Stereo Image Rectification for Simple Panoramic Image Generation

Yun-Suk Kang and Yo-Sung Ho
Gwangju Institute of Science and Technology (GIST)
261 Cheomdan-gwagiro, Buk-gu, Gwangju 500-712 Korea
Email:{yunsuk, hoyo}@gist.ac.kr

***Abstract*- In this paper, we propose an image rectification method to generate panoramic views from stereoscopic images. Due to different intrinsic and extrinsic characteristics of stereoscopic cameras, there are some mismatches in the vertical direction of the stereo images and they become major obstacles to make a panoramic image. Therefore, we need to reduce mismatches to generate more natural panoramic views. Since the distance between two cameras is relatively large for panoramic image generation, we perform two different camera calibration operations: relative calibration and non-relative calibration. After we calculate a rectifying transformation by estimating the camera parameters of the rectified images and apply it to the captured images, we can generate a panoramic image from the rectified stereo images. Experimental results show that the proposed method is useful for panoramic image generation and provides us much clearer panoramic view with fewer mismatch regions.**

I. INTRODUCTION

Since the television (TV) broadcasting service was started in 1930s, a variety of technologies related to scene capturing, image or video processing, data storage and transmission, and display devices have been developed steadily. As these technologies grow up, the demands for various multimedia formats that can provide three-dimensional (3-D) and immersive senses also have been increased. A stereoscopic image has been a challenging issue for a long time as one of the strong candidates to satisfy the user demands. The stereo image consists of two images of the same object or scene captured by two cameras at separate locations.

With this stereo image, we can generate a panoramic image, simply called panorama. Basically, the panorama is a wide view or a broad angle representation of a 3-D scene obtained by capturing or drawing. However, the panoramic image is the captured one by a panoramic camera or by merging multiple images from two or more cameras.

In order to generate the panoramic image from the stereo image, we have to consider vertical pixel mismatches between two images, which can be a crucial obstruction to generate a natural panoramic image. Unfortunately, these mismatches are unavoidable since the vertical coordinates of two camera centers are not equal. In addition, different intrinsic characteristics and different orientations of two cameras also cause the vertical mismatches.

In image rectification, we reduce these vertical pixel mismatches by relocating two camera centers, rotating two image planes, and regulating the camera intrinsic parameters. After image rectification, the vertical coordinates of two camera centers have the same value. The stereo camera also has the same orientation and the intrinsic features. Finally, we obtain the rectified stereo image that has only the horizontal displacement between two images. We can use this rectification technique to generate the panoramic image from the stereo image [1] [2] [3].

In this paper, we propose an image rectification algorithm to generate a simple panoramic image from a stereo image. Since the distance between two cameras for panoramic image generation is relatively larger than other stereo camera systems, we divide the camera calibration operation into two steps: relative and non-relative calibration. While we obtain the rotation and internal characteristics of the left camera and two camera centers from the relative calibration, we acquire the rotation and intrinsic characteristics of the right camera from the non-relative calibration operation. We then estimate new camera parameters of the right camera so that the rectification condition is satisfied. Finally, we calculate the rectification transformation from the two-dimensional (2-D) homography between the original and the estimated camera parameters and apply it to the right image.

Experimental results show that the generated panoramic image by the rectified stereo image has much smoother boundaries than the one obtained by the original images.

II. STEREO CAMERA SYSTEM

A. *Camera Model*

A pinhole camera model is explained by the camera center **C**, which is the center of the projection and the 3-D position of the camera, and the image plane *R* that has the projected image of the captured scene. The direction from the camera center perpendicular to the image plane is called the principal axis. The focal length is the distance from the camera center to the image plane.

Figure 1 shows the pinhole camera model. A point **M** in 3-D space is projected onto a point **m** in the image plane. Then, the camera operation is modeled by the linear transformation from the world coordinate system to the image coordinate system. This perspective transformation is given by a projection matrix **P** like Eq. 1.

$$\mathbf{m}=\mathbf{PM} \tag{1}$$

Equation 2 indicates that the camera projection matrix consists of the camera parameters. The camera parameters have three sets of the components which are the intrinsic parameters represented by the matrix A and the extrinsic parameters represented by the matrix R and the vector t.

$$\mathbf{P} = \mathbf{A}[\mathbf{R}|\mathbf{t}] \quad (2)$$

The intrinsic parameters form a matrix indicated in Eq. 3, where α_u and α_v represents the focal lengths in the horizontal and vertical pixels, respectively. The coordinate (u_0, v_0) is called the principal point at which the principal axis intersects the image plane. γ is the skew parameter that describes the non-orthogonality between the horizontal and vertical axes.

$$\mathbf{A} = \begin{bmatrix} \alpha_u & \gamma & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

The rotation matrix **R** and the translation vector **t** are composed of the extrinsic parameters. The 3×3 matrix **R** has the form shown in Eq. 4. The first, second, and third rows of **R** represents the direction of the horizontal, vertical, and principal axes of the camera, respectively. Therefore, these three rows must be orthogonal to one another.

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} = \begin{bmatrix} \mathbf{r}_1 \\ \mathbf{r}_2 \\ \mathbf{r}_3 \end{bmatrix} \quad (4)$$

The translation vector **t** has the three components like Eq. 5. This vector explains the distance of the camera from the world coordinate origin with considering the camera orientation. Therefore, the world coordinate system and the camera coordinate system are related through the extrinsic parameters.

$$\mathbf{t} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad (5)$$

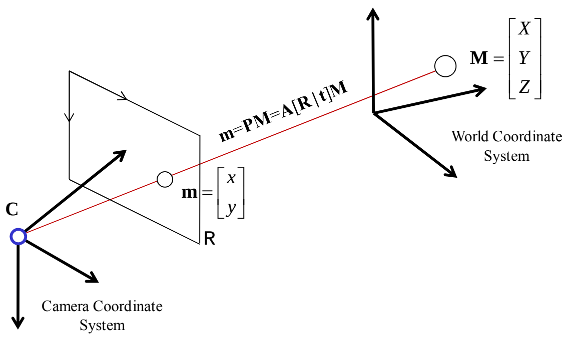


Figure 1. Pinhole camera model

B. *Stereo Image Characteristics*

We consider a stereo image captured by two cameras located apart. Figure 2 shows the geometrical model of the stereo camera system. There are two camera centers $\mathbf{C}_1$ and $\mathbf{C}_2$. The 3-D point **M** is projected onto two image planes. Then, the projected image points are $\mathbf{m}_1$ and $\mathbf{m}_2$, which represents the corresponding points on the left and right images, respectively.

The corresponding point of $\mathbf{m}_2$ in the left image is always located on the line called the epipolar line. This epipolar line is the intersecting line between the image plane and the epipolar plane which is defined by $\mathbf{C}_1$, $\mathbf{C}_2$, and **M**. Conversely, the corresponding point of $\mathbf{m}_1$ in the right image is also located on the epipolar line.

The epipole is the projected point of the other side camera center in each image. Thus, the intersections between two image planes and the line through two camera centers becomes the epipoles, and all the epipolar lines in one image pass the epipole.

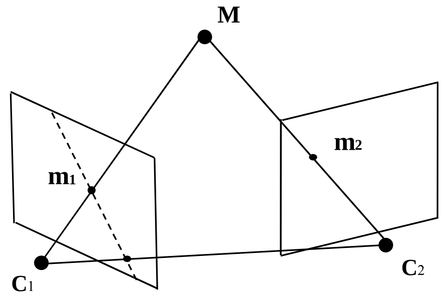


Figure 2. Geometric model of the stereo camera

However, the captured images by such stereo camera system have the mismatches since the orientation and the internal characteristics are different each other, and even two camera centers does not have the same vertical coordinates. In this case, the epipolar lines in each side are not parallel one another.

These mismatches lead to the pixel differences in the vertical direction between two images that decrease the correlation between each view as shown in Fig. 3. The mismatches also become the critical obstacles to the stereo image processing and applications.

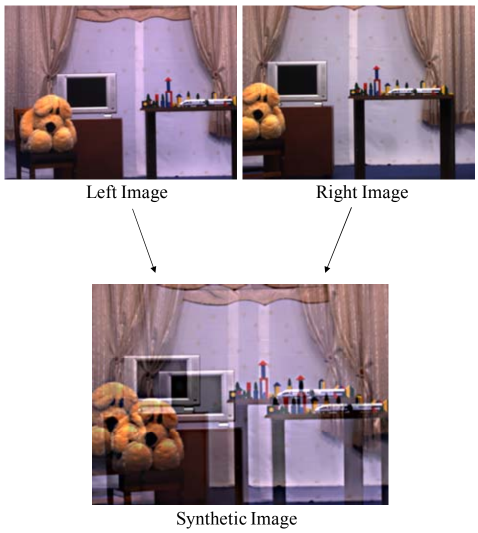


Figure 3. Mismatches in the stereo image

III. Stereo Image Rectification

In order to solve the mismatch problem in the stereo image, we perform image rectification. Image rectification is an essential preprocessing for the stereo image captured by the stereo camera system that has the aforementioned mismatches. For the given stereo image, rectification is a transformation of two image planes that makes their all epipolar lines are parallel each other as depicted in Fig. 4.

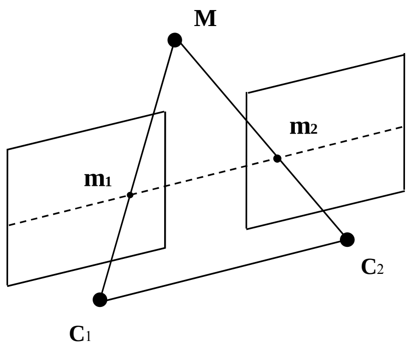


Figure 4. Geometric model of the rectified stereo camera

The rectified images have the corresponding points which have the same vertical coordinates. It means that there exist only the horizontal disparities between two images. Therefore, the images after rectification can be considered as the images captured by a correctly arranged stereo camera system as shown in Fig. 5.

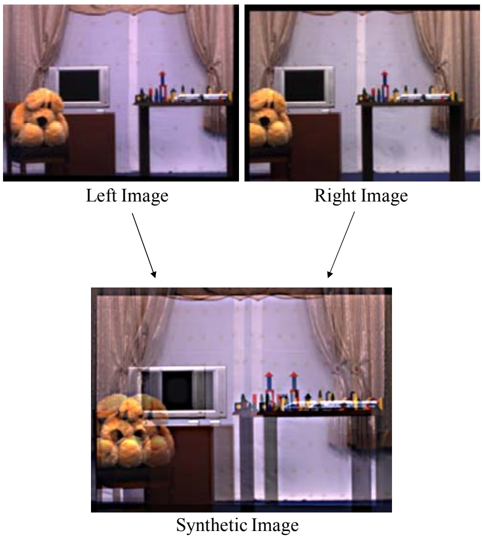


Figure 5. Rectified stereo image

There are two main categories of image rectification. One is for the calibrated case and the other is for the non-calibrated ones. Although we must know the camera parameters, the calibrated case of image rectification can give us much clear and simpler algorithms than the non-calibrated case. Rectification algorithms belong to these two categories has been studied for a long time in computer vision.

In order to rectify the stereo images, we define two new camera projection matrices by adjusting the camera intrinsic and extrinsic parameters. Then, we calculate the rectifying transformation using the relation between the original camera matrices and the new camera matrices. These new camera matrices must satisfy that all the epipolar lines are parallel each other and the corresponding points are located on the same vertical coordinates. These conditions can be achieved by rotating and shifting the original image planes and equalizing their intrinsic parameters [3].

After applying this rectification algorithm to the stereo image, we can obtain the regulated orientations, positions, and the intrinsic parameters of the cameras like Fig. 4. Hence, this rectified stereo image can be considered as the images captured by a well-arranged stereo camera. Then, the complexity of searching for the corresponding points between two images can be significantly reduced since there are disparities only in the horizontal direction. Therefore, we can expect advantages on time and accuracy when we perform a 3-D image processing and applications.

IV. Panoramic Image Generation Using Stereo Image Rectification

In this section, we introduce our proposed method to generate a panoramic image using stereo image rectification. The procedure of our method is shown in Fig. 6. We firstly capture a scene by two cameras apart which have the same orientation with a little built-in orientation mismatch. We then perform camera calibration to each camera relatively and non-relatively, respectively. The camera parameters obtained from two calibration steps become the input for image rectification. After rectifying the stereo images, we can easily generate the panoramic image.

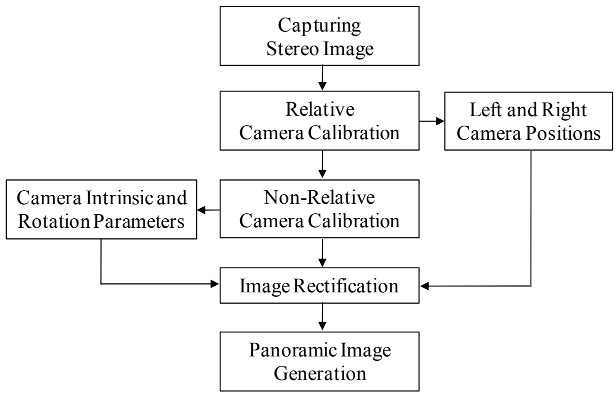


Figure 6. Procedure of the proposed method

A. *Camera Calibration*

Camera Calibration [4] is the process to extract metric information of the camera from 2-D images. In order to calibrate the camera, it is required several images of a check-patterned planar board. We capture images of this board with changing its position and orientation several times. Then, by extracting a number of feature points from each image with the real distance between these points, we can find out camera intrinsic and extrinsic parameters indicated in Eq. 3, Eq. 4, and Eq. 6. Finally, we can calculate camera projection matrices by combining the camera parameters.

When we calibrate two or more cameras, we simultaneously capture a check-patterned board by whole cameras in general. In this case, the distance between adjacent cameras has to be small enough for simultaneous capturing. For generating the panoramic image with the stereo image, however, it is required relatively large distance between two cameras. In this case, we cannot simultaneously capture images of the planar board. Therefore, we divide the calibration into two steps which are the relative and non-relative calibration.

1. Relative Calibration

Relative calibration means that we simultaneously capture images of various shapes of a check-patterned board. Figure 7 shows the relative calibration of the stereo camera. We fix the positions of two cameras and rotate the right camera so that it can capture several different images of the board.

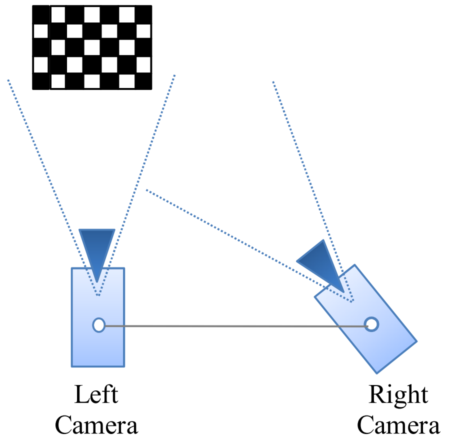


Figure 7. Relative camera calibration

As the results, we obtain the camera intrinsic, rotation and translation parameters of two cameras, which have the relation to each other. However, we choose only information of the camera positions since our required data in this camera arrangement are only the coordinates of the camera centers. Each camera center is calculated as Eq. 6. It is because the right camera is rotated to capture the planar board and to obtain only the camera center.

$$\mathbf{C}_r = -(\mathbf{A}_r\mathbf{R}_r)^{-1} \cdot (\mathbf{A}_r\mathbf{t}_r) \quad (6)$$

2. Non-Relative Calibration

After obtaining the camera centers of two cameras, we rotate back the right camera to have the original orientation which is the same as the right camera's one. Then, we capture the images of the planar board for the left camera as shown in Fig. 8. This is the non-relative calibration whose purpose is to obtain the camera intrinsic and rotation parameters of the right camera.

However, to acquire the intrinsic and rotation information of the right camera in the same condition of the left camera, we have to use the same set of the variously posed planar board used for the relative calibration. Figure 10(a) shows the relative calibration of the left camera, Fig. 10(b) also indicates the relative calibration of the right camera, and Fig.10(c) depicts the non-relative calibration of the right one.

Therefore, we can obtain the intrinsic and rotation parameters of the right camera from this non-relative calibration, and we can use the relative calibration data for the left camera. Finally, we have two camera projection matrices indicated in Eq. 7, where the subscript L and R mean the left and right cameras, and r and n mean the relative and non-relative calibrations, respectively.

$$\begin{cases} \mathbf{P}_{\mathrm{L}} = \mathbf{A}_{r,\mathrm{L}}[\mathbf{R}_{r,\mathrm{L}}| - \mathbf{R}_{r,\mathrm{L}}\mathbf{C}_{r,\mathrm{L}}] \\ \mathbf{P}_{\mathrm{R}} = \mathbf{A}_{n,\mathrm{R}}[\mathbf{R}_{n,\mathrm{R}}| - \mathbf{R}_{n,\mathrm{R}}\mathbf{C}_{r,\mathrm{R}}] \end{cases} \quad (7)$$

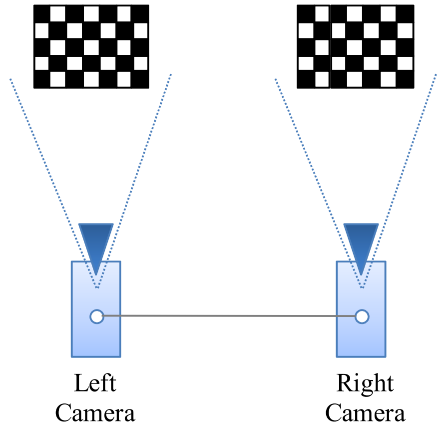


Figure 8. Non-relative camera calibration

B. Stereo Image Rectification for Panoramic Image Generation

In order to rectify the stereo image, it is required a transformation between the images from the original cameras and the rectified images from the new cameras. Therefore we have to estimate information about the new cameras.

In the proposed method, we consider the orientation and position of the left camera as the reference and estimate the right camera projection matrix based on this reference. In order words, we need a transformation for the right image whereas the left image is not transformed. Therefore, the new position of the right camera center can be defined as Eq. 8, where d is the distance between two original camera centers and $\mathbf{r}_{1,r,\mathrm{L}}$ is the horizontal direction of the left camera.

$$\mathbf{C}'_{\mathrm{R}} = \mathbf{C}_{r,\mathrm{L}} + d \cdot \mathbf{r}_{1,r,\mathrm{L}} \quad (8)$$

After defining the new right camera center, we estimate the rotation and the intrinsic parameters of the right camera. For the rotation, we simply choose the rotation matrix of the left camera. Then, we calculate the average value of each intrinsic parameter of two cameras except the skew parameter, and select these values as the new intrinsic parameters. Finally, the new projection matrix of two cameras has the form indicated in Eq. 9.

$$\begin{cases} \mathbf{P}'_{\mathrm{L}} = \mathbf{A}'[\mathbf{R}_{r,\mathrm{L}}| - \mathbf{R}_{r,\mathrm{L}}\mathbf{C}_{r,\mathrm{L}}] \\ \mathbf{P}'_{\mathrm{R}} = \mathbf{A}'[\mathbf{R}_{r,\mathrm{L}}| - \mathbf{R}_{r,\mathrm{L}}\mathbf{C}'_{\mathrm{R}}] \end{cases} \quad (9)$$

With the relation between the new projection matrix and the original one of the right camera, we define the rectifying transformation as the 2-D homography. Figure 10 shows that the point transfer via the plane. A point $\mathbf{x}_i$ in the left image is mapped onto the point $\mathbf{x'}_i$ by transformation indicated in Eq. 10 that is composed of the two projection matrices. Therefore, we can consider those two image planes in Fig. 9 as the original and new image planes, respectively. We then apply the rectifying transformation shown in Eq. 11 to the right image and obtain its rectified image.

$$\mathbf{H} = \mathbf{P}_2\mathbf{P}_1^+ \quad (10)$$

$$\begin{cases} \mathbf{T}_\mathrm{L} = \mathbf{P'}_\mathrm{L}\mathbf{P}_\mathrm{L}^+ \\ \mathbf{T}_\mathrm{R} = \mathbf{P'}_\mathrm{R}\mathbf{P}_\mathrm{R}^+ \end{cases} \quad (11)$$

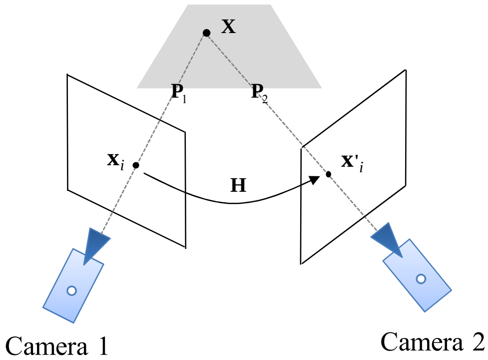


Figure 9. Point transfer via the plane

V. Experimental Results

A. Experimental Setup

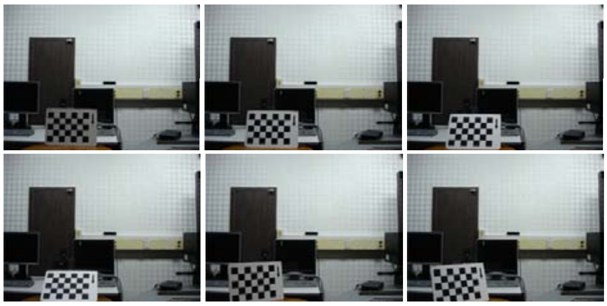
We captured test images for testing by using a digital camera whose model is SONY DSC-W1. For camera calibration, we used the check-patterned planar board shown in Fig. 10 and the Camera Calibration Toolbox for Matlab provided by Caltech [5]. We captured our test images inside from two positions with one camera mounted on a tripod by shifting it.

B. Panoramic Image Generation

In order to acquire the test images, we firstly captured the image and images of variously posed planar board of the left view as depicted in Fig 11(a) and Fig. 10(a), respectively. Then, we shifted the camera to the left side to capture the left view. For the relative calibration, we obtained the planar board images like Fig. 10(b). After that, we rotated the camera and captured the scene and images of the patterned board of the right view as indicated in Fig. 11(b) and Fig. 10(c), respectively. For capturing the same poses of the planar board, we put the board on the table which has prominence and depression on its surface to stand the board with regular angles.

(a) Relative calibration of the left camera

(b) Relative calibration of the right camera



(c) Non-relative calibration of the right camera

Figure 10. Various poses of the planar board

(a) Left image

(b) Right image

Figure 11. Captured stereo image

Figure 12. Result by the original stereo image

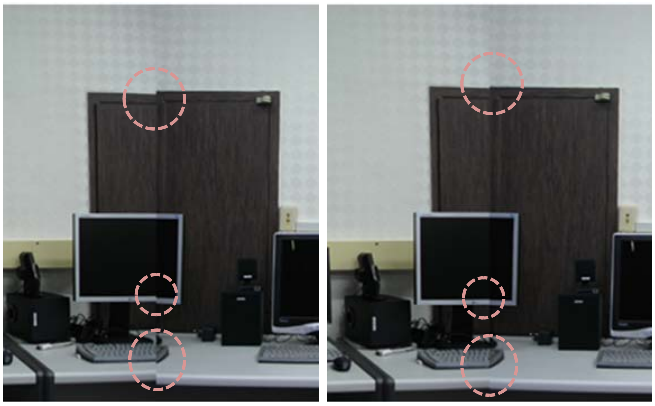
When we generated the panoramic image by using the original stereo image, however, there existed a few mismatch regions in the joining area between two images as shown in Fig. 12.

Figure 13(a) and Figure 13(b) show the rectified stereo image. In this case, holes occurred around image boundaries due to rectification, and they could be removed since the transformation considered the increasing the focal length of each camera. Therefore, the rectified stereo image was a little zoomed in.

(a) Left image (b) Right image

Figure 13. Rectified stereo image

Figure 14. Result by the rectified stereo image



(a) Original (b) Proposed

Figure 15. Improvements in mismatch regions

The generated panoramic image by the rectified stereo image is shown in Fig. 14. Although we had a few lost areas when we join two images as the original case, we could obtained much clear result compared with the one by the original stereo image. This result has less mismatch regions and gives a natural view as the panoramic image. Figure 15 shows the improvements in the mismatch regions of the original panoramic images. Because of the rectifying transformation, the rectified two images are like images captured by horizontally apart two cameras that have the same intrinsic characteristics.

VI. CONCLUSION

In this paper, we introduced a stereo image rectification method for simple generation of a panoramic image. Since the distance between two cameras is relatively long when we try to make the panoramic image, we divided camera calibration as two steps which are the relative and non-relative calibrations. Then, we obtained two camera centers and the camera parameters of the left camera from the relative calibration. From the non-relative calibration, we merely acquired the intrinsic and rotation parameters of the left camera. Finally, we calculated the rectifying transformation with these parameters and applied it to the original stereo image. The generated panoramic image by the rectified stereo image not only has fewer number of mismatch regions but also gave us more natural view.

ACKNOWLEDGMENT

This work was supported in part by ITRC through RBRC at GIST (IITA-2008-C1090-0801-0017).

REFERENCES

[1] R. Hartley, “Theory and Practice of Projective Rectification,” International Journal of Computer Vision, vol. 35, no. 12, pp. 115-127, 1999.

[2] C. Loop and Z. Zhang, “Computing Rectifying Homographies for Stereo Vision,” In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’99), pp. 125-131, 1999.

[3] A. Fusiello, E. Trucco, and A. Verri, “A Compact Algorithm for Rectification of Stereo Pairs,” Machine Vision and Application, vol. 12, no.1, pp. 16-22, 2000.

[4] Z. Zhang, “A Flexible New Technique for Camera Calibration,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22(11), pp. 1330-1334, 2000.

[5] http://www.vision.caltech.edu/bouguetj//calib_doc/index.html

[6] R. Hartley and A. Zisserman, “Multiple View Geometry in Computer Vision,” Cambridge University Press, 2003.

[7] G. Medioni and S. B. Kang, “Emerging Topics in Computer Vision,” Prentice Hall, 2004.