Executive Briefing:
Basic Visual Perception Concepts Related to 3D Movies

Phil Lelyveld
Program Manager
Consumer 3D Experience Lab
Entertainment Technology Center at USC

June 2009

Preface

1) Looking at an image on a 3D display (cinema, TV, laptop) is not the same for your eyes as looking at the real world.

> One reason relates to the mismatch between the object *appearing to be* in front of or behind the screen while the object *is in focus* onto the surface of the screen.

2) When producing 3D content, define and work within a "safe" depth budget in-front of and behind the screen.

> There are guidelines and limits for a comfortable viewing experience based on screen size and viewing distance. It would be useful for the industry to develop standards and guidelines that optimize the depth budget(s) for consistent digital workflows and consumer experiences.

3) Some effects simply don't work the same way in 3D as they do in 2D.

> We should aim to provide guidelines on what works well in 3D and what doesn't. This is the emerging language of 3D. The ETC is developing a Standard Test and Evaluation Material (STEM) reel for this purpose.

4) Audience members will vary in their response to 3D. While most will find 3D easy to watch and more engaging than 2D, research has found that a small percentage of the population will either not see the 3D effect or find it uncomfortable. Some individuals who are in that small percentage are vocal critics of 3D.

> It would be useful to provide resources to inform the public of this issue, and develop tools that individuals can use to self-identify and avoid a disappointing and possibly unpleasant experience. The ETC is researching the criteria for this self-identifying tool.

Consumer 3D Experience – basic concepts and guidelines

This is a brief executive primer on 3D movies and human perception. It is intended to cover the basic terms and concepts behind how we see 3D movies and what to watch out for when they are created and displayed. Links and references are provided at the end for those who want a more detailed overview (ref. 1 and 2).

Binocular vision

Our brain gets its visual information about the real world through our eyes. Because the eyes are approximately two inches apart, each eye "sees" and sends a slightly different signal/angle to the brain. The brain understands the difference between those two views as cues for depth, and automatically fuses those two images to get a "center" view, which was not actually seen by either eye. Hold a finger in front of your face and alternate between two eyes open and one eye shut to see this process in action.

Two key terms used by vision researchers for how our eyes capture three-dimensional information are vergence and accommodation.

Vergence is the angle one of your eyes turns relative to the other eye so that they both look at (aka converge on) the object that you want to see. When you look at the horizon the vergence is zero. When you look at something close to your face the vergence is significant.

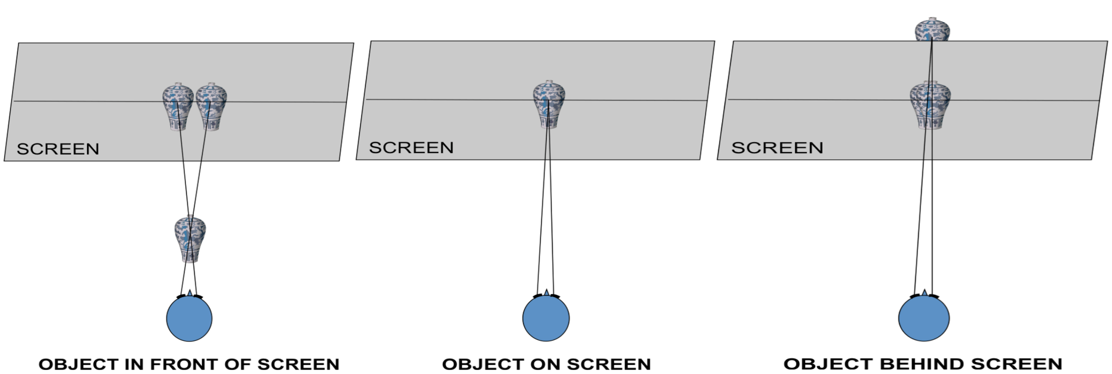
Accommodation is the act of focusing your eyes so that you see what you are looking at clearly.

Vergence-accommodation conflict

In the natural environment, the distance at which your eyes converge is the same as the distance at which your eyes should focus. This is not the case with stereoscopic 3D, where the images for both eyes are projected as two separate images on a screen.

Assume we project a 3D object that is meant to appear to be in front of the screen. Your left eye turns to look at the left-eye image of the object, and your right eye turns to look at the right-eye image of the object. Put together, your eyes converge (vergence) as if the object exists in front of the screen.

As your eyes converge, your brain sends instructions to the eyes to focus the way they normally would for a real object at that convergence distance (accommodation). But in 3D movies the "object" is in focus on the screen, which is behind this convergence point. So your brain keeps working your eyes until the "object" is in focus. That inescapable difference between how we naturally see the real world and how we see 3D movies is called the **vergence-accommodation conflict**.



The vergence-accommodation conflict also occurs if the object is meant to appear behind the screen. Only when the 3D object is meant to appear on the screen itself is there no vergence-accommodation conflict, because your eyes converge on the point where the image is indeed in focus.

An audience member's ability to deal with this vergence-accommodation conflict over the duration of a movie is impacted by how flexible the lenses of their eyes are (ref. 3) and how well his/her brain reacts to the conflict. To maximize the enjoyment of 3D for the entire audience, one guideline is to recognize that this conflict exists, and to give considerable thought to the impact on the audience before having objects jump or scenes cut rapidly and repetitively in and out of the screen.

Research has been done to quantitatively define the comfort zone for the vergence-accommodation conflict. Eye flexibility, often a function of age, is a factor (ref. 3).

Comfortable viewing and vergence-accommodation conflict

According to Prof. Martin Banks, Professor of Optometry and Vision Science at U.C. Berkeley, the vergence-accommodation conflict should be kept at less than ½ to 1/3 diopters for the majority of a 3D viewing experience to avoid discomfort and fatigue.

Diopter is a term that is widely used in vision science research. It is useful for understanding how the depth component of 3D content works. We normally think in terms of distance from the person to the screen. Diopter is the inverse of that; 1/distance (in meters) to the screen.

The practical impact of keeping the vergence-accommodation conflict less than 1/3 diopters is that for a person sitting 10 meters (32.8 feet) from the screen, the effect should come no closer than 2.31 meters (7.6 feet) from the person.

Diopters help us understand how the comfortable viewing range changes as a function of viewing distance. For the same 1/3 diopter limit;

How far the person is from the screen	How close to the person the object appears	How far in front of the screen the object appears
5 meters (16.4 feet)	1.875 meters (6.2 feet)	3.125 meters (10.2 feet)
10 meters (32.8 feet)	2.31 meters (7.6 feet)	7.69 meters (25.2 feet)
20 meters (65.6 feet)	2.61 meters (8.6 feet)	17.39 meters (57.0 feet)
Laptop distance from person to screen		
0.25 meters (9.8 inches)	0.23 meters (9.1 inches)	0.02 meters (0.8 inches)
0.5 meters (19.7 inches)	0.43 meters (16.9 inches)	0.07 meters (2.8 inches)

This table shows that the comfortable viewing range is larger for a person sitting further away from the screen than for a person sitting closer to the screen. The same 3D effect that extends from just barely in front of the screen to infinity when viewed on a laptop appears to be from 57 feet or more in front of the screen to infinity when viewed from the back of the theatre.

Perceptual distortion due to incorrect viewing angle

Your brain compensates for the distortions caused by viewing a 2D image (e.g., a painting) at an oblique angle by using the images from both of your eyes to recognize and compensate for the angle of the surface of the image.

Assume that you are sitting in the best viewing position in the theatre or your home and watching 3D content. As you move away from the centerline of the screen, either to the left or the right, the 3D object becomes increasingly distorted. Different seating positions provide a different 3D viewing experience! Your brain cannot compensate for viewing a 3D projected object at an extreme angle to the screen (ref. 4). This may be an inescapable attribute of physics and human visual perception.

Interpupillary distance (IPD)

Interpupillary distance is the lateral separation between the left and right eyes. The majority of adults have an IPD of between 5.5 and 7.0 cm. Children have a narrower IPD, with the majority greater than 4.0 cm (ref. 5).

As objects move from up close to infinitely far away, your eyes move from converging on a point to looking in parallel toward infinity.

Part of the emerging language of stereography will be establishing recommended practices for the offsets to use when establishing the deep/distance portion of the 3D image. This decision will be influenced by assumptions of both the measured distance of the offset on the screen and the viewer's distance from the screen. If the offset is too great, as might be the case for a person sitting too close to a theatre screen that is larger than the screen size anticipated by the stereographer, then the 3D effect will induce that person's eyes to diverge (e.g. turn outward in opposite directions) to see the image clearly. This is unnatural and uncomfortable, and some people are completely incapable of doing it. Yet the same offset will provide a pleasant viewing experience for someone sitting farther away from the screen in the same theatre. And the same offset in the source material, when displayed on a home theatre screen that is a fraction of the theatre screen size, will produce a shallower image. A recommended practice may be something as simple as developing a table of screen size versus minimum distance from the screen to the front row for a given screen size and telling the theatre owner to only allow audience members to sit at or further than X feet from the screen.

Depth of field

The previous section on IPD flows into the question of how to handle the distant background in 3D images. Often, filmmakers working in traditional 2D will use shallow depth of field to draw audience attention to an actor or object, while leaving the rest of the scene blurry/out of focus. Some researchers believe that people are more likely to explore a 3D image than they are a 2D image. Shallow depth of field could exacerbate eyestrain and fatigue if viewers attempt to focus on parts of the screen that they cannot bring into focus no matter how hard the brain tries to accommodate. On the other hand, increasing the depth of field in a 3D movie increases the work you are encouraging the audiences' brains and eyes to do. The director of the 3D movie Coraline made the artistic decision to selectively use shallow depth of field in the depth script (ref. 6). From the vision science perspective, research is needed on the relationship between depth of field, stereoscopic 3D imagery, and discomfort and fatigue. Part of developing the language of 3D stereography for filmmaking will be learning how to use depth of field. The language will evolve as audiences get past the novelty of 3D and become familiar with the conventions of the 3D experience.

Depth budget is the amount of depth in and out of the screen that you plan to or are able to use.

Depth script is a script/score/timeline describing how the 3D space is used over time; how to pace the action in the third dimension. It is here that you would map out where to use an extreme 3D effect relative to the ambient 3D depth value, and how to block scenes so that cuts between shots do not draw attention to the 3D effect and take the audience out of the story.

Image-pair balancing

If the displayed image-pairs are not perfectly aligned and matched, they will contain visual information beyond the parallax needed to produce the 3D effect, and will contribute to viewer discomfort and fatigue over time. Here are some image-pair balancing concerns to keep in mind.

- There may be creative or technical reasons to have the cameras 'toe in' rather than point forward in parallel. When the two images are combined, the 'toe in' introduces a keystone effect that must be corrected.
- Vertical misalignment, where one camera is slightly tilted 'up' relative to the other, and rotational misalignment, where one camera is capturing an image at a slight clockwise rotation to the other, must both be corrected.

  There is software to correct for these problems (ex. Nuke compositing software from The Foundry, Binocle.com showed prototyped tools for live 3D shoots at 2009 NAB).
- Magnification must match between the images. This is especially critical during zoom sequences.
- Illumination and color balance of the image pairs must match.
- Temporal balancing; Multiple camera-pairs used during live action filming should have lens- and camera-pair imperfections that produce images that edit together well.

Research needed

To advance the language of 3D filmmaking, accelerate the development of 3D products and tools, and expand the market for 3D content and experiences, research to answer these fundamental questions should be conducted;

- Can fundamental principles emerge so that we can produce a trailer with the tag line; *'if you find viewing this to be an unpleasant experience, then you are among the segment of the population that will not enjoy a 3D movie experience [in a theatre, at home, on a personal device, etc.], so please do not watch movies in 3D.'* We know there are people out there who, for any number of reasons, will have a bad experience or will not see the 3D effect. Anything we can do that will help those people self-identify and avoid an unpleasant experience will be extremely useful in sustaining support for 3D entertainment. Stereographers will use whatever fundamental principles emerge creatively, and build on them as they learn how to incorporate 3D into the language of storytelling.
- Will stereographers author once for all display situations, or will there be more granularity based on expected display device (ex. will there be "home version") or some other

parameter? Studios are working on digital workflow, archiving policies, etc. Recommended practices, which may emerge naturally as the industry gains more experience, will be useful.

- People have trouble resolving a 3D image as the horizontal pan rate increases. Filmmakers need more quantitative information regarding pan rate, the brain's ability to resolve the stereo image, and frame rate. Some research in this area is being done by Martin Banks and his students at UC Berkeley.
- Does increasing the amount of light that reaches the eyes significantly impact the vergence-accommodation conflict?
- How does depth of field impact the way people explore a stereoscopic 3D image? Does the impact change with time, both during the viewing of long-form content and over years of viewing experience? Research is needed on the relationship between depth of field, stereoscopic 3D imagery, and discomfort and fatigue.

Additional Resources:

1. "Foundations of the Stereoscopic Cinema" by Lenny Lipton, is available as a free pdf download at http://www.stereoscopic.org/library
2. Prof. Nick Holliman of the University of Durham, UK, has posted an excellent overview of 3D concepts, including shooting 3D images, at http://www.binocularity.org
3. Decline in accommodation with age is plotted in Fig. 3 of "Ocular and Refractive Considerations for the Aging Eye" by Kathryn Richdale, at http://www.clspectrum.com/article.aspx?article=102546
4. Marty Banks' NAB presentation slides will be posted on his website http://bankslab.berkeley.edu/ . They are currently available at www.etcenter.org/files/publications/Marty_Banks_NAB09.pdf)
5. Neil Dodgson of the University of Cambridge Computer Lab, UK, has aggregated data on the range and average of human interpupillary distance (IPD). His paper is available at http://www.cl.cam.ac.uk/~nad10/pubs/EI5291A-05.pdf .
6. Perception and the Art of 3D Storytelling, by Brian Gardner, Creative Cow Magazine, the June, 09 Stereoscopic 3D issue, at http://magazine.creativecow.net/issue/stereoscopic-3d
7. Research on visual fatigue from vergence-accommodation conflict is laid out very well in this paper; "Vergence-accommodation conflicts hinder visual performance and cause visual fatigue", by David Hoffman, Ahna Girshick, Kurt Akeley, and Martin Banks, at http://www.journalofvision.org/8/3/33/article.aspx

Author: Phil Lelyveld is an Entertainment Technology and Business Development Consultant (www.ReelWord.com). He is currently developing the Consumer 3D Experience Lab and program at USC's Entertainment Technology Center (www.etcenter.org). Phil spent 10 years developing and implementing Disney's digital media strategy as the corporate Vice President of Digital Industry Relations. He has been involved in new and emerging media for almost 20 years.

Contributors: the following contributed text, feedback, and edits to this paper:

Martin Banks is Professor of Optometry and Vision Science; Affiliate Professor of Psychology and Bioengineering, UC Berkeley. He and his team have done research into stereoscopic surface perception, virtual reality, binocular correspondence, binocular visual direction, picture perception, self-motion perception, multisensory interactions, and infant spatial vision (http://bankslab.berkeley.edu/). He frequently presents his research to both professional and lay audiences, including the recent Society for Information Display's 3D Technology Update for Display Professionals conference in Costa Mesa, CA (www.sid.org/), and the Digital Cinema Summit (ref. 4) coproduced by the ETC and SMPTE at NAB 2009.

Neil Dodgson is Reader in Graphics & Imaging in the Computer Laboratory at the University of Cambridge (UK), where he is a co-leader of the Graphics & Interaction Research Group (http://www.cl.cam.ac.uk/Research/Rainbow/). His research interests are in computer graphics, 3D display technology, human-figure animation and image processing. He is on the program committee of the annual Stereoscopic Displays and Applications (SD&A) conference (www.stereoscopic.org).

Nick Holliman is Senior Lecturer, Department of Computer Science, Durham University, Durham, England (http://www.dur.ac.uk/n.s.holliman/). His research into digital imaging covers 3D computer graphics, computer vision and visualization technologies. He specializes in interdisciplinary research investigating the theory, human factors and application of (auto-) stereoscopic 3D displays. The **Durham Visualization Laboratory**, which he founded in January 2004, supports these and other inter-disciplinary 3D research themes. It contains a wide range of 3D displays, tracking/interaction devices and 3D capture systems (http://www.dur.ac.uk/n.s.holliman/3dDurhamDisplays.html). He is cochair of the annual Stereoscopic Displays and Applications (SD&A) conference (www.stereoscopic.org).

Phil Captain 3D McNally is the Stereoscopic Supervisor for Dreamworks' Kung Fu Panda, Monsters vs Aliens, How to Train Your Dragon, and Shrek Goes Fourth. Prior to that he was Stereoscopic Supervisor for Disney's Meet the Robinsons, and has worked on other features at Disney and ILM. He is a graduate of the Royal College of Art in London and has over 15 years of 3D experience. In the 1990s he legally changed his middle name to Captain 3D.

John Merritt is CTO of The Merritt Group (http://www.merritt.com/), A Fellow of the Society of Photo-optical Instrumentation Engineers (SPIE), and cochair of the annual Stereoscopic Displays and Applications (SD&A) conference (www.stereoscopic.org).

David Wertheimer is the CEO and Executive Director of the ETC (Entertainment Technology Center) at USC. Prior to ETC, David was president of Paramount Digital Entertainment, CEO of WireBreak Networks, and worked with Steve Jobs at NeXT and Larry Ellison at Oracle.

Andrew Woods is a Research Engineer at Curtin University of Technology's Centre for Marine Science and Technology (CMST) (Australia). His research interests include Stereoscopic Imaging, Stereoscopic Video, Underwater Technology and Remotely Operated Vehicles (ROVs) (http://3d.curtin.edu.au/). He is cochair of the annual Stereoscopic Displays and Applications (SD&A) conference (www.stereoscopic.org).

Ray Zone is a leading champion of 3D (www.ray3dzone.com/). For 25 years, through his company The 3-D Zone, Ray has been converting flat art to 3-D for every conceivable application. The client list includes Warner Brothers, Walt Disney Company, A&M Records, Saban Entertainment, Galoob Toys, and many others.