Automatic Book Spine Extraction and Recognition for Library Inventory Management

Nevetha M P
Amrita Vishwa Vidyapeetham (University)
Coimbatore
Tamil Nadu - 641112
nevethaprakasam@gmail.com

Baskar A
Department of Computer Science and Engineering
Amrita Vishwa Vidyapeetham (University)
Coimbatore
Tamil Nadu - 641112
a_baskar@cb.amrita.edu

ABSTRACT

Manual inventory management in a library is by far arduous. Automation of book inspection can be achieved by using a simple camera based system that can recognize book spines in a book shelf. The book spines contain printed information such as title, author and publisher name, which can be extracted and verified with the library's database. Book spines can be segmented by detecting their rectangular boundaries which appear as straight lines. Line detection using hough transform and line segment detector may result in spurious boundaries due to the presence of long titles or graphics on the book spine. In this paper, we propose a technique to improve book spine border detection by devising set of constraints based on structural properties that can be used to filter the detected line segments so as to obtain book spine boundaries. The segmented book spines are binarized to extract the printed information such as title, author and publisher name. The text is recognized using Tesseract Optical Character Recognition Engine. The proposed algorithm was tested successfully on book shelf images with vertically oriented, uniformly inclined and multi-oriented book spines.

CCS Concepts

•**Computing methodologies** → **Image processing;** *Image segmentation;* •**Applied computing** → Optical character recognition;

Keywords

Library Inventory Management, Book Spine Segmentation, Text Recognition, Line Segment Detector, Text Binarization

1. INTRODUCTION

The task of maintaining an inventory in the library is extremely tedious and time consuming when done manually. There is a need for an automated camera based system that can do a frequent inventory of the books, so that it is always up to date so that the desired book can be retrieved easily. Sometimes, books may get misplaced in which case it might get difficult to locate the book manually. An automated system can be used to locate the misplaced book and notify its exact location. Research in the area of book spine recognition is an urgent need due to the rapid increase of books and journals each year. Some inventory systems use bar code or RFID tags which require a specialized reader, whereas a camera based system is more efficient when it comes to recognizing an entire rack of books.

Book spines are rectangular regions that contain key information about the book such as title, author name and publisher. It is the only visible part of a book when it is shelved. Information extracted from the book spine is sufficient to do an inventory of the library. Given the image of a book shelf, it must be segmented to extract he book spines so that we can use the information within the segmented region to identify the book spine. [1] proposes a method to do library inventory using a smartphone. They use hough transform to segment the book spines, but book spines are high textured regions with text and graphics which may mislead the system into identifying false borders. [2] overcomes these issues by detecting book spines using call number labels instead of the printed bibliographic description on them. But, due to extensive use, labels may wear out and the call number digits may get recognized incorrectly. Hence, we prefer a fully automated book spine extraction and text recognition system that identifies the book by its title, author and publisher name.

The proposed system takes a book shelf image as input and segments the book spines present in it by detecting the book spine boundaries. The boundaries appear as straight lines due to change in color of the book spines or narrow gaps between them. The Line Segment Detector is used for border line detection and are filtered to obtain precise book boundaries. The filtering of line segments is based on certain constraints derived from structural properties of book spines. The segmented book spines are then input to a recognition pipeline which binarizes the spine image to extract text information printed on it. Book spines may have dark text on a lighter background or light text on a darker background. We adopt a font and background color text binarization technique. Tesseract OCR Engine is used to recognize the textual content. The system is composed of two main modules as follows

WCI '15, August 10 - 13, 2015, Kochi, India

DOI: http://dx.doi.org/10.1145/2791405.2791506

- Book Spine Extraction : Boundary detection of book spines using line segment detector and filtering of line segments using structural constraints.

- Book Spine Recognition : Each extracted book spine is binarized to obtain bibliographic description such as the title, author and publisher name, which is recognized using an OCR.

The rest of the paper is organized as follows. Section 2 discusses related work. In section 3, we describe our approaches for extraction and recognition of book spines and in section 4 we present our experimental results. A conclusion is derived in section 5 with a summary of the proposed work.

2. RELATED WORK

Many automated book spine detection and recognition systems that identify book by the title were surveyed. Book spines are detected using their book boundaries that appear as dark lines due to the gaps between the two books or due to change in colour of the book spines. Techniques using sobel vertical and horizontal edge detection have been used to detect book boundaries. But these techniques only detect spines that are straight and do not detect lines that are inclined.

[4] proposed a book boundary technique based on a finite state automata model that was formulated based on structural properties of the arrangement of books in a book shelf such as that book boundaries are long near vertical edges and titles are composed of near horizontal lines. The borders were detected by global optimization of the model using dynamic programming.

Chen et al., have proposed a technique using Hough Transform [1] that can detect book spines that are densely packed and are inclined. Hough transform is applied to canny edge map filtered for long lines and book spine boundaries are detected as long lines with similar slant angle which will the highest accumulated angle in the histogram binned using hough transform.

Tsai et al.,[6] use [1] for book boundary detection and detect the text patches within the book spine using MSER and SWT. They have implemented a hybrid method of book spine recognition by matching a book spine image to the library's book spine image database using SURF features and character recognition of title as keywords to search the library book database.

Duan et al [2] proposed a system for book spine recognition by detecting the call number label which is pasted on the spine for reference. They used colour segmentation to detect the label and line segment detector to obtain borders of individual book spines. They deployed contour clustering to detect the call number and recognize the number using Tesseract OCR engine.

Talker and Moses, [5] segment book spines from any perspective by using a shape dependent active contour model to find candidate rectangular book spines. They perform a selection using spatial constraints defined for a book spine, which is formulated as maximal weighted independent set of a graph.

3. THE PROPOSED SYSTEM

The proposed system constitutes modules for book spine extraction and book spine text recognition. Fig. 1 depicts the block diagram of the proposed system. A book shelf image is given as input to the system and it segments the image into rectangular regions of book spines and recognizes the bibliographic description printed on it.

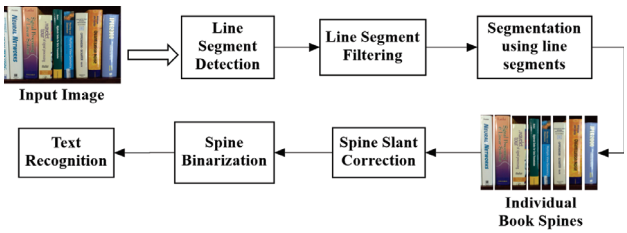


Figure 1: Block Diagram of Book Spine Extraction and Recognition System

3.1 Book Spine Extraction

The proposed work for book spine extraction is based on the facts as established in [4] which are as follows :

- The book boundaries will have long near vertical edges
- The title region will have more shorter near horizontal edges
- The background and book spine regions(excluding title region) do not have many edges.

Also, as suggested by Taira et al, we apply our border extraction algorithm to the middle part of the image. Given a book shelf image of width M and height N, we crop the image horizontally to obtain a segment of width M and height N/4. A line segment detector [7] is applied on the grayscale image. Due to the presence of textual components many broken lines will appear. We propose a selection technique that filters the obtained line segments based on constraints derived from the structural properties. The four stages of the proposed selection technique are :

- Stage 1 : All edges that are shorter than N/4 of the image are discarded.
- Stage 2 : All edges that have a slant angle of less than 15 degrees are discarded.
- Stage 3 : Edges that have a previous adjacent edge at a distance of 10 pixels are discarded.
- Stage 4 : If the edge strength in the neighborhood of an edge is low, it is discarded.

Filtered line segments at each stage serves as input for the next stage. Short broken lines that may belong to the background or the title region will get removed in the first stage. In the second stage, we remove all near horizontal edges since it may belong to the background or title region. We assume that all books will have a minimum book spine width of 10 pixels and hence we discard any adjacent lines that may appear as spurious boundaries in the third stage. In the fourth stage, we form a rectangular region of interest(ROI) of each of the edges obtained after the third filtering step. Total edge strength is calculated from the set of line segments obtained by the line segment detector. Let L, be the set of all

line segments obtained from line segment detector and L_i be the set of line segments within the ROI of line segment $x_i \in L$. The edge strength is computed as

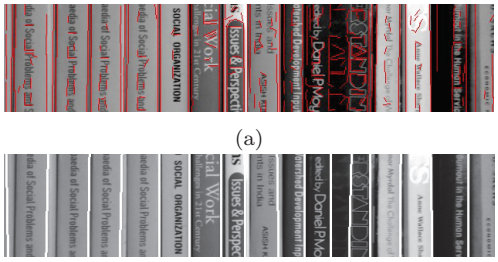
$$\sum_{x_j \in L_i} length(x_j) * sin(\theta_j) \quad (1)$$

where,

$\theta_j$is the angle with the horizontal of the line segment x_j

$length(x_j)$ is the number of pixels along the line segment x_j

The final obtained set of filtered line segments are extended to the height of the image by using the Digital Differential Analyzer(DDA) line drawing algorithm. Book spines are



(a)

(b)

Figure 2: (a)Line Segments detected by LSD. (b) Line Segments obtained after filtering.

extracted as the region sandwiched between two detected boundaries. Since the left and right boundaries of a book spine are parallel we obtain the slant of the book spine as angle made by the left boundary with the horizontal, and we rotate the segmented book spine with this angle. This is done so that the textual content in the book spine align horizontally and hence text detection is easier.

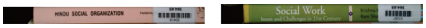
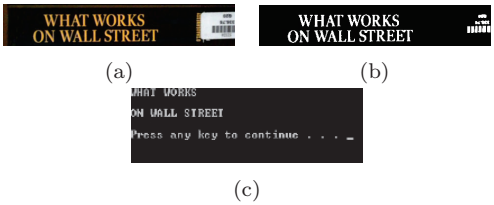


Figure 3: Book spines extracted using detected borders

3.2 Book Spine Text Recognition

After extraction of the book spine the next step is to binarize the book spine document image inorder to obtain the bibliographic information. This task is challenging due to the fact that the text may appear lighter or darker than the background. It also may have graphics or logos that may get detected with textual components. [3] a font and color independent text binarization method that extracts the text independent of the background color.

Canny edge maps from all three color spaces are obtained and merged together. Text is detected by using connected component analysis over the obtained canny edge map. Text detection is based on the observation that alphabets tend to have less than or equal to two loops inside them. The candidate text regions identified are used to calculate the threshold for binarizing the image. Intensity information from the background and foreground of detected text regions is used to calculate the desired threshold. The result is that in the binarized image, text always appears as white region. This method is used to extract the title and other information from the book spine image. The extracted textual content is segmented into lines and words by horizontal and vertical projection profile based segmentation. This method is used since it gives good results efficiently and takes very less time. Fig. 4 shows a sample result for text binarization.

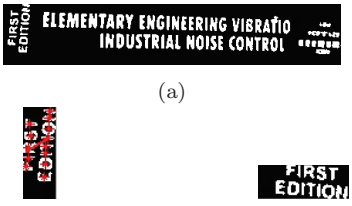


(a) (b)

(c)

Figure 4: (a)A book spine image. (b) A font and background colour independent text binarization is used to obtain text regions. (c) Text lines recognized using Tesseract OCR.

Vertical and horizontal text may simultaneously appear in the book spine. Vertical text must be rotated to horizontal for text recognition. We employ connected component analysis to determine whether the detected text is horizontal or vertical. The centroids of the connected components obtained in each segmented word is approximated to a line. If the approximated line is at an angle close to 90 degrees with the horizontal, it is rotated as shown in Fig. 5. Text Recognition is performed by Tesseract OCR engine which gives good recognition performance



(a)

(b) (c)

Figure 5: (a) A binarized book spine image with horizontal and vertical text. (b) Line approximation of connected components of a text region. (c) Vertical text rotated by angle of approximated line with the horizontal

4. EXPERIMENTAL RESULTS

In this section, we present the results obtained on testing the proposed system on book shelf images captured by a webcam in full colour format with a resolution of 640x480 pixels. The dataset includes images with vertically oriented, uniformly inclined and multi-oriented book spines. Section 4.1 discusses the results obtained for book spine extraction and section 4.2 discusses results for book spine text binarization and recognition.

4.1 Books spine extraction evaluation

The performance of the book spine segmentation technique was tested on 120 book shelf images, 40 each of vertically oriented set, uniformly inclined set and multi-oriented set. We have analyzed the performance of the proposed technique with each set having book spines with four different

(a) (b) (c) (d) (e)

Figure 6: (a),(b),(c) Border detection in vertically oriented set. (d) Border detection in uniformly inclined set. (e) Border detection in multi-oriented set

categories of sizes. The first category has book spines with uniform height and uniform width. The second category has book spines with uniform height and different width. The third category has book spines with different height and uniform width. The final category has book spines with different height and different width. Table 1, Table 2 and Table 3 contain the statistics recorded while testing the proposed segmentation technique on vertically oriented book spine set, uniformly incline book spine set and multi-oriented book spine set respectively.

Types of book spines analyzed(#images)	Total no. of book spines	Total no. of borders counted manually	No. of borders correctly identified
Uniform Height, Uniform Width(10)	104	114	101
Uniform Height, Different Width(10)	119	129	120
Different Height, Uniform Width(10)	130	140	123
Different Height, Different Width(10)	135	145	133

Table 1: Vertically Oriented Book Spine Border Detection

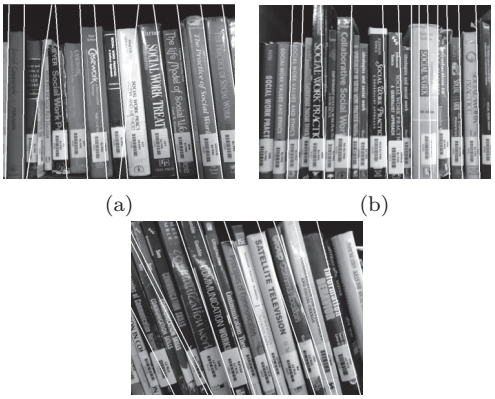
Types of book spines analyzed(#images)	Total no. of book spines	Total no. of borders counted manually	No. of borders correctly identified
Uniform Height, Uniform Width(10)	101	111	96
Uniform Height, Different Width(10)	110	120	96
Different Height, Uniform Width(10)	112	117	99
Different Height, Different Width(10)	120	130	101

Table 2: Uniformly Inclined Book Spine Border Detection

Types of book spines analyzed(#images)	Total no. of book spines	Total no. of borders counted manually	No. of borders correctly identified
Uniform Height, Uniform Width(10)	88	105	94
Uniform Height, Different Width(10)	106	126	104
Different Height, Uniform Width(10)	110	130	77
Different Height, Different Width(10)	122	142	119

Table 3: Multi-oriented Book Spine Border Detection

Although border detection is successful, there are some instances where the performance of the system deteriorates. Wear and tear on book spines introduce interference and pass off as spurious borders. Rich graphics content that contain different colours with gradients images or logos results in broken spine segmentation,i.e, a single book spine may get segmented into different parts. Multi-coloured book spine backgrounds also cause incorrect detection of borders.



(a) (b) (c)

Figure 7: (a),(b),(c) Failure in border detection due to wear and tear and graphics on book spine

Some instances of border detection failure are shown in in Fig. 7. This takes a significant toll on book spine extraction accuracy.

Fig. 8,9,10 depicts the book spine border detection statistics in bar charts. The detection rate is almost consistent in all categories analyzed.

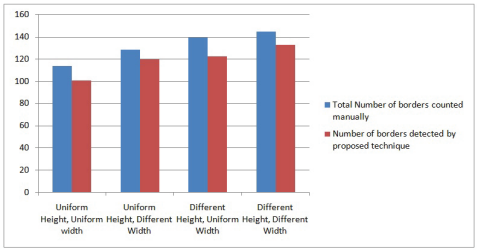


Figure 8: Vertically oriented Book Spine Border Detection Performance

4.2 Book spine text recognition evaluation

The binarization technique used, segments the text regions from the image. Fig. 11 shows results obtained by text binarization. The performance was tested on 20 book spine images. Table 4 shows the results of the book spine text recognition module tested on chosen images.

Text with regular and large font sizes were binarized and recognized successfully. Whereas, failure in proper binarization can be attributed to multi-coloured backgrounds on the

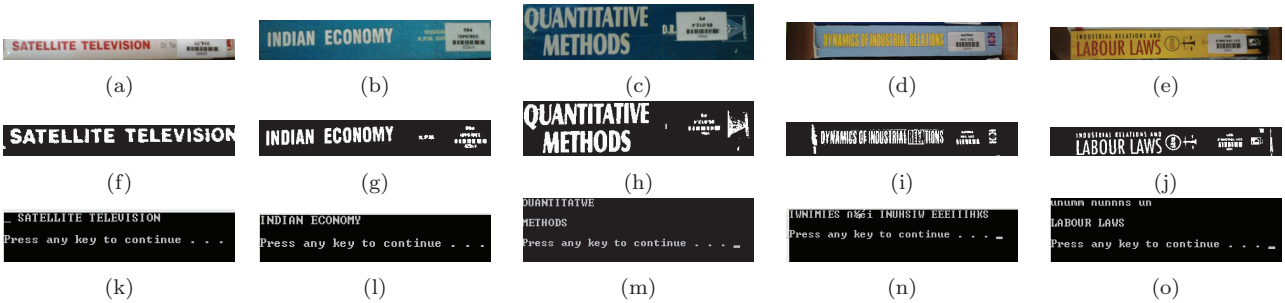


Figure 11: (a,b,c,d,e) Individual Book Spines. (f,g,h,i,j)Binarized Book Spines. (k,l,m,n,o) Text Recognized from Book Spines

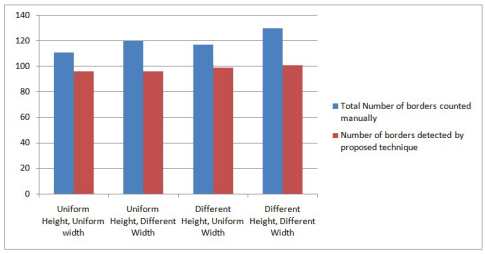


Figure 9: Uniformly Inclined Book Spine Border Detection Performance

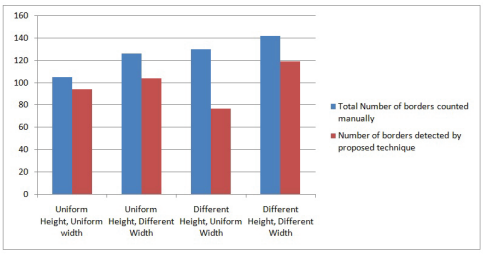


Figure 10: Multi-oriented Book Spine Border Detection Performance

book spine. Since, the binarization technique is dependent on background intensity measures, variations in intensity due to graphics and multi-coloured backgrounds interfere with the binarization process. Intensity variations caused by specular reflection over the spine also affects the binarization. Author names and publisher names that appear in small fonts fail to get binarized when there is blur in the captured image.

5. CONCLUSION

An automated book spine extraction and book spine text recognition system is proposed. The book spine extraction module works effectively on vertically oriented, uniformly inclined and multi-oriented book spines in a book shelf image. It has been tested successfully on book spines of varying sizes as well. The font and color independent text binarization technique chosen for preprocessing also serves as the text localization step. The Tesseract OCR engine recognizes the obtained textual content successfully. The indication from the obtained results is that the system shows good performance even with varying orientations, varying font sizes and varying colors of the book title. However, there is space for

Total no. of book spines analyzed	No. of book spines binarized successfully	No. of book spines recognized successfully
20	15	11

Table 4: Book Spine Text Recognition

improvement in recognition of the book title. Text binarization fails when the book spines have multiple shades or colors in the background, or the cover of the book is glossy causing specular reflections. Some book spines have rich graphic content or have text in cursive fonts which also might hinder the text binarization.

6. REFERENCES

[1] D. M. Chen, S. S. Tsai, B. Girod, C.-H. Hsu, K.-H. Kim, and J. P. Singh. Building book inventories using smartphones. In *Proceedings of the international conference on Multimedia*, pages 651–654. ACM, 2010.

[2] X. Duan, Q. Zhao, and S. Anwar. Identifying books in library using line segment detector and contour clustering. In *Computing and Convergence Technology (ICCCT), 2012 7th International Conference on*, pages 998–1003. IEEE, 2012.

[3] T. Kasar, J. Kumar, and A. Ramakrishnan. Font and background color independent text binarization. In *Second international workshop on camera-based document analysis and recognition*, pages 3–9, 2007.

[4] P. Taira, S. Uchida, and H. Sakoe. Book boundary detection from bookshelf image based on model fitting. In *International Symposium on Information Science and Electrical Engineering*, pages 601–604. Citeseer, 2003.

[5] L. Talker and Y. Moses. Viewpoint-independent book spine segmentation. In *Applications of Computer Vision (WACV), 2014 IEEE Winter Conference on*, pages 453–460. IEEE, 2014.

[6] S. S. Tsai, D. Chen, H. Chen, C.-H. Hsu, K.-H. Kim, J. P. Singh, and B. Girod. Combining image and text features: a hybrid approach to mobile book spine recognition. In *Proceedings of the 19th ACM international conference on Multimedia*, pages 1029–1032. ACM, 2011.

[7] R. G. Von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall. Lsd: a line segment detector. *Image Processing On Line*, 2(3):5, 2012.