University of Technology, VNU-HCM
Ho Chi Minh City, Viet Nam

IEEE-RIVF 2012

Addendum Contributions of the 2012 IEEE-RIVF International Conference on Research, Innovation and Vision for the Future

2012

Luong Chi Mai, Vincenzo Piuri, Laurent El Ghaoui (Eds.)

2012 IEEE-RIVF International Conference on Computing and Communication Technologies Research, Innovation and Vision for the Future

Ho Chi Minh City, Viet Nam

IEEE-RIVF 2012 Addendum Contributions

Geometrical Calibration of Stereo Images in Convergent Camera Arrangement

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

Abstract— **In this paper, we propose a geometrical calibration method of stereo image captured in a convergent camera arrangement. The stereo images captured from practical stereo convergent arrangements have an asymmetric geometry that causes not only an ambiguous convergent points but also vertical pixel mismatches between corresponding points. These characteristics make the depth estimation difficult, and also decrease the visual quality as three-dimensional (3D) images. Therefore, the proposed method calibrates the right view image based on the left view image to obtain the stereo image of a perfectly symmetric stereo geometry. The transform is calculated by using the original and estimated camera parameters of the right camera. Experimental results show that the result images of the proposed method have clear convergent points and also few vertical pixel mismatches.**

Keywords-component; geometrical calibration; stereo image; convergent camera arrangement; 3DTV

I. INTRODUCTION

Recently, a variety of broadcasting technologies, especially for three-dimensional television (3DTV), have been investigated [1]. Using two or more viewpoints for one scene, 3DTV provides more realistic and immersive sense to users than the typical TV broadcasting. In order to generate 3D contents for 3DTV, we basically require the stereo image that has two viewpoints for the scene. Then the users feel the 3D sense with stereoscopic displays. Moreover, we can reconstruct novel viewpoints' images after estimating depth information of the scene [2]. These novel viewpoints guarantee not only more natural and smooth 3D sense but also wide viewing angle.

The stereo image is obtained by two cameras. In general, there are two types for stereo camera arrangement; parallel and convergent types. In the parallel camera arrangement, two cameras are placed in the same direction with a proper distance which defines the amount of the disparity. The two cameras on the convergent camera arrangement have an angle between their optical directions at a convergent point. The disparity from the parallel camera arrangement is usually larger and more uniform than the disparity from the convergent camera arrangement.

In order to use the stereo image as the input for the stereoscopic display, it is essentially required to reduce geometric error in the images. This geometric error is caused by the different camera intrinsic and extrinsic characteristics. It is represented as the vertical pixel mismatches between corresponding points in both images. The pixel mismatches in the vertical direction decrease the visual quality of the 3D scene and cause eye fatigue. For the parallel camera arrangement case, image rectification algorithms have been proposed for stereo or multi-view image to reduce the geometric error [3] [4]. The rectified images have the intrinsic and extrinsic characteristics of an ideal parallel camera arrangement.

In order to solve these problems in convergent stereo images, we propose a geometrical calibration method that provides the stereo images from an ideal convergent camera arrangement. The ideal convergent arrangement means the perfectly symmetric stereo geometry by two internally same camera models. In this camera arrangement, the captured image has clear convergent points represented in the objects in the scene and few vertical pixel mismatches between the corresponding points in each image plane.

The transform for calibration is calculated only for the right view image based on the left view as a reference. After camera calibration, we calculate the new camera center, rotation matrix, and intrinsic matrix for the calibrated right camera. Then the transform is calculated as a homography between the original and calibrated right views. By applying the transform to the right view image, we obtain the calibrated right view image.

II. CAMERA MODEL AND STEREO GEOMETRY

A. Camera Projection

The projection of a single camera is modeled by a pinhole camera model shown in Fig. 1. The camera is placed on a point in 3D space, which is called the camera center **C**. This camera center is considered as the center of projection of the captured scene. The captured scene is projected and appeared on the image plane as an image.

The camera has its own coordinate system called camera coordinate system that has three basis axes. The *z*-axis is called the optical axis that passes perpendicular to the image plane from **C**. The focal length is the distance from **C** to the image plane in this *z*-axis' direction, and the intersecting point is called the principal point. The *x*- and *y*- axis are horizontal and vertical axes, respectively. These three axes are orthogonal one another. The line through the camera center and perpendicular to the image plane is called the optical axis.

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 2011-0030822).

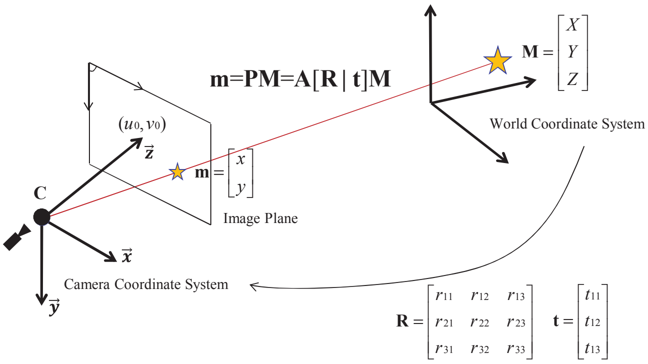

Figure 1. Camera projection.

As shown in Fig. 1, a 3D point **M** is projected to the image plane as an image point **m** by the 3x4 projection matrix **P** as Eq. 1, which is the perspective transformation.

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

Equation 2 shows the components of the camera projection matrix. These components are called the camera parameters that represent the camera's intrinsic and extrinsic characteristics. The extrinsic parameters are 3x3 rotation matrix **R** and 3x1 translation vector **t**, which are related to the camera orientation and location, respectively. By **R** and **t**, the world coordinate system relates to the camera coordinate system. The intrinsic parameters are formed as 3x3 matrix **A** indicated in Eq. 3.

$$\mathbf{P} = \mathbf{A}[\mathbf{R}|\mathbf{t}] \tag{2}$$

$$\mathbf{A} = \begin{bmatrix} \alpha_x & \gamma & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

In Eq. 3, α_u and α_v are the focal lengths in horizontal and vertical pixels. The coordinate (x_0, y_0) is the principal point and γ is the skew parameter that describes non-orthogonality between *x*- and *y*- axes. In general, γ is close to zero.

B. Epipolar Geometry

For the stereo camera located at $\mathbf{C}_1$ and $\mathbf{C}_2$, we have one 3D point **M** and two image point $\mathbf{m}_1$ and $\mathbf{m}_2$, which are the corresponding points on each view image, as shown in Fig. 2. We consider two stereo images captured by stereo pinhole cameras from two different positions like Fig. 2. Finding the corresponding points on the other viewpoints is important for 3D applications.

In this stereo camera configuration, the epipolar geometry is established. The epipolar plane is defined by the three points M, $\mathbf{C}_1$ and $\mathbf{C}_2$. Therefore, $\mathbf{m}_2$, the corresponding point of $\mathbf{m}_1$, is on the epipolar plane. We call the intersecting line between the image plane and epipolar plane the epipolar line. Therefore, the corresponding point of $\mathbf{m}_1$ is located on the epipolar line in the right image plane according to the depth.

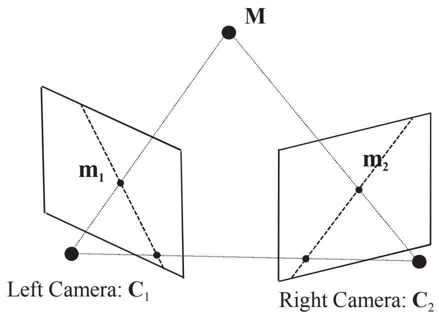

Figure 2. Epipolar geometry.

There are two epipoles that are intersecting points between each image planes and the line through two camera centers. In order words, epipoles are the projection point of each camera center to the other image plane. All the epipolar lines pass the epipole as shown in Fig. 2.

III. Geometrical Calibration Method for Convergent Stereo Image

A. Characteristics of Convergent Stereo Image

An ideal convergent camera array has the cameras which are arranged on an ideal arc in 3D space. Two image planes are placed in symmetric. Also, the intrinsic parameters of each camera are the same. Therefore, images have clear convergent points. However, a practical convergent camera array does not have such ideal conditions. Two image planes have different optical directions, and two cameras' intrinsic parameters are different each other.

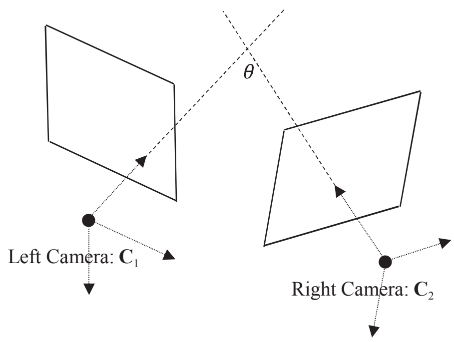

Figure 3. Practical stereo geometry.

These practical conditions of the convergent arrangement are shown in Fig. 3. The stereo image captured by this practical convergent arrangement has an ambiguity of convergent points, and also the vertical pixel mismatches between the corresponding points of each view. These conditions are not only the difficulties for 3D image processing but also the quality decrement as the 3D scene.

B. Geometrical Calibration Process

Figure 4 shows the whole procedure of the proposed method. After capturing the stereo image, we estimate the camera parameters by camera calibration [5]. By using the camera parameter, we calculate the new camera center, new rotation matrix, and new intrinsic parameters for the right view image based on the left view image as the reference view. We then compute the transform of the right view image. Applying this transform matrix to the right view image gives the calibrated right view image.

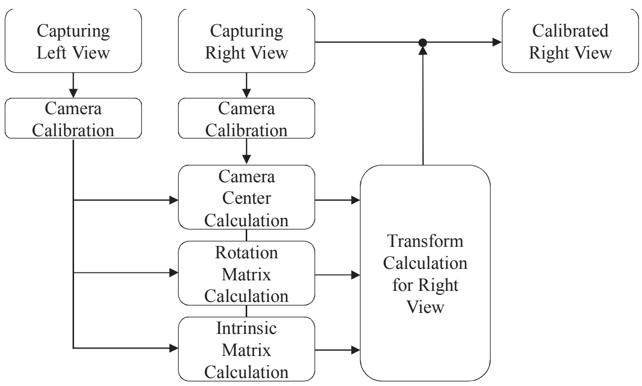

Figure 4. Procedure of the proposed method.

In order to calculate the new camera center for the right view image $\mathbf{C'}_r$, we firstly measure the distance d and angle θ between left and right cameras. Then the new camera center is calculated as Eq. 4. In this equation, we denote that $\overrightarrow{\boldsymbol{x}_l}$ and $\overrightarrow{\boldsymbol{z}_l}$ are the horizontal and optical axes of the left camera, respectively.

This new camera center for the right view is located on the plane that is defined by the horizontal and optical axes of the left camera. It also maintains the distance between two cameras and makes the calibrated stereo geometry symmetric, as shown in Fig. 5.

$$\mathbf{C'}_r = \mathbf{C'}_l + d\,\frac{\overrightarrow{z_l}+\tan((180°-\theta)/2)\cdot\overrightarrow{x_l}}{\|\overrightarrow{z_l}+\tan((180°-\theta)/2)\cdot\overrightarrow{x_l}\|} \quad (4)$$

For the next step, we calculate the new rotation matrix for the right view. The vertical axis has to be the same for both views. Therefore, we set the vertical axis direction of the right view as Eq. 5.

$$\overrightarrow{\boldsymbol{y'}_r} = (y_{l,x} \quad y_{l,y} \quad y_{l,z})^{\mathrm{T}} \quad (5)$$

Then, the optical axis of the right camera is calculated as Eq. (6). This new optical axis must be on the plane defined by $\overrightarrow{\boldsymbol{x}_l}$ and $\overrightarrow{\boldsymbol{z}_l}$. It also has to maintain the angle θ. The horizontal axis of the right camera is defined as the cross product of $\overrightarrow{\boldsymbol{y}_r}$ and $\overrightarrow{\boldsymbol{z}_r}$. The new rotation matrix of the right camera is shown in Eq. 7.

$$\overrightarrow{\boldsymbol{z'}_r} = \overrightarrow{\boldsymbol{z}_l} - \tan\,\theta \cdot \overrightarrow{\boldsymbol{x}_l} \quad (6)$$

$$\mathbf{R'}_r = \begin{bmatrix} \overrightarrow{\boldsymbol{x'}_r}^{\mathrm{T}} \\ \overrightarrow{\boldsymbol{y'}_r}^{\mathrm{T}} \\ \overrightarrow{\boldsymbol{z'}_r}^{\mathrm{T}} \end{bmatrix}. \quad (7)$$

Finally, the left rotation matrix is used for the new intrinsic matrix of the right camera because we consider the left camera as the reference camera. Therefore, the projection matrix of the calibrated right camera is defined as Eq. 8, and the calibrated stereo geometry shows the perfect symmetric characteristic as shown in Fig. 5.

$$\mathbf{P'}_r = \mathbf{A}_l[\mathbf{R'}_r| - \mathbf{R'}_r\mathbf{C'}_r] \quad (8)$$

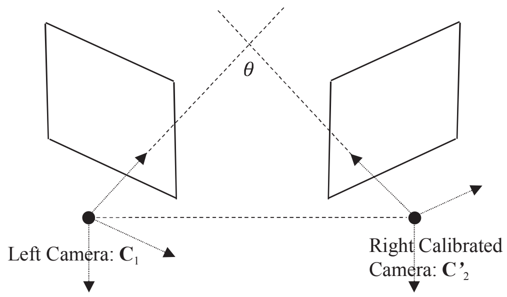

Figure 5. Calibrated stereo geometry.

C. Transform for Right View Image

We calculate the transform matrix for the right view image as the last step. By applying this transform, we obtain the calibrated right view image. This transform is calculated as a 2D homography between the original right camera and the calibrated right camera [6]. By using the camera projection matrices, the transform $\mathbf{T}_r$ is as Eq. 9, where $\mathbf{P}_r^+$ is the pseudo-inverse matrix of $\mathbf{P}_r$.

$$\mathbf{T}_r = \mathbf{P'}_r\mathbf{P}_r^+ \quad (9)$$

IV. Experimental Results

In order to test the proposed method, we captured three stereo image sequences as shown in Fig. 6. The camera model is FLEA-HICOL-CS and the image resolution is 1024x768. The original(practical) and measured angles and distances are in Table I. The measured data is based on the camera calibration results [7].

TABLE I. ANGLE AND DISTANCE

	Original		**Measured**	
	Angle (deg.)	Distance (mm)	Angle (deg.)	Distance (mm)
Sequence 1	4	65	3.81	64.8
Sequence 2	4.5	65	4.46	62.2
Sequence 3	5	50	5.12	51.7

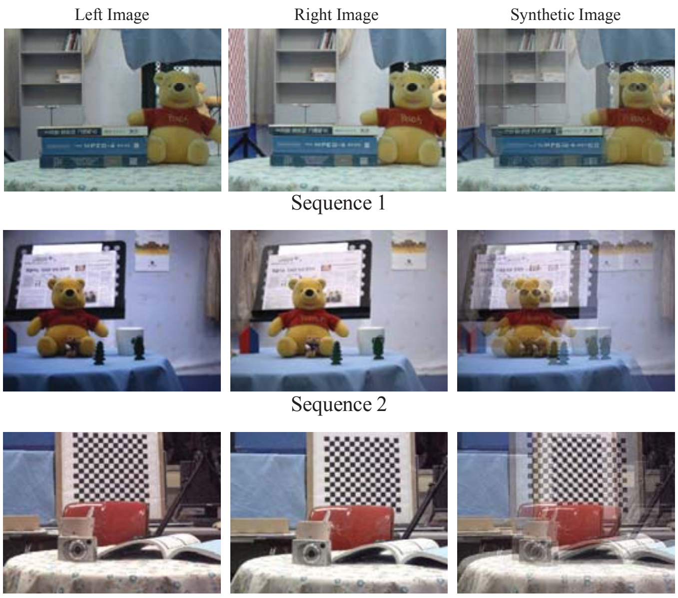

Figure 6. Captured stereo image sequences.

As shown in Fig. 6, the synthetic images show that it is hard to find the convergent points. There are also vertical pixel mismatches between corresponding points. However, the proposed method not only maintains the angle and distance between two cameras but also efficiently transform the characteristics of the right view image to have the symmetric convergent geometry.

Figure 7 shows the result images of the proposed method. The left images were not changed. The right image is calibrated by the transform and we notice that the image planes were moved and rotated. As shown in the synthetic images in Fig. 7, we can find several objects at the convergent points. Also, the vertical pixel mismatches were reduced.

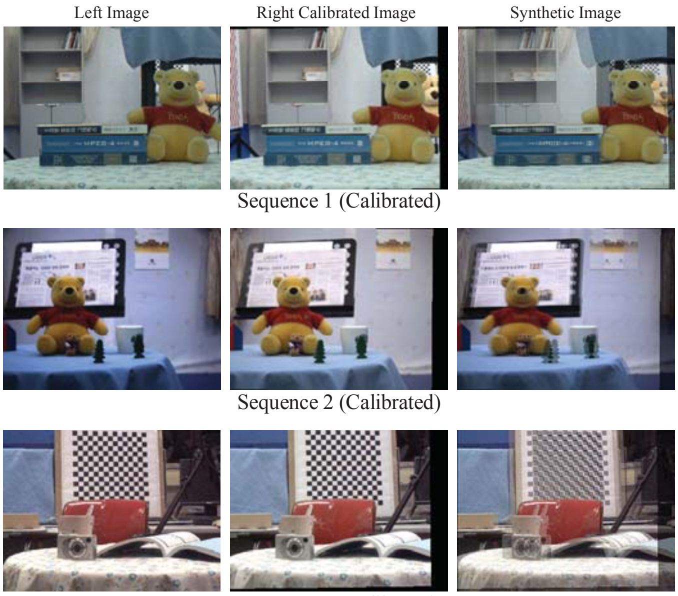

Figure 7. Calibrated stereo image sequences.

V. CONCLUSION

In this paper, we presented a geometrical calibration method for stereo image captured by a convergent arrangement. In order to reduce the ambiguity of the convergent points and the vertical pixel mismatch, we transform the right view image to obtain the calibrated right view image based on the left view. The transform is calculated by using the original and estimated camera parameters of the right camera. The experimental results show the clearness in the convergent points and also reduced vertical mismatch pixels.

REFERENCES

[1] C. Fehn, R. Barre, and S. Pastoor, “Interactive 3-DTV-Concepts and Key Technologies,” Proceeding of the IEEE, vol. 94, no. 3, pp. 524-538, March 2006.

[2] J. Sun, N.N. Zheng, and H.Y. Shum, "Stereo Matching Using Belief Propagation," IEEE Transactions of Pattern Analysis and Machine Intelligence, vol. 25, no. 5, pp. 787-800, July 2003.

[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, July 2000.

[4] Y.S. Kang and Y.S. Ho, “An Efficient Rectification Method for Parallel Multi-Camera Arrangement,” IEEE Transactions on Consumer Electronics, vol. 57, no. 8, pp. 1041-1048, Aug. 2011.

[5] Z. Zhang, “A flexible new technique for camera calibration,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no. 11, pp. 1330-1334, Nov. 2000.

[6] R. Hartley and A. Zisserman, “Multiple view geometry in computer vision,” Cambridge University Press, 2003.

[7] http://www.vision.caltech.edu/bouguetj, Camera Calibration Toolbox for MATLAB.