

Occlusion removal method of partially occluded 3D object using sub-image block matching in computational integral imaging

Dong-Hak Shin¹, Byung-Gook Lee¹ and Joon-Jae Lee^{2*}

¹Dept. of Visual Contents, Dongseo University, San69-1, Jurye2-Dong, Sasang-Gu, Busan 617-716, Korea

²Dept. of Game Mobile Contents, Keimyung University, Daemyung3-Dong Nam-Gu, Daegu 705-701, Korea

*Corresponding author: joonlee@kmu.ac.kr

Abstract: In this paper, we propose an occlusion removal method using sub-image block matching for improved recognition of partially occluded 3D objects in computational integral imaging (CII). When 3D plane images are reconstructed in CII, occlusion degrades the resolution of reconstructed images. To overcome this problem, we apply the sub-image transform to elemental image array (EIA) and these sub-images are employed using block matching method for depth estimation. Based on the estimated depth information, we remove the unknown occlusion. After completing the occlusion removal for all sub-images, we obtain the modified EIA without occlusion information through the inverse sub-image transform. Finally, the 3D plane images are reconstructed by using a computational integral imaging reconstruction method with the modified EIA. The proposed method can provide a substantial gain in terms of the visual quality of 3D reconstructed images. To show the usefulness of the proposed method we carry out some experiments and the results are presented.

©2008 Optical Society of America

OCIS codes: (100.6890) Three-dimensional image processing; (110.2990) Image formation theory; (110.6880) Three-dimensional image acquisition.

References and links

1. G. Lippmann, "La photographie integrale," C.R. Acad. Sci. **146**, 446-451 (1908).
2. A. Stern and B. Javidi, "Three dimensional image sensing, visualization, and processing using integral imaging," Proc. of IEEE **94**, 591-607 (2006).
3. B. Lee, S.-Y. Jung, S.-W. Min, and J.-H. Park, "Three-dimensional display by use of integral photography with dynamically variable image planes," Opt. Lett. **26**, 1481-1482 (2001).
4. J.-S. Jang and B. Javidi, "Improved viewing resolution of three-dimensional integral imaging by use of nonstationary micro-optics," Opt. Lett. **27**, 324-326 (2002).
5. M. Martínez-Corral, B. Javidi, R. Martínez-Cuenca, and G. Saavedra, "Multifacet structure of observed reconstructed integral images," J. Opt. Soc. Am. A **22**, 597-603 (2005).
6. D.-H. Shin, B. Lee, and E.-S. Kim, "Multidirectional curved integral imaging with large depth by additional use of a large-aperture lens," Appl. Opt. **45**, 7375-7381 (2006).
7. D.-H. Shin, S.-H. Lee, and E.-S. Kim, "Optical display of true 3D objects in depth-priority integral imaging using an active sensor," Opt. Commun. **275**, 330-334 (2007).
8. H. Arimoto and B. Javidi, "Integral three-dimensional imaging with digital reconstruction," Opt. Lett. **26**, 157-159 (2001).
9. Y. Frauel and B. Javidi, "Digital three-dimensional image correlation by use of computer-reconstructed integral imaging," Appl. Opt. **41**, 5488-5496 (2002).
10. S.-H. Hong, J.-S. Jang, and B. Javidi, "Three-dimensional volumetric object reconstruction using computational integral imaging," Opt. Express **12**, 483-491 (2004).
11. D.-H. Shin, E.-S. Kim, and B. Lee, "Computational reconstruction technique of three-dimensional object in integral imaging using a lenslet array," Jpn. J. Appl. Phys. **44**, 8016-8018 (2005).
12. S.-H. Hong and B. Javidi, "Improved resolution 3D object reconstruction using computational integral imaging with time multiplexing," Opt. Express **12**, 4579-4588 (2004).
13. S.-H. Hong and B. Javidi, "Distortion-tolerant 3D recognition of occluded objects using computational integral imaging," Opt. Express **14**, 12085-12095 (2006).

14. D.-H. Shin and H. Yoo, "Image quality enhancement in 3D computational integral imaging by use of interpolation methods," *Opt. Express* **15**, 12039-12049 (2007).
 15. H. Yoo and D.-H. Shin "Improved analysis on the signal property of computational integral imaging system," *Opt. Express* **15**, 14107-14114 (2007).
 16. B. Javidi, R. Ponce-Díaz, and S.-H. Hong, "Three-dimensional recognition of occluded objects by using computational integral imaging," *Opt. Lett.* **31**, 1106-1108 (2006).
 17. J.-S. Park, D.-C. Hwang, D.-H. Shin, and E.-S. Kim, "Resolution-enhanced 3D image correlator using computationally reconstructed integral images," *Opt. Commun.* **276**, 72-79 (2007).
 18. J.-H. Park, J. Kim, and B. Lee, "Three-dimensional optical correlator using a sub-image array," *Opt. Express* **13**, 5116-5126 (2005).
 19. D.-H. Shin, B. Lee, and E.-S. Kim, "Improved viewing quality of 3-D images in computational integral imaging reconstruction based on lenslet array model," *ETRI J.* **28**, 521-524 (2006).
 20. M. Z. Brown, D. Burschka, and S. D. Hager, "Advances in computational stereo," *IEEE Trans. Pattern Analysis and Machine Intelligence* **25**, 993-1008 (2003).
 21. J.-S. Lee, J.-H. Ko, and E.-S. Kim, "Real-time stereo object tracking system by using block matching algorithm and optical binary phase extraction joint transform correlator," *Opt. Commun.* **191**, 191-202 (2001).
 22. S.-H. Hong and B. Javidi, "Three-dimensional visualization of partially occluded objects using integral imaging," *J. Display Technol.* **1**, 354-359 (2005).
 23. J. -S. Jang and B. Javidi, "Depth and lateral size control of three-dimensional images in projection integral imaging," *Opt. Express* **12**, 3778-3790 (2004).
-

1. Introduction

The integral imaging method is attracting a great deal of interest among 3D imaging techniques as it provides full parallax and continuous viewing points [1-7]. An integral imaging system consists of two processes: pickup and display. In the pickup process, a lenslet array is used to capture the light rays emanating from 3D objects. Light rays that pass through each lenslet are recorded by using a two-dimensional image sensor, such as a charge coupled device (CCD). The captured information is known as elemental image array (EIA) which is the demagnified images of difference perspectives. In order to reconstruct 3D images from EIA, we perform the reverse process by propagating the rays from EIA through a similar lenslet array used in pickup process.

When a 3D object is reconstructed from EIA, we can use the computational reconstruction technique for some applications such as 3D visualization and depth extraction. Until now, there have been a number of computational reconstruction techniques reported [8-17]. The most noticeable computational reconstruction technique is the volumetric computational reconstruction (VCR) technique which makes use of all the elemental image information to reconstruct 3D images at any arbitrary distance [10-17]. The VCR has an advantage of freely generating the volumetric information of the reconstructed image without optical devices.

Based on VCR, a computational integral imaging (CII) has been introduced for 3D recognition of partially occluded objects [16,17]. The CII system is shown in Fig. 1. It is composed of optical pickup and VCR based on the pinhole-array model. As shown in Fig. 1(a), 3D object is picked up by a lenslet array and then recorded by a CCD camera as EIA. The VCR method based on pinhole array model is shown in Fig. 1(b). In the VCR process, the EIA is digitally reconstructed by using a computer where 3D images can be easily reconstructed at any output planes. As shown in Fig. 1(b), the EIA is mapped on the image plane inversely through each pinhole. The EIA is magnified as much as a ratio of z/g , where z is the distance between the reconstruction image plane and the virtual pinhole array and g is the distance between the EIA and the virtual pinhole array. Each magnified elemental image is overlapped each other and the reconstructed 3D images are finally produced at the reconstruction output plane. The basic principle of CII for partially occluded 3D object is to produce 3D volumetric reconstructed images and correlate with original 3D object. Therefore, the recognition performance depends on how well the 3D images are reconstructed.

In the reconstruction of a partially occluded 3D object using CII, occlusion seriously degrades the resolution of reconstructed images because it hides the 3D object to be recognized. In this paper, to overcome this problem, we propose an occlusion removal method

of partially occluded 3D object in CII. In the proposed method, EIA is transformed into sub-image domain and we remove occlusion into the sub-image array by using a sub-image block matching. Then, 3D images are reconstructed with VCR. To show the usefulness of the proposed method we carry out some experiments and the results are presented.

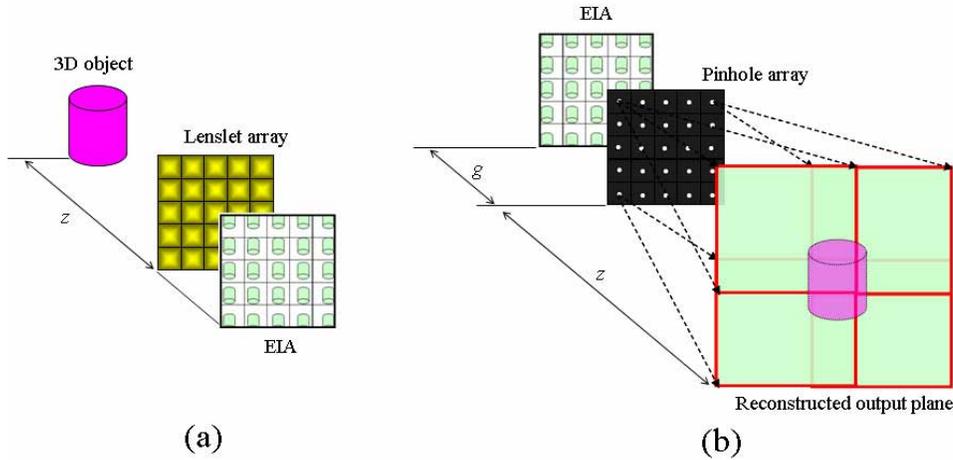


Fig. 1. Concept of integral imaging (a) Pickup (b) Computational reconstruction.

2. CII system for partial occluded object recognition

The principle of partially occluded 3D object recognition using CII is that it is possible to obtain the reconstruction of the 3D plane image of interest. The whole recognition system is composed of two steps as shown in Fig. 2 [16,17]. The first step is to extract a template image for 3D object to be recognized and the second step is to recognize target objects by using the template.

In the first step of CII system as depicted in Fig. 2(a), the recognized 3D object is picked up by a lenslet array and then recorded by a CCD camera. Here the recorded images are referred as the reference EIA in which each elemental image has particular perspective information about the 3D object. Next, the reference EIA is used to reconstruct 3D image digitally at the known distance where the 3D object was located. The reconstructed image of the reference EIA is called the template which is stored in a computer memory for the second step.

On the other hand, in the second step as shown in Fig. 2(b), target objects having occlusion and the 3D objects are recorded as the target EIA. Using VCR of Fig. 1(b), the target output image is reconstructed at the distance of the 3D object. Once an output image is obtained, the correlation process can be performed between the template and the output image. From the correlation results, 3D object recognition can be done.

When the output image is reconstructed in the second step, occlusion degrades the resolution of reconstructed images. To improve the recognition performance, we should reduce the image degradation effect by occlusion.

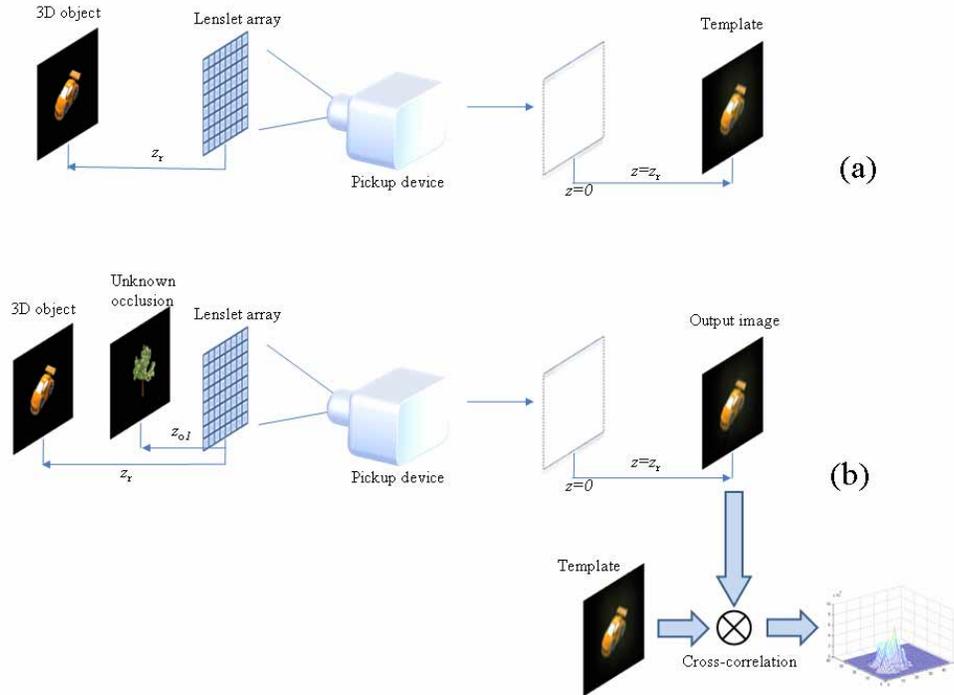


Fig. 2. Principle of CII (a) Generation of template (b) Recognition of partially occluded object.

3. Proposed occlusion removal method

In this paper, we present an occlusion removal method by enhancing the output image through the elimination of the unknown occlusion. The principle of the CII system using the occlusion removal method is shown in Fig. 3. If our system is compared with the conventional CII system, there are two additional processes to eliminate the unknown occlusion into EIA. The first process is to use a computational transform between the elemental image array and sub-image array. The second process is to remove the unknown occlusion in the sub-image array using disparity information by the sub-image block matching algorithm, which is well known in the stereo vision.

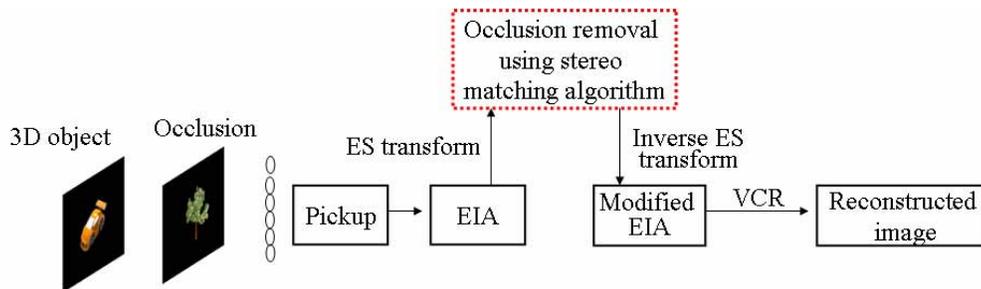


Fig. 3. Conceptual diagram for proposed occlusion removal method

3.1 Pickup process

As shown in Fig. 4, we assume that the occlusion object and 3D object of interest are located at two arbitrary distances z_o and z_r , respectively, from lenslet array. They are recorded as the EIA by using a CCD camera. The recorded EIA is shown in Fig. 5(a).

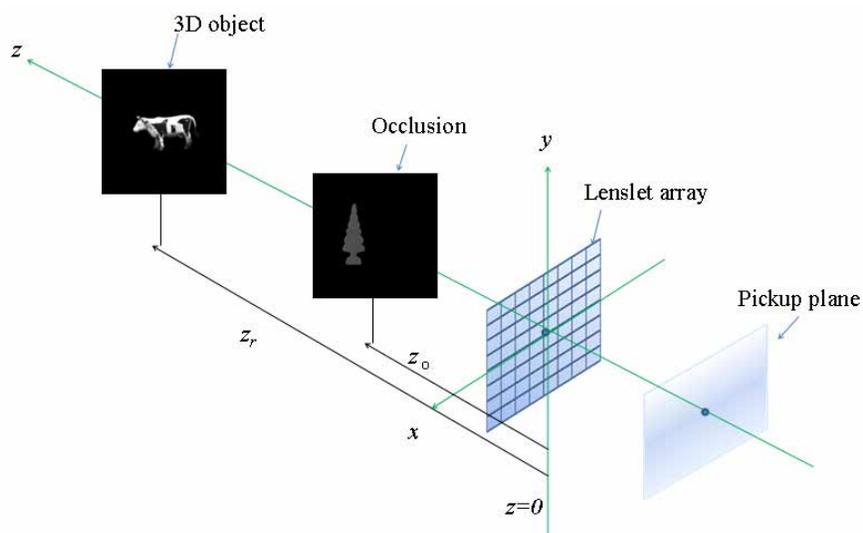


Fig. 4. Experimental setup for pickup process

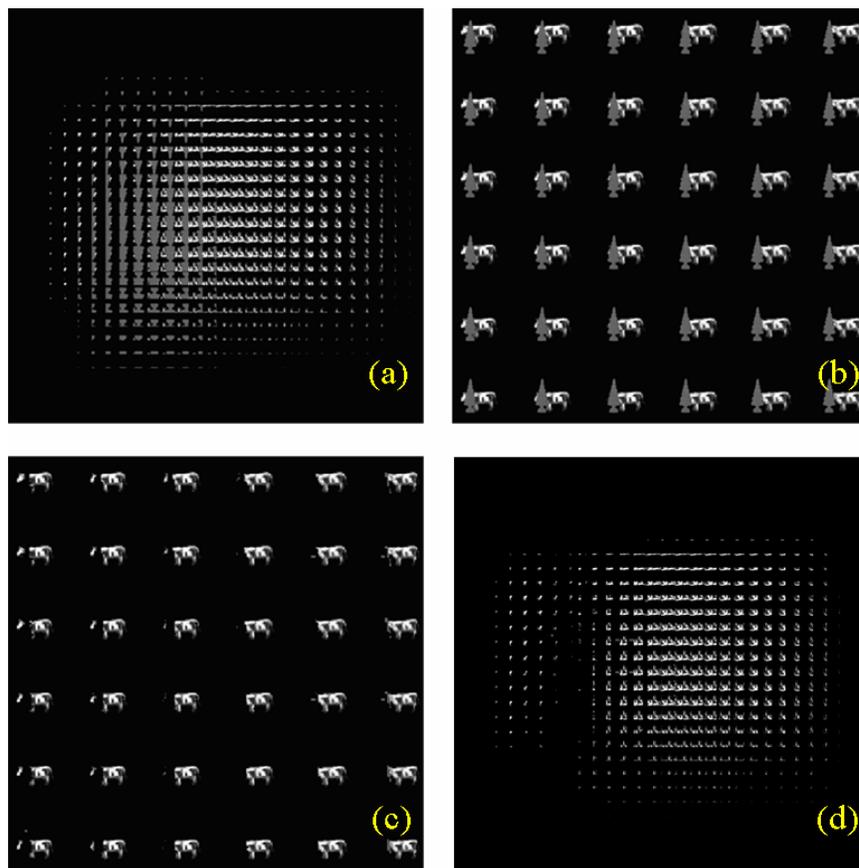


Fig. 5. (a) EIA (b) Sub-image array based on multiple pixel extraction (c) Sub-image array after occlusion removal (d) The modified EIA

3.2 Elemental-image to sub-image (ES) transform

First, we transform the recorded EIA into sub-image array for the proposed occlusion removal method. This elemental-image to sub-image transform is called ES transform in this paper. The ES transform is a kind of computational pixel recombination process [18,19]. That is, we extract the same position for all of the EIA and a collection of pixels of same position is obtained as sub-image array. This ES transform can be performed based on either single pixel extraction or multiple-pixel extraction. Figure 6 depicts the conceptual idea of ES transform based on multiple-pixel extraction. Suppose that s_x and s_y are the number of pixels for each elemental image and l_x and l_y are the number of elemental image in x and y axis, respectively. Then, the entire elemental images, which are denoted as E , become $(n_x=s_x l_x) \times (n_y=s_y l_y)$ pixels. If $(m \times m)$ pixels are collected, the m pixel-based sub-image array is calculated by

$$S(i, j) = E(t_x s_x + q_x m + r_x, t_y s_y + q_y m + r_y) \quad (1)$$

where $q_x = [i/(ml_x)]$, $q_y = [j/(ml_y)]$, $p_x = i \% (ml_x)$, $p_y = j \% (ml_y)$, $t_x = [p_x/m]$, $t_y = [p_y/m]$, $r_x = i \% m$ and $r_y = j \% m$. $[x]$ is gauss function denotes the largest integer less than or equal to the number x , and $a \% b$ is the remainder on division of a by b . From Eq. (1), we can generate a sub-image array based on arbitrary pixels from EIA. Then, the single sub-image has $(ml_x) \times (ml_y)$ pixels as shown in Fig. 5(b). As the pixel number increases, the resolution of sub-images might increase but be distorted at some degree.

To understand the effect of m pixels for ES transform, we consider a ray diagram for integral imaging system as shown in Fig. 7. Suppose that m pixels in the EIA plane are sampled at the output plane. The sampling interval of m pixels increases as the distance z increase as shown in Fig. 7. Large sampling interval ($z > z_{max}$) results in sampling crossing out of the corresponding lenslet. This may cause image distortion in sub-image array. To avoid this condition, we can calculate the maximum pixel number m_{max} for the given distance z where there is no distortion in sub-image array. The maximum pixel number can be obtained at the distance z_{max} where $n_x = M m_{max}$. It is given by

$$m_{max} = \left(m_{max,x} = \frac{g}{z} n_x = \frac{n_x}{M}, m_{max,y} = \frac{g}{z} n_y = \frac{n_y}{M} \right) \quad (2)$$

where $M = z/g$. It is seen that we generate the sub-image array with high-resolution by using this maximum pixel number.

From the experimental structure of Fig. 4, the maximum pixel number was calculated as $m_{max} = (5,5)$ when $n_x = n_y = 30$ and $M = 6$ based on the distance of occlusion. The transformed sub-image array is shown in Fig. 5(b). It is composed of 6×6 and each sub-image has 150×150 pixels which is much higher than those of the conventional sub-image transform.

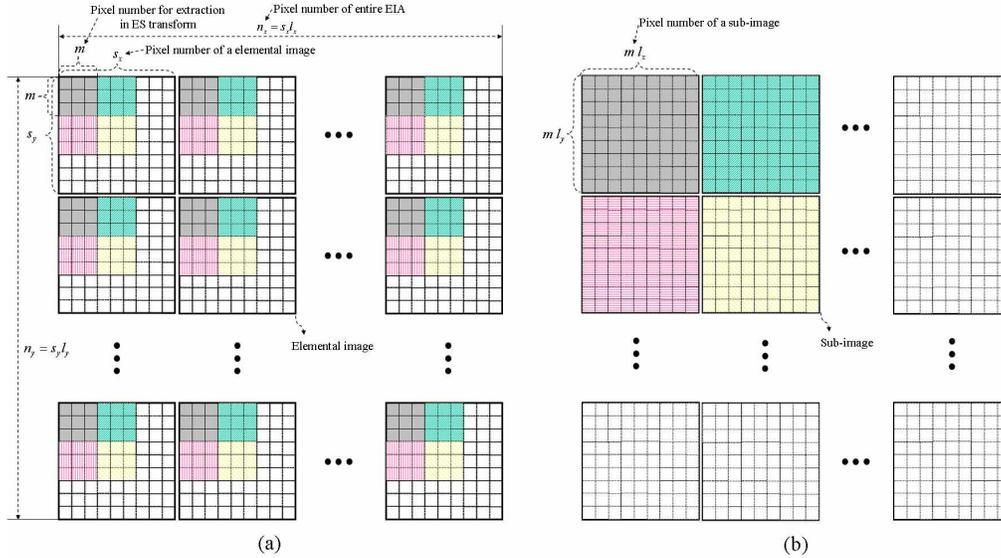


Fig. 6. ES transform based on multiple pixel extraction (a) EIA (b) Sub-image array.

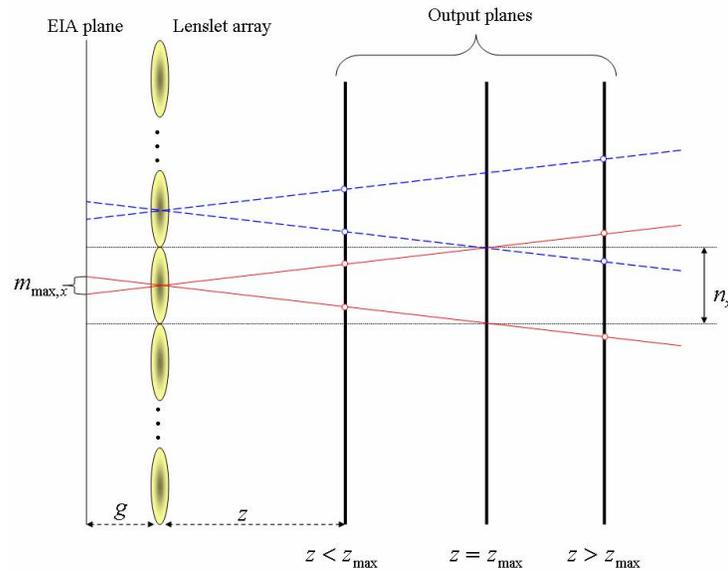


Fig. 7. Ray analysis for ES transform

3.3 Sub-image block matching scheme for sub-image array

Using the ES transform of Eq. (1), we can obtain the high-resolution sub-image array. The transformed sub-image array contains many different perspective images of 3D objects as shown in Fig. 5(b). Using this property, we may apply depth estimation algorithm to the sub-images in order to remove the occlusion. The entire process for occlusion removal is shown in Fig. 8. Among sub-images, we first select two adjacent images for the use of stereo pair. Then we apply stereo matching algorithm into two selective images for segmentation of occlusion. In this paper, we use the well-known block matching algorithm for sub-images [20,21]. This is called sub-image block matching algorithm. In general, block matching minimizes a

measure of matching error between two stereo images. The matching error between the blocks at the position (x,y) in the left image, I_L , and the candidate block at position $(x+u, y+v)$ in the reference image, I_R , is usually defined as the sum of absolute difference (SAD)

$$SAD_{(x,y)}(u,v) = \sum_{i=1}^B \sum_{j=1}^B |I_L(x+i, y+j) - I_R(x+u+i, y+v+j)| \cdot \quad (3)$$

where the block size is $B \times B$. Using SAD result, the best estimate (\hat{u}, \hat{v}) is defined to be the (u,v) which minimizes SAD. This estimate (\hat{u}, \hat{v}) calculates and compares the SADs for all the search positions $\{(x+u, y+v)\}$ in the right image I_R . That is,

$$(\hat{u}, \hat{v}) = \arg \min SAD_{(x,y)}(u,v) \cdot \quad (4)$$

After applying the sub-image block matching between two sub-images, we extract the depth map between them as shown in the right of Fig. 8. Here, we apply block matching algorithm to two adjacent sub-images in horizontal direction. Based on the extracted depth map, we can perform segmentation of occlusion and then remove it. This process is repeated for all the sub-images. As a result, we obtain the modified sub-image array with occlusion object eliminated as shown in Fig. 5(c).

To get the desired output image, the modified sub-image array is transformed back to the modified EIA of Fig. 5(d) by means of inverse sub-image transform. Here, it is seen that almost part of occlusion is removed in the sub-image array. Finally, the output image at the distance z_r is reconstructed by using the VCR method.

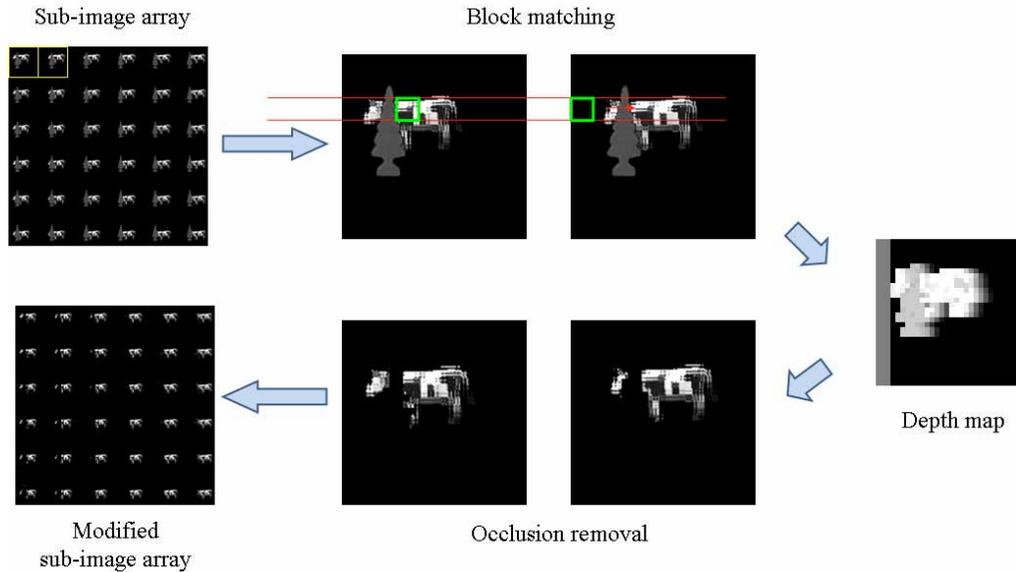


Fig. 8. Occlusion removal process in the sub-image array

4. Experiments and results

To show the usefulness of the proposed method, the computational experiments were carried out. The experimental structure is shown in Fig. 4. The distance between the lenslet array and the pickup plane was assumed to be 3 mm. The EIA of the test object with 900 by 900 pixels was computationally synthesized, in which the virtual lenslet array is assumed to have 30 by 30 lenslets. The size of each elemental image was 30 by 30 pixels. In the experiments, we

used two plane objects: 'tree' and 'cow'. The EIA was recorded when occlusion 'tree' and 3D object 'car' were located at two distances, $z_o=18$ mm and $z_r=33$ mm, from the lenslet array, respectively.

The recorded EIA was transformed into sub-image array using Eq. (1) and $m_{max}=(5,5)$. The extraction process of depth map for all sub-images was performed repeatedly using the sub-image block matching algorithm. Based on the extracted depth maps, we removed occlusion in each sub-image. We can see that the occlusion images into EIA were partially eliminated compared with the conventional method as shown in Fig. 8. After this process was completed, we obtained the modified EIA. To compare the visual quality of the reconstructed images, the output images at 18 mm and 33 mm were reconstructed by applying both the original EIA and modified EIA to the VCR method. The reconstructed images are shown in Fig. 9 of which (a) and (b) show the reconstructed images using the original EIA and our modified EIA, respectively. Here we can see that the 'tree' occlusion is largely removed in the proposed method. This shows that the proposed method removes the occlusion effectively. In addition, the image reconstructed from the proposed method is more clearly seen. This may improve the resolution of reconstructed image. Under the same conditions, we carried out the additional experiments for two different test images as shown in Fig. 10.

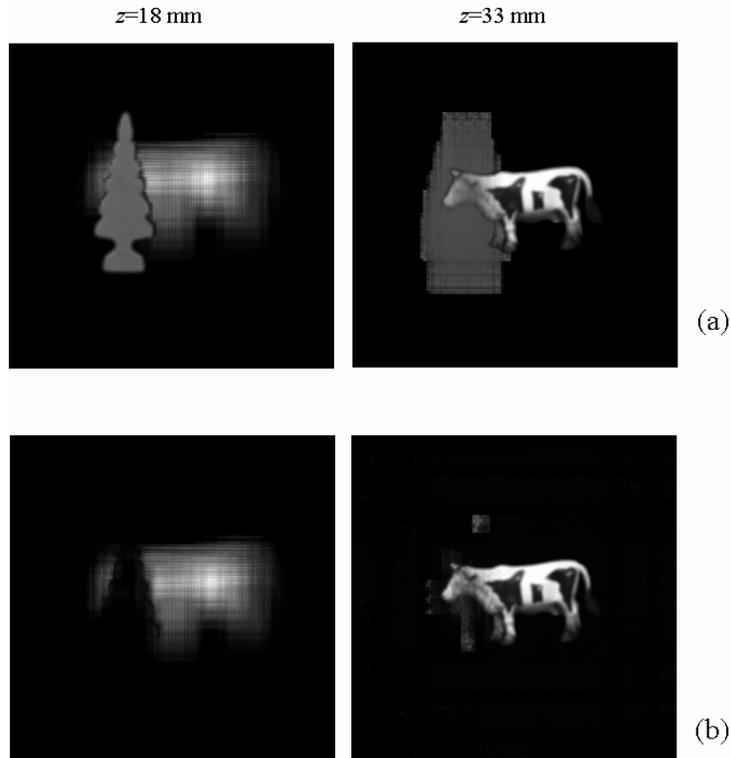


Fig. 9. Reconstructed images by using VCR process (a) Conventional method (b) Proposed method.

To evaluate the improvement numerically, we use the peak signal-to-noise ratio (PSNR) as an evaluation measure, which is well-know measure parameter in 2D image processing such as nose removal or compression. Therefore, PSNR is a good evaluation parameter since occlusion affects objects as a noise source in partially occluded object reconstruction. The PSNR is defined as

$$PSNR(O, R) = 10 \log_{10} \left(\frac{255^2}{MSE(O, R)} \right) \quad (5)$$

where O is an original image and R is the reconstructed image from VCR process using the normalization step [12]. Mean squared error (MSE) is given by

$$MSE = \frac{1}{PQ} \sum_{x=0}^{P-1} \sum_{y=0}^{Q-1} [O(x, y) - R(x, y)]^2 \quad (6)$$

where x and y are the pixel coordinates of images with the size of $P \times Q$ pixels.

For the original image and the reconstructed images as shown in Fig. 9, the PSNR was calculated. The PSNR results for three test image sets are shown in Table 1. From results of Table 1, it is revealed that the visual quality of images reconstructed from our method is better than that from the conventional method. We obtained the improvement of average 7.38 dB in PSNR from this experiment.

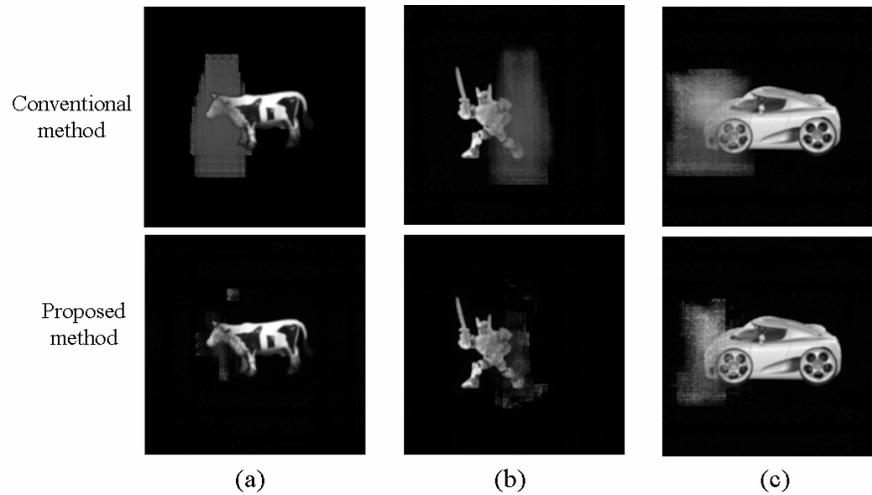


Fig. 10. Reconstructed images for three kinds of test images (a) Cow (b) Toy (c) Car

Table 1. PSNR results for three test images

Test images	Conventional method	Proposed method	Improved PSNR
Cow	20.94 dB	29.61 dB	8.67 dB
Toy	22.31 dB	29.81 dB	7.5 dB
Car	18.03 dB	24.01 dB	5.98 dB

5. Discussion and conclusion

Even though the proposed method has been successfully demonstrated, we consider that occlusion and object to be recognized were separated sufficiently in order to capture a partially occluded object. The captured information of partial occluded object affects the

performance of our block matching algorithm to extract a depth between sub-images. To improve the performance, we may avoid the limited distance range between occlusion and object [22] when a practical application system is implemented or may consider the use of curved pickup scheme to increase the distance [23]. Another issue to be considered is that the performance of the block matching algorithm might be somewhat dependent on the block size used, so that some detailed analysis with various block size is needed. Moreover, the right algorithm mostly adequate for block matching between sub-images should be also searched. These are interesting future issues to be addressed.

In conclusion, we have proposed the occlusion removal method for improved reconstruction of 3D objects that are occluded partially in CII. We employed the multiple pixel-based ES transform to get high resolution sub-images, followed by 3D depth map generation by stereo matching algorithm between sub-images to segment and eliminate occlusion object in EIA. As a result, this produces a modified EIA with occlusion eliminated. This can provide a substantial gain in terms of the image quality of 3D reconstructed images. To show the usefulness of the proposed technique, we represented some experiments and demonstrated the improvement of 3D reconstructed images. We expect that the proposed technique will aid to improve the performance of 3D recognition systems using CII.

Acknowledgments

This research was financially supported by the Ministry of Education, Science Technology (MEST) and Korea Industrial Technology Foundation (KOTEF) through the Human Resource Training Project for Regional Innovation.