Three-dimensional imaging with monocular cues using holographic stereography

Yi-Ying Pu,[1] Jian-Wen Dong,[1,2] Bing-Chu Chen,[1] Yuan-Zhi Liu,[1] and He-Zhou Wang[1,3]

[1]*State Key Laboratory of Optoelectronic Materials and Technologies, Sun Yat-sen (Zhongshan) University, 510275, Guangzhou, China*

[2]*e-mail: dongjwen@mail.sysu.edu.cn*

[3]*e-mail: stswhz@mail.sysu.edu.cn*

Received May 6, 2010; revised September 2, 2010; accepted September 3, 2010; posted September 10, 2010 (Doc. ID 128064); published September 28, 2010

Two quantitative criteria are derived to evaluate monocular cues in holographic stereograms. We find that the reconstruction has correct monocular cues when the whole scene is located in a so-called "monocular cues area" with compatible monocular and binocular cues. In contrast, incorrect monocular cues appear when the scene is in the other two areas, namely, the "visible multi-imaging area" and the "lacking information area." A pupil-function integral imaging algorithm is developed to simulate monocular observation, and a holographic printing system is set up to fabricate full-parallax holographic stereograms. Both simulations and experiments agree with the criteria.

OCIS codes: 090.0090, 090.2870, 100.0100.

Holographic stereography [1,2], which is a hybrid of holography and integral photography [3–5], is one of the promising technologies for three-dimensional (3D) display [6]. Considered as a convenient way to record dynamic 3D objects, holographic stereography has attracted extensive interest in the past decades [7,8]. Nevertheless, it is controversial whether the holographic stereogram (HS) produces natural 3D images, because it excludes information about depth and interference between neighboring points in a scene [9,10].

Human beings perceive 3D images using both physiological and psychological cues [11]. The physiological cues include binocular cues (such as convergence and stereopsis) and monocular cues (MCs) (such as accommodation and monocular movement parallax). Both binocular cues and MCs are available in daily 3D scenes. Most HSs provide correct binocular cues but incorrect MCs because of the absence of some phase information. This discrepancy sometimes causes visual fatigue [12]. Recently, several qualitative strategies have been proposed to evaluate MC effects in multiview displays [9,10,13,14]. For example, a group from Japan proposed phase-added stereograms that converged to Fresnel holograms [9], and an MIT group generated their panoramagram using controllable "wafels" [10]. However, quantitative studies and experimental demonstrations are still needed to verify the presence of MCs in the HS.

In this Letter, three kinds of regions on both sides of the HS are defined by two quantitative criteria explicitly, for the first time to the best of our knowledge. The lacking information (LI) area is next to the HS, and its size depends on the parameters of the HS. The MC and visible multi-imaging (VMI) areas depend not only on the parameters of the HS but also on the locations of the objects. The correct MCs appear when the whole scene falls in the MCs area. The criteria have been validated by a pupil-function integral imaging algorithm and optical experiment of a full-parallax HS with correct MC effects.

In a hologram, a sharp point image is observed when the eye focuses on point P [Fig. 1(a)]. Otherwise, a blurry spot is observed. So the depth perception can be estimate by naked eye. In an HS fabrication setup, there are usually a lens generating parallel beams and a beam-limit pinhole [see Fig. 4(a), for example]. The waves, radiated from a point source, become a plane wave after the lens and then a bundle of rays after the pinhole. When the eye focuses on point P [Fig. 1(b)], the image point has the same size as the pinhole (as well as the hogel). When the eye defocuses, multiple spots will be observed due to the discretization of hogels. It will cause an incorrect blur gradient and may give an incorrect depth perception. However, if the distance between two adjacent spots is less than the eye resolution, a big blurry spot can be seen, which may enhance depth perception again.

Next, the quantitative criteria are derived to evaluate MCs in the HS. Consider two adjacent hogels closed to a scene point P [Fig. 1(c)]; we have

$$\theta = 2\arctan\frac{\gamma|z_P - z_F| - h}{2z_F}. \tag{1}$$

Here $\gamma = h/|d_H - z_P|$; h is the size of the hogel; and z_F, z_P, and d_H are the distances from the eye to the focus plane,

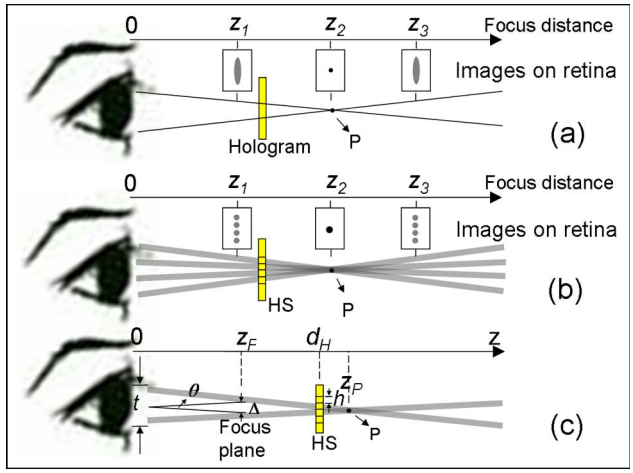

Fig. 1. (Color online) Images on retina when the eye focuses on different distances through (a) a hologram and (b) an HS. (c) Schematic diagram for the derivation of Eqs. (2) and (3).

point P, and the HS, respectively. Note that Eq. (1) is applicable when point P is on either side of the HS. When two adjacent points are undistinguishable, the angle θ is less than the lateral resolution of the naked eye (for example, 1.5′, [15]). Hence, the range of z_F is given by

$$\frac{z_P\gamma/2-1}{\gamma/2+\tan 0.75'} < z_F < \frac{z_P\gamma/2+1}{\gamma/2-\tan 0.75'}. \quad (2)$$

Moreover, in order to confirm the existence of MCs, at least two ray bundles should enter the eye. So the lateral distance t in front of the eye should be less than the pupil diameter (for example, $d_{\text{pup}} = 4$ mm [16]), yielding

$$t = z_P\gamma + h \le d_{\text{pup}}. \quad (3)$$

Equations (2) and (3) are the *quantitative* criteria to determine the properties (such as locations and sizes) of the three areas. If Eq. (3) is not satisfied, the object is located in the LI area. The number of ray bundles entering the eye is insufficient, and a distinct (sometimes discrete) image is observed wherever the eye focuses. The MC area is defined when both equations are satisfied. If the eye focuses in the MC area, it has compatible MCs and binocular cues. Furthermore, if Eq. (3) is satisfied but Eq. (2) fails, it is the VMI area of the HS. The eye perceives multiple parallax images due to the wide separation among these images.

Figure 2(a) shows three kinds of areas on both sides of an HS ($d_H = 500$ mm). Each object outside of the LI area has an MC area, in which the correct MCs are obtained. When there is more than one object in the 3D scene, one should put each object inside the intersection of the MC areas of all the objects, so that the correct MCs are obtained wherever the eye focuses in the scene. As a result, the MC area of the whole scene is no more than that of each object [see, e.g., light-blue square c in Fig. 2(b)]. We also find that the LI area can be diminished by reducing d_H [Fig. 2(b)] or h (results not shown). In principle, the LI area vanishes as h goes to zero. However, making extremely small h is challenging in the fabrication. In addition, the space–bandwidth product of a hogel should be large enough to satisfy the sampling theorem. Therefore, appropriate parameters need to be chosen so that one can fabricate HS with better MCs.

To demonstrate MCs, a method based on integral imaging is developed to simulate what the eye sees through an HS. Since the eye cannot see the entire elemental image (EI), a spatial-dependent pupil function P_{ij} is introduced [see Fig. 3(a)]:

$$P_{ij}(x,y,z_F) = \begin{cases} 1, & (x-x_{oi})^2+(y-y_{oj})^2 < a \\ 0, & \text{otherwise} \end{cases}. \quad (4)$$

Here, $a = (z_F - d_H)d_{\text{pup}}/2d_H$, $x_{oi} = x_e + z_F(x_{ui}-x_e)/d_H$, and $y_{oj} = y_e + z_F(y_{uj}-y_e)/d_H$; (x_e, y_e) and (x_{ui}, y_{uj}) are the central coordinates of the eye pupil and the (i,j)th hogel, respectively; and (x,y) is the Cartesian coordinate of the focus plane. Then the intensity on the retina can be calculated by summation and average of the EIs:

$$I(x,y,z_F) = \frac{1}{N_s(x,y,z_F)}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y} M[E_{ij}(x,y)]P_{ij}(x,y,z_F), \quad (5)$$

where E_{ij} is the unmagnified EI, $M[\cdot]$ is the magnified EI, and N_s is the superposition number of each pixel for the magnified EI in the pupil. N_x (N_y) is the number of EIs that enter the eye in the $x(y)$ direction. Here, the EIs are generated by a ray-tracing method from a 3D mesh model obtained by using an image-based modeling algorithm [17] or by a virtual computer-graphic model.

Our 3D scene contains five chess pieces. The king in the front (the bishop in the back) is at $z_F = 338(403)$ mm. The parameters of the HS are the same as those in Fig. 2, so that the whole scene falls in the MC area [light-blue square c in Fig. 2(b)]. The simulation images focusing on the king [Fig. 3(b)] and the bishop [Fig. 3(c)] are

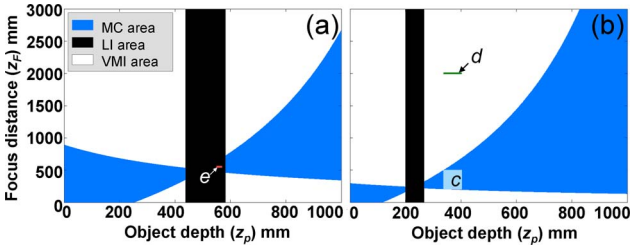

Fig. 2. (Color online) MC (light blue), LI (black), and VMI (white) areas when (a) $d_H = 500$ mm and (b) $d_H = 230$ mm. Here, $h = 0.5$ mm. The light-blue square c, green bar d, and the light-red bar e correspond to the 3D scenes in Figs. 3(b), 3(d), and 3(e).

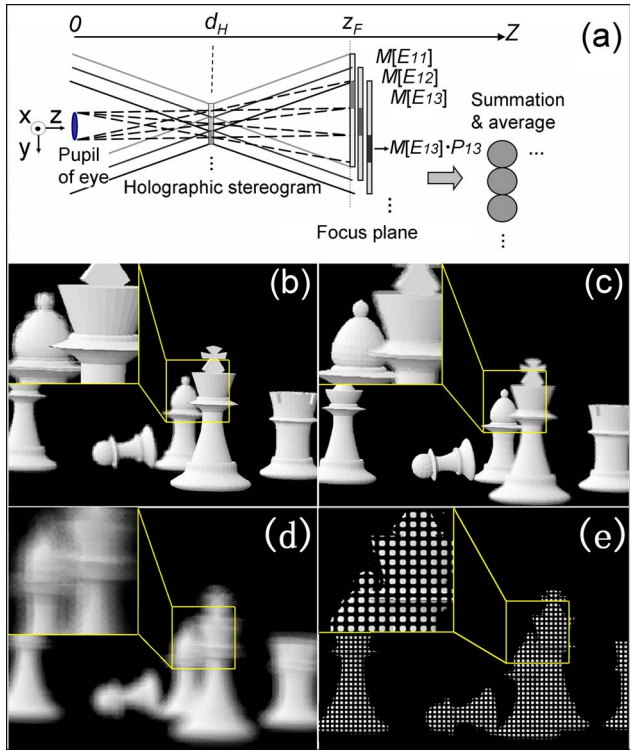

Fig. 3. (Color online) (a) Schematic diagram of the pupil-function integral imaging algorithm. (b)–(e) are four simulated images on the retina. (b) and (c) show correct MC effects when the king or the bishop is in focus, respectively. (d) shows the VMI effect, while (e) shows the LI effect.

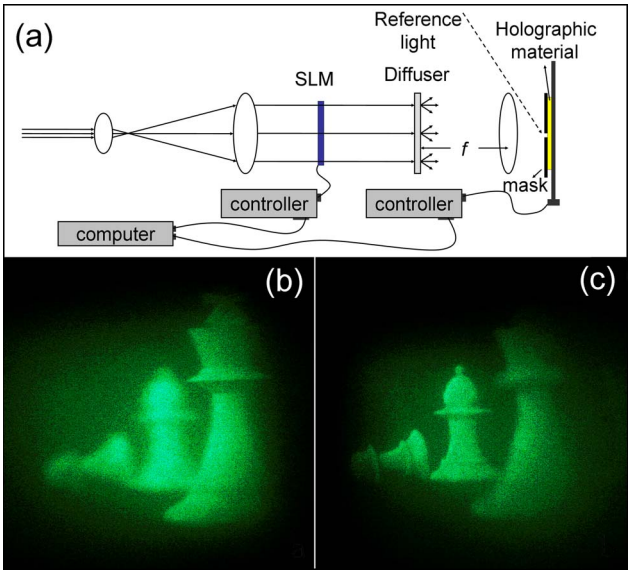

Fig. 4. (Color online) (a) Experimental setup of the HS printing system. Optical reconstruction images when the eye focuses on (b) the king and (c) the bishop. The parameters of the 3D scene and the HS are the same as those in Fig. 3(b).

presented. The bishop becomes out of focus when the king is in focus and vice versa. This defocusing effect further enhances depth sensation and produces the correct MC effect. However, when the eye focuses on the VMI area, one has a VMI effect instead of a blur effect [e.g., the horizontal edge line on the neck of the king becomes two visible lines in Fig. 3(d)]. Note that such an effect will be substantial if the object has high-spatial-frequency textures. As a result, if the scene depth is so large that some objects are located outside of the MC area, the VMI effect makes the reconstructed image appear unnatural. In addition, Fig. 3(e) shows the discrete patterns in the reconstruction when the scene [light-red bar *e* in Fig. 2(a)] falls in the LI area.

Figure 4(a) shows the HS printing system for the full-parallax HS fabrication. A diffuser is placed on the front focal plane of the lens. The HS is close to a mask with a 0.5 mm × 0.5 mm pinhole. The distance between the pinhole and the lens is less than a focal length. The laser beam (532 nm), spatial light modulator (SLM), and step motor are simultaneously controlled by a computer. In the recording process, a serial EI is uploaded sequentially onto the SLM, while the HS is translated step by step. The same scene and parameters as those in Fig. 3(b) are used. The complete HS is 50 mm × 50 mm with 100 × 100 hogels. A Nikon D6 digital camera is used to take photos focusing on different depths. Two representative pictures are shown in Figs. 4(b) and 4(c) with focal distances of 340 and 400 mm, respectively, which correspond to the cases in Figs. 3(b) and 3(c). The focal length and aperture of the camera lens are 52 mm and *f*/5, respectively. A sharp edge of the front king appears when it is in focus. As the focal distance varies, the in-focus and out-of-focus effects change smoothly and there is no VMI or discrete patterns.

In conclusion, MCs in an HS are studied. The quantitative criteria are derived to distinguish the MC area. It enables us to choose parameters to ensure that the reconstruction has the correct MCs. The pupil-function integral imaging algorithm and the optical experiment are used to demonstrate the criteria. In addition, our integral imaging algorithm is also a potential evaluation method in other multiparallax 3D displays.

This work is supported by the National Natural Science Foundation of China (NSFC) (10804131, 10874250, 11074311), the Fundamental Research Funds for the Central Universities (2009300003161450), and the Guangdong Natural Science Foundation.

References

1. R. V. Pole, Appl. Phys. Lett. **10**, 20 (1967).
2. S. A. Benton and V. M. Bove, *Holographic Imaging* (Wiley, 2008), Chap. 20.
3. G. Lippmann, Comptes Rendus **146**, 446 (1908).
4. M. Cho and B. Javidi, Opt. Lett. **33**, 2737 (2008).
5. I. Moon and B. Javidi, Opt. Express **16**, 13080 (2008).
6. K. Sato, S. Koizumi, K. Chou, and K. Takano, Proc. SPIE **6488**, 64880V (2007).
7. J. G. Wu, M. Conry, C. H. Gu, F. Wang, Z. Yaqoob, and C. H. Yang, Opt. Lett. **31**, 1265 (2006).
8. M. W. Halle, "Multiple viewpoint rendering for 3-dimensional displays," Ph.D. dissertation (Massachusetts Institute of Technology, 1997).
9. M. Yamaguchi, H. Hoshino, T. Honda, and N. Ohyama, Proc. SPIE **1914**, 25 (1993).
10. Q. Y. J. Smithwick, J. Barabas, D. E. Smalley, and V. M. Bove, Proc. SPIE **7619**, 761903 (2010).
11. V. Bruce, P. R. Green, and M. A. Georgeson, *Visual Perception: Physiology, Psychology, and Ecology* (Psychology Press, 2003), Chap. 7.
12. D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, J. Vis. **8**, 33 (2008).
13. H. Kang, T. Yamaguchi, and H. Yoshikawa, Appl. Opt. **47**, D44 (2008).
14. S. K. Kim, D. W. Kim, Y. M. Kwon, and J. Y. Son, Opt. Express **16**, 21415 (2008).
15. A. S. Patel, J. Opt. Soc. Am. **56**, 689 (1966).
16. B. Winn, D. Whitaker, D. B. Elliott, and N. J. Phillips, Invest. Ophthalmol. Visual Sci. **35**, 1132 (1994).
17. Y. Z. Liu, J. W. Dong, Y. Y. Pu, B. C. Chen, H. X. He, and H. Z. Wang, Opt. Express **18**, 3345 (2010).