Distortion Correction for Digital Cameras

YUNG-CHIEH LIN AND CHIOU-SHANN FUH

Department of Computer Science and Information Engineering
National Taiwan University
Taipei, Taiwan

Abstract. This paper describes a method to correct lens distortion for digital cameras. A simple but effective distortion model using local spatial transformation will be used. In addition, the effect of effective focal length on the distortion model will be considered, and we will discuss how to introduce this factor to enhance the model. Finally we have some theoretical analysis and experiments demonstrating that the correction improves the accuracy of image stitching.

Keywords: Camera Calibration, Lens Distortion, Image Reconstruction, Fast Fourier Transform

1 Introduction

Picture is one of the elementary contents of digital library. Some of them are taken from high-resolution digital camera. In the project of digital library[1] in National Taiwan University, Danshin files have some manuscripts. We use a Kodak DCS 460 digital camera to take pictures of them in the size of 3060×2036 pixels. For higher resolution (DPI: dots per inch) and better quality, we do not put a large object in a single frame. Therefore we have to take multiple pictures of a large object and stitch them together. However, geometric distortion is the major problem of the optical components. Particularly on the edges of the picture, pixels are distorted severely and those are the most important data when stitching.

There are many previous works [1, 2, 3] introducing camera calibration, but they did not show how to correct the distortion if the effective focal length, the distance from image to lens center, is changed. Besides, they assume the distortion is radial and use a global transformation to correct the distortion. We think the method is not suitable because a camera usually does not only have a single piece of lens. The model will not be radial when the lens set is not mounted exactly at the same axis.

In this paper, we assume our platform is well calibrated, so that the image plane is parallel to the table. In addition, there is rectilinear white grid on the table for calibration as shown in Figure 1. Therefore the calibration lines can be identified in the background of the pictures as shown in Figure 1. These lines will be used to determined the image resolution.

[1]This research was supported by the National Science Council of Taiwan, R.O.C., under Grants NSC 85-2212-E-002-077 and NSC 86-2212-E-002-025, by Mechanical Industry Research Laboratories, Industrial Technology Research Institute, under Grant MIRL 873K67BN3, by the EeRise Corporation, ACME Systems, Tekom Technologies, and Foxconn Inc.

2 Distortion Model

After setting the camera in a fixed focal length, we take a picture of the table for calibration. Our distortion model simply considers each block in the grid to be transformed by Equation (1) as stated in [4].

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} c_1 & c_2 & c_3 & c_4 \\ c_5 & c_6 & c_7 & c_8 \end{pmatrix} \begin{pmatrix} 1 \\ x \\ y \\ xy \end{pmatrix} \quad (1)$$

where (u, v) is the coordinates of the distorted picture, $c_1, \cdots, c_8$ are the coefficients of the transformation matrix, and (x, y) is the coordinates of corrected picture. We can obtain the shapes of the distorted grid from the picture, and evenly distribute the blocks in the corrected picture. Then the transformation for each block can be solved by its four points. For the incomplete blocks at edges of pictures, we use the same transformation parameters in the nearest block.

The next problem is picking up the tie-points of the calibration grid. In our implementation, we use pattern-recognition technique to find crosses in the picture. After locating the tie-points, we connect neighbor points and the shape for each block is reconstructed.

3 Distortion Center

The commonly used lenses are spherical, so the light transmitted is symmetric to the lens center. When we compute the distortion center, we assume the distortion is radial, so that the model is symmetric to the distortion center. Notice that this radial assumption is not necessary for our distortion model. The distortion center will be used only when we enhance the model in the consideration of effective focal length. In Equation (2), Δx and Δy become zero when (x_c, y_c) approaches the distortion center, because any two displacements can be canceled during the integration when they are symmetric to the distortion center.

$$\Delta x = \int_{y_c - \frac{h}{2}}^{y_c + \frac{h}{2}} \int_{x_c - \frac{w}{2}}^{x_c + \frac{w}{2}} ((u(x,y) - u(x_c, y_c)) - (x - x_c))dxdy$$
$$\Delta y = \int_{y_c - \frac{h}{2}}^{y_c + \frac{h}{2}} \int_{x_c - \frac{w}{2}}^{x_c + \frac{w}{2}} ((v(x,y) - v(x_c, y_c)) - (y - y_c))dxdy \quad (2)$$

where (u, v) is the coordinates of the distorted picture, (x, y) is the coordinates of corrected picture, (x_c, y_c) is the candidate of the image center, w and h are the width and height of the region for calculation, and $(\Delta x, \Delta y)$ is the integration of the distortion displacements. First we choose the image center for candidate, then we use a mountain climbing strategy to approach the distortion center. The following pseudo code describes the algorithm.

> $(x_c, y_c) \leftarrow$ image center
> $w \leftarrow$(image width)/2
> $h \leftarrow$(image height)/2
> compute $\Delta x, \Delta y$
> repeat until $|\Delta x|$, $|\Delta y|$ are small enough
>
> > find proper $(\delta x, \delta y)$ that $(x_c+\delta x, y_c+\delta y)$ can reduce $|\Delta x|$ and $|\Delta y|$
> > $x_c \leftarrow x_c + \delta x$
> > $y_c \leftarrow y_c + \delta y$
> > $w \leftarrow$MIN(w,image width$-w$)
> > $h \leftarrow$MIN(h,image height$-h$)
> > compute $\Delta x, \Delta y$
>
> end repeat

Since it is possible that some local minimums occur in the points other than the distortion center, the algorithm will only succeed when the distortion center is not too far from the image center.

4 Resolution

We simply determine the resolution by the background information. According to the assumption described above, there are always some calibration lines identifiable in the background. We can find them by choosing the rows of darker pixels. After fast Fourier transformation (FFT) for each row, the coefficients have the first jump around the spatial frequency value. As shown in Figure 2, the first jump around the spatial frequency value is at 28 meaning there are 28 grids horizontally with extension. Therefore we can estimate the resolution. The width of the picture is about $3 \times 28/2.54 \approx 33$ inch because the length of each square is 3 cm. According to the camera specification, there are 3060 pixels in each row, and the resolution should be $3060/33 \approx 92$ DPI.

5 Distortion Correction

Equation (3) expresses the relation between the effective focal length and the distance from the lens center to the object.

$$d^{-1} + e^{-1} = f^{-1} \quad (3)$$

where d is the distance from object to lens center, e is the effective focal length, i.e. the distance from image to lens center, and f is the focal length. Equation (4) expresses the relation between two projections (Figure 3).

$$\begin{aligned} {e_1}^{-1} + {d_1}^{-1} &= {e_2}^{-1} + {d_2}^{-1} \\ h_1/d_1 &= x_1/e_1 \\ h_2/d_2 &= x_1/e_2 \\ x_2/x_1 &= e_1/e_2 \end{aligned} \quad (4)$$

where e_i is the effective focal length, d_i is the object distance, h_i is the actual object size, x_1 is the size of camera sensor, and x_1/x_2 expresses the ratio between the distortion models. After the arrangement of Equation (4), we can obtain the ratio by Equation (5).

$$\frac{x_1}{x_2} = \frac{{x_1}^{-1} + {h_2}^{-1}}{{x_1}^{-1} + {h_1}^{-1}} \quad (5)$$

Object size h_i can be determined from the resolution determination, and x_1 can be determined from the manufacturer's specification. By scaling the distortion model with the distortion center as the symmetric point, we can obtain a suitable distortion model for other pictures. Then the distortion can be corrected by applying Equation (1) to the pictures.

For example, the width of Figure 1 is about 3cm$\times$17 = 0.51m , the width of Figure 1 is about 3cm$\times$28=0.84m, and x_1 is 0.0276m for Kodak DCS 460 digital cameras. Therefore we can estimate the distortion ratio by Equation (5):

$$\frac{x_1}{x_2} = \frac{0.0276^{-1} + 0.84^{-1}}{0.0276^{-1} + 0.51^{-1}} = 97.98\%.$$

Figure 4 is the result of the correction.

6 Experiments on Image Stitching

First we calibrate the camera by the following steps:
1. setup the table and the camera parallel to the horizontal flat;
2. take a picture of the table;
3. find the tie-points in the picture;

4. solve the coefficients of the transformation matrix for each block;
5. estimate the actual size of the picture and the position of the distortion center.

For each manuscript of Danshin files, we take its picture by the following steps:
1. take a picture of the manuscript;
2. estimate the actual size of the picture;
3. compute the distortion ratio;
4. scale the distortion model;
5. use the transformation matrix to correct the distortion.

Finally we have some experiments on image stitching by translating and rotating the images manually. Figure 5 is the stitched image of Danshin Document #13214034 without correction. We rotate the left image about 1° counterclockwise to align words to have the same orientation on the overlap. But they still cannot match well because words on the right one are shorter. If we correct them by our distortion model, there is about 0.5° rotation between the orientations of corrected images. Figure 6 shows that the correction improves the accuracy of image stitching.

7 Discussion

Our distortion model fails without the assumption that the image plane is parallel to the table. In such condition, there is a good method to calibrate the camera and table geometry in [5]. However, we cannot ignore the factor of the effective focal length if we take pictures at different distances. Here we provide a method involving this factor and show that we do not have to calibrate for each effective focal length. The better way to determine the effective focal length is reading the camera zoom setting and it should be more accurate. Since the maximum error of h_1 is the block width of the calibration grid, the error of the distortion ratio can be estimated by Equation (6).

$$E = \frac{{x_1}^{-1} + {h_2}^{-1}}{{x_1}^{-1} + {h_1}^{-1}} - \frac{{x_1}^{-1} + (h_2 + w)^{-1}}{{x_1}^{-1} + {h_1}^{-1}} \qquad (6)$$

where E is the error and w is the width of a block. In our experiments, h_i is about 1m, w is 3cm, x_1 is 27.6mm, then E is about 0.08%. The error can be reduced by thicker grid.

We implement a portable C++ program and test it on several platforms. Table 1 shows the performances of the program on different machines. The program uses multithread to do the correction, so it gains much speed on the machines with multiple processors in correction step.

References

[1] G. P. Stein, “Lens Distortion Calibration Using Point Correspondences,” *Proceedings of IEEE Computer Society Conference on Computer Vision and Patter Recognition*, Puerto Rico, pp. 602–608, 1997.

[2] H. S. Sawhney and R. Kumar, “True Multi-Image Alignment and Its Application to Mosaicking and Lens Distortion Correction,” *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Puerto Rico, pp. 450–456, 1997.

[3] M. C. Chiang and T. E. Boult, “A Public Domain System for Camera Calibration and Distortion Correction,” *Technical Report CUSS-038-095*, Columbia Univ., Dept. of Computer Science, 1995.

[4] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Addison Wesley, Reading, MA, 1993.

[5] Y. Nomura, M. Sagara, H. Naruse, and A. Ide, “Simple Calibration Algorithm for High-Distortion-Lens Camera,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 14, pp. 1095–1099, 1992.

OS	CPU	Memory	Compiler	Time(sec)
SunOS 5.5.1	UltraSPARC 250MHZ×4	256MB	g++	4: 3: 3
Linux 2.0.33	PentiumII 233MHZ×2	128MB	g++	3: 7: 5
Windows NT 4.0	PentiumII 233MHZ×1	64MB	VC++	5: 13: 6

Table 1: The execution time for one image (3060 × 2036). Time is expressed in seconds and in the order of loading, correction, and writing.

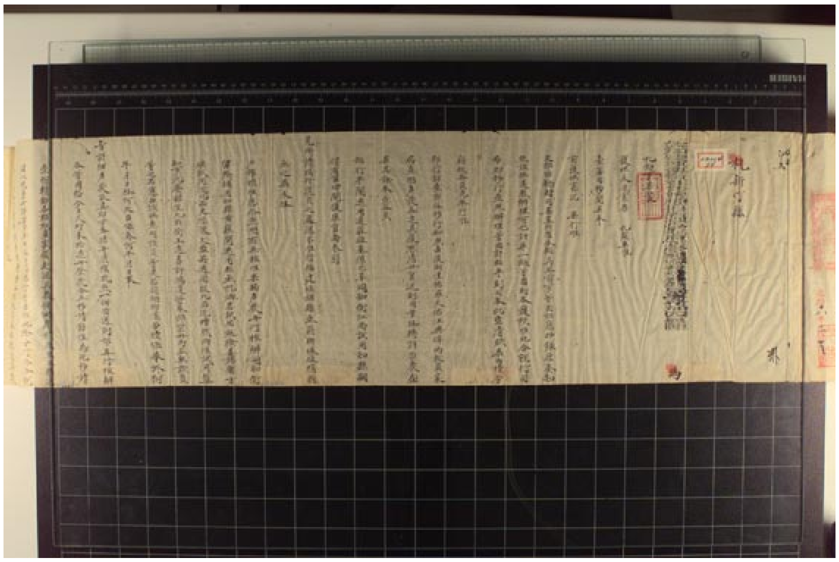

Figure 1: Right half of Danshin Document #13214034.

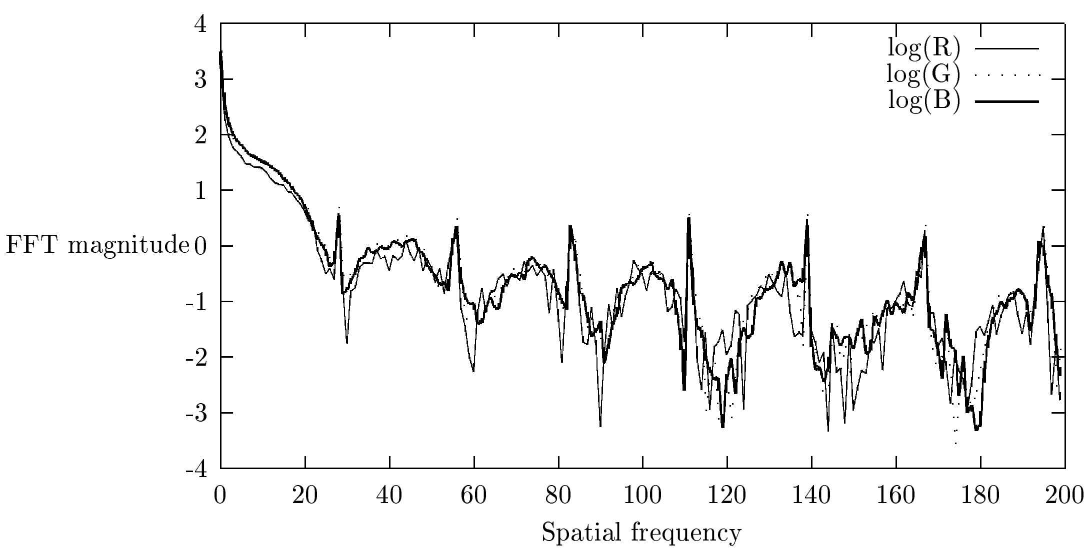

Figure 2: Results of FFT on the 200th row of Figure 1 (R, G, B for each curve). The first jump around the spatial frequency value is at 28 meaning there are 28 grids horizontally with extension.

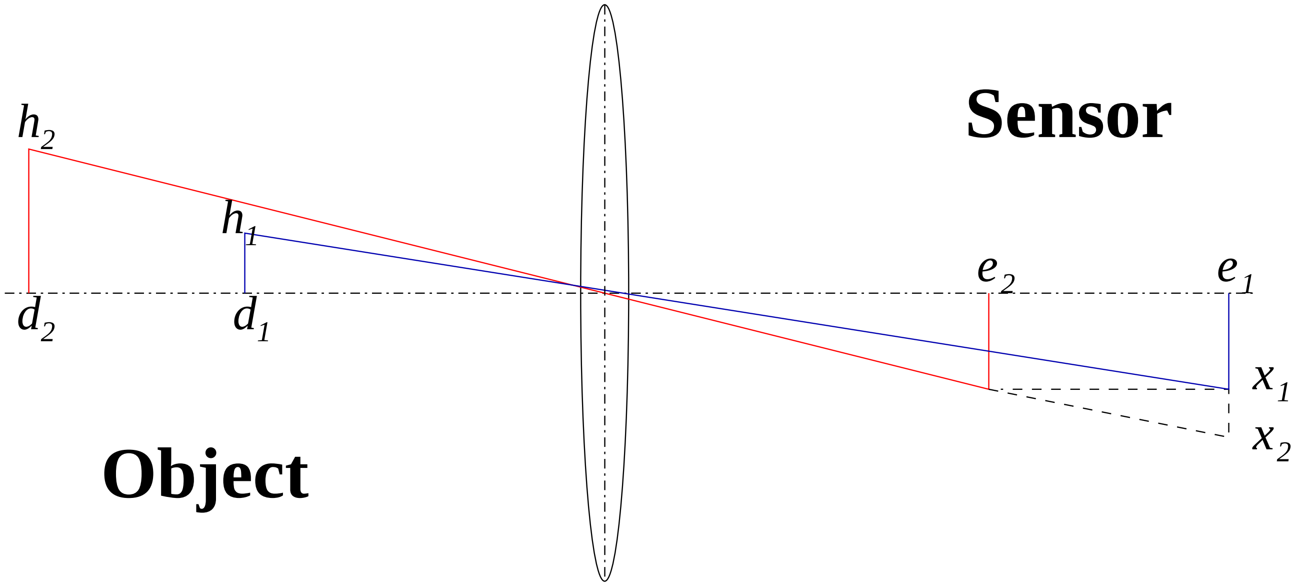

Figure 3: Camera geometry and parameters, where e_i is the effective focal length; d_i is the object distance; h_i is the actual object size; x_1 is the size of camera sensor; and x_2 is the projection size of object 2 if it is focused on the same plane as where the object 1 is.

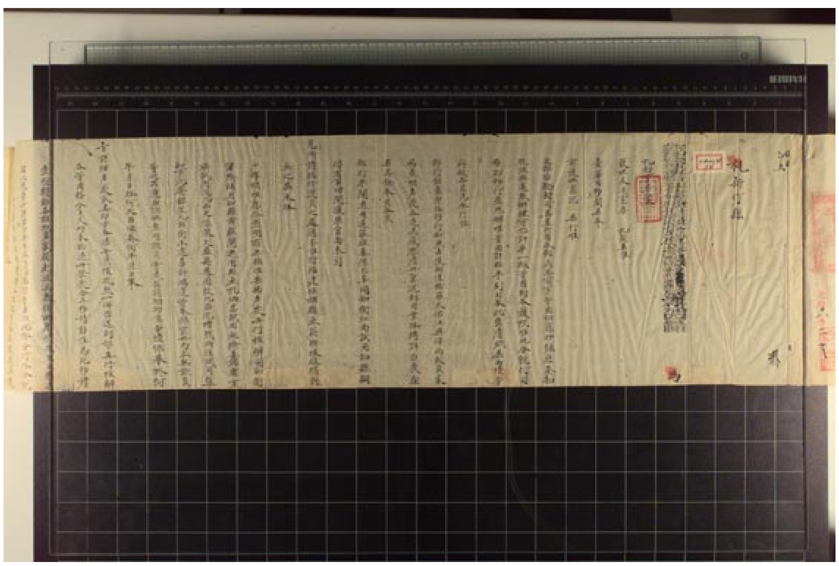

Figure 4: Right half of Danshin Document #13214034 after correction.

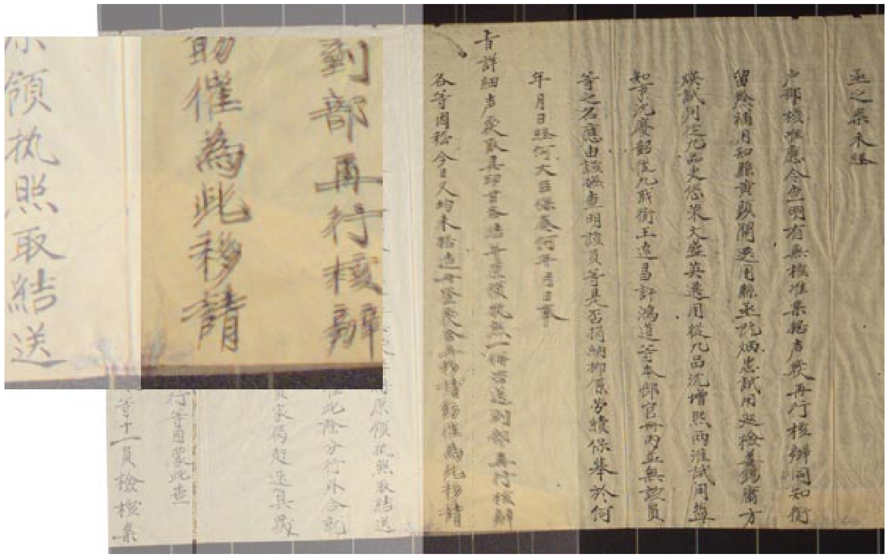

Figure 5: The stitched image of Danshin Document #13214034 without distortion correction. The inset on the left side is a magnification to show bad result without distortion correction.

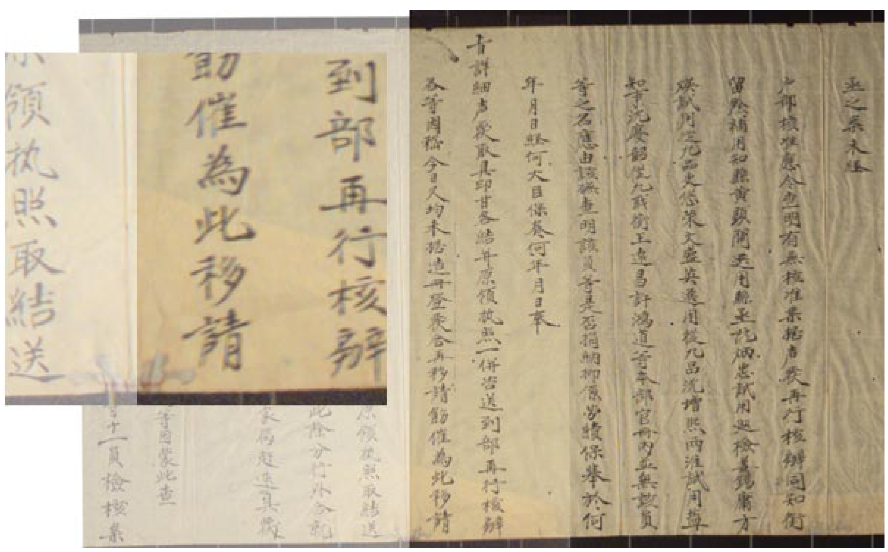

Figure 6: The stitched image of Danshin Document #13214034 with distortion correction. The inset on the left side is a magnification to show good result achieved with distortion correction.