Animation: 2D Versus 3D and Their Combined Effect

by
Kristin C. Au

Submitted to the
Department of Architecture
in Partial Fulfillment of the Requirements for the Degree of

Bachelor of Science in Architecture

at the

Massachusetts Institute of Technology

June 2014

Signature redacted

Signature of Author ______________________
Department of Architecture
May 23, 2014

Signature redacted

Certified by ______________________
Takehiko Nagakura
Associate Professor of Design and Computation
Thesis Supervisor

Signature redacted

Accepted by ______________________
Meejin Yoon
Associate Professor of Architecture
Director of the Undergraduate Architecture Program

Animation: 2D Versus 3D and Their Combined Effect

by Kristin Au

Submitted to the Department of Architecture on
on March 23, 2014 in Partial fulfillment of the
requirements for the Degree of Bachelor of Science in
Architecture

ABSTRACT

This thesis studies the differences in the perception of space and character movement between 2D and 3D animation. 2D animation is defined by elements constructed in a 2D environment while 3D animation by elements constructed in a 3D environment. Modern day animated films have been seen to mix the two forms for the sake of artistic effect, expedited production, and general convenience. Though some modern animations combine the two in the explorative quest to discover new animation forms, few films directly compare the forms to visualize the differences in their perceived qualities. Noticeably, the two animation methods differ in level of detail, dimension, realism, and artistic expression. In terms of lighting, the science of illumination dictates the 3D environment whereas in the 2D environment, lighting is an illusion created by coloring conventions. This study looks specifically at lighting as the controlling factor delineating the two forms.

Two short mixed media films were created. One film had a 3D base while the other a 2D base. A varied set of subjects were shown one of the two short films produced and asked to complete a survey. The survey measured the subject's understanding of space and character movement as seen the film. Results show that in 3D there is an enhanced understanding of spatial perception while in 2D there is a lower sensitivity to character movement.

Thesis Supervisor: Takehiko Nagakura
Title: Associate Professor of Design and Computation

ACKNOWLEDGEMENTS

A huge thank you to my advisor Professor Takehiko Nagakura who guided me through the process and constantly gave me advice. Additionally, thank you to my friends and family who supported me as I worked to complete my research and my four years at MIT.

TABLE OF CONTENTS

Abstract **3**
Acknowledgements **5**

Chapter 1 – Introduction **9**
1.1 Project Contribution 9
1.2 Thesis Organization 10

Chapter 2 – Understanding Animation **11**
2.1 The History of Animation 11
2.2 Toy Story 14
2.3 The Japanese Animation Industry 15
2.4 The Walt Disney and Pixar Contribution 17

Chapter 3 – Looking at Mixed Media **21**
3.1 Defining 2D and 3D 21
3.2 Uncanny Valley 23
3.3 Mixed Media: *Paperman* by Walt Disney Animation Studios 26

Chapter 4 – Production Pipelines **31**
4.1 Building the 3D 31
4.2 Building the 2D 35
4.3 Mixing the Media 37

Chapter 5 – Perception of Space and Character Movement **39**

Chapter 6 – Conclusion **47**
6.1 Sources of Error 47
6.2 Future Research 48

The Survey **49**
Bibliography **52**

CHAPTER 1: INTRODUCTION

Over the past 150 years, animation has developed from simple black and white line drawings to high quality three dimensional computer generated images. With the development of computer graphics, animations have become more realistic, thus making it easier for audiences to relate to storylines and empathize with characters. In the past, cartoons did not require a realistic element allowed greater flexibility to creative and artistic portrayals of objects, animals, and people. Arguably now, with higher quality computer generated images (CGI), animations could now reach new heights in the portrayal of realism using 3D techniques.

This thesis proposes to define the two types media, 2D and 3D, as well as propose a method of discovering the strengths and weaknesses of the two. There are four categories in which the two types of animation can be differentiated: scene movement, character movement, spatial understanding, and lighting. The study specifically analyzes audience perception of space and character movement between 2D and 3D animation. Lighting is used as a control factor in the creation of a mixed media so that its effects may not hinder the study. Additionally, scene movement and spatial understanding are combined into a single, generalized category.

1.1 Project Contribution

A distinction has been made between this flat 2D cartoon animation and 3D animation techniques. Modern day, perhaps in nostalgic memory to the past, or hope to discover new forms of media has shown a movement towards a form of mixed media. In this mixed media form, both 2D and 3D are used to create new effects, expedite production processes, and progress storylines, to name a few motives. The creation of mixed media, however, confuses and muddles the distinction between what seems to be the solidly defined medias of 2D and 3D. Additionally, the distinction in services and abilities provided between the two media becomes increasingly more important to understand.

1.2 Thesis Organization

There are six chapters in this thesis. Chapter 1 presents the problem and idea as well as provides the background to this project. Chapter 2 takes an in depth look at animation and where it stands today. Chapter 3 analyzes specifically mixed media and looks at Paperman produced by Walt Disney Animation Studios as a case study. Chapter 4 describes the methodology by which the 2D and 3D animation material was created. Chapter 5 explains the survey and reveals the results to audience perception of space and character movement. Chapter 6 provides final insight on the situation as well as ways to improve and grow this thesis.

CHAPTER 2: UNDERSTANDING ANIMATION

In order to delve deeper into discerning the differences between 2D and 3D animation, an understanding of its history and key moments must be reached.

2.1 The History of Animation

The beginnings of film animation dates as far back in time as 1877 when French scientist Charles- Emile Reynaud invented the praxinoscope, a device that would project strips of pictures and images and project them onto a screen. By the 1890's, the animation process had adopted a rudimentary form of frame animation. The early 1900's held many firsts for the world of animation and was the time period of many beloved characters including Felix the Cat, and Mickey Mouse came to existence (Dirks).

Even without the technological developments of the 21st century, the first animated films and segments were already varied in production techniques. Common methods included puppet animation, silhouette animation, cut out animation, early clay animation, and hand drawn frame animation (Lemay). In 1914, Winsor McCay produced one of the first film animations: *Gertie the Dinosaur*. A black and white cartoon, it was a laboriously drawn series of 10,000 line drawings (Dirks). In 1937, Walt Disney produced the first full-length cel animated feature animation that had both color and sound: *Snow White and the Seven Dwarves*. The incredible success of *Snow White* proved it possible for animations to hold top ranking positions in the film industry (Lemay). Other notable first animations included Lotte Reiniger's *The Adventures of Prince Achmed*, avant-garde for its use of single colored backgrounds and completely silhouetted scenes, Disney's *Fantasia*, and *Sesame Street* ("Chronology- History of Animation- Films from 1926 to 1946").

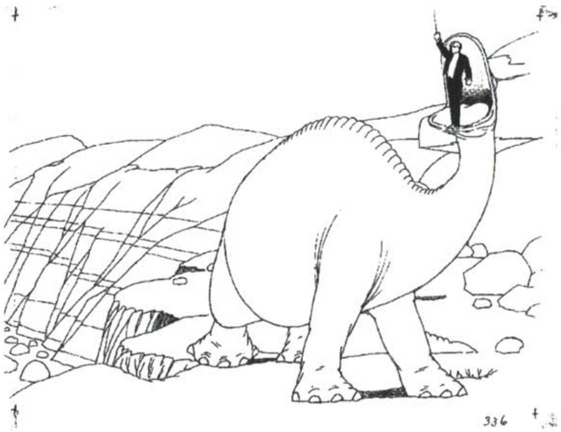


Image 1: Still frame from Gertie the Dinosaur
Source: scanned from Windsor McCay: His Life and Art, John Canemaker, 2005

Image 2: Cel of Snow White and the Seven Dwarves
Source: © Disney Enterprises Inc.

Until the mid 20th century, animations were hand produced frame-by-frame at 24 frames per second of animation. The most common technique was a layering technique called cel animation. Transparent hand drawings were overlaid on top of fixed backgrounds and photographed (Dirks). By the late 20th century, however, animators began to take advantage of newly developing computer technology. *Westworld*, produced in 1977, was the first feature film to incorporate computer graphics. Representing a robot's vision, early computer graphic engineers edited the light intensity of pixels in a photograph (Anders).

Closely following, present day Chief Creative Officer of Pixar John Lasseter made an appearance as a member of Lucasfilm's computer unit. *The Adventures of Andre and Wally B.* was produced in 1984. An fully computer generated animated short, *The Adventures of Andre and Wally B.* was the first of its kind demonstrating motion blur as well as squash and stretch techniques. Squash and stretch techniques are now known to be one of the most important techniques in animation. It denotes for stiff objects, while the object maintains the same volume, it will deform upon the encounter of external forces (Lemay).

Although not as successful in its release, *The Black* Cauldron, a 1985 Disney production, also contributed to both flat and computer based animation technology. Its production also marked the first use of the APT process. Even though APT process allowed quick transfer of drawings to cels, transferred line art would quickly fade from cels with time.

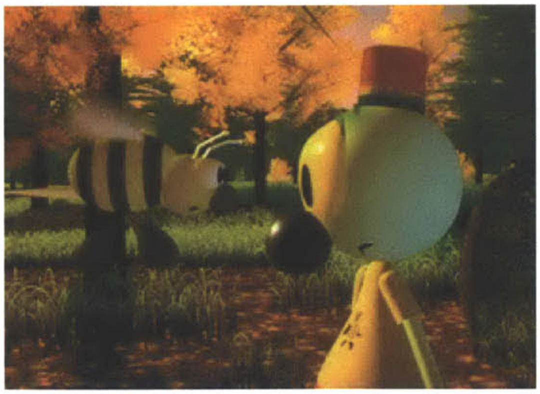


Image 3: Still frame from *The Adventures of André and Wally B.*
Source: *The Adventures of André and Wally B.*, 1984

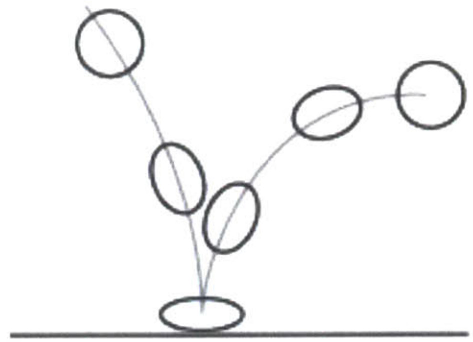


Image 4: Demonstration of the squash and stretch technique
Source: *Squash and Stretch Technique*

Next in 1986 came *Luxo Jr.*, another production by Lasseter, produced for the annual SIGGRAPH computer technology exhibit. This film was fundamental to the beginnings of Pixar. A first of its kind in many aspects, *Luxor Jr.* gave human qualities to an inanimate object, a lamp, as well as demonstrated the ability of 3D software RenderMan in surface shading, specifically self-shadowing in which an object causes itself to have shadows on itself.

In the time period just before the turn of the millennium, animations, both hand drawn and computer animated, were accepted as commercial forms of film. The early 2000's marked a period of exploration in techniques. In 1997, *Marvin the Martian in 3D*, produced by Warner Bros., was the first computer animated movie that required 3D glasses for viewing. In 2001, Square Pictures rendered *Final Fantasy: The Spirits Within*. Inspired by a video game, this feature film was the first to be made based on photorealism and motion capture for character actions. Further building on live performance motion capture techniques, the mapping of live human movements to 3D characters in computer imagery, was *The Polar Express* in 2004, also made by Warner Bros. Pictures. In 2005, *The Corpse Bride* was produced using stop motion, a technique that involved taking photos while physically moving a live object in increments (Dirks).

2.2 Toy Story

Image 5: Still scene from *Toy Story*
Source: *Toy Story*, 1995

Toy Story was the first ever computer animated feature length film that told a story using convincing computer generated characters. Produced in 1995 by John Lasseter, it was a 77 minute film that required 800,000 machine-hours just for the production of the final cut (*The Making of Toy Story*). *Toy Story* defined a brand new pipeline for animation production. Each scene required work done by eight different departments. First the art department drew the general color scheme and lighting and passed it off to the layout department who pre-visualized camera movements. The animation department then finalized and key framed each object in a scene. Next, the shader team programmed the appropriate shaders for each object. Shaders, in the field of computer graphics, define the levels of light, colors, and textures of all the objects in a given scene. Once done, the lighting team decided the final lighting of the scene, mainly adding mood lighting to help with storytelling. Finally, the completed scene is sent to the "render farm", a set of computers dedicated to rendering each frame of a scene resulting in an animation clip (*The Making of Toy Story*).

Image 6: Using Menv to animate Woody
Source: *The Making of Toy Story*, 1995

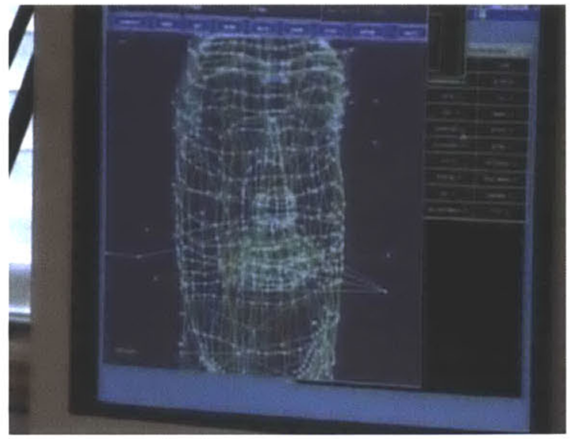


Image 7: Detecting vertices for Buzz Lightyear using a clay model
Source: *The Making of Toy Story*, 1995

The significance of the animation lay in the technology involved to produce such a film. The film was rendered with Pixar's proprietary rendering software RenderMan. Additionally, Pixar developed and used a program called Modeling Environment (Menv). This tool allowed the creation of three dimensional computer character models with built-in controls, mimicking the function of muscles and joints. With this, animators could come in and specify at what frames and in what position a model should be, a technique known as key framing. The computer would then interpolate the intermediate movement and create a smooth motion. Animators, most of whom had never used computers, were now able to create precise yet fluid animations that may have not been possible in hand drawn animations (Snider).

2.3 The Japanese Animation Industry

While John Lasseter was intent on creating a new medium for animations via computer generated imagery, a portion of the world continued with the traditional mode of production: hand drawing. Japan is well-known for its own subset of animation, commonly known as *anime*. One such studio is Studio

Ghibli, famous for its partnership with Disney and its production of animated films including *My Neighbor Totoro* in 1988, *Spirited Away* in 2001, and *Ponyo* in 2008. Films produced by Studio Ghibli are usually in the science fiction genre and demonstrate some sort of opinion on a global scale problem, for example man's industrialization versus nature in *Princess Mononoke* which was produced in 2007. For Studio Ghibli, each frame is hand-drawn using cel animation techniques. To achieve the artistic effect, cels are hand painted with watercolors (Ghibli Gabble). While other *anime* studios have begun switching to using computers to speed the production process, Hiyao Miyazaki, the director of the studio, is skeptical of computer animation, believing that hand drawn scenes are the only way to achieve the artistry that he seeks. Due to the time constraints of the modern world, however, Studio Ghibli has used CG for camera panning, digital composition, as well as some image generation (Ebert).

Other *anime* producers have been less skeptical of using technology to produce their films, in fact, Makoto Shinkai embraces the use of computers and even looks forward to future technological developments. Makoto Shinkai became first known for directing the animation *Voices of a Distant Star* in 2002 and has since produced many others including *5 Centimeters Per Second* in 2007 and *Children Who Chase Lost Voices* in 2011. Using technology to speed the process of production was vital to his success. *Voices of a Distant Star*, was a 25 minute short with impressive imagery single-handedly made by Shinkai and his fiancée in 7 months using only a Power Mac G4. He now uses other tools including a Wacom tablet, Adobe Photoshop, and After Effects. By using digital tools, his animations have achieved more sophisticated lighting and color schemes. For example global color illumination could now be reflected in shadows and refractions, mimicking colors in real life (Fenlon).

Image 8: Still frame from *Princess Mononoke*
Source: *Princess Mononoke*, 2007

Image 9: Still frame form *5 Centimeters Per Second*
Source: *5 Centimeters Per Second*, 2007

2.4 The Walt Disney and Pixar Contribution

Though Studio Ghibli and Makoto Shinkai have both been great contributors to the evolution of modern animation, their use of computer generated imagery for animation has been minimal compared to that of Walt Disney Animation Studios. Walt Disney Animation Studios began to use computer graphics beginning with *Chicken Little* in 2005. Since then they have produced *Bolt* in 2008, *Tangled* in 2010, and *Wreck-It Ralph* in 2012.

By 2010, computer animation had made leaps and bounds, but still problem spots remained especially in the simulation of hair and cloth. Both require particle system connected by a spring system. A single strand of hair, for example, is made of a chain of particles connected by springs, while a square piece of cloth is a square array of particles with spring forces between each. For *Tangled*, simulation of the main character's, Rapunzel, 70 feet of hair required the help of Disney's own software *dynamicWires*, a mass-spring system for curve dynamics. This software imitated the behavior of the hair piling on top of itself and other objects by creating spur-of-the-moment spring forces at collision times during simulation. Tangential forces had to be applied to the hair strands in order to make it seem like Rapunzel effortlessly dragged her hair behind herself (Ward).

Wreck-It Ralph also added to Disney's warehouse of techniques. Portraying different gaming dimensions, lighting and animation style had to be unique for each world. In the world of Fix-It Felix, 8-bit game style was displayed with choppy character animation movements; Hero's Duty reproduced a first person shooter game setting; and Sugar Rush contained objects that simulated the texture of candy. Because of all the glossy candy surfaces in Sugar Rush, lighting of surfaces played a large part in creating a convincing environment. Disney thus had to develop yet another software called Gummi Shader to specifically make all the candy believable (Tipton).

Image 10: DynamicWires used for hair physics
Source: *Tangled*, 2010

Image 11: Gummi Shader for glossy candy surfaces
Source: *Wreck-It-Ralph*, 2012

Contributing even more to the development of computer graphics tools for animation is Pixar Animation Studios. To date, Pixar has produced 14 computer animated feature films, each making their own contribution to computer graphics. Like Walt Disney Animation Studios, Pixar also had their fair share of hair simulation issues, but in their case, it was Sulley's fur in *Monster's Inc*, produced in 2001. Unlike *Tangled*, Sulley's fur used a spring and hinge system. This gave animators greater control over the fur's movement (Davis). For *Up*, produced in 2009, the computational problem lay with finding a way to animate over 10,000 helium balloons with their strings while lifting a house up in the air, not even thinking about the ripple effect of a single balloon bumping into another. Working with a set of balloon cluster experts, the Pixar team was able to define a set of algorithms and procedurally animate the balloon scenes (Terdiman).

Pixar's greatest successes in technological enhancement, however, lay with one of their newest films, *Brave*, produced in 2012. *Brave* posed issues in hair, cloth, and liquid simulations. While Pixar successfully produced a film that solved each of these very specific issues, they spent years to develop software that would achieve animations that can be easily recreated in real life within seconds (Terdiman).

CHAPTER 3: LOOKING AT MIXED MEDIA

Due to the rise in the creation of mixed media, there is a growing confusion in the distinction between 2D and 3D. It is possible to draw by hand an image so photorealistic that it would seem as if it were a 3D image. Similarly, it is also possible to render a 3D image with certain toon settings to make elements seem two dimensional and cartoon-like. The answer is not as simple as saying 2D consists of images drawn by hand and 3D consists of images rendered by computers. Especially with technology developing, many software programs such as Pencil, Synfig, Adobe Illustrator, and Adobe Photoshop exist to aid in the production of 2D animations. Additionally, 3D software packages such as Autodesk Maya and Autodesk 3D Studio Max oftentimes require rasterized 2D images as texture maps, backgrounds, skyboxes, bump maps, and normal maps.

3.1 Defining 2D and 3D

For the purposes of this study, clear definitions of 2D and 3D must be set. The following definitions will be used in this thesis. 2D animations will consist of elements constructed within a 2D environment while 3D animations will consist of elements mostly constructed within a 3D environment.

With this distinction made, there can now be defined differences between the two media. Beginning with 3D, all elements have some sort of volume or mass. Though lines and planes arguably have no mass or volume in the real world, it will be assumed that within 3D software packages, some volume and mass exists, though infinitesimally small. In computer generated imagery, a certain pipeline must be followed after the creation of the mass in order to produce and render an animation. Most unique to the 3D methods are the animation and lighting processes.

To animate a figure, or character, there are three main components involved. The first is the skin, which consists simply of the geometry of the model. Autodesk Maya specifically uses polygonal modeling and subdivision surfaces. From the skin, a skeleton must be created. Though called a skeleton, the skeleton is in actuality a construction of joints. These joints are connected by bone-like objects that simply serve as a visual tool to depict the hierarchy of the joints. The skeleton must be bound to the skin in order for any type of dependency to be created. The last component involved is rigging. Character rigging can be thought of as a rule set that defines how joints may move within a model. For example, wrists on a hand operate as a ball and socket joint and thus have almost complete rotational freedom. Knees, on the other hand, may only bend in a single axis.

When animating a movement, such as a walk, the 3D differs from the 2D in that the entire model must be animated. From the profile view of the character, though only the arm closest to the camera is seen, both arms must be animated in the 3D environment. The arm furthest from the camera, though not always in view of the camera, is always moving even if this movement is not immediately apparent. This is unlike 2D in which the arm that is not seen in the scene does not have to be drawn or animated.

In terms of lighting, computer generated imagery has what is called global illumination. In 2D animation, a light source has no effect on any of its elements. For example, the sunrays shining into a room through a window will not light up the bed or the walls. Instead the illumination of the room is an illusion created by the artist's coloring conventions. In the 3D setting, however, the light source is what causes elements to become visible and seen. Depending on the type of raytracer used and the number of bounces designated, the lighting within the same 3D scene can be significantly affected. Since global illumination simulates how light would hit and bounce off of objects, reflectivity plays a large role in 3D animation. For example, a white sphere placed in a room with bright green walls and some sort of light source will be tinted green as a result of the light bouncing from the green walls onto the white sphere. More advanced topics such as caustics, or the bending of light due to the change of medium that the light penetrates, are also reliant on global illumination but rely more on the actual behavior of the light as defined by mathematical algorithms.

2D animation also has two points unique to itself. Firstly, the tolerance for still frames is higher than that of 3D animation. Due to the labor intensiveness of traditional 2D animation creation, animators occasionally utilize the technique of showing a still frame for the length of several frames. Even though a still frame is shown, viewers tend to have a lower tendency of noticing that a frame is held in the 2D than in the 3D. Because of this, 3D animations must practice what is called the moving hold. For example, a 3D character that is not moving dramatically must blink, periodically relax then tense its jaw muscles, slightly flare its nostrils, and perhaps slowly sway from side to side. In contrast, a 2D character that is not moving only requires the single frame to span over the allotted number of frames in which there is not distinct character movement.

Secondly, because 2D does not have the connection to realism that 3D has, 2D animation has a higher threshold and allowance for artistic expression and style. For the reason that 2D art has had a longer and more developed history than 3D art to this date, artists have a greater familiarity with producing different types of 2D art. Within the 2D art realm, there are different medias such as watercolor, oil paint, charcoal, and color pencil to name a few. 3D on the other hand is restricted to pre-made render settings provided in the 3D software package of choice. Additionally, because 3D is based off real physical forms and sometimes gravity, there is little flexibility in drifting away from these forms. 2D on the other hand does not have these limitations and can allow a level of abstraction that cannot be recreated in 3D. Notably, however, there is ongoing research on attempting to recreate this artistic element found in the 3D environment.

3.2 Uncanny Valley

The term "uncanny valley" was coined by Japanese robotocist and Professor of Engineering at the University of Tokyo Masahiro Mori in 1970. His contribution to the study involved whether of not increased likeness of a robot to the human form would affect the relationship between the affinity felt towards the robot and the realism of the robot. His proposal stated that as a robot grew to seem more human-like, so would human acceptance grow in reaction. This upward trend would continue until the robot reaches a nearly human state in which there will be a sudden drop in human acceptance to which there will not be a recovery from until the robot reaches exact human-like replication (Borody).

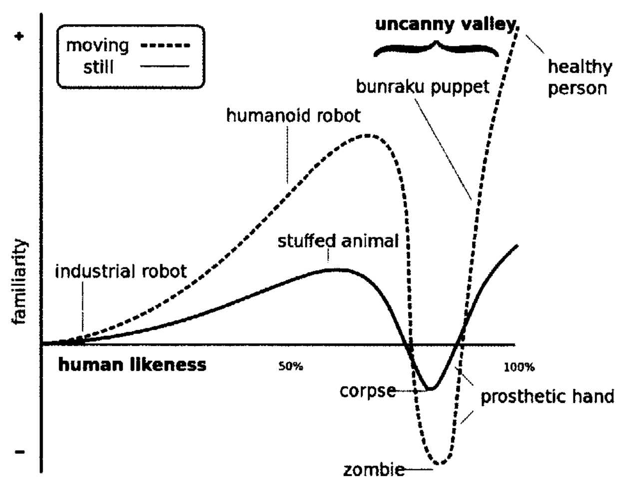


Figure 1: Graph of uncanny valley with examples
Source: *Bukimi no Tani Genshō*, 1970

Though coined and suggested by Masahiro Mori, this phenomenon was not proved by Mori. Modern research has conducted psychological tests in order to prove the existence of this phenomenon. While by now various research has more or less proved its existence, there lacks an explanation for why the uncanny valley occurs (Pollick). Nonetheless, computer generated imagery and 3D models are also thought to reflect this effect of uncanny valley.

As 3D software and technology continually grows, generated models and images have increasingly shown greater similarity to real life. As the likeness grew, so did human affinity until the point at which uncanny valley was reached. Awareness of uncanny valley in CGI was first made in 1988 with the production of Pixar's short film "Tin Toy". In this short, Billy, the first ever computer animated baby, elicited extreme discomfort from the audience, especially as

Billy chased the tin toy around the room. While the actual creation of Billy and use of Pixar's new Menv animation program was widely accredited, many found the animation to be both frightening and disturbing ("The Uncanny Valley and CGI").

The next major appearance of uncanny valley in 3D animation was seen in *Final Fantasy: The Spirits Within*. The hyper-realism of the characters and use of the new motion capture technique was unprecedented, however, many attribute its box office failure to the still mechanical movements of the body as well as the lack of expression in the eyes. Andy Jones, the animation director, is quoted to say, "It can get eerie. As you push further and further, it begins to get grotesque. You start to feel like you're puppeteering a corpse" (Kaba).

Another famous example of uncanny valley appearing in animation can be found in Warner Brothers' *The Polar Express*. Like *Final Fantasy*, *The Polar Express* was also created using motion capture techniques. While by the time of its production improvements in body movement have been made, animators and technical directors still struggled with expressing facial emotions. In motion capture, all facial movement is made by the live actor in response to an emotion. The animation, however, must be recreated by the animator and applied to the model. Thus, facial expressions may seem odd or life-less (Kaba).

Image 12: Example of uncanny valley in *Final Fantasy*
Source: *Final Fantasy: The Spirits Within*, 2001

Image 13: Example of uncanny valley in *The Polar Express*
Source: *The Polar Express*, 2004

Uncanny valley is notably more prominent in 3D animation than in 2D animation. Little research has been completed on incidences of uncanny valley in 2D animations.

3.3 Mixed Media: *Paperman* by Walt Disney Animation Studios

With a defined understanding of 2D and 3D, the form of mixed media can now also be studied. Many animation studios today have found beneficial the combination of 2D with 3D for various reasons. Reasons may include serving to expedite the production process, allowing for greater artistic effect, or enhancing a storyline. This thesis will look specifically at *Paperman* as a case study of mixed media and will briefly discuss other examples and purposes for the mixed media.

Recent Academy Award winning Best Animated Short Film, *Paperman* is a 7 minute romantic comedy produced by Walt Disney Animation Studios and directed by John Kahrs. Having a touching, yet exciting storyline, this short famously combines 2D with 3D animation. Disney pays homage to its past of classical cel animation, in which each frame of an animation is hand drawn, while simultaneously looking towards the future of 3D computer generated graphics. According to his interview, Kahrs wanted to achieve a look that was both realistic and believable, but at the same time held the magic and style of traditional Disney animations. This he achieved by overlying 2D brush strokes over a rendered 3D CG layer (Wong).

Because the 2D animation had to move with the CG, Disney had to create a new animation process in order to produce this short. First, the base 3D model and graphics had to be created. The process began as any 3D animation film would with modeling, rigging, and animation. Disney then created motion fields from the computer graphics so that the 2D portion may be mapped to the 3D. The creation of the motion fields required per-frame per-element renders such that each pixel in the image would have a 2D offset. Ultimately, this process allows the flow of the 2D portion (Wong).

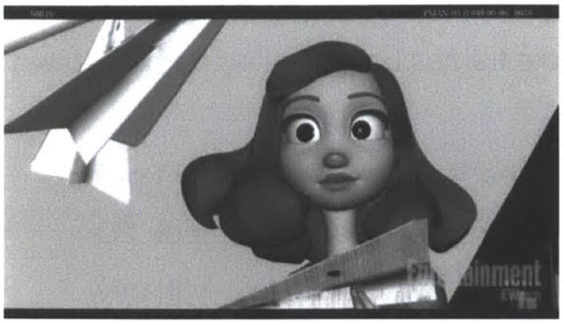


Image 14: CG animation layer
Source: *Paperman and the Future of 2D Animation*, 2012

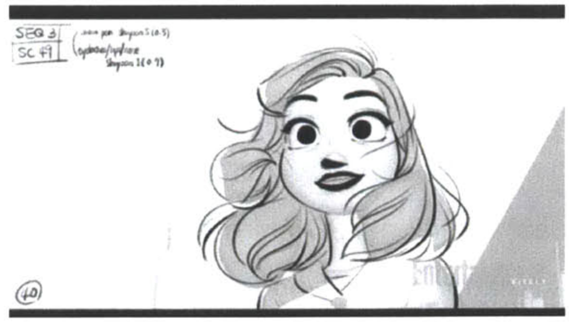


Image 15: Hand drawn key poses
Source: *Paperman and the Future of 2D Animation*, 2012

Image 16: Hand painted lighting key
Source: *Paperman and the Future of 2D Animation*, 2012

Image 17: Hand drawn hair animation layer
Source: *Paperman and the Future of 2D Animation*, 2012

Image 18: Completed hand drawn layer
Source: *Paperman and the Future of 2D Animation*, 2012

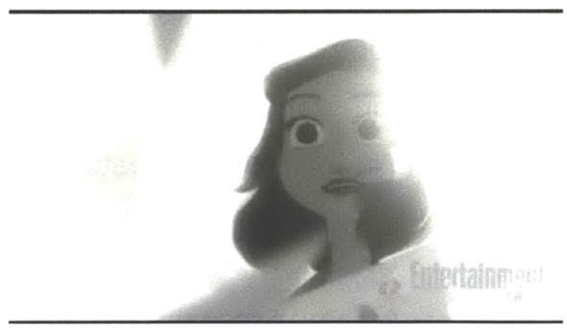


Image 19: Final composite
Source: *Paperman and the Future of 2D Animation*, 2012

Preserving the traditional style of Disney drawings, silhouette ribbons had to be created for the characters. Characters, from their silhouettes were then divided into cylindrical components. By settings these components perpendicular to the camera direction, the light feel that 2D animations have when characters move could be achieved. The offset provided a large buffer for the character movement. Further more, an applied paper texture gave the image more depth. Then using the software Meander, line artists drew 2D drawings over the CG renders to generate non-key framed drawings. Motion pasting allowed for the generation of all intermediate frames between key frames drawn by the artist (Wong).

Paperman was a short that stood out because of the homage it paid to past traditional 2D animation while using 3D technologies. In the past, 2D animation was laborious and difficult to produce. By combining it with new technology, however, this vintage style can be preserved while still maintaining the desired effects of a 3D rendered animation.

Not all forms of mixed media, however, have the same artistic intentions as *Paperman*. For previously mentioned Makoto Shinkai, the CGI component served as a way to enhance the artistry seen in his animation. Other Japanese animators use computer graphics to speed up the production process and cut costs. Computer generated imagery was generally used for cars, machines, or rigid bodies that only had simple rotation, scaling, or translational transformations. Use of 3D in 2D animations is not only limited to Japanese animations. American animations such as *Anastasia* is seen to use 3D for creation of complex architectural model. Disney's *Mulan* can also be seen to utilize 3D in programmatically creating the thousands of Hun warriors in the battle scene in the mountain battle scene.

Image 20: Architectural 3D model
Source: *Anastasia*, 1997

Image 21: Programmatically built 3D Hun warriors
Source: *Mulan*, 1998

In the beginning, there was very little integration technique of 2D with 3D as the 3D model lacked the level of details that are present in 2D drawings. Shinkai, who also uses computer graphics in his works, notes that in order for the 3D computer graphics to blend seamlessly, careful texture mapping and cel shading needs to be applied. Even so, the "too perfect" movements of the CG animation gives away what is 3D and what is 2D.

Another of Walt Disney Animation Studio's productions, *Get a Horse* is a short film that uses the mixed media as a part of the story line and as a way to enhance the plot. An interaction between 3D animated characters and characters animated in a style reminiscent of the Disney animations in 1928, the short uses the 2D world and the 3D world as two environments that the characters go back and forth between.

Image 22: 2D and 3D scene interaction
Source: *Get a Horse*, 2013

Image 23: 2D and 3D scene interaction
Source: *Get a Horse*, 2013

Mixed media has reached many new forms. Each of the films utilizing mixed media serves its own purpose. Arguably, however, there may be a misunderstanding of how to properly create mixed media without first understanding the strengths and abilities of the 2D and 3D forms. By discovering these strengths and abilities, enhanced technique can be used in the composition of these mixed media forms.

CHAPTER 4: PRODUCTION PIPELINES

To conduct the research for this thesis, two mixed media animations were created from both a 3D animation as well as a 2D animation. Both the 2D and the 3D animations depict the same sequence of scenes. The mixed media animations combine the two animations using light as the controlling factor.

4.1 Building the 3D

Autodesk Maya was used as the 3D software program to create the 3D animated base for this thesis. Following a standard 3D animation pipeline, the following tasks were completed.

Focusing first on the character model for the animation, the skin of the character was modeled using box modeling techniques. All components except for the pupils of the character are in a single mesh. The pupils are separate objects but parented to the rest of the character.

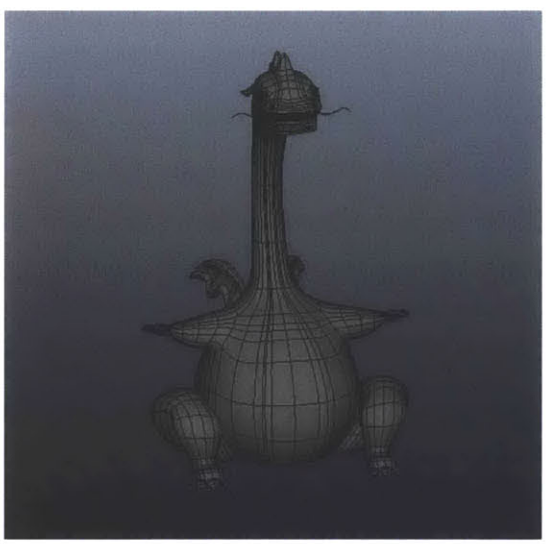


Image 24: Front perspective view of model

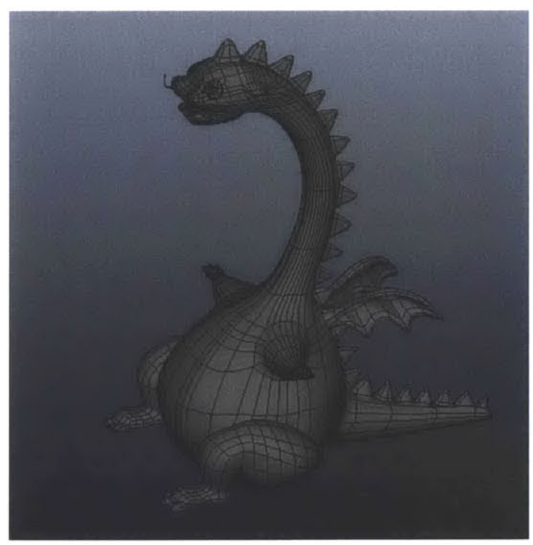


Image 25: Side perspective view of model

Next, using the skin as a base, the skeleton was built. Each leg consists of 5 joints including the hip joint. The toes are not differentiated and have a single bone and joint to control each. Each arm consists of 13 joints including the clavicle. Unlike the toes, the fingers are differentiated. Each finger consists of 3 joints that are connected to the wrist joint. Each wing consists of 10 joints. The neck and the head are composed of seven joints. To stabilize the midsection of the character, three joints line the front of the midsection of the character.

Only a single joint was used to control the direction and position of the character's face. Because blend shapes were used to control facial animation, no joints were required. Initially to bind the skin with the skeleton, the Smooth Bind tool in Maya was used. To fix and the bind weights, the Interactive Bind Weights tool was used to hand paint the weights for each of the joints. Since each of the horns were an extrusion of the mesh and did not have a joint of its own, careful attention was paid to give each of the horns the correct bind weight in order to avoid shearing when animating the model. All joints were joined to the base joint at the bottom of the spine.

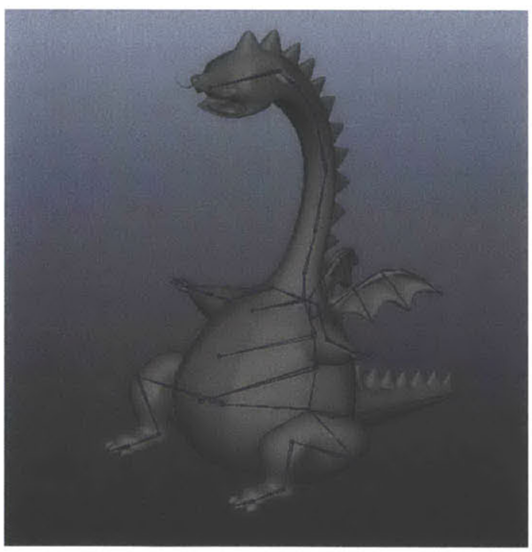


Image 26: Skeleton with the skin

Image 27: Skeleton without the skin

Once the skeleton was built, the character was rigged. A variety of IK Handlers and curve controls were used to construct the rigging of the character. IK Handlers were used to connect the top of the neck to the base the neck, the base of the neck to the base of the spine, the tip of the tail to the base of the spine, each hip to each ankle, each ankle to each toe and each wrist to each clavicle. Additionally, each wing tip has an IK Handler connecting it to the joint at the base of the neck. These IK Handlers provided enough skeletal support so that the mesh would not have severe distortions or shearing. Curve controls allows for greater flexibility and ease of animation. Curve controls were mainly used for the heel, toes, and foot, as well as each of the midsection spinal joints, and a single curve control for the waist of the model.

Image 28: Rig controls with the skin

Image 29: Rig controls without the skin

Next, texture and blend shapes were applied. For textures, lambert materials of different colors were applied to the mesh. No bump or texture maps were added to save both creation time and rendering time. Two blend shapes were created for the face. The first blend shape consisted of simple eye movement, and the second blend shape closed the mouth of the character. The character was then animated and placed into an environment.

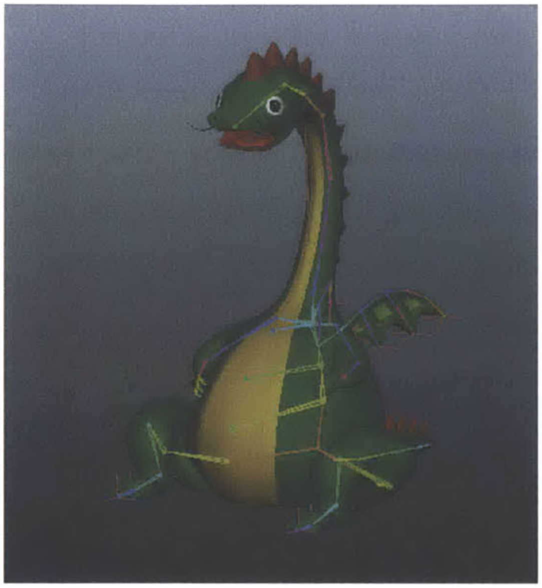


Image 30: Character with applied texture

Image 31: Character in animation

The environment built for the character was a simple room that consisted of a window and a shelf. The character was placed on the shelf. Each of the newly created objects were also created from box modeling technique. Into the scene three lights were used: an ambient light for inside the room, a sunlight to act sun passing into the room via the window, and a directional light to act as the moving sunlight. A single camera was used to survey the space. Once all items were in place, the animation was rendered.

Image 32: Entire scene

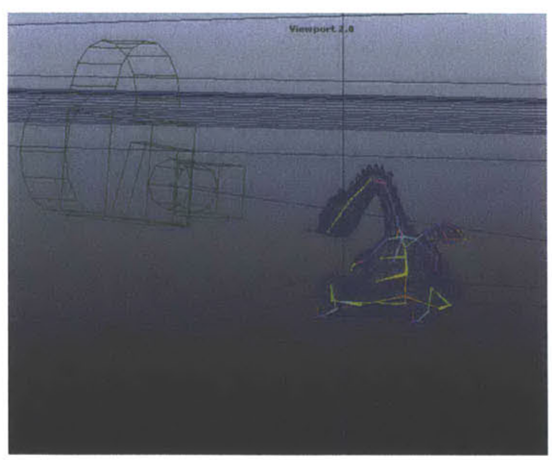


Image 33: Scene with character and camera

4.2 Building the 2D

The following software programs were used to create the 2D animated base: Adobe Illustrator, Adobe Photoshop, and Adobe Premiere Pro. The 2D animation portion can be split into two ways of creation. The first was using assets and the second was creating frame by frame hand drawn images. All elements of the 2D animation were created using assets except for the character animations. Character animations required frame by frame hand drawing. While 2D animation tools such as Pencil and Synfig are available, none of those tools were utilized.

The asset method required the drawing of individual elements in Adobe Illustrator and importing them into Adobe Premiere Pro for translation, scaling, rotation, and warping. All architecture elements including the window, wall, and shelf, required non-uniform scaling to imitate perspective changes. For extreme distortion such as that seen in the shelf, the Corner Pin tool was used. To imitate the lighting change, two versions of the same scene were created. One version was dark, as if no light had been shined on the scene, and the other version was light, as if light had been shine on the entire scene. To merge these two scenes and imitate the changing light, the Four-Point Garbage Matte tool was used.

Image 34: Using assets to build 2D

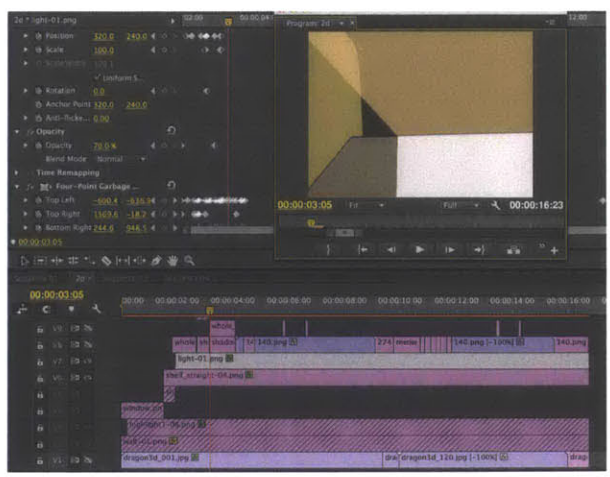


Image 35: Assembly and effects used on assets

Frame-by-frame character animation consisted of three parts: line drawing, shadows and highlights, and fill color. Each of these parts compose its own layer in Adobe Illustrator. The frame-by- frame character animation also consists of the character shadow that animates with the character. Since this is a shadow, however, there is no line drawing or fill color. Additionally, due to the labor intensiveness and time required to create each character animation frame, only one frame was drawn for every 2-5 frames, depending on the point in animation (see Chapter 6.1 Sources of Error). The frame by frame character animation could then be inserted into the asset based 2D animation in Adobe Premiere Pro.

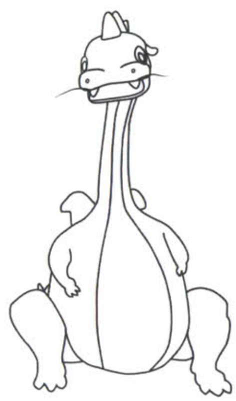


Image 36: 2D line drawing

Image 37: Shadows and highlights

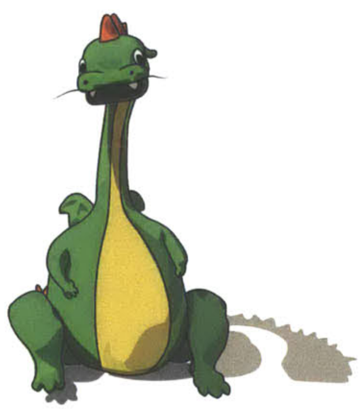


Image 38: Fill color

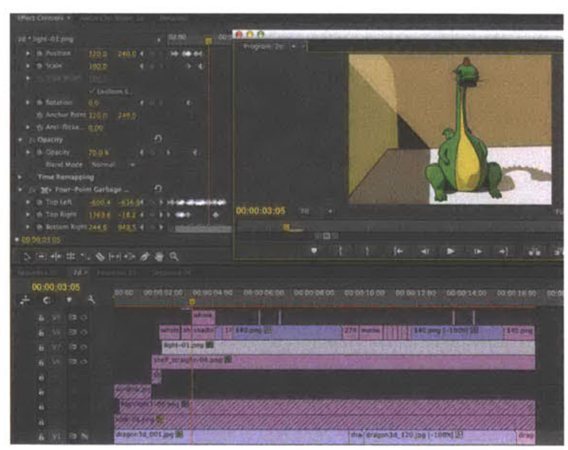


Image 39: Merge with asset animation

4.3 Mixing the Media

Once both the 3D and 2D bases were created, the two versions of the mixed media were created in Adobe Premiere Pro. The thought behind creating the mixed media animation was to carefully overlay the 2D and 3D and intersect each only where the light shines. Thus each of the two mixed media uses either 2D or 3D as its base animation. Areas where light has been shined reveal the opposite animation. To achieve this effect, the Four-Point Garbage Matte tool was used.

Image 40: 3D based mixed media

Image 41: 2D based mixed media

CHAPTER 5: PERCEPTIONS OF SPACE AND CHARACTER MOVEMENT

To measure the differences in the perception of space and character movement, a survey was conducted. Each participant was given either the 3D base video or the 2D base video to watch. Which video the respondent watched was determined on a random selection. The respondent was asked to choose a number either '1' or '2'. If '1' was chosen, then the respondent was taken to watch the 3D base video. If '2' was chosen, the respondent was taken to watch the 2D base video. No matter which video was watched, all participants were asked the same set of questions on the given animation that he or she watched. There was a total of 43 respondents. Of the 43 respondents, 18 watched the 3D base animation and 25 watched the 2D base video. Note that in the 3D base animation, a majority of the character movement was seen in the 2D animation format and vice versa.

Most of respondents were between the ages of 18-35. Of all the respondents one was under 18, one was between the ages of 35 and 50, and 6 were over the age of 50. Additionally, since mostly college students were surveyed, most of the respondents identified themselves as engineers, scientists, and students. Other popular identifications included designer, and athlete. Other self-entered identifications included business person, medical professional, and teacher. Most respondents ranked their understanding on the difference between the 2D and 3D animation pipelines at a level of 1. The ranking was such that 1 was the lowest level of understanding and 5 was the highest level of understanding. There were 12 respondents who ranked their understanding at a level of 2, 5 respondents who ranked their understanding at a level of 3, and 2 respondents who ranked their level of understanding at a level of 5.

The survey is composed of two sections. The first section asks questions to determine the respondents perception of space and scene movement. The second section asks questions that focus on character movement. In order to control results, answers were pre-made and respondents were asked to either select their answers from images or from a pre-made list. All questions required an answer.

Due to the nature of this survey, it was assumed that all respondents are watching their respective animations with the same perception and judgment. For example, Respondent A who is watching the 2D base is expected to have the same responses to the 3D base as Respondent B watching the 3D base if Respondent A were to watch the 3D base. The opposite applies to Respondent B's responses on the 2D base.

Refer to page 49 to view the survey questions and available answers in their entirety.

In the following result charts, all blue charts indicate results from the 3D based video surveys while all orange charts indicate results from the 2D based video surveys. Note that the darker, more saturated, portion of the chart indicates those who answered correctly.

Part 1, Question 1: Which of the above plans best depicts the layout of the room as seen in the animation?

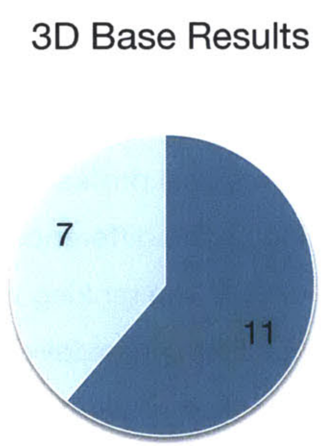


Figure 2: 3D based responses to Question 1

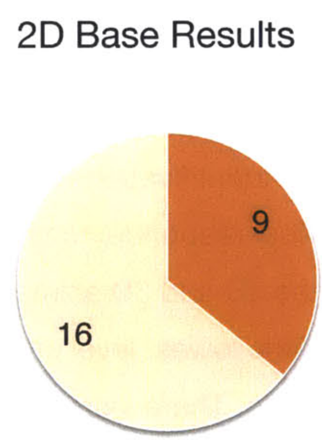


Figure 3: 2D based responses to Question 1

Out of the 18 responders on the 3D base, 11 answered correctly and 7 answered incorrectly. Of the 25 responders on the 2D base, 9 answered correctly and 16 answered incorrectly.

These results indicate that the 3D base animation better depicted the spatial layout of the room than the 2D base animation. The spatial layout was revealed prior to any lighting changes, thus indicating that 3D animations allow for enhanced spatial understanding over 2D animations.

Part 1, Question 2: Which of the above illustrations best represents the trajectory of the camera over the duration of the animation?

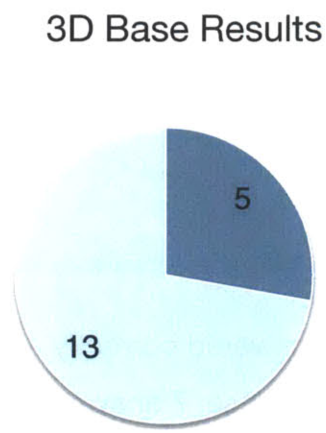


Figure 4: 3D based responses to Question 2

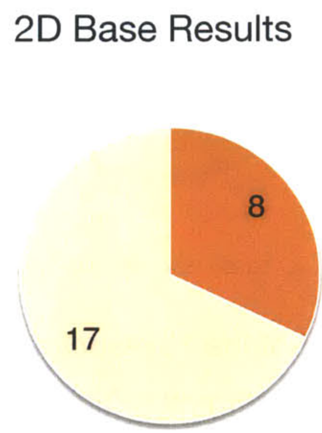


Figure 5: 2D based responses to Question 2

Out of the 18 responders on the 3D base, 5 answered correctly and 13 answered incorrectly. Of the 25 responders on the 2D base, 8 answered correctly and 17 answered incorrectly.

This question judged scene movement based on camera trajectory. Because in the 2D animation, no camera was present, the camera is based solely on the 3D representation. For 2D animation this question measured whether or not there was an understanding of scene movement as depicted by a fictitious camera. The ratio of results only slightly suggests that the 2D based animation was a better indicator of scene movement even though the 2D animation does not truly have a real camera.

Part 1, Question 3: Referring to the above image, which of the following combinations best describes the position of the camera in relation to the character after the camera has come to a resting position?

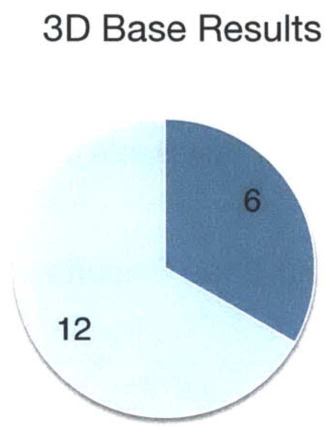


Figure 6: 3D based responses to Question 3

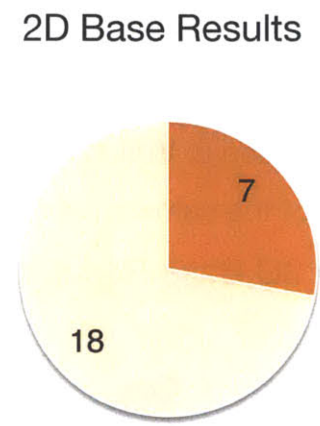


Figure 7: 2D based responses to Question 3

Out of the 18 responders on the 3D base, 6 answered correctly and 12 answered incorrectly. Of the 25 responders on the 2D base, 7 answered correctly and 18 answered incorrectly.

This question judged scene position based on camera position. Similar to Question 2 there is no camera present in 2D animation and thus the camera is based solely on the 3D representation. For 2D animation this question measured whether or not there was an understanding of scene position in relation to the viewer. The ratio of results only slightly suggests that the 3D based animation was a better indicator of scene position over the 2D animation.

Part 2, Question 4: Please indicate which of the following parts of the character moved as an individual unit at any point in the animation.

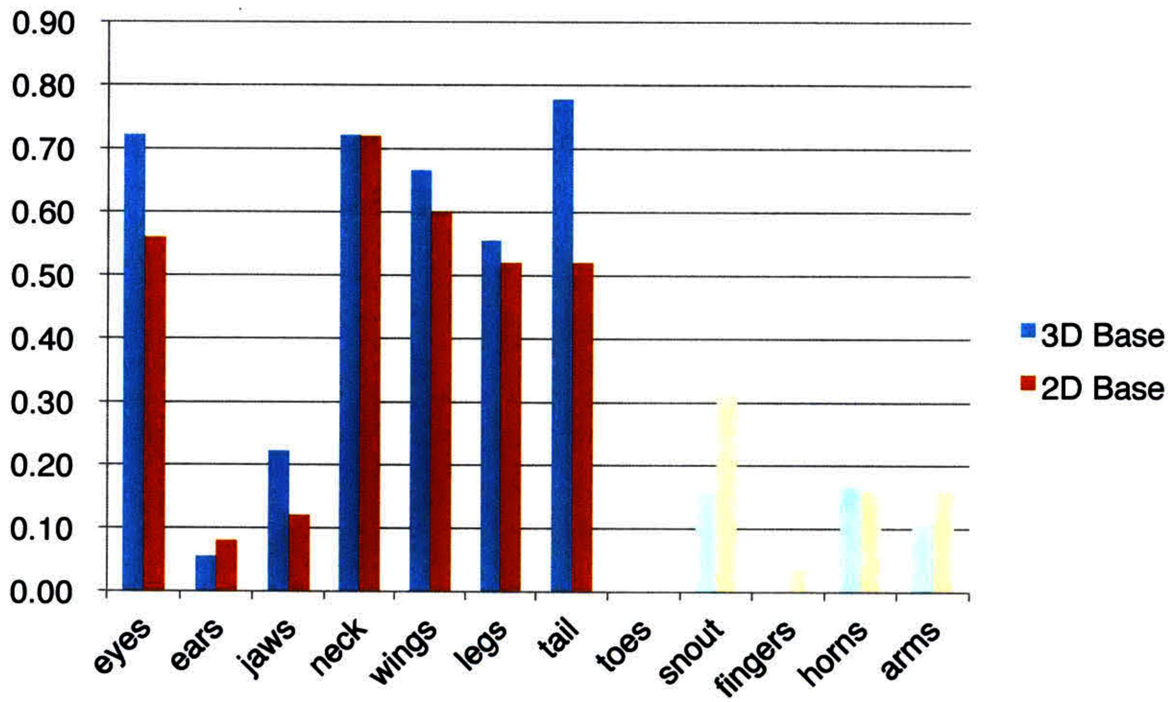


Figure 8: 3D based responses to Question 4

Again note that darker, saturated colors indicate correct answers made by the respondents. To normalize results, the ratios were taken instead of whole numbers. This question targeted the respondents ability to distinguish which of the character's body parts moved as an individual unit and not as a result of other forces. Except for the ears, viewers of the 3D base were more able to discern the correct movements and were also less likely to answer incorrectly. The neck received almost equal responses from both 3D and 2D bases most probably because the neck was the largest and most observable movement. Large differences in results are seen in the eyes and tail. These smaller movements were more noticed by the 3D base viewers. The snout received the most discrepancy of all the body parts that did not move. This may be attributed to the mesh many may have mistaken the snout for the jaw.

Because the 3D base in actuality showed most of the character movement in the 2D animation and vice versa, this shows that there is a lower sensitivity to character movement in the 3D animation than in 2D animation.

Part 2, Question 5: In the duration that the character was moving, were there any points at which the character seems to pause as if frozen?

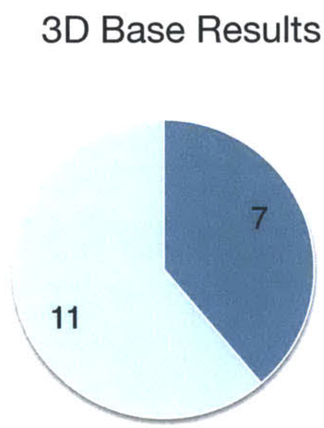


Figure 9: 3D based responses to Question 5

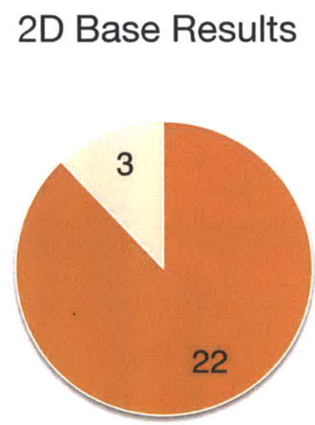


Figure 10: 2D based responses to Question 5

In the case of this question, it is difficult to say whether or not there is a correct answer as this is a matter of perception as understood by the viewers. Technically, however, the 3D base had two sets of back to back still frames at the midpoint of the animation, the first for 20 frames and the second for 28 frames. The 2D base did not have any still frames.

According to the results, out of 18 respondents, 7 perceived that the 3D base had still frames. Of the 25 respondents in the 2D base, 22 perceived that it had still frames. These results support that there is a higher tolerance in 2D animation for still frames than in 3D animation.

Part 2, Question 6: Referring to the video again, please indicate the time stamp in which you felt this pause happen.

Most respondents in both the 3D and 2D base made note that a pause was seen at 0:09-0:10 which was where the back to back frames occur.

Respondents also noted in the 2D base to have felt a pause at 0:03-0:04 of the animation which is the point before the character begins moving. Overall, respondents responded as expected.

CHAPTER 6: CONCLUSION

From the results, it can be generalized that the 3D animation allows for enhanced spatial layout understanding. 2D animation also lends itself to lower sensitivity in character movement. These results, however, are inconclusive without noting the sources of errors or indicating what future research can be conducted.

6.1 Sources of Error

The sources of error can be categorized into three areas: production quality, conduction of the survey, and the demography of the participant audience.

First, in terms of production quality, it is difficult to guarantee the quality of either the 2D or 3D animations. Because of the time limit in completing this research, the 2D animation was not created with the same frame rate as the 3D animation. For a better comparison between 2D and 3D the same frame rate should have been used. Additionally, proper 2D animation software products such as Pencil or Synfig may have produced more successful 2D animation results. In terms of the 3D animation, some shortcuts were taken to achieve the effects wanted. For example, to achieve the lighting change, the ceiling of the room was partially removed. Whether or not this affected the responses of the respondents is unknown.

The survey was conducted online using Google Forms. Because this was an online survey, the survey environment was unable to be controlled. Whether or not respondents answered truthfully or to their best ability is unknown. Since respondents were encouraged to open the video in another window, another uncontrolled factor is the number of times each respondent watched a single video. Additionally, this survey was limited to the extent and abilities of Google Forms. Because Google Forms does not have the option to not allow respondents to go back in the survey, there is an uncertainty of whether of not respondents had watched both animations. In terms of the survey content, many of the questions had a 3D focus perhaps skewing the responses in favor of either the 2D or 3D media. Survey questions should be reworded to not have the 3D focus and instead should be dimension neutral.

Most of the participants were college students studying at MIT. This demographic may have skewed responses in terms of both age and education background. It is expected that each age group would have a different perception of spatial layout and character movement. For a more accurate depiction of results, equal numbers of participants in each age range as well as participants at varying levels of education background should be surveyed.

6.2 Future Research

In addition to refining and re-conducting this research for higher levels of accuracy, future research can include a more in depth comparison of 2D and 3D animation methods. In this research light was being used as a control factor. Additionally, spatial layout and camera movement were combined into a single category. For future research, the four categories of lighting, spatial layout, camera movement, and character movement should be isolated and studied in depth.

To build off of the conducted research, using the results, a new mixed media animation should be created utilizing the found strengths in 2D and 3D animation. From this newly created mixed media animation, participants should be once again surveyed to re-evaluate if the found strengths and weaknesses can be held true.

THE SURVEY

The survey was conducted using Google Forms. Not all pages are numerically linear. The * symbol indicates a required question. Checkboxes (☐) indicate that the respondent may choose from as many of the selections as they wish while circles (**O**) indicate that only one answer may be selected. Correct answers are boxed and highlighted. The form is accessible at this link: https://docs.google.com/forms/d/1zF-HkdutkkYNAQtU9TkDi66-Z5KXBXpexp4gEY333p8/viewform?usp=send_form.

Page 1

1. Please indicate your age group: *
   - O Under 18
   - O 18-35
   - O 35-50
   - O Over 50

2. Check all of the following with which you would identify yourself. *
   - ☐ Artist
   - ☐ Athlete
   - ☐ Designer
   - ☐ Engineer
   - ☐ Scientist
   - ☐ Student
   - ☐ Socialogist
   - ☐ Other:

3. On a scale of 1-5, how well do you feel that you understand the difference between the 2D and 3D animation pipelines? Use 1 as minimal understanding and 5 as complete understanding. *
   - O 1
   - O 2
   - O 3
   - O 4
   - O 5

Page 2

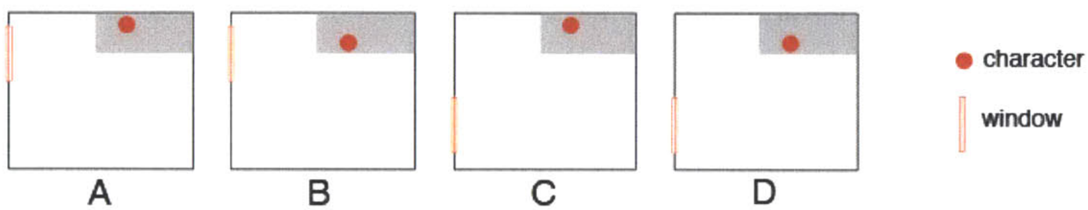
1. Choose a number: *
   - O 1 [takes respondent to Page 4]
   - O 2 [takes respondent to Page 3]

Page 3: Animated Video (2D Base)

Please watch the following animation. You will be asked 5-6 questions based on this video.

Note: It is highly recommended that you open this video in a new tab for future reference.
Video link here: https://www.youtube.com/watch?v=SPXyudgprdo

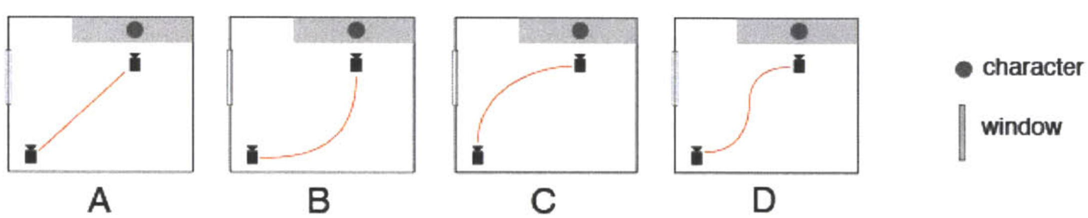
Page 4: Animated Video (3D Base)
Please watch the following animation. You will be asked 5-6 questions based on this video.
Note: It is highly recommended that you open this video in a new tab for future reference.
Video link: https://www.youtube.com/watch?v=vEP48x8k1vQ&feature=youtu.be
Page 5: Part 1.1



Question 1: Which of the above plans best depicts the layout of the room as seen in the animation? Answer based on the previous drawings. *

- ○ A
- ○ B
- ○ C
- ○ D

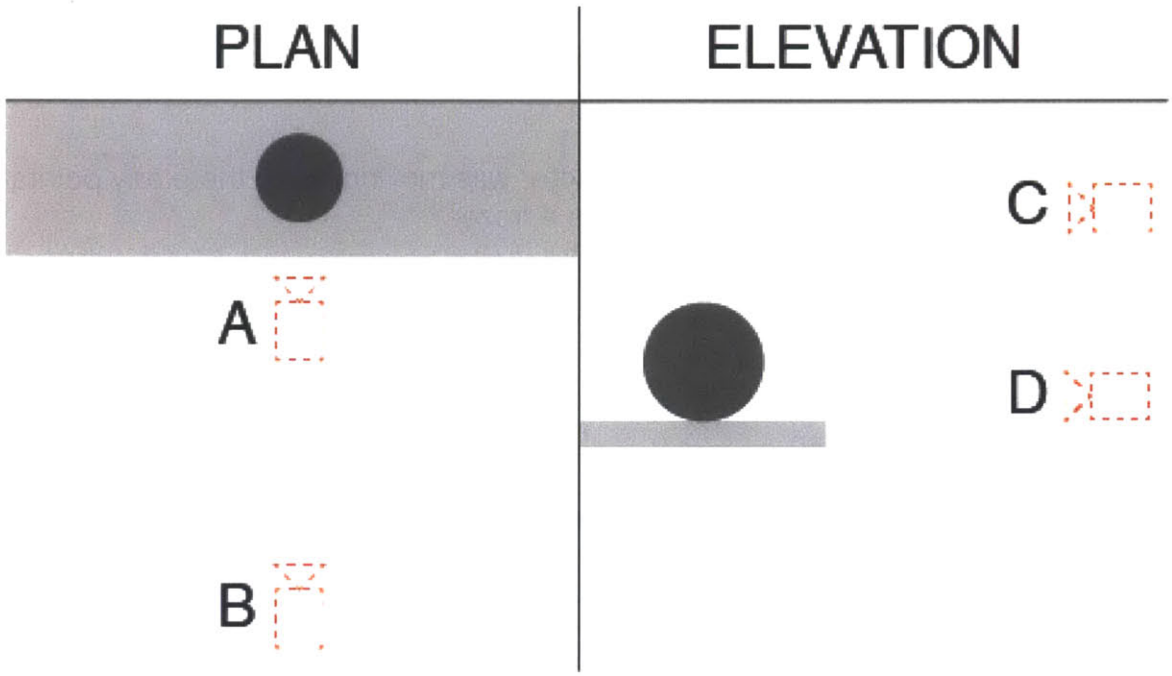
Page 6: Part 1.2



Question 2: Which of the above illustrations best represents the trajectory of the camera over the duration of the animation? Answer based on the previous drawings. *

- ○ A
- ○ B
- ○ C
- ○ D

Page 7: Part 1.3



Question 3: Referring to the above image, which of the following combinations best describes the position of the camera in relation to the character after the camera has come to a resting position? In the plan, the options for the camera are "close" and "far". In the elevation, the options for the camera are "high" and "low". Assume that the character is always the target of the camera. *

- ○ A and C
- ○ A and D
- ○ B and C
- ○ B and D

Page 8: Part 2.1

Question 4: Please indicate which of the following parts of the character moved as an individual unit at any point in the animation. Body parts moving due to the motion of other body parts should not be indicated. For example, the ears moving left because the character walks left is not a valid movement. *

- ☐ Horns
- ☐ Eyes
- ☐ Snout
- ☐ Ears
- ☐ Jaws
- ☐ Neck
- ☐ Wings
- ☐ Arms

- ☐ Fingers
- ☐ Legs
- ☐ Toes
- ☐ Tail

Page 9: Section 2.2

Question 5: In the duration that the character was moving, were there any points at which the character seems to pause as if frozen? *

- O Yes [takes respondent to Page 10]
- O No [takes respondent to Page 11]

Page 10: Section 2.3

Question 6: Referring to the video again, please indicate the time stamp in which you felt this pause happen.

Page 11: Thank you for completing this survey!

Is there anything you want me to know? Please indicate below if you found anything confusing or interesting. If you have any comments or questions, please note them below as well!

BIBLIOGRAPHY

Anders, Charlie Jane. "10 Amazing Milestones of Early Computer Animation". *io9*. Web. Nov 2013.

Borody, W. A. "The Japanese Roboticist Masahiro Mori's Buddhist Inspired Concept of "The Uncanny Valley" (Bukimi no Tani Genshō, 不気味の谷現象)". *Journal of Evolution & Technology* 23.1 (2013) 31-44. Online.

Brown, Blain. *Cinematography: Theory and Practice*. Focal Press, 2002. Print.

Calahan, Sharon. *Storytelling through Lighting, a Computer Graphics Perspective*.

"Chronology- History of Animation- Films from 1926 to 1946". *Animation Supplement*. Web. Nov 2013.

Cohen, Karl. "Milestones Of the Animation Industry In the 20th Century". *Animation World Network.* Issue 4.10. Animation World Magazine. Web. Nov 2013.

Davis, Christopher. *Simulating Clothing and Hair for the Characters in Monsters, Inc*. University of New Mexico. Web Abstract. Nov 2013.

Dirks, Tim. "Animated Films". *filmsite*. Web. Nov 2013.

Ebert, Roger. "Hayao Miyazaki Interview". *Roger Ebert Interviews*. Web. Nov 2013.

Fenlon, Wesley."2D Animation in the Digital Era: Interview with Japanese Director Makoto Shinkai." Tested. 20 Sept. 2012. Web. 10 Oct. 2013.

Fundamentals of Cinematography. University of North Carolina School of the Arts.

Ghibli Gabble. Web. Nov 2013.

Kaba, Fethi. "Hyper-Realistic Characters and the Existence of the Uncanny Valley in Animated Films". *International Review of Social Sciences and Humanities* 4.2 (2013) 188-195. Online.

Lemay, Brian. "History of Animation Timeline". *The Animated Cartoon Factory.* Web. Nov 2013.

Pollick, Frank E. "In Search of the Uncanny Valley". University of Glasgow: Department of Psychology. PDF File.

Snider, Burr. "The Toy Story Story". *Wired*. Web. Nov 2013.

Rangaswamy, Sudeep. *Visual Storytelling through Lighting*. Pixar Animation Studios. 2000.

The Making of Toy Story. Pixar.

"The Uncanny Valley and CGI- The Challenge of Creating Realistic Humans." *SimplyMaya*. Simply Maya, 19 Sept. 2013. Web. 22 Mar. 2014.

Terdiman, Daniel. "Bravely going where Pixar animation tech has never gone." *CNET.* 16 June 2012. Web. 10 Sept. 2013.

Terdiman, Daniel. "How Technology Lifts Pixar's 'Up' ". CNET. Web. Nov 2013.

Tipton, Janelle. " 'Wreck-It Ralph' Inspires Technological Innovations at Walt Disney Animation Studios". *Disney Post*. The Walt Disney Company. Web. Nov 2013.

Ward, Kelly. *Simulating Rapunzel's Hair in Disney's Tangled*. Walt Disney Animation Studios. Web Abstract. Nov 2013.

Wong, Raymond." How Disney used 3D CG to create the 2d world of 'Paperman'." DVICE. 25 February 2013. Web. 23 Sept. 2013.