Depth Distortion in Color-Interlaced Stereoscopic 3D Displays

Joohwan Kim, Paul V. Johnson, Martin S. Banks*
University of California, Berkeley, Vision Science Program, 360 Minor Hall, Berkeley CA, USA

ABSTRACT

In the conventional temporally interlaced S3D protocol, red, green, and blue are presented simultaneously to one eye and then to the other eye. Thus, images are presented in alternating fashion to the two eyes. Moving objects presented with this protocol are often perceived at incorrect depth relative to stationary parts of the scene. We implemented a color-interlaced protocol that could in principle minimize or even eliminate such depth distortions. In this protocol, green is presented to one eye and red and blue to the other eye at the same time. Then red and blue are presented to the first eye and green to the second. Using a stereoscope, we emulated the color-interlaced protocol and measured the magnitude of depth distortions as a function of object speed. The results showed that color interlacing yields smaller depth distortions than temporal interlacing in most cases and never yields larger distortions. Indeed, when color interlacing produces no brightness change within sub-frames, the distortions are eliminated altogether. The results also show that the visual system's calculation of depth from disparity is based on luminance, not chromatic signals. In conclusion, color interlacing provides great potential for improved stereo presentation.

1. INTRODUCTION

There are many ways to present stereoscopic 3D (S3D) imagery. One common protocol uses temporal interlacing to present full-resolution images to both eyes: full RGB images are presented to the left and right eyes in alternation. But this protocol yields depth distortions in which moving objects are seen at incorrect depths relative to stationary scenes. The cause is the temporal delay of the image in one eye relative to the other eye; the delay is interpreted by the visual system as a spatial disparity. The change in perceived depth, which can be quite distracting, has been called the Mach-Dvorak effect [1]. Researchers have recently proposed a variant that we will call *color interlacing* [2-4]. As illustrated in Figure 1, color-interlaced displays present G to the left eye and R and B to the right eye at one time and exchange the colors at another time. Thus, color interlacing keeps the overall luminance presented to an eye relatively constant over time. If the visual system estimates depth from disparity based on luminance information, not chromatic information [5], this protocol should reduce or even eliminate depth distortions.

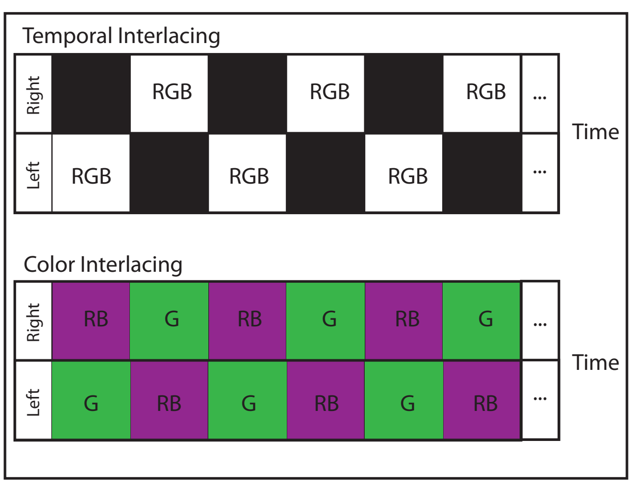

Figure 1. Temporal interlacing vs color interlacing. The upper panel represents conventional temporal interlacing and the lower panel color interlacing. In temporal interlacing, full RGB images are presented alternately to the two eyes. In color interlacing, the green channel (G) is presented to the left eye at the same moment that the red and blue channels (R and B) are presented to the right eye. The channels are reversed in the subsequent sub-frame.

Stereoscopic Displays and Applications XXIV, edited by Andrew J. Woods, Nicolas S. Holliman, Gregg E. Favalora,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 8648, 86480N · © 2013 SPIE-IS&T
CCC code: 0277-786X/13/$18 · doi: 10.1117/12.2008455

Figure 2 explains how depth distortions occur in temporal interlacing and how color interlacing could minimize such distortions. The left and right halves illustrate temporal and color interlacing, respectively.

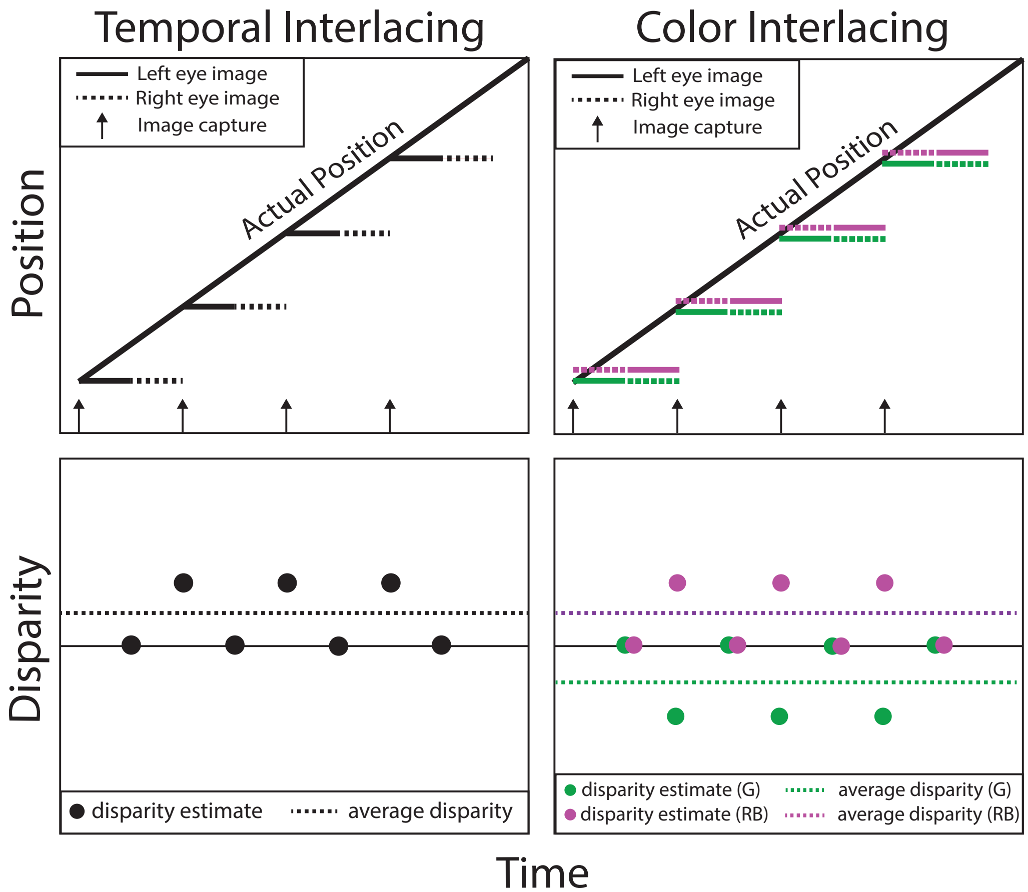

Figure 2. Depth distortions in temporal and color interlacing. The left and right sides illustrate temporal and color interlacing, respectively. The upper panels show the horizontal position of the displayed image as a function of time. A white object is moving rightward at constant speed and has zero disparity. The arrows indicate the times at which the left and right images are captured or generated. The horizontal line segments represent the sequence of image presentations in a sample-and-hold display, such as an LCD: solid lines for the left eye and dashed lines for the right. The lower panels show the disparity estimates over time. Each frame has two sub-frames, so two estimates are generated per frame. The horizontal lines represent the correct disparities of zero and the dashed lines represent the time-averaged disparities. With temporal interlacing, the average is displaced in one direction from the correct value. With color interlacing, the average is displaced in one direction for green (G) and the opposite direction for magenta (R + B).

The upper row shows how an object moving at constant horizontal speed is displayed under these protocols. Horizontal position on the screen is plotted as a function of time. The diagonal line represents the intended motion of the object. The arrows indicate the times at which the images were captured (or computer-generated). The captured or computer-generated object has zero disparity, so it should be perceived in the plane of the screen. With a sample-and-hold display (such as an LCD), the displayed images are a sequence of horizontal lines: solid for the left eye and dashed for the right. In temporal interlacing, the left-eye image (R, G, and B) appears first and the right-eye image (again, R, G, and B) second. Note that the right-eye's image is therefore delayed compared to what would happen in the real world. In color interlacing, the G component of the left-eye image appears first and at the same time as the R and B components of the right-eye image. Then the R and B components of the left-eye image appear at the same time as the G component of the right image.

The bottom row shows the disparities that would be derived from those image sequences. Estimated disparity is plotted as a function of time. The correct disparity (zero) is represented by the solid horizontal line. The dots represent the moment-to-moment disparity estimates created by matching images between the left and right eyes. Each stereoscopic frame is composed of two sub-frames, so there are two dots per frame. The disparity estimate is correct (zero) whenever the visual system matches two sub-frames from the same stereoscopic frame. But the visual system has no way of knowing where a frame ends or begins, so matches will also be made from sub-frames that are from different frames. This produces incorrect disparity estimates. In temporal interlacing, the incorrect estimates are all in the same direction (upward in this example) and this yields an error in the average disparity over time (represented by the dashed horizontal line). The average disparity for an object moving at horizontal speed *s* and captured at frequency *c* is:

$$d = \frac{s}{2c} \tag{1}$$

This equation predicts depth distortions accurately until the object speed is rather high [6].

In color interlacing, the erroneous disparity estimates are in opposite directions for green (G) and magenta (R + B), so we obtain two predictions in opposite directions. The visual system decodes retinal images into a luminance signal (basically, brightness) and two chromatic signals (basically, red-green and blue-yellow) [7]. If the system calculates disparity separately in the two chromatic channels, we expect two depth distortions in opposite directions; i.e., a green object that is seen as too near and a magenta object seen as too far. Alternatively, if the system calculates disparity in the luminance channel, depth distortions should be reduced because the luminances of the green and magenta stimuli are similar, which creates minimal luminance modulation and therefore minimal error in perceived depth.

We can determine how the visual system calculates disparity with respect to the luminance and chromatic channels by using color and temporal interlacing and measuring depth distortions with stimuli of different colors. Desaturated stimuli (e.g., gray) consist of green and magenta of about the same luminance, so if disparity is measured in luminance channels, depth distortion should be greatly reduced when such stimuli are presented with color interlacing as opposed to temporal interlacing. If disparity is measured in chromatic channels, however, we should observe depth distortions simultaneously in opposite directions with desaturated stimuli; this would be akin to color breakup but in depth rather than in position on the display screen. Saturated stimuli (e.g., green or red) are presented in similar fashion in color and temporal interlacing so depth distortions should be similar for the two types of interlacing. We measured depth distortion using a variety of saturations in order to determine whether disparity is calculated in luminance or chromatic channels and whether color interlacing reduces depth distortions relative to temporal interlacing.

2. EXPERIMENT 1: BRIGHTNESS MATCHES FOR GREEN AND MAGENTA

We first used a psychophysical technique—heterochromatic flicker photometry—to determine the luminances of green (G) and magenta (R + B) that are perceived as equally bright.

2.1 Apparatus

The apparatus was a mirror stereoscope. It consisted of a CRT (Viewsonic G225f) viewed stereoscopically via four mirrors. The left half of the CRT screen displayed the image for the left eye and the right half displayed the image for the right eye, so viewers saw a virtual stereoscopic image. The optical viewing distance was 1.18m, display resolution was 1280x960 (angular size of one pixel was 1minarc), and refresh rate was 120Hz. The field of view of the stereoscopic images was 10.7 x 16º.

2.2 Methods

We used heterochromatic flicker photometry [8] to find the luminances of green (G) and magenta (R + B) that were equally bright for each subject. Luminance channels respond to higher temporal frequencies than chromatic channels [9]. Heterochromatic flicker photometry capitalizes on this property by presenting two stimuli (in our case, G and R+B) alternately at a high rate. The relative luminances of the two colored stimuli are adjusted until minimal flicker is perceived. At this point, the two stimuli have the same brightness. Figure 3 illustrates the green and magenta stimuli over time. The average luminances of green and magenta are L_G and L_M, respectively. Both are temporally modulated with a contrast of C_C. The maximum and minimum luminances for green are therefore $L_G(1+C_C)$ and $L_G(1-C_C)$, and those for magenta are $L_M(1+C_C)$ and $L_M(1-C_C)$.

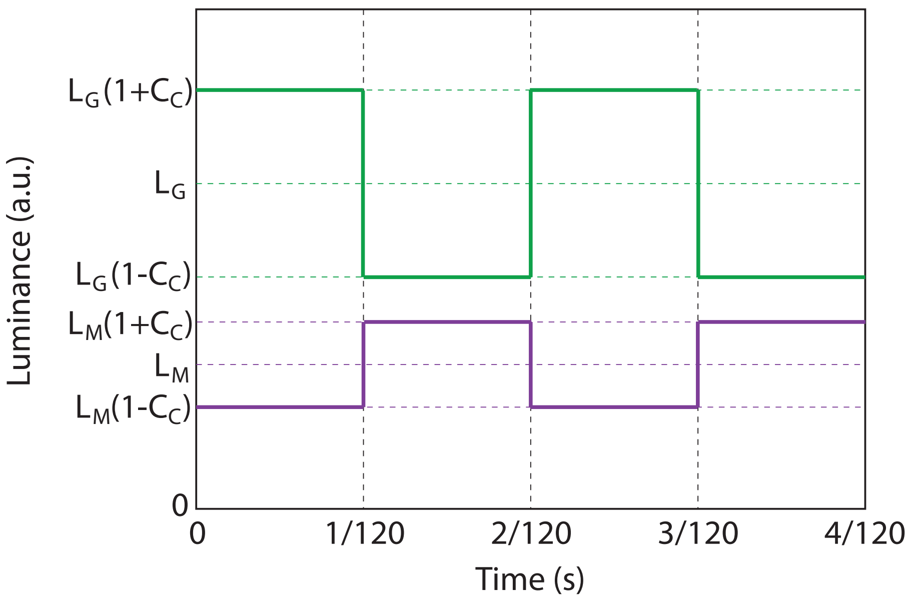

Figure 3. The green and magenta stimuli over time in heterochromatic flicker photometry. The contrasts of both stimuli are C_C. The frame rate is 120Hz resulting in a heterochromatic flicker rate of 60Hz. The two colors are presented in opposite phase. Here the green stimulus has higher luminance than magenta so the alternating stimulus will create flicker. The luminance of green or magenta is therefore adjusted until perceived flicker is eliminated.

The CRT was run at 120Hz, resulting in a heterochromatic flicker rate of 60Hz. For the alternating stimulus:

$$\begin{aligned} I_1 &= w_G L_G \left(1 + C_C\right) + w_M L_M \left(1 - C_C\right), \\ I_2 &= w_G L_G \left(1 - C_C\right) + w_M L_M \left(1 + C_C\right), \end{aligned} \tag{2}$$

where I_1 and I_2 are the perceived brightnesses of the two stimulus states, and w_G and w_M are the weighting coefficients for brightness for each color. In heterochromatic flicker photometry, one adjusts L_G and/or L_M until I_1 equals I_2. Those values L_G and L_M are then by definition equally bright.

We determined the isoluminant matches for green and magenta for each subject in two sessions. Full-screen, uniform stimuli composed of green or magenta were presented alternately at 60Hz. The sum, L_G plus L_M, was 7.2cd/m^2. In the first session, the ratio $L_G/(L_G+L_M)$ was 0, 0.2, 0.4, 0.6, 0.8, and 1. Subjects indicated after each presentation whether the stimulus appeared to flicker or not. From the data, we determined the $L_G/(L_G+L_M)$ ratio that was approximately equally bright. We then presented a smaller range of ratios to find the precise equal-brightness point. We found that we could not eliminate flicker altogether if the contact C_C was too high so we manipulated that contrast as well.

Two subjects participated. One was a 23 year-old graduate student who was unaware of the experimental hypothesis. The other was one of the authors.

2.3 Results

Figure 4 shows the results from the two subjects. In each panel, the abscissa is the proportion of green ($L_G/(L_G+L_M)$) and the ordinate is the contrast (C_C). The ratio $L_G/(L_G+L_M)$ for which C_C had to be greatest was 0.6 for subject ZJH, and 0.58 for JSK. From this, we estimated the isoluminant ratio as 0.59.

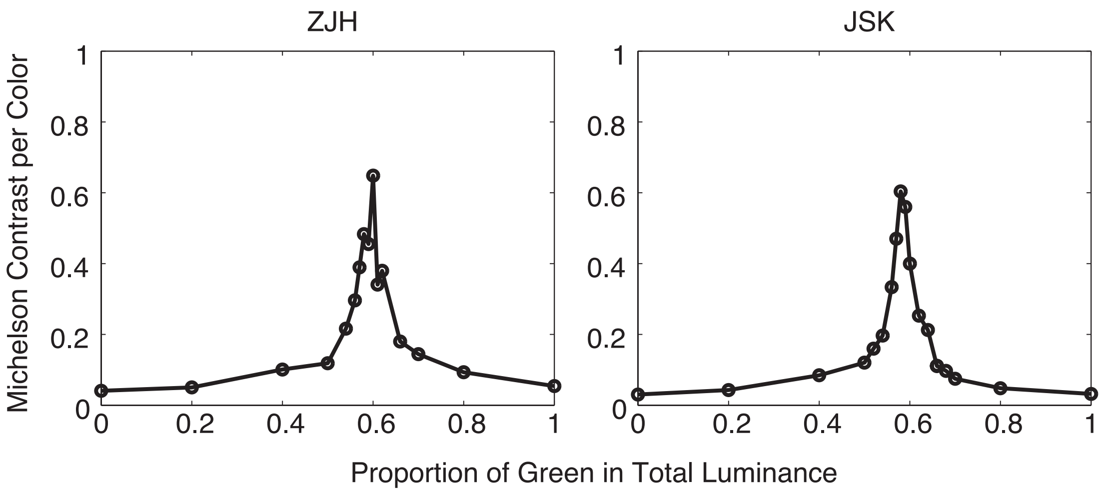

Figure 4. Heterochromatic flicker photometry results for two subjects: ZJH and JSK on the left and right, respectively. The contrast required to perceive flicker is plotted as a function of the ratio $L_G/(L_G+L_M)$. A specific value of that ratio yielded much less visible flicker, so the contrast had to be increased significantly to reach flicker threshold.

3. EXPERIMENT 2: DEPTH DISTORTION IN COLOR-INTERLACED S3D DISPLAYS

In this experiment, we measured the magnitude of depth distortions with color interlacing for stimuli of different saturations.

3.1 Apparatus

The apparatus was the same as in Experiment 1.

3.2 Methods

Depth distortions in temporally interlaced displays are caused by horizontal, not vertical, motion. Thus, left-to-right motion may appear too near and right-to-left motion too far, but up-down motion is seen correctly. To measure the magnitude of distortions, we presented ten bright 1° disks that rotated around a point in the center of the screen for 2.5sec (Figure 5a). The radius from the rotation center to each disk center was 6°. When distortion occurred, the stimulus appeared pitched top forward or top back. After each stimulus presentation, the viewer indicated which of the two pitches he/she perceived. A vertical gradient of horizontal disparity was added via a staircase procedure to find the gradient that made the stimulus appear fronto-parallel. The added gradient was our measure of the magnitude of the depth distortion.

We presented five different colors with the same luminance: saturated green, desaturated green, neutral, desaturated magenta, and saturated magenta. The ratio $L_G/(L_G+L_M)$ for the five colors was respectively 1, 0.75, 0.5, 0.25, and 0. The stimuli were presented with the color-interlaced protocol depicted in Figure 1. The neutral stimulus had the green and magenta luminances that were equally bright according to Experiment 1.

With saturated green and magenta, the color- and temporal-interlaced protocols become identical. That is, with saturated green (G only), the stimulus is presented alternately to the two eyes in both protocols with no offsetting magenta (R + B). Similarly, with saturated magenta (R + B only), the stimulus is presented alternately to the eyes with no offsetting green (G). With desaturated stimuli (desaturated green, neutral, and desaturated magenta), color interlacing provides stimulation to both eyes in every sub-frame, while temporal interlacing does not. Thus, if color interlacing reduces depth distortions, it should do so when the stimulus is not highly saturated.

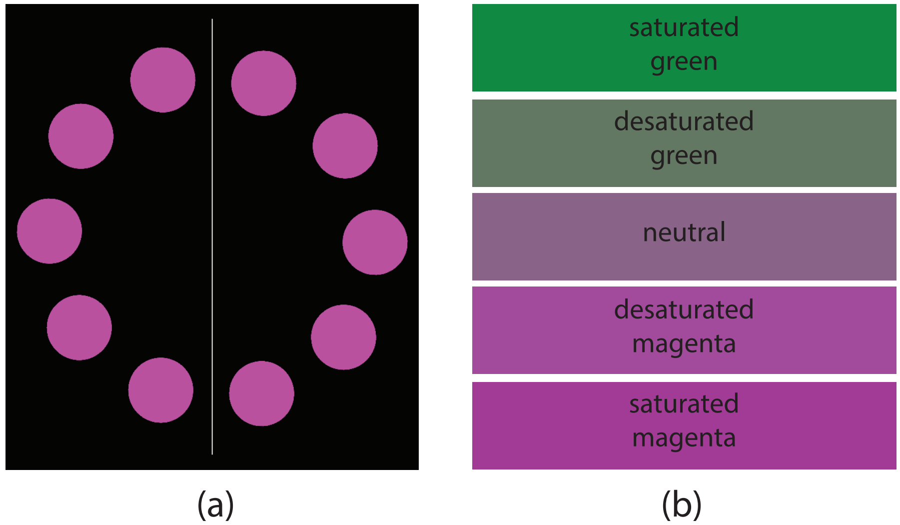

Figure 5. The stimuli in Experiment 2. (a) The rotating disk stimulus for measuring depth distortion. Ten disks rotated about the center of the display screen. The viewer indicated after each presentation whether the plane of disks appeared pitched top forward or top back. Based on the response, we added a vertical gradient of horizontal disparity to make the plane appear fronto-parallel. (b) The five colors used for in the experiment. They all had the same luminance. The ratio $L_G/(L_G+L_M)$ from top to bottom in the figure was 1, 0.75, 0.5, 0.25, and 0.

3.3 Results

Figure 6 shows the experimental results from the three subjects.

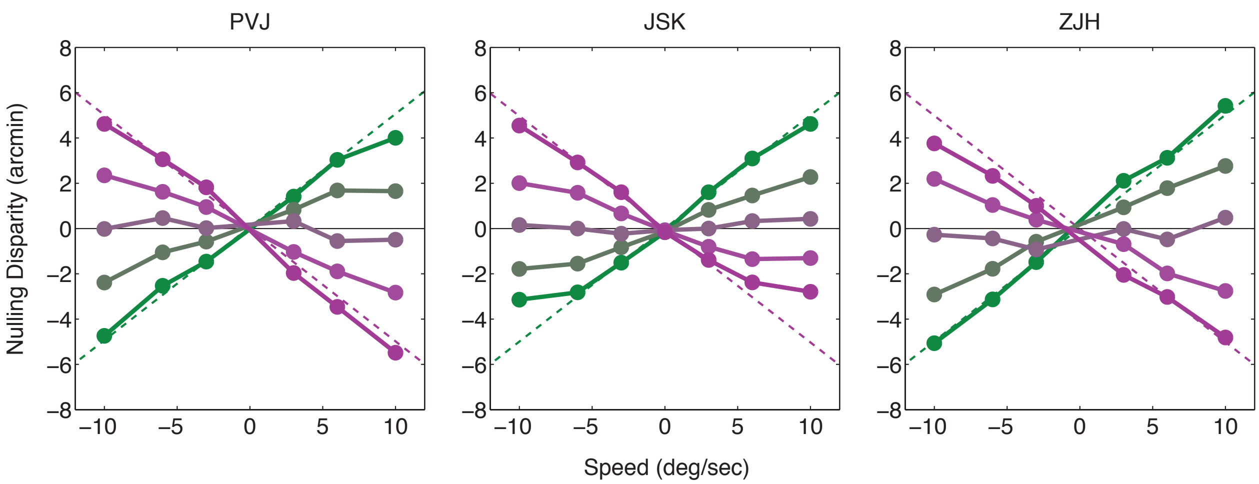

Figure 6. Results of Experiment 2. Each panel displays the results from one subject. In each panel, the abscissa is the tangential rotation speed of the rotating disk stimulus (positive values correspond to clockwise rotation) and the ordinate is the disparity at the top of the rotating disk plane that made the plane appear fronto-parallel (positive values mean that the plane appeared pitched top forward until the nulling disparity gradient was added). Dashed lines represent the depth distortions predicted by Eq. (1). Symbols represent the experimental results, the colors corresponding to the stimulus color.

The disparity gradient that made the disk plane appear fronto-parallel is plotted as a function of the tangential speed of the rotating disks. The ordinate is the nulling disparity gradient. Positive nulling disparity means the rotation plane appeared top pitched forward until the nulling disparity was added. The dashed lines show the depth distortion predicted by Eq. (1). The symbols represent the results with the five colors. The color of the lines matched the color of their corresponding conditions. All three subjects showed the same pattern of results. When the stimulus was saturated, the observed depth distortions were very similar to the ones predicted by Eq. (1). The results with temporal interlacing would have been very similar to these but for all colors (i.e., saturated or unsaturated). When the stimulus was desaturated, the observed distortions were smaller than when the stimulus was saturated. Indeed, the distortions were eliminated when the stimulus was neutral. None of the subjects observed an in-depth color breakup with desaturated stimuli in which one color appeared pitched top forward while another appeared pitched top back.

Thus, we found clear evidence that depth distortions are indeed reduced, or even eliminated, when the stimulus is a desaturated color. The results show that color interlacing could be a very useful technique for S3D displays. They also show that the visual system calculates disparity based on luminance signals rather than chromatic signals.

4. CONCLUSION

We implemented a color-interlaced protocol for S3D displays. This protocol is a variant of the common temporally interlaced technique, which is known to produce distracting depth distortions. We found that color interlacing reduces the magnitude of depth distortions when the stimulus color is not highly saturated. The distortions are eliminated altogether when the colors presented simultaneously to the two eyes (green to one eye and red and blue to the other) are equally bright. We also observed qualitatively that color interlacing reduces the visibility of flicker and motion artifacts, but that it can increase the visibility of color breakup in the plane of the display screen.

5. ACKNOWLEDGMENT

This research was supported by NIH research grant R01EY012851.

REFERENCES

[1] Burr, D. C., and Ross, J., "How does binocular delay give information about depth?," Vision Research, 19 (5), 523-532 (1979).

[2] Simon, A., and Jorke, H., "Interference filter system for high-brightness and natural-color stereoscopic imaging," SID Symposium Digest of Technical Papers 42(1), 317-319 (2011).

[3] Kim, J., Johnson, P. V., and Banks, M. S., "The visibility of temporal artifacts in stereo 3D displays," 8th Asia-Pacific Conference on Vision, 62, (2012).

[4] Kim, J., Kim, Y., Hong, J., Park, G., Hong, K., Min, S.-W., and Lee, B., "A full-color anaglyph three-dimensional display system using active color filter glasses," Journal of Information Display, 12(1), 37-41 (2011).

[5] Livingstone, M., and Hubel, D., "Segregation of form, color, movement, and depth: Anatomy, physiology, and perception," Science, 240, 740-749 (1988).

[6] Hoffman, D. M., Karasev, V. I., and Banks, M. S., "Temporal presentation protocols in stereoscopic displays: Flicker visibility, perceived motion, and perceived depth," Journal of the Society for Information Display, 19 (3), 271-297 (2011).

[7] Wyszecki, G., Stiles, W. S., [Color Science], John Wiley & Sons, New York, (1982).

[8] Walsh, J. W. T., [Photometry], Constable & Co. Ltd, London, (1958).

[9] Mullen, K. T., "The contrast sensitivity of human colour vision to red-green and blue-yellow chromatic gratings," Journal of Physiology, 359, 381-400 (1985).