Three-dimensional free view reconstruction in axially distributed image sensing

Chang-Taek Seo[1,2], Shin-Won Kang[2], and Myungjin Cho[3,*]

[1]*3D Display Components Support Center, Gumi Electronics and Information Technology Research Institute, Gumi-si, Gyeongsangbuk-do 730-853, Korea*

[2]*School of Electronics Engineering, College of IT Engineering, Kyungpook National University, Daegu 702-701, Korea*

[3]*Department of Electrical, Electronic, and Control Engineering, IITC, Hankyong National University, 327 Chungang-ro, Anseong-si, Gyonggi-do 456-749, Korea*

**Corresponding author: mjcho@hknu.ac.kr*

Received January 31, 2017; accepted April 28, 2017; posted online May 15, 2017

In this Letter, we propose a three-dimensional (3D) free view reconstruction technique in axially distributed image sensing (ADS). In typical integral imaging, free view reconstructed images can be obtained by tilting all elemental images or tilting the reconstruction plane due to large lateral perspectives for 3D objects. In conventional ADS, the reconstructed images at only a front view can be generated since the sensor is moved along with its optical axis so that it has small lateral perspectives for 3D objects. However, the reconstructed 3D images at any viewing point may be obtained because the virtual viewing camera may capture these slightly different perspectives for 3D objects. Therefore, in this Letter, we employ the virtual viewing camera to visualize the 3D images at the arbitrary viewing point. To support our proposed method, we show the experimental results.

OCIS codes: 110.0110, 110.6880.

doi: 10.3788/COL201715.081102.

Three-dimensional (3D) acquisition has been the most important technique for augmented reality or virtual reality industries. Stereo vision[1–9] can obtain 3D information of objects using disparity between two images captured by a stereo camera. However, due to lack of perspectives for 3D objects in this technique, accurate 3D information may not be generated and the number of viewing points is limited. By contrast, multi-view or super multi-view[10–24] imaging techniques can obtain more accurate 3D information because they use multiple cameras for pickup of 3D objects so that more perspectives and viewing points of 3D objects can be recorded. However, these conventional 3D imaging techniques have complicated system structures (e.g., synchronization among cameras) and require a lot of cameras. To overcome these drawbacks of conventional 3D imaging methods, axially distributed image sensing (ADS) has been proposed[25–28]. It has a very simple system structure and requires only a single camera. It moves the camera along with its optical axis to record slightly different perspectives of 3D objects.

In typical 3D imaging techniques such as multi-view imaging, super multi-view imaging, or integral imaging, the reconstructed 3D image at the arbitrary viewing point (i.e., free view reconstruction) can be generated since multiple images captured by multiple cameras have different perspectives[29,30]. On the other hand, in ADS, it may be difficult to obtain free view images due to lack of the large perspectives for 3D objects. Therefore, in this Letter, we propose a free view reconstruction method in the ADS technique. Our method utilizes elemental images with slightly different perspectives by ADS and sets a virtual camera at the arbitrary position for free view reconstruction. To support our proposed method, we show the experimental results.

The remainder of this Letter is organized as follows. First, we overview the pickup and reconstruction processes of ADS, present free view reconstruction of ADS, and discuss its limitations. Then, to support our proposed method, we show the experimental results. Finally, we conclude with summary.

Free view reconstructions for integral imaging by tilting elemental images and the reconstruction plane have been previously reported[29,30]. Since more perspectives of elemental images can be generated in integral imaging, it is easy to implement free view reconstruction[29,30]. However, in ADS, it may be difficult to obtain the reconstructed images at any viewing point due to lack of perspectives in elemental images. Therefore, in this Letter, we employ the virtual viewing camera for free view reconstruction of ADS. We set the position of the virtual viewing camera arbitrarily and then find the matching pixels from the elemental images captured by ADS.

Figure 1 illustrates the basic concept of the conventional ADS[26]. This technique has a very simple structure (e.g., one-dimensional movement of the camera) and can obtain the multi perspectives of 3D objects. In addition, since the resolution of each elemental image is the same as the camera, high-resolution reconstructed 3D images can be provided.

Moving the single camera along with its optical axis, as depicted in Fig. 1, the elemental images are recorded with slightly different perspectives and magnification ratios. Using these features of ADS, the reconstructed 3D images

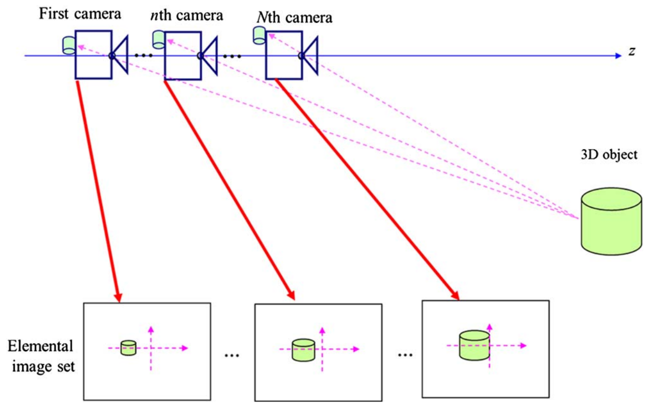

Fig. 1. Concept of pickup in ADS.

at the front viewing point can be obtained, as shown in Fig. 2(a).

The reconstruction process is based on the inverse ray projection model, as shown in Fig. 2(b)[26]. We suppose that the reconstruction plane of the 3D image (P) is at distance z. For a fixed z, the reconstruction planes are generated within the field of view of the Nth camera, as shown in Fig. 2(a). Thus, we can consider the Nth camera as the virtual viewing camera for the reconstruction planes. In addition, an image pixel on each reconstruction plane is projected back to the elemental image through each virtual pinhole. The inversely projected image pixel position is demagnified according to the distance z. Therefore, for N cameras, the reconstruction process can be implemented by

$$P(x, y, z) = \frac{1}{N}\sum_{n=1}^{N} p_n(-x/M_n, -y/M_n), \tag{1}$$

where $M_n = z_n/g = [z - (n-1)\Delta - z_1]/g$ is the magnification ratio for the nth elemental image, Δ is the distance between two adjacent cameras, g is the gap between the imaging lens and the image sensor, p_n is the matching pixel in the nth elemental image, and N is the total number of elemental images, respectively.

In fact, free view reconstruction can help 3D object recognition to be more accurate because it has more information of the 3D object than the conventional reconstruction. However, in ADS, it is not easy to obtain more information of 3D objects because of low perspectives from elemental images. Therefore, in this Letter, we employ the virtual viewing camera to produce more 3D information in ADS.

Figure 3 shows the principle of free view reconstruction for ADS using the virtual viewing camera. In our proposed method, we first define the Nth camera as the virtual viewing camera, as shown in Fig. 3. The virtual viewing camera can be shifted and rotated to the desired position. When the camera is moved, the reconstruction planes are also moved with the same parameters. For a simple calculation, we consider the movement of both the virtual

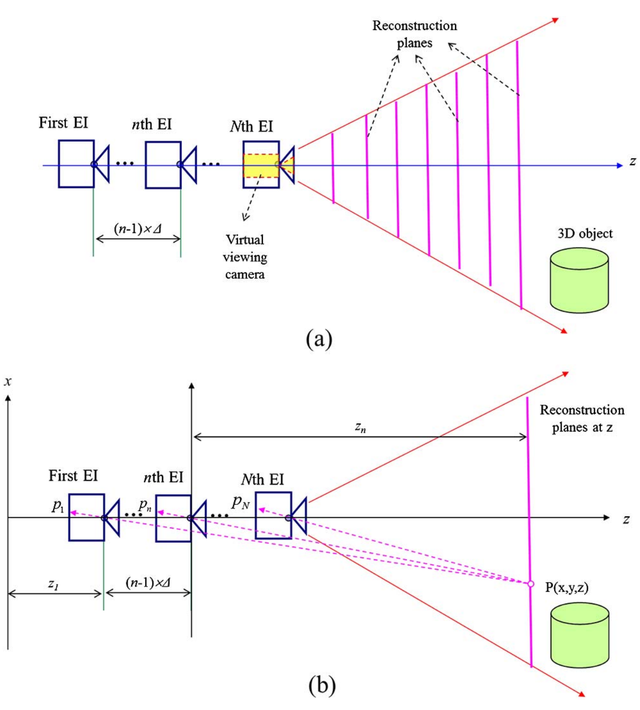

Fig. 2. Reconstruction process in ADS. (a) Multiple reconstruction planes. (b) Ray mapping between the reconstruction plane and elemental images.

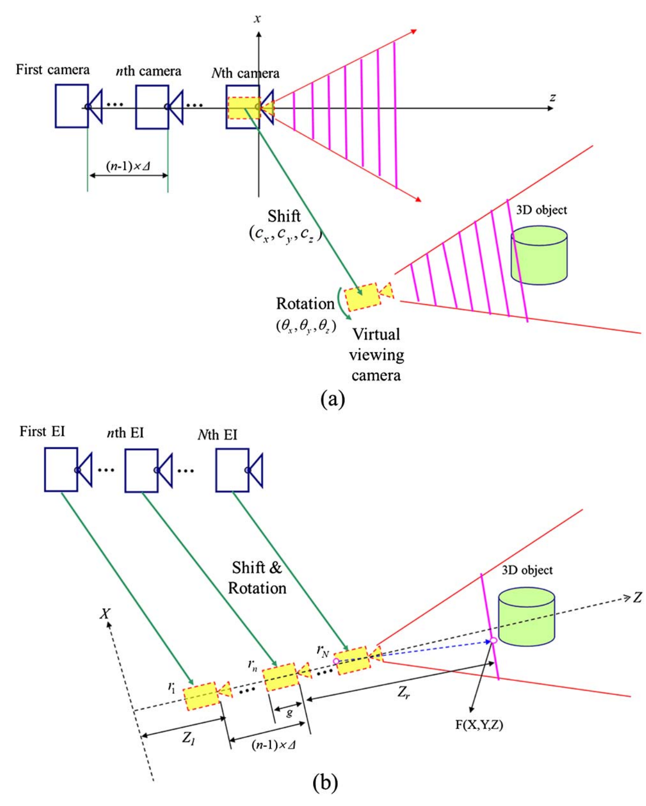

Fig. 3. (a) Movement of the virtual viewing camera for the proposed free view reconstruction. (b) Generation of the plane image from the proposed free view reconstruction.

viewing camera and the reconstruction planes. Here, the rotation and the shift matrix operators are

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = R \begin{bmatrix} x \\ y \\ z \end{bmatrix} - T, \quad (2)$$

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & -\sin\theta_x \\ 0 & \sin\theta_x & \cos\theta_x \end{bmatrix} \begin{bmatrix} \cos\theta_y & 0 & \sin\theta_y \\ 0 & 1 & 0 \\ -\sin\theta_y & 0 & \cos\theta_y \end{bmatrix} \begin{bmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (3)$$

$$T = \begin{bmatrix} s_x \\ s_y \\ s_z \end{bmatrix}, \quad (4)$$

where X, Y, Z are the destination position of the desired reconstruction plane in Cartesian coordinates, R is the rotation matrix, T is the translation matrix, θ_x, θ_y, θ_z are the rotation angles in Cartesian coordinates, and s_x, s_y, s_z are the shifting distance from the front view position, respectively.

After moving them, free view reconstruction of ADS, as shown in Fig. 3, can be implemented by finding the matching pixels between the image pixel of the moved reconstruction plane and each elemental image. The final image by free view reconstruction is generated by summation of all matched elemental images. This is given by

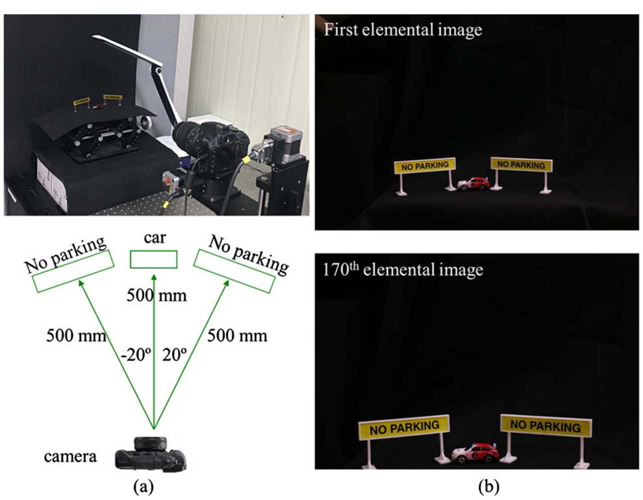

Fig. 4. (a) Experimental setup. (b) Examples of the recorded elemental images.

$$F(X, Y, Z) = \frac{1}{N} \sum_{n=1}^{N} r_n(-X/M'_n, -Y/M'_n), \quad (5)$$

where $M'_n = Z_n/g = [Z - (n-1)\Delta - Z_1]/g$ is the magnification ratio of the virtual viewing camera and r_n are the matching pixels between the virtual viewing camera and the nth elemental image.

To calculate the parallax between the free view reconstructed images, we used the distance of the center point between them (i.e., disparity) as

$$D(\theta_1, \theta_2, z_r) = z_r(\sin|\theta_1| + \sin|\theta_2|), \quad (6)$$

where θ_1, θ_2 are two free view angles for reconstruction, z_r is the reconstruction depth, and D is the disparity between two free view reconstructed images, respectively.

To demonstrate our method, we carry out the optical experiment. To capture the elemental images, we use a commercial digital camera (Nikon D3200) with a focal length $f = 50$ mm. It has 2184 (H) × 1426 (V) pixels and it is moved by uniform step $\Delta z = 1$ mm. We capture 170 elemental images to obtain the perspectives for 3D objects, as shown in Fig. 4. 3D objects are two sign boards with “NO PARKING” characters and a toy car with “6” character. Two sign boards are tilted by −20° and 20° from the optical axis, respectively. All 3D objects are located at 500 mm from the camera. In this experiment, the parallax between two free view reconstructed images with 20° left and right views can be calculated as the disparity. Thus, the parallax $D = 342.0201$ mm.

Figure 5 shows the reconstruction results with various reconstruction depths using conventional ADS reconstruction (i.e., the front view). The focused image was obtained when the reconstruction plane was at $z = 500$ mm, which is the same as the original position of the 3D object (toy car). Otherwise, the two sign boards were blurred.

Next, we moved the virtual viewing camera to the position as shown in Fig. 6. We generated the plane images according to distance z. Examples are shown in the right of Fig. 6. We can generate the right reconstructed image at the distance $z = 300$ mm. The objects were reconstructed in the center of the plane image. This implies that the proposed free view reconstruction method is more effective.

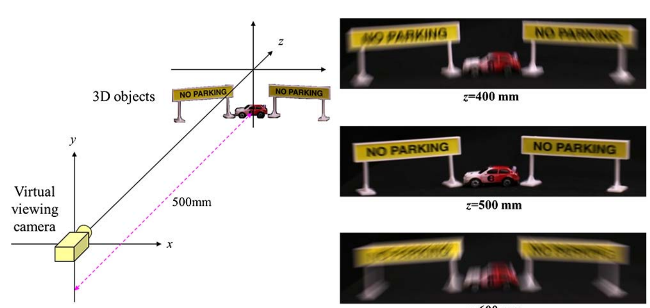

Fig. 5. Experimental results of the conventional ADS.

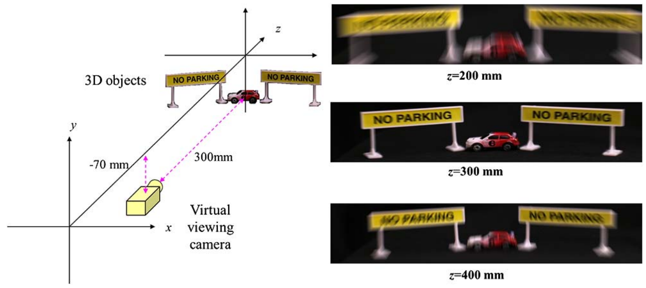

Fig. 6. Example of free view reconstruction by moving the virtual viewing camera.

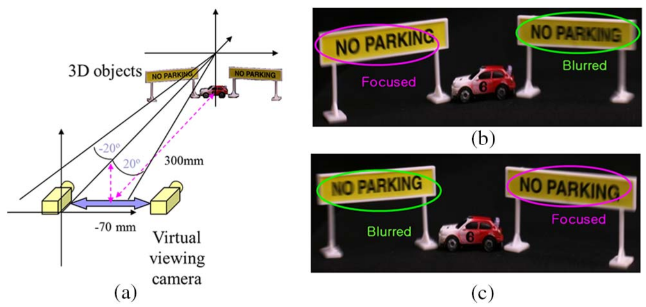

Fig. 7. Examples of free view reconstruction by moving and rotating the virtual viewing camera. (a) The free view reconstruction setup, (b) the free view reconstruction result with a tilting angle of −20°, and (c) the free view reconstruction result with a tilting angle of 20°.

In addition, the virtual viewing camera was moved and rotated as shown in Fig. 7. To do this, we can observe different perspectives of 3D objects. As shown in Fig. 7(b), only the left sign board is focused and otherwise is defocused in the free view reconstruction result with a tilting angle of −20°. Using free view reconstruction with a tilting angle of 20°, we can observe that the right sign board is focused, as shown in Fig. 7(c). From the result of Fig. 7, we prove that the proposed free view reconstruction method can generate plane images at any position and any angle of virtual viewing camera.

We propose a 3D free view reconstruction technique for ADS. In conventional ADS, only a front view reconstructed image can be obtained but the reconstructed 3D images at any viewing point can be obtained in our proposed method. Using our proposed method, more information can be achieved for 3D objects so it can be applied to various 3D applications. In future works, we will investigate solutions to solve the limitation of the viewing angle and depth range in ADS.

This work was supported by the Technological Innovation R&D Program (No. S2405402) funded by the Small and Medium Business Administration (SMBA, Korea).

References

1. F. Tombari, S. Mattoccia, L. Di Stefano, and E. Addimanda, in *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR)* (2008).
2. M. Gerrits and P. Bekaert, in *Proceedings of Canadian Conference on Computer and Robot Vision (CRV)* (2006).
3. F. Tombari, S. Mattoccia, and L. Di Stefano, in *Pacific-Rim Symposium on Image and Video Technology* (2007).
4. K. Yoon and I. Kweon, IEEE Trans. PAMI **28**, 650 (2006).
5. Z. Wang and Z. Zheng, in *IEEE Conference on Computer Vision and Pattern Recognition* (2008).
6. H. Hirschmüller, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2006, PAMI 30* (2008), p. 328.
7. F. Tombari, S. Mattoccia, and L. Di Stefano, IEEE Trans. Pattern Anal. Mach. Intell. **31**, 129 (2009).
8. D. Min, J. Lu, and M. Do, in *ICCV '11 Proceedings of the 2011 International Conference on Computer Vision (ICCV)* (2011).
9. L. De-Maeztu, S. Mattoccia, A. Villanueva, and R. Cabeza, in *International Conference on Computer Vision (ICCV)* (2011).
10. Y. Takaki and N. Nichiyom, Opt. Express **18**, 8824 (2010).
11. D. Teng, L. Liu, and Z. Pang, in *Progress in Electromagnetic Research Symposium (PIERS)* (2016).
12. N. A. Dodgson, J. R. Moore, and S. R. Lang, in *International Broadcasting Convention* (1999), p. 2.
13. Y. Takaki, Y. Urano, S. Kashiwada, H. Ando, and K. Nakamura, Opt. Express **19**, 704 (2011).
14. W. Matusik and H. Pfister, ACM Trans. Graphics (TOG) **23**, 814 (2004).
15. G. Lippmann, C. R. Acad. Sci. Paris **146**, 446 (1908).
16. S.-H. Hong, J.-S. Jang, and B. Javidi, Opt. Express **12**, 483 (2004).
17. J.-S. Jang and B. Javidi, Opt. Lett. **27**, 1144 (2002).
18. R. Ng, M. Levoy, M. Bredif, G. Duval, M. Horowitz, and P. Hanrahan, in *Stanford University Computer Science Tech Report CSTR 2005-02* (2005).
19. R. Ng, "Digital light field photography," Ph.D. dissertation (Stanford University, 2006).
20. A. Stern and B. Javidi, Proc. IEEE **94**, 591 (2006).
21. J. Hong, Y. Kim, H.-J. Choi, J. Hahn, J.-H. Park, H. Kim, S.-W. Min, N. Chen, and B. Lee, Appl. Opt. **50**, H87 (2011).
22. J.-S. Jang and B. Javidi, Opt. Lett. **27**, 324 (2002).
23. R. Martinez-Cuenca, G. Saavedra, M. Martinez-Corral, and B. Javidi, Proc. IEEE **97**, 1067 (2009).
24. M. Cho, M. Daneshpanah, I. Moon, and B. Javidi, Proc. IEEE **99**, 556 (2011).
25. M. Cho and B. Javidi, J. Disp. Technol. **9**, 56 (2013).
26. R. Schulein, M. Daneshpanah, and B. Javidi, Opt. Lett. **34**, 2012 (2009).
27. M. Cho and D. Shin, J. Opt. Soc. Korea **17**, 410 (2013).
28. D. Shin, M. Cho, and B. Javidi, Opt. Lett. **35**, 3646 (2010).
29. Y. S. Hwang, S. H. Hong, and B. Javidi, J. Disp. Technol. **3**, 64 (2007).
30. M. Cho and B. Javidi, J. Disp. Technol. **5**, 345 (2009).