

ADVANCES IN DESKTOP 3D-STEREOSCOPIC VISUALIZATION OF GEOSPATIAL DATA

Barry Bitters

Science Applications International Corporation (SAIC)
8301 Greensboro Drive, McLean, VA 22101 - bittersb@saic.com

KEY WORDS: Visualization, 3-D, Stereoscopy, Human Factors, Visual Acuity, Geospatial.

ABSTRACT:

For many years imagery has been exploited in the production of charts and maps using manual stereoscopic viewing techniques. Only recently and only to a limited extent has geospatial image data, been exploited in an automated 3D stereoscopic environment – and only since the early 1990s,. Accompanying the current interest in in-home, 3D-ready HDTVs, there is renewed interest within the geospatial community in the exploitation of geospatial products in a fully automated 3D-stereoscopic visualization environment. Both the entertainment industry and the geospatial community are confronted with a “chick-or-the-egg” dilemma – which comes first – the 3D-stereoscopic display hardware or the 3D-stereoscopic visual content. Within the geospatial community, a promising approach to this dilemma involves the simultaneous technology transfer of 3D-stereoscopic hardware from the computer gaming and entertainment industries; and at the same time performing software modifications to existing geospatial visualization software to allow visualization in either 2D or 3D-stereoscopic mode. This paper describes an active initiative to simultaneously deliver both 3D-stereoscopic viewing hardware and 3D-stereoscopic visualization software and will allow interactive and immersive exploration of very detailed cultural and natural virtual environments, very detailed social networks, and synthetic spatio-temporal environments. There have been significant engineering quandaries during the development of a capability to display both aesthetically appealing and stereographically accurate 3D-stereo visualizations – visualizations of computer generated displays. In an effort to reduce visual confusion during the 3D-stereoscopic viewing experience, significant human factors engineering has been incorporated during this development effort to ensure proper relative alignment of paired eye-points and to avoid window violations.

1. INTRODUCTION

3D stereoscopic viewing has been available in various forms for many years (Wheatstone, 1838). Early attempts at the display of 3D-stereo involved viewing gray-scale and color stereo pairs in a wide array of apparatus such as the Holmes stereoscope (Holmes 1906). Later during the early and mid 1900’s, Hollywood producers created anaglyph 3D movies that required the viewer to wear red/blue glasses. These attempts at 3D-stereo cinematography were generally considered to be poor quality, caused significant viewer discomfort and were sporadically produced during the 1900’s.

Only recently, has a 3D-stereoscopic capability been available on desktop computer systems to view, edit and manipulate geospatial data. The advent of efficient graphics processors has made it possible to view compelling 3D content at the desktop. In the early 1990’s, Stereographics Corporation (now RealD Inc.) introduced its CrystalEyes and SimulEyes product lines (Stereographics, 1997). These products allowed the viewing of 3D-stereoscopic content using very high-end personal computer systems. The initial version of the SOCETSet software (BAE Systems, 2010) was first released in 1992 and allowed interaction with geospatial data in a 3D desktop environment. Today, the SOCETSet suite of image processing tools still allows interactive editing and viewing of geospatial data, both imagery, elevation and feature data, in a virtual 3D-stereoscopic environment. These systems require a viewer to wear either polarized glasses or

shutter glasses, each providing the left and right eyes the ability to discriminated different views of the same scene.

There are very few geospatial software tools that have a 3D-stereoscopic display capability similar to that of SOCETSet. However, there are numerous software tools that can exploit 3D geospatial data but only display results on a 2-dimensional screen. The intent of this research effort is to develop affordable 3D-stereoscopic viewing capabilities to allow more geospatial software to display 3D-stereoscopic geospatial content. This requires not only software development but also requires a significant hardware engineering effort.

For the last 10 years, to a limited extent, the computer gaming industry has used desktop 3D-capable stereo displays. Further, the current trend toward in-home, 3D-stereoscopic high-definition television (HDTV) has increased the availability of 3D-ready display devices. Advances in 3D-stereoscopic displays by both the gaming industry and the in-home 3D-capable HDTV manufacturers have been instrumental in making this technology available and affordable.

If 3D-stereoscopic televisions and their desktop counterparts for 3D gaming and scientific exploration are to be successful, visual comfort must at least be comparable to those principles incorporated in conventional HDTV and computer display standards. Technical choices and compromises made throughout the entire process of stereoscopic image generation, transmission, rendering, and display can affect the overall visual comfort

experienced by the viewer. In fact, visual discomfort associated with stereoscopic displays is often cited as one of the main barriers to rapid diffusion and customer acceptance of this new technology in the marketplace. Therefore, it is imperative that aesthetics and visual comfort be considered when generating 3D-stereoscopic video content.

In most studies, visual discomfort is defined according to a combination of underlying factors and symptoms. Four main causes of visual discomfort are attributed to stereoscopic displays: anomalies of binocular vision, binocular asymmetries, accommodation-convergence mismatch (AC mismatch), and excessive binocular disparity. Anomalies of individual human visual systems that affect visual comfort have been well documented in the ophthalmologic literature, and involve only a relatively small number of viewers. The effects of various binocular asymmetries, based either on asymmetrical optical geometry (e.g., image shift, rotation, magnification) or filter characteristics (e.g., luminance, contrast, color differences, cross-talk), have been studied extensively, and reviews are available elsewhere (Ijsselstein,2005; Kooi, 2004).

Generally, in today's Hollywood 3D productions, artifacts and visual effects that contribute to visual discomfort are often removed or repaired during post-production processing. This though is a recent change. Until recently, the novelty of 3D pop-outs – objects flying out from the screen toward the audience - was considered a normal and crowd pleasing experience. Although an interesting and compelling special effect, the 3D pop-out is a mentally disturbing and visually uncomfortable cinematographic effect and one employed all too often in the past. Only recently have Hollywood production companies paid particular attention to the aesthetics and the visual comfort factors of 3D-stereo video content.

Unlike Hollywood movie productions, real-time 3D-stereoscopic video production does not include those long timelines that allow post-production processing. When generating real-time 3D-stereoscopic video content - 3-D video content to be viewed as it is produced – there are special considerations that must be incorporated in the video generation process to insure the viewing experience is not uncomfortable or detrimental to the viewer's health and well being. This paper discusses an approach to the automated, real-time generation of 3D-stereoscopic video content using readily available geospatial data – an approach that incorporates particular attention to the aesthetics and the visual comfort of 3D-stereoscopic video content.

2. BACKGROUND

Basically, modern 3D-stereoscopic viewing is based on the science and art of stereoscopy; the ability of the human visual cortex to simultaneously and automatically understand a left-eye image and a right-eye image – each from a slightly different aspect. Fundamentally, it is based on the human inter-ocular distance – the distance between the human eyes; approximately 65 millimeters. Stereoscopy exploits the human eye separation to create different views for each eye. Since different offsets as viewed from each eye are interpreted as different distances from the viewer, by producing paired images for simultaneous viewing,

it is possible to transmit different views of the same scene to each eye. The offsets of these two views are interpreted as depth cues.

The intent of modern 3D-stereoscopic display systems is to exploit depth cues from a pair of overlapping images to enhance the viewing experience. By adding a sense of depth to traditional 2-dimensional video displays, a compelling view of “reality” can result – a viewing experience that can provide an added sense of realism and a deeper appreciation of the viewed scene. 3D-stereoscopic display systems also provide a potential to reveal additional detail that otherwise would be missed in the traditional 2D visual environment. This is possible by exploiting the additional depth cues that result from viewing an environment in an enhanced third dimension of depth. Rather than seeing a flat rendition of an object, in a 3D-stereoscopic view the added dimension of depth more accurately expresses the true shape, size and surface texture of the object. As an example, viewing a 2D aerial photograph provides only limited depth cues and detailed information about elevation is not available. However, with a 3D-stereoscopic aerial photograph it is possible to perceive depth cues in the form of measurable elevation. These additional depth cues in 3D-stereoscopic content provide an analytical capability not available in 2D visual content.

There are a number of capabilities that the human visual and cognitive systems use to interpret the visual world and generate a perception of depth. Many of these perceptual cues are effective with two dimensional images as well as 3D-stereoscopic images. They are:

- **Perspective.** Objects appear to be smaller the further away they are from the viewing plane. The convergence of parallel line with distance also is a perspective cue of depth.
- **Sizes of known objects.** The elephant and the tea cup is an illustration of how our mind can understand and recognize objects based on relative size. If they appear to be the same size then the mind automatically assumes the elephant to be further away.
- **Detail.** Closer objects appear in more detail and the surface textures of closer objects are more distinct; whereas distant objects will appear to be less detailed and have less distinct surface textures.
- **Occlusion.** An object that blocks another is assumed to be in the foreground. The occlusion factor is closely aligned with size and perspective to allow the human mind to assume depth.
- **Lighting and shadows.** Closer objects are brighter while more distant objects appear dimmer. A number of other more subtle visual cues are also implied by lighting, the way a curved surface reflects light suggests the rate of curvature. Also, the mind interprets shadows as a form of occlusion.
- **Relative motion.** Objects further away seem to move more slowly than objects in the foreground. The mind uses these cues to interpret depth based on relative movement.

There are other visual cues that are generally not present in 2D images and are peculiar to 3D-stereoscopic situations. They are:

- **Binocular disparity.** This is the difference in the images projected onto the back of each individual eye (and then transmitted to the visual cortex). The disparity in the two

images transmitted to the brain is a direct result of the horizontal separation between the eyes – the inter-ocular or inter-pupillary distance. With 2D images the same image is projected on the back of the eye; whereas with 3D images each eye sees a slightly different perspective view.

- **Accommodation.** This is the ability of the eyes' lenses to change shape in order to focus. This is the muscle tension needed to change the focal length of the eye lens in order to focus at a particular depth. When viewing a stereo scene the eye is focused on the display surface.
- **Convergence.** This is the muscle tension required to rotate each eye so that each is always facing the focal point. Convergence is the inward rotation, while divergence (sometimes referred to as "vergence") is the outward rotation of the eyes, in the horizontal direction, ultimately producing fusion. This provides the ability for each eye to independently center on the corresponding image in a stereo pair while staying focused on the display surface.

In the visual cortex, a stereo pair image is transformed into a "virtual" three dimensional image. Binocular disparity and convergence cues are correct but accommodation cues are inconsistent because each eye is looking at a flat image. The visual system can tolerate this conflicting accommodation to a certain extent.

For most individuals, binocular disparity is considered to be the dominant depth cue. However, if the other depth cues are presented incorrectly they can have a strong detrimental effect on depth perception. In order to render a correct stereo pair one needs to create two images, one for each eye in such a way that when simultaneously viewed they will transmit an acceptable image to the visual cortex where the two distinct views will fused into a single image. If stereo pairs contain conflicting depth cues then one of a number of visual situations may occur:

- A single depth cue can dominant and may not be the correct or intended situation.
- Perceived depth can be exaggerated or reduced.
- Perceived depth can be uncomfortable to view and cause eye strain, headaches or even nausea.
- Perceived depth will be so extreme that the viewer will not be able to mentally fuse the images into a recognizable/understandable stereo scene.

3. Design Considerations for 3D-Stereoscopic Software

Guided by those basic principles of aesthetic and comfortable 3D-stereoscopic viewing outlined above, we have developed a strategy for the generation in real-time visualizations of geospatial data in 3D-stereo; a capability for desktop computer systems equipped with 3D-ready monitors, 3D-ready wide-screen displays or 3D-capable projected displays.

3.1 Window Violations

A window violation occurs when an object is aligned in such a way that when viewed in stereo, part of the object appears to be forward of the window frame and is cut or intersected by the window frame. When this situation occurs only part of an object appears to be within our "house" (closer than the window frame) and the rest, *impossibly* (paradoxically), must be outside somewhere.

The outer limits of a stereo view are defined by the bezel of the video display or border area of the projection screen. When displayed objects in front of the screen plane (in negative parallax) intersect the edge of the window frame. This can lead to contradictory depth cues for the viewer.

Window violation can be prevented by insuring that the entire stereo scene is behind the screen plane (in positive parallax). This situation can further be reduced by applying a narrow transparent mask to the outer limits of each images – a narrow vignette that smoothly transitions from display image to bezel. Another technique that will resolve side window violations is to crop the left edge of the left image and the right edge of the right image. This will reduce the strength of the window violation cues caused by the offset differences in image edge content.

3.2 Ghosting in Stereo Displays

Ghosting: with an anaglyph, when a part of a stereo-side image goes to the other eye, making this other eye viewing a bit of the 2 views. The major ghosts are in red eye, seeing a bit of the cyan view. These ghosting effects often occur in images with strong contrasts, but possibly over all kind of colors.

In modern 3D-stereo, ghosting occurs with polarized glasses technology where one eye sees the persistence of an image from

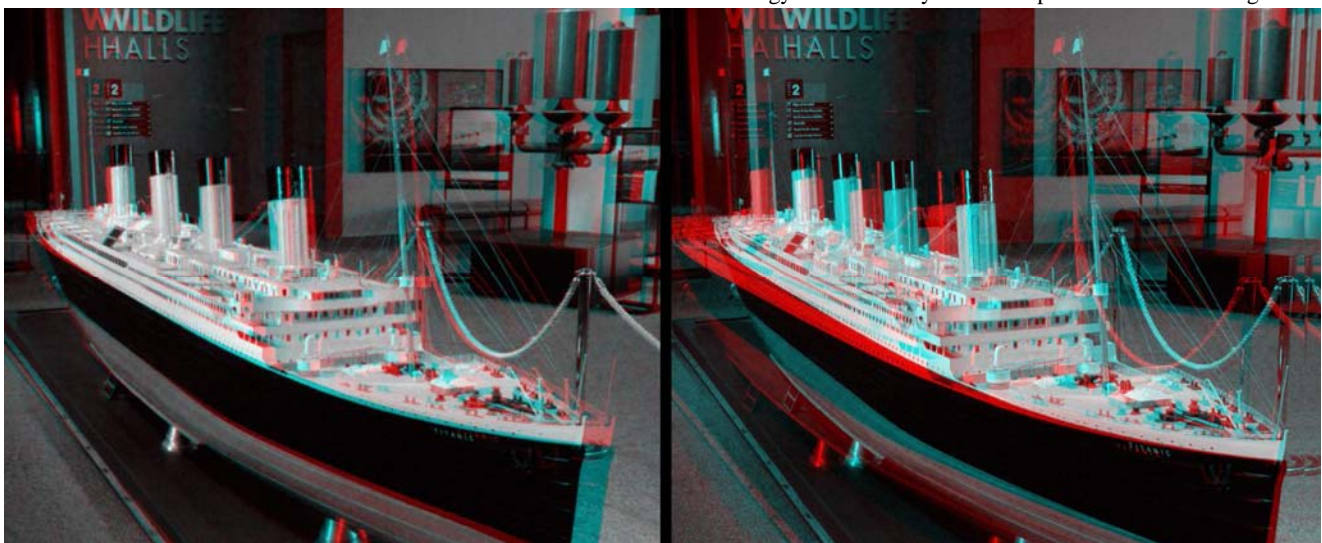


Figure 1. An example of a 3D stereo red/cyan anaglyph image (Left) with window violations and (Right) without window violations (Images courtesy of Michael Beech, with permission)

the detail displayed for the previous eye. With 120 Hz and higher refresh rates, shutter glasses will reduce often eliminate the ghosting effect.

3.3 Retinal Rivalry

While viewing 3D-stereoscopic content, when one eye sees a light area and the other eye perceives a darker area, this can often result in an uncomfortable visual effect at the location in the image when contrasting colors are perceived by each eye. This is called retinal rivalry. This effect can also occur in areas of high parallax and high contrast areas and is often accentuated by particular phosphor colors.

In a post production environment, retinal rivalry can be remedied by changing or attenuating the “offensive” color. However, in a real-time CGI environment this situation must be resolved during the image generation process. This must be achieved by careful application of graphic shader techniques to insure that each eye perceives similar colors within each stereo pair.

3.4 Audience Comfort Zone

An important concept in understanding the impact of 3D-stereoscopic displays is the audience’s “comfort zone”. This is sometimes called “Percival’s Zone”. It is the limit of separation that allows the audience to “fuse” – or view comfortably – the stereo images. Remember that the viewer’s eyes are focused on the screen but converged in front of or behind the screen. Therefore, too great a disparity between focus and convergence will begin to wear out the audience as the brain tries to direct converging and focusing muscles to the same point in space.

As mentioned above, the size of the comfort zone is related to how close the viewer is to the screen. How close the viewer sits is related to the screen size. A viewer sits closer to a home screen than a movie screen. Therefore the comfort zone on a home screen is narrower than on a movie screen. (Also, a child’s comfort zone is narrower than an adult’s.)

3.5 Stereo Parallax

Parallax is the apparent displacement or difference in the apparent position of an object viewed along two different lines of sight. In a stereo image, parallax is the distance between corresponding points in two stereoscopic images. Image pairs with positive parallax appear to be inside or behind the surface of a screen; image pairs with zero parallax appear to be on the plane of the screen; while image pairs with negative parallax appear to “pop-out” in front of a screen (also known as audience space). The range of distances in audience space from the background point producing maximum acceptable positive parallax to the foreground point producing maximum acceptable negative parallax is known as the depth budget, and may be given as a ratio of screen width.

Significant research has taken place and precise methods are now available for calculating of the stereoscopic camera baseline – the stereo base – for cinematography purposes. These methods account for the geometry of both the display/viewer and scene/camera space. However, only limited effort has been

expended to extend stereo base computations into the realm of real-time generation of 3D-stereoscopic video generation.

3.6 Viewing Distance

Viewing distance is an important consideration for visual comfort, particularly for prolonged near point tasks that are typical of desktop computing. Such tasks place demands on both the visual focusing mechanism and the alignment mechanism of the eyes. If either of these mechanisms is overly stressed for long periods of time, visual discomfort can result.

If a stereoscopic image has been created to appear properly on a screen of size X and is projected on to a screen of size 2X, then all objects in the right eye view will be physically shifted double of what they would have been on the screen. This means that parallax will be increased for objects that appear behind the screen surface. At the same time, parallax will decrease for objects that appear to pop-out of the screen. Objects will appear to have greater depth. However, to the untrained eye, this appearance of greater depth will often cause retinal strain which can result in difficulties maintaining stereo vision.

Table 1. Stereo Viewing Distance Versus Stereo Parallax

<i>Inches of Parallax</i>	<i>Distance (Viewer to Screen in %)</i>
2.5	Infinity
2	500
1.5	250
1	167
.5	125
0	100
-.5	83
-1	71
-1.5	63
-2	56
-2.5	50
-3.5	42
-5	33
-7.5	25
-8	24
-9	22
-10	20

Conversely, a stereo image projected for a smaller screen size will have apparent depth distances reduced and as a consequence appear to have less dramatic stereo effects.

Therefore, the screen size for which the image was designed must be a constant consideration during 3D-stereo image generation. For any given viewing distance and screen size, there is some range over which the eyes may converge or diverge and still maintain clear, comfortable and effective stereo vision. In Table 1, this may be measured over a person's range of accommodation and an area may be plotted to show the zone of clear single binocular vision. This is the basis for a set of viewing distance ranges versus display sizes that has been incorporated in the software design.

3.7 Stereo Base

Integral to the comfortable viewing of 3D-stereoscopic content on real-time display systems is the computation of an appropriate offset of two images in each stereo pair. In order to generate 3D stereo image pairs with adequate parallax – enough parallax to

adequately convey a representative sense of depth; the viewpoints of each image must be positioned in such a way that they will ultimately be at an appropriate distance from the projection plane – the assumed viewing distance. Further, they must contain an appropriate separation distance from each other – the stereo base. However, to ensure viewer comfort, the distance between the two viewpoints must not have excessive parallax that would otherwise cause an uncomfortable viewing experience. Figure 2 presents a graphic depiction of those computations necessary to automatically determine the stereo base of a 3D-stereoscopic scene in real-time. Equations 1-4 provide the generalized mathematical context for the interactive computation of stereo base.

This approach assumes that all objects within a scene will be in positive parallax (inside or behind the screen) and not intrude into the audience space; thus eliminating any possible windows violations. Further, it assumes that the distance of the far objects and the near objects will be computed automatically for each scene and will be adjusted to insure that there are no drastic changes in scene geometry from one scene to the next.

$$b_n = \left[\frac{a' \cdot x \cdot Sb}{a_n} \right] + Sb \quad (1)$$

$$b_f = \left[\frac{a' \cdot x \cdot Sb}{a_f} \right] + Sb \quad (2)$$

$$(b_f - b_n) / 2 = 2 * atan(Sb / 2 / (a_n - a')) \quad (3)$$

$$Sb = d \left[\frac{a_f \cdot x \cdot a_n}{a_f - a_n} \right] \frac{1}{a'} \quad (4)$$

Where:

- T1 and T2 are similar triangles
- T3 and T4 are similar triangles
- $f \approx 22.22$ mm (lens to retina distance)
- a' = Viewing distance based on screen size.
- $a_n = a'$ for zero parallax
- $(b_f - b_n) / 2 =$ Left-/right-eye image offset
- a_f is derived from current scene content.

Since the size of the target display and also its comfortable viewing distance are known, it is then merely a computational effort to solve for the stereo base. This approach will allow the automatic generation of many types of real-time 3D-stereoscopic content destined for single, pre-defined video display systems.

3.8 Data Standards and 3D-Stereoscopic Display Systems

3D-Stereoscopic standards for hardware, software and data formats are currently in a state of instability. The 3D-stereoscopic market place is currently dominated by manufacturer dictated standards. Until this situation is resolved it will be difficult to produce software and user's interfaces that will provide "universal" display of 3D-stereoscopic content. At this time the only appropriate approach to resolve this dilemma is to develop software that depends of advanced graphic card technology to provide the interface to various display hardware systems. Exploiting DirectX and OpenGL firmware interfaces

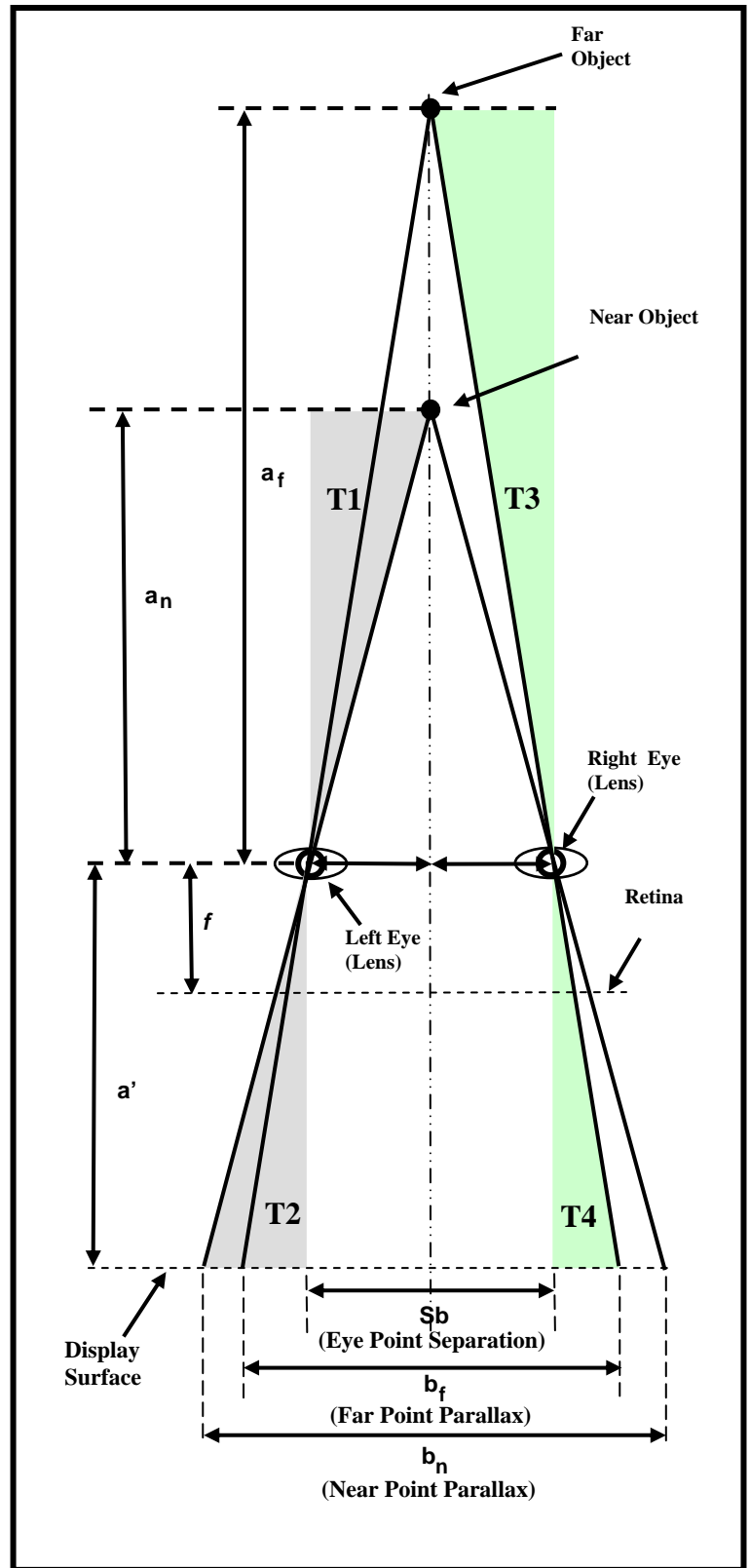


Figure 2. The mathematical relationships of stereo base (Sb) to stereoscopic eye points, stereoscopic object space, and the 3D-stereo display surface.

imbedded in professional quality graphic cards will provide a capability to display 3D-stereoscopic content on most 3D-ready display systems. Until national and international 3D-stereoscopic standards for hardware, software and data content are implemented and universally employed, there will still be situations where it will not be possible to view a specific type of 3D-stereo content on a specific display system. This is not a problem exclusive to 3D-stereo technologies. Even with the “universality” built into the Windows operating system, there will always be data formats that can’t be opened.

3.9 Human Factors – The User’s Interfaces

A major consideration in the 3D-stereoscopic software design process is the incorporation of human factors when developing a user’s interface. To make 3D stereoscopic software compatible with a variety of different hardware configurations, special attention must be paid to user controls. The software must be able to operate on a system configured with single or multiple displays. In a multi-screen configuration, stereo viewing will be dedicated to the 3D capable monitor and a menu with user controls will be display on another monitor. On a single screen configuration, both the stereo display and a floating or docked menu will appear on a single screen.

The user’s interface will consist of a menu containing interactive controls for a wide variety of 3D specific parameters. This will allow the user to interactively adjust the 3D experience without interruption. The following 3D-stereo parameters are essential in the interactive user’s menu:

- Stereoscopic Layout of Video Input
- Stereoscopic Layout of Video Output
- Parallax Adjustment – Scene Depth
- Parallax Adjustment – Background/Foreground Depth
- Screen Aspect Ratio
- Stereo Reversal (left-eye right-eye reversal)
- Stereo/Mono Switching

All of these operator controls will also be activated by user definable and intuitive hot keys – hot keys that do not conflict with standard operating system hot keys.

An intuitive user’s interface and attention to the factors discussed earlier affecting the viewer’s comfort during 3D-stereo interaction will all contribute to a rewarding, enjoyable, compelling and hopefully productive environment.

4. CONCLUSIONS AND FUTURE EFFORTS

With the current instability of hardware, software, and data standards within the 3D-stereoscopic industry it is difficult to predict stable future configurations for 3D-stereo display technology and 3D video content.

We have proposed a method for the real-time generation of aesthetically appealing 3D-stereoscopic video content. We have taken an approach that the proposed 3D-stereo image generation system should be hardware, software and video content agnostic, depending heavily on advanced graphic processor technology.

Therefore, the generation of 3D-stereo video content will be accomplished using computer animation techniques in both a DirectX and OpenGL environment. This approach has been tested extensively using still stereo images and only to a limited extent using computer generated 3D video content. Human factor have been considered in the design of the user’s interface to control the 3D viewing experience. The proposed design employs either a single or dual display configuration; in the dual display mode one display is for viewing 3D-stereo content and the second display is for controlling the 3D viewing experience. In the single screen configuration, both the stereo content and the user’s menu will display on the same screen.

The concepts described in this paper will soon be implemented in several geospatial, social networking and spatio-temporal visualization software tools to allow the generation of 3D-stereoscopic visualizations of geospatial data in real-time. The aim of this research/production effort has been to design a “strawman” approach for future development of aesthetic and user-friendly 3D-stereoscopic visualizations. Although the current effort is centered on geospatial video content, this approach to the generation of 3D-stereoscopic video content should not be limited to only the geospatial domain; but can be applied to all research and entertainment domains that might delve into 3D-stereoscopic visualization.

References

- BAE Systems. 2010. SOCET Set Software Overview. <http://www.socetgxp.com/content/products/socet-set> (accessed 07 Sep 2010).
- Beech, M. 2009. *Digital 3D Stereo Guide*. Michael Beech, Arvada, Colorado.
- Holmes, O.W. 1906. *The Stereoscope and Stereoscopic Photographs*. Underwood & Underwood, New York and London.
- IJsselsteijn, W.A., P. J. H. Seuntjens, and L. M. J. Meesters, 2005. Human factors of 3D Displays. In: *3D Videocommunication: Algorithms, Concepts, and Real-Time Systems in Human-Centred Communication*, Eds. O. Schreer, P. Kauff, and T. Sikora, John Wiley, New York, pp. 219–234.
- Jones, G.R.; D. Lee, N.S. Holliman, and D. Ezra. 2001. Controlling perceived depth in stereoscopic images. *Stereoscopic Displays and Applications Proc. SPIE* 4297A. <http://www.dur.ac.uk/n.s.holliman/Presentations/EI4297A-07Protocols.pdf> (accessed 07 Sep 2010).
- Kooi, F. and A. Toet, 2004. “Visual comfort of binocular and 3d displays,” *Displays*, **25**, pp. 99–108.
- Lambooy, M., W. IJsselsteijn and I. Heynderickx, 2007. Stereoscopic displays and visual comfort: a review. *SPIE Newsroom*. <http://spie.org/x14604.xml?ArticleID=x14604> (accessed 07 Sep 2010).
- Lipton, L., 1997. *Stereographics Developer’s Handbook*. StereoGraphics Corporation.

<http://www.cs.unc.edu/Research/stc/FAQs/Stereo/stereo-handbook.pdf> (accessed 07 Sep 2010).

Wheatstone, C., 1838. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London*, pp. 371-94.