Journal of information and communication convergence engineering
http://jicce.org

J. Inf. Commun. Converg. Eng. 15(3): 193-198, Sep. 2017 **Regular paper**

Three-Dimensional Visualization Technique of Occluded Objects Using Integral Imaging with Plenoptic Camera

Min-Chul Lee[1], Kotaro Inoue[2], Masaharu Tashiro[1], and Myungjin Cho[2*], *Member, KIICE*

[1]Department of Computer Science and Electronics, Kyushu Institute of Technology, Fukuoka 820-8502, Japan
[2]Department of Electrical, Electronic, and Control Engineering, IITC, Hankyong National University, Anseong 17579, Korea

Abstract

In this study, we propose a three-dimensional (3D) visualization technique of occluded objects using integral imaging with a plenoptic camera. In previous studies, depth map estimation from elemental images was used to remove occlusion. However, the resolution of these depth maps is low. Thus, the occlusion removal accuracy is not efficient. Therefore, we use a plenoptic camera to obtain a high-resolution depth map. Hence, individual depth map for each elemental image can also be generated. Finally, we can regenerate a more accurate depth map for 3D objects with these separate depth maps, allowing us to remove the occlusion layers more efficiently. We perform optical experiments to prove our proposed technique. Moreover, we use MSE and PSNR as a performance metric to evaluate the quality of the reconstructed image. In conclusion, we enhance the visual quality of the reconstructed image after removing the occlusion layers using the plenoptic camera.

Index Terms: Integral imaging, Occlusion removal, Plenoptic camera

I. INTRODUCTION

Integral imaging is a three-dimensional (3D) photography technique first proposed by G. Lippmann in 1908 [1]. It can provide full color, full parallax, and continuous viewing points without special viewing glasses. Integral imaging does not require a coherent light source as required in holography; hence, it is remarkable for next-generation 3D display and 3D object recognition [2–12].

Integral imaging consists of two main process stages: pickup and reconstruction. In the pickup stage, 3D object rays are captured through a lenslet array by image sensors, such as a charge-coupled device (CCD). These 3D object rays have different perspectives, where these are referred to as elemental images. In the reconstruction stage, 3D images can be displayed or reconstructed in a 3D space by back-projecting elemental images onto a lenslet array or a virtual pinhole array. In this study, we use the optical setup for the pickup process and the computational integral imaging reconstruction (CIIR) [2] for the reconstruction process. The advantages of CIIR are that it does not require any optical systems and has flexibility for data processing in a computer simulation [3–5]. Therefore, integral imaging may be an optimal solution for 3D image reconstruction and recognition because of its simple structure. In addition, the occlusion removal of integral imaging has been expected and researched for 3D object recognition and tracking [6–12].

Human eyes can observe a wide scene; however, the region of interest (ROI) can be small. Furthermore, most

Received 04 September 2017, **Revised** 11 September 2017, **Accepted** 19 September 2017
***Corresponding Author** Myungjin Cho (E-mail: mjcho@hknu.ac.kr, Tel: +82-31-670-5298)
Department of Electrical, Electronic, and Control Engineering, Hankyong National University, 327 Chungang-ro, Anseong 17579, Korea.

Min-Chul Lee and Kotaro Inoue equally contributed to this work.

Open Access **https://doi.org/10.6109/jicce.2017.15.3.193** print ISSN: 2234-8255 online ISSN: 2234-8883

parts of the ROI have some occlusions. Thus, the accident rate for robots or unmanned auto vehicles may decrease if these occlusions can be removed. Integral imaging can remove occlusions using a depth map generated from elemental images because it can record multi-view images.

The plenoptic camera invented by Ng's research [13] is cost effective and has a unique photography technique, which is different from conventional cameras. A conventional camera system can only record a 2D image in a 3D space. In contrast, the plenoptic camera can capture a light field, which has the intensity and direction of rays, and reconstruct a 3D image. In addition, reconstructing all-in-focus images and estimating the depth map by a single shot are possible with the plenoptic camera. In recent research, the depth map is estimated from low-resolution elemental images by a lenslet array. Thus, the resolution and the accuracy of the depth map are low. However, in this study, the elemental images and the depth map with a high resolution are obtained using synthetic aperture integral imaging (SAII) with a plenoptic camera.

The rest of the paper is organized as follows: Section II presents the theory of our proposed method; Section III shows the experimental results; and Section IV concludes and summarizes our method.

II. OCCLUSION REMOVAL USING INTEGRAL IMAGING WITH A PLENOPTIC CAMERA

The plenoptic camera records four-dimensional (4D) light field by locating the lenslet array in front of the image sensor [13]. A 4D light field is defined as the direction of rays in intersections between two planes. Fig. 1 illustrates the concept of a 4D light field.

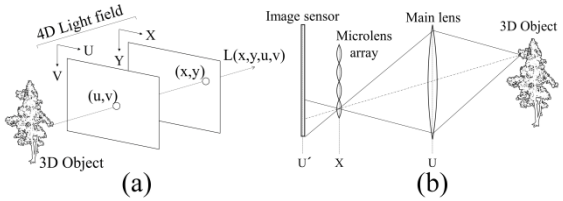


Fig. 1. 4D light field: (a) overview and (b) plenoptic camera.

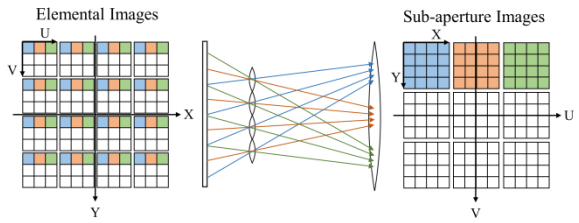


Fig. 2. Sub-aperture images.

$L(x,y,u,v)$ is depicted in Fig. 1(a) as a light field obtained from 3D objects, where rays can be presented by a straight line between the coordinates of the XY plane (x,y) and the UV plane (u,v). In a practical plenoptic camera, the XY plane is the lenslet array plane, while the UV plane is the main lens plane. Fig. 1(b) shows the XU plane of the plenoptic camera. The object rays are imaged on the image sensor through both the main lens and the lenslet array, where the X plane is determined by a system parameter of the lenslet array. However, the main lens cannot record the image (i.e., U plane). In addition, the information of the U plane is recorded as the pixel position of the elemental images on the image sensor plane U′. The plenoptic camera used in this paper is Lytro Illum and can record 40M rays.

The distance between the image sensor and the lens may be adjusted for an in-focus image in the conventional camera. However, in the plenoptic camera, the image can be reconstructed at any focal point because the virtual image sensor plane of the light field is moved. This is referred to as refocus. The basic principle of the refocus proposed in Ng's paper [14] is that sub-aperture images are shifted and summed, as depicted in Fig. 2.

Let us consider the refocus in 2D as shown in Fig. 3 for simplicity. The light field obtained from the 3D objects, L, is presented by a straight line between the coordinate of the U plane (main lens) and the coordinate of the X plane (lenslet array). In Fig. 3, the light field $L(x,u)$ is defined by a straight line between the coordinate of the U plane, u and the coordinate of the X plane, x. The light field of the X′ plane, L', can be written with L as follows when the image on the X′ plane with distance F' from the main lens is reconstructed and $\alpha = F'/F$:

$$L'(x',u) = L\left(u + \frac{x'-u}{\alpha}, u\right). \tag{1}$$

This equation can be expanded by considering the 4D light field as follows:

$$L'(x',y',u,v) = L\left(u + \frac{x'-u}{\alpha}, v + \frac{y'-v}{\alpha}, u, v\right). \tag{2}$$

The 4D light field can be transformed as a 2D image using integration with u and v. Thus, the 2D image at distance F', $E_{F'}$, can be calculated as follows:

$$E_{F'}(x',y') = \frac{1}{F'^2}\int\int L'(x',y',u,v)\,du dv. \tag{3}$$

The 2D image, $E_{F'}$, can be obtained as follows by substituting Eq. (2) into Eq. (3):

$$E_{F'}(x',y') = \frac{1}{\alpha^2 F'^2}\int\int L\left[u\left(1-\frac{1}{\alpha}\right)+\frac{x'}{\alpha}, v\left(1-\frac{1}{\alpha}\right)+\frac{y'}{\alpha}, u, v\right] du dv. \tag{4}$$

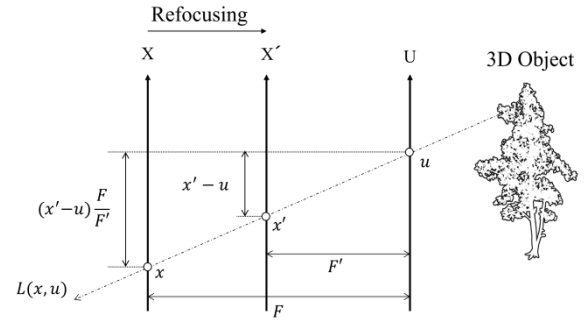


Fig. 3. Refocusing.

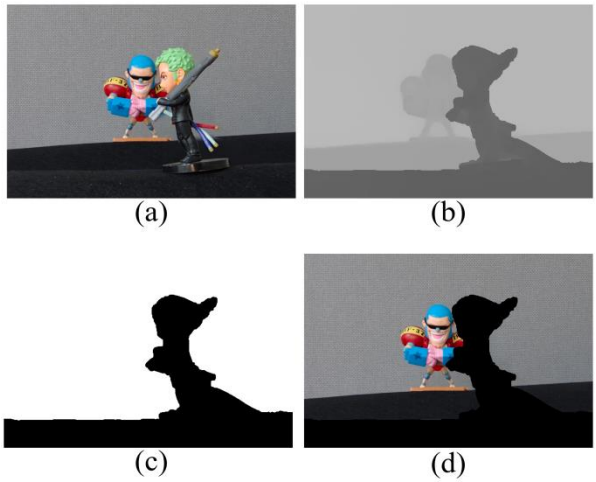


Fig. 4. Image of equations (k = 3, l = 3): (a) original image, (b) depth map, (c) occlusion layer, and (d) image after occlusion layer removal.

As depicted in Eq. (4), $E_{F'}$ is presented by magnifying and shifting the sub-aperture images with the α term. Thus, considering the uth column and the vth row sub-aperture image, $L(u, v)$, Eq. (4) as follows [14]:

$$E_{F'}(x',y')=\frac{1}{\alpha^2F'^2}\iint L^{(u,v)}\left[u\left(1-\frac{1}{\alpha}\right)+\frac{x'}{\alpha},v\left(1-\frac{1}{\alpha}\right)+\frac{y'}{\alpha}\right]dudv. \quad (5)$$

The expanding coefficient of the image, α, can be ignored in the reconstruction stage because its effect is similar in all the images.

The depth map estimation can be implemented by ray tracing, stereo matching between sub-aperture images, or contrast variation using multiple refocus images. However, the depth map estimation algorithm of the Lytro Company is not revealed. Hence, we present herein the specification of the output depth map from the Lytro camera. The depth map obtained by Lytro software has a text file and an all-in-focus image. The depth map has 16-bit grayscale. The brightness in the depth map is matched with Lambda written in the text file, where LambdaMin is Brightness 0, and LambdaMax is Brightness 2^{16}–1. Lambda is not mapped to a physical distance. Thus, we need a modification process to find the physical distance.

The occlusion removal can be performed by removing the occlusion part, which is satisfied with a certain threshold of the depth map. We assume that the farthest depth has a high brightness. Thus, the threshold can be set as follows:

$$OL_{(k,l)}(x,y)=\begin{cases}1, & D_{(k,l)}(x,y)\geq \text{threshold}\\ 0, & \text{otherwise}\end{cases}, \quad (6)$$

where $OL(k,l)$ is the kth column and lth row occlusion layer, and (x, y) is pixel position. The occlusion layer is removed from the elemental images using Eq. (6). Let the elemental image and the elemental image without occlusion be EI and $OREI$, respectively.

$$OREI_{(k,l)}(x,y)=\begin{cases}EI_{(k,l)}(x,y), & OL_{(k,l)}(x,y)=1\\ 0, & \text{otherwise}\end{cases}. \quad (7)$$

Fig. 4 shows the calculation results of Eqs. (6) and (7).

Finally, we perform the computational reconstruction using EIs. Table 1 shows the system parameters for the computational reconstruction, which is almost similar to the shift and sum of the sub-aperture images for the plenoptic camera. The numbers of the shifting pixels of each elemental image in the x and y directions S_x and S_y are obtained as follows when the reconstructed image I is located at distance d:

$$S_x=\frac{N_x pf}{c_x d}, \quad S_y=\frac{N_y pf}{c_y d}. \quad (8)$$

The reconstructed image I is presented as

$$I(x,y,d)=\frac{1}{O(x,y,d)}\sum_{k=0}^{K-1}\sum_{l=0}^{L-1}OREI_{(k,l)}\left(x-kS_x,y-lS_y\right), \quad (9)$$

where K and L are the index of the elemental images in the x and y directions, and $O(x,y,d)$ is the superposition matrix for each reconstruction depth d.

Table 1. Definition of the parameters

Parameter	Definition
EI	Elemental images
D	Depth maps
OREI	Occlusion-removed elemental images
OL	Occlusion layer
S	Number of shifted pixels for reconstruction
N	Number of pixels for each elemental image
p	Moving gap between the image sensor
f	Focal length of the image sensor
C	Size of the image sensor
d	Reconstruction distance

Table 2. Specification of LYTRO ILLUM

Parameter	Specification
Number of pixels for the image sensor	7728×5368
Number of pixels for the export image	2022×1404
Size of the image sensor (mm)	10.82×7.52
Size of the pixel (μm)	1.4
Number of microlens	540×433
Diameter of each microlens (μm)	20.0

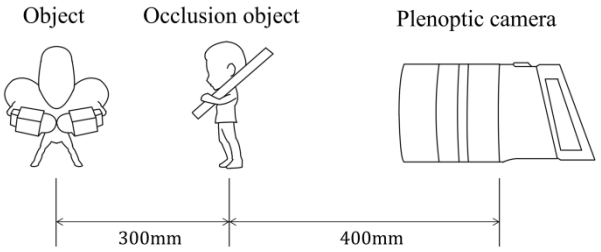


Fig. 5. Optical experiment setup.

III. EXPERIMENTS AND RESULTS

We used the Lytro Illum camera in the pickup stage of our experiments. Table 2 presents its specification. The focal length of the camera lens (*f*) is 70 mm; the moving distance between the cameras (*p*) is 10 mm; and the refocus range is set as 40–75 mm. We recorded 5(H) × 5(V) elemental images by SAII. Two 3D object scenes are used: with occlusion and without occlusion. The 3D scene without occlusion is used as the reference for calculating the mean square error (MSE) and the peak signal-to-noise ratio (PSNR). Fig. 5 shows the distance between the camera and the 3D objects. The occlusion is placed at the left shoulder of the 3D object.

Fig. 6 shows the experimental results at a reconstruction depth of d = 620 mm. Fig. 6(a) depicts the reconstructed image using the elemental images without occlusion, while Fig. 6(b) presents the reconstructed image using elemental images with occlusion. Fig. 6(c) shows the reconstruction result using our proposed method (elemental images removed the occlusion layer). Fig. 6(d)–(f) illustrate the enlarged images for Fig. 6(a)–(c), respectively. Characters "BF-37" on the shoulders of the 3D object can be recognized well in Fig. 6(c) and (f). We calculate the MSE and the PSNR between the reconstruction result and the reference image (Fig. 6(a)) to evaluate the visual quality of the reconstructed images. Fig. 7 shows the graph of the MSE and the PSNR for the visual quality evaluation. Both the MSE and PSNR values are worse in our proposed method; however, it means that the brightness is lost when

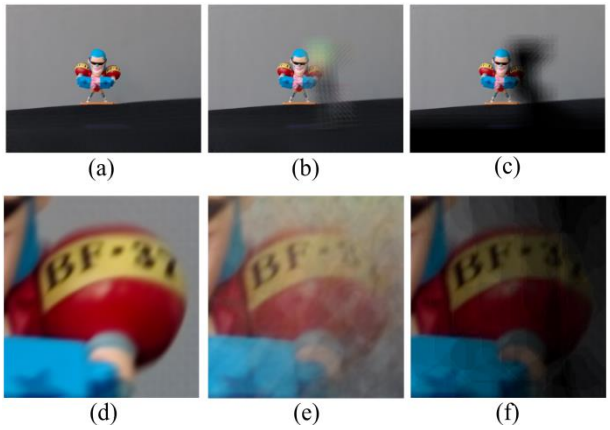


Fig. 6. Experimental results of the reconstructed images (a, b, c) and its enlarged images (d, e, f) at d = 620 mm.

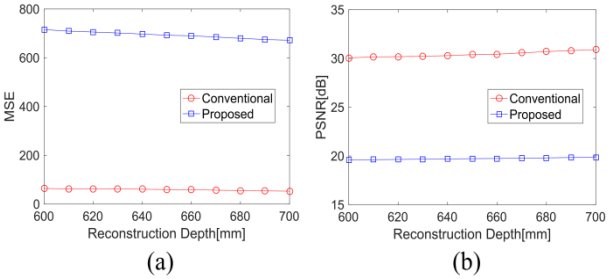


Fig. 7. MSE and PSNR results via various reconstruction depths.

the occlusion layer is removed using our method. We will study a method that can interpolate elemental or reconstructed images to improve the MSE and the PSNR.

IV. CONCLUSIONS

In this study, we proposed a new occlusion removal method of integral imaging using a plenoptic camera. We also presented a 3D visualization of the occluded objects. In conventional techniques, the depth map is estimated using elemental images recorded through a lenslet array. The resolution and accuracy of the depth map are not sufficient for occlusion removal because of the low resolution of the elemental images. However, in our proposed method, the depth map with a high resolution may be generated using a plenoptic camera, and the occlusion layers may be accurately removed. Meanwhile, the image information of the removed occlusion parts is partially lost, and the 3D reconstructed image has some distortions. We will solve this problem by studying the improved occlusion removal technique with interpolation between the elemental or reconstructed images.

ACKNOWLEDGMENTS

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. NRF-2017R1D1A1B03030343).

REFERENCES

[1] G. Lippmann, "La photographie integrale," *Comptes Rendus de l'Académie des Sciences*, vol. 146, pp. 446–451, 1908.

[2] S. H. Hong, J. S. Jang, and B. Javidi, "Three-dimensional volumetric object reconstruction using computational integral imaging," *Optical Express*, vol. 12, no. 3, pp. 483-491, 2004.

[3] B. Javidi, R. Ponce-Diaz, and S. H. Hong, "Three-dimensional recognition of occluded objects by using computational integral imaging," *Optics Letters*, vol. 31, no. 8, pp. 1106-1108, 2006.

[4] D. H. Shin and H. Yoo, "Image quality enhancement in 3D computational integral imaging by use of interpolation methods," *Optics Express*, vol. 15, no. 19, pp. 12039-12049, 2007.

[5] H. Yoo and D. H. Shin, "Improved analysis on the signal property of computational integral imaging system," *Optics Express*, vol. 15, no. 21, pp. 14107-14114, 2007.

[6] S. H. Hong and B. Javidi, "Distortion-tolerant 3D recognition of occluded objects using computational integral imaging," *Optics Express*, vol. 14, no. 25, pp. 12085-12095, 2006.

[7] M. Cho and B. Javidi, "Three-dimensional tracking of occluded objects using integral imaging," *Optics Letters*, vol. 33, no. 23, pp. 2737-2739, 2008.

[8] D. H. Shin, B. G. Lee, and J. J. Lee, "Occlusion removal method of partially occluded 3D object using sub-image block matching in computational integral imaging," *Optics Express*, vol. 16, no. 21, pp. 16294-16304, 2008.

[9] B. G. Lee and D. H. Shin, "Enhanced computational integral imaging system for partially occluded 3D objects using occlusion removal technique and recursive PCA reconstruction," *Optics Communications*, vol. 283, no. 10, 2084-2091, 2010.

[10] M. Zhang, Y. Piao, and E. S. Kim, "Occlusion-removed scheme using depth-reversed method in computational integral imaging," *Applied Optics*, vol. 49, no. 14, pp. 2571-2580, 2010.

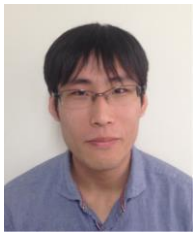
[11] J. J. Lee, B. G. Lee, and H. Yoo, "Image quality enhancement of computational integral imaging reconstruction for partially occluded objects using binary weighting mask on occlusion areas," *Applied Optics*, vol. 50, no.13, pp. 1889-1893, 2011.

[12] Z. Zhou, Y. Yuan, X. Bin, and Q. Wang, "Enhanced reconstruction of partially occluded objects with occlusion removal in synthetic aperture integral imaging," *Chinese Optics Letters*, vol. 9, no. 4, article ID. 041002, 2011.

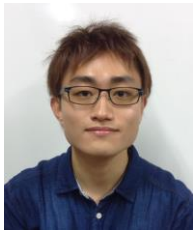
[13] R. Ng, M. Levoy, M. Bredif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," Computer Science Technical Report, vol. 2005, no. 2, pp. 1-11, 2005.

[14] R. Ng, "Digital light field photography," Ph.D. dissertation, Stanford University, CA, 2006.

Min-Chul Lee
received the B.S. degrees in Telecommunication Engineering from Pukyong National University, Busan, Korea, in 1996, and the M.S. and Ph.D. degrees from Kyushu Institute of Technology, Fukuoka, Japan, in 2000 and 2003, respectively. He is an assistant professor at Kyushu Institute of Technology in Japan. His research interests include medical imaging, blood flow analysis, 3D display, 3D integral imaging, and 3D biomedical imaging.



Kotaro Inoue
received the B.S. and M.S. degrees in computer science and electronics from Kyushu Institute of Technology, Fukuoka, Japan, in 2015 and 2017, respectively. He is currently a doctoral student at Hankyong National University in Korea. His research interests include visual feedback control, 3D display, 3D reconstruction, and 3D integral imaging.



Masaharu Tashiro
is an undergraduate student in the Department of Computer Science and Electronics, Kyushu Institute of Technology in Japan. His current research interests include 3D integral imaging.

Myungjin Cho
received the B.S. and M.S. degrees in Telecommunication Engineering from Pukyong National University, Pusan, Korea, in 2003 and 2005, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Connecticut, Storrs, CT, USA, in 2010 and 2011, respectively. Currently, he is an associate professor at Hankyong National University in Korea. He worked as a researcher at Samsung Electronics in Korea, from 2005 to 2007. His research interests include 3D display, 3D signal processing, 3D biomedical imaging, 3D photon counting imaging, 3D information security, 3D object tracking, 3D underwater imaging, and 3D visualization of objects under inclement weather conditions.