An Efficient Trinocular Rectification Method for Stereo Vision

Young Ki Baik, Jonghyun Choi and Kyoung Mu Lee

School of Electrical Engineering and Computer Science, Seoul National University,
hyunxx00@snu.ac.kr, ppolon@snu.ac.kr, kyoungmu@snu.ac.kr

Abstract In this paper, we introduce an efficient method of trinocular rectification to enhance the performance of stereo vision. In order to perform the trinocular rectification, we first obtain trinocular images by three cameras lying on a plane that faces the target objects. Then, we extract the original camera parameters, and compute the new camera parameters that transform the epipole on each image to the point at infinity. After that, we determine the homographies which relate the original camera parameters to the new ones, and decide an affine matrix which ensure 45 degrees for the epipolar line between top and left image. Finally, the trinocular rectification is finished by applying homographies and an affine transform to the obtained trinocular images. In contrast to conventional trinocular rectification methods that are mathematically and geometrically complicated, the proposed method is simple but efficient and accurate.

1 Introduction

One of the major problems of stereo vision is to establish dense and accurate correspondences. To alleviate the computational overhead, one can make the epipolar line be horizontally coincide. This process is called the image rectification and is possible by using epipolar geometry. There have been a lot of researches for this topic.

Fusiello [1] introduced a compact method with known camera parameters for the binocular rectification. Hartley et al. [6] proposed an advanced rectification method with only corresponding information between two images without known camera parameters. Also, there have been researches for minimizing the geometric distortion and computational error [7, 8, 9]. With respect to the accuracy of stereo vision, the trinocular rectification method is preferable rather than binocular one. So, Ayache [2] devised a trinocular rectification method with known camera parameters. Based on Ayache's method, An et al. [3] proposed an modified algorithm that finds the proper internal parameters. Sun [4] suggested a method using a fundamental matrix from corresponding points with unknown camera parameters. Finally, Heinrichs [5] introduced a practical algorithm based on Sun's method.

In this paper, we propose an efficient trinocular rectification method with known camera parameters. Conventional works are mathematically and geometrically complicated, ambiguous and hard to understand. On the other hand, the proposed method is intuitive, easy to understand graphically and accurate as much as the state of the arts.

This paper consists as follows. In section 2, we introduce a general rectification method. In section 3, we describe a camera model for rectification and epipolar geometry. In section 4, a relevant binocular rectification method is addressed. We propose a new trinocular rectification method in section 5. Experimental results are shown in section 6. And finally conclusions are drawn in section 7.

2 Image Rectification

In the case of 3D reconstruction using more than two images, we should find out corresponding points among images first. If the corresponding points lie on the same parallel line to one of the two image axes, we can expect to find the corresponding points quickly. In order to match the corresponding points in this manner, we have to transform the images so that they posses this property and this procedure is called by image rectification. Note that in general, the trinocular stereo is more accurate than the binocular stereo, however, the trinocular rectification procedure is more complicate than the binocular one. In this paper, we propose an efficient and simple method that can produce same or better result than the existing complicated trinocular rectification methods.

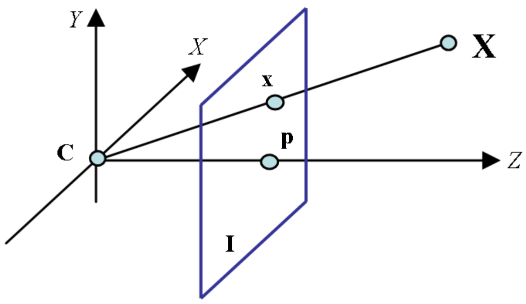

Fig. 1: Camera model

3 Camera Model and Epipolar Geometry

3.1 Camera Model

To understand the image rectification, we need to know how the real world 3D points are mapped to the image plane. In general, we use a pin-hole camera model as shown in Fig.1.

By a pin-hole camera model, we assume a center of projection $\mathbf{C}$, which is the homogeneous 3D point, and define the image plane $\mathbf{I}$. In the real world, a homogeneous 3D point $\mathbf{X}$ composed of $(X, Y, Z, 1)^T$ is mapped to a homogeneous point $\mathbf{x}$ on image plane where a line joining the point $\mathbf{X}$ to the center of projection meets the image plane. The center of projection $\mathbf{C}$ is called camera center. The line perpendicular to the image plane via the camera center is called principal axis and is represented by Z-axis in Fig.1. The imaged point in the principal axis is the principal point $\mathbf{p}$ as shown in Fig.1.

The relationship between a 3D point $\mathbf{X}$ and an imaged point $\mathbf{x}$ is as follows.

$$\mathbf{x} \cong \mathbf{PX} \tag{1}$$

In this case, projection matrix $\mathbf{P}$ is 3 by 4 linear transform matrix and it can be decomposed by an internal parameter matrix $\mathbf{K}$ and the external parameter matrix $\mathbf{R}$ and $\mathbf{t}$ as in (2).

$$\mathbf{P} = \mathbf{K}[\mathbf{R}|\mathbf{t}] = \mathbf{KR}[\mathbf{I}|\text{-}\widetilde{\mathbf{C}}] \tag{2}$$

where $\widetilde{\mathbf{C}}$ denotes an inhomogeneous 3D point equal to the camera center $\mathbf{C}$, and $\mathbf{I}$ describes the 3x3 identity matrix. If we assume that $\mathbf{KR}$ is $\mathbf{Q}$ and $\mathbf{X} = (\widetilde{\mathbf{X}}^T, 1)$, then the inhomogeneous 3D point $\widetilde{\mathbf{X}}$ is expressed as follows.

$$\widetilde{\mathbf{X}} = \widetilde{\mathbf{C}} + \lambda \mathbf{Q}^{-1}\mathbf{x}, \lambda \in \mathbf{R} \tag{3}$$

where λ denotes scale factor.

3.2 Epipolar Geometry

To explain the image rectification precisely, we should mention epipolar geometry for two pin-hole cameras. As shown in Fig.2, the world 3D point $\mathbf{X}$ is mapped to the corresponding image points $\mathbf{x}_1$ and $\mathbf{x}_2$ on the image planes $\mathbf{I}_1$ and $\mathbf{I}_2$ respectively for the each camera center $\mathbf{C}_1$, $\mathbf{C}_2$. These two points are corresponding point pair each other. In Fig.2, the line between two camera centers is baseline, and the points on the baseline which meets the each image plane are epipoles $\mathbf{e}_1$, and $\mathbf{e}_2$ respectively. Also, the line between the imaged point $\mathbf{x}$ and epipole $\mathbf{e}$ is epipolar line. By this geometrical property, we can obtain the corresponding epipolar line $\mathbf{l}$ on the other image plane. We can find the corresponding point by searching along the epipolar line on the other image plane. To alleviate the computational complexity, we can make the epipolar line parallel to one of the image plane axes. This process is called image rectification.

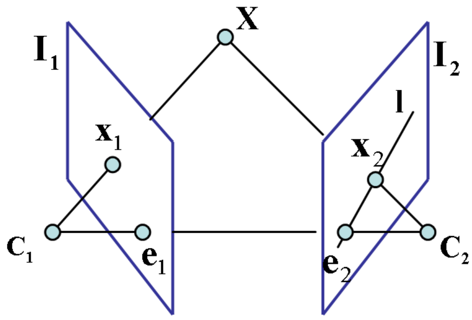

Fig. 2: Epipolar geometry

4 Binocular Rectification

In order to make the epipolar line parallel to an axis of the image plane, we should move the epipoles to infinity. In other words, principal axes of two cameras should be orthogonal to the baseline and look out the same direction. To rectify the given images, we should find the baseline and produce the new projection matrices satisfying the rectification conditions. In this procedure, we have to make sure that internal parameters of each camera are same. Here is the summary.

1. Produce new images that is parallel to the baseline and face the same direction.

2. Find the new camera matrices which have the same internal camera parameters.

Detail algorithm is as follows. First of all, we assume that the value of each camera parameters are known, where we can compute them by using real 3D object such as plane pattern or 3D rig. With known camera parameters, we can obtain the original camera matrices $\mathbf{P}_1$ and $\mathbf{P}_2$. Also, we can produce the new camera matrices $\mathbf{P}'_1$ and $\mathbf{P}'_2$ that are satisfying the rectification con-

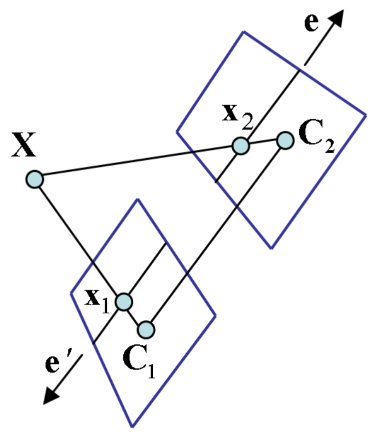

Fig. 3: Binocular rectification

straints. The detail equation for new camera matrix as in (4).

$$\mathbf{P}'_1 = \mathbf{K}'\mathbf{R}'[\mathbf{I}|\text{-}\widetilde{\mathbf{C}}_1], \mathbf{P}'_2 = \mathbf{K}'\mathbf{R}'[\mathbf{I}|\text{-}\widetilde{\mathbf{C}}_2] \tag{4}$$

where inhomogeneous camera center points $\widetilde{\mathbf{C}}_1$ and $\widetilde{\mathbf{C}}_2$ are derived from $\mathbf{P}_1$ and $\mathbf{P}_2$, respectively.

Now, the unknown values are $\mathbf{K}'$ and $\mathbf{R}'$. The internal parameter matrix $\mathbf{K}'$should be same for new projection matrices, and we can assume the specific one similar to original internal matrix $\mathbf{K}$ from $\mathbf{P}_1$, $\mathbf{P}_2$, arbitrarily. $\mathbf{R}'$ can be obtained as follows.

$$\mathbf{R}' = \begin{bmatrix} \mathbf{r}'^T_1 \\ \mathbf{r}'^T_2 \\ \mathbf{r}'^T_3 \end{bmatrix} \tag{5}$$

where $\mathbf{R}'$ is a rotation matrix which transforms the original 3D cartesian coordinates system to new one. In order to ensure the condition of rectification, $\mathbf{R}'$ is constrained as follows.

1. The new X-axis is parallel to baseline.

$$\mathbf{r}' = \frac{(\widetilde{\mathbf{C}}_1 - \widetilde{\mathbf{C}}_2)}{||\widetilde{\mathbf{C}}_1 - \widetilde{\mathbf{C}}_2||} \tag{6}$$

2. The new Y-axis is orthogonal to X-axis and arbitrary unit vector $\mathbf{k}$.

$$\mathbf{r}'_2 = \mathbf{k} \times \mathbf{r}'_1 \tag{7}$$

3. The new Z-axis is orthogonal to new X and new Y-axis simultaneously.

$$\mathbf{r}'_3 = \mathbf{r}'_1 \times \mathbf{r}'_2 \tag{8}$$

where the arbitrary unit vector $\mathbf{k}$ is set to the unit vector of Z-axis for one of the cameras. By this unit vector $\mathbf{k}$, we can obtain the new Y-axis orthogonal to the new X-axis and orizinal Z-axis.

Using the relationship between each world 3D point and the corresponding imaged point, we can come up with the relationship between unrectified and rectified images as follows.

$$\begin{cases} \mathbf{x} = \mathbf{P}\mathbf{X} \\ \mathbf{x}' = \mathbf{P}'\mathbf{X} \end{cases} \tag{9}$$

By (3), we can lead the new equations.

$$\begin{cases} \widetilde{\mathbf{X}} = \widetilde{\mathbf{C}}_i + \lambda \mathbf{Q}_i^{-1}\mathbf{x}_i \\ \widetilde{\mathbf{X}}' = \widetilde{\mathbf{C}}_i + \lambda' \mathbf{Q}_i'^{-1}\mathbf{x}'_i \end{cases} ,i\text{=1,2} \tag{10}$$

By (10), we can derive the following equation.

$$\mathbf{x}'_i = \lambda'' \mathbf{Q}'_i \mathbf{Q}_i'^{-1} \mathbf{x}_i = \lambda'' \mathbf{H}_i \mathbf{x}_i \ ,i\text{=1,2} \tag{11}$$

Finally, we are able to complete the rectification procedure when the homographies $\mathbf{H}_i$ drawn in (11) are applied to the each original image, where λ, λ' and λ'' denote scaling factors.

5 Trinocular Rectification

In the previous section, we reviewed a general binocular rectification method. In this section, we will introduce a proposed trinocular rectification method with following condition.

The center of each camera is on a line or on a plane.

If all cameras satisfy this condition, then we can rectify multiple images. However, this condition is hardly satisfied in the real world situation. So, as an alternative, we would show a new condition for the trinocular rectification.

1. Obtained images are parallel to a plane which is composed of three camera centers.
2. Internal parameters of all cameras are same.
3. The angle of diagonal epipolar line should be 45 degrees.

The reason of third constraint is that the disparity between left and right images d_{RL} should be same as the distance between upper and left images d_{RT}, which denotes $d_{RT} = d_{RL}$. By this constraint, we can expect images that are convenient for the trinocular stereo vision algorithm. Fig.4 shows the constraints.

The first step of trinocular rectification is to obtain $\mathbf{R}'$ which makes principle axis of each camera to perpendicular one for a plane composed of three camera centers. This method is similar to the binocular rectification case. The $\mathbf{R}'$ in (5) can be obtained as follows.

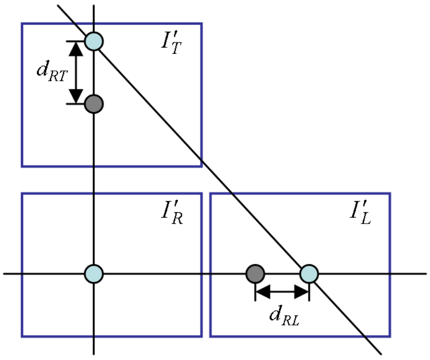

Fig. 4: Vertically and horizontally rectified images

1. The new X-axis is parallel to the baseline between left and right camera.

$$\mathbf{r}'_1 = \frac{(\widetilde{\mathbf{C}}_1 - \widetilde{\mathbf{C}}_2)}{\|\widetilde{\mathbf{C}}_1 - \widetilde{\mathbf{C}}_2\|} \tag{12}$$

2. The new Z-axis is perpendicular to a plane composed of three camera centers.

$$\mathbf{r}'_3 = \frac{(\widetilde{\mathbf{C}}_1 - \widetilde{\mathbf{C}}_3) \times (\widetilde{\mathbf{C}}_1 - \widetilde{\mathbf{C}}_2)}{\|(\widetilde{\mathbf{C}}_1 - \widetilde{\mathbf{C}}_3) \times (\widetilde{\mathbf{C}}_1 - \widetilde{\mathbf{C}}_2)\|} \tag{13}$$

3. The new Y-axis is perpendicular to new X and new Z-axis, simultaneously.

$$\mathbf{r}'_2 = \mathbf{r}'_1 \times \mathbf{r}'_3 \tag{14}$$

Using this $\mathbf{R}'$ and arbitrary $\mathbf{K}'$, we can obtain the new projection matrix $\mathbf{P}'$ as follows.

$$\mathbf{P}'_i = \mathbf{K}'\mathbf{R}'[\mathbf{I}| - \widetilde{\mathbf{C}}_i] \quad ,i\text{=left,right,top} \tag{15}$$

Additionally, applying (11), we can estimate homographies for rectification: $\mathbf{H}_{\text{left}}, \mathbf{H}_{\text{right}}, \mathbf{H}_{\text{top}}$. The result of transformation by these rectification homographies is shown in Fig.5.

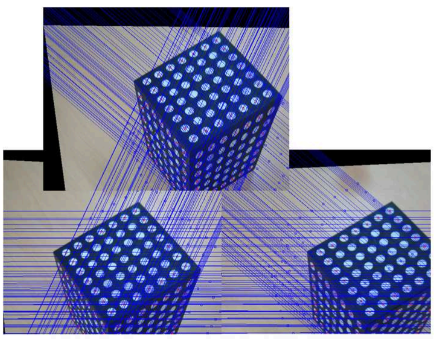

Fig. 5: Epipoles at infinity

Now, we are having rectified images parallel to a plane composed of three camera centers. Since each epipole in the images is a point at infinity, all epipolar lines are parallel to each other. So far, we can rectify images which satisfy constraints number 1 and 2. Finally, in order to rectify the trinocular images with constraint number 3, we should correct the normal triangle to the right-angled isosceles triangle as shown in the Fig.6.

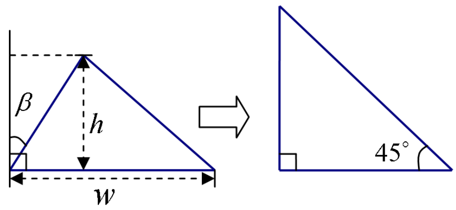

Fig. 6: Affine transform for trinocular rectification

We can draw the nomal triangle from 3 camera centers. In Fig.6, w is length between the left and right camera center, and h is distance between the top camera center and baseline from the left and right camera. β is the angle depicted in Fig.6. It should be zero after rectification. The affine transform matrix named $\mathbf{H}_{\text{affine}}$ is as follows.

$$\mathbf{H}_{\text{affine}} = \begin{bmatrix} 1 & \tan\beta & 0 \\ 0 & w/h & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{16}$$

Applying this homography to all images equivalently, we can finish the trinocular rectification.

6 Experimental Result

For experiments, we took three different pictures [resolution: 640×480] of which we knew the 3D real information as shown in Fig.7. Then we obtained camera parameters from 3D information and corresponding points on images. Applying the proposed algorithm, we could produce trinocular rectified image as shown in Fig.8.

(a) (b) (c)

Fig. 7: Before rectification (a) right (b) left (c) top

Table 1. shows a result of rectification error for all corresponding points. We set the error value of rectification to the perpendicular distance between each epipolar line and each corresponding point. The reason of which the error of R-L part is big relative to the others caused

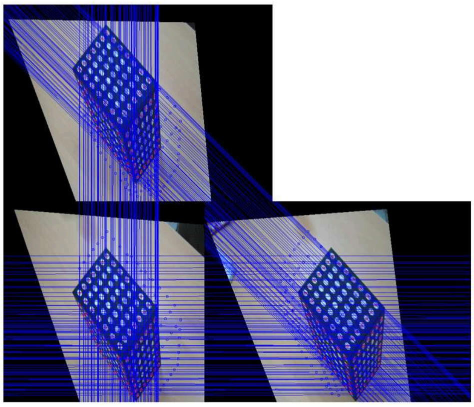

Fig. 8: After rectification

Table 1: Distance error for each corresponding point (in pixel)

	R-L	R-T	L-T
Average	0.113021	0.0758203	0.0730634
Minimal Error	0.001007	0.0007324	0.0012408
Maximal Error	0.529724	0.2337340	0.3106760

by amplifying the original error through the last affine transform. The result shows that the error value is dependent on the location of three cameras when the pictures were taken. In order to assure the result accurate, we should position the cameras to make a right-angled isosceles triangle or use more accurate camera parameter extraction algorithm [10].

7 Conclusion

In this paper, we proposed an efficient trinocular camera rectification method for trinocular stereo vision. Although previous works tried to solve this problem by using mathematically, geometrically complicated method, the proposed method is simpler and intuitive than the previous works. More than that, the accuracy of the algorithm is good or better than the previous ones.

Acknowledgment This research is supported by University IT Research Center (ITRC) supporting program by Ministry of Information and Communication and Agency for Defense Development through Korea and Image information special research center in KAIST (Korea Advanced Institute of Science and Technology).

References

[1] A. Fusiello, E. Trucco, and A. Verri.,"A compact algorithm for rectification of stereo pairs. Machine Vision and Applications", 12(1):16-22, 2000

[2] Ayache, N. and Hansen, C.,"Rectificaton of Images for Binocular and Trinocular Stereovision", Proc. 9th International Conference on Pattern Recognition, Rome, pp.11-161,1988

[3] An, L., Jia, Y., Wang, J., Zhang, X. And Li, M., "An efficient rectification method for trinocular stereovision", Proc. Int. Conf. on Pattern Recognition, vol. 4, Cambridge, pp.56-59, 2004.

[4] Sun, C., "Uncalibrated Three-View Image Rectification", Image and Vision Computing, vol.21, no. 3, pp.259-269, 2003

[5] Heinrichs, M. and Rodehorst, V.,"Trinocular Rectification For Various Camera Setups", Symposium of ISPRS Commission III - Photogrammetric Computer Vision PCV'06, pp. 43-48, Bonn, Germany, 2006

[6] R. Hartley and A. Zisserman, "Multiple View Geometry 2nd Edition," Cambridge, 2003

[7] Matousek, M., Sara, R. and Hlavac, V., "Data-optimal rectification for fast and accurate stereo vision", 2004

[8] Roy, S., Meunier, J. and Cox, I., "Cylindrical rectification to minimize epipolar distortion", Proc. Computer Vision and Pattern Recognition, San Juan, pp. 393-399, 1997

[9] Oram, D.T., "Rectification for any epipolar geometry", Proc. British Machine Vision Conf., London, pp. 653-662, 2001

[10] R.Y.Tsai, A Versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses, IEEE J. Robotics and Automation, Vol.3, No.4, pp.323-344, Aug. 1987

Young-Ki Baik: received the B.S. and M.S. degree in electronic engineering from Hong-Ik University, Seoul, Korea, in 1999 and 2004, respectively. He is currently pursuing the Ph.D. degree at Seoul National University. His research interest includes stereo vision, 3D reconstruction.

Jonghyun Choi: received the B.S. degree in electrical engineering from Seoul National University, Seoul, Korea, in 2003. He is currently pursuing the M.S. degree in Computer Vision Laboratory at Seoul National University since 2006. His research interest includes 3D reconstruction.

Kyoung-Mu Lee: received the B.S. and M.S. Degrees in Control and Instrumentation Eng. from Seoul National University, Seoul, Korea in 1984 and 1986, respectively, and Ph. D. degree in Electrical Engineering from the USC, Los Angeles, California in 1993. He was with the Samsung Electronics Co. Ltd. in Korea as a senior researcher from 1994 to 1995. On August 1995, he joined the department of Electronics and Electrical Eng. of the Hong-Ik University, and worked as an assistant and associate professor. From September 2003, he is with the school of Electrical and Computer Engineering at Seoul National University as an associate professor. His primary research interests include computational vision, image understanding, pattern recognition, man-machine interface and multimedia applications.