

5-2009

THE DESIGN, PRODUCTION AND ANALYSIS OF A REALISTIC STEREO CG SHORT FILM ON A SIX MONTH BUDGET.

Celambarasan Ramasamy
Clemson University, celu_jn31@yahoo.com

Follow this and additional works at: https://tigerprints.clemson.edu/all_theses

 Part of the [Film and Media Studies Commons](#)

Recommended Citation

Ramasamy, Celambarasan, "THE DESIGN, PRODUCTION AND ANALYSIS OF A REALISTIC STEREO CG SHORT FILM ON A SIX MONTH BUDGET." (2009). *All Theses*. 605.
https://tigerprints.clemson.edu/all_theses/605

This Thesis is brought to you for free and open access by the Theses at TigerPrints. It has been accepted for inclusion in All Theses by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.

THE DESIGN, PRODUCTION AND ANALYSIS OF A REALISTIC STEREO CG
SHORT FILM ON A SIX MONTH BUDGET.

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Fine Arts
Digital Production Arts

by
Celambarasan Ramasamy
August 2009

Accepted by:
Dr. Donald House, Committee Chair
Dr. Tim Davis
David Donar
Dr. Andrew Duchowski

ABSTRACT

The production of stereoscopic CG films poses some interesting challenges, especially for student productions that work under the severe limitations of time and resources. This is mainly due to the non availability of off the shelf production tools catering to stereoscopic CG productions. This work presents the production process of one such student produced stereoscopic short film. The production process is described in detail starting from the initial conception of the narrative plot to the actual production of the film. Finally an experimental technique of using eye tracking as a tool for finding out the effectiveness of the various stereoscopic framing techniques used in the film is presented. The feasibility of eye tracking as an effective tool for filmmakers in stereoscopic 3D to analyze the viewing behavior of the audience and to improve the film using that information is assessed. This paper first provides the basic background needed to understand the various terms related to stereoscopic 3D. Then it describes the custom stereoscopic pipeline that was implemented for the film, followed by an in depth description of the actual production process. Finally, the eye tracking experiment is described in detail and the analysis of the result is presented.

TABLE OF CONTENTS

	Page
TITLE PAGE.....	i
ABSTRACT.....	ii
LIST OF FIGURES.....	iv
CHAPTER	
I. INTRODUCTION.....	1
II. BACKGROUND.....	5
Binocular Vision and Stereopsis.....	5
Accommodation, Convergence and Disparity.....	6
Visual Field and Visual World.....	8
Stereoscopic Film: A Breakdown.....	9
III. METHODS.....	11
Stereoscopic Production Pipeline.....	11
Creative Statement.....	19
Film Production.....	22
Working with Stereoscopic 3D.....	26
Eye Tracking Study.....	31
IV. DATA ANALYSIS.....	34
Questionnaires.....	34
Analysis of Videos.....	36
V. CONCLUSION.....	53
REFERENCES.....	55

LIST OF FIGURES

Figure	Page
1 The anaglyph image shows the disparity between the left and right eye image.....	8
2 Red/Cyan anaglyph glass.....	10
3 Parallel Lens Configuration.....	12
4 Cross Eyed Lens Configuration.....	13
5 Curved stereo space.....	14
6 Focal Plane Depth of Field.....	15
7 Stereo Frame and its corresponding horizontal disparity map.....	17
8 A Stereo frame and its corresponding vertical disparity map along with a diagram showing the distribution of vertical disparity in a frame.....	19
9 Screenshots from the story reel.....	23
10 The remodeling process.....	24
11 Early tests of stereoscopic pipeline.....	26
12 Enhancing stereoscopic depth through layering of objects in the scene.....	27
13 Stereo frame with foreground object and without foreground object.....	29
14 Living Room shot before transition, Living Room shot after transition, Last shot of the film with gun in focus, Last shot of the film with gun out of focus.....	31
15 Heat map visualization with standard deviation circles.....	33
16 Test images for the first study question. The image on the left was labeled image 1 and the image on the right was labeled as image 2.....	34

List of Figures (Continued)

Figure		Page
17	Test Images for the second study question. Image on the top was labeled image 1 and the image on the bottom was labeled image 2.....	35
18	Screenshots from the animation clip in which both the convergence point and the depth of field are animated together.....	39
19	Screenshots from the animation clip having a fixed convergence point and animated depth of field.....	40
20	Screenshots from the animation clip with fixed depth of field and animated convergence point.....	41
21	Screenshots from the animation clip with animated convergence point having no depth of field.....	42
22	Screenshots from the Living Room shot in the stereoscopic film.....	46
23	Screenshots showing the focus shift between trapdoor and gun.....	47
24	Heat map of the tunnel sequence from stereoscopic and non-stereoscopic versions of the film.....	49
25	Living room shot from the first and the second viewing of the stereo film.....	50
26	‘Eye’ mode clip with flicker.....	52

CHAPTER ONE

INTRODUCTION

Over the past couple of decades, the film as a story telling medium has changed dramatically. Everything from the way the plot is handled, to the way the imagery is portrayed on screen has undergone a dramatic change. One of the important factors behind this change has been the openness of the filmmakers for adapting new technology and techniques in order to entertain the audience in interesting and innovative ways. As we trace back the history of change in cinema, it is not uncommon to see that technological innovations often open up new frontiers in story telling, challenging the filmmakers to adapt its strengths and weakness seamlessly into the narrative and ultimately managing to tell a captivating story.

These technologies and techniques often tend to have an interesting life cycle. They go through a phase where they seem to cause more pain than their worth before they 'mature' and get adapted into the world of professional film making. In their early phases they often tend to appear redundant and extraneous to some pre-existing means of film making. For any new technology to see the light of day it needs a passionate group of individuals, a mix of filmmakers, artists and technologists, who are willing to back it up through its darkest times. It is often such groups of people who in retrospect, are considered pioneers of a new age of filmmaking. The efforts by George Lucas in making the 'Star Wars' [Kline, 1999] and the story of Lucasfilm Computer Division and their evolution into Pixar Animation Studios[Paik, 2007] are notable mentions in this regard.

Often, over time, new techniques and technologies tend to grow beyond their originally intended scope. So when evaluating a new technology it is critical to not just evaluate it in its current state, as an alternative to an existing technique, but to also look at the potential future directions it might take. A prime example of this is the role of computer graphics in film making. Starting out as academic visualizations computer graphics soon developed into an efficient alternative for traditional visual effects techniques in live action films. *Jurassic Park* [Spielberg, 1993] is one of the landmark films that set the trend for the role of computer generated effects in live action film making. Before long, computer graphics evolved into a filmmaking medium of its own with some of the greatest films of our times made entirely using it. *Toy Story* [Lasseter, 1995], produced by Pixar Animation studios, was the landmark film that set the trend for other films to follow. Even though these computer generated films started out as an alternative to traditional hand drawn animation, with the amount of research going into the field and the cost of computing power going down rapidly, computer graphics is currently on the verge of attaining the 'holy grail' of photo-realism. This has lead to a number of attempts at making photo realistic computer generated films in recent times, with films like *Starship Troopers* [Sakaguchi, 2001], *The Polar Express* [Zemeckis, 2004] and *Beowulf* [Zemeckis, 2007] being some of the most notable examples. Even though these films are criticized for their current technical limitations in doing photo realistic characters, they are nonetheless precursors to an era of filmmaking in which computer generated films could prove to be a viable and convenient alternative to live

action films. If the past is anything to go by, photo-realistic computer generated films will soon evolve into a medium whose benefits would outweigh its limitations.

Amongst other things, computer graphics is also playing a major role in reviving some of the other techniques from the past that tried to make it into mainstream film making. A notable example is stereoscopic cinema. Before the advent of computer graphics and the digital age, the process of making stereoscopic films was a cumbersome task. Everything from the slight misalignment of the cameras during filming, to the errors introduced during the projection hampered the success of stereoscopic 3D as a viable mainstream story telling medium. But now computer graphics enables us to build mathematically accurate stereo rigs that are perfectly aligned. Computer graphics has greatly helped in simplifying a lot of the problems that are traditionally associated with stereoscopic 3D.

In the 1950s due to the changes in the social and economic structure of the American society, fueled by the introduction of television sets into American households, the count of cinema going audience went on the decline. At this time, there was a major push to bring stereoscopic 3D into mainstream Hollywood films in an effort to bring audiences back to the theaters, but technical difficulties limited its success at that time. Now a similar situation facing Hollywood has revived interest in stereoscopic cinema. With the help of computer graphics and digital projection techniques, stereoscopic 3D has opened up interesting challenges and opportunities for filmmakers to explore.

My goal for this thesis was to undertake the production of a realistic computer generated short film. Through this project, I got a chance to work with stereoscopic 3D

and develop stereo related production tools. The most critical challenge faced in this production was the planning needed to execute a realistic looking film within the constraints of time and resources that I had to work with. In this film, I also got to a chance to experiment with some new stereoscopic framing technique. After the completing the film, I used eye tracking to study the audience's viewing pattern of the movie to study the effectiveness of different techniques and to see what worked and what didn't work in the film.

CHAPTER TWO

BACKGROUND

Arguably stereoscopic 3D is one of the least understood forms of filmmaking that exists today. Over the years, there has been a lot of myth surrounding stereoscopic content creation in movies [Lipton, 2008]. Tracing back the history of stereoscopic films in Hollywood, around 1950s there was major interest in promoting stereoscopic films. This period saw the release of a number of stereoscopic films, *House of Wax* was promoted as the first stereoscopic film produced by a major studio [House of Wax Trailer]. Even Alfred Hitchcock experimented with stereoscopic cinematography in his film *Dial M for Murder* [Wikipedia, 2009]. But because of the sub par nature of most of the films released in stereoscopic 3D, it earned a reputation as just a gimmick. This along with the technical difficulties associated with the production of stereoscopic films prevented them from taking off in a big way.

Stereoscopic film making has always remained a black art. Before trying to understand stereoscopic films one must have a firm grasp of stereoscopic vision in general and the various terms associated with it.

Binocular Vision and Stereopsis

When we see the world around us with our two eyes, we experience binocular stereopsis. It is the ability of our brain to fuse two images of slightly different perspective, enabling us to perceive depth. It has been found that cells in certain regions of our brain seem to respond to simultaneous stimulations from both the left and the right

eye. It is widely thought that it is these cells that are responsible for the stereopsis that we experience. Depending on the amount of distance between the retinal images of a point perceived by the left and the right eye, different subsets of these cells get stimulated. Depending on the distance between the two images, one group of cells seems to be responsible for making the objects appear closer from the screen. While, another group makes the objects appear further away from the screen. But it is still unclear how the brain uses these stimulations to construct the 'true depth' that we perceive [Hubel 1995].

Stereopsis was largely an unknown phenomenon until 1838 when Sir Charles Wheatstone made his publication on the topic [Wheatstone, 1838]. The discovery along with the advent of photography made it a popular medium in the mid 1850s. Conventional cinema exhibits a number of depth cues like perspective, overlapping, motion parallax etc. So in a sense they can be perceived as three dimensional. In addition to these depth cues stereoscopic films exhibit an additional depth cue called stereopsis. This stereopsis phenomenon requires two views, so it cannot be captured in a conventional camera.

Accommodation, Convergence and Disparity

There are three important terms relating to the human visual system that are often used while talking about stereopsis,

Accommodation: The depth at which the visual field is in clear focus is called the lens accommodation.

Convergence: When we look at objects around us, both our eyes converge on the object we are looking at, enabling us to form a clear image of the object at the center of the retina in each eye. This convergence phenomenon is inherent in the way in which we experience stereopsis in the world around us.

Intraocular Distance: The term used to describe the distance between the eyes is called the intraocular distance. In the case of a stereoscopic camera setup, a term called interaxial distance is also used to describe the distance between the cameras.

Disparity: When our eyes converge on an object in the scene the images of that object on the two retinas overlap so we see that object as a single image. But the images of other objects that are either in front of or behind the convergence point do not overlap on the retina and are seen as double images. In other words some 'disparity' is introduced. Disparity is a very important cue in the perceived spatial depth of the object in stereoscopic vision. It is measured as a distance between similar image points as perceived by the left and the right eye. Angle is a good metric to measure disparity.

In the figure 1, the topmost image is a frame from the left eye camera in the movie, while the middle image is a frame from the right eye camera and the bottommost image is an anaglyph created by combining the left and the right eye images. You can see that many of the objects in the anaglyph image are double edged because of their disparity between the left and right eye images.



Fig.1. The anaglyph image (bottom) shows the disparity between the left(top) and right eye image(middle).

In a stereoscopic camera setup, both the distance to the convergence point and the inter-axial distance affect the disparity of the objects in the scene.

Visual Field and Visual World

When we perceive the world around us there are two important terms that we must learn to differentiate, the visual field and the visual world. The visual field can be

understood as a snapshot of our field of vision at any point in time [Lipton, 1982]. We can sense the visual field only when we become aware of what we are seeing. At any instance, the object that we are looking at remains in focus at the center of our visual field and the rest of the scene gradually blurs out as we move outward in the visual field. On the other hand, visual world is a term used to describe the world as we experience it everyday. Unlike the visual field, the visual world is a mental construct that is perceived as always being in focus. Our eye and brain work together to create this sharp mental image of the world around us. When our eyes move, because of the interaction with our brain, our visual world remains static whereas our visual field changes.

Stereoscopic Film: A Breakdown

Unlike conventional films, a stereoscopic movie is filmed using a 'two camera' setup. Each shot is captured from two slightly different perspectives. So finally you end up with two movies, one for each eye. There are a number of different projection techniques for creating the final stereoscopic version from these two source movies, each with its own cost/quality trade off.

Anaglyphic Stereo

This is one of the most common technique for displaying stereoscopic images and films. This is the format that I worked with during the production of this film. In this technique the left and right images are encoded into two different color channels of the image, red/cyan is a common choice. To experience the stereoscopic 3D, this combined

image is viewed with a matching filter over each eye, so in this case, a red filter for the left eye and a cyan filter for the right eye(usually in the form of a red/cyan anaglyph glasses). Figure 2 shows a pair of red/cyan anaglyphic glasses from a front view, on the left is the cyan filter and the red filter is to the right.

All the stereoscopic images in this document are meant to be viewed through a pair of red/cyan anaglyphic glasses.



Fig.2. Red/Cyan anaglyph glass.

The advantage of this technique is that it does not require any special projection techniques and can be displayed on just about any screen, all you need for viewing the image is a pair of 3D anaglyphic glasses. But the disadvantage to this technique is that depending on the exact color encoding technique used, there is usually some color distortion [Anaglyphs Method Comparison, 2005].

CHAPTER TWO

METHODS

Stereoscopic Production Pipeline

Autodesk Maya™ was the primary animation package that I used for this production. After I rendered the frames for the left and the right eye separately using Maya, I converted them into movies using Adobe After Effects™, which is a compositing package. After the individual shots are assembled this way, I used Adobe Premiere Pro™, an editing package, to assemble the movie from the individual shots. At the end of this process I had separate left and right eye movies. In order to create a red/cyan anaglyph version from these two movies, I used a software called Stereo Movie Maker [Masuji, 2008]. This software takes the left and right eye movies as input and outputs an anaglyph version of the movie. I also used an animation package called Autodesk Motion Builder™ to process the motion capture data.

Since I planned to have a considerable number of first person stereoscopic shots in the film, I started working on a custom pipeline based around motion capture data that would enable me to create the first person stereoscopic shots. To accomplish this, I implemented a custom pipeline consisting of Mel scripts, which is a scripting language in Maya, and a stereoscopic camera shader in Mentalray™, which is a powerful renderer that is part of the Maya package.

Stereoscopic Mentalray Shader

There are two main types of commonly used stereo lens configurations, namely *parallel* lens configuration and *cross-eyed* lens configuration. In a *parallel* lens configuration the lens axis of the left and right eye cameras are aligned parallel to each other and the film back of the cameras are offset horizontally with respect to the orientation of the lens. Figure 3 shows a *parallel* lens stereo configuration, both the lens of the left and the right eye cameras remain parallel to each other and only the film back in both the cameras are offset in a direction parallel to the lens axis.

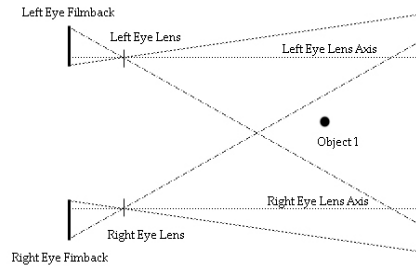


Fig.3. Parallel Lens Configuration

In a *cross-eyed* lens configuration, both the left and the right eye cameras are toed in towards each other such that their lens axes intersect, the point of intersection is called the convergence point. The *cross eyed* lens configuration works similar to the way our eyes work to perceive the world around us. There is an imaginary plane passing through the convergence point and aligned parallel to the inter-axial axis between the left and the right eye cameras. This plane is called the *screen plane*. The objects that lie along the

screen plane appear to lie on the screen or very close to the screen, in the rendered image. While the objects that lie in front of the plane appear to pop out of the screen and the objects that lie behind the convergence point appear to lie behind the screen in the rendered image. *Cross-eyed* lens configuration is a very intuitive way to understand stereoscopic imagery and it integrates very well with a ray tracer. Mentalray being a ray tracer, I decided to implement the *cross-eyed* lens configuration in the Mentalray shader. Figure 4 shows a *cross-eyed* lens configuration, the left and the right eye camera are toed in towards each other, object 1 is in front of the screen plane, so it appears to pop out in the rendered image, while object 2 which is behind the convergence point appears deeper in the rendered image.

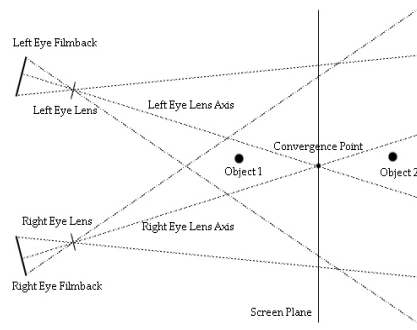


Fig.4. Cross Eyed Lens Configuration

Since the *cross-eyed* lens stereo configuration is very similar to binocular vision produced by our eyes, the locus of points having zero horizontal disparity, also called as the *Horoptyer*, in the field of binocular vision [Lipton, 1982], does not lie completely on the *screen plane*. In fact, the locus of zero horizontal disparity points is a curve passing through the convergence point, whose slope is directly proportional to the inter axial

distance between the left and the right eye cameras. In other words, in a *cross-eyed* lens stereo configuration, the stereo space is curved with respect to the physical space in which the camera is present. In computer graphics, the physical space is usually the Cartesian space of the world coordinate system. Figure 5 shows the curve that represents the locus of zero horizontal disparity points, which indicates a curvature in the stereo space.

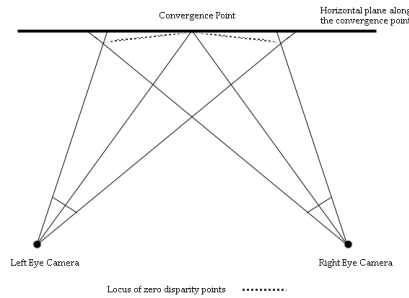


Fig.5. Curved stereo space

The shader can render out the left/right eye images separately or directly render out an anaglyph version of the stereo image. The shader takes in a set of input parameters and uses them to create the projection parameters required for the left and right eye camera for each frame. When mental ray requests the shader to return the color for a pixel, in anaglyph mode, the shader traces a ray through the requested pixel in the left eye camera and then in the same call traces another ray through the corresponding pixel in the right eye camera. The green and blue channel values from the left eye is combined and placed in the red channel of the output color. Then the whole intensity of the red channel of the output pixel is bumped up using gamma correction. The green and blue channel

from the right eye pixel is directly placed in the green and blue channels of the output color. This kind of encoding is called as Optimized Anaglyph encoding. This is just one of the methods to encode the image, there are other ways to distribute the left and right pixel intensities among three channels [Anaglyphs Method Comparison, 2005]. The left and right eye intensities of the anaglyphic image needs to be gamma corrected separately, as our perception of the red intensity is lesser than our perception of cyan (green and blue channels combined). Mental Ray renders the output image by requesting output color from the shader one pixel at a time. Please refer to the appendix for this shader's parameter list and source code.

Features

Other than creating stereoscopic images, this shader has a number of additional features that provides more creative control over the production.

Focal plane depth of field:

In addition to supporting conventional depth of field that controls the blurriness with depth, I also implemented a way to control the blurriness along the field of view of the camera. The blurriness across the field of view is controlled using a two circle setup where you give the radius of the two circles and the amount of blurriness at each circle. In Figure 6, the circle of confusion values $v1$ and $v2$ between the inner and the outer circles are interpolated to get the blurriness values at the various points on the image.

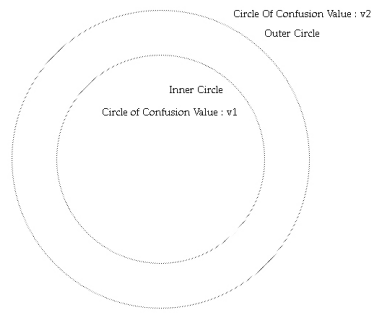


Fig.6.Focal Plane Depth of Field

Spherical lens:

The shader can also simulate spherical lens that's used to create fish eye lens effect.

Encoding Stereo Pixel Disparity:

The shader also has the ability to encode either the horizontal or the vertical pixel disparity of the stereo image in the alpha channel. In other words, along with a color at each pixel, a disparity value is also stored. To enable mental ray to encode this value, one must turn on the 'Pass custom Alpha Channel' flag in 'Custom Entites' under the Mental Ray rendering tab.

The pixel disparity is encoded in a normalized scale between 0.0 and 1.0 to make the values resolution independent. The disparity values are encoded using the following logic. In the normalized scale, the extreme values of 0.0 and 1.0 are used to indicate an invalid pixel. The maximum allowable value that can be encoded is a pixel disparity of a quarter of the image resolution. An alpha value of 0.5 indicates no pixel disparity. The range 0.0 to 0.5 indicates that the left eye pixel is to the right of the corresponding right eye pixel in the stereo image. The range 0.5 to 1.0 indicates that the left eye pixel is to the left of the corresponding right eye pixel in the stereo image.

Figure 7 shows a stereoscopic frame and its corresponding horizontal disparity map. In the horizontal disparity map, because of the alpha encoding, objects at screen depth appear mid grey, objects that pop out of the screen appear darker while the objects that lie further back appear brighter. White indicates invalid disparity for the corresponding image pixels. The teddy bear and the stand appear darker than the rest of the image as they are at screen depth and thus the objects with the least stereoscopic depth in the image.

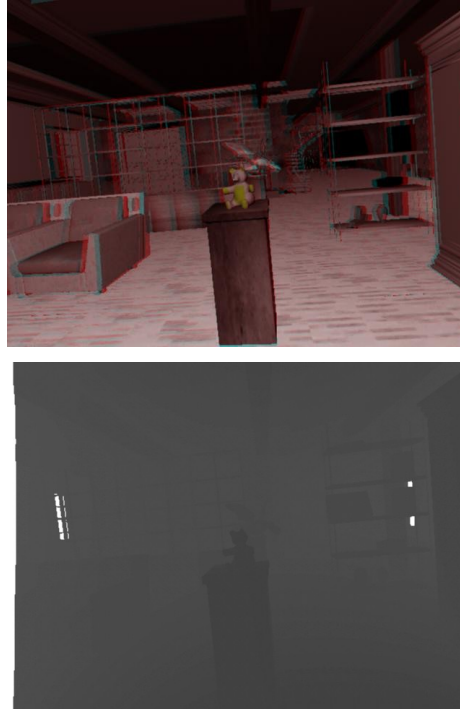
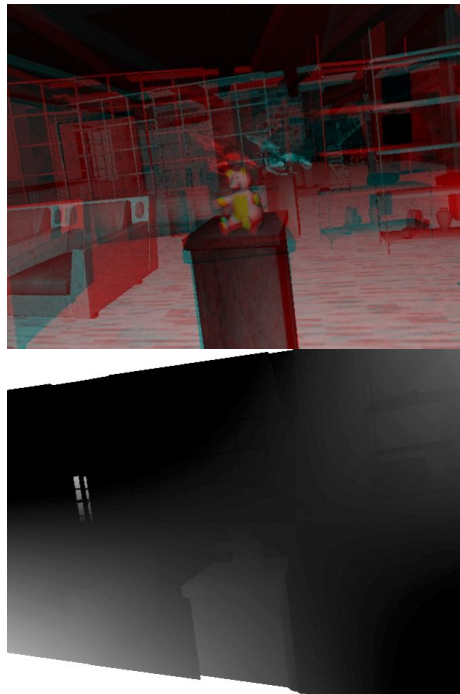


Fig.7.Stereo Frame and its corresponding horizontal disparity map

Since the shader renders stereoscopic images using a *cross-eyed* lens configuration. This setup introduces some vertical disparity in the stereo image due to perspective distortion between the two cameras, especially along the edges of the frame.

Similar to the horizontal pixel disparity, the vertical pixel disparity in the stereo image can also be encoded in the alpha channel. An alpha value of 0.5 indicates no vertical pixel disparity. The range 0.0 to 0.5 indicates that the left eye pixel is to the top of the corresponding right eye pixel in the stereo image. The range 0.5 to 1.0 indicates that the left eye pixel is to the bottom of the corresponding right eye pixel in the stereo image. Figure 8 shows a stereo frame (top) and its corresponding vertical disparity map (middle). As seen in the figure, the vertical disparity varies across the different regions of the frame. The bottommost diagram shows the vertical disparity distribution in a frame that renders a point at infinity. The vertical disparity in the different regions of the frame is labeled according to the color code of the encoding that I implemented in the shader. The vertical disparity is maximum near the four corners of the frame and it is minimum near the center of the frame and along the vertical and horizontal axis cutting the frame.



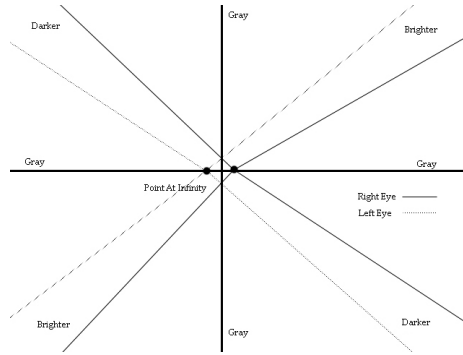


Fig.8. A Stereo frame and its corresponding vertical disparity map along with a diagram showing the distribution of vertical disparity in a frame.

Most of the shader parameters were not meant to be controlled manually. The original design plan for the shader was that only a few of the parameters, like the ones that control the quality of depth of field, were meant to be user controllable. The rest of the parameters were designed to be controlled by a Mel script.

Mel script

ctSimulateEye is the Mel script that updates the parameters of the stereoscopic shader every frame. The design of this Mel script is closely coupled with the motion capture data obtained from the actors' head. The script can operate in two modes namely the '*Eye*' mode and the '*Locked to Head*' mode.

'Eye' Mode:

In '*Eye*' mode the stereo cameras behave similar to the way in which human eyes behave on the head. In this mode the cameras fixate on a particular object in the scene even as the head of the actor moves around. This setup helps to recreate the visual field that the

character is seeing as opposed to the visual world that he or she is experiencing [Lipton, 1982]. This setup also makes it possible to create nice eye shift transition effects.

'Locked To Head' mode:

In this mode the stereo cameras point in the same direction as the head. This mode has a handheld camera look that is very common in the first person shots in movies.

Creative Statement

I wanted to create an intense experience for my audience by taking them on an emotional journey through the lens of the camera. I was inspired by the letters written by soldiers from the front-line from various wars across time. These letters are usually filled with strong emotions and grim details of dangerous situations that they encounter [Franklin Platt, 2006]. Soldiers often experience a heightened state of fear and emotion that normal people usually don't get a chance to experience everyday. I wanted to capture such a moment of fear and uncertainty that a soldier or a commando goes through when he enters an unfamiliar territory. I planned the film such that as the shots unfold the audience would experience the same dilemma that the commando faces. I wanted to take the audience on a journey of fear and uncertainty through the lens of the camera. I also wanted the camera to be another actor in the film. I wanted the audience to experience the heightened emotional state from a first person perspective by looking into a stereoscopic world. I was inspired by the way Alfred Hitchcock handled the camera in his film *The Rope* where the camera becomes another actor by capturing the on screen action without any time lapse, by avoiding cuts. In my film, I wanted to recreate the intimacy that the

camera creates in *The Rope*, as it pans from room to room [1]. I wanted the film to have a non linear narrative structure. I was inspired by the experimentation done by post war filmmakers like Alian Resnais and Federico Fellini in their use of non linear narrative film making. Resnais's film *Hiroshima Mon Amour* deserves a special mention for its bold experiments in using very brief flashback sequences inter cut into scenes to suggest the idea of a brief flash of memory [2]. Fellini's *Satyricon* which is loosely based on extant parts of Petronius's *Satyricon* is another profound example of non-linear cinema. The whole film moves with the logic of a fragmentary dream with incomprehensible shots that literally end in the middle of a sentence.

My goal is to expose the audience to an experience so realistic that they become emotionally invested and follow the action more intensely. I then begin to direct them by dropping clues that make my audience construct their own narrative interpretation of the film. I intend to expose the audience to a logical progression of shots, thereby creating a natural expectation of things to come and then suddenly, I break that expectation by introducing a twist at the end by revealing a child peeping through the trap door. Introducing something as innocent as a child into a volatile situation causes the audience to drastically realign their mental image of what is happening. This ambiguous journey reinforces the intense fears that the audience experiences. As a filmmaker I want to setup a foundation and allow the audience to build their own conclusions. The audience is as much a part of the film as the filmmaker is!

Film Production

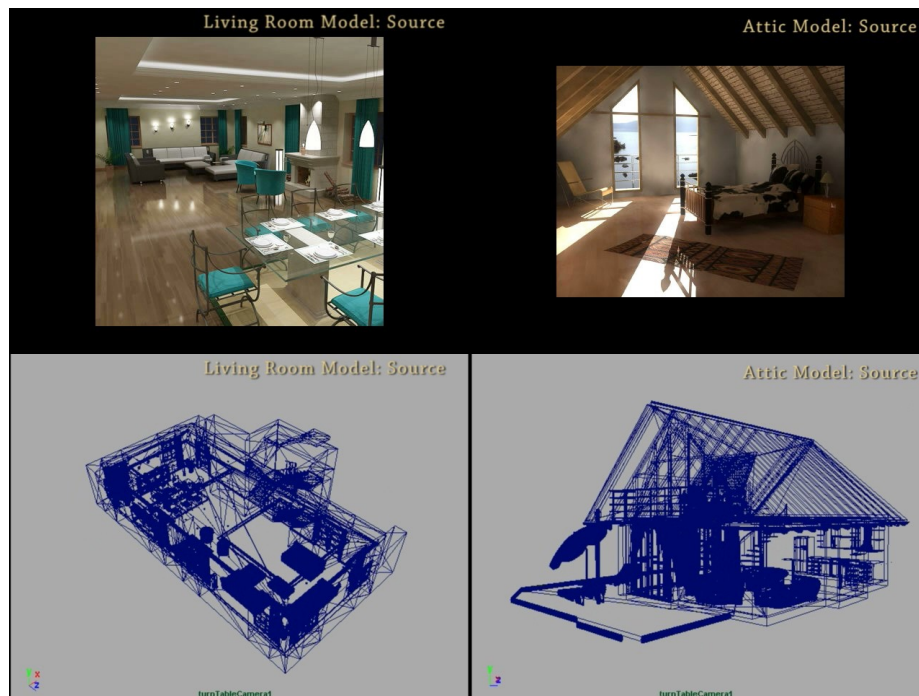
Before I started with any CG production work, I wanted to make sure that the shots I had in mind would work together as a good narrative. So I put together a rough story reel using some live action clips that I shot with my friends. Figure 9 shows a few screenshots from that reel. I was able to identify certain important elements from the reel that I wanted to replicate in the CG shots. For example, I ended up replicating the lens flares from the flashlights in the CG shot. I was also able to pickup important clues about the various aspects of the first person point of view that I wanted to replicate in the CG production. The reel also inspired me to incorporate various other elements like the staircase and the child's slippers into the CG version. The captions in the reel, as seen in the last two screen shots, provided me with critical timing information about the various story beats, that I needed while creating the CG shots.



Fig.9. Screenshots from the story reel

For the film that I had visualized, the music plays a very important role in driving the mood of the film. So, as I was putting together the reel, I purchased some royalty free music tracks and edited them into the reel to get a feel for the final finished version of the film. The shots in the story reel were just placeholders for a different set of shots that I had visualized in my mind. The reel helped to roughly work out the timing of the shots and ensure that the film worked as a cohesive cinematic piece.

As for the CG assets needed for my film, in order to get what I wanted within my limitations of time and resources, I scouted the stock model websites on the Internet searching for “virtual locations” for the film. For example, in order to create the main house where all the action was taking place, I ended up buying two separate models that had features that I liked, and then remodeled and combined them to create the final location in my film. Even though I couldn’t get the exact props and sets that I had in mind, I still had enough choices to get a combination of different things that conveyed the core essence of the shots. Figure 10 shows the original renderings of these models when they were purchased (top two images) and the various stages of the remodeling process after which the final model was created.



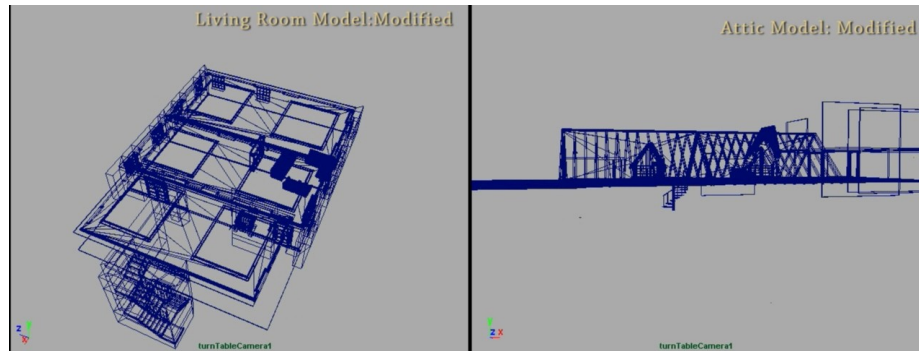


Fig.10. The remodeling process

I also purchased the various other assets like the child and the soldier as rigged models from online.

When I started with the production of the film, my first goal was to capture the ambience of the live action shots, in the computer generated shots. I planned to rely on stereoscopic 3D to create a heightened sense of involvement for the audience in the film. The stereoscopic imagery turned out to be a very crucial tool in enhancing the visual story telling in the film.

A few months before I started working on this production, I implemented the stereoscopic shader in Mentalray. The shader had support for an animatable stereo convergence point. This feature along with the support for depth of field enabled me to create shots that mimicked the visual field of the character. In the film, this technique proved really helpful in guiding the audience's eyes through certain shots. It also helps in a powerful recreation of the first person point of view.

I did some early experiments with trying to recreate the visual field of the character. Some of them worked and some didn't. Figure 11 shows some of the

screenshots from these early tests. The top two images are from one of the very first animation tests in which I got the eye shift transition to work. The bottom image is from a test animation in which I tried to place the stereo cameras into the eye socket of a model, in an attempt to capture the facial features that are sometime a prominent aspect of our visual field. I wanted to see if they would enhance the recreation of the visual field in any way. As it turned out, they didn't! This was one of those tests that didn't go as planned.

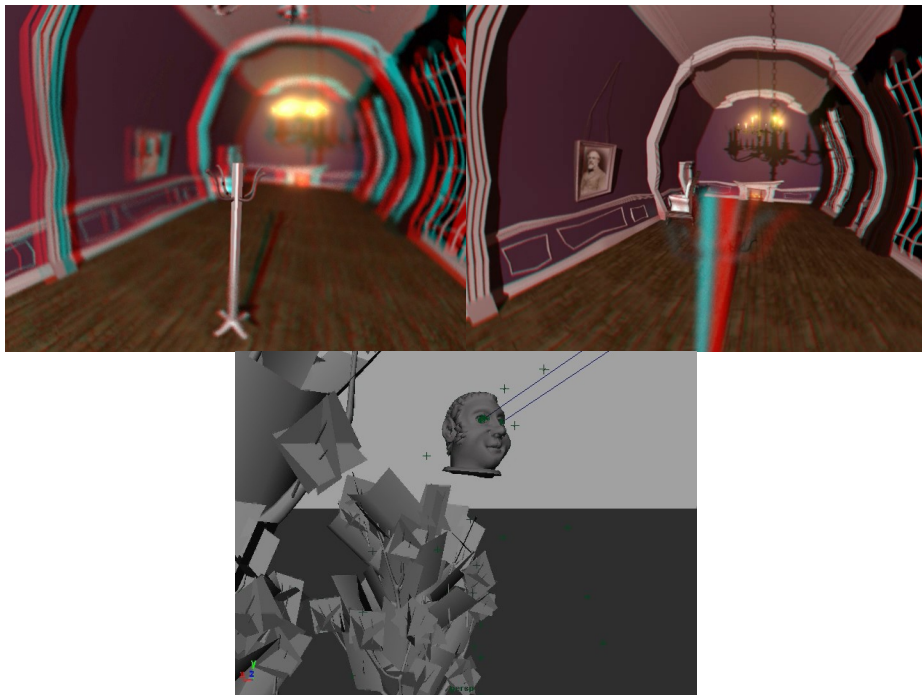


Fig.11. Early tests of stereoscopic pipeline

Working with Stereoscopic 3D

This movie is my first experience with producing a stereoscopic film. Working on the film taught me a lot about the limitations and the strengths of stereoscopic imagery. One aspect worth mentioning is the importance of layering objects in the frame in order

to maximize the stereo depth. There were instances where adding just a few objects into the shot greatly enhanced the perception of stereo in the shot.

For example, in figure 12, the perceived stereoscopic depth in the left image is poor but is greatly enhanced in the right image after additional elements are introduced into the shot. The image on the right contains two additional depth planes in the form of the two shelves. It also contains additional beams that run along the ceiling which helps to enhance the perceived stereoscopic depth in the shot.

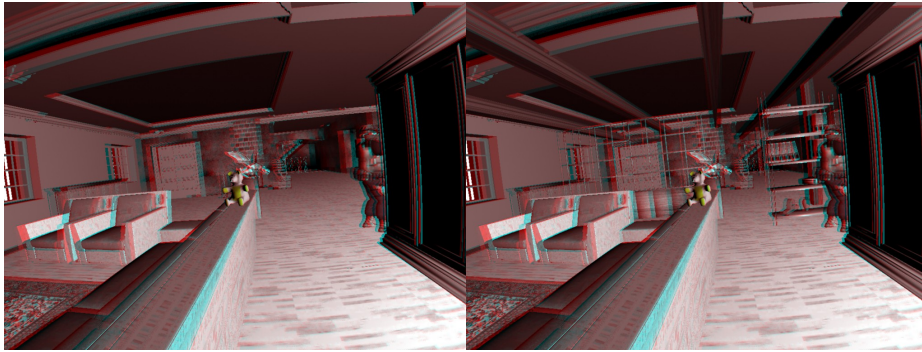


Fig.12. Enhancing stereoscopic depth through layering of objects in the scene.

The props and the dimensions of the set play a very important role in enhancing the stereo in the shots. Trying to fine-tune the stereo for the shots with a lot of camera movement turned out to be quite challenging. Even though fine tuning the stereo was the last stage in the production, when I was constructing the set, I added features like long deep hallways and pillars to help enhance the stereo.

In order to get the best results, for certain shots where the camera was moving through the scene, I animated both the inter axial distance and the distance of the

convergence point to adapt the stereo to the changing imagery on screen. In these shots, because of the camera motion, I tried to keep the pixel disparity as low as possible to make it easy on the eye. Here, I learnt to make an important stereo framing decision, in sections of the shot with some object in the foreground and the main subject further back into the scene. I tried to set the stereo to keep the main subject on the screen and slightly push the foreground object out of the screen. I did this by pushing the convergence point deeper into the scene and increasing the inter axial distance between the left and right eye cameras. The layering of objects in the frame seemed to really help enhance the stereo. But this kind of layering is not always possible, and might even start to feel cinematically awkward in certain shots. If I wanted to enhance the stereo in a shot (or section of a shot) with no prominent layering, I just pushed the entire shot deeper into the screen by moving the convergence point closer towards the camera. This encodes a sense of “volume depth” between the edges of the screen and the subject. In the figure 13, the top image shows a frame with a foreground object. Notice how the staircase grill is made to slightly pop out of the frame. In the bottom frame, as there is no background object, the entire scene is pushed deeper into the stereo volume to give it a sense of depth.



Fig.13.Stereo Frame with foreground object (left) and without foreground object (right).

In sections of the shot that contained only the floor, I tried to angle the floor diagonally across the ‘stereo cubical volume’. Interestingly, this diagonal angle helps to enhance the stereo just like layering of objects does (provided the floor has a reasonable texture on it!). As I was fine tuning the stereo in these shots, I tried to be subtle and deliberate about it, as it is very easy to disorient the viewer in stereoscopic 3D.

The living room shot is one of the shots where I tried to recreate the visual field of the character. The shot starts out with zero intraocular distance (without any stereoscopic depth) and then slowly transitions into stereo around the mid point of the shot by increasing the inter axial distance between the left and right eye cameras. Also at the beginning of the shot, there is no depth of field but as the shot transitions into stereo, depth of field is introduced. My plan was to allow the viewers to explore the shot in the first few seconds when the shot has no stereo depth, then introduce stereo to direct their eyes to the teddy bear. A combination of elements was used in the shot to direct the audiences’ eyes, first the stereo shader was set to work in ‘Eye’ mode. In this stereo mode, the object of focus always lies on the convergence point, this keeps the object centered in

the frame irrespective of the head movement. Visually, all the other objects in the frame appear to move about the center of the frame, this is similar to how we experience our visual field through our eyes. Adding depth of field around the convergence point in turn emphasizes the object of interest. There is a point in the shot when the character shifts gaze from the teddy bear in the foreground to the gargoyle statue in the background. To mimic this, both the convergence point and the depth of field are animated to shift the focus from the teddy bear on to the gargoyle statue. This creates an interesting eye shift like transition. The top left image in fig.14 shows a frame from the living room shot that has the teddy bear in focus, and the rest of the shot is blurred out. The top right image in fig.14 has the gargoyle statue in the background in focus and the teddy bear in the foreground is blurred out. A similar technique is used in the very last shot of the film to shift the viewers' attention between the child and the gun. In fig.14 the bottom left image shows a frame from the shot that has the gun in focus, notice how the background containing the trapdoor is blurred out. The bottom right frame in fig.14 shows a frame with the trap door in focus and the gun in the foreground is out of focus. In both these shots, stereo just plays its role as one of the components in guiding the audience's eyes, as opposed to just creating stereoscopic depth.

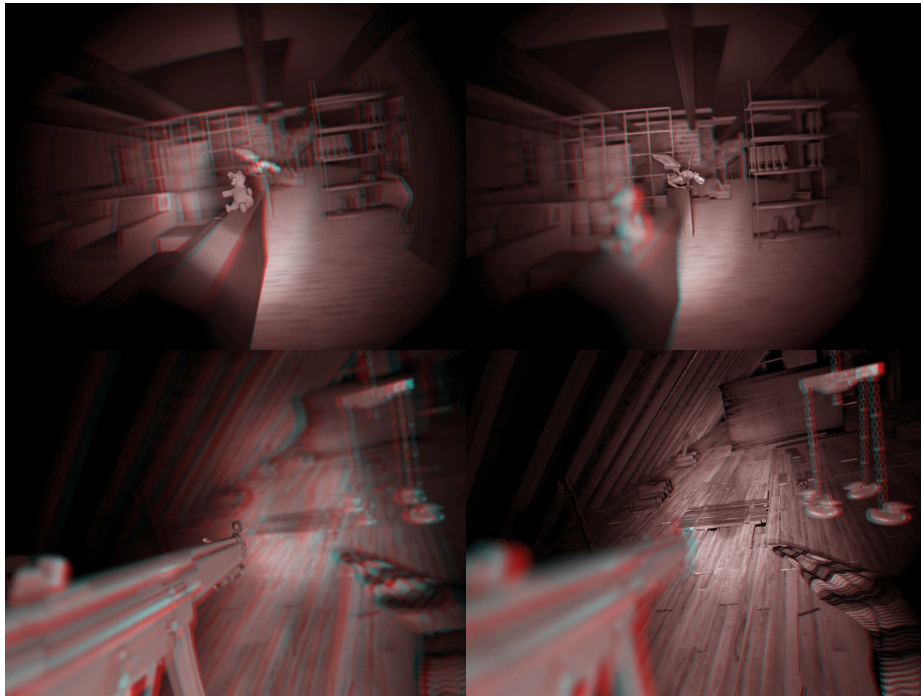


Fig.14.Living Room shot before transition(top left), Living Room shot after transition(top right),Last shot of the film with gun in focus(bottom left), Last shot of the film with gun out of focus(bottom right).

Eye Tracking Study

As the final part of my thesis work, an eye tracking study was conducted to observe the viewing pattern of the audience while watching my short film, and to perform an analysis of the gathered data from a filmmakers' point of view. Another graduate student, Brian Daugherty, implemented a movie player that communicated with the eye tracker. With this software we were able to track the audience's gaze while they were watching the movie [Daugherty, 2009]. My focus was to analyze the collected eye tracking data to study the viewing pattern of the audience and also to study the effectiveness of certain stereoscopic techniques used in the film.

In the study, each participant was first read an informational letter, then they were asked to wear red/cyan anaglyph glasses and were shown a series of stereoscopic images, based on which they were asked to answer some questions. After they finished answering those questions, they were asked to watch a series of video clips for which their gaze data was collected.

For each subject, we altered the order in which the clips were shown to compensate for any learning that might happen during the study. In the study, we also showed both the stereo and the non-stereo versions of my short film to the test subjects to see if we could identify any noticeable difference in the viewing patterns between the two versions. Again, order of presentation was randomized. We also showed a series of short clips to study the viewers' reactions to different stereoscopic framing techniques. A total of 12 subjects took the test. They were aged between 22 to 26. Three of the subjects were female. It is interesting to note that only three of the subjects had seen a stereoscopic 3D film before this study.

Analyzing the eye tracking data for the video clips was a tricky problem. Because of the visual nature of the data, it was hard to find a way to quantify it. One of the ways we decided to visualize the data was to do a heat map of the gaze data overlaid on top of the video clip. This turned out to be a good intuitive way to study the viewing behavior. The heat map was constructed by running a Gaussian blur filter over the gaze points and then the values from these blurred points were accumulated over time, resulting in a slight motion blur. Depending on the amount of gaze over a point, the color of the heat map may range from green to yellow to red, with red representing the highest

gaze intensity, followed by yellow and finally green represents the least gaze intensity.

Figure 15 shows a visualization of the heat map.

In order to do a more quantitative analysis of the data, we calculated the mean and standard deviations of the gaze points for each frame and visualized it as circles centered at the mean with radius equal to the standard deviations about the mean. This gave us a good way to visualize the viewing patterns of the audience at any point in time. The calculated standard deviations were normalized to a value between 0.0 and 1.0, where a value of 1.0 represents the width of the screen. In Figure 15, the two white circles represent the first (inner circle) and the second (outer circle) standard deviation about the mean for the gaze data in the frame.

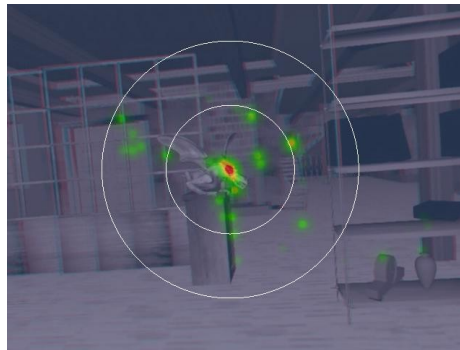


Fig.15. Heat map visualization with standard deviation circles.

CHAPTER FOUR

DATA ANALYSIS

Questionnaires

In the first part of the study, the subjects were shown a series of still images and were asked to answer two questions. For the first question, they were presented the two stereoscopic still images shown in figure 16. The two images shown in figure 16 were rendered with the same inter axial value. The image on the bottom contained additional elements, compared to the image on the top. The bottom image has two additional depth planes in the form of the two shelves and also some additional beams that run along the ceiling.

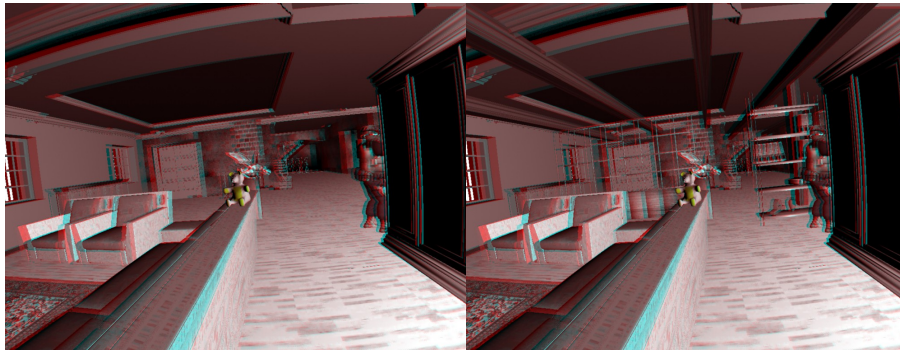


Fig.16. Test images for the first study question. The image on the left was labeled image 1 and the image on the right was labeled as image 2.

The idea was to see how the layering of objects helped to enhance the perceived stereoscopic depth of the shot. After seeing the two images the subjects were asked to answer the following question,

Which of the following options best describes the two images that were shown to you?

- i) Image 1 has more stereo depth than image 2.
- ii) Image 2 has more stereo depth than image 1.
- iii) Both the images seem to have the same stereo depth.

25% of the subjects thought that image 1 had more stereo depth than image 2. Another 25% thought that they both had the same stereoscopic depth and the rest 50% thought that image 2 had more stereoscopic depth.

Next, the participants were presented with two more stereoscopic still images as shown in figure 17. Then they were asked to answer a question designed to study the effect of stereoscopic 3D on the perceived scale of the space shown in the image.

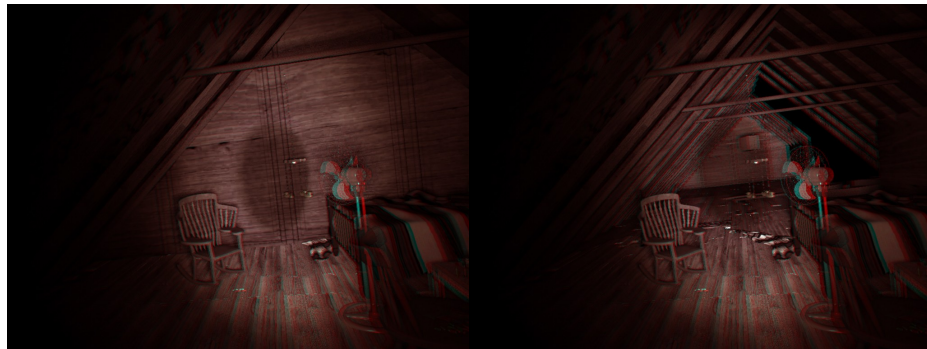


Fig.17. Test Images for the second study question. Image on the top was labeled image 1 and the image on the bottom was labeled image 2.

They were asked to answer the following question,

Compare the scale of the room shown in image 1 with the scale of the room shown in image 2. Which of the two images look close to a miniature model of the room?

- i) Image 1 looks like a miniature.
- ii) Image 2 looks like a miniature.
- iii) Both Image 1 & Image 2 look like miniatures.
- iv) Neither of them looks like miniature.

Interestingly enough each option was selected by 25% of the subjects.

Analysis of Videos

Before presenting an analysis of the eye tracking data, I first describe the accuracy of the eye tracker and its data. The eye tracker has an accuracy of about one degree of visual angle. Although the eye tracker is fairly accurate in picking up the eyes through the stereo glasses, if the subject moves his or her head out of the eye tracker's range then it starts generating invalid data. In a viewing session lasting about 20 minutes (as in our case) the subjects invariably tend to slouch or lean back from time to time, causing the eye tracker to miss their eyes. But, we were able to obtain usable data for the most part. The raw gaze data from the eye tracker was quite jittery. So, in order to get a smooth reading, the gaze points were plotted by taking the weighted average of the previous and the next five gaze points. These smoothed points were then passed through a Gaussian filter and motion blurred to construct the heat map used for the visualization. These steps

meant that the eye tracking data was not suitable for studying the minute movements of the viewer's gaze, but it was able to effectively capture the larger eye movements of the viewers within the frame.

Effectiveness of Eye Shift transition

A series of four stereo video clips designed to evaluate the effectiveness of the *eye shift* transition possible in the 'Eye' mode of the shader, were shown to the audience. The *eye shift* transition works by animating both the depth of field and the stereo convergence point at the same time. As shown in the leftmost image in figure 18, the test clip starts out with both the stereo convergence point and the depth of field fixated on a teddy bear in the foreground. As the camera moves closer to the teddy bear, both the convergence point and the depth of field are animated to bring the gargoyle statue into focus, as shown in the middle image in figure 18. In order to make this transition effective, this shot was specifically designed to keep both the teddy bear and the gargoyle statue as close to each other along the line of sight as possible. This creates an interesting transition that mimics the eye shift that happens in our visual field, hence the name *eye shift* transition. Along with this clip, there were two other clips that were designed to create conflicting visual cues by animating either one of the depth of field or the stereo convergence point while keeping the other constant. Finally there is the last clip in which only the convergence point is animated with no depth of field. The goal was to study the viewer's gaze in each of these cases and compare the effectiveness of the *eye shift* transition to the other scenarios.

The order of these clips was changed for each subject and they were mixed in with other videos to compensate for any learning that might happen during the study [Daugherty, 2009].

Eye shift Transition: Animation clip with both the convergence point and depth of field animated.

In this clip, the gaze points converge on the teddy bear in the initial section, before the transition, as shown in the topmost image of figure 18. Immediately after the transition, the gaze points spread out to scan the new background objects that come into focus, as shown in the middle image in figure 18. After a quick scan, a big portion of the gaze points come back to fixate on the gargoyle statue as shown by the bottommost image in Figure 18.

During the transition, it is interesting to observe how the gaze points get effortlessly transitioned in the image space, from the teddy bear on to the gargoyle statue as the convergence point gets animated.

The overall standard deviation for the gaze data in the clip is 0.156 screen width. In the initial section before the transition, the standard deviation is 0.109 screen width and in the section after transition, it increased to 0.175 screen width. The larger standard deviation value in the later half of the clip can be attributed to two factors:

a) A bigger portion of the scene comes into focus after the transition, freeing the eye to scan a larger area on the screen.

b) During the transition, there is a sudden shift of the scene elements in image space. This sudden movement causes the viewers gaze to disperse around the scene for a few moments. This dispersion adds to the standard deviation value.

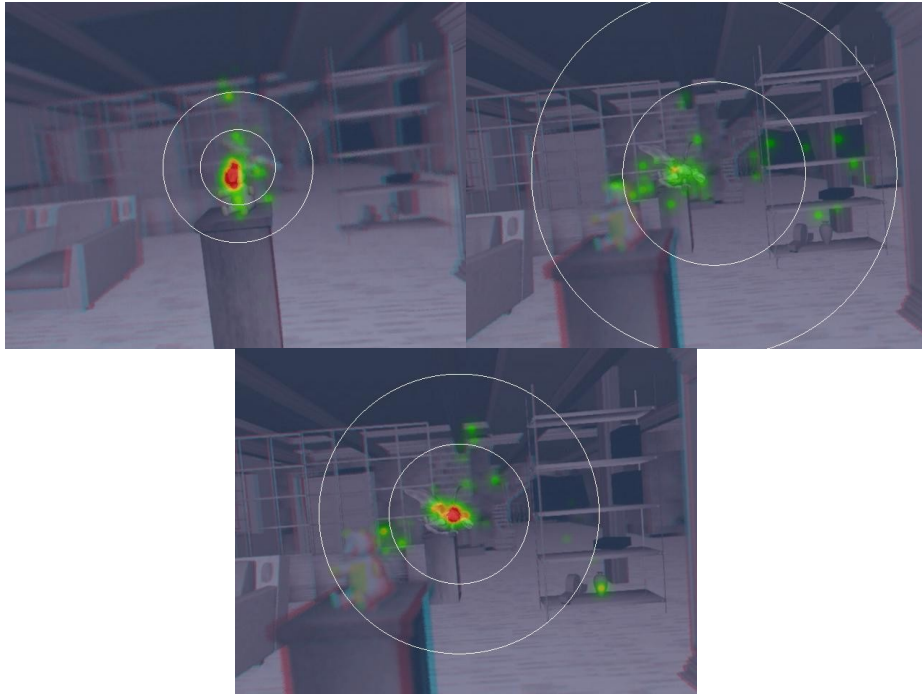


Fig.18. Screenshots from the animation clip in which both the convergence point and the depth of field are animated together.

Animation clip with fixed convergence point and animated depth of field

As shown in the top image in figure 19, in the initial portion of this clip the gaze is centered on the teddy bear and when the transition happens, only the depth of field gets animated to change the focus from the teddy bear on to the gargoyle statue, the convergence point remains on the teddy bear even after the transition. This creates a conflicting visual cue, as the blurred teddy bear remains at the center of the image after

the transition. As soon as this happens, the gaze points on the teddy bear immediately disperse into other sections of the image that come into focus, as shown in the right image in figure 19. This clip has an overall standard deviation of 0.152 screen width. The section before the transition has a standard deviation of 0.097 screen width and the section after the transition has a standard deviation of 0.21 screen width.



Fig.19.Screenshots from the animation clip having a fixed convergence point and animated depth of field.

Animation clip with fixed depth of field and animated convergence point.

As shown in top image in figure 20, the gaze is converged on the teddy bear in the initial section of the clip. During the transition, only the stereo convergence point gets animated to the gargoyle statue, while the focus remains fixed on the teddy bear. This transition also creates conflicting visual cues. There is an obvious confusion in the way the gaze data reacts to these conflicting visual cues, as shown in the bottom image of figure 20. In fact, this is one of the really bad scenarios that we tested. In the final sections of the clip, the whole frame goes out of focus. Yet this clip shows the importance

for the visual cues to work together in creating an effective eye shift transition. The overall standard deviation of this clip is 0.155 screen width. In the initial section before the transition it is 0.100 screen width and it is 0.181 screen width after transition.



Fig.20.Screenshots from the animation clip with fixed depth of field and animated convergence point.

Animation clip with moving convergence point and no depth of field

As shown in the topmost image in figure 21, in the initial portion of this clip, the gaze points converge on the teddy bear as it is the object with zero disparity and also the center of focus in the clip as it lies on the stereo convergence point. Here again after the transition, there is a dispersion of gaze points into the scene, as shown in the middle image in figure 21. But, there is a relatively high concentration of gaze points on the gargoyle statue following the initial dispersion, as shown by the bottommost image in figure 21. The gaze data for this clip is characterized by a number of quick glances around the scene, this can be attributed to the absence of depth of field. This also contributes to the high standard deviation value of 0.178 times the screen width for the entire clip, with the section before the transition having a standard deviation of 0.096 screen width and the section after the transition having a value of 0.21 screen width.

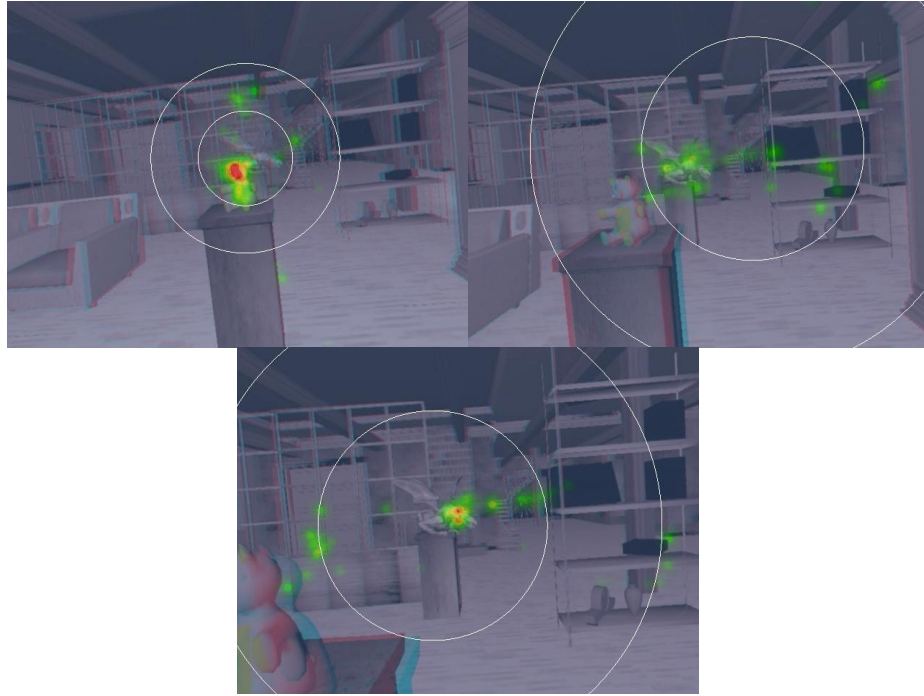


Fig.21. Screenshots from the animation clip with animated convergence point having no depth of field

<i>Animation Clip</i>	<i>Standard Deviation</i>	
	<i>Before Transitio n</i>	<i>After Transitio n</i>
Both Depth of Field and Convergence Point Animated	0.109	0.175
Fixed Convergence and Animated Depth of Field	0.097	0.210
Fixed Depth of Field and Animated Convergence Point	0.100	0.181
Animated Convergence Point with no Depth of Field	0.096	0.210

Table.1. Standard deviation values of the four different test cases.

Among the standard deviation values of all the clips in the section following the eye shift transition, the clip with both convergence point and depth of field animated has

the least standard deviation value. This shows that animating both the depth of field and the convergence point does the most effective job of holding the audience's gaze to the region around the convergence point, following the transition.

In the clip with no depth of field, the gaze points are mostly concentrated around the convergence point, which also happens to be the point of zero disparity in the frame. The viewers seem to limit their exploration of the rest of scene to quick glances around the scene while resting their gaze on the object near the convergence, as it has the minimum onscreen motion and disparity.

Analysis of the stereoscopic film.

In the study each subject was also shown both the stereoscopic and the non-stereoscopic versions of my film. The order of the movies was swapped for each subject, such that half of them saw the stereoscopic version first and the other half saw the non-stereoscopic version first. This is to compensate for any learning that might happen while viewing the different versions of the film. My goal with this study was to,

- a) Analyze the viewer's gaze data to evaluate the effectiveness of different stereoscopic techniques that I used in the film.
- b) Being a very visual film, I wanted to see how well the viewers spot the critical elements of the film.
- c) Wanted to compare and identify the differences, if any between, in the viewing pattern of the audiences between the stereoscopic and non-stereoscopic versions of the film.

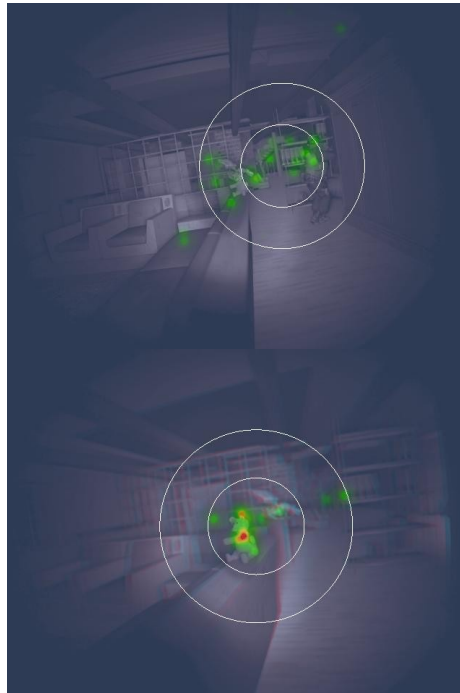
d) I also wanted to identify if there was any noticeable feature in the viewing pattern of the audience during the second viewing of the movie when compared with the first viewing.

Living Room shot

In this shot, I wanted to guide the audience's gaze through the shot by simulating the visual field of the character. In order to do that, I designed the shot so that it starts out with zero intraocular distance (without any stereoscopic depth) and then slowly transitions into stereo around the mid point of the shot, by increasing the inter axial distance between the left and right eye cameras. Also at the beginning of the shot, there is no depth of field but as the shot transitions into stereo, depth of field is introduced. My plan was to allow the viewers to explore the shot in the first few seconds when the shot has no stereo depth, then introduce stereo to direct their eyes to the 'teddy bear'. A combination of elements was used in the shot to direct the audiences' eyes, first the stereo shader was set to work in 'Eye' mode. In this stereo mode the object of focus always lies on the convergence point, visually this keeps the object centered in the frame irrespective of the head movement. All the other objects in the frame appear to move about the center of the frame, this is similar to how we experience our visual field through our eyes. Adding depth of field around the convergence point in turn emphasizes the object of interest. There is a point in the shot when the character shifts gaze from the teddy bear in the foreground to the gargoyle statue in the background. To mimic this, both the convergence point and the depth of field are animated to shift the focus from the teddy

bear on to the gargoyle statue. Visually, this creates an interesting eye shift like transition.

On analyzing the gaze data of the viewers who saw the stereoscopic version of the film first, I was able to identify the following pattern from the gaze data. As shown in topmost image in figure 22, there was a tendency for the audience to explore the scene more in the initial sections of the shot, when it starts out with zero disparity. As stereo and depth of field are introduced into the shot, their tendency to scan the scene gets reduced, as shown in the middle image in figure 22. This shot has a constant gaze concentration around the center of the frame. After the focus shifts from the teddy bear on to the gargoyle statue, there seems to be quick glances around the scene, as more of the scene comes into focus. This is noticeable in the bottommost image in figure 22.



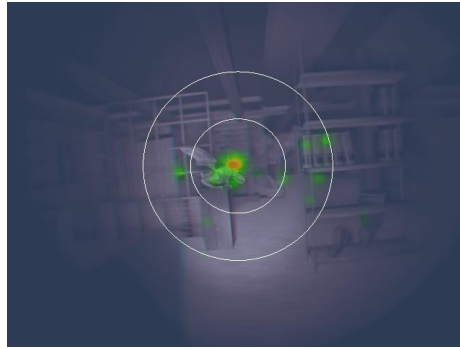


Fig.22. Screenshots from the Living Room shot in the stereoscopic film

Trapdoor to gun focus shift

I also wanted to study the viewer's gaze behavior in the very last shot of the movie when the focus shifts between the child and the gun. At the beginning of the shot, when the gun first enters the frame, the focus is on the trapdoor under which the child is hiding. Then the focus is shifted from the trapdoor to the tip of the gun by animating both the convergence point and the depth of field. After a few seconds, the focus shifts back to the trapdoor and the child is revealed. This is another shot, where I tried to mimic the visual field of the soldier. I wanted to study the viewer's reaction to this shot by analyzing their gaze data.

As soon as the gun enters the frame, even though it remains out of focus, there are some noticeable quick glances towards the gun. This is shown in the topmost image in figure 23. As soon as the focus shifts from the trapdoor on to the tip of the gun, there is a sudden dispersal of gaze points within the scene, as seen from the middle image in figure 23. At this point, there are noticeable quick glances between the gun and the trapdoor. After the focus shifts back to the trap door, the child slowly emerges from under the trap

door and as soon as the child appears, there is a very focused, highly intense gaze on the child, as shown in the bottommost image in figure 23.

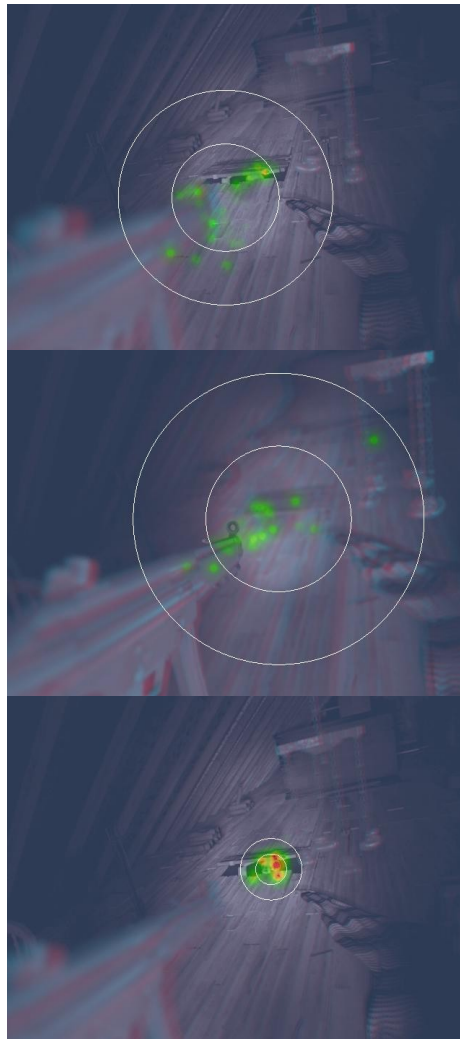


Fig.23. Screenshots showing the focus shift between trapdoor and gun

Comparison of the viewer's gaze pattern while watching the stereoscopic and the non stereoscopic versions of the film.

I compared the gaze data of the viewers, who saw the stereoscopic version of the film first, with the gaze data of the viewers who saw the non-stereoscopic version of the film first. Both their gaze patterns very similar for the most part. One noticeable difference was right at the end of the stairs sequence, when a long tunnel gets revealed. As shown in top image of figure 24, in the stereoscopic version of the film, there is a highly focused gaze intensity looking right at the other end of the tunnel. But in the non stereoscopic version of the film, the gaze data is more spread out across the frame and is not as highly focused as in the case of the stereoscopic version. The bottom image in figure 24 shows the gaze field for the non stereoscopic version of the film.



Fig.24. Heat map of the tunnel sequence from stereoscopic (top) and non-stereoscopic (bottom) versions of the film.

Analysis of the viewer's gaze pattern on watching the stereoscopic version of the film after watching the non-stereoscopic version and vice versa.

In both the versions of the film, one very noticeable characteristic of the second viewing can be seen in the shot at the beginning of the movie when the soldier walks up to the trap door and gets shot. On the second viewing, the audience's focus on the soldier increases considerably, especially around the time he gets shot, as they are aware of what is going to happen to him.

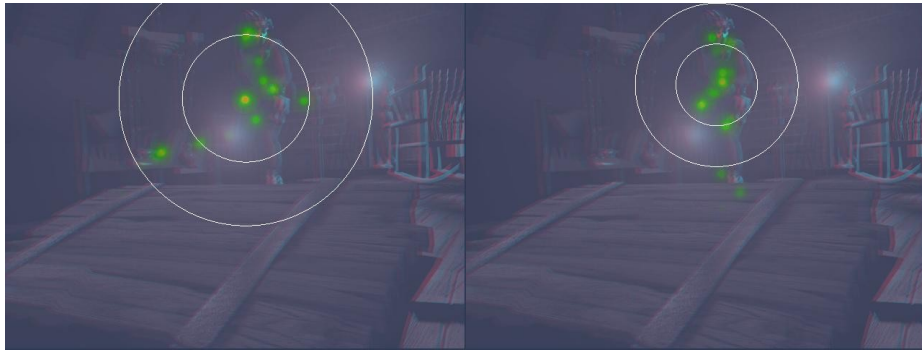


Fig.25. Living room shot from the first and the second viewing of the stereo film.

The visual nature of this film meant that it was absolutely crucial that the audience spot all the elements in the shot that I wanted them to see. The eye tracking analysis of the film enabled me to observe if I had accomplished this effectively. The analysis of gaze data helped to reassure me that the shots in the film were presented in an effective manner such that the audiences were able to pickup all the crucial elements of the film. The highly focused gaze on the child when he gets revealed at the very end shows the audience's curiosity in finding out who is hiding behind the trap door. This stands as a testament to the effectiveness of shots in helping to build up the audience's involvement in the movie. The analysis also enabled me to study the effectiveness of different techniques that I used to direct the audience's attention in the shots. Although, a separate section of the study was devoted to testing the effectiveness of the *eye shift* transition, in the movie itself, all the cues work together effectively in directing the viewers attention to the teddy bear and then on to the gargoyle statue. These two objects are the crucial elements in the shot that symbolizes the sense of ambiguity and confusion that the soldier experiences.

In the very last shot of the film, I tried to recreate the visual field of the slider by shifting the focus between the gun and the trapdoor. When the focus shifts from the trapdoor to the gun, I expected the audience's gaze to completely follow the shift to the gun. But the analysis of the eye tracking data showed that as soon as the focus shifts from the trapdoor on to the gun, the audience's gaze does not completely shift to the gun, but instead appears to go back and forth between the gun and the trap door. But since I have the trapdoor completely blurred out, this might create discomfort to the audience. Based on the analysis of the eye tracking data, I can try to create other versions of the shot that might be more conducive to the audience's viewing behavior-like having the blur on the trap door reduced to a minimum when it goes out of focus or having the blur completely removed and observe how the audiences reacts to these variations.

Stimulus response in 'Eye' mode of the shader

This clip was designed to study the effect of a stimulus, like a high intensity flicker, on the viewer's gaze in the 'Eye' mode of the shader. The clip starts out with both the convergence point and the depth of field focused on the coat hanger in the foreground. A few seconds into the clip, the focus shifts from the coat hanger to the objects in the background. This is accomplished by moving the convergence point from the coat hanger to the chair in the background.

The analysis of the data from the eye tracking study showed that in the initial portion of the clip, before the eye shift transition, there were two hot spots in the heat map, one on the coat hanger that lies near the convergence point and the other on the high

intensity flicker originating from the out of focus background. This is shown in the topmost image in figure 21. As soon as the transition happens, the gaze points disperse into the scene, scanning the new background objects that come into focus. This is shown in the middle image in figure 26. But as soon as the flickering light starts to move, all the gaze points converge on the moving light to create an intense spot on the heat map, as shown in the bottommost image in figure 26.

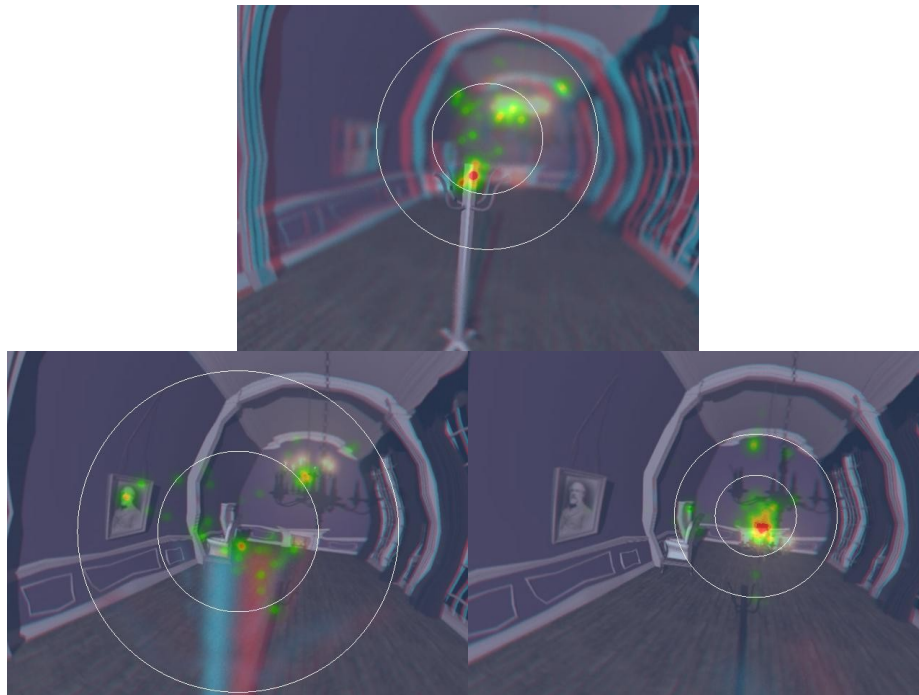


Fig.26. 'Eye' mode clip with flicker

CHAPTER FIVE

CONCLUSION

During this project, I was able to develop my short film from conception to final execution. In the process, due to the limitations on time and resources, I got a chance to try some unconventional production techniques for assembling the CG assets needed for my production. I also developed a stereoscopic production pipeline that would enable me to create realistic first person stereoscopic shots. In the film, I experimented with some new stereoscopic framing techniques for guiding the audiences' eyes through the shot. Finally, I used eye tracking to evaluate the effectiveness of the new stereo framing techniques that I tried in the film. In turn, I was able to assess the feasibility of eye tracking as a tool for film makers.

It is fairly easy to predict where the audiences are looking in a conventional dialogue driven shot. However, for visually complex shots it is not so obvious. In this case, eye tracking can be used as a tool to observe the gaze pattern of the audience to identify the regions of interest in the frame. This information can be used to improve the shot itself. Eye tracking can also be used to gauge how the viewers react to some new experimental framing technique, especially in a medium like stereoscopic 3D films where the viewers can be quite sensitive to any kind of a change. On analyzing the eye tracking data for this film, I could not find any particular element that was distracting the audience. But nonetheless, eye tracking data can be used to identify elements in the shot that might distract the audience from the flow of the movie. If a consistent viewing pattern was identified from the gaze data, it can be used to make effective stereo

transitions during cuts. Studying the viewer's gaze behavior, particularly for stereoscopic films, can be used to enhance the viewing experience of the audience.

A possible future extension of this work might be a production in which the shots are iteratively refined based on studying the audience's reaction to the shots using eye tracking. Eye tracking might be extremely valuable especially when experimenting with new cinematographic techniques in the film. With the recent increase in the number of stereoscopic productions, eye tracking provides an interesting means for optimizing the viewer experience in these films.

WORKS CITED

- KLINE, SALLY. *George Lucas: Interviews*. Mississippi: University Press of Mississippi, 1999.
- PAIK, KAREN. *To Infinity and Beyond!: The Story of Pixar Animation Studios*. Chronicle Books, 2007.
- LIPTON, LENNY. "Wrong, Wrong, Wrong: Myths of Stereoscopic Filmmaking." Lenny Lipton's Blog. 2008. Lenny Lipton. 30 March 2009 <<http://lennylipton.wordpress.com/2008/06/11/reality-check/>>.
- "House Of Wax Trailer." 1 April 2009
<http://www.sabucattrailers.com/expo2/House_Of_Wax_h.wmv>.
- DIAL M FOR MURDER. Dir. Alfred Hitchcock. Perf. Ray Milland and Grace Kelly. Film. Warner Bros. Pictures, 1954.
- LIPTON, LENNY. *Foundations of the Stereoscopic Cinema*. Van Nostrand Reinhol, 1982.
- "ANAGLYPHS METHOD COMPARISON." Know How. 2005. Peter Wimmer. 31 March 2009
<http://www.3dtv.at/Knowhow/AnaglyphComparison_en.aspx>.
- "HITCHCOCK ON CONTENT VS TECHNIQUE." Hitchcock Interview. 2008. goodfellaOS. 31 March 2009 <http://www.youtube.com/watch?v=uG43hjICE2U&feature=channel_page>
- HUBEL, DAVID H. *Eye, Brain, and Vision*. 1995. Harvard University. 04 Apr. 2009
<<http://hubel.med.harvard.edu/b36.htm#ste>>.
- WHEATSTONE, CHARLES. "Contributions to the physiology of vision. Part the first. On some remarkable and hitherto unobserved phenomena of binocular vision." *Philosophical Transactions of the Royal Society* 128 (1838): 371-94.
- JURASSIC PARK. Dir. Steven Speilberg. Perf. Sam Neill and Laura Dern. Film. Universal Pictures, 1993.

TOY STORY. Dir. John Lasseter. Perf. Tom Hanks and Tim Allen. Film. Walt Disney Pictures, 1995.

FINAL FANTASY. Dir. Hironobu Sakaguchi. Perf. Alec Baldwin and Ming-Na. Film. Chris Lee Productions, 2001.

THE POLAR EXPRESS. Dir. Robert Zemeckis. Perf. Tom Hanks and Leslie Zemeckis. Film. Castle Rock Entertainment, 2004.

BEOWULF. Dir. Robert Zemeckis. Perf. Ray Winstone and Anthony Hopkins. Film. ImageMovers, 2007.

SUTO, MASUJI. Stereo Movie Maker. Computer software. Vers. 0.93. Muttan's Homepage. 01 May 2008. 07 Apr. 2009 <<http://stereo.jpn.org/eng/index.html>>.

DAUGHERTY, BRIAN. "Ocular vergence response over anaglyphic stereoscopic videos." Thesis. Clemson University, 2009.

FRANKLIN PLATT, STUART. Letters from the Front Lines. Granville Island, 2006.