A STUDY IN PERCEIVED BELIEVABILITY

Utilizing Visual Movement to alter the Level of Detail

Louise Blom Pedersen
Aalborg University Copenhagen
May 24th 2013

Table of Contents

1 Introduction ... 4
2 Motivation ... 4
3 Preliminary Analysis ... 6
3.1 Related Work ... 6
3.1.1 Visual Attention ... 6
3.1.2 The Human Visual System ... 8
3.1.3 Detail ... 10
3.1.4 Believability ... 11
3.2 Defining Believability ... 16
3.3 Preliminary Analysis Summary ... 17
3.3.1 Final Problem Statement ... 18
3.3.2 Criteria for Success ... 18
4 Analysis ... 19
4.1 Movie Theory ... 19
4.1.1 Choice of Genre ... 19
4.1.2 Sound ... 20
4.2 Previous Research on Measuring Believability ... 21
4.3 Choice of CGI ... 22
4.3.1 Render Engine ... 23
4.4 Analysis Summary ... 23
5 Method ... 24
5.1 Variables ... 24
5.1.1 Visual Movement Variable ... 24
5.1.2 Level of Detail Variable ... 24
5.1.3 Measuring Believability Variable ... 25
6 Design ... 29
6.1 Workflow ... 29
6.2 Sketching ... 30
6.3 Rendering ... 30
6.3.1 HDRI Mapping ... 30
7 Implementation ... 31
7.1 Tank ... 31
7.1.1 UV Mapping and Texturing ... 31

7.1.2 Ambient Occlusion and Final Result 32
7.2 The Unidentified Object 33
7.2.1 Implementing HDRI and Ambient Occlusion 33
7.2.2 Polygon reduction 33
7.2.3 Low LoD Material 34
7.2.4 Final Result 35
7.3 The Footage 36
7.4 The Match move 37
7.5 Animating the tracks 38
7.5.1 Create a curve 39
7.5.2 Forming the tank treads 40
7.5.3 Grouping 40
7.5.4 Looping the animation 41
7.6 Final Rendering 41
7.7 Compositing 42
7.7.1 Adding Sound 44
7.7.2 The Narrative 44
8 Experimental Design 45
8.1 Factorial Design 45
8.2 Latin Square 46
9 Results 47
9.1 Participants 47
9.2 Overall Believability Response Rating $\dot{\mathbf{B}}$ 48
9.3 Rating the Clips 49
9.4 ANOVA 50
10 Discussion 52
11 Conclusion 54
12 Future Development 54
13 Appendix 55
13.1 Questionnaire 55
13.2 Sketches 58
13.3 Average response table 59
13.4 Average rating all statements 60
14 References 61

15 Table of figures 66
16 Table of Pictures 66
17 Table of Tables 66

1 INTRODUCTION

Computer generated imagery has become a familiar element in the film world, and despite the fact that these elements are not exactly a new invention, today's creators strive to make these elements as *realistic* as possible. Even though these elements may not be a representation of something that we know from our own world, we still accept them and thus they become believable. One reason why computer generated imagery (CGI) is often used, is that it offers a way to create what else would be limited to dolls or human dressed the part, such as the doll made for the first Alien (1979) movie or C-3PO, played by Anthony Daniels, from Star Wars (1977).Furthermore CGI offers a way to create objects for a scene that otherwise would have to be build. With CGI one can instead use e.g. green screen and later implement objects into the scene as one pleases. However, implementing CGI into live footage still require the creators to make it a part of the scene, give it life and make the viewer accept it as a part of the movies fictional world.

Implementing CGI into live footage is reflected in TV series such as Falling Skies (2010), Game of Thrones (2011), Terra Nova (2011), or movies such as, The Hobbit – an unexpected Journey (2012) or Transformers (2010) – just to mention a few. What they all have in common is a created universe where CGI creatures exist and is a part of the environment, without the viewer questioning their presence.

2 MOTIVATION

Computer generated imagery is commonly used and often there is a tendency to generate much more detail in computer graphics than the user can perceive (Reddy, 2001, Reddy, 1998). Creating these details does not only take time, but can also result in longer render time than necessary. Furthermore if these details are unnoticeable to viewer due to e.g. dim lit conditions or that the object is seen at a distance, they can be considered a waste of time. Therefore I want to investigate if it is possible to reduce the level of detail in a CGI object and thus also reduce the rendering time, without the lowering in detail is noticeable to the viewer.

One area of interest, when it comes to optimizing the field within computer graphics, is the human visual system (HVS) (Hasic et al., 2010, Ferwerda 1997). The reason why the human visual system is taken into consideration, when creating computer graphics, is first and foremost because, even though the human visual system is good it is not perfect (Chalmers et al., 2006). The HVS and its limitations are reflected e.g. when the light intensity chances, certain elements within the human eye changes, and our ability to detect detail change. These elements are known as rods and cones. Rods respond during dim light conditions, and our ability to detect detail and distinguish color is poor, while during day light cones are responsive and provide us with color vision and detail perception is greater than during darker light conditions (Ferwerda, 1996).

Another aspect of the HVS is how the human eye can fixate and concentrate on particular targets or objects, humans can selectively attend to some objects and events and thus in doing so ignore others. This may also be referred to as visual attention (Hasic et al., 2010, Styles, 2005, Yarbus, 1967). When viewing a scene our visual attention is a series of conscious and unconscious process, called top-down processing bottom-up processing, where top-down processing is task dependant, e.g. looking for a target, whereas bottom-up processing is stimuli-driven, e.g. movement of an object (James, 1981). Visual attention can be used to lead a viewer's attention away from other objects in the scene (Hasic et al., 2010, Styles, 2005, Yarbus, 1967). Hasic et al. (2010) proved this by having movement in a foreground objects which led the viewers attention away from a poorly rendered background.

All of the above mentioned means that one should be able utilize the HVS to create a lower detailed model and thus lower the render time, if viewer is focusing on *something else*. I would therefore like to like to investigate if it is possible to reduce the detail of a CGI object, used in a scene where another object attracts the viewer's visual attention, and thus maintain the overall believability of the scene. In other words, because the viewer does not notice the difference in level of detail, they viewer should still accept the CGI as a part of the scene, hence maintaining the perceived believability.

On the basis of my motivation it results in the following initial problem statement.

Is it possible to utilize visual attention toward an object in a scene and on this basis lower the detail in the remainder objects in the scene without it being noticeable and thus maintain the perception of believability?

3 PRELIMINARY ANALYSIS

In this section of the report terms used in the initial problem statement will be elaborated to get a deeper understanding, of how they will be used in relation to this experiment. In order to know to what degree the level of detail can be lowered in an object, it is important to investigate what *visual attention* is and *how* it can be utilized. Therefore previous research relating to visual attention will be looked at. Furthermore it is important to understand what is meant by *level of detail* in order to alter the detail of an object in relations to this experiment. Visual attention and level of detail are the two factors which may or may not affect the users' perception of believability and therefore it is also important to understand what believability is and how to measure it.

3.1 Related Work

Related work will be investigated within visual attention, the human visual system, detail, believability, semantic consistency and uncanny valley. Investigating and defining these terms will provide insight into what has been done before and in the end help narrowing down the initial problem statement

3.1.1 Visual Attention

Visual attention is a process in the brain that involves conscious and unconscious actions, which allows us to find and focus on relevant information quickly (Yarbus, 1967, Chalmers et al., 2006). If information is needed from many different areas of the visual environment, the eye jumps so that the relevant information fall on the fovea, these jumps are called saccades (Yarbus, 1967).

The mechanism of rapidly shift in visual attention is a part of two processes happening in the brain and it was James (1981) that was the first to describe the two processes, bottom-up and top-down, which determine where humans locate their vision (James, 1981). Top-down are task dependent, causing the eyes to focus on those parts of a scene which are necessary for the user's current task, for example looking for targets in a game. (Yarbus 1967, James, 1981).

On the other hand, bottom-up process is automatic and stimulus driven, rapidly and involuntarily shifts in attention to salient visual features of potential importance (Connor et al., 2004, James, 1981). The saliency of an object can be influenced by different visual inputs like color, size, luminance, movement and thus attract the viewer's attention (Theeuwes, 2010, James, 1981, Yarbus, 1967). Shifts of attention and eye movement are initiated toward the point with the highest salience (Nui et al., 2012). Visual salience is the relationship between an object and other objects in the scene. Objects that are highly salient and stand out from the background may immediately receive attention priority (Theeuwes, 2010).

In a study by Zhu et al. (2011), they suggest that in static images luminance contrast appear to be very important to saliency computation. While in video sequences the human visual system is more sensitive to motion than other features, and in general pay more attention to fast moving target than a static background. However, if the scene is static or contain no significant global motion, visual attention will be attracted to other visual stimuli, such as intensity, color and orientations (Zhu et al., 2011).

Sudden change in the visual environment will capture attention e.g. sudden movement in the peripheral vision will immediately be noticed (Styles, 2005). Furthermore, it will draw attention away from what you are currently looking at. Even though sudden movement will capture visual attention, once we have noticed the stimuli that captured our attention in the first place, we are able to orient our attention back to what we were previously looking at or on to another object of interest (Styles, 2005). Yarbus (1967) points out that the more information an element contains the longer the eyes will stay on that given element, however the fixation on the objects can change depending on how useful the object is to the observer (Yarbus, 1967)

In a study done by Hasic et al (2010) they utilized visual attention in the form of movement (Hasic et al., 2010). They wanted to determine if the viewer would notice the difference in rendering quality of a scene if movement was present, versus if no movement was present. Their results showed that they were able to selectively render the moving object in high quality, while reducing the quality of the remainder of the scene without the viewer being aware of the quality difference. However, they did not take into consideration how quickly a viewer can get accustomed to the moving object and perhaps start to ignore it, which would result in the viewer may noticing the quality difference. This statement is consistent with the previously mentioned statement that once the moving object has been noticed, the viewer will be able to attract attention back to other objects in the scene (Styles, 2005).

This is further supported by a study done by Folk et al. (1992) they used a top-down approach to test if an abruptly onset cue would attract the viewer's attention, despite of being instructed to solve a task at hand and ignore it. The test showed that the viewer involuntarily focused his or her attention on to the cue, affecting how quickly subjects could respond to the given task (Folk et al., 1992). Even though the abruptly onset cue did distract the viewer from the given task, they were still able to focus back on the relevant object and distract from the cue, once having noticed the cue.

Other studies suggest that semantic meaning and social relevance of elements within a scene also influence attention. A study done by Birmingham et al. (2009) shows that when asked to look at a visual scene, that included human faces, viewers most frequently fixated on the eyes above other visual stimuli. She argues that the reasoning for viewers fixating on the eye is because they communicate important social information about people. However, she also argues that repeatedly presenting images containing the same type of stimuli, one might risk overestimating how often the viewer would select that particularly stimulus (Birmingham et al, 2009). The HVS tends to suppress responses to frequently occurring features, while at the time remaining sensitive to features that deviate from the norm (Nui et al., 2012).

In a recent study made by Humphrey et al. (2012) it was shown that when a neutral background was edited to contain a single emotionally salient object and a single visually salient object, the viewers tended to fixate on the emotionally salient object more often than the visually salient object (Humphrey et. al, 2012). While visually salient refers to how much an object stands out from its neighboring objects, based on color, intensity and orientation, emotionally salient would refer to something as being of negative or positive valence (Humphrey et. al, 2012), e.g. a dead animal on a road would be seen as negative valence whereas a box of chocolate on a table may be perceived as positive valence.

Looking at previous research it is clear that visual attention can be influenced by different factors, depending on whether it is top-down or bottom-up approach. I want to investigate if the bottom-up approach can be utilized in order to lower the level of detail used in a background object in a scene and still maintain the believability of the scene. The reason for using bottom-up approach is that the experiment is not task-dependent and the object used for the experiment should affect the visual attention in a stimulus driven manner. Furthermore, it could be taken into consideration to use an emotional salient object which according to Humphrey et al. (2012) would be fixated first due to higher saliency than visual salient objects. An interesting observation is the findings from the study done by Zhu et al. (2011) and the fact that in a video sequences we should pay more attention to moving targets than a static background, however a thing to keep in mind is how long the clip presented should be, because the viewer could start to ignore the moving object once having established what it is and thus start to focus on other things in the scene. Another factor to consider is where to place the moving object, from the above research it was established that sudden movement happening in the peripheral vision should immediately be noticed, which means that the object should be noticed fast if placed in the peripheral region.

The next thing is to define level of detail, what is level of detail and how can you alter it. In order to answer this question the human visual system and how detail is perceived will have to be researched. In the next section research into what detail is in relation to the human eye will done and further how one can alter the level of detail of an object.

3.1.2 The Human Visual System

Even though the research in visual attention gave a good indication that movement in a scene should attract the viewer's attention, it is still important to define level of detail in order to alter it. However, it is first and foremost important to understand how the HVS perceives detail and on this basis *how* level of detail can be altered in relation to how it is being perceived.

The eyes ability to perceive fine detail can be described as visual acuity (Wolfe, 2009) and is measured in cycle/degree (cy/de) in grating patterns. Figure 1 shows the contrast sensitivity function which describes how the relation of cy/de affect the overall perception of contrast to a given grating pattern ranging from low (0.1 cy/de) to high (100 cy/de) (Wolfe, 2009). Furthermore, the function shows that the human eye is particularly sensitive to detail around 8 cy/de and that the sensitivity drops at around 60 cy/de, which means that the eye cannot resolve any detail smaller than this limit (Reddy, 2001).

Another important fact is that eyes ability to perceive fine detail is not uniform across the entire retina. In the peripheral field of our vision, the human eye perceive less detail, and in order to see the object in high detail we have to rotate eyes until the light from the object falls directly onto the fovea of our retina (Reddy, 2001). Perhaps this is also the reason why sudden movement happening in the peripheral field will be noticed immediately, we cannot perceive fine detail in that part of our vision, and the moving object appearing could be danger lurking, we have to shift out attention so we are able to see if clearly (Babcock et al., 2002). Velocity is also a factor that can affect our ability to perceive detail (Luebke et al., 2003).

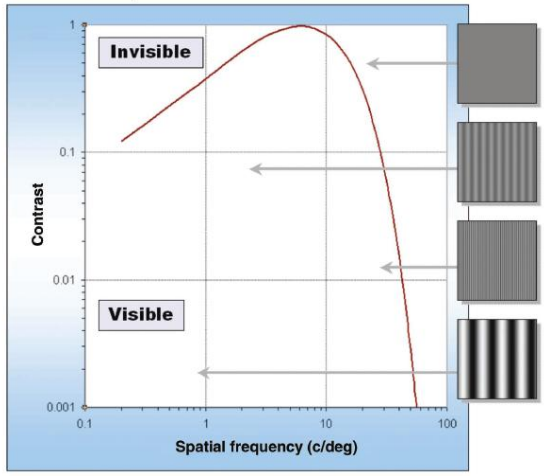


Figure 1: Contrast correlation with spatial frequency (c/deg). Furthermore, this shows four different contrast gratings, illustrating the combination of contrast and spatial frequency at certain points (Reddy, 2001)

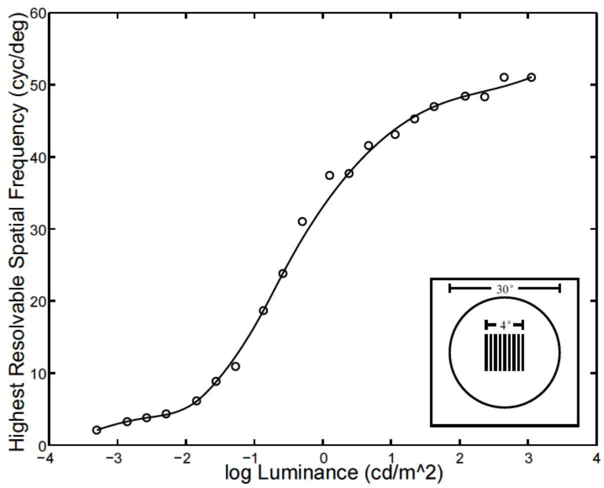


Figure 2: Luminance level correlated to the highest resolvable spatial frequency (Shaler, 1937)

Another important aspect of the human visual system is how the rods and cones behave during shifts in lighting conditions. Rods and cones are sensitive at different luminance levels, rods are responsive during low illumination, and they do not mediate color whereas cones are responsive during high illumination levels and mediate colors (Wolfe, 2009). This is also referred to as scotopic and photopic vision. This means that during photopic vision the eye perceive sharp detail and colors, while during scotopic vision no colors can be mediated and the contrast sensitivity is low (Ferwerda, 1996). Furthermore, the lower amount of luminance the lower acuity is expected. Figure 2, shows the relationship between luminance and the spatial frequency in grating patterns.

The above research has established that the human visual system and its ability to perceive detail vary depending on lighting conditions, velocity and where the object is placed in relation to our vision.

In the next section detail in relation to computer generated imagery will be researched in order to see how one can vary the level of detail in an object in relation to the HVS and its ability to detect it.

3.1.3 Detail

Often there is a tendency to generate much more detail in computer graphics than the user can perceive (Reddy, 1997). These details can take time to create and rendering power to produce. Therefore today's researchers (Reddy, 2001, Reddy, 1997, Ferwerda, 1996, Ferwerda, 1997, Hasic et al., 2010) often taken the human visual system into consideration when creating computer graphics, due to the fact there is no need to spend time on details that might go unnoticed by the user. The human visual system does not perceive all the detail in its environment and there will therefore always be a distinction in displayed and perceived detail. Since human performance only depends on perceived detail, this distinction represents wasted computation (Watson et al., 1997). If the human eye is not able to perceive all detail in its environment it must therefore mean that it is possible to create different level of detail for different objects, which would also reduce the computation cost.

The most common and maybe simplest way of varying the level of detail is by distance (Luebke, 2003), the further away and object is the less detail the human eye is able to detect. Velocity of an object can also result in different level of detail. An object moving quickly across the retina can be simplified, compared to a slow moving or still object. (Luebke et al., 2003). However, if the user attention is focused on a moving object across a stationary background, the object should be rendered in high detail, whereas the background can be simplified (Luebke et al., 2003).

Rendering works by taking a section and analyzing it, combining the information that was found in the shading group (geometry, lights and shading network) with the render settings information. As the renderer moves on to the next section, it again analyzes the situation. The more complex the model is, the longer the rendering will take (Guindon, 2009).

Methods on how to vary the complexity of a polygon based model and thus create various level of detail for an object has been suggested by David Luebke et al. (2003). The first method is polygon reduction which can be used to generate a model with the same general shape and genus as the original model, but having fewer polygons. Seeing an additional feature of this experiment is to reduce the render time, the first thing to focus on in the different level of detail should be polygon count (Luebke et al., 2003).

The reasoning for lowering the poly count in the low detailed model in this experiment is to investigate if it is possible to use a lower poly model, reduce the rendering time, without the viewer noticing the difference in detail between the models. This should be possible due the moving object in the foreground which should attract the viewer's attention, and thus the viewer will not notice the difference in detail.

Another way to add detail and still keep the polygon count low is by using texture mapping. Instead of modeling e.g. a wooden box in full detail, with a high poly count, one can portray this wooden box by using texture. Texture can be used to create the illusion of holes and cracks, instead of increasing the polygon count by modeling it. The texture can be further supported by using normal maps or bump maps, to heighten the illusion of holes, bumps and cracks (Guindon, 2009).

The illumination model used on an object can also influence the render time (Luebke et al., 2003). Using a Blinn can result in a more realistic render due to the reflection that can be produced, but it will also take longer time to render than e.g. a Lambert (Guindon, 2009).

Due to the limitations of the human visual system perceived detail in an object can vary depending on different factors such as an objects distance from a viewpoint, its velocity, and its position in the user's peripheral fields (Funkhouser, 1993, Luebke et al., 2003). Level of detail can be achieved in various manners, such as polygon count, texture mapping, and by using different illumination models.

On the basis of the research presented, the level of detail in an object can relate to *how much detail the human eye is able to detect and thus either lower of heighten the complexity of said object*. Level of detail in relation to this experiment will also refer to how much the rendering time can be reduced.

Further elaboration on how level of detail will be regulated in this project can be found in *Level of Detail Variable.* In the next section *believability* will be investigated.

3.1.4 Believability

In order to create a believable character it is not necessary to create a realistic one or a replica of something *real*. Believable does not mean an honest or reliable character, nor does it require human form in order to be believable, but one that provides the illusion of life and thus permits the audience's suspension of disbelief[1] (Bates, 1994, Livingstone 2006).

Lee et al. (2008) found, in their analysis of character believability, 5 qualities that would contribute to overall general believability: appearance, personality, intentionality, emotions, and social relations. (Lee et al, 2008). The character appearance describes all the things that are exposed to the viewer, such as gender, age, height, ethnicity etc. However, the characters appearance must do something more than simply reveal demographic information, it should also indicate other qualities such as emotion and personality (Lee et al, 2008). Though Lee et al.(2008) also state that character believability can be established mainly by one or two of the above describe believability qualities (Lee et al, 2008).

Loyall (1997) states, that one of the most important requirements for a character to be believable is personality. He further states that if to create a character that are to be believable, they must appear to have emotions and expressions of those emotions that are true to their personalities. Conveying emotion in a character has the primary means to achieve believability (Loyall, 1997). Thomas and Johnson (1981) said, "*From the earliest days, it has been the portrayal of emotions that has given the Disney characters the illusion of life*" (Thomas and Johnson, 1981, p 505). Furthermore, it helps the user to know that the character cares about what happens around them (Bates, 1994).

[1]Suspend disbelief, is when the user must be able to imagine that the world portrayed is real, without being jarred

The illusion of life or Lifelike is the key words that is often related to believability (Porter et al., 2000) and in order to create a believable character it does not necessarily mean to create a realistic one (Loyall, 1997), realistic meaning *a real* replica, but one that emits as lifelike. Lifelike does not mean have movement, but reflects intelligence, personality and emotion (Porter et al., 2000).

Another important quality for believable characters character intentionality, or that the character appears to be goal driven. It means that a believable character must act as if it has the capability to "think" and change what they are doing in response to the unfolding situation (Loyall, 1997, Lee et al., 2008, Livingstone, 2006). Livingstone (2006) gives an example in which behavior of an agent can cause *suspension of disbelief* to disappear, because the behavior is unrealistic. Two enemy guards are patrolling, after blowing up one of the guards, the other guards should naturally react to the situation, however instead of reacting he fail to pick up any of the obvious clues and just slanders by (Livingstone, 2006).

The last quality is social relations. By having a character appear to have social relation with other characters help in the viewers willingly suspend their disbelief (Lee et al., 2008) and can help in making a character seem alive (Loyall, 1997).

A good example of believable characters is derived from the first full-length computer generated animated feature film Toy Story (Porter et al., 2000). Woody's personality flows through his actions as well as his reactions. His motions, postures, movement of his brows, mouth and eyes all convey emotions, when woody respond to situations throughout the movie. The audience is convinced that Woody is lifelike, making his character believable by adding emotions, reaction and personality (Porter et al., 2000).

Another example could be Bambi (1942), the representation of the deer Bambi is no near what we know from a real dear, however Tomas and Johnson (1981) note that "*If we had drawn real deer in Bambi there would have been so little acting potential that no one would have believed the deer really existed as characters. But because we drew what people imagine a deer look like, with a personality to match, the audience accepted our drawing as being completely real" (*Thomas and Johnston, 1981, p. 332)

In I Robot (2004), this is another example with believable characters, Sonny express personality, emotion and reacts to what is happening around him. What separate I Robot from the two other examples, are the fact that CGI has been inserted into real life footage, which can require precise creation and manipulation of multiple 3D cues, such as light, color, texture, movement and sound (Prince, 1996).

Rademacher (2001) support this in his study where he explored the different aspect of CGI rendering in the context of perception of realism. He found that elements such as shadow softness and surface roughness are important factors in relation to perceived realism of CGI (Rademacher, 2001). However Porter et al (2000), points out that pure physical realism does not ensure a character's or a scene's believability (Porter et al., 2000).

It is no secret that developers of computer generated imagery, whether it is within the movie or game industry, keep aiming for it to look more and more realistic. Ferwerda (2003) describes three varieties of realism in computer graphics, being physical, photo and functional realism.

Physical realism aims to provide the same visual stimulation as the scene, which means that a model much contain accurate descriptions of the shapes, materials and illumination properties of the scene. "Scene" referrers to the "real world" and this type of realism if what you often see in the manufacturing industry. (Ferwerda, 2003).

Another type of realism is Photorealism which aims "to create an image that is indistinguishable from a photograph of a scene", and requires for the image to produce the same visual response as live action footage would do. (Ferwerda, 2003).

The last definition of realism is functional realism and according to Ferwerda (2003) this aims at providing the same visual information as the scene would (Ferwerda, 2003). In relation to CGI, the three definitions of realism can have different purposes of usage because in the entertainment industry the goal is usually not the same as it would be in an industrial usage.

In entertainment, the picture does not necessarily have to be a physical correct representation as long as the user believes it. Realism in entertainment tends to include referentially fictional representations of objects, creatures etc., which Prince (1996) refers to as "unreal" images (Prince, 1996).

Prince (1996) argues that we can perceive something *unreal* as being real in a film context, to which he uses Steven Spielberg's *Jurassic Park* as an example. The dinosaurs have no basis in any photographic reality, but never the less seemed realistic. However, since no one really knows what these creatures looked like, deciding whether or not these creatures are realistic is up to the viewer. Though, if looking at the definitions regarding realism, made by Ferwerda (2003), photorealism refers to re-creating an image that is indistinguishable from a photograph, and due to the fact that no-one have ever photographed a dinosaur it cannot be re-created in a photorealistic manner. It is hard to judge something as being realistic, when you do not know it's true nature and on these grounds I would therefore argue that evaluating whether or not *Jurassic Park* is realistic becomes a matter of perceived believability within the movies own reality. Prince (1996) points out that the dinosaurs made such a huge impact on the viewer because they seemed far more lifelike than the miniature models and stop-motion models of previous generation of film (Prince, 1996), which is one of the things that define a believable character.

This then means that something "unrealistic" or "unreal" can still be perceived as believable, because all of the action taking place within the movie still holds true to the narratives own fictional world. On the other hand, this also means that believability can be affected by realism, if the audience expects it to be realistic and it fails to meet these expectations it will thus not be believable. However, something can still be believable without it being realistic if the narrative and the content of the movie do not aim for realism, and the audience is aware, they should still accept the movies own realism and thus it should be believable.

Let's take a look at some examples, in the movie Space Jam (1996) Michael Jordan agrees to help the Looney Toons play a basketball game vs. alien slavers to determine their freedom.

This is of course far from realistic, but because the narrative is laying the ground for the movie not to be realistic the audience accepts the stories own realism and thus it will be believable.

Another example is a newer movie, Life of Pi (2013), where Pi survives a disaster at sea and goes through a journey of adventure. Here is would make no sense that roger rabbit all of a sudden showed up, both because the story is aiming for a more realistic narration, but also because the CGI in the movie is aiming for photorealism. If roger rabbit were to appear, it would break the photorealistic film universe and thus make the story perceived as not believable.

When watching a movie, the audience are introduced to the narrative, the characters, and the environment and should see these as being semantic consistent. However, if seeing just a photograph of the Looney Toons at the backetball court, would the audience still see these elements as semantic consistent? In the next section, semantic consistency will be investigated to see if it can have an effect on the perceived believability.

3.1.4.1 Semantic Consistency

Davenport et al. (2004) have conducted different studies within semantic consistency and given evidence towards that objects and their settings are processed together, and further that foreground objects provide context for the background and the background provides the context for the foreground object. These studies suggest that semantics play a large role in how a scene is perceived and how easily objects can be identified and named (Davenport et al., 2004, Davenport 2007, Henderson et al., 1999, Boyce et al., 1992, Palmer, 1975). In the studies done by Davenport et al. (2004, 2007) the participants were exposed to different pictures consisting of a background scene and a single foreground objects that was either semantically consistent or inconsistent with the scene. In other words, the pictures could e.g. contain a football player on a football field, and thus be semantic consistent, or the football player would be placed inside a church, and thus according to Davenport et al. (2004) be inconsistent. The assignment was then to name the foreground object, the background, or both and results showed that reported objects were named more frequently when they appeared with a consistent background.

Glancing at the pictures, without any background story to why a football player would be inside a church or why the pope is standing at the football fields, they seem to be semantic inconsistent. However, could a story leading up to the picture make the background and foreground object appear to be semantic consistent?

This question is also interesting in relation to believability, it may not seem believable that the pope is standing on a football field if we are only presented with the picture, however, if there had been a story leading up to the picture it may be perceived as believable and therefore semantic consistent. This should be taken into consideration when conducting the experiment, should there be a story leading up to the clip or should the participant just view the clips without a background story.

3.1.4.2 Uncanny Valley

In *Believability* it was noted that a believable character does not necessarily needs to be human, but should emit lifelike behavior in relation to one of the 5 qualifiers: appearance, personality, intentionality, emotions, and social relations. One reason to avoid using CGI humanoids in this experiment is to avoid the Uncanny Valley.

Advance in technology have enabled increased visual realism in computer graphics and as characters approach high levels of human-likeness, their appearance and behavior are more likely to be carefully examined by the audience. A facial expression that would otherwise appear as like-like can appear off and unnatural and thus make a character appear life-less (Tinwell et al, 2011.). Masahiro Mori observed that as robot's appearance became more human-like, a robot continued to be perceived as more familiar and likeable to a viewer, until a certain point was reached, he called this the "Uncanny Valley" (Tinwell et al, 2011.). Figur 3 shows the "Uncanny Valley". If creating a CGI humanoid and falling into the uncanny valley this could also affect the believability of the scene, because the viewer may rate the CGI as being creepy, strange, or out of place and thus no longer fit into the content of the scene. Furthermore, the viewer may have more expectation to CGI humanoids than non-humanoids in terms of looks, because we know how humans look and act, and if failing to meet these expectations, it may also affect the perceived believability of the scene. Therefore it has been decided, that not to create CGI humanoids to use in this experiment.

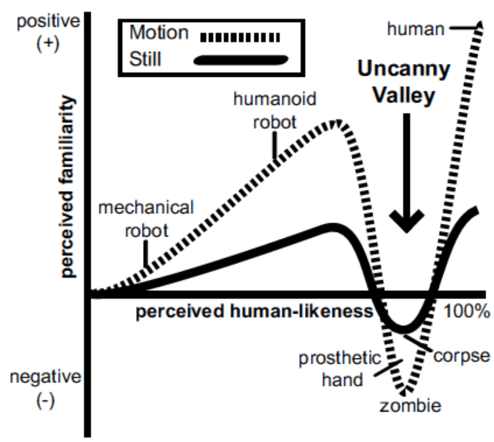


Figure 3: The Uncanny Valley (Tinwell, 2011)

3.2 Defining Believability

When creating a CGI character, it was found in *Believability* that a character should appear lifelike, in terms of one of the five qualities, which can contribute to the overall general believability. Furthermore it was argued that the narrative can affect the perceived believability of the CGI and scene content. The character's appearance can be supported by the narration and scenes content and thus contribute to the character appear believable within the scene.

Therefore a definition on believability in relation to a computer generated character will be defined as:

"A computer generated character will be perceived as believable if it appears lifelike and if there is a correlation between said element, scene content and the narrative."

Furthermore implementing CGI in real footage requires manipulation of different cues in order to make the CGI fit visually into the scene and match the remainder elements in the scene. In other words, the relationship between shading, texture, proportions and so on needs to correlate with the remaining elements.

"A computer generated element will be perceived as believable if said element correlates with the other non-CGI elements in the scene."

This definition is based on the research found in *Believability*. When using CGI in a scene with other elements, the CGI needs to correlate with the remaining elements in order to make the CGI be perceived as believable. In other words, the relationship between shading, texture, proportions and so on needs to be in agreement with the remaining elements.

How measuring believability has been done prior to this experiment will be researched in the *Analysis* and how it will be done in this project will also be elaborated.

3.3 Preliminary Analysis Summary

Computer graphics is often used in movies however, there is also a tendency to generate more detail than the user is able to perceive and can therefore go unnoticed by the viewer and thus be a waste of time. In this report I want to investigate if there is a way to reduce the time used on these different elements in a scene and thus reduce the time spent on creating detail but also the time it takes to render. The human visual system has over time become a familiar area of interest when it comes to optimization in relation to computer graphics. There are several reasons for taken the human visual system into consideration when creating computer graphics, because even though the HVS is good it is not perfect and has its limitation when it comes to perceiving detail. The eyes ability to perceive detail is measured in cycle/degree (cy/de) in grating patterns, and the further away from the fovea, which is the area in the eye that have 100% visual acuity, the lower detail sensitivity the eye has. Distance is also a factor that can affect the eyes ability to perceive detail because the further away and object is the less detail the human eye is able to detect. Velocity of an object can result in different level of detail. If the viewer's attention is focused on the moving object the background can be simplified, and vice versa. Luebke et al. (2003) describes how one can vary the complexity of an object, and thus create different level of detail. Polygon reduction, texture mapping and using different illumination models is a part of the variation.

In order to utilize visual attention to reduce the level of detail, research had to be done in order to see which visual attention methods should be used in relation to this experiment. Visual attention is a process in the brain that involves conscious and unconscious actions and the two processes is called top-down and bottom-up, and determine where humans locate their vision. I decided that I wanted to use bottom-up, which is automatic and stimulus driven. Furthermore bottom-up can be affected by salient visual features and will result in involuntarily shifts in attention. The saliency of an object can be influenced by different visual inputs like color, size, luminance and movement. Based on the researching done in *Visual Attention* I have decided to utilize *movement* as my saliency feature to catch the viewer's attention in a scene and thus redirect their vision away from other objects in the scene.

It was found that believability of a character was to give the character a life-like behavior, and is based on 5 qualifiers, which can contribute to the overall believability of the character. It was found that a CGI character does not necessarily needs to be humanoid, and it was decided to avoid creating CGI humanoids due to the uncanny valley and how this could affect the perceived believability. Another aspect taking into consideration was based on research found in S*emantic Consistency,* here it was noted that having a short narrative presented to the viewer *before* the experiment may affect them to more frequently find the CGI and scene semantic consistent and thus perceive it as believable, than if not having the short narrative presented to them.

How level of detail is varied and how believability is measured is relation to this project will be further elaborated in the next chapter Analysis.

3.3.1 Final Problem Statement

Based on the research found in the *Preliminary Analysis* the initial problem statement can be narrowed down to the following

To what degree can movement in a foreground object be utilized in a scene in order to alter the level of detail in a background object without any noticeable change and still maintain the perception of believability?

3.3.2 Criteria for Success

- The foreground object will convey the "illusion of life", in form of one or more of the 5 qualifiers mentioned
- The moving object attracts the viewers attention
- The viewer is not able to detect the differences in level of detail

The foremost concern of this experiment is that the viewer has to perceive the CGI and the scene *together* as believable. In order to deem the experiment successful the viewer should not have noticed the difference in level of detail due to the moving object in the foreground and thus still maintain the perceived believability of the scene. In order to make the foreground object be perceived as believable as well it has to convey the "illusion of life", meaning it has to act as if it has e.g. a brain, intentionality, emotions etc.

4 ANALYSIS

Level of detail and believability has been defined, furthermore it was decided that visual attention should be in form of movement. In the next section theories regarding movie genres which will lead to the choice of CGI, based on the genre chosen so that there is consistency between genre and CGI. Furthermore research regarding *how* to measure believability will be investigated.

4.1 Movie Theory

The scene and CGI used for this experiment could be kept as simple as possible and perhaps be something like the experiment that Hasic et al. (2010) did, having a ball roll across the screen, though implement into real footage and not a computer generated background. However, due to personal interests there will be put some time and effort into the CGI, this may result in the CGI not being "simple".

In the following section I will investigate some of the different movie genres to get an understanding of which elements they may contain to avoid semantic inconsistency. Once the type of genre has been chosen, a decision regarding what type of CGI used for this experiment will be determined and what methods that can be utilized in the creation of the CGI will be elaborated.

Furthermore sound is looked at. Though this is not the main element of my research I do feel that it is an important element in the overall impression of the scene.

4.1.1 Choice of Genre

The use of CGI cannot be pinned down to one specific genre of movies and it is therefore necessary to determine which type of genre that has to be used for this project.

Fowkes (2010) note that people may expectations to the different genres and in order to meet these expectations and not have it affect the believability the elements used in the genre have to be in agreement with the genre and what is expected.

When encountering wizards, crystal balls, flying brooms, fairies, magic talismans, or talking animals we are likely to think fantasy unless otherwise informed (Fowkes, 2010). The core of fantasy is magic, physical transformation (vampire, werewolves) or the ability to fly like Peter Pan (Fowkes ,2010). Fantasy can be about escaping into other worlds, like Alice in Wonderland (2010), Peter Pan (2003) or The Chronicles of Narnia (2005). Furthermore movies such as X-men (2000) or I am Number 4 (2011) can also be seen as fantasy. In X-men mutants have the ability to fly, being shape shifters or have the power manipulate fire or ice, the ladder also goes for I am number 4 (2011).

Science Fiction are stories that rise from rational and scientific principles and we expect a certain iconography like spaceships, robots, advanced technology, alien encounters etc. (Fowkes, 2010, Telotte, 2001). Science fiction has an ability to produce what might be, a synthesized new reality (Telotte, 2001), and deals with man's ability to survive any confrontation with the forces of a hostile threats posed by technological and scientific advancement (William, 1927). This is seen in movies such as War of the Worlds (2005) and

Battleship (2012), where unknown forces are invading earth, leaving the humans to fight for survival.

The core of horror is the “monster”, the predator of the dread and fear that causes an emotional response to the movie (Odell et al., 2001). The way that horror distinguishes from science fiction or fantasy is that it attempts to scare us, but also make use of certain elements, such as dark and stormy nights, monsters, vampires, ghosts ect. (Fowkes, 2010). I can already now say that horror will not be the chosen genre as I am not attempting to scare people.

When choosing a genre and appertaining CGI I have to look at the time I have at my disposal. Even though fantasy could be interesting, some of the aspects within the genre can be time consuming such as creating two worlds or physical transformation of a character.

An interesting aspect of science fiction is that it can be presented from either the human perspective or *the enemy* perspective to which the ladder leaves room for more creativity in relation to CGI. If presenting it from human perspective it could be in form of Tanks, Helicopters of other machinery that is created by humans in order to avoid humanoids (red. The Uncanny Valley). Science fiction may be “easier” to create and make believable since earth invasion is a familiar topic in movies.

Based on the discussion the choice of genre will be science fiction. Whether to portray the genre from the human or enemy perspective, and thus also choosing of CGI, will be presented in *Choice of CGI.*

4.1.2 Sound

Our perception of an environment is not only what we see, but can be influenced by other sensory input (Chalmers et al, 2008.). One input could be sound, which Bordwell et al. (2010) notes as a powerful film technique for different reasons. First off, sound can engage a distinct mood in a scene. Furthermore, sound can direct viewers’ attention within an image or help guide their attention. Sound effects supply an overall sense of a realistic environment and are seldom noticed, however, if they were missing the silence would be distracting (Bordwell et al., 2010).

Sound can affect visual perception under certain circumstances. In a study, subjects were found to incorrectly count a small number of visual stimuli when accompanied by a short beep (Shams et al., 2000). This implicate that sound feedback is important for improved perception of virtual events (Astheimer, 1993).

There are two different types of sound, diegetic sound and nondiegetic sound. Diegetic sound is sound that has a source in the story world, e.g. sound made by objects n the story, whereas nondiegetic sound is presented as sound coming from outside the story world like music added to enhance the film’s action.

Though sound is not the main focus of this project, sound could help direct the viewer’s attention towards the main element in the scene. (Bordwell et al., 2010) Therefore sound will be added to the shot in order to help guide the viewer’s attention. In the next section, previous research on how to measure believability will be looked into.

4.2 Previous Research on Measuring Believability

Judging whether or not a scene with computer graphics is *believable* is highly subjective (Mac Namee, 2004), meaning it depends on the viewer who are watching the movie or playing the game to judge whether or not what is seen is believable.

Previous attempt to measure believability has been done before within the field of e.g. AI behavior. Mac Namee (2004) measures believability by presenting two versions of an AI implementation and simply asking the user which of the two was *more* believable followed up by asking which differences, if any, were noticed by the viewer. In order to further examine the reasoning behind the participants answer to the first statement, which one they found most believable, he presents three statements, regarding believability, to which the viewer could answer on a response rating ranging from *strongly disagree* – through *disagree, no opinion, agree* – to *strongly agree* (Mac Namee, 2004). The three statements that Mac Namee (2004) presented to the user dealt with the AI characters intentionality, which refers to the characters choice of actions, intentions and behavior. As mentioned in *Believability* intentionality was one of qualifiers that would allow a character to achieve the "illusion of life".

An important observation in the study by Mac Namee (2004), was that the test participants were both novices and veterans. To the novices, the whole experience of playing a game is so new that they fail to notice the significant difference between the two versions of AI, to which this difference is apparent to most veteran. There may be several reasons why the novice user fails to notice the differences however one important reason is the knowledge they have to the performance medium being observed (Magerko, 2007). This is also something to keep in mind when choosing the test participants. Maybe some of the participants are not as familiar with the science fiction genre, as oppose to others, and will find it more believable than a participant that are highly interested in science fiction.

Magerko (2007) argues that believability of AI consist of what is expected of the agent, and that the agents performance is evaluated through a subjective lens of an individual observer's expectations of that performance (Magerko, 2007). An important aspect of Magerko's (2007) research was that the viewer should have a comparison measure, e.g. in relation to the AI characters: How lifelike is the model and how natural does the model interact with the environment (Magerko, 2007). This relates somewhat to what Mac Namee (2004) did in the second part of his test, having statements relating to believability that the participants could rate.

In relation to the above it is worth mentioning that the characters in the above was limited to humanoids, but as Livingstone (2006) points out the character does not need to be human but could in fact be any creature as long as they demonstrate life like behavior (Livingstone, 2006).

How both Mac Namee (2004) and Magerko (2007) is measuring believability is what Togelius et al. (2012) refers to as *subjective believability assessment*. According to Togelius et al. (2012) this measurement can be based on either participant's free response or on forced data retrieved through questionnaires. Having the participants report believability through a questionnaire constrains them into a specific area and yields data that can easily be used for analysis (Togelius et al, 2012). Togelius et al. (2012) suggest three types of forced questionnaires for believability assessment:

1. Boolean: participants have a single Boolean answer choice
   - *"Is this believable?"*
2. Ranking: participants are asked to answer questionnaire items given in a ranking/scaling form
   - *"How believable was that on a scale of 1-5?"*
3. Preference: participants are asked to compare the believability of two or more sections of e.g. a game or movie
   - *"Which one was more believable?"*

Togelius et al. (2012) notes that the main disadvantage of ranking questionnaires is that they do not control the subjective notion of believability across participants. This means that participants can rate the statements presented differently, because they have a diverse interpretation of the statements.

From the above there is an indication that measuring believability is a subjective manner and in order to measure believability some comparison measure may be needed. It is important to note that the viewer may have expectations and if these expectations are not me the perceived believability may be affected. How believability will be measured in relation to the experiment performed in this report will be elaborated in *Measuring Believability Variable.*

4.3 Choice of CGI

The reason for using CGI is the fact recreating the exact same scene with a dynamic object and a still object would be difficult if the camera is moving, and seeing as the scene has to be identical, the optimal choice is to use a CGI object. Due to having eliminated humanoids, it was decided to create some sort of mechanic means of transportation, and though it may not be a perceived as having emotions or a personality, it can still portray determination in form of having intentions or goals.

Another reason for creating a computer generated tank is that it may be less time consuming to animate, it does have less mobility possibilities than say a robot having two legs and two arms, but it does still have movement.

The risk of creating a tank lies in the fact that a tank is created by humans and that most people have an idea of what a tank looks like. However, there are many different types of tanks and further the appearance of a tank seems to have great variation.

The possibility of creating a tank that is more of a science fiction tank is of course there too seeing as the genre of the narration is science fiction. This gives room for some creativity, but may also affect the user's view on believability seeing as the tank may be unfamiliar in appearance but familiar in its usage. The appearance of the tank will be decided later and sketches will be created in order to have a point of origin when starting to model.

Based on the choice of CGI, the narrative will be as following.

A tank is driving on a routine drive by, while coming across an unidentified object. The soldiers make a call to the base to their commanding officer, who orders them to carefully close in and check it out. In order to relate to the Science Fiction genre, the unidentified background object will be in the form of a UFO.

4.3.1 Render Engine

The CGI will be rendered using mental ray, because it allows for features that the other render engine do not support, e.g. global illumination, HDRI and further mental ray offers better control over the quality level of the objects being rendered, by e.g. adjusting the number of samples which set the number of times mental ray samples a pixel to determine how best to anti-alias the result (Derakhshani, 2011).

4.4 Analysis Summary

In the *Analysis* different movie genres was looked at in order to meet the expectations the viewer may have regarding the different genres and to avoid semantic inconsistency. Furthermore, it was decided to implement sound, because it can help guide the viewer's attention.

Previous research on how to measure believability was investigated and it was found that *believability* is a subjective manner and in order to measure believability there should be a comparison measure, e.g. how natural does the model interact with the environment. These different measures can then be used to *measure* the perceived believability. However, it was also found that the viewer may not have the same definition of believability as the one created in relation to this experiment and furthermore, the viewers' participating could rate the different comparison measures, which in this experiment will be created as statements to rate, differently because they have a diverse interpretation of them.

In *Choice of CGI* it was decided to create a CGI tank as the foreground object, due to avoiding humanoids which was based on the research found in *The Uncanny Valley*. Furthermore it was decided that the background object should be a UFO, which related to the science fiction genre. The CGI will be rendered using mental ray, because it allows for features that the other rendering engines do not support.

5 METHOD

In this section the method used to test the experiment will be presented.

5.1 Variables

The three terms used in this project, being visual movement, level of detail and believability has been defined in relation to this experiment. In the next section these terms will be described as the three variables, two of which are independent variables and the ladder is a dependant variable.

5.1.1 Visual Movement Variable

Visual movement if one of the independent variables used for this experiment, and in order to test whether or not motion in an object will attract the viewer's attention, it will have to be compared to a sequence containing no motion. Therefore the visual movement is divided into two different conditions:

- Movement
- No movement

This means that there will be a sequence containing an object with motion and one with the CGI presented but having no movement. These two conditions will have to be measured on in proportion to level of detail.

5.1.2 Level of Detail Variable

The level of detail works as the other independent variable that may or may not affect the believability of the scene. In relation to the theories presented throughout the report, it should be possible to implement a low detailed model in the background without the viewer noticing the difference due to movement in the foreground object.

The different level of detail will be primary controlled by the render settings in mental ray, however there are some parameters that cannot be controlled by *only* tuning the render settings. For the high level of detail a higher polygon count will be used, while for the low detailed model there will be a reduction in the polygon count and there will only be applied a Blinn Material to the object.

Mental ray, as mentioned in *Render Engine,* has a number of different render options compared to the other render engines available when rendering in Maya. This means that for the high level of detail the different parameters such as global illumination, HDRI, ambient occlusion and etc. will be used, though for the low detailed model neither of these parameters will be used. The reason for only having two LoD condition is based on the hypothesis that the participants will not notice the difference in level of detail while movement in the foreground is presents and therefore there is no reason to have a "medium" level of detail. Though they should notice the different, with no movement in the foreground, there is still no reason to have three different LOD conditions s because the "no movement" condition is used to validate the hypothesis made in relation to movement in the foreground.

This means that for this experiment there are two conditions of visual attention, each two levels of detail:

1. Movement
    a. High detail
        i. Higher polygon count
        ii. Ambient Occlusion
        iii. HDRI
    b. Low Detail
        i. Lower polygon count
        ii. A material applied only
2. No movement
    a. High detail
        i. Higher polygon count
        ii. Ambient Occlusion
        iii. HDRI
    b. Low Detail
        i. Lower polygon count, e.g. by polygon reduction
        ii. A material applied only

5.1.3 Measuring Believability Variable

In relation to the definition made earlier regarding believability, it was found that

"A computer generated character will be perceived as believable if it appears lifelike and if there is a correlation between said element, scene content and the narrative."

to which lifelike referred to 5 qualifiers being appearance, personality, intentionality, emotions, and social relations. It was also found that believability can be established mainly by one or two of the above describe believability qualifiers. Therefore it has been chosen, due to the choice of CGI used in this experiment, to focus on the intentionality qualifier to make the CGI appear as "lifelike".

Another definition was presented in *Defining Believability:*

"A computer generated element will be perceived as believable if said element correlates with the other non-CGI elements in the scene."

This definition refers to the fact that CGI objects and the non-CGI objects in a scene have to correlate with each other in form of shadow, movement, lights and so on. Therefore the statement presented to the participants will also have to relate to the correlation between scene and CGI objects.

Believability will in this experiment be measured as what the subject perceive as believable, and is the dependant variable. The dependant variable is the one that will be investigated whether or not it will be affected by the two independent variables, visual movement and level of detail.

Based on the section *Measuring Believability Previous Research* it was decided to present the measure comparison in form of statement presented on a questionnaire. The approach of ranking and preference will be used.

The subject will be given a short explanation on the procedure of the test, without indicating the goal of the experiment. Furthermore there will be a short narrative attached to the clip, which will help guide the participant. The reasoning behind having a narration to guide the viewer is based on the research found in *Semantic Consistency* and *Believabilit*y, which indicate that the narrative can affect the perceived believability. After viewing each clip the participant are required to fill out a questionnaire, which will give an indication on the participants view on the clip and the overall perceived believability.

The first thing after watching the clips the participants have to evaluate different statements present regarding believability. The statements will have to be evaluated by participants on a 6-point Likert scale, ranging from *strongly disagree* to *strongly agree*. The Likert scale will be presented after each question for the participant to rate his or her opinion relating to the given question. Having an even numbered Likert scale, and thereby not including a "neutral point", will force the participant to make a decision and not sitting on the fence. In this experiment 1-3 simulate that the participant are inclined to disagree while 4-6 will incline that the participant are agreeing, the further up or down the scale the more the participant will agree or disagree with the statement presented. Because the objects used in this experiment does not portray emotions or personality the statements will deal with intentionality and if the objects are consistent with the background/surroundings, much like the experiment presented by Davenport et al. (2004). The intentionality of the objects and the consistency in this experiment will be supported by the narrative, presented to the participants at the beginning of the experiment. Furthermore a statement concerning correlation between scene movements, shadowing, light etc. will also be presented to the participant. Furthermore a statement relating to the implemented sound will be presented. Together these statements will result in a believability response rating (**$\dot{B}$**), this is calculate by the total number of rating responses in each statement divided by the total number of respondents. This response rating will be calculated for each level of detail condition in each visual movement condition. The believability response rating will give an indication of what effect the LoD and visual movement had on the perceived believability. Table 1 show an example of what a believability response rating for each visual movement condition and level of detail could look like.

Believability response rating			
Visual Movement/LoD	Low Level of Detail	High Level of Detail	Average
Movement	3	4	3,5
No movement	3,33	4,5	3,67
Average	3,17	4,25	

Table 1: example of what a believability response rating for each visual movement condition and level of detail could look like

All of the following statement will have to be rated in relation to the Likert scale seen on figure 4. Table 2 shows the statements presented to the viewer, for a full questionnaire, please see Appendix.

Statement	Explanation
1. The content of the scene fits with the narrative told before seeing the clip	This question is related to the how well the content of the scene fit the narrative given to the participant, before seeing the clips.
2. There is connection between the foreground elements and background elements in the scene	This question is related to the scene consistency, meaning the participant are told to rate how well they feel like the background element and foreground element are corresponding.
3. The implemented audio seems to fit the scene content	Because sound is to be implemented, it would seem appropriate to ask the participant if the implemented sound actually fit the scene content.
4. The tank appears as if it was really there	This question relates to the fact that there has to be some sort of similarity between the CGI objects and the remainder of the scene, meaning how the shadows falls on the ground, the movement of the CGI compared to the scene etc.
5. The tank appears to respond to the unfolding situation relating to the unidentified object	This question related to the lifelike factor that was one of the main things needed in order for the CGI to appear believable, and as mentioned there are 5 qualifiers that can affect this “lifelike” factor to which believability can be established by mainly one or two of these qualifiers. Because a tank was chosen as the CGI, it would be difficult to create a tank expressing e.g. personality or emotions unless it was a cartoonish content it would have been created, therefore for this experiment, intentionality was the chosen qualifier
6. The unidentified object appears as if was really there	
7. The tank was perceived as believable	This statement will give an indication if the tank was perceived as believable or not.
8. The unidentified object was perceived as believable	Because the object are presented in a low and high LoD model, the participants are asked to rate the perceived believability of the UFO. This can give an indication of which model they found to most believable and if there is a correlation between the overall believability rating and the believability rating for the UFO.

Table 2: Statements and explenations

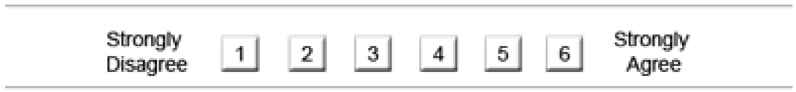


Figure 4: Likert scale used in this experiment

After watching both clips and answering both questionnaires, the participants will be asked which of the two clips he or she found most believable. This will be followed by asking the participant if they notice any difference between the clips, and to elaborate on their answer. This follow up question and elaboration is due to believability being subjective and in order to fully understand the reasoning behind the subject's choice and elaboration is needed. Furthermore this can help to establish if the participants noticed the difference in level of detail and if so if it may have had an effect of the perceived believability.

Because it was not possible to implement eye tracking, one way to see if the participants noticed the chance of the background object between the clips is to ask if they noticed any differences and if yes, they should elaborate on these chances. The last thing is a “further comments”, here the participant can explain what differences they notices and leave any other comments they must have.

The ideal scenario will be that in the two clips where the tank has no movement, the participants will notice the difference more frequently than in the clip where the tank is moving into the scene.

6 Design

This section of the report covers the planning of the workflow for this experiment.

6.1 Workflow

There are a few steps in the production of the different clips that has to be used in this experiment, and to organize the work a little a "workflow chart" has been created, see figure 5.

The first step in the production is to model the different 3D objects which will be created based on the sketches created, see *Sketching*, followed by texturing and perhaps rigging. The footage will be in accordance to the choice of genre, namely being science fiction, and should be shot a location that fit the narrative described shortly in *Choice of CGI*. The scene should be located away from any public places, due to the fact that is not a "professional" shoot which means stopping traffic or people from crossing paths with the shot can be difficult if shot in a public place. In order to make the matchmoving easier, zoom should be avoided, because zooming means that the focal length will chance. Therefore the camera movements will be limited to panning the camera, which refers to the rotation in a horizontal plane of a still camera.

Before being able to animate the CGI, the live footage has to be matchmoved. Once the animation is done, light maps in form of HDR mapping can be created, and thus render the final CGI to be used in the post-production. The post-production will be done in After Effects and Adobe Premier Pro and is the compositing of CGI into the live footage, editing the sequence in terms of e.g. color correction and grading and in the end adding sound.

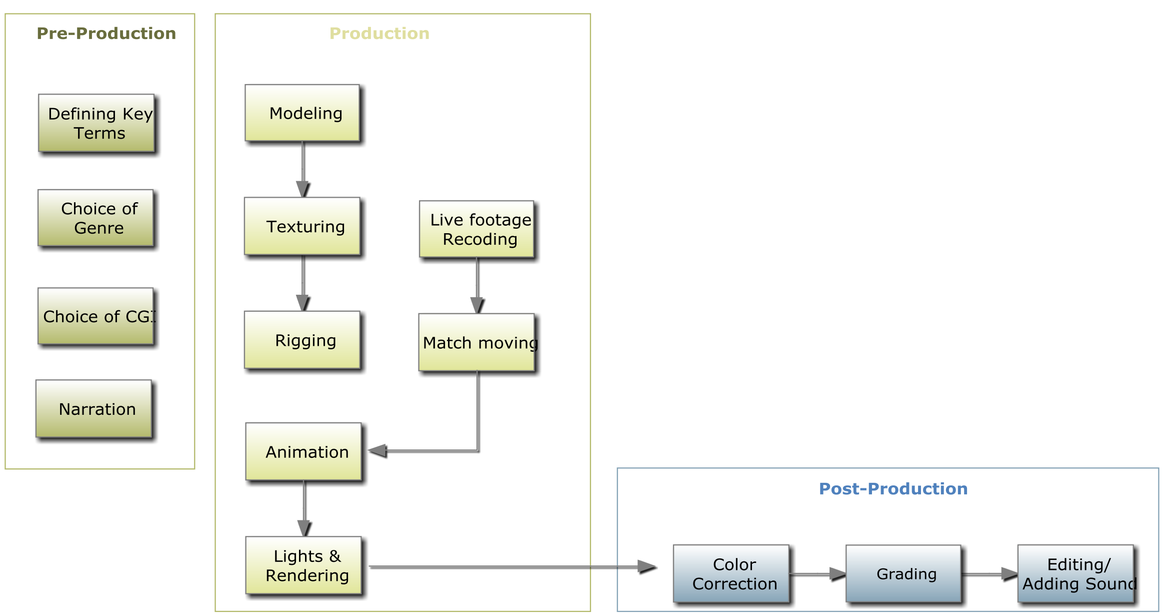
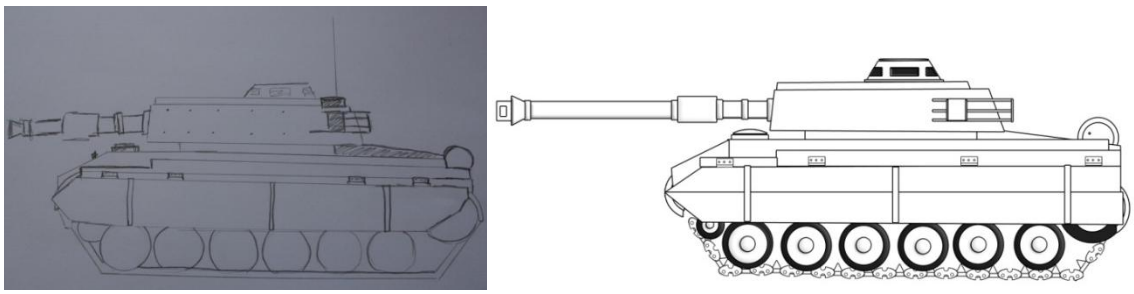


Figure 5: The Workflow followed in this experiment

6.2 Sketching

To make the modeling easier, sketches of the different views of the tank was created, being front, side and top view. These sketches were then loaded into Photoshop, where they were used as the base for creating an image plane that could be imported into Maya. These image planes will only be used as reference images, after having the models basis shape created, they will be removed. Picture 1 is an example of one of the sketches created for the tank and the image plane created on the basis of the sketch. The rest of the sketches can be seen in *Appendix*.



Picture 1: (Left) Sketch and (Right) Image plane created in Photoshop

6.3 Rendering

The rendering will be done using mental ray, because mental ray allows for e.g. global illumination and HDRI mapping.

Furthermore render layers will be used. By default a scene has at least one render layer, which have all lights and geometry included. Creating a new layer, gives the opportunity to specify which lights and objects are included in that layer. When rendering the different layers, it gives the ability to create alternative lights for each layer, use different shaders for each object, and render one layer using one render method and other layers using other render methods and so on (Lanier, 2008). This means that in one layer it is possible to have the *original* model with the *original* texture applied and in another layer the Ambient Occlusion can be applied, and then one can switch between the two different layers when having to render the different layers.

6.3.1 HDRI Mapping

The ideal would be to create a HDR image at shooting location, however, seeing as it requires equipment that I do not possess a "stand-in" HDR images will have to be used. The reasoning for using a HDR image is due to the fact that when combining CGI with real footage, it is important that the reflections on the CGI are in agreement with the surrounding scene. For this a HDR image can be used to create more realistic CGI. HDRI environment map can give an overall fill light for the scene, but also acted as a great reflection environment (Lanier, 2008).

7 IMPLEMENTATION

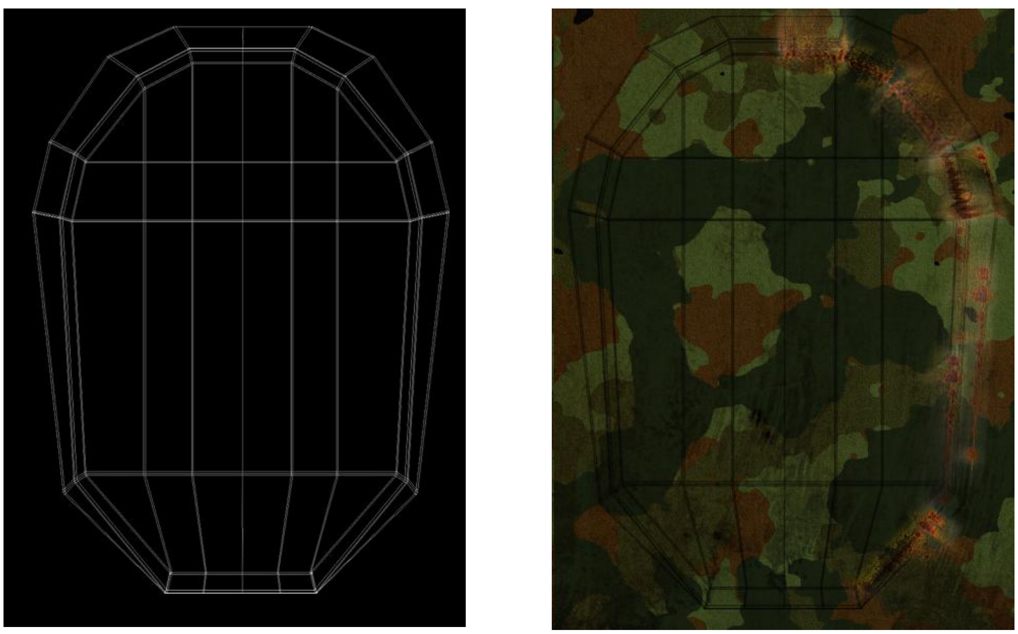
This section will cover the implementation of the different parts used for this experiment. It should be noted that is not a complete walkthrough, but will cover the basis and give an overall understanding of how the different 3D models, the recording, the matchmove and the compositing of the final footage was created.

7.1 Tank

Tank used for this experiment, as the moving object, was modeled in Autodesk Maya 2012 and based on the sketches found in *Sketching*. The sketches were used as Images Planes, to make the modeling a little easier. The model was created rather detailed, because it is the *main* object in the scene and the object that which should attract the viewer's attention. Picture 3 shows the model and the wireframe.

7.1.1 UV Mapping and Texturing

In order to apply a 2D texture to a 3D object, UV mapping of the different parts of the model is a process that has to be done. Depending on what parts of the model, different types of UV maps was applied. Picture 2 shows the UV mapping created for the top part of the tank and how the texture was created in Photoshop, based on the UV. This process was done for all of the different part of the tank.



Picture 2: to the left is the UV used for the top tank and to the right is the texture applied

7.1.2 Ambient Occlusion and Final Result

Another detail that was applied was ambient occlusion. Ambient occlusion works by the premise that when two objects or surfaces are close to each other, they reduce the amount of light at the intersection (Derakhshani, 2011). Picture 4 is an example of the ambient occlusion and picture 5 the model used in this experiment, with the ambient occlusion applied. The ambient occlusion was applied to the model in a separate layer in Maya.

Picture 3: The wireframe

Picture 4: Ambient Occlusion

Picture 5: The final model with ambient occlusion applied.

7.2 The Unidentified Object

The unidentified object was also modeled in Autodesk Maya 2012. The UFO had to be created in two versions, a low and a high level of detail model. Based on *Defining Level of Detail* the LoD was defined as the complexity of the model and could include heightening the number of polygons, the shading used, applying texture etc. Having this in mind, the high level of detail UFO was created with a higher polygon count than the low detailed model. Furthermore HDRI, Ambient Occlusion and Global illumination was applied.

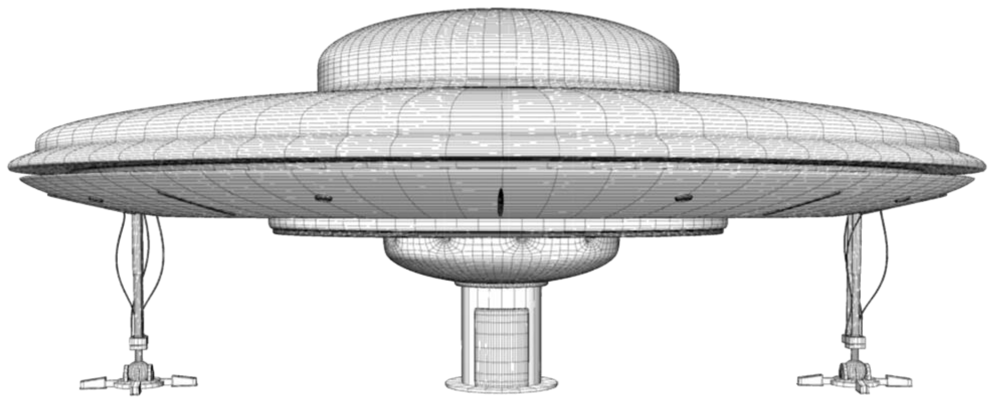
7.2.1 Implementing HDRI and Ambient Occlusion

Creating a HDRI, of the environment where the footage was recorded, requires equipment that was not obtainable therefore a substitute was used (Openfootage.net, 2010). The image, seen on picture 6, was used to create reflections in the high detailed model of "the environment". These reflection, when implemented in the footage, will help to make it look as if was really there. In order to see the reflections, from the HDRI properly, the model was applied a DGS material and rendered with Mental Ray. To further enhance the chances of the model being perceived as belonging to the scene Ambient Occlusion was applied. The final result can be seen in *Final Result*.

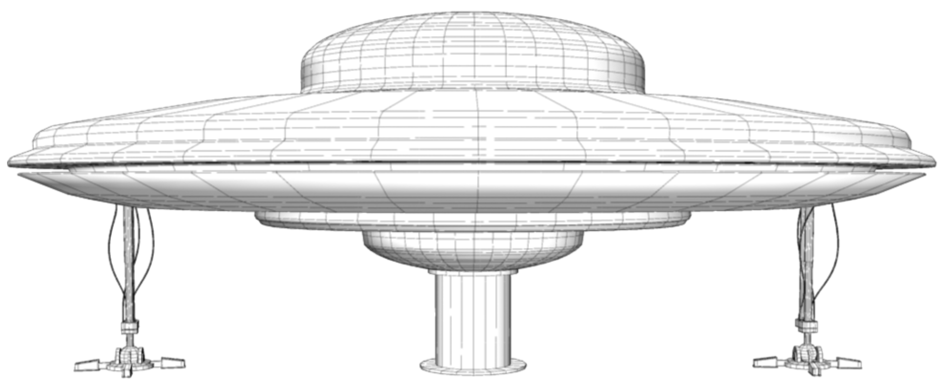
Picture 6: HDRI used (Openfootage.net, 2010)

7.2.2 Polygon reduction

As oppose to the high detailed model, which contained 60.388 faces, the low detailed model *only* consist of 12482, which is close to only 20% of the faces used for the high detail model. This reduction was obtained using polygon reduction. The reduction was done in order to make the low level of detail model lower poly and to make the rendering more efficient. Picture 7 is a render of the wireframe for the high LoD model and picture 8 is the wireframe for the low LoD. On the two pictures the difference in faces can be seen, meaning the low LoD model would be considered of lower polygon count than the high LoD.



Picture 7: Wireframe of the high LoD UFO model



Picture 8: Wireframe of the low LoD UFO model

7.2.3 Low LoD Material

For the low detail model a Blinn was applied, to still give a bit of reflection from the light sources placed in the scene. In order to create the same illusion that ambient occlusion would do, without using time on rendering a separate ambient occlusion layer, the material was darkened placed where there would otherwise be shadows. The difference between rendering an entire ambient occlusion sequence for the tank would be significantly different in rendering time compared to rendering the UFO with a darkened material.

7.2.4 Final Result

The *final* result of the low LoD UFO model can be seen on picture 9 and picture 10 shows the ambient occlusion and on picture 11 is the *final* UFO having the ambient occlusion and HDRI applied. There is a clear difference between the low LoD model and high LoD model. The high model had HDRI and ambient occlusion applied, which will result in different reflections and shadows, compared to the low LoD model. However, comparing the render time, the low LoD model sequence was significantly faster to render due to being less *complex*.

Picture 9: Final low LoD model

Picture 10: Ambient occlusion

Picture 11: Final high LoD model

7.3 The Footage

The footage used for the experiment was recorded using Canon 7D, with a EF-S 15-85 mm lens. Picture 12 shows the footage. In order to matchmove, some information regarding the recording device should be noted, which depending on the matchmove software used can vary from only needing the film back width to needing both the film back width and focal length. Providing this information to the software will result in a more correct match move.

On a 35 mm movie camera, the film back is 36x24 mm, which is also called full-frame. When shooting with the Canon 7D, the lens that is shot with will have to be multiplied with a focal length multiplier (FLM) in order for the field of view to have the same format as the referred format, namely the 35 mm (Multimedia Digest, 2012).

Picture 12: The footage

Because the sensor on the 7D is 22.3x14.9 mm, and not the "standard" 36x24 as with the 35 mm, the lens has to be multiplied with 1.6, in order to be equivalent to a 35 mm lens. For the Canon 7D, the local length will then, when using the focal length multiplier, be equal to 35x1.6 = 56, because the scene was recorded at 35 mm on the 15-85 mm lens (Multimedia Digest, 2012).

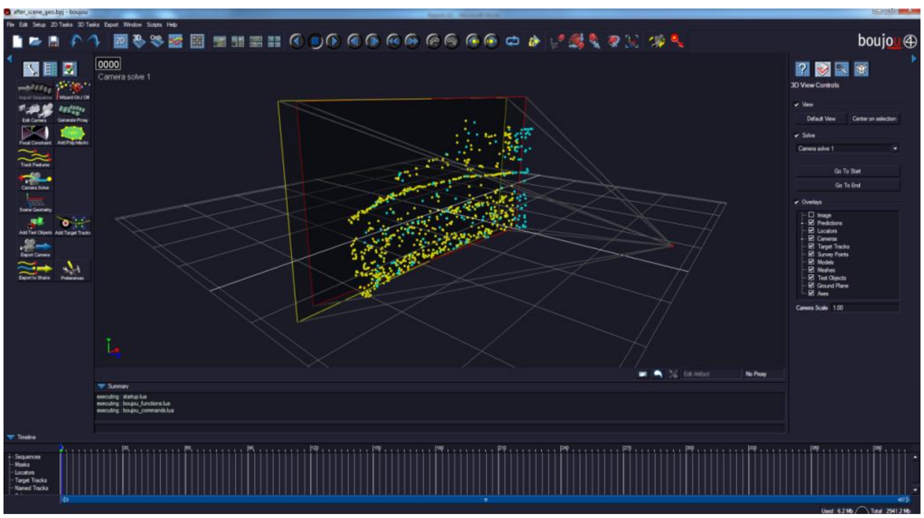
The film back width is calculated from the sensor size, being 22.3x14.9 mm and the capture aspect ratio, which is 1.7789 for the 16:9 format. This will results in a height of 22.3/1.7778 = 12.543 mm. (Multimedia Digest, 2012).

For production purposes the Canon 7D is actually not an optimal recording device, when wanting to matchmove the footage afterwards, because of its CMOS image sensor which causes rolling shutter effects. Rolling shutter means that the image is recorded not from a snapshot of a single point in time, but rather by scanning across the frame and can cause unwanted effects such as skewing of the image when moving the camera. Rolling shutter can therefore, due to these unwanted "effects" make it difficult to math move, but not impossible (The Foundry, 2013).

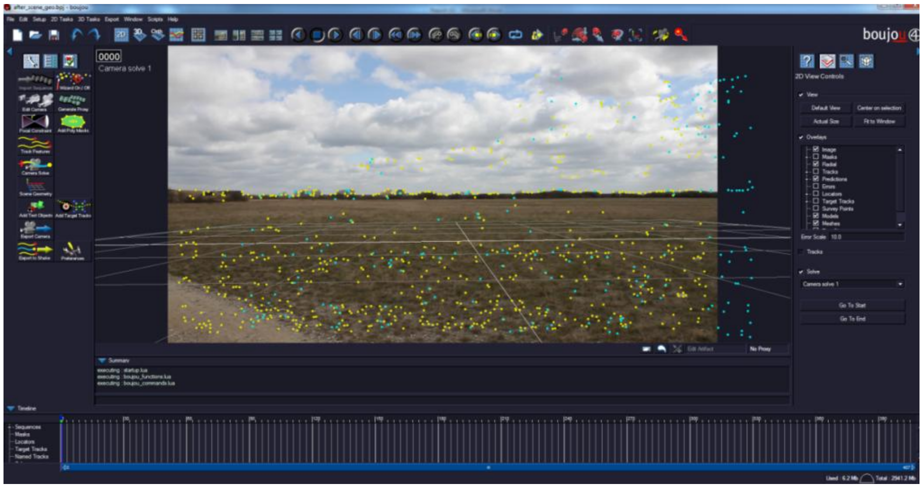
7.4 The Match move

The footage was matchmoved using Boujou 4.0, but before importing the footage it had to be exported as a PNG sequence, this was done in Adobe After Effects CS5. Information provided to the Boujou before starting tracking the shot, was film rate and the film back width which as mentioned earlier for the Canon 7D is 22.3x12.543.

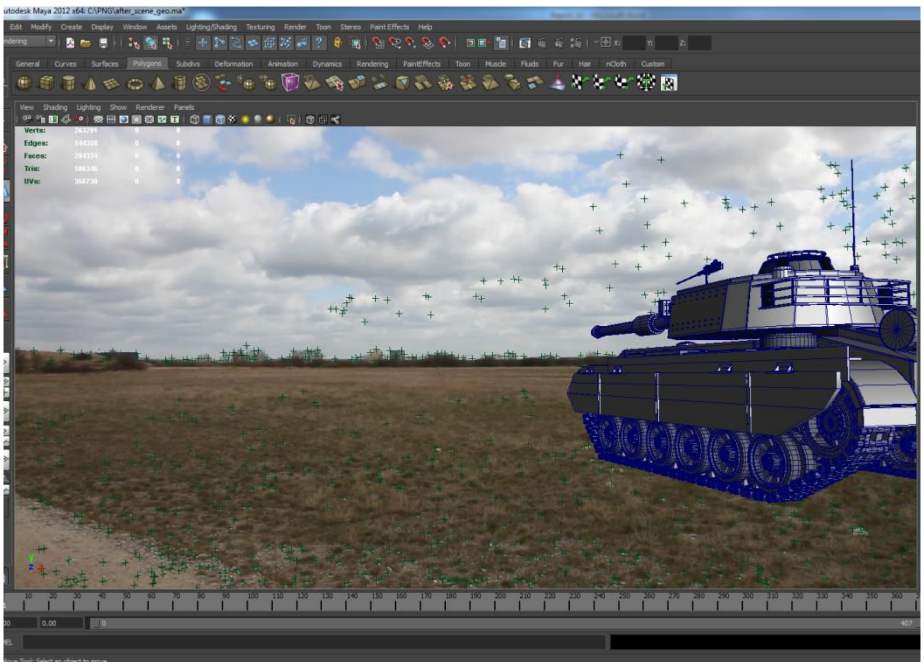
For tracking the *Track Features* was used, which did an automatic tracking of the footage. However new tracks can be added if needed or the entire shot can be tracked manually. Once the automatic tracking was done, the camera has to be re-creates into a 3D camera, this was done using *Solve for Camera*. In order for the software, and the 3D program used, to know what is up *scene geometry* has to be created, which is basically a coordinate system. For this shot a *plane parallel to z-x plane* and then adding an *origin* was used, picture 13 is a shot from Boujou's 3D view showing the tracks and the 3D camera after having added the scene geometry. Picture 14 is shows the ground plane. The scene was exported as a .ma file, in order to import it to Maya. Once imported to Maya, the tank was imported in order to test the view of the scene. Picture 15 shows how importing the 3D camera into Maya and placing a 3D object looked like.



Picture 13: The re-created 3D camera, the yellow and blue dots are the tracks.



Picture 14: Ground Plane and tracks



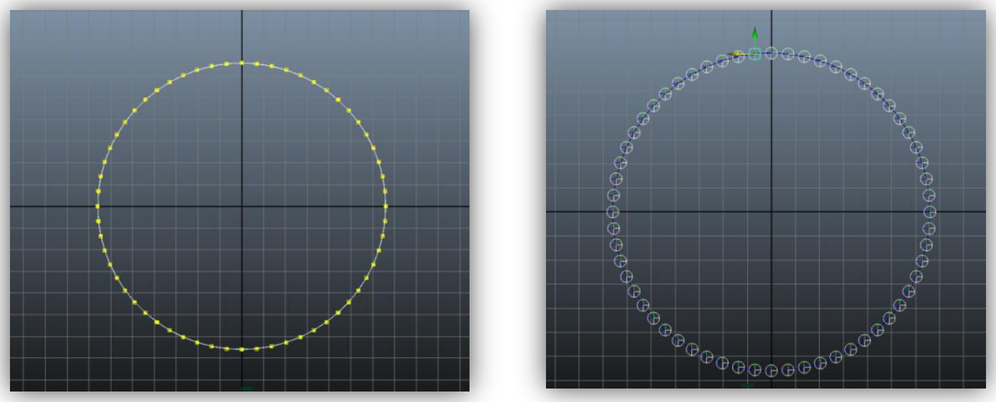
Picture 15: 3D camera and tracks imported to Maya

7.5 Animating the tracks

According to *Workflow* the animation can be done after having matchmoved the footage. This is due to the fact that the model has to move accordingly to the scene movements and be placed upon the ground, so it will appear as if it was really there. In this section the animating of the tank tread will be explained. The animation was done using a motion path, which is a curve that defines a path that the tread will follow.

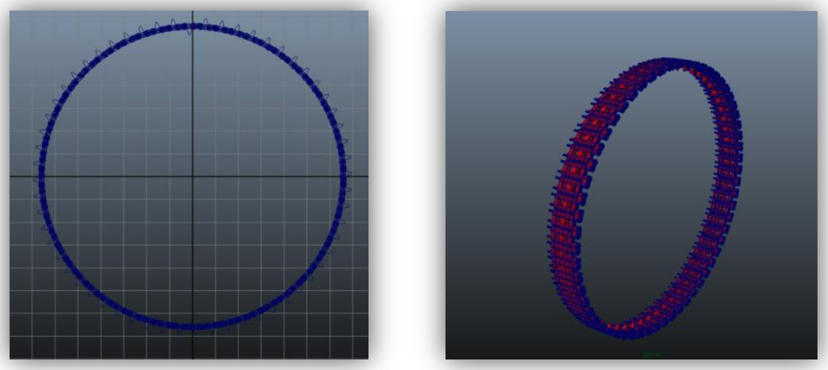
7.5.1 Create a curve

First a NURBS curve circle was created and an even number of vertices was distributed around the curve, in this case 60 because the number of vertices had to match the number of tracks on the tread. The number of vertices also indicated the angle between each track, meaning 360/60 = 6 degrees. Each track should have its own curve circle and therefore the curve was duplicated, using the *duplicate special* tool, 59 times and rotated with the same number of degrees as mentioned before. Before the tracks were attached to the curves, a joint for each track was created and attached to the motion path on each curve. Picture 16 shows the NURBS curve circle and the joint attached to it.



Picture 16: (Left) NURBS Curve and (Right) NURBS curve with joints attached

When the tracks were attached to the curve and smoothed-binded to the joints, the tracks would flip 180 degrees when reaching a certain point. This is due to the fact that the objects always want to be on positive a Y direction. This was fixed by taking the first track, scaling it in – Y, and then freezing the transformation. However, to keep them from flipping the other way around a locator was created and used to chance the up position for the tracks, a locator always has the same up position. Then, in the motion path attribute editor, the *World Up Type* was changed from *Scene Up* to *Object Up* and the locator was chosen as that object, by typing the name locator1 into the *World Up Object* spacing. This had to be done for each track and each motion path, though when reaching a certain point in the rotation, the tracks was set to “inverse up position” because the tracks was starting in the locators position that was negative. Picture 17 shows the NURBS curve having the tracks attached to it.



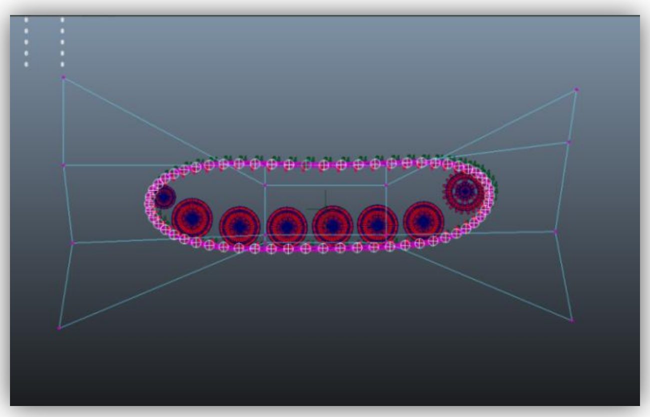
Picture 17: The NURBS curve having tracks attached to it, seen from difference angles

7.5.2 Forming the tank treads

Seeing as a tank tread is not shaped as a uniformed circle, some deformation had to be done. This was done using a lattice, with division 4-4-4. The divisions depend on what type of shape the final result should have, the more divisions the more funky shapes can be produced. Picture 18 shows the *final* shape for the tread. The more the tank treads was squished together, the more "hick-ups" occurred. Therefore some of the tracks were removed, while some were moved slightly in order to make it look even.

7.5.3 Grouping

In order to move the tracks forwards along with the wheels and the rest of the tank *Groups* were created. These groups contained the parts that should move together, e.g. the curves was grouped together under the locator, which meant that when moving the locator the curves, with the attached tracks on it, would move too. After animating the wheels, rotating 360 degrees, they were grouped under a separate locater, though with the same positioning as the other locator. The remainder of the tank was group under the same locator as the wheels.



Picture 18: Final shape of the tank treads

7.5.4 Looping the animation

After 24 frames the animation would stop, because the motion path was set to end at 24 frames. In order to loop the animation the 24 frame range were changed under graph editor, by highlighting all the joints, which then showed all the motion paths attached to the joints in the graph editor. "All" the paths were then selected and the animation was changed to *Curves* → *Post Infinity* → *Cycle*. This made the animation loop beyond 24. However, the tank treads still moved too fast, due to the original 24 frames per second. Therefore the frame end time was changed to 500, which slowed down the animation to the desired speed. Due to looping the animation, the animation continued to rotate until reaching 500, but because the footage was only 407 frames long the animation had to stop before reaching 500 frames. This was done by selecting all the joints and changing the U-value to 0 and keying it at the desired frame that the animation should stop.

7.6 Final Rendering

For rendering the 3D models, mental ray was used. Table 3 show the render settings and other specifications used for the different models. The high LoD model was rendered in 3 different layers, material with HDRI applied, ambient occlusion and a shadow layer, all from frame 1-407. The low LoD model was also rendered using batch render also from frame 1-407.

Specifications	Tank	UFO – High LOD	UFO – Low LOD
Resolution: 1280x720	✔	✔	✔
Multi-Pixel Filtering • **Gauss: 3x3**	✔	✔	
Multi-Pixel Filtering • **Box: 1x1**			✔
Raytrace/Scanline Quality • **Max sample: 2**	✔	✔	
Raytrace/Scanline Quality • **Max sample: 1**			✔
Raytracing enabled	✔	✔	
Global Illumination	✔	✔	
Ambient Occlusion Layer applied	✔	✔	
HDRI applied		✔	
DGS Material		✔	
Blinn Material	✔		✔
Texture Applied	✔		

Table 3: Render settings for the different objects

7.7 Compositing

In order to create scene consistency, making the narrative and the scene content consistent, some set extensions for the original footage was needed. The idea was to make it look like a deserted foreign place with mountains, which we obviously do not have in Denmark and hence the set extensions. The image used for this was found on the internet (Getintravel, 2013) imported into Photoshop to remove any unnecessary image parts, which included the sky and some of the foreground. The image was saved as a PNG, leaving the sky and foreground transparent, and then imported into After Effects. Before the image could be used, tracking the original footages movement was needed, in order to make the set extension follow the original cameras movement. The tracks were parented with a null-object, called Camera Movement.

In order to make the set extension match the original footages, the original footage can be seen in *The Footage* on picture 12, some color correction was done using the different effects found in After Effects under Color Correction. Picture 19 show the set extension, the original footage and set extension combined in After Effects and an example of the footage, with the low and high LOD UFO and tank, where a filter have been added.

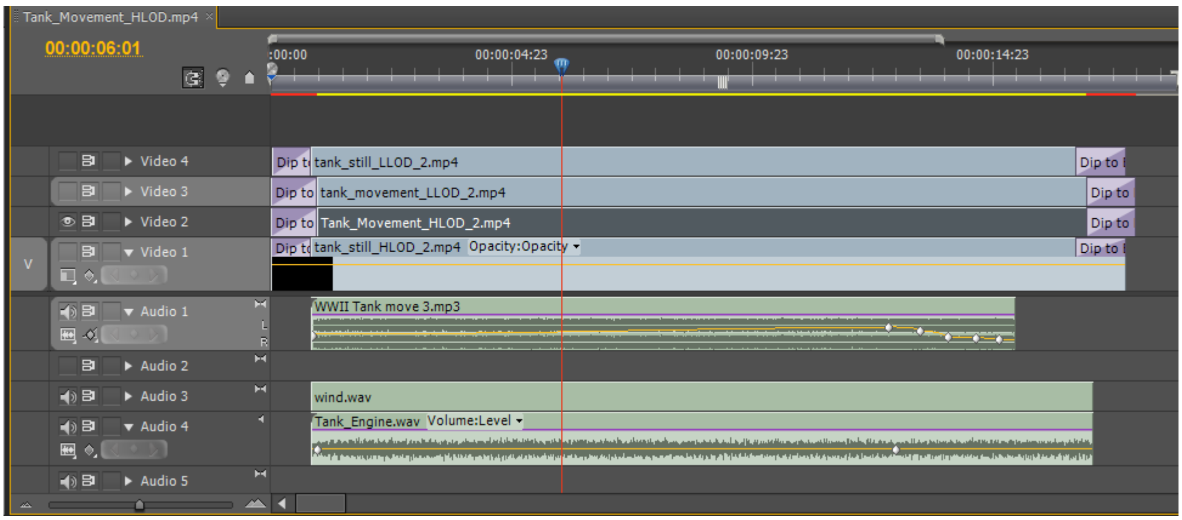
When importing the 3D objects, it is important to chance the frame rate of the 3D footage from 30 fps to 23.967 frames. The frame rate can be chanced under “Interpret Footage”, found by right-clicking on the sequence in the Project tab. The next step was to add the 3D objects and color correct them to fit the footage. The important thing was to try and match the highlights, mid tones and shadow colors found in the footage to the 3D objects. Furthermore there was added some motion blur to the moving tank. Other effects were also used in order to make the objects “fit in” better, which included e.g. some additional shadowing under the UFO. These were then parented to *Camera Movement*, in order to make them follow the camera.

The final footage was rendered as 896x504, as an Mp4 video and a filter was applied to the entire footage, for each clip, the clips can be found on the CD.

Picture 19: (one) set extension, (second) the original and set extension combined (third) low LoD model and tank implemented (fourth) high LoD model and tank implemented

7.7.1 Adding Sound

As found in *Sound* sound can help guide the viewer's attention toward the main elements in a scene, and it was therefore determined that sound should be added to the shot. The sound was added in Adobe Premiere Pro. Three different sound files were used, the files can be found on the CD. In the footage where the tank is moving, sound of a tank moving was added, while in the footage with the still shot of the tank, only the sound of an engine was added, furthermore a wind sound was added to all 4 clips. Some adjustments to the volume had to done, to which both the audio mixer and some additional key frames were placed to lower the sound in the desired placed of the audio. Picture 20 illustrate the 4 clips and added sound files.



Picture 20: Sound being implemented in Adobe Premier Pro

7.7.2 The Narrative

As mentioned earlier, the participants will be told a short story before watching the clips. This narrative was writing in Word and printed on a piece of paper for the participant to read before watching the clips, that way the participants was able to glance at the paper and be reminded of the narrative, in case they would forget.

8 Experimental Design

The choice of experiment structure depends on different aspects such as number of subjects, how many independent variables and how many different conditions each independent variable has. In the next section factorial design, which is the type of test method used for this type of experiment, will be described.

8.1 Factorial Design

When manipulating more than one independent variable in an experiment it is called a factorial design. In a factorial design, all levels of each independent variable are combined with all levels of the other independent variable (Cosby, 2011). The independent variable visual attention, which includes movement or no movement, will contain each condition of the second independent variable level of detail.

This means that for this experiment a 2x2 factorial design should be used, to which the prior number refers to number of level of detail and the ladder refers to the two levels of visual attention.

The last thing to do now is to consider whether or not the test subjects should be exposed to only one condition or multiple conditions. The first test method is called Independent Group design while the ladder is called Repeated Measure design (Cosby, 2011). The Independent Group design would require the most participants as each conditions requires independent samples, say that 10 subjects is wanted to test each conditions this would mean that a total of 40 different participant was needed. Repeated Measure design would require fewer participants, because the same subjects will participate in *all* conditions (Cosby, 2011). It is also possible to make a mixed factorial design. This would require each participant assigned to condition A1 to receive all levels of the independent variable B. If having 10 participants assigned to each "A" condition, this design would require 20 participants in total. Table 4 shows the different factorial designs (Cosby, 2011). Choosing a mixed factorial design could have the disadvantage that subjects would already have been exposed to e.g. the visual attention movement condition and thus be familiar with the object and start to look elsewhere. One way to counter this possible inherent bias is to make use of counterbalancing. To achieve this counterbalance, subjects will be randomly assigned a condition, where the order of the different clips presented within the given condition will be randomized for each test participant (Cosby, 2011). This would mean that if having enough test subjects the bias should be minimized.

With this in mind, the mixed factorial design is chosen as the best design for this experiment. This means that there will be two different level of visual attention acting as the independent group and the two different level of detail will act as the repeated measures.

A: Visual Movement Condition	B: Level of Detail Condition		
	Within Group	B1	B2
	A1	S1	S1
	A1	S1	S1

A: Visual Movement Condition	B: Level of Detail Condition		
	Between Group	B1	B2
	A1	S1	S3
	A1	S3	S4

A: Visual Movement Condition	B: Level of Detail Condition		
	Mixed Design	B1	B2
	A1	S1	S1
	A1	S2	S2

Table 4: The 3 different factorial designs: within group, between group and mixed design.

8.2 Latin Square

As mentioned in *Factorial Design* using a mixed factorial design could result in bias, due to test participants being exposed to the same the visual attention condition twice, and in order to counter this subjects should be randomly assigned a condition, where the order of the different clips presented within the given condition will be randomized for each test participant. This means that the following random assignment will apply for the experiment. In each order a minimum of 5 participants is needed, see table 5 for example of this randomization.

For the purpose of this experiment it is necessary to have at least 20 subjects which mean that each condition of visual attention will require 10 subjects. However if more than 10 subject are available it is not the maximum limit. Seeing as everyone could be exposed to a sequence containing real footage and CGI, there could be in reality be no limitation to who these clips could be shown to. However, due to experience there should be some sort of age limit, seeing as children is usually not the easiest to use for testing and to avoid any parental issues the age limit is set to 18.

Movement			
Order	Level of Detail	Level of Detail	Number of participants
1	H	L	5
2	L	H	5

No Movement			
Order	Level of Detail	Level of Detail	Number of participants
1	H	L	5
2	L	H	5

Table 5: (Top) Example of the movement randomization and (top) example of the no movement randomization, with a min. of 5 participants in each order.

9 RESULTS

In this chapter the results from the conducted test will be presented and the data analysis method used is ANOVA to test for significant difference.

The null hypotheses in this experiment are based on the problem statement:

> ***To what degree can movement in a foreground object be utilized in a scene in order to alter the level of detail in a background object without any noticeable change and still maintain the perception of believability?***

For the experiment two conditions movement and no movement and the appertaining level of detail was created. The idea was that when conducting the experiment, the participants watching the clip with *movement* would not notice the difference between the LoD, while the participants watching the clip containing *no movement* would notice the difference in LoD between the clips. In other words there should be a noticeable difference between the interaction of visual movement and level of detail and the overall believability rating $\dot{B}$, thus the null hypotheses are:

H_0: Visual movement will have no significant effect on the overall $\dot{B}$

H_0: Level of detail will have no significant effect on the overall $\dot{B}$

H_0: Visual movement and level of detail interaction will have no significant effect on the overall $\dot{B}$

To test if the above mentioned null hypotheses can be rejected the data was analyzed using excel to calculate an ANOVA on the entire dataset. The results are presented in the following section and explained.

9.1 Participants

The participants taking part in this experiment consisted of 24 adults, where 7 females and 17 males were amongst the test participants, with an average age of 25 (±3,15). There were two different groups with an appertaining visual movement condition, being movement or no movement, and additionally two different level of detail *low* and *high*. The distribution of gender between conditions can be seen in table 6. The experiment took place at *Aalborg University in Copenhagen* in a silent room, where participants were seated in from of a laptop and instructed in the experiment taking place. The participant rated on a 6-point Likert scale how familiar they considered themselves with visual effects, the average rating can be seen in table 7.

Condition	Female	Male
Movement	3	9
No Movement	4	8

Table 6: distribution of gender between conditions

Condition	Average rating
Movement	4,5 (±0,6)
No Movement	3,9 (±0,6)

Table 7: Average rating of participants acquaintance with VFX

9.2 Overall Believability Response Rating $\dot{B}$

In this experiment a mixed factorial design was used to test the connection between the two independent variables, visual movement and the level of detail. Each visual movement condition was run twice with each level of detail condition. To avoid carry over effects the level of detail condition was randomly assigned among the 12 participants, though have an equal number of participants watching each condition.

In table 8 and 9 the average believability response rating $\dot{B}$ can be seen, this value was calculated as the total mean for all participants in each visual movement condition and their appertaining LoD. The overall believability response rating shows an indication towards the high LoD object having an overall higher $\dot{\mathbf{B}}$ in both visual movement conditions than the low LoD object. This indication may also relate to why the participants more frequently rated clips containing the high LoD object to be *most believable* in both visual movement conditions compared to the low LoD or the LoD as being equally believable, see figure 6 for reference.

Movement condition/LoD	Low	High	Average
Movement	3,47 (±0,6)	3,88 (±0,67)	3,67 (±0,66)
No movement	3,11 (±0,72)	3,34 (±0,82)	3,23 (±0,77)
Average	3,29 (±0,72)	3,61 (±0,82)	

Table 8: Overall believability response rating $\dot{B}$

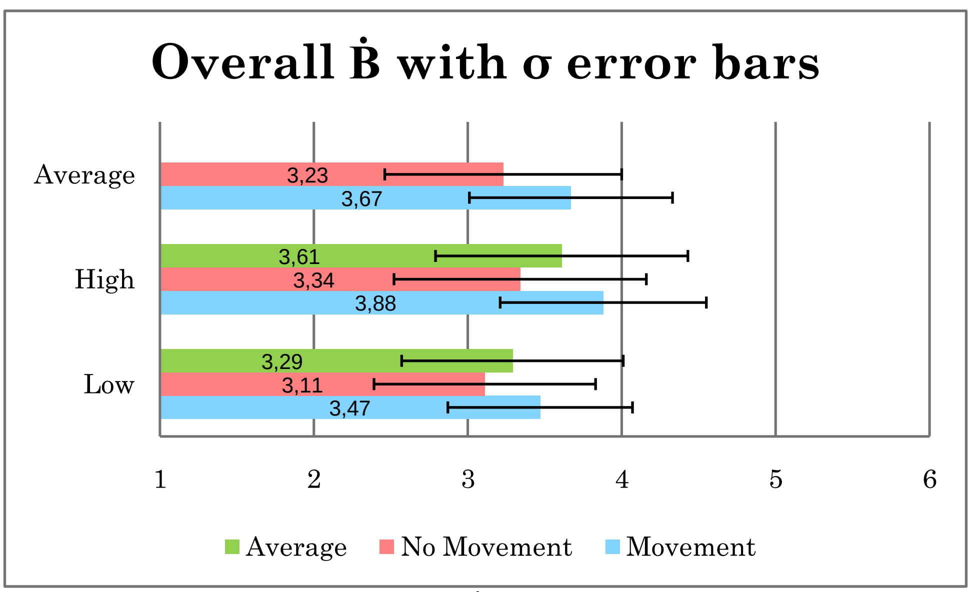


Table 9: Overall $\dot{B}$ with σ error bars

9.3 Rating the Clips

Even though participants said they did not see a difference when movement was present, see figure 7, people still rated the high LoD as being the most believable more frequently than the low LoD or both as being equally believable, see figure 6.

This was also the case when no movement was present. However unlike the movement condition, there were some who noticed a difference in the background. One test participant stated:

> *The unidentified objects seemed to be clearer in clip 2, it looked like it belonged. In clip 1 the base of the unidentified objects looked fake.*

Furthermore three test participants felt like there was a difference in the background elements, even though they had difficulties stating exactly what the difference was. This could be due to the fact that they were only exposed to the clip for such a short duration, and only noticed the difference on an unconsciously level.

In the condition where movement was present, people did not notice the camera movements before watching the second clip, which might be due to *something* taking their focus in the first clips

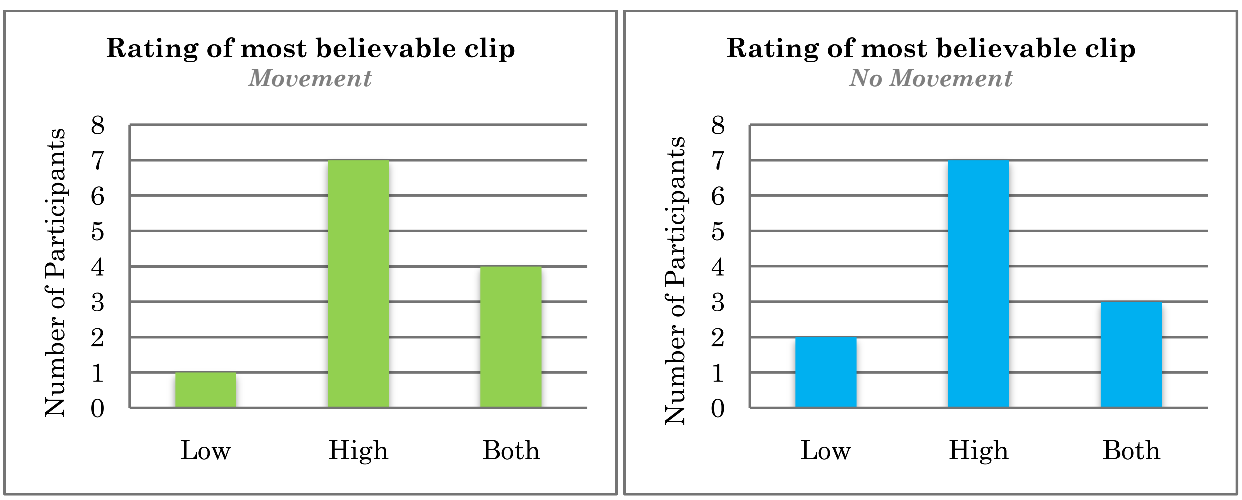


Figure 6: Number of participants rated "most believable clip" (low, high or both) for movement (left) and no movement (right)

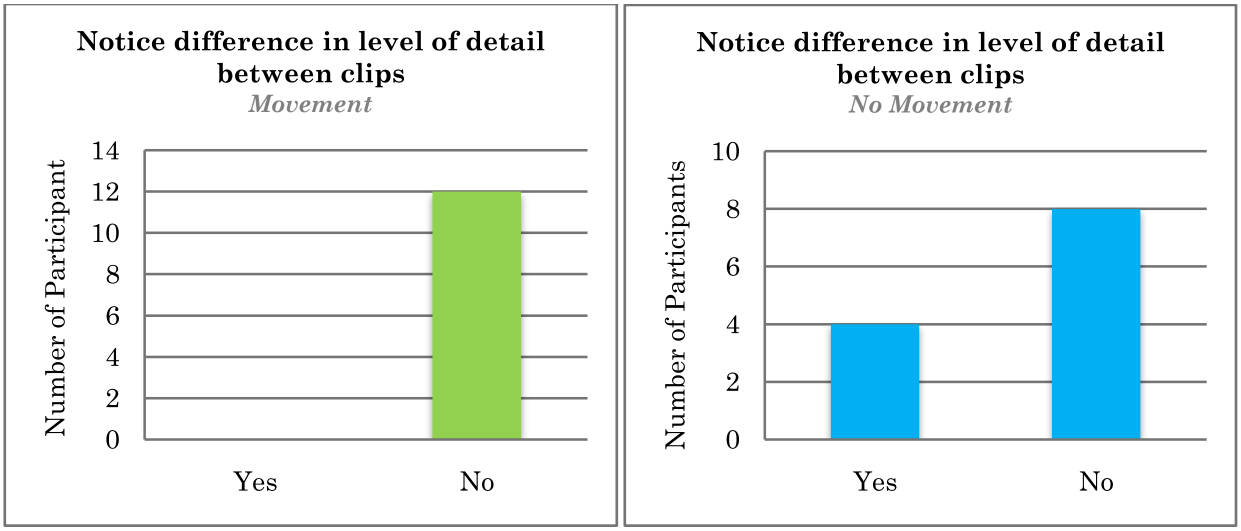


Figure 7: Number of participant stating if they noticed a difference between the clips (yes or no) in (left) movement and (right) no movement condition.

9.4 ANOVA

In this experiment a mixed design was used because there is more than one independent variable, being visual movement and level of detail, Two-Way ANOVA should be used. Initially an overall ANOVA was conducted to test for any significant difference between the independent variables.

The input used for the ANOVA test was a table containing a mean rating value for each participant in both visual movement conditions, meaning the combined rating of the 8 statements for high and low LoD, see *Appendix* for further explanation.

The overall ANOVA showed a significant difference between the conditions *movement* and no *movement* (F (4,06) = 4,67, p = 0,03613), while there were no significant difference between the level of detail (F (4,06) = 2,40, p = 0,128001) or the interaction between the variables (F (4,06) = 0,09, p = 0,66762).

At this point it is possible to reject the null hypothesis H_0: Visual movement will have no significant effect on the overall Ḃ, as there is a significant difference in the visual movement variable. Additionally we must accept the other null hypotheses made, as no significant difference was found in the level of detail variable or the interaction between the two (level of detail x visual movement).

Even though the overall ANOVA showed a significant difference between the two conditions, movement and no movement, the ANOVA does not tell where this significant difference is.

To examine which statement caused the significant difference in the overall ANOVA, a new ANOVA was conducted on each statement to compare the two variables visual movement and different level of detail, to see if they were significantly different. Statement 1: (F (4,06) = 10,08, p = 0,00451) and Statement 3: (F (4,06) = 13,580, p = 0,00623) showed a significant difference between the conditions movement and no movement.

The average rating of the two statements can be seen in table 9 and 10, which also shows are difference compared across visual movement, see *Appendix* for the rest of the statements and their average rating.

Statement 1	Low	High
Movement	4,42 (±1,16)	4,33 (±1,07)
No Movement	3,5 (±1,09)	3,42 (±0,9)

Table 10: Average rating statement 1

Statement 3	Low	High
Movement	4,42 (±0,51)	4,58 (±0,67)
No Movement	3,58 (±1)	3,75 (±0,087)

Table 11: Average rating statement 3

10 DISCUSSION

The goal of this experiment was to investigate if utilizing movement in a foreground object would attract the viewers' attention and thus one would be able to alter the level of detail in a computer generated background object without any noticeable change and therefore still maintain the overall perception of believability of the scene. As mentioned in *Criteria for Success* the optimal results would have been that the participants would have noticed a difference in LoD in the *no movement* condition, whereas the participants in the *movement* condition would not have been able to notice any difference in the LoD.

The overall believability response rating $\dot{B}$, as seen in table 8, showed an indication towards the high LoD object having an overall higher $\dot{B}$ in both visual movement conditions compared to the low LoD object. This overall higher $\dot{B}$ could also relate to why the participants more frequently rated clips containing the high LoD object to be the most believable clip of the two, see figure 6 for reference. Even though the high LoD was rated higher than the low LoD it seems as the *movement* condition has been rated as more believable when comparing it to the overall $\dot{B}$ for the *no movement* condition.

However, it was necessary to investigate whether the raw data gathered was valid and if there could be concluded anything on the test conducted, for this a 2x2 mixed factorial ANOVA was used.

At first glance the ANOVA showed that there were *only* a significant difference between the two conditions, movement and no movement ($F (4,06) = 4,67$, $p = 0,03613$), and on these grounds the corresponding null hypothesis made was rejected. To further investigate which statement caused the significant difference an ANOVA for each of statements was conducted. These findings showed a significant difference in Statement 1 and 3, relating to the narrative told before seeing the clip and the implemented sound. Comparing the average rating for Statement 1 and 3 across visual movement it is also clear that there is a difference in the ratings.

The narrative may have indicated the tank was moving *towards* the UFO, however in the no movement scene the tank was stationary. Therefore the rating of the statement *"The content of the scene fits with the narrative told before seeing the clip"* might have been negatively affected in the no movement condition.

In the movement condition the sound emitted should indicate that the tank was actually moving, and therefore engine sound and gear squeaking noise was applied, whereas the stationary tank only had the engine sound applied. Based on the mean value being higher for statement *"The implemented audio seems to fit the scene content"* in the movement condition, it is likely that the added sound can has contributed to a higher believability rating of the scene, contrary to the no movement condition. However, due to the sound differ from each other is hard to say if it was actually the sound, the movement or both affecting the participants rating.

The overall impression may have been affected depending on whether or not movement was present in the scene, however there is no conclusive evidence that supports the notion that movement draws attention away from a background object. On the other hand participants stated that they noticed a difference in the background object when no movement was present, though 1/3 of the participants noticed the difference, only 1 participant was able to exactly specify where this difference was located (red. The UFO).

There can be several season why the UFO was not noticed more often, perhaps the tank was a too dominant part of the scene composition, which could have drawn attention away from the remainder of the scene in both movement and no movement. In addition, the integration of the low detail UFO in the scene could have been too well done, making the distinction hard to notice for the participants when presenting the clips in a bottom-up approach compared to a top-down approach where the participants may have been told to look for any differences beforehand.

Seeing as the test was conducted on students who may have knowledge within the field of VFX, see table 7 for reference, this could have affected the overall believability rating, because they may feel the VXF does not live up to a certain standard. If conducting the test again, it could be considered to choose a wider target group. Another issue could be the low amount of conditions in each of my variables (red. Visual movement and LoD), one could consider expanding the conditions with more intermediate conditions and making sure the distinction between the conditions (red. low and high, movement and no movement) grew enough to be noticeable from the end point conditions. This condition would then contain no foreground object that could attract the viewers' attention and could help in determining if the difference in LoD was actually distinguishable enough for the viewer to notice or if further difference had to be applied. However, this would also mean that the questionnaire had to be reversed, seeing as the clip would not contain the foreground element for the participant to rate.

Another factor that can have affected the results is the fact that believability is a subjective manner, and my definition of believability may not be the same as the test participants. An idea could be to have a focus group meeting gathering a different group of people and ask them to discuss the word *believability* to see which key-words they would come up with in relations to movies (real and cartoons), computer games, and overall computer generated imagery and effects. These keywords could then be compared to the theory used. However, it should be kept in mind that this target group only represents a small number of the entire population.

Lastly, it is uncertain what the participants were focusing on the in clips presented to them because eye tracking was not implemented. I thought that by asking the participants if they noticed a difference would indicate what they were focusing on, however this was not the case. Therefore an idea could have been a follow-up interview asking the participant directly what they were focusing on when they saw the clip1 and clip2.

11 Conclusion

If looking at the problem statement created for this experiment,

> ***To what degree can movement in a foreground object be utilized in a scene in order to alter the level of detail in a background object without any noticeable change and still maintain the perception of believability?***

The results gained from the test conducted, showed that visual movement will influence the perceived believability. However, there is no conclusive evidence that movement attracts attention away from background objects, as the test results showed no distinction between the level of detail and believability response rating in the scene with no movement. Ultimately this means that there is proof what the participants were focusing on, and the information extracted from the questionnaire was too ambiguous to draw any definite conclusions from.

12 Future Development

If conducting the experiment again, implementing another visual movement factor, where the tank would not be present, would seem appropriate in order to establish if the tank was the reason for the participants not noticing the difference between the the LoD or if the difference was just not significant enough. Furthermore, a revision of the scene composition could be made, e.g. by conducting preliminary tests to find the most suitable positioning of the objects.

Furthermore implementing eye tracking could be a possibility in order to establish what the participants were focusing on, however if this is not doable, interviewing the participants could be an additional feature in the test. This would give an opportunity to ask the participants what they were focusing on the clips presented.

Another possibility is to simplify the scene, containing fewer elements that could distract the participants.

13 Appendix

This part of the report contains the appendix. It contains the questionnaire, the sketches used for creating the CGI tank and different tables from the test conducted.

13.1 Questionnaire

This is the questionnaire used for the experiment. The first section was the personal information that the participants had to fill out *before* seeing the clips.

Questionnaire regarding perceived believability

Before starting the test, please fill out the following questions below

---

How old are you? _______

What is your gender, please cross off? **Male** ☐ **Female** ☐

What is your occupation? ________________________

I am well familiar with visual effects:

Strongly Disagree	1	2	3	4	5	6	Strongly Agree

The sequence that you have just watched contained computer generated object, which means that they have been created in a 3D program and implemented into live footage. Please cross out the most appropriate number for each statement which corresponds closely to your desired response. Remember there is no right or wrong answers. *Required

The content of the scene fits with the narrative told before seeing the clip*

Strongly Disagree	1	2	3	4	5	6	Strongly Agree

There is connection between the foreground and background elements in the scene *

Strongly Disagree	1	2	3	4	5	6	Strongly Agree

The implemented audio seems to fit the scene content*

Strongly Disagree	1	2	3	4	5	6	Strongly Agree

The tank appears as if it was really there*

Strongly Disagree	1	2	3	4	5	6	Strongly Agree

The tank appears to respond to the unfolding situation relating to the unidentified object*

Strongly Disagree	1	2	3	4	5	6	Strongly Agree

The unidentified object appears as if it was really there*

Strongly Disagree	1	2	3	4	5	6	Strongly Agree

The tank was perceived as believable*

Strongly Disagree	1	2	3	4	5	6	Strongly Agree

The unidentified object was perceived as believable*

Strongly Disagree	1	2	3	4	5	6	Strongly Agree

Now that you have watched both clips please cross off the clip that you found to be most believable, or if they were equally believable*

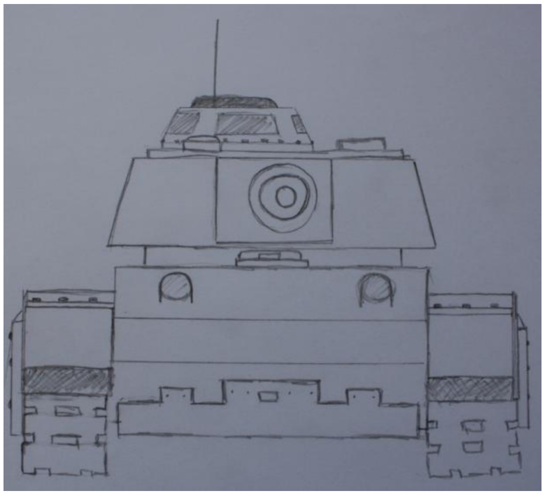
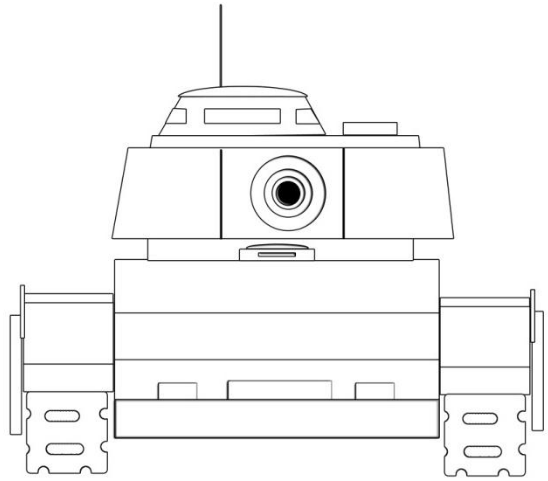
Clip 1: ☐ *Clip 2:* ☐ *Equally believable:* ☐

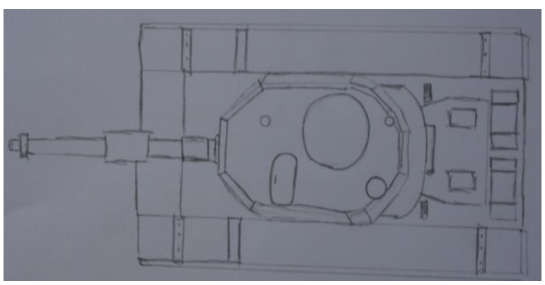
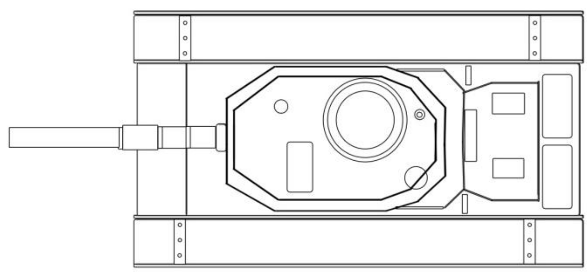
Did you notice any changes in the scene appearance between the clips, if yes please elaborate on which below?

Further Comments:

13.2 Sketches

The remainder of the sketches used to create the CGI tank.

13.3 Average response table

This is the table used to conduct the overall ANOVA, each participant have resulted in one average rating for low and high LoD.

Movement condition/LoD	Low	High
Movement	3,38	4,13
Movement	3,63	3,88
Movement	4,13	4,00
Movement	2,75	3,13
Movement	3,75	4,88
Movement	3,88	4,88
Movement	2,00	2,63
Movement	3,63	3,13
Movement	3,88	4,00
Movement	4,00	4,13
Movement	3,13	3,63
Movement	3,50	4,13
No Movement	2,63	2,63
No Movement	3,38	4,63
No Movement	1,63	2,00
No Movement	3,25	2,75
No Movement	3,50	4,00
No Movement	2,75	3,00
No Movement	3,75	3,88
No Movement	3,13	2,88
No Movement	4,38	4,50
No Movement	3,75	4,00
No Movement	2,50	2,88
No Movement	2,75	3,00

Overall ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Sample	2,351888	1	2,351888	4,672526	0,03613	4,061706
Columns	1,211263	1	1,211263	2,406432	0,128001	4,061706
Interaction	0,094076	1	0,094076	0,186901	0,66762	4,061706
Within	22,14714	44	0,503344			
Total	25,80436	47				

13.4 Average rating all statements

This table shows the average rating of all the statements.

Statements	1	2	3	4	5	6	7	8
Movement / low LoD	4,42 (±1,16)	3,17 (±0,94)	4,42 (±0,51)	2,67 (±1,15)	3,5 (±1,17)	3,42 (±1,24)	3,25 (±1,14)	2,92 (±1,38)
Movement / high LoD	4,33 (±1,07)	3,83 (±0,94)	4,58 (±0,67)	3,08 (±1,31)	3,75 (±0,87)	3,83 (±1,11)	3,75 (±0,97)	3,83 (±1,27)
No Movement / low LoD	3,5 (±1,09)	3,58 (±1)	3,58 (±1)	2,83 (±1,4)	3,58 (±0,9)	2,42 (±1,38)	2,92 (±1)	2,5 (±1,38)
No Movement / high LoD	3,42 (±0,9)	3,58 (±1,31)	3,75 (±0,87)	2,67 (±1,37)	3,58 (±1,16)	3,42 (±1,56)	3,25 (±1,22)	3,08 (±1,38)

14 REFERENCES

Agre, P. E., & Chapman, D. (1990). What are plans for? (P. Maes, Ed.) *Robotics and Autonomous Systems*, *6*(1-2), 17-34. Elsevier. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0921889005800260

Astheimer, P. (1993). What you see is what you hear-Acoustics applied in virtual worlds. *Proceedings of 1993 IEEE Research Properties in Virtual Reality Symposium*. Retrieved from http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=378256

Babcock, S., A., Lipps, M., & Pelz, B., J., (2002). How people look at pictures, before, during and after a scene capture: Buswell revisited. Human Vision and Electronic Imaging VII (January), 4662(1), pp.34-47. Retrieved from http://www.jasonbabcock.net/research/Babcock_SPIE_HVEI.pdf

Bates, J. (1994). The Role of Emotion in Believable Agents Joseph Bates. *Emotion*, *37*(April), 122-125. Carnegie Mellon University. Retrieved from http://portal.acm.org/citation.cfm?doid=176789.176803

Bordwell, D. & Thompson, K., (2010). *Film Art – An Introduction* (9th ed.) New York, NY: McGraw-Hill

Boyce, S. J., & Pollatsek, A. (1992). Identification of objects in scenes: the role of scene background in object naming. *Journal of experimental psychology Learning memory and cognition*, *18*(3), 531-543. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1534354

Birmingham, E., Bischof, W., & Kingstone, A. (2009). Get real! Resolving the debate about equivalent social stimuli. *Visual Cognition*, *17*(6), 904-924. Psychology Press, part of the Taylor & Francis Group. Retrieved from http://www.tandfonline.com/toc/pvis20/17/6-7

Chalmers, A., Debattista, K., Santos, D., & Paulo, L. (2006). Selective rendering : computing only what you see. *GRAPHITE 06 Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia*, *1*(212), 9-18. ACM. Retrieved from http://dx.doi.org/10.1145/1174429.1174431

Chalmer, A. and Ferko, A., (2008). Levels of Realism: From Virtual Reality to Real Virtuality. *Computer Graphics: Proceedings of the 24th Spring Conference, (SCCG '08), 2008*, pp.19-25. Retrieved from http://dl.acm.org/citation.cfm?id=1921272

Connor, C. E., Egeth, H. E., & Yantis, S., (2004). Visual attention: bottom-up versus top-down. *Current Biology*, *14*(19), R850-R852. Elsevier. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15458666

Cozby, C. P, & Bates, C. S., (2011). *Methods in behavioral research* (11th ed.). New York, NY: McGraw-Hill

Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science*. SAGE Publications. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/15271002

Davenport, J. L. (2007). Consistency effects between objects in scenes. *Memory cognition*, *35*(3), 393-401. Psychonomic Society Publications. Retrieved from http://mc.psychonomic-journals.org/content/35/3/393.short

Derakhshani, D., (2011) *Introducing Autodesk Maya 2012*. Hoboken, NJ: Sybex . Retrieved from http://site.ebrary.com/lib/aalborguniv/docDetail.action?docID=10513932

Ferwerda, J. A., Pattanaik, S. N., Shirley, P., & Greenberg, D. P. (1996). A model of visual adaptation for realistic image synthesis. *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques SIGGRAPH 96*, *30*(3), 249-258. ACM Press. Retrieved from http://portal.acm.org/citation.cfm?doid=237170.237262

Ferwerda, J. A., Shirley, P., Pattanaik, S. N., & Greenberg, D. P. (1997). A model of visual masking for computer graphics. *Proceedings of the 24th annual conference on Computer graphics and interactive techniques SIGGRAPH 97*,*31*(3), 143-152. ACM Press. Retrieved from http://portal.acm.org/citation.cfm?doid=258734.258818

Ferwerda, J. A. (2003). Three varieties of realism in computer graphics. (B. E. Rogowitz & T. N. Pappas, Eds.) *Proceedings of SPIE*, *5007* (2003), 290-297. Retrieved from http://link.aip.org/link/?PSI/5007/290/1&Agg=doi

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1030-1044. APA American Psychological Association. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1431742

Funkhouser, A. T, (1993, A). Database and Display Algorithms for Interactive Visualization of Architectural Models. . *Report, Berkeley University of California Computer Science Division*, 1-144. Retrieved from http://www.eecs.berkeley.edu/Pubs/TechRpts/1993/CSD-93-771.pdf

Funkhouser, A., T., and Séquin, H., C., (1993, B). Adaptive Display Algorithm for Interactive Frame Rates During Visualization of Complex Virtual Environments. *Computer graphics and interactive techniques: Proceedings of the 20th annual conference, (SIGGRAPH '93), 1993*, 247-254. Retrieved from http://dl.acm.org.zorac.aub.aau.dk/citation.cfm?id=166117

Fowkes, A. K., (2010). The Fantasy Film. Chichester, UK: Wiley-Blackwell

Getintravel (2013). Tour to the band E Amir National Park Afghanistan. Retrived May 2nd 2013 from http://www.getintravel.com/visit-band-e-amir-afghanistan-in-the-central-hindu-kush-mountains/tour-to-the-band-e-amir-national-park-afghanistan/

Guindon, M., (2009). Learning Autodesk Maya 2010: Foundation. San Rafael, CA: Sybex. Retrieved from http://proquest.safaribooksonline.com.zorac.aub.aau.dk/book/-/9781897177556

Hasic, J., Chalmers, A., & Sikudova, E. (2010). Perceptually guided high-fidelity rendering exploiting movement bias in visual attention. *ACM Transactions on Applied Perception*, *8*(1), 1-19. Association for Computing Machinery, Inc. Retrieved from http://dx.doi.org/10.1145/1857893.1857899

Henderson, J. M., Weeks, P. A. J., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology Human Perception and Performance*,*25*(1), 210-228. American Psychological Association. Retrieved from http://doi.apa.org/getdoi.cfm?doi=10.1037/0096-1523.25.1.210

Humphrey, K., Underwood, G., & Lambert, T. (2012). Salience of the lambs : A test of the saliency map hypothesis with pictures of emotive objects. *Journal of Vision*, *12*(1), 1-15. Retrieved from http://ww.w.journalofvision.org/content/12/1/22.short

James, W., (1981). *The Principles of Psychology*. Cambridge, MA: Harvard University Press. Retrieved from http://ebooks.adelaide.edu.au/j/james/william/principles/preface.html

Lanier, L., (2008). *Advanced Maya Texturing and Lighting* (2nd ed.). Indianapolis, IN: Wiley Publishing. Retrieved from http://proquest.safaribooksonline.com.zorac.aub.aau.dk/book/animation-and-3d/9780470292730?bookview=overview

Livingstone, D. (2006). Turing' s test and believable AI in games. *Computers in Entertainment*, *4*(1), 6. ACM Press. Retrieved from http://portal.acm.org/citation.cfm?doid=1111293.1111303

Loyall, A. B. (1997). Believable Agents: Building Interactive Personalities. *Fujitsu*. Citeseer. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.62.2436&rep=rep1&type=pdf

Luebke, D., Reddy, M., Cohen, J. D., Varshney, A., Watson, B., & Huebner, R. (2003). *Level of Detail for 3D Graphics*. San Francisco, CA: Morgan Kaufmann. Retrieved from http://books.google.dk/books?id=uiz6IKAVxP8C&printsec=frontcover&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false

Namee, M. B., (2004).Proactive Persistent Agents: Using Situational Intelligence to Create Support Characters in Character-Centric Computer Games. *Department of Computer Science, University of Dublin*. Retrieved from http://www.tara.tcd.ie/bitstream/2262/1200/1/TCD-CS-2004-27.pdf

Magerko, B. (2007). Measuring Dramatic Believability. *Paper from the AAAI Fall Symposium 07*. AAAI Press, 79-82. California. Retrieved from http://www.aaai.org/Papers/Symposia/Fall/2007/FS-07-05/FS07-05-014.pdf

Mastoropoulou, G., Debattista, K., Chalmers, A., & Troscianko, T. (2005). Auditory bias of visual attention for perceptually-guided selective rendering of animations. *Proceedings of the 3rd international conference on Computer graphics and interactive techniques in Australasia and South East Asia GRAPHITE 05*, *1*(212), 363. Comenius University. Retrieved from http://dl.acm.org/citation.cfm?id=1101462

Mastoropoulou, G., Debattista, K., Chalmers, A., & Troscianco, T. (2005). The influence of sound effects on the perceived smoothness of rendered animations. *Proceedings of the 2nd symposium on Applied perception in graphics and visualization APGV 05*, 9. ACM. Retrieved from http://dx.doi.org/10.1145/1080402.1080404

Multimedia Digest (2012, Novement 25th). Matchmoving in Maya Live – Canon 7D, Nikon, Film Camera [Web log post]. Retrived April 20th, 2013, from http://multimedia-digest.com/matchmoving-in-maya-live-canon-7d

Niu, Y., Todd, R. M., Kyan, M., & Anderson, A. K. (2012). Visual and emotional salience influence eye movements. *ACM Transactions on Applied Perception*, 9(3), 1-18. Retrieved from http://dl.acm.org.zorac.aub.aau.dk/citation.cfm?doid=2325722.2325726

Odell, O., & Le Blanc, M., 2001.The pocket essential horror films. Harpenden, GBR: Pocket Essentials,. Retrieved from http://site.ebrary.com/lib/aalborguniv/docDetail.action?docID=5007504

Openfootage.net (2010, September 15th). HDRI 360º Stallersattel eastern Triol [Web log post]. Retrieved May 6th 2013 from http://openfootage.net/Openfootage/Vorschau/ContactSheet_Stallersattel.jpg

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition*, *3*(5), 519-526. Springer. Retrieved from http://www.springerlink.com/index/10.3758/BF03197524

Porter, T. and Susman, G., (2000).Creating Lifelike Characters in Pixar Movies. *Communications of the ACM*, 43(1), 25-29. Retrieved from http://dl.acm.org.zorac.aub.aau.dk/citation.cfm?doid=323830.323839

Prince, S. (1996). True Lies: Perceptual Realism, Digital Images, and Film Theory. (L. Braudy, Ed.)*Film Quarterly*, *49*(3), 27-37. University of California Press. Retrieved from http://www.jstor.org/stable/1213468

Rademacher, P., Lengyel, J., Cutrell, E., & Whitted, T. (2001). Measuring the Perception of Visual Realism in Images. In S. J. Gortler & K. Myszkowski (Eds.), *Proceedings of the 12th Eurographics Workshop on Rendering Techniques* (p. 235–248). Citeseer. Retrieved from http://research.microsoft.com/en-us/um/people/cutrell/EGRW2001-Rademacher-final_version.pdf

Reddy, M., (1997). Perceptually Modulated Level of Detail for Virtual Environments. *University of Edinburgh. College of Science and Engineering. School of Informatics*. Retrieved from: http://www.era.lib.ed.ac.uk/bitstream/1842/505/2/Reddy_thesis_pdf

Reddy, M., (1998). Specification and Evaluation of Level of Detail Selection Criteria. *Journal: Virtual Reality,* 3(2), 132-143. Retrieved from: *http://link.springer.com.zorac.aub.aau.dk/article/10.1007/BF01417674*

Reddy, M. (2001). Perceptually optimized 3D graphics. *IEEE Computer Graphics and Applications*. IEEE Computer Society Press. Retrieved from http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=946633

Shlaer, S. (1937). The relation between visual acuity and illumination. *The Journal of general physiology*, *21*(3), 255-281. The Rockefeller University Press. Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2141937&tool=pmcentrez&rendertype=abstract

Shams, L., Kamitani, Y., & Shimojo, S. (2000). Illusions. What you see is what you hear. *Nature*, *408*(6814), 788. Nature Publishing Group. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=11130706

Styles, A., E., (2005). *Attention, Perception and Memory, an Integrated Introduction*. New York, NY: Psychology press.

Telotte, J. P., (2001). *Science Fiction Film*. New York, NY: Cambridge University Press. Retrieved from http://site.ebrary.com/lib/aalborguniv/docDetail.action?docID=10001857&p00=Focus+on+the+science+fiction+film

Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica*, *135*(2), 77-99. Elsevier B.V. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/20507828

The Foundry (2013). ROLLINGSHUTTER CMOS Motion Artefact Correction. Retrieved April 20th 2013 from http://www.thefoundry.co.uk/products/rollingshutter/

Thomas, F. and Johnston, O., (1981). The illusion of life: Disney Animation. New York, NY: New York: Hyperion. Retrieved from http://www.scribd.com/doc/25187588/The-Illusion-of-Life-Disney-Animation

Tinwell, A., Grimshaw, M., Nabi, D. A., & Williams, A. (2011). Facial expression of emotion and perception of the Uncanny Valley in virtual characters. *Computers in Human Behavior*, *27*(2), 741-749. Elsevier Ltd. Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S074756321000316X

Togelius, J., Yannakais N. G., Shaker, N. & Karakovskiy, S. (2012). Assessing Believability. In Phil Hingston (2nd ed.), *Believable Bots: Can Computers Play Like People?,* (pp. 215-230). Berlin, Heidelberg: Springer Berlin Heidelberg

Yarbus, L., A., 1967. *Eye Movement and Vision.* (L. A. Riggs, Trans.) New York, NY: Plenum Pres. (Original work published 1965)

Watson, B., Walker, N., Hodges, L. F., & Worden, A. (1997). Managing level of detail through peripheral degradation: effects on search performance with a head-mounted display. *ACM Transactions on Computer-Human Interaction*,*4*(4), 323-346. ACM. Retrieved from http://portal.acm.org/citation.cfm?doid=267135.267137

William, J., (1927). *Focus on the science fiction film*. Englewood Cliff, NJ: Prentice-Hall.

Wolfe, M. J., Keith, R. K., Levi, D. M, Bartoshuk, L. M., Herz, R. S., Klatzky, R. L., Lederman, S. J., and Merfeld, D. M., (2009). *Sensation & Perception (2nd ed.)*, Sunderland, MA: Sinauer Associates

Zhu, Y., Jacobson, N., Pan, H., & Nguyen, T. (2011). Motion-decision based spatiotemporal saliency for video sequences. *2011 IEEE International Conference on Acoustics Speech and Signal Processing ICASSP*. IEEE. Retrieved from http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5946658

15 TABLE OF FIGURES

Figure 1: Contrast correlation with spatial frequency (c/deg). Furthermore, this shows four different contrast gratings, illustrating the combination of contrast and spatial frequency at certain points (Reddy, 2001) ... 9
Figure 2: Luminance level correlated to the highest resolvable spatial frequency (Shaler, 1937) ... 9
Figure 3: The Uncanny Valley (Tinwell, 2011) ... 15
Figure 4: Likert scale used in this experiment ... 28
Figure 5: The Workflow followed in this experiment ... 29
Figure 6: Number of participants rated "most believable clip" (low, high or both) for movement (left) and no movement (right) ... 49
Figure 7: Number of participant stating if they noticed a difference between the clips (yes or no) in (left) movement and (right) no movement condition. ... 50

16 TABLE OF PICTURES

Picture 1: (Left) Sketch and (Right) Image plane created in Photoshop ... 30
Picture 2: to the left is the UV used for the top tank and to the right is the texture applied 31
Picture 3: The wireframe and Picture 4: Ambient Occlusion ... 32
Picture 5: The final model with ambient occlusion applied. ... 32
Picture 6: HDRI used (Openfootage.net, 2010) ... 33
Picture 7: Wireframe of the high LoD UFO model ... 34
Picture 8: Wireframe of the low LoD UFO model ... 34
Picture 9: Final low LoD model ... 35
Picture 10: Ambient occlusion and Picture 11: Final high LoD model ... 35
Picture 12: The footage ... 36
Picture 13: The re-created 3D camera, the yellow and blue dots are the tracks. ... 37
Picture 14: Ground Plane and tracks ... 38
Picture 15: 3D camera and tracks imported to Maya ... 38
Picture 16: (Left) NURBS Curve and (Right) NURBS curve with joints attached ... 39
Picture 17: The NURBS curve having tracks attached to it, seen from difference angles ... 40
Picture 18: Final shape of the tank treads ... 40
Picture 19: (one) set extension, (second) the original and set extension combined (third) low LoD model and tank implemented (fourth) high LoD model and tank implemented ... 43
Picture 20: Sound being implemented in Adobe Premier Pro ... 44

17 TABLE OF TABLES

Table 1: example of what a believability response rating for each visual movement condition and level of detail could look like ... 26
Table 2: Statements and explenations ... 27
Table 3: Render settings for the different objects ... 41
Table 4: The 3 different factorial designs: within group, between group and mixed design. . 46

Table 5: (Top) Example of the movement randomization and (top) example of the no movement randomization, with a min. of 5 participants in each order. 46
Table 6: distribution of gender between conditions 47
Table 7: Average rating of participants acquaintance with VFX 48
Table 8: Overall believability response rating Ḃ 48
Table 9: Overall Ḃ with σ error bars 48
Table 10: Average rating statement 1 51
Table 11: Average rating statement 3 51