Depth extraction of three-dimensional objects using block matching for slice images in synthetic aperture integral imaging

Joon-Jae Lee,[1] Byung-Gook Lee,[2] and Hoon Yoo[2,*]

[1]Department of Game Mobile Contents, Keimyung University, Daemyung3-Dong Nam-Gu, Daegu 705-701, South Korea

[2]Department of Visual Content, Dongseo University, San69-1, Jurye2-Dong, Sasang-Gu, Busan 617-716, South Korea

[3]Sangmyung University, 7 Hongji-dong, Jongno-gu Seoul 110-743, South Korea

*Corresponding author: hunie@smu.ac.kr

Received 14 April 2011; revised 23 August 2011; accepted 24 August 2011;
posted 29 August 2011 (Doc. ID 145951); published 3 October 2011

We describe a computational method for depth extraction of three-dimensional (3D) objects using block matching for slice images in synthetic aperture integral imaging (SAII). SAII is capable of providing high-resolution 3D slice images for 3D objects because the picked-up elemental images are high-resolution ones. In the proposed method, the high-resolution elemental images are recorded by moving a camera; a computational reconstruction algorithm based on ray backprojection generates a set of 3D slice images from the recorded elemental images. To extract depth information of the 3D objects, we propose a new block-matching algorithm between a reference elemental image and a set of 3D slice images. The property of the slices images is that the focused areas are the right location for an object, whereas the blurred areas are considered to be empty space; thus, this can extract robust and accurate depth information of the 3D objects. To demonstrate our method, we carry out the preliminary experiments of 3D objects; the results indicate that our method is superior to a conventional method in terms of depth-map quality.

OCIS codes: 100.6890, 110.6880, 110.4190.

1. Introduction

Recently, there has been great interest in three-dimensional (3D) imaging and display systems [1–7]. Among them, integral imaging can be considered as a promising solution to implement a next-generation 3D imaging and display system, because it provides a simple system structure, color images, and full parallax [5–8]. To implement a 3D imaging and display, integral imaging systems require different perspectives from 3D objects. These recorded perspective images are called elemental images, which is an important play in integral imaging.

Original integral imaging picks up perspective views using a lenslet array and a single camera [6,7]. This system has the advantages of a simple structure and rapid recording of 3D objects, because the elemental images are recorded through the lenslet array at once. However, the recorded elemental images can suffer from low-resolution quality in perspective images because a camera should cover the whole lenslet array. Also, the small aperture of each small lens in the lenslet array possibly has some distortions, such as aberrations. Thus, there are some modified versions of integral imaging. One method is using multiple elemental image set to improve the image quality of the reconstructed image [9]. Another modified method is the synthetic aperture integral imaging (SAII), where a camera is translated or a camera array is employed to obtain multiple perspective images [9–11]. Although this method needs a mechanical translation of a camera, high-resolution elemental images are obtained during

0003-6935/11/295624-06$15.00/0

the camera movement. This advantage enables the method to be applied to various applications, such as 3D object recognition [11,12].

Depth extraction of 3D objects is known to be one of the important issues in the research fields of 3D vision, object tracking, and video surveillance. Some studies for integral-imaging-based depth extraction methods have been proposed [13–18]. However, they had used low-resolution elemental images recorded from a lenslet array or a complicated algorithm in order to extract a depth map for objects.

In this paper, we propose a new computational method for depth extraction of 3D objects using SAII techniques. We first record high-resolution elemental images by moving a camera. Then, the recorded elemental images are used to generate a set of 3D slice images using a fast computational reconstruction algorithm [19]. To extract the depth of the 3D objects, we utilize a new block-matching algorithm between a reference elemental image and a set of 3D slice images. Our method is based on the fact that the focused areas in the slice images are the right location for an object, whereas the blurred areas are considered to be empty space; thus, this can extract robust and accurate depth information of the 3D objects. To demonstrate our method, we carry out preliminary experiments of the 3D objects, and the results are presented.

Section 2 gives a brief explanation for the concept of SAII. Section 3 describes the proposed depth extraction method. Section 4 presents experimental results on real 3D objects. Finally, Section 5 provides the discussion and conclusion.

2. Synthetic Aperture Integral Imaging

Figure 1 shows the principle of SAII. Basically, the SAII system consists of two parts: pickup and digital reconstruction. In the pickup part, as shown in Fig. 1(a), the intensity and direction of the rays coming from 3D objects are recorded in a certain fixed position of the camera. The recorded image is called the elemental image of SAII. By moving the camera with a two-dimensional grid structure, we can record different perspectives (elemental images) of the 3D object according to the location of the camera. In the digital reconstruction of SAII, as shown in Fig. 1(b), each elemental image is first projected inversely through the corresponding pinhole modeled from the original position of the pickup camera. Next, when an image is reconstructed on the output plane of z from the pinhole, the projected elemental image is magnified by a factor of z/g, where z is the distance between the virtual pinhole and the output plane and g is the distance between the pinhole and the elemental images, which is normally considered as the focal length f of a moving camera. Finally, the magnified elemental images are overlapped and accumulated at the corresponding pixels of the output plane. In addition, the iterative process mentioned above gives us a set of slice images for 3D objects.

3. Proposed Method

The proposed depth extraction method is shown in Fig. 2. It is divided into three subparts: SAII pickup, digital reconstruction, and depth extraction. In the SAII pickup part, 3D objects are recorded as elemental images by moving a camera in the x and y directions with a uniform moving distance (p). The recorded elemental images contain intensity and directional information of the 3D objects so that they have different perspectives according to the location of a group of image sensors for an elemental image. SAII can provide the elemental images with a large number of pixels. That is, because SAII captures each elemental image by moving the image sensor with $N_x \times N_y$ pixels at corresponding positions, the resolution of each elemental image is equal to that of the sensor.

Next, a computational reconstruction algorithm for digital reconstruction converts the recorded elemental images into a series of 3D slice images by adjusting the z value. In this paper, we use a fast reconstruction algorithm to efficiently reconstruct the 3D slice images [19]. The size of the

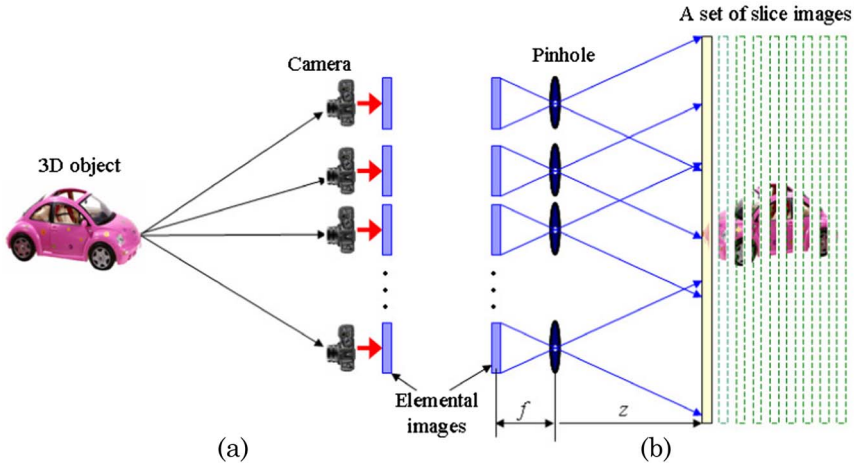

Fig. 1. (Color online) Principle of SAII.

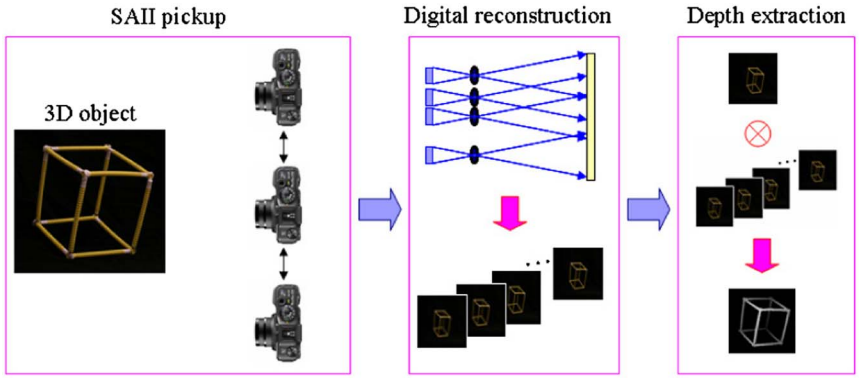

Fig. 2. (Color online) Proposed depth extraction method using slice images.

reconstructed slice images can be the same as that of elemental images. The principle of digital reconstruction is shown in Fig. 3. Hereafter, we assume that the number of pixels for projected elemental images is the same as the number of pixels for each elemental image. To have 3D slice images, the computational reconstruction algorithm generates a slice image independently by superimposing the pixels from all elemental image, as shown in Fig. 3(b). Let the total number of elemental images be $K \times L$. A slice image $I(x,y,z_0)$ located at distance z_0 is evaluated by the summation of I_{kl}, which is the intensity of the kth column and the lth row elemental image. This yields

$$I(x,y,z_0) = \frac{1}{O(x,y)} \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} I_{kl}(x - kS, y - lS), \quad (1)$$

where x and y denote the index of pixels in the x and y directions and $O(x,y)$ is the counted number of overlapped pixels at (x,y). The shifting parameter S is the number of shifted pixels in overlapping elemental images. This is given by

$$S = \frac{p}{\delta \times M}, \quad (2)$$

where δ is the size of each image sensor for a pixel and p is the distance between pickup locations in a moving camera. The parameter S can be adjusted by varying the magnification factor $M = z/g$. Using Eqs. (1) and (2), the reconstructed slice image can be obtained at distance z_0. To obtain a set of slice images, the reconstructed process can be repeated for different z values. Also, the parameters S_x and S_y mean S in the x axis and in the y axis, respectively. Putting $S = S_x = S_y$ is likely to be used in our method.

The accuracy of the extracted depth information is dependent on the step size of the z value. We can calculate the minimum step size in the z axis using the shifting parameter S of Eq. (2). In other words, because the shifting parameter should be an integer number for superimposing elemental images pixel by pixel, the minimum step between shifting distances becomes one pixel. Thus, the minimum step size Δz is obtained from two adjacent shifting parameters (S and $S+1$), and it becomes $\Delta z = pf/\delta S(S+1)$, where f is the focal length of the moving camera. Depth extraction in our method can be effective in case of using more than the minimum step size of the z value.

In the depth extraction part, depth information is determined by block matching between a reference elemental image and the reconstructed 3D slice

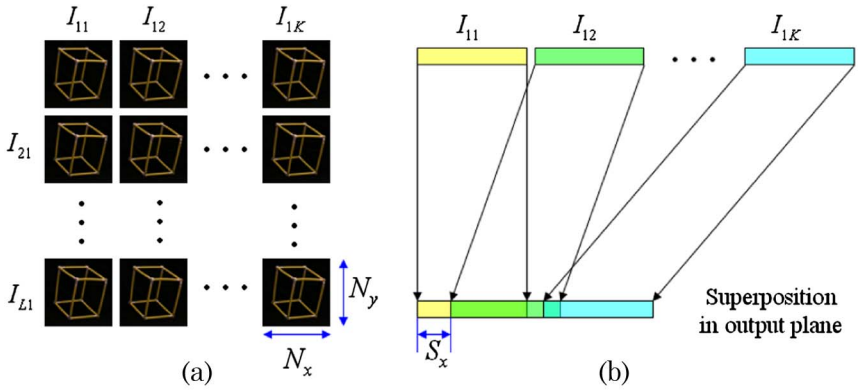

Fig. 3. (Color online) (a) Elemental images (b) principle of computational reconstruction.

images according to various distances. In fact, the reconstructed slice images consist of different content images with a focused or blurred area, which is dependent on the reconstruction distances. The focused images of a 3D object are reconstructed only at the original position of the 3D object. On the other hand, blurred images are shown out of the original position. Our depth extraction is based on the fact that the focused area in the reconstructed slice images can be detected by a block-matching algorithm using a reference image. To do so, we select one elemental image among the recorded elemental images as a reference image for the matching process. Any reference image chosen from elemental images can be used for matching the slice images, as shown in Fig. 4, because the pixel size of a reference elemental image is identical with that of each slice image.

Once slice images are generated by superimposing all recorded elemental images, most elemental images except the reference elemental image we use in the matching process for the depth extraction are discarded. In addition, the center elemental image is a good candidate for the reference elemental image, although another elemental image possibly works as the reference elemental image. This is demonstrated in Figs. 7(b) and 7(c) Here, the field of view of the reference elemental image is very important to extract the entire surface of the 3D scene. Thus, in this paper, we assume that the field of view of each elemental image that can be the reference image is wide enough to include the whole 3D scene of interest.

Our new block-matching algorithm is shown in Fig. 4. First, we select (k_x, k_y) block images in both the reference elemental image and the slice images. The block images have the same position, as shown in Fig. 4. The block-matching algorithm is then applied with two selected block images. Here, block matching minimizes a measure of matching error between the block image of the reference elemental image (B_{ref}) and a set of block images of slice images (B_z). As a metric for block matching, our method employs the sum of absolute differences (SAD):

$$\text{SAD}(x,y,z) = \sum_{i=1}^{b}\sum_{j=1}^{b} |B_{\text{ref}}(x+i, y+j) - B_z(x+i, y+j)|, \tag{3}$$

where the block size is $b \times b$.

The depth for each point (x,y) can be extracted by selecting the z value from the argument of minimum among the $\text{SAD}(x,y,z)$ results. This can be mathematically formulated as

$$\hat{z}(x,y) = \text{argmin}\text{SAD}(x,y,z). \tag{4}$$

Finally, a depth map can be obtained by calculating the depth for all points in the reference image.

4. Experiments and Results

To demonstrate our proposed method, we performed preliminary experiments for a cubic object shown in Fig. 5 and character objects shown in Fig. 8(a). For the cubic object, the nearest part of the object from the camera was located at approximately 1000 mm. For SAII pickup, a camera with 2184×1456 pixels, an $8.2\,\mu$m pitch, and a lens with a focal length of 50 mm was used. The camera was moved in 10 mm increments vertically and horizontally. We recorded 35 (7×5) elemental images. Among them, some examples are shown in Fig. 5. In our method, all elemental images should be clearly recorded to extract the exact depth information. Thus, we minimized the aperture of the imaging lens in our experiment to maximize the depth of focus. We set the minimum F number ($F\# = 32$) in the imaging lens, and this allows a depth of focus of 0.84 mm when the object is located at 1 m away from the pickup camera [20]. This means we can record many objects within a large depth of focus without a blurring effect.

After elemental images are prepared, we reconstructed a series of 3D slice images from 35 elemental images using the computational reconstruction algorithm of Eq. (1). The reconstruction process we used generates 31 slice images from the elemental

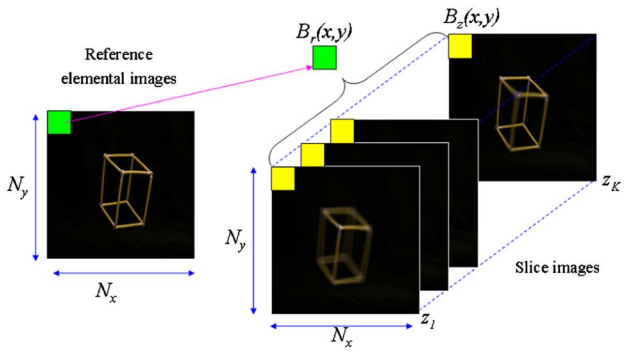

Fig. 4. (Color online) Block-matching algorithm between reference elemental image and slice images.

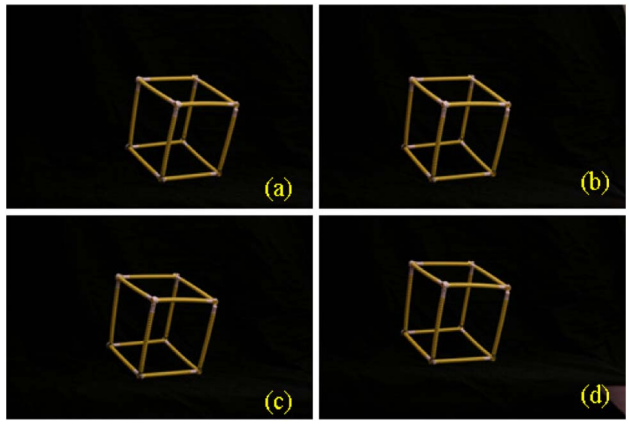

Fig. 5. (Color online) Four examples of recorded elemental images: (a) (1, 1)th elemental image, (b) (7, 1)th elemental image, (c) (4, 3)th elemental image, and (d) (1, 5)th elemental image.

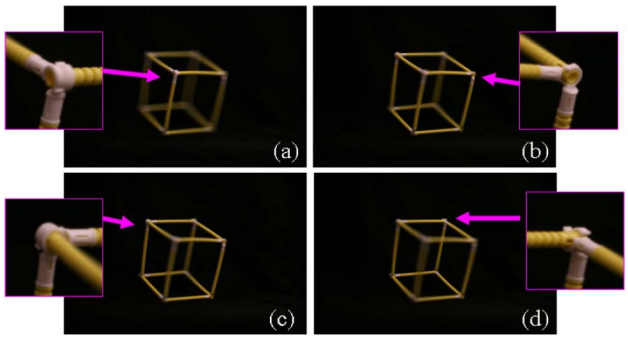

Fig. 6. (Color online) Reconstructed slice images at (a) 980, (b) 1050, (c) 1110, and (d) 1190 mm.

images in the range from 900 to 1200 mm. For examples, 4 slice images reconstructed at different distances $z = 980$, 1060, 1110, and 1190 mm are shown in Fig. 6. It can be seen that each elemental image shows both the focused areas and blurred areas. Note that the focused areas represent the right positions of the object parts.

Next, we estimated the depth of the object parts using block matching between the reference elemental image and a set of slice images. Among 35 elemental images, one elemental image was used as the reference elemental image for block matching. Also, another elemental image was selected as the reference image. For each case, the block-matching algorithm with the SAD metric was applied, as described in Eqs. (3) and (4). The block size is set to be 5×5 pixels. For comparison, a conventional method [21] is conducted. The estimated depth maps are shown in Fig. 7. It is seen that the proposed method provides a clearer depth map and more accurate depth information than the conventional method. Especially, the conventional method is very poor at the background region whose characteristic is a lack of information to detect depth. Note that the conventional stereo matching algorithm for two clear images among many perspective images largely depends on image characteristics. The depth-map image from our method agrees with the cubic structured object well. This result reveals that the proposed method can extract the 3D information of object effectively. Figures 7(b) and 7(c) indicate that choosing a reference elemental image among elemental images is not a serious problem, but the center

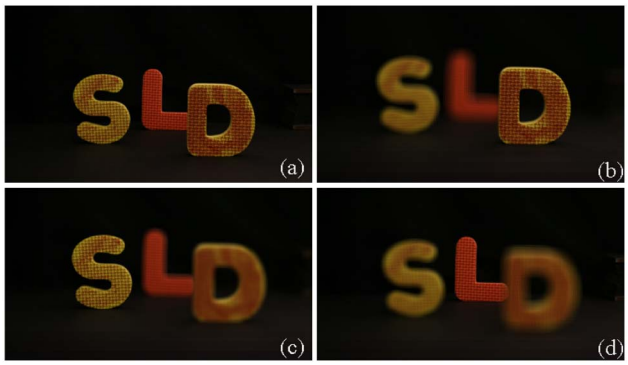

Fig. 8. (Color online) (a) Character objects and reconstructed slice images at (b) 970 (c) 1070, and (d) 1200 mm.

Fig. 9. Visual comparison for extracted depth maps of character objects extracted by the (a) conventional method and (b) proposed method.

elemental image shows slightly better performance in depth extraction than another elemental image.

In addition, we performed another experiment for multiple object case. For the experimental structure, three character objects, as shown in Fig. 8(a), are located at 970, 1070, and 1200 mm from the camera, respectively. The experimental parameters are the same as those of the above experiment with the cubic object. Some reconstructed slice images for each character object are shown in Figs. 8(b)–8(d). Each slice image includes a clear area of the corresponding object at the original distance. Also, Fig. 9 indicates that the extracted depth map from our method is superior to that from the conventional method.

5. Discussion and Conclusion

In the conventional depth extraction methods using multiple cameras, there have been some studies on cost functions for depth extraction [22]. Some cost functions, such as SAD, sum of SAD (SSAD), sum

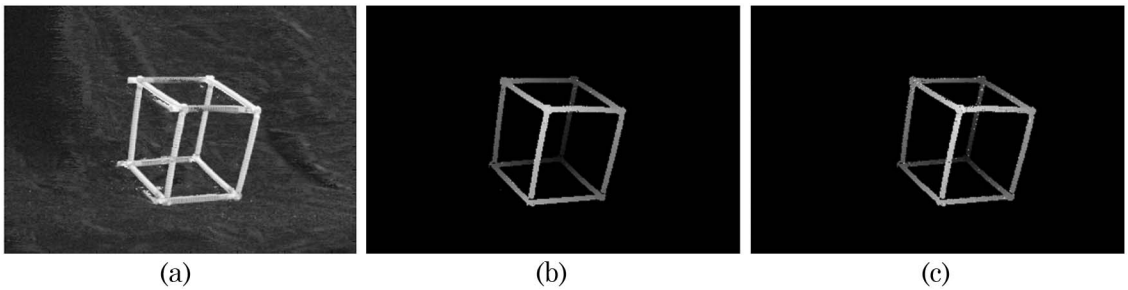

Fig. 7. Visual comparison for extracted depth maps of a cubic object extracted by the (a) conventional method, (b) proposed method with the center elemental image, and (c) another elemental image as the reference image.

of squared differences (SSD), and correlations, are well known in the literature [22]. SAD is commonly used in depth extraction from two perspective images, whereas SSAD is possibly used in depth extraction from more than two perspective images. Also, SSD and correlations can be applied to the systems using two or more images. We tried the known block-matching metrics in our method; we knew that the performances of depth extraction are almost the same in our method. Thus, we chose SAD as a cost function to determine the depth for its simplicity.

Note that our method superimposes all perspective images, resulting in depth slices. The matching process using SAD between a reference image and depth slices can be considered as a matching process using SSAD among perspective images due to the linearity property. However, if some complicated methods providing depth slices [23,24] are applied to our method, the SAD metric is not the same with the SSAD metric anymore. In addition, our method is more accurate and efficient to search the correspondence. This is because superimposing perspective images cancels some background noises and very high frequency terms, which possibly introduces a problem in detecting depth information. Thus, if we want to search correspondence between two images, our depth slices provides better matching results because of using only focused image areas.

In conclusion, we have presented a depth extraction method using new block matching for slice images in SAII. Because SAII provides high-resolution elemental images, high-resolution 3D slice images for objects are generated. To extract the depth of the 3D objects, we have proposed a new block-matching algorithm between a reference elemental image and a set of the 3D slice images. The property of the reconstructed slices images that have blurred areas and focused areas enables us to extract robust and accurate depth information of 3D objects. To demonstrate our method, we carried out the preliminary experiments of the 3D objects, and the results indicate that our method is superior to the conventional method in terms of depth-map quality, and it shows the usefulness of the proposed method.

References

1. T. Okoshi, *Three-Dimensional Imaging Techniques* (Academic, 1976).
2. S. A. Benton, *Selected Papers on Three-Dimensional Displays* (SPIE Press, 2001).
3. B. Javidi and F. Okano, *Three-Dimensional Television, Video, and Display Technologies* (Springer, 2002).
4. Y. Kim, K. Hong, and B. Lee, "Recent researches based on integral imaging display method," 3D Res. **1**(1), 17–27 (2010).
5. M. G. Lippmann, "Épreuves réversibles donnant la sensation du relief," J. Phys. **7**, 821–825 (1908).
6. J.-S. Jang and B. Javidi, "Improved viewing resolution of three-dimensional integral imaging with nonstationary micro-optics," Opt. Lett. **27**, 324–326 (2002).
7. H. Yoo and D.-H. Shin, "Improved analysis on the signal property of computational integral imaging system," Opt. Express **15**, 14107–14114 (2007).
8. D.-H. Shin, B. 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).
9. A. Stern and B. Javidi, "Three-dimensional image sensing, visualization, and processing using integral imaging," Proc. IEEE **94**, 591–607 (2006).
10. Y. S. Hwang, S.-H. Hong, and B. Javidi, "Free view 3-D visualization of occluded objects by using computational synthetic aperture integral imaging," J. Display Technol. **3**, 64–70 (2007).
11. M. Cho and B. Javidi, "Three-dimensional visualization of objects in turbid water using integral imaging," J. Display Technol. **6**, 544–547 (2010).
12. I. Moon and B. Javidi, "Three-dimensional recognition of photon-starved events using computational integral imaging and statistical sampling," Opt. Lett. **34**, 731–733 (2009).
13. J.-H. Park, S. Jung, H. Choi, Y. Kim, and B. Lee, "Depth extraction by use of a rectangular lens array and one-dimensional elemental image modification," Appl. Opt. **43**, 4882–4895 (2004).
14. J.-H. Park, Y. Kim, J. Kim, S.-W. Min, and B. Lee, "Three-dimensional display scheme based on integral imaging with three-dimensional information processing," Opt. Express **12**, 6020–6032 (2004).
15. G. Passalis, N. Sgouros, S. Athineos, and T. Theohari, "Enhanced reconstruction of three-dimensional shape and texture from integral photography images," Appl. Opt. **46**, 5311–5320 (2007).
16. 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," Opt. Express **16**, 16294–16304 (2008).
17. B.-G. Lee, H.-H. Kang, and E.-S. Kim, "Occlusion removal method of partially occluded object using variance in computational integral imaging," 3D Res. **1**(2), 6–10 (2010).
18. J.-H. Jung, K. Hong, G. Park, I. Chung, J.-H. Park, and B. Lee, "Reconstruction of three-dimensional occluded object using optical flow and triangular mesh reconstruction in integral imaging," Opt. Express **18**, 26373–26387 (2010).
19. H. Yoo and D.-H. Shin, "Fast computational integral imaging reconstruction method using a fractional delay filter," Jpn. J. Appl. Phys. **49**, 022503 (2010).
20. A. R. Greenleaf, *Photographic Optics* (MacMillan, 1950), pp. 25–27.
21. B. B. Alagoz, "Obtaining depth maps from color images by region based stereo matching algorithms," OncuBilim Algorithm Systems Labs **8**, 4 (2008) arXiv:0812.1340v2.
22. M. Z. Brown, D. Burschka, and G. D. Hager, "Advances in computational stereo," IEEE Trans. Pattern Anal. Mach. Intell. **25**, 993–1008 (2003).
23. H. Yoo, "Artifact analysis and image enhancement in three-dimensional computational integral imaging using smooth windowing technique," Opt. Lett. **36**, 2107–2109 (2011).
24. 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," Appl. Opt. **50**, 1889–1893 (2011).