Post-processing of Depth Image from Time-of-Flight Depth Camera

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

***Abstract*- In this paper, we propose a post-processing method of the depth image from a Time-of-flight (TOF) depth camera. Although the TOF depth camera can measure the accurate depth information of the scene in real-time, it has several problems to overcome. The depth image captured by the TOF depth camera has the lens radial distortion, noise at object boundaries, and depth difference with respect to the stereo camera that is used at the same time. The proposed method is composed of four steps to solve these problems and then 3D warp the depth image to the stereo color image position. Finally, we obtain the accurate disparity information at the color image position and they can be utilized as the initial data for 3D contents generation.**

I. Introduction

Various technologies related to the broadcasting system such as capturing, processing, encoding, decoding, and display devices have been gradually investigated since the TV broadcasting started. In recent years, three-dimensional TV (3DTV) or 3D video, represented as the movie Avatar, are considered as the next generation broadcasting that satisfies the increasing demand for more realistic multimedia services. Generating a 3D video content of a scene require image sequences from more than two viewpoints and depth information of the scene. Depth information indicates the range data of the scene and we can reconstruct the images at arbitrary viewpoints using these depth and color images. Then users can watch a natural and immersive 3D scene from the multiple view images on the 3D display devices [1].

There are several methods to obtain depth information of the scene. Generally, there are two types: methods based on the passive range sensors and methods based on the active range sensors. In the passive range sensor type, the depth of the scene is generally estimated by using the captured two-dimensional (2D) images. Stereo matching [2] is one of the most popular one, and there are also depth estimation methods such as shape from silhouette, focus, motion, and so on. These methods are efficient because the input data acquisition is easy. However, it is hard to obtain the high quality depth data from those methods due to some limits such as the illumination, color mismatch, occlusion, textureless regions, and so on. The accuracy of the scene's depth data is important because the quality of reconstructed arbitrary viewpoint image totally relies on the quality of depth.

The active range sensor based methods measure the scene's depth using a measuring instrument such as a depth camera and a 3D range scanner. The time-of-flight (TOF) technology is one of the most powerful one to acquire the depth data. A TOF depth camera can measure the range from the camera to objects in real time. The principle of TOF is to measure the round-trip time of emitted signal of the sensor. Therefore, we acquire more accurate depth information of the scene compared to the passive range sensors. However, TOF technology also has several problems to overcome such as low spatial resolution and noisy acquisition depending on the capturing environment.

In this paper, we introduce a post-processing method of the depth image that is captured by the TOF depth camera. In order to use the depth image of the TOF depth camera for the 3D image processing, we have to correct the inherent errors and distortion. After referring to the operation and characteristics of the TOF depth camera, we explain each step of the post-processing for the TOF depth camera image. We also capture stereo images and process them to verify the proposed method. Then, we show the experimental results and conclude this paper.

II. Time-of-Flight Depth Camera

The operation principle of the TOF depth camera is to check the round trip time for emitted light from an illumination source of the camera to the object and back to the camera. The depth camera modulates its illumination and the imaging sensor measures the phase of the returned modulated signal at each pixel. The full-phase value is corresponding to the maximum range of the captured scene. The maximum range depends on the modulation frequency of the camera [3].

There are several TOF depth camera models. The first real-time TOF depth camera is ZCAM™ that was developed by 3DV systems. It simultaneously captures color and depth images in the SD resolution at the same viewpoint. However, this ZCAM™ has an optical noise of the output image and cannot measures any black-colored objects. Also its large volume bothers to compose various types of camera arrangement.

After ZCAM™, smaller and cheaper TOF depth cameras have been developed. Figure 1 shows SR4000 developed by MESA Imaging, which is used our experiment. SR4000 is composed of the illumination cover and optical filter. The illumination cover protects the LEDs that emit the light, and the optical filter receives the arriving light emitted from the camera LEDs only. The volume of this camera is 65x65x68mm. The minimum and maximum capturing ranges are 0.3m and 5.0m, respectively. It provides its framerate up to 54 frames per second. We can obtain two types of output images; depth image and intensity (amplitude) image as shown in Fig. 2. The resolution of output image is very small as QCIF (176x144) [3].

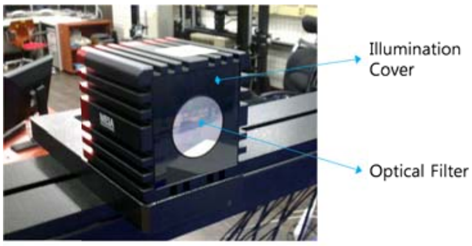


Figure 1. TOF depth camera (SR4000)

Figure 2. Output images: Depth and intensity images

However, SR4000 has some inherent distortion and errors. As shown in Fig. 2, the output image has not only a large amount of lens distortion but also boundary error. Also there is a small difference between measured depth values by the depth camera and the corresponding depth values from the stereo or multi-view camera when we use the depth camera with color cameras. Because two types of cameras have different position in *z*-direction. Therefore, we also have to correct this difference.

After correcting the aforementioned distortion and errors, 3D warping of the depth image is required for 3D contents generation because we need the measured depth information at the position of the stereo or multi-view cameras, when we use two types of cameras together.

III. POST-PROCESSING OF TOF DEPTH IMAGE

In this section, we explain each step of the proposed post-processing method of TOF depth image. Figure 3 shows the procedure of the proposed method.

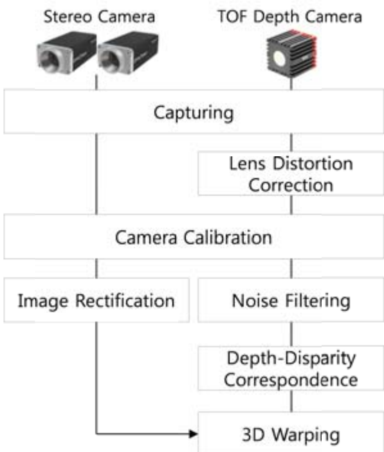


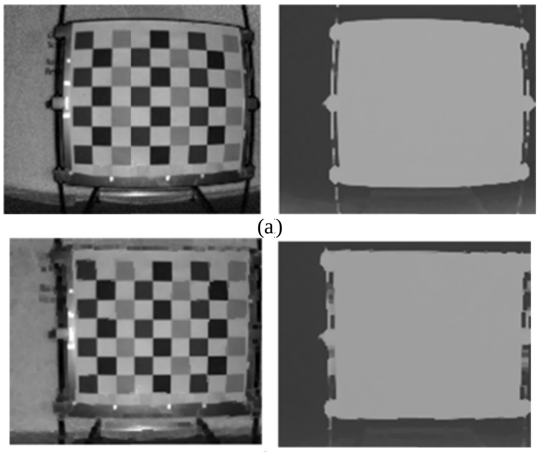
Figure 3. Procedure of the proposed method

After performing the proposed method, the TOF depth image can be used as the depth information of the scene for 3D contents generation. To verify the proposed method, we also capture stereo images of the scene with the depth camera image and process them.

A. Lens Distortion Correction

SR4000 has the inherent distortion in its capturing. The output images of SR4000 have a lens barrel distortion. This distortion causes the shape mismatch between the depth and color images. Also, we cannot expect good results when we perform any point based processing such as camera calibration.

Lens distortion correction is reconstructing the distorted image based on the distorted components. In the proposed method, we extract several points on the distorted line components in the image, and then correct the image using the calculated distortion center and distortion parameter [4]. After the reconstruction, generated holes are filled by dilation. Figure 4 shows the result images of the lens distortion correction.



(a)

(b)

Figure 4. Lens distortion correction:
(a) original, (b) corrected intensity and depth images

B. Camera Calibration

The camera parameters are composed of the intrinsic and extrinsic parameters. They are the principal information of the camera for 3D image processing. The intrinsic parameters indicate the camera internal characteristics such as the focal length and principal points. The extrinsic parameters of the camera are the rotation and translation. They represent the orientation and position of the camera in the world coordinate system.

These camera parameters are estimated by camera calibration [5]. Figure 5 shows the images of the planar checker board images for camera calibration. These images are reconstructed by lens distortion correction. By extracting all the corner points of the pattern in each image, we estimate the intrinsic and extrinsic parameters of the camera. For camera calibration, Camera calibration toolbox for Matlab is used [6].

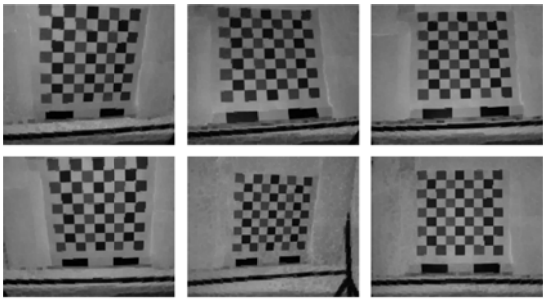


Figure 5. Lens distortion corrected pattern images for camera calibration

C. Noise Filtering

In general, noise occurs at the depth discontinuities of the depth image. This noise means the pixels that have median depth values between two objects or foreground and background. Therefore, it is required to remove these median values at the object boundaries. Object boundaries are obtained by applying edge detection algorithm to the depth image as shown in Fig. 6. We used Canny edge detector to obtain the object boundaries. Then we remove the detected pixels.

Figure 6. Depth image and the edge map

D. Depth-Disparity Correspondence

Generally, disparity that is obtained the rectified stereo image of the scene has a non-linear characteristic as shown in Fig. 7. The data shown in Fig. 7 is obtained by capturing the images of the check-pattern at several predefined positions within the background and camera as shown in Fig. 8. The x-axis means the distance from the camera position and the y-axis indicates the disparity values in pixel. However, the depth information obtained from the depth camera show the linear characteristics with respect to the distance. Also there are position differences between the rectified stereo camera and depth camera along z-axis.

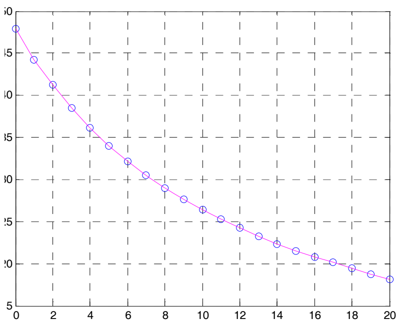


Figure 7. Disparity characteristics of the scene

Due to the different representation of the actual range and different location along z-axis, we map the depth index values to the valid disparity values at the stereo image. As shown in Fig. 8, 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 mapped to the disparity values.

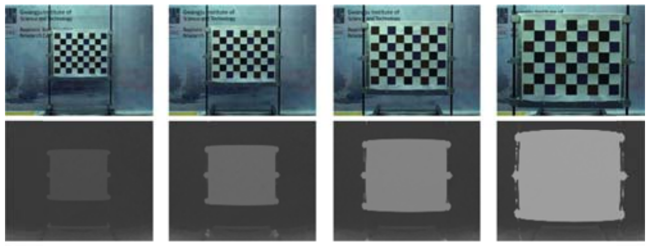


Figure 8. Images for depth-disparity correspondence

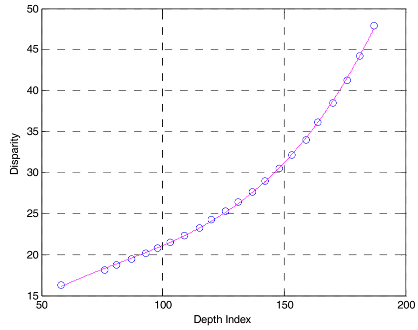
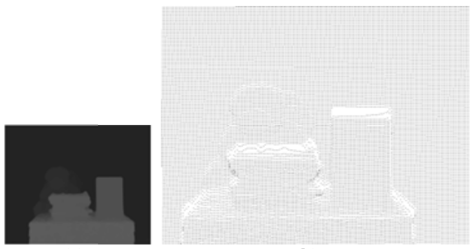


Figure 8. Depth-disparity correspondence curve

E. 3D Warping

3D warping back-projects the value of each pixel to the 3D space based on the source camera parameters and depth information. Then, the projected pixel values in the space are remapped onto the color image plane. This is also based on the target camera parameters and the corresponding depth information.

In our approach, we use the depth-disparity correspondence curve in Fig. 8 to change the depth values to the disparity values. These disparity values can be regarded as the initial disparity of the stereo or multi-view image at each pixel. Figure 9(a) shows the original depth image. The warped depth image has the same resolution as the color image as shown in Fig. 9(b).



(a) (b)

Figure 9. 3D warping of depth image

IV. Experimental Results

For the experiment of the proposed method, we capture a scene using the depth camera and stereo camera as shown in Fig. 10. The stereo camera model is Basler Pylon GigE and the depth camera is SR4000. The stereo image is captured with the resolution of 800x600, and then calibrated and rectified [7]. Figure 11 shows the rectified stereo and depth images. The distance between two color cameras is 65mm.

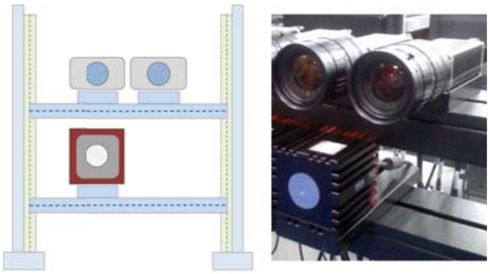
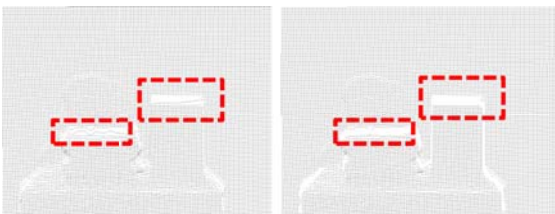


Figure 10. Camera setup

Figure 11. Captured images

The result of lens distortion correction is depicted in Fig 12. The barrel distortion in the image is corrected. Figure 13 shows the results of noise filtering. Noisy depth values at the object boundary became the error regions as shown in Fig. 13(a) when we applied 3D warping to the depth image so that each pixel warped to the color camera position. However, noise-filtered results show that these error regions are accurately removed as shown in Fig. 13(b).

Figure 12. Result of lens distortion correction



(a) (b)

Figure 13. Result of noise filtering:
(a) Before noise filtering, (b) After noise filtering

The warped depth image shown in Fig. 13 then indicates the disparity values at each pixel of the stereo image. Finally, these disparity values can be used as accurate initial information for matching cost calculation or depth map generation. The wrong pixel values are boxed in the left figure, and they are removed in the right figure by noise filtering. Then, the warped depth image indicates the disparity values at each pixel of the stereo image.

Table 1 shows the disparity values of the stereo image, warped depth image without depth-disparity correspondence matching, and warped depth image after correspondence matching. The results in Table 1 show us that the depth-disparity correspondence matching accurately corrects the depth information of the scene (Disparity positions are indicated in Fig. 11). Finally, these disparity values can be used as accurate initial information for matching cost calculation or depth map generation.

Table 1. Results of depth-disparity correspondence

Position	Stereo(True)	Before	After
1	16	16	16
2	21	25	21
3	37	40	37
4	40	41	40

V. CONCLUSION

In this paper, we explained a post-processing method for the TOF depth camera images. Since the captured images from the TOF depth camera are noisy and distorted, we have to correct to use them for 3D image processing and application. After the proposed post-processing method, the depth camera data is provided as the accurate disparity values at the color image position. These values can be used as initial disparity information for stereo matching.

ACKNOWLEDGMENT

This work was supported by the IT R&D program of MKE/KCC/KEIT. [KI001932, Development of Next Generation DTV Core Technology]

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] 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.
[3] SR4000 User Manual, Mesa Imaging AG.
[4] A. Wang, T. Qiu, and L. Shao, “A Simple Method of Radial Distortion Correction with Centre of Distortion Estimation,” Journal of Mathematical Imaging and Vision, vol. 35, no. 3, pp. 165-172, 2009.
[5] Z. Zhang, “A Flexible New Technique for Camera Calibration,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22(11), pp. 1330-1334, 2000.
[6] http://www.vision.caltech.edu/bouguetj, Camera Calibration Toolbox for MATLAB.
[7] 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.