Journal of the Optical Society of Korea, Vol. 8, No. 2, June 2004, pp. 72-78

An Improved Stereovision Scheme Using Single Camera and a Composite Lens Array

Heejin Choi, Jae-Hyeung Park, Jisoo Hong, and Byoungho Lee*

National Research Laboratory of Holography Technologies, School of Electrical Engineering, Seoul National University, Seoul 151-744, KOREA

(Received May 18, 2004)

We propose an improved stereovision scheme using a composite lens array and single camera. The use of a composite lens array makes it possible to construct a compact stereovision system and to improve the quality of the results by adopting image processing method. The principle of the proposed system is explained and the simulation and the experimental results are also presented.

OCIS codes : 110.2990, 220.2740

I. INTRODUCTION

The stereovision scheme is a depth-extraction method from two or more images of the original object [1-3]. Figure 1 shows the basic concept of the stereovision scheme. In Fig. 1, two sets of cameras are set at different positions and take different images of the original objects. The different images have different perspectives of the original objects, and hence contain three-dimensional (3D) information. There are many well-known algorithms and it is possible to extract the 3D information from the images after they have been acquired. As shown in Fig. 1, however, the conventional stereovision system uses two or more cameras and therefore requires a bulky structure. Moreover, in the conventional stereovision system there also exist some difficulties such as the need of precise calibration between different cameras and sensitivity to external interferences such as vibration.

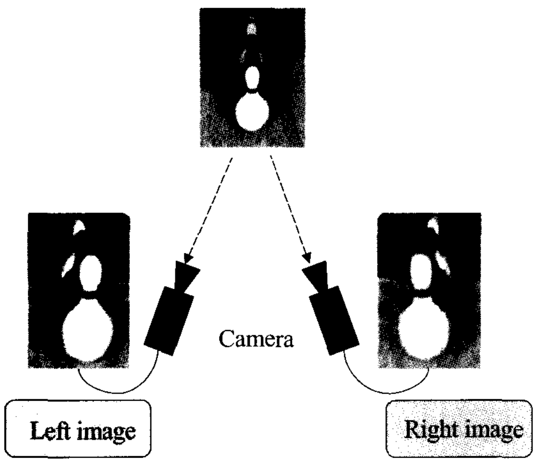


FIG. 1. The concept of the conventional stereovision scheme.

There has been proposed a method using single camera and a flat lens array to improve the conventional scheme [4,5] as shown in Fig. 2. The key idea is similar to the pickup process in integral imaging [6-9]. A flat lens array used in the previous method is composed of numbers of elemental lenses and can form multiple images all at once. Since the relative position of each elemental lens is different from the others, the acquired image, which is called the elemental image, has different perspective for each lens. Therefore, the flat lens array in the previously proposed method can replace the role of multiple cameras and reduce the complexity and problems of the conventional system remarkably.

In this paper, we propose a method to improve the

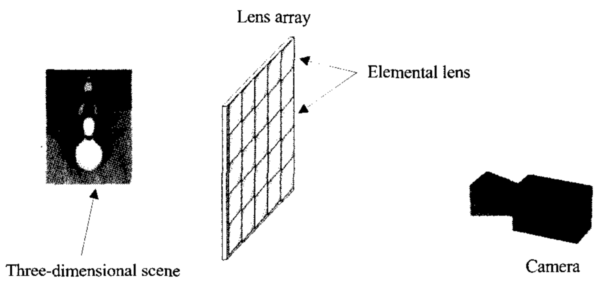


FIG. 2. The concept of the previously proposed stereovision scheme.

performance of the previously proposed method. The proposed method uses a composite lens array instead of the conventional flat lens array. In the conventional flat lens array, all the elemental lenses have same size and same focal length. However, the composite lens array is composed of two sets of elemental lenses that have different focal lengths and can form two sets of elemental images through different focal lengths [10]. Therefore, the use of a composite lens array increases the variations of information of the original object that we can acquire all at once. It is possible to use additional image processing procedures and the performance of the proposed system can be improved compared with that of the previously proposed method.

II. PRINCIPLES OF THE STEREOVISION

Before explaining the principle of the proposed method, it is necessary to introduce the basic principles of stereovision. Figure 3 shows the basic principle of the depth extraction algorithm. In Fig. 3, a point P is located with distance Z from the two lenses L_1 and L_2. The two lenses are separated with center-to-center distance B. Then, the image of P by each lens (L_1 and L_2) is located at the image plane which is far from the lenses with distance of focal length f of the two lenses. Let us define the location of each image as the distances from the center of the each lens and denote them as l_1 and l_2 respectively.

If the point P is located at x in transverse direction, l_1 and l_2 are given by

$$l_1 = \frac{f}{Z}x, \tag{1}$$

$$l_2 = \frac{f}{Z}(B-x), \tag{2}$$

and the disparity d, that is the position difference between two images from the centers of elemental

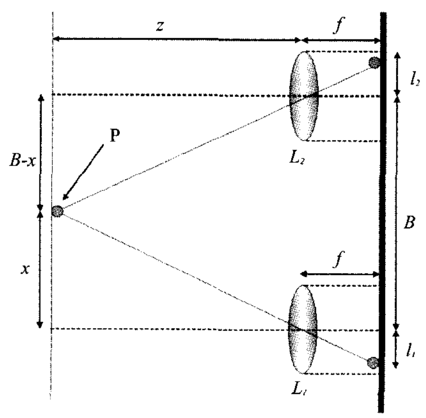


FIG. 3. Basic principles of stereovision.

lenses, can be written by

$$d = l_1 + l_2 = \frac{f}{Z}(x+B-x) = \frac{f}{Z}B. \tag{3}$$

If we modify Eq. (3) for the depth Z, it can be expressed as follows.

$$Z = \frac{f}{d}B \tag{4}$$

As shown in Eq. (4), the depth Z of arbitrary point P can be expressed as a function of disparity d, focal length of the lenses f, and the center-to-center distance B between the two lenses. Since f and B can be acquired from the parameters of the system setup, the depth of object P can be acquired from the disparity d between the images from it.

III. PARAMETERS OF STEREOVISION SCHEME USING THE LENS ARRAY

Among the various parameters of the stereovision system using the lens array, there are two important ones that are related to the focal length of the elemental lens. One is the resolution of the image and the other is the field of view. The resolution of the elemental image plays an important role in extracting the depth of the object precisely. Especially in the stereovision scheme using the lens array, the resolution of the elemental image is still more important, since the total pixels of the CCD should be divided among the elemental lenses and thus each elemental image has low resolution. Equation (1) indicates that, for a given object depth, the image magnification by the elemental lens is proportional to the focal length of the elemental lens. Hence the resolution of each elemental image increases as the focal length of the elemental lens increases. The other factor is the field of view of the elemental image. The field of view is rarely a problem in the conventional stereovision systems adopting multiple cameras. On the other hand, in the stereovision system using the lens array, all the elemental images are captured by one camera so that the field of view of each elemental image is severely restricted. The problem of the narrow field of view arises mainly when the field of view is too small to include the entire extent of the object image. In this case, we should perform disparity searching with reference to several elemental images respectively and combine the obtained disparity map to acquire the entire depth map of the object, which generally brings about large error. In the lens array method, the field of view of each elemental image depends on both the pitch and the focal length of the elemental lens. Provided the elemental lens pitch is fixed, the field of view is inversely proportional to the focal length of the elemental lens.

As we discussed above, two important parameters of stereovision have an opposite relationship with the focal length of the elemental lens. It is ideal to increase both resolution and field of view of the elemental image, but the conventional flat lens array is composed of elemental lenses of uniform focal length and it is impossible to control independently these two conditions. Therefore, a special method is required to get an elemental image with higher resolution and larger field of view than with the conventional flat lens array method.

IV. PRINCIPLES OF THE PROPOSED SCHEME USING SINGLE CAMERA AND A COMPOSITE LENS ARRAY

We propose an improved method by using a composite lens array as a substitute for a conventional lens array. The composite lens array consists of two sets of elemental lenses that have different foci. The concept of the proposed method is shown in Fig. 4.

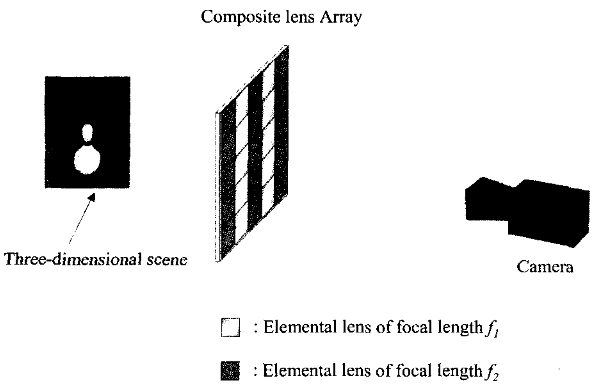


FIG. 4. The concept of the proposed method using single camera and a composite lens array.

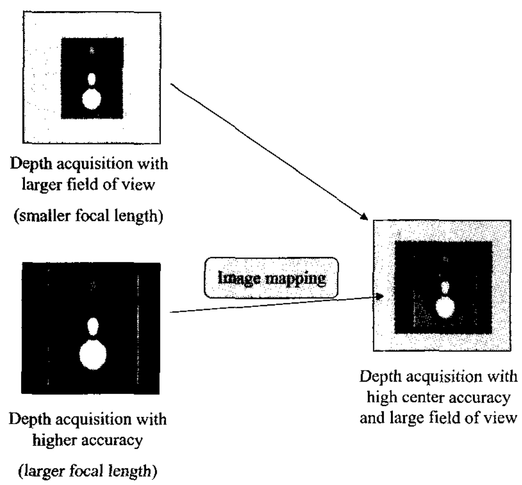


FIG. 5. The concept of the proposed image processing.

With the use of a composite lens array, it is possible to get two sets of elemental images. One set is composed of elemental images of higher resolution and the other set is composed of elemental images of larger field of view. In the proposed method, we combine these two advantageous features of two sets of elemental images to obtain a depth map with higher resolution and large field of view. The concept of proposed image processing is shown in Fig. 5.

As shown in Fig. 5, the proposed method replaces a more accurate depth extraction results with the corresponding position of the results with a larger field of view. The final result has high center accuracy and a large field of view. For most cases of stereovision, the primary object is located around the center of the image and thus the resolution of the center area of the depth extraction results is more important and needs to have higher accuracy. We can also easily recognize that the field of view of the entire image is still reserved. Although the use of a composite lens array decreases the number of elemental lenses with each focal length by half, the number of elemental lenses is still sufficiently large and hence the decrease of the number of the elemental images for each focal length does not severely affect the accuracy of depth extraction results.

V. SIMULATION RESULTS

We performed several simulations to verify the feasibility of the proposed method before the experiments. The first simulation is performed with two images that have depths of 124 mm (orange image) and 82 mm (apple image)· respectively. We assumed the composite lens array is composed of two sets of elemental lenses with focal lengths of 15 mm and 22 mm. We generated the elemental image by computer

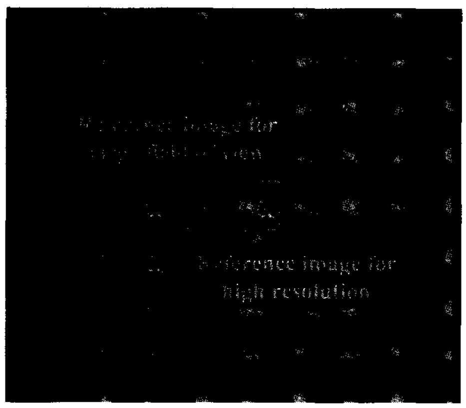


FIG. 6. Computer-generated elemental images of two objects with different depths.

and processed them to obtain the depth map. Figure 6 shows the generated elemental images.

The results of first simulation results are shown in Fig. 7. The depth acquisition result using elemental lens with focal lengths 15 mm and 22 mm are shown in Fig. 7 (a) and (b) respectively. The final result after fusing the two depth maps is shown in Fig. 7 (c). We can see that the large field of view of the depth map in Fig. 7 (a) is preserved in the resulting depth map in Fig. 7 (c) as expected. To evaluate the accuracy of the final depth map numerically, we propose a new evaluation function Error-of-Depth-per-Pixel (EDP) that is defined as follows:

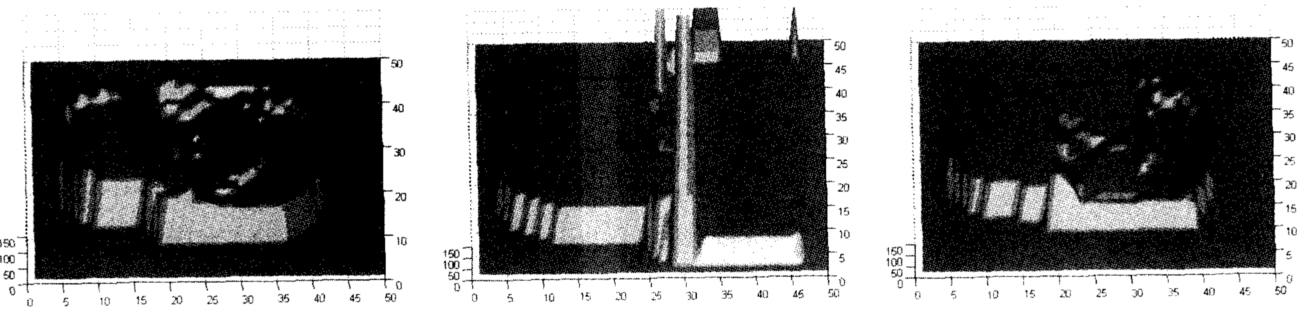
$$EDP = \frac{\sum_{Total\ pixels}(depth\ mismatch\ of\ each\ pixel)}{(Number\ of\ total\ pixels)} \quad (5)$$

Using the EDP, it is possible to compare the results numerically. In our simulation, the calculated value of EDP is 17.83 for Fig. 7 (a) and 14.78 for Fig. 7 (c). Note that it is not necessary to compare the value of EDP for Fig. 7 (b) since the depth map in Fig. 7 (b) and those in Figs. 7 (a) and (c) have different fields of view. Comparing the value of EDP, we can easily recognize that the EDP is reduced after the image processing and this means that the final result contains more accurate depth information about the object.

Another simulation is also performed with two images with depths of 226 mm and 180 mm respectively. The results of the second simulation are shown in Fig. 8. The depth acquisition result using elemental lenses with focal lengths 15 mm and 22 mm are shown in Fig. 8 (a) and 8 (b) respectively. The final result after the image processing procedure is shown in Fig. 8 (c). The EDP is also used in second simulation to compare the results. The EDP value of Fig. 8 (a) is 12.11 and Fig. 8 (c) is 7.39 and also reduced on second simulation. Therefore, it is obvious that the proposed method and image processing procedure make it possible to get better results with higher center accuracy and larger field of view.

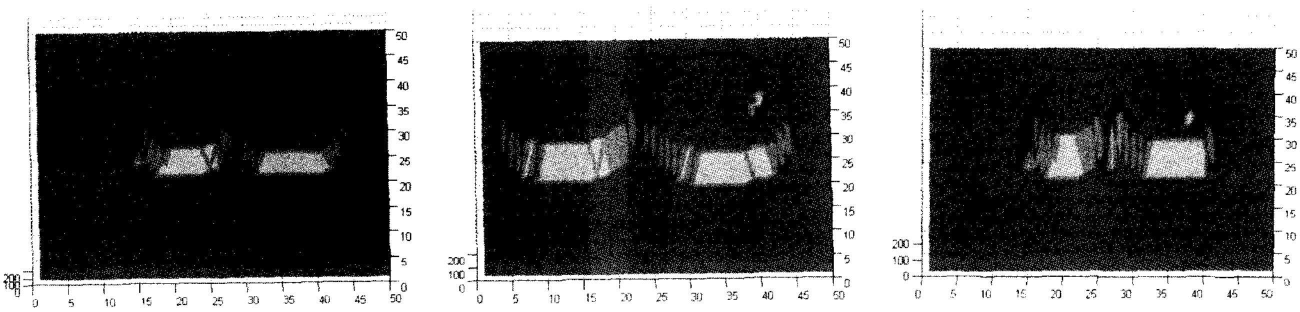
VI. EXPERIMENTAL RESULT

We performed preliminary experiments to verify that the proposed method can improve the performance of previously proposed method using the lens array. A composite lens array with two sets of elemental lenses with focal lengths of 15 mm and 22 mm is used. The elemental images from the composite lens array are



(a) (b) (c)

FIG. 7. Results of simulation of two images with depth of 124 mm and 82 mm respectively: (a) focal length 15 mm, (b) focal length 22 mm, (c) result of image processing with (a) and (b)



(a) (b) (c)

FIG. 8. Results of simulation of two images with depth of 226 mm and 180 mm respectively: (a) focal length 15 mm, (b) focal length 22 mm, (c) result of image processing with (a) and (b)

captured by a CCD camera. In the first experiment, we prepared a die and extracted its depth from the elemental images obtained using a composite lens array. The obtained elemental image is shown in Fig. 9.

We can see that the size of the die image differs in columns due to the different focal lengths of the elemental lenses. The disparity maps detected from each set of the elemental lenses with different focal lengths are shown in Fig. 10. Figure 10 (a) shows the disparity map obtained from the elemental images with small focal length and Fig. 10 (b) shows the disparity map obtained from the elemental images with large focal length. We can observe that the disparity map with small focal length has larger field of view. For the experimental results, it is impossible to compare the results by the value of EDP since we cannot know the depth of every pixel of elemental image unlike the case of the simulation results. Instead we compare the cross section of the disparity maps. Figure 11 shows the cross section of the disparity map in Figs. 10 (a) and (b). The fluctuation of the disparity map is much larger in the disparity map with small focal length that that with the large focal length. Hence we can be convinced that the error in the disparity map obtained with large focal length is smaller than that with small focal length.

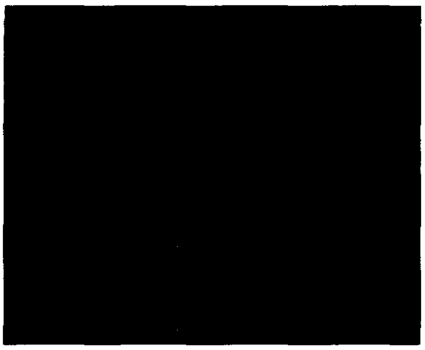
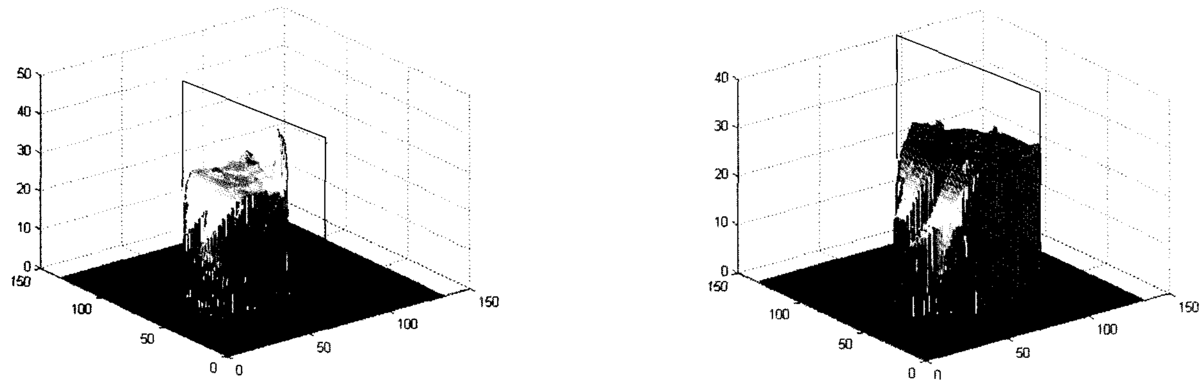


FIG. 9. Elemental images of a die captured by composite lens array.

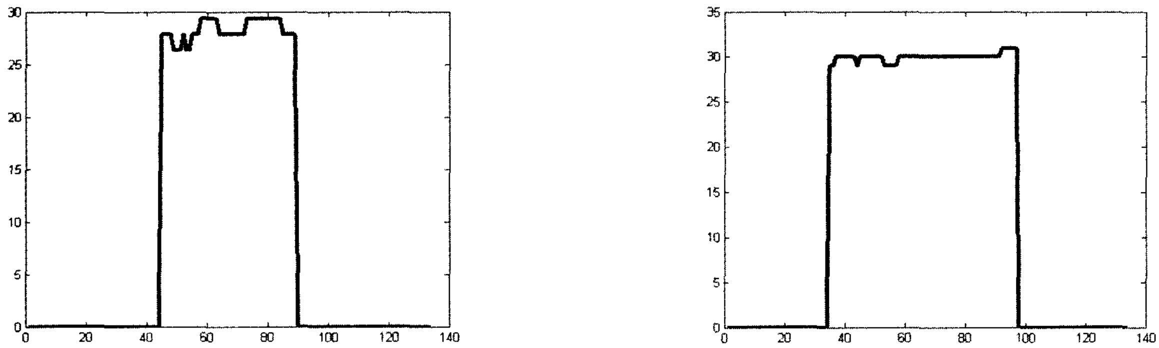
In the second experiment, we used two objects with different depths. The object is composed of two red and white spheres which are located at 340 mm and 270 mm from the composite lens array respectively. The captured elemental images are shown in Fig. 12. The elemental images of focal length 15 mm and focal length 22 mm are located in every other column.

The experimental results are shown in Fig. 13. In



(a) with small focal length (b) with large focal length

FIG. 10. Detected disparity map from the elemental images.



(a) with small focal length (b) with large focal length

FIG. 11. Cross section of the disparity map from the elemental images.

Fig. 13 (a) and (b), the depth extraction result using elemental images which are formed by elemental lenses with focal lengths 15 mm and 22 mm are shown. Figure 13 (c) shows the final result that is processed from Fig. 13 (a) and (b). We compare the results by the cross sections of Fig. 13 (a), (b), and (c). Figure 14 shows the cross sections of each figure. In Fig. 14 (a), the cross section of the results of focal length 15 mm is shown and we can recognize that the images of two spheres are not separated while the real objects are separated from each other. This is because the resolution of images from elemental lens with focal length 15 mm is not high enough to detect the separation of the two images. In Fig. 14 (c), however, the images of the two spheres are clearly separated after the image processing procedure and it is obvious the proposed method can improve the performance of the previously proposed stereovision system by adopting a composite lens array.

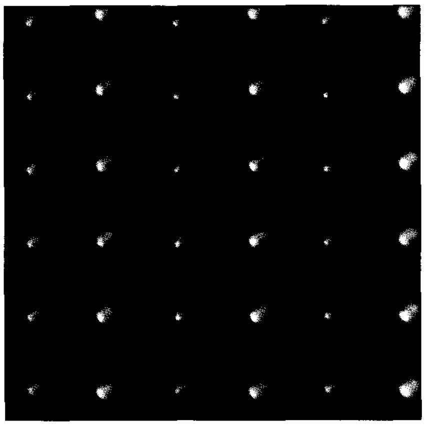


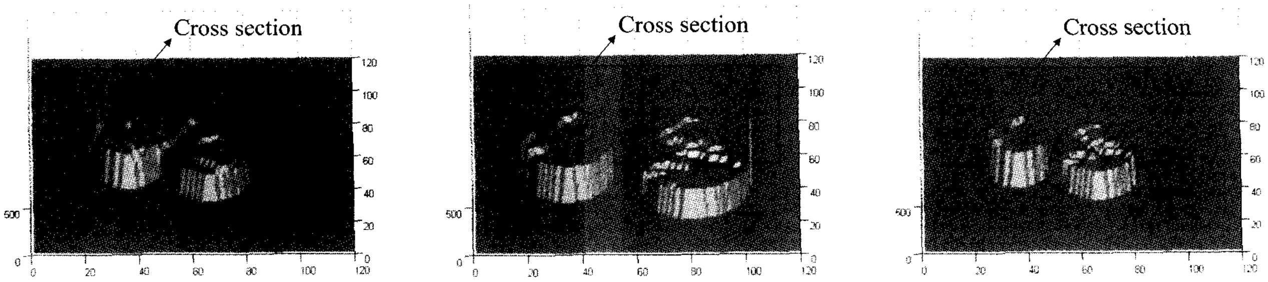
FIG. 12. Elemental images formed by the composite lens array and captured by a CCD camera.

VII. CONCLUSION

An improved stereovision scheme using single camera and a composite lens array is proposed and proved by simulation and experimental results. The use of a composite lens array makes it possible to construct a small and light-weighed stereovision system which is suitable for a mobile structure. Beyond that, the performance of the total system can be improved by adopting the proposed image processing procedure between the two sets of elemental images.

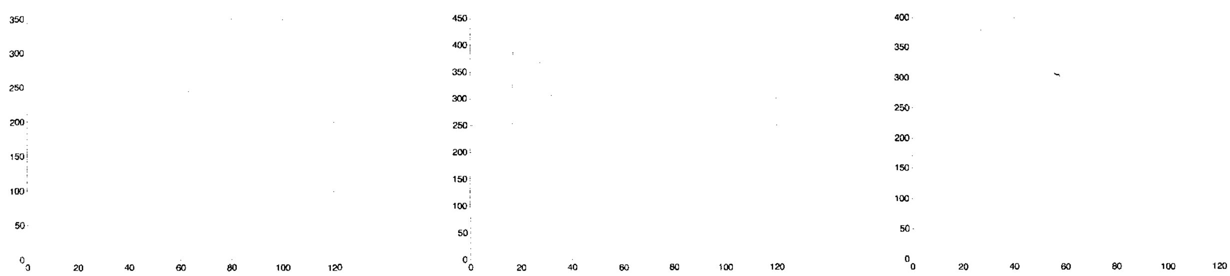
ACKNOWLEDGMENT

This work was supported by the Next-Generation



(a) (b) (c)

FIG. 13. Results of simulation of two images with depths of 340 mm and 270 mm respectively: (a) focal length 15 mm, (b) focal length 22 mm, (c) result of image processing with (a) and (b)



(a) (b) (c)

FIG. 14. Cross sections of experimental results in Fig. 10. (a) focal length 15 mm, (b) focal length 22 mm, (c) result of image processing with (a) and (b)

Information Display R&D Center, one of the 21st Century Frontier R&D Programs funded by the Ministry of Science and Technology of Korea.

*Corresponding author : byoungho@snu.ac.kr

REFERENCES

[1] D. Marr and T. Poggio, "A theory of human stereo vision," *Proc. Roy. Soc. B (London)*, pp. 301-328, 1979.

[2] S. T. Barnard and M. A. Fischler, "Stereo vision," *in Encyclopedia of Artificial Intelligence*. New York: John Wiley, pp. 1083-1090, 1987.

[3] M. Yachida, Y. Kitamura, and M. Kimachi, "Trinocular vision: new approach for correspondence problem," *Proc. International Conference on Pattern Recognition 2*, pp. 1041-1044, 1986.

[4] J.-H. Park, S.-W. Min, S. Jung, and B. Lee, "A new stereovision scheme using a camera and a lens array," *Conference on Algorithms and Systems for Optical Information Processing V, SPIE Annual Meeting, Proc. SPIE*, 4471, pp. 73-80, San Diego, CA, USA, Jul.-Aug. 2001.

[5] C. Wu, A. Aggoun, M. McCormick, and S.Y. Kung, "Depth extraction from unidirectional image using a modified multi-baseline technique," Conference on Stereoscopic Displays and Virtual Reality Systems IX, SPIE Electronic Imaging, *Proc. SPIE*, 4660, pp. 135-145, San Jose, CA, USA, May, 2002.

[6] J.-H. Park, S.-W. Min, S. Jung, and B. Lee, "Analysis of viewing parameters for two display methods based on integral photography," *Appl. Opt.*, vol. 40, pp. 5217-5232, 2001.

[7] B. Lee, S. Jung, and J. -H. Park, "Viewing-angle-enhanced integral imaging using lens switching," *Opt. Lett.*, vol. 27, pp. 818-820, 2002.

[8] S. Jung, J.-H. Park, B. Lee, and B. Javidi, "Viewing-angle-enhanced integral 3D imaging using double display devices with masks," *Opt. Eng.*, vol. 41, pp. 2389-2390, 2002.

[9] H. Choi, S.-W. Min, S. Jung, J.-H. Park, and B. Lee, "Multiple-viewing-zone integral imaging using a dynamic barrier array for three-dimensional displays," *Optics Express*, vol. 11, no. 8, pp. 927-932, 2003.

[10] H. Choi, J.-H. Park, J. Hong, and B. Lee, "Depth-enhanced integral imaging with a stepped lens array or a composite lens array for three-dimensional display," *The 16th Annual Meeting of the IEEE Lasers & Electro-Optics Society (LEOS 2003)*, Tucson, Arizona, USA, vol. 2, pp. 730-731, Oct. 2003.