

## 平積みされた雑誌カラー画像からの雑誌数の計数と異種雑誌の検出

大倉 充・松本健児\*・石井 良\*\*・濱本高志

塩野 充

岡山理科大学工学部情報工学科

\*四国情報管理センター

\*\*瀬戸内システム

(2000年11月1日 受理)

### 1. まえがき

近年、文字認識の研究は個別文字の認識から文書画像あるいは情景画像に認識対象を拡張し、画像理解の方面に進んでいる。文書画像あるいは情景画像を認識対象とした場合には、画像中の文字列の存在位置が不明となるために、処理の第一段階では、画像からの文字列抽出処理が最も重要な処理となる。情景画像からの文字列抽出および認識に関する研究では、これまでに自動車のナンバープレート<sup>1)</sup>、部屋番号<sup>2)</sup>、銀行等の看板文字<sup>3)</sup>、書籍の背文字<sup>4)</sup>等に関する研究例がある。

本研究では、情景画像の一つである平積みされた雑誌カラー画像を処理対象とし、雑誌の背文字列を基にした雑誌数の計数および雑誌中に含まれる異種雑誌の自動検出について基礎的な検討を行った。本研究の最終的な目的は、毎週、書店やコンビニエンスストアに配送され平積みされる雑誌をカメラで撮影し、雑誌の在庫管理を自動的に行うことにある。カラーの文書画像を対象とした研究では、基本的に文字列の単色性を仮定し、代表色（画像中に多く存在する“主要な色”）を求めることにより文字列の候補領域を抽出する場合が多い<sup>5,6)</sup>。本研究では、平積みされた雑誌の背文字列は単一色を有するとしているため、色情報に基づくクラスタリング等の手法は採用せず、代表色として最も画素頻度の高い色を求めることで背文字列領域の抽出を行っている。また本研究では、個別文字の切り出しおよび文字認識は行わずに、抽出された背文字列領域のマッチング処理のみで異種雑誌の判定処理を行っている。

### 2. 原画像の取得

少年誌を平積みして机の上に置き、雑誌に対して仰角約  $15^\circ$ 、距離  $0.7[\text{m}]$ 、雑誌の背文字列がフレームの中央付近に配置されるようにデジタルカメラ（DSC-700：SONY）を設置して室内で撮影を行った。コンピュータへ入力された画像は、画像サイズ縦  $480 \times$  横  $640$  画素、RGB 各  $256$  階調のカラー画像である。本研究では総雑誌数を  $8$  冊、異種雑誌数を  $1$  冊としており、背文字列は同種の色を有すること、および雑誌の背文字列の向きは同一という条件を設けている。表 1 に示す雑誌の積み上げ状態および雑誌の設置角度を設定し、総計  $28$  枚の原画像を取得した。表中、裏返しの雑誌は背文字列が見えない雑誌、逆向きの雑誌は背文字列が逆向きとなる雑誌のことである。また雑誌の設置角度は、平積みされた雑誌前面中心を原点とし、奥行き方向に左右  $15^\circ$ 、 $30^\circ$ 、 $45^\circ$  雑誌全体を回転させた状態を指す。図 1 に原画像の一例を濃淡画像で示す。これは、組番号 1 の設置角度を  $0^\circ$  とした場合の画像である。

表 1 平積みされた雑誌の状態

組番号	異種雑誌の位置	その他の条件	設置角度
1	2 冊目	—	7 種類
2	1 冊目	—	7 種類
3	2 冊目	裏返しの雑誌：4 冊目	7 種類
4	3 冊目	逆向きの雑誌：6 冊目	7 種類

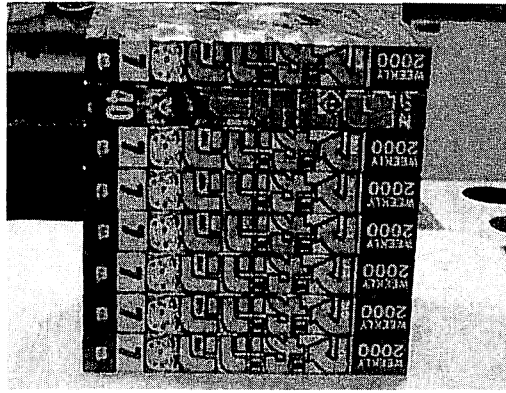


図 1 原画像の一例

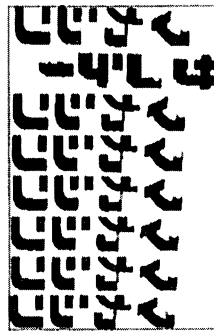


図 2 背文字列領域

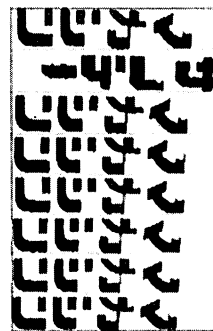


図 3 境界線分

### 3. 処理手順

#### 3.1 背文字列の抽出処理

##### 3.1.1 背文字列領域の確定処理

入力された原画像は RGB 各 256 階調のカラー画像であるが、処理の軽減のために各色 3 ビットに再量子化する。この結果、画像の総色数は 512 色となる。次に得られた画像の中央付近で最も画素頻度の高い RGB 値を有する画素のみを残した 2 値画像を作成する。雑誌が画像中央付近に配置されていることより、この処理によって背文字列の有する RGB 値と等しい RGB 値を有する画素のみが残されることになる。画像中央付近とは、画像の高さを *height*、幅を *width*、画像左上を座標原点とした場合、以下に示す座標で囲まれた矩形領域を指す。

$$\text{矩形領域の左上座標} : \left( \frac{width}{2} - 50, \frac{height}{2} - 180 \right)$$

$$\text{矩形領域の右上座標} : \left( \frac{width}{2} + 50, \frac{height}{2} - 180 \right)$$

$$\text{矩形領域の左下座標} : \left( \frac{width}{2} - 50, \frac{height}{2} + 180 \right)$$

$$\text{矩形領域の右下座標} : \left( \frac{width}{2} + 50, \frac{height}{2} + 180 \right)$$

次に、上述の処理により得られた 2 値画像に対して 2 種類の雑音除去処理を行う。まず膨張・収縮処理<sup>7)</sup>による微小領域の除去を行い、次にラベリング処理<sup>7)</sup>に基づき面積の小さな連結成分の除去を行う。本研究

では、連結成分の面積が 30 画素以下の領域を除去した。雑音除去処理後、画像中に存在する全ての連結成分を囲む外接矩形を求め、新たな画像を生成する。図 2 にこれまでの処理で得られた 2 値画像の一例を示す。

### 3.1.2 各背文字列の切り出し処理

3.1.1 項で得られた画像から各雑誌の背文字列を切り出す。切り出し処理には Hough 変換<sup>7)</sup>による直線検出アルゴリズムを適用した。本処理の目的は、背文字列間の空白領域内に各背文字列を分離するための境界線分を求めることにある。そのため、通常の Hough 変換では黒画素を注目画素として線分を求めるが、本研究では白画素を注目画素とし、さらに以下に示す条件を設定して線分を求めた。

- 雑誌が平積みされていることより、背文字列はある一定の角度方向に連続して形成されている。これより、抽出する線分も同様の角度に限定されることより、本研究では、抽出する線分の角度を水平方向を基準として $\pm 20^\circ$ の範囲に設定した。
- 本来、背文字列間に境界線となる線分が一本求められれば良いが、Hough 変換をそのまま適用すると多数の線分が得られる結果となる。そのため、線分間の距離を設定してこの問題に対処した。画像上端から下端に向かって最初に得られる線分と画像上端までの距離を求め、その距離から 5 画素引いた値を、本研究では線分間の距離とした。

これらの条件を基に以下の処理手順で線分を求めた。

- (1)  $\rho \cdot \theta$  空間への投票
- (2) 描画線分角度の決定
- (3) 線分の描画

図 3 に描画された境界線分を示す。この画像中の境界線分のみを残し、背景領域に対するラベリング処理を行うことにより各背文字列の領域を確定し、領域ごとに切り出す。なお、この時点で雑誌数の計数が行われる。

### 3.2 異種雑誌の検出処理

切り出された背文字列同士の比較を行うことにより、異種雑誌の検出を行う。まず、3.1.2 項で得られた画像に対して背文字列が存在するかどうかの判定処理を行い、背文字列が存在すると判定された場合には、背文字列領域の外接矩形を求め、余分な背景領域を除去した画像を作成する。次に画像サイズの統一を行う。上から一冊目に対応した画像を基準として、アフィン変換<sup>7)</sup>の拡大・縮小処理により他の画像のサイズを揃える。最後にマッチング処理を行う。マッチングの基準となる画像（以下、基準画像と呼ぶ）は、基本的には画像サイズの統一の際と同様に、上から一冊目を基準としてその他の画像（以下、対象画像と呼ぶ）との比較を行う。マッチング処理には画像間のユークリッド距離を採用した。ただし、ユークリッド距離は画像サイズに応じて変動するために判定時のしきい値設定が困難となることより、本研究では、求めたユークリッド距離 $\gamma$ を基にして次式により正規化を行い類似性の尺度 $D$ に変換して使用した。

$$D = \frac{R - \gamma}{R} \quad (1)$$

ここで $R$ は画像間のユークリッド距離の最大値を表わし、画像サイズを $M \times N$ 、白画素の値を 255、黒画素の値を 0 とした場合、次式で求められる。

$$R = 255 \times \sqrt{M \times N} \quad (2)$$

本研究では、次式で示される判別のためのしきい値 $T_1$ を設定し、求めた $D$ の値が $T_1$ よりも小さい値の場合に異種雑誌と判定した。

$$T_1 = ave - 0.075 \quad (3)$$

ここで $ave$ は $D$ の平均値を表わし、雑誌数を $n$ とすると次式で求められる。

$$ave = \frac{1}{n} \times \sum_{i=1}^n D_i \quad (4)$$

以下、裏返しの雑誌および逆向きの雑誌が存在する場合、基準画像が異種雑誌である場合の判定処理について述べる。

- 本処理では、上述したようにマッチング処理の前に 3.1.2 項で得られた画像中に背文字列が存在するかどうかの判定処理を行っている。そのため、裏返しの雑誌が存在する場合には、この画像中に背文字列が存在しないために、マッチング処理を行うことなしに異種雑誌と判定した。
- 基本的には  $D$  の値がしきい値  $T_1$  よりも小さい場合に異種雑誌と判定するが、判定前に、対象画像を  $180^\circ$  回転させて、再度、マッチング処理を行うことで、逆向きの雑誌が存在する場合に対処した。
- 上述したように、上から一冊目を基準としてマッチング処理を行い  $ave$  を求める。この値があらかじめ定めたしきい値  $T_2$  よりも小さい値の場合には、一冊目を異種雑誌と判定し、基準画像を上から 2 冊目の画像に変更して、再度、検出処理を行った。本研究では  $T_2=0.75$  とした。

#### 4. 検出実験

##### 4.1 実験結果

本研究で使用した 28 枚の画像では、27 枚の画像において異種雑誌の検出に成功した。検出率は 96(%)である。このことは、本研究で使用した画像データに関しては、しきい値の設定値が的確なものであったことを示していると考えられる。異種雑誌の検出例として、2 節で述べた組番号 1~4 に含まれる 4 種類の画像を図 4~7 に示し、各画像で得られた距離値を表 2~5 に示す。

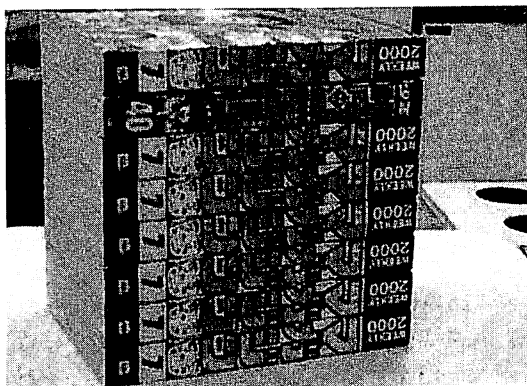


図 4 画像データ（組番号 1：設置角度右  $15^\circ$ ）

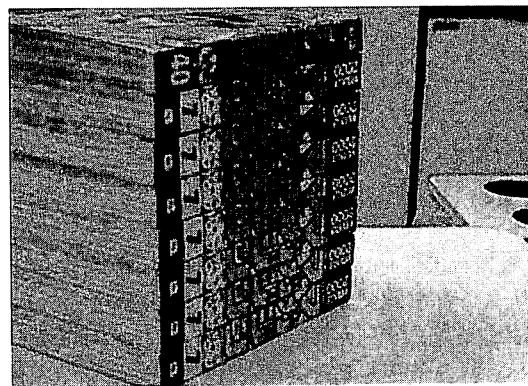


図 5 画像データ（組番号 2：設置角度右  $45^\circ$ ）

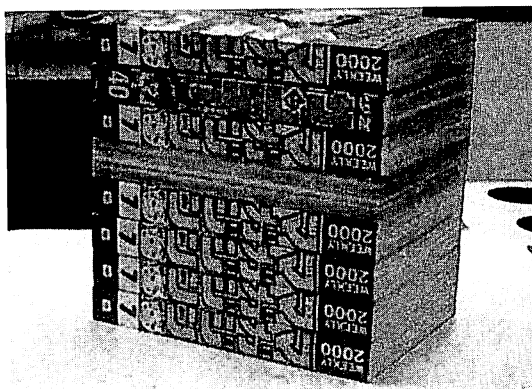


図 6 画像データ（組番号 3：設置角度左  $30^\circ$ ）

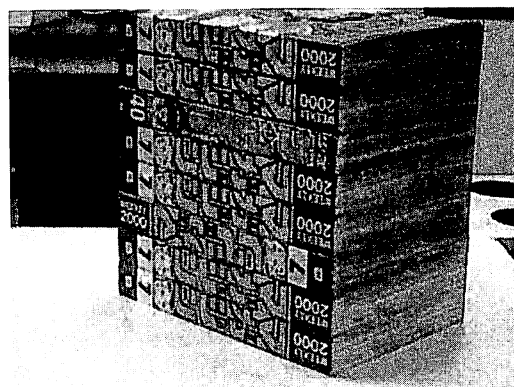


図 7 画像データ（組番号 4：設置角度左  $45^\circ$ ）

図 4 では、表 2 に示された 2 冊目の雑誌の距離値のみがしきい値  $T_1$  を下回っていることより 2 冊目の雑誌が異種雑誌と判定された。図 5 では、表 3 に示されるように基準雑誌を 1 冊目とした場合の  $ave$  がしきい値  $T_2$  を下回った結果、1 冊目が異種雑誌と判定され、更に基準雑誌を 2 冊目に変更して検出処理が行われている。その結果、得られた  $ave$  がしきい値  $T_2$  を上回り、各距離値もしきい値  $T_1$  を上回ったために、異種雑誌は 1 冊目のみという判定結果が得られた。図 6 では、表 4 に示されるように 2 冊目の雑誌の距離値のみがしきい値  $T_1$  を下回っていること、および 4 冊目に背文字列が存在しないと判定されたことより、2 冊目および 4 冊

目が異種雑誌と判定された。最後に、図 7 では、最初の判定処理では表 5 に示された 3 冊目および 6 冊目の距離値がしきい値  $T_1$  を下回っていたが、画像を  $180^\circ$  回転させて、再度、距離値を求めると 3 冊目のみがしきい値  $T_1$  を下回ったために、3 冊目のみが異種雑誌という判定結果が得られた

表 2 得られた距離値（組番号 1：設置角度右  $15^\circ$ ）

基準画像 No.	比較画像 No.	距離値
1	2	<b>0.645</b>
1	3	<b>0.906</b>
1	4	<b>0.893</b>
1	5	<b>0.880</b>
1	6	<b>0.857</b>
1	7	<b>0.837</b>
1	8	<b>0.820</b>
<i>ave</i>		<b>0.834</b>

表 3 得られた距離値（組番号 2：設置角度右  $45^\circ$ ）

基準画像 No.	比較画像 No.	距離値
1	2	<b>0.714</b>
1	3	<b>0.711</b>
1	4	<b>0.707</b>
1	5	<b>0.720</b>
1	6	<b>0.714</b>
1	7	<b>0.731</b>
1	8	<b>0.729</b>
<i>ave</i>		<b>0.718</b>
2	3	<b>0.942</b>
2	4	<b>0.926</b>
2	5	<b>0.894</b>
2	6	<b>0.880</b>
2	7	<b>0.855</b>
2	8	<b>0.844</b>
<i>ave</i>		<b>0.890</b>

表 4 得られた距離値（組番号 3：設置角度左  $30^\circ$ ）

基準画像 No.	比較画像 No.	距離値
1	2	<b>0.717</b>
1	3	<b>0.914</b>
1	4	—
1	5	<b>0.877</b>
1	6	<b>0.869</b>
1	7	<b>0.843</b>
1	8	<b>0.826</b>
平均		<b>0.841</b>

表 5 得られた距離値（組番号 4：設置角度左 45°）

基準画像 No.	比較画像 No.	距離値
1	2	0.922
1	3	0.737
1	4	0.898
1	5	0.856
1	6	0.729
1	7	0.826
1	8	0.811
ave		0.826
1	3	0.712
1	6	0.778

最後に，異種雑誌の検出に成功した雑誌数が最大の画像を図 8 に示す．上から 3 冊目，8 冊目および 11 冊目に異種雑誌があり，総雑誌数は 20 冊である．この画像では，背文字列の領域がかなり小さいために，3.1.1 項で述べた連結成分の面積が 30 画素以下の領域を雑音領域として除去処理を行うと，背文字列が雑音領域として除去されることになる．そのため，しきい値を 30 画素から 15 画素に変更し検出実験を行った（その他のしきい値は変更していない）結果，異種雑誌の検出を行うことができた．なお，雑誌を左右奥行き方向に 30° 回転させた場合にも同様の結果が得られた．

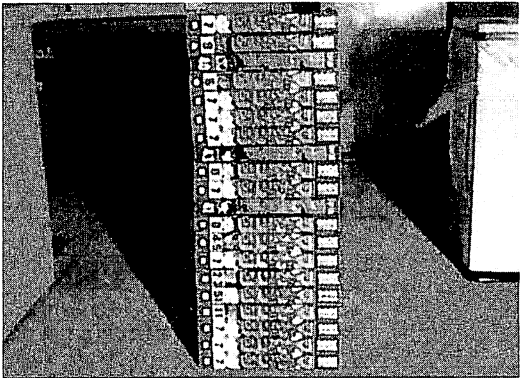


図 8 画像データ（総雑誌数 20 冊）

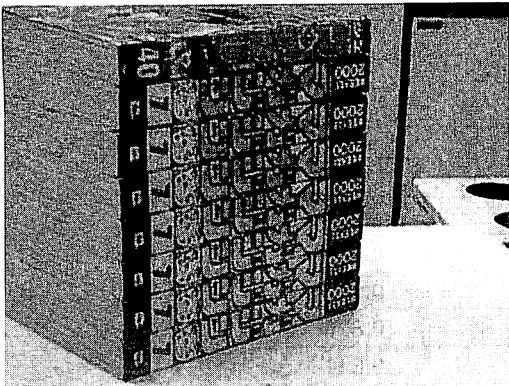


図 9 画像データ（組番号 2：設置角度右 30°）

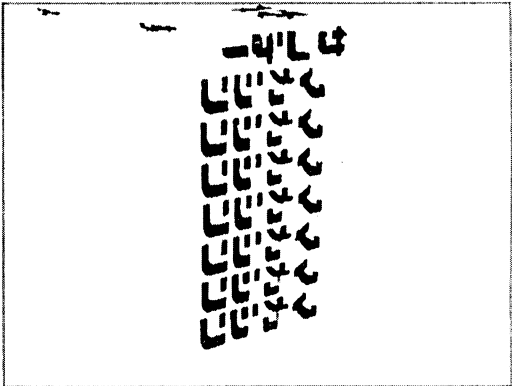


図 10 背文字列領域

#### 4.2 検出失敗に関する考察

本研究で使用した画像データ 28 枚中 1 枚の画像で異種雑誌の検出に失敗した。図 9 にその画像を示し、図 10 に背文字列領域を抽出した 2 値画像を示す。この画像は組番号 2 の右奥行き方向に 30° 雑誌を回転させた画像である。図 10 から背文字列領域以外に、背文字列と同色を有する雑誌の表紙の一部が雑音として残されていることがわかる。本研究では 3.1.1 項で述べたように、連結成分の面積が 30 画素以下の領域を雑音領域として除去処理を行っているが、この画像では、30 画素以上の雑音領域が存在する。この領域は、明らかに背文字列領域とは異なる形状および分布をしているが、本研究では、背文字列領域と雑音領域との区別処理を行っていないために、各背文字列領域の抽出に失敗することとなった。この点は今後の課題の一つとしたい。

#### 5. おわりに

本研究では、平積みされた雑誌カラー画像からの雑誌数の計数と異種雑誌の検出に関して基礎的な検討を行った。雑誌の積み上げ状態と雑誌の設置角度を変化させた 28 枚の画像データに対して良好な異種雑誌の検出結果を得た。しかし、背文字列の色情報を基に検出アルゴリズムを構築したために、最も上に位置する雑誌の表紙に同色を有する領域があり、その領域が撮影された場合には、背文字列領域の抽出処理を行うことができず、検出処理に失敗することが明らかとなった。また、雑誌の積み上げ状態が左右方向にずれた場合も同様の原因により背文字列領域の抽出処理を行うことができない。更に、本論中では述べていないが、最も上および最も下に裏返しの雑誌が存在した場合には、その雑誌が存在することが判定できずに雑誌数の計数に失敗することもある。今後は、これらの課題に対処できる頑健なアルゴリズムの構築を目指し、また雑誌数に応じたしきい値の設定法も考えたい。

#### 参考文献

- 1) 尾上博和, 塩野 充: “自動車画像からのナンバープレートの抽出とその漢字を含む全文字の切り出しと認識”, 信学論(D-II), Vol.J77-D-II, no.3, pp.483-492, 1994.
- 2) 恩田邦夫, 笹尾和仁, 青木由直: “廊下の情景画像からのドアと部屋番号の認識システム”, 信学技報 PRU88-145, 1989.
- 3) 松尾賢一, 上田勝彦, 梅田三千雄: “適応しきい値法を用いた情景画像からの看板文字列領域の抽出”, 信学論(D-II), Vol.J80-D-II, no.6, pp.1617-1626, 1997.
- 4) 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列領域の抽出”, 日本印刷学会誌, Vol.29, no.1, pp.38-45, 1992.
- 5) 仙田修司, 美濃導彦, 池田克夫: “文字列の単色性に着目したカラー画像からの文字パターン抽出法”, 信学技報, PRU94-29, 1994.
- 6) 長谷博行, 米田政明, 酒井 充, 丸山 博: “カラー文書画像中の文字領域抽出を目的とした色分割についての検討”, 信学論(D-II), Vol.J83-D-II, no.5, pp.1294-1304, 2000.
- 7) 長谷川純一, 興水大和, 中山 晶, 横井茂樹共著: “画像処理の基本技法”, 技術評論社, 1986.

## Detection of the Number and Cover Design Differences in Images of Color Magazine Piles

Mitsuru OHKURA, Kenji MATSUMOTO<sup>\*</sup>, Ryo ISHII<sup>\*\*</sup>,

Takashi HAMAMOTO and Mitsuru SHIONO

*Department of Information & Computer Engineering,*

*Faculty of Engineering,*

*Okayama University of Science,*

*Ridai-cho 1-1, Okayama 700-0005, Japan*

*\*Shikoku Information Management Center*

*\*\*Setouchi System*

(Received November 1, 2000)

We develop an algorithm to detect different magazines from images of color magazine piles. The algorithm is constructed using two procedures. One is the extraction process of character strings from each spine of the magazines and the other is a matching process using Euclidian distance as the discriminant function. The major color in the middle area of the image is chosen, and the pixels with this color are extracted as the character strings in the extraction process. The top character string is decided as a template pattern, and the distances between the template and the other character strings are measured to detect a different magazine in the matching process. We make an experiment using 28 images and obtain a 96% of detection rate.