

Computer generated hologram from full-parallax 3D image data captured by scanning vertical camera array

(Invited Paper)

Masahiro Yamaguchi*, Koki Wakunami, and Mamoru Inaniwa

*Global Scientific Information and Computing Center, Tokyo Institute of Technology, 2-12-1-I7-6,
Ookayama, Meguro-ku, Tokyo, 152-8550, Japan*

**Corresponding author: yamaguchi.m.aa@m.titech.ac.jp*

Received April 8, 2014; accepted May 15, 2014; posted online May 30, 2014

Full-parallax light-field is captured by a small-scale 3D image scanning system and applied to holographic display. A vertical camera array is scanned horizontally to capture full-parallax imagery, and the vertical views between cameras are interpolated by depth image-based rendering technique. An improved technique for depth estimation reduces the estimation error and high-density light-field is obtained. The captured data is employed for the calculation of computer hologram using ray-sampling plane. This technique enables high-resolution display even in deep 3D scene although a hologram is calculated from ray information, and thus it makes use of the important advantage of holographic 3D display.

OCIS codes: 090.0090, 090.1760, 090.2870, 110.6880.

doi: 10.3788/COL201412.060018.

1. Introduction

Holography is capable of reproducing optical 3D image that satisfies all the depth cues of human visual system, and it is expected to be a higher-quality 3D display medium in near future^[1,2]. The optical image reconstruction is also possible by light-ray reconstruction, such as a 3D display by integral imaging, but the resolution is degraded in the image distant from the display screen^[2]. One of the most important advantages of holography, which is based on wavefront reconstruction, is that high-resolution images can be reproduced even in deep 3D scene^[3].

In the electronic display of holography, the holographic fringe pattern have to be calculated by simulating the interaction and propagation of light waves based on the technology of computational holography. As a simple method called point source method have been often used for computational holography^[4], in which point light sources are defined on the object surface and the spherical waves from those point sources are superposed on the hologram plane. More advanced polygon-based method enabled high-definition 3D image reproduction using computer generated holograms (CGH)^[5]. Another approach for computational holography is the use of multiview images^[6–9]. It is based on the principle of holographic stereogram, and can be considered as light-ray reproduction if the sampling density is satisfactory high. Since a hologram is calculated from light-ray information in this approach, sophisticated rendering techniques of conventional computer graphics (CGs) can be employed, such as occlusion culling, rendering of glossy or specular surfaces, translucent or semi-translucent objects. However, the image resolution is limited if the image location is distant from the hologram plane in ray-based computation of holograms, because the sampled ray becomes sparse and the diffraction effect becomes significant in the image far from the hologram plane^[10,11]. Although

the resolution of reconstructed image is limited by the device technology in current electronic holography display, the high-resolution reproduction of deep 3D scene will be an important advantage of holographic approach in future. Therefore ray-based calculation technique that loses this advantage of holographic display, is not preferable for reproducing deep 3D scene.

To make use of the advantages of the above two approaches, we proposed a method using ray-sampling (RS) plane for computational holography^[2,10]. In this method, a hologram can be created from the light-ray information calculated by conventional CG rendering software. In addition, high-resolution image can be reproduced even in deep 3D scene by exploiting ray-wavefront conversion.

In the CGH calculation using RS plane, ray-based 3D data are employed and enables the compatibility with other 3D systems of computer graphics, image-based rendering, multiview imaging, and integral imaging. When producing holograms of pictured real objects, it is possible to apply the data from conventional 3D systems. However, holograms require much higher sampling density of light-rays. Shooting images for horizontal-parallax-only 3D is not much difficult with using horizontal camera motion or horizontal camera array. But it is considerably laborious to capture full-parallax images, and large-scale systems have been reported for the light-field acquisition^[12,13]. Integral imaging is also a promising approach to capture full-parallax light-field^[11,14–16]. An integral imaging camera can capture dense light-ray information, but the viewing zone size is limited by the size of the lens-array. Although the viewing zone achieved in the current electronic holography is small due to the limitation of device technology, it is expected to obtain light-field data with wide viewing zone for future holographic display. It will be needed to employ multiple integral imaging cameras for this purpose.

Instead of capturing full-parallax images, a method using 3D measurement and texture mapping is also possible, but the reproduction of specular or glossy surface,

and translucent or semi-translucent object requires complex modeling of 3D object.

We previously proposed a system for capturing full-parallax light-field of still-images with a scanning vertical camera array^[17,18]. A vertical array of small cameras is scanned horizontally, and the intermediate views between cameras in vertical direction are interpolated, thereby obtaining high-density light-ray information with a compact and small-scale system controlled by a single PC. In the previous report, full-parallax holographic stereograms were printed from the image data captured by this system. However, there was a limitation in the view interpolation method, in which the depth-image-based-rendering (DIBR)^[15] was used; object motion or specular reflection sometimes caused error in depth estimation, resulting that artifacts were appeared in interpolated images. In this paper, improved interpolation method is presented to reduce the influence of object motion or specular reflection.

In addition, CGHs are computed from the images captured by the system of vertical camera array scanning. The algorithm for calculating CGH is based on an RS-plane; ray-wavefront conversion is applied. Though the hologram is calculated using ray-information, high-resolution reproduction is possible even in deep 3D scene. Experimental results of holographic 3D image produced by a CGH printer^[2] is demonstrated.

2. Scanning vertical camera array system

In the system for full-parallax image capture, a vertical camera array is scanned horizontally as shown in Fig. 1(a)^[17,18]. Figure 1(b) shows the vertical array of 7 cameras controlled by a single PC, which was used in the experiment below. High-density sampling is easily realized in horizontal direction in this system, but the number of cameras in vertical direction is limited. Then the vertical parallax information is interpolated from the captured data. As only a small number of cameras are used, high-density light-field data can be acquired by relatively simple and compact implementation.

For the image acquisition in wider angular range, the vertical camera array is rotated with the horizontal motion so that the camera always points the object as shown in Fig. 1(c). In this case, the captured images have key-stone distortion and a distortion correction is applied using the images of a checkerboard pattern. A perspective projection was applied to derive the images virtually captured by cameras arranged parallel to the camera motion axis as shown in Fig. 1(c). In this manner, the ray information from the object is efficiently acquired.

3. Vertical interpolation

To obtain high-density light-ray information, the rays passing through the vertical gap between cameras are interpolated. In this system, the depth of each pixel is estimated using the captured images, and vertical views between cameras are interpolated by DIBR technique with the estimated depth information^[19–21]. The method for depth estimation is described in the next section.

Figure 2 shows the geometry of the interpolation of a vertical view. The image of virtual rendering camera is generated from the images of upper and lower cameras using the depth information of the object. It is a linear interpolation of the pixel data of corresponding points. In many cases the angular characteristics of reflection is smooth and high-frequency component is not dominant, the interpolation shown in Fig. 2 works well. If there is an occlusion, corresponding points is not found in a vertical camera pair. Then, the depths of a corresponding points in upper and lower images are different. In such case, the pixel data that has smaller depth is used from either upper or lower camera.

4. Depth estimation

For the vertical interpolation explained above, the depth map of each captured image is required. Various depth estimation techniques for multiview images can be applied, such as stereo matching or Epipolar plane image (EPI) analysis. In previous paper, EPI analysis was

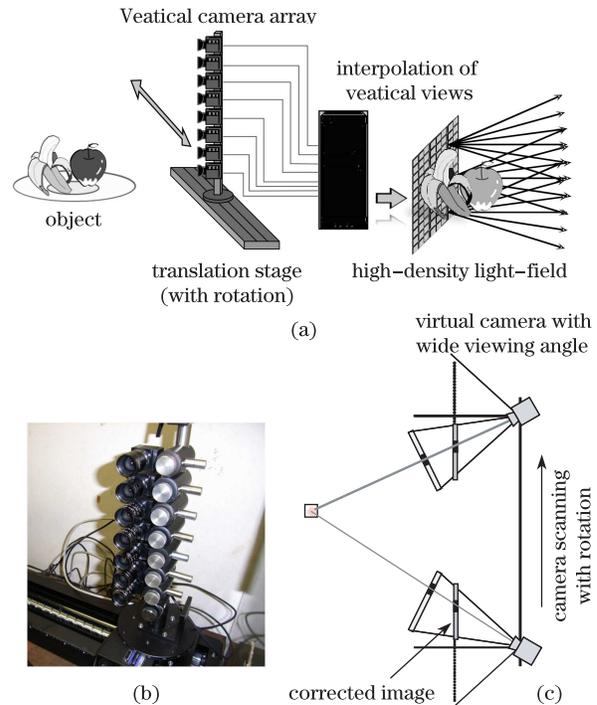


Fig. 1. (a) Scanning vertical camera array system for the acquisition of high-density light field information. (b) Experimental setup of vertical camera array. (c) Geometry of horizontal camera motion and distortion correction.

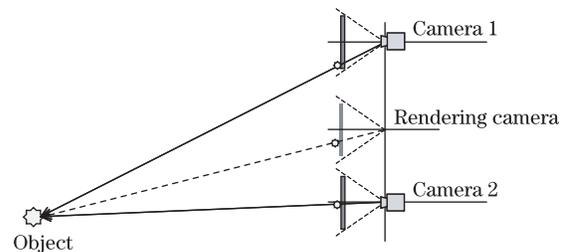


Fig. 2. Geometry for vertical-parallax interpolation. (Side view) The image of rendering camera (new camera) is interpolated from the images captured by upper (camera 1) and lower (camera 2) cameras using depth information.

applied in the horizontal direction, as the sampling rate in horizontal views is high and high-resolution EPI is obtained. However, to avoid mismatching due to possible object motion or specular reflection on the object surface, only limited range of camera positions around the target camera is used, because their effect increases if using long sequence of horizontal camera motion. But when the range of camera positions used in matching is small, i.e., the baseline is small, the depth resolution becomes low.

Therefore, an improved depth estimation is adopted in this paper. In this method, stereo matching is applied to the vertical pair of images. Since the images of different heights are captured simultaneously, the depth estimation is not affected by the object motion. The baseline of vertical camera pair is satisfactory because the purpose of depth estimation is just the interpolation between the cameras. However, when the baseline is larger, stereo matching sometimes fails. For example, if the object has specular reflection, it may cause mismatch. To prevent such error, we utilize the information of horizontal views. A small number of images captured with horizontal scanning are used for rough estimation, and the disparity search using vertical camera pair is done only in the range determined by the rough estimation. Since the baseline is small in horizontal search, the stereo matching becomes robust though the accuracy is not high. In addition, the number of images used in the horizontal direction is small, the influences of object motion, angular dependent reflection, and occlusion effect become insignificant. On the other hand, the accuracy required for this purpose is achieved by vertical camera pair.

5. Syntheses of arbitrary viewpoint images

After the interpolation of vertical views, it is possible to synthesize the images of arbitrary viewing position based on the light-ray information. Then the image data can be applied to free-viewpoint images, though only still objects are captured by this scanning system. For computing a CGH using RS plane, which is described in the next section, the projection images are calculated first where the centers of projection are the RS points. The projection images are considered to be the angular spectra of the wavefront on the RS plane, and the wavefront on the hologram plane is obtained by the technique presented below.

6. Calculation of hologram using ray-sampling plane

The method for calculating hologram using RS-plane, presented in Ref. [10], is briefly reviewed in this section. As shown in Fig. 3, an RS plane is defined near the object and the light-rays that pass through the RS points on the RS plane are computed by an arbitrary rendering technique for CG. A set of light-rays that pass an RS point is collected, and is corresponds to a projection image where the RS point is the viewpoint. Then random phase is attached to each pixel of the projection image, and a discrete Fourier transform (or fast Fourier trans-

form: FFT) of the projection image yields the wavefront in a small region around the RS point. After FFTs of the projection images for all the RS points, the wavefront on the RS plane is obtained. The wavefront propagation from the RS plane to the hologram is calculated by Fresnel diffraction theory. The wavefront propagation by Fresnel diffraction can be implemented with discrete Fresnel transform, which can also be efficiently computed using FFT. Despite that the wavefront is generated from ray information, the resolution is not degraded because the long-distance propagation is calculated with diffraction theory.

When there are multiple objects at different depth, multiple RS planes are defined as shown in Fig. 4. Then occlusion culling between the objects that are assigned to different RS planes can be implemented using mutual conversion between ray and wavefront^[22].

To apply the method to the calculation of hologram of real objects, the projection images are generated from the captured images. It can be implemented using model-based or image-based rendering. We employ image-based rendering because it can deal with complex scene that includes specular surface, iridescent colors, transparent or translucent objects, etc., without introducing complicated models. The projection images are synthesized from the full-parallax light-field data captured by the scanning vertical camera array system. Then it is possible to calculate a CGH from the captured images of real objects.

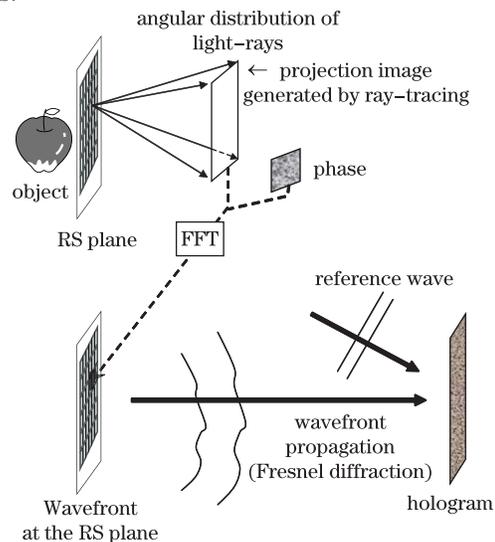


Fig. 3. Schematic diagram of CGH calculation using RS plane.

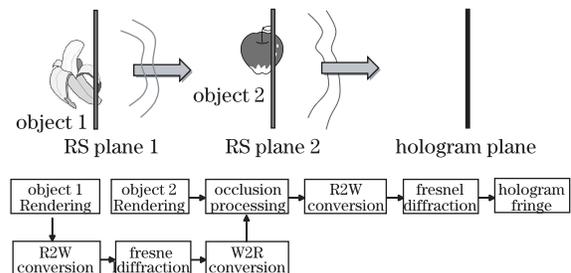


Fig. 4. Diagram of occlusion processing in CGH calculation by using R2W and W2R conversions.

7. Experimental result

In the experiment, a set of full-parallax 3D image was captured by the scanning vertical camera array system and employed for calculating CGH. Seven compact CCD cameras (Point Gray Research, Flea) were vertically arranged and set on rotation and translation stages. The vertical interval of cameras was 60 mm. The system is controlled by a PC. 577×7 images in 480×640 pixels were captured with horizontal scan of 576 mm in 40 s. The distance between the object and the camera array was about 500 mm. The camera array was rotated to point the object, where the maximum rotation angle was 60° . The distortion correction was applied as described above.

Figure 5 shows the captured images before and after the distortion correction. The background was removed by chroma-key before the distortion correction. The depth estimation was performed by stereo matching between the images of the vertical pair of cameras [Figs. 5(c) and (e)]. The block matching using sum of absolute difference (SAD) with 25×25 pixel window was used. Figure 6 shows an example of SAD for different disparity at the pixel coordinates (131, 334) of the image in Fig. 5(c). SAD became minimum when the pixel index was 80, but it was incorrect matching due to the difference of specular reflection appearance between the images of larger baseline. Stereo matching between the images where the camera interval was 6 mm was also applied with the same window size. Then the obtained corresponding pixel was 104. As the depth resolution was low for smaller baseline (6 mm) stereo, the search area was defined around the correspondent point obtained by horizontal stereo matching as shown in Fig. 6. The search area size was empirically decided as ± 5 pixels around the corresponding point obtained by horizontal stereo matching. Then the SAD became minimum at pixel index=106, which was a correct value. After the depth estimation, a 9×9 median filter was applied to remove noise in the depth image. In Ref. [20], the depth map is preprocessed by Gaussian filter to smooth out the large discontinuities. More advanced technique to obtain smooth depth map by iteration is presented in Ref. [21]. In the system of this paper, the occlusion-hole problem is not very serious compared with a system in which the separation of cameras is large. Thus simple median filtering works well in this system, in addition to the depth estimation technique integrating vertical and horizontal camera images.

Figure 7 shows the results of depth estimation by the proposed method, the vertical matching only, and the EPI analysis in horizontal direction. In EPI analysis, the depth was calculated using 50 images captured by the cameras located at ± 25 mm in horizontal direction. In the results of vertical matching only [Fig. 7(b)], there found noises and holes in the depth image. They appeared at the location of specular reflections by gold weave. There also found noises in the results of EPI analysis where the object had fine textures. This shows that the proposed method gives good depth estimation with relatively simple stereo matching operations in both horizontal and vertical directions.

Figure 8 is the result of view interpolation. The estimated images of interpolation between those of Figs. 5(c) and (e) are shown. We also capture an image at the same camera location as the interpolated image as shown in Figs. 8(a)–(c) are those obtained with Figs. 7(b) and (c), respectively. It can be confirmed that the proposed method enables satisfactory results of view interpolation. The accuracy of the interpolated image was evaluated with the differences from the image actually captured by the camera placed at the location of the rendering camera. Then the peak signal to noise ratio (PSNR) was derived. It was 28.64 dB in the proposed method, 27.05 dB in vertical stereo matching only, and 27.16 dB in horizontal EPI analysis. About 1.5 dB improvement was achieved in this example.

Finally, a CGH was calculated using the captured light-field data. The geometry for CGH calculation is shown in Fig. 9. An RS plane was defined near the object, and

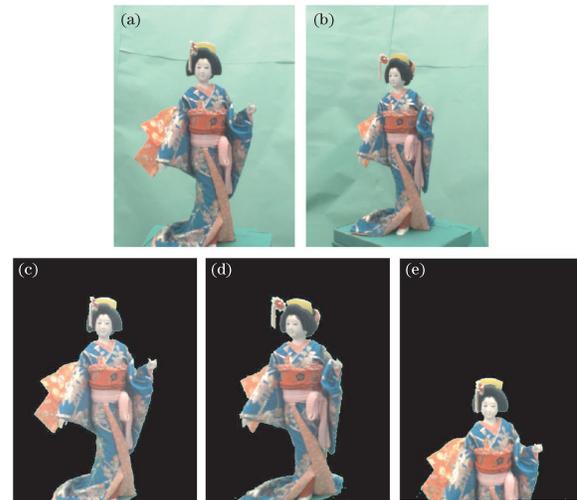


Fig. 5. Examples of images captured by the camera (a) from front and (b) right positions. (c) and (d) are the images after the distortion correction of the images of (a) and (b), respectively. (e) is the image captured from front but by the lower camera.

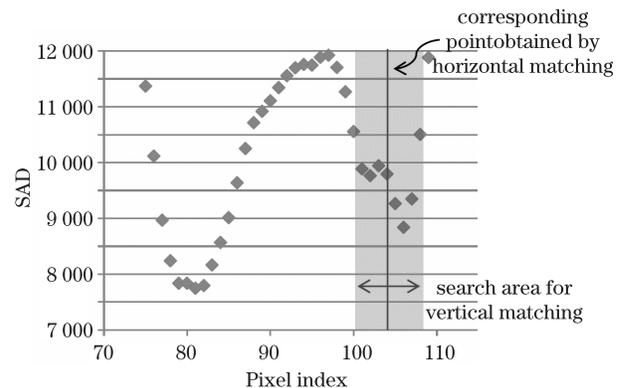


Fig. 6. An example of vertical matching point search. The horizontal axis corresponds to the vertical pixel index and vertical axis is the SAD. The correct matching point is likely the pixel index = 108, but the SAD became smaller value when pixel index = 78–82. The pixel index of corresponding point obtained by horizontal matching was 104, and correct matching point is obtained by narrowing the search area based on the horizontal matching result.

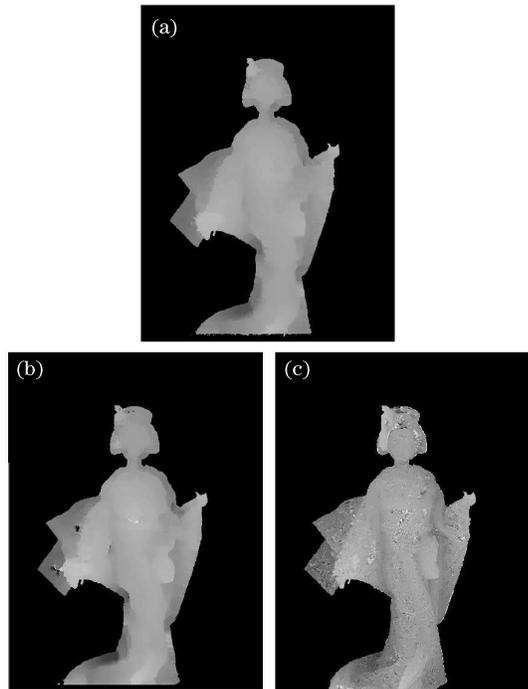


Fig. 7. Estimated depth images for Fig. 5(a). Higher intensity represents smaller depth in the depth images. (a) Proposed method, (b) vertical stereo matching only, and (c) horizontal EPI analysis.

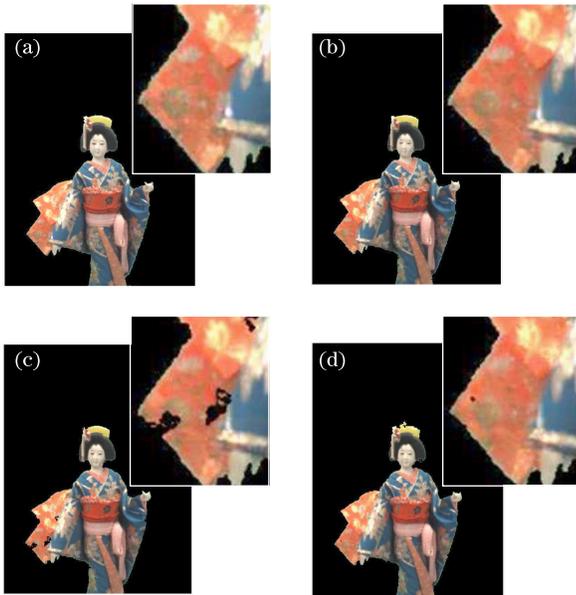


Fig. 8. (a) Image captured by a real camera at the position of interpolated camera. (b)–(d) Results of vertical interpolation. (b) Proposed method, (c) vertical stereo matching only, and (d) horizontal EPI analysis. Magnified images of the rectangular region shown in (a) are presented at upper right of all figures. Interpolation errors are observed in (c) and (d).

the distance between the RS plane and CGH was about 200 mm. The parameters for the CGH calculation are summarized in Table 1. The calculated CGH pattern was printed using the CGH printer^[3]. Figure 10(a) shows the simulated reconstructed image where the eye pupil whose diameter is assumed to be 5 mm is located at 200 mm

apart from the hologram plane. Figure 10(b) is the photograph of the image obtained by optical reconstruction. Full-parallax 3D image could be observed, but the viewing angle was small ($\cong 7.3^\circ$). The viewing angle was determined by the resolution of hologram and the hologram size. Recording of holograms with larger viewing zone is an important future issue by higher-resolution and larger-size printing.

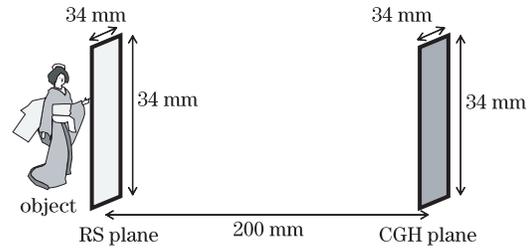


Fig. 9. Geometry for CGH calculation. The sizes of RS plane and CGH plane are both 34 mm.

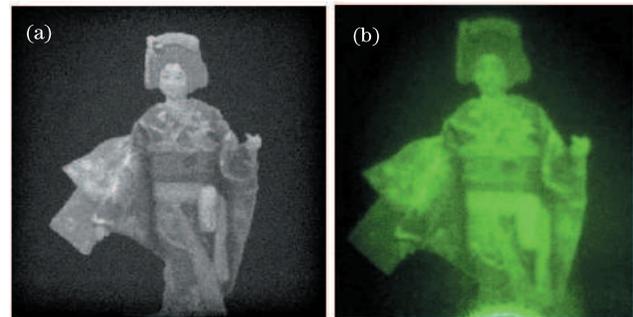


Fig. 10. Reconstructed images by (a) computer simulation and (b) experiment.

Table 1. Parameters for CGH Calculation

Total Numbers of Pixels in RS Plane and CGH	16384×16384 pixels
Number of RS Points	256×256 points
Resolution of a Ray-image	64×64 pixels
Sampling Interval of RS Points	134.4 μm
Sampling Intervals in RS Plane and CGH	2.1 μm
Sizes of RS Plane and CGH	34.4×34.4 mm
The Distance from RS Plane to CGH	200 mm

8. Conclusion

In this paper, the system for capturing high-density light-field information is presented and applied to the generation of hologram of a real 3D object. The scanning vertical camera array technique allows a small-scale system controlled by a single PC to collect high-resolution full-parallax 3D images. The method for view interpolation in vertical direction is improved with effectively using horizontal and vertical parallax information. Then the error in depth estimation is significantly reduced and interpolated images of better quality are obtained. The

angular dependent reflection of object surface, i.e., specular or glossy surfaces, can be reproduced from captured images by applying DIBR technique. The high-resolution full-parallax data is considered to be ray-information, and it is converted to wavefront for calculating CGH using RS plane. The method for the computation of hologram presented here is capable of reproducing deep 3D scene in high resolution.

Acknowledgement

This work was partly supported by the JSPS Grant-in-Aid for Scientific Research #17300032.

References

1. H. Sasaki, K. Yamamoto, Y. Ichihashi, and T. Senoh, "Image Size Scalable Full-parallax Coloured Three-dimensional Video by Electronic Holography," *Scientific Reports*, 4, Article number: 400006 (2014).
2. M. Yamaguchi, *Proc. SPIE* **8043**, 804306 (2011).
3. Y. Kim, J.-H. Park, H. Choi, J. Kim, S.-W. Cho, and B. Lee, *Appl. Opt.* **45**, 4334 (2006).
4. J. P. Waters, *Appl. Phys. Lett.* **9**, 405 (1966).
5. K. Matsushima and S. Nakahara, *Appl. Opt.* **48**, H54 (2009).
6. T. Yatagai, *Appl. Opt.* **15**, 2722 (1976).
7. A. D. Stein, Z. Wang, and J. J. S. Leigh, *Comput. Phys.* **6**, 389 (1992).
8. P. W. McOwan, W. J. Hossack, and R. E. Burge, *Opt. Commun.* **82**, 6 (1993).
9. M. Yamaguchi, H. Hoshino, T. Honda, and N. Ohya, *Proc. SPIE* **1914**, 25 (1993).
10. K. Wakunami and M. Yamaguchi, *Opt. Express* **19**, 9086 (2011).
11. K. Wakunami, M. Yamaguchi, and B. Javidi, *Opt. Lett.* **37**, 5103 (2012).
12. B. Wilburn, N. Joshi, V. Vaish, E.-V. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy, *ACM Trans. Graphics*, **24**, 765 (2005).
13. M. Brewin, M. Forman, and N. A. Davies, *Proc. SPIE* **2409**, 118 (1995).
14. J.-H. Park, M.-S. Kim, G. Baasantseren, and N. Kim, *Opt. Express* **17**, 6320 (2009).
15. N. Chen, J.-H. Park, N. Chen, J. Yeom, J.-H. Park, and B. Lee, *Opt. Express* **19**, 26917 (2011).
16. T. Mishina, M. Okui, and F. Okano, *Appl. Opt.* **45**, 4026 (2006).
17. M. Yamaguchi, R. Kojima, and Y. Ono, *Nicograph International* (2008).
18. M. Yamaguchi, *Proc. SPIE* **9042**, 90420A (2013).
19. L. Zhang, W. J. Tam, *IEEE Transactions on Broadcasting* **51**, 191 (2005).
20. C. Fehn, R. Barre, and R. S. Pastoor, *Proc. IEEE* **94**, 524 (2006).
21. W. Li, J. Zhou, B. Li, and M. I. Sezan, *IEEE Trans. Circ. Sys. For Video Technol.* **19**, 533 (2009).
22. Koki Wakunami, Hiroaki Yamashita, and Masahiro Yamaguchi, *Opt. Express* **21**, 21811 (2013).