Virtual Reality and Visual Perception by Jared Bendis

Introduction

Goldstein (2002) defines perception as a "conscious sensory experience" (p. 6) and as scientists investigate how the human perceptual system works they also find themselves investigating how the human perceptual system *doesn't* work and how that system can be fooled, exploited, and even circumvented. The pioneers in the ability to control the human perceptual system have been in the field of Virtual Realty. In <u>Simulated and Virtual Realities – Elements of Perception</u>, Carr (1995) defines Virtual Reality as "...the stimulation of human perceptual experience to create an impression of something which is not really there" (p. 5). Heilig (2001) refers to this form of "realism" as "experience" and in his 1955 article about "The Cinema of the Future" where he addresses the need to look carefully at perception and breaks down the precedence of perceptual attention as:

Sight	70%
Hearing	20%
Smell	5%
Touch	4%
Taste	1%

(p. 247) Not surprisingly sight is considered the most important of the senses as Leonardo da Vinci observed: "They eye deludes itself less than any of the other senses, because it sees by none other than the straight lines which compose a pyramid, the base of which is the object, and the lines conduct the object to the eye... But the ear is strongly subject to delusions about the location and distance of its objects because the images [of sound] do not reach it in straight lines, like those of the eye, but by tortuous and reflexive lines. ... The sense of smells is even less able to locate the source of an odour. Taste and touch, which come into contact with their objects, can only gain knowledge from this direct contact" (Kempt, 1989 p. 18). A large part of

Virtual Reality is fooling this 'less deluded' system of visual perception and this paper addresses the how visual perception is used (and abused) in practical Virtual Reality.

Wheatstone and the Stereoscope

One of the most powerful cues in depth perception is binocular disparity – the fact that the eyes each see a different view. These views are fused in the brain in the process known as stereopsis (Goldstein, 2002). Wade (1998) calls is it "one of the supremely psychological phenomena of vision" (p. 235). The idea that two eyes saw more than one was well known and observations to that effect were made by Aristotle, Euclid, and Ptolemy to name a few (Wade, 1998).

The breakthrough discovery leading towards Virtual Reality is published in 1838 by Sir Charles Wheatstone in his work "On some remarkable and hitherto unobserved, Phenomena of Binocular Vision". Wheatstone, knowing that the retinal image of a real object was the same as the retinal image of the object projected on a plane, theorized that the sensation of depth would occur if each eye was presented not with two views of the same 'real' object but with those plane projections or in his case line drawings (Wheatstone, 1838).



Figure1: Wheatstone's Skeleton Figures (Wheatstone, 1838)

To test this he made perspective line drawings (which he referred to as 'skeleton figures') each from the viewpoint of a different eye. To be able to view one image in each eye is tricky – you would have to be able to focus on the (relatively close) images and keep your eyes parallel at the same time. This is contrary to the physiology of the linked systems of accommodation and vergence. In accommodation the muscles in the eye relax when looking far and tighten when looking near. In vergence the eyes converge when looking near and diverge when looking far (Won & Mon Williams, 1998). Accommodation and vergence are connected and work together – vergence can cause accommodation and vice versa, when you focus on something near the eyes turn inward and when you turn your eyes inward your focus near. Wheatstone's original experiments were conducted using 'free-fusion' or 'free-viewing' where the eyes are trained to either look parallel or to completely cross each other in both cases while holding focus (Wade, 1987 & Ogram, 2001). This is difficult to learn. To allow others to be able to view his drawings and achieve the sensation he invented the Stereoscope.



Figure 2: Wheatstone's Stereoscope (Valyus, 1966)

Wheatstone's stereoscope is a refracting or mirror stereoscope where each eye looks forward at two mirrors mounted at 45 degree angles outwards with each mirror facing the right and left images respectively. Since mirrors are used the eyes are focusing farther away and the eyes can easily look parallel ahead (Ogram, 2001).

It is interesting to note that the stereoscope predates the discovery of photography and about six months after publishing his treatise, Wheatstone experiments with stereographic Talbotypes (an early form of photographs) (Wheatstone, 1838). Another interesting observation is made by Wade (1987) in his paper "On the late invention of the stereoscope" since both the ideas of binocular vision and perspective line drawings had been around for hundreds of years. He notes Wheatstone's invention was conceptual and not technical, a factor that often comes into play when integrating perception into Virtual Reality.

While Wheatstone's mirror stereoscope is not in common usage in modern Virtual Reality two other mirror stereoscopes are worth noting – Mirror Viewing and the View Magic viewer. Mirror viewing, uses a single mirror and is one of the simplest and most elegant (and cheapest) ways to view stereoscopic pairs. The left and right images are placed side by side, but the right image is laterally reversed (i.e. it would look right in a mirror). A mirror is placed perpendicular to and diving the images, with the shiny side towards the right. With the nose on the right side of the mirror the eyes look towards the left. The left eye is now looking at the left image and the right eye, which would also be looking at the left image (if the mirror wasn't blocking the view), is now seeing the image on the right side flipped, and stereopsis occurs. (Ogram, 2001) Mirror viewing allows the accommodation and vergence of the eyes to work naturally and the effect is very strong. The device doesn't allow for a very immersive implementation so it is often limited to use in computer visualization and can also be seen in the DK Eyewitness 3D series (Oldershaw, 1999).



Figure 3: Mirror Viewing (Ogram, 2001)

The ViewMagic viewer is basically 2 periscopes with one side pointing up and one side pointing down. The device again allows the eyes to focus naturally and in parallel (because of the added distance created by the mirrors) the images are place above and below each other (Ogram, 2001).



Figure 4: The ViewMagic Viewer

Brewster and Holmes

While Wheatstone's stereoscope proved that stereopsis can be achieved from binocular disparity of images instead of objects, his device is too big to be used in a practical setting and, to avoid ghosting, it needs to use front silvered mirrors which are delicate and costly (Judge, 1928).

In 1844 Sir David Brewster invented the closed box stereoscope known as Brewster's Stereoscope (Moragn & Symmes, 1982). Instead of mirrors, Brewster's stereoscope used wedge shaped prisms to overcome the difficulty of accommodation and vergence. In the Brewster stereoscope accommodation focuses the eyes near and vergence rotates the eyes inward towards the center. The prisms refract the light so that the image on the left 'appears' in this center and the image on the right 'appears' in the center as well. The prisms also offer some magnification as an aid in focusing (Judge, 1928).



Figure 5: Inside Brewster's Stereoscope (Valyus, 1966)

In 1861, Oliver Wendell Holmes, redesigned the viewer to be open and handheld and it became the standard for stereoscopes in the late 1800s (Ogram, 2001). Both the Brewster and Holmes stereoscopes were able to give experiences that isolated and immersed the viewer and become the foundation for the Head Mounted Displays used in Virtual Reality.



Figure 6: The Holmes Stereoscope (Valyus, 1966)

It is interesting to note that while culturally fascinating, the TrueView viewer (1933), The ViewMaster viewer (1938), and the myriad of handheld viewers made popular in the 1950s offered no new innovations as it related to visual perception. The popularity and innovation of these viewers was due to the invention of color slide film. (Morgan & Symmes, 1982).

Head Mounted Displays & Head Tracking

In 1965, Ivan Sutherland (2001) in his article on "The Ultimate Display" describes "a room within which the computer can control the existence of matter. A chair displayed in such a room would be good enough to sit in. Handcuffs displayed in such a room would be confining, and a bullet displayed in such a room would be fatal. With appropriate programming such a display could literally be the Wonderland into which Alice walked." (p.256) A year later he invented the Head Mounted Display (HMD), a direct descendant of Brewster's Stereoscope (Sutherland, 2001).



Figure 7: Sutherland's Head Mounted Display

A stereoscopic head mounted display replaced the two still images traditionally found in a stereoscope with two video displays. The displays are then mounted into a helmet to isolate the viewer and give them a more immersive experience (Iovine, 1995). Ideally, head mounted displays are 'collimated displays' where additional optics are used and the display is placed at the focal point of the optics so image appears at optical infinity (Kalawsky, 1993). These additional optics allow for a "comfortable distance of accommodation" and as "large a magnification as possible" (Kalawsky, 1993).

While the 'collimated optics' allow for comfortable accommodation it also creates a perceptual conflict since the accommodation of the eyes is constant at optical infinity while the vergence varies based on the depth of the virtual scene. The perceptual conflict has been shown to cause eyestrain, nausea, vomiting and physiological changes in the visual system (Wann & Mon-Williams, 1979).



For interactive applications, head tracking technologies are sometimes incorporated into the head mounted displays to allow the user to control their point of view. Trackers determine position in space (X, Y, & Z coordinates) as well as orientation (azimuth, elevation & roll) (Kalawsky, 1993). Head Tracking provides the perceptual system with movement cues on top of the binocular disparity and stereopsis provided by the head mounted display.

Superimposed Viewing / Passive Stereo

"Whereas individual viewing is possible with pictures mounted side by side, as in the stereoscope, the conditions of mass viewing on large screens demand that the pictures be superimposed" (Spottiswoode & Spottiswood p. 1). There are two flavors of superimposed viewing – anaglyph and polarized. This is also known as passive stereo since the viewer is wearing only filtered glasses.

Anaglyph, from the greek work anaglyphos meaning "in low relief", was invented in the mid 1800s and uses complementary-colored filters to allow the eyes to extract binocular disparity from superimposed views of both the left and right images (Morgan & Symmes, 1982). In the anaglyph the right image is projected in red and the left image is projected in cyan – simultaneously on the same screen. The viewers wear glasses with a red left lens and a cyan right lens. The left lens filters out the red image (leaving the cyan), the right lens filters out the cyan image (leaving the red) and each eye is fed the appropriate views to achieve binocular disparity and stereopsis (Ogram, 2001). Over the years various colors combinations have been used with anaglyphs including red/blue red/green red/cyan orange/blue and even yellow/blue.



Figure 9: Early Anaglyphic Projection (Morgan & Symmes, 1982)

While the colored filters cause the separation and binocular disparity they also cause a retinal rivalry – since the images fused are in two different colors and while stereopsis is achieved so is fatigue and it is not useful for long term viewing. Also do to the nature of the colors required for anaglyph images – anaglyphs tend to work best with monochromatic images (black and white) or a very particular color range (Ogram, 2001 & Zelle & Figura, 2004).



Figure 10: Anaglyphic Line Drawings (Girling, 1990)

It is interesting to note that most people assume that the 3D technology used in the 1950s, the heyday of 3D film, was the red/blue anaglyphic process. This is in fact not the case. The technology used in 1950s 3D films is the same one used today in most virtual reality display environments – the polarized system (Orgam, 2001).

Instead of projecting images of two different colors, polarization projects the images in two different orientations of polarized light. The advantage is that the eye can not tell in which direction the light is polarized but the filter can. Polarization was proposed in 1891 by J. Anderton and in his experiments he transmitted two images in which the light from each was polarized perpendicular to the other (Valyus, 1966). Polarized projection systems require the screen to be able to 'hold the polarization' and usually a 'lenticular silver screen' is used since most other surfaces will disperse the polarization. Two projectors are used each with its own polarized filter projecting onto the silver screen. The image on the screen appears doubled until the viewer wears the polarized glasses which then filters out so that the left can see the left image and the right can see the right image (Zelle & Figura, 2004).



Figure 11: Theory of Polarized Glasses (Ogram, 2001)

The effects of a polarized system is very strong – the only drawbacks are that the dark lenses dim the scene, distortions based on viewing angles, and that linear polarization requires that the viewer not tilt their head from side to side. In more advanced virtual reality systems circular polarizing filters are used instead of linear filters that allow for greater head movement. The polarized system allows the viewer to have a relaxed accommodation and a far vergence achieving binocular disparity and stereopsis without any color retinal rivalries. The biggest drawback of the polarized system is that it is not a desktop technology – requiring 2 projectors and a special screen and is not good for small scale viewing.

Time Multiplexed Viewing / Active Stereo

For small scale viewing – many virtual reality systems incorporate active stereo systems also known as shutter glasses. Shutter glasses take advantage of persistence of vision using a technique called time-multiplexing. Shutter glasses have an electronic/mechanical aperture in each eye that toggles open and close very fast so that the left eye and the right eye are never open at the same time. The idea is based on traditional film where one image is displayed at a time, but at a rate high enough that the brain fuses these stills together as motion. In this case the film displays the left images and then the right image in sync with the high speed shutters of the glasses. The effect of which is that the right eye only sees the right stills and the left eye only sees the left stills. This is done at high enough speeds that the right and left eyes end up seeing the appropriate binocular films to achieve stereopsis (Iovone, 1995).



Figure 12: Mechanical Shutter Glasses (Valyus, 1966)

The concept of shutter glasses is perceptually fascinating since the brain is able to achieve stereopsis from a time shifted binocular disparity – the right and left eyes are not being shown the images simultaneously but 'close enough'. This phenomenon was studied in 1963 by Ogle in the paper "Stereoscopic Depth Perception and Exposure delay between Images to the Two Eyes" (Ogle, 1963). Funnily enough, this technology was used commercially as early as 1922 in the Teleview system using a mechanical shutter predating the commercial use of polarization (Morgan & Symmes, 1982).



Figure 13: The Teleview System (Morgan & Symmes, 1982)

There are limitations to using mechanical apertures in glasses (Valyus, 1966) and modern shutter glasses use liquid crystal displays that toggle between clear and opaque at extremely high refresh rates – virtually eliminating all flicker from the shutter (Iovine, 1995).

While active stereo represents the most affordable, passive stereo the most group viewable, and head mounted displays the most immersive virtual reality experiences, there are several other phenomena with strong perceptual ties that are worth mentioning while their value in mainstream virtual reality is still mostly regarded as curiosities.

Chromostereoscopy

Chromostereoscopy is best demonstrated by the physiological chromostereoscopic effect in which a viewer, when presented with an image containing areas of red and blue against a black background will see the red areas closer than the blue ones. This is because of the "inherent chromatic dispersion of the eye" and causes a mild "color-depth" ordering effect. (Steenblik, 1993) This effect was fist observed by Einthoven in 1885 (Cucchi, Massari, & Piano, 2003).

This effect can be greatly exaggerated by using prismatic optics that bend red more than green and green more than blue. When a viewer wears these prismatic glasses, red items cause the eyes to converge more than blue items and an artificial disparity is formed (Bailey & Clark, 1999).



Figure 14: Chromostereoscopy and the eye (Cucchi, Massari, & Piano, 2003)

In computer graphics chromostereoscopy is used to color items based not on their natural colors but on their distance from the viewer. A viewer wearing these prismatic glasses would have a full depth effect while the viewer without the glasses would simply see an image that was colored in a unique way without seeing a double image.



Figure 15: Paris in chromostereoscopic 3D (Reynaud, Tambrin, & Timby, 2000)

ChromaDepth is a patented thin sheet plastic optic that produces this effect without the need for expensive prisms and is the standard when working with the chromostereoscopic effects (Bailey & Clark, 1999).

The Pulfrich Effect

The Pulfrich Effect is a staple of psychological demonstrations. In the demonstration, a pendulum that swings from side to side appears to swing in a circular motion when one of the eyes is covered by a dark lens. The psychological explanation is that "the much-reduced light received by the one eye is analyzed more slowly than the light received by the other eye" (Ninio, 2001 p. 55). Because of the time delay the binocular disparity suffers from a time disparity with one eye lagging behind the other – the effect is the sensation of depth of the pendulum moving

away from the viewer in one direction and towards the viewer in the other and the overall effect being circular.



Figure 16: The Pulfrich Illusion (Ninio, 2001)

To exploit this technique for virtual reality requires an understating of how stereoscopic video is acquired. Stereoscopic video is shot using two video cameras separated by the distance of the eyes, however it is also possible to shoot stereoscopic video with one camera, under a very particular circumstance. Imagine the original two video cameras on a long track and the cameras pointing not forward but to the right. If the cameras were moving at a steady rate then the second video camera would see almost the same as the first video camera but a moment later. Under those circumstances the second camera could be eliminated and replaced with the video feed from the first video camera with an added time delay (Spottiswoode & Spottiswoode, 1953). This technique has been used in still photography for taking stereoscopic images with a single camera from a train, plane, boat or anything with a lateral movement.

When a viewer watching a video clip of steady lateral motion (where the video is constantly panning) wears glasses were one lens is clear and the other is dark (called Pulfrich Glasses) the Pulfrich effect displaces the perception of the video of the darker lens which causes binocular disparity and stereopsis. In effect the right eye (the dark lens) sees what the left eye sees but a moment later becoming the 'phantom' second video camera. The speed of the lateral motion affects the strength of the depth and it goes away entirely when the motion is stopped. It has been used commercially in several 3D broadcasts as the video looks normal to those without the glasses (McAllister, 1993).



Figure 17: Pulfrich Video (Ogram, 2001)

Autostereoscopy

Autostereoscopy for some is the 'holy grail' of virtual reality systems, as it allows the viewer to achieve stereopsis without the means of special devices or glasses. And while

autostereoscopic displays are not invasive to the viewer the technology is still there but displaced closer to the image surface (Valyus, 1966).

As in all stereoscopic systems, to achieve binocular disparity and stereopsis each eye must see a different image. The two main types of autostereoscopic displays are raster barrier and lenticular and both work off from similar principles. The two stereoscopic views are thinly slices vertically and interlaced together. The 'raster barrier' or 'lenticular lens array' restricts viewing of the appropriate strips based on the angle or vergence of the eye. For this to work successfully the viewer must be at a 'sweet spot' and usually a wide range of stereoscopic pairs are combined to create more 'sweet spots' – allowing the viewer to move their head from side to side. What is critical in autostereoscopic displays is the minds ability to assemble these discontinuous strips into a whole image (Ogram, 2001).

Computer graphics specialists have been working on computer displays with lenticular sheets and raster barriers and while some results have been successful for stereoscopic viewing their limitations preclude them from being an immersive virtual reality technology. (Halle, 1997).



Figure 18: Anatomy of a Lenticular Display (Ogram, 2001)

Monocular Stereoscopy

Lastly, one of the most intriguing of perceptual ideas is that of monocular stereoscopy. This is the sensation of depth that is felt by using one eye. In perceptually ambiguous situations the brain 'fills in the blanks' and it can create the sensation of stereopsis even without binocular disparity.

To augment this sensation, Karl Ziess patented his synopter in 1907. The synopter is an aid in fooling the brain into 'filling in the blanks' by feeding both eyes the same image. The binocular disparity feeds two identical images to the brain and stereopsis occurs if the brain has an idea of what it is seeing. "The photographs seen with the synopter were perceived with a relief that, for some observers, was even more vivid that true stereoscopic relief" (Ninio, 2001 p. 173).



Figure 19: Anatomy of a Lenticular Display (Ninio, 2001)

Conclusion

As an industry, virtual reality owes its entirety to the study of visual perception as advances in the perceptual sciences offer new opportunities to fool, circumvent, and explore the perceptual system. Virtual reality demonstrates how important even the most bizarre or seemingly trivial perceptual phenomenon can be both in understanding perception and in building a new one.

Bibliography

Bailey, Michael & Clark, Dru (1999, August) Using ChromaDepth to obtain Inexpensive Singleimage Stereovision for Scientific Visualization. *Journal of Graphics Tools*, Volume 3, Number 3, p. 1-9.

Carr, Karen & England, Rupert ed. (1995) *Simulated and Virtual Realities – Elements of Perception* Briston, PA: Taylor and Francis Inc., p.5.

Cucchi, Franco & Massari, Giancarlo & Piano, Chiara (2203) La Cromostereoscopia Come Supporto Per L'analisi Degli Elementi Geomorfologici Dai Modelli Digitali Del Terreno: Alcuni Esempi a Scala Regionale. <u>http://ggaci.units.it/pubbli.html</u>

Girling, Arthur N. (1990) *Stereoscopic Drawing – A theory of 3-D Vision and Its Application to Stereoscopic Drawing*. London: Arthur N. Girling, plate 8.

Goldstein, Bruce E. (2002) Sensation and Perception: Sixth Edition. Wadsworth – Thompson Learning, p. 6, 223-234.

Halle, Michael (1997, May) Autostereoscopic Displays and Computer Graphics *Computer Graphics*, p.59-63.

Heilig, Morton (2001) The Cinema of the Future. In Packer, Randall and Jordan, Ken (Eds.) *Multimedia from Wagner to Virtual Reality* New York: W.W. Norton & Company, p. 247.

Iovine, John (1995) Step into Virtual Reality. New York: Windcrest/McGraw-Gill, p. 26, 166.

Judge, Arthur W. (1928) *Stereoscopic Photography*. Boston: American Photographic Publishing Company, p. 85.

Kalawsky, Roy (1993) Critial Aspects of Visually Coupled Systems. In Earnshaw R.A., Gigante M.A., and Jones H. (Eds.) *Virtual Reality Systems*. London: Academic Press, Harcourt Brace & Company, p. 204.

Kalawsky, Roy (1993) *The Science of Virtual Reality and Virtual Environments*. Wokingham, England: Addison-Wesley, p.109-110.

Kempt, Martin (Ed.), (1989) *Leonardo On Painting – An anthology of writings by Leonardo da Vinci with a selection of documents relating to his career as an artist.* New Haven: Yale University Press, p.18.

McAllister, David F. (1993) 3D Hardcopy. In McAllister, David F. (Ed.) *Stereo Computer Graphics and Other True 3D Technologies*. Princeton NJ: Princeton University Press, p. 30.

Morgan, Hal & Symmes, Dan (1982) *Amazing 3D*. Boston: Little, Brown & Company p. 10, 14, 15, 16, 18, 26, 32.

Ninio, Jacques (2001) *The Science of Illusions* (Franklin, Phillip Trans.), Ithaca: Cornell University Press, p.55, 173-174.

Ogle, Kenneth N. (1963) Stereoscopic Depth Perception and Exposure Delay between Images to the Two Eyes. *Journal of the Optical Society of America* Volume 53, Number 11, p.1269-1304.

Ogram, G.R. (2001) *Magical Images – A Handbook of Stereo Photography*. Stafford: Self Published, p. 47, 48, 52, 54, 60, 63, 103, 120, 121, 122, 149, 155, 173.

Oldershaw, Cally (1999) Rocks and Minerals Eyewitness 3D. England: DK Publishing, p. 2.

Reynaud, Francoise & Tambrun, Catherine & Timby, Kim (Eds.) (2000) *Paris in 3D - From stereoscopy to virtual reality 1850-2000*. Paris: Booth-Clibborn, p.272.

Robinnet, Warren & Rolland, Jannick P. (1993) A Computational Model for Stereopscopic Optics. In Earnshaw R.A., Gigante M.A., and Jones H. (Eds.) *Virtual Reality Systems*. London: Academic Press, Harcourt Brace & Company, p. 57.

Spottiswoode, Raymond & Spottiswoode, Nigel (1953) *The Theory of Stereoscopic Transmission & its application to the motion picture*. Berkley: University of California, p. 1, 50-51.

Steenblik, Richard A. (1993) Chromostereoscopy. In McAllister, David F. (Ed.) *Stereo Computer Graphics and Other True 3D Technologies*. Princeton NJ: Princeton University Press, p. 183.

Sutherland, Ivan (2001) The Ultimate Display. In Packer, Randall and Jordan, Ken (Eds.) *Multimedia from Wagner to Virtual Reality* New York: W.W. Norton & Company, p. 253, 256.

Valyus, N. A. (1966) Stereoscopy. London: The Focal Press, p. 101, 102, 106,111, 117.

Wade, Nicholas (1987) On the late invention of the stereoscope. Perception Vol. 16, p. 785-818.

Wade, Nicholas (1998) A Natural History of Vision. Cambridge: MIT Press, p. 235, 241-242.

Wann Jon P. and Mon-Williams, Mark (1997, May) Health Issues with Virtual Reality Displays: What We Do Know and What We Don't. *Computer Graphics*, p.53-56.

Wheatstone, Charles (1979) *The Scientific Papers of Sir Charles Wheatstone*. London: The Physical Society of London, p. 225, 235, 268.

Zelle, John M. and Figura, Charles (2004, March 3-7) Simple, Low-Cost Stereographics: VR for Everyone *SIGCSE* p.348-352.