AN EFFICIENT RECTIFICATION ALGORITHM FOR MULTI-VIEW IMAGES IN PARALLEL CAMERA ARRAY

Yun-Suk Kang, Cheon Lee, and Yo-Sung Ho

Gwangju Institute of Science and Technology (GIST), Korea

ABSTRACT

In this paper, we present an efficient rectification algorithm for multi-view images captured by a parallel camera array. Since conventional stereo image rectification methods did not consider three or more cameras simultaneously, we propose an algorithm to rectify multi-view images at the same time. We calculate the common baseline considering all camera positions and apply the rectifying transformation defined by camera rotations and camera intrinsic parameters. From our experiments, we have found that the proposed method rectifies multi-view images efficiently. Rectified images have uniform horizontal disparities and no vertical mismatches between adjacent views.

***Index Terms*—**Image Rectification, Multi-view Rectification, Multi-view Video, 3DTV

1. INTRODUCTION

Unlike the single-view video, the multi-view video has advantages of providing multiple viewpoints to users by capturing a three-dimensional (3D) scene from different camera positions [1]. Recently, 3DTV [2] and free viewpoint TV (FTV) [3] have been discussed as main applications of the multi-view video. We need multi-view images and the corresponding depth maps to provide 3D videos or the multi-view video as inputs for these applications.

Generally speaking, multi-view images have horizontal disparities and vertical mismatches between neighboring views. We can generate depth maps from images having horizontal disparities. If multi-view images have not vertical mismatches but disparities only in the horizontal direction, we can generate high quality depth maps much faster than the case of having vertical mismatches. In addition, we can get a clear viewpoint change of the multi-view video. Therefore, we need to align vertical mismatches among views by performing image rectification.

Image rectification is a kind of image transformation. There are numerous rectification methods using camera parameters [4, 5] or homography transformations [6, 7], however, most of them consider stereo pairs. Stereo image rectification is a process that makes epipolar lines of two images captured at different positions parallel each other. Then, vertical coordinates of all image points of two images become equal, and there remain only horizontal disparities.

However, in the case of multi-view, we hardly get acceptable rectified results with previous stereo rectification methods. Since previous methods did not consider three or more views at once. Therefore, we need to extend previous methods and to develop a new rectification algorithm for multi-view images considering all images simultaneously.

In this paper, we propose an efficient rectification algorithm for multi-view images captured by a parallel camera array. For calibrated multi-view images, we calculate the common baseline considering all camera positions. We then apply the rectifying transformation to all images and obtain rectified images. Our proposed algorithm is performed clearly, and easily convertible to different camera arrays. Since the rectifying transformation are not related to camera positions but camera rotations and camera intrinsic parameters.

2. STEREO IMAGE RECTIFICATION USING CAMERA PARAMETERS

2.1. Camera model and camera parameters

We model a pinhole camera with its camera center C, which is the center of projection of the camera, and its image plane R that has the projected image of the scene. The line through the camera center and perpendicular to the image plane is called the principal axis. The focal length is the distance between the camera center and the image plane. The line through two camera centers is the baseline.

In a stereo camera model as shown in Fig. 1(a), a point W in the real world is projected onto two image planes, respectively. They are specified as M_1 and M_2. Let **w** and **m** be their coordinates. Then, the camera operation is modeled by the perspective projection, which is a linear transformation from the world coordinate to the image coordinate. It is given by a camera projection matrix **P**.

$$\mathbf{m}=\mathbf{P}\mathbf{w}=\mathbf{A}[\mathbf{R}|\mathbf{t}]\mathbf{w} \quad (1)$$

As indicated in Eq. (1), the camera projection matrix consists of camera parameters. Camera parameters have

978-1-4244-1755-1/08/$25.00 ©2008 IEEE 3DTV-CON'08, May 28-30, 2008, Istanbul, Turkey

three components which are the intrinsic parameters represented by the matrix **A** and the extrinsic parameters represented by the matrix **R** and the vector **t**.

The intrinsic parameters describe the internal features of the camera. There are the focal lengths, the skew parameters, and the principal points in the intrinsic parameters, and they are represented as the 3×3 matrix **A**.

The extrinsic parameters are composed of the 3×3 rotation matrix **R** and the translation vector **t**, indicating the camera orientation and position, respectively. The world reference frame and the camera reference frame are related via the rotation and translation. In other words, the rotation and translation indicate the transformation between the camera reference frame and the world reference frame.

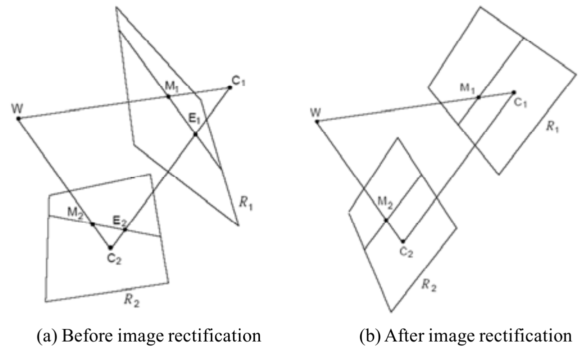

Fig. 1. Stereo camera model

2.2. Stereo image rectification

We assume that two cameras are calibrated, i.e., we already know two camera projection matrices. The line through two camera centers becomes the baseline. Then, we can get rectified images by applying the rectifying transformation to all images. As shown in Fig. 1(b), rectified images have their epipolar lines which are parallel each other and parallel to the baseline. It means that two image planes are coplanar.

The rectifying transformation can be calculated by solving relationships between original projection matrices and rectified projection matrices. We can define rectified projection matrices according to following conditions. For the rotation, the horizontal axis of each image plane is parallel to the baseline. Vertical axes are orthogonal to corresponding horizontal axes and newly defined principal axes. For the intrinsic parameters, we equalize each focal length and adjust vertical coordinates of two principal points. Then, two image planes become coplanar and have no vertical mismatches [4].

2.3. Difficulties in extending to multi-view images

It is hard to rectify multi-view images simultaneously using the stereo image rectification method since it merely considers two images at once. One possible strategy with this algorithm is that we group two adjacent views of multi-view, and rectify these image pairs several times.

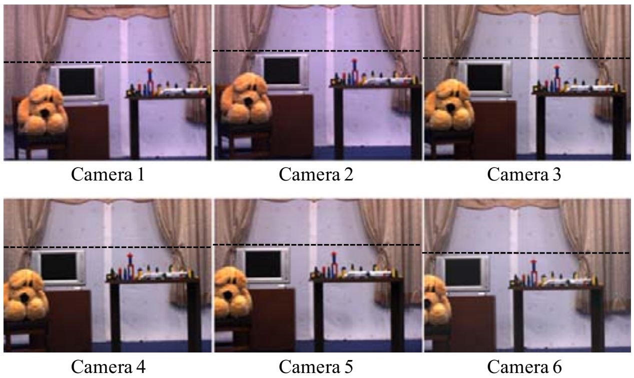

Fig. 2. Unrectified multi-view images

Figure 2 shows multi-view images captured by six cameras in parallel camera array. These six images have horizontal disparities and vertical mismatches. Horizontal disparities are caused by distances between cameras. Vertical mismatches can occur by different camera rotations or the camera misalignment.

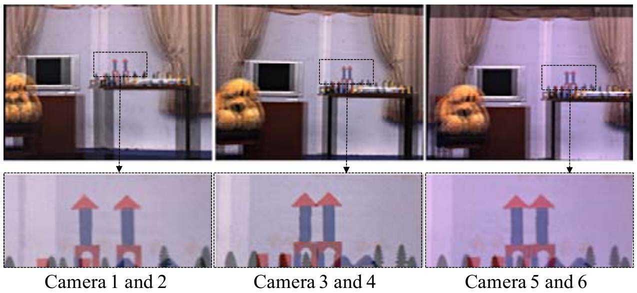

Fig. 3. No vertical mismatches in respective stereo pairs

We perform stereo image rectification three times to three pairs of two images like Fig. 3. For each stereo pair, we can see that images are aligned correctly. However, we cannot obtain overall rectified images but partially rectified images since the stereo algorithm rectifies three image pairs according to three different baselines. As shown in Fig. 4, we can notice that there are still vertical mismatches between neighboring stereo pairs.

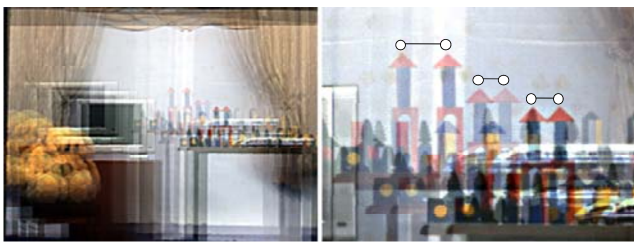

Fig. 4. Vertical mismatches among stereo pairs

3. RECTIFICATION FOR MULTI-VIEW IMAGES

In order to solve problems in applying previous stereo rectification algorithms to multi-view, M. Tanimoto and T. Fujii [8] calculate the baseline that minimizes the sum of squared distances to camera centers. They assume ideal virtual cameras on the baseline, and then obtain rectified multi-view images using transformation to virtual cameras.

In this section, we propose an efficient rectification algorithm. The proposed method is related to camera rotations and the intrinsic parameters without camera translations. It means that this method has clear procedures and we can easily extend to different camera arrays.

3.1. Calculation of the common baseline

In order to extend stereo image rectification to multi-view images, we need the common baseline. We use a simple algorithm to get the baseline. We find $N-1$ midpoints between adjacent two camera centers from N cameras. From the set of the first midpoints, we attain $N-2$ midpoints between neighboring two old midpoints in the set. We perform this process iteratively until the last two points are remained. Finally, the line through these two points becomes the common baseline. This line can be calculated easily and consider all camera positions simultaneously.

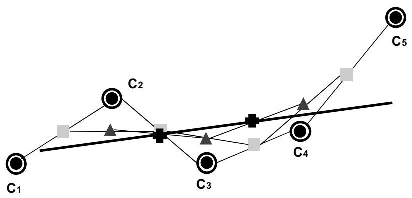

Fig. 5. Calculation of the common baseline

Figure 5 shows how to calculate the common baseline. Points C_1 to C_5 indicate camera centers and two cross marks are the last two midpoints. The line through these two points is defined as the common baseline.

3.2. Rectifying transformation

After obtaining the common baseline, we have to determine the rectifying transformation from original and rectified projection matrices. In order to get rectified projection matrices from original matrices, we consider camera rotations, and then camera intrinsic parameters.

First, we regulate camera rotations so that horizontal axes of all image planes become parallel to the baseline. Each vertical axis is perpendicular to their horizontal axis and newly defined principal axes. New principal axes are defined as the average direction of all original principal axes. Figure 6 shows the concept of regulating camera rotations.

Then, we adjust the intrinsic parameters. In order to align vertical mismatches, we replace the vertical coordinate of each principal point by their average value. We can obtain all image planes which have uniform distances between adjacent views by making distances between neighboring horizontal coordinates of all principal points identical.

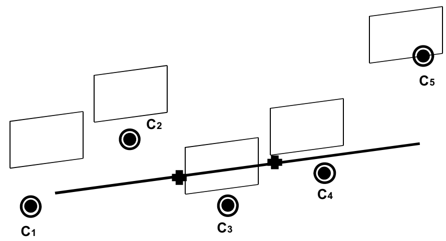

Fig. 6. Regulation of camera rotations

For the last step, we adjust all focal lengths so that each distance between the baseline and the image plane becomes equal. In order to calculate the distance from the baseline to each image plane, we use the orthogonal projection and the inner product of vectors.

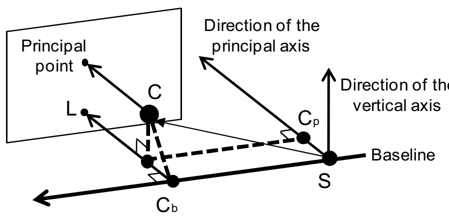

Fig. 7. Orthogonal projection of camera centers

As shown in Fig. 7, the camera center C is projected onto the baseline as C_b and on the line which is parallel to the principal axis as C_p. They can be evaluated by the orthogonal projection of the vector $\overrightarrow{CS}$. Then we can calculate the distance between the baseline and the image plane, $\overline{LC_b}$, by adding or subtracting the distance between C_p and S to the focal length. Addition and subtraction depend on the position of the camera center with respect to the baseline. It can be easily defined since the inner product tells us the angle between the vector $\overrightarrow{CC_b}$ and the direction of the principal axis. Finally, we obtain the length between the baseline and each image plane. We now adjust each focal length so that they have the length of the longest focal length. Then, rectified images become coplanar and have uniform distances between neighboring views as shown in Fig. 8.

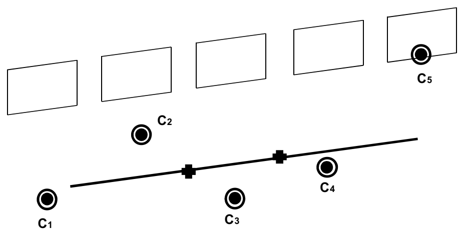

Fig. 8. Adjustment of the intrinsic parameters

4. EXPERIMENTAL RESULTS

We apply the proposed method to multi-view test sequences to confirm that the algorithm performs rectification properly. As already shown in Fig. 2, these images are captured by six cameras in parallel camera array. The camera model is FLEA-HICOL-CS. The image resolution is 1024×768 in pixels and distances between cameras are about 40 mm. Figure 9 shows that test sequences have non-uniform horizontal disparities and vertical mismatches.

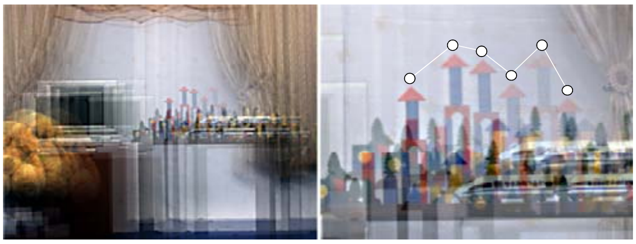

Fig. 9. Horizontal disparities and vertical mismatches

After applying the rectifying transformation, we can see that multi-view images are rectified like Fig. 10, respectively. Although there are some black regions and distortions in each image due to the transformations, we can notice that the proposed method can rectify all images correctly as presented in Fig. 11. There are no vertical mismatches among views, but uniform horizontal displacements due to camera positions and rotations.

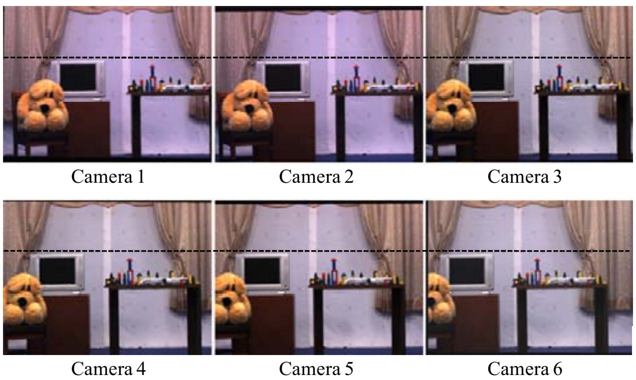

Fig. 10. Rectified multi-view images

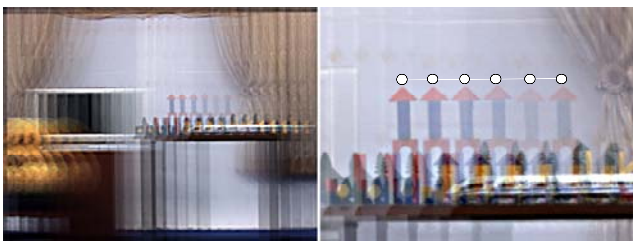

Fig. 11. Aligned in horizontal and vertical directions

5. CONCLUSION

In this paper, we have presented the efficient rectification algorithm for multi-view images in parallel camera array. The proposed method calculates the common baseline by iterative midpoint connecting algorithm, and aligns multi-view images in horizontal and vertical directions by applying the rectifying transformations. It also makes all image planes coplanar. Finally, we can get the results as though multi-view images captured in well arranged parallel camera array. From this algorithm, we can make preprocessed inputs for many multi-view video applications such as 3DTV. We can also generate depth maps much fast and correct, and enjoy a clear viewpoint change with rectified multi-view images. Our work is able to extend to rectify arbitrary camera arrays because the algorithm is independent on camera positions. We believe that our method can be widely used as a preprocessing method in the multi-view video.

ACKNOWLEDGEMENTS

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

REFERENCES

[1] ISO/IEC JTC1/SC29/WG11 N6909, “Survey of Algorithms used for Multi-view Video Coding (MVC),” Hong Kong, 2005.

[2] C. Fehn, R. de la Barre, and S. Pastoor, “Interactive 3-DTV Concepts and Key Technologies,” *Proceedings of the IEEE, Special Issue on 3-D Technologies for Imaging & Display,* vol. 94, no. 3, pp. 524-538, 2006.

[3] ISO/IEC JTC1/SC29/WG11 M8595, “FTV-Free Viewpoint Television,” Klagenfurt, 2002.

[4] 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.

[5] A. Fusiello, E. Trucco, and A. Verri, “Rectification with unconstrained stereo geometry,” *Proceedings of the Brithsh Machine Vision Conference,* Essex, pp. 400-409, 1997.

[6] C. Loop and Z. Zhang, “Computing Rectifying Homographies for Stereo Vision,” *Proc. of Computer Vision and Pattern Recognition*, pp. 125-131, 1999.

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

[8] ISO/IEC JTC1/SC29/WG11 M11014, “Utilization of Inter-View Correlation for Multiple View Video Coding,” Redmond, 2004.

[9] ISO/IEC JTC1/SC29/WG11 M11435, “Calibration and Rectifi cation Procedures for Multi-Camera System,” Palma de Mallorca, 2004.