DISPARITY MAP GENERATION FOR COLOR IMAGE USING TOF 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
Telephone: +82-62-715-2258, Fax: +82-62-715-3164
E-mail: {yunsuk, hoyo}@gist.ac.kr

ABSTRACT

We propose an effective method for disparity map generation for a color image using a time-of-flight (TOF) depth camera. After we capture the color and TOF depth images, the processed depth information is warped into color image positions as a form of disparity. Then, by applying the averaging operation in each segment, we obtain the disparity map. Experimental results show the correct and stable disparity maps, compared to stereo matching results.

***Index Terms*—** Time-of-flight camera, disparity map, depth map, 3DTV

1. INTRODUCTION

Recently, three-dimensional (3D) video contents provide more immersive and realistic video contents to users than the traditional two-dimensional (2D) videos. Users can watch 3D video contents on the 3DTV that has been actively investigated as the most powerful one of the next generation broadcasting [1]. It is expected that 3DTV can satisfy the growing demand for the realistic multimedia service. In order to generate 3D video contents, the multi-view image and depth information are required. The multi-view image is a set of images of the same scene captured by multiple cameras. We can generate depth maps representing distance information of the scene. Based on the depth maps, we can reconstruct intermediate view images of the scene at arbitrary virtual viewpoints.

In general, there are two types for acquiring depth information of the scene. One is based on passive range sensors and the other is based on active range sensors, such as stereo matching [2]. The passive range sensor based method does not require any special equipment except cameras. However, the result of stereo matching relies on the texture of input images and object placement of the scene in many times.

978-1-61284-162-5/11/$26.00 2011 © IEEE

The depth acquisition methods based on the active range sensors use special equipment for measuring the range of scene, such as well-known time-of-flight (TOF) depth cameras. The TOF depth camera emits light signals itself and then measuring the arriving back time of the signals to obtain the range data. However, in spite of its high price, it merely yields small spatial resolution images with noise.

Recently, there have been several attempts to combine the active and passive range sensor based methods to obtain the advantages and discard disadvantages of each method. Many of these attempts are consisted of stereo or multi-view camera with one TOF depth camera. The depth of the scene is obtained by the stereo matching with the data from the TOF camera [3][4]. However, these approaches also perform the stereo matching based on the global optimization that requires high complexity.

We present an effective method to generate disparity maps of the scene using a TOF depth camera with a stereo camera. After performing a few steps of preprocessing, the depth information from the TOF depth camera is warped to the segments of the color image and transformed to the disparity values. Then, we calculate the proper disparity values in each segment. Finally we obtain the disparity map by filling some unfilled region using the morphological operation.

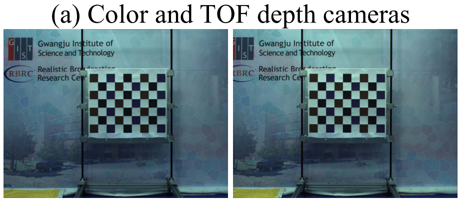
2. DATA ACQUISITION

In the proposed system, we use color video and TOF depth cameras. The video camera model we use is Basler Pylon GigE and the TOF depth camera model is MESA Imaging SR4000. Figure 1 shows the cameras and captured color and depth images.

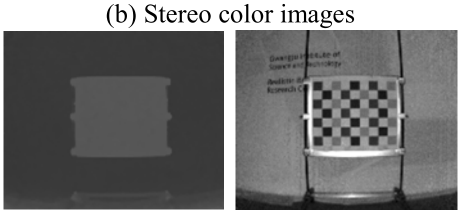
However, SR4000 has some inherent distortion and errors. As shown in Fig. 1(c), the output image has not only a large amount of lens distortion but also errors at the object boundary region. Moreover, there is a small difference between measured depth values by the depth camera and the corresponding 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 correct this difference.

(a) Color and TOF depth cameras



(b) Stereo color images



(c) Depth and intensity images

Figure 1. Cameras and captured images

After correcting the 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.

3. PROPOSED METHOD

Figure 2 shows the whole process of the proposed method. In this section, we explain each step of the proposed disparity map generation method.

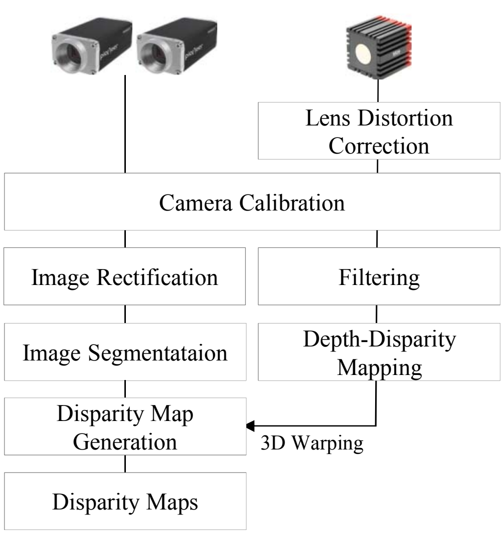


Figure 2. Flow chart of the proposed method

3.1. Camera Calibration

The camera parameters are the principal information of the camera for 3D image processing and application. They are composed of the intrinsic and extrinsic parameters. The intrinsic parameters represent the camera's internal features such as the focal length. The camera's orientation and location information are appeared in the extrinsic parameters.

Camera calibration is the process to estimate those parameters using the captured images [5]. To calibrate the camera, we extract all the corner points from each image and then calculate the intrinsic and extrinsic parameters based on the practical size of one square in the checker board. In the case of the multiple cameras, we need the same world coordinate system. Then each camera can be calibrated correctly. In the proposed method, we perform camera calibration of the depth cameras after correcting the lens distortion of the depth images. Camera calibration toolbox for Matlab [6] is used for camera calibration process.

3.2. Depth Image Processing

3.2.1. Lens Distortion Correction

Depth images captured by SR4000 have a large amount of lens radial distortion as shown in Fig. 1(b). Generally, there are two kinds lens radial distortion which are barrel and pincushion distortion. In our case, the barrel distortion is occurred due to the inherent camera problem. Because of the distortion, there exists the shape mismatch between the depth and color images. It also interferes some point based processing such as camera calibration.

Lens distortion correction is the image reconstruction of the distorted image. In the proposed method, we use a point extraction based method to correct the distortion [7]. After finding the curved straight line component in the captured image, we estimate the distortion center and the distortion parameter. With the distortion information, we can reconstruct the image from the distorted image. The holes generated after the reconstruction are filled by the dilation process. Figure 3 shows the depth and intensity images after the correction.

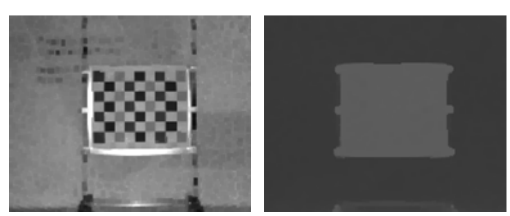


Figure 3. Results of lens distortion correction

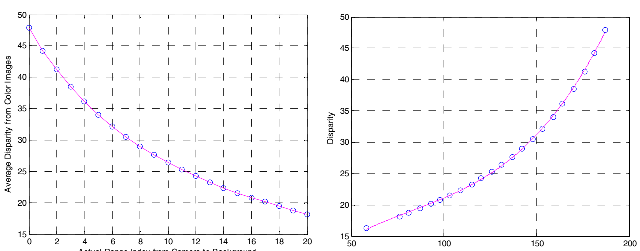
3.2.2. Depth Image Filtering

In general, the 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. The depth image also has the optical noises in the homogeneous regions.

Therefore, we perform bilateral filtering to the depth images to remove the optical noises with maintaining the boundary components. Then, object boundaries are obtained by applying edge detection algorithm to the depth image. We used Canny edge detector to obtain the object boundaries. Then we remove the detected pixels.

3.2.3. Depth-Disparity Mapping

Generally, disparity that is obtained the rectified stereo image of the scene has a non-linear characteristic as shown in Fig. 4(a). The data shown in Fig. 4(a) is obtained by capturing the images of the check-pattern at several predefined positions within the background and camera. The *x*-axis means the distance from the camera position and the *y*-axis indicates the disparity values in pixel. This disparity values are extracted from the stereo image. However, the depth information obtained from the depth camera show the linear characteristics with respect to the distance.



(a) Disparity-range curve (b) Depth-disparity mapping

Figure 4. Depth-disparity mapping

Due to the different representation of the actual range of the scene, it is required to correct the depth image. As shown in Fig. 4(b), 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 [8].

3.3. Color Image Processing

3.3.1. Image Rectification

After camera calibration, we perform multi-view image rectification to the color images to minimize the geometric error [9]. 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 5 shows the rectified result.

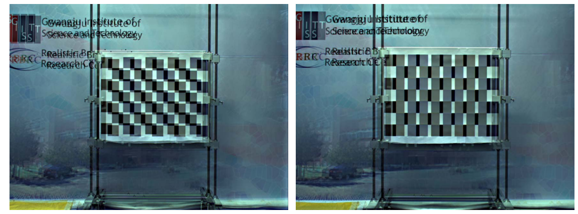

(a) Original stereo (b) Rectified stereo

Figure 5. Result of image rectification

3.3.2. Image Segmentation

In the proposed method, we divide the color image into many segments such that each segment has the similar color distribution. Figure 6 shows the original color image and its segmented image by using the mean-shift segmentation algorithm [10]. Then we assume that, each segment is relatively small, and it has the same disparity value.

(a) Original image (b) Segmented image

Figure 6. Color image segmentation

3.4. 3D Warping

3D warping backprojects the value of each pixel to the 3D space based on the camera parameters and depth information. The projected values in the space are then reprojected to the color image plane. This process is also based on the camera parameters and depth information.

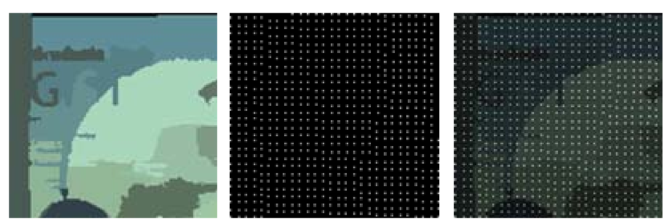

(a) Segmented (b) Warped depth (c) Disparity values

Figure 7. 3D warping results

Figure 7 shows the warped depth pixels on to the segment color image. There are holes which have no depth values due to the resolution difference between the color and depth images.

3.5. Disparity Map Generation

As shown in Fig. 7(c), there are the disparity values at some pixels and holes in each segment. In order to fill the hole with the appropriate disparity values, we first define Eq. (1)

as the candidates $d_k(x, y)$ for the disparity. Then we use Eq. (2) to calculate the disparity $d(x, y)$ in each segment.

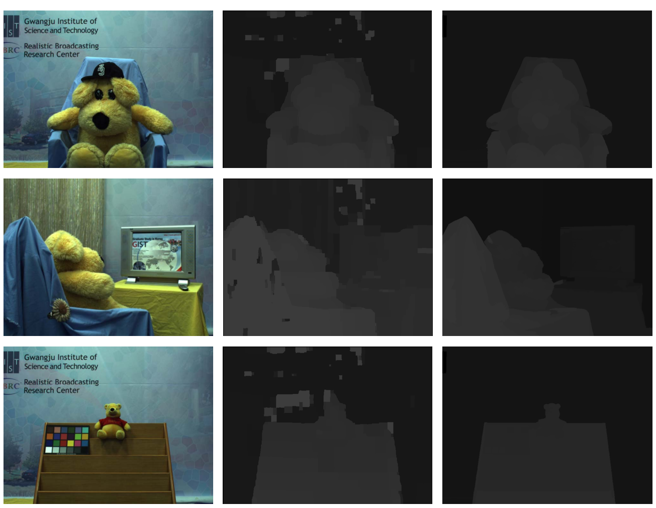
$$d_k(x,y) = \frac{1}{n}\sum_{i=-k}^{k}\sum_{j=-k}^{k} d(x+i, y+j) \qquad (1)$$

$$d(x,y) = \sum_{k} \lambda_k d_k(x,y) \qquad (2)$$

where n indicates the number of nonzero pixels of the summation process. Moreover, $d(x+i, y+j)$ is considered as zero when it is outside the current segment. λ_k is the weight according to the distance, and $\lambda_k=0$ when $d_k(x,y)=0$. We choose k according to the size of the generated holes. After that, we dilate some unfilled regions in each segment and then obtain the final disparity map.

4. EXPERIMENTAL RESULTS

For experiments, we used three sets of the color and depth images. The resolutions of color and depth images are 800x600 and 176x144, respectively. In the experiments, we used k=3. Figure 8 shows the experimental results. The images in Fig. 8(b) are obtained by using stereo matching method [2]. Figure 8(c) is the results of the proposed method.



(a) Original (b) Stereo matching (c) Proposed method

Figure 8. Generated disparity maps

As shown in Fig. 8, the disparity maps correctly represent scene depth compared with the stereo matching results. Since the disparity values are measured from the manual range data. They also have the clean object boundary regions because they are generated based on the color segment. At some homogeneous regions such as the background, the proposed method can obtain more stable disparity values than the stereo matching.

5. CONCLUSION

We presented an effective method to generate the disparity map using color and TOF depth cameras. Since the captured images from the TOF depth camera are noisy and distorted, we have to correct them. The color images are also processed to generate their disparity maps. With these data, we generated the correct disparity maps. Since we did not use the stereo matching process, the proposed method is simple and fast.

ACKNOWLEDGEMENT

This research was supported by the The Ministry of Knowledge Economy (MKE), 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] 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, Jan. 2005.

[2] 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, Dec. 2006.

[3] 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, Nov. 2008.

[4] E.K. Lee, Y.K. Jung, and Y.S. Ho, "3-D Video Generation using Foreground Separation and Disocclusion Detection," *Proc. of 3DTV Conference*, pp. 69(1-4), May 2010.

[5] Z. Zhang, "A Flexible New Technique for Camera Calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 22, no. 11, pp. 1330-1334, Oct. 2000.

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

[7] 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, July 2009.

[8] Y.S. Kang and Y.S. Ho, "High-Quality Multi-view Depth Generation using Multiple Color and Depth Cameras," *Proc. of Hot Topics in 3D*, pp. 1405-1408, July 2010.

[9] 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.

[10] D. Comaniciu and P. Meer, "Mean Shift: A Robust Approach Toward Feature Space Analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 24, no. 4, pp. 603-619, May 2002.