

COMPENSATION OF PARTLY PHOTOGRAPHED PAGE-IMAGES USING 3-D SHAPE INFORMATION

Toshifumi Nakajima, Masaaki Kashimura, Shinji Ozawa

Faculty of Science and Technology
Department of Electrical Engineering, Keio University
3-14-1 Hiyoshi Kouhoku-ku Yokohama 223-8522, Japan
E-mail: {nakaji,kassy,ozawa}@ozawa.ics.keio.ac.jp

ABSTRACT

In this paper, we propose a method in order to compensate contortion of page image which is caused by photographing a book opened and placed face-up (two pages opposite each other state) on the stage just above. For compensating contortion 3-D shape information of page surface is necessary to apply some geometrical adjustment. To obtain much higher resolution image a page needs to be photographed partly, in this situation each neighboring image necessarily has overlap area for the purpose of joining together. we put characters exist in overlap area to good use for seeking of corresponding points in stereo method by which 3-D shape information(3-D coordinate of corresponding points)of book surface can be obtained. If the assumption that no change of the shape of page cross section along the binding are satisfied, satisfactory compensated images were obtained and joining them together was well performed.

1. INTRODUCTION

Recently by the progress of computer performance, cost down of electric/magnetic recording media and wide spread of networking environment, digital libraries that had been only concept till a few years ago begin working practically [1]. In order to build an electronic library, first necessity is obtaining digital page images of book which have been collected up to now there. Second one is digitally archiving their contents efficiently. Some examples are given as following, general books and kinds of magazine are enough to be archived as text base information, rare or historical books are desired to be archived as high resolution images. Because in the former the very information of articles is most important of all other information like which font used. Contrastively, in the latter information such as the shape of characters, illustrations, impressions of paper, etc., are more important than contents of articles, which is perhaps well-known Bible[2]. However, in both case, high resolution image is beneficial. It is recommended to recognize character with OCR at

400dpi resolution and desired to bear the research of bibliography in the humanities. Third one is how to obtain digital page images. Now there are two devices of taking digital image, scanner and digital camera. the former is comparatively common and suitable for general books. But for the reason of scarcity and fragility of rare books the latter will be an optimum technique at the present time, comparison with flatbed-scanner which force to push two pages opposite each other against the contact-window. Further as respects of time and cost digital camera is suitable. Accordingly, in order to obtain digital page images of book, a book placed on the stage(See Fig.1)with its unstrained posture and its pages is photographed just above. If much higher resolution images are required, it is necessary to magnify and partly photograph that join each partial image together. Images taken in this situation not form an exact rectangle but have swollen curved top and bottom edge lines and contents of the page have been contorted. This contortion cause by 3-D shape of page give bad influence with the character recognition and make it impossible to join partial images together. Thereupon, we have proposed a method for obtaining rectangular just-front view of page images from partly photographed images of some books page, using geometrical adjustment. Partial images have geometrical contortion caused by the 3-D shape of the page [3]. It is difficult to join them together without image adjustment processing. In [3] it is impossible to compensate an image excluding the case that at least top or bottom of edge of page exists in the image, in other words [3] uses information of curvature of page' top or bottom edge for image adjustment. In our research, we make better use of overlapped area which necessarily exists for the purpose of joining partial images together and pay attention to characters which exist in overlapped area of the neighboring two partial images. A method to obtain 3-D shape information for geometrical adjustment of partially photographed images is proposed.

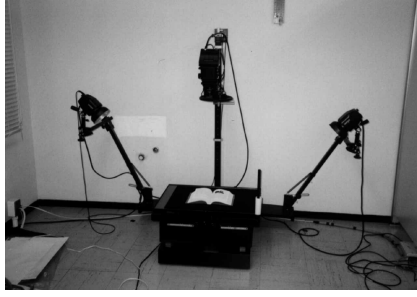


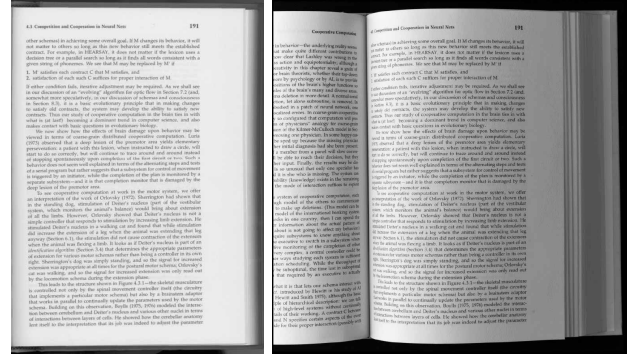
Figure 1: The image input device

2. PROBLEM OF CONTORTED IMAGE

The image input device showed in Fig.1 has a high resolution color CCD camera(1996 × 1996, 14bit/pixel RGB each)set on the central pillar and directed downward to the stage. Though obtained images are color, transformed 8bit gray level images are used in this research because of reduction of calculation time. The image plane (CCD plane) is adjusted parallel to the stage. (Note that images shown in this manuscript have smaller size and lower resolution because of the restriction caused by document file size, but full size images are used in our study in usual.) Synchronous flashlights are set to the top of slant pillars in the left side and right side. The photograph extent is adjusted by distance between camera and the stage. Photograph area can be easily set by parallel slide of the stage. Coordinates are defined as Fig.3(a). The origin of camera coordinate is located at the center of the lens. The origin of image coordinate is located on the Z axis. In general, surface of opened pages forms curved shape swelling from the central bounded valley-like part to side edges.

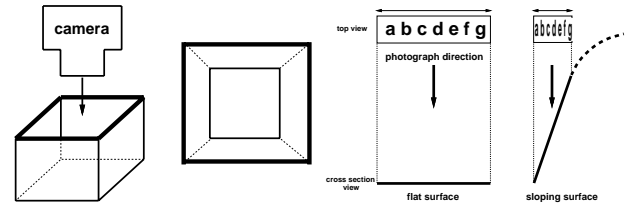
2.1. Cause of image curvature

An example image obtained with scanner is shown in Fig.2(a) and with camera in Fig.1 is shown in Fig.2(b). In Fig.2(b), page surface is not depicted as exact rectangular area but bend rectangular area which have swollen top and bottom edge lines. The degree of swelling of text line in the image has relevance to the 3-D shape of page, and found that two type of contortion exist. One is the curvature of text line along y direction and the other is the narrowness of characters width near the binding. The former is caused by the perspective transformation, when the object having 3-D shape is photographed with the camera, the nearer a part of object to camera, the more swell it become on image plane, in other words, photographing 3-D object with camera is point projection transformation to the image plane (See Fig.3(a)). the latter is caused by slope of binding part(See Fig.3(b)).



(a)scanner (b)digital camera

Figure 2: Example images



(a)perspective transformation (b)narrowness of character width

Figure 3: Causes of contortion

2.2. Perspective Transformation

As illustrated in Fig.4(a), The relation between 3-D point on the page surface and its projected corresponding point in the image is described as

$$x = f \frac{X}{Z}, \quad y = f \frac{Y}{Z} \quad (1)$$

These are the basics of perspective transformation and these mean that the 2-D coordinate (x, y) in image and 3-D coordinate (X, Y, Z) on the object is in proportion of focal length f and reciprocal of 3-D depth Z . Therefore we prepare depth axis z on image coordinates in order to treat the assumed 3-D height on image plane. After this, the adjustment method is discussed only on image plane using coordinates of (x, y, z) .

3. ADJUSTMENT METHOD FOR PARTLY PHOTOGRAPHED IMAGES BEFORE JOIN THEM TOGETHER

One of the solutions for requirement of high resolution image is that some partly photographed images are joined together in order to make an image of larger size. In this situation, the joint area of neighboring two partial images necessarily exists and 3-D shape of joint area can be obtained by applying stereo method to this area. Using this 3-D information to image adjustment, more precise compensation of swell of page surface near the joint area can be performed and result

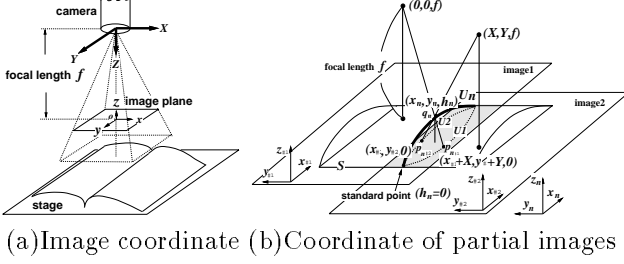


Figure 4: Definition of coordinates

image of joint process will have higher quality. When stereo method is applied to obtain 3-D shape information, corresponding points in two images of different view must be extracted. In our method, the pairs of corresponding points is extracted by matching of characters in overlapped area of neighboring two partial images. Fig.7(a), (b) show examples of partial images of a page. Though these images include top or bottom edge and number of partial images is only two, the condition is for the basic study of adjustment method. The stage in Fig.1 can be slid parallel with x and y axis, and partial images such as Fig.7 can easily be photographed. Magnified overlapped area in Fig.7(a), (b) is shown in Fig.8(a), (b) respectively. Page view of the bottom area of upper part image Fig.8(a) and the top area of lower part image Fig.8(b) swell in the opposite direction. Then the degree of swell of page surface can be calculated from page cross section obtained by applying stereo method. The flowchart of our method is shown in Fig.5. At first, rotational transformation with angle $-\theta$ is performed. By this transformation, the line on the central valley of opened pages, S in Fig.4(b) is arranged to be parallel to y axis. And characters in overlapped area are binarized and labeled along each text line in each partial image. Then the feature matching of characters in each line is performed between neighboring images and one-to-one correspondence of each character is obtained. This matching is performed line by line for all the lines in overlapped area. In this paper we assume that there is no change of the shape of page cross section along the binding direction, y direction. Each dot in Fig.9 is cross section of 3-D shape information of page surface calculated using a pair of location parameter of the center of corresponding characters in each partial image, In Fig.4(b) q_n is a point on page surface having 3-D coordinate and p_{n1} , p_{n2} are the center of corresponding characters. the height h_n and x coordinate x_n of q_n in Fig.4(b) are given by

$$h_n = f \times \left(1 - \frac{y_{20} - y_{10}}{y_2 - y_1} \right), \quad x_n = \frac{x_2}{y_2 - y_1} \times (y_{20} - y_{10}) \quad (2)$$

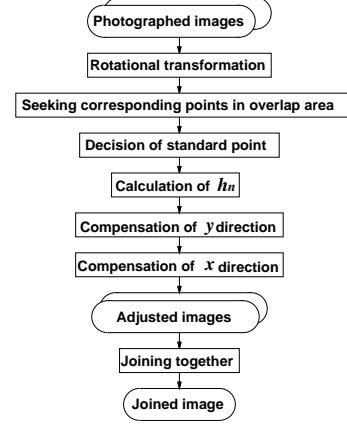


Figure 5: flowchart

In Fig.4(b) (x_1, y_1, z_1) are coordinates of p_{n1} in image1(upper part), (x_2, y_2, z_2) are coordinates of p_{n2} in image2(lower part) and $(x_{10}, y_{10}, 0)$, $(x_{20}, y_{20}, 0)$ are coordinates of the standard point in each image coordinate system. Once the standard point decided (X, Y) , coordinates between lens of different view, are easily obtained. The continuous shape of page cross section is estimated by 4D polynomial approximation. After the 3-D shape of page surface is obtained, geometrical compensation, pixel value remapping, is performed to obtain adjusted images ready to be joined together. The adjustment process consists of two parts. In the first part, y directional swelling area of page in the image caused by perspective transformation is adjusted and orthogonal projection image along z axis is obtained. Fig.6(a) is a overhead view of the page images and p_n , q_n is the same meaning in Fig.4(b), pixel sequence P_n , is remapping to pixel sequence Q_n , notice we assume that there is no change of page' height along y direction, so processing unit is pixel sequence. After processing against all pixel-sequence P_n orthogonal projection images is obtained. In the second part, the view if the 3-D bend shape of the page is unbent to flat shape is obtained by x directional geometrical transformation. Using the height h_n of each q_n , pixel sequence Q_n is horizontally shifted and remapped on appropriate position in the image of flat page area. Thick bend line in Fig.6(b) shows a example of cross section view of the page surface. In order to get the shift distance for Q_n , each horizontal remapping distance l_n between neighboring calculation points (q_n and q_{n+1}) are calculated by linear approximation, result value of each remapped pixel is calculated by linear approximation using values of nearest two points of object image. Finally, compensated images are joined together using information of corresponding points' coordinates between partial images.

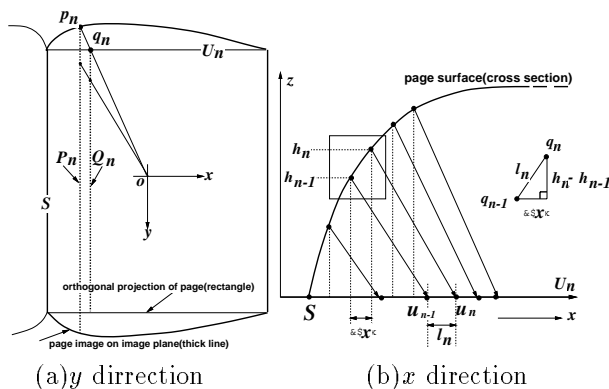


Figure 6: Illustration of Compensation

4. EXPERIMENTAL RESULTS

Partial page images of a contemporary book obtained by the device shown in Fig.1 are shown in Fig.7(a), (b). Magnified image pieces of the valley-like part in the page are also shown in Fig.8(a), (b). And calculated by stereo method and estimated curved line of page cross section is shown in Fig.9. Shape of characters are compensated well in the example of contemporary book (Fig.7(a), (b)), as the surface of the object book satisfies assumed condition in our method, good adjustment result could be obtained. Fig.10 is the result of the adjustment of areas shown in Fig.8. This image is made by blend processing two areas of image, and two areas seem to perfectly overlap each other. Fig.11 is the joined image of result of our adjustment method. In the result images, there is the shade on the left-hand side because we have not yet considered the lighting effect in our adjustment method. We are also working to develop a method to remove the shade in the results using 3-D shape information that is already obtained in first step in our present method.

5. CONCLUSION

In this paper, we presented an image adjustment method to remove contortion caused by perspective transformation etc., and showed results of the adjustment. In the case that curved line of page cross section have no change along the binding direction, satisfactory com-

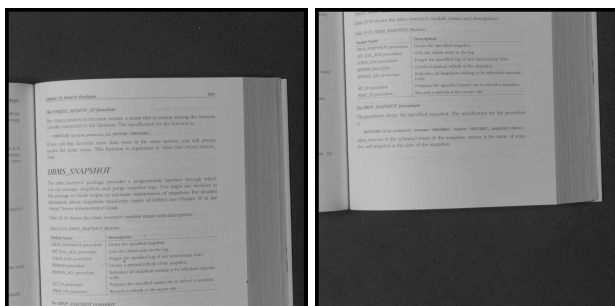


Figure 7: Partial page images

SET_UP procedure	Prepares the specified master site to refresh a
WRAP_UP procedure	Records a refresh at the master site.

(a)Upper part

SET_UP procedure	Prepares the specified master site to refresh a
WRAP_UP procedure	Records a refresh at the master site.

(b)Lower part

Figure 8: Overlapped area in Fig.7
pensated images were obtained and joining them together was well performed. After this, we are going to introduce pixel value conversion to reduce the shade in page image into this adjustment method.

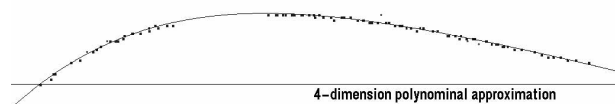


Figure 9: Page cross section

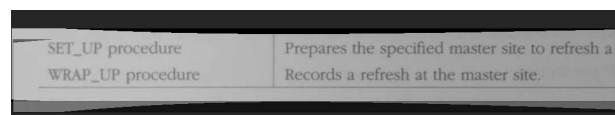


Figure 10: Result of adjustment

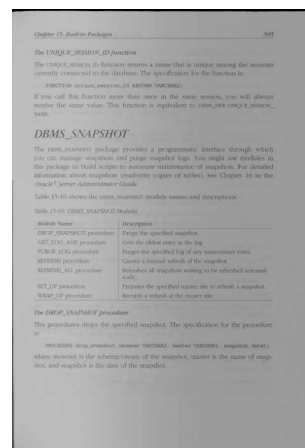


Figure 11: Result of joint

6. REFERENCES

- [1] K.Tabata: "Electronic Libraries : Their Prospect", TECHNICAL REPORT of IEICE,PRMU96-164, pp31-37(1997)(in Japanese)
- [2] M.Nagao: "Technologies for Digital Library". TECHNICAL REPORT of IEICE,PRUM96-165,pp39-44(1997)(in Japanese)
- [3] M.Kashimura,H.Saito,N.Onda,Y.Hua,S.Ozawa: "Application of Image Processing Technology to Digital Archiving of Rare Books". TECHNICAL REPORT of IEICE,PRMU97-77, pp41-48(1997)(in Japanese)