

モデル当てはめによる書棚画像解析

平 英二[†] 高山 誠悟[†] 内田 誠一^{††} 迫江 博昭^{††}

Bookshelf Image Analysis Based on Model Fitting

Eiji TAIRA[†], Seigo TAKAYAMA[†], Seiichi UCHIDA^{††}, and Hiroaki SAKOE^{††}

あらまし 本論文では画像処理による書籍管理を目的として書棚画像から各書籍の境界を検出する手法を提案する。従来法ではエッジや影からハフ変換などの直線検出法を用いて書籍境界を検出している。本手法では、そのような局所的な情報だけでなく大域的な最適性も考慮して、書棚画像の最適領域分割（各書籍の背表紙領域、書棚背景領域）を動的計画法に基づくアルゴリズムにより行い、各書籍の境界を検出する。更に最適化問題として定式化する際、書棚画像の文法モデルを組み込むことで高精度化を図っている。実験により、本手法の有効性を定性的及び定量的に確認した。

キーワード 書棚画像解析, 書籍境界検出, FSA モデル, モデル当てはめ解析, 動的計画法

1. ま え が き

図書館や書店など大量の書籍が配架され、それらの出し入れが頻繁に繰り返される場所では、書架中の書籍の管理に多大の労力が必要となる。配架情報を電子化して省力化を図る検討もなされているが、計算機内の情報と実際の配架状況の整合性を保つためには依然として定期的に書棚を検査する必要がある。書棚検査を省力化する手法としてはバーコードを利用したデータ照合法 [1] が主流であるが、1冊ずつリーダを接触させて読み取る必要があるため、依然として多くの人手と時間を要する。最近では、RFID (radio frequency identification) 技術を導入することで非接触で書籍を管理するシステムの開発が進められているが、通信可能距離の法的規制や、金属遮へいや通信電波の干渉などによる検出精度の低下などの問題がある。RFID 技術と同様に非接触な書棚管理を図る手法として、書棚を撮影して得られた画像を解析することにより各書籍の背表紙に書かれた文字を抽出し [2] ~ [4]、タイトル

や巻号などを自動認識 [5] する技術が検討されている。画像を用いることにより、配架書籍のデータベース構築の支援といった、RFID 技術だけでは不可能な新たなサービスの提供が可能となる。

本研究では、これら書籍タイトルなどの自動認識を容易にすることを目標として、書棚画像中の各書籍の境界を検出する手法を検討する。本手法の特徴は以下の3点である。

(1) 画像中の各書籍の傾きが異なっても、それらの境界を検出できる。

(2) 動的計画法 (dynamic programming: DP) に基づく大域的最適化により、画像全体として最適な境界系列を検出する。

(3) 最適化の際に有限状態オートマトン (finite state automaton: FSA) により表現された書棚画像のモデルを組み込んで利用することで、高精度化を図る。

第1点は書籍の配架状況に柔軟に対処する上で必要であり、従来法においても考慮されている。本手法が従来法と異なるのは主に第2点及び第3点においてである。従来法では、エッジ [2] ~ [4], [6] や影 [5] から検出される直線を書籍境界の候補とし、平行な近隣直線間を書籍の背表紙領域としている。しかし、照明条件や書籍のデザイン、タイトル文字の影響を考えると、エッジのような局所的な情報のみ依存するよりも、画像全体としての最適性を考慮した方が安定した境界

[†]九州大学大学院システム情報科学府, 福岡市
Graduate School of Information Science and Electrical Engineering, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka-shi, 812-8581 Japan

^{††}九州大学大学院システム情報科学研究院, 福岡市
Faculty of Information Science and Electrical Engineering, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka-shi, 812-8581 Japan

検出ができると考えられる．更に書棚画像独特の性質を表すモデルを利用することで検出精度の向上が期待できる．

以下では，本手法の原理とアルゴリズムを述べ，実際に書棚画像から書籍境界の検出実験を行い，本手法の有効性を確認する．

2. 書籍境界検出法

2.1 定式化

本論文では，処理の対象とする書棚画像をサイズ M (横) \times N (縦) の HSI カラー画像 [7] とする．図 1 (a) に書棚画像の例 (I 成分のみ) を示す．このように，実際の書棚画像の中央部を切り出したものを処理対象とする．この切出しの理由は書籍境界検出において書籍高さのばらつきや棚板の影響をなくすためである．

この書棚画像中の各書籍の境界を検出するにあたり，各横座標値 i において，傾きをもった線分を考える (図 1 (b))．もし，各線分の傾きをその位置 i での書籍の傾きと合致するように正しく推定できれば，その線分系列の中に書籍境界が含まれていると考えられる．本手法ではこの考えを利用し，書籍境界検出問題を線分系列の最適推定問題の枠組みに埋め込んで扱うことにする．ここで位置 i の線分の傾きは上辺と交差する個所の横座標値 p_i で制御できる (図 1 (c))．

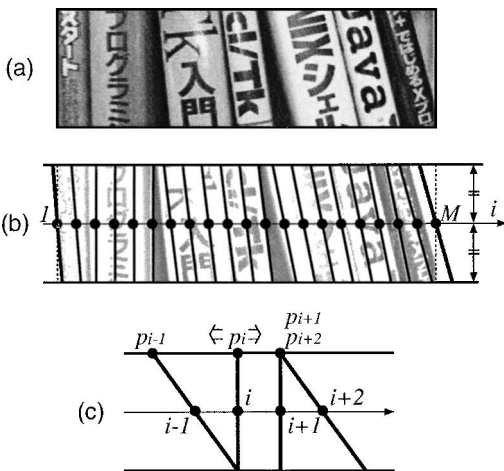


図 1 書棚画像 (I 成分) (a) と線分系列 (b) , 及び位置 i の線分の傾きを制御する変数 p_i (c)

Fig. 1 A bookshelf image (intensity image) (a), a sequence of line segments corresponding to slants of books (b), and the control variable p_i at position i (c).

よって以下では線分系列の傾きの最適推定問題を系列 $p_1, \dots, p_i, \dots, p_M$ の最適推定問題に置き換えて考える．

書籍境界を自動的に検出するためには以上の考えに加え， M 本の線分のうち，いずれが求めるべき書籍境界に当たるかを判定する仕組みが必要となる．また，精度向上のためには，1 冊の書籍の背表紙領域内部では傾き変化は起きないといった，書棚画像独特の性質を組み込むことも重要と考えられる．

そこで本手法においては，書棚画像の性質を FSA モデルとして表現し，それを上述の線分系列の最適推定問題に組み込んで利用することを考える．図 2 (a) のように書棚画像を水平方向に走査して観察すると，書籍境界，背表紙 (spine)，背文字，及び書棚背景が，ある規則性をもって繰り返し出現していることが分かる．この規則性は，図 2 (b) に示した 6 状態 ($Q = \{q_1, \dots, q_6\}$) とその間の遷移からなる FSA モデルで表現できる．なお，任意の状態が初期状態並びに最終状態になり得るものとする．この FSA モデルは以下のような書籍画像の性質に対応している．

- 左境界 (q_3) と右境界 (q_2) の間には背表紙領域 (q_4, q_5) がある．
- 背表紙領域 (q_4, q_5) の間には背文字領域 (q_6) がある．

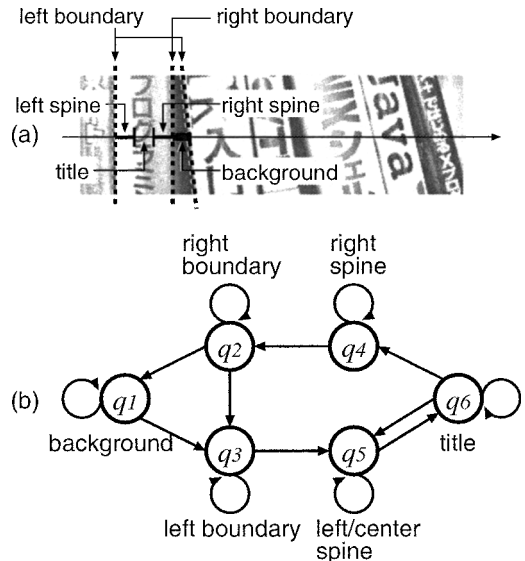


図 2 書籍画像の水平方向の変化 (a) とその FSA 表現 (b)

Fig. 2 State transition of a bookshelf image (a) and its FSA representation (b).

- ・一つの書籍中に背文字 (q_6) が背表紙領域 (q_5) を挟んで繰り返すことがある。
- ・書籍間、すなわち右境界 (q_2) と左境界 (q_3) の間には、書棚背景 (q_1) が入ることもある。

この FSA モデルを線分系列の最適推定問題に組み込むために、各線分をその傾きだけでなく状態すなわち属性も併せて考える。以下では位置 i の線分の状態を $s_i \in Q$ で規定し、各線分を $(p_i, s_i|i)$ で表現する。よって解くべき問題は $2M$ 個の変数 p_i, s_i ($i = 1, \dots, M$) の最適推定問題となる。最適解において $s_i \in \{q_2, q_3\}$ を満たす線分 $(p_i, s_i|i)$ が求めるべき書籍境界となる。

この最適推定問題において、FSA モデルは制約条件として働く。例えば線分 $(p_i, s_i|i)$ において s_i が q_4 (右背表紙部) ならば、直後の線分 $(p_{i+1}, s_{i+1}|i+1)$ における s_{i+1} は q_4 (右背表紙部) 若しくは q_2 (右境界) に制限される。更に、傾き変化、すなわち p_i と p_{i-1} の関係についても状態に応じた制約を課する。具体的には、各書籍の背表紙領域内部に対応する状態 q_4, q_5, q_6 では傾き変化を許さないように p_i, p_{i-1} を制御すれば、精度の向上を期待できる。

以上をまとめると、境界検出問題は次のように定式化される。

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^M f(p_i, s_i | i) \\ \text{w.r.t.} \quad & p_i, s_i \quad (i = 1, 2, \dots, M) \end{aligned} \quad (1)$$

ここで $f(p_i, s_i|i)$ は、位置 i の線分について p_i 及び s_i の値が妥当であるほど小さな値を返す関数である。関数 $f(p_i, s_i|i)$ の詳細は 2.3 で述べる。目的関数 (1) の最小化は次の三つの制約条件のもとで行う。第 1 の条件は傾きの許容範囲条件

$$|i - p_i| \leq W \quad (2)$$

である。ここで W は正の定数である。本論文では $W = N/2$ (垂直方向を基準に $\pm 45^\circ$ の範囲内) とし、それ以上傾いている書籍は対象外とする。第 2 の条件は前述のように s_1, \dots, s_M が図 2 (b) の FSA モデルに従うという条件である。第 3 の条件は状態ごとに傾き変化の許容範囲を定める次の条件である。

$$p_i = \begin{cases} p_{i-1} + \{0, 1, 2\} & \text{if } s_i \in \{q_1, q_2, q_3\} \\ & \text{and } s_{i-1} = s_i \\ p_{i-1} + 1 & \text{otherwise} \end{cases} \quad (3)$$

すなわち、隣り合う線分の傾きが変化するのは、両線

分が同一状態にあり、かつ、それが書棚背景 (q_1) 及び書籍境界 (q_2, q_3) のとき起こり得るものとする。更に傾きが変化する際は p_i と p_{i-1} の間隔はたかだか 2 とし、線分系列中の線分同士が両端以外では交差することがないように制御している。

2.2 DP アルゴリズム

線分を規定する横座標値 i, p_i をそれぞれ横軸、縦軸とする平面を各状態 q_1, \dots, q_6 について用意すると、 $(p_i, s_i|i)$ の系列は、図 3 のようにこの平面上で一つの経路を成す。この経路上の 1 点が 1 本の線分に対応している。またこの経路の遷移は、状態間の遷移については図 2 の FSA モデル、状態内の遷移については条件 (3) に従うことになる。図 4 はこれらの遷移規則をまとめたもので、実線が状態内遷移、破線が状態間遷移を表している。よって解くべき問題は、図 4 の遷移規則に従う経路のうち目的関数 (1) を最小にするものを探索する、いわゆる最適経路問題に帰着する。

よく知られているように、現在の状態に依存して次の状態が定まるような最適経路問題は DP により効率的に解くことができる。本手法の具体的な DP アルゴリズムを図 5 に示す。FSA による制御は連続音声認識に例があり [8]、それを参考にして解析アルゴリズムを作成した。本アルゴリズムでは次のようなワークエリアを用いる。

- ・ $g(p_i, s_i|i)$: 座標 $(p_i, s_i|i)$ に至るまでの最小累積コストを格納する。

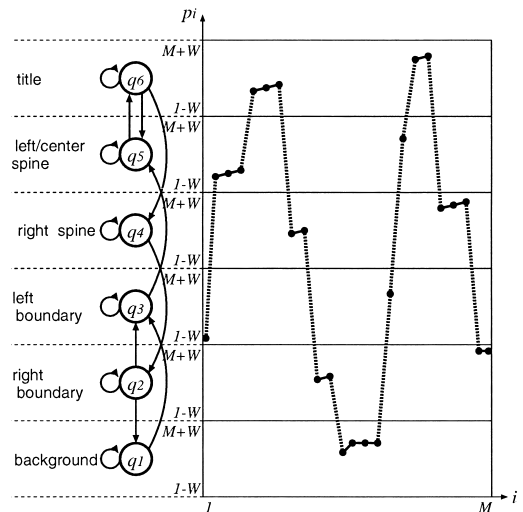


図 3 DP 計算空間
Fig. 3 DP calculation space.

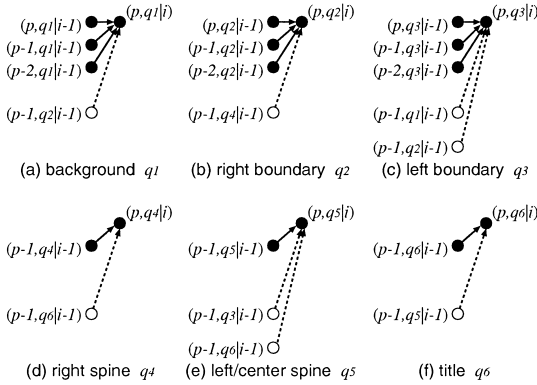


図 4 状態別の遷移パス

Fig. 4 Transition paths according to states of FSA model.

```

/* Initialization: */
1 for all  $s_1 \in Q$  do
2   for  $p_1 := 1 - W$  to  $1 + W$  do
3      $g(p_1, s_1|1) := f(p_1, s_1|1)$ 
/* DP Recursion: */
4 for  $i := 2$  to  $M$  do
5   for all  $s_i \in Q$  do
6     for  $p_i := i - W$  to  $i + W$  do begin
7        $g(p_i, s_i|i) := f(p_i, s_i|i)$ 
          +  $\min_{p_{i-1}, s_{i-1}} g(p_{i-1}, s_{i-1}|i-1)$ 
8        $bp(p_i, s_i|i) := \operatorname{argmin}_{p_{i-1}, s_{i-1}} g(p_{i-1}, s_{i-1}|i-1)$ 
9     end
/* Backtracking: */
10  $(p_M^{\text{opt}}, s_M^{\text{opt}}|M) := \operatorname{argmin}_{\substack{p \in [M-W, M+W] \\ s \in Q}} g(p, s|M)$ 
11 for  $i := M$  downto 2 do
12  $(p_{i-1}^{\text{opt}}, s_{i-1}^{\text{opt}}|i-1) := bp(p_i^{\text{opt}}, s_i^{\text{opt}}|i)$ 

```

図 5 DP アルゴリズム

Fig. 5 DP algorithm.

- $bp(p_i, s_i|i)$: 上記 $g(p_i, s_i|i)$ を与える経路が座標 $(p_i, s_i|i)$ に至る前に経由した直前の座標を格納する.

ここで図 5 の Step 7 の最小値選択は図 4 の遷移規則に従う p_{i-1}, s_{i-1} の組合せについて行い、その際に最小値を与える組合せを Step 8 で $bp(p_i, s_i|i)$ に格納する. 例えば, $s_i = q_3$ の場合には図 4 (c) より分か

るように 5 通りの組合せが存在し, $s_i = q_6$ の場合には同図 (f) より 2 通りの組合せが存在する. p_i, s_i の最適値をそれぞれ $p_i^{\text{opt}}, s_i^{\text{opt}}$ で表すと, $p_i^{\text{opt}}, s_i^{\text{opt}}$ は Step 10–12 のバックトラック処理により求められる. 前述のように, このうち $s_i^{\text{opt}} \in \{q_2, q_3\}$ を満たす線分 $(p_i^{\text{opt}}, s_i^{\text{opt}}|i)$ が求めるべき書籍境界となる.

2.3 線分評価関数の定義

関数 $f(p_i, s_i|i)$ は位置 i の線分について p_i 及び s_i の値が妥当であるかを評価する関数で, 妥当性が高いほど小さな値を返すものとする. この関数の設計のために, 書棚画像を特にエッジ特徴に着目して観察した結果, 次のことが分かった.

- 書籍境界では縦方向のエッジが長く直線的に検出されやすい.
- 横方向のエッジは背文字や背表紙のデザインによるものが多い.
- 書棚背景や背表紙ではエッジが検出されにくい.

そこで, 書籍境界部の評価値 $f(p_i, q_2|i), f(p_i, q_3|i)$ については線分上に縦方向 (垂直方向を基準に $\pm 45^\circ$ の範囲内) のエッジが多いほど小さくなるよう設計する. ただし, エッジ方向が線分の傾きと大きく異なるエッジは無視する. また, 連続するエッジは書籍境界としての信頼性が高いとして, その連続数に応じた重みを付加する. 背文字部の評価値 $f(p_i, q_6|i)$ については線分上に横方向 (水平方向を基準に $\pm 45^\circ$ の範囲内) のエッジが多いほど小さくなるよう設計する. 書棚背景部の評価値 $f(p_i, q_1|i)$ については線分上にエッジが少ないほど小さくなるよう設計する. 類似して背表紙部の評価値 $f(p_i, q_4|i), f(p_i, q_5|i)$ についても線分上にエッジが少ないほど小さくなるよう設計するが, 表紙デザインの影響を低減するために横方向のエッジは無視する. 本研究で設計した各評価値の詳細は付録に示す.

なお, 位置 i の線分と位置 $(i-1)$ の線分とで傾きが異なる場合, すなわち $p_i \neq p_{i-1} + 1$ のときは線分間にただか 1 画素幅のすき間が生じる (例えば図 1 (c) の p_i と p_{i-1} の場合). そのすき間についても線分上と同様に評価する必要がある. 具体的には, その隙間を埋める線分について状態 s_i としての妥当性を評価し, 評価値 $f(p_i, s_i|i)$ に加える^(注1).

(注1): よって関数 f は正確には $f(p_i, p_{i-1}, s_i|i)$ と表記すべきであるが, 簡素化のため文中では $f(p_i, s_i|i)$ と表記している.

3. 実 験

3.1 前 処 理

本実験ではデジタルカメラで撮影した 640×480 の書棚画像 60 枚を使用する．図 6(a) に書棚画像の例 (I 成分のみ) を示す．各画像には以下の前処理を施す．

(1) 画像の切出し

各書棚画像 (640×480) に対して, 240 行目を中心とする 640×100 の領域を固定的に切り出す (図 6(b)). 図 5 の境界検出はこの切出し画像の $[1+W, 640-W]$ の範囲を対象とする．ここで $W = N/2 = 50$ であり, したがって $M = 540$ として図 5 の処理を行う．切出しの目的は前述のように書棚画像中の書籍の高さのばらつきや棚板の影響をなくするためである．また計算時間の低減効果もある．

(2) エッジ検出

評価関数 $f(p_i, s_i|i)$ の計算に備え, 切出し後の画像に対して次の手順でエッジ画像を生成する．まず画像の HSI 各プレーンそれぞれについて 5×5 の Canny フィルタ [9] を用いてエッジ強度及び方向を求める．次に各プレーンそれぞれのエッジ強度は固定しきい値で 2 値化する．このしきい値は書籍境界付近のエッジなるべく欠落しないようにすることを基本方針として実験的に定めた．なお, 本実験では全プレーンにおいて同一のしきい値を用いた．この後, 各プレーンのエッジ画像それぞれから縦方向エッジ画像と横方向エッジ画像を作成する．ここで縦方向エッジ画像とはエッジ方向が垂直方向を基準に $\pm 45^\circ$ の範囲内であるエッジのみを集めた画像であり, 横方向エッジ画像とはそれ以外のエッジを集めた画像である．更に, 各プレーンの縦方向エッジ画像については連結成分のラベリングを行い, 連続画素数を各エッジ画素に与えておく．ここで連続画素数とは各エッジが属する 8-連結成分の面積を表す．最後に, 3 プレーンの横方向エッジ画像については重ね合わせて統合する．一方, 3 プレーンの縦方向エッジ画像についても重ね合わせて統合するが, 統合の際に複数のプレーンでエッジが重複する場合, 最大の連続画素数を与えるプレーンのエッジ方向を統合後の方向とする．これは連続画素数が多いほどエッジ方向の信頼性は高いと考えられるためである．図 6(c), (d) に縦方向エッジ画像及び横方向エッジ画像の例を示す．

3.2 従 来 法

本実験ではハフ変換を用いた境界検出法 [6] と本手

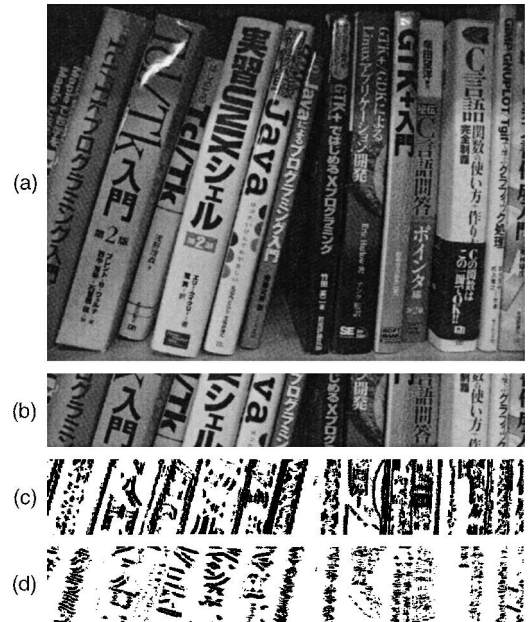


図 6 書棚画像 (a), 切出し処理後の画像 (b), 縦方向エッジ画像 (c), 横方向エッジ画像 (d) の例

Fig. 6 (a) An example of bookshelf images. (b) The middle part of (a). (c) The vertical edge image of (b). (d) The horizontal edge image of (b).

法との比較を行う．前者の実装に際し, できるだけ忠実に再現するようにしたが, 文献 [6] において詳細が不明な部分については極力性能が向上するように配慮した．こうして実装した従来法のアルゴリズムの概略を以下に示す．

- HSI^(注2)カラー画像からエッジを検出し, 各プレーンにおけるエッジ強度を各画素ごとに比較し, 最大強度をその画素のエッジ強度とする．その後, 3.1 で用いた 2 値化しきい値で 2 値化する．更に横方向のエッジを除去した後, 連結成分のラベリングを行い, 連続画素数を各エッジ画素に与えておく．
- 前ステップで作成したエッジ画像に対してハフ変換を行う．ただし, 長く連続したエッジは書籍境界としての信頼性が高いとして, ハフ平面へ投票する際, 連続画素数に応じて 1~4 の重みを付加する．

(注2): 参考とした広井らの手法 [6] では HSY 表色系が用いられている．本研究においてその HSY 表色系の利用を検討したが, その場合の結果は HSI 表色系を用いた場合に比べ劣位であった．

- ・ハフ平面上で投票数の局所的極大点を検出して書籍境界の候補とする．ただし各極大点について，自身の投票数を上回る極大点が近傍^(注3)に存在する場合，不要な境界候補とみなし除去する．
- ・しきい値以上の投票数を有する極大点についてハフ逆変換を行い，書籍境界として検出する．そのしきい値は，前ステップで検出された極大点を投票数の降順で並べたとき，次位との差が最も大きくなる極大点の投票数とする．

以下，上記手法を従来法 A と呼ぶ．また投票数しきい値の決定法として上記最終ステップの適応的決定処理を行わず，単にすべての画像について一定のしきい値を採用する場合を従来法 B と呼ぶ．なお，3.1 の切出し処理を従来法で行うとエッジの連続性による重み付けの効果が薄くなり，背文字から検出されるエッジの影響により結果が悪くなったため，従来法では切出し処理を行わなかった．

3.3 結果と考察

3.1 の手順で前処理された書棚画像 60 枚について書籍境界検出実験を行った．本手法の画像 1 枚当りの計算時間は，前処理と DP アルゴリズムを合わせた全体で約 3.0 秒 (Xeon 1.7 GHz) であった．

図 7 に境界検出結果の例を示す．図 7 の左から順に原画像，従来法 A による結果，本手法による結果である．検出された書籍境界は黒線で示してある．また本手法による結果は，切出し画像に対して検出された境界を上下に延長して示してあり，横じまの入った領域は書棚背景として推定されたことを示している．なお，本手法ではある書籍の左境界 (若しくは右境界) として連続した複数の直線が検出されることがあるが，本実験ではそれらの直線を集約して 1 本の境界とみなす．すなわち，図 7 の本手法による結果に見られるような帯状の直線はそれ全体で 1 本の境界を表すものとする．

図 7(a), (b), (c) を見ると，本手法により画像中のほとんどの書籍の境界が適切に検出できているのが分かる．特に同図 (b) の矢印が示す境界は境界付近のエッジが弱く，従来法 A では検出できていないのに対し，本手法では局所的かつ網羅的に状態別の評価を行うため，エッジが弱くとも検出できており，本手法の境界検出能力が従来法よりも優れていることが確認できる．また同図 (c) の矢印が示す境界は，背文字の存在及びその傾きなど，境界を表すエッジ以外の情報を考慮したことにより検出できた結果であり，FSA モデ

ルと大域的最適化を組み合わせるといふ本手法の考え方が境界検出の精度向上に有益であることを示唆している．

一方，図 7(d) のように本手法により求めるべき境界が欠落したり，誤って境界が検出される場合もいくつか見られた．前者の多くは表紙色の似た書籍が隣り合うことにより境界付近でエッジが検出できなかったことに起因した．後者は主に不要なエッジの発生によるものであった．特に背文字や照明の映り込みによる光沢の右端若しくは左端部から検出されたエッジが境界の誤検出の原因となっていた．また，傾きが異なる書籍の間に表紙の一部が映っているとき，その表紙部と書棚背景の境目や表紙イラストから検出されたエッジにより不要な境界が検出される場合も見られた．

次に書棚画像 60 枚について，欠落した書籍境界 (false negative) の本数と書籍境界として誤検出された直線 (false positive) の本数を測定した．境界の欠落及び誤検出の判定は以下の 3 条件に従って目視により行った．

[条件 1] 書籍境界が存在すべき個所がない (false negative)

[条件 2] 境界とは大きく異なる (およそ 10 画素以上) 位置にある (false positive)

[条件 3] 境界付近 (数画素程度) にあるが，背表紙のタイトル文字と交差するほど傾き及び位置が不適切である (false positive)

前述したように本手法では帯状の直線が検出されるが，その一部が条件 2, 3 を満たす場合は誤検出とみなす．図 8 は false negative と false positive の本数の関係を示したグラフである．図 8 中の実線は従来法 B において投票数しきい値を変動させた場合の結果である．しきい値を大きくするほど誤検出は減るが欠落が増え，逆にしきい値を小さくすれば欠落は減るが誤検出が増える．また “ ” は従来法 A による結果で，“ ” は本手法による結果である．本手法は境界の欠落が非常に少なく，全境界の 96.8% を検出できた．従来法 A において欠落した境界の本数は本手法の約 4 倍，従来法 B に至っては 300 本の誤検出を許してもいまだ 2 倍程度あり，本手法の境界検出能力が従来法よりも優れていることが確認できる．一方，境界の誤検出は本手法の方がわずかに多く (約 1.3 倍)，過剰に分割してしまう

(注 3): 近傍とみなす領域のサイズは ρ 方向及び θ 方向それぞれについて実験的に最適化した．



図 7 書籍境界検出例：原画像（1成分のみ）（左），従来法 A による検出結果（中央），本手法による検出結果（右）．検出結果において黒線は書籍境界，横じまは書棚背景を示す．

Fig. 7 Examples of book boundary detection. For each image, the original image (left), the result by the conventional method (middle), and the result by the proposed method (right) are shown. Detected book boundaries are shown as vertical lines, and background areas are shown as horizontal hatchings.

傾向が見られた．境界の誤検出の更なる低減については今後の課題である．

最後に誤検出による境界を間に含むことなく左右境界がともに適切に検出できた書籍の数を測定した．ただし，背表紙全体が画像内に収まっていない書籍は無視するものとする．書籍の検出率を表 1 に示す．この表における「no slant」「uniform slant」「nonuniform

slant」は書籍の並び方の複雑さに対応し，それぞれ「すべての書籍がすき間なく直立している」（例えば図 7(a) のような状態），「すべての書籍がすき間なく一定の角度で傾いている」（同図 (b)），「各書籍がそれぞれ異なる角度で傾いている」（同図 (c)），といった場合の結果を示している．表 1 より，本手法では原画像中の書籍の並び方が複雑になっても検出率に大きな

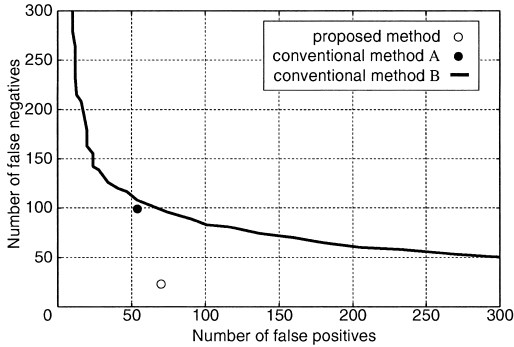


図 8 欠落した境界と誤検出された境界の本数

Fig. 8 Number of false positives and false negatives.

表 1 書籍の検出率 (%)
Table 1 Book detection rate (%).

image(#books/#images)	detection rate
no slant (231/20)	90.0
uniform slant (197/20)	88.8
nonuniform slant (182/20)	91.2
total (610/60)	90.0

影響はないことが確認できる．検出を誤った主たる原因は境界検出の失敗によるものであり，したがって前述のように背文字や光沢から検出されたエッジによる境界誤認と表紙色の似た書籍が隣り合うことによる境界エッジの欠落であった．

4. む す び

本論文では書棚画像中の書籍のタイトル文字認識を支援することを目標として，書籍境界検出法を提案した．本手法では，動的計画法に基づいたアルゴリズムにより画像全体としての最適性を考慮した境界検出を行っている．更に最適化の際に書棚画像の FSA モデルを利用することで高精度化を図っている．デジタルカメラで撮影した書棚画像に対して，96.8%の書籍境界検出率及び 90.0%の書籍検出率が得られた．従来法との比較も行い，本手法の有効性を確認した．

今後は実験で確認された問題点（境界の誤検出と欠落）を検討し，手法の改善を図る．具体的な改善策としては，(1) 事後処理として近隣境界間の位置関係を調べて不要な境界を除去する，(2) 目的関数の改良として色情報などを用いた新しい特徴を追加する，(3) 書棚画像の詳細な性質をモデルに反映させる，ことが挙げられる．特に色情報を用いる場合は (4) 局所的評価を線分単位から領域単位に拡張した方が効果がある

と予想される．以上に加え，次の段階となる背文字の抽出やその認識の検討も行っていく予定である．

謝辞 本研究の一部は文部科学省科学研究費補助金 (No.14780293) によった．

文 献

- [1] 喜多智慧夫，“小規模図書館における総合管理の機械化—短大・ニュー図書 II の場合”；情報の科学と技術，vol.41, no.7, pp.588–595, July 1991.
- [2] 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，vol.29, no.1, pp.38–45, Jan. 1992.
- [3] 石川幸博，大倉 充，塩野 充，橋本禮治，“書棚画像からの書籍の背文字領域の抽出について”；岡山理科大学紀要，vol.32A, pp.163–171, March 1997.
- [4] Y. Akiyama and M. Ito, “Book recognition from color images of book shelves,” Proc. IAPR Workshop on MVA'98, pp.106–110, Chiba, Japan, Nov. 1998.
- [5] 澤木美奈子，村瀬 洋，萩田紀博，“劣化推定に基づいた辞書の自動選択による本棚画像中の文字認識”；映情学誌，vol.54, no.6, pp.881–886, June 2000.
- [6] 広井敬輔，藤山雅晶，三浦竜哉，伊藤 稔，“書棚画像からの書籍境界の抽出”；画像の認識・理解シンポジウム講演論文集，vol.1, pp.271–276, July 2000.
- [7] 画像処理標準テキストブック編集委員会監修，画像処理標準テキストブック，画像情報教育振興協会，東京，1997.
- [8] 迫江博昭，藤井浩美，吉田和永，巨理誠夫，“フレーム同期化，ビームサーチ，ベクトル量子化の統合による DP マッチングの高速化”；信学論 (D)，vol.J71-D, no.9, pp.1650–1659, Sept. 1988.
- [9] 徐 剛，辻 三郎，3次元ビジョン，共立出版，東京，1998.

付 録

ここでは本研究で実際に設計した線分評価関数について記す．いずれの評価関数の値域も $[0, 1]$ である．

はじめに書籍境界部の評価値 $f(p_i, q_2|i)$, $f(p_i, q_3|i)$ について説明する．線分上の縦方向エッジについて，連続している部分ごとにカウントし，そのカウント値を x_1, x_2, \dots, x_K で表すものとする．線分上のすべての縦方向エッジが連続している場合は $K = 1$ となる．ただし，線分の傾きとエッジ方向のずれがしきい値以上となる縦方向エッジは無視する．そのしきい値は 3.1 で得られるエッジ画像に対して実験的に定め， 9° とした．こうして得られた x_1, x_2, \dots, x_K を用いて，書籍境界部の評価値を次式で与えた．

$$\begin{aligned}
 f(p_i, q_2|i) &= f(p_i, q_3|i) \\
 &= 1 - \frac{1}{N \log N} \sum_{k=1}^K x_k \log x_k
 \end{aligned} \tag{A-1}$$

ただし，線分上にエッジが存在しなければ評価値は 1

とする．ここで $\log x_k$ はエッジの連続数に応じた重みである．大きな x_k , すなわち長く連続したエッジが存在する線分ほど評価値は小さくなる．

背文字部の評価値 $f(p_i, q_6|i)$ は次式で与えた．

$$f(p_i, q_6|i) = \max \left\{ 0, 1 - \frac{2y}{N} \right\} \quad (\text{A}\cdot 2)$$

ここで y は線分上の横方向エッジの個数とする．横方向エッジが縦方向に長く連続することはなく，線分上の半数程度の画素が横方向エッジであれば，その線分は背文字部であると予想できる．

書棚背景部の評価値 $f(p_i, q_1|i)$ は，書籍境界部と同様の評価法を縦方向エッジの代わりに非エッジ画素を対象として行うことで与えた．また，背表紙部の評価値 $f(p_i, q_4|i)$, $f(p_i, q_5|i)$ は書籍境界部と同様の評価法を縦方向エッジの代わりに非エッジ画素及び横方向エッジを対象として行うことで与えた．

(平成 15 年 3 月 5 日受付, 7 月 14 日再受付)



迫江 博昭 (正員:フェロー)

昭 41 九工大・制御卒．昭 43 同大学院修士課程(電子)了．同年日本電気(株)入社．平元九大に移り，現在，システム情報科学研究院知能システム学部門教授．工博．音声認識及び画像パターン解析の研究に従事．IEEE ASSP Senior Award, 本会業績賞, 本会米沢ファウンダーズメダル受賞記念特別賞及び論文賞(社)日本発明協会・内閣総理大臣発明賞, 九工大嘉村記念賞各受賞．IEEE, 日本音響学会, 日本情報処理学会, 映像情報メディア学会各会員．



平 英二 (学生員)

平 12 九大・工・電気情報卒．平 14 同大学院修士課程了．現在，同大学院博士後期課程在学中．パターン認識及び最適化手法に関する研究に従事．



高山 誠悟 (正員)

平 13 九大・工・電気情報卒．平 15 同大学院修士課程了．同年松下電器産業(株)入社．在学中，画像処理・理解に関する研究に従事．



内田 誠一 (正員)

平 2 九大・工・電子卒．平 4 同大学院修士課程(情報)了．セコム(株)勤務を経て，平 11 同大学院システム情報科学研究科博士後期課程了．現在，同大学院システム情報科学研究院知能システム学部門助教授．工博．画像パターンの解析・認識，及び音声の認識・合成に関する研究に従事．IEEE, 日本音響学会, 日本情報処理学会, 映像情報メディア学会各会員．