Time-multiplexed dual-focal plane head-mounted display with a liquid lens

Sheng Liu and Hong Hua*

3D Visualization and Imaging Systems Laboratory, College of Optical Sciences, University of Arizona, 1630 East University Boulevard, Tucson, Arizona 85721, USA

**Corresponding author: hhua@optics.arizona.edu*

Received November 18, 2008; revised March 1, 2009; accepted April 17, 2009; posted May 4, 2008 (Doc. ID 103869); published May 21, 2009

Focus cues play a significant role in three-dimensional (3-D) depth perception. Conventional stereoscopic displays, however, lack the ability to correctly render these cues, because they present a pair of stereoscopic images on a fixed image plane while forcing the eyes to converge at different distances to view objects at different depths. Using a fast liquid-lens device, we present the design and implementation of a time-multiplexed dual-focal plane display that is capable of rendering correct or near-correct focus cues as well as other depth cues, such as occlusion and shading for a 3-D scene. The focus range of the dual focal planes can vary from infinity to as close as 8 diopters. Two driving mechanisms are proposed to render near-correct focus cues and their effects on image sharpness, flicker, and brightness are compared. © 2009 Optical Society of America

OCIS codes: 120.2040, 120.2820, 120.4570, 220.1080, 330.1400, 330.7322.

Focus cues (accommodation and retinal blur) of a 3-D scene play a significant role in 3-D depth perception [1]. Conventional stereoscopic displays, however, lack the ability to correctly render these cues, as they present a pair of stereoscopic images on a fixed image plane while forcing the eyes to converge at different distances to perceive depth. These displays thus force an unnatural decoupling of the accommodation and convergence cues and induce incorrect retinal blurs that do not vary with the rendered depth of a virtual scene. Many researchers have carried out studies to investigate the adverse effects of such unnatural viewing conditions [1–4]. Their results suggest that unfaithful focus cues are associated with a variety of visual artifacts, such as the underestimation of the rendered depth of a 3-D scene and visual fatigue after a prolonged exposure to stereoscopic environments.

There have been emerging research interests in developing 3-D displays with correct or near-correct focus cues. Classified as one of the true 3-D displays, a volumetric display can render correct focus cues by portraying millions of voxels within a physical volume [5,6]. The practical development of such technology, however, faces many technical challenges, such as low computational efficiency, limited rendering volume, and difficulty of correctly rendering view-dependent lighting effects such as occlusion, specular reflection, and shading. Multifocal plane displays render near-correct focus cues for virtual objects in different depths by presenting perspective images on multiple discrete focal planes. Each of the focal planes is responsible for rendering 3-D objects with a depth range centered around it, and these discrete focal planes together render a volume of 3-D objects with near-correct focus cues specific to a given viewpoint. Rolland *et al.* presented a theoretical framework of a spatially multiplexed multifocal plane head-mounted display (HMD) by using a stack of 14 planar displays as the image source [7], and Akeley *et al.* recently demonstrated a three-focal plane display prototype by optically splitting a large liquid-crystal display panel into three zones located at different depths [8]. An alternative to the spatially multiplexed scheme is the time-multiplexed approach by the usage of active optical elements, in which 3-D objects of different depths are rendered sequentially on a single 2-D focal plane whose focal distance is adjusted in synchronization with the depth of the objects being rendered. If a high enough rate can be achieved, all objects would appear to have near-correct focus cues without visible flicker. Schowengerdt and Seibel [9] demonstrated a dual-focal plane retinal scanning display using a deformable membrane mirror (DMM) device. In this design, a 2-D image is generated on a pixel-by-pixel basis by raster scanning a laser beam across the visual field, and the focus cues of each pixel are rendered by defocusing the laser beam through the DMM.

Our approach to a time-multiplexed multi-focal-plane display is achieved by modulating the focal distance of a 2-D display time-sequentially with a liquid lens [10]. It requires no moving parts for rendering focus cues and uses conventional microdisplay and graphics hardware. Mainly owing to the low speed of the liquid-lens device, the preliminary results in a dual-focal plane prototype failed to demonstrate faithful focus cues at a frame rate higher than 5 Hz [10]. We have now adopted a faster liquid lens (Arctic 314, Varioptic, Inc.), with a response speed of about 9 ms, to increase the frame rate of a dual-focal plane display to up to 37.5 Hz. We have also explored an alternative rendering mechanism that reduces artifacts and improves depth cue accuracy. Moreover, in order to compensate the reduced clear aperture from 3 mm (Arctic 320) to 2.5 mm (Arctic 314), we have made modifications to the original design [10]. As shown in Fig. 1(a), the liquid lens is offset from the center of the radius curvature O of the spherical mirror by Δ; thus the exit pupil of the HMD is magnified by $m_p = R/(R+2\Delta)$ to the size of the clear aperture of the liquid lens. The focal distance is specified by the

0146-9592/09/111642-3/$15.00

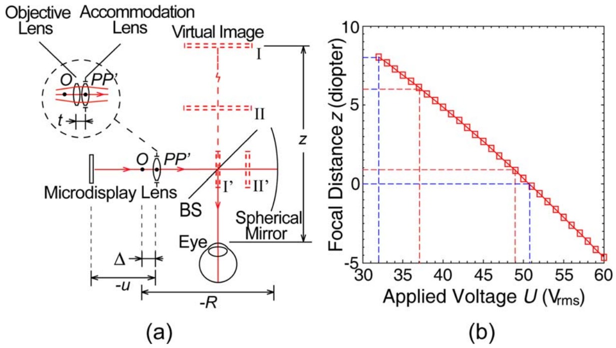

Fig. 1. (Color online) (a) Schematic design of a time-multiplexed dual-focal plane display. (b) Simulation of the focal distance versus the applied voltage of the liquid lens.

distance z from the virtual image to the exit pupil of the HMD, which is given as

$$z = \frac{-R(u+\Delta+u\Delta\varphi)}{2(u+\Delta+u\Delta\varphi)+R(1+u\varphi)} + \frac{\Delta R}{R+2\Delta}. \quad (1)$$

For a detailed description of the parameters in Eq. (1), please refer to [10].

Based on the schematic analysis, we built a proof-of-concept monocular prototype using off-the-shelf components. The accommodation lens is a fast liquid lens (Arctic314) which varies optical power from −5 to 20 diopters with the application of an ac voltage from 32 V rms to 60 V rms [11]. The other optical parts are the same as those in our previous implementation [10]. The spacing t between the singlet objective and the liquid lens is 6 mm, the offset Δ is 6 mm, and the object distance $-u$ is 34 mm. As a result, the prototype demonstrates a 24° diagonal field of view with an exit pupil of 3 mm.

Given the dependence of the optical power φ_A upon the voltage U applied to the liquid lens [11], Fig. 1(b) plots the focal distance z as a function of the voltage U applied to the liquid lens derived from Eq. (1). To create a flicker-free appearance of 3-D objects rendered sequentially on multiple focal planes, the speed requirements for the liquid lens, the microdisplay, and the graphics card are proportional to the number of focal planes [10]. Limited by the 75 Hz frame rate of the graphics hardware, our current prototype is feasible for dual-focal planes at up to 37.5 Hz. As suggested by the black (blue online) dashed lines in Fig. 1(b), the dual-focal planes can be positioned as far as 0 diopter or as close as 8 diopters to the viewer by applying ac voltages between 51 V rms and 32 V rms. To clearly demonstrate focus cues, the front and back focal planes in the prototype were setup at 6 diopters and 1 diopter, respectively, with a large dioptric spacing of 5 diopters. It is worth mentioning that depth-weighted blending algorithms may be implemented to render 3D objects between two adjacent focal planes, whereas less dioptric spacing may be preferred for rendering pseudocorrect focus cues [4,8].

As illustrated in Fig. 2(a), the liquid lens is driven by a square wave, with a period T, fast-switching between 49 V rms and 37 V rms to temporally multiplex the focal planes at 1 diopter and 6 diopters. In

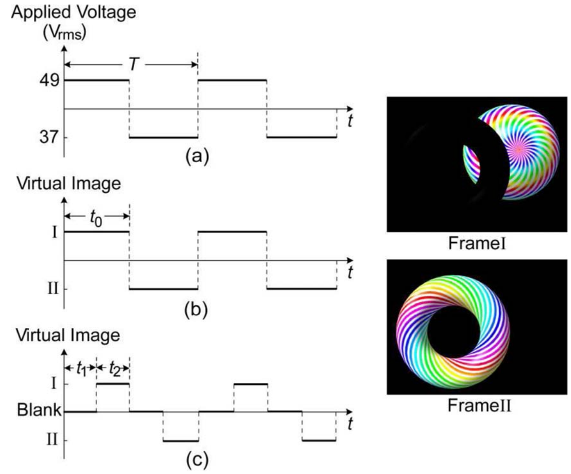

Fig. 2. (Color online) (a) Driving mechanism of the liquid lens, (b) dual-state image rendering mechanism, and (c) tristate image rendering mechanism for a time-multiplexed dual-focal plane display. The right insets show the image frame I of the occluded sphere on the far focal plane and the image frame II of the torus on the near focal plane.

synchronization with the liquid lens, two frames of images (I and II), corresponding to far and near objects, respectively, are rendered sequentially as shown in Fig. 2(b). Correct occlusion can be portrayed by creating a stencil mask of the near objects in Frame II, superimposed with the far objects in Frame I. In this rendering method, the duration t_0 of both the far and near frames is one half of the period T. Limited by the 75 Hz frame rate of the graphics card, the minimum value of t_0 is 13.3 ms, and the highest refresh rate of the dual-focal plane display is $f=1/(2t_0)=37.5$ Hz to complete the rendering of both far- and near-focal states.

By adopting the driving mechanism in Figs. 2(a) and 2(b), Figs. 3(a) and 3(b) demonstrate the experimental results of the dual-focal plane display at 37.5 Hz. Three real bar-type resolution targets were placed at 6 diopters (large size), 3 diopters (middle size), and 1 diopter (small size) to visualize the transition of focus cues from near to far distances. Photographs of the dual-focal plane display were taken by a camera mounted at the eye location with an exposure of 0.5 s and a speed of $f/3.2$. Two virtual objects, a sphere and a torus, were rendered sequentially at 1 diopter and 6 diopters, respectively. As shown in Fig. 3(a), when the camera was focused on the bar target at 6 diopters, the torus appears to be in focus, while the sphere shows noticeable out-of-focus blurring. Figure 3(b) demonstrates the case when the camera was focused on the sphere at 1 diopter. In the accompanying video (Media 1), the virtual objects were animated in such a way that they both move along the visual axis at a constant speed from either 6 diopters to 1 diopter or vice versa, synchronized with the voltage applied to the liquid lens. Although it shows some minor noticeable flicker, the video clearly demonstrates correct correspondence of focus cues for the two virtual objects matching with the focus setting change of the camera.

In the current prototype, as the response speed of the liquid lens is about 9 ms [11], we expect longitu-

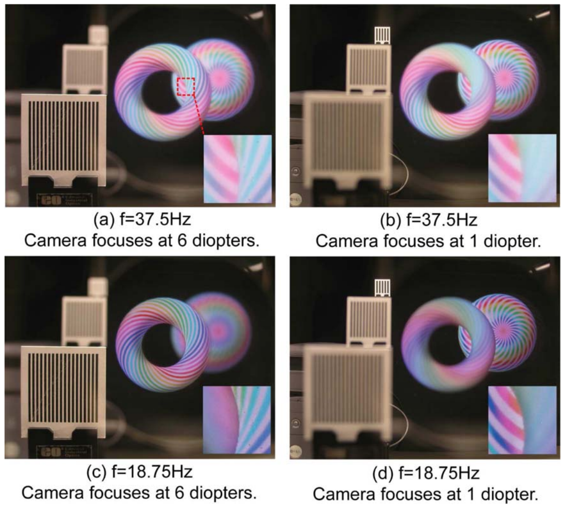

Fig. 3. (Color online) Photographs of the dual-focal plane display (a), (b) with the speed of f=37.5 Hz (Media 1) and (c), (d) of f=18.75 Hz (Media 2). The camera was focused at 6 diopters in (a) and (c), and at 1 diopter in (b) and (d). Insets, zoomed-in images at the boundary of the front torus and back sphere.

dinal shifts of the focal planes during the settling time of the liquid lens when its driving signal is switched between the two states. As a result, the rendering mechanism shown in Fig. 2(b) may yield not only minor image blur but also less accurate depth representations. Such artifacts are noticeable in the insets of Figs. 3(a) and 3(b), which shows the zoomed-in image of the area marked by dashed box in Fig. 3(a). The supposedly in-focus virtual objects (i.e., the torus and the sphere in Figs. 3(a) and 3(b), respectively) appear to be slightly blurred, and the supposedly out-of-focus objects (i.e., the sphere and the torus, respectively) appear to be less blurred than their real counterparts. Liquid lenses with a faster response speed can reduce such artifacts and render more accurate focus cues at high speed.

Alternatively, we experimented with another mechanism for image rendering. As shown in Fig. 2(c), a blank frame, with a duration t_1, is inserted to lead the rendering of each actual image frame whose duration is reduced to $t_2=t_0-t_1$ to maintain in synchronization with the liquid lens. Limited by the 75 Hz graphics card, the minimum value for both t_1 and t_2 is 13.3 ms, and the highest refresh rate of the dual-focal plane display is $f=1/(2t_1+2t_2)=18.75$ Hz. Figures 3(c) and 3(d) demonstrate the experimental results by adopting the rendering mechanism in Fig. 2(c). Compared with Figs. 3(a) and 3(b), the front torus is in sharp focus only when the camera was focused at 6 diopters, as shown in Fig. 3(c), and vice versa for the back sphere at 1 diopter, as shown in Fig. 3(d). The insets of Figs. 3(c) and 3(d) clearly demonstrated the improved focus cues. Meanwhile, the occlusion cue becomes more prominent with a sharper boundary between the near torus and far sphere. Owing to the shortened duration of image frames, one compromise, however, is the decrease of the relative brightness level quantified by $B=t_2/(t_1+t_2)$. Knowing $t_1=t_2=13.3$ ms, the relative brightness level in Figs. 3(c) and 3(d) is $B=0.5$, half of that in Figs. 3(a) and 3(b) with $B=1$. Another compromise is the increased flicker, which is more noticeable in the video demonstration at 18.75 Hz (Media 2) than the one at 37.5 Hz.

For future developments, a faster liquid lens along with higher speed graphics hardware and displays is beneficial for producing accurate focus cues at a flicker-free rate. Researchers also proposed overshooting methods to drive the liquid lens with decreased time-to-depth-of-field in an auto-focusing imaging system [12]. Other active optical technologies, such as high-speed DMM and liquid-crystal lenses, should also be considered to increase the frame rate of a time-multiplexed dual-focal plane display.

In this Letter, by adopting a fast liquid-lens device, we demonstrated a proof-of-concept time-multiplexed dual-focal plane display with the capability of rendering near-correct focus cues as well as other depth cues such as occlusion and shading. The focal distance of the dual-focal plane display can be presented within a wide range, from infinity to as close as 8 diopters. We further compared the effects of two rendering mechanisms: The first method yields a higher refresh rate (e.g., $f=37.5$ Hz) and brighter image ($B=1.0$) but reduced image sharpness and focus cue accuracy, and the second method produces sharper images and more accurate focus cues but with compromised speed (e.g. $f=18.75$ Hz) and image brightness ($B=0.5$). In a preliminary usability study [10], we observed matching accommodative responses of the viewer's eye to the addressable focal planes of the proposed system operated at a vari-focal mode. The results suggest potentially improved depth perception, compared with conventional stereoscopic displays with a single focal plane.

This work was funded by National Science Foundation (NSF) awards 05-34777 and 06-44446.

References

1. S. J. Watt, K. Akeley, M. O. Ernst, and M. S. Banks, J. Vision **5**, 834 (2005).
2. J. P. Wann, S. Rushton, and M. Mon-Williams, Vision Res. **35**, 2731 (1995).
3. T. Inoue and H. Ohzu, Appl. Opt. **36**, 4509 (1997).
4. D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, J. Vision **8**, 1 (2008).
5. A. Sullivan, SID Int. Symp. Digest Tech. Papers **34**, 1531 (2003).
6. G. E. Favalora, J. Napoli, D. M. Hall, R. K. Dorval, M. G. Giovinco, M. J. Richmond, and W. S. Chun, Proc. SPIE **4712**, 300 (2002).
7. J. P. Rolland, M. W. Krueger, and A. Goon, Appl. Opt. **39**, 3209 (2000).
8. K. Akeley, S. J. Watt, A. R. Girshick, and M. S. Banks, ACM Trans. Graphics **23**, 804 (2004).
9. B. T. Schowengerdt and E. J. Seibel, J. Soc. Inf. Disp. **14**, 135 (2006).
10. S. Liu, D. W. Cheng, and H. Hua, in 7th IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR, 2008), pp. 33–42.
11. http://www.varioptic.com.
12. O. Jacques-Sermet, Photonics Spectra, November 2007, pp. 60–62.