A MODEL-BASED BOOK BOUNDARY DETECTION TECHNIQUE FOR BOOKSHELF IMAGE ANALYSIS

Eiji TAIRA †, *Seiichi UCHIDA* ‡, *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

ABSTRACT

Several systems based on bookshelf image analysis have been studied for automating bookshelf inspection in libraries and bookstores. In conventional systems, book boundaries are firstly detected for extracting individual book by using some popular line detection technique, such as Hough transform. Their detection accuracies, however, are sometimes insufficient. In this paper, we propose a novel book boundary detection technique using a bookshelf model for higher detection accuracy. The model is designed for describing the structural properties of bookshelves and represented as a finite state automaton, each of whose states corresponds to a component of bookshelves, such as boundary and title. With this model, the book boundary detection problem is formulated as a model-based optimization problem where states and local slant angles at all horizontal positions were variables to be optimized. The globally optimal solution is efficiently searched for using a dynamic programming based algorithm. The effectiveness of the proposed technique was shown by several experiments.

1. INTRODUCTION

Inspection of bookshelves for managing libraries and bookstores is a time-consuming and tedious task because there are enormous books in the bookshelves and many of those books are moved frequently and repeatedly. Bar code systems are often utilized for semi-automating the bookshelf inspection. Those systems, however, are still largely based on manual tasks.

Several systems based on bookshelf image analysis have been studied for full-automating the bookshelf inspection [1]–[5]. The goal of those systems is to extract individual book information, such as title written on the spine of each book, from bookshelf images. For this goal, the following four steps are assumed in the conventional bookshelf image analysis systems; (i) detection of book boundaries, (ii) segmentation of individual book spine regions, (iii) extraction of title characters on each book spine region, and (iv) recognition of book titles. Among those processes, the book boundary detection process (i) is the essential front end and therefore its performance is crucial.

In conventional bookshelf image analysis systems, book boundaries are detected by some popular line detection technique. For example, Hough transform on the edge map of a bookshelf image is employed for detecting contiguous and straight edges corresponding to book boundaries. Unfortunately, the detection accuracies of the conventional techniques are sometimes insufficient. This is because book boundaries are often missed due to the lack of boundary edges caused by ill lighting conditions and the contiguity of books of the same color. In addition, spurious boundaries are often detected due to the edges around title characters (or illustrations) on book spine regions.

In this paper, we propose a model-based book boundary detection technique. For higher detection accuracy, a finite state automaton (FSA) model is newly employed, which describes the structural properties of bookshelves. Roughly speaking, each state of this FSA model corresponds to a component of bookshelves, such as boundary and title. With this model, the book boundary detection problem is expressed as a model fitting problem to identify the state (i.e., the component) at each horizontal position. Finally, the horizontal positions identified as the boundary state are detected as the positions where book boundaries exist.

For the practical boundary detection, it should also be taken into account that books have their individual slants and therefore their boundaries are also slanted non-uniformly. For coping with the non-uniformly slanted books, the slant angle at each horizontal position should be estimated optimally. Consequently, the book boundary detection problem is expressed as

This work was supported in part by the Ministry of Education, Culture, Sports, Science and Technology in Japan under a Grant-in-Aid for Scientific Research No.14780293.

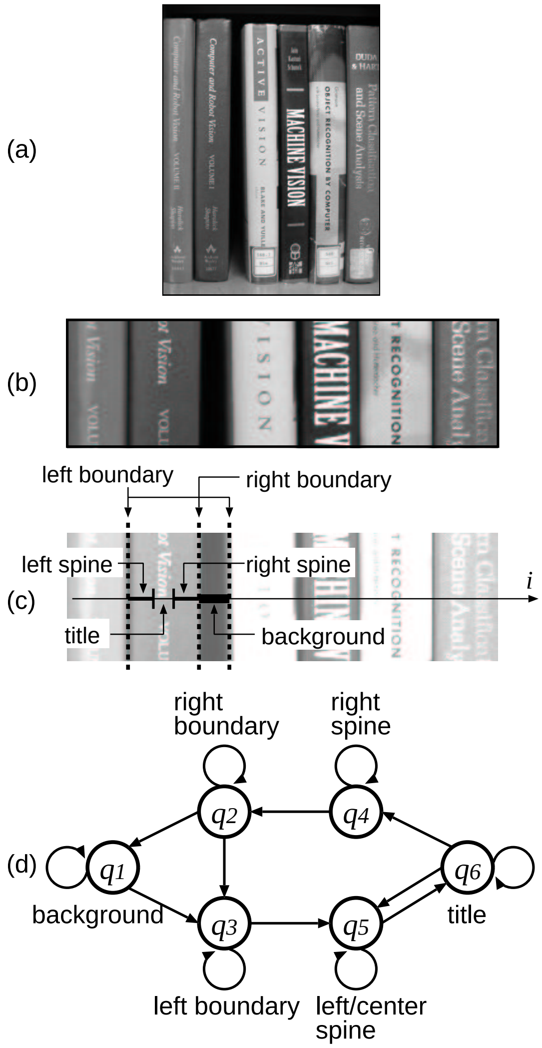

Fig. 1. (a) An original bookshelf image (HSI color image actually). (b) The middle part of (a). (c) The structural properties of (b). (d) An FSA representation of (c).

a model-based optimization problem of both the state and the local slant angle at each horizontal position. The globally optimal solution is efficiently searched for using a dynamic programming (DP) based algorithm.

Recently, bookshelf inspection systems using radio frequency identification (RFID) tags is developed. Those systems can be also realize full-automatic and contact-free inspection. Incorporation of bookshelf image analysis systems into the RFID inspection system is promising because bookshelf image analysis systems will compensate several drawbacks of the RFID tags, such as interference between tags, shield by metals, and narrow scope (due to weak wave power restricted by some law).

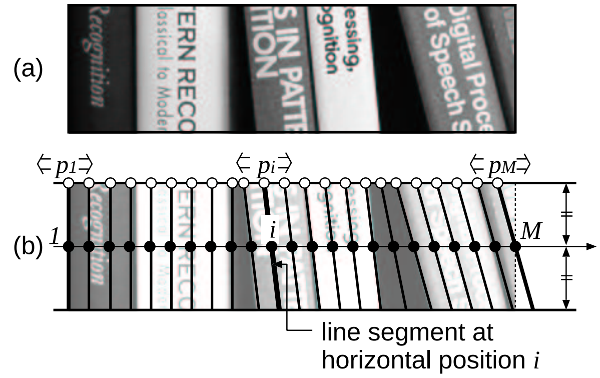

Fig. 2. (a) Non-uniformly slanted books. (b) A sequence of line segments representing the slants of books.

2. MODEL-BASED BOOK BOUNDARY DETECTION TECHNIQUE

2.1. Problem formulation

Figure 1(a) shows an example of an original bookshelf image. As shown in Fig. 1(b), we use the middle part (M (width) $\times$ N (height)) of the original image for disregarding bookshelf boards and the difference in the heights of the books.

As shown in Fig. 1(c), there are the following structural properties in bookshelf images.

- A book spine region always lies between two parallel neighboring boundaries.
- A book title region lies around the center of each book spine region.
- A bookshelf background region sometimes lies between two neighboring books.

Thus, it seems that the order of the components of bookshelves, i.e., boundary, spine, title, and bookshelf background, is naturally governed by some rules. In fact, as shown in Fig. 1(d), the rules can be represented by an FSA model composed of 6 states ($\boldsymbol{Q} = \{q_1, \ldots, q_6\}$), each of which corresponds to one of those components. With this model, the book boundary detection problem is expressed as a model fitting problem, i.e., a model-based optimization problem of the sequence $s_1, \ldots, s_i, \ldots, s_M$ where $s_i \in \boldsymbol{Q}$ denotes the state identified at horizontal position i.

For the practical boundary detection, it should also be taken into account that books have their individual slants and therefore their boundaries are also slanted non-uniformly (Fig. 2(a)). For coping with the non-uniformly slanted books, the slant angle at each horizontal position should be estimated optimally. This estimation problem can be expressed as an optimal estimation problem of the sequence $p_1, \ldots, p_i, \ldots, p_M$ where p_i is the variable to represent the local slant angle at horizontal position i. As shown in Fig. 2(b), $p_i \in [1, 2, \ldots, M]$ denotes the horizontal position of the top end of the line segment which passes through the center of the ith column. For practical simplicity, we use p_i instead of some real-valued angle.

Finally, for the detection of the book boundaries, we should consider a model-based optimization problem of both s_i and p_i at each horizontal position i. Specifically, this problem is formulated as the following maximization problem,

$$\begin{array}{ll} \textbf{maximize} & \displaystyle\sum_{i=1}^{M} f(p_i, s_i \mid i), \\ \textbf{w.r.t.} & p_i, s_i \ (i = 1, 2, \ldots, M), \end{array} \qquad (1)$$

where $f(p_i, s_i|i)$ is a criterion function to evaluate p_i and s_i at horizontal position i. The details of $f(p_i, s_i|i)$ will be described in Section 2.3. The maximization problem (1) is solved under three constraints. The first constraint is a condition that the state sequence $s_1, \ldots, s_M$ is governed by the state transition rule of the FSA model (Fig. 1(d)). The second constraint is employed to limit maximum slant angle and defined as

$$|i - p_i| \le W \qquad (2)$$

where W is a positive integer specifying the range of compensable slant angles. The third constraint is employed to limit the interval between p_i and p_{i-1} and defined as

$$p_i = \begin{cases} p_{i-1} + 1 & \textbf{if } s_i \in \{q_4, q_5, q_6\} \\ & \quad \textbf{or } s_{i-1} \in \{q_4, q_5, q_6\}, \\ p_{i-1} + \{0, 1, 2\} & \textbf{otherwise}. \end{cases} \qquad (3)$$

With this constraint, angle fluctuations are not allowed if $s_i \in \{q_4, q_5, q_6\}$ or $s_{i-1} \in \{q_4, q_5, q_6\}$. This means that a book spine region has its fixed slant. On the other hand, at book boundaries and bookshelf background regions, the angle fluctuations are allowed, while its degree is limited.

2.2. Solution by DP

Considering a sequential optimization process of p_i and s_i $(i = 1, \ldots, M)$, the process has the Markov property. This is because under the definition of the maximization problem and the three constraints of Section 2.1, only two values p_{i-1} and s_{i-1} are necessary for determination of p_i and s_i, and the other past values (e.g., p_{i-2} and s_{i-2}) are not necessary. It is well known that the optimization problem with the Markov property can be solved efficiently using DP.

/* Initialization: */
1 **for all** $s_1 \in \boldsymbol{Q}$ **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** $s_i \in \boldsymbol{Q}$ **do**
6 **for** $p_i := i - W$ **to** $i + W$ **do begin**
7 $g(p_i, s_i|i) := f(p_i, s_i|i) + \max_{p_{i-1}, s_{i-1}} g(p_{i-1}, s_{i-1}|i-1)$
8 $bp(p_i, s_i|i) := (p_{i-1}, s_{i-1}|i-1)$ which gives the maximum at step 7
9 **end**
/* Backtracking: */
10 $(p_M^{\text{opt}}, s_M^{\text{opt}}|M) := \operatorname{argmax}_{p,s} g(p, s|M)$
11 **for** $i := M$ **downto** 2 **do**
12 $(p_{i-1}^{\text{opt}}, s_{i-1}^{\text{opt}}|i-1) := bp(p_i^{\text{opt}}, s_i^{\text{opt}}|i)$

Fig. 3. DP algorithm.

Figure 3 shows a DP algorithm for the optimization problem formulated in Section 2.1. The value $g(p_i, s_i|i)$ is the maximum (i.e., optimal) cumulated value of $f(p_k, s_k|k)$ up to $k = i$. Step 7 is so-called DP recursion, and its two variables p_{i-1} and s_{i-1} are restricted by the FSA model (Fig. 1(d)) and the constraints (2) and (3). The optimized p_i and s_i, denoted as p_i^{opt} and s_i^{opt} respectively, are obtained by the backtracking procedure Step 10–12. Finally, if s_i^{opt} equals to q_2 or q_3, the line segment at horizontal position i is detected as a book boundary, which is slanted at the angle specified by p_i^{opt}. Note that the DP algorithm in the proposed technique required about 2.4 sec on a PC (Xeon 1.7GHz) for an image of size 640×100.

2.3. Design of criterion function

The criterion function $f(p_i, s_i|i)$ is designed based on the following observations.

- Long contiguous edges with near-vertical directional feature are often detected around book boundaries.

- Most edges with near-horizontal directional feature are detected around the title characters and the illustrations on book spine regions.

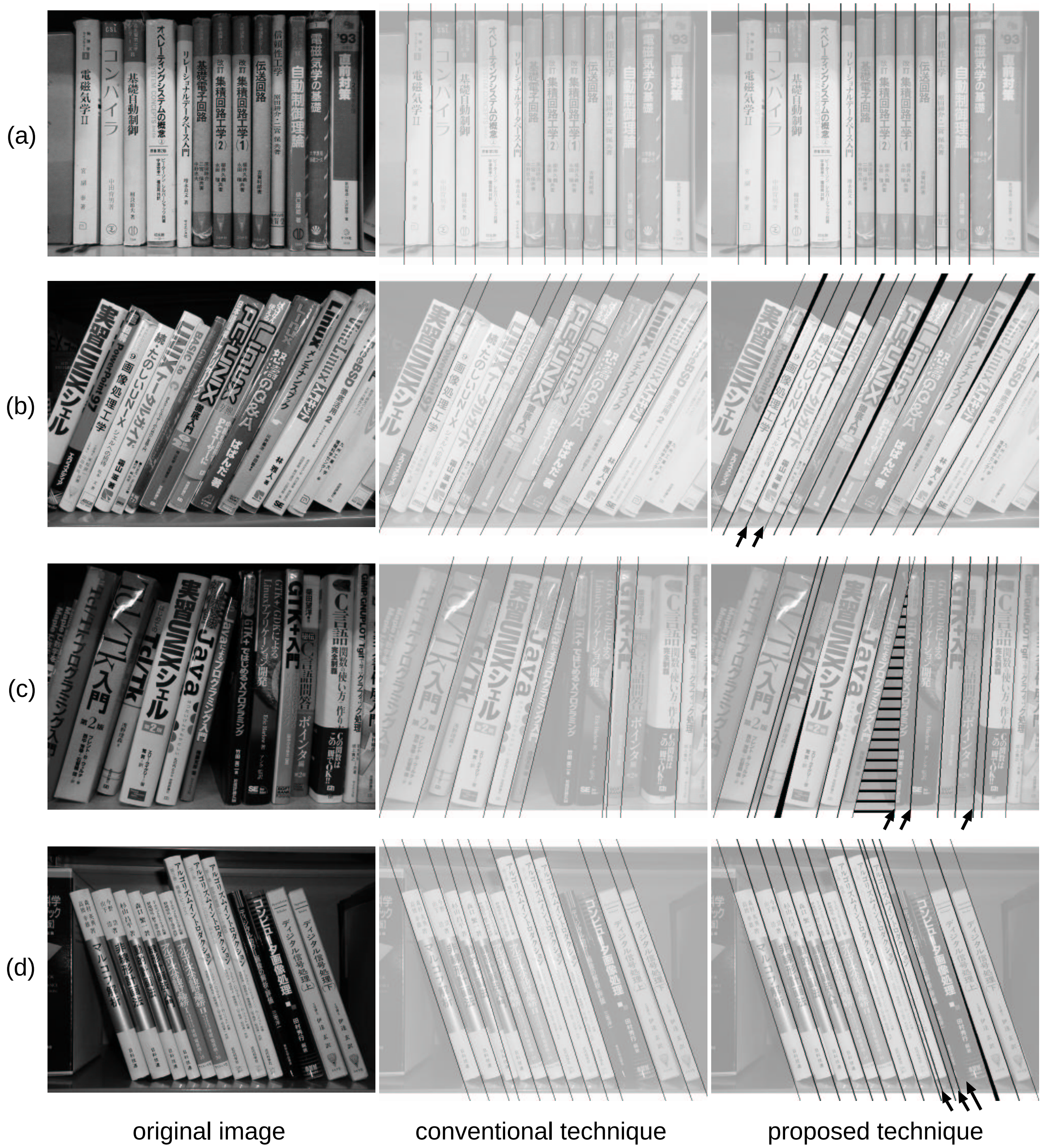

Fig. 4. Examples of book boundary detection. For each image, the original image (left), the result by the conventional technique A (middle), and the result by the proposed technique (right) are shown. The book boundaries detected are shown as black lines, and the bookshelf background regions detected by the proposed technique are shown as horizontal hatchings. In (b) and (c), the boundaries correctly detected only by the proposed technique are indicated by arrows. In (d), the boundary missed and the spurious boundaries detected by the proposed technique are indicated by arrows.

- Edges are rarely detected in bookshelf background regions and book spine regions.

Thus, the criterion function for the book boundary (i.e., $f(p_i, q_2|i)$ and $f(p_i, q_3|i)$) are designed to take larger value if there are many and/or long near-vertical edges on the line segment at horizontal position i. Since long contiguous edges show the book boundary more clearly than other edges, each edge is given a weight according to the number of the edges connected to it. The criterion function for the title (i.e., $f(p_i, q_6|i)$) is designed to take larger value if there are many near-horizontal edges on the line segment. The criterion function for the bookshelf background (i.e., $f(p_i, q_1|i)$) is designed to take smaller value if there are many edges on the line segment. The criterion function for the book spine (i.e., $f(p_i, q_4|i)$ and $f(p_i, q_5|i)$) are designed to take smaller value if there are many near-vertical edges on the line segment. For suppressing the effect of the illustrations, near-horizontal edges are disregarded.

3. EXPERIMENTAL RESULTS

3.1. Sample images

Experiments for evaluating the proposed technique qualitatively and quantitatively were conducted. For the experiments, 60 color (HSI) bookshelf images of size 640×100 were prepared, each of which was clipped from an original image of size 640×480 as noted in Section 2.1. Then the following edge detection procedure was performed on each bookshelf image. Firstly, Canny edge detector was performed on three color components (i.e., hue, saturation and intensity) independently. Then the edge map of each component was binarized, and the three edge maps were combined to form one edge map.

3.2. Book boundary detection technique using Hough transform [5]

The proposed technique was compared with the following conventional technique [5].

1. Hough transform is performed on the edge map prepared in Section 3.1. At the voting to Hough parameter space, near-horizontal edges, which may not correspond to any book boundary, are disregarded and a weighting operation for emphasizing long contiguous near-vertical edges is employed.

2. A thresholding operation is performed to find local maxima in the Hough parameter space. Then the line corresponding to each local maximum is detected as a book boundary.

This conventional technique has two versions; in the first version (hereafter called "conventional technique A"), the threshold to find local maxima is adaptively decided at a value depending on the result of voting to the Hough parameter space. In the second version (hereafter called "conventional technique B"), the threshold is fixed at a manually determined value regardless of the result of voting.

3.3. Qualitative Evaluation

Figure 4 shows several results by the conventional technique A and the proposed technique ($W = 50$). The book boundaries detected are shown as black lines. Figures 4(a), (b), and (c) show that the proposed technique (right column) can detect most book boundaries correctly. Especially in (b) and (c), it is shown that the proposed technique can detect all boundaries of slanted books while the conventional technique A fails.

As shown in Fig. 4(d), the proposed technique sometimes misses some boundaries (false negative) and detects spurious boundaries (false positive). The former is mostly because the colors of adjacent books are similar each other and therefore boundary edges are lacked entirely. The latter mostly resulted from contiguous edges around title characters and areas reflecting some light.

It is worth to note that the proposed technique not only can detect book boundaries but also can segment a bookshelf image into individual book spine regions and bookshelf background regions, whereas the conventional technique can only detect book boundaries. For example, the bookshelf background regions identified by the proposed technique are shown as horizontal hatchings in Fig. 4. Furthermore, since the proposed technique can also detect title regions, it is expected that the proposed technique can be extended to the next steps, i.e., extraction of title characters and recognition of book titles.

3.4. Quantitative Evaluation

In order to ensure the superiority of the proposed technique against to the conventional techniques, quantitative evaluations were conducted using the 60 bookshelf images of Section 3.1. Figure 5 shows the number of false negatives and the number of false positives. In the graph, "" and "" show the results by the proposed technique and the conventional technique A, respectively. The evaluation of the conventional technique B was conducted while changing the threshold to find local maxima, and therefore its result is shown as a curve line in the graph.

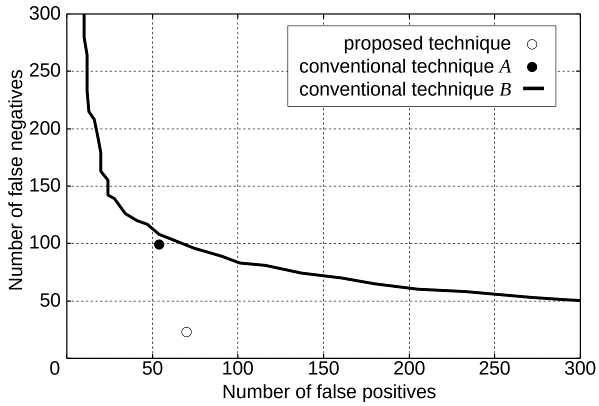

Fig. 5. Number of false positives and false negatives.

Table 1. Book detection rate (%)

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

The number of false negatives by the proposed technique was 3.2% (= 23/727) of all book boundaries and far less than 13.6% (= 99/727) of the conventional technique A. This number was less than the half of that by the conventional technique B even when 300 false positives were allowed. On the other hand, the number of false positives by the proposed technique was slightly more than that by the conventional technique A. The reduction of the false positives is remained as future work.

Finally, we measured the number of books whose spine regions could be correctly identified by the proposed technique. Table 1 shows the rate of correctly detected books. In the table, "no slant" means that all books stand straight closely (like Fig. 4(a)), "uniform slant" means that all books have the same slant closely (like Fig. 4(b)), and "non-uniform slant" means that books have their individual slants (like Fig. 4(c)). This evaluation shows that the slants of books do not influence the book detection accuracy. Needless to say, the main cause of misdetection was the failure of boundary detection.

4. CONCLUSIONS

A novel book boundary detection technique using a bookshelf model was proposed in this paper. The model was represented as a finite state automaton based on the structural properties of bookshelves. The book boundary detection problem was formulated as a model-based optimization problem where states and local slant angles at all horizontal positions were variables to be optimized and then solved efficiently using a dynamic programming based algorithm. Experimental results show that the proposed technique has superior detection performance than the conventional techniques. The results also show that the reduction of spurious boundaries is the remaining problem.

5. REFERENCES

[1] Z. Otake, T. Nagao, T. Agui and M. Nakajima, "Extraction of character regions from bookshelf images," *Bulletin of the Japanese Society of Printing Science and Technology*, vol. 29, no. 1, pp. 38–45, Jan. 1992. (in Japanese)

[2] Y. Ishikawa, M. Ohkura, M. Shiono and R. Hashimoto, "A study of extraction of character regions from bookshelf images," *Bulletin of the Okayama University of Science*, vol. 32A, pp. 163–171, March 1997. (in Japanese)

[3] Y. Akiyama and M. Ito, "Book recognition from color images of book shelves," *Proceedings of the IAPR Workshop on Machine Vision Applications (MVA'98)*, Nov. 1998, pp. 106–110.

[4] M. Sawaki, H. Murase and N. Hagita, "Recognition of characters in bookshelf images using automatic dictionary selection based on estimated degradation," *Journal of the Institute of Image Information and Television Engineers*, vol. 54, no. 6, pp. 881–886, June 2000. (in Japanese)

[5] K. Hiroi, M. Fujiyama, T. Miura and M. Ito, "Book boundary extraction from book shelf images," *Proceedings of the Meeting of Image Recognition and Understanding (MIRU2000)*, vol. I, July 2000, pp. 271–276. (in Japanese)