Character Recognition of Thick Books using a Stereo Vision System

Atsushi Kawarago*, Atsushi Yamashita*, Toru Kaneko* and Kenjiro T. Miura*
Department of Mechanical Engineering
Shizuoka University

Abstract

In this paper, we propose a character recognition method for thick books with curved surfaces using a stereo vision system. Characters in images of thick book's pages acquired with an image scanner are difficult to recognize because they are deformed under the influence of curved surface. In our method, the 3D shape reconstruction of the book's surface is executed from the result of the stereoscopic measurement by putting the book upward, and a flat surface is recovered from the curved surface to improve the character recognition rate. The validity of the proposed method is shown through experiments.

1 Introduction

The digitization of books that are provided only in print is an important subject. Usually, images of books' pages are acquired with an image scanner and character recognition is executed. However, when scanning thick books, the midst areas become curved shapes and page surfaces cannot be in contact completely. For this reason, dark shadows, distortions and blurs come into existence at the midst areas, and the quality of images becomes insufficient for character recognition. Therefore, it is necessary to measure the shape of their surfaces to rectify images. In order to solve this problem, Wada et al. proposed the method that can recover the shape of unfolded book's surface from scanner images using their shading information[1]. This method is suitable for an image scanner, although it requires many calibration data because this method greatly depends on the condition of the illumination and the quality of book's material. Ukida et al. developed a method that can recover the shape in practical time by adopting the eigenspace approach[2]. These methods assume that the book surface is cylindrical and the cross section shape is uniform. Therefore, the reconstructed shape is the 2D cross section shape. However, when the scanline of the camera is not parallel nor perpendicular to the edge of the book surface, this assumption does not always approved in the case of thick books like dictionaries or conference proceedings.

In this paper, we propose a new character recognition technique based on the 3D shape reconstruction by using a stereo vision system (Figure 1). As to the issue of the curved surface, the 3D shape of book's surface is measured through triangulation and the 3D shape of the curved surface is reconstructed. The image of curved surface is transformed to that of flat surface to recognize characters with high reliability. As to the issue of the lighting condition, a correlation-based character recognition method for gray scale images is employed in order to improve the robustness for the shadows of images. Therefore, our proposed method does not depend on the lighting condition and the paper's material.

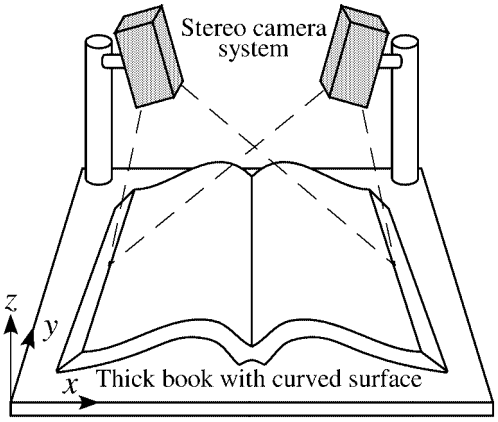

Figure 1: Proposed system using stereo camera.

Moreover, we can save time to prepare new pages by using a stereo camera system. Once the book is put upward, the digitalization can be executed while turning over pages in our method, while the book has to be replaced for every sheet when pages are entered through the image scanner.

2 Character Recognition with a Stereo Camera System

2.1 Outline of Our Method

The procedure of our method is divided into the off-line processing and the on-line processing. The former consists of 1) stereo camera calibration and 2) character template construction. The latter consists of 1) 3D measurement, 2) shape reconstruction, 3) image transformation, and 4) character recognition.

In the off-line processing, a camera calibration is executed for calculating extrinsic and intrinsic parameters by using images of a planer pattern. If the cameras' optic axes parallel[3], two images from camera are similar and the resolution of the book's curved area is not high in both cameras' images. Therefore, the camera configuration is set in a way that the convergent squint can be done.

The character templates for character recognition are created manually from characters that are on flat area of the book. The characters are classified into

*Address: 3-5-1 Johoku, Hamamatsu-shi, Shizuoka 432-8565, Japan.
E-mail: f0130019@ipc.shizuoka.ac.jp, yamashita@ieee.org, tmtkane@ipc.shizuoka.ac.jp, ktmiura@eng.shizuoka.ac.jp

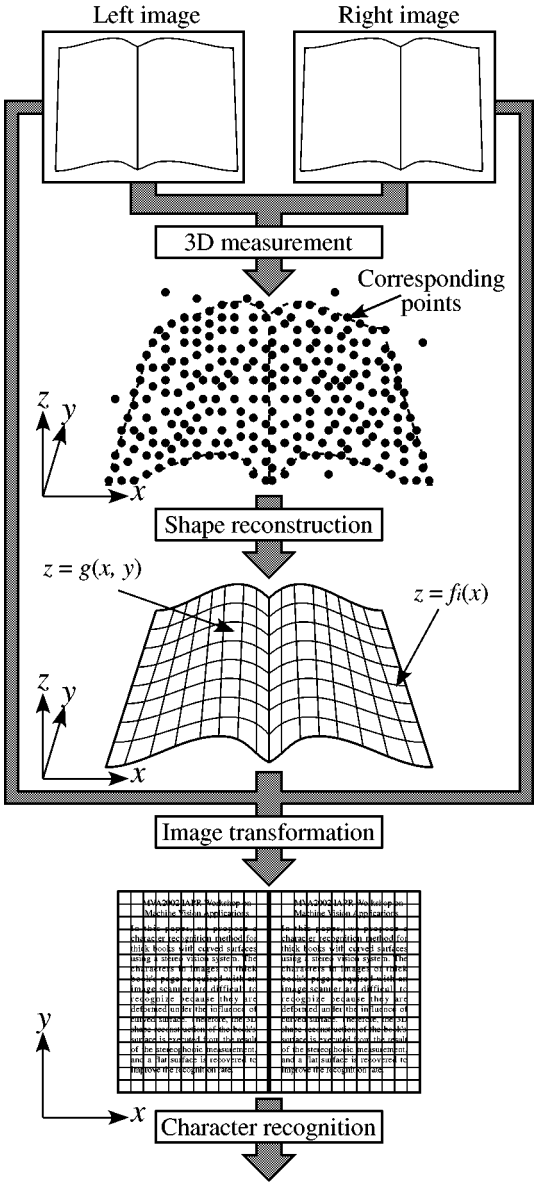

Figure 2: Overview of our method.

groups by their width and height, and these groups are utilized for recognizing characters fastly and precisely by the template matching.

In the on-line processing, the corresponding points are detected from the left image and the right image. The 3D positions of these points are measured through triangulation, and the 3D shape of the book surface is reconstructed. The two original images of curved surfaces are transformed to those of flat surfaces by maintaining the distance between points, and the clear regions of two images are combined. Finally, the characters on the transformed image are recognized. The overview of the on-line processing is shown in Figure 2.

2.2 3D Measurement

In the on-line processing, the 3D measurement of book's surface is carried out at first. The brightness of the left image and the right image differs from each other, because of the difference of the cameras' directions. Therefore, the detection of the stereo corresponding points is performed on gray scale stereo images by using the normalization cross correlation (NCC) method that is a kind of the template matching. This is because this method is robust against the change of brightness and it does not need thorny thresholding procedure.

The template is cut off from the left image and the region of the right image similar to the template is searched with the NCC. The size and the search area vary adaptively according to the correlation value to reduce the computation time. In the case, however, that characters and figures do not exist in that area, the NCC does not work well. Therefore, the corresponding points in these areas are excluded from the 3D measurement.

2.3 Shape Reconstruction

The results of the 3D measurement involve the detection error of the corresponding points, and only discrete points on the book surface are measured. Therefore, we interpolate the region where corresponding points can not be detected, and express the shape of the book surface as an approximate expression.

We define z coordinate value of the curve surface as the function of (x, y). At first, $z_{y=y_i}$ (z value in the case of $y = y_i$) is expressed as polynomial of x from all measured corresponding points whose y-coordinates are near y_i.

$$\begin{aligned} z_{y=y_i} &= f_i(x) \\ &= \sum_{j=0}^{n_i} a_{i,j} x^j, \end{aligned} \tag{1}$$

where $a_{i,j}$ and n_i are a coefficient and a degree of the polynomial respectively in the case of $y = y_i$.

The corresponding points whose distances from the calculated polynomial are larger than a given threshold value are regarded as the outliers, and the polynomial excluding these outliers is recalculated. The polynomials $f_i(x)$ $(i = 0, \ldots, m)$ are generated for left and right pages at regular intervals of y-coordinates. After that, a curve surface z $(y_i < y \leq y_{i+1})$ is expressed as follows:

$$\begin{aligned} z &= g(x, y) \\ &= \frac{1}{y_{i+1} - y_i} \Big\{ (y_{i+1,j} - y) \sum_{j=0}^{n_i} a_{i,j} x^j \\ &\quad + (y - y_i) \sum_{j=0}^{n_{i+1}} a_{i+1,j} x^j \Big\}, \end{aligned} \tag{2}$$

To ensure the continuity and smoothness of the 3D reconstructed surface, the fitting coordinate value $(x, y, g(x, y))$ is used as substitute for the measured coordinate value (x, y, z).

The pixel values at corresponding points on the surface of the reconstructed book can obtained from the pixel values of the original images. The pixel values at arbitrary $(x, y, g(x, y))$ are interpolated from the coordinate values of neighborhood corresponding points and the pixel values of the original images.

2.4 Image Transformation

The curvature correction is executed in such a way that the images of curved surfaces are transformed to

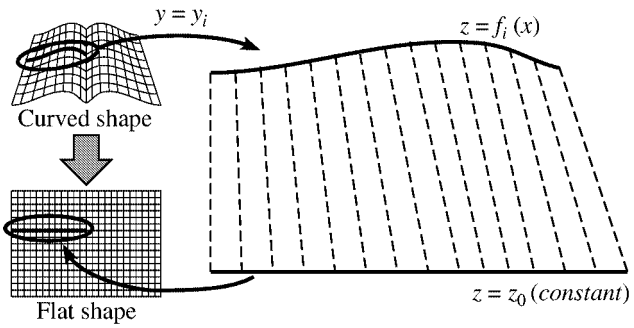

Figure 3: Recoverd 3D shape.

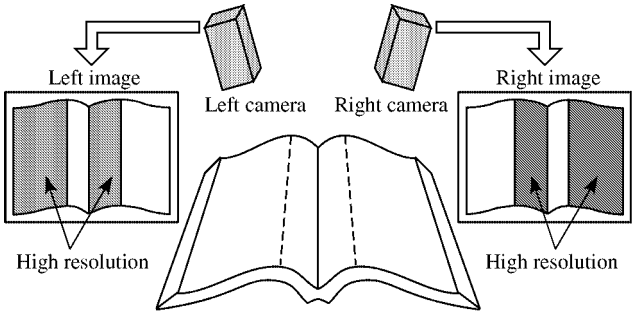

Figure 4: Combination of two images.

those of surfaces by maintaining the distance between points.

The approximating curve $z = f_i(x)$ is transformed to the straight line $z = z_0$ where z_0 is a constant value. The distance between two points fitted on an approximating curve $f_i(x)$ are precisely kept (Figure 3). After all approximating curves are transformed to the straight lines, the points that are not on $f_i(x)$ are transformed by linear interpolation. This curvature correction process is applied to the left and right images, and two transformed images are obtained.

The image in which a book surface is flat is created by combining the clear regions of two images where the resolution is higher. The resolution of the image becomes high when the directions of normal vectors of the surface are close to be parallel to the optic axis of the camera. Therefore, the region whose resolution is higher than that of other camera is combined (Figure 4). In our method, one page is divided into two region by the line whose z coordinate value is the largest.

2.5 Character Recognition

The character segmentation and the character recognition are executed automatically.

In the process of the character segmentation, the areas where only characters exist are extracted because processing targets in this study are characters. This is done by using the characteristics that English texts are printed as horizontal writing. Next, the centerlines along character strings are detected and the characters are segmented along the centerlines. The character strings are divided into regions each of which contains only one character by considering the width and the height of character templates that are generated in the off-line processing.

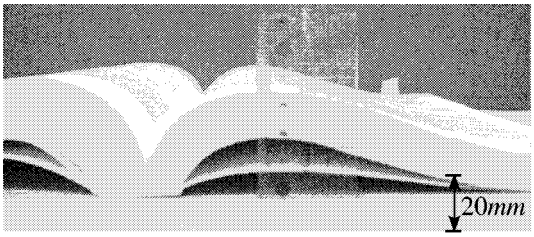

Figure 5: A book with curved surface.

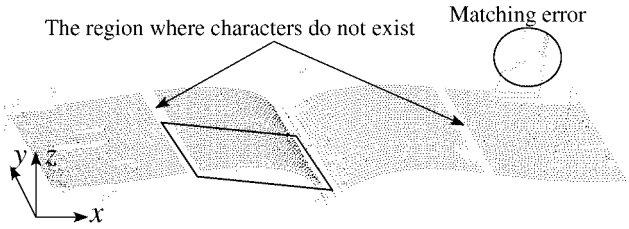

Figure 6: Result of 3D measurement.

After that, the template matching between the transformed gray scale image and the character templates is carried out to recognize characters. The NCC method is adopted for template matching because of the robustness against the change of brightness.

Characters of thick books can be recognized through the above procedure.

3 Experiments

We used a book of conference proceedings for our experiments (Figure 5). Experimental results are shown in the following.

An example of a result that indicates the 3D measurement of the book surface is shown in Figure 6. The corresponding points can be detected in most areas except in the midst area of each page because the text of this book is in two columns. Figure 7 is a result of the 3D shape reconstruction that is indicated as a quadrangle area in Figure 6. In most cases, the maximum error of the reconstructed shape is within 1.0mm.

Figure 8(a) and (b) show an acquired image of stereo pair, and Figure 8(c) shows a result of a combined image after transformation. From these results, it can be said that the shape of characters is transformed to be suitable for character recognition.

In order to verify the validity of the proposed method quantitatively, the comparison with some OCR software that is available commercially was made. Table 1 shows the recognition error rate in each case.

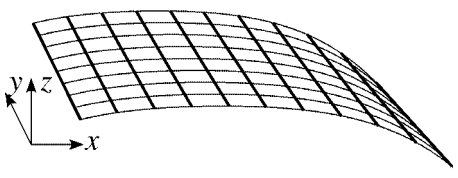

Figure 7: Result of 3D shape reconstruction.

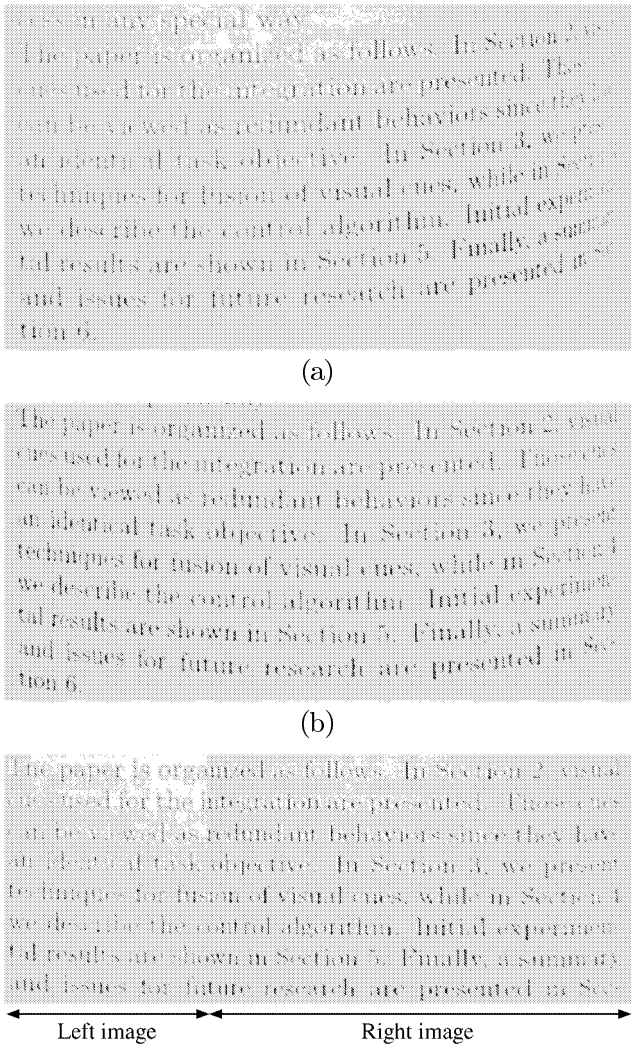

Figure 8: Result of image transformation. (a) Left image before transformation. (b) Right image before transformation (c) Combined image after transformation.

The recognition area is limited only the endmost left and right words where the curvature of the surface is very large. Case 1 uses the OCR software with the left image shown in Figure 8(a). Case 2 uses the OCR with right image shown in Figure 8(b). Case 3 uses our proposed method with the transformed image shown in Figure 8(c). Comparing Case 1 with 3, the recognition error rate reduces greatly.

Table 2 shows the result of the average recognition error rate in the case of curved books like Figure 5 in the whole area. Strategy A uses the OCR software with the original images (without the transformation that correct curved surfaces), Strategy B uses the OCR with the transformed image, Strategy C uses the NCC method for character template matching with the original images, and Strategy D uses the NCC with the transformation (our proposed method). By comparing Strategy A with B, and C with D, we can confirm the effect of the transformation. By comparing A with C, and B with D, we can recognize the advantage of the NCC, too. By comparing A with D, it can be verified that the recognition error rate reduces greatly and

Table 1: Error rate of character recognition (curved area).

Case	Image	Error rate
1	Left image	60.8 %
2	Right image	47.3 %
3	Transformed image	16.2 %

Table 2: Average error rate of character recognition (whole area).

Strategy	Process	Error rate
A	OCR + original images	16.3 %
B	OCR + transformed image	12.6 %
C	NCC + original images	15.8 %
D	NCC + transformed image	7.0 %

that the combination of the NCC and the transformation is very effective for character recognition of thick books. From these experimental results, the validity of our proposed method has been shown.

4 Conclusions

In this paper, we propose a character recognition method for thick books with curved surfaces using a stereo vision system. The 3D measurement of the book's surface is executed through triangulation, and the 3D shape of the surface is reconstructed. The curvature correction is done and the flat surface is recovered to improve the recognition rate by maintaining the distance between points. The validity of the proposed method is shown through experiments.

In these experiments, the main reasons of character recognition error are the remained inaccuracy of the 3D reconstruction and of the character miss-extraction. Therefore, as the future works, the accuracy of the camera calibration must be improved and the character template construction must be done more precisely.

References

[1] Toshikazu Wada, Hiroyuki Ukida and Takashi Matsuyama: "Shape from Shading with Interreflections Under a Proximal Light Source: Distortion-Free Copying of an Unfolded Book," *International Journal of Computer Vision*, Vol.24, No.2, pp.125–135, 1997.

[2] Hiroyuki Ukida, Katsunobu Konisho, Toshikazu Wada and Takashi Matsuyama: "Recovering Shape of Unfolded Book Surface from a Scanner Image using Eigenspace Method," *Proceedings of IAPR Workshop on Machine Vision Applications* (*MVA* 2000), pp.463–466, 2000.

[3] Atsushi Kawarago, Atsushi Yamashita, Toru Kaneko and Kenjiro T. Miura: "Character Recognition for Books with Curved Surface using a Stereo Vision System," *The Journal of The Institute of Image Information and Television Engineers*, Vol.56, No.5, pp.858–862, 2002 (in Japanese).