Shape Reconstruction and Image Restoration for Non-Flat Surfaces of Documents with a Stereo Vision System

Atsushi Yamashita Atsushi Kawarago Toru Kaneko Kenjiro T. Miura
Department of Mechanical Engineering, Shizuoka University
yamashita@ieee.org

Abstract

In this paper, we propose a shape reconstruction and image restoration method for paper documents with curved surfaces or fold lines by 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. Therefore, 3-D shape reconstruction of the book's surface is executed from the result of the stereoscopic measurement by putting the book upward, and an image of a flat surface is recovered from the curved or folded surface. The validity of the proposed method is shown through experiments.

1 Introduction

The digitization of documents 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, the midst areas become curved shapes and page surfaces cannot be in contact completely when scanning thick books. 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. The same is true of non-flat documents such as a sheet of paper with folds.

In order to solve this problem, SFS (Shape-from-shading) method that can recover the shape of unfolded book's surface from scanner images is proposed [1]. This method assumes that the book surface is cylindrical and the cross section shape is uniform. Therefore, the reconstructed shape is the 2-D 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. Moreover, in the case of rare books [2], there is a possibility of damaging them when they are pushed on a scanner by force.

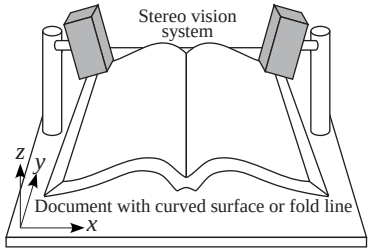


Figure 1. Overview of our system.

Furthermore, there is an another problem that it will take time to digitalize numbers of pages if a book is placed downward and scan it. Therefore, a shape recovery method of book surface using two shade images under the condition that a camera is placed above books [3] is proposed. Distortion recovery methods with a camera and a laser range finder [4] or a light projector [5, 6], and methods that need no auxiliary device [7, 8] are also proposed. However, these methods only use one camera and an image is acquired from a certain direction. Therefore, acquired images are distorted very much in the region where the angle between the optic axis of the camera and the normal vector of the book surface is large. If they are transformed into the plane surface images, the resolution of images is not high.

Therefore, we propose a new shape reconstruction and image restoration method based on the 3-D measurement by using a stereo vision system (Figure 1). A new image that has no distortion and high resolution can be generated by correcting and merging two images from two cameras whose directions of optic axes are different from each other.

2 Distortion recovery method

The procedure of distortion recovery method is divided into three steps; 1) 3-D measurement, 2) shape reconstruction, and 3) image restoration. At first, corresponding points are detected from the left and right images, and the 3-D positions of these points are measured through triangula-

tion (**3-D measurement**). Next, the 3-D shape of the document's surface is reconstructed by using NURBS (Non-Uniform Rational B-Spline) curve representation (**shape reconstruction**). Finally, 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 (**image restoration**).

2.1 3-D measurement

The 3-D measurement of document's surface is carried out at first with a stereo vision system. Corresponding points between the left and right images can be detected in the almost region of document's surface, because there are thick texture on it.

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. This is because the NCC method is robust against the change of brightness and it does not need thorny thresholding procedure. The correlation value C of NCC is calculated as follow:

$$C = \frac{\sum_{j=1}^{N}\sum_{i=1}^{M}(I_l(i,j)-\mu_l)(I_r(i,j)-\mu_r)}{MN\sigma_l\sigma_r}, \qquad (1)$$

where $I_{r,l}(i,j)$ is the pixel value of the left and right gray-scale image at pixel (i,j), $\mu_{l,r}$ and $\sigma_{l,r}$ are average and standard deviation of pixel value of templates, and $M \times N$ is a template size, respectively.

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 method. The size and the search area vary adaptively according to the correlation value to reduce the computation time by using a coarse-to-fine search and the epipolar constraints. However, NCC does not work well in the case that characters and figures do not exist in that area. Therefore, the corresponding points in these areas are excluded from the 3-D measurement to avoid searching incorrect corresponding points and to omit redundant calculation at the same time. These areas can be detected by using $\sigma_{l,r}$ values, because these areas have a little change in texture. In addition, $\sigma_{l,r}$ will be always calculated when C is calculated. Therefore, there is no additional computation time.

After detecting corresponding points, the 3-D coordinate values of these points on document's surface is calculated through triangulation. The points whose distances from neighbors are very large are eliminated as outliers.

2.2 Shape reconstruction

The results of the 3-D measurement involve the detection error of the corresponding points even after eliminating outliers, and only discrete points on the document's surface are measured. Therefore, we interpolate the region where corresponding points can not be detected, and express the shape of the document's surface as an continuous expression.

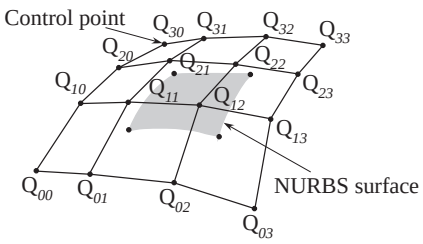


Figure 2. NURBS curve.

However, the gradient of document's surface in the binding area of books or the folded line of paper documents is discontinuous. Therefore, the points where the gradients are discontinuous are extracted as inflection points at first. Then, inflection lines are detected as the set of inflection points.

After searching inflection lines, NURBS curves [9] express the regions surrounded by the inflection lines. A NURBS curve is generated by setting control point Q_{ij}, weight w_{ij}, and knot vector $\mathbf{T}$ (Figure 2). The coordinate values of control points are decided from average of coordinate values of corresponding points in grids with certain sizes. Weights and knot vectors are decided according to the condition that edges of each NURBS curve are connected smoothly from each other.

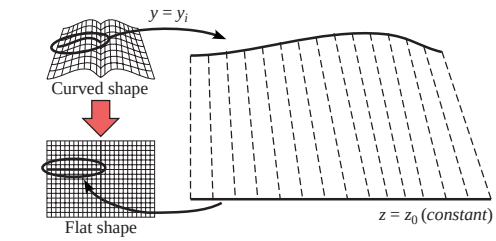
2.3 Image restoration

The images of curved surfaces are transformed to those of surfaces by maintaining the distance between points.

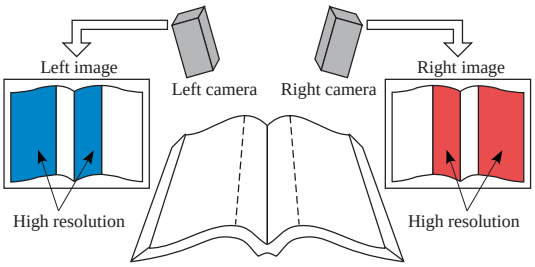
The curved line on the NURBS curve surface whose y coordinate value is y_i is transformed to the straight line $z = z_0$ along x-axis where z_0 is a constant value (Figure 3(a)). After one curved line is transformed to the straight one, the same procedure is repeated by changing the value of y_i in a certain discrete intervals. Finally, the curved lines whose y coordinate values are not y_i are transformed to the straight one by the linear interpolation of neighbor area.

The image in which document's 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 3(b)).

In this way, the textures on curved surfaces of paper documents are transformed to the flat ones.



(a) Distortion recovery.



(b) Combination of two images.

Figure 3. Image restoration.

3 Experiments

To verify the effectiveness of our proposed method, paper documents with curved surfaces and fold lines are used in the experiment. The resolution of each image is 2400 $\times$ 1800 pixel, and an image of two pages (A4 paper) can be gained by one shot.

An example of a result with a book of conference proceedings with curved surface is shown in Figure 4. Figure 4(a) and (b) indicate acquired images of stereo pair, and Figure 4(c) shows the result of 3-D measurement. 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 4(d) indicates the result of 3-D shape reconstruction. The inflection line can be detected between left and right pages, and two NURBS curves can describe the book's surface. Figure 4(e) shows an example of cross-sectional form of the book that is measured with a slide caliper. The difference between the actual shape and the measured shape is very small, and in most cases, the maximum error of the reconstructed shape is within 1mm.

Figure 5(a) and (b) show a part of acquired images of stereo pair, and Figure 5(c) shows a result of a combined image after image restoration.

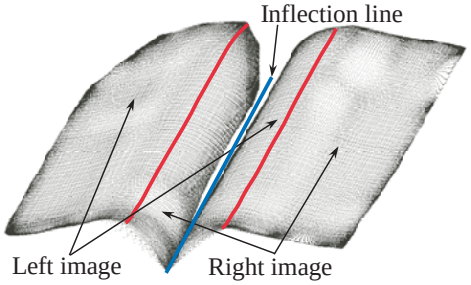
In order to verify the validity of the proposed method quantitatively, the comparison with some OCR software that is available commercially was made. Table 1 shows the recognition error rate in each image. Comparing the result of corrected images with those of the original images, the character recognition error rate reduces greatly.

Figure 6 shows the result of a paper with fold lines. From

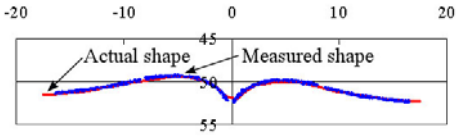
(a) Left image.

(b) Right image.

(c) Result of 3-D measurement.



(d) Result of shape reconstruction.



(e) Accuracy of 3-D measurement.

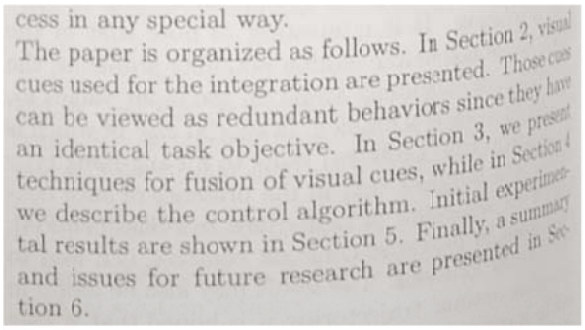
Figure 4. Result of book with curved surface.

this result, it is verified that an image restoration can be executed correctly even when there are fold lines on the images.

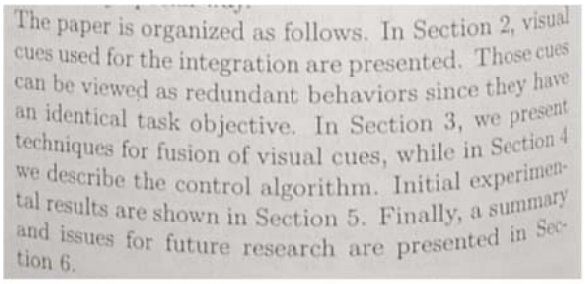
About the computation speed, the great portion of total computation time is spent on the detection of corresponding points. In our system, more than 200 corresponding points can be detected in 1 second.

4 Conclusions

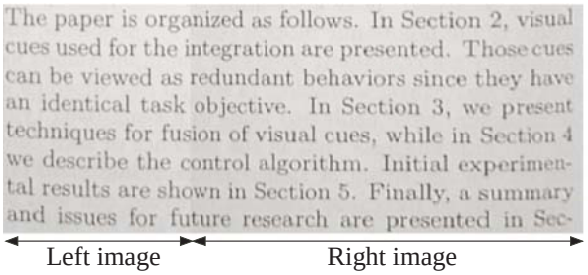
In this paper, we propose a distortion recovery method for paper documents with curved surfaces or folding lines by using a stereo vision system. The 3-D measurement of the paper surface is executed through triangulation, and the 3-D shape of the surface is reconstructed with the NURBS



(a) Left image before transformation.



(b) Right image before transformation.



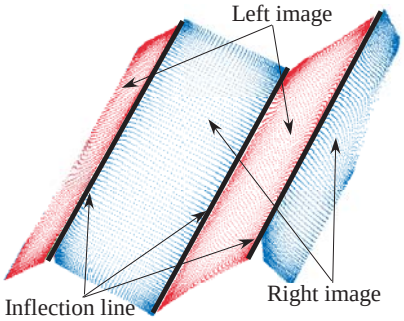
(c) Combined image after transformation.

Figure 5. Result of image restoration.

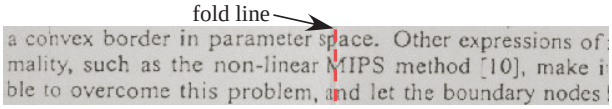
curves. The curvature correction is done and the flat surface is recovered to improve the recognition rate by maintaining the distance between points. The validity of the proposed method is shown through experiments.

Table 1. Character recognition error rate.

Image	Recognition region	Error rate
Left image	Whole area	22.1%
	Curved area	60.8%
Right image	Whole area	27.4%
	Curved area	57.3%
Corrected image	Whole area	0.3%
	Curved area	1.4%



(a) Result of shape reconstruction.



(b) Result of image restoration.

Figure 6. Result of a paper with fold lines.

References

[1] T. Wada, H. Ukida and T. 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] M. Kashimura, T. Nakajima, N. Onda, H. Saito and S. Ozawa: "Practical Introduction of Image Processing Technology to Digital Archiving of Rare Books," *Proceedings of the 9th International Conference on Signal Processing Applications and Technology*, pp.1025–1029, 1998.

[3] S. I. Cho, H. Saito and S. Ozawa: "Shape Recovery of Book Surface Using Two Shade Images Under Perspective Condition," *Transactions of the Institute of Electrical Engineers of Japan*, Vol.117-C, No.10, pp.1384–1390, 1997.

[4] T. Amano, T. Abe, T. Iyoda, O. Nishikawa and Y. Sato: "Development of The Next Generation Document Reader –Eye Scanner–," *Proceedings of the SPIE*, Vol.4669, pp.250–258, 2002.

[5] M. Pilu: "Undoing Paper Curl Distortion Using Applicable Surfaces," *Proceedings of the 8th IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp.67–72, 2001.

[6] M. S. Brown and W. B. Seales: "Document Restoration Using 3D Shape: A General Deskewing Algorithm for Arbitrarily Warped Documents," *Proceedings of 8th IEEE International Conference on Computer Vision*, pp.367–375, 2001.

[7] H. Cao, X. Ding and C. Liu: "A Cylindrical Surface Model to Rectify the Bound Document Image," *Proceedings of 9th IEEE International Conference on Computer Vision*, pp.228–233, 2003.

[8] M. S. Brown and Y.-C. Tsoi: "Unidistorting Imaged Print Materials using Boundary Information," *Proceedings of the 6th Asian Conference on Computer Vision*, pp.551–556, 2004.

[9] L. Piegl and W. Tiller: The NURBS Book 2nd Edition, Springer-Verlag, New York, 1997.