社団法人　電子情報通信学会
THE INSTITUTE OF ELECTRONICS,
INFORMATION AND COMMUNICATION ENGINEERS

信学技報
TECHNICAL REPORT OF IEICE.

動的計画法による書棚画像からの書籍境界検出

平　英二† 高山　誠悟† 内田　誠一†† 迫江　博昭††

† 九州大学大学院システム情報科学府
†† 九州大学大学院システム情報科学研究院
〒 812–8581 福岡市東区箱崎 6–10–1
E-mail: †{taira,takayama,uchida,sakoe}@human.is.kyushu-u.ac.jp

あらまし　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により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全体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，書棚背景領域等) を文法制御を組み込んだ動的計画法に基づくアルゴリズムにより行う．実験により本手法の有効性を確認する．
キーワード　書籍境界検出，書棚画像， 動的計画法

Book Boundary Detection in Bookshelf Image Using Dynamic Programming

Eiji TAIRA†, Seigo TAKAYAMA†, Seiichi UCHIDA††, and Hiroaki SAKOE††

† Graduate School of Information Science and Electrical Engineering, Kyushu University
†† Faculty of Information Science and Electrical Engineering, Kyushu University
6–10–1 Hakozaki, Higashi–ku, Fukuoka–shi, 812-8581 Japan
E-mail: †{taira,takayama,uchida,sakoe}@human.is.kyushu-u.ac.jp

Abstract　For the automatic management of book arrangement in bookshelves, several systems to extract book information, such as title, from bookshelf images have been studied. In this paper, we propose a technique for detecting boundaries of the books in the bookshelf image. In the present technique, the book boundary detection are formulated as an optimal slant estimation problem under a FSA model of the bookshelf images. The globally optimal solution is searched for by a dynamic programming-based algorithm. The effectiveness of the present technique was shown by experiments.
Key words　book boundary detection, bookshelf image, dynamic programming

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を必要とする．配架情報を電子化し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．こうした現状に対して管理の自動化を図る手法として，書棚に配架された各書籍のタイトルや巻号などの自動認識技術が検討されている [1]– [3]．

本研究では，これら書籍タイトル等の自動認識を容易と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いても，それらの境界を検出できる．

(2)　動的計画法 (DP) に基づくアルゴリズムにより，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 (FSA) 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．

第一の特徴は書籍の配架状況に柔軟に対処する上で必要であり，従来の手法においても考慮されている．本手法が従来の手法と主に異なるのは第二および第三の特徴である．従来の手法では，書籍境界をエッジ [1], [2] や影 [3] として検出し，平行な近隣境

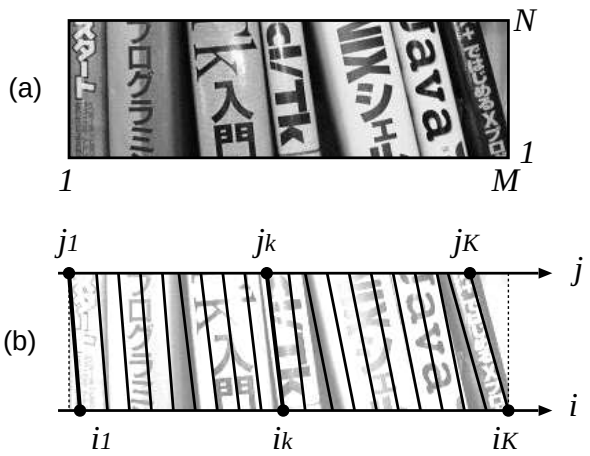

図1　書棚画像 (I成分)(a) と線分系列 (b)

界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，タイトル文字フォントの影響を考えると，エッジのような局所的な情報のみに依存するよりも，画像全体としての最適性も考慮した方が安定した境界検出ができると考えられる．また，このように最適化の枠組で境界検出問題を扱うことで，第三の特徴であるモデルの組み込みも可能となり，結果として高い検出精度を期待できる．

以下では，本手法の基本原理を述べ，実際に書棚画像から書籍境界の検出実験を行い，そ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．この切り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．

この書棚画像において，図1(b) のような K 本の線分の系列を考える．第 k 番目 $(k = 1, 2, \ldots, K)$ の線分は上下端の水平座標 (i_k, j_k) の組で表現され，また第 k 線分と第 $k+1$ 線分の間隔は後述するように1画素程度とする．ここでもし，各線分の傾きをその位置の書籍の傾きに合致させることができれば，線分系列のうちに書籍境界が含まれる．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傾き変化の最適推定問題として扱う．第 k 線分の傾きは (i_k, j_k) により一意に定まるので，実際には (i_k, j_k) の最適化問題となる．ところで，(i_k, j_k) を単純に最適推定しても，どの線分が求めるべき書籍境界に当るかは自明ではない．また，精度向上のためには，1冊の書籍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，それを上述の傾き変化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) のように書棚画像の任意の水平断面を観察すると，書籍境界，背表紙 (spine) ，タイトル，および背景が，ある規則性をもって繰り返し出現していることがわかる．この規則は，図2(b) に示した6状態 ($\boldsymbol{S} = \{S_1, \ldots, S_6\}$) とその間の遷移からなるFSAモデルで表現できる．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．

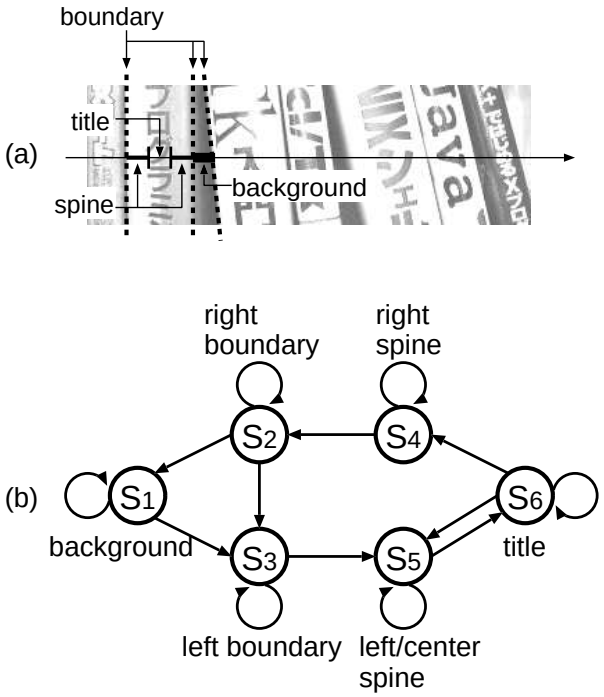

図2　書籍画像の水平方向の変化 (a) とそのFSA表現 (b)

- 右境界 (S_2) と左境界 (S_3) の間には背表紙領域 (S_4, S_5) がある．
- 背表紙領域 (S_4, S_5) の間には背文字領域 (S_6) がある．
- 1つの書籍中に背文字 (S_6) が背表紙領域 (S_5) を挟んで繰り返すことがある．
- 書籍間，すなわち左境界 (S_2) と右境界 (S_3) の間には，書棚背景 (S_1) が入ることもある．

このFSAモデルを傾き変化の最適推定問題に組み込むために，各線分を $(i_k, j_k|s_k)$ と表現し，その傾きだけでなく状態すなわち属性も併せて考える．その結果，解くべき問題は $(i_1, j_1|s_1), \ldots, (i_k, j_k|s_k), \ldots, (i_K, j_K|s_K)$ の最適推定問題となる．最適解において $s_k \in \{S_2, S_3\}$ を満たす (i_k, j_k) が求めるべき書籍境界となる．

最適化において，このFSAモデルは制約条件として働く．例えば第 k 線分の状態 s_k が S_4(右背表紙部) ならば第 $k+1$ 線分の状態 s_{k+1} は S_4(右背表紙部) もしくは S_2(右境界) に制約される．さらに，傾き変化についても状態に応じた制約を課する．具体的には，書籍内部では傾き変化が起らない点を利用し，状態 S_4, S_5, S_6 では傾き変化を許さないように制御すれば，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．

$$\begin{aligned} &\textbf{minimize} \quad \sum_{k=1}^{K} f(i_k, j_k | s_k) \\ &\textbf{w.r.t.} \quad i_k, j_k, s_k \ (k = 1, 2, \ldots, K) \end{aligned} \tag{1}$$

ここで，$f(i,j|s)$ は線分 (i,j) の状態が $s \in \boldsymbol{S}$ として妥当であるほど小さな値を返す関数である．関数 $f(i,j|s)$ の詳細は2.3節で述べる．最小化の際の制約条件は，s_k と s_{k+1} が図2(b)のFSAモデルに従うという条件と，傾き変化を許す状態を定める次の条件である．

$$\left.\begin{aligned} i_k &= i_{k-1} + \Delta i \\ j_k &= j_{k-1} + \Delta j \end{aligned}\right\} \quad \textbf{if} \quad s_k \in \{S_1, S_2, S_3\} \ \textbf{and} \ s_{k-1} = s_k \tag{2}$$

$$\left.\begin{aligned} i_k &= i_{k-1} + 1 \\ j_k &= j_{k-1} + 1 \end{aligned}\right\} \quad \textbf{otherwise} \tag{3}$$

ここで，$\Delta i, \Delta j \in \{0, 1\},\ \Delta i + \Delta j > 0$ とする．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 (S_1) および書籍境界 (S_2, S_3) のとき起こりうるものとする．なお，この条件より，第 k 線分と第 $k+1$ 線分の間隔は高々1 画素となる．

2.2 DPアルゴリズム

線分の上下端の水平座標値 j, i をそれぞれ縦軸，横軸とする平面を各状態について用意すると，$(i_k, j_k | s_k)$ の系列は，図3のようにこの平面上で1つの経路を成す．この経路上の1点が，1つ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 (2),(3) に従うことになる (図4)．従って解くべき問題は，このような遷移を行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，こうした最適経路問題はDPに基づくアルゴリズムにより効率的に解くことができる．具体的なDPアルゴリズムを図5に示す．ここでは次のようなワークエリアを用いている．

- $g(i,j|s)$：座標 $(i,j|s)$ に至るまでの最小累積コストを格納する．
- $bp(i,j|s)$：上記 $g(i,j|s)$ を与える経路が座標 $(i,j|s)$ に至る前に経由した直前の状態および座標を格納する．

ここで，図5のStep 8,9の $(i', j'|s')$ は座標 $(i,j|s)$ に至る直前に経由可能な座標を表す．例えば，$s = S_3$ の場合には図4(3)よりわかるように5通りの $(i', j'|s')$ が存在し，$s = S_6$ の場合には同図 (6) より2通りの $(i', j'|s')$ が存在する．

このアルゴリズムが終了した後，$i = M$ もしくは $j = M$ を満たす座標 $(i,j|s)$ のうち $g(i,j|s)$ の値が最小となるものを探索する．その座標を $(i_K, j_K|s_K)$ とし，そこから $bp(i,j|s)$ を用いてバックトラック処理を行うことで，順次最適経路 $(i_{K-1}, j_{K-1}|s_{K-1}), \ldots, (i_k, j_k|s_k), \ldots, (i_1, j_1|s_1)$ を得る．前述のように，このうち $s_k \in \{S_2, S_3\}$ を満たす (i_k, j_k) が求めるべき書籍境界となる．

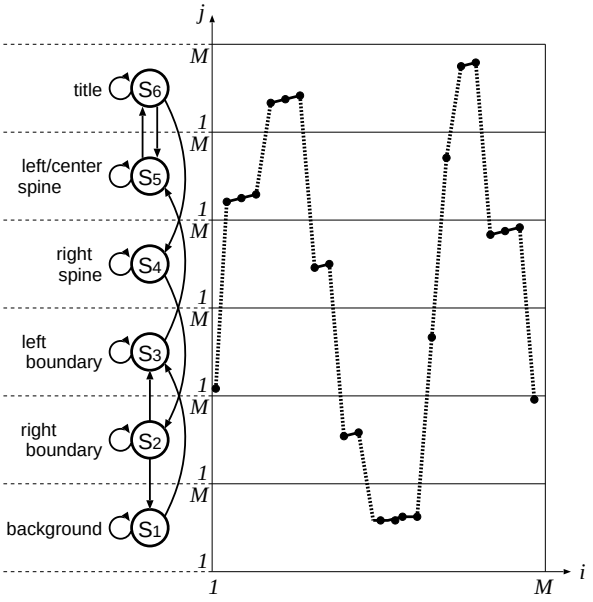

図3　DP 計算空間

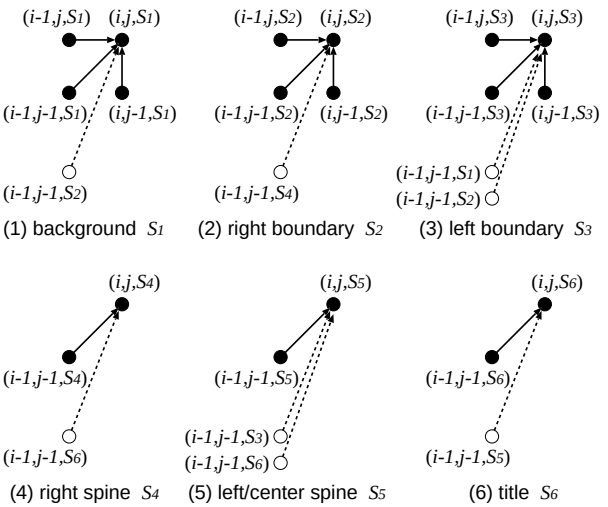

図4　状態別の遷移パス

2.3 特徴抽出

関数 $f(i,j|s)$ は線分 (i,j) におけるエッジの重なり具合によりその線分の状態が s であることの妥当性を評価する．書棚画像から検出されるエッジ画像を観察すると次のような特徴をもつことがわかる．

1. 書籍境界では長い縦方向エッジが検出されやすい．
2. 横方向エッジは背文字や背表紙のデザインによるものが多い．
3. 書棚背景や背表紙ではエッジが検出されにくい．

従って，書籍境界部 $f(i,j|S_2), f(i,j|S_3)$ は線分 (i,j) 上に縦方向エッジが多いほど小さい値を与える．その際，連続する縦方向エッジは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．逆に背景部 $f(i,j|S_1)$ はエッジが少ない

/* Initialization: */
1 **for all** $s \in \boldsymbol{S}$ **do begin**
2 **for** $j = 1$ **to** M **do** $g(1, j|s) := f(1, j|s)$
3 **for** $i = 1$ **to** M **do** $g(i, 1|s) := f(i, 1|s)$
4 **end**
/* DP Recursion: */
5 **for** $i = 2$ **to** M **do**
6 **for all** $s \in \boldsymbol{S}$ **do**
7 **for** $j = 2$ **to** M **do begin**
8 $g(i, j|s) := f(i, j|s) + \min_{(i',j'|s')} \{g(i', j'|s')\}$
9 $bp(i, j|s) := \operatorname{argmin}_{(i',j'|s')} \{g(i', j'|s')\}$
10 **end**

図5 DPアルゴリズム

ほど小さい値を与える．類似して背表紙部$f(i, j|S_4), f(i, j|S_5)$もエッジが少ないほど小さい値を与えるが，表紙デザインの影響を低減するために横方向エッジは無視する．タイトル部$f(i, j|S_6)$は，以上とは異なる性質を持つものとして，横方向エッジが多いほど小さい値を与える．

3. 実　　験

3.1 前　処　理

本実験ではディジタルカメラで撮影したサイズ640×480の書棚画像50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) 画像の切り出し

各書棚画像に対して図6(b)のように画像中央部(サイズ640×100)を切り出す．図5の境界検出はこの切り出し画像を対象とする．これは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ためである．また計算時間を低減させる効果もある．

(2) エッジ検出

評価関数$f(i, j|s)$の計算に備え，切り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[2]．まず画像のHSIプレーンそれぞれについて5×5のキャニーフィルタを用いてエッジ検出を行う．エッジは$\pm 45°$を境界として縦横2方向に量子化する．エッジ画像は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．最後にエッジ画像には2値化を施す．図6(c)にエッジ画像の例を示す．上が縦方向エッジ，下が横方向エッジを表す．

3.2 書籍境界検出実験

以上の手順で前処理された書棚画像50枚について実際に書籍境界検出実験を行った．画像1枚当りの計算時間は，前処理とDPアルゴリズムを併せた全体で約2.8秒(Xeon 1.7GHz)であった．

図7に境界検出結果の例を示す．図7の検出結果(右)は黒線が検出された書籍境界を表し，横縞の入った領域は書棚背景と

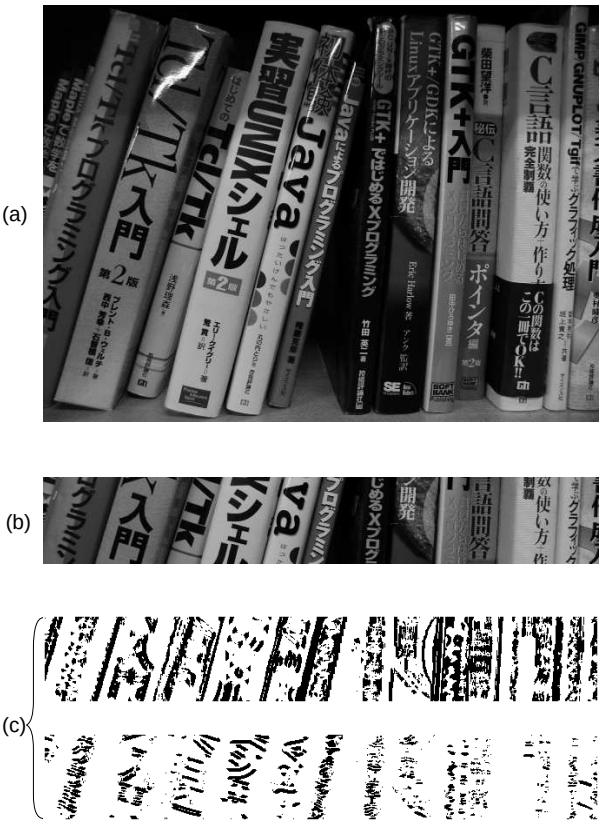

図6 書棚画像(a)，切り出し処理後の画像(b)，エッジ画像(c)(上：縦方向エッジ，下：横方向エッジ)の例

して推定されたことを表す．図7(a),(b),(c)では画像中のほとんどの書籍の境界が適切に検出できており，本手法の有効性が確認できる．特に同図(c)の矢印が示す境界上ではエッジがあまり検出されなかったにもかかわらず境界が検出された．これは大局的最適化とFSAモデルを組み合わせるという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)のように誤って境界が検出されたり，背表紙部を書棚背景に取り違えるといった現象も見られた．前者は，背文字の右端もしくは左端部の連続した縦方向エッジを境界と見なしたことに起因する．後者は，切り出し時に背文字が欠落したり，背表紙上の背文字のコントラストが低いなどの理由により，文字エッジが検出されなかったことが主たる原因と考えられる．

次に本手法の性能を定量的に評価するために，左右境界が共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「すべての書籍が隙間なく直立している」(例えば図7(a)のような状態)，「すべての書籍が隙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，本手法では入力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影響はないことが確認できる．検出を誤った主たる原

表1　書籍の検出率(%)

image(#samples/#books)	detection rate
no slant (16/205)	79.0
uniform slant (17/182)	79.7
nonuniform slant (17/174)	78.2
total (50/561)	79.0

因は，背文字エッジによる境界誤認，背文字エッジの欠落，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．さらに書棚画像の文法モデルを利用することで高精度な境界検出を実現した．実験により本手法の有効性を確認した．

今後の課題としては，実験で確認された問題点を検討し，手法の改善を図る．具体的には，(1)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，(2)背文字部のエッジの欠落を抑える工夫を施す，(3)背文字部および背表紙部はある程度の幅を持つといった詳細な性質をモデルに反映させる，ことが挙げられる．

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] 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抽出に関する研究,” テレビジョン学会技術報告, Vol. 15, No. 42, pp. 27–32, Jul. 1991.

[2] 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I, pp. 271–276, Jul. 2000.

[3] 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,” 映像情報メディア学会誌, Vol. 54, No. 6, pp. 881–886, Jun. 2000.

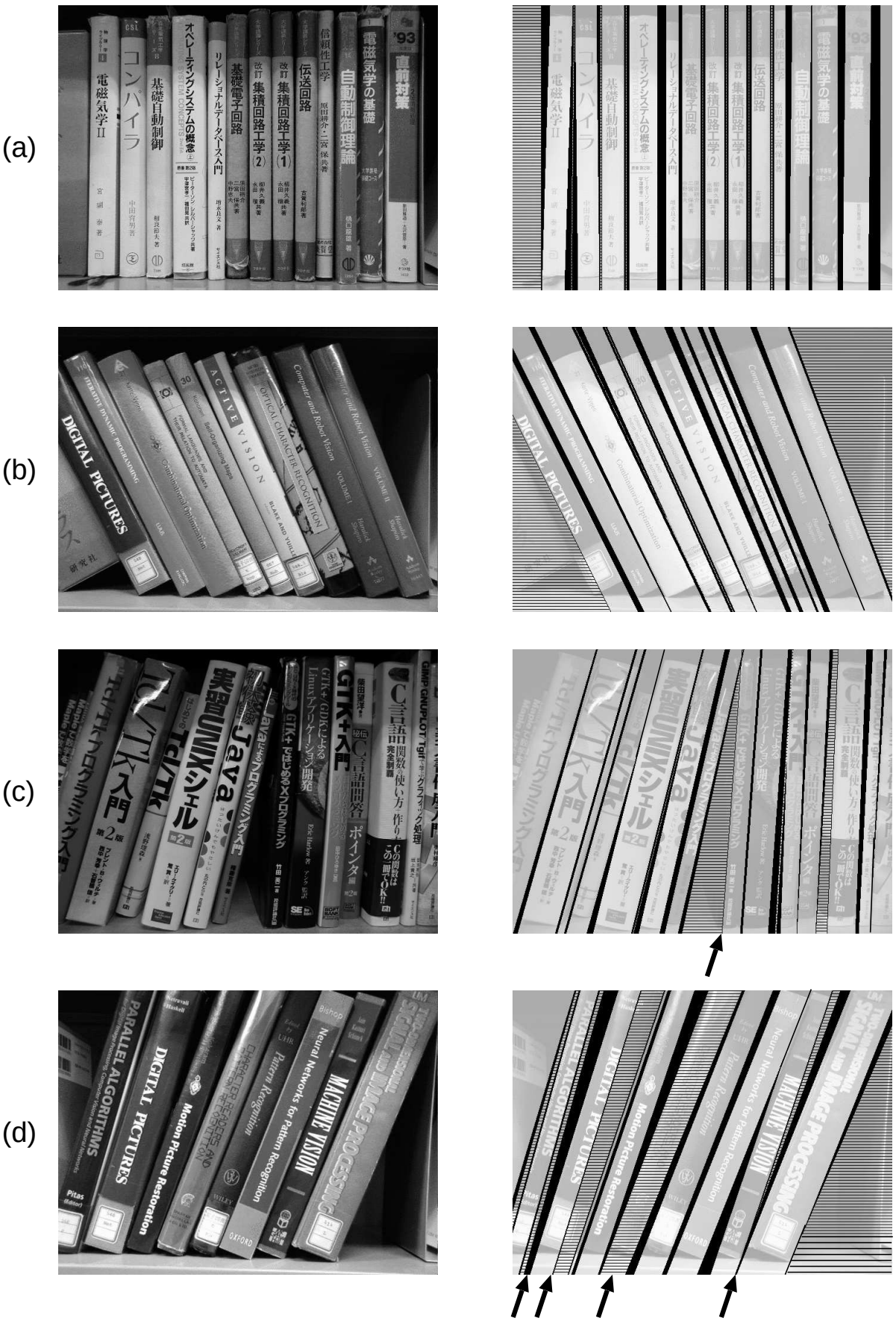

図7　書籍境界検出例：原画像 (I 成分のみ)(左)，検出結果 (右)．検出結果において黒線は書籍境界，横縞は書棚背景を示す．切り出し画像に対して検出された境界を上下に延長して示してある．