Generation of Multi-view Images Using Stereo and Time-of-Flight Depth Cameras

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

Abstract- **We propose a multi-view image generation method using a low-resolution time-of-flight (TOF) depth camera and a stereoscopic camera. The TOF depth camera obtains the scene's range information in real-time using the emitted light signal. Then, it quantizes the measure range and provides depth images of the scene. However, the TOF depth camera has not only too small resolution but also noise and lens distortion. In order to utilize the TOF depth image with the high-resolution color images, we have to reduce noise and distortion, and also overcome the resolution difference. The proposed method generates disparity maps for a stereo color image using the TOF depth data. We reduce the noise and lens distortion in the TOF depth image and then warp the TOF depth values to the one color image position. The color image is segmented, and the warped TOF depth values are called the segmented regions as the disparity. After generating the disparity map for the other view, we obtain the multi-view image by synthesis of intermediate view images. Experimental results show the generated disparity maps and reconstructed virtual view images verify that the proposed method efficiently generate the disparity maps and the multi-view image.**

Keywords- Depth generation, Multi-view, ToF depth camera, 3DTV

I. INTRODUCTION

In recent years, three-dimensional TV (3DTV) is one of the most popular multimedia services. 3D video contents for 3DTV can satisfy the demand for more realistic and immersive contents [1]. In order to generate 3D video contents, at least two viewpoint images are required. However, images from multiple viewpoints can provide us more realistic and natural 3D feeling than the stereo images. To make the multi-view image from the stereo image, scene's depth information is essential. Using the depth map that represents the distance information of the corresponding color view image, we can synthesize intermediate view images between two original images. Since the quality of the synthesized view is dependent on the depth quality, it is important to generate accurate depth information of the scene.

In general, the depth of the scene can be obtained by two types of depth acquiring methods. One is a passive method that uses only captured color images. Stereo matching is one of the most popular techniques to obtain the depth information of the scene [2]. It requires two viewpoint images and calculates the disparity value for each pixel. Although we can obtain depth without any special equipment or measuring instruments except cameras, it is difficult to obtain an accurate depth especially for the occluded regions and textureless regions.

The second is an active method that uses measuring instruments to measure the distance from the camera to the scene. The time-of-flight (TOF) depth camera calculates and quantizes the depth of the scene, and then provides the depth image using the light signal in real-time. However, it has some problems such noise, distortion, and too small resolution.

Recently, some approaches attempts to combine those two methods to complement each other [3] [4]. Many of them uses stereo or multi-view camera with a TOF depth camera. The TOF depth information is transferred to the color image position and then used as an initial disparity for stereo matching, generally.

In this paper, we explain a multi-view generation from the stereo image with a TOF depth camera. After enhancing the captured images, the TOF depth values are 3D warped to each segment of the one color image. Then, we fill each segment with the proper disparity values. Finally we obtain the multi-view image from stereo image and its disparity maps.

II. CHARACTERISTICS OF TIME-OF-FLIGHT DEPTH CAMERA

Figure 1 shows the TOF depth camera SR4000 that is used in our experiment. It is composed of the illumination cover, which emits the light signal to the scene, and the optical filter, that receives the arriving emitted signal. The minimum and maximum capturing ranges are 0.3m and 5.0m, respectively.

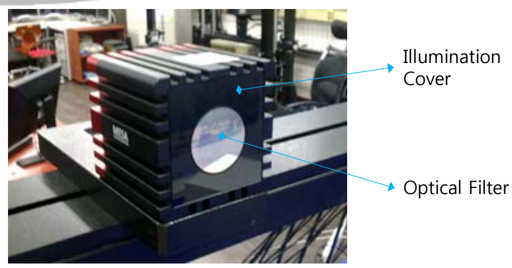
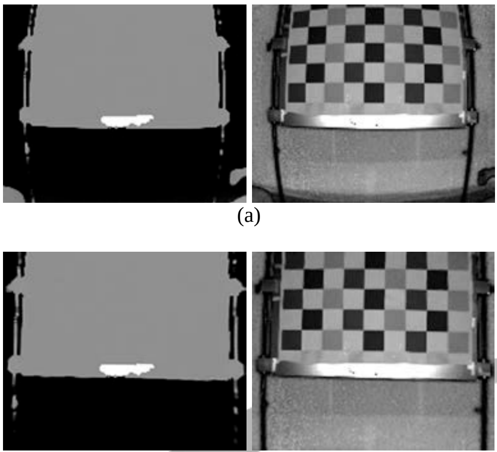


Figure 1. TOF depth camera (SR4000)

Figure 2 shows the output images from the TOF depth camera. There are two types of output images which are the depth and intensity images. Each pixel of the depth image represents the depth index from 0 to 255 which are quantized scene's depth. The intensity image is similar to the gray image of the scene.

However, as shown in Fig. 2(a), there are inherent noise and distortion in the TOF depth image. There are a large amount of the radial distortion and a boundary noises. In the proposed method, we correct this lens distortion before all processes [5]. Figure 2(b) shows the corrected depth camera images.



(a)

(b)

Figure 2. Lens distortion correction

III. Proposed Multi-view Generation Method

In this section, we explain our proposed method to generate multi-view image from the captured stereo image using the TOF depth image. After obtaining camera parameters by camera calibration, we enhance the quality of captured images. Then, the TOF depth image is warped to the color image and used as an initial disparity values for disparity filling. After the disparity map of one color view is generated, the disparity map of the other view is then obtained. Finally, we can generate the multi-view image from the stereo image. Figure 3 shows the whole process of the proposed method.

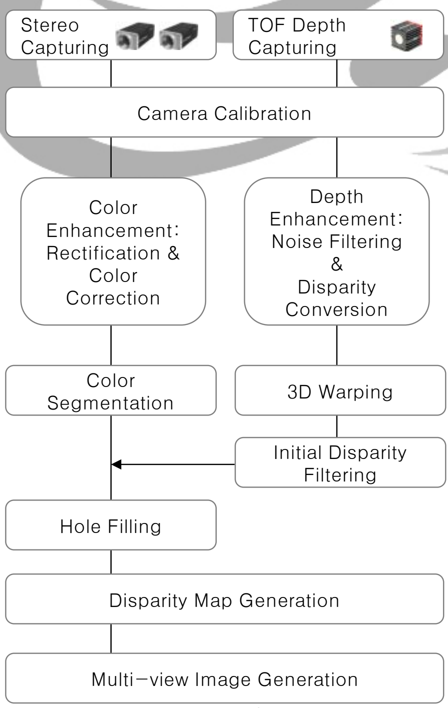
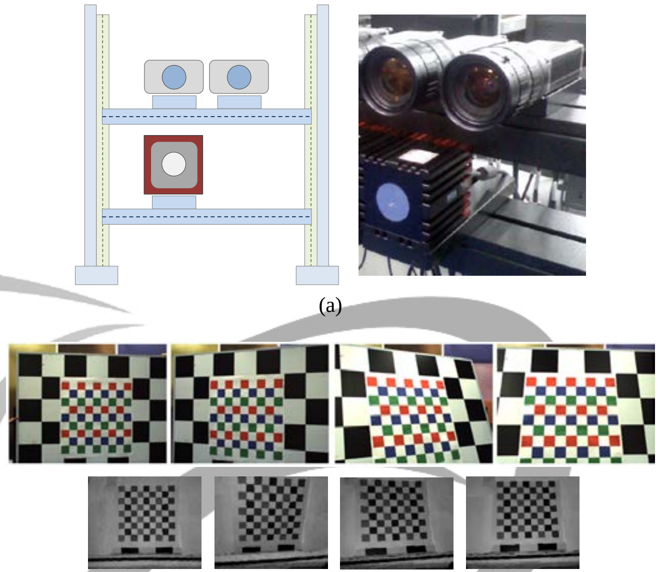


Figure 3. Whole procedure of the proposed method

A. *Capturing and Camera Calibration*

Figure 4(a) shows the camera arrangement for capturing scene. Two color cameras are arranged in parallel and one TOF depth camera is located under the one color camera. To calculate the camera parameters of three cameras, we used several grid-pattern images of each camera shown in Fig. 4(b), respectively. We used Camera calibration toolbox for MATLAB [6].



(a)

(b)

Figure 4. Camera arrangement and camera calibration

B. *Enhancement for Captured Images*

After obtaining camera the parameters by camera calibration, we perform stereo image rectification to the color images [7]. Stereo image rectification minimizes the geometric error in the stereo image, which means the vertical pixel mismatches as shown in Fig. 5. Also we perform color correction to the color images using Macbeth color chart to reduce color inconsistent problem. Since these two problems can deteriorate the quality of depth and also disturb the natural 3D feeling. The processed stereo color image is regarded as two images captured by two cameras that have an equal internal characteristics and the camera interval only in the horizontal direction.

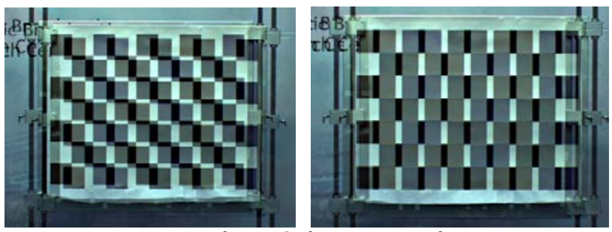


Figure 5. Before and after image rectification

Then, one of color images is segmented by using mean-shift segmentation algorithm [8], as showin in Fig. 6. The segmentation results provide an accurate object boundary information for hole filling.

For the TOF depth image, there is a boundary noise that has median depth values between the background and foreground object. These noisy pixel values can be an error when the TOF depth image is warped to the color image positions. In order to solve this problem, we use the shock filtering to the TOF depth image. The shock filtering flattens homogeneous regions, also reduces the boundary median values.

Figure 6. Color segmentation

Then, each pixel of the TOF depth image is changed to the disparity values for the stereo image. The original pixel values of the TOF depth image is depth index information which means the range information. However, for disparity map generation, we require the disparity value for each depth index value. Therefore, we calibrate the space for scene capturing, and pre-calculate the mapping curve from the depth index to the disparity.

C. *Multi-view Generation*

In order to generate multi-view image from the stereo image, we firstly calculate the disparity map for each view image. In the proposed method, we have TOF depth camera data for an initial disparity. The enhanced TOF depth image is warped to the color image position using 3D warping technique. 3D warping is composed of two steps. The first step is a back-projection of the pixel value based on its depth information and camera parameters.

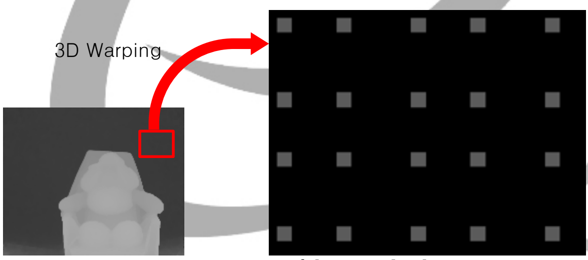


Figure 7. 3D warping of the TOF depth image

Then, the second step is a projection of the back-projected pixel values to the target image plane using also its depth information and camera parameters. After the 3D warping, the TOF depth image is located the color image plane with many holes due to the resolution difference. Figure 7 shows the 3D warping result.

For the next step, we fill the holes shown in Fig. 5 in each color segment. Since the color image is segmented based on the color, we can obtain the accurate boundary regions. In order to filter the pixels that have wrong initial values, we compare the pixel's color value to the others color value that is located on as much as disparity value. If the color difference is larger than the threshold value, we regard that the pixel has wrong initial disparity and we remove that pixel in the segment.

In order to fill the correct disparity value, we consider the three cases as follows:

Case 1: Only one initial disparity value in the segment
Case 2: More than one initial disparity values in the segment
Case 3: No initial disparity value in the segment

For Case 1, we fill the segment with the only initial disparity value. For Case 2, we calculate the disparity $d(x, y)$ as Eq. (1).

$$d(x,y) = \sum_k \lambda_k d_k(x,y) \tag{1}$$

Where $d_k(x, y)$ is defined as Eq. (2)

$$d_k(x,y) = \frac{1}{n}\sum_{i=-k}^{k}\sum_{j=-k}^{k} d(x+i, y+j) \tag{2}$$

n is the number of nonzero pixel, k is the window size, and the weighting factor λ_k according to the distance from the current pixel is defined as Table 1.

Table 1. Weighting factor with conditions

Conditions	λ_1	λ_2	λ_3
$d_1 \neq 0$, $d_2 = d_3 = 0$	1.0	0	0
$d_1 \neq 0$, $d_2 \neq 0$, $d_3 = 0$	0.7	0.3	0
$d_1 \neq 0$, $d_2 \neq 0$, $d_3 \neq 0$	0.65	0.25	0.1

For all segments in the color image, we fill most holes by using Eq. (1). However, if there is lack of disparity pixels in the segment, there can be unfilled regions in the segment. In this case, we perform in-segment dilation operation iteratively.

After considering Case 1 and Case 2, we fill segments in the Case 3. The segments in the Case 3 have no initial disparity values. In other words, the size of segment is very small, or the segment is located on image borders. In this case, we just perform dilation iteratively until the unfilled regions are filled.

For the last step, we generate the second disparity map that is corresponding to the other view's color image. Each pixel of this second disparity map is filled from the first disparity map using its own disparity values. Also, the second disparity map has holes due to the occluded regions. Therefore, we perform the segmentation to the other view's color image, and we verify the disparity values by using the same method explained before. Finally the wrong disparity values are filtered, and holes are filled by in-segment dilation. Then, the multi-view image is generated by using the stereo image and the disparity maps.

IV. Experimental Results

For experiments of the proposed method, we captured two sets of stereo image with TOF depth images by using the camera system shown in Fig. 4. The resolution of the color and TOF depth image is 1024x768 and 176x144, respectively. The horizontal distance and vertical distance between cameras are 6.5cm and 7.5cm, respectively. Figure 8 shows the captured stereo color and TOF depth images after image enhancement.

Figure 8. Captured and enhanced images

As shown in Fig. 5, stereo images have few vertical pixel mismatches. These color images are segmented, and each segment is filled by the 3D warped TOF depth image. In our experiments, we generated left disparity map first, then generated right disparity map. Figure 9 shows the generated disparity maps. The disparity maps have accurate object boundaries. Also, there are stable homogeneous regions since the disparity values are based on the TOF depth data.

Figure 9. Generated disparity maps

Finally, we obtained the multi-view image from the stereo image as shown in Fig. 10. The three intermediate view images shown in Fig. 10 are generated by using the stereo image and the calculated stereo disparity maps shown in Fig. 9.

V. Conclusion

In this paper, we presented an effective method to generate the multi-view image using stereo color and TOF depth cameras. Enhancing methods reduce the noise and distortion in the TOF depth image, and also decrease the color and geometric errors in the stereo image. The TOF depth data is used to stereo disparity generation, and finally we generate the multi-view image from the stereo image by synthesizing the intermediate viewpoint images. Since we did not perform the stereo matching, the proposed method is simple and it can be easily adopted for 3D many applications.

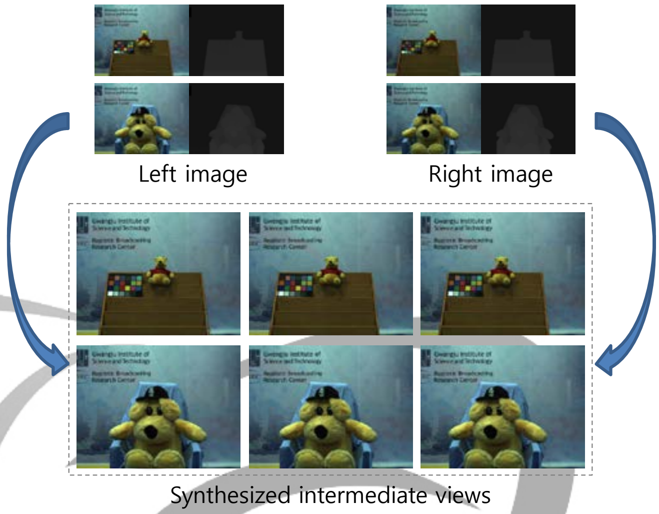


Figure 10. Synthesized intermediate views

Acknowledgment

This research is supported by MCST and KOCCA in the CT Research & Development Program 2012.

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

[4] 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, 2009.

[5] 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.

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

[7] Y.S. Kang and Y.S. Ho, "Efficient Stereo Image Rectification Method using Horizontal Baseline," Lecture Notes in Computer Science, vol. 7087, pp. 301-310, 2011.

[8] D. Comaniciu and P. Meer, “Mean Shift: A Robust Approach Toward Feature Space Analysis,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 4, pp. 609-619, 2002.