HIGH-QUALITY MULTI-VIEW DEPTH GENERATION USING MULTIPLE COLOR AND DEPTH CAMERAS

Yun-Suk Kang and Yo-Sung Ho

Gwangju Institute of Science and Technology (GIST)
{yunsuk, hoyo}@gist.ac.kr

ABSTRACT

In this paper, we propose a high-quality multi-view depth generation method using multiple color and depth cameras. After we capture low-resolution depth maps by three TOF cameras, the depth information is warped into color image positions and used as the initial disparity value. By applying the stereo matching using belief propagation with the initial disparity information, we have obtained more accurate and stable multi-view disparity maps, compared to those results without the initial disparity information.

Keywords— Multi-view depth generation, multiple cameras, TOF camera, 3DTV

1. INTRODUCTION

Since the first TV broadcasting started, various technologies related to the broadcasting system such as video capturing and processing, encoding and decoding, communications, and display device manufacturing have been steadily investigated. Recently, three-dimensional TV (3DTV) and free viewpoint TV (FTV) have been investigated as the next generation broadcasting systems that can satisfy the demand for more realistic multimedia services. In order to generate 3D video contents, image sequences from multiple viewpoints and depth information are generally required. Using the depth information, we can synthesize the intermediate view images of the scene. It enables us to feel a natural 3D view through the images of the scene from multiple viewpoints [1, 2].

It is important to generate the scene's accurate depth information because the intermediate view synthesis is dependent on the depth. There are numerous methods to acquire depth information of the scene. Generally, there are two types; one is based on passive range sensors and the other is based on active range sensors. In the passive range sensor type, stereo matching is the most popular method [3]. It is efficient because it only requires a stereo or multi-view image. However, some difficulties such as textureless or occluded regions have been remained.

The most popular one in the active range sensor type is the time-of-flight (TOF) depth camera that obtains the depth information of the scene in real time. The principle of depth acquisition of TOF is based on measuring the arriving time of emitted infrared signal from the sensor. However, TOF has several problems to overcome such as low spatial resolution and noisy acquisition depending on the capturing environment.

There were proposals of systems that combine a TOF depth camera and video cameras [5-9]. In these approaches, they can take the advantages and discard the disadvantages of passive and active methods. The depth information from the passive method can be enhanced by using the depth obtained from the active method.

In this paper, we introduce the multi-view depth generation method based on five video cameras and three TOF depth cameras. We explain the camera system we used to capture the scene's color and depth images. Then, we describe the proposed method. The depth images are warped into the corresponding video camera positions after the preprocessing steps, and then used as the initial disparity for stereo matching. After showing the experimental results, we conclude this paper.

2. CAMERA CONFIGURATION

In the proposed method, we integrate five video cameras and five TOF depth cameras [10, 11]. There are two rows to amount the cameras. The fundamental setup is that the first row has five video cameras and the second row has three depth cameras at the center. They are all arranged in one-dimensional (1D) parallel setup. Figure 1 shows our multi-depth camera system.

This camera system captures the color images from five viewpoints and low-resolution depth images from three viewpoints. Figure 2(a) shows the multi-view color image. Figure 2(b) is the output images of three depth cameras, which are the depth images and intensity images, respectively.

This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency) (NIPA-2010-(C1090-1011-0003)).

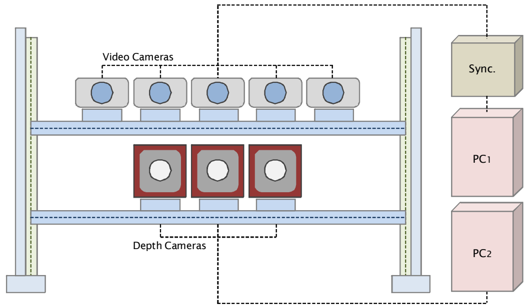
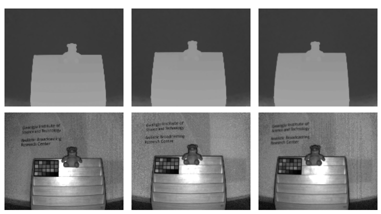


Fig. 1. Multiple color and depth cameras

View 1 View 2 View 3 View 4 View 5

(a) Color images (640x480)



View 2 View 3 View 4

(b) Depth and intensity images (176x144)

Fig. 2. Captured images

We used three depth cameras because the inherent problem of the depth camera we used. In order to operate multiple depth cameras simultaneously, we have to use the different modulation frequencies for the emitted infrared signals. Since the depth camera provides three different modulation frequencies, the number of depth cameras is limited to three.

3. PROPOSED METHOD

In this section, we explain each step of the proposed method. Figure 3 shows the whole procedure to generate multi-view depth information.

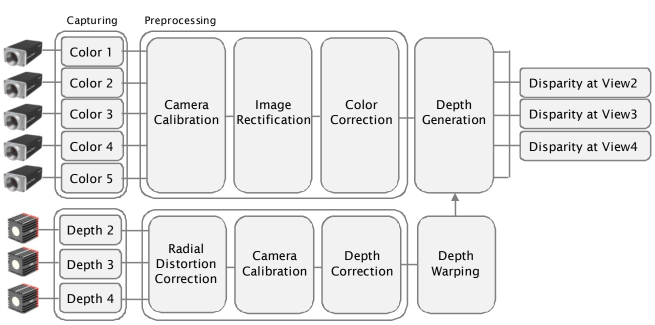
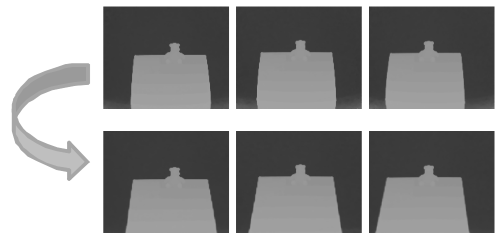


Fig. 3. Procedure of the proposed method

3.1. Preprocessing for Depth Images

3.1.1. Lens Distortion Correction

The depth images from the depth camera have the lens radial distortion as shown in Fig. 2(b). For efficiency, this lens distortion has to be corrected. Therefore, we corrected the lens distortion and the result is shown in Fig. 4 [12]. Camera calibration is then performed using the radial distortion corrected images.

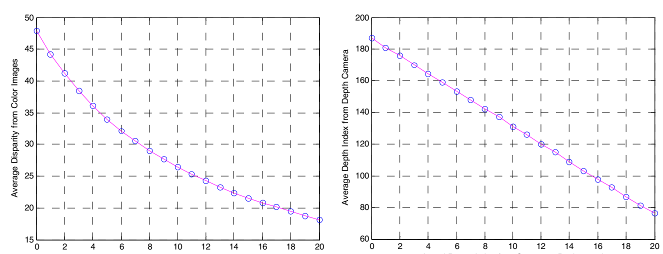


View 2 View 3 View 4

Fig. 4. Result of lens distortion correction

3.1.2. Depth Correction

Generally, depth information of the scene has a non-linear characteristic as shown in Fig. 5(a). Figure 5(a) shows the relationship between the actual range and the disparity obtained from the color images. However, the relationship between the range and the depth index from the depth camera is almost linear as shown in Fig. 5(b). The data shown in Fig. 5 is obtained by capturing the images of the check-pattern at several predefined positions within the background and camera as shown in Fig. 6. There is a regular interval between two adjacent predefined positions.



(a) Disparity-range curve (b) Depth index-range curve

Fig. 5. Relationship between the actual range and its representation

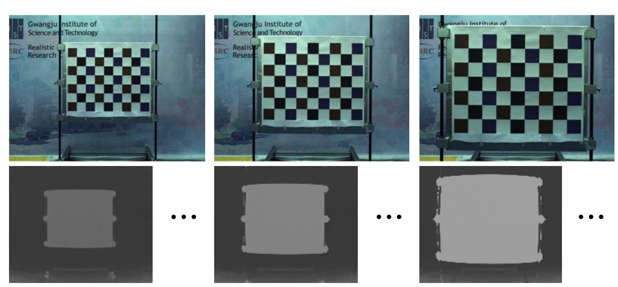


Fig. 6. Images for depth calibration

Due to the different representation of the actual range of the scene, it is required to correct the depth image. As shown in Fig. 6, we estimate a cubic curve between the disparity values and depth indexes. By using this curve, the depth index of each pixel of the depth image is corrected and has the non-linear range characteristic that is obtained from the color images.

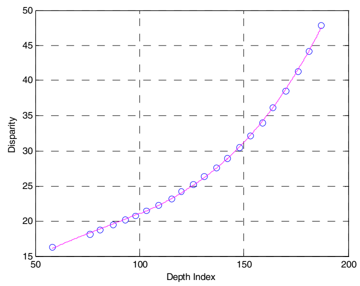
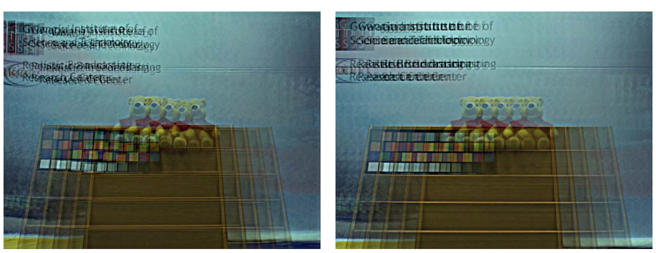


Fig. 6. Depth value correction

3.2. Preprocessing for Color Images

3.2.1. Multi-view Image Rectification

After camera calibration, we perform multi-view image rectification to the color images to minimize the geometric error [13]. This geometric error existed in the multi-view color image can not only deteriorate the depth quality but also interfere with the natural 3D view. Therefore, we minimize the geometric error by performing the multi-view image rectification and then obtain the rectified multi-view color image and rectified camera parameters. Figure 7 shows the rectified result.



Original Rectified

Fig. 7. Result of multi-view image rectification

3.2.2. Multi-view Color Correction

In order to decrease the color mismatch problem among multiple color views, we capture the images of Macbeth color chart and use them for multi-view color correction [14]. The color chart based method provides high accuracy than the other color correction methods.

3.3. Multi-view Depth Generation

After performing the preprocessing steps to the color and depth images, we generate the multi-view depth information of the scene. For high-quality depth generation, each pixel value of the preprocessed depth images from the depth cameras are used as the initial disparity for stereo matching. Therefore, the depth image at each view is warped to its corresponding position of the color view by 3D warping. Then, each pixel value of the warped depth images represents the disparity of that pixel as shown in Fig. 8.

Fig. 8. Initial disparity for view 2, view 3, and view 4

Then, we generate the three-view disparity map by using stereo matching with the initial disparity value. For the data cost calculation, we use two adjacent (left and right) reference views. The initial disparity value for each pixel is decided in three different ways. If there is a nonzero value for a pixel, it is directly used. Otherwise the initial disparity value is decided as the average of the neighboring pixel values [5].

4. EXPERIMENTAL RESULTS

We used five color video cameras and three TOF depth cameras for testing as shown in Fig. 9. By using these cameras, we captured five-view color images and three-view depth images. The resolution of the color images is 640×480 and the depth images have 176×144 pixels.

Fig. 9. Camera configuration

Figure 10(a) and Fig. 10(b) show the generated multi-view disparity maps using the previous stereo matching method [4] and the proposed method, respectively. Especially for the homogeneous regions such as the bookstand, the previous method cannot obtain the accurate disparity values. Some background regions also have wrong disparity values. However, the results of the proposed method have the correct disparity values of the scene without those problems.

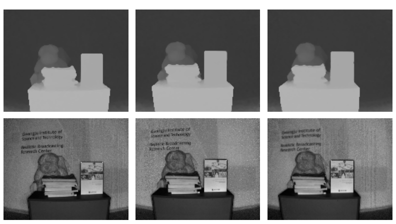
The other experimental results are shown in Figure 11. We also notice that the previous method fails to obtain the correct disparity information in some parts of the scene. However, the results of the proposed method shown in Fig. 11(e), using the initial disparity shown in Fig. 11(d), have the accurate and stable depth information.

(a) Result without the depth cameras: view 2, view 3, and view 4

(b) Result by the proposed method: view 2, view 3, and view 4

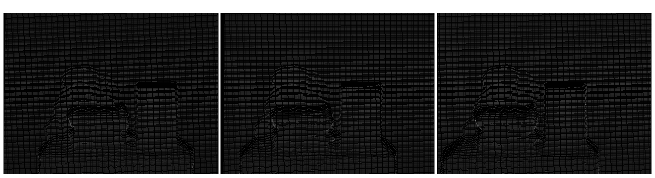
Fig. 10. Multi-view disparity map

(a) Five-view color images



(b) Three-view depth and intensity images

(c) Result without the depth cameras: view 2, view 3, and view 4



(d) Initial disparity: view 2, view 3, and view 4

(e) Result by the proposed method: view 2, view 3, and view 4

Fig. 11. Results from the other test images

5. CONCLUSION

In this paper, we proposed a high-quality multi-view depth generation method using multiple color and depth cameras. After capturing five color images and three depth images, we performed preprocessing for the color images and depth images, respectively. Then, each depth image is warped to its corresponding color image position and used as the initial disparity information for stereo matching. The proposed method provided the high-quality three-view disparity map that has the accurate and stable depth information even at the homogeneous or weak textured regions.

6. REFERENCES

[1] A. Smolic and P. Kauff, "Interactive 3D Video Representation and Coding Technologies," Proceedings of the IEEE, Spatial Issue on Advances in Video Coding and Delivery, vol. 93, no. 1, pp. 99-110, 2005.

[2] ISO/IEC JTC1/SC29/WG11 M8595, "FTV-Free Viewpoint Television," 2002.

[3] 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, 2003.

[4] P.F. Felzenszwalb and D.P. Huttenlocher, "Efficient Belief Propagation for Early Vision," International Journal of Computer Vision, vol. 70, no. 1, pp. 41-54, 2006.

[5] Y.S. Ho and Y.S. Kang, "Multi-view Depth Generation using Multi-Depth Camera System," International Conference on 3D Systems and Application (3DSA), May 2010 (Accepted).

[6] S.A. Gudmundsson, H. Aanaes, and R. Larsen, "Fusion of Stereo Vision and Time-of-Flight Imaging for Improved 3D Estimation," Int'l Journal of Intelligent Systems Technologies and Applications, vol. 5, no. 3, pp. 425-433, 2008.

[7] J. Zhu, L. Wang, R. Yang, and J. Davis, "Fusion of Time-of-Flight Depth and Stereo for High Accuracy Depth Maps," Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pp. 231-236, June 2008.

[8] K.D. Kuhnert and M. Stommel, "Fusion of Stereo-Camera and PMD-Camera Data for Real-time Suited Precise 3D Environment Reconstruction," Proc. of the IEEE/RSJ Int'l Conference on Intelligent Robots and Systems, pp. 4780-4785, March 2006.

[9] B. Bartczak and R. Koch, "Dense Depth Maps from Low Resolution Time-of-Flight Depth and High Resolution Color Views," Proc. of 5th International Symposium on Visual Computing, pp. 1-12, Nov. 2009.

[10] http://www.baslerweb.com, Basler Pylon GigE Camera piA I1900-32gm/gc.

[11] http://www.mesa-imaging.ch, SwissRanger SR 4000.

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

[13] Y.S. Kang and Y.S. Ho, "Geometrical Compensation for Multi-view Video in Multiple Camera Array," Proc. of Int'l Symposium on Electronics and Marine, pp. 83-86, Sept. 2008.

[14] N. Joshi, B. Wilburn, V. Vaish, M. Levoy, and M. Horowitz, "Automatic Color Calibration for Large Camera Arrays," in UCSD CSE Technical Report CS2005-0821, 2005.