

Efficient Stereo Image Rectification Method Using Horizontal Baseline

Yun-Suk Kang and Yo-Sung Ho

School of Information and Communications
Gwangju Institute of Science and Technology (GIST)
261 Cheomdan-gwagiro, Buk-gu, Gwangju 500-712, Republic of Korea
{yunsuk, hoyo}@gist.ac.kr

Abstract. In this paper, we propose an efficient stereo image rectification method using the horizontal baseline. Since the stereo camera is generally manually arranged, there are geometric errors due to the camera misalignment and the differences between the camera internal characteristics. Although the conventional calibration-based stereo image rectification method is simple, it has an opportunity to provide the results that have some visual distortion such as image skewness. Therefore, the proposed method calculates the baseline for stereo image rectification, which is parallel to the horizontal line in the real world. Using this baseline, we estimate the camera parameters and the rectification transform. By applying the transform to the original images, we obtain the rectified stereo images. Experimental results show that the results of the proposed method provide the better rectified stereo image without visual distortion.

Keywords: Image rectification, stereo image, stereo camera, 3DTV.

1 Introduction

Three-dimensional (3D) TV provides us more realistic video contents than the current two-dimensional (2D) television broadcasting. Since the input signal of 3DTV is composed of more than single viewpoint images or videos, users can watch the scene with immersive feeling. In recent years, much research on 3DTV and 3D content generation has been investigated to satisfy the increasing demands for realistic multimedia services in the world [1].

In order to generate 3D contents for 3DTV, at least two view images are required basically. Two cameras, called the stereo camera, capture a 3D scene or object in the real world from two different positions. Users watch this stereo image with 3D sense with stereoscopic displays. Moreover, from this stereo image, we can estimate the scene's depth information using stereo matching [2], and also generate novel view images based on the depth.

However, there is a constraint to use stereo images for 3D applications. Two image planes of the stereo camera determine their epipolar geometry that satisfies the epipolar constraint between two images. Epipolar constraint is that a point in one image has its corresponding points in the other image along an epipolar line.

Therefore, if the epipolar lines in each image plane are parallel, the corresponding points have the same vertical coordinates. In other words, there is no vertical pixel difference between two images. In this case, the stereo image has only horizontal displacement. The visual quality of the image as the 3D contents increases and also the stereo matching process becomes very simple [3]. Unfortunately, the practical stereo image captured by a manually arranged stereo camera does not have the parallel epipolar lines. There are not only position and orientation differences but also internal parameter differences between two cameras.

In order to solve these problems in stereo images, we perform image rectification. Image rectification is rotation and movement of two image planes that makes epipolar lines parallel each other. The rectified stereo image is then considered as the images captured by two physically-equal cameras with only horizontal camera interval.

Image rectification has been studied for long time. There are two categories; one is based on the image features [4] [5], and the other is calibrated case [6] [7]. Recently, rectification has been extended to cover the multiple views [8]. In general, the result of image feature based rectification has some visual distortion such as image skewing. It is also influenced by the extracting features. While the calibration based image rectification gives more stable results and rectified camera parameters which is essential information for 3D applications. However, the stereo cameras have to be calibrated before the rectification, and the reliability of calibration is also influence on the results.

In this paper, we explain a stereo image rectification method using the horizontal baseline. We introduce the stereo geometry briefly in Section 2. In section 3, we explain the proposed method. After scene capturing and camera calibration, we calculate the horizontal baseline for rectification. Using this baseline, we estimate the rectification transform. By applying this transform to the original stereo image, we obtain the rectified stereo image. After showing the experimental results in Section 4, we conclude this paper in Section 5.

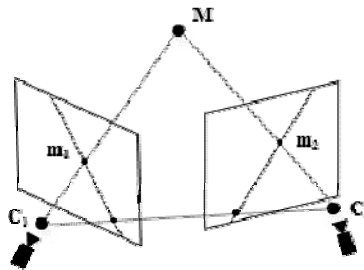


Fig. 1. Stereo geometry

2 Stereo Geometry

Figure 1 shows the geometry of stereo image. There are two camera centers C_1 and C_2 , and the object at the point M in the 3D space. These three points are in the world

coordinate system, and two cameras also have their own coordinate systems. Each camera coordinate system has three axes; vertical, horizontal, and principal axes. The principal axis is also called the optical axis, which indicates the optical ray direction of the camera.

By operating two cameras, the object located on \mathbf{M} is projected to the image points \mathbf{m}_1 and \mathbf{m}_2 in each image plane. The corresponding point of \mathbf{m}_1 in the right image plane has to be on the epipolar line. The epipolar line of the right image plane is defined as the intersecting line between the right image plane and the plane described by \mathbf{M} , \mathbf{C}_1 , and \mathbf{C}_2 , which is called the epipolar plane.

Figure 2 shows the geometry of the rectified stereo image that has the parallel epipolar lines in each image plane. All the points in the Fig. 2 excluding \mathbf{M} are changed and also each image plane is rotated and moved. Therefore, two image planes and epipolar lines are parallel to the line through \mathbf{C}'_1 and \mathbf{C}'_2 which is called the baseline. In this case, two image points \mathbf{m}'_1 and \mathbf{m}'_2 have the same vertical coordinates. It means that there is no vertical pixel displacement between two corresponding points.

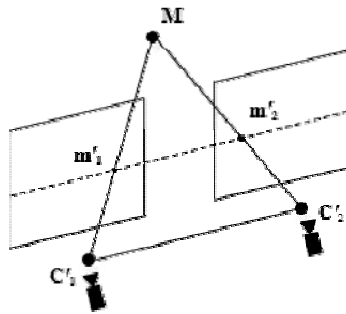


Fig. 2. Rectified stereo geometry

3 Proposed Stereo Image Rectification Method

In this section, we explain our proposed method to rectify the stereo image. In the conventional method based on the camera parameters [6], the baseline for rectification is determined as the line through \mathbf{C}_1 and \mathbf{C}_2 in Fig. 1. Then, \mathbf{C}'_1 and \mathbf{C}'_2 are the same as \mathbf{C}_1 and \mathbf{C}_2 after rectification. If this baseline is not parallel to the horizontal line in the real world, the rectified image can be skewed with respect to the users' view.

Therefore, the proposed method calculates the baseline that is parallel to the horizontal line in the real world. Figure 3 shows the procedure of the proposed method. After scene capturing, we estimate the camera parameters by camera calibration [9]. By using these camera parameters, we calculate the baseline which is parallel to the real horizontal line, and then we estimate the camera parameters of the

rectified stereo image based on the baseline. Finally, we obtain the rectified stereo image by applying the rectification transform to the captured images. This transform is computed using both of the original and estimated camera parameters.

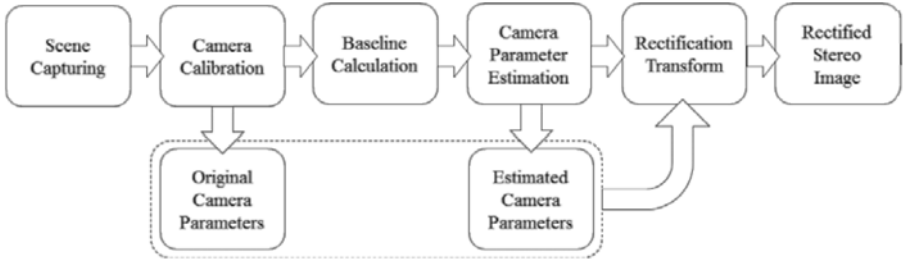


Fig. 3. Procedure of the proposed method

3.1 Baseline Calculation

After obtaining camera the parameters by camera calibration, we calculate the baseline. Baseline calculation begins with the initial line which is obtained by connecting the two camera centers. From this initial line, we can calculate the baseline. The baseline must satisfy the following two conditions. First, this baseline and the initial line are on the same plane that has its normal vector as the direction of the new principal axis. The new principal axis is determined as the direction orthogonal to both of the initial line and the average direction of all the original vertical axes. It means that the baseline can preserve the orientation of the camera array which is obtained based on camera positions.

The second condition is that the baseline is parallel to the horizontal line in the real world. It guarantees that the rectified stereo image according to this baseline does not have the skew problem. In order to obtain such a baseline, we use a line image projection algorithm that requires an image containing a short and non-tilted line like Fig. 4(a). Through the line image projection, we can measure the slope of the initial line, and then we can calculate a suitable correction vector to make the baseline parallel to the real horizontal line.

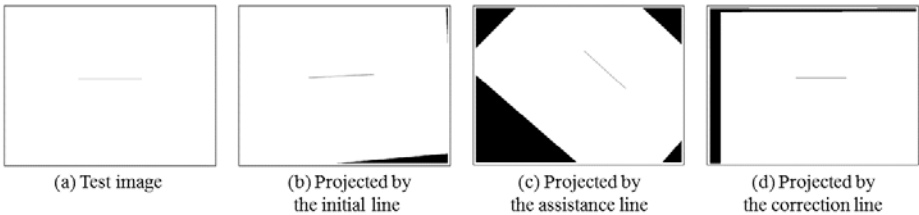


Fig. 4. Line image projection

In order to measure the slope of the initial line, we assume that the line image is left of the stereo view. We then project this image so that the horizontal axis of the image plane is parallel to the initial line. As a result, the projected image has the line tilted as the slope of the initial line as indicated in Fig. 4(b). We can measure the slope of this line by counting the number of pixels between the start-point and the end-point of the line. This measured value means the slope of the initial line.

After measuring the slope of the initial line, we need a correction vector to obtain the baseline that satisfies the second condition. Figure 5 shows how to calculate the correction vector. The cross product between the new principal axis and the initial line vector makes the orthogonal vector. The correction vector is then calculated as the sum of the initial line vector and the orthogonal vector.

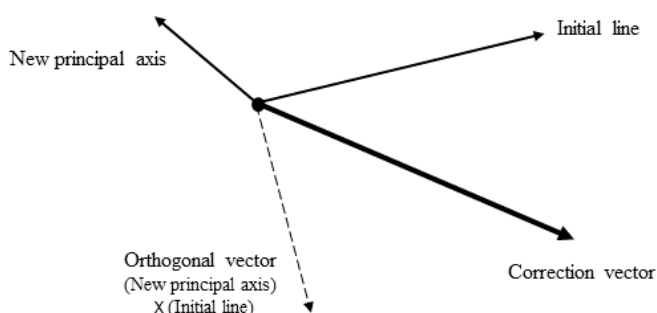


Fig. 5. Correction vector calculation

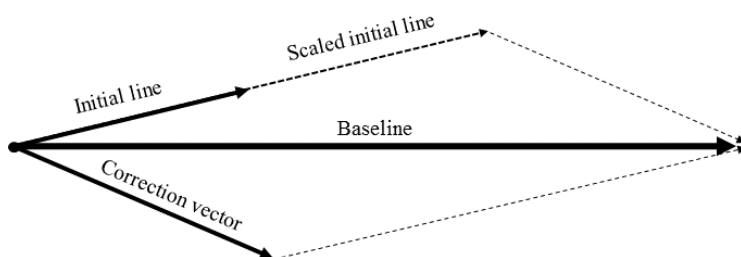


Fig. 6. Baseline calculation

In order to calculate the baseline, we measure the slope of the correction vector. By using the line image projection again, the projected image in accordance with the correction vector has an opposite slope to the initial line like Fig. 4(c). Finally, we can calculate the baseline which is parallel to the real horizontal line by summing the initial line vector and the correction vector with a proper scale factor. Let i and c be

the slopes of the initial line vector and the correction vector, respectively. The scale factor s is defined as the ratio of i and c . The baseline vector $\vec{\mathbf{b}}$ is then calculated as Eq. 1 where $\vec{\mathbf{i}}$ and $\vec{\mathbf{c}}$ mean the initial line vector and the correction vector, respectively. Figure 4(d) shows the baseline that is parallel to the real horizontal line. This process is indicated in Fig. 6.

$$\vec{\mathbf{b}} = \begin{cases} s \cdot \vec{\mathbf{i}} + \vec{\mathbf{c}} & (s = c/i, \text{ if } c > i) \\ \vec{\mathbf{i}} + s \cdot \vec{\mathbf{c}} & (s = i/c, \text{ if } i > c) \end{cases} \quad (1)$$

3.2 Camera Parameter Estimation

After calculating the baseline, we estimate the rectified camera parameters. We firstly find the new camera centers. In the proposed method, the left camera center is considered as the reference and we estimate the new camera center of the right camera. Then the new camera center of the right camera is defined as a point that is apart with the user-input camera distance along the direction of the baseline.

After that, we consider the camera rotation matrices. We estimate each camera rotation matrix that satisfies the following conditions. The horizontal axis of every image plane becomes parallel to the baseline vector. All the principal axes are defined in common as the direction perpendicular to both of the baseline vector and the average of all the original vertical axes. Then, the vertical axis of each image plane is orthogonal to both of the new principal axis and the baseline vector. Thus, the rotation matrix for the rectified stereo camera \mathbf{R}' has the form shown in Eq. 2, where $\vec{\mathbf{b}}$ and $\vec{\mathbf{v}}_a$ mean the directions of the baseline vector and the average of all the original vertical axes, respectively.

$$\mathbf{R}' = \begin{bmatrix} \vec{\mathbf{b}}^T \\ ((\vec{\mathbf{b}} \times \vec{\mathbf{v}}_a) \times \vec{\mathbf{b}})^T \\ (\vec{\mathbf{b}} \times \vec{\mathbf{v}}_a)^T \end{bmatrix} \quad (2)$$

Then, we estimate the common camera intrinsic parameters. The focal length and the principal point are obtained as the averages of their original values, respectively. The same focal length of each camera makes all image planes coplanar. There are also uniform horizontal displacement between corresponding points and few vertical mismatches in pixels between corresponding points due to the same principal point of each camera. Finally, we obtain the rectified camera projection matrices which are composed of the estimated camera parameters like Eq. 3.

$$\mathbf{P}'_k = \mathbf{A}' [\mathbf{R}' | \mathbf{t}'_k] = \mathbf{A}' [\mathbf{R}' | -\mathbf{R}' \mathbf{C}'_k] \quad (3)$$

3.3 Rectification Transform

For the last step, we can generate the rectified stereo image by calculating and applying the rectification transform. We consider the epipolar geometry for each viewpoint. Then, we use the point-to-point mapping between images of the original and estimated cameras called the 2-D homography \mathbf{H} [10]. Finally, the transform for k -th image is obtained by using this homography like Eq. 4. By applying this transform to each image, we can obtain the rectified stereo image.

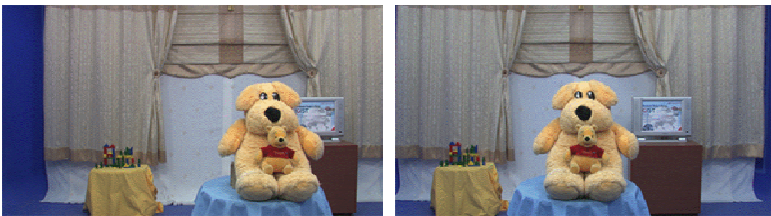
$$\mathbf{T}_k = \mathbf{H}_{\pi,k} = \mathbf{P}'_k \mathbf{P}_k^+ \quad (4)$$

4 Experimental Results

For experiments, we captured two sets of stereo image. We used two types of cameras; one provides 1024x768 and the other provides full HD (1920x1080) resolution. The first test images are shown in Fig. 7(a), the distance between two cameras is about 6.5cm. The second test images are shown in Fig. 7(b), the distance between two cameras is 40cm in this case. Figure 8(a) and Fig. 8(b) show that the captured images have the practical stereo geometry which is shown in Fig. 1. There were vertical pixel displacement between corresponding points, and also we notice the camera rotation difference.



(a) Yut-game(1024x768)



(b) Bear(1920x1080)

Fig. 7. Captured images

Figure 9 shows the rectified result by the conventional method [6]. We notice that the rectified images are skewed. This skewness is due to the skewed baseline connecting the original C_1 and C_2 . Especially for the second image set, although there is a little geometrical misalignment in the original images, the rectified results have more visual distortion.

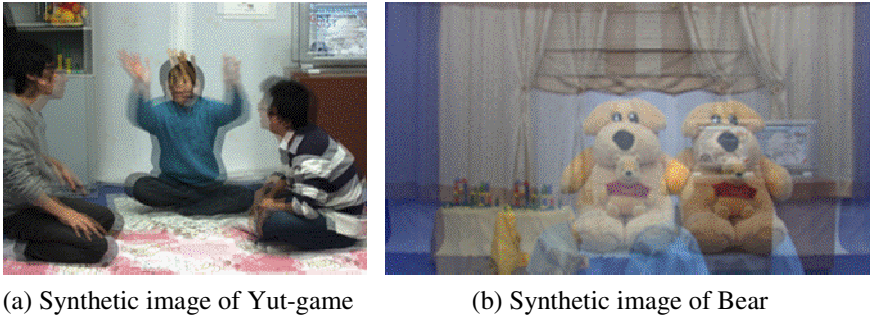
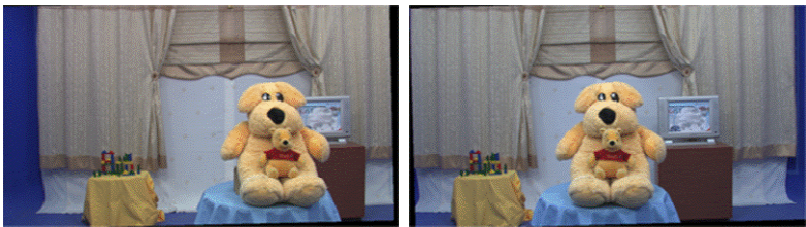


Fig. 8. Synthetic images for the captured stereo images



(a) Yut-game



(b) Bear

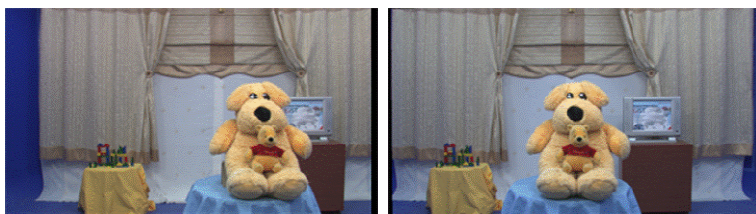
Fig. 9. Skewed result images

Figure 10 shows the rectified stereo images and their synthetic images by the proposed method. As shown in Fig. 10, the result images are not only rectified well but also almost parallel to the horizontal line in the real world. It is because the

baseline for rectification is calculated to be parallel to the real horizontal line. Also, as shown in Fig. 10(c) and Fig. 10(d), there are few pixels of vertical difference between corresponding pixels.



(a) Yut-game



(b) Bear



(c) Synthetic image of Yut-game



(d) Synthetic image of Bear

Fig. 10. Results by the proposed method

5 Conclusion

In this paper, we presented a stereo image rectification method using the horizontal baseline. The proposed method avoids that the rectified images become skewed due to the miscalculated baseline. The baseline in the proposed method is calculated to be parallel to the horizontal line in the real world using the initial and correction vectors calculation. Therefore, the experimental results show that the results from the proposed method have less geometrical misalignment without the visual distortion compared to the conventional method.

Acknowledgements. This research was supported by the MKE(Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency) (NIPA-2011-(C1090-1111-0003)).

References

1. Smolic, A., Kauff, P.: Interactive 3D Video Representation and Coding Technologies. Proc. of IEEE, Spatial Issue on Advances in Video Coding and Delivery 93(1), 99–110 (2005)
2. Sun, J., Zheng, N.N., Shum, H.Y.: Stereo Matching Using Belief Propagation. IEEE Transactions on Pattern Analysis and Machine Analysis (PAMI) 25(5), 787–800 (2003)
3. ISO/IEC JTC1/SC29/WG11 M12030: Comments on Input and Output Format of MVC (2005)
4. Hartley, R.: Theory and Practice of Projective Rectification. International Journal of Computer Vision 35(2), 115–127 (1999)
5. Loop, C., Zhang, Z.: Computing Rectifying Homographies for Stereo Vision. In: Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 125–131 (1999)
6. Fusiello, A., Trucco, E., Verri, A.: A Compact Algorithm for Rectification of Stereo Pairs. Machine Vision and Application 12(1), 16–22 (2000)
7. Kang, Y., Lee, C., Ho, Y.: An Efficient Rectification Algorithm for Multi-view Images in Parallel Camera Array. In: Proc. of 3DTV Conference 2008, pp. 61–64 (2008)
8. Kang, Y., Ho, Y.: Geometrical Compensation for Multi-view Video in Multiple Camera Array. In: Proc. of International Symposium ELMAR, pp. 83–86 (2008)
9. Camera Calibration Toolbox for Matlab,
<http://www.vision.caltech.edu/bouguetj>
10. Hartley, R., Zisserman, A.: Multiple View Geometry in Computer Vision. Cambridge University Press (2003)