Correction of Geometric and Photometric Distortion of Document Images Using a Stereo Camera System

Yusuke Suzuki[1], Atsushi Yamashita[1,2], and Toru Kaneko[1]
1 Department of Mechanical Engineering, Shizuoka University
3-5-1 Johoku, Hamamatsu-shi, Shizuoka 432-8561, Japan
2 Department of Mechanical Engineering, California Institute of Technology
{f0630042, tayamas, tmtkane}@ipc.shizuoka.ac.jp

Abstract

In this paper, we propose a new method for correcting distorted document images using a stereo camera pair. Geometric and photometric distortions may occur in document images of thick book pages because of its curved surfaces. The proposed method acquires document shape by stereo measurement to correct geometric distortions in images. By considering normal vectors of document surfaces, optimum-resolution images are generated as a combination of a stereo image pair. Moreover, the method corrects photometric distortions by analyzing lighting and reflecting conditions. The validity of the proposed method was shown through experiments.

1. Introduction

In recent years, digitization of existing documents such as books is performed actively, and flatbed scanners are widely used to acquire book page images. However, space between the spine of the book and a scanner creates shadings and distortions in cases of thick books. This problem cannot be always solved by pressing the book on a scanner. In addition, from the viewpoint of document preservation, pressing is unfavorable. A method which considers three-dimensional shape of documents is needed to solve the problem.

In order to solve the above problem, methods which use *Shape from Shading* are proposed to restore shape of documents [1] [2]. These are simple and easy methods with a flatbed scanner. However, a lot of standard data have to be prepared, for example, illumination conditions, and reflection properties of documents.

Alternative systems using a camera instead of a flatbed scanner are proposed as follows.

A method using a camera and *Shape from Shading* is proposed [3]. This method acquires document images with a camera installed above the document. By using a camera, damage of documents is avoided, and shading is reduced. However, standard data such as reflection properties are necessary as same in [1].

Another distortion correction method using a camera is proposed [4]. In this method, shadow information is not needed. This method calculates three-dimensional shapes of documents from calibrated system parameters, and corrects distortions using three-dimensional geometric models of documents. However, it cannot fill in missing intensity information due to projection.

A method to perform video mosaicing of the curved document surface is proposed [5]. This method restores document shapes by *Structure from Motion* using a mobile camera. After reconstruction, this method corrects distortions, and makes high-resolution video mosaic images. It is a simple and easy method that does not need a special device, however, improvement of the accuracy of three-dimensional reconstruction around the binding area is desired.

A method with a stereo camera system is proposed [6]. This method acquires document shapes by stereo measurement, and corrects distortions. In addition, this method can make an optimum-resolution image by combination of stereo images. However, there are differences of the brightnesses between stereo images, and combined images are not natural. Consideration of shading is also needed.

In this paper, we propose a document digitization method using a stereo vision system which provides three-dimensional measurement of document surfaces. Distortions are corrected by using the three-dimensional measurement result. In addition, considering geometric relations between cameras and document surfaces, we make an optimum-resolution image from the stereo image pair. Furthermore, specular reflection is considered. By combination of images acquired in different lighting conditions, highlight is removed. Afterwards, shading correction is carried out considering properties of light reflection. The problem that the brightnesses of images are different between stereo images is solved by this process.

2. Method Outline

An overview of our system is shown in Figure 1. Two cameras are set above the document, and two light sources are set in left and right of the document. The three-dimensional coordinate system is defined as shown in Figure 1. Geometric relations of cameras and light sources are calibrated before image acquisitions. The reflectance of the document surface is assumed to be constant in whole pages.

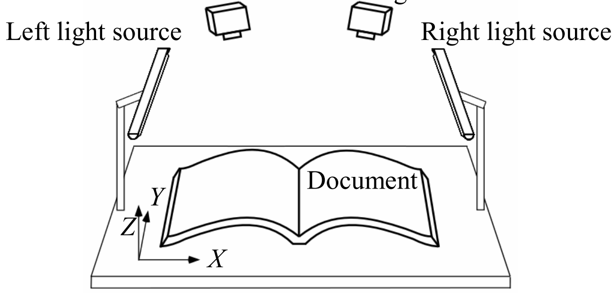


Figure 1. Overview of our system.

First, the document shape is acquired by stereo measurement, and geometric distortions are corrected. Next, the optimum-resolution image is generated by combination of the stereo image pair. These processes are based on our previous work [6]. Highlight detection is also executed. If highlight is detected, it will be removed by image combination. Finally, by shading correction, differences of the brightnesses between stereo image pair are corrected.

3. Geometric Distortion Correction

3.1. Three-dimensional measurement

The document shape is acquired by stereo measurement. We use two cameras whose directions of optic axes are converging at the document surface. By this configuration, even if one image is distorted drastically, the other image may not be distorted very much. Therefore, we can get clear texture about the whole document surface.

3.2. Shape reconstruction

3.2.1. Detection of inflection points

We approximate shapes of documents to NURBS curves. Inflection points have to be detected before modeling because curved surfaces can not be connected smoothly at inflection points. We detect inflection points from stereo measurement results. Then, the inflection lines of document (the binding area of books or the folding line of paper documents) will be detected by performing the Hough transformation of inflection points projected on the X-Y plane.

3.2.2. Approximation by NURBS curve

After inflection line detection, the document shape is expressed by NURBS curves. A NURBS curve is generated by setting control points Q_{ij}, control point weights, and knot vectors (Figure 2). The three-dimensional coordinate values of control points are determined from the average of three-dimensional coordinate values of measured points in grids of a certain size. Weights and knot vectors are determined according to the condition that edges of each NURBS curve are connected smoothly to each other.

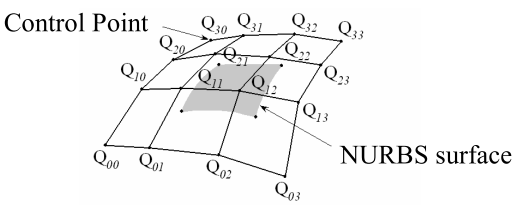


Figure 2. NURBS curve.

3.3. Geometric distortion correction

Based on the reconstructed shape by NURBS curve, geometric distortion correction of the document image is performed. In this paper, we consider the section that is perpendicular to the inflection line (Figure 3), and perform geometric distortion correction of the document image by flattening a curved surface on a plane.

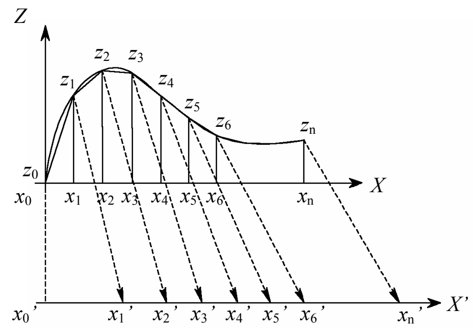


Figure 3. Geometric distortion correction.

In Figure 3, the curved line between i and i+1 is approximated to a straight line. The x coordinate value after correction is calculated by Equation (1).

$$x_i' = \sum_{j=0}^{i-1} \sqrt{\left(z_{j+1} - z_j\right)^2 + \left(x_{j+1} - x_j\right)^2} \qquad (1)$$

By fitting an image pixel of position x to position x', geometric distortions are corrected. To correct whole geometric distortions of the document image, the same procedure is repeated by changing the y coordinate value.

4. Combination of Images

We use two cameras whose directions of optic axes are converging at the document surface. The resolution of the image becomes optimum when the directions of normal vectors of the surface are close to be parallel to the optic axis of the camera. Therefore, the regions whose resolutions are higher than those given by other camera are combined (Figure 4).

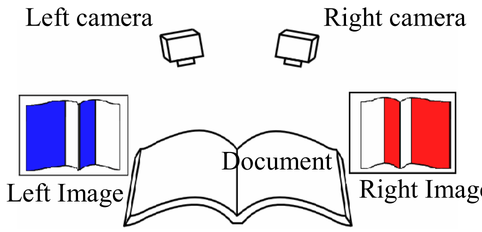


Figure 4. Making of optimum-resolution image.

5. Photometric Distortion Correction

5.1. Removal of highlight

In removal of highlight, directions of specular reflection vector **R** and cameras' eye vectors **V** are considered. By the reflection model of Phong, the specular reflection light intensity $I_h(x,y,z)$ is calculated by Equation (2),

$$I_h(x,y,z) = I_{in}(x,y,z) \cdot k_s \cdot \cos^n \psi \qquad (2)$$

where $I_{in}(x,y,z)$ is the intensity of the incident light to point P(x,y,z) on the document, k_s is the specular reflectance of the surface, ψ is the angle that **R** and **V** make, and n is a power which models the specular reflected light. Parameters k_s and n are determined from properties of materials.

In Equation (2), as ψ increases, I_h decreases. Therefore we define a certain threshold angle ψ_0, and detect highlight by using the condition $\psi<\psi_0$.

After the highlight detection, by combination of images that is acquired by alternating lighting of the left and right light source, highlight is removed.

5.2. Shading calculation

The model of light reflection is shown in Figure 5. The illumination is given by a line source, and diffusion reflection is considered in this process. Light intensity $I(x,y,z)$ at point $P(x,y,z)$ is calculated by Equation (3),

$$I(x,y,z)=k_d I_q \int_0^L \frac{\cos\theta}{r^2}\cos\varphi \; dl \qquad (3)$$

where k_d is the diffusion reflectance of the document surface, I_q is the intensity of the line source, L is the length of the line source, dl is the infinitesimal length of divided line source, θ is the angle that is made by the ray vector and the normal vector of the line source, r is the distance from the line source to point P, and φ is the angle that is made by the ray vector and the normal vector of the document surface.

The reflectance of the document surface is assumed to be constant in one page. The intensity of the line source is also assumed to be constant. Therefore, in Equation (3), k_d and I_q are constant. Only the integral part changes with positions and shapes of documents. As the integral part changes, the reflection light intensity changes according to Equation (3), which causes shading.

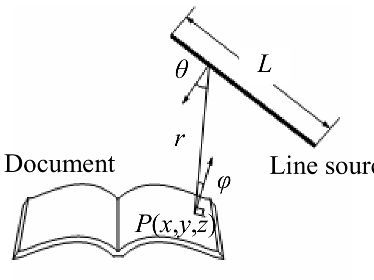


Figure 5. Model of light reflection.

5.3. Image correction

By Equation (3), the reflection light intensity of the document surface is calculated, and shading is corrected as follows. At first, we calculate the reflection light intensity $I(x,y,z)$ at each pixel (i,j) of the document image. The relations between (i,j) and (x,y,z) were calibrated beforehand. Next, shading is expressed by calculating the ratio $I(i,j)/I_{max}$, where I_{max} is the maximum intensity. Afterwards, shading is corrected by Equation (4),

$$C'_k(i,j)=C_k(i,j)\times\frac{I_{\max}}{I(i,j)} \qquad (4)$$

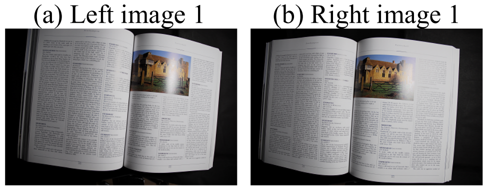
where $C_k(i, j)$ is the pixel value of the document image (k=1:red, k=2:green, k=3:blue). By Equation (4), pixel value $C_k(i,j)$ becomes larger in shading region, and shading is corrected.

6. Experiment

Two stereo image pairs acquired with digital camera are shown in Figure 6 (1936*1296pixels). In the experiment, one stereo image pair is mainly processed and the other stereo image pair is used for filling up the highlight.

The stereo measurement result is shown in Figure 7, where measured points are expressed in black points. After the stereo measurement, inflection points are detected from the measurement result.

(a) Left image 1 (b) Right image 1



(c) Left image 2 (b) Right image 2

Figure 6. Acquired image.

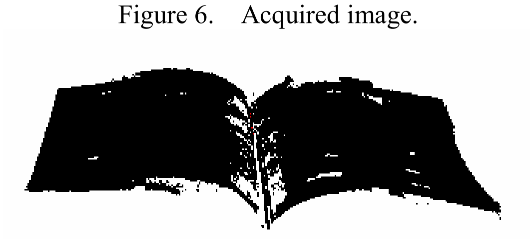
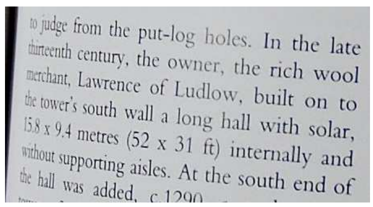


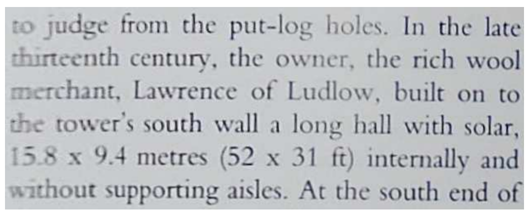
Figure 7. Result of stereo measurement.

The geometric distortion correction result is shown in Figure 8. Curved character strings are corrected to be straight by flattening the surface on a plane (Figure 8(b)). In addition, the resolution of flattened image is optimized by combining left and right images (Figure 8(c)).



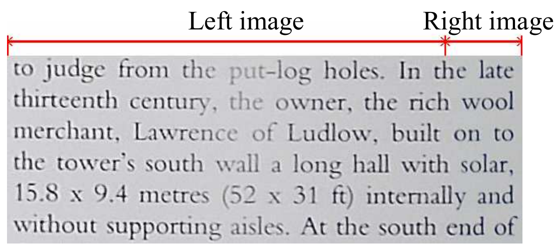
to judge from the put-log holes. In the late thirteenth century, the owner, the rich wool merchant, Lawrence of Ludlow, built on to the tower's south wall a long hall with solar, 15.8 x 9.4 metres (52 x 31 ft) internally and without supporting aisles. At the south end of the hall was added, c.1290

(a) Before correction (right image).



to judge from the put-log holes. In the late thirteenth century, the owner, the rich wool merchant, Lawrence of Ludlow, built on to the tower's south wall a long hall with solar, 15.8 x 9.4 metres (52 x 31 ft) internally and without supporting aisles. At the south end of

(b) After correction (right image).



to judge from the put-log holes. In the late thirteenth century, the owner, the rich wool merchant, Lawrence of Ludlow, built on to the tower's south wall a long hall with solar, 15.8 x 9.4 metres (52 x 31 ft) internally and without supporting aisles. At the south end of

(c) After combination.

Figure 8. Result of distortion correction.

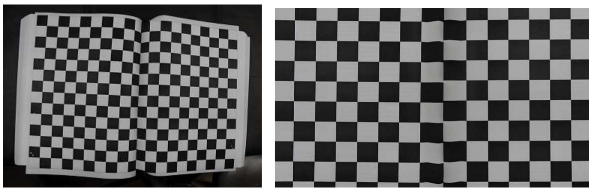
After geometric distortion correction, we execute photometric distortion correction. Figure 9 shows a result of highlight removal. Figure 9(a) and Figure 9(b) are distortion-free images of “Left image 1” and “Left image 2” in Fig. 6, respectively. By combination of these images, highlight is removed as shown in Figure 9(c).

Shading is also corrected by calculating lighting and reflecting conditions on the document surface.

(a) Image 1. (b) Image 2. (c) Image 1+2.

Figure 9. Result of highlight removal.

To verify the validity of the geometric distortion correction, an experiment result with a grid pattern image was given (Figure 10). Geometric distortions were corrected, and the grid pattern image was corrected to fit to a flat image. The comparison of actual values and measured values is shown in Table 1, where measured values are the average of the measurement result of the grid pattern.



(a)Before correction. (b)After correction.

Figure 10. The result of a grid pattern image.

Table 1. The quantitative evaluation with a grid pattern.

	Actual value	Measured value	Error rate
Grid pitch	20.0mm	20.7mm	3.5%
Intersection angle	90.0deg	89.7deg	0.37%

The comparison with a commercially available OCR software was made. The recognition rates are shown in Table 2, where (a), (b), (c) are corresponding to each image of Figure 8. Comparing the results of corrected images with the original image, the character recognition rate rises greatly. The result for “Flat image” is also shown in Table 2. It is the image of the same region captured in a flat state. The recognition rates of “Final result” and “Flat image” are almost equal. From these results, the validity of the proposed method was verified.

Table 2. Character recognition rate (in curved area).

Image	Recognition rate
(a) Before correction	39.6%
(b) After correction	82.2%
(c) Final result	97.0%
(d) Flat image	97.0%

The final result is shown in Figure. 11. Geometric and photometric distortions were corrected, and optimum-resolution image was generated.

Figure 11. Final result.

7. Conclusion

We proposed a method for correcting geometric distortions of document images based on stereo measurement results. The method also corrects shading and replacing low-resolution regions and highlight regions by image combination.

As a future work, we extend this system to handle documents which have more complicated shapes.

Acknowledgments

This research was partially supported by the Ministry of Education, Culture, Sports, Science and Technology, Grant-in-Aid for Young Scientists (B), 17700182, Japan.

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] Chew Lim Tan, Li Zhang, Zheng Zhang, and Tao Xia: “Restoring Warped Document Images through 3D Shape Modeling”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.28, No.2, pp.195-208, 2006.

[3] Seong Ik Cho, Hideo Saito and Ozawa Shinji: “Shape Recovery of Book Surface Using Two Shade Images Under Perspective Condition”, *Proceedings of The Institute of Electrical Engineers of Japan* C, Vol.117-C, No.10, pp.1384-1390, 1997.

[4] Michael S. Brown and W. Brent Seales: “Image Restoration of Arbitrarily Warped Documents”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.26, No.10, pp.1295-1306, 2004.

[5] Tomokazu Sato, Akihiko Iketani, Sei Ikeda, Masayuki Kanbara, Noboru Nakajima, and Naokazu Yokoya: “Mobile Video Mosaicing System for Flat and Curved Documents”, *Proceedings of the 1st International Workshop on Mobile Vision*,, pp.78-92, 2006

[6] Atsushi Yamashita, Atsushi Kawarago, Toru Kaneko, and Kenjiro T. Miura: “Shape Reconstruction and Image Restoration for Non-Flat Surfaces of Documents with a Stereo Vision System”, *Proceedings of the 17th International Conference on Pattern Recognition*, Vol.1, pp.482-485, 2004.