変換用データの一例

記憶する。ここで傾きと切片に一定のしきい値を設けておき、検出しようとする直線の傾き、切片とすでに検出された直線の切片、傾きとの差がそれぞれしきい値以内であれば、この直線の検出は行わない。本研究では、傾きのしきい値を $2^\circ$ 、切片のしきい値を6画素とした、この方法では近接した2本の平行なエッジ成分は一本しか抽出されないこともあるが、その後の処理には影響しない。

### 3.5 書籍領域の決定

Hough 変換により求められた直線群（図6参照）より各書籍領域の決定を行い、書籍を1冊ごとに分離する。領域の決定は以下の条件によって行った。1本の基準直線を定め対象直線を順に移動させてゆき、その対象直線に対して、

- (1) 傾きの差が定められたしきい値の範囲内である。
- (2) 切片の差があるしきい値の範囲内である。

の2つの条件を満たせば基準直線と対象直線の1組を1冊の書籍のエッジであると決定する。これを全ての検出直線に対して行う。基本的な条件はこれだけあるが、その他に対象直線が(1)、(2)の条件を満たしても、その対象直線に近接した直線により平行な直線があればそちらを対象直線にするなどの細かい条件がある。上述した処理によって書棚画像から各書籍領域の抽出を行った一例を図7に示す。

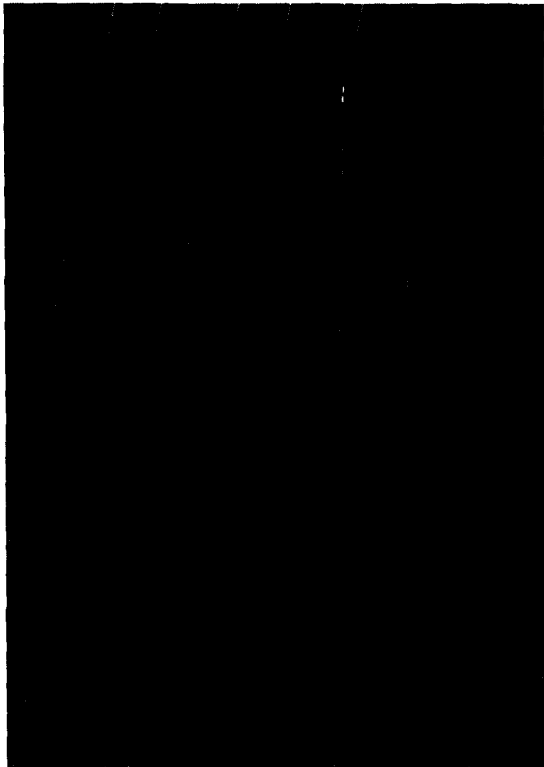


図6 Haugh 変換により求められた直線群

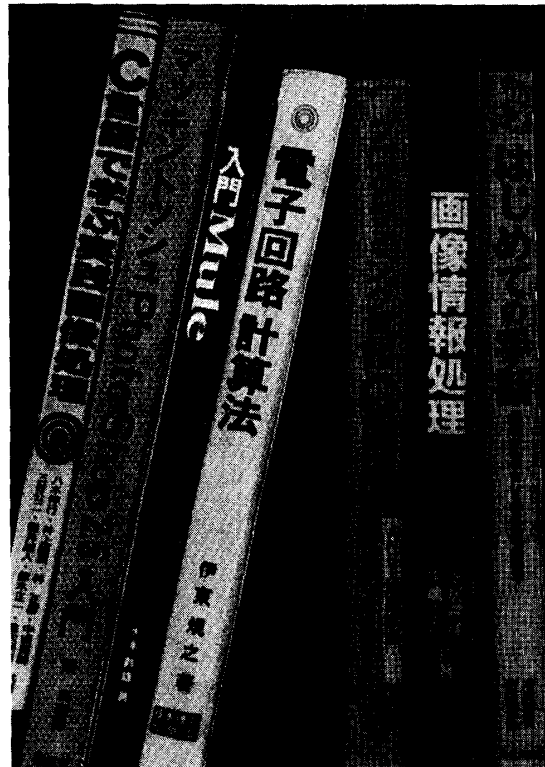


図7 書籍の抽出例

#### 4. 背文字領域の抽出

本章では3章で抽出された書籍領域に対して、HSI変換による入力画像の分解処理を用いた背文字領域の抽出について述べる。また、本章で述べる各処理は、抽出された各書籍領域に対して行う。

##### 4.1 アフィン変換による傾き補正

これまでに決定された書籍領域に対して、エッジ直線の傾きによりアフィン変換を用いて回転させ、書籍領域が直立するように補正を行う。図8に傾き補正を行った画像例を示す。

##### 4.2 HSI変換による分解画像の作成

書籍の背文字領域は様々な色が使われている場合が多く、1冊の書籍に対して何色も用いられている場合も少なくない。このような画像を2値化する際には、固定しきい値法ではうまく2値化できない場合があるので、局所的な濃度値の平均を用いた動的なしきい値決定法が多く用いられるが、本研究ではカラー情報を有効に利用するためにHSI変換を用いたカラー分解による2値化法<sup>2)</sup>を用いる。本研究では、まず入力画像を色相により60°ごと等間隔で6枚に分解した(図9参照)。また分解する際に境界値付近の色相値を持つ領域が

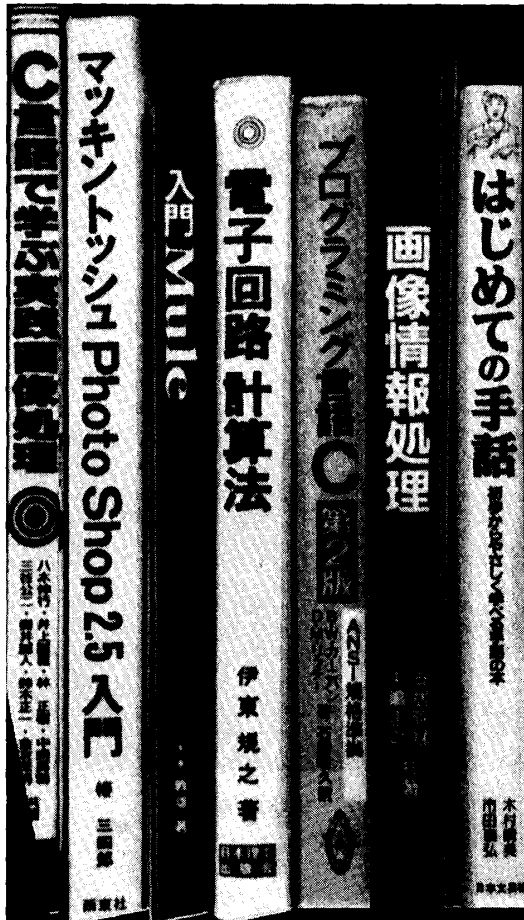


図8 傾き補正を行った画像例

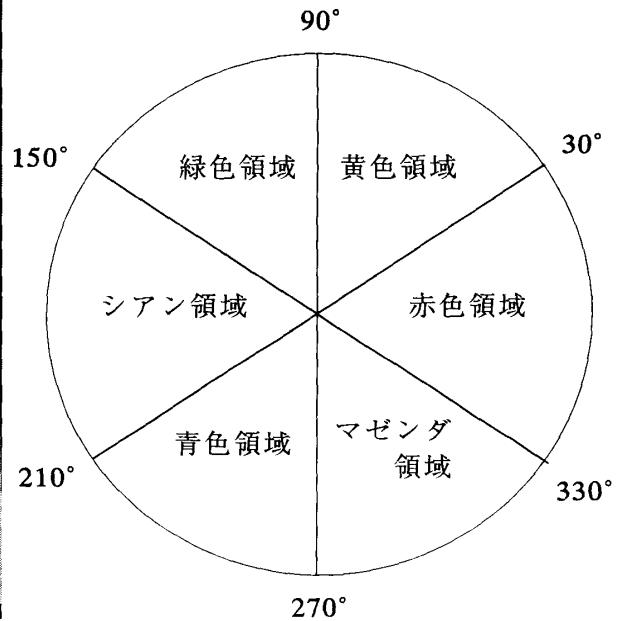


図9 色相分解

2つの画像に分解されないように隣接する色相領域に対して20°の重なりをもたした。さらに彩度値が25(%)以下の領域は無彩色領域として明度値により2値化し、その白黒反転画像も用いた。本研究では上述した計8枚の画像に入力画像を分解した。

#### 4.3 背文字候補領域の抽出

分解された全ての2値画像に対してラベリング処理を行い、求められたそれぞれの領域に対して外接矩形枠を求める。求められた矩形枠の大きさが $3 \times 3$ 画素より大きく、 $80 \times 80$ 画素より小さい領域を文字候補領域とする。図10に抽出された背文字候補領域の一例を示す。

#### 4.4 背文字領域の抽出

求められた背文字候補領域の内、定められたしきい値より大きい幅を有する外接矩形枠の個数に対して縦方向の周辺分布を求める。求められた周辺分布の値が定められたしきい値以上の幅を持っているならば、その範囲の始まりと終わりのX座標値により文字範囲を定める(図11参照)。定められた文字範囲に外接矩形枠が90(%)以上含まれる領域を全て文

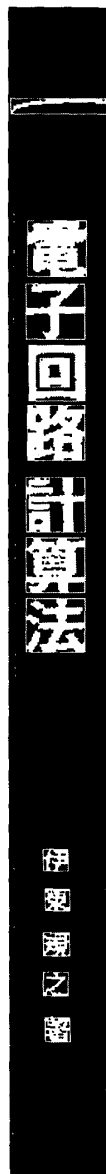


図10 分解画像（青色領域）の背文字候補領域の一例

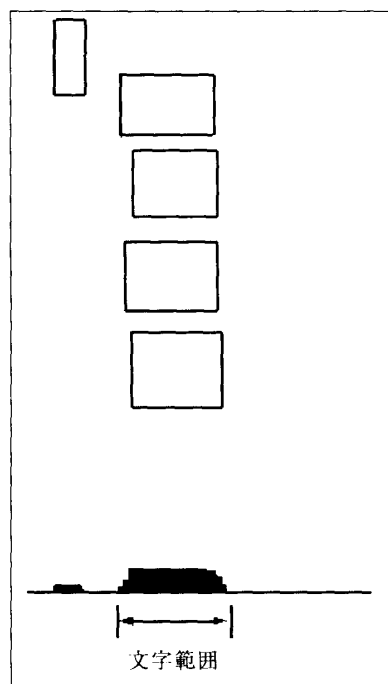


図11 文字範囲



図12 抽出された背文字領域の一例

字領域とし、これらの外接矩形枠の個数に対して次に横方向の周辺分布を求める。求められた周辺分布の値をもとに孤立した小領域を除き、それ以外の外接矩形枠の領域を背文字部分領域とする。最後に全分解画像から得られた背文字部分領域に対して論理和を求め、得られた各領域を最終的な背文字領域と決定する。図12に抽出された背文字領域の一例を示す。

## 5. 抽出実験

書棚の写真をイメージスキャナ GT6000 (EPSON) を用いて、パーソナルコンピュータ Power Macintosh 8100/80AV に入力した後、ワークステーション NWS-5000VI (SONY)

に画像を転送して処理を行った。使用言語はC言語である。

### 5.1 書籍領域の抽出実験

書籍領域の抽出実験結果の判定は目視で行った。背表紙内の全ての文字を含み、書籍のエッジと求められた1組の直線の平均の傾きが違和感のない程度( $1^{\circ}\sim 2^{\circ}$ )であれば領域抽出成功とした。ただし、エッジが隣の書籍の背文字領域にかかっていたり、領域が書籍より著しく大きく抽出されているものは抽出失敗とした。また、画面から背文字の一部でもはみ出している書籍は判定対象から除いた。抽出結果は23枚の書棚画像、書籍数124に対して、抽出成功数は116、抽出率は93.5(%)であった。主な抽出失敗原因は、黒っぽい書籍が隣り合っている場合、および写真撮影の際に書籍の側面が写り込み、濃度勾配が一定値に集中せず、Hough変換によるエッジ検出が正確に行われなかった場合に生じた。

### 5.2 背文字領域の抽出実験

背文字領域抽出実験結果の判定も目視で行った。本研究では用いた画像サイズより、抽出可能文字サイズを $30\times 30$ 画素(タイトルなどは十分に抽出可能なサイズ)以上とした。抽出結果は、抽出成功した書籍数116における文字数1227に対して、抽出成功数は1103であり、抽出率は89.9(%)であった。主な抽出失敗は、抽出対象文字の線幅が狭い場合にアフィン変換などによる画像のボケによって色相値にばらつきが現われ、背文字候補領域が不連続となる場合に生じた。また、彩度値が低い場合にも同様に色相値にばらつきが現われ抽出失敗が生じた。これらの問題に対する対策としては入力画像の解像度を高くすることが考えられる。また、画像分解の際に彩度値が低い場合に、動的しきい値決定法を用いて2値化を行う方法も検討したい。

## 6. む す び

本研究では書棚画像からの書籍の背文字領域の抽出に関する検討を行った。書籍領域の抽出に関しては、得られた実験結果より、黒っぽい書籍が隣り合った場合にエッジの抽出が困難であることがわかった。この点については、照明条件なども含めて検討が必要であると考えられる。背文字領域の抽出に関しては、解像度が不足していることが抽出失敗の原因である場合が多かった、今後書籍の領域抽出を行った後、求められた領域をもとにして、カメラのズーム機能を利用し、背文字領域の抽出を行う方法についても検討を行う予定である。さらに、抽出した背文字領域に対して、領域の統合、1文字ごとの切り出しを行い、タイトル、著者名、ロゴマーク等の認識も行いたい。

### 参考文献

- 1) 大竹善二郎, 長尾智晴, 安居院猛, 中嶋正之: “書棚画像からの書籍の背文字領域の抽出”, 日本印刷学



- 会誌, 第29巻第1号 (1992).
- 2) 松尾賢一, 上田勝彦, 梅田三千雄: “色彩情報を用いた情景画像からの文字列抽出”, 信学技報, PRU94-92 (1994).
- 3) 高木幹雄, 下田陽久: “画像解析ハンドブック”, 東京大学出版社 (1991).

## A Study of Extraction of Character Regions from Bookshelf Images

Yukihiro ISHIKAWA, Mitsuru OHKURA\*,  
Mitsuru SHIONO\* and Reiji HASHIMOTO\*\*

*Graduate School of Engineering,*

*\*Department of Information & Computer Engineering,*

*\*\*Department of Electronic Engineering,*

*Okayama University of Science,*

*Ridai-cho 1-1, Okayama 700, Japan*

(Received October 7, 1996)

In this paper, we examined the extraction of character regions from bookshelf images to assist in making a book database system. The extraction algorithm has two processes, 1) an extraction of the book images from the bookshelf image, 2) an extraction of character regions from the book images. In the first stage, the bookshelf image is divided into several sectors based on the direction of density gradient of the image pixels. Then, by applying the Hough translation, the straight lines composing the edge lines of the book images can be detected. In the second process, we use the peripheral distribution of the book image to detect the width and height of each character. Using 23 bookshelf images containing 124 books with 1227 characters, we obtained an extraction rate of book images of 93.5 (%) and an extraction rate of character regions of 89.9 (%).