Stereoscopic vision and the design of stereoscopic displays

Joshua M. Cobb
Corning Tropel Corporation, 60 O'Connor Road, Fairport, NY 14450

ABSTRACT

Since the discovery of stereopsis in the 19^{th} century the display of stereoscopic information has taken on a variety of forms. From the original Wheatstone stereoscope to the current technology of head-mounted displays and parallax barrier displays, the advantages of binocular vision have been exploited to produce a different kind of display experience.

Designing optical systems that take advantage of stereopsis creates a different set of constraints and image artifacts. This paper reviews some of the methods of creating a stereo image and highlights some of the unique system design considerations.

Keywords: Stereoscopic, autostereoscopic, stereopsis

1. INTRODUCTION

Human beings have been trying to create a display that mimics what their eyes see and their brains interpret since cavemen scratched stick figures on rock walls. We have come a long way since then, but we still have not approached the full gamut of the visual perception system. Suppose we would like to develop a virtual reality display that would present to the viewer an experience that did not fall short of the perceptual detection system that has evolved. What would the specifications be? What metrics is the design judged by? Let us consider a few rough, conservative numbers.

Without rotating the head, we have a horizontal field of view of over 120 degrees. The resolution across the field of view is much higher in the region of the fovea than in the periphery, but we are constantly moving our eyes around the scene to integrate it in high resolution. Some have presented data showing human eye resolution for a single eye as fine as 20 seconds of arc[1]. However, stereo acuity has been shown to be, on average, 10 seconds of arc and at the maximum, 2 seconds of arc[2]. If we consider a display that would present 120 degrees by 90 degrees, and each eye must be able to see an angular pixel size of 2 seconds of arc then each frame of the display would need about 70 Gigapixels. At this time, there are no electronic display technologies that can present this resolution economically. Resolution and field of view are only two parameters that describe the perceptual gamut. Others include contrast, cross-talk, color gamut, brightness, distortion, alignment, binocular overlap, accommodation point, and accommodation range. Therefore, designing a stereoscopic display requires adopting a set of compromises that lie within the perception gamut. Where does the designer make the trades that will produce the best image experience? This is a very complex problem where the optimized solution space depends on human factors that vary from one person to another. Many people have performed experiments to find an optimum solution space for some of these human factors. It is the purpose of this paper to present some of the compromises that can be made, but not to propose the optimum solution as this optimum solution depends on the application and the subject that is using the display.

This paper will discuss the metrics in a stereoscopic display that allow the viewer to have a comfortable experience. This paper will then review some of the methods that have been used in displays to exploit stereopsis and what compromises the designers of these displays have chosen. Examples of stereo pair images can be found at Terry Wilson's website: http://www.terryfic3d.com.

International Optical Design Conference 2006, edited by G. Groot Gregory,
Joseph M. Howard, R. John Koshel, SPIE Vol. 6342, 63420V,
 · doi: 10.1117/12.692261

2. STEREOSCOPIC METRICS

2.1 Resolution and Field of view

A large field of view can create an immersive viewing experience that is very impressive. However, with a finite number of pixels to fill this field the detriments of low resolution will eventually outweigh the benefits of immersion. The metric that describes how sharp an image appears is angular resolution which is measured in minutes of arc per pixel. Stereoscopic displays are generally in the range of 1-5 minutes of arc per pixel. In the author's personal experience of viewing displays a resolution of about 2 minutes of arc per pixel creates a very sharp image. 2 arc minutes is the resolution one sees when viewing an ordinary LCD screen from about 400 mm. This angular resolution then sets the field of view for a given number of pixels in the display. For example, if the display resolution is XGA at 1024 by 768, the field of view is 34 degrees by 25.6 degrees for an angular resolution of 2 minutes per pixel. Some pupil imaging displays (especially head mounted displays) have compromised the stereoscopic overlap to achieve a larger field of view. For example, each eye may see a 34 degree field of view, but the field that is common to both eyes may be smaller. This can give a more immersive experience without sacrificing angular resolution. A large field of view can also increase the fusional range[3], that is, the amount of disparity that can be fused.

2.2 Accommodation-Convergence Mismatch

There are many cues from a scene that our visual perception system uses to give us a sense of depth. These include the retinal image size, perspective, shading, and others that can be perceived on a flat screen with one eye[4]. Two very strong depth cues that can not be presented on a conventional flat display are accommodation and convergence. In the real world, we change our focus as we look at objects at different depths. Our eyes also converge at different angles to objects at different depths. Convergence is what is varied within a stereoscopic display and is the depth cue that is linked to stereopsis. Many have documented the uncomfortable stereoscopic viewing experience that happens when the viewer is focused and converged at different distances[5].

Research has also been performed to find the optimum accommodation distance that allows the largest range of comfortably viewed disparity[6]. Displays that create a virtual image through optics can take advantage of this by defining the focus point of the viewer to be more distant and thus can present a wider range of disparity.

Another metric that is tied into this phenomenon is the brightness of the display. As the display gets brighter, the pupil contracts and the eye's depth of focus increases, thus reducing the effect of accommodation-convergence mismatch.

2.3 Cross-talk

In real life, the left eye's image and the right eye's image are completely separate. Stereoscopic displays each have a unique way of feeding one image to the left eye and its stereo complement to the right eye. Some of these methods result in cross-talk, that is, the left eye sees a dim image intended for the right eye and vice-versa. The amount of crosstalk that is acceptable to still maintain fusion has been shown to be quite large,[7] however, crosstalk also affects how long a user can view the display without discomfort. Some research has shown that cross-talk must be kept below 0.3 percent for the ghosting to be imperceptible to the viewer.[8]

2.4 Alignment

The alignment of the stereo pair images is one of the most important considerations in creating a comfortable display. The eyes have wide freedom to move horizontally, but they have very little tolerance for vertical disparities between images. The most important consideration for horizontal alignment is that the eyes should never have to point beyond parallel. This is sometimes termed "wall-eyed". The variation in vertical disparities can be caused by many factors. A vertical shift of one image with respect to the other is an obvious cause of vertical disparity, but so too is any distortion of the images. Distortion changes the image position as a function of where it is in the field of view. Since stereo pair images have common objects that have horizontal disparity, even if there is the same distortion in each eye's image, vertical disparity will result. It may be possible to align a portion of the field of view, but not the entire field of view. A similar effect can be seen with a rotation misalignment of the images. When the images are not aligned with respect to rotation some parts of the field can be fused and some can not. Human factors research has documented that left to right vertical misalignment should be below 10 arc-minutes or the viewer will experience significant discomfort.[9]

Another common but obvious misalignment is presenting the image in reverse-stereo, that is the right image goes to the left eye and the left image goes to the right eye. This happens quite often to novice users as reverse stereo will present a stereo image that does not look quite right.

Alignment is critical because even a badly aligned image may be able to be partially fused. The viewer will see a stereo image and think that it is a good image. However, after a short period of time eye strain can set in. As the image alignment improves, a noticeable improvement in the stereo quality and comfort will be experienced.

2.5 Other metrics and overall image quality

We can perceive stereo even with a terrible stereo image. One of the most common methods of displaying stereo images is in anaglyph form where one eye sees a red image and the other sees a blue image. In this case, the color rendition and brightness of the image are severely reduced and the cross-talk can be high. Some can fuse stereo images with large amounts of vertical disparity. There is a large range of differences in how people see stereo, but stereo images that are well aligned, have low cross-talk, and are bright and uniform are easier for more people to see and can be viewed for a longer period of time without discomfort.

3. METHODS OF STEREOSCOPIC DISPLAY

3.1 Anaglyph

Stereo presentation in anaglyph has the advantage that it is one of the easiest methods available. Both left and right images can be presented on the same screen at the same time and are altered only by color. Colored filters prevent the eyes from seeing the incorrect image, and it is very important that the filters are precisely matched to the image or significant cross-talk will be seen. Anaglyphs have been a common way of showing 3-D motion pictures, and recently they have been the principle method of showing a relatively new art form called phantograms. Examples of these can be seen at http://www.terryfic3d.com.[10] Phantograms are stereo photographs that are taken in a precise fashion and then digitally altered to make it appear that the image is sitting on top of the page when viewed from the correct angle.

3.2 Shutter Glasses

High quality stereo can be viewed on a conventional cathode ray tube (CRT) display by modulating the left and right images at a very fast frequency. The user then wears liquid crystal shutter glasses that are synchronized to the display. Each eye sees an image at half the display frequency. These stereo displays can have high angular resolution and a wide field of view if the monitor is big enough. The brightness is reduced by a factor of two, and if the frequency of the CRT is not high enough there is noticeable flicker. These displays have another disadvantage in that the user is focused within about three feet and the amount of disparity must be reduced to alleviate the convergence-accommodation mismatch. It should be noted that liquid crystal displays are not fast enough to present this type of image to the viewer without creating substantial cross-talk. There is some cross-talk with CRT displays due to the persistence of the phosphors, but this is negligible on better CRT displays that run at about 250 Hz.

3.3 Polarized glasses

Polarized glasses can give a very similar stereo experience in comparison to the shutter glasses. The advantage to the polarized glasses is that they are generally lighter and more comfortable than shutter glasses. Polarized glasses are used in high quality motion picture stereo projection and require a polarization preserving screen. If the polarization is linear, a rotation of the head while viewing can cause cross-talk. This cross-talk can be reduced by circularly polarizing each stereo channel. This is done with the Stereographics Z-screen which modulates the polarization of a CRT. High quality stereo can also be viewed with dual electronic business projectors by using polarizing filters and a polarization preserving screen. An advantage to projection is that the viewer's focus can be more distant on a larger screen thus reducing convergence-accommodation mismatch.

3.4 Parallax barriers/Lenticular barriers

A number of autostereoscopic displays are available that employ either a parallax barrier or a lenticular screen to create a viewing zone where each eye sees a different image. In these cases, the resolution of the display is traded for a "look-around" feel. The minimum number of views that is presented is a stereo pair, but multiple views can be shown which will create viewing locations for multiple users. Columns of pixels on a liquid crystal display (LCD) screen are obscured from the view of one eye by a barrier, but the other eye can see around the barrier. In this way, the stereo pair is sliced up and interleaved to be displayed on a single screen with a barrier. The barrier can be placed over the LCD screen, or in the backlight illuminator of the LCD. One downside to these displays is the decreased spatial resolution especially when operating in 2D. It can sometimes be difficult to read text because of the lower pixel count. Sharp has developed a product that addresses this in that the parallax barrier is comprised of another LCD screen.[11] This way the parallax barrier can be turned off in 2D mode and thus allowing full resolution.

3.5 Headmounted displays/Pupil imaging displays

The previously mentioned displays do not incorporate any optics that would create a virtual image of the display. Head mounted and pupil imaging displays are different. In the case of the head mounted displays, a micro-display (such as a Liquid Crystal on Silicon (LCOS) or Organic Light Emitting Diode (OLED) device) is imaged by optics in front of each eye. The advantage to these systems is that very large fields of view can be created within a compact volume. One disadvantage is that the resolution of the micro display devices is usually low. Nintendo brought this technology to market in a video game system called "Virtual Boy" in the mid-1990's. Although it was a true stereoscopic video game, the image quality suffered due to a small angular resolution, a small field of view, and monochrome display.

Headmounted displays have a variety of unique design challenges. Higher resolution systems usually require either multiple micro-displays (three (one for each color) for each eye) or larger micro-displays. Both of these lead to larger, heavier systems that degrade the viewer's comfort. In addition, if the display is mounted to the viewer's head then any movement of the head will result in a corresponding movement of the image. This can be compensated by a head tracker and subsequent changing of the image data, but it adds cost and weight to the system.

The size of the pupil is another consideration in the design of pupil imaging systems. The average pupil of the eye varies from 2 mm – 6 mm, but if the pupil of the display is too small it may be difficult for the viewer to find the best viewing location. Also, if the field of view is substantial, the user is continually moving his eyes around the scene. The pupil size must be large enough to accommodate this. Of course, the image quality across the pupil does not have to be corrected completely since only a portion of the viewing pupil is used at any one time, however, large aberrations across the pupil can manifest themselves in the form of a "swimming" image as the user's head moves.[12]

Field of view versus pupil size is another compromise in the design of pupil imaging systems. A larger field of view requires a larger pupil size so that the viewer's eyes don't walk out of the pupils when viewing the extreme fields of view. However, designing a system with a large field and large pupil usually requires a more complicated optical system that adds weight and cost to the head mounted display.

Pupil imaging displays can be very bright because there is no ambient light to reduce the contrast. The amount of light needed is much smaller than a projection display or LCD screen, and so there are many choices for illumination.

There have been several experimental pupil imaging displays that use large, thin-film-transistor LCD screens as the image generators. These have the advantage of wide fields of view while maintaining high angular resolution due to the large number of pixels in the displays.[13]

4. USE OF STEREOSCOPIC DISPLAY TECHNOLOGY

Because of the difficulty in capturing and displaying stereoscopic images, the technology has been historically used by niche markets and hobbyists. This is changing due to the availability of three-dimensional content, digital capture, and new display technology. One area of particular interest is in optical design and mechanical design. Current software packages render a three-dimensional image through the application programming interface (API) called Open-GL. This API and others (such as Microsoft's DirectX) are used to shade and render three-dimensional objects with the correct perspective. All of the data needed to create a stereo-pair image is already calculated by these API's. NVIDIA Corporation has taken advantage of this by writing a stereoscopic driver for NVIDIA graphics cards that converts any Open-GL or DirectX application into a stereo pair output[14]. In this way, software that uses these API's can be presented in stereo-3D with minimal software changes. Several mechanical design software packages have already supplied the option of displaying their Open-GL applications in stereo-3D. Stereoscopic technology is also being used in high end designs, oil and gas exploration, medical imaging, and video games.

5. CONCLUSIONS

Content for displaying stereoscopic images is now widely available and this has generated multiple technologies for viewing them. Each technology has its benefits and compromises. The perception of three-D varies widely among people and developing a stereoscopic display requires the designer to trade one feature for another. Since the effects of these features are very dependent on human factors it is difficult to know if the correct trade has been made. It is often useful to build several different prototypes and see which is best for the particular application.

6. REFERENCES

[1] R. N. Clark, "Notes on the Resolution and Other Details of the Human Eye," www.clarkvision.com/imagedetail/eye-resolution.html

[2] Christopher Tyler, "Spatial limitations of human stereoscopic vision," SPIE Vol. 120, Three-Dimensional Imaging, 1977

[3] Elaine Jin, Michael Miller, Serguei Endrikhovski, Cathleen Cerosaletti, "Creating a comfortable stereoscopic viewing experience: Effects of viewing distance and field of view on fusional range," SPIE-IS&T/Vol. 5664, Stereoscopic Displays and Virtual Reality Systems XII, 2005 p.11

[4] Lenny Lipton, *Foundations of the Stereoscopic Cinema*, Van Nostrand Reinhold Company, inc. 1982, pp 53 – 61.

[5] L. Gooding, ME Miller, J Moore, S. Kim, "The effect of viewing distance and disparity on perceived depth," SPIE Vol 1457 Stereoscopic Displays and Applications II, 1991, P. 259.

[6] John O. Merritt, "Stereoscopic Display Application Issues", SC060, Electronic Imaging Science & Technology, 2003

[7] Huang, KuoChung; Yuan, Jy-Chyi; Tsai, Chao-Hsu; Hsueh, Wen-Jean; Wang, Nai-Yueh, "How crosstalk affects stereopsis in stereoscopic displays", [5006-44], SDA 2003

[8] Nick Holliman, "3D Display Systems", draft report, to be published in "Handbook of Optoelectronics", IOP Press, ISBN 0 7503 0646 7, in Press

[9] John O. Merritt, "Stereoscopic Display Application Issues", SC060, Electronic Imaging Science & Technology, 2003

[10] Terry Wilson, http://www.terryfic3d.com

[11] Sharp Systems of America, http://www.sharp3d.com/technology/howsharp3dworks/

[12] J. Cobb, D Kessler, J. Agostinelli, , "Optical Design of a monocentric autostereoscopic immersive display," Proc. Int. Optical Design Conf., 2002.

[13] J. M. Cobb, "Autostereoscopic desktop display: an evolution of technology",[5664A-15], Stereoscopic Displays and Virtual Reality Systems XII, 2005

[14] Nvidia Corporation, http://www.Nvidia.com