

Optical Engineering

OpticalEngineering.SPIEDigitalLibrary.org

Generation of binary holograms for deep scenes captured with a camera and a depth sensor

Thibault Leportier
Min-Chul Park

Generation of binary holograms for deep scenes captured with a camera and a depth sensor

Thibault Leportier^{a,b} and Min-Chul Park^{a,b,*}

^aKorea University of Science and Technology, Department of HCI and Robotics, Hwarangno 14-gil 5, Seongbuk-gu, Seoul 136-791, Republic of Korea

^bKorea Institute of Science and Technology, Center of Opto-Electronic Materials and Devices, Hwarangno 14-gil 5, Seongbuk-gu, Seoul 136-791, Republic of Korea

Abstract. This work presents binary hologram generation from images of a real object acquired from a Kinect sensor. Since hologram calculation from a point-cloud or polygon model presents a heavy computational burden, we adopted a depth-layer approach to generate the holograms. This method enables us to obtain holographic data of large scenes quickly. Our investigations focus on the performance of different methods, iterative and noniterative, to convert complex holograms into binary format. Comparisons were performed to examine the reconstruction of the binary holograms at different depths. We also propose to modify the direct binary search algorithm to take into account several reference image planes. Then, deep scenes featuring multiple planes of interest can be reconstructed with better efficiency. © The Authors. Published by SPIE under a Creative Commons Attribution 3.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: 10.1117/1.OE.56.1.013107]

Keywords: digital holography; binary hologram; direct binary search; bidirectional error diffusion.

Paper 161706 received Oct. 31, 2016; accepted for publication Dec. 15, 2016; published online Jan. 13, 2017.

1 Introduction

Most technologies proposed for 3-D display like stereoscopic or multiview systems do not provide sufficient visual depth cues for observers to perceive depth in natural conditions. Then, even if 3-D stimuli can be induced, internal conflicts such as mismatch between accommodation and vergence may induce visual fatigue and discomfort.¹ Holography is a technique that has received attention for its ability to reconstruct complex optical fields in natural viewing conditions.² However, several technical bottlenecks still need to be overcome for recording and displaying digital holograms in real time and with a high quality.

The two main approaches to obtain holographic data are optical capture and numerical computation. For optical methods, 3-D information of real objects can be acquired by generating an interference pattern between an object wave and a reference wave. However, recording conditions require a dark room and high stability. Different techniques such as phase-shifting holography³ and optical scanning holography⁴ (OSH) have been demonstrated for capturing complex holograms, but only small objects can be recorded due to the limitations imposed by the capturing setup.

In the past few years, efforts have been invested in the computation of holograms by numerical methods. One approach is to represent an object by a point-cloud⁵⁻⁷ or a polygonal mesh.^{8,9} The advantage of the point-cloud method is that it is a natural approach to compute diffraction patterns since diffraction theory is well established for a point source. The approach based on polygon representation allows the reduction of computation time, but it requires a more complex algorithm to compute the propagation formula between nonparallel planes. An alternative method for fast computation is to generate a hologram by combining an intensity

image and a depth map.¹⁰⁻¹² Since the optical propagation of a plane can be computed efficiently with fast Fourier transformation, the depth map information can be used to decompose the image of the scene into multiple layers that can be propagated separately to the hologram plane. The computation can be done in parallel for the different layers, and computation time can, therefore, be very fast. The main drawback is the low resolution in depth.

For optical reconstruction of digital holograms, data can be displayed with spatial light modulator devices (SLMs). However, data format should be adapted to the used device. Indeed, complex information cannot be loaded directly into the SLMs. It is of high interest to convert the holographic data into binary format to exploit the full bandwidth of the SLM. In addition, the binary format enables to reduce the storage burden and is suitable for hologram printing. The most basic technique to convert a complex hologram into a binary hologram is to perform a threshold operation, but the quality of the reconstruction is low.¹³ A method relying on bidirectional error diffusion (BERD) was proposed to convert complex data into a phase-only hologram.¹⁴ It can be adapted for the generation of binary holograms. Iterative methods such as the direct binary search (DBS) algorithm were also reported.^{15,16}

In this paper, we focused our investigations on the generation of binary holograms from data acquired from a Kinect sensor. We used the depth layer approach to compute quickly a hologram of large real objects. Then, we examined the performance of a threshold method, BERD, and DBS algorithms to convert the holographic data into a binary format. Comparison of the image planes obtained from the binary holograms and the original hologram were performed for various reconstruction distances including the main focusing plane of the object. We show that the mean square error (MSE) between the images obtained from the DBS method and the original hologram is optimal for the

*Address all correspondence to: Min-Chul Park, E-mail: minchul@kist.re.kr

reconstruction distance used as reference in the application of the DBS procedure. Then, we propose to modify the DBS algorithm to take into account multiple reference planes and therefore increase the efficiency when multiple objects are located at different depths.

2 Hologram Generation from Real Objects

The scene used in this study was captured with a Kinect V2 sensor. It included the figurine of a robot of size $40 \times 15 \times 12 \text{ cm}^3$. The device allows the processing of data with different approaches. Point cloud and mesh representations of the scene can both be generated.

The Kinect sensor features both an RGB camera and an IR sensor. Then, three different outputs are obtained: an RGB image of the scene, an IR image, and a depth map. Since the camera and the IR sensor are spatially separated, it is first necessary to calibrate the data for the depth map to match the RGB image.¹⁷ For simplicity, we decided to use the IR image instead of the RGB image because it already matches the depth map pixel by pixel. Different representations of the object under interest are shown in Fig. 1.

The advantage of the Kinect sensor is that it enables to reconstruct a 3-D model of a large real scene when other traditional optical techniques to generate real object holograms are generally restricted to small objects. However, the spatial resolution is limited. We chose the depth layer approach to generate the hologram from the intensity image and the depth map since this approach is fast and is well adapted to the format of the data captured.

The principle of the depth layer approach is to separate all pixels, layer by layer, that correspond to a given depth value. Since the depth map is represented by 256 values, we have 256 depth layers. The depth range of the scene can be controlled by setting the actual distance corresponding to values 0 and 255. The distances between the hologram and the different layers are then defined in a linear way between the minimum and maximum values. Then, the propagation of each layer to the hologram plane can be handled by a numerical method such as the Fresnel propagation equation or the convolution approach.¹⁸ Illustration of the computation of a

hologram with the depth layer-based approach is shown in Fig. 2.

In our case, we selected only the depth layers including the robot so that we could focus our efforts to reconstruct the large object. In practice, we, therefore, used only layers 14 to 22. Once the hologram is computed, numerical reconstruction can be performed. Optical reconstruction can be realized directly by complex modulation with a single SLM¹⁹ or multiple display devices,²⁰ but it is not an easy task. Then, an additional step is necessary to convert the data into a suitable format for display with an SLM. The focus of our study is the conversion into a binary hologram. Since the complex nature of the hologram will be lost after conversion, a zeroth order and twin image will become visible in the reconstruction plane. We, therefore, started by converting the complex hologram into an off-axis hologram by multiplying the data with a spatial carrier. The angle was set at 1 deg to ensure the spatial separation of the twin image and object in the reconstruction plane of the binary holograms.

3 Conversion to Binary Format

The simplest way to convert a hologram into binary format is to apply a threshold on the hologram. For instance, all values for which the real part of the hologram is negative can be put to zero, whereas positive values are set to one. This operation is easy and fast to implement but may lead to strong noise and distortions in the reconstruction plane. We are going to investigate the performance of two additional methods that were demonstrated for generation of a binary hologram.^{4,14,15}

3.1 Bidirectional Error Diffusion

BERD is a noniterative method. Tsang and Poon¹⁴ demonstrated its potential for generation of phase-only holograms. The same principle can be used for the generation of a binary hologram. The principle of the method is to scan each individual pixel of the hologram and to put it at the desired value. For instance, a pixel can be put to zero if the original value is negative, and to one otherwise. The difference with the threshold approach is that after each change, the error between the original and the new values is computed and

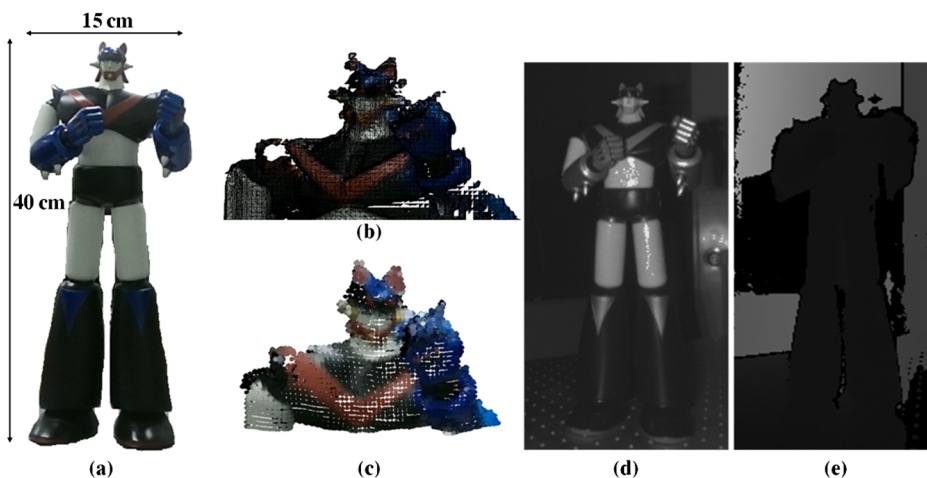


Fig. 1 (a) The object of interest used in this study is a robot figurine. (b) Data captured by the Kinect sensor can be represented with a mesh model or a (c) point cloud. For the depth-layer approach, we used (d) the intensity image and (e) the depth map.

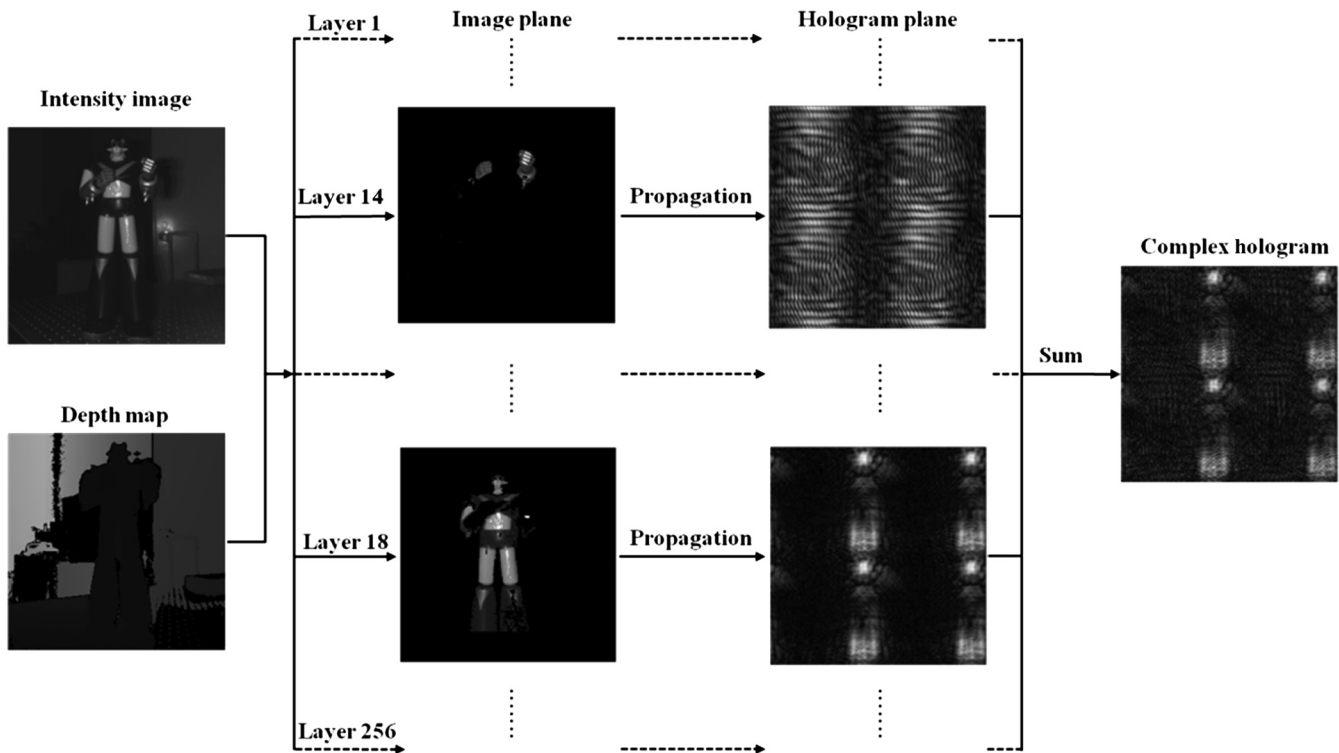


Fig. 2 Principle of hologram computation based on a depth layer-based method.

diffused to the neighboring pixels that have not yet been processed. The update of the pixel $\text{holo}_{\text{orig}}(i, j)$ to a value $\text{holo}_{\text{bin}}(i, j)$ leads to an error $E(i, j) = \text{holo}_{\text{orig}}(i, j) - \text{holo}_{\text{bin}}(i, j)$. When a line of pixels is scanned from left to right, the diffusion of error is performed as follows to the surrounding pixels:

$$\begin{cases} \text{holo}_{\text{orig}}(i, j + 1) = \text{holo}_{\text{orig}}(i, j + 1) + w_1 \times E(i, j) \\ \text{holo}_{\text{orig}}(i + 1, j - 1) = \text{holo}_{\text{orig}}(i + 1, j - 1) + w_2 \times E(i, j) \\ \text{holo}_{\text{orig}}(i + 1, j) = \text{holo}_{\text{orig}}(i + 1, j) + w_3 \times E(i, j) \\ \text{holo}_{\text{orig}}(i + 1, j + 1) = \text{holo}_{\text{orig}}(i + 1, j + 1) + w_4 \times E(i, j) \end{cases} \quad (1)$$

The constants w_1 to w_4 are weighting coefficients set, respectively, to 7/16, 3/16, 5/16, and 1/16 in accordance with the

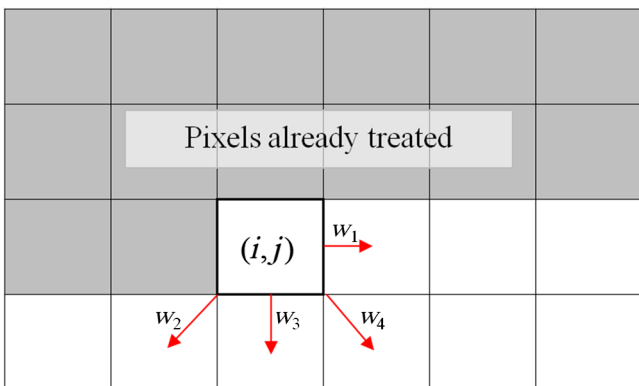


Fig. 3 The BERD algorithm consists of updating the pixels one by one and to diffuse the error to the pixels that have not been treated yet.

previous studies related to BERD. It was reported that the performance of the method was greatly improved when the scanning direction alternated from left to right and from right to left between each pixel line. Note that when the scan direction is from right to left, Eq. (1) should be slightly modified to diffuse the error in the other direction. BERD procedure is illustrated in Fig. 3. Since error is diffused to the neighboring pixels by scanning the hologram row by row, the first and last columns, as well as the last row, cannot be processed. A column of zeros and a line was, therefore, inserted at the left, the right, and the bottom of the hologram matrix for the application of the algorithm, and removed from the final binary hologram at the end of the procedure.

The procedure is fast and enables to obtain a binary hologram that can be used to reconstruct the object in a very similar way to the one obtained with the original complex hologram.

3.2 Direct Binary Search

DBS algorithm is an iterative method where the complex optical field Im_{ref} obtained by propagating the complex hologram from a given distance z_0 is used as a reference. For the initial condition, a random binary pattern is generated and the complex optical field in the image plane Im_{bin} is computed by numerical propagation of this pattern. The difference between the two complex fields presenting $M \times N$ pixels is then quantified by computing their mean MSE.¹⁵ In order to preserve the 3-D information, the MSE was computed by considering the complex field and not the amplitude only. We used the following equation:

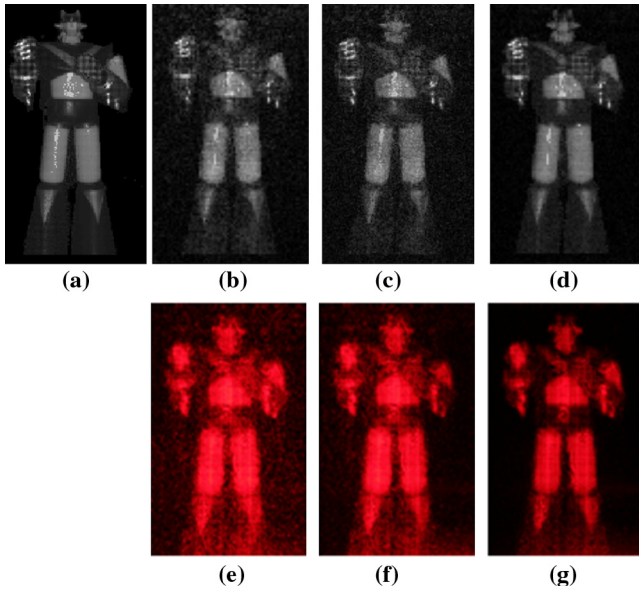


Fig. 4 (a) Numerical reconstructions of the original hologram, and numerical and optical reconstructions of the binary holograms generated with the (b)–(e) threshold method, (c)–(f) BERD, and DBS (d)–(g) algorithms.

$$\text{MSE} = \frac{1}{M \cdot N} \frac{\sum_{n=-\frac{M}{2}}^{\frac{M}{2}-1} \sum_{m=-\frac{N}{2}}^{\frac{N}{2}-1} |\text{Im}_{\text{ref}}(n, m) - k \cdot \text{Im}_{\text{bin}}(n, m)|^2}{\sum_{n=-\frac{M}{2}}^{\frac{M}{2}-1} \sum_{m=-\frac{N}{2}}^{\frac{N}{2}-1} |\text{Im}_{\text{ref}}(n, m)|^2}, \quad (2)$$

with

$$k = \frac{\sum_{n=-\frac{M}{2}}^{\frac{M}{2}-1} \sum_{m=-\frac{N}{2}}^{\frac{N}{2}-1} \text{Im}_{\text{ref}}(n, m) \cdot \text{conj}[\text{Im}_{\text{bin}}(n, m)]}{\sum_{n=-\frac{M}{2}}^{\frac{M}{2}-1} \sum_{m=-\frac{N}{2}}^{\frac{N}{2}-1} |\text{Im}_{\text{bin}}(n, m)|^2}.$$

In Eq. (2), (M, N) are the number of pixels of the image, Im_{ref} and Im_{bin} are the image fields obtained, respectively, from the reference and binary holograms, and $\text{conj}[\cdot]$ is the operator to take the conjugate. Note that the MSE is computed only on a given region of interest (ROI) of the image plane containing the object to decrease the computation time.

Every pixel of the binary pattern is toggled one by one (from 0 to 1 or from 1 to 0) and the updated image field is computed each time and compared with the reference. If the MSE is improved, the new value is kept for the binary pattern. Otherwise, the pixel is switched back to its original value. When all pixels have been tested once, the procedure is repeated from the first pixel for a new iteration. The loop stops when any change in the binary pattern induces a degradation of the MSE. In practice, no significant changes were observed in the MSE after five iterations.

This method not only gives the best result, but also offers a flexibility that cannot be found in noniterative approaches. Indeed, during the application of the algorithm, the field Im_{bin} is computed by propagation of the binary pattern. Parameters such as the reconstruction distance or the wavelength can be set differently from the original hologram in the propagation formula. Fourier and Fresnel binary holograms can both be generated,¹⁵ and the parameters can be modified to fit the desired configuration of the display

system. The main drawback is the long computation time despite various improvements that can be made to speed up the process.²¹ In order to lessen the computational burden, the best solution is to restrict the computation of the MSE between Im_{bin} and Im_{ref} to the ROI of the image where the scene is located.

4 Experimental Results

The performance of the different methods to convert the hologram into binary format was first evaluated by visual comparison of the reconstructed images. The depth map was set so that the values 0 and 255 corresponded, respectively, to a distance of 400 and 1000 mm. For this depth range, the torso of the robot was in focus for a reconstruction distance of 442 mm. Numerical reconstructions of the complex and binary holograms obtained by the threshold method, BERD, and DBS are shown in Fig. 4. Optical reconstructions were also performed with an LCoS (Syndiant Co., model SYL2061) presenting a resolution of 1024×600 pixels with a pitch of $9.4 \mu\text{m}$. The laser was emitting at 633 nm.

It is clear by observing both numerical and optical reconstructions that the best results are obtained with the DBS algorithm, at the price of long computation times. For the two noniterative methods, we see that the noise level is significantly lower in the image computed with the BERD approach.

In terms of computation time, the noniterative methods can potentially be considered for real-time applications since the binary patterns of 600×600 pixels were obtained in 0.21 s with the threshold method, and 0.23 s with the BERD algorithm. The computation time is more difficult to estimate with the DBS algorithm because it depends also on the size of the ROI selected to compute the MSE. In our case, we selected a region of 240×140 pixels centered on the robot. The computation time was around 8 h 20 min in total to complete 5 iterations. The hardware used was an Intel Core i7-3770 CPU with RAM 32 GB.

For a quantitative comparison, we computed the MSE between the optical fields reconstructed numerically from the binary holograms and the reference field reconstructed from the complex hologram. We examined the MSE in different focusing planes, and for two different depth ranges (400 to 450 mm and 400 to 1000 mm). The plane on which the torso of the robot was in focus was considered as the reference plane since it was the plane that was used as reference in the application of the DBS algorithm. The reference reconstruction distances were 403 and 442 mm, respectively, for the two depth ranges. Results are shown in Fig. 5.

Quantitative estimation of the MSE confirms the visual evaluation. DBS gives the best results and BERD shows better performance than the simple threshold method. Increasing the depth range in this case was equivalent to increasing the spatial gap between each layer. We see (Fig. 5) that when the depth range was lowered to 400 to 450 mm, performance of all algorithms was degraded. With the decrease of depth resolution, it was more difficult to reproduce the optical field accurately with a binary pattern.

It is interesting to see that for the DBS method, the MSE is optimal in the reference plane while the MSE is only gradually degrading in far field with the two other methods. This is expected because the principle of the DBS algorithm is to

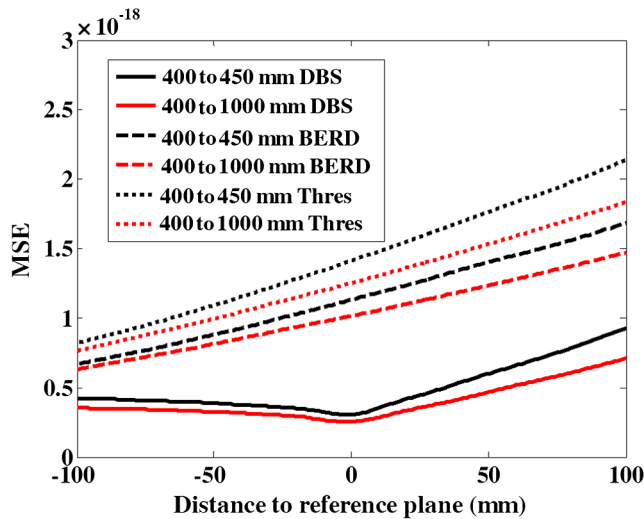


Fig. 5 Evolution of the MSE computed in different focusing planes for the different methods (DBS, BERD, and threshold).

minimize the MSE in a given reconstruction plane. For the BERD and threshold methods, there is no specific reference plane since the conversion into binary format is obtained directly from the values of the hologram.

5 Extension of DBS Algorithm to Multiple Reference Plane

The choice of reference plane in the DBS algorithm enables choosing a plane of interest, in which the reconstruction will be optimized. However, since a single plane is used as a reference in the original implementation of the algorithm, this may not be the best solution for deep scenes.

We propose to modify the DBS algorithm by using several reference planes instead of a single one. Each time a pixel is toggled in the binary pattern, reconstruction is performed in parallel at multiple depths, and MSE is computed between the images and the corresponding reference plane obtained from the complex hologram. Then, the switch in the pixel value of the binary pattern is kept only if the sum of the MSE computed in the different plane decreases. The principle of the modified algorithm is illustrated in Fig. 6 in the case in which two reference planes were used.

In order to compare the influence of the choice of reference plane on the reconstructed image, we applied the DBS algorithm several times on the same hologram (with depth range 400 to 1000 mm) with different configurations. Since the object was focused at 442 mm, we chose arbitrarily to examine the planes located at $\sim \pm 100$ mm to see if we could extend the performance of the DBS algorithm on a further depth range. For a single reference configuration, we used, respectively, the reconstruction distances of 350, 442, and 550 mm. For the double reference plane configuration, we used the planes 350 and 550 mm. The MSE, noted MSE_1 and MSE_2 in Fig. 6, was computed separately for the two reconstruction planes. We chose to use the same ROI for the two planes, a rectangular area of 240×140 pixels centered on the robot.

The use of multiple reference plane results in an increase of computation time. Then, we did not consider the case of more than two reference planes. However, since the DBS algorithm was implemented in an efficient way,²¹ the computation time was not doubled. The total computation time for five iterations was around 8 h 20 min with a single reference plane, and 10 h 50 min with two reference planes.

The performances were compared by computing the MSE between the image planes reconstructed from the different

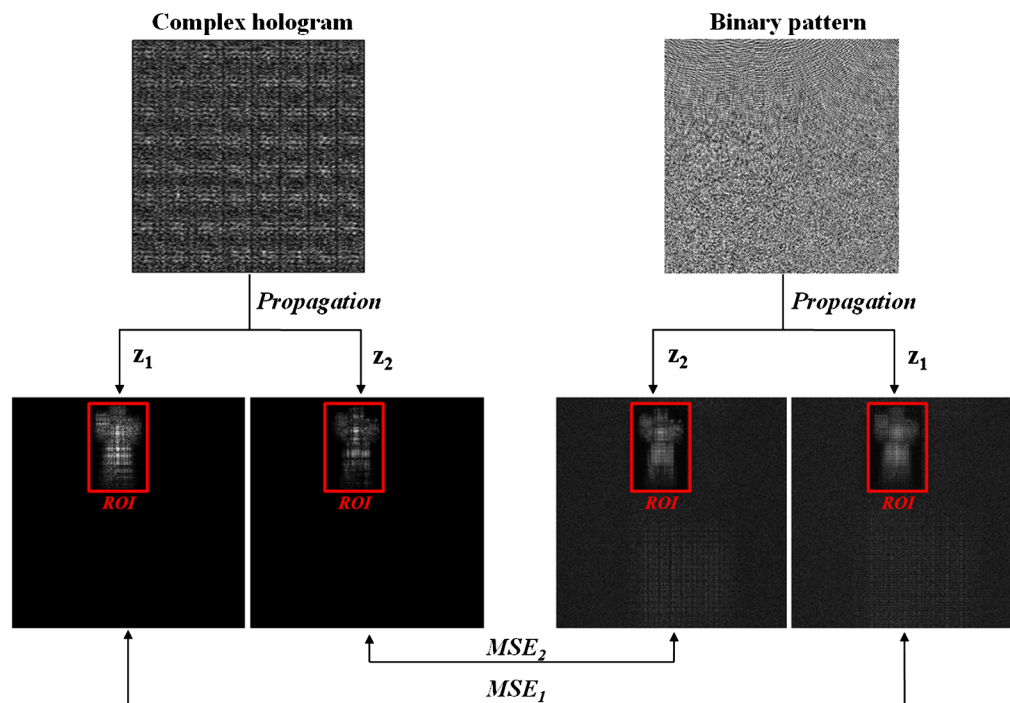


Fig. 6 Principle of the DBS algorithm with multiple reference planes. z_1 and z_2 correspond, respectively, to a propagation distance of 350 and 550 mm, whereas the object was focused at 442 mm.

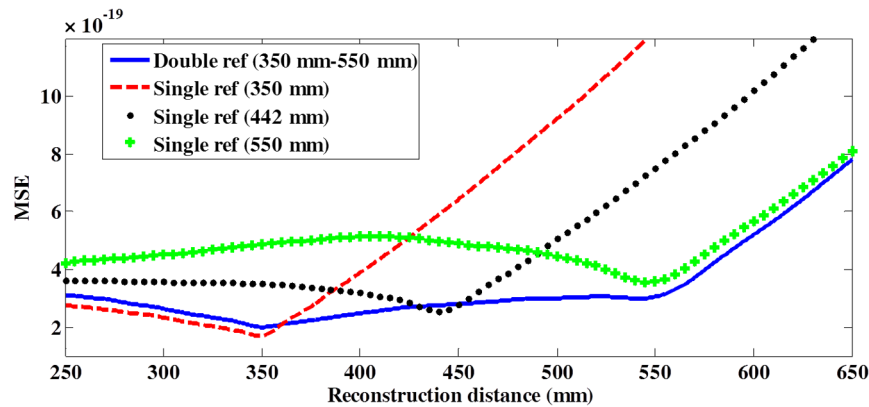


Fig. 7 Influence of the choice of reference plane on the performance of DBS algorithm.

binary holograms and the original hologram as a function of the reconstruction distance. The results are presented in Fig. 7.

We see that with a single reference plane, the best result is obtained at the desired reconstruction distance. However, the image quality is degraded out of the reference plane, especially at far distance. Using two different reference planes, the quality of the reconstruction obtained with the binary hologram can be enhanced on a further depth of field. It is, therefore, of high interest for the reconstruction of deep scenes. In future work, we intend to study further the influence of choice of reference planes to see how far we can extend the field of view. In addition, the change of pixel during the modified DBS procedure was kept in our experiment only if the sum of MSE_1 and MSE_2 was improved. Different criteria could be considered to improve further our proposed method.

6 Discussion and Conclusion

We demonstrated the generation of a binary hologram from a real scene using data captured from a Kinect camera system. The spatial resolution of the camera was low, but large objects could be captured. Since both intensity and depth map are available, the Kinect is well adapted to generate holograms in a fast way through the depth layer-based method. Different approaches can be adopted to convert the holograms into a binary format. It is not surprising to see that the iterative method with a long computation time showed the best performance. For fast computation, the BERD method that was demonstrated for the generation of phase-only holograms can be successfully adapted to the case of binary holograms and present results significantly better than the basic threshold method. We also investigated the influence of the reference planes in the application of the DBS algorithm. A single reference plane enables to optimize the binary pattern for a specific reconstruction distance, but a greater depth of field could be obtained by using several references during the procedure.

In future work, we intend to study further the DBS procedure by testing the use of different ROIs for the different reference planes. It can also be of high interest to exploit the flexibility of the DBS algorithm to investigate the generation of binary holograms for RGB display. Since parameters such as reconstruction distance and wavelength can be set

differently from the parameters of the original complex hologram, we expect to be able to reduce chromatic aberrations.

Acknowledgments

This work was supported by the Cross-Ministry Giga KOREA Project of the Ministry of Science, ICT, and Future Planning, Republic of Korea (ROK) [GK 16C0100, Development of Interactive and Realistic Massive Giga-Content Technology].

References

1. M. Lambouij et al., "Visual discomfort and visual fatigue of stereoscopic displays: a review," *J. Imaging Sci. Technol.* **53**(3), 030201 (2009).
2. S. Reichelt et al., "Depth cues in human visual perception and their realization in 3D displays," *Proc. SPIE* **7690**, 76900B (2010).
3. I. Yamaguchi and T. Zhang, "Phase-shifting digital holography," *Opt. Lett.* **22**(16), 1268–1270 (1997).
4. Y. S. Kim et al., "Speckle-free digital holographic recording of a diffusely reflecting object," *Opt. Express* **21**(7), 8183–8189 (2013).
5. R. H. Y. Chen and T. D. Wilkinson, "Computer generated hologram from point cloud using graphics processor," *Appl. Opt.* **48**(36), 6841–6850 (2009).
6. H. Niwase et al., "Real-time electroholography using a multiple-graphics processing unit cluster system with a single spatial light modulator and the InfiniBand network," *Opt. Eng.* **55**(9), 093108 (2016).
7. T. Shimobaba, N. Masuda, and T. Ito, "Simple and fast calculation algorithm for computer-generated hologram with wavefront recording plane," *Opt. Lett.* **34**(20), 3133–3135 (2009).
8. H. Nishi, K. Matsushima, and S. Nakahara, "Rendering of specular surfaces in polygon-based computer-generated holograms," *Appl. Opt.* **50**(34), H245–H252 (2011).
9. D. Im et al., "Accelerated synthesis algorithm of polygon computer-generated holograms," *Opt. Express* **23**(3), 2863–2871 (2015).
10. J. S. Chen, D. P. Chu, and Q. Smithwick, "Rapid hologram generation utilizing layer-based approach and graphic rendering for realistic three-dimensional image reconstruction by angular tiling," *J. Electron. Imaging* **23**(2), 023016 (2014).
11. J. S. Chen and D. P. Chu, "Improved layer-based method for rapid hologram generation and real-time interactive holographic display applications," *Opt. Express* **23**(14), 18143–18155 (2015).
12. S. Lee et al., "Digital hologram generation for a real 3D object using by a depth camera," *J. Phy. Conf. Ser.* **415**, 012049 (2012).
13. P. Tsang et al., "Computer generation of binary Fresnel holography," *App. Opt.* **50**(7), B88–B95 (2011).
14. P. W. M. Tsang and T. C. Poon, "Novel method for converting digital Fresnel hologram to phase-only hologram based on bidirectional error diffusion," *Opt. Express* **21**(20), 23680–23686 (2013).
15. T. Leportier et al., "Converting optical scanning holograms of real objects to binary Fourier holograms using an iterative direct binary search algorithm," *Opt. Express* **23**(3), 3403–3411 (2015).
16. J. P. Allebach, "DBS: retrospective and future directions," *Proc. SPIE* **4300**, 358–376 (2001).
17. D. Pagliari and L. Pinto, "Calibration of Kinect for Xbox one and comparison between the two generations of microsoft sensors," *Sensors* **15**(11), 27569–27589 (2015).

18. U. Schnars and W. P. O. Juptner, "Digital recording and numerical reconstruction of holograms," *Meas. Sci. Technol.* **13**(9), R85–R101 (2002).
19. J. P. Liu et al., "Complex Fresnel hologram display using a single SLM," *Appl. Opt.* **50**(34), H128–H135 (2011).
20. R. Tudela et al., "Full complex Fresnel holograms displayed on liquid crystal devices," *J. Opt. A: Pure Appl. Opt.* **5**(5), S189–S194 (2003).
21. B. K. Jennison, J. P. Allebach, and D. W. Sweeney, "Efficient design of direct-binary-search computer-generated holograms," *J. Opt. Soc. Am. A* **8**(4), 652–660 (1991).

Thibault Leportier received his BS and MS degrees in physics from the Institut d'Optique Graduate School, France, in 2010 and 2013,

respectively. He is a PhD candidate at the Korea University of Science and Technology (UST). His current research interests include holography, 3-D display, and spatial light modulators.

Min-Chul Park received his PhD degree in information and communication engineering from Tokyo University in 2000. He was an associate professor at Tokyo University in 2005. He is currently a principal research scientist at KIST and a professor at UST. His research focuses on 3-D image processing and display, 3-D human factors, and human–computer interaction. He is a member of SPIE.