COMPUTER GRAPHICS
AND DIGITAL VISUAL EFFECTS

A Project
Presented
to the Faculty of
California State University, Chico

In Partial Fulfillment
of the Requirements for the Degree
Master of Arts
in
Interdisciplinary Studies
Applied Computer Graphics
Communication Design

by
© Russell W. Robbins 2014
Spring 2014
COMPUTER GRAPHICS
AND DIGITAL VISUAL EFFECTS

A Project

by

Russell W. Robbins

Spring 2014

APPROVED BY THE DEAN OF THE GRADUATE SCHOOL
AND VICE PROVOST FOR RESEARCH:

_________________________________
Eun K. Park, Ph.D.

APPROVED BY THE GRADUATE ADVISORY COMMITTEE:

__________________________
Sharon A. Barrios, Ph.D.
Graduate Coordinator

__________________________
Clarke H. Steinback, Ph.D., Chair

__________________________
Jennifer Meadows, Ph.D.
PUBLICATION RIGHTS

No portion of this project may be reprinted or reproduced in any manner unacceptable to the usual copyright restrictions without the written permission of the author.
TABLE OF CONTENTS

PAGE

Publication Rights ........................................................................................................ iii
List of Figures .............................................................................................................. vi
Abstract ..................................................................................................................... vii

CHAPTER

I. Introduction ................................................................................................................. 1
   Scope of the Project................................................................................................. 2
   Significance of the Project .................................................................................... 3
   Limitations of the Project ..................................................................................... 6

II. Review of Related Literature .................................................................................. 8
   Related Media ......................................................................................................... 8
   History, Concepts and Paradigms ........................................................................ 10
   Overview of Digital Elements ................................................................................ 22
   Previsualization ...................................................................................................... 29
   Live-action Shooting .............................................................................................. 34
   Matchmoving .......................................................................................................... 40
   Traveling Mattes ..................................................................................................... 42
   Rotoscopying .......................................................................................................... 47
   Animation ................................................................................................................ 49
   Environment Creation ............................................................................................. 59
   Light Matching ....................................................................................................... 62
   Rendering ................................................................................................................ 80
   Compositing ............................................................................................................. 82

III. Methodology .......................................................................................................... 87
   Plan ......................................................................................................................... 87
   Previsualization ...................................................................................................... 90
   Shooting .................................................................................................................. 92
<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matchmoving</td>
<td>95</td>
</tr>
<tr>
<td>Models</td>
<td>97</td>
</tr>
<tr>
<td>Backdrop Creation</td>
<td>98</td>
</tr>
<tr>
<td>Pulling the Matte</td>
<td>104</td>
</tr>
<tr>
<td>Animation</td>
<td>107</td>
</tr>
<tr>
<td>Lighting</td>
<td>110</td>
</tr>
<tr>
<td>A Description of Passes Rendered</td>
<td>118</td>
</tr>
<tr>
<td>Compositing</td>
<td>131</td>
</tr>
<tr>
<td>Creating the Application</td>
<td>137</td>
</tr>
<tr>
<td>Results</td>
<td>141</td>
</tr>
<tr>
<td>IV. Summary, Conclusions, and Recommendations</td>
<td>146</td>
</tr>
<tr>
<td>Recommendations</td>
<td>148</td>
</tr>
<tr>
<td>Summary</td>
<td>152</td>
</tr>
<tr>
<td>References</td>
<td>154</td>
</tr>
<tr>
<td>Appendices</td>
<td></td>
</tr>
<tr>
<td>A. Online Survey</td>
<td>167</td>
</tr>
<tr>
<td>B. Definition of Terms</td>
<td>171</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.  An Early Storyboard</td>
<td>89</td>
</tr>
<tr>
<td>2.  Previsualization in 3D Software</td>
<td>92</td>
</tr>
<tr>
<td>3.  A Set Schematic</td>
<td>93</td>
</tr>
<tr>
<td>4.  Constructed Set</td>
<td>94</td>
</tr>
<tr>
<td>5.  Early Light Matching Test</td>
<td>113</td>
</tr>
<tr>
<td>6.  3D Elements and their Passes</td>
<td>120</td>
</tr>
<tr>
<td>7.  Double Premultiplication Problem</td>
<td>133</td>
</tr>
<tr>
<td>8.  Pre-rendered Alpha Sequence</td>
<td>136</td>
</tr>
<tr>
<td>9.  Organizational DVD Flowchart</td>
<td>140</td>
</tr>
<tr>
<td>10. Online Survey Results</td>
<td>144</td>
</tr>
</tbody>
</table>
Because of the competitive nature of the visual effects industry for film, digital artists can expect a high demand for technical expertise as well as a demand for efficient practices. Artists’ works are required to satisfy an industrial need for both aesthetics and cost-effectiveness. Since there are so many different topics that well-rounded artists need to master, it is fitting that practical, critical thinking should be encouraged about the efficient use of such disciplines.

For this project, extensive research was conducted of ten major, industrial principles that are used to add digital effects to live-action footage. This research, in turn, was applied in the creation of a video clip that combined the many concepts outlined in
the research. The video clip was rendered in 24 different ways so that distinctions could be drawn between variations of techniques. The 24 renders were then incorporated into a comparative computer application. This application encourages critical thinking from users by asking them to choose the most efficient combination of techniques, much the same way that a visual effects supervisor might do in an industrial setting.

It is hoped that the outcome of the project will yield a formation of opinion amongst users about the kinds of methods that are most efficient in an industrial context. This opinion could improve the way beginning visual effects artists approach future projects. An online survey has shown that the program has helped clear up preconceptions about these techniques and solidify opinions regarding the techniques.
CHAPTER I

INTRODUCTION

Advances in the technology of the visual effects industry have inaugurated a new level of production quality and speed. Visual effects are elements, captured optically or created digitally, that are added to live-action footage in an environment separate from the camera. These days, computer technology has allowed for more visual effects—in digital form—to be added to films in a much shorter time than ever before. Motion picture companies now demand their artists to efficiently produce more effects at a higher level of quality than in years past. For an artist, efficiency is paramount to retaining a job in the effects industry. A variety of techniques for matching CG, or computer graphics, elements with live-action elements has been discovered by the industry. The artist who can accomplish these tasks in less time has a better chance for success in this field (Dunlop, Malcolm & Roth, 2008).

There are relatively few sources of information that provide clear, side-by-side comparisons of common visual effects methods; nonexistent are utilities which can allow beginning visual effects artists to quickly sample various methods and view their results. A computer application that offers a comparative analysis of integration techniques can be used as a tool to help beginning artists compare, contrast, and decide for themselves which methods are best for use in the industry.
Scope of the Project

This software would be best deployed in an environment of learning. Presented as a class supplement or industrial orientation tool, this program is meant to address concerns of students of the visual effects field. A student may be someone who wants to produce independent films with self-created visual effects; or, a student may be a person just beginning a career in the visual effects field and needs insight into the decision-making process.

The project is presented as a computer program that acts as a decision list generator for the creation of specific visual effects. It has an introduction that provides directions on the purpose and operation of the program by presenting a fictional scenario from the visual effects industry. The scenario is designed to draw users into the program so they may feel entertained and compelled to complete the scenario. Users are offered a mission to complete a single, complex shot for a movie by adding specific visual effects to it. Users can view videos in the program to inform them of the advantages and disadvantages of each choice. In each category, participants are prompted to select a method that will provide the highest quality for the best price. A budget is set for participants to try to stay within. Users view images of each method to see how they compare with one another, regarding production time and quality. Once the choices are complete, the program tabulates the choices and displays the correct, completed video that reflects the decisions made by the user. Viewers are then invited to see a chart of
popular decisions made by their peers. The chart provides feedback so users can compare their decisions to those human test subjects who have contributed to the data.

Significance of the Project

The purpose of this project is to investigate, analyze and share the process of adding visual effects to live-action video. The research obtained is then used to synthesize a tool to encourage critical thinking about digital visual effects, its methods, and the industry ideal of balancing quality and efficiency. Finally, the project will recommend specific solutions and general guidelines for integration in future projects.

The project is important because it does not simply recite information to learners, like other forms of presentation do. Instead, it allows users to immediately apply information and let it inform decisions that can possibly be made in a real-world setting. With this tool, participants instantly put themselves into the shoes of visual effects producers and directors. Users can quickly see several major integration techniques in juxtaposition, and then see the results of selecting one over the other. Such an experience provides valuable insight into the thinking of people in the industry and can nurture a better understanding of production decisions.

There are many sub-concepts that compose the multilayered subject of digital effects. In any given shot, dozens of separate elements may be filmed, animated, digitally painted, scanned, or algorithmically simulated and placed into a composition. As well, the generated elements may need to be processed to seamlessly combine with any live-action video that will be part of the shot. This concept, known as plate-matching,
may involve the color-correction, motion tracking, light-matching, and the blending of any specific element that must be combined with another.

Most of these major concepts can be achieved on a personal computer with readily-accessible programs. For each of these concepts, more than one technique may be employed to realize the full potential of the concept. It is these techniques that will be explored. Several concepts that have two or more different techniques for accomplishing the same goal will be evaluated in this project. Each technique within the major concepts will be tested to gauge the time involved and how effective the technique is for maintaining the desired illusion.

Such an exploration will result in the production of a computer application that explores several major functions of digital effects and provides a comparative analysis of two or more techniques in that can be used to create the same effect. The application will illustrate advantages and disadvantages of specific effects techniques. This program will help beginning and amateur digital effects artists understand industry-standard concepts in combining computer graphics with live action video. The program will also assist artists in making decisions to maximize efficiency in their effects workflow. This will be done by providing first-hand results of various decisions in the creation of a low-budget “effects” shot.

The computer program, entitled Soccer Mom: VFX Supervisor for a Day, will be used to:

- Gather data about popular choices via human test subjects.
- Offer information on each proposed choice in the form of video “e-mails” that the user watches.
• Share the production findings of each element to be tested, as reported by fictional vendors, staff members, and hourly figures attached to each choice.

• Provide comparisons of the user’s choice with the choices of other users.

The presentation format makes the project unique. The software offers an immersive role-playing experience that engages users. The program prompts players to take the information they have learned in the included videos and immediately make decisions with it, reinforcing those concepts as users quickly see the results.

Also unique in this application is its broad, central range of subjects that are brought together in one venue for analysis. Exploration is conducted not just by testing the techniques of one concept, like chroma screen color, but also by the combining of all four interactive concepts (chroma keying, light matching, shadows, and backdrop creation) as a single unit. It is this breadth of scope that best simulates an actual working environment as described in the literary review section of this paper.

It is hoped that a result of the project will be the stimulation of opinion for users about the kinds of methods that are most efficient in the visual effects industry. It is possible that users may be made to be more excited about integration techniques that they previously were. The act of role-playing the part of the visual effects supervisor (or VFX supervisor, as is often used in the industry) may intrigue viewers and create a realization that such techniques are possible for them to learn, and that knowledge about those techniques could enable them to assume the role as director or VFX supervisor one day. It is also possible that users may lose the desire to pursue further learning in integration techniques. There may be a realization that more discipline is necessary to achieve good results than previously imagined. Either way, this program provides a way to “place”
users into an industrial scenario before a significant amount of time and money is spent on training and software.

Limitations of the project

The techniques employed in the content of the project needed to be accomplished in a realistic, achievable way. Since there were neither the funds, manpower, nor the expertise that large visual effects companies have, testing concepts were customized to those that a small, boutique-style shop might use. Boutique shops often use off-the-shelf software and basic studio equipment to accomplish their goals (Cohen, 2005). At first, many more tests were proposed that could have been possible with additional funds, equipment, facilities, software, and time. It was necessary to trim the scope of the tests to allow the maximum value to fit into a modest budget and into a reasonable timetable. Several other factors also influenced the decision to limit the tests. The software interface, to be created in Flash, needed to be detailed enough without being too large. Limiting the total size of the program to 9GB was a good way to keep the application portable. It was possible to offer four major concepts for analysis with 24 distinct video clips to reflect the choices that users could make.

Each concept was carefully tailored to match the resources that were available. A relatively inexpensive set of chroma-key curtains were purchased to provide a temporary backdrop on the shooting set. Hand tools, plywood, and PVC pipe provided the materials needed for construction of green and blue screen panels. Family and friends helped assemble the set and had important roles to play in the actual production. Software
that was already purchased was available to create the other graphic elements and integrate the composite.

Definition of Terms

Many terms that may be foreign to readers appear in this text. Important terms that have not been explained in the body of the text are further described in the “definition of terms” heading of Appendix B.
CHAPTER II

REVIEW OF RELATED LITERATURE

To know how this project relates to existing works, it is appropriate to examine media that are similar to “Soccer Mom: VFX Supervisor for a Day.” The comparative application is part role-playing video game, part behind-the-scenes documentary, and part educational computer program. Outlined below are examples of media that are similar to this project.

Related Media

The role-playing aspect of this program is similar to video games that require players to immerse themselves into a role and obey the limitations of that role. *Wing Commander IV*, a 1996 space simulator and role-playing video game, was an influence over format. Players often received video messages from characters and were prompted to take action (“Wing Commander IV,” 1996). This program does the same. Next to each choice in the interface menu, there is an optional, clickable button that provides detailed information in the form of a “video e-mail” from a fictional director. These video clips were produced with the purpose of reinforcing directions, quickly summarizing the involved techniques, and supplying perspective about each choice. The videos also help preserve the illusion that users are still actively playing a role in the program.
The aforementioned video e-mails are also what makes this project similar to an educational computer program. These video clips provide helpful information that prompts users for action. This kind of exercise is unique but is somewhat similar to the *Hollywood Camera Work* video series, in that the scope of concepts is broad and addresses many of the same issues as this project addresses (“Hollywood Camera Work,” 2009).

Another aspect of the project emulates the behind-the-scenes documentary style of popular collectable movie DVDs. Documentaries on the *Star Wars: Revenge of the Sith* bonus section offer a “featurette” entitled *Within a Minute: the Making of Episode III*. While the purpose of such documentaries is to increase DVD sales for collectors, the work still engages beginning VFX artists. The project seeks to do the same thing, except that the initial purpose is to encourage critical thinking about techniques. In the program, users receive actual “behind the scenes” information of the video production featured in the application. This information is disguised as the results of the role-playing choices that users of the program make.

Not only is it worthwhile to examine related media for this project, but it is also important to review the literature that informed each of the employed techniques. It is the literature that contributed the most to this endeavor and is described fully in this text.

The content of the project encompasses a multitude of techniques, creating 24 distinct works that shall be compared side-by-side. The research that went into the techniques addresses the major techniques used by boutique-style visual effects facilities
as outlined by the limitations section of the introduction. Some techniques for creating assets are common through all 24 works while other techniques are presented as choices. Ten major concepts were researched to create the project, so it is fitting that all ten concepts will be reported upon herein. The information presented here came from specialty books, technical journals, white papers, magazine articles, and documentary films. These concepts will also be used to illustrate the changing paradigms of the visual effects industry and demonstrate the need for efficiency.

History, Concepts and Paradigms

As long as there has been cinema, there have been visual effects, and there have been visual effects artists to help movie directors realize their storytelling visions. As cinema developed through the advancement of technology, visual effects artists also had to advance their own knowledge and skills. With each passing discovery in the realm of cinematic illusion, artists have found that those discoveries made their jobs easier and harder simultaneously. Just as early artists found themselves working hard to master the technology of optical effects, today’s visual effects artists work even harder to compete in the expanding field of digital effects. There are many fields of visual effects for cinema that artists and supervisors can participate in. Indeed, it behooves any visual effects person to have a wide knowledge of each of the major concepts.

The changes that visual effects for cinema have undergone also have been slow, but steady. Starting at the dawn of the twentieth century, visual effects were built from simple film camera tricks to complex projection techniques, achieving a dynamic, yet stable production model that has endured for the last 50 years.
(Dunlop, et al, 2008). This foundation eventually led the way to digitally-based filmmaking, which represents one of several paradigm shifts in the long history of visual effects.

As will be shown, three major ways of thinking have developed in the visual effects industry that reflect how technological advancements have changed the experience, skill set and expectations of visual effects artists and supervisors.

The first major paradigm involved using the camera to compose special effects. Since the invention of celluloid film in 1889 by Henry Reichenbach (Alfred, 2011), the scene was set for a rapid series of camera inventions that allowed for motion pictures to be made. By 1895, Thomas Edison and W.K.L. Dickson invented the Kinetograph and Kinetoscope for the filming and viewing of short “peep show” style films (Rickitt, 2000). These very short film clips were viewed through slots, a feature soon improved upon by the Lumiére Brothers’ Cinematograph that same year (Rickitt, 2000). The Cinematograph was the complete package of the day, serving as camera, developer and projector. Both inventions, supported by the advancements of other innovators, quickly captured the public’s imagination and encouraged others to improve upon cinema technology.

Even then, the nascent field of cinema had special effects. The first known special effects shot was made in 1895 with Edison’s Kinetograph and featured a stop-action beheading in Mary, Queen of Scots (Rickitt, 2000). Although the development of “trick shots” during that time was slow, pioneers in filmmaking made their mark on the industry. In those days, illusions of this type were called “special effects,” and
involved the use of only the camera as the means of creating the effect. Later, as will be shown, techniques for creating effects would make use of tools other than the camera. These techniques would later be referred to as “visual effects.”

One of the first filmmakers to truly realize the potential of special effects for film was theater magician George Méliès. After purchasing and modifying his own camera system, Méliès took special effects to a new level. Trailblazing key concepts like stop-action, double exposures, variable camera speeds, dissolves, forced perspective, and the split-screen process, Méliès created scores of films with effects by manipulating everything that appeared in front of the camera. These types of effects, called “in-camera” effects, represent the way special effects were made before the mid-1920s (Rickitt, 2000).

As Méliès used these techniques in *Indian Rubber Head* and *A Trip to the Moon* in 1902, others caught on and used the same concepts (Rickitt, 2000). In 1903, Edwin S. Porter produced *the Great Train Robbery*, a short film that used many of the same techniques as Méliès, but in a less obvious way. Porter used effects as tools to help the story along, using subtle techniques to “ground” the setting (Rickitt, 2000). In one such scene, a double-exposed train passes outside a window to help establish the setting.

Soon, other ways of producing content appeared. Animation came onto the scene in 1906 with Stuart Blackton’s *Humorous Phases of Funny Faces*, allowing drawings to come to life and setting the stage for a whole genre of entertainment and films (Rickitt, 2000). The potential for special effects, using animation as a source,
would be realized in years to come in stop-motion films like *Lost World* in 1925 and *King Kong* in 1936. Stop-motion, where miniature models of monsters or creatures are animated frame-by-frame, would be pioneered by Willis O’Brien and Ray Harryhausen over the next fifty years (Rickitt, 2000).

Other developments of in-camera techniques involved placing live actors into scenes that were impractical or impossible to create or gain access to. The first of these techniques was called the “glass shot” where artists like Norman O. Dawn, who developed the method in 1907, painted scenery onto glass while production crews waited. The glass was positioned in front of the camera so the painted scenery aligned with real background elements, forming a seamless composite. Dawn later expanded the technique to allow the film to be rewound so paintings could be added at a later date (Rickitt, 2000).

Eugene Shuftan, in 1926, created a related method for combining models or paintings with live action by using a technique later known as the Shuftan process. During filming of Fritz Lang’s *Metropolis*, Shuftan would use a mirror, rotated to a forty-five degree angle to the camera, so the model’s or painting’s reflection could be seen from the camera. The reflective backing was then scraped away in the portions of the image where live action elements would appear. Miniatures and performers could be filmed at the same time, much like with the Dawn process (Anderson, 1998).

One might say that a second paradigm solidified, gradually, as new tools for combining elements were developed. The bipack camera, rear screen projection, optical printer, rotoscope, and the traveling matte all helped create an atmosphere that
allowed visual effects artists to plan and shoot elements in stages. No longer were filmmakers forced to film all elements at once, on the same negative. This greatly reduced the risk of ruining the original negative with problem elements or misregistration (Rickitt, 2000).

The bipack camera was a device that helped reduce risk. Created in the 1920s, this camera had the capability of loading a pre-developed roll of film and creating a second roll that essentially copied the first roll, with changes. Special black and white silhouettes, called mattes, were painted on glass to restrict the exposure of the second roll of film. The film was copied with the silhouettes obscuring part of the original, leaving the obscured portions unexposed. The copy was then rewound and a reverse silhouette created, allowing the live action to be filmed onto the unexposed portion of the copy. This method allowed more flexibility for directors, who could take pre-developed film and design a matte shot around it instead of having to plan both shots to great detail (Rickitt, 2000).

One disadvantage to using the bipack method was that it produced a lower-quality, second generation image of the copied footage. This problem is the same problem that the rear-projection technique has. Used for both matte painting shots and for outdoor simulation, rear projection allowed pre-shot footage to be projected behind actors. Attempted as early as 1913, rear projection developed slowly but enjoyed great success from the 1930s, on. Even though a loss in generation led to a slight loss in quality, improved projector technology brightened backdrops and produced nearly seamless composites (Rickitt, 2000). Rear projection, now improved
in digital form, is still used in film and television (Failes, 2013). Its capability of rendering color-corrected, in-camera composites are more cost-effective than other compositing measures, although digital projectors, like other rear-projection methods, have quality limits (Taylor, 2007).

Another method for optical compositing is the traveling matte. In the days of Meliès, parts of the camera were simply “masked off” when two elements needed to be brought together. This worked reasonably well as long as the added element did not overlap the first element. Careful masks had to be made to insure that this did not occur; otherwise, the effect could be ruined. What was needed was a way to extract a moving element—like an actor—and place that element into a scene with a different background (Rickitt, 2000). Color processes and treatments were developed to allow this to take place.

An early version of the traveling matte method was used in bipack cameras. Later, in the 1930s, a special compositing tool called the optical printer was developed by Linwood Dunn and implemented into the moviemaking workflow (Rickitt, 2000). Matte combinations could be made away from the movie set and composited in a post-production setting. Optical transitions, split-screen effects, titles, and complex matte shots could be created without reloading film into bipack cameras. The optical printer expanded the amount of control possible with compositing and improved the overall image quality of effects (Rickitt, 2000).

Some devices, designed to help filmmakers create a certain type of production, were later adapted to further the cause of special effects. One such tool
was the rotoscope, invented in 1917 by Max Fleischer to help animators create realistic movement in their characters. The machine also allowed visual effects artists to create their own hand-drawn traveling mattes. Although slow, rotoscoping had the advantage of creating very precise mattes when other processing methods were impractical or impossible (Rickitt, 2000).

Traveling mattes, miniatures, projection, and varied processes evolved and improved through the years as research and funds were committed to each film. As each advancement changed cinema, whether in profound or minor ways, the special effects companies also changed over time. Even though the optical framework did not change much in 50 years (Dunlop, et al, 2008), visual effects companies adapted to changing paradigms and conditions all the way up to the digital revolution.

Once visual effects for film were well-established in the early years of cinema, it became clear that more than just one person was needed to create quality illusions for movies. Movie studios sprang into being and larger numbers of special effects were called for. An assembly-line approach to effects developed that placed the processing of shots in the final phase of production. Separate departments were created that removed the burden of effects from the cinematographer’s shoulders (Dunlop, et al, 2008).

All through the technology advances of “talkies”, the development of color film, and the onset of television, visual effects were created by large studios like RKO, Paramount, Warner Brothers, 20th Century Fox, and MGM. These studios had the money to finance their own visual effects departments, where artists and
technicians enjoyed stable careers and worked with the same co-workers year to year (Dunlop, et al, 2008).

Eventually, the economy of the 1960s took its toll on the movie industry, leading to a change of approach to visual effects. In an effort to win audiences back from television, Hollywood created more spectacular films with elements that were more reality-based. Crashes with real vehicles and location shooting became the norm, largely eliminating the need for traditional visual effects artists (Rickitt, 2000).

The 1970s brought a slow buildup of “disaster” movies like Airport and the Towering Inferno, which found a new niche in entertainment. The required special effects helped keep artists working (Rickitt, 2000). Finally, in 1977, a breakthrough occurred; two major films, George Lucas’ Star Wars and Steven Spielberg’s Close Encounters of the Third Kind ushered in a new era for optical visual effects. The sheer number of effects in Star Wars alone required the creation of a specialized company just for visual effects. Lucas’ Industrial Light and Magic, the first facility of its kind, has set the standard for quality visual effects in cinema (Rickitt, 2000). Soon, a core of “effects houses” sprang up to supply the large studios’ needs for optical effects; Robert Abel & Associates, Entertainment Effects Group, and Apogee Productions were some of the many companies that worked to sate the world’s expanding special effects appetite.

Even with the creation of computerized camera motion-control techniques, the basic methods of creating visual effects remained the same until the digital age. Through the 1970s and 1980s, cinema had seen a slow development of the computer
as a tool for aiding in production. Motion control and primitive, wireframe computer displays represented the limit of the computer’s role in movies until the early eighties. Then, *TRON*, *Star Trek II: the Wrath of Khan*, and *the Last Starfighter* all featured sequences that were created with raster-based computer graphics that showed much greater texture and detail than in the previous decade (Dirks, n.d.).

It wasn’t long before the first tentative steps were made to scan footage from film, create special effects within the computer, and export the results back out to film again. In 1988, Mike Boudry and his team at Computer Film Company pioneered the technology to produce *the Fruit Machine*, a film where a man changes into a dolphin (Rickitt, 2000).

Digital visual effects, as they are now known, were fully realized with James Cameron’s *the Abyss* in 1989 and *Terminator 2: Judgment Day* in 1991 (Ling, 2003). These films reflected a leap of faith into the realm of computer-generated effects that paid off. They showed that computer graphics were feasible in creating compelling, realistic effects that could be achieved in no other way. In 1993, Steven Spielberg’s *Jurassic Park*, a landmark film featuring computer-generated dinosaurs, solidified digital effects as a powerful tool in the modern age (Rickitt, 2000).

These films, and others, paved the way to the prevailing paradigm of the early twenty-first century. This new century has witnessed a change in production mindset, brought about by the digitization of the filmmaking workflow. Now, developments in camera technology have allowed entire films to be shot digitally, without the need for the scanning of film stock (Harris, 2002). This new “filmless”
media lets directors instantly review the current shot and mix it with other digital elements without even leaving the set (Harris, 2002). The new paradigm relies on the advancement of technology and gives filmmakers a nearly unlimited range of possibility. Nearly any effect that a storyteller can imagine is now possible to achieve digitally (Rickitt, 2000). The digital workflow has also reduced the time it takes to complete special effects (Kaufman, Nov. 2006). Digital formats allow footage to be placed immediately into post-production without waiting for film to be developed. As result, more shots can be completed in less time (Cohen, 2007). This has led to a notion among some filmmakers that digital effects are “easy” and any shooting mistakes can be “fixed in post” (Delicata, 2008).

Three main approaches to visual effects have emerged since the advent of digital technology in filmmaking. The first is a practical approach with only basic digital effects, like wire removal and cleanup (Cohen, 2007). The second is a synthesis of the old and new techniques; to avoid the flat “CG look” a combination of live action, miniatures and CG is used (Cohen, 2007). The third method is the wholesale employment of digital scenery in a “digital backlot” approach, where nearly the only elements shot practically are the actors on a green screen stage (Kaufman, 2006, May 1). Movies like 300, Sky Captain, World of Tomorrow, Speed Racer and the Star Wars prequels are examples of digital backlots.

Because digital technology has streamlined the moviemaking workflow, moviemakers demand higher-quality visual effects that are made more quickly and in greater frequency (Cohen, 2007). This need is so great that studios and large effects
companies have been unable to supply all the shots needed for films, especially for major productions (Cohen, 2007). To compensate, producers will enlist the help of smaller, “boutique” visual effects companies to fill in the gaps. Boutiques generally have a small staff and use inexpensive software to do a wide range of effects (Skweres, 2003).

Such a staff should have, depending on the size of the boutique, artists who have a strong aesthetic sense and also have a strong grasp of the digital tools being used (Dunlop, et al, 2008). Indeed, artists need to be multidisciplined and embody a prolific range of skills (“The Core Skills of VFX”, 2011).

The boutique model is more relevant in this study because they exemplify the paradigm of “faster and cheaper”. By no coincidence, the boutique model has been chosen in this text as the template in which the subject matter of this project has been produced. This project is being implemented at such a time in history where the available resources are also capable of producing quality content that can, conceivably, compete with larger studios (Cohen, 2005).

There are two paths that a visual effects artist can follow. The first path leads to large effects companies and requires a narrower concentration of skills and specialization to ensure ongoing productivity. The second path, the one that may lead to a boutique environment, mandates a more generalized skillset; artists need a much wider range of functionality to contribute in small visual effects shops (“Sony Pictures IPAX,” n.d.).
Boutiques also use mostly commercially-available programs to create their effects instead of costly motion-control, motion capture, miniatures or labor-intensive coding. The main differences between large-to-midsized effects facilities and the boutique company is that the staffs are generally much smaller, the operating expenses are much lower, no proprietary software or coding is developed, and the required skill set for artists is much broader. Large effects companies have large staffs where each artist does a narrow range of tasks. On the other hand, if an artist prefers to work for a smaller boutique shop, the spectrum of tasks performed encompasses a wider gamut. The boutique artist may need to have skills that range from 3D modeling all the way to computer animation and even compositing (Masters, 2013).

As a prerequisite for employment in the industry, many companies, large or small, demand a certain degree of training and education from artists, even from beginning artists. It is true that many large facilities have their own training centers, like Imageworks’ PAX program (“Sony Pictures IPAX,” n.d.). and Industrial Light and Magic’s Studio Talent Group (“Lucasfilm Recruiting,” n.d.). Conversely, boutiques usually do not offer training. The training becomes, largely, the responsibility of the artist. As such, there are many avenues for education and training. For those who can afford it, art schools like Gnomon 3D, Full Sail, and Savannah College offer comprehensive training from industry professionals and help students build a strong portfolio or “demo reel” (Massion, 2008).
For those who cannot afford expensive art schools, other means are available for training. State and community colleges, workshops, specialty books, and free online tutorials are all viable means to increase a student’s skill in the area of visual effects.

According to research, there are at least ten major skills that can be expressed with the boutique model of production. Each concept can be demonstrated with consumer-level equipment and software and has already been shown to be competitive with the larger studio models (Hetherington, 2006). As well, each discipline can be learned from available books, DVDs, online tutorials, and classroom settings. The major concepts include: previsualization, set preparation, matchmoving, traveling mattes, rotoscoping, animation, environment creation, light matching, rendering and compositing.

Overview of Digital Elements

The aforementioned concepts are means of recording, creating and manipulating electronic images that are managed with a computer system. Because each idea can be used to create separate digital elements, it is important to know what these elements actually are and what comprises them. Whichever of the ten methods were used to acquire them, all digital elements can be understood by students in terms of pixels, channels, image files, resolution, and finally, the types of elements that are to be manipulated in the above terms.

Pixels are the foundation of electronic images. The word “pixel” is actually a joining of “picture” and “element” (Wright, 2006). Digital images are
constructed of an array of points that are laid out in a grid pattern. These pixels each have their own sets of information about brightness and color (Brinkmann, 2008).

The color of each pixel can be described as existing in layers, or channels. There are three channels that make up a pixel: red, green and blue. It is a combination of these colors at varying intensities that give each pixel its appearance and allow an image to have a full range of color. Channels of an entire image can be manipulated in a wholesale manner; separation of an entire color channel gives greater control over the final image (Brinkmann, 2008).

In addition to the red, green and blue channels, also known as the RGB channels, there can be a fourth channel that describes the opacity of what is depicted. This channel is called an alpha channel. Commonly used in CG images, alpha channels store information about the opacity of certain areas of the image. Developed in 1977 by Ed Catmull and Alvy Ray Smith, the embedded alpha channel was designed to be included in the rendering process so an outline could be used by a compositing program to layer a rendered object over a preexisting background element (Smith, 1995). The rendering process would later be adapted, in 1984, to include generating an alpha channel and then “premultiplying” or premixing the color information with the alpha channel, to prepare the image for compositing and thereby saving time in compositing (Porter & Duff, 1984).

The pixels and channels of digital images are stored, of course, in an image file. There are many types of image files that can be manipulated, depending upon their usage. A digital movie clip may be stored as a string of still images, where
each frame of the clip is saved into separate, numbered files called an image sequence. Among others, popular image formats include TIFF, TARGA, SGI (Mitchell, 2004), and the more recent OpenEXR (Brinkmann, 2008).

The size of each image file impacts the cost in computer resources that it takes to manipulate the image. The larger the file size, the more memory is needed to perform the tasks required. If images are unnecessarily large, a computer’s memory resources could be taxed and the process slowed. As well, images that are too small may not carry the level of detail that a digital artist needs to complete the work.

A file’s size depends upon its resolution. Resolution is defined as the amount of data that makes up an image (Brinkmann, 2008). Resolution can describe the number of pixels wide the image is and how many pixels high it is. In addition, an image’s pixels are designed to hold discrete bits of data to represent colors. Images whose pixels contain more bits for color are said to have a greater bit depth, or color resolution (Brinkmann, 2008).

Many common image types typically support only 8 bits per channel. Each of those bits are represented as whole-number integers and provide a decent number of colors, but are not suitable for cinema (Brinkmann, 2008). Each channel of an 8-bit format image operates in an intensity scale and is expressed as integers from 0 to 255 (Wright, 2006). Red, for example, is expressed (255,0,0) because the red channel’s intensity is at the full 255 level, while the green and blue channels are at the zero level.
Images with higher bit depth per channel have smoother gradations between colors and generally look more realistic. Higher-end formats exist that represent images with 16-integer or 32-integer bits per channel for a very high bit depth. Such file types take up much more storage space but allow a higher flexibility when artists implement color corrections (Wright, 2006). Color corrections are known to create undesirable artifacts in lower bit-depth images.

Other high-end formats work on a floating-point system and provide an even greater range of color. 16- and 32-bit “floating” point formats offer a much wider gamut than integer formats because floating values are not discrete integers. Floating formats possess colors with intensities described in decimals between zero and one. A bluish green may be expressed as (0, 0.6, 0.5). These decimals can be refined as needed to best describe the color at hand. A pixel with the values of (.95, 0.553, 0.3989) could exist. Not only are these colors more precise than integer colors, but they also include values that can exceed 1.0 or even go into negative numbers. This feature is valuable for creating images that reveal detail at varying exposure levels (Wright, 2006).

Images in feature films have at least 16 bits per channel and frequently go even higher (Brinkmann, 2008). One of the most popular formats used in VFX today is the OpenEXR format which features a 32 bit-per-channel, floating point data type. Images with 32 bits per channel and floating point values are used at Industrial Light and Magic, but 16-bit floating images are also used frequently in the industry (“Industrial Light and Magic,” n.d.). 16 bit Open EXR files are known as HALF files.
Because images with greater bit depths cause increases in file size and computing time, float data is not used as often as integer data; however, the 16-bit Open EXR file type has allowed a greater color precision than integer files while keeping the file size the same as 16-bit integer files (Wright, 2006).

In digital images, the pixels of each channel are arranged into image files with a specific spatial and color resolution. The pixels, channels and files can be arranged to form an element, which is thought of as a piece of a larger composite that is made of several, even dozens, of elements made from varying sources. There are four major types of elements: digitized live-action footage, flat digital paintings or photographs (known as 2-D or two-dimensional elements), 3D graphic elements, and 2 ½ D elements.

First, live-action video is recorded. The foundation of most visual effects shots is the combining of footage of real people with electronic imagery of some kind. In digital effects, all elements have to be captured digitally or converted into a digital form that can be manipulated with a computer. For some productions, this means using a film scanner to capture the frames of live-action footage that was originally created with film (Rickitt, 2000). For others, live action footage can be obtained in a native, electronic format by shooting the scene with digital video cameras. While today’s cameras do not yet produce the full dynamic color range of film, technology is rapidly improving and just beginning to equal the resolution and shadow detail that film provides (Sudhakaran, 2013). The flexibility that digital
cameras offer encourages more and more films to be shot entirely as digital video (Harris, 2002).

Next, 2D digital paintings are a regular source of digital content. This type of element generally is used for static environmental backdrops, set extensions, or texture maps for 3D objects (Rickitt, 2000). Digital paintings are seldom created from a paint program’s native tools “from scratch”, but usually are made from on-set photos as a base. Features can be easily copied and manipulated, layer by layer, while other paint tools are used to augment those basic elements (Rickitt, 2000).

Backdrops are usually created at a very high resolution; 2k (2,048 pixels X 1,556 pixels) is the minimum resolution for feature film backdrops, while some productions have created backdrops as high as 20k. These extra-high resolution backdrops are made to add flexibility, in the event that a director may request unexpected panning or zooming (“Tips and Tricks,” 2006). Set extensions involve modifying existing live-action footage and adding elements to fit the scene. A set extension might include adding a haunted mansion to a grassy hillside where none existed originally, or perhaps to add another floor to a single-story building. Shots with digital matte paintings, either as background elements or set extensions, usually are produced with “locked off” camera footage, where the camera does not move. However, set extensions are being used increasingly with footage from a moving camera, made possible through motion-tracking technology (Rickitt, 2000).

Thirdly, 3D graphics programs build elements that are difficult or impossible to create with conventional means. 3D models are built in virtual space,
textured and placed before virtual cameras and 3D lights and then animated. The result is rendered for compositing as high-resolution files like PNG or TARGA, but the versatile Open EXR format has also been used extensively in the industry (Brinkmann, 2008). 3D graphics have been used to create all manner of elements, including fantastic objects like spaceships, castles and robots; as well, accurate historical representations have been made, like the Titanic and extinct animals. 3D simulations are used frequently to create fire, smoke and water. Creatures of countless types have been built and animated, even humans, entirely in a virtual 3D environment.

The downside of creating elements in a 3D environment is the computational cost. Productions can slow dramatically while time is spent rendering complex 3D scenes. As an alternative, some elements can be created for visual effects by the means of “image based rendering” or IBR. Otherwise known as “2 ½ D,” IBR is, essentially, a combination of photography, simple 3D geometry, and 3D animation (Huston, 2001). Artists use the 2 ½ D method to load images into 3D virtual space and map them to primitive polygonal structures. Virtual cameras and lights are positioned as needed. These “cardboard cutouts” give the illusion of complex elements without the long render times (Huston, 2001). Characters can be replicated hundreds of times in a shot where the camera moves across the landscape. This feat is enabled by mapping the actor’s separate element onto a 3D “cutout” and judiciously placed in a 3D scene.
Now that a basic overview of digital elements has been outlined, ten core skills that create and wield those elements can be discussed. As previously stated, major fields in digital visual effects can be separated into:

1. Previsualization
2. Live-action shooting
3. Matchmoving
4. Traveling matte creation
5. Rotoscoping
6. Animation
7. Environment creation
8. Light matching
9. Rendering
10. Compositing

These ten fields were chosen for study because they are the areas that can be accomplished with a modest budget and consumer-level equipment and software. These methods are also viable in a boutique visual effects environment, which is the most feasible approach to this project.

Previsualization

No visual effect has been completed without a good plan. This concept has been true since the early days of cinema. Even as early as 1930, a special technique was used in Howard Hughes’ film *Hell’s Angels*. This technique for visualizing and
planning the film, called “storyboarding”, helped save time and production costs (Taylor, 2007).

Still in widespread use today, storyboarding involves creating a series of drawings that outline how the scene, sequence or shot will occur visually (Wellins, 2005). Alfred Hitchcock used detailed drawings to generate storyboards for action and camera movement. Such a technique helped Hitchcock communicate his plans for his films to others (Ferster, 1998). This method proved valuable for filmmaking and evolved to benefit Walt Disney Studios. Not only were storyboards created for feature films, but the boards were also filmed and soundtrack was added. The result was a rough movie called an “animatic” or “Leica reel”. Disney animators found that Leica reels gave them a more complete picture of the films than storyboarding alone (Ferster, 1998).

Director George Lucas used current technology to map out complex scenes. In 1977, Lucas used World War II footage of airplanes to cut together a rough demonstration of battling space vehicles for Star Wars (Gregoire, 2003). In Return of the Jedi, Lucas employed small video cameras and miniatures to plan the speeder bike pursuit, helping him make decisions and cut expenses (Gregoire, 2003).

Francis Ford Coppola took the preplanning of films even further. Using storyboards and soundtracks, he also added captured video tap footage from the main cameras. In doing so, Coppola funneled these assets to a special RV called the “Silverfish” near the studio sets and was able to use editing equipment to begin splicing One from the Heart together. The elements were only proxies of the real film
footage but still allowed Coppola to reorganize and make changes without having to wait for film to be developed (Ferster, 1998).

The field of visual effects benefited from the new workflow that Coppola pioneered. By the late eighties, miniature video cameras on scaled-down versions of sets helped directors visualize and plan visual effects shots (Fink, 2003). Foam-core cutouts, video, paper dolls, and action figures were used in conjunction with photos and hand-drawn set depictions to help directors explain their visions. Pantomiming also helped describe complex shots as production members moved through shots and held models of airplanes (Fink, 2003).

Eventually, computer processing power allowed directors to try 3D graphics as a planning tool for projects. Known today as previsualization or “previs”, the use of computer graphics to lay out movie scenes has been indispensable in the industry (Fink, 2003).

Previsualization with computer graphics has many uses. The first major use is for planning. It is a design tool for figuring out how a sequence plays. CG previs also helps plan how a difficult visual effect might be executed (Desowitz, 2008). It is a decision-making tool for directors for shooting and digital effects. The entire film can be evolved at the same time—since all elements are digital—using previsualization as a foundation and overlaying proxy elements and sound tracks as they are finished (Dickrueter, 2007).

The second major use of previsualization is to communicate to producers all that goes into a particular shot. This is important so key executives can have an
idea how much the visual effects of a shot are going to cost. The predictive abilities of previsualization are used this way to obtain budget approval, especially on expensive sequences (Fink, 2003).

Another important use for previsualization is for logistical planning. In 3D previs, digital versions of sets are built with a computer. These digital sets include cameras, lights and props and are built to see if everything fits in a given space (Cowell, 2009). That way, directors can see if there might be any setup problems for actors and equipment. Even camera lenses can be simulated to exactly match the look that the real camera may produce (Katz, 2003).

Not only can shots be planned ahead of time and away from the movie set, but previsualization data can also be used directly on the set. Laptop computers with the previs files are brought to the director’s side during shooting to give the director more control and confidence (Fisher, 2003). Laptops can also be used to control, through the previsualizaiton files, the motion-control camera rigs that make some complicated shots possible. Motion-control rigs allow the perfect repetition of camera moves so various elements can be filmed independently of one another. The motions generated in a previsualized, virtual camera can be exported and loaded into a motion-control system and used to control the motion of the camera.

When preparing to create previsualization for a shot, scene or sequence, visual effects artists must take into account the real-life setting that might be recreated in virtual space. An example of this might involve placing live-action elements, like actors, into a computer-generated environment. To accomplish a seamless blending,
many complex tasks must be done to prepare all elements for compositing. These tasks are best performed in the previsualization phase first to see how well the proposed effects will work ("Q&A: Daniel Gregoire," 2006).

The first step in recreating the real-world set in 3D is to visit the live-action shooting location and survey the area. The set is measured carefully to see where major features are located. Sketches of the area are made to show a top view, side elevation, and locations of lights (Aithadi, Larsen, Pescosolido, & Wegerhoff, 2006). Extensive reference photos are taken of the set, including any structures, the sky, and lighting conditions ("Q&A: Daniel Gregoire," 2006). Rough panoramas are often created from photos in the previsualization phase to simulate backgrounds (Aithadi, et al, 2006).

Next, the collected data and photos are used to create a virtual copy of the live-action set. Computer graphics are utilized to reconstruct all elements of the set, including the camera, dolly equipment, lights, reflectors, and any physical elements that exist on the set ("Q&A: Daniel Gregoire," 2006). Any green screen or blue screen walls are also built. These computer models are low-resolution versions of their real counterparts and placed in the virtual scene. A dome or part of a sphere may be created as well to map environmental photos onto, serving as an instant backdrop to show where more distant features of the set are.

Once the set is rebuilt in 3D space, the director and the previs artist “block in” actors’ movements, set pieces, and camera angles to best describe the director’s vision for the shot. After animation, previsualization artists can go so far as to
simulate many of the compositing tasks that eventually will need to be done with the real live-action footage. Simple matte extraction, marker removal, rotoscoping, motion tracking and compositing backgrounds may all be tested in the previsualization stage ("Q&A: Daniel Gregoire," 2006).

As stated earlier, previsualization renders are also useful during the shooting of the actual live components. Because the previsualization renders are digital and can be stored on a laptop, they can be placed into a digital timeline that represents the entire sequence. This non-linear timeline can inform the director about how effective the shoot is at describing the same action that is contained in the previs files. If changes are needed, the previsualization artist can make basic changes and re-insert those changes into the timeline, saving the time and trouble that it takes to move real set pieces around (Dickreuter, 2007).

If having the previsualization files on the set helps the director and crew save time and resources, how much more does having an actual visual effects supervisor (or VFX supervisor, for short) on the set help the production? In charge of the decision-making and execution of visual effects for a film ("Mara Bryan," n.d.), the VFX supervisor works with both the previsualization renders and the crew on the live-action set. Increasingly, it has been beneficial for a production to have the VFX supervisor on the set to adapt to changes that a director may make (Fisher, 2003).

Live-action Shooting

There are many factors involved in the shooting and setup that a VFX supervisor needs to keep in mind. In fact, these factors are the same whether they are
needed by large visual effects teams or by small one- or two-person visual effects boutique teams. In addition to helping the director understand the technical requirements of the shot, visual effects people on the set are there to monitor distances, light angles, camera details, screen positioning, and reference data (Fisher, 2003).

Distances on the set are important to measure for different reasons. Any real object on the set should be measured, especially if that object will interact with a CG object (Dobbert, 2005). Objects should also be measured in case they have to be reconstructed as 3D objects or as “shadow casting” objects or “camera mapping” objects, which will be described in a later section of this paper. It is also important to know the distance of real objects from the camera for purposes of “matchmoving,” a process that matches a 3D camera to the movements of the real one. Measurements can usually be made with a measuring tape, but, sometimes, tall sets and high ceilings make measuring a challenge. In this case, laser rangefinders are often used to quickly capture and update linear measurements (Aithadi, et al, 2006).

Special instruments are used, as well, to measure light angles and intensities. Inclinometers come in handy for taking angle readings on artificial lights or the sun. An inclinometer, which can be homemade from fishing line, a small weight and a pre-made angle chart template, provides the angles needed for recreating the lighting in a 3D program (Aithadi, et al, 2006). Light meters are used to test the evenness of lighting across colored screens (Fisher, 2007).
Visual effects supervisors must gather as much information about the camera as possible, including position and lens information. Camera position is important so effects artists can match the motion of the real camera to a virtual one. The height and tilt of the camera must be documented to make the jobs of artists easier. Without such information, the angle and height must be guessed at, often slowing down an artist’s workflow (Dobbert, 2005). The “film back”, or the point where light from the camera lens falls on the light-sensitive chip in the camera, must also be noted. Regarding digital video cameras, the film back is simply the size of the light-gathering chip. In conjunction with the camera’s focal length, or the distance from the lens to the film back, a film back measurement yields the total measurement of the camera’s field of view, or FOV. The numbers describing the FOV are then used to create a virtual camera that exactly matches the real camera. This data is also used to help match any movement that the real camera undergoes during the shoot (Dobbert, 2005).

After the camera measurements are complete, the VFX supervisor may ask that the cameraman shoots a few seconds of a distortion grid (Aithadi, et al, 2006). A distortion grid is a simple sheet of paper that shows a grid of horizontal and vertical lines (Dobbert, 2005). As one might guess, a distortion grid measures the amount of distortion inherent in the lens of the camera. The grid, in distorted footage, typically indicates a bending inward or outward of the lines as they travel away from the center of the picture. Lenses that have a longer focal length, as well as extreme wide-angle lenses, tend to produce the most lens distortion (Dobbert, 2005).
Distortion causes trouble both for aesthetics and for camera matchmoving. As the computer analyzes the movement of the scene, an artifact known as “squirming” can develop. The distorted reference points lead to a slipping of the tracking points (Brinkmann, 2008). If lens distortion is discovered, steps can be taken to correct the distortion of the actual footage. Algorithms can be run on the original to “undistort” the image by using the distortion grid that was filmed during the shoot (Aithadi, et al, 2006).

If there are plans to introduce CG elements and combine them with the live-action video, a light probe is a crucial tool that helps to re-create the lighting conditions that were present on shooting day. Once a video take has been completed, a visual effects technician will employ a simple chrome ball on a stick, called a light probe. The technician will simply hold the ball in the camera’s view and face the camera with it. Another technician would then take pictures of the ball. Photos of this type may be used in post-processing to sample the colors of lights, simulate reflected light in the scene, or to create specialized images that help light CG objects for compositing (Birn, 2006).

As millions of film fans know, a visual effects shot is often completed with the use of blue screens or green screens. These screens, usually made of a stiff fabric, provide a false backdrop for actors and can be removed, or “keyed out” from the footage in postproduction. This allows another background to replace the original. Since it is important for the visual effects artists to work with clean and well-lit blue
or green screen footage, steps must be taken to ensure the correct setup of the shoot (Fisher, 2007).

To prepare a colored screen for filming, production crews need to make sure that the screen is lit evenly. Lights should be positioned to light the screen with no “hot spots”, shadows, changes in color, or changes in brightness (Brinkmann, 2008). A good way of checking a screen for variation in brightness is to take luminance readings with a light meter (Fisher, 2007).

In an effort to light a screen evenly, a visual effects supervisor may find that overlighting is often a difficult problem to overcome. Even though steps may be taken to constrain the primary screen lighting from falling on actors, indirect lighting from the screens may spill into the scene. When extra spill light bleeds onto actors or props, keying software may have difficulty in producing a complete matte (Brinkmann, 2008). To mitigate green or blue spill, technicians can use backlighting on the foreground actors (Brinkmann, 2008). The screen lighting can also be underexposed slightly for less spill (Fisher, 2007). The simplest solution, however, is to make sure that the foreground elements are positioned as far from the colored backing as possible. If this is not sufficient, a reduction in the size of the screen itself might also restrict the amount of spill on the subject (Brinkmann, 2008).

The choice of which color screen to use on the set also may influence the quality of the shot. Blue screens have advantages over green, and vice versa. The most obvious factor in choosing a color of screen is the colors that the actors may be wearing. It makes sense that a costume with blue in it should not be worn in front of
a blue screen; the result may be the keying out of important body parts. Otherwise, more subtle, yet important, factors influence the choice of color of backdrops. While green may produce a smoother, more reflective result than its noisier, grainier blue counterpart, blue backing tends to favor mattes for blonde hair and is more forgiving with regard to spill in outdoor shoots (Brinkmann, 2008). Red screens are also used occasionally, but the inherent red component of human skin make such a scenario unlikely for a shoot with actors. Red screens are used mostly with miniatures or large models (Brinkmann, 2008).

Some visual effects shots call for the camera to be in motion. To make sure that any computer generated elements match the motion of the real camera, the shot must also be tracked, or analyzed to determine the camera’s motion. Many times, there are plenty of natural tracking references to allow for the computer to generate a matchmove, especially on a set that does not have much green screen. At other times, when large, flat expanses of colored backdrop limit the number of trackable features, additional markers must be placed. Without the markers, matchmove software has difficulty detecting parallax in that kind of shot. Many types of markers may be used, including green dots, white X-shaped masking tape, or even small LED lights. These markers are attached strategically to areas of the green or blue backdrop with care. Too many markers could make the job of rotoscope artists harder. Each marker must be removed by a paint program in postprocessing so that a clean key might be created.
Matchmoving

Once the principal photography of the video has been completed, video footage can now be prepared for post-processing. Often, the first step, after receiving the footage from the camera and placing it into a digital pipeline, is to process the footage and analyze the motion of the camera. Called *matchmoving*, this process produces a blank scene in 3D animation software that is populated with only a virtual camera and light—but, this camera matches exactly the motion of the real world camera used on the set. When other 3D objects are added to the scene, those objects will appear to “move” when rendered through the matchmoved camera just like with the real-world camera (Dobbert, 2005).

Generally speaking, matchmoving has been around for some time. While, at its beginnings, matchmoving involved only non-digital camera control, it still provided convincing results. Camera control, or “motion control” involves a system for allowing an exact repetition of a camera’s movements on the set. This system let filmmakers create composite shots while the camera was moving, rather than forcing them to “lock off” the camera and stop it from moving. Through the years, various techniques were employed to store the camera motion for repetition: in 1914, it was a block-and-tackle winch with markers; in 1949, it was a strip of film; In 1951, a phonograph record; in 1957, an anti-aircraft machine gun computer was modified; in 1971, a stereo tape recorder stored motion; in 1977, a microchip system was used for motion control in *Star Wars* (Rickitt, 2000). Today, motion control has evolved to fully include computer control, where a personal computer can store motions, replay
them, and use them in other applications like 3D software, where a virtual camera for 3D scenes can be created, based from the motion control data.

While motion a control system is one way to gather data for virtual cameras, software-based matchmove programs are also available—and in widespread use—for producing a virtual camera that can match the movements of a normal camera, without the use of costly motion control rigs. Matchmoving is based on the concept of photogrammetry, a process that allows the mapping of a three-dimensional space by using only photographs of that space. A vital ingredient in this process is the parallax, or perspective shift, between two or more photos that are being analyzed. In the case of matchmove software, an entire image sequence of film footage is processed for perspective shift. Modern matchmove programs can automatically pick out features to track, usually small, high-contrast points in the footage, and analyze these features for parallax shift (Dobbert, 2005).

Matchmove programs can do several functions with the data it receives from processing an image sequence. First, it can produce tracking points that help describe the 3D space that it has reconstructed from the sequence. These tracking points, if accurate, can allow 3D artists to place geometry to match certain features from the original footage. Real objects, like benches and boxes that were in the original footage, can often be represented by these tracking points. 3D geometry can be built around them to establish spatial relationships between virtual and real objects. Additionally, a virtual camera can be created to exactly match the movement
of the real camera. Translation, tilt and roll movements are constructed from the real footage.

Traveling Mattes

Matchmoving is one of those processing steps that comes early in the production pipeline. Another step that comes early is the production of the traveling matte. We know that the basis of compositing digital elements together is that each element can be viewed as a layer, as though pixels have been painted on panes of glass. These layers are stacked in order, usually from furthest away to the closest in relative distance from the camera. Each layer has visible components and transparent components. The live-action video layer is no exception to this concept. In order to view other elements that are positioned behind the video layer, certain parts of the video layer must be made transparent. This is done by creating negative areas in film or video called mattes.

Through the years, mattes have been created in many different ways. As mentioned earlier, the early days of film saw only a limited way to combine elements together—by using a simple masking method and compositing actors over large unexposed areas of film. What producers wanted, however, was a way to refine this matting method so that actors would not be restrained to a small area of movement. To do this, the matte would have to perfectly match the shape of the actor and change shape to accommodate the actor’s movements. A silhouette of the actor on film would make these traveling mattes possible.
The first widely-used method for creating traveling mattes involved a bipack camera and high-contrast film. The Williams process, invented in 1918, involved copying the foreground element, usually actors in front of a black curtain, onto high-contrast film. The copy was usually sufficient to create a decent silhouette of actors. Another copy was made, using the bipack camera to produce an image of the background material with the silhouette applied. The result was a copy of the background footage with an unexposed area where the actor was supposed to be. The film was then rewound and the original foreground footage was pressed against the copy. Reexposed to light, only the silhouette part of the copy would accept any image; the actor was copied onto the unexposed portion, completing the composite (Rickitt, 2000).

The late 1920s saw the development of the Dunning-Pomeroy method, which still made good use of the bipack camera. The difference between this method and the Williams process was that the composite was created entirely during studio shooting. The background footage, previously developed, bleached and treated with a special orange dye, was loaded into the bipack camera. The Dunning-Pomeroy method used a system for the absorption of colors to create traveling mattes. During the shooting of foreground elements, the subjects were lit with orange light and a large blue screen was brought in as a backdrop. In the camera, light from the blue screen was absorbed by the orange areas of the background footage, so it never arrived onto the film. The result was an effective black-and-white, compositied print (Rickitt, 2000).
Although effective for black-and-white films, the Dunning-Pomeroy method could not be used for color film productions. Further refinement was needed to improve the matting process for color film. In the 1940s, a new process, invented by Larry Butler, emerged that used the red and blue layers of Technicolor film to create mattes (Sawicki, 2011). Much like the Dunning-Pomeroy method, a large blue screen was used as a backdrop for foreground elements. After the foreground had been shot and printed in full color, the red and blue components are then separated and printed onto black and white film. Several steps of duplication and combination are needed to unite these two components into a single matte. Mattes and original footage were then combined by using a special machine called an optical printer, which eventually replaced the bipack camera for compositing (Rickitt, 2000).

Other methods for the matting of film emerged that improved the quality of the result. These processes involved using expensive cameras with the limited use of lenses. Prisms and multiple film stocks made the process complex and cumbersome. A new method was pursued that could create a better matte but still use more conventional filming techniques. Introduced by Petro Vlahos in 1959 for Ben Hur, the blue screen color difference process allowed high-quality mattes to be made from conventional film and cameras (Gates, 2013). During development, blue and green filters were used to obtain male and female mattes. Each color record was developed normally except for the blue record, which was screened with the female matte. The male matte was used to make a copy of the background footage, leaving an unexposed area where the foreground elements were to appear. The matted
foreground and background clips were then sandwiched together in the optical printer and the background’s unexposed areas were exposed to the matted foreground. Thus, the blank areas on the background footage was filled in with the foreground subject. This process improved matte creation because fewer copies of the original footage were necessary; the quality improved with less degradation, so transparent materials could be seen with much more clarity (Rickitt, 2000).

It was found that Petro Vlahos’ blue screen color difference process was so successful that it dominated the field of visual effects for decades. The years saw improvement in the process, including those first steps into electronic matte creation as designed in 1978, predictably, by Petro Vlahos. His desire was to do for the television industry what was done for the motion picture industry. Vlahos’ system, called UltiMatte, processed the analog red, blue and green video signals separately to combine a foreground image with a new backdrop, using many of the same techniques as with the film version. Later, Newsmatte was designed to produce digital mattes for video cameras that would not provide the separated signals required for UltiMattte to work. This system was used widely in news broadcasts to key out blue backgrounds for weather broadcasts, among other venues (Foster, 2013).

It was around the same time, in 1977, that Ed Catmull and Alvy Ray Smith developed the integrated matte channel for computer-generated graphics. Perfected by Porter and Duff in the form of premultiplication, digital compositing was poised to replace the old film processes in cinema. Computer graphics began to see screen time in movies like TRON, the Last Starfighter, and Star Trek II.
Ultimatte began producing its Cine Fusion software for computers in 1993. This product was a commercially-licensed, digital compositing program and immediately was put into use in many productions. *Cliffhanger* was the first motion picture to employ the use of Cine Fusion, which became one of the cornerstones of the digital revolution (Foster, 2012).

Electronic mattes, called alpha channels, are the foundation of digital compositing. The invention of the integrated alpha channel would eventually allow computer programs to save keyed-out digital, live-action sequences with an embedded alpha channel. In essence, this is a built-in matte that takes the place of film mattes, which were created as a separate strip of film. After blue screen or green screen processing, the alpha channel of the produced matte could be saved with the other color channels. This saves a step in the digital compositing process.

The ability to produce non-destructive, instant copies of elements in the computer promotes composites that were very difficult with older methods. One non-destructive approach is a two-part “perfect” matte that is generated by combining two matte channels and then multiplying it with the original footage. First, a soft extraction is created from the original footage. The settings are lowered so that the edges of the matte are softer and more detailed. Mattes like these are often incomplete, allowing unwanted holes to appear. A second matte is then created, a “hard” channel, that clamps the edges tightly and keeps holes filled. The edges, however, are unrefined and often lose detail, especially with hair or blurred edges.
These two mattes are then combined into a third matte, one that possesses qualities of both (Wright, 2006).

Rotoscoping

Techniques like these are employed today with modern compositing software. Such software, like Adobe After Effects or the Foundry’s *Nuke* (among many, many others) contain tools for hand-made mattes as well as tools for processing color separations. Often used to refine chroma-keyed mattes, these manual mattes are created in the process of rotoscoping. As described earlier, rotoscoping has been around for many years. It was invented in 1919 by Max Fleischer as an animation apparatus that allowed animators to add images to live-action film. In those days, rotoscoping was a very different, and difficult, process than today. After live-action film was developed, it was run through a downward-facing projector and displayed on paper, where it could be advanced a frame at a time. Drawings could be made over the actors in the film or new characters could be added to live footage. The new animation work could then be filmed and composited over the original onto a new strip of film (Rickitt, 2000).

The technique was adapted to serve Alfred Hitchcock’s *The Birds* in 1963. Rotoscoping was used directly to create mattes for birds that could not be filmed in front of colored backing. Painstakingly, Disney technicians processed footage of birds in flight, hand-painting holdout mattes over images projected from live action footage. These images were then re-photographed and combined with a matte painting of a small town. This process allowed artists to add desired elements to a
shot, but also to remove unwanted problems. When puppetry was used in a film, control rods and wires could be painted out. “Garbage mattes” could also be produced, creating a small square around a single portion of the shot that would exclude everything else. A model spaceship, for example, with blue backing could be rotoscoped to include only the blue backing but leave out equipment and lights that might be sitting in the shot nearby. The blue could then be processed normally (Rickitt, 2000).

Such techniques, achievable then with many artists over several weeks, today can be completed by a single artist over a few days (Bratt, 2011). Many of the same concepts about rotoscoping are still true, even in digital form. Thought of as “animation in reverse,” rotoscoping involves an awareness of motion and the knowledge of what to look for to solve problems. As in the past, digital rotoscope artists strive to create clean mattes and separations in the least amount of time (Bratt, 2011). The main difference is that today, artists have an expanded toolset that allows greater speed. Shapes are created within the computer in the form of splines, or vector-based curves that can be adjusted with a small number of control points. In the past, matte shapes had to be re-inked for every frame of the sequence. Now, splines can be keyframed and partially automated to match the motion of objects. Garbage mattes, wire removal, and general cleanup are made easier with digital tools. New features include cloning, motion blur, and motion tracking to help stabilize hand-drawn mattes (Seymour, 2011, October).

Understood by the rotoscope artist are many of the principles of animation
that can be discovered from the study of real life. “Squash and stretch,” “secondary action,” and “anticipation” are disciplines that both rotoscope artists and animators are aware of and put to great use. Animators, however, must expand their application of such principles to a much greater degree than rotoscope artists. No matter what kind of production it is, the fundamental principles of animation guide the artists’ every keyframe to create believable performances (Bratt, 2011).

While rotoscoping may be thought of as an intense study and reproduction of reality, it may also be considered a subset of the general field of animation. Like rotoscoping, animation was hand-drawn, one frame at a time, inked and photographed onto film. Unlike rotoscoping, animation does not necessarily need to conform to live-action footage. Humorous cartoons were born as a result of this flexibility, depicting characters in funny and impossible situations that delight audiences to this day. It is the fusion of live-action and animation that is the focus of our discussion here. It is appropriate to include animation history in this analysis.

Animation

The history of animation for visual effects, like with rotoscoping, is a long one. It started not long after the invention of cinema; in 1906, Stuart Blackton created the first animated film, made from chalkboard drawings and cutout figures. Most people are aware that motion pictures are made up of a sequence of individual images that are fed through a projector and viewed as though it were a continuous scene. Innovators like Stuart Blackton took advantage of the fact that movies did not have to
be shot all at once—each individual frame of a movie could be moved through a camera and shot one frame at a time. In 1906, Blackton produced the first short animated film, *Humorous Phases of Funny Faces*, by using a blackboard and changing the pictures incrementally for each frame of film. (Rickitt, 2000)

This technique was used by French caricaturist Emile Cohl from 1908 to 1910 to produce dozens of animated films (“Emile Cohl,” 2009) and led to several animation companies to be formed in New York by 1914 (Crandol, 2001).

Soon after, in 1914, Winsor McCay dazzled audiences with his highly developed animated film, *Gertie the Dinosaur*. In this short film, McCay produced animation that appeared to interact with a real human—McCay—creating one of the early live-action/animation composites. From there, animated cartoons were produced on celuloid sheets, assembly-line style, to satisfy growing demand for animation in cinema (Simmon, 1999).

As the practice of cel animation spread, other forms of animation came into being, as well, to aid filmmakers with their artistic vision. Stop-motion is one such medium. Animations were created with miniature figures and models. While filming, a movie camera is stopped and part of the scene in front of it is changed slightly. The camera records a single frame, is stopped again and more changes are made. This is done for the entire length of the sequence, providing the illusion that the subject matter is moving of its own accord.

In the turn of the century, when the technique was first tried, stop motion was mainly done with toys and used for children’s productions or comedies. Later, in
1914, Willis O’Brien expanded the art as he began experiments with stop-motion. O’Brien animated several smaller films with stop-motion, and then was hired to help create *The Lost World* in 1925. After two years, O’Brien had created a movie with animated clay dinosaurs that seemed to interact with live-action people and locations. Using the Williams process, O’Brien screened out the backgrounds of his dinosaurs and inserted them into shots of a street full of people. The techniques, although much less developed than later attempts, succeeded in creating a lasting impression on audiences (Rickitt, 2000).

O’Brien surpassed the achievements of *The Lost World* in 1933, when he and his staff finished *King Kong*. This production used the method of rear-projection to include animated figures behind real actors. As well, a miniature rear-projection technique was devised to project live action subjects amongst the animated figures. It was O’Brien’s work in *King Kong* that inspired Ray Harryhausen to study animation and, eventually, make improvements to the process. His technique, called Dynamtion, placed a live-action rear projection behind the puppets and a glass pane in front of them. The glass pane would be painted black in areas where more live action was to appear (Gilsdorf, 2013). Harryhausen designed this system to require less labor and save money for lower-budget productions, like the 1953 film *Beast from 20,000 Fathoms*. Harryhausen animated a multitude of fantasy films over the next several decades, including *Mysterious Island*, *7th Voyage of Sinbad*, *Jason and the Argonauts*, and *Clash of the Titans*, just to name a few (Rickitt, 2000).
Animation continued advancing and finding new ways to combine with live-action footage. In the 1920s, Walt Disney created one of his first animated series in the form of “Alice Comedies” featuring a live-action girl who has adventures in a cartoon world. Cel animation eventually was used in full-length feature films alongside live-action films. As well, the two genres of films continued to intersect as animation/live-action hybrids through the years. Fantasia (1940), Song of the South (1946), Mary Poppins (1964), Bedknobs and Broomsticks, (1971) and Pete’s Dragon (1977) were some of Walt Disney’s premeire offerings that combined cartoons and real actors. Arguably, the success of such hybrid films culminated with Who Framed Roger Rabbit in 1988 (Bernstein, 2003).

Who Framed Roger Rabbit employed a complex system for making live actors share space with cartoons. Most of the scenes were live-action and the cartoons were added later. The difficulty was in manipulating the environment to accommodate the mass and actions of animated characters. Wires and air cannons were used to move objects that were knocked over. Even real objects, supposedly carried by cartoons, were animated using robotic armatures that were operated by remote control. Plastic dummies of characters were moulded and placed into the set so actors could deliver lines and keep a realistic eye line (Bernstein, 2003).

After principal photography, animators and visual effects artists went to work. Animators used photographs from the live action to rough in animated characters for approval. 82,000 frames of animation were drawn and painted to match the live action (Rickitt, 2000). Many of the characters were done in passes for the
animation cameraman. To achieve a rounded look that more closely matched the environment, cartoons were created with a color pass that had no shadows, a matte pass to let live-action components exist in front of and behind the characters, a darker shadow pass, and a lighter highlight pass. These passes were shot separately and then combined by optical printer with the live-action footage. The passes were used to cover up the telltale wires and equipment that was used to simulate the interaction of cartoons with the real life environment (Bernstein, 2003).

The *Who Framed Roger Rabbit* film was a technical achievement for its day, unsurpassed until the 1990s when *Rescuers Down Under* first used a computer-aided system for compositing animated components. The CAPS system, or Computer Animation Post-Production System, was useful in aiding the inking process of animation cels. Each frame of animation was drawn by hand, then digitally scanned and colored within the computer, and exported to film (Rickitt, 2000). Subsequent Disney films saw the use of CAPS combine 3D animation with traditional animation. With it, films like *Beauty and the Beast* and *The Lion King* boasted smooth camera motion, CG environments, and crowd replication (Robertson, 2002). CAPS was used for the final time on the animated film *The Little Match Girl* in 2006. The system was discontinued as Disney Studios began to produce feature films with 3D animation, a venture that cost less labor and money than traditional 2D animation (“Company News”, 2004).

While the aforementioned 3D animation created a revolution in film as we know it, it had a slow gestation period in film that began in the 1970s. In 1973,
Westworld featured computer graphics in the pixelated vision of the movie’s robotic antagonist. Despite the limited computing resources and very long rendering times, the film was well-received and served as the launch pad for the digital revolution in film. Despite the success of Westworld, computer graphics retained a minor role in film of the 70s. Productions like Futureworld, Star Wars and Alien used computer graphics modestly for brief scenes of computer displays and movie titles (Price, 2013).

1982 was a year where film began to see substantial contribution from computer graphics. In Paramount’s Star Trek II, the Wrath of Khan, a team of programmers and artists at the new Lucasfilm special effects division produced a 67-second animated sequence. In the animation, a barren planet is transformed into a beautiful earthlike paradise. Inspired by the Voyager flyby animations created by JPL’s Jim Blinn, the “Genesis Demo” simulated a spacecraft’s point-of-view in a spiraling journey over the transforming planet below. Several innovations were unveiled in this production, including fire made with particles, a fractal mountain-generating algorithm, a custom texture-mapping program, and a paint program for creating the planetary textures. Jim Blinn’s procedure for creating “bump maps” was also used, forming the illusion of a bumpy, cratered surface without adding extra geometry (Smith, 1982).

Whereas the CG in Star Trek II mostly stood alone in dedicated scenes, it was Disney’s TRON, also released in 1982, that dared combine its computer graphics with live action footage. In an attempt to breathe new life into Disney Studios’
productions, animators Steven Lisberger and Donald Kushner introduced *TRON* as an adventure inside a supercomputer. In the movie, live actors interacted with computer-generated environments and characters of an electronic world. It took four companies to create the 15 minutes of animation and another 15 minutes of CG backgrounds. Digital Effects, Robert Abel & Associates, Information International, Inc. (Triple I), and Mathematical Applications Group, Inc. (MAGI) worked to bring *TRON* to life (Carlson, 2013).

Geometry was actually a challenging issue for the creators of *TRON*. Polygon modeling and animation was the preferred method for Triple I, but the dense polygon structure and textures were difficult for fledgling graphics computers to keep up with. It was a very slow process to plot motions and work with 3D objects this way. Just as difficult was the painstaking need for the corners of each polygon to be plotted in 3D views. The larger and more complex objects, like the solar sailer, took weeks to complete (Conlan, 1986).

One method for generating CG images in *TRON* was MAGI’s use of Synthavision. This technique used simple shapes and textures to create objects that were much easier for the computers of the day to process. Synthavision was actually developed in 1967 by MAGI as a way to study the path-tracing of nuclear radiation into objects. It was soon discovered that this software could be adapted to trace the path of light from its source and into a virtual camera, thus creating the first raytracing renderer for film. In concert with MAGI’s solid primitive model system, high-quality images were created for *TRON* that featured lightcycles and other
vehicles. The model system, unhindered by dense polygons, allowed MAGI’s shots to be completed efficiently (Carlson, 2013).

Developments in CG for film began to pick up steam. In 1984, *The Last Starfighter* showcased computer graphics generated by the Cray X-MP supercomputer. This computer was much faster than the VAX computers used in *TRON* and allowed filmmakers to process a large amount of polygons while using a vector wireframe version of models to plan motion. Better lighting could be applied to scenes. After the success of *The Last Starfighter*, machines continued to improve. The Pixar computer arrived in 1985 with four processors for display purposes. Coupled with software advances, the Pixar computer lowered costs. The next year, the Reyes (Renders Everything You Ever Saw) computer was revealed by Lucasfilm, promising an even better performance than the Cray or the Pixar (Conlan, 1986).

James Cameron’s *The Abyss* was next to break new ground on the CG front. In 1989, live actors were made to interact with a computer-generated “pseudopod” or tentacle of seawater that acted under alien control. A new animation program called Alias allowed the lead ILM animator, Steve Williams, to help advance the tools of the trade and create stunning effects for the pseudopod. B-splines were used in the animation and modeling of the creature instead of the more prominent polygon method of the day. B-splines are curves that are mathematically calculated to span two points in space, also called control vertices or CVs. Where many polygons—which are made of straight lines—are arranged to simulate a curve with a multitude of vertices, the same curve can be achieved with a B-spline and only a few
vertices. The B-spline method helped create the organic look of the psuedopod in *the Abyss*, even helped Williams to animate it with the manipulation of the control vertices. Facial scan data was also used to prompt the psuedopod to mimic the faces of actors (Barr, 2012).

*Terminator 2: Judgment Day* continued the use of SGI computers and Alias software in 1991. With lessons learned from creating *the Abyss*, director James Cameron sought to bring this *Terminator* sequel to life with a made-for-CG character, a liquid metal antagonist that could change shape at will. Even with the success of *the Abyss*, *T2* was still a leap of faith for Cameron. The movie would be a marriage of live actors, computer graphics, and puppetry—the first of its kind. The computer graphics shots, though conceptually validated through *the Abyss*, remained very costly and still mostly experimental (Ling, 2003).

Industrial Light and Magic’s CG team went to work, using many of the same B-spline animation methods used in *the Abyss*. Since the character was made of B-spline patches, the model was built in a series of pieces. When animated, cracks between the pieces would appear. Programs were written to sew the pieces back together on every frame. Another program created C2 continuity between pieces, which means that the patch curves are made to be identical where the curves meet (Wernecke, 1994). This smoothed the joints so the model appeared to be seamless (Cameron, 2003).

*Jurassic Park* in 1993 featured photoreal dinosaurs and opened the eyes of directors to the possibilities of CG in film. Alias, the same modeling program used for
the Abyss and Terminator 2, was also used to create the B-splines for dinosaur models. Softimage was used for animation because it had inverse kinematics, a system for aligning the joints of a CG character so that endpoints, like a tail or foot, can remain fixed while the rest of the character moves (Barr, 2012). Never before had a film achieved such realism with the combination of actors and CG characters. Even compositing tools had advanced to a degree where film could be digitized, processed for bluescreen extractions, and combined with rendered 3D animation for realistic composites (Failes, 2013).

Animation tools evolved even more for the introduction of Toy Story in 1995. Having reached a critical mass of technological improvement, computer graphics were now ready to be used to create the world’s first full-length 3D animated film. Computer modeling, texturing, lighting, animation and rendering had matured with the help of Industrial Light and Magic, Lucasfilm, Disney and a new company called Pixar. Under the guidance of president Ed Catmull and director John Lasseter, Toy Story showcased a fusion of technology and filmmaking grammar. Live-action techniques were used for lighting and camera motion, while each character could have hundreds of controls for animation and timing. Around fifty characters received the benefits of new tools for inverse kinematics, walk cycles, facial animation, deformations, and texturing. A new tracking system for updating changes also aided workflow, helping secure a true “integration of Hollywood and Silicon Valley” (Robertson, 1995).
Environment Creation

Filmmakers have realized that CG effects do not only encompass the addition of computer-generated characters and objects into a live-action scene. The reverse is also true of visual effects: live action characters can be added to a fully digital background, or “digital backlot,” with equal effectiveness. In 1996, *Mission: Impossible* was one of the first films to incorporate a partially digital environment in a scene near the end of the film. Tom Cruise could be seen hanging from a computer-generated bullet train while being chased by a helicopter. The train and helicopter were both CG and were lit to match a live-action plate. The entrance to a tunnel was painted digitally. Tom Cruise and John Voight performed on a bluescreen stage and were composited onto the back of the CG train (Vaziri, 1996).

The extensive use of digital assets on *Mission: Impossible* encouraged the further development of digital backlots. George Lucas increased the use of the technique by orders of magnitude with *Phantom Menace* and later with *Attack of the Clones* and *Revenge of the Sith*. Matte paintings and CG backdrops served as backgrounds for actors filmed in front of green screens and blue screens. Matchmovers analyzed camera movements and created virtual cameras to match real cameras, allowing digital backdrops to replace moving backgrounds from the set (Bushkin, 2005).

While *Mission: Impossible* and *Star Wars* prequels used these techniques to make dangerous or improbable scenes come to reality, digital backlots also aided in
the production of films that required tighter budgets, yet provided stylized hyper-
realism that directors wanted (“On the Virtual Backlot,” n.d.). This was a practice of
creating films entirely with actors in front of colored screens. Films like *Sky Captain,*
_World of Tomorrow,* _Sin City,* _300,* and _Speed Racer* could now be made without
building expensive practical sets or visiting costly locales (Payne, 2008).

Four basic methods are used for the creation of digital backlots, each with
its own ability to meet a production’s needs. The key ingredient is the presence of
parallax, or the apparent displacement of an object that comes with the positional
change of the observer. If the camera is moving, parallax will need to be simulated to
maximize the feeling of depth (Hanson, 2008).

A fully 3D method, where all parts of the background are built in the
computer to full resolution, allows a great amount of flexiblity for productions with a
large number of shots. Buildings and objects can be re-used and rearranged to suit the
needs of the director. The wholesale use of detailed 3D models, however, costs more
time for modeling, lighting and for rendering than the other methods (Hanson, 2008).

The set-extension method, while not as flexible as the fully 3D method,
offers more realism because it combines live-action footage with CG. Camera
tracking software is used to add objects that sit in the live-action environment. CG
structures can be added to real landscapes or large indoor settings. This method is
also cheaper and the most widely-used because it can replace costly set construction
with computer graphics (Hanson, 2008).
A third method is the nodal pan or pan and tile technique. Still photos are created, stitched into a panorama and used in establishing shots. These establishing shots are created entirely within the computer, using virtual cameras to create shots that pan, zoom or tilt. More complex camera moves are avoided because of the lack of parallax in panoramas (Hanson, 2008).

The fourth approach includes a concept called “2 ½ D.” The name implies a hybrid of flat, two-dimensional images and 3D methods. One method, called camera projection, creates an “animated digital painting” by mapping a digital painting onto low-resolution 3D geometry. Once a virtual camera is placed into the scene, the digital painting now appears to have depth and parallax as the camera moves (Prince, 2012). In another method, digital photographs or video footage can be introduced into a virtual 3D space and mapped onto flat polygons, also called sprites or cards. These sprites are placed amongst other elements in the 3D scene and help satisfy the graphical needs of the shot while still keeping computing resources low (Hanson, 2008).

One key practice that is important to know in 2 ½ D or 3D backlot design is the practice of conserving computing resources. Baking textures helps artists accomplish such a goal, allowing them to populate a scene with many 3D objects while preserving reasonable render times. Complicated scenes with a large amount of geometry can use up available computer memory. At render time, when textures, lights, shadows, and bounced light need to be calculated, all that processing slows a computer’s CPU to a crawl. The complexity of such a scene, perhaps of a large city
with hundreds of buildings, may be too much for a renderer to even attempt. A process called *baking* or *light mapping* helps smooth the process out. An object, already supplied with its own color maps, can be set up in a scene with all the lighting, shadows and global illumination desired by the director. The object is rendered once; in doing so, all the information about the lighting is captured and saved into a new texture image. The image is then re-mapped onto the original object. Shadows, highlights and bounced light are now incorporated into the texture of the object itself. The lighting scenario is no longer needed for that object and resource-hungry illumination can be turned off. The scene may now have a chance at being rendered with many objects, each re-mapped with baked texture maps (Birn, 2006).

A combination of any of these techniques can be used to complete a digital backlot. Distant backdrops, where no parallax is revealed during camera motion, generally are created from large, high-resolution images. When objects appear a little closer, they can be mapped onto flat polygons, the aforementioned card method. Closer still, low-polygon count objects may be used with quality textures. Even closer, larger-polygon objects may be used with a higher level of detail. The highest-detailed objects are generally reserved for the foreground, where they must share space with real actors and should have the maximum readability (Hanson, 2008).

**Light Matching**

Whether using CG graphics, cel animation, puppets, or live actors, the most important goal of visual effects is that all assets work together to form a
convincing composite. The elements in play must appear to occupy the same space as one another. The light that appears to fall on all objects and characters in a given scene also needs to appear identical. Color influence, direction, and softness should match across all aspects of the shot. Shadows and bounced light need to be distributed appropriately throughout the composite. The matching of light has always been a concern in visual effects, even from the early days. Alfred Hitchcock used indoor studio lighting to match outdoor locations. In one such effect, called a match patch, Tippi Hedren in *The Birds* crosses a San Francisco street. As she passes behind a billboard, the shot changes imperceptibly from the location in San Francisco to a studio mock-up of the San Francisco street. Hitchcock was known to prefer the controlled environment of studios rather than on-location filming (Bouzereau, 2000).

Conversely, outdoor lighting has been used to match lighting “for free,” meaning that the same light is used for both elements. Outdoor colored screens can be erected and used to get the outdoor lighting for live actors. At times, instead of blue or green screens, simple rotoscoping is used to replace backdrops. In Steven Spielberg’s *Indiana Jones and the Kingdom of the Crystal Skull*, rotoscoping was a preferred method for matte creation because it was easier to match the available lighting without bluescreen spill (McGorry, 2008).

There are two major kinds of light that artists strive to match with live-action video. They are *direct light* and *indirect light*. Direct light comes from a specific source, like the sun or a light bulb. This light has little diffusion or scattering from the atmosphere and lets objects form solid, directional shadows. Indirect light
reaches objects from other nearby objects, the ground, or a cloudy sky. This is light that bounces from direct light sources and forms softer, diffused shadows. Sometimes there are no shadows at all, only slightly darker areas on the ground.

Direct light can be created in a virtual space—with the intention of lighting CG objects to match real lighting—with the aid of 3D lights. Different kinds of CG lights exist that help with the task of simulating real light. Directional or distant lights provide light like the sun, with parallel rays that make the scene appear that the light source is far away. With spotlights, the light spreads out from one location in a cone shape, projecting a circle of light that can have soft or hard edges. Point lights simulate a candle or light bulb, casting light in all directions. Area lights simulate a light source that emits from a specific area, rather than simply from a location, like point lights and spotlights, or from a ubiquitous general direction, like directional lights (Derakhshani, 2008).

With these lights, information is relayed from the shaders and textures of object polygons to the renderer, which processes this information and creates the final images. As well as the color of objects, specular highlights are also placed according to the position of lights. Specularity, controlled by shaders, defines a hot spot on surfaces that can be seen from certain angles. Specular highlights are not true reflections, but are examples of view-dependent shading, based on the angle between the camera, the light, and the surface (Birn, 2006). Reflections of the environment can be simulated with reflection maps, which are photos or images that influence a shader to appear to reflect surroundings. Reflection maps, however, only provide reflections
of the image. No other objects outside the reflection map are visible in these kind of reflections (Derakhshani, 2008).

**Raytracing**

If a director wants 3D objects to exhibit realistic reflections or refractions, an algorithm known as *raytracing* makes this possible. Raytracing was developed in 1968 by IBM’s Arthur Appel to improve shading for architectural drawings and industrial design (Appel, 1968). It was subsequently refined by Mathematics Application Group, Inc. in 1971 and Turner Whitted in 1979 (Whitted, 1979).

Raytracing is capable of rendering 3D objects with true reflections and refractions by emulating the path that light takes between objects. Conceptually, the algorithm works backwards from reality. The render engine uses the camera to analyze every pixel in a scene with simulated rays. Rays emanate from the camera and encounter a CG object. A single ray will strike a polygon of the object and return information about the object’s color. If the surface has reflective attributes, a new reflection ray is fired from the point of intersection. Its angle of departure is based on the *normal* of the polygon. A polygon normal is simply a line perpendicular to the polygon and describes the exact direction that a polygon faces (Birn, 2006).

The reflected ray propagates until it reaches another 3D object, reaches its bounce limit or exits the scene. That ray returns information about the color of the geometry that it hits next and adds it to the color information of the first object. This extra information builds the reflection from pixels that each ray encounters, enabling multiple bounces if desired (Whitted, 1979).
Raytracing is also a widely-used algorithm for generating accurate shadows. A second set of rays can be fired through the camera for calculating shadows. These rays encounter 3D geometry in the pixels of the scene in the same way as reflection rays. This time, however, a new ray is spawned at the point of contact and is aimed directly at lights in the scene. If the ray makes it from the primary pixel to the light without being interrupted by other geometry, the pixel is shaded as fully lit. If the ray is blocked by geometry, then the pixel will be rendered in shadow. Modern “distributed raytracing” makes use of rays that are averaged over time, lens area, or light area, assuming that the light in question has a surface area. This distribution enables the renderer to generate soft shadows, as well as motion blur, depth of field, translucency, and fuzzy reflections (Cook, et al, 1984).

**Shadow Mapping**

An alternative method for realistic shadows is shadow mapping. It was first used in film for the 1986 animated short, *Luxo, Jr.*, directed by John Lasseter (Reeves, Salesin, & Cook, 1987). Invented in 1978 by Lance Williams, shadow mapping was created to include smooth surface patches for shadowing. Previously, shadows were restricted to planar polygons. Williams’ algorithm for shadow creation made use of a computer’s framebuffer, or a temporary place in the computer’s memory that stores images before they are displayed on a monitor. It is a subset of the framebuffer, the z-buffer, which is the key to displaying the objects correctly. The z-buffer stores values of pixels according to how far away from the camera they are, represented as depth values or z-values. The z-buffer passes the depth values along to
the renderer and tells it which geometry blocks the view of other geometry (Williams, 1978).

With shadow mapping (or depth map shadowing, as it is sometimes called), shadows are created in two stages. The first stage is a preliminary render from a light’s point of view. Only the depth values of pixels are recorded in the z-buffer while color information is left out. The result is called a depth map and is kept in the z-buffer until rendering is complete. In the second stage, the camera renders the scene, comparing its own view with that of the depth map. All pixels found to be blocked from the light are rendered in shadow (Williams, 1978).

Shadow mapping is generally much faster than raytracing because it does not have to emit rays and test each pixel for blocked light that way. The image stored in the z-buffer has already done the work. The renderer needs only to check the depth map to see where shadows fall. However, because the depth map is an image with a fixed resolution, care must be taken to make sure the depth map has enough resolution to create quality shadows. If the resolution of the depth map is too low, jagged edges may appear in the shadows. If resolution is too great, the act of rendering could crash the computer system. That is why certain lights are more ideal for shadow mapping than others. Shadow maps for directional lights must be very high-resolution images to provide quality shadows and are too cumbersome for most systems to handle. Spotlights are most ideal for shadow mapping because they have adjustable cone angles and can be positioned to maximize a shadow map’s limited resolution (Birn, 2006).
Transparency is ignored completely by shadow maps. Depth maps are rendered without the added burden of color or opacity values, so light is simply cut off when geometry is encountered. Therefore, an object with transparent shaders will not cast shadows that respect those shaders. In other words, a polygon with a transparency map of a tree applied to it will not produce a tree-shaped shadow. The shadows only obey the shapes described by the depth map (Derakhshani, 2008).

Despite the drawbacks of depth-mapped shadows, shadow mapping is the favored method in creating shadows for film production (Birn, 2006). Its algorithm makes it the fastest and most efficient, even with the need to adjust the shadow size to fit the scene. Raytracing is easier to use and creates better-quality shadows, but its higher computational needs tend to cost more in render time.

**Ambient Occlusion**

There exists another method for adding soft, shadow-like features to a CG scene. Ambient occlusion seeks to add realism by simulating the natural darkening around objects. When an object is lit by indirect light, soft, fuzzy shadows appear near the intersections of geometry. This happens when the ambient light, which largely comes from all directions, is partially occluded by geometry. Hayden Landis, Ken McGaugh and Hilmar Koch from Industrial Light and Magic devised a way to use raytracing and generate ambient occlusion in the 2001 film, *Pearl Harbor*. CG planes and ships were rendered with a separate ambient occlusion “pass” and composited later. Ambient occlusion, now used everywhere in the VFX industry, helps provide added realism for little computational cost (Seymour, 2011, January).
Ambient occlusion works by sampling the area with rays that are shot from every pixel in the scene. Rays propagate in a hemisphere, testing if there is nearby geometry. The more geometry that the rays strike within a certain distance, the more darkening will occur at that particular pixel. In an occlusion pass, surface colors, specularity, reflections and transparency are all ignored; only black and white values are rendered. This grayscale pass can then be combined with other passes in the compositing process (Birn, 2006).

If matching the light from reality via direct lighting is challenging, certainly the art of matching \textit{indirect} lighting is even more so. There are many methods for matching the indirect lighting of a 3D scene with live-action lighting. Three different methods will be explored that have widespread use in the visual effects industry. These are global illumination techniques for simulating light that bounces from other objects, creating soft, diffuse images that are realistic and pleasing. These methods are the \textit{final gather}, \textit{photon mapping}, and \textit{light dome} techniques.

\textbf{Radiosity}

Global illumination for computer graphics first emerged in 1984 with the development of \textit{radiosity}. Radiosity is a \textit{view-independent} lighting method, unlike raytracing, which depends upon the camera’s point of view. In a 3D scene that seeks to simulate bounced light, geometry is subdivided into smaller polygons before rendering. These patches, in effect, each become its own light source. Polygons are subdivided more in higher-contrast areas. The total amount of light energy is
calculated from the main light source. Taking into account how much light is “absorbed” from other surfaces in the scene, the light emitted from polygons is attenuated, based upon those surface attributes. The resulting solution for each polygon is stored for use in rendering. Once the radiosity solution is stored, the camera can access it quickly, no matter what the camera’s point of view is (Rusnell, 2006).

While radiosity provided accurate results for indirect illumination, it was not without limitations. Radiosity is best used for scenes where the lights and objects do not move. Otherwise, the radiosity solution would have to be re-calculated for every frame. This action consumes computing resources and increases rendering time, especially if there are also high-polygon-count objects in the scene. With an increased number of polygons comes an increased complexity in the radiosity solution, which is found by analyzing the attributes of each polygon (Birn, 2006).

Radiosity improved with stochastic sampling, developed in 1986 by Robert Cook to help with rendering artifacts. This was a modification made to distributed raytracing and could send out rays in a random fashion rather than the old uniform pattern. Stochastic sampling is performed with a randomizing technique called Monte Carlo (Cook, 1986). Distributed raytracing could now be combined with radiosity to form a more complete global illumination method. Raytracing could be used to simulate specular highlights, which was unavailable with conventional radiosity. Shadows, caustics, and reflections could now be added to global illumination solutions (Jensen, 1996).
The new technology of radiosity was first used in the feature film, *Casino* in 1995. A digital backlot was created, rebuilding the 1970s Las Vegas strip. Lightscape Technologies worked with artists from Matte World Digital to adapt radiosity algorithms and produce realistic digital matte paintings of the nighttime Vegas. These paintings perfectly matched live-action footage filmed on location (Rickitt, 2000).

**Photon Mapping**

In 1996, Henrik Wann Jensen of the Technical University of Denmark devised a new way to create global illumination. To solve radiosity’s problem of long render times with high-geometry scenes, *photon mapping* was introduced as a way to simplify the illumination process. As well as indirect illumination, photon mapping is also capable of reproducing shadows, reflections, glossy diffuse surfaces, and caustic effects (Jensen, 1996).

Photon mapping occurs in two steps. The first step is for lights to fire photons towards all objects in a scene. In computer lighting, a photon is a discrete construct that has size, intensity and direction. Unlike raytracing, photons come from light sources instead of the camera. Photons encounter geometry and test the surface to see if they can be reflected, absorbed, or transmitted. The surface’s attributes, along with a stochastic sampling algorithm called *Russian Roulette*, inform the photon about what to do next. These instructions are stored in a *photon map*, a rough representation of the entire scene that outlines where all photons have hit objects and where *shadow photons* have been spawned to indicate where shadows lay. Shadow
photons are instrumental in saving time in the rendering process. Shadows are already recorded in these areas so the renderer does not have to calculate them. (Jensen, 1996).

There are two kinds of photon maps that are created separately to save computing resources. The *global photon map* is what was described above. Global photon maps record the behavior of photons that have low density and high energy. The separately-generated *caustic photon map*, on the other hand, is concerned only with photons that have low energy and high density. Caustics are a focused version of global illumination that forms bright, distinct shapes that reflect from or transmit through geometry. Caustic photons are projected only toward objects with high specularity. Because of the high density of photons needed for realistic caustics, the caustics map is stored separately (Jensen, 1996).

In the second step, raytracing is used in conjunction with the photon map to render direct lighting, shadows, caustics, and indirect illumination. Only reflections are rendered without the photon map, needing only the Monte Carlo process and importance sampling to create specular reflections (Jensen, 1996). Importance sampling is a way of limiting the sampling of a 3D scene to mostly the important areas (Owen, 2013).

**Final Gather**

To achieve good results with photon mapping’s indirect illumination, a high number of very small photons are needed, requiring longer rendering times. Otherwise, the result is often a splotchy, grainy appearance to the final render. Photon
mapping, as well as radiosity, can call upon a process called *gathering* or *final gathering* to smooth the result without having to add so many photons (Birn, 2006). Final Gather was introduced in 1992 by Mark Reichert as a way to speed up the radiosity process (Pharr & Humphreys, 2010) and was implemented in photon mapping as well. When indirect light is being rendered, raytracing is used to shoot primary rays from the camera and sample the scene. A ray will come into contact with geometry, consult the global photon map and see if needs to reflect, transmit or be absorbed. If the photon map indicates reflection at the point of contact, a new set of rays, called *final gather rays*, are fired in a hemisphere from that point. The new rays return information about the colors they find and add them to the first point of contact.

At times, a primary ray will encounter a spot that the photon map gives low importance to. This is an area where there have not been many photon hits when the map was created. In this case, no new final gather rays are fired; instead, the color values are averaged with the nearest photon hits to the point of contact. This method saves render time and computing power because secondary rays are not used.

Final gather smoothes out rougher photon map renders and usually requires only one ray bounce to provide good results. Final gather can also be used without a photon map, using raytracing to generate final gather rays and add only a single bounce. This method is not as complete as final gather with photon maps because the photon maps aid with multi-bounce solutions. However, without the extra
burden of photon maps, final gather can provide adequate results in an efficient manner (Birn, 2006).

The disadvantage to final-gather-only solutions is the occurrence of flickering in the final render. Flickering happens when final gather computes a solution between frames that differ significantly from the prior frame. Motion of 3D geometry or cameras tend to cause this, spawning final gather rays that did not exist in prior frames. A solution to the flickering problem includes a method for recording the final gather points and projecting them at regular intervals along the camera path. Each final gather map is used to interpolate the lighting solution between the intervals so that the differences between frames can be averaged (“3ds Max Help,” 2012).

Global illumination was used extensively in 2004 for the Dreamworks animated film Shrek 2. A custom solution was devised that combined radiosity, final gathering, and irradiance caching while minimizing costly raytracing. Photon mapping was considered too expensive in render time to handle the hundreds of thousands of frames needed for the feature film. Dreamworks’ new process met the aesthetic needs of the movie while providing an efficient workflow (Tabellion, et al, 2004).

Radiosity, photon mapping and final gather are techniques known as biased methods. This means that the results are convincing but not necessarily physically accurate. These methods enlist the help of “cheats”, like interpolation and averaging to reduce render times. A newer development in biased rendering is the
point cloud method. Pioneered in 2004 by Michael Bunnell, the point cloud method could be used to create ambient occlusion without the use of raytracing. Further development by Pixar’s Per Christiansen and Industrial Light and Magic’s Chistophe Hery allowed point clouds to be used in *Pirates of the Carribean, Dead Man’s Chest*. With this system, special colored disks are placed on geometry that is lit by direct illumination. These disks, called *surfels*, store information about their size, location, colors, and the direction they face. Color bleeding and ambient occlusion is calculated based upon the approximate radiance of the surrounding surfels (Sandys, 2010).

Other, more render-intensive techniques are available that are more physically accurate, but also take more time to produce renders. These methods are known as *unbiased methods*, also called brute force raytracing methods. Path tracing, bi-directional path tracing, and Metropolis light transport are all examples of brute force renderers that are physically accurate but generally cost more in render time (“GI Methods,” 2010). The brute-force method involves sending far more sample rays into a 3D scene than conventional raytracing or biased systems, so the need for manipulation of settings is much less for artists. Coupled with faster computers available today, brute force techniques are gaining popularity in film productions (Seymour, 2013).

**Light Dome Rigs**

A third lighting method used in visual effects is worth mentioning. In the light dome technique, a special program uses a digital backdrop image as a guide to generate conventional lights and place them into a 3D scene. The lights, usually
spotlights or point lights, are an inexpensive substitute for global illumination. They are arranged in a spherical array around the scene and used to match the colors in the backdrop.

Light dome rigs do not offer true global illumination, but provide environmental lighting that creates a sense that light is arriving from all around the scene. This technique is popular in the visual effects industry because it can achieve a “fake” yet adequate lighting scenario that saves render time (Van Gool, et al, 2004). Advantages of the light dome method include flexibility and speed. Once the lights have been created, they are no longer constrained to the values taken from the backdrop image. They can be adjusted individually to suit the artist’s needs. The technique is also faster than true global illumination. Neither bouncing photons nor final gather rays are needed. Color bleeding and caustics are not present, however. Additional lights will need to be placed near objects to simulate light that bounces from other objects (Bloch, 2010).

Further time may also be saved with the use of the spinning light trick. Once generated and placed into a scene, lights are parented each to a null object and rotated around it like a centrifuge. Spinning lights make use of motion blur to multiply their intensities and blur the light around the axis of rotation. A common rotation rate is two rotations per frame. Light appears to multiply with each higher number of motion blur passes; a setting of 5 will make 10 lights appear the same as 50 lights (Bloch, 2010).
Image-based Lighting

As described, all of these techniques have in common the ability to use HDRI images to generate realistic lighting. This process is called image based lighting or IBL. With IBL, global illumination is powered by a sphere that surrounds a 3D scene. Indirect light is cast into the scene by this sphere. Invented in 1998 by Paul Debevec, IBL was devised as a way to cast real-world light into a synthetic scene. It is a means of using the actual lighting from a live-action scene and allowing 3D objects to be lit by that light. The radiance values from the live-action set are recorded into a special image format and used to re-create the lighting solution virtually. To do this, the image is typically mapped onto a large sphere that encompasses the entire scene (Debevec, 1998). Brute force raytracing, photon mapping and final gather can each be used to send samples from the sphere’s normals and into the scene. Single-bounce and multi-bounce solutions are possible with IBL. Even fake global illumination, as described earlier as a light rig, can be generated with the same general idea as IBL: a high-dynamic range image is sampled and used to create point lights or spotlights (Bloch, 2010).

The use of HDR images are crucial in image-based lighting. HDR stands for high dynamic range. These images represent a much broader range of colors and detail than conventional 8-bit images can contain. Where an 8-bit digital photograph can show detail in shadows but not in bright areas (and vice versa), an HDR image can retain superior detail in both bright areas and dark areas. While an 8- or 16-bit image only has discrete, integer levels of color and intensity, an HDR image’s color levels
can have ultra-precise values called floating point values, as mentioned earlier. It is this precision that lets a 3D application generate realistic-looking global illumination that can closely match real-world light values (Bloch, 2010).

In many VFX productions, artists seek to “place” 3D objects into an environment by lighting those objects with the same light that was present during a film shoot. If a CG robot was created with the intent that it would need to cross a busy city street and talk to live-action people, the robot would need to be lit to match the available lighting from the shoot. Unfortunately, simply using the filmed footage as a global illumination light source will not yield the complexity needed to simulate real-world lighting. An HDR image needs to be created that captures the entire range of radiance from that city street.

A standard method for capturing the needed data involves a digital camera and a light probe. A light probe is a simple mirrored ball that is placed on the set, usually in the corresponding location where planned CG objects might appear. These chrome balls have been used for years for obtaining reflection maps. A photograph is taken of the sphere, “unwrapped” in postprocessing and used in 3D software as a reflection map (Debevec, 1998). These maps help cast real world features into reflections and also assist artists with the placement of light sources to match lighting from the set.

The use of the light probe, in HDR photography, is very similar to the collection of reflection map data. The process is essentially the same, but with a few differences. Instead of a single photograph, a series of light probe exposures are
taken, each with a different shutter speed. A wide gamut of exposures are taken with the camera in a fixed position. The exposures will each contain a low dynamic range of values, but collectively span a much wider range. The shutter speed is adjusted to provide photos where subjects are nearly completely washed out with overexposure, on one end of the series, or barely visible from underexposure on the other end of the series. The exposures in between these two extremes, in total, contain the rest of the dynamic range of available light (Debevec, 1998).

These exposures must be combined to form an HDR image. Programs are available that can take the original photos and align them with one another. Care must be taken during shooting to avoid moving elements, like cars, people, moving branches, and the like. Movement will cause the exposures to be mismatched and create a “ghosting” phenomenon.

Once ghosting is corrected and the images aligned, the program can now combine the data from all exposures into one of the many HDR file formats. These are generally large 32-bit files that are encoded with floating-point values in the pixels. Some formats store pixels as 32 bits per channel. The Radiance format, developed in 1987 by Greg Ward, uses only 8 bits per channel, but adds a fourth channel to store luminance information. This channel instructs the other channels how to calculate the real floating-point values (Ward, 2010). Other formats have improved upon Radiance, like Tiff LogLuv and Open EXR (Bloch, 2010). Open EXR, an open-source format developed by Industrial Light and Magic, provides efficient
compression, alpha channel capability, flexibility, and accuracy for use in HDR files or general film production (Brinkmann, 2008).

In image-based lighting, light from the HDR image must be cast from all directions to light the virtual scene. A sphere of geometry is placed in the scene so that all other geometry is enclosed within it. Then, the HDR image is spherically mapped to this geometry so the image’s real-world values can drive the global illumination solution.

Rendering

Visual effects artists are concerned with high-quality renderings of 3D objects for combination with live-action elements. An efficient and established practice for accomplishing this is to render discrete passes of various characteristics. Not only is the matted live-action green screen footage used a separate pass, but 3D objects are treated as separate elements, as well. Objects can be rendered separately from other objects and combined at a later time, allowing for each object to be manipulated independently of the other objects. A rendered object may, for example, be made more transparent, or blurred, or even change color without contaminating the other elements with the same changes (Birn 2006).

Furthermore, separate render passes help save on computer memory. Complex scenes with many objects and textures can be cumbersome to manage within the 3D application. Renders can go much slower or even cause a system crash. Rendering objects separately in passes eases the computer’s workload (Birn 2006).

Multipass rendering is not limited to the separation of one object from
another. One can separate the object’s *attributes* from each other, as well. An object may have a complex surface that has color, shininess, and reflections. There may be bounced light that comes from global illumination. There may be ambient occlusion present that adds darkness to cracks and crevices. Even the shadows that the objects cast or receive qualify as attributes. Each one of these characteristics can be rendered separately from one another. This is done to add the maximum amount of control to the compositing process. Each render is imported into a compositing program so they can be fine-tuned to the director’s artistic vision. If an object is too shiny, the specular component can be turned down without having to re-render the entire object. Multi-pass rendering and compositing adds control while saving extra time that may have been spent waiting for new renders (Birn, 2006).

Color, specularity, reflection, occlusion, and shadow passes are only some of the pass types available for use in compositing. A *depth pass* is a grayscale image that represents the distances of objects from the camera. This map is generated from the z-buffer, the same place that stores the information for shadow maps. The depth pass can be placed into a compositing program with the other passes and used to blur the composite only in places that are further away from the camera. The scene can be tinted a different color to simulate the thickening of the atmosphere that comes with distance, or an increase in dust or darkness (Brinkmann, 2008).

A *mask pass* or *I.D. pass* is a simple way of separating another pass into smaller parts. If a scene is rendered in color, for example, but separate passes are desired for many individual objects, an I.D. pass will be created to set up the
separations instead of rendering each object one at a time. A compositing program may have a feature that creates mattes from simple colors, like red, blue, green, yellow, et cetera. An I.D. pass will be made where each object to be separated will be rendered with a flat, separate colors like these. Now, a rendered pass with only one built-in alpha channel can potentially have three or four mattes when the pass is combined with the I.D. render. A matte will be created for all green objects, as directed by the I.D. pass, another for all red objects, and so on (Birn, 2006).

Compositing

Compositing is the step in visual effects production that makes all elements and passes come together. A compositing program refers to the locations in computer memory where elements are kept and creates previews that artists can use. Artists work with the previews to arrange composites with the tools that are available in the program. Nodes or layers, each containing an element, can be manipulated and finally rendered out into new image sequence or movie file.

The concept of digital compositing emerged with the invention of the integrated alpha channel by Ed Catmull and Alvy Ray Smith in 1977. As outlined earlier, alpha channels in CG renders are akin to live-action footage that has been chroma-keyed, where the subject matter has been separated from its green screen background. The alpha channel made it possible for computer-generated elements to be easily overlaid onto other elements within the realm of compositing software. It is this domain that will be described below.
Compositing software has the task of stacking images on top of one another and blending them together into a composite. It is not enough simply for the program to lay the images over each other. The right areas of key images must be made transparent so the lower layers can show through. Transparency can be created in a multitude of ways in a compositing package, but each method involves the use of black and white images in one form or another. Transparency can be interpreted from simple 8-bit grayscale images that can represent 256 values of gray. The color values of one image are multiplied by the color values of the other. The grayscale image is used as a stencil to make certain areas of the normal image transparent. The darker pixels are in the grayscale image, the more transparent the corresponding pixels will be in the normal image. The grayscale image now serves as the alpha of the normal image (Smith, 1995).

With integrated alpha channels, a CG program renders the subject’s color values in the normal three channels—red, blue and green—but also saves a grayscale component as a fourth channel, or alpha channel. File types like Targa, PNG, and OpenEXR are designed to accept the fourth channel when saving. The alpha channel produces black-and-white values that correspond to the shapes of rendered 3D objects, and also describe semi-transparent areas with grayscale values. Antialiasing, which is the softening of edge “jaggies” with subpixel transparencies, is handled by the alpha channel. Half-transparent pixels along edges are created in the compositing step when the compositor refers to the alpha channel’s gray pixels that reside along the edges (Smith, 1995).
Not long after Catmull and Smith developed the integrated alpha channel, Tom Porter and Tom Duff went a step further and created the *premultiplied alpha channel* in 1984 (Porter & Duff, 1984). Like a conventional, straight image, a render with premultiplication yields a color image with an embedded alpha channel. However, a premultiplied render finds the semitransparent pixels in the alpha channel and modifies the actual color pixels with it, as well. Before saving the render to disk, the render engine adds the backdrop color to where the semitransparent pixels correspond to in the RGB channels. This process is known as *premultiplying* or *matting with color*. To be exact, this does not make the actual pixels of the RGB channels transparent; the alpha channel does that later, in the compositing program. Premultiplying “treats” the semitransparent areas with the color of the background, according to the transparency values of the alpha channel. Since artists who render in passes usually allow black or white as the background, the added color is also, usually, black or white (Christiansen, 2013).

The addition of the background color to the RGB values essentially saves computer time in multiplying the color channels with the alpha image. Since this has already been done at render time to produce the black matting, the compositing program no longer has to do it. The program can quickly subtract the black pixels upon import and make the element available for compositing. Then, when a composite goes to output, the black is brought back in to balance semitransparent areas (Christiansen, 2013).
A common problem with working with premultiplied images is that artifacts can appear with manipulation of the color channels. Black has now been introduced as an integral part of the color channels to represent partially transparent pixels. When a compositing program “expects” this added black in the pixels, it can take steps to process the black so it contributes to transparency. If a program does not expect a premultiplied image, it performs the normal operation of multiplying the color channels by its alpha channel to achieve transparency. When the program does this to an image that has already been premultiplied, the black pixels along partially transparent edges do not respond correctly. These pixels, in fact, appear darker than intended and form a dark halo in areas where there was supposed to be semi-transparent pixels. The color values no longer blend with the background but are now conspicuous artifacts. This kind of problem also appears when making color corrections to a premultiplied layer, even when the compositing application is made to expect premultiplied images. A change in brightness, for example, affects the black pixels all the way to the edge of the element, sometimes making them brighter or darker than they should be (Brinkmann, 2008).

The easy solution for double-multiplication is to make sure that the compositing program expects the premultiplied image. This way, the image will not be multiplied by its alpha on import because this has already been done. The black edge pixels will be blended properly with the background (Brinkmann, 2008).

Halos that come from color changes require a bit more finesse. In many programs, an “unpremultiply” node is available to invert the premultiplication process.
before the color correction is made. An unpremultiply node acts in reverse, *dividing* the values of all pixels in the element by its alpha channel. This allows color corrections to be made without fringing. Once the corrections are finished, the element can now be re-multiplied by its alpha channel for compositing (Wright, 2006).

Multiplication is only one of several blending methods available in compositing programs. Mathematical equations govern how these nodes function. An “add” node sums values of layers, “subtract” displays differences of values, and so forth. “Multiply” and “add” were two methods that were used often in this project. “Screen” was also put to good use. “Screen” inverts the values of both images before multiplying them together. In so doing, the bright areas of the first image tend to brighten the darker areas of the second, without affecting the bright areas of the second image. This operator is often used to composite reflection passes over glass (Brinkmann, 2008).
CHAPTER III

METHODOLOGY

For the sake of efficiency and flow, the format of this section is written in first-person style. What follows is a narrative of how this project was conducted. Because of its length, the first-person style may be preferred for clarity and comfort for the reader.

Most productions that need visual effects begin with a story. A script is written, footage is acquired, and then the proper techniques are enlisted to add visual effects. While attention to effects arrives to service the story, in the instance of this research, the story was tailored to showcase the major effects techniques.

Plan

The requirements that I set for myself influenced the storyline. There were many concepts that I was interested in exploring. A platform was needed where all the concepts could come together in an interesting way. I wanted to compare light matching techniques and to simulate indirect light. It seemed simpler to avoid harsh shadows and overspill from green screens. I did not have easy access to a studio with controlled lighting, so I elected to use soft, natural lighting from the outdoors under a cloudy sky, or use “magic hour” lighting from dawn or dusk.

I also wanted to demonstrate matchmoving, so I knew the camera needed to be in motion. There was a sequence in Minority Report where the camera moved around Tom Cruise’s character as he watched a holographic video. The camera viewed the scene
through the hologram, back at Cruise. I liked the depth of the shot and was inspired to try a similar technique.

I also wanted to try to match the motion of a CG object with a real person. *Star Wars, Episode III* and *Matrix Revolutions* had actors who appeared to interact with computer-generated vehicles. I decided to try and create a vehicle that a person could appear to climb aboard and fly away upon, all in the same shot, without cutting away.

Finally, I wanted to explore chroma key concepts and replace the background behind the talent with a new scene. The theme of this production was leaning toward a futuristic science fiction scenario, so a *Star Wars*-style cityscape seemed to be a good candidate for a digital backlot experiment.

The concepts that I wanted to investigate led me to create a simple story that would support such ambitions. “Soccer Mom” is a story about a woman and her teenaged daughter who live in a large city, in the year 2043. They each live busy lives in their own technologically-enhanced bubbles. The mother surrounds herself with holographic displays to support her own business while the daughter is firmly anchored in her world of soccer practice and video wrist-phones. While waiting at the platform for their own, separate trams to work and school, both retreat into their techno-worlds. When the daughter’s tram arrives on the platform first, she is forced to “break” into the mother’s holographic rhythm in order to say “good-bye” for the day. The daughter tosses her soccer ball into her mother’s holographic screens, dissolving them and bringing her mother back to reality. The two embrace and say their good-byes. After a bit of fussing from Mother, the daughter boards the tram and glides away, waving. The message is
simple: as always, it is good to break from our busy lifestyles, even briefly, and reconnect with the important people in our lives.

Using this framework, I designed a scene with the intent to test several ways of incorporating live actors with computer graphics. My limited budget forced me to take a minimalist approach regarding live-action. The resources I had available in the digital realm, however, allowed me to compensate for the lack of funds with digital assets. I chose to shoot the live action outdoors with both blue and green screens behind the actresses, in two separate takes. I created a digital storyboard, as detailed in Figure 1, with Bauhaus Mirage to outline the shot.

![Figure 1. An early storyboard depicts a child boarding a flying bus.](image-url)

At first, I planned to use a small bus, loaned by the pastor of our church, to carry the daughter away. That idea later proved to be untenable, so I decided to design a small,
wheeled platform that could be pulled manually through the set. The only other real object in this scene would be the soccer ball and a modest bench that Mother would be seated on. The rest of the environment would be digital. Railings, displays, buildings, the background, and an art-designed tram replacement would be created and combined with the live-action video. Like the live-action component, several “takes” or renders, would be needed to test CG methods for efficiency. Three tests for light matching, two tests for shadows, and two tests for the digital backdrop eventually were planned and executed. Each combination of tests would be produced and offered for comparison in the interactive application.

Previsualization

After the plan was in place, I was ready to begin the previsualization stage of production. I used illustrations, photographs of the shooting site, and simple 3D geometry to outline the production. Everything was recreated in 3D space, including the camera, dolly and track. I also built virtual models of the colored screen background and frames, floorboards, partition, and a tram stand-in object. Tracking markers, to be created later by simply applying white tape in “+” shapes, were placed on the 3D colored screens in regular intervals. These were to serve as reference points for matchmove software, both in the practical set and in the virtual set. The set was also re-created from reality via measurements; even the environment around the set was simulated by creating a 3D panoramic image that wrapped around the entire scene.

The previsualization phase also included some rudimentary animation of simple 3D characters, representing the live actors. Crude reference video was shot of the
two actresses. They acted out the main motions in the shot, as directed, while the camera moved around them as planned. I used this footage as reference for the animation of the characters in the previsualization scene. Slowly, I built up set changes and timing adjustments that helped me get ready for shooting day. I animated the virtual camera and adjusted its motion until I was satisfied. A final schematic of the set was printed out for reference and to aid in the building of the set. Please see Figures 2, 3 and 4 for illustrations of previsualization, set schematic and the constructed set.

Set construction began. Based on the previsualization diagram, a list of materials was created. I had already scouted a suitable location—a church’s outdoor basketball court. It had everything I needed: a large, flat area, powerful floodlights, electric outlets, restrooms, benches, and easy access for vehicles. I bought large sheets of colored muslin fabric from an online film equipment outlet. These were to serve as the chroma-keyable backdrops. I designed and built two large frames to allow me to stretch the cloth tightly and hold the sheets upright. Plywood panels from a local hardware store served as floorboards and partitions. One side of each panel was painted green for the green screen test and the opposite side was painted blue for the blue screen test. The panels were to be reversed when it was time to switch tests. I built the rolling tram with lawnmower wheels and removable side panels, also dual-coated with blue and green paint. A simple camera dolly was built with plywood and skateboard wheels. The dolly’s wheels sat on stiff rubber hoses that acted as dolly rails. The dolly could glide smoothly along those rubber hoses when the outside temperature was warm enough.

Other essential materials were collected. Various toolkits were designed and
Figure 2. Previsualization in 3D software includes building a virtual set.

assembled. One such kit included HDRI photography tools, like a chrome ball on a stick with one side painted gray, a similar white foam ball, and a crude inclinometer for noting the sun angle. Spare PVC supplies, tools, marker tape, tripods, stands, paint, and a host of other supplies were carefully anticipated and obtained for the video shoot.

Shooting

Finally, my little team got together to shoot the video. The actual set was built according to previsualization specifications. Careful measurements were taken of the camera height and various distances between objects. Shooting was done in the predawn hours so we could get soft lighting. I directed and operated the camera.
Figure 3. A set schematic, made from previsualization files, helped guide the construction of the outdoor set.

A friend served to push the dolly along its track. There was a person who operated the timer and called out crucial event cues for the talent. The actresses from the previsualization session were present, in costume, to act out their roles.

After the action was complete, the camera continued filming while the assistant held the light probe in various locations on the set. This provided usable
reference material for matching the lighting conditions of shooting day. The light probe’s chrome hemisphere allowed me to note which way the light was coming from and what may be contributing to the reflections in the scene. The light probe’s gray hemisphere helped me see how pronounced the shadows were and what color the light was. Finally, a white light probe, made from a foam ball painted with gesso, showed the light color in a way that could be directly sampled and matched.

I had hoped to shoot both the green and blue shots on the same day, but by the time we finished the green shoot, broke down the set and set up the blue screens and panels, the sun was already too high to provide the flat lighting desired. A rescheduling was required in a new site that would have shade much longer during the day. I found a location at an abandoned shopping mall. We set up our blue set next to a high wall and

Figure 4. Constructed set on shooting day.
completed the blue shoot in the same manner as the green shoot. Overall, the lighting was nearly the same if not a little better quality. The foggy morning diffused the sunlight and provided just a hint of rim light that illuminated the talent. I would have to make slight adjustments to my tests in postproduction, as will be explained.

I shot the video on a Sony Handycam with 4:2:0 AVCHD format video, at a resolution of 1920 x 1080i. After a satisfactory take was recorded, I set up a digital camera and took close-up exposures of the reflective light probe ball. I recorded between 7 and 9 exposures of the ball while changing the shutter speed on each exposure. These images would later be combined into an HDR image for use in postprocessing.

Later, I transferred the video from the camera and into the computer. Using Adobe Premiere CS4, I selected the proper clips from both green and blue shoots and trimmed them for length. I also created proxy clips of lower resolution to aid in the matchmove and animation processes. The lower-resolution proxies saved RAM and allowed those processes to be accomplished faster.

**Matchmoving**

In the matchmoving phase of production, the features of the original video are analyzed so that a 3D camera can be produced to match the motion of the live video camera. If this is not done, the 3D objects can not appear to “move” with the motion of the video camera, but merely appear to float in space in front of the live action video. Boujou 3 is a program that creates a 3D camera and imports it into a variety of 3D programs.
Initially, the matchmove software must search the live action footage for trackable features and then solve for the new camera’s motion. Automatic tracking is an aspect of Boujou 3 that makes tracking easy. I imported my proxy version of the footage into Boujou and let the computer do its work. Hundreds of features were found. These were simply high-contrast groups of pixels that the program could follow through the entire shot.

When this process was complete, I needed to prepare the shot for calibration, which is the finalizing of the data and solution to create the new camera. There were many variables that would affect the camera solution if not addressed. Primarily, there was extensive motion of people and objects that did not contribute to the camera solution. To make sure that Boujou ignored these aspects, I used built-in tools to mask out everything that was moving. Upon further inspection, I realized that the wind was blowing the green screen slightly. This would also hinder the camera solution, so I created more masks to eliminate this contamination.

At this point, calibration could begin. Boujou analyzed the features that were left over from masking and calculated a preliminary camera motion. Reference points, called tracks, were also included in a 3D space that marked the positions of features. I could inspect the tracks and camera motion and see how well it matched the real camera. The motion seemed to match well; however, there was further calibration to perform before conducting the final solution. I had to make sure the final camera export was the right scale and that the new camera was pointing in the right direction.

I provided the computer with more information that helped with these important values. In Boujou, this was called “adding scene geometry”. In essence, it was
simply the act of selecting tracks and applying descriptions to them. First, I selected all the tracks that appeared to lie on the ground plane, according to the feature tracking. When compared to the default ground plane given in Boujou’s 3D simulation, those points appeared to lie tilted at least 45 degrees from the Boujou ground plane. With those tracks selected, I instructed Boujou to view these points as lying on the ground. The entire scene pivoted and made the correction. Similar adjustments were needed to describe the x-y plane (the rear wall) and the camera’s height.

The final information needed was that for establishing the correct scale. All I required was one set of measurements of known length. I found two tracks that rested on corners of the partition that hid the tram. Consulting my previsualization notes, I found the correct length between the corners and entered the new parameters. The calibration was now complete.

The program then exported a camera for use in Maya 2008, along with a parented rectangular polygon with the live-action footage mapped to it. This was a guide that Maya created to gauge how well the camera had been matchmoved, and to further animate objects to match the video.

Models

Meanwhile, 3D models were built and textured. My attention was first turned to the background buildings that made up the CG cityscape. Some artistic exploration was necessary to create buildings that would match the theme. I turned to science fiction movies like *Star Wars, Episode III, Star Trek* and *Astro Boy* for design inspiration. I elected to use repeating horizontal and vertical lines that were broken up by sweeping curves or direction-changing diagonals. To me, these shape choices seemed to represent
the cold emotional distance that people may have from one another, and how that
dominant, regular pattern can be “broken up” by brief moments of warmth. Textures
added to the emotional contrast, as well, contributing metallic grays and blues that were
occasionally interrupted by playful shapes and a few reds, browns and golds.

The buildings were created with arbitrary measurements in Lightwave 9.3. Exact size relationships with the foreground objects were not necessary, since the
cityscape was to be rendered as an image that would exist far behind the foreground
models. Lightwave had a familiar interface to me and I already knew how to use it for
modeling. I created images for those buildings in Adobe Photoshop CS2 and used UV
mapping to lay the textures onto the models.

Next, I built the foreground models that the actresses were to appear to
interact with. I built these to scale in Lightwave and exported them into Maya. Maya has
a better system for rendering multiple passes of an object without having to generate
whole new scene files. I initially built the models with 4-sided polygons, but I had to
convert all of them into three-sided polygons, otherwise known as tripling, so the models
could export into Maya. I used the tracking points that came across from Boujou to help
scale the models properly and position them for lineup with live action footage. A
lengthy process ensued for applying textures to each of these new models.

Backdrop Creation

At that time, I had enough digital assets to begin my first test. I began to
produce a large image of the futuristic cityscape that was to serve as a static backdrop.
Since the backdrop was to represent buildings that were far away in 3D space, I felt that
the lighting scenario did not necessarily need to be exactly the same as the lighting of the foreground objects. I decided to select a single HDR image from shooting day and create a backdrop that would serve both green and blue scenes. This decision was made to reduce the number of tests in the project. An all-purpose backdrop that could accommodate both blue and green video clips was necessary to keep work time and file storage manageable.

I chose the HDR image that had a more natural look. This was the image produced from the green shoot, in the church basketball court. There were more colors present in this image than with the blue HDRI, taken at the mall site. There were also more interesting cloud formations in the green take, which might provide a more realistic variation of bounced light on my CG objects.

I needed a guide to help me make sure the backdrop looked convincing. I used the Internet to search for images of skyscrapers and cityscapes. Specifically, I needed a sample to show me how a city might look under predawn or evening conditions, or if there were other conditions that might match the video shoots. I took a preliminary still image from the green footage, keyed out the green backdrop with After Effects, and arranged various skyscraper images behind the still image. I found an image of the New York City skyline at dawn and found it to be a good match to the live-action footage. I used this image as a reference that helped me create the sky, gauge the fuzziness of receding buildings, the colors, light intensities, sign illumination, and shadow depth.

My first test was to see which was more efficient, a “2 ½D” method for creating the backdrop, or a 3D method. The 2 ½D method, as explained before, involves creating simple, flat images of buildings and importing them into an editing program like
Adobe After Effects. The images could then be manipulated in a 3D space, allowing for placement, resizing, and rotation. The only difference was that the images would not be 3D models but would remain like flat “cards”. The images would be layered and copied until a complete city image was produced. The advantage to this method was that fewer computational resources would be required to create the image. The 3D method involved using 3D models of buildings together in one scene and rendered only once, without extensive manipulation afterwards.

To accomplish the 2 ½ D test, I still needed the 3D models of buildings that I already created. I imported all 15 models into a new scene in Lightwave and created spotlights to provide accent lighting to each building. This created a little more realism and helped establish a sense of scale. Copies of buildings were placed near each primary building to act as shadow casters and reflection casters. These secondary buildings were made invisible to all cameras even though their shadows and reflections were present.

The buildings were arranged in a row and a new camera was assigned to each model. One exposure of each building was not going to be enough to provide the variety that most people expect when viewing a city. The same image of a building, viewed from the same angle, multiplied by dozens would be a dead giveaway that the image was contrived. If several angles were captured of each building and used separately, it was much more likely that those images may be interpreted as separate buildings and not as mere copies. To capture multiple angles quickly, the buildings were animated to rotate 360 degrees over seven frames. This yielded 6 different images of each building from 6 angles. Rendered with bounced light from the HDRI image, occlusion, shadows and
spotlights, 90 unique images of buildings were made available for editing in compositing software.

With the image of New York to guide me, I used After Effects to arrange hundreds of instances of the 90 cards that came into the scene. By grouping and layering, I could move dozens of buildings at once, clone the groups, and quickly move the new groups to a new area, resize it, reverse it, or slightly change their colors. I blurred the buildings that were further away and slightly desaturated their colors, lightening them as they gained distance from the camera. Fake shadows, while not very accurate, were used judiciously to add variety and to enhance depth cues. A large, 4680 X 1648 image was exported for use in the 3D scene with the matchmoved camera.

The 3D building experiment did not have as many steps as the 2 ½ D test, but it did not lack complexity. The primary issue was in rendering dozens of 3D buildings, complete with large texture sizes, global illumination, occlusion, and shadows together. Preliminary experiments showed that, even with many of these features turned off, my computer did not have enough RAM to even begin to render all the buildings together. An alternative was needed. I used a process called “baking” as a shortcut around these problems.

Baking is like taking a picture of an object from all sides at once. In my 3D program, I could create a special camera that could record the data like this and let me use it as an image. This camera captured all the textures from multiple UV maps, bounced light, reflections, spotlights, shadows and combine them into a large Atlas-style UV map. An Atlas map divides images into discrete zones and appears to “unfold” the textures of an object into a flat square. My baking camera would not capture ambient
occlusion with the color channels, so a second exposure was needed to create an Atlas map for occlusion.

Ambient occlusion, or simply occlusion, has been mentioned several times in this text without much explanation. Occlusion passes were used liberally throughout the entire project and deserve clarification. As explained in the literary review portion of this paper, occlusion is the adding of dark areas to a render to simulate soft shadows. This is usually done in a separate pass and composited with other passes. Direct light sources are not needed in this type of pass because it is assumed that the sky is the source of the illumination. In reality, a well-lit environment typically has a large amount of light bouncing around the scene. Especially outdoors, a large amount of light is “bouncing” from the sky, softening the shadows and giving the appearance that light is coming from everywhere at once. This phenomenon is called ambient lighting. With ambient lighting comes a natural darkening around some areas where not as much light is scattered. This is called ambient occlusion. It occurs in corners and edges of objects (Birn, 2006).

Ambient occlusion is also achievable in 3D graphics and is often used in production. An occlusion pass can be made separate and can replace soft shadows created by a direct light source, such as an area light. Faster than photon mapping or area lighting, the process uses ray tracing to test pixels and see if they are occluded from the sky. The result is a soft, white image that has subtle, dark pixels in occluded or partially occluded areas. This image, or image sequence, can be laid over the color pass and “multiplied” with the color pass. Multiplication is simply the color values of each channel multiplied by the intensity of each pixel of the occlusion pass. Multiplying leaves no
effect in the white areas but leaves the composite darker in the dark areas of the occlusion pass (Birn, 2006).

In After Effects, I brought both the baked color and occlusion Atlas maps in and combined them by using a “multiply” blend node. This effectively used the occlusion map to darken the color map along seams and corners. After export, the new Atlas map was put into my 3D program and used to replace the previous textures. Now, I could turn off global illumination, all lights and shadows because those features had already been textured onto the building. All I had to do was turn up the illumination of the texture so it could be seen without the aid of external lights.

I went through and repeated this same process with all 15 buildings, capturing the results of all lighting and remapping those results as textures. As a result of this, I could multiply all the buildings to the volume that I needed to populate my futuristic city. I tried to arrange the buildings to match the 2 ½ D version; soon I found that the 3D city looked more realistic because the perspectives of the buildings were true. In the 2 ½ D version, the perspectives of those buildings did not line up with each other, since each building was photographed separately.

As a final step in producing the 3D cityscape image, I decided to use my 3D program to obtain a depth map of the entire image. A depth map is a grayscale image that represents the distance that objects lie from the camera. Darker objects are closer to the camera while lighter ones are further away. This image was taken into After Effects and used to modify the composite of the original 3D image that was rendered by the 3D software. The depth map allowed me to accurately blur the render in places where
buildings faded into the distance. It also helped me change the color gradually on those buildings that were further away, contributing to the appearance of atmospheric haze.

A few more subtle touches made my cityscape passes complete. First, the images needed to be imported and mapped onto large polygons in my matchmoved Maya scene. The 2½ D image had its own render layer, as did the 3D image. The large polygons were simple curved structures that filled in the camera’s view. When rendered with this camera, a pass would be produced that would match the motion of all the other passes. In addition, I added lens flares to each version to simulate the external lights that might be seen on the outside of buildings. Finally, I added a large lens flare to one of the taller buildings to create the appearance of the rising sun, reflecting in the glass windows.

When it came time to do the same to the blue screen version, the reflection did not make sense. The light angle of the sun was much too high to warrant a rising sun reflection, so it was omitted from the blue take effects. No lens flares, outside of building ornamentation, were produced in the blue take’s cityscape component.

Pulling the Matte

Now that the matchmoving and backdrop creation were finished, I decided to implement treatment of the live action footage and key out the green and blue backgrounds. This would make those backgrounds transparent and allow me to place the live action video element over the new backdrop and whatever other elements that I wished to place behind it.

The first step was to prepare the footage for keying. This involves making it easier for the keying program to identify and process the indicated color for removal.
Arbitrarily, I began work on the green footage and made an attempt to limit the number of green shades that the computer would have to process. Doing so would facilitate a cleaner matte and limit the amount of additional removal time needed. In a technique known as creating a “garbage matte,” I attempted to eliminate the extra green pixels that lay well outside the areas of interest. I used Bauhaus Mirage to paint out tracking markers before keying. This was a slow process, taking more than 12 hours to process all 1,280 frames of the footage and remove all the markers. In the blue version, this extra time was not needed as much. The markers used during the shoot were actually blue instead of the white color used in the green take. It took considerably less time to do the rest of the garbage matte. The footage was placed into After Effects, where green boxes were added to the shot to help mask off the larger, unoccupied areas of the colored screen. A new image sequence was then created from this treatment and used to replace the original footage. A similar process was used for the blue footage, as well.

With the new, treated video in place, I then used After Effects and a plug-in called Keylight to remove the improved colored backing. In doing so, I used a two-pass technique that featured a soft matte which provided soft edges to hair and complex objects. This matte was more sensitive but left “holes” in some areas. The second pass was a hard-edged matte and was added to provide solid cores in the interiors of actors and objects. These two mattes were combined into a hybrid matte that functioned well. This matte simultaneously had both solid cores and soft edge detail.

The technique provided an effective base but was far from complete for my needs. I needed to create “helper” mattes and use rotoscoping to correct problems that Keylight could not handle. Objects with finer details, like the walker used with the
moving tram or the bench that Kelly sat on, required extra help to make their mattes smoother. Rotoscoping was the only way to improve these mattes, isolating those problem components and issuing new keying parameters upon those components. After some difficulty in hand-rotoscoping those areas, I finally decided to use a combination technique where I could use After Effects’ motion tracker to guide my rotoshapes. A rotoshape is a vector-based spline, created in After Effects, that lets the user adjust its anchor points over time to mask off an area. By using the motion tracker, I could cut my rotoscoping time in half. With a rotoshape following the motion of the camera, I could essentially lock down the shape so it didn’t have to be moved so much. This technique, in tandem with a color-coded organizational process, let me produce a large number of rotoshapes in a reasonable amount of time.

By the time that the green sequence was finished with the matte creation process, 39 separate rotoshapes were used with 8 chroma-keying instances. Combined together, all the mattes formed a “super matte.” The live action video was then rendered as a premultiplied image sequence in 32 bit, Targa format with a built-in alpha channel.

The blue take was not as tricky as green, but it presented challenges, as well. I decided to use blue tracking markers instead of white, like in the green take. This proved to save time in postprocessing; however, I overlooked a problem during the blue shoot regarding the amount of wrinkles present on the blue muslin sheets. Behind the bench area, overtightening of one of the screws produced a dense fan of wrinkles in the fabric. This turned out to be a problem in the chroma keying phase when the background needed to be removed. The wrinkles presented a shift in hues that was beyond those of the rest of the backdrop. The result was a bulgy, uneven matte line along the bench’s contours that
occurred where the wrinkles intersected the bench. To fix this problem, it was necessary to rotoscope a shape around the bench to mask out the unsightly bulges.

58 rotoshapes and seven active keying instances were used on the blue take. The keyed-out sequence was rendered in Targa format just like the green footage, providing a testable, premultiplied counterpart that would later be juxtaposed with it.

Once both takes were processed and new, 960 X 540 premultiplied renders of each were completed, I was free to begin production on the many CG passes that went into the final composite. My plan was to use the green screen take do full testing of my main concepts and then adjust the scene files and assets to fit into the blue take. Secondary elements would also be created, outside of testing, and would also be re-used for the blue take.

Thus, I began a lengthy routine for preparing multiple layers of each object for pass rendering. Since the plan was to render separate passes for diffuse, reflection, specular, illumination and occlusion channels on nearly every object, many separate layers in Maya’s interface needed to be created. The number of layers soon soared into dozens as the production progressed. The amount of RAM required was soon beyond what my computer could efficiently provide. At this stage, several minutes sometimes passed before I could even access a different layer than the one I was currently using.

Animation

Nevertheless, the project continued. Models populated the scene. Soon, I was ready to do animation work. The previsualizational called for the animation of the holographic screens above Kelly’s head. These screens were to respond to Kelly’s hands
as she moves them around. Kelly had no guide as to how to move her hands. During the shoot, she moved her hands as though moving windows around in her smart phone, resizing them, and advancing lists of pictures by repeatedly tracing a path with her finger. She dutifully continued this motion until Casey tossed the soccer ball into her lap to get her attention, dissolving the screens and interrupting her workflow.

To simulate the five or six displays around Kelly’s head, I built square polygons in my 3D program and curved them slightly to provide a “windshield” experience. Some of the screens were to move from behind Kelly and slide in front of her. Since all the displays would rotate with Kelly around a common center, I needed to find a measurement in real space that coincided with Kelly’s pivot point. Boujou, when generating feature tracks in the live-action video, provided a cloud of feature tracks that imported into the Maya interface. After analyzing the 3D scene from several vantage points, I found that one of these tracking points rested exactly in the right place for a good pivot point. I parented my holographic screens to this pivot point and used it to anchor all the other movements of the holographics.

It took some time and finesse to match the movements of the screens with the movements of Kelly’s hands. Many of the motions needed extra attention in Maya’s graph editor, an interface that allows users to adjust the motions of objects while still keeping those motions smooth. Multiple layers were also necessary to make some of those elements appear that her hands passed in front of certain screens while allowing other screens to pass in front of the hands.

At one point, a soccer ball, tossed by Casey toward Kelly, triggers a dissolving effect that makes all the screens procedurally melt away. Using Adobe After
Effects, I created an animated texture of a circular pattern of digital static that expands and forms a white solid color. Imported into Maya and mapped to each of the screens in the transparency channel, the static made those screens transparent at the right time.

Similar animation work was done to Casey’s wrist phone hologram, a glowing image of another actress which was mapped to another polygon. Motion needed to be created to keep it “attached” to Casey’s wrist as she moved.

The most difficult animation work was certainly that of the computer-generated tram which enters screen left and slows to a stop at the main platform. To make the animation work look genuine, it needed to exactly match the motion of the practical rolling platform from shooting day. Before animation work began, I decided that it would be best to build an exact CG replica of the tram’s handrail—a simple aluminum walker found in a friend’s garage—to act as a stand-in object. The CG tram’s flight path carried it into the scene from far beyond where the live tram was in reality. The CG walker, which was parented to the CG tram, needed to be replaced with the live walker as soon as it was visible.

It was easiest to simply animate the tram and walker together and orchestrate the walker handoff at a later time. This would actually be done by using rotoscoping while the walker was traveling behind the actress. A simple matte was drawn that covered up the real walker until it passed behind Casey. Another matte was made that hid the 3D walker after it passed behind Casey. The result was a 3D walker, parented to a 3D tram, that passes behind Casey and is replaced by the real walker.

Animation of the tram was a separate matter. With the live action video overlaid into Maya 2008, it was a fairly simple process to hand-keyframe the 3D tram to
match the movements of the live-action tram. Without object-tracking software, the results were not as precise as I hoped, but certain methods were employed to minimize the appearance of slight mismatches. Casey’s shoes and leggings were fairly dark, so I textured the tram’s floor to more closely match those features, hiding problem areas of the animation. It took about the same amount of time to do the blue take’s matching as did the green take.

**Lighting**

With the main render layers in place and the tram fully animated, it seemed safe to begin the light-matching test that I had in mind for the foreground objects. Each foreground object had its own set of layers for rendering various passes. These objects included the large railings, the coffin-shaped recycler unit, the service pads in the near background, the tram, and the holographic displays. Since the displays had already been finished, there was no need to bother with them. I was only interested in the diffuse pass renders of the foreground objects and how to light them realistically. Three tests were planned: a final gather test, a photon mapping test, and a light dome test. The first two involved global illumination, which calculates the bouncing of light between CG objects. The third test, light dome, was considered “fake” global illumination and utilized an array of point lights to simulate environmental lighting.

The purpose of all three lighting methods, at least, in my project, was to realistically light CG objects to match the lighting of the live action video clips. The lighting needed to match the real-world lighting that we recorded on both shooting days. The blue take and the green take were recorded in similar, but not identical, lighting
conditions. Because of this fact, all three tests would have to be conducted on each of the takes.

As with the rest of the tests, I elected to start with the green video. The first step was to prepare the virtual scene for testing. A large ball, with inward-facing polygons, was created that encompassed the entire CG scene. This sphere was to serve as the actual light source that would power the global illumination in the shot. A special image was to be mapped to this ball to give the proper contrast to the illumination.

The creation of this high-dynamic range image, or HDR image, was actually done much earlier in the project. The description of this process was saved until now for illustration purposes. It involved the use of two computer programs to process the normal pictures of the light probe and composite them into the HDR image. Picturenaut from HDR Labs served to align the photos and merge them into an HDR format. HDR Shop was used to unwrap the final composite into a square image.

The new image was equipped with enough color data to allow changes to be made in exposure levels but still avoid the clipping to white or to black that can occur with low-dynamic range images. Exposure levels could now be increased or decreased by a large amount and the image could still retain a satisfactory amount of detail in the light or dark regions.

Once the HDRI image was created and successfully mapped to the scene-wide sphere, testing could begin on the lighting scenario. During the blue and green shoots, quick photos were also taken of a chrome ball, its reverse side which was painted a neutral gray, and a white ball which was brought in separately. Saved and cropped, these photos were brought into Maya and arranged side-by-side. Next to the photos, I built
three spheres that matched the size and color of each photographed ball. To ensure proper matching of colors, scans were made of each ball and color values were recorded (except for the chrome ball, which represented reflection values). Figure 5 depicts how photos of the actual light probes were brought into 3D software and placed next to 3D spheres for light matching.

When using global illumination, realism is achieved by calculating and simulating bounced light. The bounced light tends to “bleed” the color of one object to an adjacent object. This phenomenon is highly desirable because it mimics the physics of the real world and provides additional weight and presence to CG objects. Similar steps need to be taken in the combination of live-action objects and characters to a composite. Video elements do not cast shadows or bounced light onto 3D objects unless further steps are taken to ensure this.

A method similar to the 2 ½D process used in the cityscape backdrop was employed to provide further light contribution to CG objects. 3D polygons were built to provide further light contribution to CG objects. 3D polygons were built to accompany the geometry in the render layers. The method used was called image-based rendering, or IBR. In this procedure, the image of the live-action video was projected through the virtual camera lens and onto the new polygons. A second image was also projected. This image was a black-and-white silhouette of the green take’s matte channel and served as the polygon’s transparency attribute. The polygons retained the color of the original footage and also were transparent where they needed to be. When the polygons were
placed near one of the railings and made invisible, the computer could still render the bounced light and occlusion contribution from the live action characters. Kelly could appear to walk across the CG set and cast bounced light onto the 3D railing.

Methodically, the global illumination values and the color values of the HDRI sphere were adjusted to try to match the photos. It was highly iterative process, requiring constant render testing to compare my results with the photos. The final gather test was first. Before anything else could be done, an IBL sphere needed to be created. As described earlier, IBL stands for image-based lighting and works on the principle of an image, wrapped around a sphere, acting as the main light source. Using the method of raytracing, the 3D program establishes rays, or vectors, that emanate from the camera and analyze every pixel visible the 3D scene. The rays told the 3D renderer what color to make pixels based on where that ray would hit next. Since most rays would bounce from

*Figure 5.* Early light matching test shows 3D spheres placed next to light probes.
a 3D object and hit the IBL sphere overhead, most of the pixels would render with some type of sampling from the sphere. Final gather uses raytracing to add more importance to the other 3D objects in the scene, too, so bounced rays from the camera could make 3D objects light each other, as well. I continued iteration while changing the final gather preferences and the image-based lighting preferences until my 3D spheres matched the photos of the light probes. I then rendered all foreground geometry in only the diffuse channel with no reflections, shadows, or specular highlights. These other channels would be rendered separately and imported into After Effects alongside the diffuse renders.

The second test involved the addition of photon mapping to the image-based lighting scenario. Instead of shooting rays through the camera, directional vectors are released from the light sources in the scene. These vectors, called photons, bounce from surface to surface in the environment, testing to see if each can be absorbed, transmitted or reflected. When the assigned number of bounces from each photon has been reached, the position, intensity, and direction of that photon is recorded and saved into a photon map. In the case of my 3D scene, photons were released from the IBL sphere that surrounded the scene and shot inward, toward the center. A photon map was created to describe the diffuse hits on all geometry, irrespective of whether it could be seen by the camera or not.

Then, in the final gather step, rays were shot through the camera lens to analyze every visible pixel in the scene and compare those pixels with the photon map. There are many more rays present than there are hits from the photon map. Unavoidable gaps in the map, if not reconciled, could produce a final render that flickers. To correct the disparity, the raytracer also searches nearby pixels and “gathers” any photon hits
found in proximity. The information from these nearby photon hits are averaged, and a
new set of rays are emitted randomly from the reconciled pixel. These rays return
information about what other light has landed on that spot. The values of the photon map
and final gather are combined for every pixel and rendered.

I found that somewhat large photons worked best for my project. With
181,820 photons emanating from the IBL sphere, a radius setting of 0.03 worked best.
Each photon hit was visible on the foreground geometry as a splotch of light, about the
size of a fist (relative to the size of the railing). The radius setting described the total size
of each photon. From software documentation, it was not possible to know how to judge
the radius setting except by trial and error. With final gather turned off, it required dozens
of single-frame renders to discover that the photons should appear to just overlap each
other. The settings of 0.03 radius with 181,820 photons gave me just the right balance of
smoothness (once final gather was added) and accuracy while maintaining a flicker-free
image sequence. The final render came out a bit brighter than the final-gather-only render
and revealed that a second bounce of light contributed to the light solution. It was a
simple matter to transfer the settings from the green take lighting to the blue take and
render all assets.

The final test for creating the diffuse pass of foreground geometry was called
the “light dome” or “light rig” method (Bloch, 2010, Chapter 7, Section 6). It has been
said that this was a worthwhile technique to try because it yielded positive lighting results
without having to wait for photon mapping or final gather. Instead of indirect lighting,
like in the latter two techniques, the light dome technique involves the creation of direct
lighting sources in a 3D program and “faking” the soft lighting of global illumination.
Essentially, an HDR image is evaluated and an array of point lights are generated to match the values of the image. The lights are then imported and placed in various locations but at a fixed radius from the 3D scene’s origin, effectively forming a sphere of lights around the 3D space.

My first step was the evaluation of the HDR image. I used a program created by Binary Alchemy called HDR Light Dome, version 1.3.05. After loading my image into the program, I was prompted to apply color from the image to special “points” that would be laid out over the image. These points were assigned RGB values and luminance values as sampled from the HDR image. By default, the program offered 642 points and laid them out in a fixed pattern over the image. If left alone, 642 points would generate that many lights and cause a real strain on the render engine. It needed to be optimized so the number of lights could be reduced, but still give a soft result.

The process of optimizing the number of points took three steps to complete. On the first attempt, step 1 reduced the number of points to 142; however, step 2, apparently a more detailed step, took over three hours to process. Eventually, the program reported an error and aborted the optimization.

On the second try, I was prompted by the program to lower the number of points that step 1 analyzes. The default was set for 225 points, so I rounded that down to 200 and tested the effects. The remaining optimization steps finished running, successfully leaving me with 121 points. This still seemed like too many lights to work with. Indeed, upon further research, I found a reference that inferred a reduction to a range between 40 and 50 lights (Bloch, 2010).
HDR Lightdome also had a feature that helped reduce the number of lights. I left the settings at their defaults: lights with luminances lower than .08 would be deleted. This action brought my total light count down to 60. At the risk of losing too much detail from the lighting solution, I accepted this number and proceeded to export the light rig to Maya. HDR Lightdome actually created a special script that was saved as a separate file.

When I loaded this script into Maya and ran it, 60 point lights were generated in a sphere around my 3D objects. Point lights shine in all directions and do not need to be adjusted rotationally. Falloff was also disabled on each of these lights, so the distance from objects was not a concern. Initial test renders gave quite good results.

With shadows turned on for each light, I expected a grainy result in the renders. In my research, I was warned that blotchiness could occur and a special “spinning lights” technique would need to be employed to eliminate the effect. This technique calls for each of the lights to be rotated twice around an off-center pivot point for every frame of animation. As long as motion blur is enabled, the blotchiness would be flattened and minimized.

My expectations appeared to be unfounded. I did not experience the predicted grainy renders. I found my renders to be very similar to the photon mapping and final gather techniques. I used the light dome technique to render the remaining foreground passes.

Many passes of each piece of geometry were isolated and rendered separately. The HDRI-lit diffuse passes and the ambient occlusion passes were not the only passes of geometry made. Specular and reflection passes were also created and rendered. Each foreground element, including the mid-ground building, both railings, the recycler unit,
the floating tram, and the armlike service pads behind them, (see Figure 6 for an outline of the passes) all had multiple passes to maximize control in the compositing step.

A Description of Passes Rendered

Ten levels of renders needed to be done for this project. Excluding the live-action layer, each level represents a different distance from the camera. Like paintings on glass, each level lays on top of one another to form the entire picture. Nearly all renders produced a built-in alpha channel to allow compositing, with the exception of I.D. passes.

Identity passes, also called I.D. passes or matte passes, are simple three-color renders of important zones in a 3D object. With the help of a compositing program, these colors can be made into matte channels that are similar to green screens or blue screens. Parts of a render can be isolated using their corresponding I.D. colors as guides. The first level, containing concepts that exist the farthest away in space, is the cityscape backdrop. This level is made up of three separate renders. The first, as described earlier, is the color backdrop with the buildings, clouds and sky. Two versions of this backdrop were created for testing, a 3D version and a 2 ½ D version.

Accompanying this backdrop were two more renders of the lights that belong to various buildings. The first of these renders was a distant light render that contained lights which were far away and little more than blurry halos on distant buildings. The second “light” render was sharper, brighter and contained lights that were closer to the camera. These two renders were done separately to simulate depth.

Level two, overlaid on top of the backdrop level, was the distant tram render, where three trams slide along the horizon. These trams were rendered without passes,
allowing 2 ½ D cards with actor footage to ride along with the geometry. Because these objects were distant and would not be viewed in detail, a single render with no shadows and a direct lighting solution was sufficient.

The third level was the mid-ground building, called “Telequark” for its blue signage. This building was unique from the others because it was much closer to the camera and lent itself to adding further perspective to the composition. It also included glows, a large spotlight in the front, and a window with video footage mapped to it. This level required five passes to fully realize my artistic vision. I chose the final gather global illumination method to render the diffuse pass outside of testing. Since the building was still somewhat far away, it seemed right to maximize efficiency by choosing the faster indirect lighting method. There were no reflection nor specular passes here. Instead, certain important elements of the building were isolated. Along with a diffuse pass and an occlusion pass, I also added a spotlight pass, a window pass and an I.D. pass. The spotlight pass was simply a normal diffuse pass (with no global illumination) of the building, lit only by a single accent light that was placed at the front of the building. There was a texture on the geometry in the shape of a large “41” on the wall there. I wanted there to be a spotlight that illuminated this feature, but not interfere with the global illumination scenario. Using a “screen” blend technique, the spotlight could be overlaid and manipulated separately.

The window pass was simply the window polygons with live-action footage mapped to them. All geometry was made invisible in this render except the window. The separate render made it easy to make the window brighter or dimmer as needed.
Figure 6. 3D elements and their passes for the green take of *Soccer Mom: VFX Supervisor for a Day.*
The I.D. pass, as explained earlier, allowed the compositing program to separate other renders into zones so I could manipulate them individually. In this render, I changed the colors of certain zones to red, blue and green. The window geometry was red, a side wall was green, and the Telequark neon sign was blue. Everything was made flat with no other lights, no darker corners or textures. The window color let me add columns to the window and the blue color let me brighten the neon sign and add a glow.

Level 4 contained two renders of the “service pads” as I call them. When designing the platform, I wanted some geometry beyond my floating tram to catch the moving shadows from the tram. A small pad or two, adorned with a couple of vertical rods, seemed right to add to the depth and provide a little more high-tech feel to the composition. I could imagine a technician or robot perched on one of these little pads to make some fast repairs to the floating taxis. I felt that only a diffuse pass and an occlusion pass would be fine to achieve the weather-vane look of the service pads.

The fifth level was a complex one. The floating tram was being asked to do a lot of things. It had a CG duplicate of the real walker from the video footage. This geometry would be swapped out for the real footage once it entered view. The tram functioned as a shadow catcher for the raytracing/shadow mapping test; it had illuminated lights on the front, a holographic display on the rear, and even responded to bounced light from the live action video.

To produce all of these intended effects, nine renders were required. I needed maximum control over this complex element, so every major channel was rendered separately. The diffuse, or color channel, of the tram was rendered without shadows or global illumination. Essentially, it was pure, flat color. This pass also included a rear
holographic display that turned from red to green when the actress stepped aboard. I used an area light to add extra illumination to the tram’s polygons that were near the display. This simulated spill light that came from the display.

Other tram passes were fairly straightforward. The walker was done as a separate pass and needed little extra modification. The small illuminated panels on the front of the tram were rendered in an illumination pass and were used to adjust the brightness of the panels and to create glowing halos through the compositing program. Specular and I.D. passes were also rendered with little difficulty. The I.D. pass was extremely helpful in separating zones for manipulation in the editing program.

Three passes were created with the aforementioned 2 ½D casting method, wherein live-action footage of actresses were projected onto polygons and placed near the tram. This technique was also used with the 3D railing. The live-action projection was used to add a light contribution to each of the three passes. These passes were global illumination, ambient occlusion, and reflection passes. The reflection pass was simple to create and cast realistic reflections of the video footage onto the tram, especially in the spot where an actress steps aboard it.

The global illumination pass had the same effect, only that it added a bit of spill light from the actress’ red shirt onto geometry. This render, like the Telequark render, was made using the final gather algorithm. Only one bounce was needed to generate the desired effect.

The ambient occlusion pass, measuring the partial blocking of rays near intersection points of geometry, presented a different challenge. What I desired was a slight shadow to appear across the tram’s geometry as the actress crossed the threshold
and boarded the tram. I placed a polygon horizontally over the area where she would have stood in reality. On this polygon, I attempted to map live-action into the occlusion shader that Maya used to create occlusion. The process proved to be rich and difficult. Many research sessions on the Internet were needed to find the right procedure. Only the extensive juggling of shader nodes, connecting and reconnecting of channel links, and perseverance finally saw success. The shader eventually recognized the transparency channel of the live action footage and forced occlusion into the correct shape. A soft, subtle dark spot could be seen boarding the tram along with the real actress.

It may be asked, “why not just render a shadow using a direct light source and a normal transparency map?” The answer was in the testing of my actual shadow pass, which comes into relevancy at this point. The nature of my shadow test was to query about the efficiency of one shadow-generating method over another, using a direct light source. Ambient occlusion does not use this technique, only an indirect method, as though light existed everywhere. If I were to use a direct light source to cast the actress’ shadow onto the tram, I could only do it with a raytraced light source. My shadow test involved comparing a raytraced method to a shadow mapping method, which does not allow transparencies. My shadow test would be invalidated.

This discussion brings us to the ninth and final pass of the tram level, the shadow pass itself. This pass is where my shadow testing took place. It was designed to add a subtle flavor to the tram and provided a sense of weight to the CG railing. As the tram passes by the railing, the railing’s shadow falls across the tram. I limited the testing to just the tram level to help mitigate the time spent on this test. As well, the tram was the
only moving CG object in the foreground. It was the only object that could benefit from the precision of directed shadows (shadows for artistic purposes).

I created a fresh render layer in Maya and added the ingredients to conduct the raytracing test for shadows. The floating tram, the platform railing, and a direct light source went into the layer. This was supposed to be an outdoor scene under a cloudy sky. To do the test, I needed to suggest that there was a modicum of sunshine present. The green video footage implied that diffused light was coming from behind the camera’s position, while, in the blue video, the light emanated from a higher angle. I desired subtle shadows to fall across the tram through the rounded triangles of the railing. Although I liked that angle and effect, it did not make much sense for hard shadows to come from that direction. To compensate, I decided to create softer shadows and make them more subdued than in a direct daylight environment. So, even though the CG light source did not match the light from the actual video shoots, I exercised creative license and placed this directional light source to get my desired effect.

I chose a directional light because it closely approximated the near-parallel rays of the sun. I turned on ray tracing and set the light to a moderate intensity, about 1/3. I then set the layer’s attributes to render shadows only. Initial test renders of the shadows seemed grainy, so I set the number of rays to 28 to improve the quality. These rays were fired from the camera onto every visible pixel of the scene and bounced toward the light, testing to see if that pixel lies in the shadow of an object. The light angle setting, used to increase or decrease the softness of the shadows, needed to be set low, to about 4 degrees. I wanted a definite shadow pattern to fall across the tram and 4 degrees kept the
shadow from losing its shape. I would later fade and blur the shadow renders in After Effects.

No other changes were needed with the ray traced shadow tests. This shadow pass was rich and detailed. It could only be viewed through its alpha channel, since it was made up of black pixels with a transparent background, which also appeared black. Once the render was complete, I was able to move on to the shadow mapping test. I duplicated the raytracing layer and changed the settings on the main light. For raytracing, I had used a directional light with parallel rays. If I used the same kind of light for shadow mapping, the required settings for a smooth shadow may have needed too much RAM to complete. Instead, I changed the light to a spotlight and turned on shadow mapping. Viewing the 3D set from the light’s position, I changed the light’s cone angle, or angle of illumination, to encompass the railing and tram. At the distance that the spotlight was from the subject matter, I found that the shadows were very similar in shape and intensity to the shadows from the raytracing test.

To finish the test, I wanted to get the depth-mapped shadows to visually match the ray traced shadows as closely as possible. To do this, I decided to use Mental Ray’s built-in plugin to override Maya’s depth map lights. According to research, Mental Ray’s rendering algorithm can provide better results than Maya’s depth maps (Derakhshani, 2008). I set the depth map’s resolution to 500, increased the map’s samples to 400, and gave the shadows a softness of .005. Maya’s documentation suggested that the scene’s resolution, 940, would be an ideal starting place for the shadow resolution. After some experimentation, I found that 500 met the needs of my experiment and saved a little render time in the process. The softness value of 0.005 appears negligible, but it helped
compensate for the lower map resolution just enough to keep the shadows from appearing pixelated.

Like the raytraced layer, I rendered out the finished shadow mapping layer. I used the OpenEXR file type to help keep the file sizes low (79 KB vs. 2 MB in .tga format). In the compositing program, both premultiplied sequences were brought in, blurred and reduced to 35% opacity. This achieved a subtle presence and matched the soft lighting scenario.

The previous narrative of the shadow testing concludes the description of the tram layers. Two of the remaining foreground levels were rendered in mostly the same way as one another, one railing occupying the space behind the actresses, and one railing occupying the space in front of the actresses. In the forward level, there was also geometry representing a squat concrete structure called a “recycler,” which has no apparent use except to display the time, date and implores viewers to have a nice day. Both levels of geometry, front and rear, were rendered virtually the same way. They both contained diffuse, occlusion, specular, reflection, and I.D. passes. The diffuse passes were lit with the three image-based lighting methods as described earlier. The final gather, photon mapping, and light dome methods were each performed on both levels of geometry, and placed with the other passes for testing in the compositing program. The main difference between the two levels was that a display screen pass was needed for the recycler’s display screen.

Between the two railing levels was where the live action video and the holographic displays were to be placed. The live action video was already finished at this point. The two versions were already premultiplied and placed into separate compositions
for testing. Animation had to be done twice on the moving displays to match the actress’ hand movements for each take. The render requirements were the nearly the same for both green and blue takes. However, since the actress’ hand movements were different between the two, I felt the need to make changes. The green take displays called for more of a modal window kind of display. Kelly moved her hands in the air as though moving windows around and resizing them. During the blue shoot, Kelly did likewise, except that the majority of her hand movements mimicked the scrolling motion that one may use with a smart phone or tablet. She motioned with her index finger as through scrolling through a list of photos. For the blue take, I created two scrolling windows to match Kelly’s finger movements. These were not in the green take.

For each color take, eight or nine separate renders were needed to achieve the effect of the holographic displays that glowed around Kelly. There were several screens of different sizes. Not only did I render the semitransparent pixels for the screens themselves, but also I rendered the glowing blue edges of those screens. There was a set that went in front of Kelly’s body and a set for behind. In the compositing software, one or two elements were moved from the rear to the front. This proposed a bit of a challenge, but was solved by copying that layer and applying a hand-drawn rotoscope matte to it at the right moment. This manual mask was drawn around Kelly’s hand and arm when the screen was supposed to fall behind her. The mask was withdrawn once that screen was moved to the front.

Other displays were created in separate passes. At a certain point in the action, a glowing 3D polyhedron appears, grows in size and spins gently in space. The words “Encapsulation 01” appears near the shape. The meaning is just pseudo-tech speak but
seemed to fit the science fiction atmosphere of the shot. The shape and text both had their own render passes. To augment those passes and make them appear as projected images, secondary passes were made to establish glowing halos around those items.

There arrives, finally, a termination to these holographic elements when Casey tosses a soccer ball at her mother, disrupting the images and dissolving them. I created an animated transparency map that was to fall across each of these elements. With After Effects, I made a moving, black-and-white pattern that resembled TV static, and then used an expanding circle shape to dissolve the static over 15 frames. This was rendered as an animated .jpeg image sequence and mapped onto each of the holographic shapes. A little time offset on each shape provided a smooth dissolve effect that dissipated the shapes at varying intervals.

The last display was a holographic video-phone image that appeared to be projected from Casey’s wrist. The image was taken from live-action footage of an actress shot before a green screen. The actress, whose name is Christy, acted for the camera as though she was speaking to a friend into it. I took this footage, used a chroma key process to remove the backdrop, and imported this into After Effects. Here, I added a flickering effect and semitransparent video scan lines to the footage. The result was rendered out as a black-and-white image sequence and mapped to a polygon that was animated to match Casey’s wrist. Although its presence was brief, that small amount of animation added a good deal of technological character to the scene as a whole.

The final level to be rendered was a pass for adding occlusion to the floor. This process, as discussed earlier, mimics the natural darkening that occurs at the intersection of geometric structures. This darkening occurs naturally in the live-action
video and needed to be reproduced in my 3D objects. The most noticeable area of this
dimming is on the asphalt surface where actress’ feet make contact and where the bench
sits. When the 3D geometry is added to the composite, like the railings and the recycler,
it becomes evident that there is no darkening on the asphalt around these new structures.
To supply this new occlusion, I created a new layer in Maya and added a large, flat
polygon to occupy the area where the asphalt would be.

I found that a simple preset, hidden in one of Maya’s render layer submenus,
was all I needed to do to set the ambient occlusion correctly. The preset turned on
raytracing and turned off final gather and global illumination, which were not needed.
Also, a simple shader network was loaded onto each object that occupied the render
layer. This layer was only supposed to represent the darkening upon the platform floor
that would be caused by presence of the recycler and the railings. Those items were made
invisible to the camera but were allowed to contribute to the ray tracing solution. For
each ray fired through the camera at a pixel in the scene, 16 more rays were released from
each pixel in the camera’s view. These rays sampled the sky and tested to see if there was
any occlusion from nearby geometry. The shading information from each pixel was
returned to the camera and the pass was rendered.

After the pass was finished, I could bring it into After Effects and combine it
with the live action video level. I could control the opacity of the occlusion shadows to
precisely match the natural occlusion cast by the actresses and real objects. This worked
very well, although further refinement was needed. The floor occlusion layer cast some
shading where it did not belong. Shadows crossed in front of the bench and in front of
actress’ legs. Extensive rotoscoping was needed to create masks and isolate the bench and legs where the occlusion occurred.

Most of the other foreground levels had their own occlusion layers, as well. Five more occlusion passes where needed for different foreground elements. Three of these passes, one for the Telequark building, one for the service pads, and one for the front railing geometry, could be accomplished with default settings just like the floor. The service pad occlusion pass was made by including the tram in the scene as well. The tram passes close to the service pads and likely would add occlusion to it. These passes were rendered quickly and needed no major adjustments.

The remaining two occlusion passes, belonging to the tram and to the rear railing geometry, needed a bit more finesse. These items would appear to be in close proximity to actresses. I wanted the live action footage to influence the occlusion of the railing and the tram, darkening them slightly as they moved closer to those objects. This objective would require a transparency component to the solution. Like the reflection passes for live action video, the alpha channels would need to be borrowed from the chroma-keyed live video and used to create transparency in the occlusion.

Unfortunately, the built-in occlusion shader used in the floor layer would not accommodate my plans for the transparency of the live-action footage. I would need to create a custom shader in Maya to get what I needed. After viewing many tutorials on the subject, a complex array of texture nodes were needed (Keller, 2009). In essence, the new shader network used final gather as a replacement for the standard occlusion shader. This shader employed a secondary transparency node to allow final gather to “see” my live-action transparency channel. I imported the same stand-in geometry that I used for the
reflection passes of the railings and tram. To these, I applied the new shader network and turned on final gather. I was able to render soft occlusion in the shape of the actresses, adding significant realism to the composite. The occlusion followed the actresses as they moved around the set, casting on the railing and tram. This was done for both the tram layer and for the rear railing layer.

Compositing

The creation of over forty CG passes (eighty if the renders for the other color take is taken into account) was only one part of the total production process. Much more was to be done if all assets were to function together successfully. The compositing phase of the project was equally important—and time consuming—as the creation of any live-action or CG element. My compositing software provided the means and the tools to blend these elements together in a seamless way. There were several noteworthy techniques that bear mentioning, including premultiplication, track matting, I.D. matting, and blending.

Nearly every element that was composited together had in common a built-in alpha channel that controlled the opacity of the element. Even the live-action portion, once the chroma key work was finished, was exported with an alpha channel. To save computational resources, all elements came premultiplied with their alpha channels and were matted with color. Premultiplication saved time because the elements’ color channels were already multiplied with their alpha channels, so the compositing program was not burdened with this additional task. Elements with the multiplication process
already finished could be combined in a “normal” or “over” process, sometimes called A over B. This is the most common process in compositing.

There were instances where the mistake was made of premultiplying an element and then multiplying it again for a specific purpose. Notably, in the case of ambient occlusion passes, I made the error of rendering the alpha information along with the color information. It may be remembered that ambient occlusion supplies extra darkness in corners and intersections of objects. The actual render is a grayscale image sequence that must be multiplied with the image underneath it. The act of multiplying takes the values of every pixel in the occlusion pass and multiplies those values with the values of pixels lying beneath. The occlusion pass darkens where its own pixels are dark while leaving pixels alone in places where its own pixels are pure white.

My occlusion passes were rendered with a premultiplied alpha channel that made the background transparent. This was a mistake. I still had to blend the occlusion pass with the layer underneath with a “multiply” operator to obtain the desired effect. The multiplication had already been done, however, but with its own alpha channel. Multiplying it again, with the underlying layer, produced a white outline around the objects that were to receive ambient occlusion. The value of semitransparent pixels along the edges were calculated to the wrong product and then subtracted incorrectly, leaving a white halo around objects. Figure 7 illustrates this phenomenon.
This is a common problem, but is often misunderstood. Many new artists would erode the matte by a few pixels to get rid of the fringe, but in so doing would degrade the quality of the composite. The correct solution is to undo the premultiplication process so the multiplication only is done once, and with the correct pixels (Brinkmann, 2008). In my case, that did not apply because my render included a black background. Proper occlusion usage demands that the backgrounds are rendered white so composited background pixels will be unaffected. I was required to re-render all my occlusion passes so the backgrounds were white. I did the re-render without any premultiplication added and produced an occlusion pass that worked without the white fringe.

Of the myriad of potential blending techniques for elements, I only had need of a few. “Over,” of course was the most used, which is simply A over B. “Multiply” was also used extensively, albeit already done at render time by premultiplying the renders from the 3D software. The live action video, as well, was premultiplied after the

**Figure 7.** Double premultiplication problem and corrected image.
matte extraction was done, leaving an image sequence with the background already removed. “Multiply” was used in the main composite only for the ambient occlusion layers. As reported above in the description of premultiplying, the occlusion renders were not premultiplied and had no alpha channel. “Multiply” was the only way to use the occlusion pass to darken parts of the composite as needed.

“Add” was used for specular passes. Specularity is the reflection of highlights that appear to bounce from shiny surfaces. An “add” node sums the pixels of the specular pass with the underlying layer. This blend mode produced bright specular highlights that stood out better than by using a regular “over” or “normal” mode.

A “screen” blend mode was used with illumination passes of the tram’s glowing panels, as well as other glows. These areas needed to be bright, but an “add” node made these too bright, clipping many color values to white. “Screen” gave more subtle, attenuated results than “add” by inverting pixel values of both layers, multiplying them and then inverting them back again. “Screen” also gave good results with most reflection passes.

Some reflections, however, did not benefit as much from “screen” mode. Even the attenuated nature of “screen” sometimes gave results that filled in and flattened the dark areas underneath the reflection pass. The “overlay” node solved the problem by allowing the lower layer to be unaffected by the reflection pass in areas below 50% gray. Areas above 50% could receive the effects of the “screen” operator. Reflections could now be seen in bright areas of the composite while darker areas could retain their original detail.
There were many occasions when rendered layers needed extra help in accommodating the live action video. Occlusion layers that affected background layers would cast darkness along the feet of the actresses. Hands and arms blocked holographic screens that should have appeared in front of them. The live soccer ball disappeared behind more screens and behind occlusion layers. These and other problem areas were handled with the use of trackmattes. These were simple nodes inside my compositing program that essentially used another layer and multiplied it with the problem area. Trackmattes were always black-and-white sequences that quickly let me change its parameters for multiplication. With them, I could rotoscope an area and draw masks as needed. These masks could screen out areas and allow the live action to be seen through them. I could also use pre-rendered black and white sequences from my 3D software to create new masks (see Figure 8). This method was helpful in removing the foreground from the live action layer that was supposed to be the empty space beyond the forward edge of the railing.

There were times when I did not know how many masks I would need. Some objects were complex and it was important for me to get them right. The tram, for example, that the daughter rides out of the scene was a particular challenge. It had multiple textures and surfaces that interacted with live actors. I needed shadow passes, specularity, varying degrees of global illumination, occlusion, and reflections that included the live actors in them. There needed to be a non-shadowed diffuse pass, lighted panels and glowing halos. A holographic sign on the rear of the tram changed colors. The complexity of the tram made it necessary to assume maximum control over the top,
front and sides. This would require a separate pass for each attribute and for each side of the tram. This would have taken too much time to create all those passes; instead, a single I.D. pass would let me split the object up into zones and let me create new masks from it.

In this pass, I turned off all but the simplest shading, making the top of the tram green and the edges red. The jet-like engine and the half-sphere on the bottom of the tram were shaded blue. I rendered the entire sequence with only the tram and its three colors. Then, I used this pass in my compositing program to create three new masks for the tram. After Effects could identify each color channel and make a new mask for each. Now, I could render the rest of the passes for the entire model and use the I.D. sequence to separate each pass into zones that I could manipulate independently of one another. This method
worked so well and saved me so much time that I used it on the railings, the recycler, and the windows of the Telequark building.

These and other, more minor methods were used to arrange and polish the elements of both the blue and green versions of the composition. In both versions, separate composites were rendered with the different testable concepts in place. A composite was rendered for each shadow type, each light-matching technique, and for each backdrop type. There were 12 combinations of methods for the blue version and for the green version, giving a total of 24 separate movie files that represented every possible combination of the above techniques. Watermarks were also added to each movie file to differentiate the combinations.

Creating the Application

I created the application interface, entitled Soccer Mom: VFX Supervisor for a Day as a platform for presenting all 24 renders in a comparative way. To help users understand these comparisons, they are prompted to engage in the editing process—at least the editing thought process—by choosing which “ingredients” will comprise the main categories of elements. A component of this experience is a viewing of the finished render with the user’s choices, along with a graph that indicates what percentile that their particular combination fell into, when juxtaposed with choices made by peers. In other words, participants can compare their choices to those made by others.

I used an old version of Adobe Flash (this version was actually Macromedia Flash MX) to design the interface with the intention of creating an attractive, interesting format of presenting the content. The first screen made an introduction to the program,
explaining the goals that users were to engage in. Players were invited to go to
“orientation” and then return and begin their first day on the job as a visual effects
supervisor. Orientation actually was a video file that explained the premise of the
program in the form of a situational narrative. In the video, the director of the fictional
visual effects company “Sweet and Dill Pixels” explains that the usual visual effects
supervisor has left with his work unfinished, and it was the user’s job to finish the work.
The work in question was a single clip of live action video that was to have several
different forms of computer graphics added to it.

After viewing orientation, the player is then directed to the main screen, which
is the decision list generator screen. There, users can see all four decisions that they have
to make. Radio buttons allow them to select their choices. Near each choice, participants
can view further video files that explain the nature of each decision. Next to each video
file is a button that lets the viewer see samples of the element in question as rendered
with each technique. The live action video section allows players to view samples of the
blue take and green take, with or without their premultiplied alpha channels, side-by-side
or alone. Each choice also indicates how much time it will take to create. The user is
prompted to select the choice that has the best balance between aesthetics and cost-
effectiveness. Along with the live-action decision, players also were to choose methods
for completing the diffuse pass for CG objects, the shadow type on the CG tram, and for
creating a cityscape backdrop for the entire scene. A glossary is also provided to help
explain some of the more complex terms in the videos.

The player can take as much time as desired to complete the decisions. Once
satisfied, the user can then submit the choices for final review before approval. Here,
players learn if their choices are within budget restraints, and can re-do the choices if they wish. Either way, participants may move forward to the final submission screen. Here, they have a chance to see a preview of the choices before ordering the crew to continue. This is where users can see one of the 24 renders that corresponds to their choices. The preview is loaded into the default web browser available on the computer used for playback. If all looks well, players can finalize their decisions by clicking the “send choices to the crew” button.

The final payoff is a video that plays after the last button is clicked. The video is one of three that can possibly play. One represents the results if the green take is selected; the second plays if blue is selected and stays within budget; the third starts if blue is selected but goes beyond the budget. All three are a general summary of the entire project and add several narratives that explain how other effects in the shot were achieved by the crew. See figure 9 below for an organizational diagram of the comparative application.

All references to crew members in the videos are fiction. Every function of the Sweet and Dill Pixels postproduction crew was actually performed by me. A method for data collection was needed to provide figures for the choice comparison scoring and for the eventual recommendation of combined techniques. The project itself proved to be the most likely data collection tool. It was modified slightly to conform to human test subjects requirements and deployed with a linked, online survey. The survey served as a collection tool for the data generated by my program and added questions regarding the effectiveness of the program.
Figure 9. Organizational DVD flowchart for Soccer Mom: VFX Supervisor for a Day.
From a population of 500 computer graphics students in 25 colleges and universities across the United States, 66 respondents used the program and participated in the online survey. The survey, held from September 20, 2013 to March 29, 2014, was launched with an initial E-mail campaign of 388 requests from 10/20/2013 to 12/10/2013. An additional 42 requests were sent between 1/16/2014 and 2/28/2014. On February 12, 43 follow-up E-mails were sent to colleges that had accepted the initial request during the previous semester. A total of 473 E-mail requests were sent for this survey effort. It generated a response rate of 13.2%, yielding a margin for error of ±11.25%, at a 95% confidence level.

Online tools were used in the implementation of the survey and the evaluation of the data. The survey was created with the Survey Monkey online creation tool and hosted on Survey Monkey’s website. It can be viewed at http://www.surveymonkey.com/s/RXT7ZW6. A copy of the survey has also been supplied in Appendix B. Also used was an online calculator for computing margin of error. It was found at the American Research Group website at http://americanresearchgroup.com/moe.html.

The collected data, in turn, was used to help formulate a recommendation and help support my conclusions in the Summary, Recommendations and Conclusions section of this paper.

Results

This section will address the results of testing the techniques of each of the four concepts—screen color, light matching, shadow type, and backdrop style—and
report the amount of time that it took to complete each method. Subjective sampling is also incorporated into these results, taken from an online survey.

My observations revealed variations of completion time between techniques and, by extension, variations of cost in labor. Described below is an itemized report of each test’s results.

The time it took to process the green screen footage was 57 hours and 16 minutes, requiring 39 rotoscoped masks and 8 active keying instances. The blue footage took 80 hours and 33 minutes. The blue footage needed 58 rotoshapes and 7 keying instances to complete the extraction.

For the light matching tests, final gather took the least amount of time, requiring 14 hours to complete. The light dome technique finished second at 33 hours and 15 minutes. Photon mapping took the longest amount of time, needing 37 hours and 15 minutes to complete.

The shadow tests revealed that shadow mapping took 45 minutes to complete while raytracing took two hours and 45 minutes to finish.

For the backdrop tests, it was found that the 2 ½ D method took 17 hours to do while the 3D method took only 10 hours.

These figures were placed into the interactive application and presented to users of the program in an online survey. Participants of the survey were asked to evaluate the methods and choose one for each concept. They were asked to pick the highest quality at the best value. As part of the role-playing scenario, users were told that the director needed the post-production budget to stay within $5,800, and that there would be career-changing consequences if that budget was exceeded.
65% of users chose the green footage over blue when taking into account the quality of the resulting matte, the performance of the actresses, how pleasing the composition was, and overall cost value. 46% of the participants who chose green stated that they did so because it was a good value, or balance of cost and quality.

When participants were asked to compare light-matching systems, parameters included realism, pleasantness, quality, and value. From all three light matching tests, photon mapping drew the most favorability at 45%, while final gather scored 42% and the light dome method scored 13%. Of those who selected photon mapping, 37% stated that they chose it because it looked the most realistic. Photon mapping also drew a higher pleasantness ratio than the other tests, 31%. Of those who chose final gather, on the other hand, 31% stated it was for the value of the technique. Another 31% chose it for its realism. The majority of those who chose light dome also chose it for its realism.

During the shadow type test, more users selected raytracing over shadow mapping as the preferred shadow technique. Raytracing scored 72%, 54% of its supporters stating that they selected raytracing for its realism. The shadow mapping technique, which earned 28% of the vote, primarily was supported for its quality vs. cost value, as indicated by 70% of its voters.

The backdrop style test revealed a 70% support of the 3D backdrop style. Of those who selected this style, 48% chose it for its realism. The 2 ½ D method, which won 30% favorability, did so mostly on the merit of its pleasantness score of 40%. The backdrop topic drew comments that 2 ½ D had “errors” and looked “muddier,” “fake,” and that the perspective disparities were “distracting.” For the whole breakdown of voting statistics for all of the four concepts, please see Figure 10 below.
As part of the online survey, participants were also asked to indicate how they were affected by the application. A broad range of questions were asked about how players felt about live action/CG integration, both before starting the program and after completing it. Questions were asked about what interest the player had in visual effects and how important live action/CG integration was to them. A follow-up series of questions came next, asking if, and how, their preconceptions about visual effects may have changed after using the application. To obtain the maximum amount of input on the answers, users were allowed to mark all answers that applied. Correlations were discerned by averaging each choice independently against the total number of respondents. These included questions about interest, perception of difficulty, and the use of CG passes.

**Online Survey Results - Soccer Mom: VFX Supervisor for a Day**

Percentages reflect the reasoning of those who made that particular choice.

**Color Tints**

<table>
<thead>
<tr>
<th></th>
<th>Votes</th>
<th>Performance</th>
<th>Aesthetics</th>
<th>Matte Quality</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td>43</td>
<td>0%</td>
<td>12%</td>
<td>40%</td>
<td>46%</td>
</tr>
<tr>
<td>Blue</td>
<td>23</td>
<td>16%</td>
<td>25%</td>
<td>65%</td>
<td>13%</td>
</tr>
</tbody>
</table>

**Light Matching**

<table>
<thead>
<tr>
<th></th>
<th>Votes</th>
<th>Realism</th>
<th>Aesthetics</th>
<th>Quality</th>
<th>just chose cheapest</th>
<th>just chose expensive (wrote in)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Photon Mapping</td>
<td>30</td>
<td>37%</td>
<td>31%</td>
<td>3%</td>
<td>17%</td>
<td>8%</td>
</tr>
<tr>
<td>Film Gather</td>
<td>26</td>
<td>31%</td>
<td>19%</td>
<td>19%</td>
<td>31%</td>
<td>6%</td>
</tr>
<tr>
<td>Light Dome</td>
<td>9</td>
<td>56%</td>
<td>0%</td>
<td>1%</td>
<td>0%</td>
<td>8%</td>
</tr>
</tbody>
</table>

**Shadow Type**

<table>
<thead>
<tr>
<th></th>
<th>Votes</th>
<th>Realism</th>
<th>Aesthetics</th>
<th>Quality</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raytracing</td>
<td>48</td>
<td>54%</td>
<td>26%</td>
<td>0%</td>
<td>29%</td>
</tr>
<tr>
<td>Shadow Mapping</td>
<td>19</td>
<td>15%</td>
<td>15%</td>
<td>0%</td>
<td>79%</td>
</tr>
</tbody>
</table>

**Backdrop Style**

<table>
<thead>
<tr>
<th></th>
<th>Votes</th>
<th>Realism</th>
<th>Aesthetics</th>
<th>Quality</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D method</td>
<td>46</td>
<td>46%</td>
<td>27%</td>
<td>15%</td>
<td>19%</td>
</tr>
<tr>
<td>2.1D method</td>
<td>19</td>
<td>15%</td>
<td>46%</td>
<td>20%</td>
<td>25%</td>
</tr>
</tbody>
</table>

From a population of 500 computer graphics students in 25 colleges and universities across the United States, 66 respondents used the program and participated in the online survey, yielding a margin for error of ± 11.25%, at a 95% confidence level.

**Figure 10.** Online survey results for *Soccer Mom: VFX Supervisor for a Day*.
Of those who said they had no interest at all in visual effects before trying the program, 20% reported that their interest in the subject had increased slightly. 34% of users stated that the program had increased their interest somewhat. 26% of users reported that they, before using the program, thought that they did not have an aptitude for combining live action video with computer graphics. After using the application, 19% of participants reported that the concepts did not seem as difficult as before. 17% indicated that live action/CG integration seemed harder than they originally thought. 34% of users reported an increased awareness of the concept of rendering CG elements in passes, and that they would like to learn more about it.
CHAPTER IV

SUMMARY, CONCLUSIONS AND RECOMMENDATIONS

Initially, this researcher found that the green take, final gather, shadow mapping and the 3D backdrop methods worked together as the best combination of quality and efficiency. This allowed for a comfortable margin in the budget while giving results that looked “good enough.” The majority of those polled voted differently, instead favoring photon mapping and raytracing as the lighting and shadowing methods of choice.

The light matching test results were nearly split evenly between photon mapping and final gather at 45% to 42%, respectively. Statistics indicate that photon mapping was selected for its realism and pleasantness. Several voters wrote in that they chose photon mapping because it looked the best and, combined with their other choices, allowed their decision lists to still come in under budget.

Final gather received its votes on the merit of realism and value, suggesting that quality differences between photon mapping and final gather were not high enough to warrant the more expensive photon mapping choice.

Raytracing was preferred over shadow mapping for its realism and pleasantness. It is not difficult to see why; shadows are crisper and more well-defined,
especially in small nooks and crevices. Despite its longer render time, raytracing took the day. Shadow mapping was selected for its value. It took much less time to render this shadow type. Combined with its low-key presence in the composite, 70% of those who chose shadow mapping did so because it was “good enough.”

Even though it took less time to complete than the 2 ½ D backdrop method, the 3D method won out over 2 ½ D not because of its value, but mostly because of its realism and pleasing look. Concerning the backdrop, realism and aesthetics were more important than value to survey takers.

On the other hand, the colored screen test was dominated by value. Winning 65% of the vote, the take with the green backdrop was preferred over the blue take. There were various considerations for each, but value was the determining factor for most of those who chose green. 46% of those who picked green factored in the quality of the matte with the time involved, but were not particularly bothered by the composition or by the performance of the actresses. It was reported on several occasions that a green selection allowed players to upgrade other departments, like light matching. The blue take, which had a slightly different lighting scenario and performance, was favored by 35% of users for its matte quality, performance, and aesthetics. This take also took quite a bit longer to process than green, so its cost in labor was much higher than green.

From these findings, it may be concluded that cheapest solution is not always the most preferred. Participants seemed to choose the cheaper option if they had to, but if they could upgrade another option in the process, they would. Users of the program tended to push the boundaries of what was affordable, cutting in some areas that are not as important to supplement other areas that are.
Obtaining a conclusion about technical choices was not the only goal of this project. The computer application was designed to serve as a tool to encourage critical thinking about video/CG integration so users could better decide for themselves which techniques are best for a given scenario. Statistics have shown that the program caused a definite shift in participants’ understanding of the process of CG integration. The application succeeded in generating more interest, clarifying difficulty levels, and inspiring students to consider the use of passes in future work.

Recommendations

One of the goals of this text was to provide scene-specific recommendations of integration techniques and, where possible, general recommendations as well. These proposals are a product of study, experiential observations, and subjective data obtained from surveyed participants of the program.

The colored screen endorsement came with ease. While the blue take had some nice features, like the natural rim light and better actress performance, the process time involved in removing the blue background was too high. The amount of rotoscoping needed to remove artifacts was so much greater than green that blue’s inherent benefits were far outweighed by its detriments. Green’s color saturation proved to apply more evenly across its screens and was much easier to key out. As supported by the survey, a green screen will be recommended for outdoor video shoots like this one.

The light-matching recommendation was not as simple as the colored screens. Since the survey results between photon mapping and final gather were close, final gather will still be recommended in this text. Photon mapping took more than twice as long as
final gather and is only slightly richer in color. While photon mapping does have superior quality and pleasantness, as reflected in the survey, it seems risky to select photon mapping simply because other choices can make it affordable. Final gather seems “good enough” for a scene like this. The lesser time spent in production seems to validate this assertion.

It was true that the raytracing method for generating shadows gave crisp, accurate shadows that respected small corners and crevices. In most situations, it would seem desirable to use quality shadows like this, even with the extended render times, instead of shadow mapping. The survey certainly supported this sentiment, indicating that 72% of users favored the raytracing technique. It is logical to believe that, because the shadow component is a relatively small part of the greater share of labor (between .75 hours and 2.75 hours), that it really would not matter which choice was made. The more expensive choice could be made with minimal risk.

However, this thinking could be flawed in larger productions. With multiple shots in production, a two-hour difference could be multiplied with the number of shots. In this particular usage of shadows, where dark areas were dimmed down and blurred for a more subtle effect, the shadow mapping method might be the better choice because of its speed. In the case of this production, the goal was to gain the highest quality and still remain under budget. The raytracing choice, overall, still deserves recommendation for small, low-shot-count productions.

It was easy to recommend the 3D backdrop method over the 2 ½ D method. Voters preferred it by a wide margin for its realism, quality and faster production time. The 2 ½ D technique involved a lengthy process of rendering single frames of each
building separately from 6 different angles, compositing the occlusion pass for each angle, saving each result, and then bringing each premultiplied still render into the compositing program. The 2½D method might be recommended for constructing a cityscape from photographs, but the 3D method was certainly preferred in this particular production.

During production, in a winding process of trial and error, it was found that certain rules of thumb helped move the production along faster. Many of these workflow discoveries came about with the production of the blue take and with its corresponding effects handling. The hard-won findings from the green take helped generate these new guidelines, which were implemented in the blue take for a positive effect.

Previsualization is very important for complex shoots for CG compositing. When possible, a previs should be done for the whole shoot, including the visual effects. A mock-up of important elements should be constructed in 3D space that will help keep things straight in the mind of the director and act as a “dress rehearsal.” Previs artists should build crude models of everything on the proposed set—even actors—and use real world measurements, if possible. Known measurements of the set and props should be used so production designers can figure out the measurements of items that have not been thought of yet. Major elements should be produced, even crudely, for mock compositing to see if those elements fit properly together.

At the video or film shoot, visual effects supervisors should be present to take down measurements of the set. They will come in handy in ways that artists may never expect. Not only are measurements important for matchmoving and establishing the scale
of 3D objects, but they can also pinpoint where real things are so 3D objects can be animated to match them.

After the shoot, the video clip should be trimmed for length and a proxy video should be created before any processing is done. A proxy video is a watered-down version the original, in that it is rendered at lower resolution to make processing faster. Proxy videos are used for matchmoving and element-matching. It is easy to rush into the postprocessing before the video is trimmed, but if artists start to produce rendered assets and the video is trimmed later, it is very difficult to align the asset’s frames with the new video. A matchmoved camera will be out of sync with the footage. Instead, artists should wait for the clip to be trimmed and made into a proxy, first.

Assets should be kept well-organized. Render artists should create a digital folder for each separate element and put sub-folders for each pass inside the first folder. The actual image sequence should be placed inside the pass folder. The omission of this step could result in time lost in finding assets for compositing or for the replacement of such files.

Finally, as learned from the survey results, the production as a whole should be treated like a construction job, where the results of one component’s creation can influence the choices of the next component. It should be kept in mind that the quality of the backdrops in the live-action shoot can greatly affect the rest of the post production process. If money can be saved in one section, much more play room can be had for squeezing in those more render-intensive methods.
Summary

The study of past and present issues with the integration of live actors and computer graphics has shown us the need for efficiency in the visual effects industry. Detailed descriptions of each area inform us of the factors to be considered and choices that must be made. As a review of available literature suggests, there is a wealth of descriptive publication of all ten major concepts, but there are also relatively few sources that seek to encourage the weighing of those concepts together through a single application. Readers have learned about the creation of one such application. The narrative of this genesis recounts the planning and execution of the major concepts outlined in the literary review. The text also describes how the concepts were incorporated as content into the interactive application.

Readers learned that the application allows users to make informed conclusions as to which practices are best used for integrating live action and computer graphics. The application’s testing results were reported upon in this text, reflecting the most preferred techniques among those tested and explores the possible reasons why those preferences exist. In addition, the effectiveness of the application was explored in the context of influencing users’ interests and opinions about integration. General and specific recommendations were proposed concerning the most efficient methods of integration.
REFERENCES
REFERENCES


Debevec, P. (1998). Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range


Retrieved March 17, 2014, from http://www.thetimes.co.uk/tto/arts/film/article2426783.ece


1. Thank you for taking the time to complete this survey. Please paste the code from your "results" page here.


2. Why did you choose the take color that you chose?

☐ I liked the performance of the actresses more
☐ It was a better value for the company—good balance of aesthetics and cost
☐ I liked the lighting angle more
☐ My color choice provided a cleaner matte
☐ I couldn’t decide which I liked more, just went for cheapest

Other (please specify)


3. Why did you choose the global illumination method that you chose?  

Global Illumination Choices

☐ It had the least number of artifacts
☐ It looked the most realistic
☐ It had more aesthetic value
☐ I really couldn’t tell, I just went with the cheapest

Other (please specify)
4. Why did you choose the shadow method that you chose?

☐ It looked the most realistic
☐ I liked the way it looked more
☐ I wanted to stay within budget and the differences were not important
☐ I couldn’t tell the difference, just went with the cheapest

Other (please specify)

5. Why did you choose the CG backdrop method that you chose?

☐ I liked the way my choice looked—seemed nicer
☐ My choice seemed more realistic
☐ My choice had fewer artifacts
☐ Both choices seemed fine—I just went with the cheaper one

Other (please specify)

6. Before you first used “Soccer Mom: VFX Supervisor for a Day,” did you have any preconceptions about the field of visual effects, especially in combining live-action video with CG elements? Please mark all that apply.

☐ Not interested at all in visual effects
☐ Visual effects is a mysterious field. I could never learn how to do such things.
☐ This subject is mildly interesting. It might be an interesting hobby to pursue.
☐ I think it would be cool to learn how to do visual effects, but I don’t think I have an aptitude for it
☐ I’m studying 3D graphics but never really thought about integration of CG and video.
☐ I’m studying 3D graphics but thought it was really hard to integrate CG and video.
☐ I’m studying video production but never really thought about adding CG elements to my work
☐ I’m studying video production but thought it was really hard to add CG elements to my work.
☐ I’m very interested in integrating live action video and 3D graphics and I’m actively studying this.
☐ I’m learning other forms of graphic work but was thinking of learning visual effects, too

Other (please specify)
7. AFTER you first used “Soccer Mom: VFX Supervisor for a Day,” did your preconceptions about visual effects change? Did anything in the content of the program influence your ideas or attitudes about visual effects or about live action/CG integration? Please mark all that apply.

- My interest has been slightly piqued. There are some cool things going on, but not that cool.
- If I had any interests in visual effects or integration, those interests are now gone forever.
- This program has increased my interest somewhat.
- Maybe I can do some of these things after all. It doesn’t seem that hard.
- Visual effects are much harder than I originally thought.
- I think it’s possible for me to include live action video with my 3D animation productions.
- I don’t like the thought of including live action video with my 3D animation productions.
- I’m wondering if I can add 3D graphics to my video productions.
- This is awesome, I intend on adding 3D graphics to my video productions!
- There’s no way I’m adding 3D graphics to my videos. Way too much hassle!
- I see how rendering 3D graphics in passes is important. Maybe I should learn more about it.
- I don’t really agree with the whole “rendering in passes” concept. I’m just gonna do it all in one pass.
- I’m definitely going to render in several passes and try my hand at compositing.

Other (please specify):

8. Which section of “Soccer Mom: VFX Supervisor for a Day” had the most important role in helping you form opinions about visual effects?

- Orientation
- Blue vs. Green Take Selection
- Global Illumination for Diffuse Pass
- Shadow Type
- Background Creation
- Summary (displayed at the end. If you missed it, you can go back and view it)

Other (please specify):

APPENDIX B
DEFINITION OF TERMS

1. **4:2:0 AVCHD format** – Digital video format created primarily for high-definition consumer camcorders. 4:2:0 refers to the format’s chroma subsampling, or the ratio of luminance values (4) to its blue minus luminance values (2) and its red minus luminance values (0).

2. **Antialiasing** – the smoothing of edge artifacts in digital images by introducing semitransparent pixels along edges.

3. **Anticipation** – the motion of an animated character that acts as a precursor to that character’s main motion. Anticipation lets viewers know that a motion is about to occur, usually by applying motion in the opposite direction first.

4. **Atlas Map** – An image used for the UV texturing of a 3D object. Atlas maps are derived from the “unwrapping” of an object’s faces and by the arrangement of these faces into discernible zones for manipulation. The map is then re-applied to the model in the same way that it was unwrapped.

5. **B-spline Patches** - In 3D graphics, a smooth curve, or basis-spline, can be created between two or more points. The locations of the points govern the depth of each curve. Smooth surfaces are created between curves to form a patch.

6. **C2 Continuity** – the smoothness of transition between two end control points of joined curves.

7. **CG** – Computer Graphics

8. **Cone angle** – an angle that defines the radius of the limited area projected by a 3D spotlight. The radius of the projected light is calculated from the center of the light’s projection and a user-defined cone angle.

9. **Demo Reel** – A collection of works by a single author for the purpose of communicating the skill level of that author. Demo reels are often produced on DVDs or uploaded to websites and allow potential employers to evaluate the work.

10. **Diffuse Pass** – a separate rendering of a 3D object’s diffuse attributes. Diffuse attributes describe basic coloring and shading, but excludes reflections, shadows and specular highlights.
11. **Falloff** – also called *decay*, falloff describes how light intensity diminishes over the distance from the light.

12. **Fractal** – describes a partially-random method for generating 3D geometry or textures. Fractal techniques use non-regularity to produce the same randomness on all scales.

13. **ILM** – Abbreviation for Industrial Light and Magic, a large visual effects production company.

14. **Irradiance Caching** – in global illumination, the storing of calculated light values from specific points on a surface, spread out over a distance. To save render time, surface colors are interpolated between these calculated points.

15. **Lens Flare** – in computer graphics, a simulation of light that scatters through a real camera lens, causing visible artifacts in the shape of streaks, starbursts, rings or hexagons. Simulated lens flares can be used to create a sense of realism.

16. **Luminance** – brightness of a light source, measured in candelas per square meter.

17. **Nodal Pan** – a method for creating a digital scene that involves a simple horizontal camera rotation on a tripod. A nodal pan does not create parallax, so all digital elements can be combined into a single image to pan across.

18. **Pivot Point** – in 3D graphics, the local center of a distinct object. All rotations are executed with the object’s pivot point at the center.

19. **Pixelated** – describes an image that has a small number of large pixels (picture elements) and has the attributes of a low-resolution image format: edges with artifacts and loss of detail.

20. **Post-processing** – the manipulation of video or computer graphics after they have been initially produced.


22. **Radiance** – the measure of how much radiation falls within a given radius in a specific direction.

23. **Radio Buttons** – In a graphic interface, a radio button is a simple element that allows users to make selections. This is represented by a clickable graphic, usually in the shape of a small circle. Upon clicking, the circle will be filled in, indicating the selection.
24. **RAM** – Abbreviation for Random Access Memory. This is a temporary storage area for memory in a computer system, crucial for allowing robust programs to work smoothly.

25. **Raster-based image**– also called a bitmap image. Raster-based images and textures are made from a fixed number of pixels. Changes in size will affect the overall quality of the image, as opposed to vector-based graphics, which are mathematically calculated and scalable.

26. **REYES** – Abbreviation for Renders Everything You Ever Saw, a rendering system that uses micropolygons to represent curved surfaces and save time with little raytracing.

27. **Secondary Action** – In animation, secondary action is added to the main action of a character to add extra life. Secondary action supports and emphasizes the main action.

28. **Spill Light** – Unwanted light that bounces from objects and shines on live actors. Blue or green screens sometimes cast spill light that can make matte production difficult.

29. **Squash and Stretch** – Major animation principle for adding weight to an object or character as it moves. As an object may stretch along one axis, it will squash along another, providing a sense that the object’s volume is preserved.

30. **Straight Image** – In video compositing, an image that has not been premultiplied by its integrated alpha channel, which describes opacity. A straight image’s color information has also not been affected by the alpha channel.

31. **UV Mapping** – A method for applying textures onto a 3D object that involves establishing texture coordinates. These coordinates may be moved along a relative axial grid, in the U and V directions, to distort a texture to fit the model.

32. **VFX** – Abbreviation for visual effects.

33. **Video Tap** – a small device, inside a movie camera’s viewfinder, that would allow the director to see the same view as the camera operator and record it to videotape. This let directors create intermediate rough cuts without waiting for the film to be developed.

34. **Wireframe** – a display style that allows 3D models to be viewed without textures, saving computational resources while modeling and animation is completed. This style displays only the points and connecting lines of a model, providing a view that resembles a frame made from wires.