Dynamic focusing in head-mounted displays

Jannick P. Rolland[a], Myron W. Krueger[b], and Alexei A. Goon[a]

[a]School of Optics, Univ. of Central Florida, Orlando FL 32816-2700

[b]Artificial Reality Corporation, Vernon CT 06066

ABSTRACT

In stereoscopic virtual environment systems, vergence eye movements are required but the absence of the need to accommodate is not consistent with real-world vision. Ideally, virtual objects would be displayed at the appropriate distances from the viewer and natural, concordant accommodation and vergence would be required. Based on optical principles and human vision, we investigate the feasibility of a novel display based on multiple focal plane arrays to provide these cues. We then briefly discuss some design approaches to focusing at multiple depth planes.

Key words: stereoscopic head-mounted displays, accommodation and vergence, multiple depth-planes focusing.

1. INTRODUCTION

Existing head-mounted displays (HMDs) are focused at a fixed distance. Perhaps surprisingly, the vast majority of deployed virtual reality systems present the same images to both eyes. Such *biocular* systems require neither a change in accommodation nor vergence. In elite systems that can afford two separate graphics generator and thus a distinct image for each eye, vergence eye movements are required but the absence of the need to accommodate is not consistent with real-world vision. Such systems, referred to as stereoscopic HMDs, are investigated in this paper. Thereafter, they are simply referred to as HMDs.

Current HMDs are limited by a combination of low resolution, narrow field of view, and poor ergonomic designs. A comprehensive discussion of trade-offs in designing HMDs is given in Kocian (1988).[1] Even the highest resolution HMDs, however, would not yield sharp 3D objects at multiple depth of presentation as a result of existing conflicts between accommodation and vergence. Thus, while we shall continue the quest for sharper displays and more ergonomic designs, it is necessary to also ask whether conflicts of accommodation and vergence can be overcome.[2-3]

Conflicts of accommodation and vergence in HMDs results from the inability to provide realistic accommodation cues.[3-6] While accommodation is a weak cue in itself for the perception of depth compared to other cues (e.g. occlusion, head motion parallax, stereopsis),[7] when we look around in the real world, it is not all in focus at once.[8-9] In an earlier paper, we highlighted potential problems that HMDs could place upon the accommodation/vergence system.[4] Studies of binocular stress have followed.[6] Thus, the requirement for accommodation as well as for vergence is a necessary ingredient for the synthetic representation of realistic scenes.

No existing HMD provides the viewer with the capability to change accommodation while viewing objects at different distances. Images are typically presented on flat surfaces (i.e. screens)--either on CRTs or on liquid crystal displays. The optics is designed to image the screens at a sufficient distance from a user in order for the user to focus on the screens' optical images. In order for the imaging optic to form virtual images, the displays are positioned either at the object focal point or between the first lens and the object focal point of the imaging optics. In the former case, the virtual images are collimated. In the latter case, the virtual images are located between the near point of accommodation or punctum proximum (i.e. 250 mm) and the furthest point of accommodation or punctum remotum (i.e. infinity) of the user.

With HMDs for pilots, for example, collimated imagery is typically preferred because of the need to visualize information in the far field. In contrast, augmented reality systems for medical or engineering applications are likely to require visualization in the near field to allow manipulation of objects. In this case, uncollimated virtual images

Part of the IS&T/SPIE Conference on The Engineering Reality
of Virtual Reality 1999 • San Jose, California • January 1999
SPIE Vol. 3639 • 0277-786X/99/$10.00

are preferred to minimize conflicts between accommodation and vergence as well as for maximizing the accuracy of rendered depth in systems with no eye-tracking capability.[10-11]

While the expedient of a fixed focal distance can be made to work successfully for various applications, many real tasks require constant switching between different focal distances. For example, operating a car requires navigating in the far field, looking for targets that may be located either in the far or near field, and simultaneously attending to instrumentation in the near field. Similarly, as we walk around a scene we are constantly moving our eyes around it looking out for obstacles which may be at different distances. We suppose that in some extreme cases such as athletics, out-of-focus mode may in fact be a common mode of operation. Athletes may never have time to converge and accommodate on particular objects. There are other tasks such as the close examination of three-dimensional objects that may involve, to the contrary, many small focus adjustments.

In designing technology, we must ask both whether and to what extent perturbations in a fundamental mode of operation does affect task performance and how a failure to provide a natural visual experience may cause difficulties for the visual system over a long term exposure. For a system with no capability to account for the human visual system natural accommodation, instant focus is combined with natural vergence delays.[9-12] A lack of vergence in a HMD would lead to diplopic images (i.e. double vision).[13-14] In general, a lack of delay and out-of-adjustment viewing as the visual system adapts is unrealistic and will definitely lead to an unnatural viewing, possibly eye strain, and some form of accommodation discomfort as well.[5-6] Given that current virtual reality exposure is measured in minutes rather than hours, there is no reason to be alarmist in the short run. However, if we contemplate a future in which virtual reality is a ubiquitous feature of video games, such viewing conditions may conceivably cause changes to the visual system over years of use. Therefore, it is incumbent on us to design displays that incorporate these cues as realistically as possible to ensure safety as well as to see if the viewing experience can be made more believable.

In order to provide natural accommodation cues in HMDs we propose a focal solution. The general idea is not novel. Marran and Schor have previously outlined various focal solutions for virtual reality systems.[15] Mon-Williams et al. suggest varying focal depth either by using an oscillating lens or by adjusting the image depth based on the user's gaze point.[3] The closest work to what we propose is that by Dolgoff who proposed a display device that provides two planes of accommodation, one for the foreground and one for the background.[16]

We propose a novel approach to the problem, however, by suggesting to make the HMD miniature displays themselves multi-focal. This can be achieved using an array of focal planes upon which the pixels would be presented according to the distance of the simulated object. The approach is attractive in that it does not require any moving elements or eyetracking capability. The approach comes from observing that in HMD optical systems where the transverse magnification is greater than one, a small movement of the actual screens towards the viewer's eye or away from it leads to a large displacement in depth of the optical images. The axial magnification in optical systems varies indeed as the square of the transverse magnification. Therefore, it is possible to envision a thick display system that could adjust the depth of the displayed objects on a plane-by-plane basis. The principle of operation of a multiple planes array for focusing in depth is shown in Fig.1.

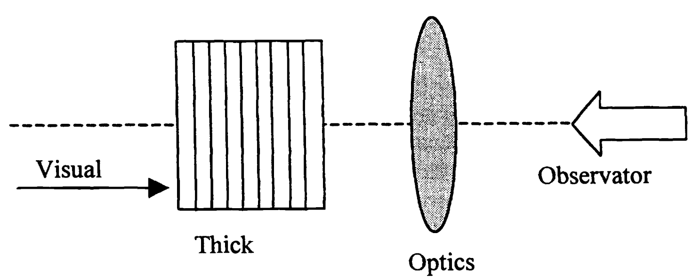

Fig. 1. Principle of multi-focus planes HMD.

Many implementation approaches are possible using variations on current or easily foreseeable technology. Before looking at specific options, we present a simple, yet comprehensive investigation of the feasibility and system requirements for providing accommodation in HMDs through multi-focal plane arrays in the same manner that is available in the real world. We first derive the relationship between the focal length f of the imaging optics, the punctum proximum Lm of accommodation, and the range of focusing dx of the miniature display near the focal-point object of the optics. We then provide a framework to compute how many planes are required as well as the interplane spacing requirements within a range of focusing corresponding to accommodation spanning from infinity to Lm. Given a focusing capability, we then establish the display resolution requirement as a function of stereo-depth acuity. Finally, we present conceptual ideas of how the multi-planes focusing can be achieved with no moving parts and no need for an eyetracking capability.

2. RANGE OF MULTI-PLANES FOCUSING

We first establish what thickness dx of the multi-planes display is needed to cover a given range of accommodation.

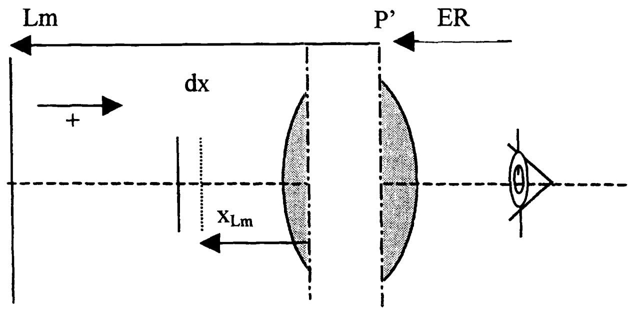

Fig. 2. Basic layout of the imaging optics in a head-mounted display.

Let's consider the imaging equation given by

$$\frac{n'}{x'} = \frac{n}{x} - \frac{n}{f} \quad , \tag{1}$$

where n and n' are the indices of refraction in object and image space of the imaging optics, respectively; x and x' are the distances of the object (i.e. the miniature display) and the image (i.e. the virtual image plane) with respect to the principal planes P and P', respectively; and f is the focal length of the imaging optics. Let's denote as x_{Lm} the value of x that corresponds to an image located at the punctum proximum Lm. To focus, x varies from f to x_{Lm} and x' varies accordingly from infinity to L_m-ER where ER is the eyerelief measured from P' for simplicity and generality as shown in Fig. 2. The focal plane in object space is the reference plane from which the range of focusing dx is computed. By manipulation of Eq.1, dx is given by

$$dx = x_{Lm} - f = -\frac{f^2}{f + L_m - ER} \quad . \tag{2}$$

If P' is located close to the last optical surface of the optics, |ER| equal 25 mm will allow wear of a wide variety of eye glasses. The shorter the focal length and the larger the value of L_m, the less ranging is required. The range of focusing dx ranges from about 0.2 mm to 26 mm given focal lengths between 15 mm and 90 mm. 3D focal plane arrays of such sizes can currently be implemented in various materials.

3. NUMBER OF PLANES FOR COMPLETE RANGE FOCUSING

The minimum number of planes required for focusing from a nearest plane to infinity is determined by the available range of accommodation and the depth-of-focus of the human visual system on each side of a plane of fixation. The depth of focus dL, a function of the visual acuity η, the size d of the pupil of the user's eye, and the distance L of accommodation, is given by

$$| dL_{\pm} | = | \frac{\eta \cdot L^2}{d \pm \eta \cdot L} | \quad , \tag{3}$$

where dL_+ and dL_- denote the distal and the proximal depth of focus with respect to the fixation plane, respectively. Note that L is negative in Eq.2, yielding $|dL_+| > |dL_-|$ as also observed in Fig. 3. Based on a value of visual acuity of one arcmin, Eq. 3 and Eq. 2 combined yield 27 planes from infinity to 0.5m. A few planes are represented in the figure. Based on this theoretical prediction, 14 planes only would be needed because every other plane can be selected as a plane of accommodation.

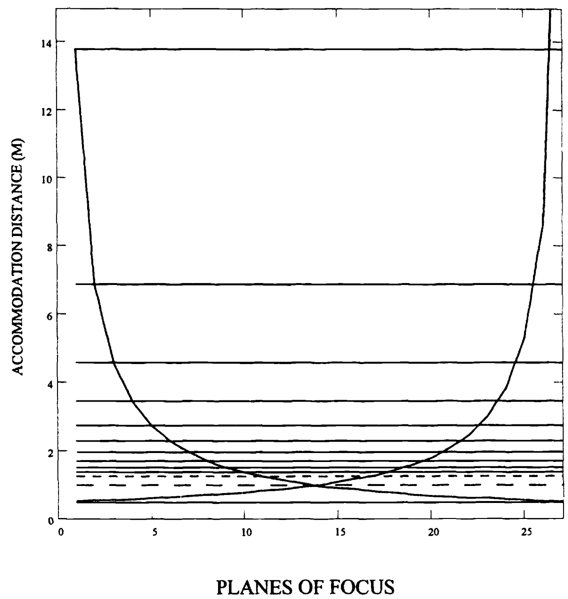

Fig. 3. Location of the planes of fixation for accommodation based on the computed depth of focus planes for the human visual system. A visual acuity of 1arcmin is assumed. Both schemes starting from L_m equal 0.5 m to infinity (i.e. red solid curve) and from infinity to 0.5 m (i.e. solid blue curve) were considered. In both cases, we find that a total number of 27 planes is required for a range of accommodation from infinity to 0.5 m. A few of these planes are represented as horizontal lines in the figure.

4. PLANES INTERSPACING

Using Eq.1 the values of L are mapped back in object space to dx values. The interplane spacing is given by the consecutive differences of the computed dx values. The interplane spacing as a function of the plane number is shown in Fig. 4 for three values of the focal length 15 mm, 50 mm, and 90 mm. The interspacing is quasi-constant and a constant spacing chosen to be that computed for plane number one can serve as a practical solution.

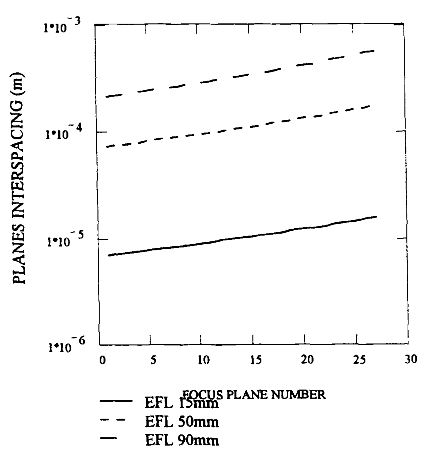

Fig. 4. Planes' interspacing as a function of the focus plane number for three different values of the focal length.

While the smallest interplane spacing is about 10 microns for an extremely short focal length, typical values are an order of magnitude larger. Interspacing in the order of 100 microns is more readily feasible.

5. RESOLUTION REQUIREMENT OF THE DISPLAYS

We now ask how many depth units can be resolved within the range of accommodation imposed by the depth of focus of the human visual system around a fixation plane. The number of resolvable units sets requirements for the resolution of the displays. A HMD user has stereoscopic information available from the disparate images provided to the two eyes. Binocular disparity can be defined as the angular disparity η between any two object points in the field of view.[17] Simple geometry yields an expression for η_+ behind the fixation plane (i.e. distal) and η_- in front of the fixation plane (i.e. proximal) given by

$$\eta_{\mp} = \mp \frac{\Delta l \cdot i}{(L \pm \Delta l) L} \quad , \tag{4}$$

where i is the interocular distance, and Δl is the resolvable depth at a given fixation distance L taken to be as negative following our sign convention shown in Fig. 2. Given a value of η, L, and i, similarly to that computed for accommodation, we solve for Δl in Eq. 4 to yield Eq. 5. Δl_+ and Δl_- on the distal and proximal side, respectively, are given by

$$|\Delta l_{\pm}| = \frac{|\eta| \cdot L^2}{i \pm |\eta \cdot| L} \quad . \tag{5}$$

Also note that according to sign conventions, η is positive and negative on the distal and the proximal sides, respectively. Thresholds for stereoacuity η vary widely between users in the extreme between about 2 arcsec and 130 arcsecond.[18] Both extreme values are considered in addition to a more typical value of 30 arcsec. It has been shown that η is constant over distance.[17] We shall next consider the distal value of Δl_+ in computing the resolvable depth units. L is recursively given by

$$L_k = L_{k-1} + \Delta l_{k-1} = L_{k-1} + \frac{|\eta| \cdot L^2{}_{k-1}}{i + |\eta| \cdot L_{k-1}} \quad . \tag{6}$$

For computational purpose, the interocular distance i is set to an average value of 65 mm. Using Eqs. (5) and (6), the number of distances N resolved in depth as a function of η for three values of observer interocular distance is shown in Fig. 5. This number is constant for all planes of accommodation previously computed and these resolvable distances in depth are also found to be equally spaced in the space of the display device.

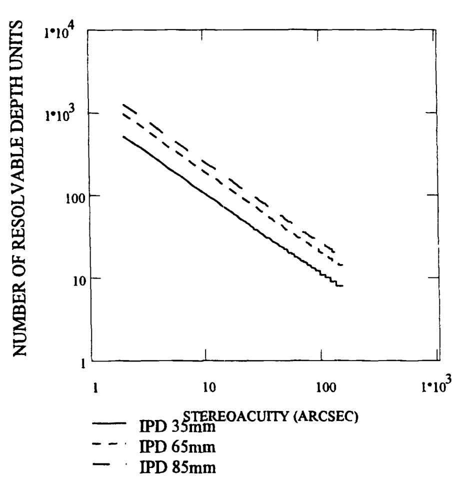

Fig.5. Number of resolvable depth units as a function of stereoacuity for three values of the user interocular distance.

The number of resolvable depth units between the discrete set of planes of accommodation shown in Fig. 3 imposes some requirement on the resolution of the display device. Based on the smallest resolvable depth around any fixation plane, the required display resolution p is given by

$$p = \frac{i}{2(L + \Delta l)(1 + \frac{L}{f})} \left(\frac{|\eta| L^2}{i - |\eta| L} \right) \quad . \tag{7}$$

The values of p as a function of the accommodation distance for three focal lengths at a stereo acuity of 30 arcsec, as well as a function of stereoacuity for a focal length of 50 mm are shown in Fig. 6 (a) and (b), respectively. Except at the highest resolution of 2 arcsec or for a short focal length (e.g. 15 mm) where the resolution requirement is quite stringent (i.e. one micon), a display resolution of about 5 microns is required. Such resolutions are achievable with current technology.

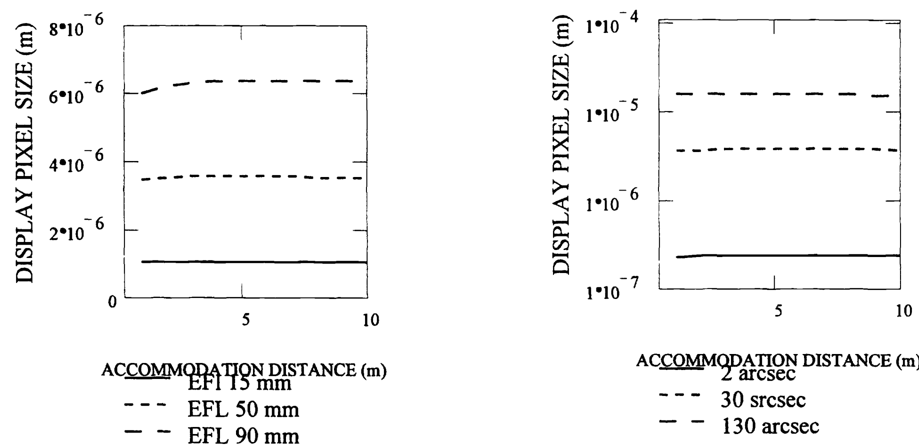

Fig. 6. Display resolution requirements: (a) Plot of the display resolution required as a function of the distance of accommodation L for three values of the focal length: 15 mm, 50 mm, and 90 mm. The stereoacuity threshold $\eta$is 30 arcsec in this computation. (b) Plot of the display resolution required as a function of the distance of accommodation L for three values of the stereo acuity threshold η: 2 arcsec, 30 arcsec and 130 arcsec; In this case the focal length is set to 50 mm.

6. POSSIBLE DESIGN APPROACHES TO MULTI-PLANES FOCUSING

Multi-planes focusing can be achieved by laminating a number of flat panel displays. We propose that this approach could be used in a HMD to provide ranging. In this case, each pixel would be displayed on only one of the planes: the one determined by its location in depth from the viewer. While it might seem that scanning multiple displays would present a problem, this is not so. As part of the calculation of a rendered computer-graphics image, a Z-buffer is created which has a distance value for each pixel currently displayed. Instead of ignoring this construct during the display process, it can be used to determine which of the display planes each pixel is written to as it is scanned. In all other planes, the visually aligned pixels would remain transparent.

Alternatively, the laminated planes could be simple binary devices that would switch between transparent and opaque in sequence. The pixels appropriate to each plane would be projected onto it when it was opaque. Recently, Allan Sullivan from Dimensional Media Associates has proposed a volumetric display for medical image display with 16 laminated planes.[19] Sullivan planned to use a single laser to illuminate the pixels on each plane in turn. This use of laser scanning would necessarily limit the resolution. However, spatial light modulators could provide the display speed if not the contrast ratio.

Another approach is the use of a true volumetric display using for example a homogeneous crystal, gas, or even liquid and a three dimensional scanning scheme.[20-22] Elizabeth Downing used intersecting laser beams to excite dye molecules at points in a small volume.[21] While such vector scanning is slow, faster techniques can be imagined. For instance, it would be possible to scan in one dimension with a sheet of laser light and then to use a number of laser beams to scan the illuminated sheet in x and y in parallel.

Regardless of the specific implementation, it is necessary to provide realistic accommodation and vergence cues if we expect virtual reality display systems to improve to the point that the question of whether a scene is real or virtual becomes an interesting one.

7. CONCLUSION

In order to resolve conflict of vergence and accommodation in HMDs, we propose to make the miniature displays multi-focal. We presented a theoretical feasibility study to add multi-planes focusing capability to head-mounted displays. Under a range of HMD parameters considered, the range of focusing to accommodate from infinity to 0.5m goes from about 0.2mm to 26mm. The minimum number of planes within this range is 14 for a standard visual acuity of 1 arcmin and a 4 mm pupil diameter. While 14 is the minimum theoretical value, it remains an experimental question of investigation how many planes can be adopted for various applications. The framework we layed out will allow to compute the number of planes required. Under most stringent conditions imposed by the theoretical study, the interplane spacing is found to be constant and may be as small as 10 microns but more typically about 100 microns. Finally, stereoacuity imposes that the transversal resolution of the display be in the order of 5 microns. Based on this investigation, we conclude that adding multi-planes focusing to HMDs may be challenging but nevertheless realizable with today's technology.

ACKNOWLEDGMENTS

We thank Edward Rinulducci at the University of Central Florida, and Bernard Peuchot at the University of Blaise Pascal in Clermont Ferrand, France, for stimulating discussion about this investigation. This work is supported in part by a First Award 1-R29-LM06322-01A1 from the National Institute of Health, The Office of Naval Research grant N00014-97-1-0654, and DARPA grant DAMD17-94-J-4113.

REFERENCES

1. D.F. Kocian, "Design considerations for virtual panoramic display (VPD) helmet systems," Armstrong Aerospace Medical Research Laboratory, Visual Display Systems Branch, Wright Patterson Airforce Base, Dayton, Ohio 45433-6573 (1988).
2. C. Roumes, J. Plantier, and J.P. Menu, "L'avenir est-il au visual de casque binoculaire?" Proc. of AGARD 517, 11-1:11-9 (1992).
3. J.Wann, S. Rushton, and M. Mon-Williams, "Natural problems in the perception of virtual environments," Vision Research, 35, 2731-2736 (1995).
4. W. Robinett, and J.P. Rolland, "A computational model for the stereoscopic optics of a head-mounted display,"" *Presence: Teleoperators and Virtual Environments* 1(1), 45-62 (1992).
5. E. Peli, "Real vision and virtual reality," *Optics and Photonics News*, July, 28-34 (1995).
6. M. Mon-Williams, J.P. Wann, and S. Rushton, "Binocular vision in a virtual world: visual deficits following the wearing of head-mounted displays," Ophthal. Physiol. Opt. 13, 387-391 (1993).
7. H.A. Sedgwick. Space Perception, in Handbook of Perception and Human Performance **21**, 43 (1986).
8. W.C Gogel, and J.D. Tietz, "Relative cues and absolute distance perception," Perception and Psychophysics, 28, 321-328 (1980).
9. G. Westheimer, The eye as an optical instrument, in Handbook of Perception and Human Performance **4**, 10 (1986).
10. J.P. Rolland, W.Gibson, and D. Ariely, "Towards quantifying depth and size perception in virtual environments," *Presence: Teleoperators and Virtual Environments* 4(1), 24-29 (1995).
11. L.Vaissie, J.P. Rolland, and G.M. Bochenek, "Study of eyepoint location in head-mounted displays," Technical Report TR98-001, University of Central Florida (1998).
12. P.E. Hallett, eye movements, Handbook of Perception and Human Performance **10**, 18 (1986).
13. A. Arditi, Binocular Vision, in Handbook of Perception and Human Performance **23**, 32 (1986).
14. J. Siderov, and R.S. Harwerth, "Precision of stereoscopic depth perception from double images," Vision Res. 33(11), 1553-1560 (1993).

15. L. Marran and C.Schor, "Multiaccommodative stimuli in VR systems: Problems and solutions," Human Factors, **37**, 382-388 (1997).
16. E. Dolgoff, "Real-depth imaging," SID'97 Digest, 269-272 (1997).
17. R.W. Reading, *Binocular Vision: Foundations and Applications*, 110-128 (1983).
18. A. Lit, and j.p. Finn, "Variablity of depth-discrimination thresholds as a function of observation distance," J. Opt. Soc.of Am. A. 66(7), 740-742 (1976).
19. A. Sullivan, DARPA Biomedical Computing Workshop, IMAGE Conference, Phoenix (1997).
20. J. D. Lewis, C.M. Verber, and R.B. MacGHee, "A true three-dimensional display," IEEE Trans. on Electron Devices 18(9), 724-732 (1971).
21. E. Downing, L. Hesselink, J. Ralston, and R. Macfarlane, "A three-color, solid-state, three-dimensional display," Science **273**, 1185-1189 (1996).
22. I. I. Kim, E. Korevar, and H. Hakakha, "Three-dimensional volumetric display in rubidium vapor," *Proc. of SPIE* 2650, 274-284 (1996).