

Beyond Stereo: An Exploration of Unconventional Binocular Presentation for Novel Visual Experience

Haimo Zhang^{1,2}, Xiang Cao¹, Shengdong Zhao²

¹Microsoft Research Asia
5 Danling Street, Beijing, China
xiangc@microsoft.com

²National University of Singapore
13 Computing Drive, Singapore
{z-haimo, zhaosd}@comp.nus.edu.sg

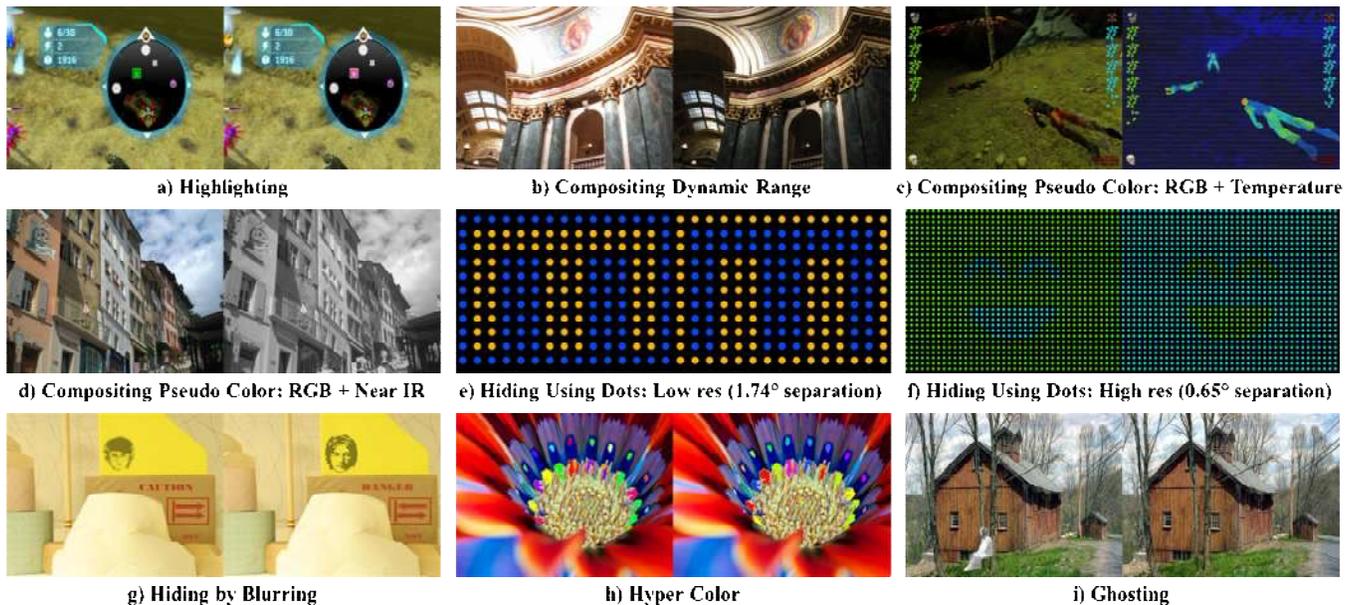


Figure 1. Effects Explored in the Study

ABSTRACT

Human stereo vision processes the two different images seen by the two eyes to generate depth sensation. While current stereoscopic display technologies look at how to faithfully simulate the stereo viewing experience, we took a look out of this scope, to explore how we may present binocular image pairs that differ in other ways to create novel visual experience. This paper presents several interesting techniques we explored, and discusses their potential applications according to an informal user study.

Author Keywords

Binocular vision; special effect; stereo.

ACM Classification Keywords

H.5.2. [Information interfaces and presentation (e.g., HCI)]; User Interfaces – Screen design (e.g., text, graphics, color);

INTRODUCTION

The binocular vision system of two human eyes allows us to see the world in two slightly different perspectives with

optical parallax, and our brains process such stereo images to allow us to perceive depth. Numerous stereoscopic display technologies have been created to present prerecorded or synthesized stereo images to our two eyes to simulate 3D sensations, and they are increasingly accessible to the general public. However, these technologies are not necessarily limited to presenting stereo images, but are technically capable of independently displaying an arbitrary pair of images to two eyes. This opens up the question, can we present pairs of images that differ from each other in ways other than stereoscopic, and what does it buy us if we were to deliberately exploit such unconventional binocular vision?

Some cognitive scientists have studied such non-stereo binocular vision in order to understand human perception, especially in terms of binocular rivalry [2, 3, 5, 8], where two dissimilar images are presented to study which features are more dominant in visual perception. Differing from these scientific works, we sought to answer this question from an engineering perspective, i.e. how we can constructively exploit human's binocular vision system in unconventional ways to present novel visual experience for practical applications.

To explore this, we proposed a set of non-stereo binocular presentation techniques inspired by cognitive science literature, and conducted an informal study to validate the viability of these techniques as well as collecting subjective

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI'12, May 5–10, 2012, Austin, Texas, USA.

Copyright 2012 ACM 978-1-4503-1015-4/12/05...\$10.00.

accounts of the actual visual experience. We first describe our study procedure, and then detail the design and findings of each technique.

STUDY PROCEDURE

We used a Vuzix VR920 binocular head-mounted display (HMD). It produces an image of 32° diagonal visual angle at 4:3 aspect ratio, focused at 9 feet, for each eye. The overlap between the two eyes’ views was set to 100%. Each participant adjusted the tiltable section of the device to clearly see the entire display. Participants looked with both eyes unless otherwise instructed.

For each presentation technique we tested (see Figure 1 for some examples), six static image pairs were used¹. As we expect our techniques to be integrated with stereo images in practice, where applicable our stimuli also included regular stereo cues. The stimuli were presented to the two eyes at 30Hz refresh rate each. Each image pair was presented twice with the left/right views swapped; and as a baseline comparison, for each pair of images, we also used the same device to present to both eyes an average image that is equally alpha-blended between the image pair. The stereo cues were kept intact in all three conditions however. Upon participants’ request, all stimuli may be freely revisited.

Upon presentation of stimuli, participants were asked to describe in their own words what they were seeing in as much detail as possible. Only when they could not discover the special visual effects by themselves did we provide hints for regions of interest, however without suggesting the actual effect expected. The study was audio recorded for further analysis. Six people (2 females) aged between 24 and 35 participated. All had normal or corrected-to-normal vision, normal stereo vision, and normal color vision. The participants were tested for eye dominance using the Miles test [6] with several repetitions: 3 participants (male) were generally right-eye dominant, while the rest were generally left-eye dominant.

TAXONOMY

Inspired by cognitive science research, we considered three general dimensions along which we can produce non-stereo difference between the pair of images: Color, Sharpness, and Semantic Content. As we are interested in practical applications, we also consider the following four general categories of visual effects: Highlighting, Compositing, Hiding, and Wowing. Table 1 illustrates this design space, and summarizes the techniques we explored and their attributes by these two criteria, to be detailed in the following sections. Due to limited space, for each technique we only reprint one example image pair in the paper, while the full set of image pairs we tested can be found in the supplementary material of this paper.

		Color	Sharpness	Semantics
Highlighting	Color Highlighting	√		
	Dynamic Range	√		
Compositing	Pseudo Color	√		
	Color dot pattern	√		
Hiding	Hiding by Blurring		√	√
	Hyper color	√		
Wowing	Ghosting			√

Table 1. Application domains and production procedures.

HIGHLIGHTING

The highlighting effect aims at making certain regions of interest more noticeable to the viewer, and currently includes one technique: color highlighting. This is produced by painting the region with two different colors (90-180 hue difference, 0-13% saturation difference, and 0-66% brightness difference). For example, in Figure 1a, the square icon in the circular game map differs between violet in one image, and green in the other.

Study revealed that regions of highly saturated color pairs that differ in hue were particularly prominent, as long as the size of the region is above a certain threshold ($\approx 0.44^\circ$ in view angle). Participants described the effect either positively (4/6, i.e. 4 of 6 participants) as “highlighted”, “blinking”, or negatively (2/6) as “distracting”, “annoying”. More interesting descriptions include “fluorescent”, “bright and colorful emissive light”, and “floating”. 2 participants were reminded of special optical material they have seen, such as lenticular or refractive films, as these materials may also display view-dependent colors. Eye dominance had a noticeable effect, where different sensations were reported when shown inverted pairs. On the other hand, these sensations were unstabilized, as typical in binocular rivalry [2]. Unsurprisingly, in the baseline image where the two views were averaged, the regions of interest became even less prominent because averaging two contrasting colors results in desaturation.

The power of color highlighting has thus been confirmed. One thing worth noting is that neither single color in the contrasting pair needs to be prominent from the surroundings on its own, so highlighting can be achieved without compromising the image composition.

The highlighting effect could be theoretically suggested by the perceived shininess of a surface [7]: in real world, some “shiny” materials with specular reflection or refraction properties display drastically different hues and brightness from different view angles. Our proposed highlighting technique produces a similar color rivalry [5] effect that resonates with this principle of perceived shininess, thus highlighting the rivaling regions as “shiny” areas.

COMPOSITING

The compositing effect aims at presenting two images of the same scene, however are complementary in terms of information spectrum along a certain dimension. We expect that the human perception system may be able to composite such information to receive a higher bandwidth than is possible with a single view. We explored such compositing effects along two different dimensions:

¹ Some test images based on:
http://pages.cs.wisc.edu/~csverma/CS766_09/HDR1/hdr.html
http://ivrg.epfl.ch/supplementary_material/cvpr11/
<http://vision.middlebury.edu/stereo/data/scenes2006/>
 and Vuzix JPS Viewer:
http://www.vuzix.com/support/downloads_drivers.html

Compositing Dynamic Range

Compositing dynamic range concerns a pair of photographs taken at different exposures, each missing part of the illumination range of the scene. Figure 1b illustrates.

When shown image pairs with different exposures, participants were able to describe details that are only available in one of the two images, where the corresponding region in the counterpart image is subject to overexposure or underexposure due to limited dynamic range of the camera. This may be explained by contour dominance [8], where the rich contours in one image effectively suppressed the more uniform over/under exposed regions in the other image.

Comparatively, although the baseline average image also incorporates these features, they were less prominent, as averaging reduced the overall image contrast. Interestingly, swapping the left and right eye images also resulted in a perceived change of global brightness; and two participants further perceived change in light source, although they could not determine the nature of that change. The global brightness was biased towards the dominant eye.

Compositing Pseudo Colors

In pseudo color images, pixel values do not represent true visible light intensities, but some other physical channels such as temperature or near infrared (NIR) response. By letting the viewer composite ordinary RGB images with such pseudo color images, we expect that they would be able to make sense of the complementary nature of different channels. Examples are given in Figure 1c and 1d.

For RGB-temperature image pairs, all participants reported seeing bright human figures. It is hard for them to see the human figures' actual color, but they could identify the green color of the background land. One participant described the effect as the contour "shaking". When shown the naïve average between the two images, participants responded that more colors could be observed on the human figure, but that the background color is less obvious. For RGB-NIR image pairs, all participants reported that the boundaries of plants, brightly colored blankets, and sign boards are "bright", "confusing", and sometimes (2/6) "floating", which "does not fit well" into the scene. When shown the averaged baseline, participants reported that the color is not as vivid as in the previous case, but the feeling of "unfitness" also disappeared. The effect of eye dominance is mainly on the overall perceived saturation: when the gray scale NIR image is shown to the dominant eye, participants reported reduced saturation as compared to viewing the other way round.

The response from the participants meets our expectation: in the RGB-temperature image pair, textured background of RGB image suppressed the almost uniform background of temperature map; the strong edges of the human figure in the temperature map suppressed the perception of its normal color; and in RGB-NIR image pairs, large luminance difference of plants and painted sign boards between RGB and NIR images creates strong response that the participants could not overlook. This effect could also be exploited in real-time vegetation inspection applications, by feeding RGB video and NIR video to each eye, while

letting the human brain to work out the distracting areas as the possible contours of vegetation.

Besides the contour dominance factor, the aforementioned highlighting effect may also play a role here: regions with rivaling colors would appear bright and catches more attention. The result shows that humans are able to effectively incorporate multi-spectrum visual information through binocular vision and make sense of them. Essentially, by leveraging the "computation power" of the human visual perception system, we may alleviate or eliminate the need for computers to algorithmically fuse such information [4], which is not always possible or straightforward to do.

HIDING

The hiding effect aims at a seemingly counterintuitive goal: to turn visible information in monocular images invisible in the binocular view. In other words, we attempt to hide some information from the viewer when both eyes are open, while revealing it once s/he closes one of the eyes. This may provide a lightweight mechanism for switching between information layers. For example, in video games, users keep both eyes open to see the regular game view, but may occasionally close one eye to access additional information such as player statistics. This is achieved without active sensing the user's eye movement.

Hiding Using Color Dot Patterns

Research in binocular color fusion [5] has discovered several possible outcomes when human brain attempts to fuse two different colors presented to two eyes, ranging from stable uniform fused color, to color sensation that vary both in space and in time. Regardless of the outcome of the color fusion, it is usually difficult for human to determine which eye is seeing which. This suggests that if we present a pair of different colors to both eyes, it may become indistinguishable from the sensation of the same pair presented with left and right eye color swapped. Therefore, by rendering a shape in one eye using a foreground and a background color, and rendering the same in the other eye but with the foreground and background color swapped, it becomes possible that in binocular views at any point the user see the result of color fusion, which are more or less consistent regardless of which color comes from each eye. Thus the information would become invisible to the viewer.

However, such a technique could not work if there exists a visual contour between the colors in either view. As explained previously in [8], such contours are sensed individually by each eye, thus cannot be eliminated by binocular vision. In order to eliminate this, we convert the shape into a dot grid pattern (Figure 1e and 1f), so that it is only encoded by the color contrast but not contours. We generated four test image pairs with two levels of grid resolution (low/high) and two color schemes (complementary/similar colors).

Participants reported that they see dots constantly changing color. Although not seeing clearly to be certain, they were able, after intentionally viewing for some time, to describe the correct pattern in the higher resolution image pairs. They were more confused and uncertain when seeing the low resolution patterns. The minimum time from the onset

of the pattern till the participants voluntarily guessed the correct or similar pattern is 10 seconds in the low resolution and 1 second in the high resolution. This period might be the sweet spot to hide information temporarily.

Hiding by Blurring

According to Fahle [3], higher spatial frequency (sharp features) is more dominant than low spatial frequency (blur features) when they are presented to the two eyes respectively. To hide information by this principle, we can create image pairs that show different semantic information in corresponding regions but in different level of sharpness. Contour dominance results in the sharp information masking the blurred information when both eyes are open, while the blurred information becomes visible only when the other eye is closed. We can apply this principle to multiple regions of the image pair, so that each image contains regions that can be either masking the other image or masked by the other, naturally supporting not one, but two individual views that can be revealed by closing either one of the eyes. Figure 1g illustrates one of the image pairs we tested, which shows this technique for both textual and graphical information and in alternate directions between the two eyes.

In general, most participants (5/6) reported seeing the sharper of the text rivalries. With the faces there was more variance between the participants, as 3 of them seeing the sharp face, and others seeing the alternate or mixed information. We suspect this was because the higher-level perception mechanism involved in recognizing faces interacted with our technique. The averaged baseline image appeared confusing and unrecognizable to all participants. When asked to use their single eyes to view the images, the participants were able to recognize all information except for the “NOV” / “DEC” texts, which might be due to the small size of the stimuli. After revelation of the effect, all participants found this effect interesting and fun.

Since no excessive attention is paid to the regions with hidden information, this effect might also be used to hide text or simple graphics to uninformed viewers, and only recognizable by informed users.

WOWING

Wowing effects aim at creating surprising sensations that are not necessarily useful for productivity applications but can facilitate compelling experience in applications such as cinema or gaming. The two effects we explored are as follows:

Hyper Color

We specifically tried to create the effect of “impossible colors” [1] by showing different colors in each eye, as shown in figure 1h. However, participants’ responses were not colors that are yellowish blue or greenish red. Instead, 3 participants saw inhomogeneous color patches that change smoothly over time. Other responses include “fluorescent light”, “jittering color patches”, “bright outline” and “shiny and unstable positions”. Based on the same principle as color highlighting, here we are addressing a different

application context: we name it “hyper color” and see potential enlargement of color vocabulary in binocular visualizations.

Ghosting Effect

We showed two image pairs where an object is presented only in one eye’s image but not the other, thus giving it a ghostly appearance, as illustrated in figure 1i. All participants reported the effect to be similar to transparency. Two participants described temporal change in transparency while the same image pair is presented. One participant explicitly used the word “ghost” in the description without prompting. All participants reported difference in transparency between the ghost effect, its left-right swapped version, and the baseline. Further, this effect demonstrates the temporal fluctuation of perceived transparency, which could not be experienced with transparency rendered in monocular static images.

The ghosting effect partially lies in the unstable perception found in binocular rivalry [2], in which the dominant perception alternates between the two eyes’ views, which explains the fluctuation of transparency of the “ghostly” object in the image pair.

CONCLUSION

We have explored several unconventional binocular presentation techniques to create new visual experience. In the future, we would like to further explore how these and other potential interesting effects could be applied in various aspects of human-computer interaction, as well as quantifying the parameters of the techniques and providing a more systematic presentation vocabulary.

REFERENCES

1. Billock, V.A. and Tsou, B.H. Seeing Forbidden Colors. *Scientific American* 302, (2010), 72-77.
2. Blake, R. A Primer on binocular rivalry, including current controversies. *Brain and Mind* 2, 1 (2001), 5-38.
3. Fahle, M. Binocular rivalry: suppression depends on orientation and spatial frequency. *Vision Research* 22, 7 (1982), 787-800.
4. Goshtasby, A.A. and Nikolov, S. Image fusion: Advances in the state of the art. *Information Fusion* 8, 2 (2007), 114-118.
5. Jung, Y.J., Sohn, H, Lee, S., Ro, Y.M. and Park, H.W. Quantitative measurement of binocular color fusion limit for non-spectral colors. *Optics Express* 19, 8 (2011), 7325-7338.
6. Roth, H.L., Lora, A.N. and Heilman, K.M. Effects of monocular viewing and eye dominance on spatial attention. *Brain* 125, Pt 9 (2002), 2023-35.
7. Tyler, C.W. Sensory processing of binocular disparity. *Vergence eye movements: Basic and clinical aspects*. Butterworths, London, UK, 1983.
8. Xu, J.P., He, Z.J. and Ooi, T.L. Surface boundary contour strengthens image dominance in binocular competition. *Vision Research* 50, 2 (2010), 155-170.